

O'REILLY®

Google BigQuery

Всё о хранилищах данных, аналитике
и машинном обучении



Валиappa Лакшманан
Джордан Тайджани

Google BigQuery: The Definitive Guide

*Data Warehousing, Analytics, and
Machine Learning at Scale*

Valliappa Lakshmanan and Jordan Tigani

Beijing • Boston • Farnham • Sebastopol • Tokyo

O'REILLY®

Google BigQuery

Всё о хранилищах данных, аналитике
и машинном обучении

Валиappa Лакшманан
Джордан Тайджани



Санкт-Петербург • Москва • Екатеринбург • Воронеж
Нижний Новгород • Ростов-на-Дону
Самара • Минск

2021

ББК 32.973.233-018
УДК 004.65
Л19

Лакшманан Валиappa, Тайджани Джордан

Л19 Google BigQuery. Всё о хранилищах данных, аналитике и машинном обучении. — СПб.: Питер, 2021. — 496 с.: ил. — (Серия «Бестселлеры O'Reilly»). ISBN 978-5-4461-1707-9

Вас пугает необходимость обрабатывать петабайтные наборы данных? Познакомьтесь с Google BigQuery, — системой хранения информации, которая может консолидировать данные по всему предприятию, облегчает интерактивный анализ и позволяет реализовать задачи машинного обучения. Теперь вы можете эффективно хранить, запрашивать, получать и изучать данные в одной удобной среде.

Валиappa Лакшманан и Джордан Тайджани научат вас работать в современном хранилище данных, используя все возможности масштабируемого, безсерверного публичного облака.

16+ (В соответствии с Федеральным законом от 29 декабря 2010 г. № 436-ФЗ.)

ББК 32.973.233-018
УДК 004.65

Права на издание получены по соглашению с O'Reilly. Все права защищены. Никакая часть данной книги не может быть воспроизведена в какой бы то ни было форме без письменного разрешения владельцев авторских прав.

Информация, содержащаяся в данной книге, получена из источников, рассматриваемых издательством как надежные. Тем не менее, имея в виду возможные человеческие или технические ошибки, издательство не может гарантировать абсолютную точность и полноту приводимых сведений и не несет ответственности за возможные ошибки, связанные с использованием книги. Издательство не несет ответственности за доступность материалов, ссылки на которые вы можете найти в этой книге. На момент подготовки книги к изданию все ссылки на интернет-ресурсы были действующими.

ISBN 978-1492044468 англ.

Authorized Russian translation of the English edition of Google BigQuery: The Definitive Guide ISBN 9781492044468 © 2020 Valliappa Lakshmanan and Jordan Tigani. This translation is published and sold by permission of O'Reilly Media, Inc., which owns or controls all rights to publish and sell the same.
© Перевод на русский язык ООО Издательство «Питер», 2021
© Издание на русском языке, оформление ООО Издательство «Питер», 2021
© Серия «Бестселлеры O'Reilly», 2021

ISBN 978-5-4461-1707-9

Оглавление

Отзывы о книге «Google BigQuery. Всё о хранилищах данных, аналитике и машинном обучении»	13
Предисловие	15
Для кого написана эта книга?.....	16
Условные обозначения	16
Использование примеров программного кода	17
Благодарности	17
От издательства	18
Глава 1. Что такое Google BigQuery?	19
Архитектуры обработки данных	19
Система управления реляционными базами данных.....	20
Фреймворк MapReduce.....	22
BigQuery: бессерверный распределенный движок SQL.....	23
Работа с BigQuery.....	25
Анализ наборов данных.....	25
ETL, EL и ELT	26
Эффективная аналитика	28
Простота управления.....	30
История появления BigQuery.....	31
Что позволило создать BigQuery?	34
Отделение вычислений от хранилища	35
Хранилище и сетевая инфраструктура.....	36
Управляемое хранилище.....	37
Интеграция с платформой Google Cloud	39
Безопасность и соответствие требованиям	40
Выводы	41

Глава 2. Основы запросов	42
Простые запросы.....	44
Извлечение записей с помощью SELECT	44
Создание псевдонимов столбцов с помощью AS	46
Фильтрация с WHERE	48
SELECT *, EXCEPT, REPLACE.....	49
Подзапросы с WITH	50
Сортировка с ORDER BY.....	50
Агрегирование.....	51
Агрегирование с GROUP BY	51
Подсчет записей с COUNT	52
Фильтрация сгруппированных значений с HAVING	53
Поиск уникальных значений с DISTINCT	54
Краткое руководство по массивам и структурам.....	55
Создание массивов с помощью ARRAY_AGG.....	57
Массив структур STRUCT	59
Кортежи	60
Работа с массивами	60
Развертывание массива	61
Соединение таблиц	62
Основы соединения таблиц	63
Оператор внутреннего соединения INNER JOIN.....	66
Оператор перекрестного соединения CROSS JOIN.....	67
Оператор внешнего соединения OUTER JOIN.....	69
Сохранение и совместное использование	70
История запросов и кеширование	70
Сохранение запросов	72
Представления и общедоступные запросы.....	73
Выводы	74
Глава 3. Типы данных, функции и операторы.....	75
Числовые типы и функции.....	76
Математические функции	77
Стандартное вещественное деление.....	78

Функции SAFE	78
Сравнение	79
Точные десятичные вычисления с NUMERIC	80
Тип BOOL	81
Логические операции	81
Условные выражения	83
Обработка NULL с помощью COALESCE	84
Явное и неявное приведение типов	85
Использование COUNTIF, чтобы избежать приведения логических значений	87
Строковые функции	88
Интернационализация	89
Формирование и парсинг строк	91
Функции для обработки строк	91
Функции преобразования	92
Регулярные выражения	92
Краткие итоги по строковым функциям	94
Операции со значениями TIMESTAMP	95
Парсинг и форматирование отметок времени	95
Извлечение календарных данных	97
Арифметические операции с временными метками	98
DATE, TIME и DATETIME	99
Функции для работы с географическими координатами	100
Выводы	101
Глава 4. Загрузка данных в BigQuery	104
Основы	104
Загрузка из локального источника	105
Корректировка схемы	112
Копирование в новую таблицу	116
Управление данными (DDL и DML)	116
Эффективная загрузка данных	118
Федеративные запросы и внешние источники данных	121
Как использовать федеративные запросы	121
Когда использовать федеративные запросы и внешние источники данных	125

Интерактивное исследование и запрос данных из Google Sheets	132
Запросы SQL для выборки данных из Cloud Bigtable	141
Передача и экспорт данных.....	147
Служба передачи данных Data Transfer Service.....	147
Экспортирование журналов Stackdriver	153
Использование Cloud Dataflow для чтения/записи в BigQuery	154
Перемещение локальных данных.....	159
Методы миграции данных	159
Выводы	162
Глава 5. Разработка с BigQuery	163
Программный доступ	163
Доступ к BigQuery через REST API.....	163
Google Cloud Client Library	171
Доступ к BigQuery из инструментов исследования данных	188
Блокноты в Google Cloud Platform	188
Работа с BigQuery, pandas и Jupyter	193
Работа с BigQuery из R.....	198
Cloud Dataflow	199
Драйверы JDBC/ODBC.....	202
Внедрение данных из BigQuery в Google Slides (в G Suite)	203
Bash-скрипты для BigQuery.....	205
Создание наборов данных и таблиц	205
Выполнение запросов	208
Объекты BigQuery	210
Выводы	212
Глава 6. Архитектура BigQuery	213
Архитектура высокого уровня	213
Жизненный цикл запроса	213
Обновление BigQuery	218
Система обработки запросов (Dremel).....	219
Архитектура Dremel.....	221
Выполнение запроса	226
Хранилище.....	241
Хранение данных	241

Метаданные	248
Выводы	258
Глава 7. Оптимизация производительности и затрат	259
Принципы производительности	259
Ключевые составляющие производительности.....	260
Управление затратами.....	260
Измерение производительности и поиск проблем	262
Определение скорости выполнения запроса с помощью REST API.....	263
Определение скорости выполнения запроса с помощью BigQuery Workload Tester	265
Выявление проблем в рабочих нагрузках с помощью Stackdriver	267
Чтение плана запроса	269
Увеличение скорости выполнения запросов	274
Минимизация ввода/вывода	275
Кеширование результатов предыдущих запросов	280
Эффективное выполнение соединений	284
Исключение перегрузки рабочих серверов	293
Использование приближенных функций агрегирования.....	296
Оптимизация хранения данных и доступа к ним.....	299
Минимизация сетевых издержек	300
Выбор эффективного формата хранения	303
Секционирование таблиц для уменьшения объема сканирования	313
Кластеризация таблиц на основе ключей с большой мощностью множества	316
Случаи использования, нечувствительные ко времени	321
Пакетные запросы	321
Загрузка файлов	323
Выводы	324
Контрольный список	324
Глава 8. Продвинутые запросы	326
Множественные запросы	326
Параметризованные запросы.....	327
Пользовательские функции SQL	332
Повторное использование частей запросов	337

Продвинутый SQL	341
Работа с массивами	342
Оконные функции	351
Метаданные таблиц	356
Язык определения данных и язык манипулирования данными	360
За пределами SQL	365
Пользовательские функции на JavaScript	366
Скрипты	367
Продвинутые функции	375
Геоинформационная система BigQuery	375
Полезные статистические функции	383
Алгоритмы хеширования	385
Выводы	389
Глава 9. Машинное обучение в BigQuery	390
Что такое машинное обучение?	390
Формулировка задачи машинного обучения	391
Типы задач машинного обучения	392
Построение регрессионной модели	396
Выбор метки	396
Выбор признаков в наборе данных	397
Создание обучающего набора данных	401
Обучение и оценка модели	402
Получение прогнозов с помощью модели	404
Исследование весов модели	407
Более сложные регрессионные модели	409
Создание модели классификации	414
Обучение	415
Оценка	416
Прогнозирование	417
Выбор порога	418
Настройка механизма машинного обучения в BigQuery	420
Управление делением данных	420
Балансировка классов	422
Регуляризация	422

Кластеризация методом k-средних.....	423
Выбор признаков для кластеризации.....	424
Кластеризация пунктов проката велосипедов	425
Кластеризация.....	426
Исследование кластеров.....	427
Принятие решений на основе данных	429
Рекомендательные системы	430
Набор данных MovieLens.....	430
Разложение матрицы	432
Получение рекомендаций	434
Включение информации о пользователях и фильмах.....	436
Нестандартные модели машинного обучения в GCP.....	443
Настройка гиперпараметров	444
AutoML	448
Поддержка TensorFlow.....	450
Выводы	453
Глава 10. Администрирование и безопасность BigQuery.....	455
Защищенность инфраструктуры	455
Управление идентификацией и доступом	457
Идентификация.....	457
Роль	458
Ресурс.....	461
Администрирование BigQuery.....	462
Управление заданиями	462
Авторизация пользователей	463
Восстановление удаленных записей и таблиц	463
Непрерывная интеграция/непрерывное развертывание	464
Экспорт биллинга — получение информации о расходах.....	467
Оперативные панели, мониторинг и журналы аудита	470
Доступность, восстановление после отказа и шифрование	471
Зоны, регионы и объединения регионов.....	471
BigQuery и обработка отказов	472
Сохранность, резервное копирование и восстановление после аварий.....	476
Конфиденциальность и шифрование.....	477

Соответствие требованиям законодательств	478
Местоположение данных.....	478
Ограничение доступа к подмножествам данных.....	480
Удаление всех сделок, связанных с конкретным физическим лицом	483
Предотвращение потери данных	487
СМЕК.....	488
Защита от утечки данных.....	490
Выводы	491
Об авторах	493
Об обложке	494

Отзывы о книге «Google BigQuery. Всё о хранилищах данных, аналитике и машинном обучении»

Эта книга будет полезна организациям, которые переходят от применения устаревших технологий хранения корпоративных данных к использованию Google Cloud. Лак и Джордан подробно описывают BigQuery, чтобы вы могли не только использовать эту технологию для хранения корпоративных данных и бизнес-аналитики, но и выполнять SQL-запросы для получения потоков данных в масштабе реального времени, обращаться к BigQuery из кластеров Hadoop и Spark и использовать машинное обучение для автоматической классификации и получения прогнозов на основе данных.

Томас Курьян, генеральный директор Google Cloud

Иногда в мире технологий появляется какое-то программное обеспечение или сервис, которые все в корне меняет. Технология BigQuery кардинально изменила способ представления корпоративных данных. Будучи изначально предназначенной для работы с гигантскими наборами данных, BigQuery стала одной из лучших платформ для анализа и изучения данных. «Стандартный SQL», который был анонсирован в июне 2016 года, является одной из самых понятных, полных и функциональных реализаций SQL за все время. К числу наиболее функциональных особенностей, кроме всего прочего, относятся: поддержка глубоко вложенных данных, пользовательские функции на JavaScript и SQL, геопространственные данные, интегрированное машинное обучение и доступ к данным по URL-адресам. Едва ли вы найдете источник информации о BigQuery лучше, чем книга Джордана и Лака — людей, которые знают о BigQuery гораздо больше многих других.

Ллойд Табб, сооснователь и технический директор Looker

Даже при том, что я пользуюсь BigQuery уже больше семи лет, из этой книги я узнал много нового! Она содержит бесценную информацию о лучших приемах и методах, а сложные идеи объясняются простым языком. Примеры кода помогают увидеть применение теории на практике, благодаря чему книга получилась интересной и увлекательной. Вне всяких сомнений, она станет одним из лучших справочников для пользователей BigQuery.

Грэм Полли, управляющий консультант Servian

Благодаря BigQuery вы сможете обрабатывать большие объемы данных быстрее и дешевле. Эта платформа поможет вам собрать все данные в одном месте и быстро ознакомиться с ними. В книге подробно описываются ключевые компоненты BigQuery. Два выдающихся сотрудника Google — Лак Лакшманан и Джордан Тайджани — познакомят вас с основами BigQuery, а также с другими весьма сложными темами, такими как машинное обучение. Я давний поклонник BigQuery, и как пользователь этого инструмента могу сказать, что он сделает вашу жизнь с большими данными проще. Я испытал истинное наслаждение, читая эту книгу, а теперь это увлекательное путешествие в BigQuery начинается для вас!

Михаил Берлянт, первый вице-президент по технологиям Viant Inc.

Предисловие

Успех предприятий все больше зависит от данных, а ключевым компонентом информационной стратегии любого предприятия является хранилище данных — центральное хранилище интегрированных данных, стекающихся из всех подразделений компании. Обычно аналитики использовали хранилище данных для формирования аналитических отчетов. Но теперь оно все чаще используется для отображения информации в панелях мониторинга (дашбордах) в режиме реального времени, выполнения специализированных запросов и формирования рекомендаций по принятию решений с помощью прогнозной аналитики. Растущие бизнес-требования к углубленной аналитике, оптимизации затрат, гибкости и самообслуживанию доступа к данным заставляют многие организации переходить на использование облачных хранилищ данных, таких как Google BigQuery.

В этой книге мы отправимся в глубины BigQuery — бессерверное, легко-масштабируемое и недорогое корпоративное хранилище данных, доступное в Google Cloud. Отсутствие инфраструктуры дает предприятиям возможность сосредоточиться на анализе данных и находить ценные идеи, используя хорошо знакомый язык SQL.

Работая над BigQuery, мы стремились создать платформу, которая предлагает передовые возможности, использует преимущества многих замечательных технологий, доступных в современных облачных окружениях, и поддерживает проверенные временем технологии, актуальные и сейчас. Например, главное преимущество Google BigQuery — это бессерверная вычислительная архитектура, которая отделяет вычисления от хранилища. Такой подход позволяет разным уровням архитектуры функционировать и масштабироваться независимо друг от друга, а также дает разработчикам баз данных гибкость при разработке и развертывании. Уникальной чертой BigQuery является встроенная поддержка машинного обучения и геопространственного анализа. В сочетании с Pub/Sub, Cloud Dataflow, Cloud Bigtable, Cloud AI Platform и многими сторонними

компонентами платформа BigQuery способна взаимодействовать и с традиционными, и с современными системами в широком диапазоне требований к пропускной способности и задержкам. Наконец, BigQuery поддерживает ANSI-стандарт SQL, колоночную оптимизацию и федеративные запросы — ключевые элементы самостоятельного исследования данных, востребованные многими пользователями.

Для кого написана эта книга?

Эта книга адресована аналитикам, инженерам, а также специалистам по обработке и анализу данных, желающим использовать BigQuery для извлечения информации из больших наборов данных. Дата-аналитики могут взаимодействовать с BigQuery, используя SQL и инструменты мониторинга, такие как Looker, Data Studio и Tableau. Дата-инженеры могут интегрировать BigQuery в конвейеры, написанные на Python или Java, и использовать такие фреймворки, как Apache Spark и Apache Beam. Специалисты по обработке и анализу данных могут создавать модели машинного обучения в BigQuery, запускать модели TensorFlow для обучения на данных в BigQuery и делегировать выполнение распределенных массивных вычислений платформе BigQuery из блокнота Jupyter.

Условные обозначения

В данной книге используются следующие типографские обозначения:

Курсив

Используется для обозначения новых терминов, адресов URL и электронной почты, имен файлов и расширений имен файлов.

Моноширинный шрифт

Применяется для оформления листингов программ и программных элементов внутри обычного текста, таких как имена переменных и функций, баз данных, типов данных, переменных среды, операторов и ключевых слов.

Моноширинный жирный

Обозначает команды или другой текст, который должен вводиться пользователем.

Моноширинный курсив

Обозначает текст, который должен замещаться фактическими значениями, вводимыми пользователем или определяемыми из контекста.



Так выделяются советы и предложения.



Так обозначаются советы, предложения и примечания общего характера.



Так обозначаются предупреждения и предостережения.

Использование примеров программного кода

Вспомогательные материалы (примеры кода, упражнения и т. д.) доступны для загрузки по адресу <https://github.com/GoogleCloudPlatform/bigquery-oreilly-book>.

Если у вас возникнут вопросы технического характера по использованию примеров кода, направляйте их по электронной почте bookquestions@oreilly.com.

Эта книга написана для того, чтобы помочь вам решать необходимые задачи. В целом вы можете использовать все примеры кода из этой книги в своих программах и в документации. Вам не нужно обращаться в издательство за разрешением, если вы не собираетесь воспроизводить существенные части программного кода. Например, если вы разрабатываете программу и используете в ней несколько отрывков программного кода из книги, разрешение не требуется. Однако в случае продажи или распространения примеров из этой книги вам необходимо получить разрешение от издательства O'Reilly. Если вы отвечаете на вопросы, цитируя данную книгу или примеры из нее, получение разрешения не требуется. Но при включении существенных объемов программного кода примеров из этой книги в вашу документацию вам необходимо будет получить разрешение издательства.

Благодарности

Нам (Лаку и Джордану) очень повезло с рецензентами — Эллиот Броссард (Elliott Brossard), Эван Джонс (Evan Jones), Грэм Полли (Graham Polley), Ребекка Уорд (Rebecca Ward) и Тиган Тигани (Tegan Tigani) внимательно прочитали каждую главу этой книги и внесли многочисленные предложения. Эллиот помогал нам писать более простые и понятные запросы SQL. Опыт Эвана

пригодился, когда мы работали над описанием особенностей использования BigQuery в Google Finance. Грэм помог нам взглянуть на многие аспекты, касающиеся стоимости и регионализации, с точки зрения клиента. Ребекка снабжала нас фактами, а Тиган позаботилась о том, чтобы книга была написана простым и понятным языком. Нам также помогали многие сотрудники Google, каждый в своей сфере компетенций: Чед Дженнингс (Chad Jennings), Харис Хан (Haris Khan), Миша Брукман (Misha Brukman), Даниэль Гундрум (Daniel Gundrum), Моша Пашумански (Mosha Pashumansky), Амир Хормати (Amir Hormati) и Мингге Денг (Mingge Deng). Любые ошибки, оставшиеся неисправленными, — это только наша вина.

Спасибо нашим семьям, товарищам по команде и руководителям — Рочану Голани (Rochana Golani) и Судхиру Хасбе (Sudhir Hasbe) за поддержку. Мы получили большое удовольствие от работы с нашими редакторами в издательстве O'Reilly: Николь Таше (Nicole Taché) и Кристен Браун (Kristen Brown). Благодаря усилиям Боба Рассела (Bob Russell), нашего литературного редактора, текст получился намного лучше. Идея написать эту книгу принадлежит Саптарши Мукерджи (Saptarshi Mukherjee) — именно он подтолкнул нас к совместной работе над новой книгой о BigQuery. Наконец, мы хотели бы поблагодарить пользователей BigQuery (и конкурентов!) за то, что помогли нам сделать BigQuery лучше, а также команду разработчиков BigQuery, воплотивших это волшебство в жизнь.

Весь гонорар за эту книгу мы перечислим местной организации United Way of King County (<https://www.uwkc.org>). Мы советуем и вам принять участие в работе местной благотворительной организации, которая будет оказывать безвозмездную помощь в решении самых сложных локальных проблем.

От издательства

Ваши замечания, предложения, вопросы отправляйте по адресу comp@piter.com (издательство «Питер», компьютерная редакция).

Мы будем рады узнать ваше мнение!

На веб-сайте издательства www.piter.com вы найдете подробную информацию о наших книгах.

Что такое Google BigQuery?

Архитектуры обработки данных

Google BigQuery — это бессерверное масштабируемое хранилище данных со встроенным механизмом запросов. Механизм запросов способен выполнять запросы SQL на терабайтах данных за считанные секунды, а на петабайтах — за считанные минуты. Чтобы получить такую производительность, вам не придется организовывать и поддерживать какую-либо инфраструктуру и создавать или перестраивать индексы.

У BigQuery масса поклонников. Пол Ламер (Paul Lamere), инженер из компании Spotify, был рад наконец-то рассказать о том, как его команда использует BigQuery для анализа больших наборов данных: «Google BigQuery — это *просто бомба*», — написал он в Твиттере в феврале 2016 года (<https://twitter.com/plamere/status/702168809445134336>). «Из 2,2 миллиарда записей я смог вычленил 20 тысяч меньше чем за минуту». Масштаб и скорость — это лишь две из примечательных особенностей BigQuery. Не менее важно отсутствие необходимости поддерживать инфраструктуру, поскольку простота бессерверных вычислений способна открыть новые способы работы.

Компании все активнее принимают решения, основанные на анализе данных, и поощряют открытую культуру, когда доступность данных не ограничивается подразделениями. Предлагая средства для реализации гибкости и открытости, BigQuery играет важную роль в ускорении внедрения инноваций. Например, недавно представитель Twitter сообщил (https://blog.twitter.com/engineering/en_us/topics/infrastructure/2019/democratizing-data-analysis-with-google-bigquery.html), что с помощью BigQuery им удалось сделать анализ данных доступнее, предоставив некоторые часто используемые таблицы сотрудникам Twitter из разных подразделений (в частности, отделы проектирования, управления финансами и маркетинга).

Для Alpega Group, компании-разработчика программного обеспечения для логистики, инновационный потенциал и гибкость BigQuery стали залогом

успеха. Компания преодолела барьер, мешавший быстрому получению результатов анализа информации, и сумела обеспечить своих клиентов аналитикой практически в реальном времени. Поскольку в Alpega Group избавились от необходимости содержать кластеры и необходимую для этого инфраструктуру, ее небольшая техническая команда теперь может свободно заниматься разработкой программного обеспечения и средствами обработки данных. «Это стало для нас потрясающей новостью», — сказал ведущий архитектор компании Аарт Вербеке (Aart Verbeke, <https://cloud.google.com/customers/alpega>). «При традиционном подходе нам пришлось бы устанавливать, настраивать, разворачивать и размещать каждый отдельный компонент. Здесь же мы просто подключаемся к сервису и используем его по мере необходимости».

Представьте, что вы руководите сетью пунктов проката оборудования. Цена зависит от срока аренды, поэтому, чтобы правильно подсчитать сумму, которую должен заплатить клиент, вам нужна следующая информация:

1. Где было арендовано оборудование.
2. Когда оно было арендовано.
3. Куда его вернули.
4. Когда его вернули.

Возможно, вы вносите эту информацию в базу данных каждый раз, когда клиент возвращает оборудование.¹

На основании этого набора данных вы можете определить количество ежемесячных арендных платежей за последние 10 лет. Или, может быть, вы хотите ввести доплату за возврат оборудования в другой пункт проката, и вам нужно выяснить, на какую долю аренды повлияет это нововведение. Желание знать ответ на подобные вопросы возникает весьма часто, но важно, чтобы вы сами умели найти ответ, если вы хотите принимать решения, основанные на данных.

Какую архитектуру вы могли бы использовать? Рассмотрим некоторые возможные варианты.

Система управления реляционными базами данных

Информацию о сделках можно записывать в реляционную базу данных с оперативной обработкой транзакций (OnLine Transaction Processing, OLTP), например MySQL или PostgreSQL. Одним из ключевых преимуществ таких баз данных является поддержка запросов на языке структурированных запросов (Structured Query Language, SQL) — вашим сотрудникам не придется использовать высоко-

¹ На самом деле запись должна создаваться в момент получения оборудования в аренду, чтобы можно было узнать, не сбежал ли клиент с ним. Однако на данном этапе об этом пока рано волноваться!

уровневые языки программирования, такие как Java или Python, чтобы получать ответы на возникающие вопросы. Они могут просто писать запросы, которые можно отправить серверу базы данных:

```
SELECT
  EXTRACT(YEAR FROM starttime) AS year,
  EXTRACT(MONTH FROM starttime) AS month,
  COUNT(starttime) AS number_one_way
FROM
  mydb.return_transactions
WHERE
  start_station_name != end_station_name
GROUP BY year, month
ORDER BY year ASC, month ASC
```

Не обращайтесь пока внимания на синтаксис, подробнее мы поговорим о SQL-запросах позже, а сейчас давайте сосредоточимся на преимуществах и недостатках баз данных OLTP.

Во-первых, обратите внимание, что язык SQL позволяет не просто получать исходные данные, хранящиеся в столбцах таблиц, — предыдущий запрос анализирует отметку времени и извлекает из нее год и месяц. Он также выполняет агрегирование (подсчитывает число строк), фильтрацию (выбирает сделки, в которых пункт, где было взято оборудование, не совпадает с пунктом его возвращения), группировку (по году и месяцу) и сортировку. Важным преимуществом SQL является возможность конкретизировать, что именно мы хотим получить, и позволить программному обеспечению базы данных определить оптимальный способ выполнения запроса.

К сожалению, базы данных OLTP довольно неэффективно выполняют подобные запросы. Основной упор в них делается на согласованность данных; дело в том, что базы данных позволяют считывать данные, даже когда в них одновременно производится запись. Это достигается за счет блокировок, обеспечивающих сохранение целостности данных. Чтобы фильтрация по полю `station_name` выполнялась эффективно, необходимо создать *индекс* для поля с названием пункта проката. Только если название пункта проката индексируется, база данных будет выполнять специальные операции с хранилищем для оптимизации поиска, правда, при этом увеличение скорости чтения достигается за счет уменьшения скорости записи. Если название пункта проката не индексируется, фильтрация по этому полю будет работать довольно медленно. И даже если для названия пункта проката создать индекс, этот конкретный запрос все равно будет выполняться медленно из-за операций агрегирования, группировки и сортировки. Базы данных OLTP не созданы для таких ситуативных¹ запросов, требующих выполнить обход всего набора данных.

¹ В этой книге под «ситуативными» мы будем понимать запросы, написанные без какой-либо предварительной подготовки базы данных с применением таких возможностей, как индексы.

Фреймворк MapReduce

Поскольку базы данных OLTP плохо подходят для ситуативных запросов и запросов, требующих обхода всего набора данных, специализированные виды анализа, требующие такого обхода, можно запрограммировать на языках высокого уровня, таких как Java или Python. В 2003 году Джефф Дин (Jeff Dean) и Санджай Гемават (Sanjay Ghemawat) заметили, что они и их коллеги из Google реализуют сотни таких специализированных вычислений для обработки больших объемов исходных данных. Для решения этой проблемы они разработали абстракцию, позволяющую выражать вычисления в виде двух шагов: функции *map* (отображения), которая обрабатывает пару ключ/значение и генерирует набор промежуточных пар ключ/значение, и функции *reduce* (свертки), которая объединяет все промежуточные значения, связанные с тем же промежуточным ключом.¹ Эта парадигма, известная как *MapReduce*, оказала существенное влияние на фреймворки и привела к разработке Apache Hadoop.

Экосистема Hadoop начиналась как библиотека, написанная на Java, но теперь пользовательский анализ в кластерах Hadoop обычно выполняется с использованием Apache Spark (<http://spark.apache.org/>). Spark поддерживает программы, написанные на Python или Scala, а также позволяет выполнять специальные SQL-запросы к распределенным наборам данных.

То есть чтобы узнать количество платежей за аренду, можно организовать следующий конвейер данных:

1. Периодически экспортировать записи в текстовые файлы в формате CSV (Comma-Separated Values — значения, разделенные запятыми) в распределенной файловой системе Hadoop (Hadoop Distributed File System, HDFS).
2. Для выполнения ситуативного анализа написать программу Spark, которая:
 - а) выгружает данные из текстовых файлов в «DataFrame»;
 - б) выполняет SQL-запрос, подобный запросу, представленному в предыдущем разделе, за исключением того, что имя таблицы заменяется именем кадра данных DataFrame;
 - в) экспортирует результаты обратно в текстовый файл.
3. Выполнить программу Spark в кластере Hadoop.

На первый взгляд архитектура может показаться простой, тем не менее, в ней есть ряд скрытых недостатков. Для сохранения данных в HDFS необходим достаточно большой кластер. Еще одна немаловажная особенность архитектуры MapReduce, которой часто пренебрегают, заключается в том, что обычно

¹ Jeffrey Dean and Sanjay Ghemawat, «MapReduce: Simplified Data Processing on Large Clusters», OSDI '04: Sixth Symposium on Operating Systems Design and Implementation, San Francisco, CA (2004), pp. 137–150. Доступна по ссылке <https://research.google.com/archive/mapreduce-osdi04.pdf>.

вычислительным узлам требуется доступ к локальным для них данным. Соответственно, файловая система HDFS должна быть разбита на сегменты между вычислительными узлами кластера. Поскольку объемы данных и потребности в анализе быстро растут независимо друг от друга, часто происходит так, что кластеры испытывают нехватку или переизбыток вычислительных мощностей.¹ То есть чтобы выполнять программы Spark в кластере Hadoop, сотрудникам вашей организации придется стать экспертами в управлении, мониторинге и настройке кластеров Hadoop. Это может не входить в ваши планы.

BigQuery: бессерверный распределенный движок SQL

А теперь представьте, что у вас есть возможность выполнять SQL-запросы как в системе управления реляционными базами данных (Relational Database Management System, RDBMS), эффективно выполнять обход распределенных наборов данных как в MapReduce и не обременять себя поддержкой инфраструктуры. Это третий вариант, и именно эти особенности делают BigQuery таким привлекательным. BigQuery — это бессерверная служба, позволяющая выполнять запросы, не требуя поддерживать свою инфраструктуру. Она дает возможность анализировать большие наборы данных и агрегировать их в считанные секунды или минуты.

Мы не призываем вас верить нам на слово. Попробуйте сами. Перейдите по ссылке <https://console.cloud.google.com/bigquery> (зарегистрируйтесь в Google Cloud Platform и, если потребуется, выберите свой проект), скопируйте и вставьте следующий запрос в окно,² а затем щелкните на «Run query» («Выполнить»):

```
SELECT
  EXTRACT(YEAR FROM starttime) AS year,
  EXTRACT(MONTH FROM starttime) AS month,
  COUNT(starttime) AS number_one_way
FROM
  `bigquery-public-data.new_york_citibike.citibike_trips`
WHERE
  start_station_name != end_station_name
GROUP BY year, month
ORDER BY year ASC, month ASC
```

¹ Облачная платформа Google Cloud Dataproc (управляемое предложение Hadoop) решает эту проблему по-другому. Благодаря высокой пропускной способности в центрах обработки данных Google, кластеры Cloud Dataproc могут быть привязаны к конкретной работе — данные хранятся в облачном хранилище Google Cloud Storage и передаются по сети по запросу. Такое возможно только при достаточно высокой пропускной способности, сопоставимой с пропускной способностью дисков. Не пытайтесь повторить.

² Для удобства копирования и вставки все фрагменты кода и запросов, которые приведены в этой книге, включая запрос в этом примере (https://github.com/GoogleCloudPlatform/bigquery-oreilly-book/blob/master/01_intro/queries.txt), доступны в репозитории GitHub (<https://github.com/GoogleCloudPlatform/bigquery-oreilly-book>).

Когда мы запустили его, пользовательский интерфейс BigQuery сообщил, что запрос обработал 2.51 Гбайт данных и на это потребовалось примерно 2.7 секунды, как показано на рис. 1.1.

Query editor

HIDE EDITOR

```

1 SELECT
2   EXTRACT(YEAR FROM starttime) AS year,
3   EXTRACT(MONTH FROM starttime) AS month,
4   COUNT(starttime) AS number_one_way
5 FROM
6   `bigquery-public-data.new_york_citibike.citibike_trips`
7 WHERE
8   start_station_name != end_station_name
9 GROUP BY year, month
10 ORDER BY year ASC, month ASC

```

Run query

Save query

Save view

More

This query will process 2.51 GB when run.

Query results

SAVE RESULTS

Query complete (2.704 sec elapsed, 2.51 GB processed)

Job information

Results

JSON

Execution details

Row	year	month	number_one_way
1	2013	7	815324
2	2013	8	970474
3	2013	9	1007799

Рис. 1.1. Выполнение запроса для определения количества платежей за прокат в веб-интерфейсе BigQuery

Арендуемое оборудование — это велосипеды, таким образом, предыдущий запрос подводит ежемесячный итог проката велосипедов в Нью-Йорке по большому набору данных. Сам набор данных находится в открытом доступе (то есть любой желающий может запросить эти данные) и выпущен в Нью-Йорке в рамках инициативы «Открытый город». Результат этого запроса показывает, что в июле 2013 года в Нью-Йорке велосипеды арендовались 815 324 раза.

Обратите внимание на несколько моментов. Во-первых, вы смогли выполнить запрос к набору данных, который уже есть в BigQuery. Все, что должен сделать

ответственный за проект, в котором размещены данные, — это предоставить вам¹ доступ к этому набору данных для «просмотра». Вам не нужно запускать кластер или входить в него — вы просто отправляете запрос сервису и получаете результаты. Сам запрос написан на SQL:2011, который поддерживает синтаксис, хорошо знакомый аналитикам данных. Мы продемонстрировали пример обработки гигабайтов данных, но вообще сервис легко масштабируется и способен агрегировать терабайты и петабайты данных. Такая масштабируемость возможна, потому что сервис распределяет обработку запросов между тысячами рабочих узлов практически мгновенно.

Работа с BigQuery

BigQuery — это хранилище данных с определенной степенью централизации. Запрос, продемонстрированный в предыдущем разделе, был применен к единственному набору данных. Однако преимущества BigQuery становятся еще более очевидными при объединении наборов данных из совершенно разных источников или при обращении к данным, хранящимся вне BigQuery.

Анализ наборов данных

Данные о прокате велосипедов поступают из Нью-Йорка. А можно ли объединить их с данными о погоде Национального управления океанических и атмосферных исследований США, чтобы узнать, как дождливая погода влияла на прокат велосипедов?²

```
-- Дождливая погода влияла на прокат велосипедов?
WITH bicycle_rentals AS (
  SELECT
    COUNT(starttime) as num_trips,
    EXTRACT(DATE from starttime) as trip_date
  FROM `bigquery-public-data.new_york_citibike.citibike_trips`
  GROUP BY trip_date
),

rainy_days AS
(
  SELECT
    date,
    (MAX(prcp) > 5) AS rainy
  FROM (
```

¹ Не конкретно вам. Это общедоступный набор данных, и владелец набора данных дал это разрешение всем аутентифицированным пользователям. Вы можете быть менее щедрыми в отношении своих данных и предоставлять к ним доступ только тем, кто находится в вашем домене или в вашей команде.

² Этот код можно загрузить из репозитория GitHub с примерами для книги.

```

SELECT
  wx.date AS date,
  IF (wx.element = 'PRCP', wx.value/10, NULL) AS prcp
FROM
  `bigquery-public-data.ghcn_d.ghcnd_2016` AS wx
WHERE
  wx.id = 'USW00094728'
)
GROUP BY
  date
)

SELECT
  ROUND(AVG(bk.num_trips)) AS num_trips,
  wx.rainy
FROM bicycle_rentals AS bk
JOIN rainy_days AS wx
ON wx.date = bk.trip_date
GROUP BY wx.rainy

```

Пока оставим в стороне синтаксис запроса. Обратите внимание только на строки, выделенные жирным шрифтом. Здесь мы соединяем набор данных о прокате велосипедов и набор данных о погоде, взятый из совершенно другого источника. Результаты запроса подтверждают предположение, что жители Нью-Йорка слабаки — в дождливые дни берут велосипеды в прокат на 20% реже.¹

```

Row num_trips rainy
1  39107.0  false
2  32052.0  true

```

Что означает возможность совместного использования наборов данных и отправления запросов в контексте предприятия? Разные подразделения вашей компании могут хранить свои наборы данных в BigQuery и легко обмениваться ими с другими подразделениями компании и даже с партнерскими организациями. Бессерверные технологии BigQuery помогают устранить разрозненность подразделений и оптимизировать сотрудничество.

ETL, EL и ELT

Традиционный подход к работе с хранилищами данных включает три этапа: извлечение, преобразование и загрузку (Extract, Transform, Load; ETL), когда исходные данные извлекаются из источника, преобразуются и затем загружаются в хранилище данных. В действительности BigQuery поддерживает собственный высокоэффективный формат колоночного хранения,² что делает методологию

¹ Имейте в виду, что оба автора живут в Сиэтле, где дождь идет 150 дней в году.

² Более подробно колоночный формат хранения описывается в разделе «История появления BigQuery» ниже.

ETL особенно привлекательной. Конвейер обработки данных, обычно реализованный на основе Apache Beam или Apache Spark, извлекает необходимые исходные данные (поточковых данных или пакетных файлов), преобразует извлеченные данные, подготавливая их к очистке или агрегированию, а затем загружает их в BigQuery, как показано на рис. 1.2.

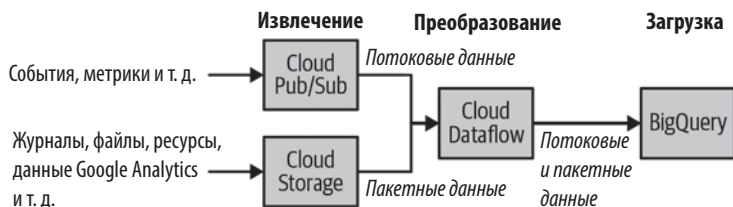


Рис. 1.2. Эталонная архитектура ETL в BigQuery использует конвейеры Apache Beam, выполняемые в Cloud Dataflow, и может обрабатывать как потоковые, так и пакетные данные с одним и тем же кодом

Хотя создание конвейера ETL в Apache Beam или Apache Spark весьма распространено, его можно создать исключительно в BigQuery. Так как BigQuery отделяет вычисления от хранилища, SQL-запросы BigQuery можно выполнять для файлов в формате CSV (а также JSON и Avro), которые хранятся в исходном виде в облачном хранилище Google Cloud Storage; эта возможность называется *федеративным запросом*. С помощью федеративных запросов можно извлекать данные, выполняя запросы SQL к Google Cloud Storage, преобразовывать данные внутри этих запросов SQL и сохранять результаты в собственных таблицах BigQuery.

Если преобразование не требуется, BigQuery может напрямую загружать стандартные форматы, такие как CSV, JSON или Avro, в свое хранилище — рабочий процесс EL (извлечение и загрузка). Такой подход, когда данные загружаются непосредственно в хранилище, используется потому, что хранение данных в собственном хранилище обеспечивает наибольшую эффективность запросов.

Мы настоятельно советуем реализовать процесс EL, если это возможно, и использовать процесс ETL, только если необходимы преобразования. По возможности выполните все необходимые преобразования в запросе SQL и сохраните весь конвейер ETL в BigQuery. Если преобразования трудно реализовать исключительно в SQL или если конвейер должен передавать данные в BigQuery по мере их поступления, создайте конвейер Apache Beam и выполняйте его бессерверным способом с использованием Cloud Dataflow. Другое преимущество реализации конвейеров ETL в Beam/Dataflow заключается в лучшем объединении таких конвейеров с системами непрерывной интеграции (CI) и модульного тестирования, потому что они содержат программный код.

Помимо процессов ETL и EL, BigQuery позволяет выполнять извлечение, загрузку и преобразование. Идея состоит в том, чтобы извлекать и загружать исходные

данные «как есть» и использовать представления BigQuery для преобразования данных во время работы. Процесс ELT особенно удобен, если схема исходных данных постоянно меняется. Например, вы можете реализовать определение необходимости исправления конкретной временной метки с учетом местного часового пояса. Процесс ELT может пригодиться для создания прототипов и позволяет организации начать извлекать информацию из данных, не принимая потенциально необратимых решений на ранних этапах.

Весь этот винегрет может сбивать с толку, поэтому мы подготовили краткую сводку в табл. 1.1.

Таблица 1.1. Краткое описание процессов, примеры архитектур и сценарии, в которых они могут использоваться

Процесс	Архитектура	Когда используется
EL	Извлечение данных из файлов в Google Cloud Storage. Загрузка в собственное хранилище BigQuery. Можно вызывать из Cloud Composer, Cloud Functions или запланированных запросов	Пакетная загрузка архивных данных. Периодическая загрузка файлов журналов по расписанию (например, один раз в день)
ETL	Извлечение данных из Pub/Sub, Google Cloud Storage, Cloud Spanner, Cloud SQL и т. д. Преобразование данных с использованием Cloud Dataflow. Имеет конвейер Dataflow для записи данных в BigQuery	Когда требуется проверить качество, выполнить преобразование или обогатить исходные данные перед загрузкой в BigQuery. Когда загрузка данных должна происходить непрерывно, то есть если сценарий требует потоковой передачи. Когда необходимы интеграция с системами непрерывной интеграции/непрерывной доставки (CI/CD) и модульное тестирование всех компонентов
ELT	Извлечение данных из файлов в Google Cloud Storage. Хранение в BigQuery данных в формате, близком к исходному. Преобразование данных в процессе работы с использованием представлений BigQuery	При использовании экспериментальных наборов данных, когда еще неясно, какие преобразования необходимо применить, чтобы сделать данные пригодными для использования. С любыми промышленными данными, когда преобразование можно выразить в SQL

Процессы в табл. 1.1 перечислены в том порядке, в каком мы рекомендуем их применять.

Эффективная аналитика

Преимущества использования хранилища напрямую вытекают из видов анализа, которые можно выполнять с хранящимися в нем данными. Основной способ

взаимодействия с BigQuery — выполнение запросов SQL, а поскольку BigQuery является движком SQL, вы можете использовать широкий спектр инструментов бизнес-аналитики (Business Intelligence, BI), таких как Tableau, Looker и Google Data Studio, для реализации разных видов анализа, визуальных диаграмм и отчетов на основе данных, хранящихся в BigQuery. Например, кликнув на «Explore in Data Studio» («Исследовать в Data Studio») в веб-интерфейсе BigQuery, можно быстро создать график помесечного изменения проката велосипедов, как показано на рис. 1.3.

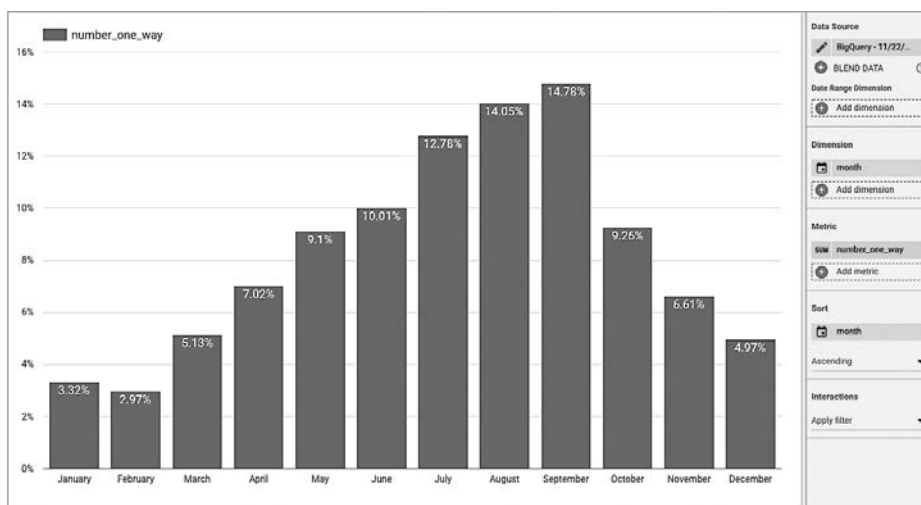


Рис. 1.3. Статистика в Data Studio проката велосипедов в зависимости от месяца; почти 15% проката в Нью-Йорке приходится на сентябрь

BigQuery предлагает полноценную поддержку SQL:2011, включая массивы и сложные соединения. В частности, поддержка массивов позволяет хранить в BigQuery иерархические данные (такие, как записи JSON) без преобразования вложенных и повторяющихся полей в плоские структуры. Помимо поддержки SQL:2011, BigQuery имеет несколько расширений, которые позволяют использовать этот сервис не только в качестве хранилища данных. Одним из таких расширений является поддержка широкого спектра геопространственных функций, позволяющих проверять в запросах местоположение, а также соединять таблицы на основе критериев расстояния или совпадения.¹ Поэтому BigQuery является полезным инструментом для проведения описательной аналитики.

Другое расширение BigQuery — поддержка в стандартном SQL создания моделей машинного обучения и выполнения пакетных прогнозов. Мы подробно рас-

¹ Например, для оценки расстояния, которое должен преодолеть клиент, чтобы купить продукт.

смотрим возможности машинного обучения в BigQuery в главе 9, но суть в том, что они позволяют обучать модели BigQuery и делать прогнозы, не экспортируя данные из BigQuery. Преимущества безопасности и локальности данных в этом случае огромны. То есть BigQuery — это хранилище данных, которое поддерживает не только описательную, но и прогнозную аналитику.

Идеология хранилища подразумевает возможность хранения данных разных типов. И действительно, BigQuery может хранить самые разные данные: числовые и текстовые данные как само собой разумеющееся, а также геопространственные и иерархические данные. BigQuery дает возможность хранить данные в упорядоченном виде, но это вовсе не обязательно, схемы могут быть очень разнообразными и сложными. Сочетание запросов с учетом пространственного местоположения, иерархических данных и машинного обучения делает BigQuery эффективным решением, не ограниченным рамками обычного хранилища и бизнес-аналитики.

BigQuery может принимать не только пакетные, но и потоковые данные. Данные можно передавать в BigQuery напрямую через REST API. Часто пользователи, которым требуется преобразовать данные — например, проводя вычисления с временным окном, — используют конвейеры Apache Beam, выполняемые сервисом Cloud Dataflow. И даже при передаче потоковых данных в BigQuery вы можете запросить их. Наличие общей инфраструктуры запросов для архивных (пакетных) и текущих (потоковых) данных открывает широкие возможности и упрощает многие процессы.

Простота управления

При разработке BigQuery немало внимания уделялось тому, чтобы обратить внимание пользователей на результаты анализа, а не на инфраструктуру. При вводе данных в BigQuery вам не нужно думать о разных типах хранилищ или о поиске золотой середины между скоростью работы и стоимостью услуги; хранилище полностью управляемое. На момент написания этой книги стоимость хранения автоматически снижается, если таблица не обновлялась в течение 90 дней.¹

Мы уже говорили, что в BigQuery не требуется настраивать индексы; ваши запросы SQL могут фильтровать результаты по любому столбцу, а BigQuery позаботится об их планировании и оптимизации. Более того, мы советуем писать запросы максимально понятно и разборчиво, а выбор стратегии оптимизации оставить для BigQuery. В этой книге мы будем говорить о настройке производительности, но в BigQuery она сводится в основном к логике и выбору соот-

¹ Все цены актуальны на момент написания этой книги, но вы обязательно должны сами свериться с соответствующими правилами и ценами (<https://cloud.google.com/bigquery/pricing>), так как они могут измениться.

ветствующих функций SQL. Вам не придется заниматься администрированием базы данных и решать такие задачи, как репликация, дефрагментация или восстановление после сбоев; обо всем этом позаботится BigQuery.

Запросы автоматически распределяются по тысячам машин и выполняются параллельно. Вам не нужно ничего предпринимать, чтобы запустить этот процесс. Машины уже подготовлены для обработки различных этапов заданий; вам не придется их настраивать.

Отсутствие необходимости поддерживать инфраструктуру уменьшает число проблем, связанных с безопасностью. Данные в BigQuery автоматически шифруются как при хранении, так и при передаче. BigQuery заботится о безопасности многопользовательских запросов и изоляции заданий. Вы сможете организовать общий доступ к своим наборам данных с помощью сервиса Google Cloud Identity and Access Management (IAM), а также применять к наборам данных (а также таблицам и представлениям в них) различные меры безопасности, в зависимости от того, нужна ли вам открытость, возможность аудита или конфиденциальность.

В других системах обеспечение надежности, гибкости, безопасности и производительности инфраструктуры часто отнимает много времени. Учитывая, что BigQuery практически избавляет от необходимости администрирования, организации, использующие BigQuery, считают, что это дает их аналитикам дополнительное время, чтобы сосредоточиться на получении информации из своих данных.

История появления BigQuery

В конце 2010 года директор подразделения Google в Сиэтле пригласил нескольких инженеров (один из которых является автором этой книги) и поставил перед ними задачу: создать коммерческую онлайн-платформу для хранения и анализа данных (data marketplace). Мы постарались найти лучший способ создания такой платформы. Основной проблемой были объемы данных, потому что мы не хотели предоставлять простую ссылку для скачивания. Платформа данных не сможет существовать, если людям придется загружать терабайты данных, чтобы начать с ней работу. Что бы вы сделали для того, чтобы не заставлять пользователей начинать работу с загрузки наборов данных?

Представляем вашему вниманию принцип Джима Грея (Jim Gray), первого исследователя в области баз данных ([https://en.wikipedia.org/wiki/Jim_Gray_\(computer_scientist\)](https://en.wikipedia.org/wiki/Jim_Gray_(computer_scientist))¹): «При работе с “большими данными”, — сказал Грей, — нужно помещать вычисления в данные, а не данные в вычисления». И уточнил:

¹ Статья в Википедии на русском языке: https://ru.wikipedia.org/wiki/Грей,_Джим. — *Примеч. пер.*

Другая ключевая проблема состоит в том, что по мере увеличения наборов данных становится все труднее получить их по FTP или выполнить поиск с помощью `grep`. Петабайт данных очень трудно передать по FTP! Поэтому рано или поздно вам понадобятся индексы и параллельный доступ к данным, и именно в этом вам помогут базы данных. Для анализа можно скачать данные, но точно так же можно направить запрос к данным. Вы можете переместить то или другое — запросы или данные. Но часто эффективнее перемещать запросы, а не данные.¹

Наша платформа данных не требовала бы от пользователей загружать наборы данных на компьютеры, если бы мы позволили им перенести свои вычисления в данные. Нам не пришлось бы давать ссылку для скачивания, потому что пользователи смогли бы работать со своими данными, не перемещая их.²

Мы, сотрудники Google, которым было поручено создать платформу данных, решили отложить этот проект и сосредоточиться на создании вычислительно-го механизма и системы хранения в облаке. Дав пользователям возможность что-то делать с данными, мы вернулись и добавили функции, характерные для коммерческой онлайн-платформы.

Какой язык пользователи должны использовать, чтобы писать вычисления для передачи в данные, хранящиеся в облаке? Мы выбрали SQL как обладающий тремя ключевыми характеристиками. Во-первых, SQL — универсальный язык, позволяющий не только разработчикам, но и широкому кругу людей задавать вопросы и решать задачи с использованием своих данных. Такая простота чрезвычайно важна. Во-вторых, SQL является «реляционно полным» языком, то есть на нем можно выразить любые вычисления с участием данных. Язык SQL не только прост и доступен. Он еще и очень функциональный. И наконец, что очень важно для выбора языка облачных вычислений, SQL не является «полным по Тьюрингу» в одном важном аспекте: код на этом языке всегда завершается.³ Как следствие, вычисления на SQL можно размещать, не беспокоясь, что кто-то напишет бесконечный цикл и монополизировать все вычислительные ресурсы вычислительного центра.

¹ Статья *Jim Gray on eScience: A Transformed Scientific Method* из сборника *The Fourth Paradigm: Data-Intensive Scientific Discovery*, под редакцией Тони Хей (Tony Hey), Стюарта Тенсли (Stewart Tansley) и Кристин Толле (Kristin Tolle), Microsoft, 2009, xiv. Доступна по ссылке <https://oreil.ly/M6zMN>.

² В настоящее время BigQuery позволяет экспортировать таблицы и результаты в Google Cloud Storage, поэтому мы все-таки создали ссылку для скачивания! Но BigQuery не просто ссылка — в большинстве случаев BigQuery предполагают обработку данных на месте.

³ В SQL есть ключевое слово `RECURSIVE`, но, как и многие движки SQL, BigQuery не поддерживает его и предлагает более эффективные способы работы с иерархическими данными, поддерживая массивы и вложение.

Затем нам нужно было выбрать движок SQL. В Google было несколько движков для работы с данными, в том числе очень популярные. Самый продвинутый движок назывался Dremel; он широко использовался в Google и был способен обрабатывать терабайты журналов за считанные секунды. Движок Dremel часто выбирали люди, занимавшиеся созданием конвейеров MapReduce и желавшие получить возможность задавать вопросы о своих данных.

В 2006 году инженер Андрей Губарев, который устал ждать завершения работ над MapReduce, создал движок Dremel. В научной литературе росла популярность колоночных хранилищ, и он быстро придумал формат колоночного хранения (рис. 1.4), который мог бы обрабатывать формат сериализации Protocol Buffers (Protobufs), повсеместно распространенный в Google.



Рис. 1.4. Колоночные хранилища могут сократить объем данных, считываемых запросами, которые обрабатывают все строки, но не все столбцы

Колоночные хранилища в целом хорошо подходят для аналитики, но в Google они оказались особенно полезными для анализа журналов, потому что многие команды работают с типами Protobuf, насчитывающими сотни тысяч столбцов. Если бы Андрей использовал типичное хранилище записей, пользователям пришлось бы читать файлы построчно и принимать огромное количество данных в полях, которые они все равно собирались отбросить. Сохраняя данные столбец за столбцом, Андрей сделал так, что если пользователю требовалось всего несколько тысяч полей из журнала Protobufs, он мог прочитать только часть общего объема данных. Это было одной из причин, почему Dremel справлялся с обработкой терабайтов журналов за считанные секунды.

Другая причина, по которой Dremel мог обработать данные так быстро, заключалась в том, что его механизм запросов использовал распределенные вычисления. Движок был способен распределить вычисления между тысячами рабочих процессов, структурируя вычисления как дерево, с фильтрами в листьях и агрегированием ближе к корню.

К 2010 году Google каждый день сканировала с помощью Dremel уже петабайты данных, и многие сотрудники использовали их в той или иной форме. Это был идеальный инструмент для нашей команды, начинавшей разработку коммерческой платформы данных.

Когда команда доработала движок Dremel, добавила к нему систему хранения, реализовала автоматическую настройку и предложила сторонним пользовате-

лям, стало понятно, что разработка облачной версии Dremel, возможно, даже более интересная задача, чем та, которая была поставлена изначально. Команда переименовала себя в «BigQuery», по аналогии с «Bigtable» — базой данных NoSQL, также созданной в Google.

В Google движок Dremel используется для анализа файлов, которые хранятся в Google Colossus — хранилище файлов. Команда BigQuery добавила систему хранения с абстракцией таблицы. Эта система хранения играла ключевую роль, делая платформу BigQuery простой и быстрой в использовании, потому что позволяла использовать такие важные особенности, как свойства ACID (Atomicity, Consistency, Isolation, Durability — атомарность, согласованность, изолированность, долговечность) транзакций, а также автоматическую оптимизацию, благодаря чему пользователям не требовалось управлять файлами.

Первоначально служба BigQuery сохраняла свою связь с Dremel и была ориентирована на сканирование журналов. Однако по мере увеличения числа клиентов, желавших создавать хранилища данных и выполнять все более сложные запросы, в BigQuery была добавлена улучшенная поддержка соединений и расширенные возможности SQL, включая аналитические функции. В 2016 году Google добавила поддержку стандартного SQL в BigQuery, что позволило пользователям выполнять запросы, используя стандартный язык SQL вместо неудобного диалекта «DremelSQL», использовавшегося вначале.

На начальном этапе BigQuery не была хранилищем данных, но со временем стала именно им. В этой трансформации есть свои преимущества и недостатки. Преимуществом является то, что BigQuery разрабатывалась для решения задач с данными, даже при том, что эта сторона платформы плохо вписывается в модели хранилищ данных. То есть BigQuery — это больше чем простое хранилище данных. С другой стороны, до недавнего времени в BigQuery отсутствовали некоторые особенности, свойственные хранилищам данных, такие как язык определения данных (Data Definition Language, DDL; например, оператор CREATE) и язык манипулирования данными (Data Manipulation Language, DML; например, оператор INSERT). Тем не менее BigQuery фокусируется на развитии в двух направлениях: добавление отличительных особенностей, которые может предложить Google, и превращение в отличное облачное хранилище данных.

Что позволило создать BigQuery?

С точки зрения архитектуры BigQuery принципиально отличается от хранилищ данных, таких как Teradata или Vertica, а также от облачных хранилищ данных, таких как Redshift и Microsoft Azure Data Warehouse. BigQuery — первое хранилище данных с поддержкой горизонтального масштабирования, поэтому единственным ограничением скорости и масштабирования является количество оборудования в дата-центре.

В этом разделе описываются некоторые компоненты, делающие BigQuery успешным и уникальным проектом.

Отделение вычислений от хранилища

Во многих хранилищах данных вычисления и хранение осуществляются на одном и том же физическом оборудовании. Это значит, что для увеличения объема хранилища может потребоваться увеличить вычислительную мощность. Или для увеличения вычислительной мощности вам также понадобится увеличить емкость хранилища.

Не было бы никаких проблем, будь у всех одинаковые потребности в данных; можно было бы найти золотую середину между вычислительной мощностью и объемом хранилища. Но на практике то один, то другой фактор оказывается ограничивающим. Некоторые хранилища имеют ограниченную вычислительную мощность и могут работать медленно в часы пиковых нагрузок. Другие хранилища имеют ограниченную емкость, из-за чего администраторы вынуждены решать, какие данные выбрасывать.

Отделив вычисления от хранилища, как это сделано в BigQuery, вам никогда не придется отбрасывать ненужные данные. Вы, может, и не придадите этому особого значения, тем не менее, доступ к данным с максимальной точностью невероятно эффективен. Вы можете решить выполнить вычисления как-то иначе, а значит, вам придется вернуться к исходным данным и запросить их. Но у вас не получится сделать это, если вы удалите исходные данные из-за нехватки места. Возможно, вы захотите разобраться, почему некоторые агрегированные значения странно себя ведут. Но вы не сможете это узнать, если удалили данные, которые участвовали в агрегировании.

Неменьший эффект дает масштабирование вычислений. Ресурсы BigQuery измеряются в слотах. Один слот соответствует примерно половине ядра процессора (мы подробно рассмотрим слоты в главе 6). BigQuery использует слоты как абстракцию количества доступных физических вычислительных ресурсов. Запросы выполняются слишком медленно? Просто добавьте слоты. Больше людей хотят создавать отчеты? Добавьте больше слотов. Хотите сократить свои расходы? Уменьшите число слотов.

Поскольку BigQuery — это многопользовательская система, управляющая большими пулами аппаратных ресурсов, она позволяет распределять слоты и на уровне запросов, и на уровне пользователей. Вы можете зарезервировать оборудование для своего проекта или организации или выполнять запросы в общем пуле. Благодаря такому механизму совместного использования ресурсов, BigQuery может выделять очень большие вычислительные мощности для обработки ваших запросов. Если вам потребуется больше вычислительной мощности, чем доступно в общем пуле, вы сможете приобрести дополнительные ресурсы с помощью интерфейса бронирования BigQuery Reservation API.

Некоторые клиенты BigQuery резервируют десятки тысяч слотов, то есть когда они выполняют запросы по одному, эти запросы могут обрабатываться десятками тысяч процессорных ядер одновременно. С некоторыми допущениями о количестве тактов процессора, необходимых для обработки одной записи, довольно легко подсчитать, что, имея подобные мощности, можно обрабатывать миллиарды или даже триллионы записей в секунду.

Некоторые клиенты BigQuery имеют петабайты данных, но ежедневно используют лишь относительно небольшую их часть. Другие хранят всего несколько гигабайтов данных, но выполняют сложные запросы, используя тысячи процессоров. Не существует универсального решения. К счастью, разделение вычислений и хранилищ позволяет BigQuery удовлетворить широкий спектр потребностей клиентов.

Хранилище и сетевая инфраструктура

В отличие от других облачных хранилищ данных, запросы в BigQuery обрабатывают данные, находящиеся в основном на вращающихся дисках в распределенной файловой системе. Большинство конкурирующих систем вынуждено кешировать данные на вычислительных узлах, чтобы получить хорошую производительность. BigQuery, напротив, использует две уникальные системы, разработанные в Google: файловую систему Colossus (https://cloud.google.com/files/storage_architecture_and_challenges.pdf) и сеть Jupiter (<https://cloudplatform.googleblog.com/2015/06/A-Look-Inside-Google-Data-Center-Networks.html>), обеспечивающие быстрый доступ к данным, независимо от их физического местоположения в вычислительном кластере.

Структура сети Google Jupiter основана на конфигурации, в которой менее производительные (и, следовательно, более дешевые) коммутаторы размещены так, чтобы обеспечить пропускную способность, для которой в противном случае потребовался бы гораздо более производительный коммутатор. Эта топология коммутаторов, наряду с централизованным программным стеком и настраиваемым аппаратным и программным обеспечением, обеспечивает пропускную способность в один петабит в пределах вычислительного центра. Это эквивалентно 100 000 серверов, обменивающихся данными со скоростью 10 Гбит/с, и это означает, что BigQuery может работать без необходимости объединять вычисления и хранилище. Если машины, на которых размещены диски, находятся в другом конце вычислительного центра, в отдалении от машин, на которых выполняются вычисления, они будут работать так же быстро, как если бы они размещались в одной стойке.

Быстрая сетевая структура дает два преимущества: она позволяет быстро считывать с диска и перемещать их между этапами обработки. Как уже было упомянуто, разделение вычислений и хранилищ в BigQuery позволяет любой машине в вычислительном центре получать данные с любого диска хранилища. Однако для этого

требуется передавать входные данные, необходимые для выполнения запросов, с очень высокой скоростью. Подробнее это описано в главе 6, а пока достаточно отметить, что для выполнения сложных распределенных запросов обычно требуется перемещать большие объемы данных между компьютерами на промежуточных этапах. Без быстрой сети, соединяющей вычислительные узлы, перемещение данных станет узким местом и может значительно замедлить обработку запросов.

Сетевая инфраструктура обеспечивает не только высокую скорость, но и высокую производительность. Вычислительные центры Google связаны магистральной сетью под названием B4 (<https://www.usenix.org/conference/atc15/technical-session/presentation/mandal>), которая управляется программным обеспечением, распределяющим пропускную способность между пользователями и обеспечивающим высокое качество обслуживания приоритетных операций. Это очень важно для высокопроизводительной параллельной обработки запросов.

Однако быстрой сети недостаточно, если дисковая подсистема работает медленно или недостаточно хорошо масштабируется. Для поддержки интерактивных запросов считывание данных с дисков должно быть достаточно быстрым, чтобы ими можно было наполнить доступную пропускную способность сети. В Google используется распределенная файловая система Colossus, способная координировать сотни тысяч дисков, постоянно выполнять перебалансировку старых, «холодных» данных и равномерно распределять новые данные по диску.¹ Это означает, что эффективная пропускная способность составляет десятки терабайт в секунду. Сочетая эту эффективную пропускную способность с эффективными форматами данных и хранилищем, BigQuery имеет возможность запрашивать таблицы с петабайтами данных за считанные минуты.

Управляемое хранилище

Система хранения в BigQuery основана на идее использования абстракции таблицы вместо абстракции файла при работе со структурированным хранилищем. Некоторые облачные системы обработки данных с открытым исходным кодом предлагают пользователям формат файла, что позволяет управлять размерами файлов и обеспечивать согласованность схемы. Хранилища позволяют создавать файлы соответствующего размера для хранения статических данных, но, как известно, поддерживать оптимальный размер файлов для данных, которые со временем меняются, очень сложно. Точно так же трудно поддерживать согласованность схемы при наличии большого количества файлов со схемами, предусматривающими самоописание (например, Avro или Parquet), — обычно каждое обновление программного обеспечения в системах, производящих эти файлы, приводит к изменениям схемы. BigQuery гарантирует согласованность схем для всех данных, хранящихся в таблицах, и обеспечивает надежный переход

¹ Узнать больше о Colossus можно по ссылкам <http://www.pds.org/pds-discs17/slides/PDSW-DISCS-Google-Keynote.pdf> и <https://www.wired.com/2012/07/google-colossus/>.

архивных данных. Абстрагируя форматы данных и размеры файлов, BigQuery может обеспечить непрерывную работу и быструю обработку запросов.

Еще одно преимущество BigQuery в управлении собственным хранилищем: возможность увеличения скорости работы без усилий со стороны конечного пользователя. Например, улучшения в форматах хранения могут автоматически применяться к пользовательским данным. Аналогично, улучшения в инфраструктуре хранения немедленно отражаются на работе системы. Поскольку хранилищем управляет только BigQuery, пользователям не нужно беспокоиться о резервном копировании или репликации. Все, от обновлений и репликации до резервного копирования и восстановления, выполняется системой управления хранилищем автоматически.

Одним из ключевых преимуществ работы со структурированным хранилищем на уровне абстракции таблицы (а не файла) и простого управления хранением этих таблиц является возможность для BigQuery поддерживать соответствующие функции, например DML. Вы можете отправить запрос, обновляющий или удаляющий строки в таблице, и положиться на BigQuery в выборе наиболее подходящего способа изменения хранилища для представления этой информации. Операции в BigQuery поддерживают свойства ACID; то есть изменения, произведенные в ходе выполнения запросов, либо подтверждаются полностью, либо отменяются. Будьте уверены, что ваши запросы никогда не будут пересекаться с промежуточным состоянием другого запроса, а запросы, запущенные после завершения другого запроса, никогда не пересекутся со старыми данными. У вас есть возможность точно настроить хранилище с помощью директив, управляющих хранением данных, но они работают на уровне абстракции таблиц, а не файлов. Например, вы можете управлять сегментированием и кластеризацией таблиц (мы подробно рассмотрим эти возможности в главе 7) и тем самым улучшить производительность и/или уменьшить стоимость запросов к этим таблицам.

Управляемое хранилище строго типизировано, то есть данные проверяются на входе в систему. Поскольку управление хранилищем осуществляет платформа BigQuery и пользователи могут взаимодействовать с этим хранилищем только через ее API, она предусматривает, что базовые данные не изменятся за ее пределами. То есть BigQuery может гарантировать, что при чтении любых данных, находящихся в управляемом хранилище, не возникнет никаких ошибок проверки. Эта гарантия также подразумевает надежность схемы, которая может пригодиться для выяснения, как правильно запрашивать таблицы. Помимо улучшенной производительности запросов, наличие надежной схемы помогает понять, какие данные у вас есть, потому что схема в BigQuery содержит не только информацию о типе, но также аннотации и описания таблиц о том, как можно использовать поля.

Одним из недостатков управляемого хранилища является сложность получения прямого доступа к данным и их обработки с использованием других платформ. Например, если бы данные были доступны на уровне абстракции файлов, вы могли бы запустить задание Nadoop на наборе данных из BigQuery. BigQuery

решает эту проблему, предоставляя структурированный параллельный API для чтения данных. Этот API позволяет читать на полной скорости из заданий Spark или Nadoop, а также предлагает дополнительные возможности, такие как создание проекций, фильтрация и динамическая перебалансировка.

Интеграция с платформой Google Cloud

Платформа Google Cloud следует принципу «разделения ответственности», согласно которому небольшое количество высококачественных и узкоспециализированных продуктов тесно интегрируются друг с другом. Поэтому, сравнивая BigQuery с другими продуктами баз данных, важно учитывать всю облачную платформу Google (Google Cloud Platform, GCP).

Существует множество разных продуктов GCP, которые пополняют список достоинств BigQuery или помогают понять, как пользоваться BigQuery. Мы подробно поговорим о многих из этих продуктов позже, но не забывайте об общем принципе разделения ответственности:

- StackDriver, инструмент мониторинга и аудита журналов, помогает понять, как лучше использовать BigQuery в вашей организации.
- Cloud Dataproc обеспечивает возможность чтения, обработки и записи в таблицы BigQuery с помощью программ Apache Spark.
- Федеративные запросы позволяют BigQuery извлекать данные, хранящиеся в Google Cloud Storage, Cloud SQL (реляционная база данных), Bigtable (база данных NoSQL), Spanner (распределенная база данных) или Google Drive (электронные таблицы).
- Google Cloud Data Loss Prevention API (<https://cloud.google.com/dlp>) помогает управлять конфиденциальными данными и дает возможность редактировать или маскировать личную информацию (Personally Identifiable Information, PII) в таблицах.
- Другие API машинного обучения добавляют новые возможности в анализ данных, хранящихся в BigQuery; например, API Cloud Natural Language может идентифицировать людей, места, эмоциональную окраску и многое другое в произвольном тексте (например, в отзывах клиентов), размещенном в каком-либо столбце таблицы.
- AutoML Tables и AutoML Text могут создавать высокоэффективные модели машинного обучения на основе данных, хранящихся в таблицах BigQuery.
- Cloud Catalog позволяет обнаруживать данные, хранящиеся в вашей организации.
- Cloud Pub/Sub можно использовать для передачи потоковых данных, а Cloud Dataflow — для преобразования и загрузки их в BigQuery. Также Cloud Dataflow можно использовать для выполнения потоковых запросов.

И конечно же, потоковые данные внутри самой платформы BigQuery можно запрашивать интерактивно.¹

- Data Studio предоставляет диаграммы и информационные панели (дашборды), отображающие данные из BigQuery. Сторонние инструменты, такие как Tableau и Looker, тоже поддерживают возможность получения данных из BigQuery.
- Cloud AI Platform дает возможность обучать сложные программы машинного обучения на данных, хранящихся в BigQuery.
- Cloud Scheduler и Cloud Functions позволяют планировать или запускать запросы BigQuery как часть более крупных рабочих процессов.
- Cloud Composer позволяет координировать задания BigQuery с задачами, которые необходимо выполнить в Cloud Dataflow или в других фреймворках обработки данных, в Google Cloud или локальных гибридных облаках.

В совокупности BigQuery и экосистема GCP обладают возможностями, охватывающими некоторые другие продукты баз данных от сторонних поставщиков облачных решений; их можно использовать как хранилище аналитики, а также как систему ETL, озеро данных (запросы к файлам) или источник BI. Далее мы более подробно расскажем, как можно использовать все аспекты платформы BigQuery.

Безопасность и соответствие требованиям

Интеграция с GCP не ограничивается простыми взаимодействиями с другими продуктами. Сквозные функции, предлагаемые платформой, обеспечивают постоянную безопасность и соответствие требованиям.

Самое быстрое оборудование и самое современное программное обеспечение бесполезны, если вы не сможете доверять им свои данные. Модель безопасности BigQuery тесно интегрирована с остальной частью GCP, что позволяет получить целостное представление о безопасности ваших данных. Для назначения конкретных разрешений отдельным пользователям или их группам BigQuery использует систему контроля доступа Google IAM. Также BigQuery тесно связана с управлением виртуальным частным облаком Google (Virtual Private Cloud, VPC), способным защитить вас от внешних попыток доступа к данным вашей организации или их передачи третьей стороне. Элементы управления IAM и VPC проектировались для работы со всеми продуктами Google Cloud, поэтому вам не придется беспокоиться о том, что какие-то продукты могут создать слабое звено в системе безопасности.

¹ Разделение ответственности заключается в том, что Cloud Dataflow лучше подходит для непрерывной рутинной обработки, а BigQuery — для специальной интерактивной обработки. Cloud Dataflow и BigQuery способны обрабатывать как пакетные, так и потоковые данные, и Cloud Dataflow позволяет выполнять запросы SQL.

BigQuery доступна во всех регионах, где есть Google Cloud, благодаря чему вы можете обрабатывать данные в любом выбранном месте. На момент написания этой книги в Google Cloud имелось более двух десятков вычислительных центров по всему миру, и компания продолжает открывать новые. Если у вас есть свои причины хранить данные в Австралии или Германии, вы легко сможете это сделать. Просто создайте свой набор данных с кодом региона Австралии или Германии, и все ваши запросы к данным будут выполняться в этом регионе.

Некоторые организации предъявляют еще более строгие требования к размещению данных, не ограничиваясь требованием к выбору региона, где должны храниться и обрабатываться данные. В частности, многие хотят иметь гарантии, что их данные не будут скопированы или иным образом не покинут определенную географическую область. GCP имеет элементы управления географической областью, которые применяются ко всем продуктам; вы можете создать систему «управление службами VPC» (VPC service controls), которая запрещает перемещение данных за пределы выбранной области. Если у вас настроены подобные системы, пользователи не смогут копировать или экспортировать данные в другие сегменты Google Cloud Storage, находящиеся в других регионах.

Выводы

BigQuery — это масштабируемое хранилище данных, обеспечивающее быстрые бессерверные вычисления больших наборов данных с помощью SQL-запросов. Пользователи ценят масштаб и скорость BigQuery, но руководители компаний часто более высоко оценивают возможность перехода на более качественный уровень, который дает возможность выполнять специальные запросы с помощью бессерверных вычислений и позволяет принимать решения на основе данных во всех подразделениях компании.

Для загрузки данных в BigQuery можно использовать конвейер EL (часто применяется для периодической загрузки файлов журналов), конвейер ETL (когда необходимо обогащение данных или контроль качества) или конвейер ELT (для исследовательской работы).

Платформа BigQuery предназначена для аналитической обработки данных в реальном времени (OnLine Analytical Processing, OLAP) и предоставляет полноценную поддержку SQL:2011. Высокая скорость BigQuery достигается за счет инновационных инженерных решений, таких как использование колоночного хранилища, поддержка вложенных и повторяющихся полей, а также отделение вычислений от хранения, о котором Google продолжает публиковать статьи. BigQuery является частью экосистемы GCP инструментов анализа больших данных и тесно интегрируется как с элементами инфраструктуры (обеспечивающими, например, безопасность, мониторинг и ведение журналов), так и с элементами обработки данных и машинного обучения (такими, как потоковая передача, Cloud DLP и AutoML).

Основы запросов

Платформа BigQuery — это прежде всего хранилище данных, то есть она обеспечивает постоянное хранение структурированных и полуструктурированных данных (например, объектов JSON). В этом постоянном хранилище поддерживаются четыре основные CRUD-операции:

Create (создание)

Добавляет новые записи. Осуществляется с помощью операций загрузки, SQL-оператора `INSERT` и интерфейса потоковой вставки. С помощью SQL-операторов, которые являются частью поддерживаемого в BigQuery языка определения данных (Data Definition Language, DDL), также можно создавать объекты баз данных, такие как таблицы, представления и модели машинного обучения. Ниже мы рассмотрим примеры каждой из перечисленных возможностей.

Read (чтение)

Извлекает записи. Осуществляется с помощью SQL-оператора `SELECT` и массового считывания API.

Update (изменение)

Изменяет существующие записи. Осуществляется с помощью SQL-операторов `UPDATE` и `MERGE`, которые являются частью поддерживаемого в BigQuery языка манипулирования данными (Data Manipulation Language, DML). Обратите внимание на то, что, как уже было отмечено в главе 1, платформа BigQuery является аналитическим инструментом и не предназначена для частых обновлений.

Delete (удаление)

Удаляет существующие записи. Осуществляется с помощью SQL-оператора `DELETE`, который также является частью поддерживаемого в BigQuery языка DML.

BigQuery — это инструмент анализа данных, и большинство запросов, которые вам предстоит написать, будут вышеупомянутыми операциями чтения. Чтение и анализ данных осуществляются с помощью оператора `SELECT`, о котором пойдет речь в этой главе. Операции создания, изменения и удаления данных мы рассмотрим в следующих главах.

ЧТО ТАКОЕ УСТАРЕВШИЙ SQL?

Долгое время BigQuery поддерживала только ограниченное подмножество языка SQL с некоторыми расширениями Google. Все потому, что BigQuery была основана на движке SQL (под названием Dremel), использовавшемся внутри компании Google. Изначально этот движок создавался для обработки журналов, хранящихся в формате Protocol Buffers (Protobufs).¹ Поскольку Dremel создавался не как универсальный движок SQL, он мог использовать диалект SQL (сейчас его называют *устаревшим SQL*), хорошо подходящий для работы с форматом Protobufs, использующимся для хранения иерархических структур. Например, устаревший SQL различает записи (иерархическая структура, включающая все сообщения в журнале) и строки (срез структуры).² Соответственно инструкция `COUNT(*)` в Dremel подсчитывает количество значений, не равных `NULL` в большинстве повторяющихся полей. Несмотря на то что это значительно облегчало написание запросов определенных типов, к диалекту Dremel нужно было привыкнуть, потому что это был нестандартный SQL.

В этой книге мы сосредоточимся на стандартном SQL. Пользовательский интерфейс BigQuery в облачной консоли Google Cloud Platform (GCP) по умолчанию использует стандартный SQL, а новые возможности не переносятся в устаревший SQL. Тем не менее некоторые инструменты и пользовательские интерфейсы все еще используют устаревший SQL. Если вы столкнетесь с подобным инструментом, просто добавьте в запрос первую строку `#standardsql`, как показано в следующем примере:

```
#standardsql
SELECT DISTINCT gender
FROM `bigquery-public-data`.new_york_citibike.citibike_trips
```

Если BigQuery обнаружит в запросе первую строку `#standardsql`, механизм обработки запросов будет интерпретировать последующий код как стандартный SQL, даже если сам клиент ничего не знает о стандартном SQL.

¹ Это формат данных, пользующийся большой популярностью в Google, потому что он обеспечивает эффективное хранение, не зависящее от языка программирования. Впоследствии структура формата была открыта для всеобщего использования; см. <https://developers.google.com/protocol-buffers/>.

² Более подробную информацию о Dremel можно найти на странице: <https://ai.google/research/pubs/pub36632>.

Простые запросы

BigQuery поддерживает диалект SQL, совместимый с SQL:2011 (<https://www.iso.org/standard/53681.html>). Если спецификация неоднозначна или имеет пробелы, BigQuery следует требованиям, принятым в существующих движках SQL. Есть также области, для которых спецификации отсутствуют, например машинное обучение; в таких случаях BigQuery определяет свой собственный синтаксис и семантику.

Извлечение записей с помощью SELECT

Оператор SELECT позволяет извлекать из таблицы значения указанных столбцов. Например, рассмотрим набор данных по прокату велосипедов в Нью-Йорке (https://bigquery.cloud.google.com/table/bigquery-public-data:new_york.citibike_trips) — он содержит несколько столбцов, касающихся проката велосипедов, включая продолжительность поездки и пол человека, арендовавшего велосипед. Вот как с помощью оператора SELECT можно извлечь значения столбцов (строки, начинающиеся с двух дефисов или с #, являются комментариями):

```
-- простая выборка
SELECT
  gender, tripduration
FROM
  `bigquery-public-data`.new_york_citibike.citibike_trips
LIMIT 5
```

Результаты должны выглядеть примерно так:

Row	gender	tripduration
1	male	371
2	male	1330
3	male	830
4	male	555
5	male	328

Полученный в итоге набор имеет два столбца (`gender` и `tripduration`), следующих в том же порядке, что и в операторе SELECT, и пять записей, потому что мы ограничили их количество в последней строке запроса. BigQuery распределяет задачу выборки записей между несколькими рабочими процессами, каждый из которых может читать свой сегмент (или часть) исходного набора данных, поэтому, выполнив предыдущий запрос еще раз, можно получить другой набор из пяти записей.

Обратите внимание, что оператор `LIMIT` ограничивает только объем отображаемых данных, а не объем данных, которые должен обработать движок запросов. Обычно клиенты BigQuery платят за объем данных, обработанных их запросами, и, как правило, это означает, что чем больше столбцов прочтает ваш запрос, тем больше вы заплатите. Количество обработанных записей обычно определяется общим размером таблицы, хотя есть способы оптимизировать и этот показатель (о чем мы поговорим в главе 7). Вопросы эффективности и ценообразования мы рассмотрим в следующих главах.

Значения извлекаются из следующего источника:

```
bigquery-public-data.new_york_citibike.citibike_trips
```

Здесь `bigquery-public-data` — это идентификатор проекта, `new_york_citibike` — набор данных, а `citibike_trips` — таблица. Идентификатор проекта определяет принадлежность постоянного хранилища, связанного с набором данных и его таблицами. Владелец `bigquery-public-data` оплачивает расходы на хранение набора данных `new_york`. Стоимость запроса оплачивается проектом, в рамках которого был запущен запрос. Выполнив предыдущий запрос, вы должны будете оплатить его стоимость. Наборы данных обеспечивают управление идентификацией и доступом (Identity and Access Management, IAM). Человек,¹ создавший набор данных `new_york_citibike` в BigQuery, сделал его общедоступным, поэтому мы смогли перечислить таблицы (https://bigquery.cloud.google.com/dataset/bigquery-public-data:new_york) в наборе данных и выполнить запрос к одной из них. Таблица `citibike_trips` содержит все поездки на арендованных велосипедах. Идентификатор проекта, название набора данных и имя таблицы разделены точками. Обратные апострофы в данном случае используются в качестве экранирующего символа из-за присутствия дефиса (`-`) в идентификаторе проекта (`bigquery-public-data`) — без них дефисы интерпретировались бы как операторы вычитания. Большинство разработчиков просто заключают всю строку в обратные апострофы, например:

```
-- простая выборка
SELECT
  gender, tripduration
FROM
  `bigquery-public-data.new_york_citibike.citibike_trips`
LIMIT 5
```

Такой способ проще, но при этом теряется возможность использовать имя таблицы (`citibike_trips`) в качестве псевдонима. Поэтому лучше выработать привычку заключать в обратные апострофы только имя проекта и не использовать дефисы при определении имен наборов данных и таблиц.

¹ Этим «человеком» в данном случае является один из членов команды открытых наборов данных Google Cloud Platform. Загляните в проект Google Cloud Public Datasets (<https://cloud.google.com/public-datasets/>), чтобы узнать о других общедоступных наборах данных.

Долгое время мы рекомендовали хранить таблицы в BigQuery в денормализованной форме (то есть помещать все необходимые данные в одну таблицу, чтобы избежать необходимости соединения нескольких таблиц). Однако, благодаря усовершенствованию сервиса, эта рекомендация потеряла свою актуальность. Теперь хорошей производительности можно добиться даже с использованием звездообразной схемы.

В табл. 2.1 перечислены три ключевых компонента имени ``bigquery-public-data`.`new_york_citibike.citibike_trips``.

Таблица 2.1. Основные объекты BigQuery и их описание¹

Объект BigQuery	Имя	Описание
Проект	<code>bigquery-public-data</code>	Владелец хранилища, связанного с набором данных и его таблицами. Проект также регулирует использование всех других продуктов GCP
Набор данных	<code>new_york_citibike</code>	Наборы данных — это контейнеры верхнего уровня, которые используются для организации и управления доступом к таблицам и представлениям. Пользователь может иметь несколько наборов данных
Таблица/представление	<code>citibike_trips</code>	Таблица или представление должны принадлежать набору данных, поэтому перед загрузкой данных в BigQuery необходимо создать хотя бы один набор данных ¹

Различие между этими тремя компонентами будет иметь решающее значение позже, когда речь пойдет о географическом местоположении, доступе к данным и их совместном использовании.

Создание псевдонимов столбцов с помощью AS

По умолчанию имена столбцов в наборе результатов совпадают с именами столбцов в таблице, откуда извлекаются данные. Однако с помощью AS именам столбцов можно присвоить свои псевдонимы:

```
-- Определение псевдонимов для имен столбцов
SELECT
  gender, tripduration AS rental_duration
FROM
  `bigquery-public-data`.`new_york_citibike.citibike_trips`
LIMIT 5
```

Этот запрос вернул следующие результаты (у вас конкретные значения могут отличаться):

¹ См. <https://cloud.google.com/bigquery/docs/datasets-intro>.

Row	gender	rental_duration
1	male	432
2	female	1186
3	male	799
4	female	238
5	male	668

Псевдонимы могут быть полезны при преобразовании данных. Например, без псевдонима следующий оператор:

```
SELECT
  gender, tripduration/60
FROM
  `bigquery-public-data`.new_york_citibike.citibike_trips
LIMIT 5
```

присвоит второму столбцу в наборе результатов имя автоматически:

Row	gender	f0_
1	male	6.183333333333334
2	male	22.166666666666668
3	male	13.833333333333334
4	male	9.25
5	male	5.466666666666667

Вы можете дать второму столбцу осмысленное имя, добавив псевдоним в запрос:

```
SELECT
  gender, tripduration/60 AS duration_minutes
FROM
  `bigquery-public-data`.new_york_citibike.citibike_trips
LIMIT 5
```

Этот запрос даст примерно такие результаты:

Row	gender	duration_minutes
1	male	6.183333333333334
2	male	22.166666666666668
3	male	13.833333333333334
4	male	9.25
5	male	5.466666666666667

Фильтрация с WHERE

Чтобы найти случаи, когда клиент брал велосипед в прокат меньше чем на 10 минут, можно отфильтровать результаты, возвращаемые оператором `SELECT`, добавив предложение `WHERE`:

```
SELECT
    gender, tripduration
FROM
    `bigquery-public-data`.new_york_citibike.citibike_trips
WHERE tripduration < 600
LIMIT 5
```

Как и ожидалось, теперь в набор результатов попали только записи, соответствующие поездкам, длившимся меньше 600 секунд:

Row	gender	tripduration
1	male	178
2	male	518
3	male	376
4	male	326
5	male	516

Предложение `WHERE` может содержать логические выражения. Например, вот как можно найти поездки, совершенные женщинами и длившиеся 5–10 минут:

```
SELECT
    gender, tripduration
FROM
    `bigquery-public-data`.new_york_citibike.citibike_trips
WHERE tripduration >= 300 AND tripduration < 600 AND gender = 'female'
LIMIT 5
```

Также можно использовать ключевые слова `OR` и `NOT`. Например, следующее предложение `WHERE` поможет отобрать клиентов не женского пола (то есть мужчин и клиентов, не указавших свой пол):

```
WHERE tripduration < 600 AND NOT gender = 'female'
```

Для управления порядком вычислений можно использовать круглые скобки. Найти женщин, совершивших короткие поездки, а также всех мужчин можно так:

```
WHERE (tripduration < 600 AND gender = 'female') OR gender = 'male'
```


Предложение **WHERE** работает со столбцами в таблице, указанной в предложении **FROM**; то есть внутри **WHERE** нельзя ссылаться на псевдонимы, указанные в **SELECT**. Иначе говоря, следующий запрос нельзя использовать для выбора только поездок, длившихся менее 10 минут:

```
SELECT
  gender, tripduration/60 AS minutes
FROM
  `bigquery-public-data`.new_york_citibike.citibike_trips
WHERE minutes < 10 -- ВНУТРИ WHERE НЕЛЬЗЯ ССЫЛАТЬСЯ НА ПСЕВДОНИМЫ
LIMIT 5
```

Вместо этого нужно повторить преобразование внутри предложения **WHERE** (более удачные альтернативы мы рассмотрим позже):

```
SELECT
  gender, tripduration / 60 AS minutes
FROM
  `bigquery-public-data`.new_york_citibike.citibike_trips
WHERE (tripduration / 60) < 10
LIMIT 5
```

SELECT *, EXCEPT, REPLACE

Из соображений экономии средств и производительности (которые мы подробно рассмотрим в главе 7) лучше выбирать только нужные столбцы. Однако если вы захотите выбрать все столбцы в таблице, то можете использовать оператор **SELECT ***:

```
SELECT
  *
FROM
  `bigquery-public-data`.new_york_citibike.citibike_stations
WHERE name LIKE '%Riverside%'
```

Здесь для поиска пунктов проката, которые в своем названии имеют слово **Riverside**, в предложении **WHERE** используется оператор **LIKE**.

Выбрать столбцы, исключив некоторые из них, можно с помощью оператора **SELECT EXCEPT**:

```
SELECT
  * EXCEPT(short_name, last_reported)
FROM
  `bigquery-public-data`.new_york_citibike.citibike_stations
WHERE name LIKE '%Riverside%'
```

Этот запрос вернет те же результаты, что и предыдущий, за исключением двух столбцов (**short_name** и **last_reported**).

Чтобы выбрать все столбцы и изменить значения в одном из них, можно воспользоваться оператором `SELECT REPLACE`. Например, вот как можно увеличить на 5 число свободных велосипедов:

```
SELECT
  * REPLACE(num_bikes_available + 5 AS num_bikes_available)
FROM
  `bigquery-public-data`.new_york_citibike.citibike_stations
```

Подзапросы с WITH

Уменьшить повторяемость кода и получить возможность использовать псевдонимы можно с помощью подзапроса:

```
SELECT * FROM (
  SELECT
    gender, tripduration / 60 AS minutes
  FROM
    `bigquery-public-data`.new_york_citibike.citibike_trips
)
WHERE minutes < 10
LIMIT 5
```

Внешний оператор `SELECT` работает с данными, возвращаемыми внутренним подзапросом, заключенным в круглые скобки. Поскольку псевдоним определяется во внутреннем запросе, внешний запрос может использовать его в своем предложении `WHERE`.

Запросы с круглыми скобками могут быть довольно сложными для чтения, поэтому лучше использовать предложение `WITH` и заключать в него код, который в иных условиях определял бы подзапрос:

```
WITH all_trips AS (
  SELECT
    gender, tripduration / 60 AS minutes
  FROM
    `bigquery-public-data`.new_york_citibike.citibike_trips
)
SELECT * from all_trips
WHERE minutes < 10
LIMIT 5
```

В BigQuery предложение `WITH` действует как именованный подзапрос и не создает временных таблиц. Мы будем называть `all_trips` «промежуточным источником» («from_item») — это не таблица, но из него можно делать выборку.

Сортировка с ORDER BY

Управлять порядком следования записей в наборе результатов можно с помощью `ORDER BY`:

```
SELECT
    gender, tripduration/60 AS minutes
FROM
    `bigquery-public-data`.new_york_citibike.citibike_trips
WHERE gender = 'female'
ORDER BY minutes DESC
LIMIT 5
```

По умолчанию строки в результатах не упорядочены. При заданном значении столбца записи по умолчанию сортируются в порядке возрастания значений в этом столбце. Запросив вывести строки в порядке убывания и ограничив их число пятью, мы получили пять самых долгих поездок, совершенных женщинами:

Row	gender	minutes
1	female	250348.9
2	female	226437.93333333332
3	female	207988.71666666667
4	female	159712.05
5	female	154239.0

Обратите внимание, что мы упорядочиваем по значениям в столбце `minutes` — псевдониму. Поскольку предложение `ORDER BY` выполняется после `SELECT`, в нем можно использовать псевдонимы.

Агрегирование

В предыдущем разделе был показан пример, в котором мы преобразовывали секунды в минуты делением на 60, мы обрабатывали каждую запись в таблице и преобразовывали ее. Аналогично можно применять функции агрегирования ко всем записям, чтобы получить в результате только одну запись.

Агрегирование с GROUP BY

Вот как можно найти среднюю продолжительность поездок, совершенных мужчинами:

```
SELECT
    AVG(tripduration / 60) AS avg_trip_duration
FROM
    `bigquery-public-data`.new_york_citibike.citibike_trips
WHERE
    gender = 'male'
```

Этот запрос вернет следующий результат:

Row	avg_trip_duration
1	13.415553172043886

Таким образом, средняя продолжительность поездки из числа совершенных мужчинами в Нью-Йорке составляет примерно 13.4 минуты. Однако поскольку набор данных постоянно обновляется, вы можете получить другой результат.

А что можно сказать о поездках, совершенных женщинами? Вы можете выполнить предыдущий запрос дважды, один раз для мужчин, а другой для женщин, но было бы слишком нерационально анализировать набор данных во второй раз, просто изменив предложение `WHERE`. Вместо этого можно воспользоваться предложением `GROUP BY`:

```
SELECT
    gender, AVG(tripduration / 60) AS avg_trip_duration
FROM
    `bigquery-public-data`.new_york_citibike.citibike_trips
WHERE
    tripduration is not NULL
GROUP BY
    gender
ORDER BY
    avg_trip_duration
```

Этот запрос вернет следующие результаты:

Row	gender	avg_trip_duration
1	male	13.415553172043886
2	female	15.977472148805207
3	unknown	31.4395230232542

Теперь агрегированные значения вычисляются для каждой группы отдельно. Выражение `SELECT` может включать в себя поле, по которому выполняется группировка (`gender`), и функцию агрегирования (`AVG`). Обратите внимание, что на самом деле в наборе данных имеется три пола: мужской (`male`), женский (`female`) и неизвестный (`unknown`).

Подсчет записей с `COUNT`

Чтобы узнать, сколько записей участвовало в вычислении средних значений в предыдущем примере, можно добавить `COUNT()`:

```

SELECT
  gender,
  COUNT(*) AS rides,
  AVG(tripduration / 60) AS avg_trip_duration
FROM
  `bigquery-public-data`.new_york_citibike.citibike_trips
WHERE
  tripduration IS NOT NULL
GROUP BY
  gender
ORDER BY
  avg_trip_duration

```

Этот запрос дал следующие результаты:

Row	gender	rides	avg_trip_duration
1	male	35611787	13.415553172043886
2	female	11376412	15.977472148805207
3	unknown	6120522	31.4395230232542

Фильтрация сгруппированных значений с HAVING

Отфильтровать результаты после группировки можно с помощью предложения **HAVING**. Вот как можно узнать, представители какого пола совершают поездки длительностью дольше 14 минут:

```

SELECT
  gender, AVG(tripduration / 60) AS avg_trip_duration
FROM
  `bigquery-public-data`.new_york_citibike.citibike_trips
WHERE tripduration IS NOT NULL
GROUP BY
  gender
HAVING avg_trip_duration > 14
ORDER BY
  avg_trip_duration

```

Этот запрос дал следующие результаты:

Row	gender	avg_trip_duration
1	female	15.977472148805209
2	unknown	31.439523023254203

Обратите внимание, что предложение **WHERE** можно использовать для фильтрации данных по полу или продолжительности поездки, но его нельзя применить

для фильтрации по средней продолжительности, потому что она вычисляется только после группировки элементов (попробуйте и убедитесь сами!).

Поиск уникальных значений с DISTINCT

Какие гендерные значения присутствуют в столбце `gender` в наборе данных? Чтобы это выяснить, можно использовать `GROUP BY`, но есть более простой способ — `SELECT DISTINCT`:

```
SELECT DISTINCT
  gender
FROM
  `bigquery-public-data`.new_york_citibike.citibike_trips
```

Этот запрос вернет набор результатов, содержащий всего четыре строки:

Row	gender
1	male
2	female
3	unknown
4	

Почему четыре? Что это еще за четвертая строка? Давайте посмотрим:

```
SELECT
  bikeid,
  tripduration,
  gender
FROM
  `bigquery-public-data`.new_york_citibike.citibike_trips
WHERE gender = ""
LIMIT 100
```

Вот результаты этого запроса:

Row	bikeid	tripduration	gender
1	null	null	
2	null	null	
3	null	null	
...			

В данном случае отсутствие значений в столбце `gender` указывает на отсутствие или плохое качество данных. Мы обсудим отсутствующие данные (значения

NULL) и способы их учета и преобразования в главе 3, а пока просто имейте в виду: если вы захотите отфильтровать значения NULL в предложении WHERE, используйте операторы IS NULL или IS NOT NULL, потому что другие операторы сравнения (=, !=, <, >) при применении к NULL будут возвращать NULL и поэтому никогда не будут соответствовать условию в предложении WHERE.

Теперь вернемся к запросу, определяющему уникальные значения в столбце gender с помощью DISTINCT. Важно отметить, что DISTINCT влияет на все результаты, возвращаемые оператором SELECT, а не только на столбец gender. Чтобы представить это наглядно, добавим второй столбец в список SELECT:

```
SELECT DISTINCT
  gender,
  usertype
FROM
  `bigquery-public-data`.new_york_citibike.citibike_trips
WHERE gender != ''
```

На этот раз набор результатов содержит шесть строк, то есть по одной строке для каждой уникальной пары значений gender и usertype (Subscriber или Customer), существующей в наборе данных:

Row	gender	usertype
1	male	Subscriber
2	unknown	Customer
3	female	Subscriber
4	female	Customer
5	male	Customer
6	unknown	Subscriber

Краткое руководство по массивам и структурам

В этом разделе мы предлагаем вашему вниманию краткое руководство по массивам. Оно необходимо, чтобы в следующей главе мы могли познакомить вас с многочисленными типами данных и функций на небольших наглядных наборах данных. Сочетание массивов (квадратные скобки в запросе) и UNNEST дает возможность поэкспериментировать с запросами, функциями и типами данных.

Например, если вы хотите узнать, как работает строковая функция SPLIT, попробуйте следующее:

```

SELECT
  city, SPLIT(city, ' ') AS parts
FROM (
  SELECT * from UNNEST([
    'Seattle WA', 'New York', 'Singapore'
  ]) AS city
)

```

А вот результаты этого запроса:

Row	city	parts
1	Seattle WA	Seattle
		WA
2	New York	New
		York
3	Singapore	Singapore

Возможность жестко закодировать массив значений в самом запросе SQL позволяет экспериментировать с массивами и типами данных и избавляет от необходимости искать подходящий набор данных или ждать завершения долгих запросов. Более того, такие запросы обрабатывают 0 байт и, следовательно, за них не приходится платить.¹

Другой способ поэкспериментировать с набором значений основан на использовании UNION ALL для объединения однострочных операторов SELECT:

```

WITH example AS (
  SELECT 'Sat' AS day, 1451 AS numrides, 1018 AS oneways
  UNION ALL SELECT 'Sun', 2376, 936
  UNION ALL SELECT 'Mon', 1476, 736
)

```

```

SELECT * from example
WHERE numrides < 2000

```

Этот запрос вернет две записи из небольшого встроенного набора данных, в которых число поездок (numrides) меньше 2000:

Row	day	numrides	oneways
1	Sat	1451	1018
2	Mon	1476	736

¹ Все цены актуальны на момент написания этой книги, но вы обязательно должны сами свериться с соответствующими правилами и ценами (<https://cloud.google.com/bigquery/pricing>), так как они могут измениться.

В следующей главе мы используем такие встроенные наборы данных с жестко заданными значениями, чтобы проиллюстрировать различные аспекты поведения некоторых типов данных и функций.

Цель этого раздела — кратко познакомить вас с массивами и структурами, чтобы потом мы могли использовать их в наглядных примерах. Более подробно мы рассмотрим эти понятия в главе 8, поэтому сейчас можете просто бегло просмотреть оставшуюся часть раздела.

Создание массивов с помощью ARRAY_AGG

Рассмотрим определение числа поездок по полу и году:

```
SELECT
  gender
  , EXTRACT(YEAR from starttime) AS year --
  , COUNT(*) AS numtrips
FROM
  `bigquery-public-data`.new_york_citibike.citibike_trips
WHERE gender != 'unknown' and starttime IS NOT NULL
GROUP BY gender, year
HAVING year > 2016
```

Этот запрос вернет следующие результаты:

Row	gender	year	numtrips
1	male	2017	9306602
2	male	2018	3955871
3	female	2018	1260893
4	female	2017	3236735



Зачем нужны начальные запятые в выражении `SELECT`? Стандартный SQL (по крайней мере, на момент написания этой книги) не поддерживает закрывающую запятую, поэтому, перенося запятую в начало следующей строки, мы получаем возможность переупорядочивать или комментировать строки и имеем следующий рабочий запрос:

```
SELECT
  gender
  , EXTRACT(YEAR from starttime) AS year
  -- эта строка закомментирована , COUNT(1) AS numtrips
FROM etc.
```

Уверяем вас, начальные запятые очень скоро станут привычным делом и помогут ускорить процесс разработки.¹

¹ Интересное исследование влияния начальных запятых на успех разработки проектов можно найти по ссылке <https://oreil.ly/mFZKh>.

А можно ли получить временной ряд числа поездок для каждого пола за эти годы, то есть вот такой результат?

Row	gender	numtrips
1	male	9306602
		3955871
2	female	3236735
		1260893

Да, можно. Но для этого нужно создать массив с числом поездок. Представить этот массив в SQL можно с помощью типа ARRAY, используя ARRAY_AGG:

```
SELECT
  gender
  , ARRAY_AGG(numtrips order by year) AS numtrips
FROM (
  SELECT
    gender
    , EXTRACT(YEAR from starttime) AS year
    , COUNT(1) AS numtrips
  FROM
    `bigquery-public-data`.new_york_citibike.citibike_trips
  WHERE gender != 'unknown' and starttime IS NOT NULL
  GROUP BY gender, year
  HAVING year > 2016
)
GROUP BY gender
```

Обычно при группировке по полу вычисляется одно скалярное значение для группы, например AVG(numtrips), чтобы найти среднее число поездок за все годы. ARRAY_AGG позволяет собрать отдельные значения и поместить их в упорядоченный список, или массив ARRAY.

Тип ARRAY можно использовать не только для результатов запросов. BigQuery может принимать иерархические форматы, такие как JSON, поэтому входные данные могут содержать массивы JSON, например:

```
[
  {
    "gender": "male",
    "numtrips": [
      "9306602",
      "3955871"
    ]
  },
  {
```

```

    "gender": "female",
    "numtrips": [
      "3236735",
      "1260893"
    ]
  }
]

```

Если попытаться создать таблицу, загрузив такой файл JSON, в результате получится таблица со столбцом `numtrips` типа `ARRAY`. Массив — это упорядоченный список элементов, отличных от `NULL`; например, `ARRAY<INT64>` — это массив целых чисел.



Технически элементы `NULL` допустимы в массивах, если не пытаться сохранить их в таблице. То есть следующий запрос *не будет* выполнен, потому что вы пытаетесь сохранить массив `[1, NULL, 2]` во временную таблицу, хранящую результаты:

```

WITH example AS (
  SELECT true AS is_vowel, 'a' as letter, 1 as position
  UNION ALL SELECT false, 'b', 2
  UNION ALL SELECT false, 'c', 3
)
SELECT ARRAY_AGG(IF(position = 2, NULL, position)) as
positions from example

```

Но следующий запрос выполнится, потому что промежуточный массив с элементом `NULL` не сохраняется:

```

WITH example AS (
  SELECT true AS is_vowel, 'a' as letter, 1 as position
  UNION ALL SELECT false, 'b', 2
  UNION ALL SELECT false, 'c', 3
)
SELECT ARRAY_LENGTH(ARRAY_AGG(IF(position = 2, NULL,
position))) from example

```

Массив структур **STRUCT**

`STRUCT` — это группа полей, следующих в определенном порядке. Полям можно давать имена (если этого не сделать, BigQuery присвоит им свои имена), и мы советуем использовать их для улучшения читаемости запросов:

```

SELECT
  [
    STRUCT('male' as gender, [9306602, 3955871] as numtrips)
    , STRUCT('female' as gender, [3236735, 1260893] as numtrips)
  ] AS bikerides

```

Этот запрос вернет следующие результаты:

Row	bikerides.gender	bikerides.numtrips
1	male	9306602
		3955871
2	female	3236735
		1260893

Кортежи

Мы могли бы опустить ключевое слово **STRUCT** и имена полей, и в этом случае получили бы кортеж, или неименованную структуру. BigQuery присваивает произвольные имена безымянным столбцам и структурированным полям в результате запроса; поэтому

```
SELECT
  [
    ('male', [9306602, 3955871])
    , ('female', [3236735, 1260893])
  ]
```

вернет результат:

Row	f0_._field_1	f0_._field_2
1	male	9306602
		3955871
2	female	3236735
		1260893

Как видите, отсутствие псевдонимов для полей делает запросы нечитаемыми и неудобными для сопровождения. Не делайте так, разве что ради эксперимента.

Работа с массивами

Имея массив, можно определить его длину и извлечь из него отдельные элементы:

```
SELECT
  ARRAY_LENGTH(bikerides) as num_items
  , bikerides[ OFFSET(0) ].gender as first_gender
FROM
  (SELECT
```

```
[
  STRUCT('male' as gender, [9306602, 3955871] as numtrips)
, STRUCT('female' as gender, [3236735, 1260893] as numtrips)
] AS bikerides)
```

Этот запрос вернет следующие результаты:

Row	num_items	first_gender
1	2	male

Отсчет смещений начинается с нуля, поэтому `OFFSET(0)` даст первый элемент в массиве.¹

Развертывание массива

В запросе

```
SELECT
[
  STRUCT('male' as gender, [9306602, 3955871] as numtrips)
, STRUCT('female' as gender, [3236735, 1260893] as numtrips)
]
```

оператор `SELECT` вернет единственную строку, содержащую массив, поэтому оба пола являются частью одной строки (обратите внимание на столбец «Row»):

Row	f0_gender	f0_numtrips
1	male	9306602
		3955871
	female	3236735
		1260893

`UNNEST` — это функция, которая возвращает элементы массива в виде строк, поэтому `UNNEST` можно применить к массиву результатов, чтобы получить строку, соответствующую каждому элементу массива:

```
SELECT * from UNNEST(
[
  STRUCT('male' as gender, [9306602, 3955871] as numtrips)
, STRUCT('female' as gender, [3236735, 1260893] as numtrips)
])
```

¹ Также можно использовать оператор `ORDINAL(1)`, отсчитывающий смещения, начиная с единицы. Более подробно массивы рассматриваются в главе 8.

Этот запрос вернет следующие результаты:

Row	gender	numtrips
1	male	9306602
		3955871
2	female	3236735
		1260893

Обратите внимание, что **UNNEST** на самом деле является промежуточным источником (**from_item**) — к нему можно применить оператор **SELECT**. Также можно выбрать только части массива. Например, можно получить только столбец **numtrips**:

```
SELECT numtrips from UNNEST(  
[  
  STRUCT('male' as gender, [9306602, 3955871] as numtrips)  
  , STRUCT('female' as gender, [3236735, 1260893] as numtrips)  
)
```

Этот запрос вернет следующие результаты:

Row	numtrips
1	9306602
	3955871
2	3236735
	1260893

Соединение таблиц

Схемы хранилищ данных часто включают первичную таблицу «фактов», которая содержит события, и сопутствующие «размерности» с дополнительной, редко меняющейся информацией. Например, схема розничной торговли может иметь таблицу фактов «Продажи» и сопутствующие таблицы «размерностей» — «Товары» и «Клиенты». При использовании схемы такого типа большинству запросов потребуется выполнять операцию **JOIN**, например, чтобы вернуть названия всех продуктов, приобретенных конкретным клиентом.

BigQuery поддерживает все распространенные типы соединений из реляционной алгебры: внутренние, внешние, перекрестные, антисоединения, полусоединения

и антиполусоединения. Иногда ради увеличения производительности лучше отказаться от JOIN, но в целом BigQuery может эффективно соединять таблицы практически любого размера. В главе 7 мы расскажем, как оптимизировать производительность JOIN, а пока познакомимся только с основными операциями JOIN.

Основы соединения таблиц

В главе 1 мы рассмотрели пример соединения таблиц из двух разных наборов данных, созданных двумя разными организациями. Вернемся к этому примеру:

```
WITH bicycle_rentals AS (  
  SELECT  
    COUNT(starttime) as num_trips,  
    EXTRACT(DATE from starttime) as trip_date  
  FROM `bigquery-public-data`.new_york_citibike.citibike_trips  
  GROUP BY trip_date  
)  
  
rainy_days AS  
(  
  SELECT  
    date,  
    (MAX(prcp) > 5) AS rainy  
  FROM (  
    SELECT  
      wx.date AS date,  
      IF (wx.element = 'PRCP', wx.value/10, NULL) AS prcp  
    FROM  
      `bigquery-public-data`.ghcn_d.ghcnd_2016 AS wx  
    WHERE  
      wx.id = 'USW00094728'  
  )  
  GROUP BY  
    date  
)  
  
SELECT  
  ROUND(AVG(bk.num_trips)) AS num_trips,  
  wx.rainy  
FROM bicycle_rentals AS bk  
JOIN rainy_days AS wx  
ON wx.date = bk.trip_date  
GROUP BY wx.rainy
```

В главе 1 мы просили не обращать внимания на синтаксис, но теперь пришло время с ним разобраться.

Первый оператор WITH извлекает количество поездок за день из таблицы citibike_trips в промежуточный источник (from_item) с именем bicycle_

`rentals`. Это не таблица, а нечто, из чего можно делать выборку. Поэтому будем называть это промежуточным источником. Второй промежуточный источник называется `rainy_days` и создается на основании наблюдений глобальной сети исторических климатологических данных (Global Historical Climate Network, GHCN). Этот промежуточный источник отмечает, был ли день дождливым или нет, в зависимости от того, зафиксировала ли метеостанция 'USW00094728' в Нью-Йорке хотя бы пять миллиметров осадков.

Итак, теперь у нас есть два промежуточных источника. Рассмотрим их по отдельности:

```
WITH bicycle_rentals AS (  
  SELECT  
    COUNT(starttime) as num_trips,  
    EXTRACT(DATE from starttime) as trip_date  
  FROM `bigquery-public-data`.new_york_citibike.citibike_trips  
  GROUP BY trip_date  
)  
SELECT * from bicycle_rentals LIMIT 5
```

Вот как выглядит промежуточный источник `bicycle_rentals`:

Row	num_trips	trip_date
1	31287	2013-09-16
2	22477	2015-12-30
3	37812	2017-09-02
4	54230	2017-11-15
5	25719	2013-11-07

А так выглядит промежуточный источник `rainy_days`:

Row	date	rainy
1	2016-10-11	false
2	2016-12-13	false
3	2016-09-28	false
4	2016-01-25	false
5	2016-05-24	false

Теперь мы можем соединить эти промежуточные источники по условию, что `trip_date` в одном совпадает с `date` во втором:


```

SELECT
  bk.trip_date,
  bk.num_trips,
  wx.rainy
FROM bicycle_rentals AS bk
JOIN rainy_days AS wx
ON wx.date = bk.trip_date
LIMIT 5

```

Этот запрос создаст таблицу, в которой столбцы из двух таблиц соединены по дате:

Row	trip_date	num_trips	rainy
1	2016-07-13	55486	false
2	2016-04-25	42308	false
3	2016-09-27	61346	true
4	2016-07-15	48572	false
5	2016-05-20	52543	false

Имея такую таблицу, легко найти среднее число поездок в дождливые и погожие дни.

Соединение, показанное выше, называется *внутренним соединением*, и этот тип соединения используется, если не указан никакой другой тип.

Вот как можно получить соединение:

- Создать два промежуточных источника. Это может быть что угодно — две таблицы, два подзапроса, массива или оператора WITH, — из чего можно сделать выборку.
- Определить условие соединения. Условие соединения не обязательно должно быть условием равенства; подойдет любое логическое условие, использующее два промежуточных источника.
- Выбрать желаемые столбцы. Если в обоих промежуточных источниках есть столбцы с одинаковыми именами, следует использовать псевдонимы (bk, wx в предыдущем примере запроса), чтобы четко указать, из какого промежуточного источника будет взят каждый столбец.
- Если требуется любое другое соединение, кроме внутреннего, нужно указать тип соединения.

Единственное требование к таким соединениям состоит в том, что все наборы данных, используемые для создания промежуточных источников, должны находиться в одном регионе BigQuery (все общедоступные наборы данных BigQuery находятся в регионе США).

Оператор внутреннего соединения INNER JOIN

Есть несколько типов соединений. **INNER JOIN** (внутреннее соединение), или просто **JOIN**, оператор из предыдущего примера, создает общий набор строк для выборки:

```
WITH from_item_a AS (  
    SELECT 'Dalles' as city, 'OR' as state  
    UNION ALL SELECT 'Tokyo', 'Tokyo'  
    UNION ALL SELECT 'Mumbai', 'Maharashtra'  
)  
  
from_item_b AS (  
    SELECT 'OR' as state, 'USA' as country  
    UNION ALL SELECT 'Tokyo', 'Japan'  
    UNION ALL SELECT 'Maharashtra', 'India'  
)  
  
SELECT from_item_a.*, country  
FROM from_item_a  
JOIN from_item_b  
ON from_item_a.state = from_item_b.state
```

Первый промежуточный источник содержит список городов, а второй определяет страну, в которой находится штат/область. Соединение двух источников дает набор данных с тремя столбцами:

Row	city	state	country
1	Dalles	OR	USA
2	Tokyo	Tokyo	Japan
3	Mumbai	Maharashtra	India

Напомним, что условие соединения не обязательно должно быть проверкой на равенство. Можно использовать любое логическое условие, хотя по возможности лучше использовать условие равенства, потому что BigQuery вернет ошибку, если не сможет эффективно выполнить **JOIN**.

Например, в бизнесе может быть правило, согласно которому доставка в другую страну осуществляется за доплату. Чтобы получить список стран, за доставку в которые нужно будет заплатить, можно отправить такой запрос:

```
SELECT from_item_a.*, country AS surcharge  
FROM from_item_a  
JOIN from_item_b  
ON from_item_a.state != from_item_b.state
```

И получить следующие результаты:

Row	city	state	surcharge
1	Dalles	OR	Japan
2	Dalles	OR	India
3	Tokyo	Tokyo	USA
4	Tokyo	Tokyo	India
5	Mumbai	Maharashtra	USA
6	Mumbai	Maharashtra	Japan

Обратите внимание, что каждый раз, когда выполняется условие соединения, в набор результатов добавляется строка. Поскольку есть две строки, в которых штат/область не совпадают, мы получаем две строки для каждой строки в исходном промежуточном источнике `from_item_a`. Если для какой-либо строки условие соединения не выполняется, данные из нее не попадут в результат.

Оператор перекрестного соединения CROSS JOIN

CROSS JOIN — это перекрестное соединение, или декартово произведение, не имеющее условия соединения. Объединяются все записи из обоих промежуточных источников. Такое соединение мы получили бы, если бы условие в соединении INNER JOIN всегда оценивалось как истинное.

Представьте, что вы организовали турнир и у вас есть таблица победителей в каждом отдельном состязании и таблица, содержащая призы за каждое состязание. Вы сможете вручить каждому победителю приз, соответствующий состязанию, только выполнив внутреннее соединение INNER JOIN:

```
WITH winners AS (  
  SELECT 'John' as person, '100m' as event  
  UNION ALL SELECT 'Hiroshi', '200m'  
  UNION ALL SELECT 'Sita', '400m'  
)  
  
gifts AS (  
  SELECT 'Google Home' as gift, '100m' as event  
  UNION ALL SELECT 'Google Hub', '200m'  
  UNION ALL SELECT 'Pixel3', '400m'  
)  
  
SELECT winners.*, gifts.gift  
FROM winners  
JOIN gifts
```

Этот запрос вернет следующие результаты:

Row	person	event	gift
1	John	100m	Google Home
2	Hiroshi	200m	Google Hub
3	Sita	400m	Pixel3

Хотя если вы решите наградить каждым призом каждого победителя (то есть вручить все три приза каждому победителю в любом состязании), вы можете сделать перекрестное соединение CROSS JOIN:

```
WITH winners AS (  
  SELECT 'John' as person, '100m' as event  
  UNION ALL SELECT 'Hiroshi', '200m'  
  UNION ALL SELECT 'Sita', '400m'  
)  
  
gifts AS (  
  SELECT 'Google Home' as gift  
  UNION ALL SELECT 'Google Hub'  
  UNION ALL SELECT 'Pixel3'  
)  
  
SELECT person, gift  
FROM winners  
CROSS JOIN gifts
```

Этот запрос вернет запись для каждой возможной комбинации:

Row	person	gift
1	John	Google Home
2	John	Google Hub
3	John	Pixel3
4	Hiroshi	Google Home
5	Hiroshi	Google Hub
6	Hiroshi	Pixel3
7	Sita	Google Home
8	Sita	Google Hub
9	Sita	Pixel3

И хотя мы написали

```
SELECT from_item_a.*, from_item_b.*
FROM from_item_a
CROSS JOIN from_item_b
```

мы могли бы выразить то же самое иначе:

```
SELECT from_item_a.*, from_item_b.*
FROM from_item_a, from_item_b
```

Поэтому перекрестное соединение иногда называют *соединением через запятую*.

Оператор внешнего соединения OUTER JOIN

Допустим, у нас есть победители, для которых не было предусмотрено призов, и призы для состязаний, которых нет в нашем турнире:

```
WITH winners AS (
  SELECT 'John' as person, '100m' as event
  UNION ALL SELECT 'Hiroshi', '200m'
  UNION ALL SELECT 'Sita', '400m'
  UNION ALL SELECT 'Kwame', '50m'
),
gifts AS (
  SELECT 'Google Home' as gift, '100m' as event
  UNION ALL SELECT 'Google Hub', '200m'
  UNION ALL SELECT 'Pixel3', '400m'
  UNION ALL SELECT 'Google Mini', '5000m'
)
```

Во внутреннем соединении INNER JOIN (по столбцу состязаний event) победитель в беге на 50 метров не получает приза, а приз для победителя в беге на 5000 метров остается невостребованным. Как мы уже отмечали, в перекрестном соединении CROSS JOIN каждый победитель получает все призы. Внешнее соединение OUTER JOIN определяет результат при невыполнении условия соединения. В табл. 2.2 перечислены разные типы соединений и их результаты.

Таблица 2.2. Типы соединений и их результаты

Синтаксис	Что произойдет	Результат												
SELECT person, gift FROM winners INNER JOIN gifts ON winners.event = gifts.event	Останутся только строки, соответствующие условию	<table><tr><th>Row</th><th>person</th><th>gift</th></tr><tr><td>1</td><td>John</td><td>Google Home</td></tr><tr><td>2</td><td>Hiroshi</td><td>Google Hub</td></tr><tr><td>3</td><td>Sita</td><td>Pixel3</td></tr></table>	Row	person	gift	1	John	Google Home	2	Hiroshi	Google Hub	3	Sita	Pixel3
Row	person	gift												
1	John	Google Home												
2	Hiroshi	Google Hub												
3	Sita	Pixel3												

Таблица 2.2 (окончание)

Синтаксис	Что произойдет	Результат																		
SELECT person, gift FROM winners FULL OUTER JOIN gifts ON winners.event = gifts.event	Останутся все строки, даже не соответствующие условию соединения	<table> <tr><th>Row</th><th>person</th><th>gift</th></tr> <tr><td>1</td><td>John</td><td>Google Home</td></tr> <tr><td>2</td><td>Hiroshi</td><td>Google Hub</td></tr> <tr><td>3</td><td>Sita</td><td>Pixel3</td></tr> <tr><td>4</td><td>Kwame</td><td><i>null</i></td></tr> <tr><td>5</td><td><i>null</i></td><td>Google Mini</td></tr> </table>	Row	person	gift	1	John	Google Home	2	Hiroshi	Google Hub	3	Sita	Pixel3	4	Kwame	<i>null</i>	5	<i>null</i>	Google Mini
Row	person	gift																		
1	John	Google Home																		
2	Hiroshi	Google Hub																		
3	Sita	Pixel3																		
4	Kwame	<i>null</i>																		
5	<i>null</i>	Google Mini																		
SELECT person, gift FROM winners LEFT OUTER JOIN gifts ON winners.event = gifts.event	Оставит всех победителей, но некоторые призы будут отброшены	<table> <tr><th>Row</th><th>person</th><th>gift</th></tr> <tr><td>1</td><td>John</td><td>Google Home</td></tr> <tr><td>2</td><td>Hiroshi</td><td>Google Hub</td></tr> <tr><td>3</td><td>Sita</td><td>Pixel3</td></tr> <tr><td>4</td><td>Kwame</td><td><i>null</i></td></tr> </table>	Row	person	gift	1	John	Google Home	2	Hiroshi	Google Hub	3	Sita	Pixel3	4	Kwame	<i>null</i>			
Row	person	gift																		
1	John	Google Home																		
2	Hiroshi	Google Hub																		
3	Sita	Pixel3																		
4	Kwame	<i>null</i>																		
SELECT person, gift FROM winners RIGHT OUTER JOIN gifts ON winners.event = gifts.event	Оставит все призы, но некоторые победители будут отброшены	<table> <tr><th>Row</th><th>person</th><th>gift</th></tr> <tr><td>1</td><td>John</td><td>Google Home</td></tr> <tr><td>2</td><td>Hiroshi</td><td>Google Hub</td></tr> <tr><td>3</td><td>Sita</td><td>Pixel3</td></tr> <tr><td>4</td><td><i>null</i></td><td>Google Mini</td></tr> </table>	Row	person	gift	1	John	Google Home	2	Hiroshi	Google Hub	3	Sita	Pixel3	4	<i>null</i>	Google Mini			
Row	person	gift																		
1	John	Google Home																		
2	Hiroshi	Google Hub																		
3	Sita	Pixel3																		
4	<i>null</i>	Google Mini																		

Сохранение и совместное использование

Веб-интерфейс BigQuery позволяет сохранять и обмениваться запросами. Это удобно: вы можете отправить коллегам ссылку на текст запроса, а те — немедленно выполнить его. Но имейте в виду, что если у кого-то есть ваш запрос, это не значит, что он гарантированно сможет выполнить его, потому что у него может быть недостаточно прав для доступа к вашим данным. Как открывать и ограничивать доступ к вашим наборам данных, мы обсудим в главе 10.

История запросов и кеширование

Следует отметить, что с целью аудита и кеширования BigQuery сохраняет историю отправленных вами запросов (независимо от успеха их выполнения), как показано на рис. 2.1.

История включает все запросы, отправленные сервису, а не только те, которые были отправлены через веб-интерфейс. Кликнув на любом из запросов, вы полу-

чите текст запроса и возможность открыть запрос в редакторе, чтобы вы могли изменить и повторно запустить его. Кроме того, архивная информация включает объем данных, обработанных запросом, и продолжительность выполнения. На момент написания этой книги история ограничивалась 1000 запросов и шестью месяцами хранения.

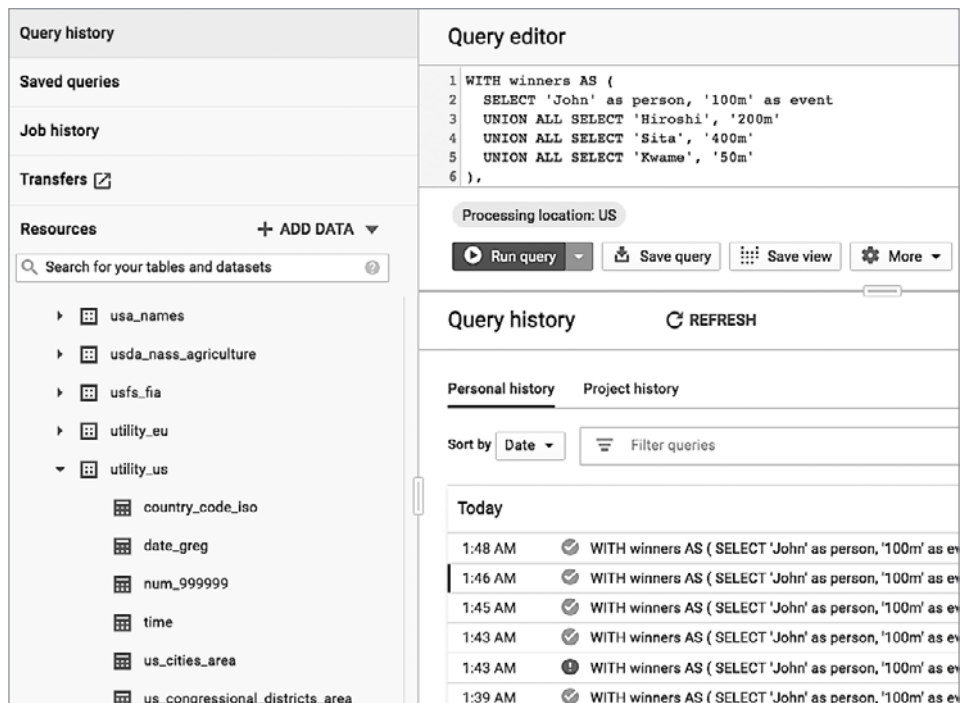


Рис. 2.1. История запросов, отправленных в BigQuery, доступна в разделе «История запросов»

Фактические результаты запроса сохраняются во временной таблице, которая хранится примерно 24 часа. В течение этого периода вы также сможете просматривать результаты запроса из веб-интерфейса. Ваша личная история доступна только вам.

Администраторы проекта, из которого извлекал данные ваш запрос, тоже увидят текст вашего запроса в истории проекта.

Эта временная таблица также используется в качестве кеша, на случай, если в сервис поступит точно такой же текст запроса, не включающий динамические элементы, такие как `CURRENT_TIMESTAMP()` или `RAND()`. Запрос, в ответ на который возвращаются кешированные результаты, обслуживается бесплат-

но, но имейте в виду, что алгоритм определения повторяющихся запросов выполняет простое посимвольное сравнение строк — даже дополнительный пробел может быть расценен как признак несовпадения, и запрос будет выполнен заново.

Сохранение запросов

Любой запрос можно сохранить, загрузив его в редактор запросов. Нужно кликнуть «Save query» («Сохранить запрос») и присвоить ему имя, как показано на рис. 2.2. После этого BigQuery сообщит URL текста запроса.

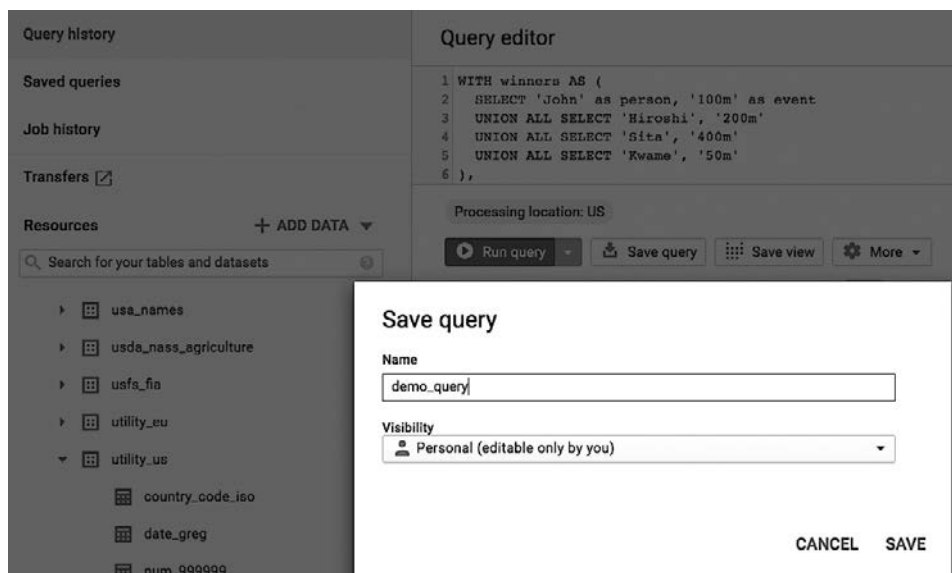


Рис. 2.2. Сохранить запрос можно, щелкнув на кнопке «Save query» («Сохранить запрос») в веб-интерфейсе

Сохраненный запрос можно сделать общедоступным, и в этом случае любой, кто введет в браузере URL запроса, будет перенаправлен на страницу с предварительно заполненным текстом запроса.

Передавая запрос другому человеку, вы передаете только текст запроса; право доступа к данным при этом не передается. Разрешение на использование набора данных для выполнения запроса должно предоставляться независимо, с использованием элементов управления IAM (см. главу 10). Кроме того, в отличие от большинства функций BigQuery, возможность сохранять и обмениваться запросами доступна только в веб-интерфейсе. На момент написания этой книги

REST API и клиентские библиотеки, поддерживающие такую возможность, отсутствовали.

Список сохраненных запросов доступен в пользовательском интерфейсе. Вы можете в любой момент отключить общий доступ к ссылкам и сделать текст запроса снова закрытым, как показано на рис. 2.3.

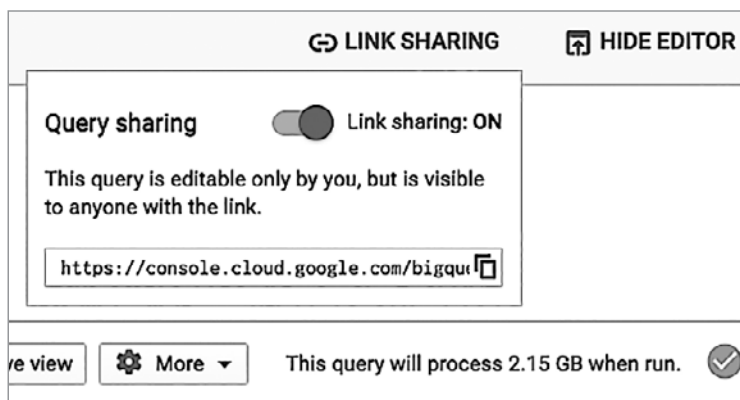


Рис. 2.3. Вы можете в любой момент отключить общий доступ к ссылкам и сделать текст запроса снова закрытым

Представления и общедоступные запросы

Одним из преимуществ общедоступных ссылок на запросы (в отличие от простой пересылки текста запроса по электронной почте) является доступность последней версии запроса после внесенных вами исправлений. Это может пригодиться, если вы пишете пробную версию запроса, чтобы другие могли изучить его, изменить, а затем запустить.

Текст запроса может содержать синтаксические ошибки; его все равно можно сохранить и поделиться ссылкой на незаконченный или шаблонный запрос, который должен выполняться конечным пользователем. Это удобно при совместной работе с коллегами над каким-либо проектом.

Если предполагается, что человек, которому вы передаете запрос, будет использовать все или подмножество возвращаемых им результатов, тогда вам лучше сохранить свой запрос как представление и отправить коллеге ссылку. Другое преимущество представлений состоит в том, что они помещаются в наборы данных и поддерживают возможность управления разрешениями через механизм IAM. Представления также можно материализовать.

Авторизованные представления и их динамическую фильтрацию на основе учетных записей пользователей мы рассмотрим в главе 10.

Выводы

В этой главе мы рассмотрели некоторые аспекты поддержки SQL:2011 в BigQuery: выборку записей (`SELECT`), определение псевдонимов для имен столбцов (`AS`), фильтрацию (`WHERE`), использование подзапросов (круглые скобки и `WITH`), сортировку (`ORDER`), агрегирование (`GROUP`, `AVG`, `COUNT`, `MIN`, `MAX` и т. д.), группировку элементов (`HAVING`), выборку уникальных значений (`DISTINCT`) и соединение таблиц (`INNER/CROSS/OUTER JOIN`). Также язык запросов поддерживает массивы (`ARRAY_AGG`, `UNNEST`) и структуры (`STRUCT`). Кроме того, в этой главе вы узнали, как просмотреть историю выполненных запросов (тексты запросов, а не их результаты) и ее доступность пользователю, отправившему запрос, и администраторам проекта. Также вы узнали, что запросами можно поделиться с другими, передавая им ссылки.

Типы данных, функции и операторы

В некоторых запросах к набору данных с информацией о прокате велосипедов, описанных в предыдущих главах, мы делили продолжительность поездки на 60, потому что это значение имеет числовой тип. Попытка разделить пол (значение в столбце `gender`) на 60 не была успешной, потому что пол — это строка. Спектр доступных функций и операций может ограничиваться в зависимости от типа данных, к которым они применяются.

BigQuery поддерживает несколько типов данных для хранения чисел, строк, значения времени, географических координат, структурированных и полуструктурированных данных:

INT64

Это единственный целочисленный тип. Он может представлять числа в диапазоне примерно от 10^{-19} до 10^{19} . Для представления вещественных чисел используется тип `FLOAT64`, а логических значений — тип `BOOL`.

NUMERIC

Тип `NUMERIC` обеспечивает точное представление вещественных чисел, содержащих до 38 цифр и 9 десятичных знаков после запятой, и подходит для точных расчетов, например, в области финансов.

STRING

Этот тип данных представляет последовательности символов Юникод переменной длины. А для представления последовательностей однобайтных символов (не Юникод) переменной длины используется тип `BYTES`.

TIMESTAMP

Представляет абсолютное значение времени.

DATETIME

Представляет календарные дату и время. Также доступны отдельные типы `DATE` и `TIME`.

GEOGRAPHY

Тип `GEOGRAPHY` представляет точки, линии и полигоны на поверхности земли.

STRUCT и ARRAY

См. описание этих типов в главе 2.

Числовые типы и функции

Как уже было сказано, существует только один целочисленный тип (`INT64`) и только один тип для представления вещественных чисел (`FLOAT64`). Оба типа поддерживают типичные *арифметические операции* (+, -, /, * — для сложения, вычитания, деления и умножения соответственно). То есть долю случаев, когда велосипед брался в прокат для поездки в один конец, можно узнать, просто разделив значение в одном столбце на значение в другом:

```
WITH example AS (  
  SELECT 'Sat' AS day, 1451 AS numrides, 1018 AS oneways  
  UNION ALL SELECT 'Sun', 2376, 936  
)  
SELECT *, (oneways/numrides) AS frac_oneway from example
```

Этот запрос вернет следующие результаты:

Row	day	numrides	oneways	frac_oneway
1	Sat	1451	1018	0.7015851137146796
2	Sun	2376	936	0.3939393939393939

Для целочисленного типа, помимо арифметических, поддерживаются также *битовые* операции (<< и >> для сдвига влево и вправо, & и | для поразрядного И и ИЛИ, и т. д.).

Для манипулирования данными можно использовать *функции*. Функции выполняют действия со своими входными значениями. Как и в других языках программирования, функции в SQL инкапсулируют многократно используемую логику, скрывая сложности их реализации. В табл. 3.1 перечислены некоторые типы функций.

Таблица 3.1. Типы функций

Тип функций	Описание	Пример
Скалярные	<p>Принимают один или несколько параметров и возвращают единственное значение.</p> <p>Скалярную функцию можно использовать везде, где тип возвращаемого ею значения считается допустимым</p>	<p><code>ROUND(3.14)</code> вернет 3 — значение типа <code>FLOAT64</code>, то есть функцию <code>ROUND</code> можно использовать везде, где допустимо значение типа <code>FLOAT64</code>.</p> <p><code>SUBSTR("hello", 1, 2)</code> вернет "he"; является примером скалярной функции, принимающей три параметра</p>
Агрегатные	<p>Функции, выполняющие вычисления с коллекциями значений и возвращающие единственное значение.</p> <p>Агрегатные функции часто используются в предложении <code>GROUP BY</code> для выполнения вычислений с группами записей</p>	<p><code>MAX(tripduration)</code> вычислит максимальное значение в столбце <code>tripduration</code>.</p> <p>К числу агрегатных функций также относятся <code>SUM()</code>, <code>COUNT()</code>, <code>AVG()</code> и т. д.</p>
Аналитические	<p>Аналитические функции, выполняют вычисления с коллекциями значений и возвращают коллекции, содержащие по одному результату для каждого входного значения.</p> <p>Для определения набора записей, к которым применяется аналитическая функция, используется окно</p>	<p>К аналитическим функциям относятся: <code>row_number()</code>, <code>rank()</code> и др.</p> <p>Мы познакомимся с ними в главе 8</p>
Табличные	<p>Функции, возвращающие набор результатов, который можно использовать в предложении <code>FROM</code></p>	<p>Чтобы сделать выборку из массива, к нему можно применить функцию <code>UNNEST</code></p>
Пользовательские	<p>Реализации этих функций определяются пользователем.</p> <p>Пользовательские функции можно писать на SQL (или JavaScript), и они могут возвращать данные любых типов, перечисленных выше</p>	<pre>CREATE TEMP FUNCTION lastElement(arr ANY TYPE) AS (arr[ORDINAL(ARRAY_LENGTH(arr))]);</pre>

Математические функции

Для округления результатов запроса, вычисляющего долю случаев проката велосипедов для поездки в один конец, мы могли бы использовать одну из множества встроенных математических функций (https://cloud.google.com/bigquery/docs/reference/standard-sql/mathematic_functions), работающих с целочисленными и вещественными типами:

```

WITH example AS (
  SELECT 'Sat' AS day, 1451 AS numrides, 1018 AS oneways
  UNION ALL SELECT 'Sun', 2376, 936
)
SELECT *, ROUND(oneways/numrides, 2) AS frac_oneway from example

```

Этот запрос вернет следующие результаты:

Row	day	numrides	oneways	frac_oneway
1	Sat	1451	1018	0.7
2	Sun	2376	936	0.39

Стандартное вещественное деление

Операция деления завершается ошибкой, если знаменатель равен нулю или происходит переполнение результата. Вместо предварительной проверки знаменателя на равенство нулю перед делением лучше использовать специальную функцию деления, если знаменатель может получить нулевое значение, как это возможно в предыдущем примере. Вот как выглядит улучшенная версия запроса:

```

WITH example AS (
  SELECT 'Sat' AS day, 1451 AS numrides, 1018 AS oneways
  UNION ALL SELECT 'Sun', 2376, 936
  UNION ALL SELECT 'Wed', 0, 0
)
SELECT
  *, ROUND(IEEE_Divide(oneways, numrides), 2)
AS frac_oneway from example

```

Функция `IEEE_Divide` соответствует стандарту, разработанному в Институте инженеров по электротехнике и электронике (Institute of Electrical and Electronics Engineers, IEEE), и при попытке деления на ноль возвращает специальное число с плавающей точкой, которое называется «не число» (Not-a-Number, NaN).

Также попробуйте выполнить предыдущий запрос со стандартным оператором деления, используя `SAFE_DIVIDE` (см. далее).¹ Напомним, что для удобства все примеры запросов из этой книги доступны в репозитории GitHub (<https://www.github.com/GoogleCloudPlatform/bigquery-Oreilly-book/>).

Функции SAFE

Любую скалярную функцию можно заставить возвращать `NULL` вместо ошибки, добавив перед ней префикс `SAFE`. Например, следующий запрос вызовет ошибку, потому что логарифм отрицательного числа не определен:

```
SELECT LOG(10, -3), LOG(10, 3)
```

¹ Стандартный оператор деления вызывает ошибку деления на ноль, а `SAFE_DIVIDE` возвращает `NULL` для записи при попытке деления на ноль.

Но если добавить префикс **SAFE** перед **LOG**, как показано ниже:

```
SELECT SAFE.LOG(10, -3), SAFE.LOG(10, 3)
```

этот запрос вернет **NULL** для выражения **LOG(10, -3)**:

Row	f0_	f1_
1	<i>null</i>	2.095903274289385

Префикс **SAFE** можно использовать с математическими функциями, строковыми функциями (например, функция **SUBSTR** обычно вызывает ошибку, если начальный индекс представлен отрицательным числом, но возвращает **NULL**, если вызывается как **SAFE.SUBSTR**) и функциями времени. Однако ее можно применять только к скалярным функциям — ее нельзя использовать с агрегатными, аналитическими или пользовательскими функциями.

Сравнение

Сравнение выполняется с помощью операторов **<**, **<=**, **>**, **>=** и **!=** (или **<>**). В операциях упорядочения считается, что **NULL** и **NaN** меньше действительных чисел (включая **-inf**). Однако сравнение с **NaN** всегда возвращает ложное значение, а сравнение с **NULL** всегда возвращает **NULL**. Эта особенность может приводить к, казалось бы, парадоксальным результатам:

```
WITH example AS (  
  SELECT 'Sat' AS day, 1451 AS numrides, 1018 AS oneways  
  UNION ALL SELECT 'Sun', 2376, 936  
  UNION ALL SELECT 'Mon', NULL, NULL  
  UNION ALL SELECT 'Tue', IEEE_Divide(-3,0), 0 -- это -inf,0  
)  
SELECT * from example  
ORDER BY numrides
```

Этот запрос вернет следующие результаты:

Row	day	numrides	oneways
1	Mon	<i>null</i>	<i>null</i>
2	Tue	–Infinity	0
3	Sat	1451.0	1018
4	Sun	2376.0	936

Однако если попытаться выбрать только дни, насчитывающие меньше 2000 поездок:

```
SELECT * from example
WHERE numrides < 2000
```

вместо трех результатов останется только два:

Row	day	numrides	oneways
1	Sat	1451.0	1018
2	Tue	–Infinity	345

Объясняется это просто: предложение `WHERE` оставило только те записи, для которых сравнение вернуло истинное значение — сравнение `NULL` с числом `2000` дает в результате `NULL`, а не истинное значение.

Обратите внимание, что операторы `&` и `|` используются в BigQuery только для поразрядных операций. Символ `!`, как в операторе `!=`, означает НЕ, но его нельзя использовать как самостоятельный оператор — нельзя записать выражение `!gender`, чтобы вычислить логическое отрицание пола, как это возможно в других языках. Для определения неравенства можно использовать оператор `!=` или `<>`, но будьте последовательны и старайтесь использовать один оператор, `!=` или `<>`.

Точные десятичные вычисления с NUMERIC

Типы `INT64` и `FLOAT64` создавались для скорости и гибкости вычислений, но их точность ограничивается возможностями 64-битного представления двоичных чисел в памяти компьютера. В большинстве случаев это ценный компромисс, но финансовые и бухгалтерские программы часто требуют точных вычислений с числами в десятичном представлении.

Тип данных `NUMERIC` в BigQuery предоставляет 38 цифр для представления чисел, причем девять из них — это цифры после десятичной запятой. В памяти значения этого типа занимают 16 байт и могут точно представлять десятичные дробные числа, что делает этот тип пригодным для финансовых расчетов.

Например, представьте, что вам нужно вычислить сумму трех платежей. Разумеется, хотелось бы получить точный результат. Однако при использовании значений `FLOAT64` постепенно накапливаются крошечные различия между двоичным (в памяти) и десятичным представлением чисел:

```
WITH example AS (
  SELECT 1.23 AS payment
  UNION ALL SELECT 7.89
  UNION ALL SELECT 12.43
)
SELECT
  SUM(payment) AS total_paid,
  AVG(payment) AS average_paid
FROM example
```


Посмотрите, что вернул этот запрос:

Row	total_paid	average_paid
1	21.549999999999997	7.183333333333334

В финансовых и бухгалтерских вычислениях такие ошибки могут накапливаться и усложнять сведение баланса.

А теперь посмотрите, что дает простая замена типа `payment` на `NUMERIC`:

```
WITH example AS (  
  SELECT NUMERIC '1.23' AS payment  
  UNION ALL SELECT NUMERIC '7.89'  
  UNION ALL SELECT NUMERIC '12.43'  
)  
SELECT  
  SUM(payment) AS total_paid,  
  AVG(payment) AS average_paid  
FROM example
```

Проблема решена. Теперь сумма платежей вычисляется точно (среднее значение не имеет точного представления даже при использовании типа `NUMERIC`, потому что это периодическая дробь):

Row	total_paid	average_paid
1	21.55	7.183333333

Обратите внимание, что в BigQuery значения типа `NUMERIC` должны включаться в запросы в виде строк (`NUMERIC '1.23'`); иначе для их представления будет использован менее точный тип с плавающей точкой.

Тип BOOL

Логические (или булевы) переменные могут иметь два значения, `True` (истина) и `False` (ложь). Поскольку язык SQL не различает регистр символов, логические значения можно записывать как `TRUE`, `true` и т. д.

Логические операции

Помните, мы говорили о том, что предложение `WHERE` может включать логические выражения с операторами `AND`, `OR` и `NOT` и круглыми скобками для управления порядком вычислений? Проиллюстрируем эти возможности на примере следующего запроса:

```
SELECT
    gender, tripduration
FROM
    `bigquery-public-data`.new_york_citibike.citibike_trips
WHERE (tripduration < 600 AND gender = 'female') OR gender = 'male'
```

К логическим переменным можно применять операторы сравнения:

```
WITH example AS (
    SELECT NULL AS is_vowel, NULL as letter, -1 as position
    UNION ALL SELECT true, 'a', 1
    UNION ALL SELECT false, 'b', 2
    UNION ALL SELECT false, 'c', 3
)
SELECT * from example WHERE is_vowel != false
```

Этот запрос вернет следующие результаты:

Row	is_vowel	letter	position
1	true	<i>a</i>	<i>1</i>

Однако для сравнения с предопределенными константами часто проще использовать оператор IS, как показано в следующем примере:

```
WITH example AS (
    SELECT NULL AS is_vowel, NULL as letter, -1 as position
    UNION ALL SELECT true, 'a', 1
    UNION ALL SELECT false, 'b', 2
    UNION ALL SELECT false, 'c', 3
)
SELECT * from example WHERE is_vowel IS NOT false
```

Этот запрос вернет следующие результаты:

Row	is_vowel	letter	position
1	<i>null</i>	<i>null</i>	<i>-1</i>
2	true	<i>a</i>	<i>1</i>

Обратите внимание, что эти два запроса вернули разные результаты. Операторы сравнения (=, !=, < и др.) возвращают NULL при сравнении с NULL, а оператор IS — нет.



Значения NULL обычно представляют отсутствующие значения или значения, которые не были выбраны. Они не имеют конкретного значения и не являются нулями, пустыми строками или пробелами. Будьте внимательны, если набор данных содержит значения NULL, потому что сравнение с NULL всегда

возвращают NULL, и предложение WHERE будет отфильтровывать значения NULL. Используйте оператор IS, чтобы проверить равенство с NULL.

Код получается проще, если использовать логические переменные:

```
WITH example AS (  
    SELECT NULL AS is_vowel, NULL as letter, -1 as position  
    UNION ALL SELECT true, 'a', 1  
    UNION ALL SELECT false, 'b', 2  
    UNION ALL SELECT false, 'c', 3  
)  
  
SELECT * from example WHERE is_vowel
```

Этот запрос вернет тот же результат, что и при использовании выражения is_vowel IS TRUE:

Row	is_vowel	letter	position
1	true	a	1

И помните, такая удобочитаемость в значительной степени зависит от выбора имени логической переменной!

Условные выражения

Логические значения можно использовать не только в предложении WHERE. Многие запросы можно упростить, используя условные выражения в SELECT. Допустим, вам нужно рассчитать стоимость каждого товара в каталоге с использованием наценки и налоговой ставки, соответствующих этому товару. Если в каталоге нужная информация отсутствует, можно задать наценку или ставку налога по умолчанию. Сделать это можно с помощью функции IF:

```
WITH catalog AS (  
    SELECT 30.0 AS costPrice, 0.15 AS markup, 0.1 AS taxRate  
    UNION ALL SELECT NULL, 0.21, 0.15  
    UNION ALL SELECT 30.0, NULL, 0.09  
    UNION ALL SELECT 30.0, 0.30, NULL  
    UNION ALL SELECT 30.0, NULL, NULL  
)  
SELECT  
    *, ROUND(  
        costPrice *  
        IF(markup IS NULL, 1.05, 1+markup) *  
        IF(taxRate IS NULL, 1.10, 1+taxRate)  
    , 2) AS salesPrice  
FROM catalog
```

Этот запрос вернет верное значение salesPrice для всех товаров, кроме тех, для которых базовая цена неизвестна:

Row	costPrice	markup	taxRate	salesPrice
1	30.0	0.15	0.1	37.95
2	null	0.21	0.15	null
3	30.0	null	0.09	34.34
4	30.0	0.3	null	42.9
5	30.0	null	null	34.65

В первом параметре функция IF принимает условие для оценки. Если условие истинно, она возвращает второй параметр, в противном случае — третий. Поскольку эта функция вызывается в операторе SELECT, она применяется к каждой записи.

Обработка NULL с помощью COALESCE

А как быть, если наценку и налог можно включить в стоимость, только если отсутствует какое-то одно значение и не более? То есть если налоговая ставка не указана в каталоге, можно применить 10% налог по умолчанию, но только если указана наценка на товар.

Самый удобный способ продолжать вычислять выражения в списке, пока не будет получено значение, отличное от NULL, — использовать COALESCE:

```
WITH catalog AS (
  SELECT 30.0 AS costPrice, 0.15 AS markup, 0.1 AS taxRate
  UNION ALL SELECT NULL, 0.21, 0.15
  UNION ALL SELECT 30.0, NULL, 0.09
  UNION ALL SELECT 30.0, 0.30, NULL
  UNION ALL SELECT 30.0, NULL, NULL
)
SELECT
  *, ROUND(COALESCE(
    costPrice * (1+markup) * (1+taxrate),
    costPrice * 1.05 * (1+taxrate),
    costPrice * (1+markup) * 1.10,
    NULL
  ),2) AS salesPrice
FROM catalog
```

Этот запрос вернет следующие результаты (от предыдущего результата он отличается только последней строкой):

Row	costPrice	markup	taxRate	salesPrice
1	30.0	0.15	0.1	37.95
2	null	0.21	0.15	null
3	30.0	null	0.09	34.34

Row	costPrice	markup	taxRate	salesPrice
4	30.0	0.3	null	42.9
5	30.0	null	null	null

COALESCE сокращает вычисления, как только получит первый результат, отличный от NULL, то есть после получения результата, отличного от NULL, последующие выражения не вычисляются. Следовательно, заключительное значение NULL в COALESCE не требуется, но оно делает намерение более ясным.

BigQuery поддерживает функцию IFNULL — как более простой вариант COALESCE, — когда имеется только два входных выражения. Выражение IFNULL(a, b) аналогично выражению COALESCE(a, b) и возвращает b, если a равно NULL. Иначе говоря, IFNULL(a, b) — это то же самое, что и IF(a IS NULL, b, a).

Первый запрос в этом разделе об условных выражениях можно упростить так:

```
SELECT
  *, ROUND(
    costPrice *
    (1 + IFNULL(markup, 0.05)) *
    (1 + IFNULL(taxrate, 0.10))
  , 2) AS salesPrice
FROM catalog
```

Явное и неявное приведение типов

Рассмотрим следующий набор данных, в котором количество часов, отработанных каждым сотрудником, хранится в виде строки, чтобы можно было учесть причины отсутствия на работе (это пример плохой организации схемы, но вы уж простите нас):

```
WITH example AS (
  SELECT 'John' as employee, 'Paternity Leave' AS hours_worked
  UNION ALL SELECT 'Janaki', '35'
  UNION ALL SELECT 'Jian', 'Vacation'
  UNION ALL SELECT 'Jose', '40'
)
```

Теперь предположим, что нам понадобилось узнать общее количество отработанных часов. Следующий запрос для этого не годится, потому что clock_worked имеет строковый, а не числовой тип:

```
WITH example AS (
  SELECT 'John' as employee, 'Paternity Leave' AS hours_worked
  UNION ALL SELECT 'Janaki', '35'
  UNION ALL SELECT 'Jian', 'Vacation'
  UNION ALL SELECT 'Jose', '40'
)
SELECT SUM(hours_worked) from example
```

Нам нужно явно преобразовать `hours_worked` в значение типа `INT64` перед агрегированием. Такое преобразование называется *явным приведением типа* и выполняется с использованием функции `CAST()`. Если приведение невозможно, BigQuery вернет ошибку. Чтобы возвращалось значение `NULL`, используйте `SAFE_CAST`. Например, следующий запрос вернет ошибку:

```
SELECT CAST("true" AS bool), CAST("invalid" AS bool)
```

А теперь попробуем использовать `SAFE_CAST`:

```
SELECT CAST("true" AS bool), SAFE_CAST("invalid" AS bool)
```

Этот запрос вернет результат:

Row	f0_	f1_
1	true	null

Неявное преобразование называется *неявным приведением типа* и выполняется автоматически, когда в выражение, требующее один тип данных, передается другой тип. Например, если в ситуации, где требуется значение типа `FLOAT64`, использовать значение типа `INT64`, целое число будет приведено к типу числа с плавающей точкой. В BigQuery неявное приведение выполняется только для преобразования типа `INT64` в `FLOAT64` и `NUMERIC`, а также для преобразования типа `NUMERIC` в `FLOAT64`. Любые другие преобразования должны выполняться с явными приведениями типа с помощью `CAST`.

В задаче вычисления общего количества отработанных часов не все строки `hours_worked` можно преобразовать в целые числа, поэтому следует использовать `SAFE_CAST`:

```
WITH example AS (  
  SELECT 'John' as employee, 'Paternity Leave' AS hours_worked  
  UNION ALL SELECT 'Janaki', '35'  
  UNION ALL SELECT 'Jian', 'Vacation'  
  UNION ALL SELECT 'Jose', '40'  
)  
SELECT SUM(SAFE_CAST(hours_worked AS INT64)) from example
```

Этот запрос вернет следующий результат:

Row	f0_
1	75

Если бы эту проблему можно было решить простым изменением схемы и все записи содержали бы только числа, пусть и в виде строк, для вычислений можно было бы использовать простую функцию `CAST`:

```

WITH example AS (
  SELECT 'John' as employee, '0' AS hours_worked
  UNION ALL SELECT 'Janaki', '35'
  UNION ALL SELECT 'Jian', '0'
  UNION ALL SELECT 'Jose', '40'
)
SELECT SUM(CAST(hours_worked AS INT64)) from example

```

Использование COUNTIF, чтобы избежать приведения логических значений

Рассмотрим следующий набор данных:

```

WITH example AS (
  SELECT true AS is_vowel, 'a' as letter, 1 as position
  UNION ALL SELECT false, 'b', 2
  UNION ALL SELECT false, 'c', 3
)
SELECT * from example

```

Вот результаты этого запроса:

Row	is_vowel	letter	position
1	true	<i>a</i>	1
2	false	<i>b</i>	2
3	false	<i>c</i>	3

Теперь предположим, что вам нужно найти общее количество гласных. Возможно, вы захотите использовать вот такой простой запрос:

```
SELECT SUM(is_vowel) as num_vowels from example
```

Но из этого ничего не выйдет (попробуйте, и вы убедитесь сами!), потому что SUM, AVG и другие подобные функции не могут оперировать логическими значениями. Перед агрегированием вам придется привести логические значения к типу INT64, например, так:

```

WITH example AS (
  SELECT true AS is_vowel, 'a' as letter, 1 as position
  UNION ALL SELECT false, 'b', 2
  UNION ALL SELECT false, 'c', 3
)
SELECT SUM(CAST (is_vowel AS INT64)) as num_vowels from example

```

Этот запрос вернет следующий результат:

Row	num_vowels
1	1

Однако приведения типов желательно избегать при любой возможности. Другой, более ясный подход основан на использовании оператора **IF** с логическими значениями:

```
WITH example AS (  
  SELECT true AS is_vowel, 'a' as letter, 1 as position  
  UNION ALL SELECT false, 'b', 2  
  UNION ALL SELECT false, 'c', 3  
)  
SELECT SUM(IF(is_vowel, 1, 0)) as num_vowels from example
```

Но есть еще более удобное решение, основанное на использовании **COUNTIF**:

```
WITH example AS (  
  SELECT true AS is_vowel, 'a' as letter, 1 as position  
  UNION ALL SELECT false, 'b', 2  
  UNION ALL SELECT false, 'c', 3  
)  
SELECT COUNTIF(is_vowel) as num_vowels from example
```

Строковые функции

При обработке данных часто требуется обработка строк, поэтому BigQuery предоставляет целую библиотеку встроенных строковых функций (https://cloud.google.com/bigquery/docs/reference/standard-sql/string_functions), например:

```
WITH example AS (  
  SELECT * from unnest([  
    'Seattle', 'New York', 'Singapore'  
  ]) AS city  
)  
SELECT  
  city  
  , LENGTH(city) AS len  
  , LOWER(city) AS lower  
  , STRPOS(city, 'or') AS orpos  
FROM example
```

Этот запрос вычисляет длину строки, преобразует все символы в нижний регистр и находит местоположение подстроки в столбце **city**:

Row	city	len	lower	orpos
1	Seattle	7	seattle	0
2	New York	8	new York	6
3	Singapore	9	singapore	7

Подстрока "or" присутствует в строках "New York" и "Singapore", но отсутствует в строке "Seattle".

Чаще всего при работе со строками используются функции `SUBSTR` и `CONCAT`. `SUBSTR` извлекает подстроку, а `CONCAT` объединяет входные строки. Следующий запрос находит позицию символа `@` в адресе электронной почты, извлекает имя пользователя и объединяет с названием города, в котором он живет:

```
WITH example AS (  
  SELECT 'armin@abc.com' AS email, 'Annapolis, MD' as city  
  UNION ALL SELECT 'boyan@bca.com', 'Boulder, CO'  
  UNION ALL SELECT 'carrie@cab.com', 'Chicago, IL'  
)  
  
SELECT  
  CONCAT(  
    SUBSTR(email, 1, STRPOS(email, '@') - 1), -- имя пользователя  
    ' from ', city) AS callers  
FROM example
```

Вот как выглядят результаты этого запроса:

Row	callers
1	armin from Annapolis, MD
2	boyan from Boulder, CO
3	carrie from Chicago, IL

Интернационализация

Строковые значения в BigQuery являются строками Юникод, поэтому не следует опираться на особенности английского языка. Например, в японском языке отсутствует понятие «верхнего регистра», а преобразования кодировки UTF-8, используемой по умолчанию, в последовательность байтов недостаточно для таких языков, как тамильский. Например:

```
WITH example AS (  
  SELECT * from unnest([  
    'Seattle', 'New York', 'சிங்கப்பூர்', '東京'  
  ]) AS city  
)  
SELECT  
  city  
  , UPPER(city) AS allcaps  
  , CAST(city AS BYTES) as bytes  
FROM example
```

Этот запрос работает совсем не так, как предполагалось:

Row	city	allcaps	bytes
1	Seattle	SEATTLE	U2VhdHRsZQ==
2	New York	NEW YORK	TmV3lFlvcms=
3	சிங்கப்பூர்	சிங்கப்பூர்	4K6a4K6/4K6Z4K+N4K6V4K6q4K +N4K6q4K+C4K6w4K+N
4	東京	東京	5p2x5Lqs

BigQuery поддерживает три разных способа представления строк: в виде массивов символов Юникод, в виде массивов байтов и в виде массива кодовых пунктов Юникод (INT64):

```
WITH example AS (
  SELECT * from unnest([
    'Seattle', 'New York', 'சிங்கப்பூர்', '東京'
  ]) AS city
)
SELECT
  city
  , CHAR_LENGTH(city) as char_len
  , TO_CODE_POINTS(city)[ORDINAL(1)] as first_code_point
  , ARRAY_LENGTH(TO_CODE_POINTS(city)) as num_code_points
  , CAST (city AS BYTES) as bytes
  , BYTE_LENGTH(city) as byte_len
FROM example
```

Обратите внимание, что для одних и тех же строк функции CHAR_LENGTH и BYTE_LENGTH могут возвращать разные значения и что количество кодовых пунктов всегда совпадает с количеством символов:

Row	city	char_len	first_code_point	num_code_points	bytes	byte_len
1	Seattle	7	83	7	U2VhdHRsZQ==	7
2	New York	8	78	8	TmV3lFlvcms=	8
3	சிங்கப்பூர்	11	2970	11	4K6a4K6/4K6Z4K +N4K6V4K6q4K +N4K6q4K+C4K6w4K +N	33
4	東京	2	26481	2	5p2x5Lqs	6

Из-за этих различий важно знать, какие столбцы могут содержать текст на разных языках, и учитывать языковые различия при использовании строковых функций.

Формирование и парсинг строк

Парсинг строк можно выполнять простым приведением к типу `INT64` или `FLOAT64`, но для их формирования потребуется использовать `FORMAT`:

```
SELECT
  CAST(42 AS STRING)
  , CAST('42' AS INT64)
  , FORMAT('%03d', 42)
  , FORMAT('%5.3f', 32.457842)
  , FORMAT('%5.3f', 32.4)
  , FORMAT('**%s**', 'H')
  , FORMAT('%s-%03d', 'Agent', 7)
```

Вот результат этого запроса:

Row	f0_	f1_	f2_	f3_	f4_	f5_	f6_
1	42	42	042	32.458	32.400	**H**	Agent-007

Функция `FORMAT` действует подобно функции `printf` в языке C (<http://www.cplusplus.com/reference/cstdio/printf/>) и принимает те же спецификаторы формата. Некоторые из спецификаторов продемонстрированы в предыдущем примере. Хотя функция `FORMAT` также принимает даты и временные метки, для преобразования данных этого типа в строки лучше использовать специализированные функции `FORMAT_DATE` и `FORMAT_TIMESTAMP`, которые учитывают региональные настройки.

Функции для обработки строк

В конвейерах извлечения, преобразования и загрузки (Extract, Transform, Load; ETL) часто приходится работать со строками, поэтому было бы целесообразно кратко познакомиться с соответствующими вспомогательными функциями в BigQuery:

```
SELECT
  ENDS_WITH('Hello', 'o') -- true
  , ENDS_WITH('Hello', 'h') -- false
  , STARTS_WITH('Hello', 'h') -- false
  , STRPOS('Hello', 'e') -- 2
  , STRPOS('Hello', 'f') -- 0, если не найдено
  , SUBSTR('Hello', 2, 4) -- отсчет позиций начинается с 1
  , CONCAT('Hello', 'World')
```

Вот результат этого запроса:

Row	f0_	f1_	f2_	f3_	f4_	f5_	f6_
1	true	false	false	2	0	ello	HelloWorld

Обратите внимание, как работает `SUBSTR()`. Первый параметр — это начальная позиция (отсчет позиций начинается с 1), а второй параметр — желаемое количество символов в подстроке.

Функции преобразования

Другой набор функций, которые могут пригодиться для работы со строками:

```
SELECT
  LPAD('Hello', 10, '*') -- дополнение символами * слева
, RPAD('Hello', 10, '*') -- дополнение справа
, LPAD('Hello', 10) -- дополнение пробелами слева
, LTRIM(' Hello ') -- удаление пробелов слева
, RTRIM(' Hello ') -- удаление пробелов справа
, TRIM(' Hello ') -- удаление пробелов по обе стороны
, TRIM('***Hello***', '*') -- удаление символов * по обе стороны
, REVERSE('Hello') -- переворачивание строки
```

Взгляните на результат этого запроса:

Row	f0_	f1_	f2_	f3_	f4_	f5_	f6_	f7_
1	*****Hello	Hello*****	Hello	Hello	Hello	Hello	Hello	olleH

Регулярные выражения

Регулярные выражения предлагают гораздо более удобную семантику, чем вспомогательные функции. Например, `STRPOS` и другие могут найти только определенные символы, тогда как `REGEXP_CONTAINS` позволяет реализовать более эффективный поиск.

Например, следующий запрос определяет наличие в поле `column` почтового индекса США (краткая форма которого содержит пять цифр, а длинная форма имеет дополнительные четыре цифры, отделяемые от основной части дефисом или пробелом):

```
SELECT
  column
, REGEXP_CONTAINS(column, r'\d{5}(?:[-\s]\d{4})?') has_zipcode
, REGEXP_CONTAINS(column, r'^\d{5}(?:[-\s]\d{4})?$') is_zipcode
, REGEXP_EXTRACT(column, r'\d{5}(?:[-\s]\d{4})?') the_zipcode
```

```

, REGEXP_EXTRACT_ALL(column, r'\d{5}(?:[-\s]\d{4})?') all_zipcodes
, REGEXP_REPLACE(column, r'\d{5}(?:[-\s]\d{4})?', '*****') masked
FROM (
  SELECT * from unnest([
    '12345', '1234', '12345-9876',
    'abc 12345 def', 'abcde-fghi',
    '12345 ab 34567', '12345 9876'
  ]) AS column
)

```

Вот результаты этого запроса:

Row	column	has_ zipcode	is_zipcode	the_zipcode	all_zipcodes	masked
1	12345	true	true	12345	12345	*****
2	1234	false	false	<i>null</i>		1234
3	12345-9876	true	true	12345-9876	12345-9876	*****
4	abc 12345 def	true	false	12345	12345	abc ***** def
5	abcde-fghi	false	false	<i>null</i>		abcde-fghi
6	12345 ab 34567	true	false	12345	12345	***** ab *****
					34567	
7	12345 9876	true	true	12345 9876	12345 9876	*****

и несколько замечаний к ним:

- Регулярному выражению `\d{5}` соответствует любая строка, состоящая из пяти десятичных цифр.
- Вторая часть выражения, заключенная в круглые скобки, ищет необязательную (обратите внимание на `?` после скобок) группу `(?:)` из четырех десятичных цифр (`\d{4}`), которая отделена от первых пяти цифр дефисом или пробелом (`\s`).
- Наличие в строке `\d`, `\s` и других подобных последовательностей символов может создавать проблемы, поэтому мы добавили перед строкой с регулярным выражением символ `r` (от англ. *raw* — необработанная), что заставляет BigQuery интерпретировать ее буквально.
- Второе выражение иллюстрирует, как найти точное соответствие: мы просто задали команду, чтобы начало и конец найденного фрагмента совпадали с началом (^) и концом (\$) исследуемой строки.
- Извлечь часть строки, совпавшую с регулярным выражением, можно с помощью `REGEXP_EXTRACT`. Эта функция возвращает `NULL`, если в строке нет со-

впадений с регулярным выражением, и только первое совпадение, если их несколько.

- `REGEXP_EXTRACT_ALL` возвращает все найденные совпадения. Если совпадений нет, она вернет пустой массив.
- `REGEXP_REPLACE` замещает каждое совпадение указанной строкой.

Синтаксис регулярных выражений в BigQuery соответствует синтаксису, поддерживаемому библиотекой RE2 с открытым исходным кодом, разработанной в компании Google (<https://github.com/google/re2>). Справочную информацию по синтаксису, используемому этой библиотекой, можно найти на странице <https://github.com/google/re2/wiki/Syntax>. Иногда регулярные выражения трудно читаются, но они предлагают богатые возможности, которые стоит освоить.¹

Краткие итоги по строковым функциям

При анализе данных настолько часто приходится выполнять операции со строками, что вам определенно стоит иметь хотя бы общее представление об имеющихся возможностях. Более полную информацию о синтаксисе можно найти в документации BigQuery (https://cloud.google.com/bigquery/docs/reference/standard-sql/string_functions). В табл. 3.2 перечислены имеющиеся функции, разбитые на соответствующие категории.

Таблица 3.2. Категории строковых функций

Категория	Функции	Примечания
Представление	<code>CHAR_LENGTH</code> , <code>BYTE_LENGTH</code> , <code>TO_CODE_POINTS</code> , <code>CODE_POINTS_TO_STRING</code> , <code>SAFE_CONVERT_BYTES_TO_STRING</code> , <code>TO_HEX</code> , <code>TO_BASE32</code> , <code>TO_BASE64</code> , <code>FROM_HEX</code> , <code>FROM_BASE32</code> , <code>FROM_BASE64</code> , <code>NORMALIZE</code>	Функция <code>NORMALIZE</code> , например, приводит разные символы Юникод к единой эквивалентной форме
Формирование и парсинг	<code>FORMAT</code> , <code>REPEAT</code> , <code>SPLIT</code>	Функция <code>FORMAT</code> имеет синтаксис, аналогичный синтаксису функции <code>printf</code> в языке C: выражение <code>format("%03d", 12)</code> вернет <code>012</code> . Для преобразований региональных настроек используйте <code>FORMAT_DATE</code> и пр.

¹ Освоить регулярные выражения вам поможет книга Джеффри Фридла (Jeffrey Friedl) *Mastering Regular Expressions*, O'Reilly (<http://shop.oreilly.com/product/9781565922570.do>). (Издание на русском языке: *Фридл Джеффри. Регулярные выражения*. СПб.: Питер, 2018. — Примеч. пер.)

Категория	Функции	Примечания
Вспомогательные	ENDS_WITH, LENGTH, STARTS_WITH, STRPOS, SUBSTR, CONCAT	Функция LENGTH действует подобно CHAR_LENGTH для строк и BYTE_LENGTH для последовательностей байтов
Преобразование	LPAD, LOWER, LTRIM, REPLACE, REVERSE, RPAD, RTRIM, TRIM, UPPER	По умолчанию функции *TRIM удаляют пробельные символы Юникод, но при необходимости можно явно указать удаляемые символы
Регулярные выражения	REGEXP_CONTAINS, REGEXP_EXTRACT, REGEXP_EXTRACT_ALL, REGEXP_REPLACE	Описание синтаксиса регулярных выражений в BigQuery можно найти по адресу https://github.com/google/re2/wiki/Syntax

Операции со значениями TIMESTAMP

Тип **TIMESTAMP** (отметка времени) представляет абсолютное значение времени, независимо от региона (часового пояса). То есть метка времени 2017-09-27 12:30:00.45 (Sep 27, 2017, at 12:30 UTC) соответствует времени 2017-09-27 13:30:00.45+1:00 (в часовом поясе, опережающем нулевой меридиан на час):

```
SELECT t1, t2, TIMESTAMP_DIFF(t1, t2, MICROSECOND)
FROM (SELECT
  TIMESTAMP "2017-09-27 12:30:00.45" AS t1,
  TIMESTAMP "2017-09-27 13:30:00.45+1" AS t2
)
```

Этот запрос вернет следующий результат:

Row	t1	t2	fo_
1	2017-09-27 12:30:00.450 UTC	2017-09-27 12:30:00.450 UTC	0

Парсинг и форматирование отметок времени

Платформа BigQuery проявляет особую лояльность, когда дело доходит до парсинга отметок времени. Дата и время в строковом представлении времени могут отделяться символом Т или пробелом в соответствии со стандартом ISO 8601 (<https://www.iso.org/iso-8601-date-and-time-format.html>). Аналогично месяц, день, час и т. д. могут иметь или не иметь начальные нули. Однако лучше ис-

пользовать каноническое представление, показанное в предыдущем абзаце. Как можно судить по этому примеру, временные отметки способны представлять только четырехзначные годы; годы до нашей эры нельзя представить с помощью `TIMESTAMP`.

Для парсинга строки с отметкой времени не в каноническом формате можно использовать `PARSE_TIMESTAMP`:

```
SELECT
  fmt, input, zone
  , PARSE_TIMESTAMP(fmt, input, zone) AS ts
FROM (
  SELECT '%Y%m%d-%H%M%S' AS fmt, '20181118-220800' AS input, '+0' as zone
  UNION ALL SELECT '%c', 'Sat Nov 24 21:26:00 2018', 'America/Los_Angeles'
  UNION ALL SELECT '%x %X', '11/18/18 22:08:00', 'UTC'
)
```

Этот запрос вернет следующий результат:

Row	fmt	input	zone	ts
1	%Y%m%d-%H%M%S	20181118-220800	+0	2018-11-18 22:08:00 UTC
2	%c	Sat Nov 24 21:26:00 2018	America/Los_Angeles	2018-11-25 05:26:00 UTC
3	%x %X	11/18/18 22:08:00	UTC	2018-11-18 22:08:00 UTC

В первом примере для создания отметки времени из указанной строки используются спецификаторы формата для года, месяца, дня и т. д. Во втором и третьем примерах используются предопределенные спецификаторы для часто встречающихся форматов представления даты и времени.¹

И наоборот, для вывода отметки времени в любом желаемом формате можно использовать `FORMAT_TIMESTAMP`:

```
SELECT
  ts, fmt
  , FORMAT_TIMESTAMP(fmt, ts, '+6') AS ts_output
FROM (
  SELECT CURRENT_TIMESTAMP() AS ts, '%Y%m%d-%H%M%S' AS fmt
  UNION ALL SELECT CURRENT_TIMESTAMP() AS ts, '%c' AS fmt
  UNION ALL SELECT CURRENT_TIMESTAMP() AS ts, '%x %X' AS fmt
)
```

Этот запрос вернет следующий результат:

¹ Полный список спецификаторов формата можно найти в документации (https://cloud.google.com/bigquery/docs/reference/standard-sql/timestamp_functions#supported_format_elements_for_timestamp).

Row	ts	fmt	ts_output
1	2018-11-25 05:42:13.939840 UTC	%Y%m%d-%H%M%S	20181125-114213
2	2018-11-25 05:42:13.939840 UTC	%c	Sun Nov 25 11:42:13 2018
3	2018-11-25 05:42:13.939840 UTC	%x %X	11/25/18 11:42:13

Для получения системного времени, соответствующего моменту выполнения запроса, в предыдущем примере используется функция `CURRENT_TIMESTAMP()`. В обеих функциях, `PARSE_TIMESTAMP` и `FORMAT_TIMESTAMP`, параметр, определяющий часовой пояс, является необязательным — в его отсутствие используется часовой пояс UTC.

Извлечение календарных данных

Имея отметку времени, можно получить информацию, соответствующую григорианскому календарю. Например, вот какую информацию можно получить о Дне перемирия:¹

```
SELECT
  ts
  , FORMAT_TIMESTAMP('%c', ts) AS repr
  , EXTRACT(DAYOFWEEK FROM ts) AS dayofweek
  , EXTRACT(YEAR FROM ts) AS year
  , EXTRACT(WEEK FROM ts) AS weekno
FROM (
  SELECT PARSE_TIMESTAMP('%Y%m%d-%H%M%S', '19181111-054500') AS ts
)
```

Вот результат этого запроса:

Row	ts	repr	dayofweek	year	weekno
1	1918-11-11 05:45:00 UTC	Mon Nov 11 05:45:00 1918	2	1918	45

Здесь предполагается, что неделя начинается в воскресенье и дни, предшествующие первому воскресенью года, относятся к неделе с порядковым номером 0. Однако подобный порядок принят не везде. Поэтому если вы находитесь в стра-

¹ Согласно информации из статьи https://en.wikipedia.org/wiki/Armistice_Day (https://ru.wikipedia.org/wiki/День_перемирия), перемирие, положившее конец Первой мировой войне, было подписано 11 ноября 1918 года в 5:45 утра. Зимой 1918 года Франция находилась в часовом поясе UTC; см. <https://www.timeanddate.com/time/zone/france/paris>.

не (например, в Израиле), где неделя начинается в субботу, вы можете указать другой день, с которого начинается неделя:

```
EXTRACT(WEEK('SATURDAY') FROM ts)
```

Количество секунд, прошедших с начала эпохи Unix (1 января 1970 года), нельзя получить с помощью EXTRACT. Для этой цели существуют специальные функции преобразования в отметки времени от начала эпохи Unix:

```
SELECT
  UNIX_MILLIS(TIMESTAMP "2018-11-25 22:30:00 UTC")
  , UNIX_MILLIS(TIMESTAMP "1918-11-11 22:30:00 UTC") -- неверный результат
  , TIMESTAMP_MILLIS(1543185000000)
```

Этот запрос вернет следующий результат:

Row	f0_	f1_	f2_
1	1543185000000	-1613784600000	2018-11-25 22:30:00 UTC

Обратите внимание, что во втором преобразовании происходит переполнение и оно возвращает отрицательное число, но ошибки при этом не возникает.

Арифметические операции с временными метками

Временные метки позволяют прибавлять и вычитать интервалы времени. Также можно найти разницу между двумя метками времени. Во всех этих функциях необходимо указать единицы измерения интервала:

```
SELECT
  EXTRACT(TIME FROM TIMESTAMP_ADD(t1, INTERVAL 1 HOUR)) AS plus_1h
  , EXTRACT(TIME FROM TIMESTAMP_SUB(t1, INTERVAL 10 MINUTE)) AS minus_10min
  , TIMESTAMP_DIFF(CURRENT_TIMESTAMP(),
    TIMESTAMP_SUB(CURRENT_TIMESTAMP(), INTERVAL 1 MINUTE),
    SECOND) AS plus_1min
  , TIMESTAMP_DIFF(CURRENT_TIMESTAMP(),
    TIMESTAMP_ADD(CURRENT_TIMESTAMP(), INTERVAL 1 MINUTE),
    SECOND) AS minus_1min
FROM (SELECT
  TIMESTAMP "2017-09-27 12:30:00.45" AS t1
)
```

Этот запрос вернет метки времени, соответствующие моментам, спустя час и за 10 минут до указанной метки времени, а также разницу в секундах между указанной меткой времени и моментами времени, спустя одну минуту и за минуту до указанной метки времени:

Row	plus_1h	minus_10min	plus_1min	minus_1min
1	13:30:00.450000	12:20:00.450000	60	-60

DATE, TIME и DATETIME

BigQuery имеет еще три типа для представления времени: `DATE`, `TIME` и `DATETIME`. Тип `DATE` можно использовать, если не нужна высокая точность и достаточно знать только дату события. Тип `TIME` можно использовать для представления времени суток и выполнения арифметических операций со временем. С помощью `TIME`, например, можно ответить на вопрос: «Который час будет через восемь часов после начала события?» `DATETIME` отображает `TIMESTAMP` в определенном часовом поясе, поэтому он может пригодиться, если часовой пояс события известен, и вам не понадобится вручную пересчитывать часовой пояс.

Для `DATETIME` доступны аналоги большинства функций, поддерживаемых типом `TIMESTAMP`. То есть вы можете использовать `DATETIME_ADD`, `DATETIME_SUB`, `DATETIME_DIFF`, а также `PARSE_DATETIME` и `FORMAT_DATETIME`. Из `DATETIME` также можно извлекать календарную информацию. Оба типа полностью совместимы — вы можете извлечь `DATETIME` из `TIMESTAMP` и преобразовать `DATETIME` в `TIMESTAMP`:

```
SELECT
  EXTRACT(DATETIME FROM CURRENT_TIMESTAMP()) as dt
  , CAST(CURRENT_DATETIME() AS TIMESTAMP) as ts
```

Этот запрос вернет следующий результат:

Row	dt	ts
1	2018-11-25T07:03:15.055141	2018-11-25 07:03:15.055141 UTC

Обратите внимание, что каноническое представление `DATETIME` включает букву `T`, отделяющую дату от времени, тогда как в представлении `TIMESTAMP` для этой цели используется пробел. Тип `TIMESTAMP` также явно включает часовой пояс, тогда как в `DATETIME` часовой пояс задается неявно. Но в большинстве случаев `DATETIME` и `TIMESTAMP` в BigQuery являются взаимозаменяемыми.

Тип `DATE` просто представляет часть значения `DATETIME` (или `TIMESTAMP`, приведенного к некоторому часовому поясу), соответствующую дате, а `TIME` — часть, соответствующую времени. Поскольку многие события в реальном мире отмечаются в определенную дату (то есть могут отмечаться несколько раз в течение дня), многие таблицы в базах данных хранят только дату `DATE`, что дает возможность непосредственно анализировать и форматировать даты. Тип `TIME` обычно используется только для представления части `DATETIME`.

По этим причинам мы советуем использовать только `TIMESTAMP` и `DATE`. Однако у типа `TIMESTAMP` есть одна неприятная особенность. Метки времени хранятся в BigQuery в виде восьмибайтных значений с точностью до микросекунды. Это означает, что с помощью типа `TIMESTAMP` можно представить любую микросекунду любого года в диапазоне от 0 до 9999. В некоторых других базах данных

(например, MySQL) значения типа `TIMESTAMP` хранятся в четырехбайтном представлении, а `DATETIME` — в восьмибайтном. В этих системах тип `TIMESTAMP` способен представлять только значения времени эпохи Unix (1970–2038 годы), это означает, что с помощью этого типа нельзя сохранить дату рождения 60-летнего человека или дату истечения срока 30-летней ипотеки. Поэтому, даже при том что тип `TIMESTAMP` с успехом можно использовать в BigQuery, вы не сможете применить ту же схему в MySQL, а это может затруднить перемещение запросов и данных между BigQuery и MySQL.

Функции для работы с географическими координатами

Более подробно функции для работы с географическими координатами мы рассмотрим в главе 8, где познакомим вас с их дополнительными особенностями. А в этом разделе мы дадим лишь краткое введение.

Тип `GEOGRAPHY` можно использовать для представления точек, линий и полигонов на поверхности земли (то есть без высоты над уровнем моря). Поскольку поверхность земли неровная, точки можно представить только в координатах усредненного эллипсоида. В BigQuery географические координаты точек и вершин линий и полигонов представлены как лежащие на эллипсоиде WGS84 (https://en.wikipedia.org/wiki/World_Geodetic_System). Фактически это тот же эллипсоид, который используется глобальной системой определения местоположения (Global Positioning System, GPS), поэтому вы можете брать значения долготы и широты, сообщаемые большинством датчиков, и использовать их непосредственно в BigQuery.

Самый простой географический элемент — точка, определяемая долготой и широтой. Например,

```
ST_GeogPoint(-122.33, 47.61)
```

представляет точку с координатами 47.61° с. ш. и 122.33° з. д. — Сиэтл, штат Вашингтон, США.

В числе общедоступных наборов данных в BigQuery есть таблица с полигонами, соответствующими штатам и территориям в США. Благодаря этому можно написать запрос, определяющий, в каком штате находится географическая точка:

```
SELECT
  state_name
FROM `bigquery-public-data`.utility_us.us_states_area
WHERE
  ST_Contains(
    state_geom,
    ST_GeogPoint(-122.33, 47.61))
```

Как и ожидалось, этот запрос вернул следующий результат:

Row	state_name
1	Washington

Чтобы определить, попадает ли заданная точка в границы штата (хранятся в столбце `state_geom` в наборе данных BigQuery), запрос использует функцию `ST_Contains`. Географические функции в BigQuery соответствуют спецификации SQL/MM 3 (<https://oreil.ly/9AgOe>) и аналогичны тем, что предоставляет библиотека PostGIS (<https://oreil.ly/x8kNM>) для Postgres.

Выводы

В заключение этой главы в табл. 3.3 перечислены типы данных, которые поддерживает BigQuery.

Таблица 3.3. Типы данных, поддерживаемые в BigQuery

Тип данных	Примеры поддерживаемых функций и операторов	Примечания
INT64	Арифметические операции (+, −, /, * для сложения, вычитания, деления и умножения соответственно)	Примерно от 10^{-19} до 10^{19}
NUMERIC	Арифметические операции	до 38 цифр и 9 десятичных знаков после запятой; подходит для финансовых вычислений
FLOAT64	Арифметические операции, а также IEEE_DIVIDE	Соответствует стандарту IEEE-754, если одно из значений равно NaN или $\pm\infty$
BOOL	Условные инструкции. MIN, MAX. Не поддерживает SUM, AVG и др. (логический тип необходимо предварительно привести к типу INT64)	Имеет два значения: True и False. Язык SQL не различает регистр символов, поэтому логические значения можно записывать как TRUE, true и т. д.
STRING	Для операций со значениями этого типа используются специальные строковые функции (https://cloud.google.com/bigquery/docs/reference/standard-sql/string_functions), такие как CONCAT, LENGTH и др.	Строки состоят из символов Юникода и имеют переменную длину

Таблица 3.3 (продолжение)

Тип данных	Примеры поддерживаемых функций и операторов	Примечания
BYTES		Последовательности символов переменной длины. Тип BYTES поддерживает также многие строковые операции
TIMESTAMP	<code>CURRENT_TIMESTAMP()</code> представляет «текущее время». Из значения отметки времени можно извлечь месяц, год, день недели и т. д. Арифметические операции с отметками времени поддерживаются с использованием специальных функций (https://cloud.google.com/bigquery/docs/reference/standard-sql/timestamp_functions). Арифметические операторы не поддерживаются	Абсолютный момент времени с точностью до микросекунды, соответствует стандарту ISO 8601. Это рекомендуемый способ хранения времени в BigQuery
DATE	<code>CURRENT_DATE()</code> представляет текущую дату в часовом поясе UTC, тогда как <code>CURRENT_DATE("America/Los_Angeles")</code> представляет текущую дату в часовом поясе Лос-Анжелеса (Los Angeles). По аналогии с типом <code>TIMESTAMP</code> , поддерживает арифметические операции с использованием специальных функций (https://cloud.google.com/bigquery/docs/reference/standard-sql/date_functions)	2018-3-14 (или 2018-03-14) — 14 марта 2018 года, независимо от часового пояса (https://cloud.google.com/bigquery/docs/reference/standard-sql/data-types#timestamp-type). Поскольку этот тип представляет разные 24-часовые интервалы в разных часовых поясах, для представления абсолютного момента времени лучше использовать тип <code>TIMESTAMP</code> . В этом случае вы сможете получить значение типа <code>DATE</code> из значения <code>TIMESTAMP</code> относительно определенного часового пояса
DATETIME	Те же, что и для типа <code>DATE</code>	2018-03-14 3:14:57 или 2018-03-14T03:14:57.000000, как и <code>DATE</code> , не зависит от часового пояса. В большинстве случаев лучше использовать <code>TIMESTAMP</code>
TIME	Те же, что и для типа <code>DATETIME</code> , но, в отличие от последнего, не содержит даты	Не зависит от конкретной даты или часового пояса. Представляет значения в диапазоне от 00:00:00 до 23:59:59.999999

Тип данных	Примеры поддерживаемых функций и операторов	Примечания
GEOGRAPHY	Поддерживает топологические операции через специальные функции (https://cloud.google.com/bigquery/docs/reference/standard-sql/geography_functions)	Точки, линии и полигоны на поверхности земли (то есть без высоты над уровнем моря). Использует представление формы Земли в виде эллипсоида WGS84; это тот же эллипсоид, который используется глобальной системой определения местоположения (Global Positioning System, GPS). Самый простой географический элемент — точка, определяемая долготой и широтой
STRUCT	Позволяют ссылаться на поля по именам	Упорядоченная коллекция полей. Имена полей являются необязательными, то есть структуру можно определить как <code>STRUCT<INT64, STRING></code> или <code>STRUCT<id INT64, name STRING></code>
ARRAY	Позволяют ссылаться на поля с помощью числовых смещений, объединять элементы в массивы или разворачивать массивы, чтобы получить их элементы по одному	Упорядоченный список непустых элементов, например: <code>ARRAY<INT64></code> . Массивы массивов не поддерживаются, но эту проблему можно решить, создав массив структур <code>STRUCT</code> , где каждая структура может содержать массив, например: <code>ARRAY<STRUCT<ARRAY<INT64>>></code> (о массивах рассказывалось в главе 2)

Все типы данных, кроме массивов и структур, можно использовать в предложениях `ORDER BY` и `GROUP BY`.

Загрузка данных в BigQuery

В предыдущей главе мы написали такой запрос:

```
SELECT
  state_name
FROM `bigquery-public-data`.utility_us.us_states_area
WHERE
  ST_Contains(
    state_geom,
    ST_GeogPoint(-122.33, 47.61))
```

Мы также узнали, что точка с координатами (-122.33, 47.61) находится в штате Вашингтон. Откуда взялись данные для `state_name` и `state_geom`?

Обратите внимание на предложение `FROM` в запросе. Владелец проекта `bigquery-public-data` уже загрузили информацию о границах штатов в таблицу с именем `us_states_area`, находящуюся в наборе данных `utility_us`. Поскольку они открыли доступ к набору данных `utility_us` всем аутентифицированным пользователям BigQuery (платформа поддерживает и более строгие ограничения), мы смогли обратиться к таблице `us_states_area` в этом наборе данных.

Но как они загрузили данные в BigQuery? В этой главе мы рассмотрим разные способы загрузки данных в BigQuery и начнем с самых основ.

Основы

Такие данные, как сведения о границах штатов в США, изменяются редко,¹ и эти изменения настолько незначительны, что в большинстве приложений их можно просто игнорировать. На языке хранилищ данных это называется *медленно меняющимся измерением*. На момент написания этой книги последнее

¹ Шесть–восемь раз в десять лет — см. <https://oreil.ly/Merow>.

изменение границ штатов в США произошло 1 января 2017 года и затронуло 19 домовладельцев и одну заправочную станцию.¹

Таким образом, информация о границах штатов относится к типу данных, которые обычно загружаются только один раз. Аналитики обращаются к одной и той же таблице и игнорируют тот факт, что данные могут со временем меняться. Например, розничная фирма может учитывать только то, в каком штате сейчас находится магазин, чтобы гарантировать применение правильной ставки налога к товарам, продаваемым в этом магазине. Поэтому при изменении границ штатов, например, в результате заключения договора между штатами или из-за изменения фарватера реки, владельцы набора данных могут решить записать в таблицу более актуальные данные. Можно не учитывать тот факт, что после обновления запросы могут возвращать результаты, немного отличающиеся от результатов, полученных до обновления.

Однако игнорировать влияние времени на правильность данных можно не всегда. Если данные о границах штатов использует кадастровая компания, которая должна отслеживать право собственности на земельные участки, или аудиторская фирма, которая должна проверять правильность уплаты определяемых штатом пошлин за поставки, сделанные в разные годы, то им нужна возможность получить актуальную на тот момент информацию о границах штатов. Поэтому, даже при том что первая часть этой главы посвящена выполнению разовой загрузки, подумайте, а не лучше ли запланировать периодическое обновление данных и дать пользователям возможность получать информацию о версии данных.

Загрузка из локального источника

Правительство США выпускает «систему показателей» для колледжей, чтобы помочь потребителям сравнить стоимость и предполагаемую ценность высшего образования. Давайте для иллюстрации загрузим эти данные в BigQuery. Исходные данные доступны на сайте *catalog.data.gov*. Для удобства они также доступны в репозитории GitHub этой книги в виде архива *04_load/college_scorecard.csv.gz* (<https://github.com/GoogleCloudPlatform/bigquery-oreilly-book/>). Мы загрузили файл в формате CSV с сайта *data.gov* и сжали его с помощью открытой утилиты *gzip*.



Зачем мы сжали файл? Исходный, несжатый файл весит около 136 Мбайт, а в сжатом формате — всего 18 Мбайт. Мы собираемся отправить файл в BigQuery по сети, поэтому имеет смысл оптимизировать передачу. Команда загрузки в BigQuery способна обрабатывать сжатые файлы, но не спо-

¹ См. <https://abc7ny.com/news/border-of-north-and-south-carolina-shifted-on-january-1st/1678605/> и <https://www.nytimes.com/2014/08/24/opinion/sunday/how-the-carolinas-fixed-their-blurred-lines.html>.

собна загружать части сжатого файла параллельно. Мы могли бы значительно ускорить загрузку, передавая в BigQuery разделяемый файл, либо в несжатом формате CSV, но находящийся в облачном хранилище Cloud Storage (чтобы минимизировать накладные расходы на передачу по сети), либо в таком формате, как Avro, когда каждый блок сжат и весь файл можно разделить между несколькими рабочими узлами.

Разделяемый файл может загружаться несколькими рабочими узлами, начиная с разных его частей, но для этого необходимо, чтобы рабочие узлы могли «найти» предсказуемую точку в середине файла, не читая его содержимое с самого начала. Сжатие файла целиком с помощью gzip не дает такой возможности, но это возможно с помощью поблочного сжатия, например Avro. Поэтому использование сжатого разделяемого формата, такого как Avro, является несомненным преимуществом. Однако если у вас есть файлы CSV или JSON, которые можно разделить только в несжатом виде, вы должны определить, стоит ли высокая скорость передачи по сети увеличенного времени загрузки.

В Cloud Shell можно пролистывать сжатый файл с помощью `zless`:

```
zless college_scorecard.csv.gz
```



Вот пошаговая инструкция:

1. Откройте Cloud Shell в браузере: <https://console.cloud.google.com/cloudshell>.
2. В окне терминала введите: `git clone https://github.com/GoogleCloudPlatform/bigquery-oreilly-book`.
3. Перейдите в папку с файлом: `cd bigquery-oreilly-book/04_load`.
4. Введите команду `zless college_scorecard.csv.gz` и затем, нажимая клавишу пробела, перелистайте данные. Введите букву `q`, чтобы выйти.

Файл содержит строку заголовка с именами столбцов. Каждая следующая строка содержит одну запись с данными.

Чтобы загрузить данные в BigQuery, сначала создайте набор данных с именем `ch04`:

```
bq --location=US mk ch04
```

Инструмент командной строки `bq` предлагает удобный способ взаимодействия с сервисом BigQuery в Google Cloud Platform (GCP), хотя все, что можно сделать с помощью `bq`, также можно сделать с помощью REST API. Многие операции можно выполнить, используя GCP Cloud Console. Здесь мы предлагаем создать (`mk`) набор данных с именем `ch04`.

Наборы данных в BigQuery играют роль папок верхнего уровня и используются для организации и управления доступом к таблицам, представлениям и моделям машинного обучения. Набор данных создается в текущем проекте,¹ и именно этот проект будет нести затраты, связанные с хранением таблиц в этом наборе данных (затраты на выполнение запросов относятся на счет проекта, выполнявшего запрос).

Мы также указали, что набор данных должен быть создан на территории США (это предполагается по умолчанию, поэтому данный параметр можно было опустить). Выбирая регион для размещения набора данных, можно указывать несколько регионов (например, `US` или `EU`) и конкретные регионы (например, `us-east4`, `europa-west2` и `australia-southeast1`).² Будьте внимательны при выборе региона для загрузки данных: на момент написания этой книги запросы не могли соединять таблицы в разных регионах. Здесь мы будем использовать объединение нескольких регионов `US`, чтобы наши запросы могли соединять таблицы из общедоступных наборов данных, расположенных в Соединенных Штатах.

Далее, из каталога, содержащего копию репозитория GitHub, загрузим данные из файла в таблицу в BigQuery:

```
bq --location=US \  
  load \  
  --source_format=CSV --autodetect \  
  ch04.college_scorecard \  
  ./college_scorecard.csv.gz
```

Здесь мы предлагаем утилите `bq` загрузить набор данных, сообщив ей, что исходные данные имеют формат CSV и нам хотелось бы, чтобы она автоматически определили схему (то есть типы данных отдельных столбцов). Затем мы указываем, что таблица с именем `college_scorecard` должна быть создана в наборе данных `ch04` и что данные должны загружаться из файла `college_scorecard.csv.gz`, находящегося в текущем каталоге.

Запустив команду, мы столкнулись с проблемой:

```
Could not parse 'NULL' as int for field HBCU (position 26) starting at location 119459103
```

¹ Назначается в раскрывающемся списке в консоли GCP Cloud Console или последней командой `gcloud init`. Как правило, проект соответствует задаче или небольшой команде.

² Актуальный список можно найти по ссылке <https://cloud.google.com/bigquery/docs/locations>.

³ Не удалось преобразовать в `int` значение `'NULL'` в поле HBCU (столбец 26), начиная с местоположения 11945910. — *Примеч. пер.*

ЗАГРУЗКА ИЛИ ПОТОКОВАЯ ПЕРЕДАЧА?

Загрузка данных в BigQuery не предполагает никаких затрат, но вам нужно будет заплатить за хранение данных после загрузки.¹ Если ваш тариф с фиксированной оплатой, придется заплатить за загрузку данных в BigQuery сверх фиксированного тарифа. Поэтому если вам не нужно поддерживать высокую актуальность данных в вашем хранилище, дешевле будет настроить передачу данных из Cloud Storage по расписанию (о которой мы поговорим далее в этой главе). Если попутно необходимо выполнять преобразования, для ежедневной загрузки данных в BigQuery можно использовать Cloud Composer или Cloud Functions.

Команда `bq` просто вызывает REST API, предоставляемый службой BigQuery. Поэтому загрузить данные можно многими другими способами — все они основаны на обращении к одному и тому же REST API. Также имеются клиентские библиотеки для нескольких языков, включая Java, Python и Node.js, — они предоставляют возможность выгрузки данных программным способом. Использование клиентских библиотек мы рассмотрим в главе 5.

Если вам нужны данные, близкие к реальному времени, используйте для их загрузки в BigQuery потоковый способ передачи (<https://cloud.google.com/bigquery/streaming-data-into-bigquery>). Несмотря на то что за потоковую передачу взимается дополнительная плата, мы рекомендуем вам использовать именно этот способ, если данные требуется загружать часто и постоянно поддерживать их в актуальном состоянии. Мы не советуем загружать данные с использованием небольших и часто выполняемых заданий (например, раз в минуту). Такие частые обновления таблиц могут привести к значительной фрагментации и высоким издержкам на изменение метаданных, в результате чего запросы к таким таблицам будут выполняться медленно, пока BigQuery не выполнит этап оптимизации в какой-то момент в будущем. В отличие от частых загрузок небольшими порциями, при потоковой передаче строки пакета сначала накапливаются на стороне сервера в течение некоторого времени, а затем записываются в хранилище, что позволяет ограничить фрагментацию и сохранить высокую производительность запросов. Потоковые данные доступны для запросов сразу, тогда как данные, загружаемые обычным способом, становятся доступны только через некоторое время. Более того, при частых загрузках небольшими порциями любые ограничения пропускной способности или операции резервного копирования в системах, производящих файлы, могут привести к неожиданным задержкам в получении данных.

¹ Как отмечалось в предыдущих главах, все упоминания о стоимости услуги верны на момент написания этой книги, но вы обязательно должны сами свериться с соответствующими правилами и ценами (<https://cloud.google.com/bigquery/pricing>), так как они могут измениться.

Это вызвало сбой задания загрузки со следующей ошибкой:¹

```
CSV table encountered too many errors, giving up. Rows: 591; errors: 1.2
```

Проблема в том, что, опираясь на анализ большей части данных в файле CSV, механизм автоопределения схемы в BigQuery предположил, что 26-й столбец (с именем `HBCU`) должен иметь целочисленный тип, но в 591-й строке в файле в этом поле встретился текст `NULL` — обычно это означает, что данный колледж не ответил на соответствующий вопрос.³

Есть несколько способов решения этой проблемы. Например, можно отредактировать сам файл данных, если известно, как должно выглядеть значение, чтобы не возникало ошибки. Другой подход — явно определить схему для каждого столбца и, в данном случае, изменить тип столбца `HBCU`, объявив его строковым, чтобы значение `NULL` стало допустимым. Также можно предложить BigQuery игнорировать несколько неверных записей, передав параметр, например `--max_bad_records = 20`. Наконец, можно сообщить программе загрузки в BigQuery, что этот конкретный файл использует строку `NULL` для обозначения пустых значений (обычно в CSV пустые значения представлены как пустые поля).

Воспользуемся последним способом, потому что он выглядит наиболее подходящим:⁴

```
bq --location=US \  
  load --null_marker=NULL \  
    --source_format=CSV --autodetect \  
    ch04.college_scorecard \  
    ./college_scorecard.csv.gz
```

Полный список параметров команды `bq load` можно получить, набрав `bq load --help`. По умолчанию `bq load` добавит новые данные в таблицу. Мы в данном случае хотим заменить существующую таблицу, поэтому добавим параметр `--replace`:

¹ Алгоритм автоматического определения схемы постепенно совершенствуется и обрабатывает все больше и больше универсальных случаев, поэтому у вас эта ошибка может не возникнуть. Однако автоматическое определение едва ли когда-нибудь будет работать идеально. Независимо от того, какой аспект схемы определился неправильно, наша задача состоит в следующем: взять за основу схему, которая была определена автоматически, и внести в нее необходимые коррективы, как будет показано далее в этом разделе.

² В таблице CSV выявлено слишком много ошибок. Записей: 591; ошибок: 1. — *Примеч. пер.*

³ Целочисленный столбец может содержать значения `NULL`, но в файле эти значения определены нестандартным способом. BigQuery интерпретирует текст `NULL` как строку, поэтому загрузка не удалась.

⁴ Строка `NULL` в файле обозначает отсутствие данных для этого поля, и ей должно соответствовать значение `NULL` в нашей таблице BigQuery.

```
bq --location=US \
  load --null_marker=NULL --replace \
  --source_format=CSV --autodetect \
  ch04.college_scorecard \
  ./college_scorecard.csv.gz
```

Также, чтобы добавить новые данные в существующую таблицу, можно указать параметр `--replace=false`.

Стоит отметить, что разовые загрузки можно выполнять из веб-интерфейса BigQuery. Перейдите в свой проект, и вам будет предложена кнопка для создания набора данных (`ch04` в нашем случае); выберите набор данных, и вам будет предложена кнопка для создания таблицы. Затем следуйте инструкциям, чтобы выгрузить файл в таблицу BigQuery. Однако на момент написания этой книги веб-интерфейс позволял загружать данные из локального файла, только если его размер не превышал 10 Мбайт и 16 000 строк. То есть этот способ не подходит для набора данных с оценочными параметрами колледжей, если предварительно не выгрузить его в облачное хранилище Google Cloud Storage.

Даже если вы не собираетесь (или не можете) использовать веб-интерфейс для загрузки данных, все равно посмотрите на созданную таблицу с помощью веб-интерфейса, чтобы проверить правильность сведений о таблице, а также схемы, которая была определена автоматически. В веб-интерфейсе также можно отредактировать некоторые детали о таблице, даже после ее создания. Например, можно указать, что срок действия таблицы автоматически истекает через определенное количество дней, добавить столбцы или ослабить ограничение, требующее обязательного наличия действительного значения в столбце, чтобы в нем можно было сохранять значения `NULL`.



Задать срок хранения таблицы также можно с помощью инструкции `ALTER TABLE SET OPTIONS`, например:

```
ALTER TABLE ch04.college_scorecard
SET OPTIONS (
  expiration_timestamp=
    TIMESTAMP_ADD(CURRENT_TIMESTAMP(), INTERVAL 7 DAY),
  description="College Scorecard table that expires
    seven days from now"
)
```

Узнать больше можно по ссылке https://cloud.google.com/bigquery/docs/reference/standard-sql/data-definition-language#alter_table_set_options_statement.

Независимо от того, как была загружена таблица, любой, у кого есть доступ к набору данных, где находится таблица, сможет выполнить запрос к ней. По умолчанию вновь созданный набор данных доступен только тем, у кого есть разрешение соответствующего уровня доступа к проекту. Однако вы можете дать

право на доступ к нему¹ конкретным лицам (по их учетным записям Google), доменам (например, xyz.com) или группам Google. Использование сервиса управления идентификацией и доступом (Identity and Access Management, IAM) для совместного использования наборов данных мы обсудим в главе 10. Но в данный момент обратиться к таблице смогут только те, кто имеет доступ к проекту, содержащему набор данных:

```
SELECT
  INSTNM
  , ADM_RATE_ALL
  , FIRST_GEN
  , MD_FAMINC
  , MD_EARN_WNE_P10
  , SAT_AVG
FROM
  ch04.college_scorecard
WHERE
  SAFE_CAST(SAT_AVG AS FLOAT64) > 1300
  AND SAFE_CAST(ADM_RATE_ALL AS FLOAT64) < 0.2
  AND SAFE_CAST(FIRST_GEN AS FLOAT64) > 0.1
ORDER BY
  CAST(MD_FAMINC AS FLOAT64) ASC
```

Этот запрос извлекает название колледжа (**INSTNM**), долю поступивших абитуриентов и другую информацию о колледжах, для которых средний балл SAT выше 1300, а доля поступивших абитуриентов меньше 20%, что вполне подходит под определение «элитных» колледжей. Также запрос фильтрует колледжи, оставляя только те, в которых доля студентов, чьи родители не имеют высшего образования, превышает 10%, и сортирует их в порядке возрастания среднего дохода семьи, тем самым отбирая элитные колледжи, которые принимают учащихся, находящихся в неблагоприятном культурном или экономическом положении. Запрос также извлекает средний заработок студентов через 10 лет после поступления:

Row	INSTNM	ADM_RATE_ALL	FIRST_GEN	MD_FAMINC	MD_EARN_WNE_P10	SAT_AVG
1	University of California — Berkeley	0.1692687830816	0.3458005249	31227	64700	1422
2	Columbia University in the City of New York	0.06825366802669	0.2504905167	31310.5	83300	1496

¹ На момент написания книги эта возможность не поддерживалась в «новом» веб-интерфейсе; чтобы воспользоваться ею, требовалась утилита командной строки bq.

Row	INSTNM	ADM_RATE_ALL	FIRST_GEN	MD_FAMINC	MD_EARN_WNE_P10	SAT_AVG
3	University of California — Los Angeles	0.17992627069775	0.3808913934	32613.5	60700	1334
4	Harvard University	0.05404574677902	0.25708061	33066	89700	1506
5	Princeton University	0.06521516568269	0.2773972603	37036	74700	1493

Но давайте рассмотрим сам запрос. Обратите внимание, что в предложении `WHERE` выполняется несколько операций приведения типа:

```
SAFE_CAST(ADM_RATE_ALL AS FLOAT64)
```

Если отбросить приведение типа, запрос вернет ошибку:

```
No matching signature for operator > for argument types: STRING, INT64.1
```

Если для приведения к типу вещественных чисел использовать простую функцию `CAST`, она выдаст ошибку в записи со строковым значением в этом поле (`PrivacySuppressed`), которое нельзя преобразовать в число с плавающей точкой:

```
Bad double value: PrivacySuppressed; while executing the filter...2
```

Это связано с тем, что механизм автоматического определения схемы идентифицировал столбец `ADM_RATE_ALL` не как числовой, а как строковый, поскольку в некоторых записях данные закрыты по соображениям конфиденциальности и вместо них вставлен текст `PrivacySuppressed`. В действительности столбец со средним семейным доходом (`MD_FAMINC`) тоже определен как строковый (для колледжей, соответствующих нашим критериям, в нем содержатся только числа), поэтому для преобразования можно использовать простую функцию `CAST`.³

Корректировка схемы

В реальных наборах данных вам неизбежно потребуется выполнить некоторые преобразования и очистку перед загрузкой данных в BigQuery. Далее в этой

¹ Нет соответствующей сигнатуры для оператора `>` и типов аргументов: `STRING`, `INT64`.

² Неверное вещественное число: `PrivacySuppressed`; при выполнении фильтрации...

³ Строки сортируются в лексикографическом порядке. Если данные хранятся как строки, то значение «100» будет меньше, чем «20», по той же причине, по какой «abc» при сортировке помещается перед «de». Когда выполняется численная сортировка, значение 20 окажется меньше 100, как и ожидалось.

главе мы рассмотрим создание более сложных конвейеров обработки данных, однако простой способ заключается в использовании инструментов Unix для замены уведомлений о том, что данные закрыты по соображениям конфиденциальности, на NULL:

```
zless ./college_scorecard.csv.gz | \
  sed 's/PrivacySuppressed/NULL/g' | \
  gzip > /tmp/college_scorecard.csv.gz
```

Здесь мы используем строковый редактор `sed` (string editor — строковый редактор) для замены всех вхождений `PrivacySuppressed` значением `NULL`, сжимаем результат и записываем его во временную папку. После этого вместо оригинального файла можно загрузить очищенный файл.

После загрузки очищенного файла BigQuery правильно определит тип для на-много большего числа столбцов, но `SAT_AVG` или `ADM_RATE_ALL` не попадут в их число; тип этих столбцов по-прежнему определяется как строковый. Это связано с тем, что алгоритм автоматического определения схемы просматривает не все строки в файле, а только некоторую часть. Поскольку в большом количестве строк `SAT_AVG` имеет пустое значение (менее 20% колледжей сообщают балл SAT), алгоритм не смог надежно определить тип поля и пошел по самому безопасному пути, присвоив ему строковый тип.

Поэтому не рекомендуется автоматически определять схему файлов в промышленном окружении — вы будете зависеть от того, какие данные выбраны для анализа. В промышленном окружении указывайте тип данных для столбца во время загрузки.

Механизм автоопределения можно использовать, чтобы не начинать писать схему с нуля. Вот как можно отобразить схему таблицы в текущем ее виде:

```
bq show --format prettyjson --schema ch04.college_scorecard
```

Схему можно сохранить в файл:

```
bq show --format prettyjson --schema ch04.college_scorecard > schema.json
```

Теперь можно открыть файл со схемой в текстовом редакторе (если у вас нет особых предпочтений, воспользуйтесь значком пера в Cloud Shell, чтобы открыть редактор по умолчанию) и изменить тип нужных столбцов. В частности, измените типы четырех столбцов, указанных в предложении `WHERE` (`SAT_AVG`, `ADM_RATE_ALL`, `FIRST_GEN` и `MD_FAMINC`), на `FLOAT64`:

```
{
  "mode": "NULLABLE",
  "name": "FIRST_GEN",
  "type": "FLOAT64"
},
```

Также измените (пока) тип столбца T4APPROVALDATE на STRING, потому что он имеет нестандартный формат даты:¹

```
{
  "mode": "NULLABLE",
  "name": "T4APPROVALDATE",
  "type": "STRING"
},
```

После корректировки можно загрузить данные с обновленной схемой:

```
bq --location=US \
  load --null_marker=NULL --replace \
  --source_format=CSV \
  --schema=schema.json --skip_leading_rows=1 \
  ch04.college_scorecard \
  ./college_scorecard.csv.gz
```

Поскольку мы передаем свою схему, нужно сообщить BigQuery, что она должна игнорировать первую строку в файле CSV (с именами столбцов).

После загрузки можно повторно выполнить запрос из предыдущего раздела:

```
SELECT
  INSTNM
  , ADM_RATE_ALL
  , FIRST_GEN
  , MD_FAMINC
  , MD_EARN_WNE_P10
  , SAT_AVG
FROM
  ch04.college_scorecard
WHERE
  SAT_AVG > 1300
  AND ADM_RATE_ALL < 0.2
  AND FIRST_GEN > 0.1
ORDER BY
  MD_FAMINC ASC
```

Обратите внимание, что теперь, поскольку SAT_AVG, ADM_RATE_ALL больше не являются строковыми столбцами, запрос получился намного чище, так как больше не требуется преобразовывать строки в числа с плавающей точкой. Они больше

¹ Файл содержит даты в формате D/M/YYYY, тогда как стандартным для дат считается формат YYYY-MM-DD (соответствует стандарту ISO 8601). Механизм автоматического определения может просмотреть несколько записей и определить, представляет ли значение 12/11/1965 дату 12 ноября 1965 г. или 11 декабря 1965 г., но крайне нежелательно, чтобы BigQuery делала подобные предположения. Конвейер преобразования, который мы создадим ниже в этой главе, преобразует даты в стандартный формат. А пока давайте просто будем рассматривать эти значения как строки.

не являются строками, потому что мы решили считать недоступными данные, закрытые по соображениям конфиденциальности, в процессе извлечения, преобразования и загрузки (ETL).

АВТОМАТИЗАЦИЯ СОЗДАНИЯ СХЕМЫ

Мы еще не познакомились с метаданными таблиц (мы это сделаем в главе 8), тем не менее вы уже сейчас можете автоматизировать создание схемы, используя язык SQL. Вот запрос для получения схемы всех таблиц в наборе ch04:

```
SELECT
    table_name
    , column_name
    , ordinal_position
    , is_nullable
    , data_type
FROM
    ch04.INFORMATION_SCHEMA.COLUMNS
```

Затем можно использовать функцию `TO_JSON_STRING`, чтобы преобразовать схему в формат JSON и избежать необходимости переходить в командную строку:

```
SELECT
    TO_JSON_STRING(
        ARRAY_AGG(STRUCT(
            IF(is_nullable = 'YES', 'NULLABLE', 'REQUIRED') AS
mode,
            column_name AS name,
            data_type AS type)
        ORDER BY ordinal_position), TRUE) AS schema
FROM
    ch04.INFORMATION_SCHEMA.COLUMNS
WHERE
    table_name = 'college_scorecard'
```

Она возвращает строку JSON в виде:

```
[
  {
    "mode": "NULLABLE",
    "name": "INSTNM",
    "type": "STRING"
  },
  {
    "mode": "NULLABLE",
    "name": "ADM_RATE_ALL",
    "type": "FLOAT64"
  },
  ...
]
```

Копирование в новую таблицу

Загруженная таблица содержит много столбцов, которые нам не нужны. Из исходной таблицы можно создать более специализированную таблицу с помощью инструкции `CREATE TABLE` и заполнить ее только интересующими нас столбцами:

```
CREATE OR REPLACE TABLE ch04.college_scorecard_etl AS
SELECT
  INSTNM
  , ADM_RATE_ALL
  , FIRST_GEN
  , MD_FAMINC
  , SAT_AVG
  , MD_EARN_WNE_P10
FROM ch04.college_scorecard
```

Благодаря использованию надежного конвейера ETL и своевременному принятию решений, последующие запросы будут более понятными и краткими. Однако для организации процесса ETL необходимо приложить дополнительные усилия (определить типы данных и указать схему) и, возможно, принять необратимые решения (например, позже не получится вернуть недоступный столбец, потому что он не отбирался и был скрыт из соображений конфиденциальности или удален). Далее в этой главе мы покажем, как конвейер ELT в SQL может помочь отложить принятие необратимых решений.

Управление данными (DDL и DML)

Зачем описывать управление данными в главе, посвященной загрузке данных? Дело в том, что обычно загрузка данных является лишь частью задачи управления данными. Если данные загружены по ошибке, вам может потребоваться удалить их. Иногда также возникает необходимость удалить данные во исполнение нормативных актов.



Мы советуем вам попробовать все команды и запросы, описанные в этой книге, но не пытайтесь выполнять приведенные в этом разделе, потому что вы потеряете свои данные!

Самую простую возможность полностью удалить таблицу (или представление) дает веб-интерфейс BigQuery. Однако удаление также можно выполнить с помощью утилиты `bq`:

```
bq rm ch04.college_scorecard
bq rm -r -f ch04
```

Первая команда удаляет одну таблицу, а вторая рекурсивно (-r) и без запроса подтверждения (-f, force — принудительно) — набор данных ch04 и всех таблиц в нем.

Также таблицу (или представление) можно удалить с помощью SQL:

```
DROP TABLE IF EXISTS ch04.college_scorecard_gcs
```

Также можно указать определенное время в будущем, когда истечет срок действия таблицы, выполнив инструкцию ALTER TABLE SET OPTIONS:

```
ALTER TABLE ch04.college_scorecard
  SET OPTIONS (
    expiration_timestamp=TIMESTAMP_ADD(CURRENT_TIMESTAMP(),
                                         INTERVAL 7 DAY),
    description="College Scorecard expires seven days from now"
  )
```

Инструкции DROP TABLE и ALTER TABLE, как и инструкция CREATE TABLE, являются примерами инструкций языка определения данных (Data Definition Language, DDL).

Из таблицы можно удалять определенные записи, например:

```
DELETE FROM ch04.college_scorecard
WHERE SAT_AVG IS NULL
```

А с помощью INSERT можно добавить записи в существующую таблицу, вместо того чтобы полностью ее заменять. Например, вот как можно добавить новые значения в таблицу college_scorecard:

```
INSERT ch04.college_scorecard
(INSTNM
 , ADM_RATE_ALL
 , FIRST_GEN
 , MD_FAMINC
 , SAT_AVG
 , MD_EARN_WNE_P10
 )
VALUES ('abc', 0.1, 0.3, 12345, 1234, 23456),
       ('def', 0.2, 0.2, 23451, 1232, 32456)
```

Для копирования значений из одной таблицы в другую можно использовать подзапросы:

```
INSERT ch04.college_scorecard
SELECT *
FROM ch04.college_scorecard_etl
WHERE SAT_AVG IS NULL
```

Инструкции DELETE, INSERT и MERGE являются примерами инструкций языка манипулирования данными (Data Manipulation Language, DML).



На момент написания этой книги BigQuery не поддерживала SQL-инструкцию `COPY`. Если вам понадобится скопировать всю таблицу, используйте команду `bq cp`:

```
bq cp ch04.college_scorecard
somedb.college_scorecard_copy
```

Стоимость этой операции не будет включена в счет, но вам придется дополнительно оплачивать хранение новой таблицы. Команда `bq cp` поддерживает возможность добавления данных в конец (добавьте параметр `-a` или `--append_table`) и замену (добавьте параметр `--noappend_table`) существующей таблицы.

Также можно использовать идиоматический для стандартного SQL способ — инструкцию `CREATE TABLE AS SELECT` или `INSERT VALUES`, в зависимости от того, существует ли таблица, куда должны добавляться данные, или нет. Однако команда `bq cp` действует быстрее (потому что копирует только метаданные таблицы) и выполняется бесплатно.

Эффективная загрузка данных

BigQuery может загружать данные из файлов CSV, однако это не самый эффективный и выразительный способ (например, из файлов CSV нельзя загружать массивы и структуры). Если у вас есть такая возможность, лучше экспортировать данные в другой формат. Но какой?

Наиболее эффективным и выразительным является формат Avro (<https://avro.apache.org/>). Это двоичный формат с самоописанием. Данные в этом формате разбиваются на блоки, и все блоки сжимаются. Благодаря этому можно параллельно загружать данные в формате Avro и экспортировать в этот формат. Кроме того, так как блоки сжимаются, файлы получаются меньше, чем объем данных в них. Теперь о выразительности: формат Avro является иерархическим, может представлять вложенные и повторяющиеся поля и поддерживается платформой BigQuery. Организовать сохранение подобных данных в файлах CSV намного сложнее. Поскольку файлы Avro включают самоописание, при их использовании не требуется указывать схему.

Единственным недостатком формата Avro является его непригодность для чтения человеком. Если для вас важны удобочитаемость и выразительность, используйте для хранения данных строковый формат JSON.¹ Формат JSON поддерживает возможность хранения иерархических данных, но требует представления двоичных данных в кодировке base-64. Однако формат JSON предлагает более широкие возможности, чем CSV, потому что имя каждого поля повторяется в каждой строке.

¹ Строковый формат JSON часто называют форматом jsonl (JSON lines), в котором каждая запись находится в отдельной строке.

Формат Parquet стал поддерживаться в BigQuery относительно недавно. Он основан на оригинальном формате Google Dremel ColumnIO,¹ и так же, как Avro, является двоичным, блочно-ориентированным и компактным и способен представлять иерархические данные. Однако, в отличие от формата Avro, в котором данные хранятся построчно, запись за записью, формат Parquet хранит данные по столбцам. Колоночные файлы лучше подходят для чтения подмножества столбцов, однако в процессе загрузки данных требуется читать все столбцы, поэтому колоночные форматы несколько менее эффективны при загрузке данных. Тем не менее колоночный формат делает Parquet более предпочтительным, чем Avro, для федеративных запросов, о которых мы поговорим позже. Еще один открытый колоночный формат — Optimized Row Columnar (ORC). Формат ORC сравним с Parquet по производительности и эффективности.

Влияние сжатия и перемещения через Google Cloud Storage

При использовании форматов без внутреннего сжатия, таких как CSV и JSON, следует подумать о сжатии файлов с помощью gzip. Сжатые файлы передаются быстрее и занимают меньше места, но они медленнее загружаются в BigQuery. Чем медленнее ваша сеть, тем более выгодно сжатие.

Если вы работаете в медленной сети, или у вас много файлов, или они очень большие, многопоточную загрузку данных можно настроить с помощью `gsutil cp`. После сохранения всех данных в облачном хранилище Google Cloud Storage вы сможете загрузить их в BigQuery оттуда с помощью команды `bq load`:

```
gsutil -m cp *.csv gs://BUCKET/some/location
bqload ... gs://BUCKET/some/location/*.csv
```

В этом эксперименте представлены разные плюсы и минусы, связанные со сжатием и размещением данных с оценочными параметрами колледжей в облачном хранилище Cloud Storage перед вызовом команды `bq load`. В табл. 4.1 приводятся результаты, полученные в ходе эксперимента. Ваши результаты, конечно, будут отличаться и зависеть от характеристик вашей сети и самих загружаемых данных.² Поэтому проведите аналогичные исследования у себя и выберите метод, обеспечивающий наилучшую эффективность.

В случае с размещением файла в облачном хранилище Google Cloud Storage вам придется заплатить за хранение данных, по крайней мере, до завершения задания загрузки в BigQuery. Однако затраты на хранение, как правило, незначительны,

¹ См. https://blog.twitter.com/engineering/en_us/a/2013/dremel-made-simple-with-parquet.html. В главе 6 мы обсудим формат Capacitor — внутренний формат хранилища BigQuery, который является преемником ColumnIO.

² Поэкспериментируйте, запуская сценарии `load_*.sh` в папке `04_load` из репозитория GitHub с примерами для этой книги.

поэтому для этого набора данных и сетевого подключения (см. табл. 4.1) лучшим вариантом является размещение сжатых данных в облачном хранилище с последующей загрузкой их оттуда. Несмотря на то что несжатые файлы загружаются в BigQuery быстрее, время передачи файлов по сети уменьшает все преимущества более быстрой загрузки.

Таблица 4.1. Плюсы и сжатия, и размещения данных с оценочными параметрами колледжей в облачном хранилище Cloud Storage перед вызовом команды `bq load`

Сжатие	Размещение в GCS	Размер GCS	Время передачи по сети	Время загрузки в BigQuery	Общее время
Да	Нет	—	—	105 с	105 с
Нет	Нет	—	—	255 с	255 с
Да	Да	16 Мбайт	47 с	42 с	89 с
Нет	Да	76 Мбайт	139 с	28 с	167 с

На момент написания этой книги размер загружаемых сжатых файлов CSV и JSON ограничивался 4 Гбайт, потому что платформа BigQuery должна распаковывать файлы сразу на рабочих машинах, имеющих ограниченный объем памяти. Если ваши наборы данных имеют больший размер, разделите их на несколько файлов CSV или JSON. Такое разделение может обеспечить некоторый параллелизм при загрузке, но, в зависимости от размеров файлов, размеры таблиц могут получиться не самыми оптимальными, пока BigQuery не решит оптимизировать хранилище.

Цены и квоты

BigQuery не взимает плату за загрузку данных. Прием данных осуществляется рабочими машинами, не входящими в кластер, который предоставляет слоты для запросов. Как результат, скорость выполнения запросов (даже к таблице, куда добавляются новые данные) не снижается из-за того, что данные продолжают поступать в систему.

Загрузка данных выполняется быстро. Запросы к таблице либо будут отражать наличие всех данных, загруженных с помощью операции `bq load`, либо нет. Результаты запроса не могут быть получены по частичному срезу данных.

Недостаток загрузки данных с использованием «бесплатного» кластера заключается в том, что время загрузки может стать непредсказуемым и узким местом из-за уже выполняемых заданий. На момент написания книги количество заданий загрузки ограничивалось 1000 на таблицу и 100 000 на проект в день (https://cloud.google.com/bigquery/quotas#load_jobs). В случае файлов CSV и JSON ячейки

и строки ограничены объемом 100 Мбайт, а для формата Avro размер блоков ограничивался объемом 16 Мбайт. Размер файла не может превышать 5 Тбайт. Если ваш набор данных имеет больший размер, разбейте его на несколько файлов, каждый размером менее 5 Тбайт. Также имейте в виду, что в одном задании загрузки можно передать не больше 15 Тбайт данных, разбитых максимум на 10 миллионов файлов. Задание загрузки должно выполняться не более чем за шесть часов, иначе оно будет отменено.

Федеративные запросы и внешние источники данных

Вы можете использовать BigQuery без предварительной загрузки данных. Данные можно оставить на месте, указать их структуру и использовать BigQuery только как механизм обработки запросов. Выше мы обсуждали запросы, обрабатывая которые BigQuery обращалась к собственному хранилищу, а в этом разделе мы познакомим вас с «федеративными запросами», обращающимися к «внешним источникам данных», и объясним, когда может потребоваться использовать такие запросы.

В настоящее время поддерживаются следующие внешние источники данных: Google Cloud Storage, Cloud Bigtable, Cloud SQL и Google Drive. Как вы узнаете, все эти источники являются внешними по отношению к BigQuery, но находятся в границах Google Cloud. Это обязательное условие, потому что в противном случае сетевые издержки и меры безопасности сделают запросы либо медленными, либо неосуществимыми.

Как использовать федеративные запросы

Запрос данных из внешних источников осуществляется в три этапа:

1. Создание определения таблицы с помощью `bq mkdef`.
2. Создание таблицы с помощью `bq mk` и передача определения внешней таблицы.
3. Выполнение запроса к таблице как обычно.

Так же как в случае, когда данные находятся во внутреннем хранилище, выполнить запрос можно либо через веб-интерфейс, либо с использованием программного интерфейса. В веб-интерфейсе достаточно выполнить перечисленные выше шаги, чтобы создать таблицу, но при этом обязательно нужно указать, что будет использоваться внешняя таблица (External table), а не внутренняя (Native table), как показано на рис. 4.1.

The image shows a web interface for creating a table in BigQuery. On the left, there is a 'Dataset name' field with a dropdown menu currently showing 'ch04'. On the right, there is a 'Table type' section with a question mark icon and two radio button options: 'Native table' and 'External table'. The 'External table' option is selected.

Рис. 4.1. Внешнюю таблицу можно создать в веб-интерфейсе, следуя процедуре «Create Table» («Создание таблицы») и выбрав тип таблицы «External table» («Внешняя таблица»)

Решив использовать интерфейс командной строки, вы должны сначала создать определение таблицы командой `bq mkdef`. Так же как `bq load`, эта команда имеет параметр `--autodetect`:

```
bq mkdef --source_format=CSV \
  --autodetect \
  gs://bigquery-oreilly-book/college_scorecard.csv
```

Эта команда выведет файл определения таблицы в стандартный вывод. Обычно этот вывод перенаправляется в файл, после чего используется для создания таблицы с помощью `bq mk`:

```
bq mkdef --source_format=CSV \
  --autodetect \
  gs://bigquery-oreilly-book/college_scorecard.csv \
  > /tmp/mytable.json
bq mk --external_table_definition=/tmp/mytable.json \
  ch04.college_scorecard
```

После выполнения этих двух этапов можно отправить запрос к таблице `college_scorecard`, в точности как было показано в предыдущем разделе, за исключением того, что запросы будут выполняться к файлу CSV, хранящемуся в облачном хранилище Google Cloud Storage, — данные не будут перемещаться во внутреннее хранилище BigQuery.

Символы подстановки

Многие платформы больших данных, такие как Apache Spark, Apache Beam и другие, распределяют выходные данные по сотням файлов, присваивая им имена, например `course_grades.csv-00095-of-00313`. При загрузке таких файлов было бы удобно иметь возможность не перечислять их по отдельности.

И это возможно: в параметре пути, который передается в команду `bq mkdef` (и `bq load`), можно использовать символы подстановки, чтобы обеспечить соответствие сразу нескольким файлам:

```
bq mkdef --source_format=CSV \
  --autodetect \
  gs://bigquery-oreilly-book/college_* \
  > /tmp/mytable.json
```

Эта команда создаст таблицу на основе всех файлов, имена которых соответствуют шаблону.

Временная таблица

Также можно совместить все три этапа (`mkdef`, `mk` и `query`), передав параметры определения таблицы вместе с запросом при таком подходе определение таблицы будет использоваться только на время запроса:

```
LOC="--location US"
INPUT=gs://bigquery-oreilly-book/college_scorecard.csv

SCHEMA=$(gsutil cat $INPUT | head -1 | awk -F, '{ORS=","}{for (i=1; i <= NF; i++){
print $i":STRING"; }}' | sed 's/,,$//g'| cut -b 4- )

bq $LOC query \
  --external_table_definition=cstable::${SCHEMA}@CSV=${INPUT} \
  'SELECT SUM(IF(SAT_AVG != "NULL", 1, 0))/COUNT(SAT_AVG) FROM cstable'
```

В предыдущем запросе определение внешней таблицы включает имя временной таблицы (`cstable`), два двоеточия, определение схемы, символ `@`, определение формата (CSV), знак равенства и URL файла (или файлов) с данными в Google Cloud Storage. Если файл определения таблицы уже есть, можете указать его:

```
--external_table_definition=cstable::${DEF}
```

Можно указать файл со схемой в формате JSON и выполнять запросы к файлам в формате JSON, Avro и других поддерживаемых форматах, хранящимся в Cloud Storage, Cloud Bigtable и прочих поддерживаемых источниках данных.

Несмотря на определенное удобство, производительность федеративных запросов оставляет желать лучшего. Поскольку файлы CSV хранятся построчно, а сами строки расположены в произвольном порядке, все преимущества эффективности, которая обычно ожидается от BigQuery, теряются. Кроме того, BigQuery не может оценить, сколько данных нужно просканировать перед выполнением запроса.

Загрузка и выполнение запросов к данным в форматах Parquet и ORC

Как уже отмечалось, Parquet и ORC являются колоночными форматами данных. Следовательно, федеративные запросы к данным в этих форматах имеют

лучшую производительность, чем к данным в других, строчных форматах, таких как CSV или JSON (тем не менее запросы будут выполняться медленнее, чем к данным, находящимся в собственном хранилище Capacitor внутри BigQuery).

Поскольку Parquet и ORC имеют встроенное описание (то есть схема неявно определена в самих файлах), при их использовании можно создавать определения таблиц без указания схемы:

```
bq mkdef --source_format=PARQUET gs://bucket/dir/files* > table_def.json
bq mk --external_table_definition=table_def.json <dataset>.<table>
```

Как и в случае с внешними таблицами, созданными из файлов CSV, запрос к этой таблице можно выполнить точно так же, как к любой другой таблице в BigQuery.

Несмотря на то что файлы Parquet и ORC обеспечивают лучшую производительность, чем построчные форматы, для них действуют все те же ограничения для внешних таблиц.

Загрузка и выполнение запросов к разделам Hive

Apache Hive (<https://hive.apache.org/>) позволяет читать данные, записывать их и управлять ими в хранилище на основе Apache Hadoop с использованием знакомого SQL-подобного языка запросов. Cloud Dataproc в Google Cloud позволяет программному обеспечению Hive работать с распределенными данными, хранящимися в разделах Hive в облачном хранилище Google Cloud Storage. Типичным шаблоном миграции в общедоступное облако является перенос локальных рабочих нагрузок Hive в Cloud Dataproc и добавление новых рабочих нагрузок, реализованных с использованием поддержки федеративных запросов в BigQuery. При таком подходе текущие рабочие нагрузки в Hive действуют «как есть», а для новых рабочих нагрузок доступны все преимущества бессерверного и масштабируемого движка запросов, предлагаемого платформой BigQuery.

Чтобы загрузить разделы Hive в Google Cloud Storage, нужно указать режим разбивки на разделы в Hive в команде `bq load`:

```
bq load --source_format=ORC --autodetect \
  --hive_partitioning_mode=AUTO <dataset>.<table> <gcs_uri>
```

Идентификаторы URI таблиц Hive в Cloud Storage должны включать префиксы путей к таблицам с заменой любых ключей разделов подстановочными символами. То есть если ключом раздела для таблицы Hive является поле с именем `datetamp`, URI в Cloud Storage должен иметь следующую форму:

```
gs://some-bucket/some-dir/some-table/*
```

Это верно, даже если пути ко всем файлам начинаются с:

```
gs://some-bucket/some-dir/some-table/datestamp=
```

На момент написания этой книги режим `AUTO` мог определять следующие типы: `STRING`, `INTEGER`, `DATE` и `TIMESTAMP`. Также можно было потребовать, чтобы ключи разделов определялись как строки (это может пригодиться для поиска):

```
bq load --source_format=ORC --autodetect \  
  --hive_partitioning_mode=STRINGS <dataset>.<table> <gcs_uri>
```

Как и в случае с файлами `CSV` в `Google Cloud Storage`, для выполнения федеративных запросов к разделам `Hive` требуется создать файл с определением таблицы и указать те же параметры, что и в команде `load`:

```
bq mkdef --source_format=ORC --autodetect \  
  --hive_partitioning_mode=AUTO <gcs_uri> > table_def.json
```

После создания файла с определением таблицы можно выполнить сам запрос, точно так же, как к данным, хранящимся в файлах `CSV`.

Кроме формата `ORC` поддерживаются также данные в других форматах. Например, создать определение таблицы для данных, хранящихся в строчном формате `JSON`, можно так:

```
bq mkdef --source_format=NEWLINE_DELIMITED_JSON --autodetect --  
hive_partitioning_mode=STRINGS <gcs_uri> <schema> > table_def.json
```

Обратите внимание, что в предыдущей команде автоматически определяются ключи разделов, но не их типы, потому что мы явно указываем, что они должны интерпретироваться как строки, а не как типы данных других столбцов, явно указанные в схеме.

В начале этого раздела мы говорили о том, что запросы к разделам `Hive` часто используются для миграции в облако, когда уже имеются значительные рабочие нагрузки в `Hive`, но при этом можно создавать новые рабочие нагрузки в `BigQuery`. В отличие от фреймворка `Apache Hive`, который поддерживает полный доступ (чтение и запись) к данным, `BigQuery` предоставляет доступ к внешним таблицам только для чтения. Более того, даже при том что `BigQuery` может обрабатывать данные (например, из `Hive`), изменяемые во время выполнения федеративного запроса, в настоящее время она не поддерживает такие концепции, как чтение данных в определенный момент времени. Из-за этих ограничений, накладываемых на внешние таблицы в `BigQuery`, со временем лучше переместить данные в собственное хранилище `BigQuery` и переписать рабочие нагрузки `Hive` в `BigQuery`. Когда данные находятся в собственном хранилище `BigQuery`, появляется возможность использовать язык `DML`, потоковую передачу, кластеризацию, копирование таблиц и многое другое.

Когда использовать федеративные запросы и внешние источники данных

Запросы к внешним источникам выполняются медленнее, чем к данным, хранящимся в `BigQuery`, поэтому в долгосрочной перспективе обычно не реко-

мендуется применять федеративные запросы для часто используемых данных. Однако в некоторых случаях федеративные запросы могут очень пригодиться:

- При выполнении исследований с целью определить, как лучше преобразовать исходные данные перед их загрузкой в BigQuery. Например, подтверждение фактических рабочих нагрузок может определять преобразования в рабочих таблицах. Кроме того, внешние источники данных можно рассматривать как промежуточные и использовать федеративные запросы для преобразования данных и их записи в рабочие таблицы.
- При хранении данных в Google Sheets, если электронная таблица будет редактироваться в интерактивном режиме, и использовании федеративных запросов только тогда, когда результаты этих запросов должны отражать действительные данные в этой электронной таблице.
- При хранении данных во внешнем источнике и когда запросы к ним выполняются нечасто. Например, данные могут храниться в Cloud Bigtable, если эти данные в основном используются для потоковой загрузки большого объема с малой задержкой и большинство запросов к данным могут выполняться с использованием ключевых префиксов.

Для больших, относительно стабильных и хорошо известных наборов данных, которые будут периодически обновляться и часто запрашиваться, собственное хранилище BigQuery — лучший выбор. Далее в разделе мы рассмотрим детали реализации каждой из этих ситуаций, начав с исследований с использованием федеративных запросов.

Исследования с использованием федеративных запросов

Автоопределение схемы — удобная функция. Чтобы выполнить свою функцию, механизм автоопределения выбирает несколько сотен записей из входных файлов и на их основе определяет типы столбцов. Он не защищен от ошибок и может их допускать, если только вы не используете форматы файлов с самоописанием, такие как Avro, Parquet или ORC. Чтобы убедиться, что конвейер ETL работает корректно, вы должны проверить каждую строку и убедиться, что для каждого столбца выбран правильный тип. Например, возможно, что столбец содержит целые числа, за исключением нескольких записей, в которых встречаются числа с плавающей точкой. Если это так, то механизм автоопределения может ошибиться и определить тип столбца как целочисленный, поскольку вероятность выбора одной из строк, содержащих значение с плавающей точкой, довольно мала. Вы не узнаете о проблеме, пока не выполните запрос, который выполнит сканирование значения в этом столбце.

Лучше всего использовать форматы файлов с самоописанием — в этом случае вам не придется беспокоиться о том, как BigQuery интерпретирует данные. Если без файлов в формате CSV или JSON не обойтись, рекомендуем вам явно

указать схему. Схему можно указать в сопроводительном файле JSON, а также передать ее в командной строке `bq mkdef` в виде строки в следующем формате:

```
FIELD1:DATATYPE1, FIELD2:DATATYPE2, ...
```

Если вы не уверены в качестве своих данных, укажите для всех столбцов тип `STRING`. Обратите внимание, что этот тип данных используется по умолчанию, поэтому описание схемы в командной строке упрощается до

```
FIELD1, FIELD2, FIELD3, ...
```

Но зачем интерпретировать все данные как строки? Даже если вы считаете, что некоторые поля являются целочисленными, а другие содержат числа с плавающей точкой, это предположение лучше проверить практически. Определите все поля как строковые и узнайте, какие преобразования необходимо выполнить при запросе данных, и вы обнаружите ошибки.

Из первой строки в файле CSV можно извлечь имена столбцов и создать строку схемы нужного формата:¹

```
INPUT=gs://bigquery-oreilly-book/college_scorecard.csv
SCHEMA=$(gsutil cat $INPUT | head -1 | cut -b 4- )
```

Решив указать схему, мы должны потребовать пропустить первую строку и решить пустые строки в файле. Сделать это можно, пропустив определение таблицы через строчный редактор `sed`:²

```
LOC="--location US"
OUTPUT=/tmp/college_scorecard_def.json
bq $LOC \
  mkdef \
    --source_format=CSV \
    --noautodetect \
    $INPUT \
    $SCHEMA \
    | sed 's/"skipLeadingRows": 0/"skipLeadingRows": 1/g' \
    | sed 's/"allowJaggedRows": false/"allowJaggedRows": true/g' \
    > $OUTPUT
```

Мы указали, что работаем в США и хотим сохранить вывод (определение таблицы) в папку `/tmp`.

На данный момент у нас есть таблица, которую мы можем запросить. Обратите внимание на два обстоятельства: эта таблица определена во внешнем источнике

¹ Первый символ в этом конкретном файле — «маркер порядка следования байтов» (`\u00eff`), поэтому мы удаляем первые несколько байтов с помощью команды `cut: cut -b 4-`.

² Полный сценарий называется `load_external_gcs.sh` и находится в репозитории GitHub с примерами для этой книги.

данных, поэтому мы можем начать запрашивать данные, не ожидая их поступления; и все столбцы являются строковыми — мы не внесли никаких необратимых изменений в исходные данные.

А теперь приступим к исследованию данных и для начала попробуем выполнить приведение типа:

```
SELECT
  MAX(CAST(SAT_AVG AS FLOAT64)) AS MAX_SAT_AVG
FROM
  `ch04.college_scorecard_gcs`
```

Запрос выдаст следующее сообщение об ошибке:

```
Bad double value: NULL1
```

Это явно указывает, что нам нужно обработать нестандартный способ кодирования отсутствующих данных в файле. В большинстве файлов CSV отсутствующие данные кодируются как пустая строка, но в этом примере они кодируются строкой NULL.

Исправить проблему можно, выполнив проверку перед приведением типа:

```
WITH etl_data AS (
  SELECT
    SAFE_CAST(SAT_AVG AS FLOAT64) AS SAT_AVG
  FROM
    `ch04.college_scorecard_gcs`
)
SELECT
  MAX(SAT_AVG) AS MAX_SAT_AVG
FROM
  etl_data
```

Обратите внимание, что мы начали с предложения WITH, содержащего все операции ETL, которые необходимо выполнить с набором данных. Теперь, немного познакомившись с данными и попробовав выполнить предыдущий запрос, мы поняли, что нам нужна многократно используемая функция для очистки числовых данных:

```
CREATE TEMP FUNCTION cleanup_numeric(x STRING) AS
(
  IF ( x != 'NULL' AND x != 'PrivacySuppressed',
      CAST(x as FLOAT64),
      NULL )
);

WITH etl_data AS (
  SELECT
```

¹ Неверное значение типа double: NULL. — *Примеч. пер.*


```

INSTNM
, cleanup_numeric(ADM_RATE_ALL) AS ADM_RATE_ALL
, cleanup_numeric(FIRST_GEN) AS FIRST_GEN
, cleanup_numeric(MD_FAMINC) AS MD_FAMINC
, cleanup_numeric(SAT_AVG) AS SAT_AVG
, cleanup_numeric(MD_EARN_WNE_P10) AS MD_EARN_WNE_P10
FROM
`ch04.college_scorecard_gcs`
)

SELECT
*
FROM
etl_data
WHERE
SAT_AVG > 1300
AND ADM_RATE_ALL < 0.2
AND FIRST_GEN > 0.1
ORDER BY
MD_FAMINC ASC
LIMIT 10

```

Теперь можно экспортировать очищенные данные (обратите внимание на SELECT *) в новую таблицу (обратите внимание на CREATE TABLE), выполнив следующий запрос:

```

CREATE TEMP FUNCTION cleanup_numeric(x STRING) AS
(
  IF ( x != 'NULL' AND x != 'PrivacySuppressed',
      CAST(x as FLOAT64),
      NULL )
);

CREATE TABLE ch04.college_scorecard_etl
OPTIONS(description="Cleaned up college scorecard data") AS

WITH etl_data AS (
  SELECT
    INSTNM
    , cleanup_numeric(ADM_RATE_ALL) AS ADM_RATE_ALL
    , cleanup_numeric(FIRST_GEN) AS FIRST_GEN
    , cleanup_numeric(MD_FAMINC) AS MD_FAMINC
    , cleanup_numeric(SAT_AVG) AS SAT_AVG
    , cleanup_numeric(MD_EARN_WNE_P10) AS MD_EARN_WNE_P10
  FROM
    `ch04.college_scorecard_gcs`
)

SELECT * FROM etl_data

```

Этот код также можно перенести в сценарий, удалив оператор CREATE TABLE, выполнить запрос с помощью команды bq query и передать ей параметр --destination_table.

ELT в SQL для экспериментов

Во многих организациях аналитиков данных гораздо больше, чем инженеров. Поэтому потребности групп, занимающихся анализом данных, часто существенно превосходят возможности инженеров по доставке этих данных. В таких случаях может быть полезно умение аналитиков самостоятельно создавать экспериментальные наборы данных в BigQuery, чтобы решать свои задачи.

При таком подходе организация может использовать информацию о фактических рабочих нагрузках для определения приоритетных направлений, на которых должны сосредоточиться инженеры. Например, как инженер данных, вы, возможно, еще не знаете, какие поля нужно извлекать из файла журнала. Поэтому вы можете настроить внешний источник данных для экспериментов и позволить аналитикам напрямую запрашивать исходные данные из Google Cloud Storage.

Если исходные файлы журналов хранятся в формате JSON, причем каждая запись имеет свою структуру, из-за того что журналы поступают из разных приложений, аналитики могут определить сообщения в журнале как один строковый столбец и использовать `JSON_EXTRACT` и функции манипулирования строками для извлечения необходимых данных. В конце месяца вы можете проанализировать журналы запросов BigQuery и выяснить, к каким полям в действительности выполнялись обращения и как производились такие обращения, а затем построить конвейер для загрузки этих полей в BigQuery.

Например, вы можете экспортировать журналы аудита BigQuery из Stackdriver в формате JSON со всеми журнальными сообщениями во вложенном столбце с именем `protopayload_auditlog.metadataJson`. Вот запрос для подсчета журнальных сообщений с корневым элементом `tableDataRead` и использования счетчика для ранжирования наборов данных по количеству обращений к каждому:

```
SELECT
  REGEXP_EXTRACT(protopayload_auditlog.resourceName,
    '^projects/[^\s]+/datasets/([^\s]+)/tables') AS datasetRef,
  COUNTIF(JSON_EXTRACT(protopayload_auditlog.metadataJson, "$.tableDataRead")
    IS NOT NULL) AS dataReadEvents,
FROM `ch04.cloudaudit_googleapis_com_data_access_2019*`
WHERE
  JSON_EXTRACT(protopayload_auditlog.metadataJson, "$.tableDataRead")
    IS NOT NULL
GROUP BY datasetRef
ORDER BY dataReadEvents DESC
LIMIT 5
```

Метод `JSON_EXTRACT` принимает имя столбца (`protopayload_auditlog.metadataJson`) в первом параметре и `JSONPath`¹ во втором.

¹ Описание грамматики `JSONPath` можно найти по адресу <https://restfulapi.net/json-jsonpath/>.

Если исходные данные находятся в системе управления реляционными базами данных (СУРБД), их можно периодически экспортировать в файл со значениями, разделенными табуляцией (Tab-Separated Values, TSV), и пересылать его в Google Cloud Storage. Например, вот как будет выглядеть соответствующая команда, если вы используете базу данных MySQL с именем `somedb`:

```
mysql somedb < select_data.sql | \  
gsutil cp - gs://BUCKET/data_$(date -u "+%F-%T").tsv
```

Файл `select_data.sql` мог бы содержать запрос, извлекающий наиболее свежие записи (в данном случае за последние 10 дней):

```
select * from my_table  
where transaction_date >= DATE_SUB(CURDATE(), INTERVAL 10 DAY)
```

При наличии таких периодически экспортируемых файлов аналитик легко может начать запрашивать данные с помощью федеративных запросов. После того как ценность набора данных будет подтверждена, можно организовать периодическую и/или потоковую загрузку с использованием конвейера передачи данных.

Однако на практике такое решение подходит не всегда, потому что оно не предполагает возможности изменения информации в базе данных. Если данные, которым больше 10 дней, обновятся, дампы TSV не будут синхронизированы. В реальной жизни передача данных в файлах TSV целесообразна только для небольших наборов данных (порядка нескольких гигабайтов), когда поля оригинальной базы данных не требуют преобразования или исправления перед их использованием для аналитических запросов.

Для тех, кому требуется организовать синхронизацию оперативной базы данных и BigQuery, существует несколько сторонних компаний, сотрудничающих с Google, каждая из которых имеет комплекс коннекторов и вариантов преобразования.¹ Эти инструменты поддерживают возможность выборки изменившихся данных (Change Data Capture, CDC) и передачи изменений из базы данных в таблицу BigQuery.

Внешние запросы в Cloud SQL

В настоящее время BigQuery поддерживает не только федеративные, но и внешние запросы. Федеративный запрос позволяет запрашивать внешний источник данных с помощью BigQuery, в то время как внешний позволяет выполнить запрос во внешней базе данных и объединить результаты с данными из BigQuery.

¹ К числу этих компаний относятся Aloomo, Informatica и Talend. Полный и актуальный список партнеров BigQuery можно найти на странице <https://cloud.google.com/bigquery/partners/>.

В настоящее время поддерживаются базы данных MySQL и Postgres в Cloud SQL (сервис управляемых реляционных баз данных в Google Cloud).

В BigQuery есть возможность предварительно настроить ресурс подключения и предоставить пользователям разрешение на его использование. После этого ресурс подключения можно использовать для вызова функции `EXTERNAL_QUERY`:

```
SELECT * FROM EXTERNAL_QUERY(connection_id, cloud_sql_query);
```

В этом примере `connection_id` — имя ресурса подключения к базе данных, созданного в BigQuery с помощью веб-интерфейса, REST API или инструмента командной строки.

Производительность внешних запросов зависит от скорости внешней базы данных, и поскольку в работу вовлекается промежуточная временная таблица, внешние запросы действуют обычно медленнее, чем запросы, которые выполняются исключительно в Cloud SQL или в BigQuery. Тем не менее возможность запрашивать данные, размещенные в СУРБД, в режиме реального времени без их перемещения может быть очень полезна, потому что позволяет избежать создания ненужных ETL, планирования и управления.

Допустим, нам понадобилось создать отчет о подарочных картах клиентов, не совершавших в последнее время каких-либо покупок. Дата последнего заказа, сделанного каждым клиентом, доступна в Cloud SQL и обновляется в режиме реального времени. Баланс каждой подарочной карты, выпущенной магазином, доступен в BigQuery. Мы можем объединить результат внешнего запроса с данными о заказах в Cloud SQL с данными о балансах подарочных карт в BigQuery и создать актуальный отчет, не перемещая никаких данных:

```
SELECT
  c.customer_id
  , c.gift_card_balance
  , rq.latest_order_date
FROM ch04.gift_cards AS c
LEFT OUTER JOIN EXTERNAL_QUERY(
  'connection_id',
  '''SELECT customer_id, MAX(order_date) AS latest_order_date
  FROM orders
  GROUP BY customer_id''' ) AS rq ON rq.customer_id = c.customer_id
WHERE c.gift_card_balance > 100
ORDER BY rq.latest_order_date ASC;
```

Интерактивное исследование и запрос данных из Google Sheets

Google Sheets — это часть G Suite, набора инструментов, способствующих увеличению производительности и совместной работе, созданный в Google Cloud. Google Sheets предлагает средства для создания, просмотра, редактирования

и публикации электронных таблиц. Электронная таблица содержит табличные значения в отдельных ячейках; некоторые из этих значений являются данными, а некоторые — результатом вычислений на основе значений из других ячеек. Google Sheets вывела электронные таблицы в онлайн — несколько человек могут одновременно править одну и ту же электронную таблицу, а доступ к такой таблице можно получить с различных устройств.

Загрузка данных из Google Sheets в BigQuery

Google Sheets — это внешний источник данных, поэтому загрузка и извлечение данных из электронной таблицы Google Sheets выполняются как федеративные запросы; они действуют подобно федеративным запросам к файлам CSV из Google Cloud Storage. Мы создаем определение таблицы в BigQuery, ссылающееся на данные в Google Sheets, а затем выполняем запросы к этой таблице, как если бы она была собственной таблицей BigQuery.

Чтобы проиллюстрировать вышесказанное на примере, создадим электронную таблицу Google Sheets и выполним запрос к ней. Откройте в веб-браузере страницу <https://sheets.new>. Вы увидите пустую электронную таблицу.

Введите следующие данные (или загрузите соответствующий файл CSV (<https://oreil.ly/ckBA5>) из GitHub и выберите в меню электронной таблицы пункт **File** ► **Import** (Файл ► Импорт), чтобы импортировать данные в Google Sheets):

Student	Home state	SAT score
Aarti	KS	1111
Billy	LA	1222
Cao	MT	1333
Dalia	NE	1444

Затем в GCP Cloud Console перейдите в раздел BigQuery, создайте набор данных (если необходимо) и таблицу, выбрав источник данных Drive (Google-диск) и указав его URL, а также то, что это электронная таблица Google Sheet. Запросите автоматическое определение схемы, как показано на рис. 4.2.

После этого вы сможете выполнить запрос к электронной таблице как к любой другой таблице BigQuery:

```
SELECT * from advdata.students
```

Попробуйте изменить данные в электронной таблице, и вы убедитесь, что возвращаемые результаты отражают текущее состояние таблицы (результаты федеративных запросов к внешним наборам данных не кешируются).

Create table

Source

Create table from: Drive Select Drive URI: https://docs.google.com/spreadsheets/d/1Mac0a_CblRnD2BjwYG File format: Google Sh...

Destination

Project name: cloud-training-demos Dataset name: advdata Table type: External table

Table name: students

Schema

Auto detect
☒ Schema and input parameters

Schema will be automatically generated.

Advanced options ▼

Create table Cancel

Рис. 4.2. Диалог «Create table» (Создать таблицу) позволяет выбрать в качестве внешнего источника данных электронную таблицу из Google Sheets

Несмотря на возможность выполнять SQL-запросы к электронной таблице, как было показано выше, вы вряд ли захотите это делать, потому что обычно удобнее использовать средства интерактивной фильтрации и сортировки, встроенные в Google Sheets. Например, можно кликнуть на кнопке **Explore** (Анализ данных) и ввести запрос на естественном языке «average SAT score of students in KS» («средний балл SAT студентов из Канзаса»), который вернет результаты, показанные на рис. 4.3.

Есть несколько вариантов использования интеграции Google Sheets и BigQuery:

- заполнение электронных таблиц данными из BigQuery;
- исследование содержимого таблиц BigQuery с помощью Sheets;
- запрос данных из Sheets с использованием SQL.

Рассмотрим все эти варианты по порядку.

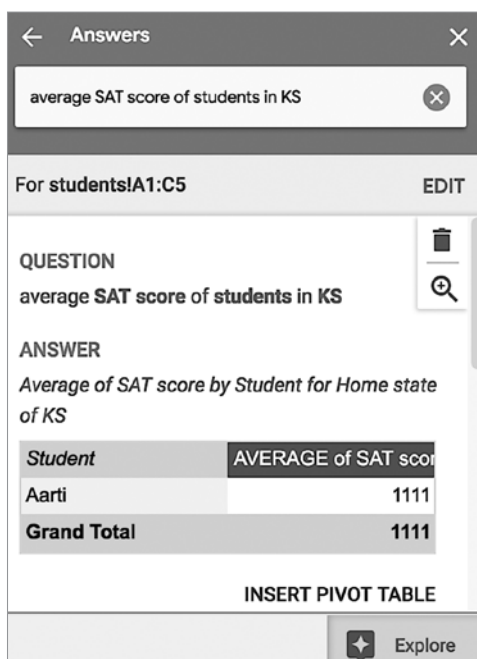


Рис. 4.3. Запрос на естественном языке в Google Sheets

Заполнение электронных таблиц данными из BigQuery

Ресурс подключения к BigQuery в Google Sheets позволяет запрашивать таблицы BigQuery¹ и заполнять электронные таблицы результатами запросов. Это может очень пригодиться для обмена данными с нетехническими пользователями. В большинстве организаций почти все офисные работники знают, как читать/интерпретировать электронные таблицы. Им не нужно быть в курсе всех тонкостей BigQuery или SQL, чтобы использовать Google Sheets и работать с данными на листе.

В Google Sheets выберите пункт меню **Data** ▶ **Data Connectors** ▶ **BigQuery** (**Данные** ▶ **Подключения к данным** ▶ **BigQuery**), выберите свой проект и напишите запрос, который заполнит электронную таблицу данными из таблицы BigQuery с оценочными параметрами колледжей:

```
SELECT
  *
FROM
  ch04.college_scorecard_etl
```

¹ На момент написания этой книги существовали ограничения на размер таблиц BigQuery.

Исследование содержимого таблиц BigQuery с помощью Sheets

Одна из причин, по которым вы можете захотеть заполнить электронную таблицу Google Sheets данными из таблицы BigQuery, — интерфейс, привычный для бизнес-пользователей, которые умеют создавать диаграммы, формулы и сводные таблицы. Например, на основе данных с оценочными параметрами колледжей очень просто создать формулу для их ранжирования по увеличению среднего дохода выпускников:

1. В новом столбце введите формулу:

```
=ArrayFormula(IF(ISBLANK(D2:D), 0, F2:F/D2:D))
```

Обратите внимание, что после этого он заполнится отношениями значений в столбце F к значениям в столбце D, то есть увеличением дохода.

2. В меню **Data** (Данные) создайте фильтр для нового столбца и отключите учет пустых и нулевых значений.
3. Отсортируйте электронную таблицу в порядке уменьшения значений в этом столбце.

Выбрав первые несколько строк записей, можно быстро создать диаграмму, демонстрирующую лучшие колледжи с точки зрения улучшения экономического положения учащихся, как показано на рис. 4.4.

Помимо интерактивных средств создания диаграмм, Google Sheets предлагает также возможность применить технологии машинного обучения для дальнейшего изучения данных.

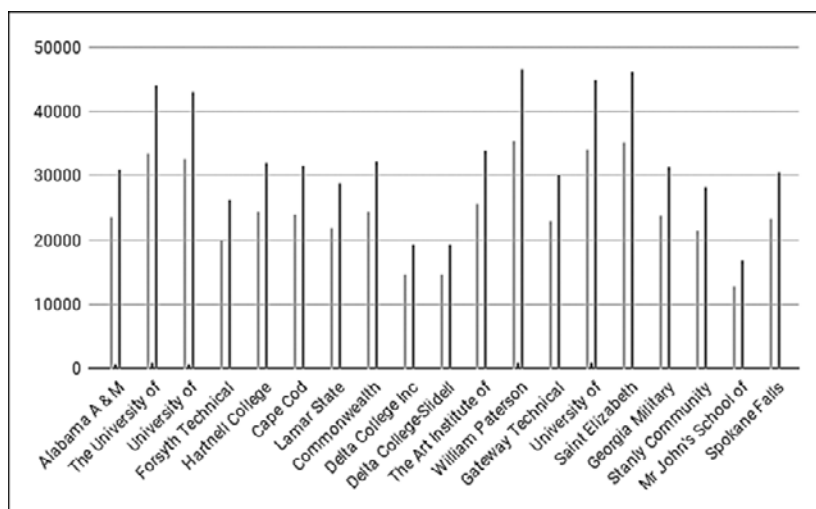


Рис. 4.4. Диаграмма с колледжами, окончание которых обеспечивает наибольшее улучшение экономического положения

В Google Sheets кликните на **Explore** (Анализ данных) и обратите внимание на диаграммы, созданные автоматически с помощью механизмов машинного обучения.¹ Например, автоматически сгенерированная диаграмма, показанная на рис. 4.5, отражает поразительное неравенство.

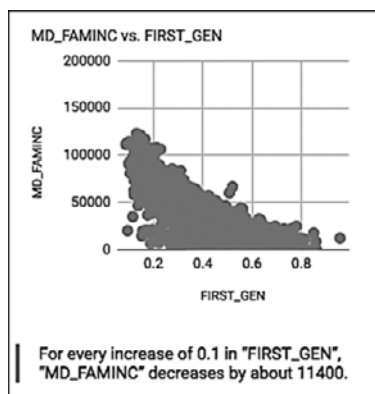


Рис. 4.5. Google Sheets автоматически генерирует диаграмму, позволяющую понять, что колледжи с наибольшей долей студентов первого поколения также имеют наибольшую долю студентов из небогатых семей; на каждые 10% увеличения числа учащихся первого поколения средний доход семьи уменьшается на 11 400 долларов

На рис. 4.6 показана диаграмма, созданная автоматически, которая иллюстрирует соотношение между средним баллом SAT и долей студентов первого поколения.

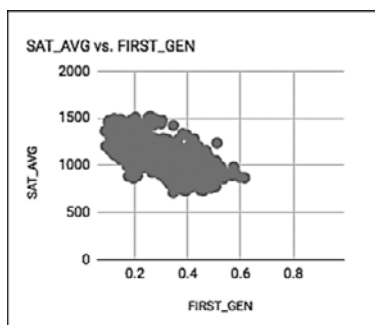


Рис. 4.6. Колледжи, имеющие наибольшую долю студентов первого поколения, как правило, имеют более низкий средний балл SAT

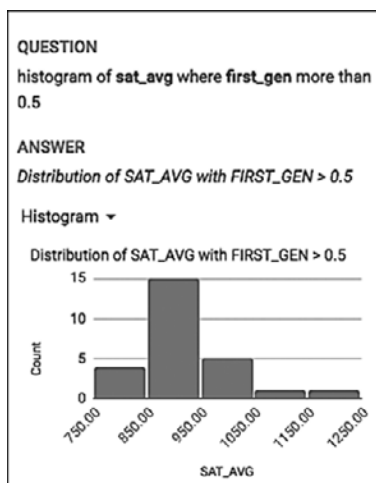


Рис. 4.7. В Google Sheets можно получить нужные графики, написав запрос на естественном языке

¹ Из-за постоянных изменений и улучшений в продуктах вы можете увидеть несколько иные графики.

Можно даже отправить запрос и создать диаграмму на естественном языке. Введите «histogram of sat_avg where first_gen more than 0.5» (гистограмма по sat_avg, где first_gen больше 0.5) в поле **Ask a question** (Задать вопрос), и вы получите ответ, показанный на рис. 4.7.

Исследование содержимого таблиц BigQuery в Google Sheets

В предыдущем разделе мы загрузили в Google Sheets всю таблицу BigQuery, но эта опция была доступна только потому, что наш набор данных с оценочными параметрами колледжей был небольшим. Загрузить большую таблицу BigQuery в Google Sheets невозможно.

Google Sheets позволяет извлекать, анализировать, визуализировать большие наборы данных, а также обмениваться ими даже в форме BigQuery Data Sheet. Чтобы опробовать эту возможность, создайте новый документ Google Sheets и выберите в меню пункт **Data ▸ Data Connectors ▸ BigQuery Data Sheet** (Данные ▸ Подключения к данным ▸ BigQuery Data Sheet).

Выберите свой облачный проект (через который осуществляется оплата услуг) и перейдите через меню к таблице в Data Sheet, которую хотите загрузить, выбрав **bigquery-public-data ▸ usa_names ▸ usa_1910_current ▸ Connect**. Эта таблица содержит почти шесть миллионов строк и слишком велика, чтобы ее можно было загрузить целиком. В таких случаях BigQuery выступает в роли хранилища данных, отображаемых в Google Sheets.

В отличие от варианта, когда в Google Sheets загружается вся таблица (как в предыдущем разделе), из Data Sheet в пользовательский интерфейс загружаются только первые 500 строк. Их можно считать ознакомительным фрагментом набора данных. Другое отличие заключается в возможности редактирования: когда в Google Sheets загружена вся таблица, вы работаете с копией данных и можете править содержимое ячеек и сохранять измененную электронную таблицу. Но когда данные фактически хранятся в BigQuery, ячейки недоступны для редактирования — пользователи могут фильтровать и рассматривать BigQuery Data Sheet с разных сторон, но не могут править данные. Когда пользователи выполняют фильтрацию или пытаются получить сводную информацию, эти действия применяются ко всей таблице BigQuery, а не только к ознакомительному фрагменту, который отображается в Sheets.

Для примера давайте создадим сводную таблицу, кликнув на **Pivot table** (Сводная таблица). В редакторе сводной таблицы выберите поле **state** в списке **Rows** (Строки) и **year** в списке **Columns** (Столбцы). В списке **Values** (Значения) выберите поле **number** и попросите Sheets суммировать по **COUNTUNIQUE** и отобразить в формате **Default** (по умолчанию), как показано на рис. 4.8.

Как видно на рис. 4.8, мы получили таблицу количества уникальных имен детей в каждом штате с разбивкой по годам.

	A	DA	DB	DC	DD
1	COUNTUNIQUE				
2	state	2013	2014	2015	2016
3	AK	43	42	42	36
4	AL	157	147	148	152
5	AR	112	114	111	101
6	AZ	197	200	195	191
7	CA	517	533	513	517
8	CO	165	164	169	161
9	CT	121	125	122	122
10	DC	52	50	51	54
11	DE	48	49	44	49
12	FL	323	333	331	341
13	GA	245	250	241	238
14	HI	53	51	51	50
15	IA	122	122	122	111
16	ID	75	75	72	73
17	IL	278	273	283	274
18	IN	193	203	202	197
19	KS	114	116	114	115
20	KY	158	154	152	149
21		160	155	153	150
22		183	183	192	187
23	MD	175	176	171	172
24	ME	60	57	60	59

Рис. 4.8. Создание сводной таблицы на основе данных из BigQuery Data Sheet

Соединение данных из Google Sheets с большими наборами из BigQuery

BigQuery и Google Sheets способны хранить и предоставлять доступ к табличным данным. Однако BigQuery — это прежде всего хранилище аналитических данных, а Google Sheets в первую очередь является интерактивным документом. Как было показано в предыдущих разделах, знакомство с Sheets, возможностями исследования и построения диаграмм делает загрузку данных BigQuery в Sheets очень функциональной возможностью.

Однако существует техническое ограничение на размер наборов данных BigQuery, которые можно загрузить в Sheets. Например, в BigQuery хранится информация о вопросах, ответах и пользователях сайта Stack Overflow. Даже если использовать BigSheets, эти петабайтные наборы данных слишком велики,

чтобы их можно было загрузить непосредственно в Google Sheets. Однако есть возможность писать запросы, соединяющие небольшие наборы данных в Sheets с гигантскими наборами данных в BigQuery и обрабатывающие результат. Проиллюстрируем эту возможность на примере.

В предыдущем разделе мы создали электронную таблицу с оценочными параметрами колледжей. Предположим, что у нас еще нет данных в BigQuery. Мы могли бы создать таблицу в BigQuery, используя электронную таблицу в качестве источника, и присвоить итоговой таблице имя `college_scorecard_gs`, как показано на рис. 4.9.

Рис. 4.9. Создание таблицы в BigQuery с помощью электронной таблицы Google Sheets в качестве источника

Теперь можно выполнить запрос в BigQuery, который объединит эту сравнительно небольшую таблицу (7700 строк) с массивной таблицей, содержащей данные с сайта Stack Overflow (10 миллионов строк), чтобы выяснить, какие колледжи чаще всего встречаются в профилях пользователей Stack Overflow:

```
SELECT INSTNM, COUNT(display_name) AS numusers
FROM `bigquery-public-data`.stackoverflow.users, ch04.college_scorecard_gs
WHERE REGEXP_CONTAINS(about_me, INSTNM)
GROUP BY INSTNM
ORDER BY numusers DESC
LIMIT 5
```

Этот запрос вернет следующие результаты:¹

Row	INSTNM	numusers
1	Institute of Technology	2364
2	National University	332
3	Carnegie Mellon University	169
4	Stanford University	139
5	University of Maryland	131

Первые две строки выглядят подозрительными,² а представительство университетов Карнеги-Меллон и Стэнфорд на Stack Overflow выглядит вполне достоверным.

Этот запрос возвращает достаточно небольшой набор данных, чтобы его можно было загрузить непосредственно в Google Sheets, выполнить интерактивную фильтрацию и построить диаграммы. Как видите, возможность SQL-запросов данных Sheets из BigQuery можно с успехом использовать для соединения небольшого, формируемого человеком (в Google Sheets) набора данных с большими корпоративными наборами данных (в BigQuery).

Запросы SQL для выборки данных из Cloud Bigtable

Cloud Bigtable — это сервис управляемых баз данных NoSQL, способный хранить петабайты данных. Cloud Bigtable предназначен для использования в ситуациях, когда важна низкая задержка (порядка миллисекунд), высокая пропускная способность (миллионы операций в секунду), репликация для обеспечения высокой доступности и бесшовная масштабируемость (от гигабайтов до петабайтов). Благодаря своим свойствам, Cloud Bigtable нашел широкое применение в финансовой сфере (сверка расчетов и анализ торговых операций, обнаружение мошенничества с платежами и т. д.), в приложениях интернета вещей (Internet of Things, IoT — для централизованного хранения и обработки данных, поступающих от датчиков в масштабе реального времени) и в рекламе (торги, размещение и поведенческий анализ в масштабе реального времени). Хотя сам сервис Cloud Bigtable доступен только в GCP, он поддерживает Apache HBase API с открытым исходным кодом, что позволяет легко переносить рабочие нагрузки в гибридную облачную среду.

¹ Это медленный запрос, потому что выполняет сопоставление с регулярным выражением 77 миллиардов раз.

² Скорее всего, строки представляют несколько колледжей, таких как Национальный университет Сингапура (National University of Singapore), Национальный университет Ирландии (National University of Ireland), Массачусетский технологический институт (Massachusetts Institute of Technology), Технологический институт Джорджии (Georgia Institute of Technology) и т. д.

Запросы NoSQL по ключевым префиксам

Сервис Cloud Bigtable предлагает высокопроизводительную поддержку запросов для поиска записей или их наборов, соответствующих определенному ключу-строке, префиксу ключа или диапазону префиксов. Cloud Bigtable требует наличия в вашем проекте экземпляра, состоящего из одного или нескольких логических кластеров, но использует этот кластер только для вычислений (а не для хранения) — сами данные хранятся в Colossus, а узлы должны знать лишь о расположении диапазонов записей в Colossus. Поскольку данные не хранятся на узлах Cloud Bigtable, кластер Cloud Bigtable легко можно масштабировать вверх и вниз без выполнения дорогостоящей миграции данных.

В финансовом анализе обычным явлением является сохранение временных рядов в Cloud Bigtable по мере их поступления в масштабе реального времени и поддержка запросов с малой задержкой к этим данным с использованием ключевой строки (например, все запросы на покупку акций GOOG за последние 10 минут, если они были). Это позволяет создавать информационные панели (дашборды), выбирающие последние данные, автоматически формирующие оповещения и выполняющие некоторые действия, основываясь на недавних событиях. Cloud Bigtable также позволяет быстро получить диапазон данных (например, все заказы на покупку акций GOOG в любой день), что необходимо для финансовой аналитики и отчетности. Сами алгоритмы прогнозирования должны быть обучены на архивных данных (например, на временных рядах стоимости акций GOOG за последние пять лет), и это возможно, потому что фреймворки машинного обучения, такие как TensorFlow, могут считывать и записывать данные как из, так и в Cloud Bigtable. Эти три рабочие нагрузки (оповещение в режиме реального времени, создание отчетов и машинное обучение) могут выполняться с одними и теми же данными, причем кластер может увеличиваться и уменьшаться с изменением рабочей нагрузки из-за отделения вычислений и хранения.

Все три перечисленные рабочие нагрузки связаны с получением цен на акции Google. Cloud Bigtable обеспечит эффективный поиск записей, если ключ, с которым хранятся данные из временного ряда, имеет форму `GOOG#buy#20190119-090356.0322234`, то есть включает имя безопасности и отметку времени. Затем все запросы, и за предыдущие 10 минут, и за последние пять лет, будут связаны с запросом записей, попадающих в диапазон префиксов.

А если мы захотим выполнить специальный анализ всех данных в Cloud Bigtable и для удовлетворения нашего запроса потребуется извлечь не только подмножество записей, то есть если наш запрос не является фильтром по префиксу ключа? Тогда парадигма NoSQL в Cloud Bigtable будет неприменима и нам лучше прибегнуть к специальным возможностям SQL-запросов, которые предлагает BigQuery, учитывая при этом, что результаты будут возвращены с более высокой задержкой.

Специальные SQL-запросы к данным в Cloud Bigtable

Как мы уже знаем, BigQuery может напрямую извлекать данные из файлов в определенных форматах (CSV, Avro и т. д.), хранящихся в Google Cloud Storage, интерпретируя их как внешние источники данных, однако точно так же BigQuery может напрямую извлекать данные из Cloud Bigtable. Так же как в случае с Cloud Storage, извлекать данные из Cloud Bigtable можно с использованием постоянной или временной таблицы. Постоянная таблица может использоваться вместе с набором данных, частью которого она является; временная таблица действительна только на период обработки запроса и поэтому не может применяться для других целей.

Таблица из Cloud Bigtable отображается на таблицу в BigQuery. В этом разделе для иллюстрации мы используем временные ряды с данными о продажах. Запустите сценарий `setup_data.sh` (https://github.com/GoogleCloudPlatform/bigquery-oreilly-book/tree/master/04_load/bigtable), который можно найти в репозитории GitHub с примерами для этой книги, чтобы создать экземпляр Cloud Bigtable, заполненный некоторыми демонстрационными данными. Этот сценарий создаст экземпляр Cloud Bigtable с кластером, поэтому не забудьте удалить экземпляр, когда закончите.

Сначала используем веб-интерфейс BigQuery, чтобы создать внешнюю таблицу в BigQuery, ссылающуюся на данные в Cloud Bigtable, как показано на рис. 4.10. Местоположение определяется строкой в формате `https://googleapis.com/bigtable/projects/[PROJECT_ID]/instances/[INSTANCE_ID]/tables/[TABLE_NAME]`, где `PROJECT_ID`, `INSTANCE_ID` и `TABLE_NAME` определяют проект, экземпляр и таблицу в Cloud Bigtable.¹

Данные в Cloud Bigtable состоят из записей, каждая из которых имеет ключ и данные, связанные с ключом и организованные в семейства столбцов в виде пар ключ/значение, где ключ — это имя семейства столбцов, а значение — набор связанных столбцов.

Cloud Bigtable не требует, чтобы каждая запись включала каждое семейство столбцов и каждый столбец, относящийся к этому семейству; фактически наличие или отсутствие определенного столбца тоже можно считать данными. Благодаря этому BigQuery позволяет создать таблицу, привязанную к данным в Cloud Bigtable, без явного указания имен столбцов. Если вы сделаете это, BigQuery обеспечит доступ к значениям в семействе столбцов как массиву столбцов, каждый из которых будет доступен как массив значений, записанных с разными временными метками.

¹ Если после этого вы запустите сценарий `setup_data.sh` из репозитория GitHub, тогда `project_id` будет представлять уникальный идентификатор вашего проекта, `instance_id` будет ссылаться на `bqbook-instance`, а `table_name` — на `logs-table`.

Create Table

Source Data
☒ Create from source
 ☐ Create empty table

Repeat job

Select Previous Job ?

Location

Google Cloud Bigtable

https://googleapis.com/bigtable/projects/cloud-training-d ?

File format

Cloud Bigtable

Destination Table

Table name

ch04

logs ?

Table type

External table ?

Column Families ?

Column Family and Qualifiers ?	Type ?	Encoding ?	Only Read Latest ?
sales +	STRING	TEXT	TRUE X
sales.itemid	STRING	TEXT	TRUE X
sales.price	FLOAT	TEXT	TRUE X
sales.qty	INTEGER	TEXT	TRUE X

Add Family

Edit as Text

Options

Ignore unspecified column families ☒ ?

Read row key as string ☒ ?

Create Table

Рис. 4.10. Создание внешней таблицы в BigQuery, ссылающейся на данные в Cloud Bigtable¹

Часто имена столбцов известны заранее, и в таких случаях лучше указать известные столбцы в определении таблицы. В нашем примере мы знаем схему каждой записи в таблице `logs-table` в Cloud Bigtable:

- Ключ записи — идентификатор хранилища, за которым следует отметка времени каждой транзакции.
- Семейство столбцов с именем `sales` для хранения сделок в реестре.
- В семействе столбцов `sales` хранятся:
 - идентификатор товара (строка);
 - цена, по которой был продан товар (число с плавающей точкой);
 - количество единиц товара, купленных в рамках этой сделки (целое число).

¹ На момент написания этой книги данная возможность была доступна только в «старом» веб-интерфейсе <https://bigquery.cloud.google.com/>, являющемся частью GCP Cloud Console (<https://console.cloud.google.com/bigquery>).

Обратите внимание: как показано на рис. 4.10, мы указали всю эту информацию в разделе **Column Families** (Семейства столбцов) определения таблицы.

Cloud Bigtable обрабатывает все данные как обычные байтовые строки, поэтому схема (string, float, integer) больше предназначена для BigQuery, чтобы мы могли избежать необходимости преобразовывать типы значений в запросах. По этой же причине (чтобы избежать приведения типов) мы указали, что ключ записи должен интерпретироваться как строка. После создания таблицы в BigQuery каждый столбец в Cloud Bigtable будет отображаться в столбец в BigQuery соответствующего типа:

sales.price	RECORD	NULLABLE	Describe this field...
sales.price.cell	RECORD	NULLABLE	Describe this field...
sales.price.cell.timestamp	TIMESTAMP	NULLABLE	Describe this field...
sales.price.cell.value	FLOAT	NULLABLE	Describe this field...

После создания таблицы BigQuery можно запустить старый добрый SQL-запрос, чтобы найти общее количество проданных товаров `itemid 12345`:

```
SELECT SUM(sales.qty.cell.value) AS num_sold
FROM ch04.logs
WHERE sales.itemid.cell.value = '12345'
```

Улучшение производительности

Федеративный запрос к данным, хранящимся в Google Cloud Storage, обрабатывается рабочими узлами BigQuery. Напротив, федеративный запрос к данным, хранящимся в Cloud Bigtable, обрабатывается кластером Cloud Bigtable. Поэтому скорость обработки второго запроса зависит от емкости кластера Cloud Bigtable и его загруженности в период обработки запроса.

Как и в случае с любым аналитическим запросом, скорость выполнения запроса зависит также от количества строк, которые необходимо прочитать, и объема извлекаемых данных. BigQuery пытается ограничить объем извлекаемых данных, читая только семейства столбцов, на которые ссылается запрос, а Cloud Bigtable распределяет данные по узлам, чтобы воспользоваться преимуществами распределения префиксов ключей по всему набору данных.



Если ваши данные часто обновляются или вам нужен точечный поиск с низкой задержкой, Cloud Bigtable обеспечит более высокую производительность для запросов с фильтрацией по диапазону префиксов ключей. Вам может показаться, что у BigQuery более высокая производительность, чем у Cloud Bigtable, за счет поддержки специального точечного поиска данных в Cloud Bigtable, не ограниченных ключами записей. Однако этот подход часто разочаровывает, и вам обязательно следует сравнить его с вашей рабочей нагрузкой перед выбором архитектуры для практического использования.

BigQuery хранит данные в порядке следования столбцов, оптимизированном для сканирования таблиц, тогда как Cloud Bigtable хранит данные в порядке следования записей, оптимизированном для коротких операций чтения и записи. Запросы к внешним данным, хранящимся в Cloud Bigtable, не дают преимуществ, которые предлагает внутреннее хранилище BigQuery, и выполняются эффективно, только если они читают подмножество строк, но никак не в случае полного сканирования таблицы. Поэтому будьте внимательны и используйте федеративные запросы BigQuery, выполняющие фильтрацию по ключу Bigtable; иначе им придется каждый раз читать всю таблицу Cloud Bigtable.

В вашем распоряжении имеется настраиваемый параметр — количество узлов в кластере Cloud Bigtable. Если вы собираетесь регулярно отправлять запросы SQL к данным в Cloud Bigtable, следите за использованием процессора Cloud Bigtable и при необходимости увеличивайте количество узлов Cloud Bigtable.

Так же как и в случае с федеративными запросами через Google Cloud Storage, подумайте, насколько выгодно настроить конвейер ETL для анализа данных из Cloud Bigtable, — возможно, будет быстрее и проще извлечь данные из Cloud Bigtable с помощью федеративного запроса и загрузить их в таблицу BigQuery для дальнейшего анализа и преобразований. Этот подход, показанный на рис. 4.11, позволит вам выполнять анализ в среде, не зависящей от загруженности Cloud Bigtable. Анализ данных во внутренней таблице BigQuery может выполняться тысячами машин, а не гораздо меньшим кластером. Поэтому аналитические запросы в BigQuery выполняются быстрее (при условии, что их нельзя выполнить с использованием префиксов ключей), чем федеративные запросы к внешней таблице. Недостатком же является дублирование извлеченных данных в Cloud Bigtable и в BigQuery. Тем не менее стоимость хранения обычно невысока, а преимущества масштабирования и скорости могут в достаточной мере компенсировать недостатки.

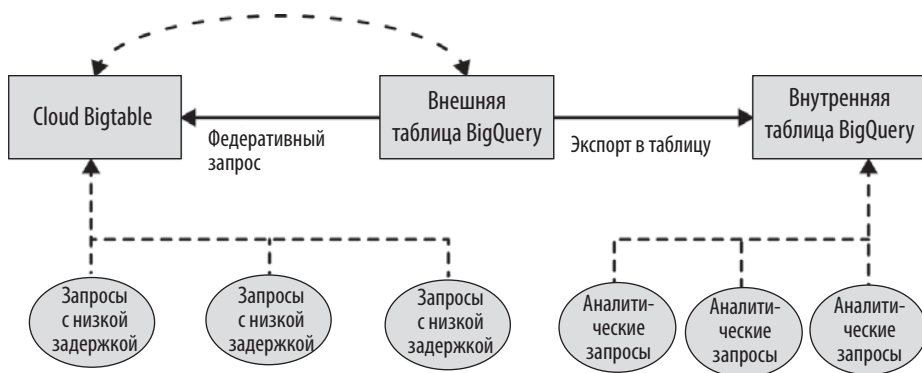


Рис. 4.11. Экпортируйте выбранные таблицы во внутреннюю таблицу BigQuery с помощью федеративного запроса и выполняйте анализ данных с ее использованием

Можно также запланировать периодическую загрузку данных во внутренние таблицы BigQuery. Мы рассмотрим этот подход в следующем разделе.



Если вы запустили экземпляр Cloud Bigtable ради эксперимента, удалите его сейчас, чтобы не увеличивать свои расходы.

Передача и экспорт данных

До сих пор мы рассматривали загрузку данных как разовую операцию и избегали перемещения данных с помощью федеративных запросов. В этом разделе мы разберем доступные сервисы для периодической передачи данных в BigQuery из разных источников.

Служба передачи данных Data Transfer Service

Служба передачи данных BigQuery Data Transfer Service позволяет запланировать периодическую загрузку данных из разных источников в BigQuery. Как и многие другие возможности BigQuery, служба BigQuery Data Transfer Service доступна из веб-интерфейса, инструмента командной строки или через REST API. Чтобы вы могли сами попробовать на примере, мы покажем способ с использованием командной строки.

После настройки передачи данных BigQuery автоматически будет загружать данные по указанному вами расписанию. Но если с исходными данными возникнет какая-то проблема, вы сможете инициировать обратную передачу данных для восстановления после любых сбоев или аварий. Такая обратная передача называется *обновлением*, и вы сможете запустить ее из веб-интерфейса.

Служба Data Transfer Service поддерживает загрузку данных из многих приложений типа «программное обеспечение как услуга» (Software as a Service, SaaS), таких как Google Ads, Google Play, Amazon Redshift и YouTube, а также из Google Cloud Storage. Далее мы посмотрим, как настроить обычную загрузку файлов из облачного хранилища Cloud Storage, и попутно отметим отличия от передачи данных из набора данных SaaS на примере отчетов о каналах на YouTube.

Локальность данных

Как мы уже говорили в этой главе, наборы данных в BigQuery создаются в определенном регионе (например, `asia-northeast1`, которому соответствует столица Японии Токио) или в объединенном регионе (например, `EU`).¹ Когда вы настраи-

¹ Список регионов для наборов данных BigQuery можно найти по ссылке <https://cloud.google.com/bigquery/docs/locations>, а список регионов для корзин Cloud Storage — по ссылке <https://cloud.google.com/storage/docs/bucket-locations>.

ваете службу Data Transfer Service в наборе данных, она будет обрабатывать и размещать данные в том же регионе, где хранится целевой набор данных BigQuery.

Если ваша корзина Cloud Storage находится в том же регионе, что и набор данных BigQuery, за передачу данных платить не нужно. Передача данных между регионами (например, из хранилища Cloud Storage в одном регионе в набор данных BigQuery в другом) повлечет за собой сетевые сборы, независимо от того, происходит ли передача посредством загрузки, экспорта или передачи данных.

Для создания заданий на передачу и запись данных в целевой набор данных вы должны включить службу BigQuery Data Transfer Service (это можно сделать в веб-интерфейсе BigQuery), и вам должен быть присвоен статус `bigquery.admin`.

Настройка таблицы назначения

Служба передачи данных не может создавать новые таблицы, автоматически определять схему и т. д. Поэтому вы должны заранее подготовить таблицу с соответствующей схемой. Если вы записываете все данные в секционированную таблицу, при создании схемы этой таблицы укажите тип столбца секционирования как `TIMESTAMP` или `DATE`. Более подробно о секционировании мы поговорим в главе 7.

Здесь мы проиллюстрируем этот процесс на наборе данных с оценочными параметрами колледжей. Он хранится в объединенном регионе US, поэтому вы тоже должны создать набор данных в объединенном регионе US, если захотите попробовать повторить.

Выполните следующий запрос в BigQuery:

```
CREATE OR REPLACE TABLE
ch04.college_scorecard_dts
AS
SELECT * FROM ch04.college_scorecard_gcs
LIMIT 0
```

Это пример инструкции DDL. Она сохранит результаты запроса `SELECT` (не содержащие ни одной строки и потому не влекущие никаких расходов) в виде таблицы с именем `college_scorecard_dts` в наборе данных `ch04`.

Создание заданий для передачи данных

Выполните в терминале следующую команду, чтобы создать задание для передачи данных:

```
bq mk --transfer_config --data_source=google_cloud_storage \
--target_dataset=ch04 --display_name ch04_college_scorecard \
--params='{"data_path_template":"gs://bigquery-oreilly-book/college_*.csv",
"destination_table_name_template":"college_scorecard_dts", "file_format":"CSV",
"max_bad_records":"10", "skip_leading_rows":"1", "allow_jagged_rows":"true"}'
```

СОЗДАНИЕ ТАБЛИЦ НА SQL

Операторы DDL позволяют создавать и изменять таблицы и представления в BigQuery, используя стандартный синтаксис SQL-запросов. Например, следующий запрос создаст новую таблицу с именем `ch04.college_scorecard_valid_sat` и заполнит ее записями из `ch04.college_scorecard_gcs`, в которых столбец `SAT_AVG` содержит допустимое значение:

```
CREATE TABLE
  ch04.college_scorecard_valid_sat
AS
SELECT * FROM ch04.college_scorecard_gcs
WHERE LENGTH(SAT_AVG) > 0
```

DDL-инструкция `CREATE TABLE` вернет ошибку, если таблица уже существует. Но на этот случай есть другие инструкции: `CREATE OR REPLACE` (заменяет существующую таблицу) и `CREATE IF NOT EXISTS` (оставляет существующую таблицу как есть).

Также можно создать пустую таблицу без использования оператора `SELECT` и с требуемой схемой:

```
CREATE TABLE ch04.payment_transactions
(
  PAYEE STRING OPTIONS(description="Id of payee"),
  AMOUNT NUMERIC OPTIONS(description="Amount paid")
)
```

Возможность выполнять DDL-запросы из командной строки BigQuery или с помощью REST API позволяет создавать таблицы программно.

Эта команда указывает, что источником данных является облачное хранилище Google Cloud Storage (при передаче данных из YouTube Channel, например, источником данных будет `youtube_channel` (<https://cloud.google.com/bigquery/docs/youtube-channel-transfer>)), а местом назначения — набор данных `ch04`. Отображаемое имя (параметр `--display_name`) определяет удобочитаемое имя для использования в различных пользовательских интерфейсах для ссылки на задание.

В случае с YouTube таблицы назначения автоматически сегментируются во время импортирования и получают соответствующие имена. Однако при передаче данных из Cloud Storage нужно явно отразить это в имени целевой таблицы. Например, если указать `mytable_{run_time|"%Y%m%d"}` в шаблоне имени таблицы назначения (параметр `destination_table_name_template`), имя таблицы будет начинаться с `mytable` и к нему будет добавляться время выполнения задания с использованием указанных параметров форматирования даты и времени.¹ Под-

¹ Список доступных параметров форматирования можно найти в документации BigQuery с описанием форматирования столбцов DATETIME (<https://cloud.google.com/bigquery/docs/reference/standard-sql/functions-and-operators#supported-format-elements-for-datetime>).

держивается удобный ярлык `ytable_{run_date}`. Он использует дату в формате `YYYYMMDD`. Также можно указать смещение по времени. Например, вот как можно присвоить таблице имя на основе отметки времени со смещением на 45 минут после выполнения:

```
{run_time+45m|"%Y%m%d"}_mytable_{run_time|"%H%M%S"}
```

В результате получится имя таблицы в формате `20180915_mytable_004500`.

Сами параметры зависят от источника данных. В случае передачи файлов из Google Cloud Storage необходимо указать следующее:

- Путь к исходным данным, возможно, с групповыми символами.
- Шаблон имени таблицы назначения.
- Формат файла. Служба передачи данных из Cloud Storage поддерживает все форматы данных, которые поддерживаются федеративными запросами (CSV, JSON, Avro, Parquet и т. д.). Для формата CSV можно дополнительно указать специфические параметры, например количество пропускаемых заголовочных строк.

Для передачи данных из YouTube Channel следует указать параметры `page_id` (в YouTube) и `table_suffix` (в BigQuery).

Запустив команду `bq mk`, как в примере выше, вы получите URL как часть рабочего процесса OAuth2; передайте требуемый токен, выполнив вход через браузер, и задание на передачу будет создано.

Службу передачи данных Data Transfer Service также можно запустить из веб-интерфейса, как показано на рис. 4.12.

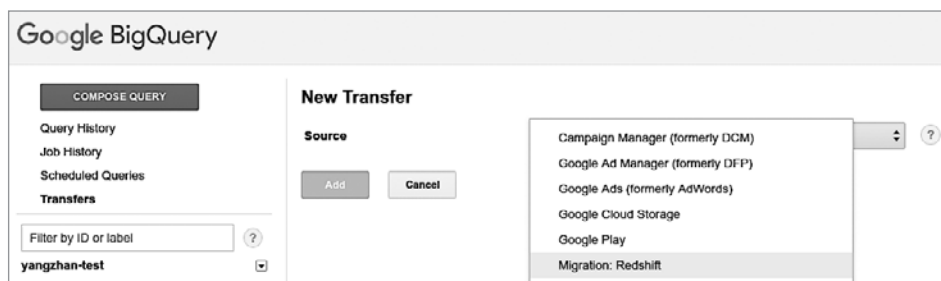


Рис. 4.12. Инициировать передачу данных также можно из веб-интерфейса

Обратите внимание, что здесь мы не определили расписание; по умолчанию задание будет запускаться каждые 24 часа, начиная с «текущего момента». Однако веб-интерфейс BigQuery позволяет скорректировать расписание передачи, как показано на рис. 4.13.

Рис. 4.13. Определение расписания для задания передачи данных в веб-интерфейсе

Стоимость передачи данных зависит от источника. На момент написания этой книги передача данных из YouTube Channel стоила 5\$ за канал в месяц, а передача данных из Cloud Storage была бесплатной. Однако поскольку служба Data Transfer Service использует задания загрузки для загрузки данных из Cloud Storage в BigQuery, на нее распространяются все ограничения для заданий загрузки, установленные в BigQuery (https://cloud.google.com/bigquery/quotas#load_jobs).

Выполнение запросов по расписанию

BigQuery поддерживает возможность выполнения запросов по расписанию и сохранения результатов в таблицах BigQuery. В частности, можно запланировать выполнение федеративного запроса для извлечения данных из внешнего источника данных, их преобразования и загрузки в BigQuery. Поскольку такие запросы могут включать операторы DDL и DML, с их помощью можно создавать сложные рабочие процессы исключительно на SQL.

Чтобы создать запрос, выполняемый по расписанию, откройте диалог, кликнув **Schedule Query (Запланировать запрос)** в веб-интерфейсе BigQuery, как показано на рис. 4.14.¹

Запросы, выполняемые по расписанию, основаны на службе передачи данных Data Transfer Service, поэтому у них есть много общего с заданиями на передачу данных. Например, вы можете указать таблицу назначения, используя те же

¹ На момент написания книги эта возможность была доступна только в «классическом интерфейсе».

параметры (например, `run_date` и `run_time`), что и в заданиях для службы Data Transfer Service (см. предыдущий раздел).



Рис. 4.14. Создание запроса, выполняемого по расписанию, в веб-интерфейсе BigQuery

Копирование наборов данных между регионами

BigQuery поддерживает возможность копирования наборов данных по расписанию между регионами с использованием службы передачи данных. В веб-интерфейсе Data Transfer Service выберите в поле **Source** (Источник) значение **Cross Region Copy** (Копирование между регионами), а также укажите имя исходного набора данных, из которого требуется скопировать таблицы, и целевой набор данных, как показано на рис. 4.15.

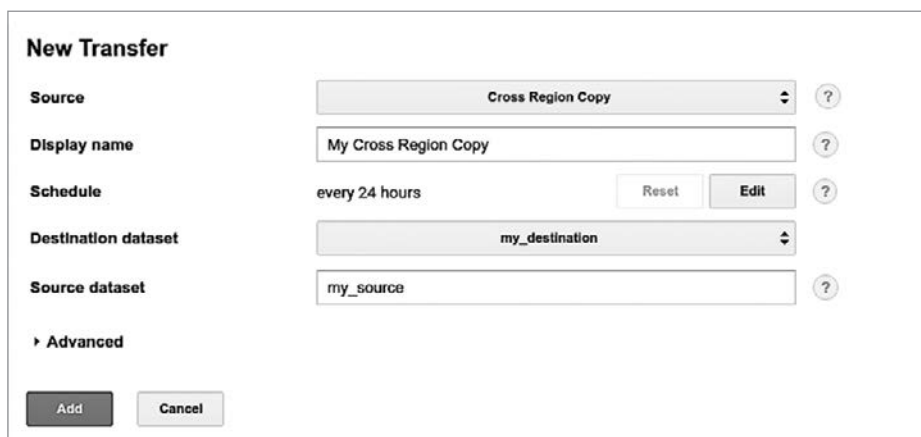


Рис. 4.15. Чтобы инициировать копирование набора данных между регионами, в веб-интерфейсе Data Transfer Service укажите, что источником является копия из другого региона

Так как оба набора данных, исходный и целевой, являются наборами данных BigQuery, инициатор должен иметь права доступа на передачу данных, получение списка таблиц в исходном наборе данных, просмотр исходного набора данных и редактирование целевого набора данных.

Копирование наборов данных между регионами также можно инициировать командой `bq mk`, указав значение `cross_region_copy` в качестве источника данных.

Экспортирование журналов Stackdriver

Журналы виртуальных машин (VM) и служб GCP¹ можно хранить и анализировать с помощью Stackdriver Logging. То есть Stackdriver Logging служит универсальным средством представления всех действий, выполнявшихся в вашей учетной записи GCP. Поэтому иногда бывает полезно экспортировать журналы Stackdriver и Firebase в BigQuery. Сделать это можно с помощью командной строки, REST API или веб-интерфейса, как показано на рис. 4.16.

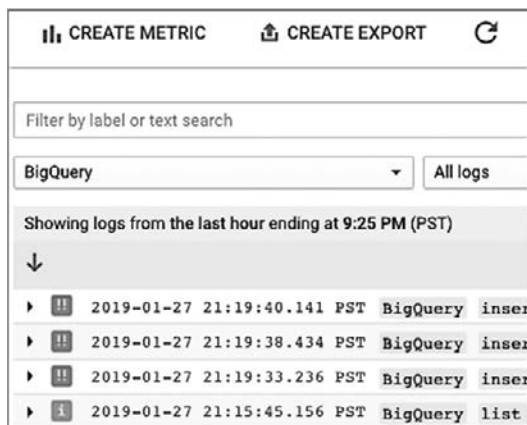


Рис. 4.16. Чтобы просмотреть журналы заданий загрузки в BigQuery, созданные в предыдущем разделе, перейдите в облачной консоли в GCP Cloud Console в раздел Stackdriver

Чтобы экспортировать все журналы из службы BigQuery, кликните **Create Export** (Экспортировать) в верхней части страницы Stackdriver Logs Viewer (<https://console.cloud.google.com/logs/>) и введите следующую информацию:

- Выберите **BigQuery** и **All Logs** (Все журналы), чтобы увидеть все журналы из BigQuery. Видите свои недавние действия?
- Укажите имя приемника, например `bq_logs`.

¹ А также виртуальных машин и служб, действующих в Amazon Web Services.

- Укажите службу-приемник BigQuery, чтобы экспортировать данные в BigQuery.
- Укажите целевой набор данных, куда следует экспортировать журналы: `ch04`.

Давайте посмотрим, какие записи в журнале генерируют запросы. Перейдите в веб-интерфейс BigQuery и выполните следующий запрос:

```
SELECT
  gender, AVG(tripduration / 60) AS avg_trip_duration
FROM
  `bigquery-public-data`.new_york_citibike.citibike_trips
GROUP BY
  gender
HAVING avg_trip_duration > 14
ORDER BY
  avg_trip_duration
```

Затем в том же веб-интерфейсе BigQuery UI выполните следующий запрос (не забудьте изменить дату соответствующим образом):

```
SELECT protopayload_auditlog.status.message
FROM ch04.cloudaudit_googleapis_com_data_access_20190128
```

В результате вы получите список сообщений из журнала BigQuery, включая сообщение о чтении результатов предыдущего запроса. В зависимости от даты в фильтре, вы также должны увидеть записи, соответствующие ранее выполненным операциям.

Обратите внимание на следующие возможности механизма экспорта:

- Схема и даже имя таблицы были взяты из Stackdriver. Мы лишь указали целевой набор данных.
- Данные обновились почти мгновенно. Это пример работы потокового буфера — Stackdriver обновляет таблицы BigQuery в реальном времени (хотя обычно запросы BigQuery возвращают данные, «устаревшие» на несколько секунд).



Чтобы избежать увеличения платы за этот потоковый конвейер, перейдите в раздел Stackdriver в консоли и удалите приемник.

Использование Cloud Dataflow для чтения/записи в BigQuery

Как мы уже говорили, BigQuery поддерживает федеративные запросы из таких источников, как Google Sheets. Служба передачи данных Data Transfer Service

поддерживает такие источники, как Google Ads и YouTube. Службы Stackdriver Logging и Firestore позволяют экспортировать свои данные в BigQuery.

А как быть тем, кто использует такие продукты, как MySQL, которые не предлагают возможность экспорта и не поддерживаются службой Data Transfer Service? В таких случаях можно воспользоваться сервисом Cloud Dataflow. Cloud Dataflow — это управляемая служба GCP, которая упрощает выполнение конвейеров данных, построенных с использованием Apache Beam API, и заботится о таких тонкостях, как производительность, масштабирование, доступность, безопасность и соответствие законодательству, чтобы пользователи могли сосредоточиться на программировании, а не на управлении кластерами серверов. Служба Dataflow позволяет организовать преобразование и обогащение данных в потоковом (в масштабе реального времени) и в пакетном (архивном) режимах с использованием одного и того же кода в потоковом и пакетном конвейерах.

Использование шаблона Dataflow для прямой загрузки данных из MySQL

Несмотря на возможность писать конвейеры Cloud Dataflow (мы сделаем это в разделе «Создание задания Dataflow» ниже), для многих общих потребностей можно использовать готовые шаблоны конвейеров Dataflow, доступные в GitHub (<https://github.com/GoogleCloudPlatform/DataflowTemplates>). Посмотрев список имеющихся шаблонов, можно заметить шаблон копирования из Jdbc в BigQuery, как нельзя лучше соответствующий нашим требованиям и подходящий для передачи данных из MySQL в BigQuery.


Откройте GCP Cloud Console и перейдите в раздел **Cloud Dataflow**. Затем выберите **Create job from template** (Создать задание из шаблона), выберите **Jdbc to BigQuery** (Из Jdbc в BigQuery) и заполните форму информацией об исходной таблице в базе данных MySQL и целевой таблице в BigQuery, как показано на рис. 4.17.

Кликнув **Run job** (Выполнить задание), вы запустите задание Dataflow. Оно выполнит указанный вами JDBC-запрос и вынесет полученные записи в BigQuery.


Создание задания Dataflow

Если у вас есть источник данных в формате, который не поддерживается ни федеративными запросами, ни службой Data Transfer Service, ни механизмом экспорта, ни предопределенными шаблонами Dataflow, вы можете написать свой конвейер Dataflow для загрузки этих данных в BigQuery.


Несмотря на то что работу с файлами CSV в Google Cloud Storage поддерживают и федеративные запросы, и служба Data Transfer Service, мы будем использовать эти файлы, чтобы показать, как организовать конвейер Dataflow. Код использует Apache Beam API, и его можно написать на Python, Java или Go. Здесь мы используем Python.


Dataflow
← Create job from template

Job name
Must be unique among running jobs. Use lowercase letters, numbers, and hyphens (-).

Cloud Dataflow template 
A pipeline that reads from a Jdbc source and writes to a BigQuery table.

Required Parameters

Regional endpoint 
Choose where to deploy Cloud Dataflow workers and store metadata for the job.

Jdbc connection URL string
Url connection string to connect to the Jdbc source. E.g. jdbc:mysql://some-host:3306/sampledb

Jdbc driver class name
Jdbc driver class name. E.g. com.mysql.jdbc.Driver

Jdbc source SQL query.
Query to be executed on the source to extract the data. E.g. select * from sampledb.sample_table

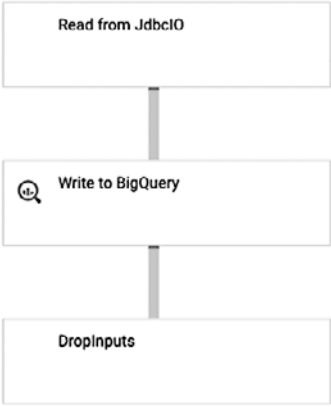
BigQuery output table
BigQuery table location (<project>.<dataset>.<table_name>) to write the output to. The table's schema must match the source query schema.

GCS paths for Jdbc drivers
Comma separate GCS paths for Jdbc drivers. E.g. gs://<some-bucket>/driver_jar1.jar,gs://<some_bucket>/driver_jar2.jar

Temporary directory for BigQuery loading process
Example: gs://my-bucket/my-files/temp_dir

Temporary Location
Path and filename prefix for writing temporary files. ex: gs://MyBucket/tmp

Optional parameters



```

graph TD
    A[Read from JdbcIO] --> B[Write to BigQuery]
    B --> C[DropInputs]
  
```

Рис. 4.17. Создание задания Dataflow из шаблона для передачи данных из MySQL в BigQuery

Основная задача кода — извлечь исходные данные, преобразовать их, оставив и очистив нужные поля, и загрузить в BigQuery:

```
INPATTERNS = 'gs://bigquery-oreilly-book/college_*.csv'
RUNNER = 'DataflowRunner'
with beam.Pipeline(RUNNER, options = opts) as p:
    (p
     | 'read' >> beam.io.ReadFromText(INPATTERNS, skip_header_lines=1)
     | 'parse_csv' >> beam.FlatMap(parse_csv)
     | 'pull_fields' >> beam.FlatMap(pull_fields)
     | 'write_bq' >> beam.io.gcp.bigquery.WriteToBigQuery(bqtable, bqdataset,
schema=get_output_schema())
    )
```

Этот код создает конвейер Beam и определяет, что тот будет выполняться службой Cloud Dataflow. Также в RUNNER можно указать `DirectRunner` (выполняется на локальном компьютере) и `SparkRunner` (выполняется службой Apache Spark в кластере Hadoop, например Cloud Dataproc в GCP).

Сначала конвейер читает все файлы, соответствующие шаблону ввода в INPATTERNS. Файлы могут находиться на локальном диске или в Google Cloud Storage. Затем конвейер построчно передает данные из текстовых файлов следующему этапу, на котором к каждой строке применяется метод `parse_csv`:

```
def parse_csv(line):
    try:
        values = line.split(',')
        rowdict = {}
        for colname, value in zip(COLNAMES, values):
            rowdict[colname] = value
        yield rowdict
    except:
        logging.warn('Ignoring line ...')
```

Метод `parse_csv` разбивает строку по запятым и преобразует значения в словарь, ключами в котором служат имена столбцов, а значениями — значения из ячеек.

Затем этот словарь передается методу `pull_fields`, который извлекает интересные нас данные (столбец INSTNM и несколько числовых полей) и преобразует их:

```
def pull_fields(rowdict):
    result = {}
    # поля должны быть строковыми
    for col in 'INSTNM'.split(','):
        if col in rowdict:
            result[col] = rowdict[col]
        else:
            logging.info('Ignoring line missing {}'.format(col))
    return result
```

```
# числовые поля
for col in \
'ADM_RATE_ALL,FIRST_GEN,MD_FAMINC,SAT_AVG,MD_EARN_WNE_P10'.split(','):
    try:
        result[col] = (float) (rowdict[col])
    except:
        result[col] = None
yield result
```

Словари с извлеченными полями последовательно пересылаются в BigQuery. Для приемника BigQuery (`beam.io.gcp.bigquery.WriteToBigQuery`) требуется определить имя таблицы, имя набора данных и схему в следующем виде:

```
INSTNM:string,ADM_RATE_ALL:FLOAT64,FIRST_GEN:FLOAT64,...
```

Если таблица BigQuery отсутствует, она будет создана. Если таблица есть, новые записи будут добавлены в нее без уничтожения старых. Есть и другие варианты: например, таблицу можно заменить.

Если запустить эту программу на Python,¹ она запустит задание Dataflow, которое прочитает файл CSV, проанализирует его по строкам, извлечет необходимые поля и запишет преобразованные данные в BigQuery.

Мы продемонстрировали программу Dataflow на примере пакетного конвейера (то есть объем входных данных ограничен), однако один и тот же конвейер можно использовать для анализа, преобразования и записи данных, получаемых в потоковом режиме (например, из Cloud Pub/Sub), как это часто имеет место во многих системах журналирования и в IoT. То есть подход на основе Dataflow дает возможность преобразовывать данные сразу и загружать их в BigQuery.

Обратите внимание, что Dataflow использует механизм потоковой передачи для загрузки данных в BigQuery, независимо от того, в каком режиме запущено задание — в потоковом или пакетном. Потоковая передача обеспечивает оперативное появление данных в потоковом буфере и возможность их получения даже во время записи. Однако здесь есть свои недостатки: в отличие от заданий загрузки BigQuery, за потоковую передачу нужно заплатить. Напомним, что загрузка данных в BigQuery может быть бесплатной, но по соображениям производительности существуют ограничения на количество заданий загрузки. Потоковая передача позволяет избежать ограничений и квот, налагаемых на задания загрузки, без ущерба для производительности запросов.

Использование потокового API напрямую

Мы представили Apache Beam в контексте Cloud Dataflow как способ извлечения, преобразования и загрузки данных в BigQuery в потоковом режиме, но это не единственный фреймворк обработки данных, способный записывать

¹ См. блокнот `04_load/dataflow.ipynb` в репозитории GitHub с примерами для книги.

данные в BigQuery. Если ваша команда ближе знакома с Apache Spark, создание конвейера ETL в Spark и его выполнение в кластере Hadoop (например, Cloud Dataproc в GCP) является жизнеспособной альтернативой Dataflow. Это связано с наличием клиентских библиотек для разных языков, а BigQuery поддерживает потоковый API.

Более подробно клиентскую библиотеку и потоковую передачу мы рассмотрим в главе 5, а пока посмотрите на фрагмент, иллюстрирующий загрузку данных с использованием Streaming API из Python:

```
# создать массив кортежей и переслать данные, когда они станут доступны
rows_to_insert = [
    (u'U. Puerto Rico', 0.18,0.46,23000,1134,32000),
    (u'Guam U.', 0.43,0.21,28000,1234,33000)
]
errors = client.insert_rows(table, rows_to_insert) # запрос к API
```

По мере появления новых данных вызывается метод `insert_rows()` клиента BigQuery. Этот метод, в свою очередь, вызывает метод `tabledata.insertAll` из REST API. Данные сохраняются в потоковом буфере BigQuery и сразу же становятся доступны для запросов, хотя для того, чтобы данные стали доступны для экспорта, может потребоваться до 90 минут.

Перемещение локальных данных

В главе 1 мы говорили, что одним из ключевых факторов, на котором основана вся работа BigQuery, является отделение вычислений от хранилищ в сети с петабитной пропускной способностью. BigQuery лучше всего работает с наборами данных, которые находятся в вычислительном центре и за брандмауэром Google Cloud — если бы BigQuery пришлось читать данные с использованием линий связи общедоступного интернета или более медленного сетевого соединения, она была бы не так эффективна. Поэтому для эффективной работы BigQuery важно, чтобы данные находились в облаке.

BigQuery — это масштабируемая аналитическая платформа, которая рекомендуется для хранения структурированных данных, за исключением случаев, когда они предназначены для использования в масштабе реального времени. То есть если BigQuery — это место для хранения всех структурированных данных, используемых для анализа, то как можно переместить локальные данные в BigQuery?

Методы миграции данных

Если у вас хорошая сеть и имеется высокоскоростное соединение с Google Cloud, вы можете использовать команду `bq load` для загрузки данных в BigQuery. Как говорилось в этой главе, лучше, чтобы загружаемые данные уже хранились

в Google Cloud Storage. Для копирования данных из локального хранилища в облачное можно использовать инструмент командной строки `gsutil`.

При копировании многих файлов, особенно больших, в облачное хранилище Google Cloud Storage используйте параметр `-m`, чтобы включить многопоточность. Она позволяет инструменту `gsutil` копировать файлы параллельно:

```
gsutil -m cp /some/dir/myfiles*.csv gs://bucket/some/dir
```

Поскольку есть вероятность, что сбор данных продолжится, перемещение данных часто является не разовым, а непрерывным процессом. Одним из способов решения является использование службы Cloud Function для автоматического запуска `bq load` всякий раз, когда файл появляется в облачном хранилище.¹ Если файлы поступают часто (и имеют небольшие размеры), вместо Cloud Storage лучше использовать Cloud Pub/Sub,² чтобы хранить входящие данные в виде сообщений, которые будут обрабатываться конвейером Cloud Dataflow и передаваться непосредственно в BigQuery.

Эти три подхода — `gsutil`, Cloud Function и Cloud Dataflow — перечислены в первых трех строках табл. 4.2 и могут успешно использоваться при достаточно хорошем сетевом соединении.

Таблица 4.2. Рекомендованные методы миграции данных для разных ситуаций

Какие данные требуется перенести	Рекомендуемый метод миграции
Относительно небольшие файлы	<code>gsutil cp -m</code> <code>bq load</code>
Периодическая загрузка (например, раз в сутки) файлов в BigQuery	<code>gsutil cp</code> вызов <code>bq load</code> из Cloud Function
Потоковая загрузка сообщений в BigQuery	Отправка данных в Cloud Pub/Sub с последующим использованием Cloud Dataflow для потоковой загрузки в BigQuery. В этом случае часто требуется реализовать конвейер на Python, Java, Go или каком-то другом языке. Как вариант, можно использовать Streaming API посредством клиентской библиотеки. Подробнее этот способ будет описан в главе 5
Разделы Hive	Перенос рабочей нагрузки Hive в Cloud Dataproc Запрос разделов Hive как внешних таблиц
Петабайты данных или при использовании низкоскоростного сетевого соединения	Transfer Appliance <code>bq load</code>

¹ Как это сделать программно, мы покажем в главе 5.

² Служба шины сообщений — см. <https://cloud.google.com/pubsub/>.

Какие данные требуется перенести	Рекомендуемый метод миграции
Из одного региона в другой или из другого облачного окружения	Cloud Storage Transfer Service
Дамп MySQL	Шаблоны для Dataflow, которые можно настроить и запустить
Из Google Cloud Storage, Google Ads, Google Play, Amazon Redshift и пр. в BigQuery	BigQuery Data Transfer Service Настроить передачу в BigQuery. Все функции Data Transfer Service действуют аналогично
Stackdriver Logging, Firestore и т. п.	Эти инструменты предоставляют возможность экспорта в BigQuery. Настройка выполняется непосредственно в этих инструментах (Stackdriver, Firestore и т. п.)

Перенос данных с использованием `gsutil` для размещения данных в облачном хранилище и последующим вызовом `bq load` легко осуществим при наличии нескольких небольших наборов данных, но если у вас имеется много больших наборов данных, то это значительно сложнее. С увеличением объема данных растет количество ошибок. Поэтому при переносе больших наборов данных необходимо уделять внимание деталям, например проверять контрольные суммы при извлечении и передаче, настраивать брандмауэры, чтобы они не блокировали передачу и не удаляли пакеты, исключить возможность просачивания конфиденциальных данных, обеспечивать их шифрование и защиту от потери во время и после миграции.

Другая характерная проблема `gsutil` заключается в том, что не всегда возможно выделить достаточную полосу пропускания для передачи данных, потому что она часто стоит слишком дорого и мешает обычным операциям передачи данных в корпоративной сети.

В случаях, когда невозможно скопировать данные в Google Cloud из-за их большого объема или сетевых ограничений, оцените возможность использовать Transfer Appliance. Это стоечный сервер хранения большой емкости, который вы получаете, заполняете данными и отправляете обратно в Google Cloud или одному из авторизованных партнеров. Transfer Appliance удобно использовать для переноса больших объемов данных (от сотен терабайт до петабайт), которые нельзя передать через ваше сетевое соединение.

Если ваши данные хранятся не локально, а в другом общедоступном облаке (например, в корзине Amazon Web Services Simple Storage Service), для их переноса можно использовать Cloud Storage Transfer Service. Типичный случай, когда приложение выполняется в Amazon Web Services, а анализ его журналов производится в BigQuery. Служба Cloud Storage Transfer Service также прекрасно подходит для передачи больших объемов данных между регионами в Google.

Служба BigQuery Data Transfer Service автоматизирует загрузку данных в BigQuery из таких приложений, принадлежащих Google, как YouTube, Google

Ads и т. д. Другие инструменты вроде Stackdriver Logging и Firestore предлагают возможность экспорта в BigQuery.

Вы можете самостоятельно организовать миграцию данных, но вряд ли у сотрудников вашего IT-подразделения имеется достаточно опыта, потому что миграция часто является единовременной задачей. Иногда для выполнения миграции данных выгоднее воспользоваться услугами авторизованного партнера GCP.¹

Выводы

Инструмент командной строки `bq` предоставляет единую точку входа для взаимодействия с сервисом BigQuery в GCP. Как только ваши данные окажутся в облачном хранилище Google Cloud Storage, вы сможете выполнить однократную загрузку данных с помощью команды `bq load`. Она поддерживает автоматическое определение схемы, но также может использовать вашу схему. В зависимости от характера потребления ресурсов вашим заданием загрузки (вычислительных и ввода/вывода), иногда бывает полезно сжать данные.

Данные можно оставить на месте, указать их структуру и использовать BigQuery только в роли движка запросов. Такие данные называют внешними наборами данных, а запросы к внешним наборам данных называют федеративными запросами. Используйте федеративные запросы для исследовательской работы или там, где данные в основном используются во внешнем формате (например, для запросов с низкой задержкой в Cloud Bigtable или интерактивной работы в Sheets). Функция `EXTERNAL_QUERY` позволяет в масштабе реального времени подключаться к базам данных MySQL и Postgres и не вызывает перемещения данных. Для больших и относительно стабильных наборов известных данных, которые будут редко обновляться и часто запрашиваться, лучше использовать собственное хранилище BigQuery. Федеративные запросы также могут пригодиться для организации процесса извлечения, загрузки и преобразования (Extract, Load, Transform — ELT), когда особенности данных еще недостаточно изучены.

Есть возможность организовать передачу данных по расписанию из различных платформ в BigQuery. Другие инструменты также поддерживают возможность экспорта своих данных в BigQuery. Для простой загрузки данных можно использовать Cloud Function; для продолжительной потоковой загрузки используйте Cloud Dataflow. Также можно запланировать периодические запросы (включая федеративные) и загружать данные в таблицы с их помощью.

¹ См. <https://cloud.google.com/bigquery/providers/>. Когда мы писали эту книгу, GCP объявила о намерении приобрести Aloomo, поставщика услуг облачной миграции, — см. <https://cloud.google.com/blog/topics/inside-google-cloud/google-announces-intent-to-acquire-aloomo-to-simplify-cloud-migration>.

Разработка с BigQuery

До сих пор для взаимодействия с BigQuery мы использовали в основном веб-интерфейс BigQuery и инструмент командной строки `bq`. В этой главе мы рассмотрим способы программного взаимодействия с сервисом. Они могут пригодиться для создания сценариев или автоматизации задач, связанных с BigQuery. Программный доступ к BigQuery также может понадобиться при разработке приложений, информационных панелей (дашбордов), научной графики и моделей машинного обучения, для которых BigQuery является одним из инструментов.

Для начала мы познакомимся с клиентскими библиотеками BigQuery, которые позволяют программно запрашивать таблицы и ресурсы BigQuery и управлять ими. Но даже имея программный доступ к BigQuery с помощью этих низкоуровневых API, вы должны знать о настройках и абстракциях более высокого уровня, доступных в определенных средах (блокноты Jupyter и сценарии командной оболочки). Эти настройки, о которых мы расскажем во второй половине главы, довольно просты в использовании, помогают должным образом обрабатывать ошибки и избавляют от массы шаблонного кода.

Программный доступ

Для программного доступа к BigQuery рекомендуется использовать клиентскую библиотеку Google Cloud Client Library на выбранном вами языке программирования. Знание REST API поможет понять, что происходит внутри, после отправки запроса в службу BigQuery, но клиентская библиотека BigQuery более практична. Поэтому вы можете лишь бегло пролистать раздел о REST API.

Доступ к BigQuery через REST API

Чтобы послать запрос в сервис BigQuery, достаточно отправить прямой HTTP-запрос на сервер, потому что BigQuery, как и все облачные сервисы Google,

поддерживает традиционный интерфейс JSON/REST (<https://cloud.google.com/bigquery/docs/reference/rest/v2/>).

JSON/REST — это архитектурный стиль проектирования распределенных служб, запросы к которым не имеют состояния (то есть сервер не поддерживает состояния сеанса или контекст; вместо этого вся необходимая информация передается в самом запросе), причем запросы и ответы имеют текстовый формат JSON. Поскольку протокол HTTP не поддерживает состояния, службы REST особенно хорошо подходят для предоставления услуг через интернет. Сообщения в формате JSON отображаются непосредственно в объекты в памяти на таких языках, как JavaScript и Python.

Программные интерфейсы REST API создают иллюзию того, что объекты, на которые они ссылаются, являются статическими файлами и для них поддерживаются операции Create (создать), Read (прочитать), Update (изменить), Delete (удалить) — CRUD, прямо соответствующие глаголам HTTP. Например, чтобы создать таблицу в BigQuery, нужно послать HTTP-запрос POST, чтобы исследовать таблицу — HTTP-запрос GET, чтобы изменить ее — HTTP-запрос PATCH, а чтобы удалить — HTTP-запрос DELETE. Есть также такие методы, как Query, которые не имеют точного аналога в наборе операций CRUD, поэтому их часто называют методами в стиле вызова удаленных процедур (Remote Procedure Call, RPC).

Все идентификаторы URI в BigQuery начинаются с префикса <https://www.googleapis.com/bigquery/v2>. Обратите внимание на использование протокола HTTPS — это означает, что запросы шифруются перед передачей в сеть. Элемент v2 в URI — это номер версии. Некоторые сервисы Google часто меняют номер своей версии, но, в отличие от них, вот уже несколько лет BigQuery придерживается версии v2 и, вероятно, продолжит эту традицию в обозримом будущем.¹

Управление набором данных

Интерфейс REST предполагает отправку HTTP-запросов на определенные адреса URL. Комбинация метода HTTP-запроса (GET, POST, PUT, PATCH или DELETE) и адреса URL определяет выполняемую операцию. Например, чтобы удалить набор данных, клиент должен отправить HTTP-запрос DELETE на URL (включающий идентификатор набора данных и название проекта, в котором он содержится):

```
.../projects/<PROJECT>/datasets/<DATASET>
```

Здесь под многоточием (...) подразумевается <https://www.googleapis.com/bigquery/v2>. Все URL, относящиеся к BigQuery REST API, являются путями, откладываемыми относительно этого префикса.

¹ Другие API, особенно отличные от REST API, такие как gRPC, будут иметь другой префикс.



После ввода URL в адресную строку веб-браузера браузер отправит HTTP-запрос GET на этот адрес. Чтобы отправить HTTP-запрос DELETE, нужен специальный инструмент, позволяющий указать метод HTTP. Одним из таких инструментов является `curl`; чуть позже мы покажем, как его использовать.

Мы знаем, что должны послать запрос DELETE на этот URL, потому что это указано в документации с описанием BigQuery REST API (<https://cloud.google.com/bigquery/docs/reference/rest/v2/>), как показано на рис. 5.1.

Datasets		
For Datasets Resource details, see the resource representation page.		
Method	HTTP request	Description
URLs relative to https://www.googleapis.com/bigquery/v2 , unless otherwise noted		
delete	DELETE /projects/projectId/datasets/datasetId	Deletes the dataset specified by the datasetId value. Before you can delete a dataset, you must delete all its tables, either manually or by specifying deleteContents. Immediately after deletion, you can create another dataset with the same name.
get	GET /projects/projectId/datasets/datasetId	Returns the dataset specified by datasetId.
insert	POST /projects/projectId/datasets	Creates a new empty dataset.
list	GET /projects/projectId/datasets	Lists all datasets in the specified project to which you have been granted the READER dataset role.

Рис. 5.1. В описании BigQuery REST API говорится, что отправка HTTP-запроса DELETE на URL `/projects/<PROJECT>/datasets/<DATASET>` приведет к удалению набора данных, если он пустой

Управление таблицами

Удаление таблицы точно так же выполняется отправкой HTTP-запроса DELETE на URL:

```
.../projects/<PROJECT>/datasets/<DATASET>/tables/<TABLE>
```

Обратите внимание, что оба запроса используют метод HTTP DELETE, но у них разный адрес URL. Безусловно, не каждый, кто может отправить HTTP-запрос,

сможет удалить набор данных или таблицу. Запрос будет выполнен, только если он включает токен доступа и этот токен (рассмотрим его чуть ниже) представляет соответствующую авторизацию в BigQuery или Cloud Platform.¹

Еще один тип метода HTTP — HTTP GET — позволяет получить список всех таблиц в наборе данных, если послать его на URL:

```
.../projects/<PROJECT>/datasets/<DATASET>/tables
```

Чтобы получить список таблиц в наборе данных, достаточно иметь право на чтение — полный доступ к BigQuery (позволяющий, например, удалять таблицы) не требуется, хотя, конечно, более широкие полномочия (например, в BigQuery) также предоставляют более широкие возможности.

Мы можем попробовать выполнить такие запросы в командной оболочке Unix:²

```
#!/bin/bash
PROJECT=$(gcloud config get-value project)
access_token=$(gcloud auth application-default print-access-token)
curl -H "Authorization: Bearer $access_token" \
      -H "Content-Type: application/json" \
      -X GET
"https://www.googleapis.com/bigquery/v2/projects/$PROJECT/datasets/ch04/tables"
```

Токен доступа позволяет приложению получить учетные данные по умолчанию. Это временные учетные данные, которые выдаются при входе в Google Cloud Software Development Kit (SDK). Токен доступа помещается в заголовок HTTP-запроса, и поскольку мы хотим получить список таблиц в наборе данных `ch04`, мы отправляем командой `curl` GET-запрос на URL:

```
.../projects/$PROJECT/datasets/ch04/tables
```

ИСПОЛЬЗОВАНИЕ SQL ВМЕСТО КЛИЕНТСКОГО API

В этой главе рассказывается о том, как организовать программный доступ к BigQuery через клиентский API. Но иногда ту же информацию проще и удобнее получить с помощью SQL-запросов. Например, создавать и удалять таблицы можно с помощью операторов `CREATE TABLE` и `DROP TABLE` соответственно. Язык SQL позволяет оставаться в рамках инструмента, который вы регулярно используете

¹ В частности, должен быть разрешен доступ к <https://www.googleapis.com/auth/bigquery> или <https://www.googleapis.com/auth/cloud-platform>.

² Это код из файла `05_devel/rest_list.sh`, который можно найти в репозитории GitHub с примерами для этой книги (<https://github.com/GoogleCloudPlatform/bigquery-oreilly-book>). Его можно выполнить в любой системе, где установлены Cloud SDK и `curl` (например, в Cloud Shell). Поскольку в этой главе мы еще ничего не создали, я использую набор данных (`ch04`), который мы загрузили в предыдущей главе.

для исследования и анализа данных, без необходимости внедрять инструменты программирования в рабочий процесс.

Например, получить список таблиц в наборе данных из главы 4 можно с помощью SQL-запроса к представлению `INFORMATION_SCHEMA`:

```
SELECT
    table_name, creation_time
FROM
    ch04.INFORMATION_SCHEMA.TABLES
```

Более подробно об использовании представления `INFORMATION_SCHEMA` и SQL мы поговорим в главе 8. А пока в табл. 5.1 перечислим соответствия между функциями клиентского API и SQL.

Таблица 5.1. Функции клиентского API и их альтернативы в SQL

Клиентский API	Альтернатива в SQL
Создание таблицы (или представления: достаточно заменить <code>TABLE</code> на <code>VIEW</code>)	<code>CREATE TABLE</code> <code>CREATE TABLE IF NOT EXISTS</code> <code>CREATE OR REPLACE TABLE</code>
Изменение таблицы (или представления)	<code>ALTER TABLE SET OPTIONS</code> <code>ALTER TABLE IF EXISTS SET OPTIONS</code>
Изменение данных в таблице	<code>INSERT INTO</code> <code>DELETE FROM</code> <code>UPDATE</code> <code>MERGE</code>
Удаление таблицы (или представления)	<code>DROP TABLE</code>
Метаданные набора данных	Запрос к: <code>INFORMATION_SCHEMA.TABLES</code> <code>INFORMATION_SCHEMA.TABLE_OPTIONS</code> <code>INFORMATION_SCHEMA.COLUMNS</code> <code>INFORMATION_SCHEMA.COLUMN_FIELD_PATHS</code>
Метаданные заданий	Запрос к: <code>INFORMATION_SCHEMA.JOBS_BY_USER</code> <code>INFORMATION_SCHEMA.JOBS_BY_PROJECT</code> <code>INFORMATION_SCHEMA.JOBS_BY_ORGANIZATION</code>

Во многих ситуациях SQL может оказаться лучшей альтернативой с точки зрения инструментов и знаний о нем.

Выполнение запросов

Иногда недостаточно просто отправить HTTP-запрос GET на URL BigQuery, вместе с запросом требуется передать больше информации. В таких случаях API требует, чтобы клиент выполнил HTTP-запрос POST и отправил в его теле JSON-запрос.

Например, чтобы выполнить SQL-запрос в BigQuery и получить результаты, нужно отправить HTTP-запрос POST на URL:

```
.../projects/<PROJECT>/queries
```

и послать в нем примерно такой JSON-запрос:

```
{
  "useLegacySql": false,
  "query": "${QUERY_TEXT}"
}
```

В данном случае QUERY_TEXT — это переменная, в которой хранится текст запроса:

```
read -d '' QUERY_TEXT << EOF
SELECT
  start_station_name
  , AVG(duration) as duration
  , COUNT(duration) as num_trips
FROM `bigquery-public-data`.london_bicycles.cycle_hire
GROUP BY start_station_name
ORDER BY num_trips DESC
LIMIT 5
EOF
```

Мы использовали синтаксис встроенных документов (<http://tldp.org/LDP/abs/html/here-docs.html>), поддерживаемый в Bash, чтобы указать, что строка EOF отмечает точку, в которой заканчивается наш запрос.

Теперь в вызове curl нужно указать метод POST и переменную с текстом запроса:¹

```
curl -H "Authorization: Bearer $access_token" \
  -H "Content-Type: application/json" \
  -X POST \
  -d "$request" \
  "https://www.googleapis.com/bigquery/v2/projects/$PROJECT/queries"
```

Здесь \$request — это переменная, хранящая сообщение в формате JSON (с текстом запроса).

¹ См. файл 05_devel/rest_query.sh в репозитории GitHub с примерами для книги.

В ответ возвращается сообщение JSON, содержащее схему набора с результатами и пять записей, каждая из которых является массивом значений. Вот схема, которая была получена:

```
"schema": {
  "fields": [
    {
      "name": "start_station_name",
      "type": "STRING",
      "mode": "NULLABLE"
    },
    {
      "name": "duration",
      "type": "FLOAT",
      "mode": "NULLABLE"
    },
    {
      "name": "num_trips",
      "type": "INTEGER",
      "mode": "NULLABLE"
    }
  ]
},
```

А вот так выглядит первая запись:

```
{
  "f": [
    {
      "v": "Belgrove Street , King's Cross"
    },
    {
      "v": "1011.0766960393793"
    },
    {
      "v": "234458"
    }
  ]
},
```

Символ *f* обозначает *field* (поле), а *v* — *value* (значение). Каждая запись — это массив полей, и каждое поле имеет значение. Эта запись означает, что наибольшее количество поездок было зафиксировано от пункта проката на Белгроув-стрит (Belgrove Street), средняя продолжительность поездки составила 1011 секунд, а общее количество поездок — 234 458.

Ограничения

В данном примере запрос успевает выполниться в течение периода ожидания по умолчанию (можно указать более длительное время ожидания), но что, если

запросу потребуется больше времени? Давайте смоделируем эту ситуацию, искусственно уменьшив время ожидания и отключив кеш:¹

```
{
  "useLegacySql": false,
  "timeoutMs": 0,
  "useQueryCache": false,
  "query": "\${QUERY_TEXT}"
}
```

Теперь ответ не содержит записей с результатами. Вместо этого мы получили расписку в виде `jobId`:

```
{
  "kind": "bigquery#queryResponse",
  "jobReference": {
    "projectId": "cloud-training-demos",
    "jobId": "job_gv0Kq8nWzXikuBwoxsKMctJIVbX4",
    "location": "EU"
  },
  "jobComplete": false
}
```

Теперь предполагается, что мы попытаемся получить состояние `jobId` с помощью REST API, отправив запрос GET на URL:

```
.../projects/<PROJECT>/jobs/<JOBID>
```

Такой опрос следует продолжать, пока в ответ не будет получено значение `true` в `JobComplete`. После этого можно получить результаты запроса, отправив запрос GET на URL:

```
.../projects/<PROJECT>/queries/<JOBID>
```

Иногда результаты запроса имеют слишком большой объем для отправки в одном HTTP-ответе. В таких случаях результаты предоставляются фрагментами. Напомним, однако, что REST — это протокол без сохранения состояния, и сервер не поддерживает контекст сеанса. Поэтому в действительности результаты сохраняются во временной таблице, которая продолжает существовать в течение 24 часов. Клиент может просматривать эту временную таблицу, используя маркер страницы, который служит закладкой для каждого запроса GET для получения результатов запроса.

В дополнение ко всем этим сложностям добавьте возможность сбоя сети и необходимость повторных попыток, и станет ясно, что REST API довольно сложен в обращении. Поэтому, даже при том что REST API доступен на любом языке, на

¹ См. `04_devel/rest_query_async.sh` в репозитории GitHub с примерами для книги.

котором можно запрограммировать отправку запросов веб-службам, мы обычно рекомендуем использовать API более высокого уровня.

Google Cloud Client Library

Для программного доступа к BigQuery рекомендуется использовать клиентскую библиотеку Google Cloud Client Library для BigQuery. На момент написания этой книги клиентская библиотека была доступна для семи языков программирования: Go, Java, Node.js, Python, Ruby, PHP и C++. Все они имеют очень удобный интерфейс для разработчика и следуют соглашениям и стилю программирования, характерным для своего языка программирования.

Установить клиентскую библиотеку BigQuery для Python можно с помощью `pip` (или `easy_install`):

```
pip install google-cloud-bigquery
```

Чтобы использовать библиотеку, сначала создайте экземпляр клиента (он позаботится об аутентификации (<https://cloud.google.com/docs/authentication/production>) с помощью токена доступа при непосредственном вызове REST API):

```
from google.cloud import bigquery
bq = bigquery.Client(project=PROJECT)
```

В параметре `project` конструктору `Client` передается глобально уникальное имя проекта, со счета которого будут оплачиваться операции, выполненные с помощью объекта `bq`.



Весь программный код на Python из этого раздела можно найти в блокноте https://github.com/GoogleCloudPlatform/bigquery-oreilly-book/blob/master/05_devel/bigquery_cloud_client.ipynb. Используйте его в качестве источника фрагментов кода, чтобы испытать их в среде Python по своему выбору.

Документация с описанием API клиентской библиотеки BigQuery доступна по адресу: <https://googleapis.github.io/google-cloud-python/latest/bigquery/reference.html>. Мы не можем охватить весь API в одном разделе, поэтому настоятельно рекомендуем открыть документацию в браузере и обращаться к ней в процессе чтения следующего раздела. Читая фрагменты кода на Python, обращайтесь к документации, чтобы узнать, что делают те или иные методы, вызываемые в них.

Управление набором данных

Чтобы получить информацию о наборе данных с использованием клиентской библиотеки BigQuery, можно воспользоваться методом `get_dataset`:

```
dsinfo = bq.get_dataset('bigquery-public-data.london_bicycles')
```

Если имя проекта не указано, используется проект, переданный в конструктор `Client`; например:

```
dsinfo = bq.get_dataset('ch04')
```

Этот вызов вернет объект с информацией о наборе данных, созданном в предыдущей главе.

Информация о наборе данных. Объект `dsinfo` имеет множество атрибутов с информацией о наборе данных. Например, инструкции

```
print(dsinfo.dataset_id)
print(dsinfo.created)
```

выведут

```
ch04
2019-01-26 00:41:01.350000+00:00
```

а инструкция

```
print('{} created on {} in {}'.format(
    dsinfo.dataset_id, dsinfo.created, dsinfo.location))
```

выведет для набора данных `bigquery-public-data.london_bicycles`:

```
london_bicycles created on 2017-05-25 13:26:18.055000+00:00 in EU
```

Также с помощью объекта `dsinfo` можно проверить привилегии доступа к набору данных. Например, вот как можно определить, каким ролям предоставлен доступ `READER` к набору данных `london_bicycles`:

```
for access in dsinfo.access_entries:
    if access.role == 'READER':
        print(access)
```

Этот код выведет следующее:

```
<AccessEntry: role=READER, specialGroup=allAuthenticatedUsers>
<AccessEntry: role=READER, domain=google.com>
<AccessEntry: role=READER, specialGroup=projectReaders>
```

Именно потому, что всем аутентифицированным пользователям предоставлен доступ к набору данных (см. первую строку в предыдущем примере), мы смогли в предыдущих главах обращаться к нему.

Создание набора данных. Создать набор данных с именем `ch05` можно так:

```
dataset_id = "{}.ch05".format(PROJECT)
ds = bq.create_dataset(dataset_id, exists_ok=True)
```

По умолчанию набор данных создается в регионе US. Чтобы создать его в другом регионе, например в EU, создайте локальный объект `Dataset` (здесь мы дали ему имя `dsinfo`), установите его атрибут `location`, а затем вызовите метод `create_dataset` для объекта клиента, передав ему этот набор данных (вместо `dataset_id`, как в предыдущем фрагменте кода):

```
dataset_id = "{}.ch05eu".format(PROJECT)
dsinfo = bigquery.Dataset(dataset_id)
dsinfo.location = 'EU'
ds = bq.create_dataset(dsinfo, exists_ok=True)
```

Удаление набора данных. Удалить набор данных `ch05` из проекта, указанного при создании `Client`, можно так:

```
bq.delete_dataset('ch05', not_found_ok=True)
```

Чтобы удалить набор данных из другого проекта, нужно указать имя этого набора данных, квалифицированное именем проекта:

```
bq.delete_dataset('{}.ch05'.format(PROJECT), not_found_ok=True)
```

Изменение атрибутов набора данных. Чтобы изменить информацию о наборе данных, достаточно изменить атрибуты локального объекта `dsinfo` и затем вызвать метод `update_dataset` объекта клиента, чтобы переслать изменения в BigQuery:

```
dsinfo = bq.get_dataset("ch05")
print(dsinfo.description)
dsinfo.description = "Chapter 5 of BigQuery: The Definitive Guide"
dsinfo = bq.update_dataset(dsinfo, ['description'])
print(dsinfo.description)
```

Первая инструкция `print` в предыдущем фрагменте выведет `None`, потому что набор данных `ch05` был создан без описания. После вызова `update_dataset` набор данных в BigQuery получит новое описание:

```
None
Chapter 5 of BigQuery: The Definitive Guide
```

Теги, привилегии доступа и другие атрибуты набора данных изменяются аналогично. Например, вот как можно предоставить доступ к набору данных `ch05` одному из наших коллег:

```
dsinfo = bq.get_dataset("ch05")
entry = bigquery.AccessEntry(
    role="READER",
    entity_type="userByEmail",
    entity_id="xyz@google.com",
)
```

```

if entry not in dsinfo.access_entries:
    entries = list(dsinfo.access_entries)
    entries.append(entry)
    dsinfo.access_entries = entries
    dsinfo = bq.update_dataset(dsinfo, ["access_entries"]) # обращение к API
else:
    print('{} already has access'.format(entry.entity_id))
print(dsinfo.access_entries)

```

Здесь мы создаем запись для пользователя, и если пользователь пока не имеет никаких привилегий на доступ к набору данных, мы получаем текущий набор записей доступа, добавляем новую запись в конец и посылаем информацию о наборе данных с обновленным списком.

Управление таблицами

Получить список таблиц в наборе данных можно вызовом метода `list_tables` объекта клиента:

```

tables = bq.list_tables("bigquery-public-data.london_bicycles")
for table in tables:
    print(table.table_id)

```

Этот код сообщит о двух таблицах в наборе данных `london_bicycles`:

```

cycle_hire
cycle_stations

```

Получение свойств таблиц. В предыдущем фрагменте кода мы извлекали `table_id` из объекта `table`. Помимо `table_id`, в этом объекте имеются также другие атрибуты, характеризующие таблицу: количество записей, описание, теги, схема и многое другое.



Количество записей в таблице является частью ее метаданных, которые можно получить из самого объекта таблицы. В отличие от полноценного запроса с `COUNT(*)`, получение количества записей таким способом не требует оплаты услуг BigQuery:

```

table = bq.get_table(
    "bigquery-public-data.london_bicycles.cycle_stations")
print('{} rows in {}'.format(table.num_rows,
    table.table_id))

```

Этот фрагмент выведет:

```
787 rows in cycle_stations
```

Например, вот как с помощью объекта `table` можно найти в схеме столбцы, имена которых содержат определенную подстроку (`count`):

```

table = bq.get_table(
    "bigquery-public-data.london_bicycles.cycle_stations")
for field in table.schema:
    if 'count' in field.name:
        print(field)

```

В результате этот фрагмент выведет:

```

SchemaField('bikes_count', 'INTEGER', 'NULLABLE', '', ())
SchemaField('docks_count', 'INTEGER', 'NULLABLE', '', ())

```

Конечно, вместо поиска вручную проще использовать INFORMATION_SCHEMA (как описывается в главе 8) или Data Catalog.

Удаление таблиц. Удаление таблицы производится аналогично удалению наборов данных, и при желании можно игнорировать ошибку, если таблица не существовала перед удалением:

```

bq.delete_table('ch05.temp_table', not_found_ok=True)

```

ВОССТАНОВЛЕНИЕ УДАЛЕННЫХ ТАБЛИЦ

Для восстановления удаленных таблиц можно использовать возможность BigQuery «Time Travel». Она позволяет в течение двух дней восстановить таблицу, если та была случайно удалена. Например, чтобы восстановить версию таблицы, которая существовала в определенное время не более семи дней назад, можно создать ее копию, указав отметку времени:

```

bq --location=US cp ch05.temp_table@1418864998000
ch05.temp_table2

```

Здесь 1418864998000 — это отметка времени (число секунд, прошедших с начала эпохи).

Но имейте в виду, что моментальные снимки удаляются, если после удаления в наборе данных была создана таблица с таким же идентификатором или если был удален вмещающий ее набор данных. Учитывайте это в своем рабочем процессе — вы сэкономите массу нервов, минимизировав вероятность создания таблиц с одинаковыми именами из разных программных приложений. Например, выбирая имена для наборов данных, вы можете использовать организационные границы (например, «accounting» — «бухгалтерия») или названия приложений и рабочих нагрузок (например, «shipping» — «отгрузка»). Также можно предотвращать создание таблиц из своих приложений, создавая их извне, до запуска этих приложений.

Создание пустых таблиц. Создание пустой таблицы выполняется аналогично созданию набора данных, и при желании можно игнорировать ошибку, генерируемую в случае, если таблица с таким именем уже существует:

```
table_id = '{}.ch05.temp_table'.format(PROJECT)
table = bq.create_table(table_id, exists_ok=True)
```

Изменение схемы таблицы. Обычно редко приходится создавать абсолютно пустые таблицы. Чаще создаются таблицы со схемой и с несколькими записями. Поскольку схема является частью набора атрибутов таблицы, схему пустой таблицы можно изменить, подобно тому как мы изменяли контроль доступа для набора данных. Для этого нужно получить объект таблицы, изменить его локально, а затем отправить измененный объект в BigQuery:

```
schema = [
    bigquery.SchemaField("chapter", "INTEGER", mode="REQUIRED"),
    bigquery.SchemaField("title", "STRING", mode="REQUIRED"),
]
table_id = '{}.ch05.temp_table'.format(PROJECT)
table = bq.get_table(table_id)
print(table.etag)
table.schema = schema
table = bq.update_table(table, ["schema"])
print(table.schema)
print(table.etag)
```

Чтобы предотвратить гонку за ресурсами, BigQuery снабжает таблицу тегом при каждом обновлении. То есть метод `get_table` возвращает объект таблицы с атрибутом `etag`. При попытке выгрузить измененную схему с использованием `update_table` эта попытка будет успешной, только если атрибут `etag` измененной таблицы совпадет с атрибутом `etag` таблицы на сервере. Возвращаемый объект таблицы будет иметь новое значение в атрибуте `etag`. Это поведение можно изменить и принудительно обновить схему, присвоив атрибуту `table.etag` значение `None`.

Если таблица пустая, схему в ней можно изменить как угодно. Но когда в таблице есть данные, любые изменения схемы должны быть совместимы с существующими данными в таблице. Вы можете добавить новые поля (если они определены как `NULLABLE`) и ослабить ограничение `REQUIRED` до `NULLABLE`.

После выполнения предыдущего фрагмента можно открыть веб-интерфейс BigQuery и убедиться, что вновь созданная таблица имеет правильную схему, как показано на рис. 5.2.

Вставка записей в таблицу. После создания таблицы со схемой в ней с помощью клиента можно добавить записи. Записи оформляются в виде кортежей Python с полями, следующими в том же порядке, что и соответствующие поля в определении схемы:

```
rows = [
    (1, u'What is BigQuery?'),
    (2, u'Query essentials'),
]
errors = bq.insert_rows(table, rows)
```


temp_table		COPY TABLE	DELETE TABLE
Schema	Details	Preview	
Field name	Type	Mode	Description
chapter	INTEGER	REQUIRED	
title	STRING	REQUIRED	

Рис. 5.2. Схема вновь созданной таблицы

Если все записи добавились успешно, `errors` будет содержать пустой список. А теперь взгляните, что получится, если в поле `chapter` передать нецелочисленное значение:

```
rows = [
    ('3', u'Operating on data types'),
    ('wont work', u'This will fail'),
    ('4', u'Loading data into BigQuery'),
]
errors = bq.insert_rows(table, rows)
print(errors)
```

Вы получите сообщение с полем `reason` (причина), содержащим строку `'invalid'` (недопустимое значение), с полем `index=1` (вторая запись; счет начинается с 0) и с полем `location=chapter`, определяющим имя столбца:

```
{'index': 1, 'errors': [{'reason': 'invalid', 'debugInfo': '', 'message':
    'Cannot
convert value to integer (bad value):wont work', 'location': 'chapter'}]}
```

Поскольку BigQuery интерпретирует каждый запрос как атомарный, ни одна из трех записей не будет добавлена в таблицу. Для других записей вы получите сообщения с полем `reason`, содержащим текст `'stopped'`:

```
{'index': 0, 'errors': [{'reason': 'stopped', 'debugInfo': '', 'message': '',
    'location': ''}]}
```

В веб-интерфейсе BigQuery, как показано на рис. 5.3, можно заметить, что в потоковом буфере находятся две записи, вставленные ранее, но они не отражаются в общем числе записей в таблице.

temp_table

QUERY TA

Schema

Details

Preview

Description

None

Table info

Table ID	cloud-training-demos:ch05.temp_table
Table size	0 B
Number of rows	0
Created	Mar 3, 2019, 10:46:40 AM
Table expiration	Never
Last modified	Mar 3, 2019, 10:48:23 AM
Data location	US

Streaming buffer statistics

Estimated size	41 B
Estimated rows	2
Earliest entry time	Mar 3, 2019, 10:48:00 AM

Рис. 5.3. Вновь вставленные записи находятся в потоковом буфере и пока не отражаются в общем числе записей в таблице, в разделе «Table info» («Информация о таблице»)



Вставка записей в таблицу является потоковой операцией, поэтому метаданные таблицы обновляются не сразу, например:

```
rows = [
    (1, u'What is BigQuery?'),
    (2, u'Query essentials'),
]
print(table.table_id, table.num_rows)
errors = bq.insert_rows(table, rows)
print(errors)
table = bq.get_table(table_id)
print(table.table_id, table.num_rows) # МЕТАДААННЫЕ НЕ ИЗМЕНИЛИСЬ
```

Значение в `table.num_rows` в этом фрагменте кода *не* будет отражать обновленное количество записей. Кроме того, потоковые вставки, в отличие от заданий загрузки, выполняются не бесплатно.

Однако если выполнить запрос, он определит присутствие двух записей в потоковом буфере:

```
SELECT DISTINCT(chapter) FROM ch05.temp_table
```

Он покажет, что в таблице присутствуют две записи:

Row	Chapter
1	2
2	1

Создание пустой таблицы со схемой. Вместо того чтобы создавать таблицу и затем изменять ее схему, можно просто передать схему при создании таблицы:

```
schema = [
    bigquery.SchemaField("chapter", "INTEGER", mode="REQUIRED"),
    bigquery.SchemaField("title", "STRING", mode="REQUIRED"),
]
table_id = '{}.ch05.temp_table2'.format(PROJECT)
table = bigquery.Table(table_id, schema)
table = bq.create_table(table, exists_ok=True)
print('{} created on {}'.format(table.table_id, table.created))
print(table.schema)
```

Получившаяся в результате таблица будет иметь требуемую схему:

```
temp_table2 created on 2019-03-03 19:30:18.324000+00:00
[SchemaField('chapter', 'INTEGER', 'REQUIRED', None, ()), SchemaField('title',
'String', 'REQUIRED', None, ())]
```

Созданная таблица не содержит записей, и этот прием удобно использовать, когда предполагается выполнить потоковую вставку записей в таблицу. Но что, если у вас уже есть файл с данными и вы хотите создать таблицу, инициализировав ее содержимым этого файла? В этом случае гораздо удобнее использовать задания загрузки. Кроме того, в отличие от потоковой вставки, BigQuery не взимает плату за загрузку.

Клиент BigQuery на Python поддерживает три метода загрузки данных: из `DataFrame` библиотеки `pandas`, из URI и из локального файла. Рассмотрим все три способа по порядку.

Загрузка из объекта `DataFrame` библиотеки `pandas`. `pandas` (<https://pandas.pydata.org/>) — это библиотека с открытым исходным кодом, которая предлагает структуры данных и инструменты анализа на языке программирования Python. Клиентская библиотека BigQuery для Python поддерживает прямую загрузку данных из объектов `DataFrame`, находящихся в памяти. Поскольку объекты `DataFrame` могут конструироваться из структур данных в памяти и обеспечивают широкий спектр преобразований, использование библиотеки `pandas` обеспечивает наиболее удобный способ загрузки данных из приложений на Python. Например, вот как можно создать `DataFrame` из массива кортежей:

```
import pandas as pd
data = [
    (1, u'What is BigQuery?'),
```

```
(2, u'Query essentials'),
]
df = pd.DataFrame(data, columns=['chapter', 'title'])
```

После создания экземпляра `DataFrame` данные из него можно загрузить в таблицу BigQuery, как показано ниже:¹

```
table_id = '{}.ch05.temp_table3'.format(PROJECT)
job = bq.load_table_from_dataframe(df, table_id)
job.result() # заблокирует программу в ожидании ответа
print("Loaded {} rows into {}".format(job.output_rows,
tblref.table_id))
```

Задания загрузки могут выполняться довольно долго, поэтому функция `load_table_` возвращает объект задания, завершение которого можно периодически проверять вызовом метода `job.done()` или заблокировать выполнение программы до его завершения вызовом `job.result()`.

Если таблица уже существует, задание загрузки добавит новые записи в имеющуюся таблицу, при условии, что загружаемые данные соответствуют схеме. Если таблица не существует, будет создана новая таблица со схемой, полученной на основе информации из `DataFrame`.² Это поведение можно изменить, определив конфигурацию загрузки:

```
from google.cloud.bigquery.job \
    import LoadJobConfig, WriteDisposition, CreateDisposition
load_config = LoadJobConfig(
    create_disposition=CreateDisposition.CREATE_IF_NEEDED,
    write_disposition=WriteDisposition.WRITE_TRUNCATE)
job = bq.load_table_from_dataframe(df, table_id,
job_config=load_config)
```

Комбинация флагов `CreateDisposition` и `WriteDisposition` управляет поведением задания загрузки, как показано в табл. 5.2.

Таблица 5.2. Влияние флагов `CreateDisposition` и `WriteDisposition` на поведение задания загрузки

CreateDisposition	WriteDisposition	Поведение
CREATE_NEVER	WRITE_APPEND	Добавляет записи в конец существующей таблицы
	WRITE_EMPTY	Добавляет записи в конец таблицы, но только если она в данный момент пустая. В противном случае возвращает ошибку, сообщающую о попытке добавления повторяющихся записей

¹ Для этого потребуется библиотека `ruarrow`. Если у вас ее еще нет, установите ее командой `pip install ruarrow`.

² Поскольку `pandas` по умолчанию сортирует имена столбцов в алфавитном порядке, таблица BigQuery получит схему со столбцами, следующими в алфавитном порядке, а не в том порядке, в каком они указаны в кортежах.

CreateDisposition	WriteDisposition	Поведение
CREATE_IF_NEEDED	WRITE_APPEND	Создает новую таблицу, используя указанную схему, если имеется. Добавляет записи в конец существующей или новой таблицы. Это поведение по умолчанию, если параметр <code>job_config</code> не был передан
	WRITE_EMPTY	Создает новую таблицу, используя указанную схему, если имеется. Требуется, чтобы таблица была пустой, если она уже существует. В противном случае возвращает ошибку, сообщающую о попытке добавления повторяющихся записей
	WRITE_TRUNCATE	Создает новую таблицу, используя указанную схему, если имеется. Удаляет все записи, уже имеющиеся в таблице, то есть задание полностью затирает старые данные

Загрузка из URI. Загрузить данные в таблицу BigQuery можно непосредственно из файла с известным Google Cloud URI. Кроме резервных копий в Cloud Datastore URL с адресами в Cloud Bigtable, поддерживаются также подстановочные символы Google Cloud Storage.¹ То есть мы можем загрузить файл в формате CSV с оценочными параметрами колледжей, который мы использовали в предыдущей главе:

```
job_config = bigquery.LoadJobConfig()
job_config.autodetect = True
job_config.source_format = bigquery.SourceFormat.CSV
job_config.null_marker = 'NULL'
uri = "gs://bigquery-oreilly-book/college_scorecard.csv"
table_id = '{}.ch05.college_scorecard_gcs'.format(PROJECT)
job = bq.load_table_from_uri(uri, table_id, job_config=job_config)
```

При этом можно устанавливать все параметры, которые мы рассмотрели в главе 4 (имеющие отношение к загрузке данных), используя флаги `JobConfig`. Здесь мы используем автоопределение схемы, указываем формат файла — CSV и то, что в файле используется нестандартный способ обозначения значений NULL.

После отправки задания можно заблокировать программу в ожидании его завершения, как было показано в предыдущем разделе, но можно опрашивать состояние задания каждые 0.1 секунды и извлечь информацию о таблице только после завершения этого задания:

```
while not job.done():
    print('.', end='', flush=True)
    time.sleep(0.1)
print('Done')
table = bq.get_table(tblref)
print("Loaded {} rows into {}".format(table.num_rows, table.table_id))
```

¹ Полный список поддерживаемых URI можно найти на странице <https://cloud.google.com/bigquery/docs/reference/rest/v2/jobs#configuration.load.sourceUris>.

У нас этот код успел вывести семь точек, представляющих состояние ожидания, после чего получил количество загруженных записей (7175).

Загрузка из локального файла. Чтобы загрузить данные из локального файла, создайте объект файла и передайте его в вызов `load_table_from_file`:

```
with gzip.open('../04_load/college_scorecard.csv.gz') as fp:
    job = bq.load_table_from_file(fp, tblref, job_config=job_config)
```

В остальном этот способ ничем не отличается от загрузки из URI.

Копирование таблиц. Скопировать таблицу из одного набора в другой можно с помощью метода `copy_table`:

```
source_tbl = 'bigquery-public-data.london_bicycles.cycle_stations'
dest_tbl = '{}.ch05eu.cycle_stations_copy'.format(PROJECT)
job = bq.copy_table(source_tbl, dest_tbl, location='EU')
job.result() # blocks and waits
dest_table = bq.get_table(dest_tbl)
print(dest_table.num_rows)
```

Обратите внимание, что мы копируем данные из таблицы `cycle_stations` в набор данных (`ch05eu`), созданный в регионе EU. Также отметьте, что мы стараемся копировать таблицы только в пределах одного региона.

Извлечение данных из таблицы. Данные из таблицы можно экспортировать в файл в Google Cloud Storage, используя метод `extract_table`:

```
source_tbl = 'bigquery-public-data.london_bicycles.cycle_stations'
dest_uri = 'gs://{}/tmp/exported/cycle_stations'.format(BUCKET)
config = bigquery.job.ExtractJobConfig(
    destination_format =
        bigquery.job.DestinationFormat.NEWLINE_DELIMITED_JSON)
job = bq.extract_table(source_tbl, dest_uri,
    location='EU', job_config=config)
job.result() # блокирует программу в ожидании ответа
```

Если таблица достаточно большая, данные будут разделены на несколько файлов. На момент написания этой книги извлечение было возможно в следующие форматы: CSV, Avro и строчный JSON. Так же как при копировании таблиц, старайтесь экспортировать данные в корзину, находящуюся в том же регионе, что и набор данных.¹ Конечно, после экспорта таблицы Google Cloud Storage будет взимать плату за хранение получившихся файлов.

Просмотр записей в таблице. В веб-интерфейсе BigQuery есть возможность просмотреть содержимое таблицы бесплатно. Та же возможность доступна

¹ Создать корзину в регионе EU можно командой `gsutil mb -l EU gs://some-bucket-name`.

через REST API с использованием метода `tabledata.list`, и следовательно, через Python API.

Вот как можно вывести пять произвольных записей из таблицы `cycle_stations`:¹

```
table_id = 'bigquery-public-data.london_bicycles.cycle_stations'
table = bq.get_table(table_id)
rows = bq.list_rows(table,
                     start_index=0,
                     max_results=5)
```

Если опустить параметры `start_index` и `max_results`, можно получить все записи, имеющиеся в таблице:

```
rows = bq.list_rows(table)
```

Конечно, таблица должна быть достаточно маленькой, чтобы поместиться в памяти. Большую таблицу можно разбить на страницы и обрабатывать ее по частям:

```
page_size = 10000
row_iter = bq.list_rows(table,
                        page_size=page_size)
for page in row_iter.pages:
    rows = list(page)
    # выполнить необходимые действия с записями ...
    print(len(rows))
```

Также есть возможность ограничить число извлекаемых столбцов. Например, вот как можно извлечь столбец `id` и все столбцы с именами, содержащими подстроку `count`:

```
fields = [field for field in table.schema
          if 'count' in field.name or field.name == 'id']
rows = bq.list_rows(table,
                    start_index=300,
                    max_results=5,
                    selected_fields=fields)
```

Полученные записи можно отформатировать так, чтобы выделить для столбцов поля фиксированной ширины, например, в 10 символов:

```
fmt = '{!s:<10} ' * len(rows.schema)
print(fmt.format(*[field.name for field in rows.schema]))
for row in rows:
    print(fmt.format(*row))
```

¹ Какие пять строк мы получим — неизвестно, потому что BigQuery не гарантирует сохранения записей в определенном порядке. Цель параметра `start_index` состоит в том, чтобы мы могли получить «следующую страницу» из пяти строк, указав `start_index = 5`.

Этот код выведет следующее:

id	bikes_count	docks_count
658	20	30
797	20	30
238	21	32
578	22	32
477	26	36

Выполнение запросов

Главным достоинством клиентской библиотеки Google Cloud Client Library является поддержка запросов. Основные сложности, связанные с разбиением на страницы, повторными попытками и т. д., обрабатываются автоматически.

Первым шагом, конечно же, является создание строки с кодом SQL, который должна выполнить BigQuery:

```
query = """
SELECT
    start_station_name
    , AVG(duration) as duration
    , COUNT(duration) as num_trips
FROM `bigquery-public-data`.london_bicycles.cycle_hire
GROUP BY start_station_name
ORDER BY num_trips DESC
LIMIT 10
"""
```

Этот запрос найдет 10 самых популярных пунктов проката велосипедов в Лондоне согласно общему числу велосипедов, взятых на прокат, и сообщит название каждого пункта проката, среднюю продолжительность поездок, начатых от этого пункта проката, и общее количество таких поездок.

Пробный прогон. Перед фактическим выполнением запроса можно произвести пробный прогон, чтобы оценить объем данных, который будет обработан запросом:¹

```
config = bigquery.QueryJobConfig()
config.dry_run = True
job = bq.query(query, location='EU', job_config=config)
print("This query will process {} bytes."
      .format(job.total_bytes_processed))
```

Этот код выведет следующее:

This query will process 903989528 bytes.

¹ Это тот же API, который используется веб-интерфейсом BigQuery для отображения оценки.

У вас могут получиться другие результаты, потому что содержимое этой таблицы постоянно обновляется и дополняется.



Пробный прогон не оплачивается. Пробные прогоны можно использовать для проверки синтаксиса запроса во время разработки или тестирования, обнаружения необъявленных параметров и проверки схемы набора с результатами без фактического выполнения запроса. Если вы пишете приложение, отправляющее запросы в BigQuery, вы можете использовать пробный прогон для ограничения финансовых затрат. Вопросы оптимизации производительности и затрат мы рассмотрим в главе 7.

Иногда невозможно определить, какой объем данных придется обработать, без фактического выполнения запроса. В таких случаях пробный прогон возвращает либо ноль, либо оценку максимально возможного значения. Это происходит в двух ситуациях: при запросе федеративной таблицы (когда данные хранятся вне BigQuery; см. главу 4) и при запросе кластерной таблицы (см. главу 7). Для федеративной таблицы пробный прогон вернет оценку в 0 байт, а для кластерной таблицы BigQuery попытается вычислить наихудший сценарий и сообщит полученный результат. Но в любом случае при фактическом выполнении запроса вам будет выставлен счет только за фактически обработанный объем данных.

Выполнение запроса. Чтобы выполнить запрос, достаточно начать цикл по поддерживаемому объекту задания. Задание будет запущено, и в ходе итераций по его поддерживаемому с помощью цикла `for` из него будут извлекаться страницы результатов:

```
job = bq.query(query, location='EU')
fmt = '{!s:<40} {:>10d} {:>10d}'
for row in job:
    fields = (row['start_station_name'],
              (int)(0.5 + row['duration']),
              row['num_trips'])
    print(fmt.format(*fields))
```

Получив запись, из нее можно извлечь значение любого из столбцов, используя псевдонимы, указанные в `SELECT` (взгляните, например, как определен столбец `num_trips`).

Ниже показаны результаты этого запроса в отформатированном виде:

Belgrove Street, King's Cross	1011	234458
Hyde Park Corner, Hyde Park	2783	215629
Waterloo Station 3, Waterloo	866	201630
Black Lion Gate, Kensington Gardens	3588	161952
Albert Gate, Hyde Park	2359	155647
Waterloo Station 1, Waterloo	992	145910
Wormwood Street, Liverpool Street	976	119447
Hop Exchange, The Borough	1218	115135
Wellington Arch, Hyde Park	2276	110260
Triangle Car Park, Hyde Park	2233	108347

Создание экземпляра DataFrame. Выше в этом разделе мы показали, как загрузить данные в таблицу BigQuery из объекта DataFrame. Аналогично можно выполнить запрос и сохранить результаты в DataFrame, используя BigQuery в роли распределенного и масштабируемого механизма в процессе исследования данных:

```
query = """
SELECT
    start_station_name
    , AVG(duration) as duration
    , COUNT(duration) as num_trips
FROM `bigquery-public-data`.london_bicycles.cycle_hire
GROUP BY start_station_name
"""

df = bq.query(query, location='EU').to_dataframe()
print(df.describe())
```

Для вывода параметров распределения значений числовых столбцов в этом примере используется функция `describe()` из библиотеки:

	duration	num_trips
count	880.000000	880.000000
mean	1348.351153	27692.273864
std	434.057829	23733.621289
min	0.000000	1.000000
25%	1078.684974	13033.500000
50%	1255.889223	23658.500000
75%	1520.504055	35450.500000
max	4836.380090	234458.000000

Итак, всего имеется 880 пунктов проката и от каждого, в среднем, начинается 27 692 поездок, хотя есть пункт проката, от которого началась только одна поездка, и пункт проката, от которого начались 234 458 поездок. Медианное число поездок для одного пункта проката составило 23 658 поездок, и большинство пунктов обслуживают от 13 033 до 35 450 поездок.

Параметризованные запросы. Запросы не обязательно должны быть статическими строками — их можно параметризовать и задавать значения параметров в момент создания задания для выполнения запроса. Вот пример запроса, который находит общее количество поездок, продлившихся больше определенного времени. Фактический порог `min_duration` будет указываться во время запуска запроса:

```
query2 = """
SELECT
    start_station_name
    , COUNT(duration) as num_trips
FROM `bigquery-public-data`.london_bicycles.cycle_hire
WHERE duration >= @min_duration
GROUP BY start_station_name
ORDER BY num_trips DESC
LIMIT 10
"""
```

Символ @ определяет `min_duration` как параметр запроса. Запрос может включать сколько угодно таких именованных параметров.



Определение запросов с использованием механизма форматирования строк — крайне плохая идея. Подобные запросы, как показано ниже, могут дать возможность проведения атак вида «инъекция SQL»:

```
query2 = """
SELECT
    start_station_name
    , COUNT(duration) as num_trips
FROM `bigquery-public-data`.london_bicycles.cycle_hire
WHERE duration >= {}
GROUP BY start_station_name
ORDER BY num_trips DESC
LIMIT 10
""".format(min_duration)
```

Мы настоятельно рекомендуем использовать параметризованные запросы, особенно когда требуется включить в запрос пользовательский ввод.

Прежде чем запустить запрос с именованными параметрами, вам придется определить эти параметры и передать их в параметре `job_config`:

```
config = bigquery.QueryJobConfig()
config.query_parameters = [
    bigquery.ScalarQueryParameter('min_duration', "INT64", 600)
]
job = bq.query(query2, location='EU', job_config=config)
```

Здесь мы указали, что запрос должен подсчитать число поездок длительностью больше 600 секунд.

Как и прежде, для извлечения полученных записей требуется обойти содержимое задания в цикле. Каждая запись при этом будет иметь вид словаря, отображающего имена столбцов в значения:

```
fmt = '{!s:<40} {:>10d}'
for row in job:
    fields = (row['start_station_name'],
              row['num_trips'])
    print(fmt.format(*fields))
```

Этот код выведет следующее:

Hyde Park Corner, Hyde Park	203592
Belgrove Street, King's Cross	168110
Waterloo Station 3, Waterloo	148809
Albert Gate, Hyde Park	145794
Black Lion Gate, Kensington Gardens	137930
Waterloo Station 1, Waterloo	106092
Wellington Arch, Hyde Park	102770

Triangle Car Park, Hyde Park	99368
Wormwood Street, Liverpool Street	82483
Palace Gate, Kensington Gardens	80342

В этом разделе мы показали, как программно выполнять операции в BigQuery, будь то управление таблицами или наборами данных, запросы или потоковая вставка. Программные интерфейсы, особенно Google Cloud Client Library, — это то, что вы будете использовать при создании приложений, которым необходим доступ к BigQuery.

Однако в некоторых конкретных случаях есть возможность использовать абстракции более высокого уровня. Мы познакомимся с ними в следующем разделе.

Доступ к BigQuery из инструментов исследования данных

Блокноты произвели революцию в подходах к исследованию данных. Они являются примером *грамотного программирования* — парадигмы программирования, которая была представлена легендой информатики Дональдом Кнuthом (Donald Knuth). Согласно этой парадигме компьютерный код смешивается с заголовками, текстом, графиками и т. д. Благодаря этому блокнот служит одновременно исполняемой программой и интерактивным отчетом.

Jupyter — наиболее популярная платформа для блокнотов на разных языках программирования, включая Python. Блокнот в Jupyter представляет собой интерактивный веб-документ, в котором можно вводить и выполнять код. Вывод кода встраивается непосредственно в документ.

Блокноты в Google Cloud Platform

Чтобы создать блокнот в Google Cloud Platform (GCP), запустите виртуальную машину глубокого обучения Deep Learning Virtual Machine и получите URL-адрес Jupyter. Это можно сделать в разделе **AI Factory** (Фабрика искусственного интеллекта) в GCP Cloud Console или воспользоваться инструментом командной строки `gcloud`:¹

```
#!/bin/bash

IMAGE--image-family=tf-latest-cpu
INSTANCE_NAME=dlvm
MAIL=google-cloud-customer@gmail.com # CHANGE THIS

echo "Launching $INSTANCE_NAME"
```

¹ Это сценарий `05_devel/launch_notebook.sh` в репозитории GitHub с примерами для этой книги.

```

gcloud compute instances create ${INSTANCE_NAME} \
  --machine-type=n1-standard-2 \
  --scopes=https://www.googleapis.com/auth/cloudplatform,
https://www.googleapis.com/auth/userinfo.email \
  ${IMAGE} \
  --image-project=deeplearning-platform-release \
  --boot-disk-device-name=${INSTANCE_NAME} \
  --metadata="proxy-user-mail=${MAIL}"

echo "Looking for Jupyter URL on ${INSTANCE_NAME}"
while true; do
  proxy=$(gcloud compute instances describe ${INSTANCE_NAME} 2> /dev/null | grep
dot-datalab-vm)
  if [ -z "$proxy" ]
  then
    echo -n "."
    sleep 1
  else
    echo "done!"
    echo "$proxy"
    break
  fi
done

```

Открыв полученный URL в браузере (или щелкнув на ссылке **Open JupyterLab** (Открыть JupyterLab) в разделе **Notebooks** (Блокноты) в GCP Cloud Console), вы окажетесь в Jupyter. Щелкните на кнопке, чтобы создать блокнот для Python 3, и вы сможете испытать фрагменты кода.

В окружении Cloud AI Factory Notebook уже установлена клиентская библиотека Google Cloud Client Library для BigQuery, но если вы находитесь в какой-то другой среде Jupyter, установите эту библиотеку сами и загрузите необходимые расширения, выполнив следующий код в ячейке для кода:

```

!pip install google-cloud-bigquery
%load_ext google.cloud.bigquery

```

Строки в Jupyter Notebook, начинающиеся с восклицательного знака (!), выполняются с использованием командной оболочки, а строки, начинающиеся со знака процента (%), запускают расширения, которые также называют *magics*. То есть строка `pip` в предыдущем фрагменте выполняется из командной строки, а строка `load_ext` загружает BigQuery Magics.

Вы можете скопировать репозиторий с примерами для этой книги, щелкнув на значке Git (выделен на рис. 5.4) и клонировав репозиторий <https://github.com/GoogleCloudPlatform/bigquery-oreilly-book>, как показано на рис. 5.4.

Найдите и откройте блокнот *05_devel/magics.ipynb*, чтобы проверить код из этого раздела. Измените переменную `PROJECT` в блокноте, присвоив ей имя своего проекта. Затем выберите в меню пункт **Run ▶ Run All Cells** (Запустить ▶ Запустить все ячейки).

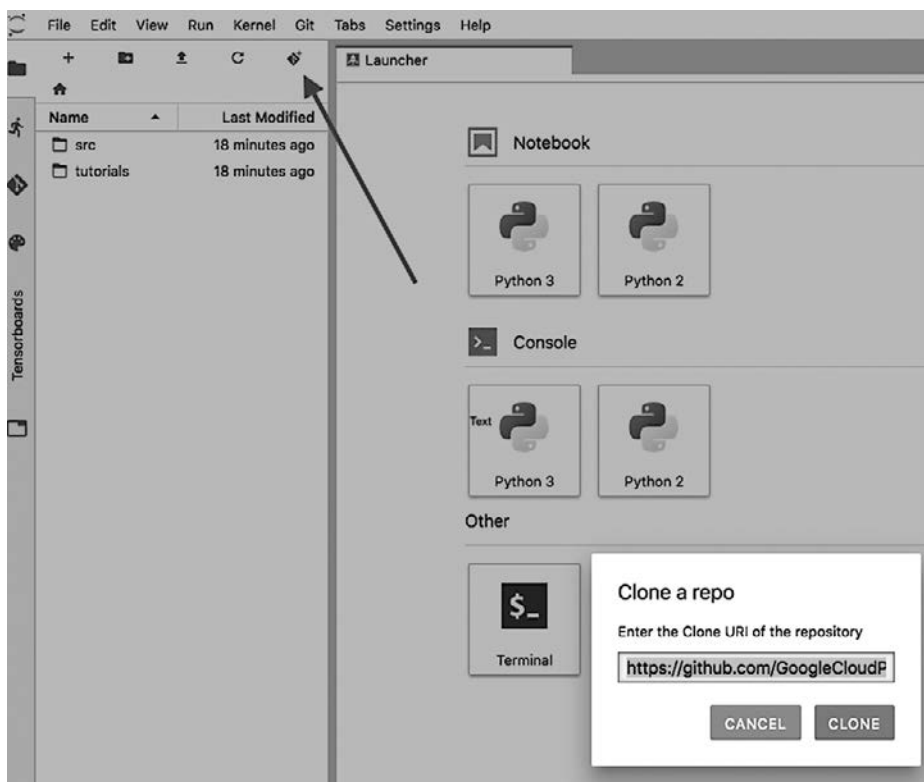


Рис. 5.4. Щелкните на значке Git (на который указывает стрелка), чтобы клонировать репозиторий

Jupyter Magics

Расширения BigQuery для Jupyter значительно упрощают выполнение запросов в блокноте. Например, чтобы выполнить запрос, достаточно ввести `%bigquery` в верхнюю часть ячейки:

```
%bigquery --project $PROJECT
SELECT
  start_station_name
  , AVG(duration) as duration
  , COUNT(duration) as num_trips
FROM `bigquery-public-data`.london_bicycles.cycle_hire
GROUP BY start_station_name
ORDER BY num_trips DESC
LIMIT 5
```

Если запустить ячейку с этим кодом, она выполнит запрос и отобразит красиво отформатированную таблицу с пятью строками, как показано на рис. 5.5.

Run a query			
[2]:	<pre>%%bigquery --project \$PROJECT SELECT start_station_name , AVG(duration) as duration , COUNT(duration) as num_trips FROM `bigquery-public-data`.london_bicycles.cycle_hire GROUP BY start_station_name ORDER BY num_trips DESC LIMIT 5</pre>		
[2]:	start_station_name	duration	num_trips
	0	Belgrove Street , King's Cross	1011.076696 234458
	1	Hyde Park Corner, Hyde Park	2782.730709 215629
	2	Waterloo Station 3, Waterloo	866.376135 201630
	3	Black Lion Gate, Kensington Gardens	3588.012004 161952
	4	Albert Gate, Hyde Park	2359.413930 155647

Рис. 5.5. Результат запроса, красиво отформатированный и встроенный в документ

Выполнение параметризованного запроса

Чтобы выполнить параметризованный запрос, добавьте параметр `--params` в команду вызова расширения, как показано на рис. 5.6. Фактические значения для параметров запроса хранятся в переменной Python, которая обычно определяется где-то в другом месте в блокноте.

Run a parameterized query			
[3]:	<pre>PARAMS = {"num_stations": 3}</pre>		
[4]:	<pre>%%bigquery --project \$PROJECT --params \$PARAMS SELECT start_station_name , AVG(duration) as duration , COUNT(duration) as num_trips FROM `bigquery-public-data`.london_bicycles.cycle_hire GROUP BY start_station_name ORDER BY num_trips DESC LIMIT @num_stations</pre>		
[4]:	start_station_name	duration	num_trips
	0	Belgrove Street , King's Cross	1011.076696 234458
	1	Hyde Park Corner, Hyde Park	2782.730709 215629
	2	Waterloo Station 3, Waterloo	866.376135 201630

Рис. 5.6. Так в блокноте выполняются параметризованные запросы

В предыдущем примере в переменной `PARAMS` определяется параметр `num_stations` (количество станций), который используется в запросе SQL для ограничения количества записей в результате.

Сохранение результатов в объекте DataFrame

Чтобы сохранить результаты запроса в объекте `DataFrame`, требуется указать имя переменной (например, `df`), которая ссылается на экземпляр `DataFrame`:

```
%%bigquery df --project $PROJECT
SELECT
  start_station_name
  , AVG(duration) as duration
  , COUNT(duration) as num_trips
FROM `bigquery-public-data`.london_bicycles.cycle_hire
GROUP BY start_station_name
ORDER BY num_trips DESC
```

Переменную `df` можно использовать как любой другой объект `DataFrame`. Например, можно получить статистические характеристики для числовых столбцов в `df`:

```
df.describe()
```

```
[7]: df.plot.scatter('duration', 'num_trips');
```

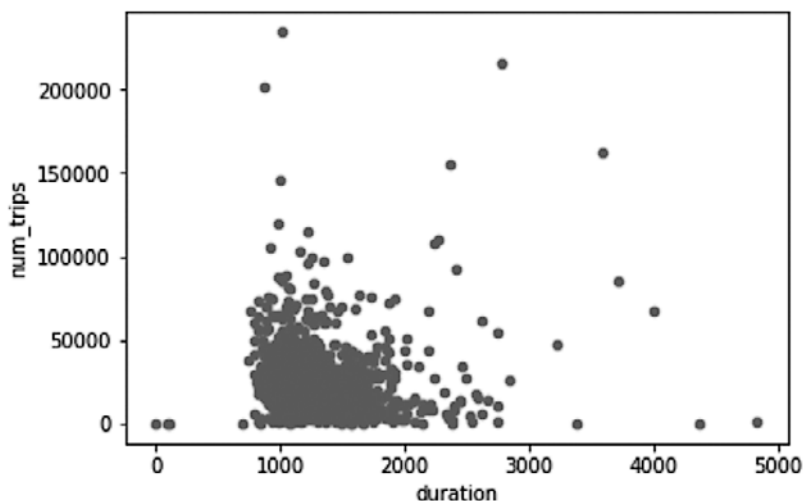


Рис. 5.7. Вывод диаграммы рассеяния с использованием DataFrame с данными, полученными из запроса к BigQuery

Также можно использовать команды построения графиков, доступные в библиотеке `pandas`, чтобы, например, нарисовать диаграмму рассеяния средней

продолжительности поездок и количества поездок по всем пунктам проката, как показано на рис. 5.7.

Работа с BigQuery, pandas и Jupyter

В разделах выше мы уже рассказали, как связаны между собой библиотеки Google Cloud Client Library для BigQuery и pandas. Поскольку pandas де-факто считается стандартом в сфере анализа данных на Python, будет полезно объединить все эти возможности и проиллюстрировать с их помощью типичный рабочий процесс обработки данных.

Представьте, что в службу поддержки поступили жалобы клиентов на неисправные велосипеды в некоторых пунктах проката. Мы решили провести выборочную проверку некоторых из них. Как выбрать объекты для выборочной проверки? Можно, конечно, довериться жалобам клиентов, где указаны такие пункты, но, как правило, больше всего жалоб будет поступать со ссылкой на наиболее популярные пункты проката — просто потому, что ими пользуется больше клиентов.

Мы посчитали, что если кто-то берет велосипед в прокат и возвращает его в тот же пункт менее чем на 10 минут, то, скорее всего, у велосипеда имеется какая-то неисправность. Назовем такой факт сорвавшейся поездкой (с точки зрения клиента, это действительно так). Мы могли бы провести выборочную проверку пунктов проката с наибольшей долей сорвавшихся поездок.

Чтобы найти долю сорвавшихся поездок, можно выполнить запрос к BigQuery с помощью Jupyter Magics и сохранить результат в `DataFrame` с именем `badtrips`, как показано ниже:

```
%%bigquery badtrips --project $PROJECT

WITH all_bad_trips AS (
SELECT
    start_station_name
    , COUNTIF(duration < 600 AND start_station_name = end_station_name)
      AS bad_trips
    , COUNT(*) as num_trips
FROM `bigquery-public-data`.london_bicycles.cycle_hire
WHERE EXTRACT(YEAR FROM start_date) = 2015
GROUP BY start_station_name
HAVING num_trips > 10
)
SELECT *, bad_trips / num_trips AS fraction_bad FROM all_bad_trips
ORDER BY fraction_bad DESC
```

Выражение `WITH` подсчитывает количество поездок, продолжительность которых менее 600 секунд и для которых начальный и конечный пункт проката совпадают. Группируя результаты по столбцу `start_station_name`, мы получаем общее количество поездок и сорвавшихся поездок для каждого пункта проката. Внеш-

ний запрос вычисляет желаемую долю и связывает ее с пунктом проката. В результате мы получаем следующее (показаны только первые несколько строк):

start_station_name	bad_trips	num_trips	fraction_bad
Contact Centre, Southbury House	20	48	0.416667
Monier Road, Newham	1	25	0.040000
Aberfeldy Street, Poplar	35	955	0.036649
Ormonde Gate, Chelsea	315	8932	0.035266
Thornfield House, Poplar	28	947	0.029567
...			

Очевидно, что пункт проката в верхней строке таблицы выбивается из общего ряда. Всего 48 поездок было совершено от пункта проката Southbury House, и 20 из них сорвавшиеся! Впрочем, мы можем подтвердить этот результат с помощью `pandas`, запросив статистику из `DataFrame`:

```
badtrips.describe()
```

Эта инструкция вернет следующие результаты:

	bad_trips	num_trips	fraction_bad
count	823.000000	823.000000	823.000000
mean	75.074119	11869.755772	0.007636
std	70.512207	9906.268656	0.014739
min	0.000000	11.000000	0.000000
25%	41.000000	5903.000000	0.005002
50%	62.000000	9998.000000	0.006368
75%	91.500000	14852.500000	0.008383
max	967.000000	95740.000000	0.416667

Глядя на результаты, можно заметить, что значение `fraction_bad` находится в диапазоне от 0 до 0.417 (посмотрите на значения `min` и `max`), но неясно, насколько уместно это соотношение, потому что пункты проката также довольно сильно различаются по числу поездок. Например, количество поездок колеблется от 11 до 95 740.

Построим диаграмму рассеяния, возможно, на ней мы увидим какую-нибудь четкую тенденцию:

```
badtrips.plot.scatter('num_trips', 'fraction_bad');
```

Результаты показаны на рис. 5.8.

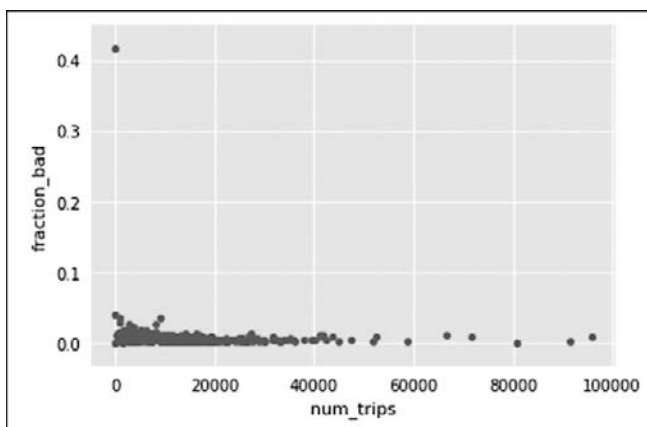


Рис. 5.8. На этой диаграмме видно, что более высокие значения `fraction_bad` наблюдаются в пунктах проката с низкими значениями `num_trips`

На диаграмме видно, что более высокие значения `fraction_bad` наблюдаются в пунктах проката с низкими значениями `num_trips`, но общая тенденция искажается аномальным значением 0.4. Давайте немного увеличим масштаб и добавим прямую наилучшего приближения, используя пакет построения графиков `seaborn`:

```
import seaborn as sns
ax = sns.regplot(badtrips['num_trips'], badtrips['fraction_bad']);
ax.set_ylim(0, 0.05);
```

Как показывает диаграмма на рис. 5.9, между долей сорвавшихся поездок и популярностью пункта проката есть четкая связь.

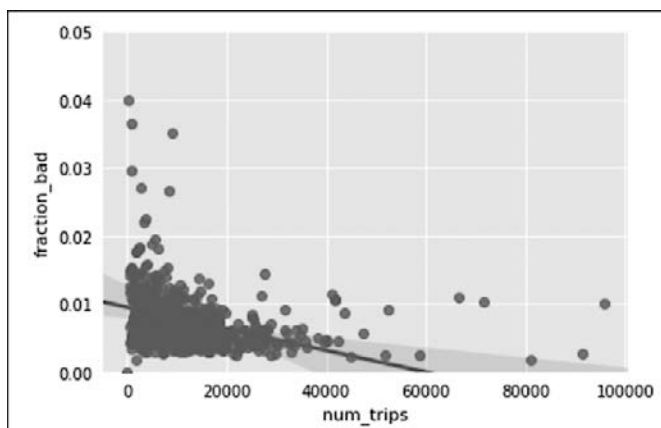


Рис. 5.9. Ясно видно, что более высокие значения `fraction_bad` наблюдаются в пунктах проката с низкими значениями `num_trips`

Поскольку более высокие значения `fraction_bad` наблюдаются в пунктах проката с низкими значениями `num_trips`, мы должны направить проверяющих вовсе не на пункты проката с высокими значениями `fraction_bad`. Итак, как выбрать пункты проката для выборочной проверки?

Можно выбрать пять худших из наиболее популярных пунктов проката, пять из пользующихся меньшей популярностью и т. д. Для этого разделим пункты проката на четыре группы по значению `num_trips`, а затем найдем пять худших пунктов в каждой из них. Этот алгоритм реализует следующий код:

```
stations_to_examine = []
for band in range(1,5):
    min_trips = badtrips['num_trips'].quantile(0.2*(band))
    max_trips = badtrips['num_trips'].quantile(0.2*(band+1))
    query = 'num_trips >= {} and num_trips < {}'.format(
        min_trips, max_trips)

    print(query) # квантиль
    stations = badtrips.query(query)
    stations = stations.sort_values(
        by=['fraction_bad'], ascending=False)[:5]
    print(stations) # 5 худших
    stations_to_examine.append(stations)
print()
```

Первая группа, охватывающая пункты проката между 20 и 40 процентилями:

```
num_trips >= 4826.4 and num_trips < 8511.8
```

	start_station_name	bad_trips	num_trips	fraction_bad
6	River Street, Clerkenwell	221	8279	0.026694
9	Courland Grove, Wandsworth Road	105	5369	0.019557
10	Stanley Grove, Battersea	92	4882	0.018845
12	Southern Grove, Bow	112	6152	0.018205
18	Richmond Way, Shepherd's Bush	126	8149	0.015462

Последняя группа, охватывающая пункты проката между 80 и 100 проценталями:

```
num_trips >= 16509.2 and num_trips < 95740.0
```

	start_station_name	bad_trips	num_trips	fraction_bad
25	Queen's Gate, Kensington Gardens	396	27457	0.014423
74	Speakers' Corner 2, Hyde Park	468	41107	0.011385
76	Cumberland Gate, Hyde Park	303	26981	0.011230
77	Albert Gate, Hyde Park	729	66547	0.010955
82	Triangle Car Park, Hyde Park	454	41675	0.010894

Обратите внимание, что в первой группе список завершает пункт проката с долей сорвавшихся поездок 0.015, а в последней — с долей 0.01. Эти малые значения могут вас успокоить, но объективно разница составляет 50%.

Теперь можно задействовать возможности `pandas`, чтобы объединить полученные данные с BigQuery API и записать эти пункты проката обратно в BigQuery:

```
stations_to_examine = pd.concat(stations_to_examine)
bq = bigquery.Client(project=PROJECT)
tblref = TableReference.from_string(
    '{}.ch05eu.bad_bikes'.format(PROJECT))
job = bq.load_table_from_dataframe(stations_to_examine, tblref)
job.result() # blocks and waits
```

Мы сохранили список кандидатов для проверки в хранилище, но нам также необходимо передать эти данные нашей команде. Лучшим форматом для этого является карта, и ее можно создать на Python, если знать широту и долготу наших пунктов проката. А они известны — географические координаты пунктов проката хранятся в таблице `cycle_stations`:¹

```
%bigquery stations_to_examine --project $PROJECT
SELECT
  start_station_name AS station_name
  , num_trips
  , fraction_bad
  , latitude
  , longitude
FROM ch05eu.bad_bikes AS bad
JOIN `bigquery-public-data`.london_bicycles.cycle_stations AS s
ON bad.start_station_name = s.name
```

Вот получившиеся результаты (показаны не все записи):

station_name	num_trips	fraction_bad	latitude	longitude
Ormonde Gate, Chelsea	8932	0.035266	51.487964	-0.161765
Stanley Grove, Battersea	4882	0.018845	51.470475	-0.152130
Courland Grove, Wandsworth Road	5369	0.019557	51.472918	-0.132103
Southern Grove, Bow	6152	0.018205	51.523538	-0.030556
...				

Имея информацию о местоположении, можно построить карту с помощью пакета `folium`:

```
import folium
map_pts = folium.Map(location=[51.5, -0.15], zoom_start=12)
for idx, row in stations_to_examine.iterrows():
    folium.Marker( location=[row['latitude'], row['longitude']],
                    popup=row['station_name'] ).add_to(map_pts)
```

¹ В этом и заключается смысл создания хранилищ данных: объединить корпоративные данные в централизованное хранилище, чтобы аналитик мог взять любые корпоративные данные и соединить их.

В результате получится красивая интерактивная карта, показанная на рис. 5.10, которую наша команда может использовать для проверки выбранных нами пунктов проката.

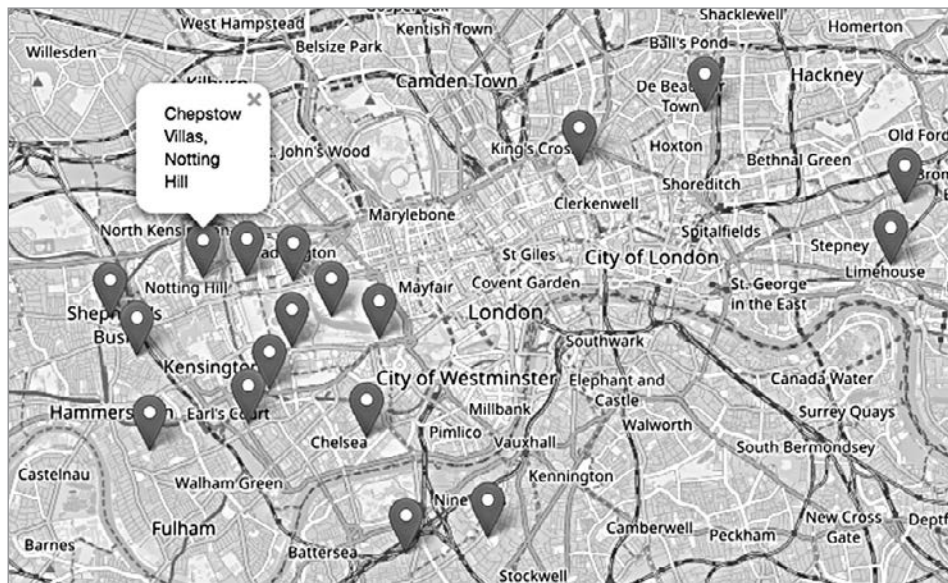


Рис. 5.10. Интерактивная карта с пунктами проката, подлежащими проверке

Мы смогли эффективно интегрировать BigQuery, pandas и Jupyter для анализа данных. Мы использовали BigQuery для вычисления агрегированных значений по миллионам поездок, pandas — для получения статистических характеристик и пакеты Python, такие как folium, — для визуализации результатов.

Работа с BigQuery из R

Python — один из наиболее популярных языков для исследования данных, но ему ничуть не уступает R — давно существующий язык программирования и программная среда для статистических расчетов и построения графиков.

Чтобы использовать BigQuery из R, установите библиотеку bigrquery из CRAN:

```
install.packages("bigrquery", dependencies=TRUE)
```

Вот простой пример на языке R, посылающий запрос к набору данных london_bicycles:

```
billing <- 'cloud-training-demos' # имя вашего проекта
sql <- "
SELECT
```

```

    start_station_name
    , AVG(duration) as duration
    , COUNT(duration) as num_trips
FROM `bigquery-public-data`.london_bicycles.cycle_hire
GROUP BY start_station_name
ORDER BY num_trips DESC
LIMIT 5
"

tbl <- bq_project_query(billing, sql)
bq_table_download(tbl, max_results=100)
grid.tbl(tbl)

```

Здесь вызовом `bq_project_query` создается запрос к BigQuery, который выполняется с помощью `bq_table_download`.

Также язык R можно использовать в блокнотах Jupyter. Окружение `conda` для Jupyter¹ имеет расширение для поддержки R, которое можно загрузить, как показано ниже:

```

!conda install rpy2
%load_ext rpy2.ipynthon

```

Чтобы выполнить линейную регрессию для прогнозирования количества держателей для велосипедов в пунктах проката в зависимости от их местоположения, можно сначала заполнить R DataFrame из BigQuery:

```

%%bigquery docks --project $PROJECT
SELECT
    docks_count, latitude, longitude
FROM `bigquery-public-data`.london_bicycles.cycle_stations
WHERE bikes_count > 0

```

а потом использовать Jupyter Magics для R по аналогии с Jupyter Magics для Python. Проще говоря, вот как можно использовать Jupyter Magics для R, чтобы выполнить линейную аппроксимацию (`lm`, от англ. *linear modeling*) количества держателей по данным в DataFrame:

```

%%R -i docks
mod <- lm(docks ~ latitude + longitude)
summary(mod)

```

Cloud Dataflow

В главе 4 мы представили Cloud Dataflow как инструмент для загрузки данных в BigQuery из MySQL. Cloud Dataflow — это управляемая служба, предназначенная для выполнения конвейеров, написанных с использованием Apache Beam.

¹ На момент написания этой книги образ PyTorch для экземпляра Notebook в GCP был скомпилирован с поддержкой `conda`.

Она может очень пригодиться в анализе данных, потому что дает возможность выполнять преобразования, которые трудно реализовать на языке SQL. На момент написания этой книги конвейеры Beam можно было писать на Python, Java и Go, причем подход с использованием Java считался наиболее совершенным.

Чтобы лучше представить, где эта возможность может быть полезна, рассмотрим распределение продолжительностей аренды велосипедов в отдельном пункте проката (рис. 5.11).

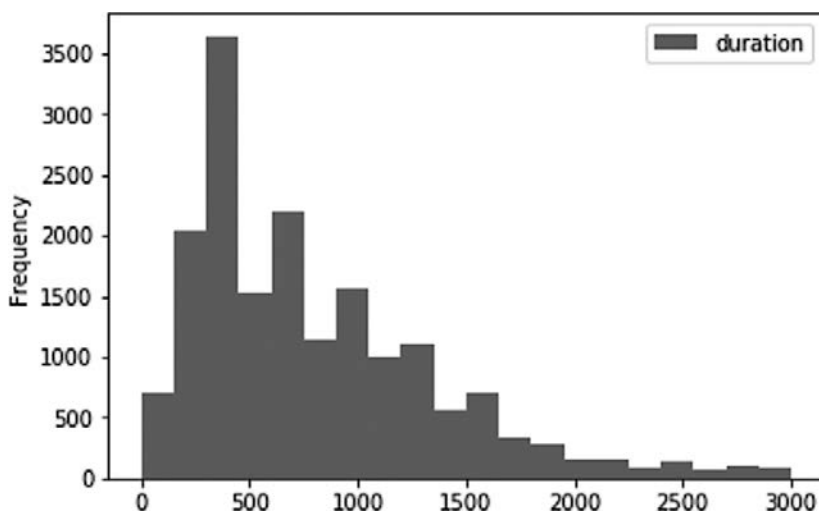


Рис. 5.11. Распределение продолжительностей аренды велосипедов в отдельном пункте проката

Как показано на рис. 5.11, столбцу с $x = 1000$ соответствует значение $y = 1500$; это означает, что примерно 1500 поездок длились около 1000 секунд.

В таблице BigQuery доступны конкретные значения продолжительности, но иногда полезно подогнать эти значения к теоретическому распределению, чтобы было проще моделировать и изучать влияние изменения цен и доступности. В Python, если предположить, что значения продолжительности хранятся в массиве `duration`, параметры аппроксимации гамма-распределением (https://en.wikipedia.org/wiki/Gamma_distribution) довольно легко вычислить с помощью пакета `scipy`:

```
from scipy import stats
ag,bg,cg = stats.gamma.fit(df['duration'])
```

А теперь представьте, что вам нужно обойти все пункты проката и для каждого рассчитать параметры гамма-распределения длительностей поездок. Решить эту задачу на SQL сложно, но Python с ней легко справляется, соответственно,

можно написать задание Dataflow, которое вычислит параметры гамма-распределения, используя для этого целый кластер компьютеров.

Сначала конвейер выполнит запрос¹, чтобы получить массив продолжительностей поездок для каждого пункта проката, передаст полученные записи методу `compute_fit`, а затем сохранит полученные результаты в таблицу `station_stats` в BigQuery:²

```
opts = beam.pipeline.PipelineOptions(flags = [], **options)
RUNNER = 'DataflowRunner'
query = """
    SELECT start_station_id, ARRAY_AGG(duration) AS duration_array
    FROM `bigquery-public-data.london_bicycles.cycle_hire`
    GROUP BY start_station_id
    """

with beam.Pipeline(RUNNER, options = opts) as p:
    (p
     | 'read_bq' >> beam.io.Read(beam.io.BigQuerySource(query=query))
     | 'compute_fit' >> beam.Map(compute_fit)
     | 'write_bq' >> beam.io.gcp.bigquery.WriteToBigQuery(
         'ch05eu.station_stats',
         schema='station_id:string,ag:FLOAT64,bg:FLOAT64,cg:FLOAT64')
    )
```

Метод `compute_fit` — это функция на Python, которая принимает словарь, соответствующий входной записи в BigQuery, и возвращает словарь, соответствующий желаемой выходной записи:

```
def compute_fit(row):
    from scipy import stats
    result = {}
    result['station_id'] = row['start_station_id']
    durations = row['duration_array']
    ag, bg, cg = stats.gamma.fit(durations)
    result['ag'] = ag
    result['bg'] = bg
    result['cg'] = cg
    return result
```

Вычисленные значения затем записываются в требуемую таблицу.

После запуска задания Dataflow можно проследить за его выполнением в GCP Cloud Console, как показано на рис. 5.12, и понаблюдать, как оно автоматически масштабируется для параллельной обработки пунктов проката.

Когда задание Dataflow завершится, можно запросить таблицу, получить статистику для пунктов проката и построить график с параметрами гамма-рас-

¹ Стоимость обработки запроса включается в стоимость выполнения задания Dataflow.

² См. блокнот `05_devel/statfit.ipynb` в репозитории GitHub с примерами для этой книги.

пределения (см. графики в блокноте в репозитории GitHub (https://github.com/GoogleCloudPlatform/bigquery-oreilly-book/blob/master/05_devel/statfit.ipynb)).

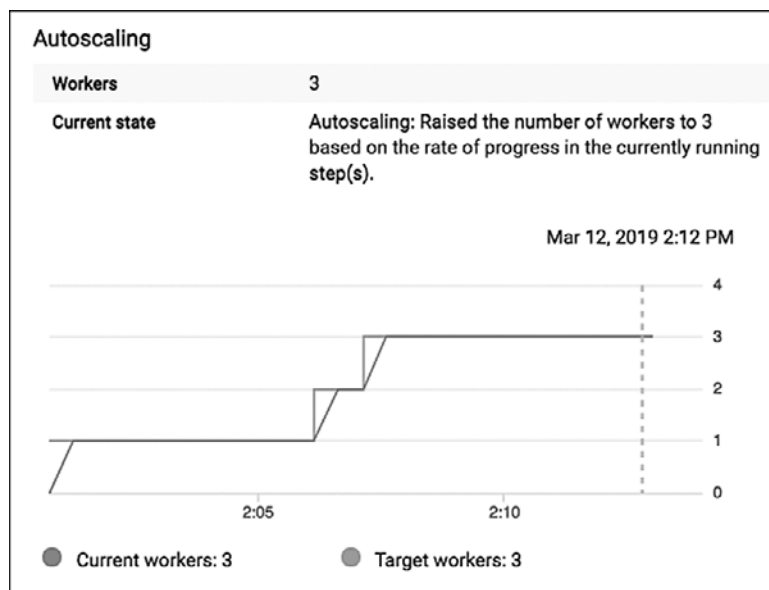


Рис. 5.12. Задание Dataflow распараллеливается и выполняется в кластере, размер которого автоматически подбирается в зависимости от скорости выполнения каждого шага

Драйверы JDBC/ODBC

Поскольку BigQuery является хранилищем для структурированных данных, в приложениях на Java или .NET может быть удобно использовать не зависящие от конкретной базы данных API, такие как Java Database Connectivity (JDBC) и Open Database Connectivity (ODBC), для взаимодействия с драйвером базы данных BigQuery. Такой подход не рекомендуется для новых приложений — лучше использовать клиентскую библиотеку. Однако если у вас есть устаревшее приложение, которое в настоящее время подключается к базе данных, и в него нежелательно вносить существенные изменения, чтобы получить возможность взаимодействия с BigQuery, применение драйвера JDBC/ODBC может быть оправданным.

Компания Simba, партнер Google, предоставляет драйверы ODBC и JDBC, способные выполнять запросы в BigQuery на SQL Standard.¹ Чтобы установить

¹ См. <https://cloud.google.com/bigquery/partners/simba-drivers/> и <https://www.simba.com/drivers/bigquery-odbc-jdbc/>.

драйверы для Java, нужно загрузить ZIP-файл, распаковать его и поместить все распакованные файлы JAR в папку, находящуюся в пути поиска классов (classpath) вашего Java-приложения. Чтобы задействовать драйвер в приложении на Java, обычно требуется изменить файл конфигурации и указать информацию о соединении. Есть несколько вариантов настройки аутентификации и строки подключения в приложениях на Java.



Мы настоятельно рекомендуем использовать клиентские библиотеки (для выбранного вами языка) вместо драйверов JDBC/ODBC, потому что функциональность, предоставляемая драйвером JDBC/ODBC, предлагает лишь подмножество полных возможностей BigQuery. При использовании драйверов JDBC/ODBC вам будут недоступны: поддержка крупномасштабного приема (то есть многие методы загрузки, описанные в предыдущей главе), крупномасштабного экспорта (то есть перемещение данных будет происходить медленно) и вложенных/повторяющихся полей (что не позволит использовать многие оптимизации производительности, которые мы рассмотрим в главе 7). Кроме того, при разработке новых систем, использующих драйверы JDBC/ODBC, вы неизбежно столкнетесь со множеством сложных технических проблем.

Внедрение данных из BigQuery в Google Slides (в G Suite)

Для управления проектами BigQuery, загрузки данных и выполнения запросов можно использовать Google Apps Script (<https://developers.google.com/apps-script/advanced/bigquery>). Он может пригодиться для автоматизации заполнения Google Docs, Google Sheets или Google Slides данными из BigQuery.

Для примера рассмотрим создание пары слайдов с анализом данных о прокате велосипедов в Лондоне. Для начала откройте страницу <https://script.google.com/create>, чтобы создать новый сценарий. Затем в меню **Resources** (Ресурсы) выберите пункт **Advanced Google Services** (Дополнительные службы Google) и включите флажок **BigQuery API** (дайте имя проекту, если такая опция будет предложена).

Полный текст сценария для Apps Script можно найти в репозитории GitHub с примерами для этой книги (https://github.com/GoogleCloudPlatform/bigquery-oreilly-book/blob/master/05_devel/bq_to_slides.gs). Скопируйте текст сценария и вставьте его в текстовый редактор. Затем измените значение `PROJECT_ID`, выберите функцию `createBigQueryPresentation` и щелкните на кнопке запуска, как показано на рис. 5.13.

Получившаяся электронная таблица и набор слайдов появятся в Google Drive (URL можно найти, выбрав пункт меню **View** ▶ **Logs** (Вид ▶ Журналы)). Набор слайдов будет выглядеть так, как показано на рис. 5.14.

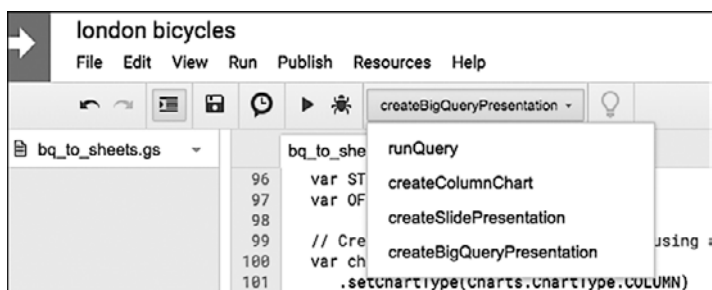


Рис. 5.13. Порядок использования Google Apps Script для создания презентации на основе данных из BigQuery

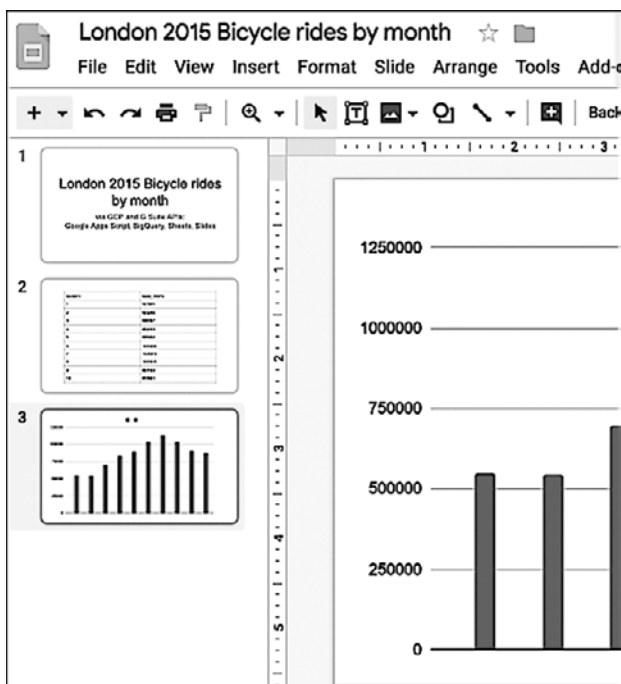


Рис. 5.14. Набор слайдов, созданных с помощью Google Apps Script

Вот как выглядит определение функции `createBigQueryPresentation`:

```
function createBigQueryPresentation() {
  var spreadsheet = runQuery();
  Logger.log('Results spreadsheet created: %s', spreadsheet.getUrl());
  var chart = createColumnChart(spreadsheet); // ДОБАВЛЕНО
  var deck = createSlidePresentation(spreadsheet, chart); // НОВАЯ
  Logger.log('Results slide deck created: %s', deck.getUrl()); // НОВАЯ
}
```

Фактически она вызывает три другие функции:

- `runQuery`, чтобы выполнить запрос и сохранить результаты в электронной таблице Google Sheets;
- `createColumnChart`, чтобы создать диаграмму на основе данных в электронной таблице;
- `createSlidePresentation`, чтобы создать набор слайдов в Google Slides.

Метод `runQuery()` использует клиентскую библиотеку Apps Scripts, чтобы вызвать BigQuery и получить результаты:

```
var queryResults = BigQuery.Jobs.query(request, PROJECT_ID);
var rows = queryResults.rows;
while (queryResults.pageToken) {
  queryResults = BigQuery.Jobs.getQueryResults(PROJECT_ID, jobId, {
    pageToken: queryResults.pageToken
  });
  rows = rows.concat(queryResults.rows);
}
```

Затем создается электронная таблица и полученные записи добавляются в лист. Две другие функции используют код Apps Scripts для построения графиков и создания набора слайдов и добавления данных в этот набор.

Bash-скрипты для BigQuery

Инструмент командной строки `bq`, входящий в комплект Google Cloud Software Development Kit (SDK (<https://cloud.google.com/sdk/>)), обеспечивает удобный способ выполнения операций в BigQuery из командной строки. По умолчанию SDK устанавливается на виртуальные машины и кластеры Google Cloud. Также есть возможность загрузить и установить SDK в локальной среде разработки или эксплуатации.

Инструмент `bq` можно использовать для взаимодействия со службой BigQuery при написании Bash-скриптов или обращаясь к командной оболочке из программ на других языках программирования и тем самым избавиться от необходимости использовать клиентскую библиотеку. Обычно `bq` используется для создания и проверки существования наборов данных и таблиц, выполнения запросов, загрузки данных в таблицы, заполнения таблиц и представлений, а также проверки состояния заданий. Давайте рассмотрим каждую из этих операций.

Создание наборов данных и таблиц

Создать набор данных можно командой `bq mk`, передав ей местоположение набора данных (например, `US` или `EU`). Также можно указать срок действия таблицы. Дополнительно рекомендуется передать описание набора данных:

```
bq mk --location=US \  
  --default_table_expiration 3600 \  
  --description "Chapter 5 of BigQuery Book." \  
  ch05
```

Проверка существования набора данных

Команда `bq mk` в предыдущем примере выдает ошибку, если набор данных уже существует. Чтобы создать набор данных, только если он еще не существует, необходимо сначала получить список существующих наборов данных, запустив команду `bq ls`, и проверить, нет ли в этом списке набора данных с заданным именем:¹

```
#!/bin/bash  
bq_safe_mk() {  
  dataset=$1  
  exists=$(bq ls --dataset | grep -w $dataset)  
  if [ -n "$exists" ]; then  
    echo "Not creating $dataset since it already exists"  
  else  
    echo "Creating $dataset"  
    bq mk $dataset  
  fi  
}  
  
# так можно вызывать эту функцию  
bq_safe_mk ch05
```

Создание набора данных в другом проекте

Пример в предыдущем разделе создаст набор данных `ch05` в проекте по умолчанию (указанном при входе в виртуальную машину или при запуске `gcloud auth` с использованием Google Cloud SDK). Чтобы создать набор данных в другом проекте, нужно кроме имени набора данных указать имя проекта:

```
bq mk --location=US \  
  --default_table_expiration 3600 \  
  --description "Chapter 5 of BigQuery Book." \  
  projectname:ch05
```

Создание таблицы

Таблицы создаются подобно наборам данных, за исключением того, что в команду `bq mk` нужно добавить параметр `--table`. Следующая команда создаст таблицу с именем `ch05.rentals_last_hour`, со сроком действия 3600 секунд и с двумя столбцами: `rental_id` (строковый) и `duration` (числовой, с плавающей точкой):

¹ Это решение, как вы понимаете, подвержено проблеме «гонки данных», если кто-то еще успел создать набор данных после того, как ваш сценарий проверил его отсутствие, и перед фактическим созданием.

```
bq mk --table \  
  --expiration 3600 \  
  --description "One hour of data" \  
  --label persistence:volatile \  
  ch05.rentals_last_hour rental_id:STRING,duration:FLOAT
```

Параметр `--label` можно использовать для добавления дополнительных характеристик; Data Catalog поддерживает возможность поиска таблиц по характеристикам: здесь `persistence` — это ключ, а `volatile` — метка.

Сложные схемы

Чтобы создать таблицу с более сложной схемой, которую нелегко выразить в виде строки с перечислением полей через запятую, можно указать файл в формате JSON, как описано в главе 4:

```
bq mk --table \  
  --expiration 3600 \  
  --description "One hour of data" \  
  --label persistence:volatile \  
  ch05.rentals_last_hour schema.json
```

Копирование наборов данных

Инструмент командной строки предлагает самый эффективный способ копирования наборов данных. Например, следующая команда скопирует таблицу из набора данных `ch04` в набор `ch05`:

```
bq cp ch04.old_table ch05.new_table
```



Копирование таблиц может занять некоторое время, но иногда желательно, чтобы сценарий дождался завершения задания, прежде чем продолжить работу. Самый простой способ реализовать такое ожидание — использовать команду `bq wait`:

```
bq wait --fail_on_error job_id
```

Эта команда будет ждать, пока задание не завершится, а следующая будет ждать не более 600 секунд:

```
bq wait --fail_on_error job_id 600
```

Если было запущено только одно задание, параметр `job_id` можно опустить.

Загрузка и вставка данных

В главе 4 мы подробно рассмотрели прием загрузки данных в таблицу с использованием `bq load`. Вернитесь к ней, если вдруг вы забыли.

Чтобы вставить записи в таблицу, запишите их в файл в строчном формате JSON и используйте команду `bq insert`:

```
bq insert ch05.rentals_last_hour data.json
```

В этом примере файл `data.json` содержит записи, соответствующие схеме вставляемой таблицы:

```
{"rental_id": "345ce4", "duration": 240}
```

Извлечение данных

Извлечь данные из таблицы BigQuery в один или несколько файлов в Cloud Storage можно с помощью команды `bq extract`:

```
bq extract --format=json ch05.bad_bikes gs://bad_bikes.json
```

Выполнение запросов

Чтобы выполнить запрос, передайте его команде `bq query`:

```
bq query \
  --use_legacy_sql=false \
  'SELECT MAX(duration) FROM \
  `bigquery-public-data`.london_bicycles.cycle_hire'
```

Также можно передать запрос через стандартный ввод:

```
echo "SELECT MAX(duration) FROM \
`bigquery-public-data`.london_bicycles.cycle_hire" \
| bq query --use_legacy_sql=false
```

Ввод запроса в одну строку с экранированием кавычек и т. д. может оказаться утомительным. Для удобочитаемости можно воспользоваться возможностью Bash считывать многострочный текст в переменную:¹

```
#!/bin/bash
read -d '' QUERY_TEXT << EOF
SELECT
  start_station_name
  , AVG(duration) as duration
  , COUNT(duration) as num_trips
FROM `bigquery-public-data`.london_bicycles.cycle_hire
GROUP BY start_station_name
ORDER BY num_trips DESC
LIMIT 5
EOF
bq query --project_id=some_project --use_legacy_sql=false $QUERY_TEXT
```

¹ См. сценарий `05_devel/bq_query.sh` в репозитории GitHub с примерами для этой книги.

Этот код считывает в переменную `QUERY_TEXT` многострочный текст, заканчивающийся словом `EOF`, и затем передает эту переменную в команду `bq query`.

Предыдущий код также иллюстрирует явное указание проекта, запрос по которому должен быть оплачен.

Не забудьте передать параметр `--use_legacy_sql = false`, потому что по умолчанию `bq` использует нестандартный диалект `SQL`, который мы рассматриваем в этой книге.

НАСТРОЙКА ФЛАГОВ В `.BIGQUERYRC`

Если вы предпочитаете использовать инструмент командной строки `bq` в интерактивном режиме, вам может пригодиться возможность определить типичные флаги и параметры, такие как `--location`, в `$BIGQUERYRC/.bigqueryrc` или в `$HOME/.bigqueryrc`, если переменная окружения `$BIGQUERYRC` не определена. Вот пример содержимого файла `.bigqueryrc`:

```
--location=EU
--project_id=some_project
[mk]
--expiration=3600
[query]
--use_legacy_sql=false
```

При наличии этого файла ресурсов все команды BigQuery будут вызываться с параметром `--location=EU`, а плата — взиматься с проекта `some_project`. Кроме того, все команды `bq mk` будут вызываться с параметром `--expiration=3600`, а все команды `bq query` — с параметром `--use_legacy_sql=false`. Переопределить эти значения можно, только явно указав соответствующие параметры в командной строке.

Определяя файл ресурсов для BigQuery, помните, что любые сценарии, которые вы пишете или запускаете, будут по-разному работать на компьютерах, где есть этот файл (как правило, на компьютерах для разработки) и где его нет (обычно на производственных компьютерах). Это может привести к путанице. Как показывает наш опыт, любое увеличение производительности, обусловленное наличием файла ресурсов, сводится на нет повышенной сложностью отладки при использовании сценариев на разных компьютерах. Впрочем, у вас может быть другой опыт.

Предварительный просмотр данных

Для предварительного просмотра таблицы используйте `bq head`. В отличие от запроса `SELECT *`, за которым следует `LIMIT`, эта команда действует детерминированно и не влечет дополнительных затрат на оплату услуг BigQuery.

Следующая команда вернет первые 10 записей:

```
bq head -n 10 ch05.bad_bikes
```

А эта команда — следующие 10 записей:

```
bq head -s 10 -n 10 ch05.bad_bikes
```

Обратите внимание, что BigQuery не гарантирует определенного порядка записей в таблицах, поэтому рассматривайте эту команду как инструмент, возвращающий произвольный набор записей.

Создание представлений

Создавать простые и материализованные представления из запросов можно командой `bq mk`. Например, следующая команда создаст представление с именем `rental_duration` в наборе данных `ch05`:

```
#!/bin/bash
read -d '' QUERY_TEXT << EOF
SELECT
    start_station_name
    , duration/60 AS duration_minutes
FROM `bigquery-public-data`.london_bicycles.cycle_hire
EOF
bq mk --view=$QUERY_TEXT ch05.rental_duration
```

BigQuery позволяет отправлять запрос как к представлениям, так и к таблицам, но, в отличие от таблиц, представления действуют подобно подзапросам — обращение к представлению влечет вставку полного текста представления в вызывающий запрос. Материализованные представления сохраняют результаты в таблице, которая затем используется при обработке запросов к этому представлению. BigQuery сама следит за актуальностью материализованного представления. Более детально обычные и материализованные представления мы рассмотрим в главе 10. Чтобы создать материализованное представление, достаточно заменить параметр `--view` в предыдущем фрагменте на `--materialized_view`.

Объекты BigQuery

Мы уже видели, как использовать команду `bq ls --dataset` для получения списка наборов данных в проекте. Однако, как показано в табл. 5.3, есть и другие объекты, списки которых можно получить.

Таблица 5.3. Команды и соответствующие им списки объектов

Команда	Список каких объектов возвращает
<code>bq ls ch05</code>	Таблицы в наборе данных <code>ch05</code>
<code>bq ls -p</code>	Все проекты
<code>bq ls -j some_project</code>	Все задания в указанном проекте

Команда	Список каких объектов возвращает
<code>bq ls --dataset</code>	Все наборы данных в проекте по умолчанию
<code>bq ls --dataset some_project</code>	Все наборы данных в указанном проекте
<code>bq ls --models</code>	Модели машинного обучения
<code>bq ls --transfer_run \</code> <code>--filter='states:PENDING' \</code> <code>--run_attempt='LATEST' \</code> <code>projects/p/locations/l \</code> <code>/transferConfigs/c</code>	Запущенные передачи, находящиеся в состоянии ожидания
<code>bq ls --reservation_grant \</code> <code>--project_id=some_proj \</code> <code>--location='us'</code>	Забронированные слоты в указанном проекте

Получение дополнительной информации

В табл. 5.4 приводятся примеры получения дополнительной информации об объекте BigQuery с помощью команды `bq show`.

Таблица 5.4. Команды BigQuery и возвращаемая ими информация

Команда	О каком объекте возвращает дополнительную информацию
<code>bq show ch05</code>	О наборе данных <code>ch05</code>
<code>bq show -j some_job_id</code>	Об указанном задании
<code>bq show --schema ch05.bad_bikes</code>	Схему таблицы <code>ch05.bad_bikes</code>
<code>bq show --view ch05.some_view</code> <code>bq show --materialized_view ch05.some_view</code>	Об указанном представлении
<code>bq show --model ch05.some_model</code>	Об указанной модели
<code>bq show --transfer_run \</code> <code>projects/p/locations/l/transferConfigs/c/runs/r</code>	О запущенной передаче

Обратите внимание, что можно получить список заданий с помощью `bq ls` и проверить их состояние с помощью `bq show`.

Изменение

Изменить некоторые аспекты уже созданных таблиц, наборов данных и т. д. можно с помощью `bq update`:

```
bq update --description "Bikes that need repair" ch05.bad_bikes
```

Командой `bq update` можно изменить определение обычного или материализованного представления:

```
bq update \  
  --view "SELECT ..."\  
  ch05.rental_duration
```

и даже объем бронирования (о слотах и бронировании мы расскажем в главе 6):

```
bq update --reservation --location=US \  
  --project_id=some_project \  
  --reservation_size=2000000000
```

Выводы

В этой главе мы рассмотрели три вида клиентских библиотек BigQuery:

- REST API — доступный из программ, написанных на любом языке, позволяющем реализовать взаимодействие с веб-сервером;
- клиент Google API, предлагающий автоматически генерируемые библиотеки для многих языков программирования;
- библиотека BigQuery Client Library, предлагающая удобные средства доступа к BigQuery из многих популярных языков программирования.

Мы рекомендуем использовать клиентскую библиотеку BigQuery Client Library, если она доступна для вашего языка. В противном случае используйте клиентскую библиотеку Google API. REST API подойдет, только если вы работаете в среде, где недоступна даже библиотека Google API.

Существует пара высокоуровневых абстракций, которые упрощают программное взаимодействие с BigQuery и используются в двух популярных средах: в блокнотах Jupyter и сценариях командной оболочки. Мы довольно подробно рассмотрели поддержку Jupyter и pandas в BigQuery и показали, насколько мощное и расширяемое окружение для сложных процессов обработки данных можно построить, используя комбинацию этих инструментов. Мы также коснулись темы интеграции с языком программирования R и набором G Suite и рассмотрели многие из возможностей инструмента командной строки `bq`. Наконец, мы узнали, как взаимодействовать с BigQuery при помощи Bash-скриптов.

Архитектура BigQuery

Платформа BigQuery старается подстраиваться под размеры ваших наборов данных и работать так быстро, как того требует ваш бизнес. Взаимодействие с платформой должно казаться волшебством. Но беда с подобными «волшебными» вещами состоит в том, что, столкнувшись с проблемой, вы не знаете, как ее исправить.

В этой главе мы подробно расскажем о внутреннем устройстве BigQuery, опишем обобщенную архитектуру платформы и механизм запросов Dremel и представим подробную информацию о метаданных хранилища. Как BigQuery управляет безопасностью, доступностью и аварийным восстановлением, мы расскажем в главе 10. Как минимум эта глава удовлетворит ваше любопытство. Однако если что-то пойдет не так, как вы ожидаете, эта глава сможет помочь вам лучше понять процесс, а также разобраться в том, как исправить или обойти проблему.

Архитектура высокого уровня

BigQuery — это крупномасштабная распределенная система, выполняющая одновременно сотни тысяч задач в десятках взаимосвязанных микросервисах в нескольких зонах доступности в каждом регионе Google Cloud. В этом разделе мы представим упрощенное описание связей между высокоуровневыми элементами. Подробное описание всех компонентов может потребовать отдельной книги, и мы рискуем потерять большинство читателей к тому моменту, когда закончим описывать транскoder хранилища, а остальные захлопнут книгу задолго до того, как мы перейдем к тупиковому прокси (stubby proxy; да, он действительно существует, и нет, в нем нет ничего таинственного).

Жизненный цикл запроса

Чтобы понять, как работает BigQuery, рассмотрим, что происходит, когда вы отправляете запрос. Не будем пока углубляться в фактическое выполнение

запроса — мы поговорим об этом в следующем разделе, а просто пройдемся по компонентам высокого уровня.

На рис. 6.1 показан упрощенный поток управления при выполнении запроса. Назначение каждого блока, изображенного на диаграмме, мы рассмотрим далее в этой главе.

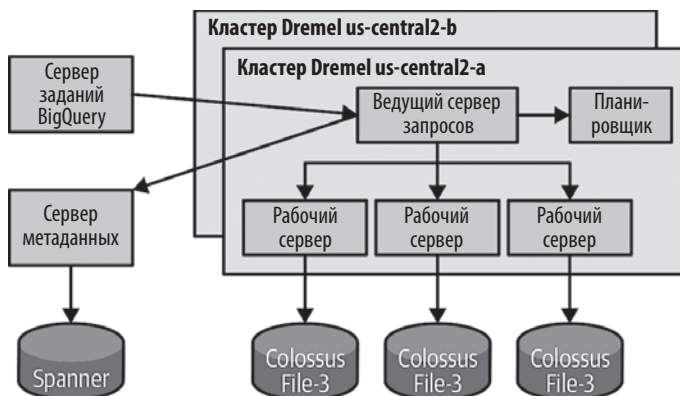


Рис. 6.1. Упрощенное представление потока обработки запроса в системе BigQuery

Для начала посмотрим, что происходит, когда вы выполняете самый простой из SQL-запросов: `SELECT 17`. Этот запрос даже не требует чтения каких-либо данных, он просто возвращает одно значение.

Шаг 1: HTTP POST

Клиент посылает HTTP-запрос POST конечной точке BigQuery. Обычно этот запрос формируется библиотекой или драйвером Java Database Connectivity (JDBC), но на базовом уровне отправить такой запрос можно с помощью `curl` или любого другого инструмента, позволяющего отправлять HTTP-запросы без их обработки (см. главу 5).

Вот как выглядит запрос, который передается в сеть:

```
POST /bigquery/v2/projects/bigquery-e2e/jobs HTTP/1.1
User-Agent: curl/7.30.0
Host: www.googleapis.com
Accept: */*
Authorization: Bearer <redacted>
Content-Type: application/json
Content-Length: 126
{'configuration': {'query': {'query': 'SELECT 17'}}
```

Здесь есть несколько важных составляющих: во-первых, это глагол HTTP, в данном случае POST. Он используется, потому что мы собираемся изменить

состояние системы, создав задание для обработки запроса. Второй — токен авторизации **Authorization**. Это токен OAuth2, идентифицирующий вас. Последний важный элемент — тело запроса с текстом в формате JSON, которое сообщает, что требуется выполнить запрос **SELECT 17**. Нетрудно догадаться, что в таком HTTP-запросе есть множество вариантов, которые можно отправить; за дополнительной информацией обращайтесь к документации с описанием BigQuery API (<https://cloud.google.com/bigquery/docs/reference/rest/>).

На рис. 6.2 более подробно показан путь запроса через систему. В следующих нескольких разделах мы расскажем, что делается на каждом этапе и зачем это необходимо.

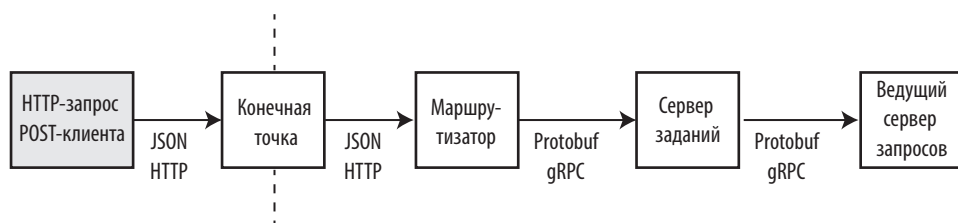


Рис. 6.2. Путь запроса до начала его фактической обработки

Шаг 2: маршрутизация

HTTP-запрос POST преодолевает просторы интернета¹ и попадает в конечную точку REST <http://www.googleapis.com/bigquery/v2/projects/bigquery-e2e/jobs>. Этот адрес обслуживает сервер Google Front-End (GFE), сервер того же типа, который обслуживает поисковую систему и другие продукты Google. В данном случае GFE должен найти сервис BigQuery, который обслужит ваш запрос.

BigQuery — это глобальный сервис. Как он определяет, в какой регион направить запрос? Существует ряд подсказок, которые помогают ему понять, куда отправить запрос. Часть URL указывает, какой облачный проект отвечает за оплату запроса. Некоторые проекты ограничивают регион, в котором им разрешено выполнять запросы. Если ваша организация настроила проект для работы только в Австралии, ваш запрос будет отправлен в Австралию. Некоторые проекты связаны бронированием по фиксированной цене. Если у вас настроено бронирование в определенном месте, запрос будет отправлен туда.

Если бронирование не предусмотрено и вы не указали в имени задания, в каком регионе должен обрабатываться запрос, маршрутизатор проанализирует запрос и выяснит, какие наборы данных в нем задействованы. Наборы данных привязаны к местоположениям, поэтому BigQuery определит регион набора данных

¹ Многие корпоративные клиенты используют прямые линии связи, поэтому их запросы никогда не оказываются в общедоступном интернете.

и перешлет запрос туда. Если вас интересует производительность или вы хотите контролировать местоположение результатов, то можете указать регион, куда должен быть направлен запрос, заполнив поле ссылки на задание.

В нашем примере мы не даем маршрутизатору никаких подсказок, и в запросе не используются никакие наборы данных. В этом случае маршрутизатор передаст запрос в регион по умолчанию US.

Маршрутизатор преобразует тело JSON HTTP-запроса в Protocol Buffers (<https://developers.google.com/protocol-buffers/>) (Protobufs) — формат сериализации, не зависящий от платформы и языка, который используется для связи практически между всеми службами Google.

Шаг 3: сервер заданий

Сервер заданий BigQuery отвечает за отслеживание состояния запроса. Поскольку сетевое соединение между клиентом и сервером BigQuery может быть неустойчивым и на обработку некоторых запросов может уходить несколько минут или даже часов, сервер заданий предполагает асинхронную работу.

Сервер заданий выполняет авторизацию, чтобы убедиться, что вызывающему абоненту разрешено выполнить запрос, плата за который взимается с указанного проекта. Это важно для того, чтобы никто не смог воспользоваться ресурсами за ваш счет. Авторизация доступа к фактическим таблицам откладывается до начала запроса.

Сервер заданий отправляет запрос правильному серверу обработки запросов. Каждый проект обычно имеет первичную и вторичную зоны доступности. Если первичная зона недоступна, запросы направляются во вторичную.

Если процессы резервного копирования или аварийного переключения запускаются редко, они, как правило, работают не так как должны, когда в них возникает необходимость. Решение состоит в том, чтобы регулярно запускать такие процессы восстановления после отказа, чтобы на них можно было положиться. Примерно раз в неделю BigQuery сталкивается с необходимостью аварийного переключения вычислительного кластера в том или ином регионе. Проблемы могут быть обусловлены целым рядом различных причин, таких как сбой сетевого коммутатора, снижение производительности зависимой службы или необычно большая длина очереди.

Когда происходит переключение вычислительного кластера, обрабатывающего запросы (то есть кластер зоны), все использующие его проекты переходят на применение своих вторичных кластеров. Состояние репликации тщательно контролируется, поэтому любые вновь загруженные или перебалансированные данные, пока не поступившие во вторичный кластер, будут извлекаться из любого другого места, где их можно найти. В редких случаях, когда оперативные копии данных отсутствуют, BigQuery сообщит об ошибке отсутствия данных.

ПРОЕКТ И ПЕРЕБАЛАНСИРОВКА ДАННЫХ

В BigQuery можно соединить любые две таблицы, находящиеся в одном регионе, если пользователь имеет доступ к обеим таблицам. Это создает проблему для сервера, потому что если таблицы физически расположены в разных местах, запрос будет обрабатываться медленно и стоить дорого (для Google, не для пользователя), потому что для этого потребуется переместить много данных по сети.

Асинхронно BigQuery постоянно решает сложную задачу оптимизации: как обеспечить максимально близкое расположение друг к другу таблиц, участвующих в соединениях. При ее решении важно также учитывать емкость различных вычислительных кластеров и кластеров хранения, топологию сети, зоны доступности и то, где в настоящий момент находятся данные.

Если внутренний балансировщик решит, что проект требуется переместить, он запустит репликацию данных в другую зону доступности или даже в другой регион (когда допускается хранение данных в объединении регионов и в настройках проекта указано, что данные должны храниться в одном регионе, такое перемещение выполняться не будет). То есть данные всегда будут доступны для выполнения соединений, и кластеры с конечной емкостью не столкнутся с проблемой нехватки места даже при значительном увеличении нагрузки.

Шаг 4: механизм обработки запросов

Порядок обработки запросов подробнее будет описан в следующем разделе, поэтому здесь мы рассмотрим процесс в общем. Запросы передаются ведущему серверу запросов (Query Master), который отвечает за выполнение запроса в целом. Ведущий сервер запросов обращается к серверу метаданных, чтобы определить, где физически находятся данные и как они секционированы. На этом этапе происходит исключение ненужных секций, то есть если запрос читает не все секции, возвращаются метаданные только активных секций.

Определив, какой объем данных потребуются обработать, и получив возможность составить предварительный план запроса, ведущий сервер запросов запрашивает слоты у планировщика. Слот — это поток выполнения на рабочем сервере сегмента; обычно слоту соответствует половина ядра процессора и около 1 Гбайта оперативной памяти, но это не строгая мера, потому что слоты могут увеличиваться или уменьшаться, если им необходимо больше или меньше ресурсов, а также по мере обновления компьютеров в центре обработки данных Google.

Планировщик решает, как распределить работу среди сегментов (shard). В ответ на запрос слотов он возвращает адреса сегментов, которые будут обрабатывать запрос. Затем ведущий сервер запросов отправляет запрос каждому из сегментов Dremel для параллельной обработки. Подробнее о порядке обработки запросов рассказывается в разделе «Выполнение запроса».

Шаг 5: возврат результатов запроса

Когда рабочие серверы сегментов завершают выполнение запроса, результаты делятся на две части. Первая страница результатов сохраняется в распределенной реляционной базе данных Spanner (<https://ai.google/research/pubs/pub39966>) вместе с метаданными запроса. Данные в Spanner размещаются в том же регионе, где выполнялся запрос. Остальные данные записываются в Colossus,¹ распределенную файловую систему Google. Запросы с небольшим объемом результатов вообще не записываются на диск и возвращаются очень быстро.

BigQuery API поддерживает возможность повторного подключения. То есть он способен обеспечить синхронную работу, но если время ожидания истекло, вызывающий клиент имеет возможность восстановить соединение. С этой целью сервер заданий возвращает клиенту идентификатор задания, по которому тот сможет найти свое задание и получить результаты. Клиенты BigQuery — `bq.py` (библиотеки облачных клиентов) и драйверы Open Database Connectivity (ODBC)/JDBC формируют этот протокол, что обеспечивает надежное получение результатов конечными пользователями.

Результаты хранятся в BigQuery в течение 24 часов; функционально они эквивалентны таблице и доступны для запросов, как если бы были таблицей. Хранимый объем результатов ограничен 10 Гбайт для обычных запросов `SELECT`. Чтобы сохранить больше данных, можно использовать операторы `CREATE TABLE AS SELECT` или `INSERT`, которые не имеют ограничений по размеру.

Обновление BigQuery

Сервис BigQuery может обновляться оперативно, без простоев. На самом деле обновления происходят постоянно, по меньшей мере один раз в неделю. Обычно обновления выполняются медленно в течение нескольких дней, начиная с одной зоны в одном регионе в первый день, и в каждый последующий день охват увеличивается. Цель такого решения в том, чтобы проблемы, в случае их появления, коснулись как можно более узкого круга инфраструктуры.

Более того, BigQuery может обновляться даже без сбоев в обработке запросов. Перед началом обновления часть сегментов отключается (то есть они перестают принимать новые задания). Затем производится обновление этих сегментов. Сегменты спроектированы для достижения максимальной отказоустойчивости и, как правило, в них могут возникать лишь небольшие сбои во время выполнения запросов. Когда ведущие серверы запросов завершают обработку своих заданий, они также могут обновиться. Так как сбой в ведущем сервере запро-

¹ См. https://cloud.google.com/files/storage_architecture_and_challenges.pdf. Файловая система Colossus доступна клиентам Google Cloud в виде Google Cloud Storage (см. <https://cloud.google.com/storage/>).

сов может привести к перезапуску запросов, такие запросы могут выполняться значительно дольше.

Серверы заданий достаточно легко обновить, потому что они хранят свое состояние в Spanner; после перезапуска они могут продолжить работу с того момента, на котором остановились. Наконец, маршрутизаторы почти не имеют состояния, поэтому могут обновляться практически в любой момент. Планировщики имеют по одному ведущему серверу на кластер Dremel, поэтому при их обновлении сначала обновляются резервные блоки, а затем они переходят в режим аварийного переключения.

Система обработки запросов (Dremel)

Система Dremel была создана в 2006 году инженером, который устал ждать выполнения своих заданий в MapReduce. Dremel быстро завоевала популярность в Google. В какой-то момент 80% сотрудников Google так или иначе оказались ее активными пользователями.

Первоначально Dremel имела древовидную топологию. Запросы входили в корень, разветвлялись и отправлялись листьям, каждый из которых выполнял свою часть запроса. Результаты возвращались обратно вверх по дереву к корню. В настоящее время движок Dremel больше не использует структуру в виде фиксированного дерева, тем не менее кластеры Dremel по-прежнему часто называют «деревьями».

Следующий запрос легко вычислить с помощью дерева выполнения:

```
SELECT
  COUNT(*)
    , start_station_name
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire
GROUP BY 2
ORDER BY 1 DESC
LIMIT 10
```

Запрос просто сканирует таблицу и выполняет агрегирование. Сканирование может быть выполнено в листьях, агрегирование — в узлах, находящихся выше в дереве, а окончательное объединение — в корне. Если вы чего-то не поняли, не волнуйтесь: мы подробно опишем работу движка ниже.

В 2010 году архитектура Dremel была изменена для поддержки динамического построения планов выполнения. Дерево отлично подходит для обработки запросов определенных типов, а именно запросов сканирования-фильтрации-агрегирования, подобных примеру выше, но оно плохо подходит для более сложных запросов. Если запрос требует выполнения операции JOIN или имеет вложенные подзапросы, для его обработки придется выполнить несколько проходов по

дереву. Более того, разные проходы по дереву обрабатывают разные объемы данных и, следовательно, должны масштабироваться по-разному.

Вот пример запроса, который нельзя обработать с использованием простого статического дерева:

```
SELECT
  COUNT(*)
  , starts.start_station_id as point_a
  , ends.start_station_id as point_b
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire starts,
  `bigquery-public-data`.london_bicycles.cycle_hire ends
WHERE
  starts.start_station_id = ends.end_station_id
  AND ends.start_station_id = starts.end_station_id
  AND starts.start_station_id <> ends.start_station_id
  AND starts.start_date = ends.start_date
GROUP BY 2, 3
ORDER BY 1 DESC
LIMIT 10
```

Этот запрос находит пункты проката велосипедов в Лондоне, между которыми совершается большинство поездок в течение одного дня. Поскольку этот запрос выполняет соединение таблиц, для его обработки необходимы дополнительные уровни в дереве выполнения.

Современная архитектура Dremel, называемая Dremel X (потому что это 10-я версия), создает динамический план запроса, который может иметь любое количество уровней и даже изменяться во время выполнения. На рис. 6.3 показана упрощенная схема обработки запроса в Dremel X. Обратите внимание, что выполнение все еще может выглядеть как дерево, но из-за шага переупорядочения между каждым этапом обработки запроса могут добавляться дополнительные уровни.

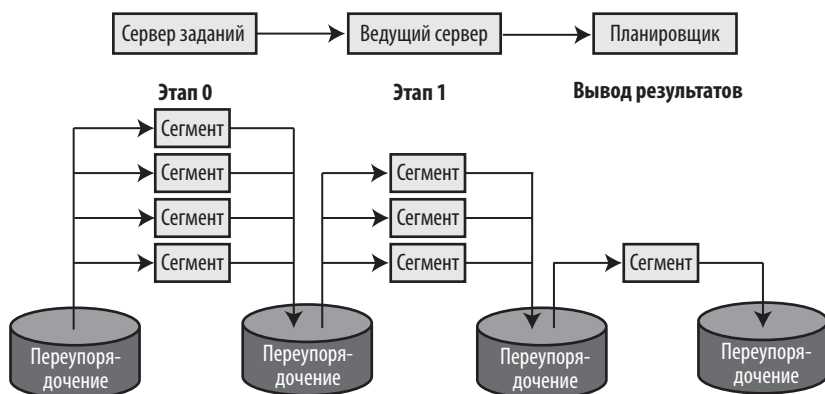


Рис. 6.3. Поток данных в Dremel X при обработке запроса с двумя этапами выполнения и одним этапом вывода результатов

Архитектура Dremel

Движок запросов делится на три части: ведущий сервер запросов, планировщик и рабочий сервер сегмента. Ведущий сервер запросов отвечает за планирование запросов (определяет, какие операции необходимо выполнить), планировщик отвечает за выделение слотов (определяет, какие серверы доступны для выполнения задания), а рабочие серверы сегментов отвечают за фактическую обработку запросов (выполнение задания). В этом разделе мы подробно опишем каждый из этих трех компонентов.

Ведущий сервер запросов

Ведущий сервер запросов отвечает за выполнение запроса. Первым делом он анализирует запрос и определяет таблицы, включенные в запрос, а также то, какие фильтры применяются к каждой таблице. Затем ведущий сервер получает метаданные таблиц от сервера метаданных, который возвращает местоположение файлов таблиц.

Фильтры необходимы для исключения ненужных секций. В BigQuery, если таблица секционирована по столбцам, а затем фильтруется по этим столбцам, можно избежать сканирования любых данных, отбрасываемых фильтром. Эти лишние данные исключаются из обработки. Для удаления ненужных разделов все фильтры передаются серверу метаданных перед выполнением запроса; фильтры секций могут передаваться до уровня базы метаданных, чтобы та вернула только соответствующие фильтрам местоположения файлов секций.

Среди прочих файлов сервер метаданных возвращает специальный метафайл. Этот метафайл определяет расположение файлов таблицы и то, как они отображаются в значениях полей. Чтобы извлечь эту информацию, ведущий сервер запросов выполняет еще один запрос Dremel к этому метафайлу.

После того как ведущий сервер запросов получит информацию, он будет знать, какой объем работы потребуется выполнить для обработки запроса. Это важно, потому что для планирования запроса планировщик должен знать, сколько слотов запланировать.

Однако прежде чем ведущий сервер запросов сможет запланировать запрос, ему необходимо сделать один важный шаг: создать план запроса. Планы запросов в BigQuery являются динамическими, но обработка запроса начинается с первоначального плана, который описывает выполнение запроса; обычно первоначальный план усложняется по мере необходимости. План запроса делит запрос на этапы, в каждом из которых выполняется набор операций. Пример плана запроса можно увидеть, посмотрев детали выполнения в веб-интерфейсе BigQuery.

После создания первоначального плана запроса ведущий сервер запросов обращается к планировщику, чтобы получить слоты для выполнения запросов. Планировщик выбирает рабочие серверы сегментов и возвращает их адреса

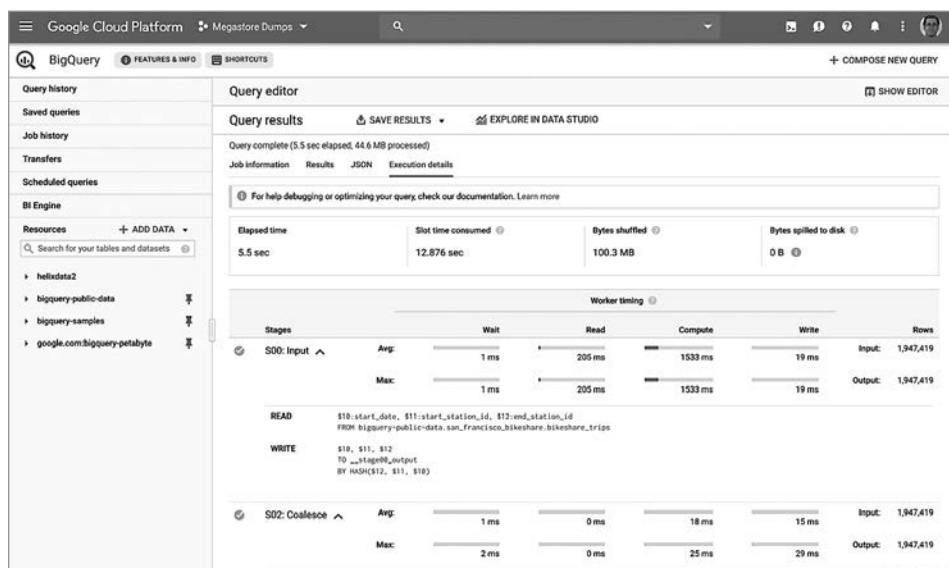


Рис. 6.4. План запроса в веб-интерфейсе BigQuery

ведущему серверу запросов. Затем ведущий сервер запросов отправляет единицы работы (обычно по одному файлу) в сегменты. Выполнение происходит параллельно в слотах, выделенных планировщиком. Если свободных слотов недостаточно, ведущий сервер запросов будет ждать, пока некоторые из рабочих серверов сегментов выполнят текущую работу, а затем запросит у планировщика выделить дополнительные слоты. Планировщик может в любой момент увеличить или уменьшить число слотов, выделенных для обработки запроса.

После того как рабочие серверы сегментов начнут возвращать результаты, ведущий сервер запросов снова обратится к планировщику и запросит слоты для выполнения второго и последующих этапов обработки запроса. После завершения последнего этапа ведущий сервер запросов вернет результаты серверу заданий.

Планировщик

Планировщик BigQuery отвечает за выделение слотов для обработки запросов. Слот — это единица работы, которая обычно соответствует обработке одного файла (этапы чтения) или *приемнику* (shuffle sink), подготавливающему данные к выполнению последующих этапов. Приемник — это временное хранилище для промежуточных результатов. Слоты — это потоки выполнения на рабочих серверах сегментов. В одной задаче на рабочем сервере сегмента может быть запущено множество потоков выполнения.

Один запрос может выполняться всего одним или миллионами слотов, в зависимости от объема обрабатываемых данных и от того, как данные размещены

физически. Запрос, обрабатывающий петабайт данных, может использовать 10 миллионов слотов. Очевидно, что не все 10 миллионов слотов можно запустить одновременно, поэтому планировщик выделит столько слотов, сколько сможет, а остальные выделит позже. Если посмотреть на график выполнения большого запроса (рис. 6.5), можно увидеть, что недостающие слоты активируются со временем. Количество входных единиц — это количество запланированных слотов, а высота графика «активных» слотов соответствует количеству выделенных слотов.

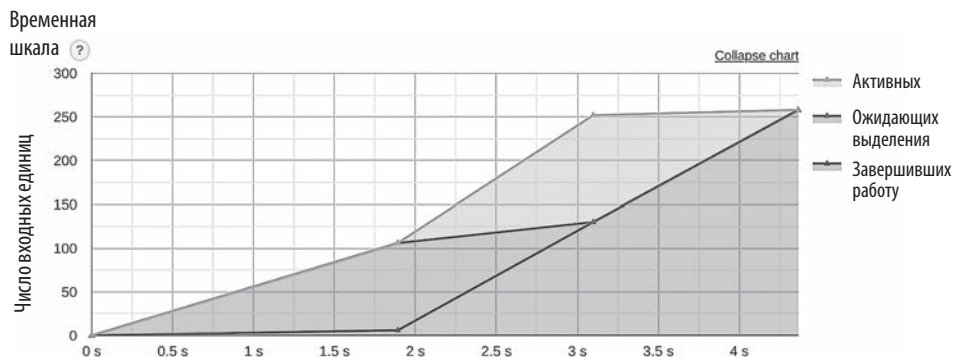


Рис. 6.5. График планирования слотов в BigQuery для обработки запроса

Планировщик играет роль арбитра, решающего, кому выделять ресурсы. На момент написания этой книги обычный пользователь BigQuery мог использовать до 2000 слотов, просто выполнив запрос. Однако этот объем ресурсов не гарантируется; если доступных слотов недостаточно, чтобы выделить каждому желающему до 2000 слотов, планировщик пропорционально сократит количество слотов, выделенных всем рядовым пользователям. Предположим, что общий пул составляет 100 000 слотов и все слоты уже используются 50 разными пользователями. Если появится новый пользователь, желающий выполнить запрос, для обработки которого необходимо 2000 слотов, количество слотов, выделенных каждому из предыдущих пользователей, будет уменьшено на 39 ($2000/51$) и новый пользователь получит те же 1961 слот, что и все остальные.

Планировщик проводит политику «справедливого» распределения ресурсов между пользователями с одинаковым приоритетом и запросами из одного и того же проекта. Если я запущу запрос, который использует 2000 слотов, а затем запущу второй запрос, пока выполняется первый, то первый запрос потеряет половину своих слотов, и обработка каждого запроса продолжится с использованием 1000 слотов. Если количество слотов, выделенных проекту, уменьшится до 1900 из-за высокой общей нагрузки на систему, каждому запросу будет выделено по 950 слотов.

Планировщик может в любой момент отменить запущенные слоты, чтобы передать ресурсы пользователю с более высоким приоритетом или обеспечить

их справедливое распределение. Каждая единица работы в BigQuery является атомарной и идемпотентной, то есть ее можно запустить, остановить и снова запустить. Это свойство также помогает при выходе из строя рабочих серверов сегментов; если рабочий сервер не ответит достаточно быстро, выполняемая им работа просто передается другому серверу и запрос продолжает обрабатываться как обычно, но, возможно, для завершения его обработки потребуются несколько дополнительных сотен миллисекунд. Если по какой-то причине одну и ту же единицу работы выполняют несколько рабочих серверов, результаты, выданные сервером, занявшим второе место, будут просто отброшены.

Некоторые пользователи BigQuery покупают «зарезервированные» слоты. Это означает, что они имеют на них преимущественное право. Такие пользователи гарантированно получают зарезервированное число слотов, когда они им потребуются. Они платят фиксированную сумму за доступ к этим слотам и могут запустить столько запросов, сколько захотят. Если их запросы потребуют больше слотов, чем зарезервировано, части этих запросов будут поставлены в очередь, в ожидании освобождения ресурсов.

Пользователь, который платит по фиксированному тарифу, может разделить свой проект на подразделы и назначить проекты каждому из этих подразделов. Например, если вы приобрели 5000 слотов, вы сможете связать проект А с подразделом «BI» (для анализа данных), а проект В — с подразделом «ETL». Обратите внимание, что названия BI и ETL ничего не значат для BigQuery; это просто обозначения, которые подсказывают сотрудникам вашей организации, для чего они используются. Затем вы можете выделить 4000 слотов для подраздела «BI», чтобы обеспечить одновременное выполнение множества запросов, а оставшиеся 1000 слотов выделить подразделу «ETL». Если окажутся занятыми все слоты, пользователи подраздела «BI» будут ограничены 4000 слотов, а пользователи подраздела «ETL» будут ограничены 1000 слотов. Если текущие рабочие нагрузки в подразделе «BI» используют только 2000 слотов, пользователи подраздела «ETL» смогут использовать оставшиеся 3000 слотов (и наоборот).

Рабочий сервер сегмента

Рабочий сервер сегмента, как нетрудно догадаться, отвечает за фактическую обработку запроса. Рабочий сервер сегмента (Worker Shard) — это задача, выполняемая в Borg,¹ — системе управления контейнерами компании Google, которая позволяет движку Dremel запускать тысячи задач в контейнерах, не беспокоясь об управлении оборудованием или инфраструктурой. Сам рабочий сервер сегмента способен параллельно выполнять несколько подзадач; каждая из них представляет единицу планирования, то есть вышеупомянутый слот.

¹ См. <https://ai.google/research/pubs/pub43438>. Система Borg была создана по образу и подобию Kubernetes и доступна в Google Cloud через Kubernetes Engine (см. <https://cloud.google.com/kubernetes-engine/>).

Рабочий сервер сегмента предоставляет интерфейс вызова удаленных процедур (Remote Procedure Call, RPC) для выполнения небольшой части одного этапа обработки запроса. Интерфейс RPC сообщает рабочему серверу сегмента, какую часть запроса выполнять и какие данные использовать. Если посмотреть на план выполнения в веб-интерфейсе BigQuery (см. рис. 6.4), можно заметить фрагмент кода SQL, который выполняется в сегменте. Большая часть этого фрагмента будет выглядеть как обычный код на SQL, но исходная и целевая таблицы могут показаться незнакомыми, особенно это касается этапов в середине обработки запроса.

Источниками данных для обработки запроса являются либо файлы в файловой системе Colossus, представляющие таблицы, указанные в запросе, либо результаты предыдущих этапов. Как правило, одному входному файлу соответствует один поток выполнения (слот), и рабочий сервер сегмента выполняет свою часть работы по обработке запроса, а затем записывает результат в указанное место назначения.

Местом назначения обычно является файловая система в памяти. Исключение составляют случаи, когда запрос требует сохранения большого объема данных, обычно это последний этап. В таком случае местом назначения будет файловая система Colossus. Файловая система в памяти обеспечивает кратковременное хранение данных между этапами запроса и позволяет выполнять их переупорядочение между этапами.

Переупорядочение

Переупорядочение — важная часть обработки данных в любой распределенной системе. В BigQuery переупорядочение позволяет распределять потоки данных между несколькими *приемниками*. Например, механизм переупорядочения может записать все, что начинается с «А», в приемник 1, а все, что начинается с «В», в приемник 2. На следующем этапе рабочий сервер сегмента может прочитать данные из приемника 1 и узнать, что у него есть доступ ко всем данным, начинающимся с «А», а другой рабочий сервер сегмента может прочитать данные из приемника 2 и узнать, что у него есть доступ ко всем данным, начинающимся с «В».

Количество сегментов, участвующих в выполнении этапа, в значительной степени зависит от количества приемников, куда выполняет запись механизм переупорядочения. Как правильно выбрать число приемников? Это сродни черной магии; BigQuery динамически изменяет количество приемников во время обработки запроса в зависимости от объема и формата вывода, как показано на рис. 6.6. Такое динамическое поведение способствует обработке запросов без риска исчерпать память. Тем не менее чем правильнее BigQuery оценит количество приемников с самого начала, тем быстрее будет обработан запрос.

Размер файловой системы в памяти ограничен; если за один этап требуется переупорядочить слишком много данных, они записываются на диск. Скорость работы с оперативной памятью на несколько порядков выше скорости работы с дисками. Каждый раз, когда запрос должен записать переупорядоченные результаты на диск, скорость его обработки значительно снижается. Заметить это

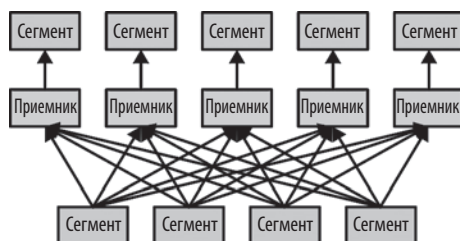


Рис. 6.6. Переупорядочение результатов, полученных четырьмя сегментами, и их распределение по пяти приемникам

можно, посмотрев статистику в плане обработки запроса после его завершения, как показано на рис. 6.7. Обратите внимание, что это одна из областей, которая продолжает активно развиваться, поэтому проблема снижения производительности со временем станет менее актуальной. Вы можете обойти эту проблему, разделив запрос на два (или более) запроса и обрабатывая в каждом разные диапазоны данных, а затем объединяя результаты.

Query complete (22.0 sec elapsed, 569 MB processed)			
Job information	Results	JSON	Execution details
? For help debugging or optimizing your query, check our documentation. Learn more			
Elapsed time	Slot time consumed ?	Bytes shuffled ?	Bytes spilled to disk ?
22.0 sec	51.361 sec	3.32 MB	0 B ?

Рис. 6.7. Окно «Execution details» (Сведения о выполнении) отображает объемы переупорядоченных байтов и байтов, записанных на диск

Выполнение запроса

Теперь, познакомившись с компонентами движка запросов Dremel, давайте обсудим, как они координируют обработку запроса. Сначала рассмотрим очень простой запрос, а затем перейдем к более сложному.

Запрос сканирования-фильтрации-подсчета

Самый простой полезный запрос — запрос сканирования-фильтрации-подсчета, то есть запрос, который читает содержимое таблицы, применяет фильтр и подсчитывает количество результатов. Например:

```
SELECT COUNT(*) as c
FROM `bigquery-public-data`.new_york_taxi_trips.tlc_yellow_trips_2017
WHERE passenger_count > 5
```

Этот запрос подсчитывает количество поездок в такси «Yellow Cab» города Нью-Йорка в 2017 году с числом пассажиров больше пяти. Когда запрос выполнится, можно открыть вкладку **Execution details** (Сведения о выполнении), чтобы увидеть план запроса, как показано на рис. 6.8.

Job information

Results

JSON

Execution details

For help debugging or optimizing your query, check our documentation. Learn more

Elapsed time

0.8 sec

Slot time consumed

2.432 sec

Bytes shuffled

81 B

Bytes spilled to disk

0 B

Worker timing

Stages		Wait	Read	Compute	Write	Rows
S00: Input	Avg:	<div><div></div></div> 1 ms	<div><div></div></div> 186 ms	<div><div></div></div> 133 ms	<div><div></div></div> 2 ms	Input: 113,496,874
	Max:	<div><div></div></div> 1 ms	<div><div></div></div> 307 ms	<div><div></div></div> 180 ms	<div><div></div></div> 3 ms	Output: 9
	Avg:	<div><div></div></div> 1 ms	<div><div></div></div> 0 ms	<div><div></div></div> 7 ms	<div><div></div></div> 3 ms	Input: 9
	Max:	<div><div></div></div> 1 ms	<div><div></div></div> 0 ms	<div><div></div></div> 7 ms	<div><div></div></div> 3 ms	Output: 1

Show debug panel

Рис. 6.8. План запроса типа «сканирование-фильтрация-подсчет»

План выполнения может дать нам много информации, но пока оставим эту тему и вернемся к ней в главе 7, где рассмотрим планы выполнения более подробно.

Обратите внимание, что на рис. 6.8 есть два этапа выполнения: S00 и S01. Первый — это вход; этап чтения данных из Colossus. Второй этап — этап вывода, в ходе которого происходит объединение результатов и их возврат пользователю.

Ведущий сервер запросов определяет объем данных, задействованных в запросе, и делит их на фрагменты (обычно один фрагмент соответствует одному файлу). В данном случае имеется девять файлов. Затем ведущий сервер запросов запрашивает у планировщика девять слотов. Поскольку в системе есть достаточное количество свободных слотов, планировщик возвращает информацию о девяти различных рабочих серверах сегментов. Вооружившись этой информацией, ведущий сервер запросов посылает запросы всем девяти рабочим серверам.

После получения запроса рабочими серверами сегментов начинается этап 0.

Этап 0. На этапе 0 каждый рабочий сервер сегмента должен прочитать файл, оставить только поездки с более чем пятью пассажирами и затем подсчитать их. Поскольку результаты вычисляются параллельно, ни один из сегментов не владеет полным объемом информации, необходимой для вычисления общего результата; поэтому промежуточные результаты должны быть переданы на следующий этап.

КАК УЗНАТЬ, СКОЛЬКО СЛОТОВ БЫЛО ИСПОЛЬЗОВАНО?

Когда на этапе ввода выполняется полное агрегирование (возвращается только одна запись), общее число записей на выходе этапа будет равно числу его входов. Когда количество входов велико, оно будет больше количества выделенных слотов. Следовательно, допущение, что для каждого входа выделяется отдельный слот, будет неверным (обратите внимание, что количество записей не связано напрямую с количеством входов). Но для такого маленького запроса почти наверняка будет доступно девять слотов, поэтому мы видим, что для обработки запросов использовано девять слотов. Чтобы подтвердить это, можете посмотреть на информацию из статистики запроса, воспользовавшись инструментом командной строки bq:

```
bq --format=prettyjson show -j <my_job_id> \
  | grep completedParallelInputs
```

Эта команда вернет следующие результаты:

```
"completedParallelInputs": "9",
"completedParallelInputs": "1",
```

Они подтверждают, что первый этап имел девять входов (файлов), обрабатываемых параллельно, а второй — один вход.

Обратите внимание, что по этим результатам можно также определить, когда увеличение числа слотов может улучшить производительность, потому что вы увидите большое время ожидания. Если на ожидание слотов тратится время, значит, увеличение их числа сократит время ожидания и ускорит обработку запроса.

Если развернуть первый этап (S00), как показано на рис. 6.9, можно увидеть проделанную работу.

Stages		Wait	Read	Compute
✓ S00: Input ^	Avg:	1 ms	186 ms	133 ms
	Max:	1 ms	307 ms	180 ms
READ	\$1:passenger_count FROM bigquery-public-data.new_york_taxi_trips.tlc_yellow_trips_2017 WHERE greater(\$1, 5)			
AGGREGATE	\$20 := COUNT_STAR()			
WRITE	\$20 TO __stage00_output			

Рис. 6.9. Работа, выполненная на первом этапе обработки запроса типа «сканирование-фильтрация-подсчет»

Первые две части читают из таблицы число пассажиров в поездке, применяют фильтр и подсчитывают результаты. Последняя часть гласит: «Записать результаты в выход `stage_00`». Это место в памяти, которое будет служить источником данных для последующих этапов. `COUNT_STAR()`, — это внутренний оператор, подсчитывающий количество записей.

После вычисления этих промежуточных результатов и записи их в назначенную область вывода каждый сегмент возвращает управление ведущему серверу запросов.

После выполнения этапа 0. Когда первый рабочий сервер сегмента вернет результат после выполнения этапа 0, ведущий сервер запросов может запланировать этап 1. Для этого он вновь обращается к планировщику и запрашивает дополнительные слоты; в данном случае требуется только один слот, потому что необходимо подсчитать всего лишь сумму девяти значений. Следующий этап может начаться до завершения этапа 0, потому что он будет продолжать читать выходные данные из `stage_00`, пока файл не закроется. Но в данном случае это не дает большого ускорения, потому что на этапе 1 не так много работы.

Этап 1. На этапе 1 выполняется простая работа; считывается девять входных данных и вычисляется их сумма, как показано на рис. 6.10.

S01: Output ^		Avg:	1 ms	0 ms
		Max:	1 ms	0 ms
READ	\$20 FROM __stage00_output			
AGGREGATE	\$10 := SUM_OF_COUNTS(\$20)			
WRITE	\$10 TO __stage01_output			

Рис. 6.10. Работа, выполненная на втором этапе обработки запроса типа «сканирование-фильтрация-подсчет»

Вот почему этап 1 выполнялся всего одну миллисекунду. После того как этап 1 вычислит сумму девяти значений, читая их из файла в памяти, он запишет результаты в окончательный вывод.

Когда этап 1 завершится, ведущий сервер запросов прочитает результат из вывода этапа 1 и вернет его серверу заданий. Запрос обработан, и теперь пользователь еще на шаг ближе к моменту, когда узнает, что в 2017 году было выполнено более трех миллионов поездок на такси с числом пассажиров больше пяти.

Когда ведущий сервер запросов вернет результат серверу заданий, тот сможет вернуть их клиенту. Поскольку объем результатов небольшой, они записываются в Spanner, чтобы клиент смог получить их, когда ему это потребуется.

Запрос сканирования-фильтрации-агрегирования

Следующий наиболее простой тип запросов — запросы сканирования-фильтрации-агрегирования. Запросы этого типа можно выполнить за один обход данных. Чтобы показать, что происходит в действительности, используем новую таблицу, содержащую один миллиард журнальных записей о просмотре страниц из Википедии. Обратите внимание, что если вы попытаете выполнить этот запрос у себя, это может влететь вам в копеечку, поэтому для многих из вас, вероятно, лучше ознакомиться с нашими результатами в книге. Вот первый запрос, который мы выполним:

```
SELECT title, COUNT(title) as c
FROM `bigquery-samples.wikipedia_benchmark.Wiki1B`
WHERE title LIKE "G%o%g%l%e"
GROUP BY title
ORDER BY c DESC
```

Он ищет страницы с буквами «G», «o», «o», «g», «l» и «e», которые следуют в таком порядке, подсчитывает количество просмотров каждой страницы и возвращает их в порядке популярности:

Row	title	c
1	Google	2904
2	Google_Chrome	1302
3	Google_Wave	623
4	Google_Translate	561
5	Google_AdSense	426

Этап 0. Первый этап отправляет этот запрос множеству рабочих серверов сегментов для параллельной обработки, каждый из которых читает столбец `title` и отфильтровывает записи, не совпадающие с условием в предложении `WHERE`. Затем выполняется частичное агрегирование. На рис. 6.11 показано, как выглядит план выполнения.

Шаг **READ (ЧТЕНИЕ)** читает столбец `title` (остальные столбцы не нужны) и применяет фильтр. Шаг **AGGREGATE (АГРЕГИРОВАНИЕ)** подсчитывает число записей, миновавших фильтр. Обратите внимание, что это количество не является окончательным: это только количество заголовков, соответствующих фильтру, встретившееся в текущем файле.

Stages		Wait	Read	Compute	Write	Rows
✓ S00: Input ^	Avg:	1869 ms	1931 ms	2211 ms	8 ms	Input: 1,249,541,131
	Max:	2556 ms	3716 ms	2837 ms	219 ms	Output: 28,693
READ \$1:title FROM bigquery-samples.wikipedia_benchmark.Wiki1B WHERE like(\$1, '%%o%g%l%e')						
AGGREGATE GROUP BY \$30 := \$1 \$20 := COUNT(\$1)						
WRITE \$30, \$20 TO __stage00_output BY HASH(\$30)						

Рис. 6.11. Работа, выполненная на первом этапе обработки запроса типа «сканирование-фильтрация-агрегирование»

Шаг **WRITE (ЗАПИСЬ)** записывает результат в `__stage00_output`, как и предыдущий запрос, но на этот раз он делает еще кое-что: добавляет директиву **BY HASH**, которая требует выполнить переупорядочение по значению. Переупорядочение используется для отправки результатов в разные корзины. Эти корзины распределяются между рабочими серверами сегментов, поэтому все сегменты, встретившие определенное значение, отправят свои результаты в одну и ту же корзину. В данном случае переупорядочение происходит по значению в столбце `title`, то есть каждый раз, когда встречается заголовок «Google», он попадает в одно и то же место, а когда появляется заголовок «Good night Seattle» (спокойной ночи, Сиэтл), он попадает в другое место. Эта маршрутизация обеспечивает возможность вычисления общих результатов.

Поскольку некоторые запросы могут обрабатывать миллиарды и триллионы различных значений (вы увидите это в следующем запросе), вместо создания одной корзины для каждого значения мы применяем математическую хеш-функцию (именно это означает инструкция **BY HASH**) и используем результат как название корзины. Одному и тому же входному значению всегда будет соответствовать одно и то же выходное значение, что важно для организации обработки результатов на одном рабочем сервере сегмента. Кроме того, несколько разных входных значений могут оказаться в одной корзине, что уменьшает количество уникальных корзин.

Столбец **Rows (Записи)** на рис. 6.11 тоже представляет определенный интерес. В отличие от предыдущего запроса, здесь мы не видим, сколько слотов использовано, потому что каждый рабочий сервер сегмента производит несколько выходных записей. Количество выходных записей соответствует количеству уникальных значений, встреченных каждым рабочим сервером сегмента. То есть если каждый рабочий сервер встретит 100 разных значений и всего было задействовано 200 сегментов, тогда общее количество выходных записей будет равно 20 000 (100 * 200). В нашем случае произведение числа использованных слотов на число уникальных значений оказалось равным 28 693.

ЧТО ТАКОЕ ХЕШ (HASH)?

Термин *хеш* (*hash*) часто встречается при описании работы BigQuery. Этот термин хорошо знаком специалистам в области информатики, но он не так распространен в других сферах. Хеширование — это метод деления неизвестного распределения данных на фиксированное количество корзин. В параллельных распределенных системах хеширование помогает разделить работу между параллельными рабочими процессами. Одно из ключевых свойств хеш-функции — одинаковые данные имеют одинаковый хеш и всегда оказываются в одной и той же корзине.

Предположим, что у вас есть данные о продаже всей недвижимости в Сиэтле и вы хотите узнать, какие дома продавались чаще всего в последнее десятилетие. Чтобы проделать такую работу в одиночку, потребуется уйма времени, поэтому предположим, что вы выбрали девять своих друзей себе в помощь. Для начала вы можете разделить записи по годам и распределить их между вами и вашими друзьями — каждый будет читать данные за один год и подсчитывать частоту продажи каждого дома.

Но как объединить результаты? Вы не можете просто взять наиболее часто продаваемые типы домов за каждый год, потому что дом, который продавался каждый год, встретится каждому из вас только один раз, но это может быть тот дом, который вы ищете. Между тем дом, который в один год был продан трижды, мог не продаваться в другие годы. Очевидно, что объединение результатов — сложная задача, и решить ее можно с помощью хеш-функции.

Вам нужно найти способ разделить между собой не только исходные данные, но и дома, чтобы каждый раз, когда продавался дом 601 N на 34-й улице, эта запись оказывалась у одного и того же человека. Тогда он сможет сообщить о своих самых продаваемых домах и игнорировать все остальные.

Для решения задачи распределения можем применить хеш-функцию к адресу. Хеш-функция, которую мы здесь используем, проста: она складывает цифры в адресе и возвращает последнюю цифру результата. То есть для адреса «931 Крокет-авеню» она получит сумму 13 и вернет последнюю цифру «3». Для адреса «2444 Вторая авеню» она получит сумму 14 и вернет последнюю цифру «4». Соответственно, каждый из 10 друзей выберет число от нуля до девяти и получит для анализа дома, адреса которых хешируются в это число. Если к вам присоединится еще один человек, чтобы помочь, и вам потребуется распределить данные по 11 корзинам вместо 10, вы можете разделить число суммы на 11 и взять остаток (именно эту операцию представляет взятие последней цифры в случае с 10 корзинами). Эта операция называется *взятием модуля*, или взятием остатка от деления.

Этап 1. На следующем этапе выполняется чтение из переупорядоченных результатов этапа 0 и производится окончательное агрегирование. На рис. 6.12 представлена соответствующая часть информации о выполнении запроса.

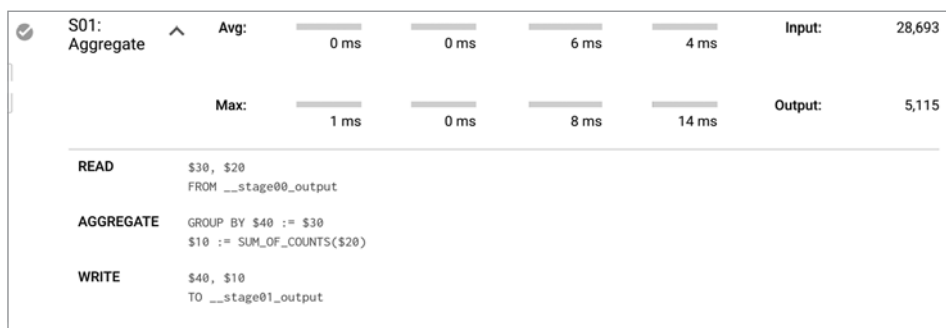


Рис. 6.12. Работа, выполненная на первом этапе обработки запроса типа «сканирование-фильтрация-агрегирование»

Поскольку входные данные переупорядочиваются, мы можем выполнить окончательное агрегирование параллельно. То есть, поскольку все промежуточные результаты подсчета значений в столбце `title`, соответствующих «Google», были отправлены в одну и ту же корзину, мы можем вычислить общее количество для слова «Google», просто подсчитав число значений в корзине «Google». Другой рабочий сервер может вычислить итоговый результат для «Google_Chrome», подсчитав число значений в корзине «Google_Chrome».

Всего получилось 5115 выходных записей, что соответствует общему числу всех записей. Мы не использовали ограничение, поэтому все они будут возвращены пользователю.

Этап 2. Этап 2 чрезвычайно прост: он просто читает 5115 значений и сортирует их, как показано на рис. 6.13.

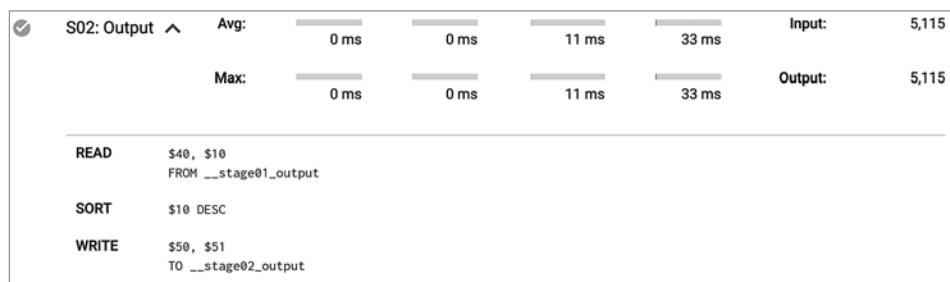


Рис. 6.13. Работа, выполненная на втором этапе обработки запроса типа «сканирование-фильтрация-агрегирование»

Операция сортировки выполняется единственным рабочим сервером сегмента, если, конечно, число значений не очень велико, в противном случае используется алгоритм распределенной сортировки.

Запрос сканирования-фильтрации-агрегирования с большим объемом результатов

Теперь посмотрим, что произойдет, если попытаться выполнить тот же запрос, но без фильтра. Он обнаружит миллионы разных значений в столбце `title` (в таких случаях говорят, что набор данных имеет большую мощность множества по столбцу `title`), поэтому если выбрать слишком маленькое число корзин для хеширования, небольшому количеству рабочих серверов придется проделать большой объем работы и для выполнения потребуется много времени.

Попробуем снова выполнить тот же запрос, но без фильтрации по столбцу `title`, чтобы определить популярность всех страниц в Википедии:

```
SELECT title, COUNT(title) as c
FROM `bigquery-samples.wikipedia_benchmark.Wiki1B`
GROUP BY title
ORDER BY c DESC
```

Для обработки этого запроса потребуется значительно больше времени, потому придется подсчитать количество намного большего числа страниц. Из-за отсутствия фильтрации он вернет более 280 миллионов записей.

Теперь в разделе с информацией о выполнении запроса появится 15 этапов (0–9, A–E).

Этап 0. Этап 0 идентичен предыдущему случаю, за исключением отсутствия фильтра, поэтому он вернет 1 205 625 714 значений. Этапы с 1-го по 8-й — новые. На рис. 6.14 показано, как выглядит этап 1.

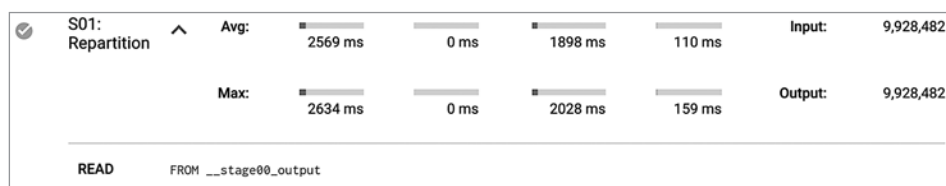


Рис. 6.14. Работа, выполненная на первом этапе обработки запроса типа «сканирование-фильтрация-агрегирование» с большой мощностью множества включает повторное разделение

Рабочие серверы, вовлеченные в выполнение этапа 1, читают данные из входов, но ничего и никуда не записывают. Это обусловлено тем, что хеш-корзины стали слишком большими и их требуется разбить на сегменты. По сути, это означает, что BigQuery выбрала несколько слишком маленьких хеш-корзин и теперь должна перераспределить данные по большему числу корзин. Это позволит избежать осложнений на более поздних этапах.

Распределенная сортировка. Мы также видим, что появились новые этапы — В и С (они имеют буквенные имена, потому что вставлены в исходный план запроса после этапа 0 из-за переупорядочения), как показано на рис. 6.15.

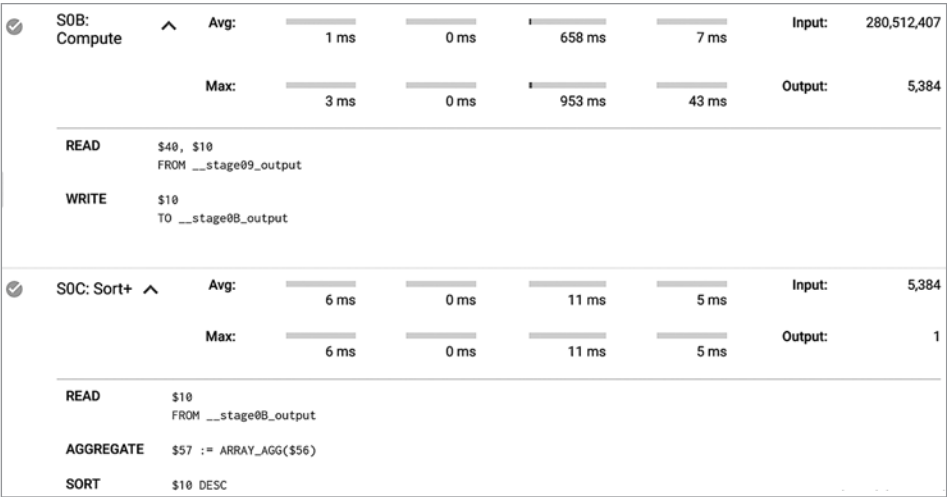


Рис. 6.15. Работа, выполненная на этапах В и С обработки запроса типа «сканирование-фильтрация-агрегирование» с большой мощностью множества

Как уже говорилось выше, этот запрос возвращает 280 миллионов записей в отсортированном порядке. 280 миллионов значений — это слишком много для сортировки на одном узле.

Представьте, что кто-то вынул все страницы из словаря и предложил вам с друзьями собрать их в правильном порядке. С этой целью, как вариант, можно дать каждому человеку диапазон букв. Например, вы можете взять буквы от А до В, следующий — от Г до Е, и т. д. Тогда каждый сможет просмотреть страницы и выбрать только те, которые содержат словарные статьи, начинающиеся с буквы в его диапазоне. После того как все соберут свои страницы, вы сможете отсортировать свою небольшую стопку, а затем объединить со стопками, собранными вашими друзьями. Это один из способов распределенной сортировки.

Однако проблема в том, что вы можете не знать заранее, как распределены страницы. Если один человек получил только букву «Э», у него будет очень мало работы, потому что не так много слов, начинающихся с этой буквы.

Этап В вычисляет распределение, или *точки разделения*, для распределенной сортировки. Мы в нашем примере сортируем страницы по количеству просмотров (вычисляемому значению). На основании объема данных выбирается степень параллелизма (то есть число «друзей», которые будут вам помогать). В данном случае у нас есть 5384 рабочих сервера сегментов (см. вторую строку,

последний столбец на рис. 6.15). BigQuery просматривает данные, определяет примерные границы каждой 1/5384 части данных и затем выводит их. Этап С просто получает эти значения и помещает их в массив в одну запись, чтобы использовать позже.

Этап D выполняет фактическое разбиение данных. Его можно сравнить с группой друзей, просматривающих и отыскивающих страницы, которые соответствуют их диапазонам. Мы читаем точки разделения на этапе С, а итоговый подсчет производим на этапе 9 (мы пропустили описание этапа 9, потому что он похож на этапы, которые мы видели раньше). Эти итоговые значения записываются в 5384 корзины (обратите внимание, что здесь этого не видно), которые отображаются на диапазон значений, как показано на рис. 6.16.



Рис. 6.16. Работа, выполненная на этапе D обработки запроса типа «сканирование-фильтрация-агрегирование» с большой мощностью множества

Этап Е выводит выходные данные. Он выполняет распределенную сортировку и записывает окончательные результаты. Поскольку каждый рабочий сервер уже подсчитал промежуточные результаты, которые отображаются в неперекрывающиеся диапазоны, на этапе Е можно просто отсортировать значения локально и вывести их. Окончательные результаты будут находиться в разных файлах, и эти файлы необходимо прочитать в определенном порядке, соответствующем порядку сортировки общих результатов, как показано на рис. 6.17.



Рис. 6.17. Работа, выполненная на этапе Е обработки запроса типа «сканирование-фильтрация-агрегирование» с большой мощностью множества

Запросы со всенаправленными соединениями (JOIN)

До сих пор все запросы, которые мы рассматривали, были направлены только к одной таблице. А как выполняется соединение таблиц? В BigQuery поддерживается два типа соединений: *всенаправленные* и *хешированием*. Мы уже видели, как хеширование используется для агрегирования, поэтому начнем со всенаправленных соединений, потому что они проще. На момент написания этой книги всенаправленные соединения можно было использовать, если одна из таблиц имеет небольшой объем: не более 150 Мбайт.

При выполнении всенаправленных соединений меньшая таблица пересылается всем рабочим серверам. Если для обработки запроса задействовано 100 рабочих серверов, обрабатывающих большую таблицу, тогда меньшая таблица пересылается каждому из этих 100 рабочих серверов. Это довольно грубый способ объединения, но его преимущество в том, что его можно выполнить за один проход через большую таблицу и он не требует переупорядочения.

Чтобы понять суть, представьте, как происходит соединение таблиц. Записи в двух или более таблицах сопоставляются с ключом, поэтому рабочему серверу нужно просмотреть все совпадающие значения в них. То есть если запись в таблице слева соответствует ключу «123», ее необходимо сопоставить со всеми записями, соответствующими тому же ключу «123» в таблице справа (или, если таких записей нет, об этом тоже нужно знать). Чтобы выполнить это сопоставление, в одном и том же месте нужно получить записи из таблиц слева и справа, соответствующие определенному ключу. При всенаправленном соединении для этого просто производится рассылка одной из таблиц повсюду, поэтому все сопоставления происходят в одном месте.

В этом примере используем образец набора данных GitHub, который содержит информацию обо всех операциях фиксации состояния (коммитах) репозиториях в GitHub, имевших место в истории этого сервиса. Выполним соединение двух таблиц: таблицы `commits`, содержащей информацию о каждой операции фиксации, и таблицы `languages`, содержащей информацию об используемых языках программирования. Запрос, который мы рассмотрим, выбирает изменения в GitHub, произведенные авторами книги, и сортирует их по количеству байтов кода, записанных в репозиторий, на каждом языке программирования. Вот как выглядит сам запрос:

```
WITH
repo_commits AS (
  SELECT repos AS repo_name, author.name AS author
  FROM `bigquery-public-data.github_repos.commits` c, c.repo_name repos
  WHERE author.name IN ("Valliappa Lakshmanan", "Jordan Tigani")
  GROUP BY repos, author),
repo_languages AS (
  SELECT lang.name AS lang, lang.bytes AS lang_bytes, repos.repo_name AS
                                                                    repo_name
  FROM `bigquery-public-data.github_repos.languages` repos, repos.LANGUAGE
                                                                    AS lang
)
```

```
SELECT lang, author, SUM(lang_bytes) AS total_bytes
FROM repo_languages
JOIN repo_commits USING (repo_name)
GROUP BY lang, author
ORDER BY total_bytes DESC
```

Выполнив его, мы получили следующие результаты:¹

Row	lang	author	total_bytes
1	Jupyter Notebook	Valliappa Lakshmanan	78900202
2	Python	Valliappa Lakshmanan	33742613
	...		
8	Jupyter Notebook	Jordan Tigani	153243
9	Python	Jordan Tigani	134409
	...		

В этом запросе используются некоторые достаточно продвинутые приемы, такие как разворачивание массивов с помощью операторов `CROSS JOIN` и `WITH`, что позволяет сделать запрос более удобочитаемым (см. главу 8). Ключевая часть запроса:

```
SELECT lang, author, SUM(lang_bytes) AS total_bytes
FROM repo_languages
JOIN repo_commits USING (repo_name)
GROUP BY lang, author
ORDER BY total_bytes DESC
```

Эта часть запроса выполняет соединение таблицы `languages` с таблицей `commits` по столбцу с названием репозитория.

Если открыть план этого запроса, в нем практически все остается таким же, кроме двух новых типов этапов: *coalesce* и *join+*. Этап слияния (*coalesce*) очень прост, он показан на рис. 6.18.

Этап слияния добавляется динамически, когда BigQuery обнаруживает, что одна из таблиц в соединении имеет небольшой объем. Он собирает все данные из таблицы на одном узле, чтобы потом разослать ее. Следует заметить, что предыдущие этапы могли бы получить большую таблицу и отфильтровать ее,

¹ Замечание от Джордана: если вам это интересно, но вы не хотите выполнять запрос, у Лака (моего соавтора) объем изменений в GitHub намного больше, чем у меня. Замечание от Лака: это связано прежде всего с расточительным форматом хранения блокнотов Jupyter.

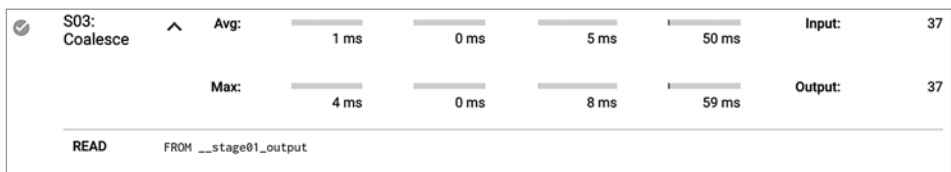


Рис. 6.18. Этап слияния (coalesce) в запросе со всенаправленным соединением

превратив в таблицу меньшего размера (как это делал наш запрос, осуществлявший фильтрацию фиксации изменений по именам авторов этой книги). Слияние не меняет количества записей; на этом этапе просто все записи перемещаются в одно место.

Другой новый этап, join+, показан на рис. 6.19.

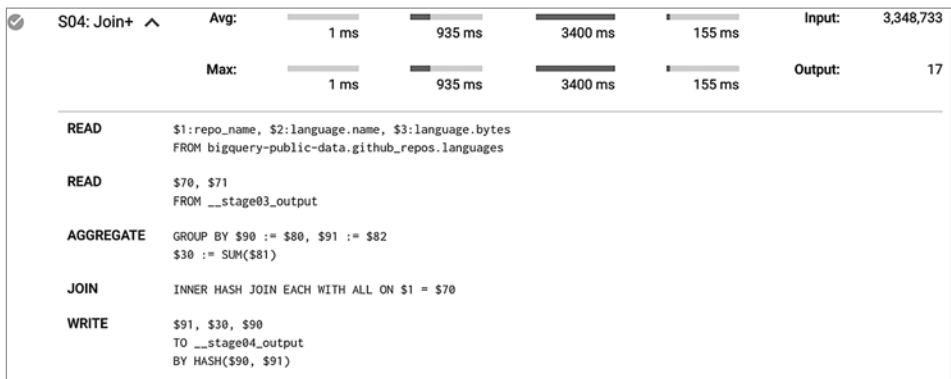


Рис. 6.19. Этап join+ в запросе со всенаправленным соединением

Этот этап называется join+, а не join, потому что он выполняет сразу две операции: соединение и агрегирование. Здесь мы видим два оператора READ, один из которых читает таблицу слева, а другой — справа. В данном случае слева находится таблица `languages`, а справа — рассылаемая таблица, полученная на предыдущем этапе слияния. На момент написания этой книги всенаправленное соединение можно было идентифицировать только одним способом — по тексту «EACH WITH ALL» (КАЖДЫЙ СО ВСЕМИ). Этот текст означает, что нужно взять каждую запись из таблицы слева и сопоставить ее со «ВСЕМИ» (ALL) записями из таблицы справа.

Запросы с соединениями хешированием

Второй вид соединений — соединение хешированием. Этот способ намного более дорогостоящий в вычислительном отношении. Соединение хешированием про-

изводится путем хеширования обеих сторон соединения, благодаря чему записи, содержащие один и тот же ключ, попадают в одну корзину. Это тот же процесс хеширования, который мы видели, когда говорили о запросах с агрегированием, но в данном случае он применяется к обеим таблицам. Поскольку процесс хеширования посылает все эквивалентные значения в одну корзину, один рабочий сервер может выбрать корзину и получить всю информацию, необходимую для выполнения соединения ключей в ней.

Для иллюстрации соединения хешированием используем предыдущий запрос, но прокомментируем выражение фильтрации, чтобы соединение выполнялось для полных таблиц. Обе таблицы в этом случае получатся слишком большими, чтобы уместиться в памяти, поэтому вместо всенаправленного соединения будет выполняться соединение хешированием. Ниже показан измененный текст запроса:

```
WITH
repo_commits AS (
  SELECT repos AS repo_name, author.name AS author
  FROM `bigquery-public-data.github_repos.commits` c, c.repo_name repos
  -- WHERE author.name IN ("Valliappa Lakshmanan", "Jordan Tigani")
  GROUP BY repos, author),
repo_languages AS (
  SELECT lang.name AS lang, lang.bytes AS lang_bytes, repos.repo_name
                                     AS repo_name
  FROM `bigquery-public-data.github_repos.languages` repos, repos.LANGUAGE
                                     AS lang
)

SELECT lang, author, SUM(lang_bytes) AS total_bytes
FROM repo_languages
JOIN repo_commits USING (repo_name)
GROUP BY lang, author
ORDER BY total_bytes DESC
LIMIT 100
```

Он возвращает похожие результаты:

Row	lang	author	total_bytes
1	C	Eric Dumazet	2917514359851
2	C	Russell King	2878666474184
3	C	Thomas Gleixner	2876903624978

План этого запроса выглядит практически идентично плану запроса со всенаправленным соединением. За исключением того, что исчез этап объединения, появилась пара дополнительных этапов переупорядочения и механизм запросов точнее определяет количество требуемых корзин. Существует также небольшое отличие в этапе `join+`, которое показано на рис. 6.20.

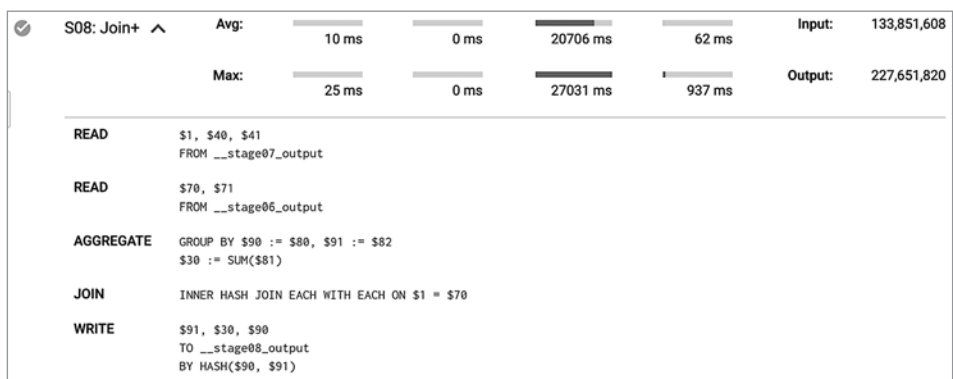


Рис. 6.20. Этап join+ для запроса с соединением хешированием

Здесь мы видим, что выполняется соединение **EACH WITH EACH** (КАЖДЫЙ С КАЖДЫМ), а не **EACH WITH ALL** (КАЖДЫЙ СО ВСЕМИ). Это означает, что каждая запись слева соединяется с каждой соответствующей записью справа, что требует их предварительного извлечения. Если посмотреть на предыдущие этапы, также можно увидеть, что входные данные были переупорядочены (**HASH**), как обсуждалось в разделе, посвященном запросам сканирования-фильтрации-агрегирования.

Хранилище

Одним из секретов успеха любой системы управления базами данных является эффективное хранение. Многие ключевые возможности, которые обеспечивают высокую скорость работы BigQuery, опираются на эффективное хранение данных. От аппаратного обеспечения (с гигантской распределенной файловой системой) до форматов файлов (обеспечивающих колоночный способ хранения) стек технологий хранения в BigQuery, включая данные и метаданные, оптимизирован для максимальной скорости анализа.

Хранение данных

BigQuery хранит эксабайты данных, распределенных по миллионам физических дисков в десятках регионов. Главная задача системы хранения нижнего уровня — обеспечить быстрый доступ ко всем этим распределенным данным и возможность соединения любых двух таблиц, для чего те должны находиться в одном месте.

Один из секретов успеха крупномасштабной аналитики заключается в том, что наибольшее увеличение производительности дает усовершенствование систе-

мы хранения. В этом разделе мы расскажем, как работает система хранения в BigQuery и что обеспечивает высокую скорость ее работы.

Когда вы загружаете данные в BigQuery, они передаются механизму хранения *Saracitor* и сохраняются в файловой системе *Colossus*. *Colossus* кодирует данные, используя алгоритм *стирающего кодирования* (*erasure encoding*), что обеспечивает их сохранность, даже если значительное число дисков выйдет из строя. Записи в единственный кластер *Colossus* достаточно, чтобы обеспечить устойчивость данных с очень большим количеством девяток.¹

Для обеспечения надежности и доступности данных они реплицируются в другую зону доступности в том же регионе. На практике это означает копирование данных в другой вычислительный центр с другой системой электропитания и сетевым оборудованием. Вероятность одновременного отключения нескольких зон доступности очень мала. Но что произойдет, если весь регион будет обесточен, например, из-за нападения Годзиллы или, что менее вероятно, из-за землетрясения или другого стихийного бедствия? Если вы используете объединенные регионы BigQuery (например, US или EU), BigQuery сохранит другую копию данных в реплике вне региона; благодаря этому сохранится возможность восстановить данные в случае крупного бедствия. Более подробно доступность и аварийное восстановление мы обсудим в главе 10.

Физическое хранилище: Colossus

Все свои данные BigQuery хранит в *Colossus* — распределенной системе хранения, используемой в Google. *Colossus* — это потомок файловой системы Google (*Google File System*, *GFS*) — новаторской крупномасштабной системы распределенного хранения, разработанной в Google. *Colossus* решает ряд задач масштабируемости, гибкости и надежности в *GFS* за счет более гибкой системы метаданных и отсутствия единых точек отказа.

Colossus управляет большим количеством дисков на большом количестве серверов, объединенных в единую файловую систему. Если у вас есть десятки или сотни тысяч дисков, десятки из этих дисков будут каждый день выходить из строя. Поэтому главная задача состоит в том, чтобы не потерять никаких данных (как минимум в течение нескольких миллионов лет). Чтобы избежать потери данных при выходе дисков из строя, их нужно записать несколько раз. В *Colossus* этот процесс называется *кодированием* (*encoding*).

Простейший вид кодирования называется *кодированием с реплицированием* (*replicated encoding*). В кодировании с реплицированием (рис. 6.21) просто создается несколько копий данных. Сколько копий нужно? Это зависит от желаемого уровня безопасности. Создав две копии, можно потерять данные, если два диска выйдут из строя. Три копии могут обеспечить достаточную степень безопасно-

¹ Имеется в виду надежность на уровне 99.999...%.

сти, если в организации практикуется политика быстрой замены оборудования. Вероятность того, что все три диска выйдут из строя в один и тот же день, равна примерно одному к 100 миллионам. Надежность распределенного хранилища определяется частотой сбоев и скоростью замены вышедших из строя устройств. Мы не можем управлять частотой выхода оборудования из строя, но нам вполне под силу разработать процесс его быстрой замены. При быстрой замене дисков вероятность потери данных можно оценить как раз в 10 миллиардов лет или около того. Конечно, жизнь намного сложнее, дисков очень много, и каждый из нас хотел бы быть уверенным в том, что потеря данных будет происходить крайне редко.



Рис. 6.21. Кодирование с реплицированием, когда все фрагменты файлов сохраняются в трех разных местах

Однако репликация файлов — дорогое удовольствие из-за необходимости хранить полные копии данных. Чтобы уменьшить объем хранимых данных, многие распределенные файловые системы используют так называемое *стирающее кодирование* (erasure encoding), или *кодирование Рида—Соломона* (Reed—Solomon encoding). При использовании стирающего кодирования на других дисках хранятся математические функции данных, и за счет сложности восстановления экономится дисковое пространство, как показано на рис. 6.22. В зависимости от выбора способа кодирования данных, можно добиться гораздо большей стойкости, чем позволяет кодирование с реплицированием, и при этом хранить меньше одной полной дополнительной копии. Однако есть одна проблема: если основная копия недоступна, для восстановления данных может понадобиться прочесть несколько дополнительных дисков.



Рис. 6.22. Стирающее кодирование, в котором дополнительные «закодированные» данные можно использовать для восстановления информации в случае потери оригинальных фрагментов

Большую часть данных BigQuery хранит, применяя стирающее кодирование и используя достаточное число блоков восстановления, чтобы обеспечить на несколько порядков более высокую надежность, чем позволяет добиться кодирование с трехсторонней репликацией. Кодирование с репликацией обычно

позволяет читать данные быстрее — если вы не получите ответ от первой реплики в течение короткого времени, вы можете отправить запрос на чтение одной из других копий. При использовании стирающего кодирования, если нет возможности прочитать основную копию, можно запустить процесс чтения с восстановлением, но это может потребовать выполнения дополнительных операций чтения. Если одна из них выполняется медленно, можно попробовать восстановить другие фрагменты, но для этого снова потребуются больше операций чтения.

Способ кодирования существенно влияет на производительность BigQuery, потому что большое значение имеет хвостовая задержка. Диски постоянно выходят из строя; скорее всего, в BigQuery в любой момент времени есть несколько неисправных дисков, поэтому чтение из таблиц на них потребует выполнить чтение с восстановлением. Обратите внимание, что Colossus использует несколько механизмов, помогающих минимизировать потери производительности, вызванные операциями чтения с восстановлением, таких как кеширование восстановленных данных. Многие запросы в BigQuery предполагают чтение из сотен тысяч файлов и более; то есть практически каждый запрос будет попадать в длинный хвост задержки Colossus. Хорошо, что Colossus — очень быстрая файловая система.

ЧТО ТАКОЕ КОЛОНОЧНЫЙ ФОРМАТ ХРАНЕНИЯ?

Традиционные базы данных хранят данные в виде записей, то есть записи в файле расположены последовательно, друг за другом. Если у вас есть несколько записей и вы хотите сохранить их в файле, такой формат хранения по записям выглядит наиболее очевидным. У таких файлов с записями есть несколько замечательных свойств: если записи имеют фиксированную длину, их можно пропускать, просто добавляя известное смещение. Файлы с записями также удобны, когда записи читаются целиком и по одной.

На рис. 6.23 показано хранилище с записями; записи следуют в файле друг за другом.

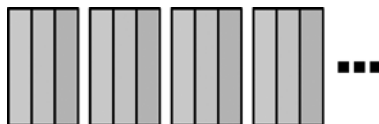


Рис. 6.23. Хранилище с записями, каждая из которых состоит из трех столбцов

Однако большинству запросов не требуется читать записи целиком; часто запросам нужны только некоторые столбцы из таблицы. Если файл хранит данные по записям, вам придется прочитать всю запись, даже если нужен только один столбец. Кроме того, файлы с записями обычно плохо поддаются сжатию; лучший способ уменьшить объем данных, читаемых с диска, — сжать их. Сжатие дости-

гается за счет кодирования повторяющихся данных. Данные в пределах одной записи редко повторяются. Представьте таблицу со столбцами «Имя клиента», «Страна», «Номер телефона» и «Идентификатор клиента». Записи в такой таблице будут содержать очень мало избыточной информации. Номер мало что говорит о стране и еще меньше — об идентификаторе клиента.

А если повернуть таблицу на 90 градусов? То есть хранить данные не в виде *последовательности записей*, а в виде *последовательности столбцов*. Столбец с названиями стран может содержать всего несколько различных значений, и большинство из них могут повторяться. Все идентификаторы клиентов могут начинаться с «0000». И номера телефонов могут иметь общие префиксы.

На рис. 6.24 показано колоночное хранилище; данные в файле хранятся по столбцам.

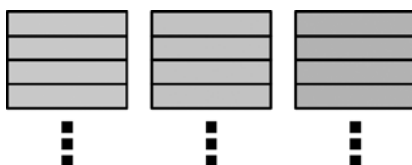


Рис. 6.24. Колоночное хранилище с тремя столбцами

Храня данные по столбцам, вы также получаете преимущество, когда запросу требуется прочитать только несколько столбцов; если в таблице 100 столбцов, а запрос читает только три, ему потребуется прочитать только 3% данных. Большинство запросов читает небольшую часть столбцов в таблице, поэтому одно-временное чтение по столбцам может значительно повысить производительность.

Одной из причин, по которой колоночные форматы хранения не использовались до появления распределенных файловых систем, является особенность физического размещения данных на диске. Если запросу нужно прочитать два столбца, ему придется последовательно просмотреть эти столбцы. Для чтения структурированных данных диск должен прочитать первые несколько значений в столбце А, а затем найти место, где хранится столбец В, прочитать первые несколько значений из него и затем вернуться к столбцу А, чтобы прочитать следующие несколько значений. Поиск местоположения столбца в файле обходится довольно дорого, и он мешает работе распространенных алгоритмов упреждающего чтения, используемых в дисковом оборудовании и операционных системах.

Специфические особенности распределенных файловых систем помогают преодолеть этот недостаток; как правило, чтение выполняется большими полосами и есть возможность читать несколько полос параллельно, потому что они хранятся на разных дисках. Например, если запросу нужно прочитать столбцы А и G, будет запущена процедура чтения столбца А на одном диске и одновременно чтение столбца G на другом.

Формат хранения: Saracitor

Формат хранения данных так же важен, как и способ хранения физических байтов. В BigQuery решили создать свой формат колоночного хранения — Saracitor.

Saracitor — это формат второго поколения, используемый в BigQuery; первым был базовый формат колоночного хранения. При разработке формата Saracitor был учтен опыт, накопленный за восемь лет работы механизма распределенных запросов по эксабайтам данных. Формат развивался параллельно с механизмом запросов, что способствовало еще большему увеличению производительности — глубокое понимание формата учитывалось в ходе развития механизма запросов, и наоборот. Как результат, Saracitor обладает возможностями и механизмами оптимизации, реализованными на основе десятилетнего опыта эксплуатации крупномасштабных систем анализа в Google.

Parquet и Optimized Row Columnar (ORC) — еще два популярных колоночных формата хранения с открытым исходным кодом; однако в BigQuery решили не использовать их. Причина не в синдроме неприятия чужой разработки; когда был создан формат Saracitor, Parquet еще находился в зачаточном состоянии, а ORC имел очень низкую популярность. Кроме того, в формате хранения Saracitor имеется ряд оптимизаций, связывающих его с движком запросов Dremel в BigQuery, что позволяет быстро развивать его и добавлять новые возможности.

Одной из ключевых особенностей формата Saracitor является *словарное кодирование* (dictionary encoding): для полей, имеющих относительно небольшую мощность (несколько уникальных значений), он сохраняет словарь в заголовке файла. Например, предположим, что таблица содержит песни, воспроизводимые музыкальным автоматом, и таких песен относительно немного. Словарь в таком файле мог бы выглядеть, как показано на рис. 6.25.

Словарь	
0	Hey Jude
1	Michelle
2	Here Comes the Sun

Рис. 6.25. Словарное кодирование в Saracitor

Вместо полных названий Saracitor может хранить в словаре только смещения, что делает его намного компактнее. Это будет выглядеть, как показано на рис. 6.26, где первый столбец является закодированным названием песни, а второй столбец — другим полем данных (возможно, именем клиента, запросившего воспроизведение песни).

Словарное кодирование дает еще одно преимущество, которое используется при фильтрации. Предположим, что вы ищете записи с названиями песен,

содержащими слово «Sun». Это относительно дорогостоящий фильтр, потому что он должен искать совпадения в любом месте в строке.

Данные

0	xc*
1	c8!
1	8ec
0	7h!
2	a7c
1	c-%

Рис. 6.26. Два столбца в Capacitor, первый из которых имеет формат словарного кодирования

Обычно такой фильтр выполняет поиск в каждой записи в таблице, пытаясь найти значения, соответствующие предикату. В Capacitor, однако, можно просто проверить соответствие элементов словаря предикату и создать таблицу истинности с результатами, как показано на рис. 6.27.

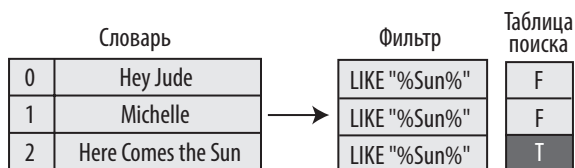


Рис. 6.27. Словарное кодирование увеличивает эффективность работы фильтра

Таблица поиска — это массив со значениями истинности предиката. Сканируя все записи, мы можем просто записать индексы в таблицу соответствия. Например, если закодированное значение было равно «1», мы бы взяли значение со смещением 1 из таблицы поиска и увидели, что предикат оценивается как ложный; с другой стороны, в смещении 2 мы обнаружим истинное значение и сохраним эту запись.

Для еще большей экономии места Capacitor выполняет *кодирование с учетом длины серии* (run-length encoding). То есть если значение «2» появляется пять раз подряд, вместо последовательности «2,2,2,2,2» можно сохранить «2:5». Для длинных последовательностей одинаковых значений это может дать существенную экономию.

Но что, если записи упорядочены так, что в них отсутствуют длинные последовательности одинаковых значений? Для решения этой проблемы в Capacitor используется хитрый трюк — он просто переупорядочивает записи, чтобы по-

лучить возможность закодировать их более компактно. Записи в BigQuery не упорядочены, и нет никакой гарантии, что какие-то записи будут идти после каких-то других записей (рис. 6.28).

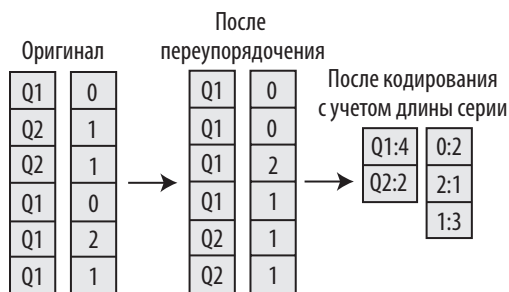


Рис. 6.28. BigQuery переупорядочивает записи, чтобы добиться более компактной формы хранения

Вычисление наилучшего порядка — NP-полная задача,¹ поэтому BigQuery применяет набор эвристик, которые обеспечивают хорошее уплотнение, но выполняются за короткое время.

Эти примеры — всего лишь две оптимизации в Capacitor, которые увеличивают производительность BigQuery.

Метаданные

Метаданные — это данные с информацией о хранимых данных. Поэтому они называются «мета» данными. К метаданным относятся схемы, размеры полей, статистические характеристики и физическое местоположение данных.

Эффективное управление метаданными почти так же важно, как и эффективное управление самими физическими данными. Фактически многие ограничения в BigQuery, например количество таблиц, на которые можно сослаться в запросе, или количество полей, которые может иметь таблица, обусловлены ограничениями в системе метаданных.

¹ NP-задачами в информатике называют задачи, правильность решения которых можно эффективно проверить, но найти правильное решение очень сложно. Вычисление наилучшего порядка — это NP-задача, потому что легко можно проверить, отсортирован список или нет, — достаточно просмотреть список только один раз, — но поиск наилучшего порядка относится к классу задач, которые называются NP-сложными. Если вы обнаружите быстрый алгоритм решения любой одной NP-сложной задачи, это будет означать, что существуют быстрые алгоритмы решения всех NP-сложных задач, потому что они в некотором смысле эквивалентны. На практике NP-сложные задачи обрабатываются с помощью эвристик, поскольку поиск единственно правильного решения был бы слишком неэффективным.

Таблица с метаданными в BigQuery имеет три слоя, только два из которых непосредственно видны пользователю. Внешний слой — это набор данных, включающий таблицы, модели, подпрограммы и т. д., с единым набором прав доступа (подробнее об этом в главе 10). Следующий слой — это таблица, содержащая схему и статистику ключей. Внутренний слой — это набор хранилищ с данными о физическом местоположении данных. Понятие наборов хранилищ недоступно пользователю, и информация о них от него скрыта.

Наборы хранилищ

Набор хранилищ — это атомарная единица данных, созданная в ответ на задание загрузки, потоковое извлечение или запрос на языке управления данными (Data Manipulation Language, DML). Наборы хранилищ позволяют вносить изменения в таблицы BigQuery в полном соответствии с принципами ACID; то есть эти изменения являются атомарными (Atomic, они происходят одновременно или вообще не происходят), согласованными (Consistent, после фиксации становятся доступными везде), идемпотентными (Idempotent, избавляют от необходимости беспокоиться о множественных фиксациях в случае ошибок или нарушений в работе сети) и долговременными (Durable, после фиксации изменения не будут потеряны).

Базовое физическое хранилище для BigQuery является неизменяемым. После закрытия файла его нельзя изменить. Наборы хранилищ также неизменяемы; после фиксации они никогда не изменяются снова. На рис. 6.29 показано, как будет выглядеть таблица с тремя наборами хранилищ.

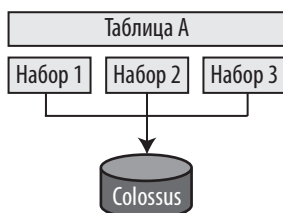


Рис. 6.29. Таблица с тремя наборами хранилищ

Свой жизненный цикл наборы хранилищ обычно начинают в состоянии **PENDING** (ожидание), затем переходят в состояние **COMMITTED** (зафиксировано) и, наконец, в состояние **GARBAGE** (мусор). В состоянии **PENDING** в набор хранилищ активно записываются данные, и данные в этом наборе недоступны пользователю ни в каком виде. После записи данных набор хранилищ переходит в состояние **COMMITTED**, что делает данные в нем доступными для запросов. Когда набор хранилищ больше не нужен, он помечается как **GARBAGE**, то есть его можно утилизировать по истечении некоторого периода.

Возврат к предыдущей версии

На момент написания этой книги BigQuery поддерживала возможность возврата к устаревшим данным, хранившимся в течение срока до семи дней, то есть возможность прочесть состояние таблицы в любой точке этого временного окна. Эта опция может пригодиться, если вы случайно удалили что-то или хотите иметь возможность выполнить повторяющийся запрос к изменяющейся таблице.

Для этого BigQuery запоминает моменты времени, когда происходили изменения в наборах хранилищ. Если вернуться к моменту времени, предшествующему моменту фиксации, набор хранилищ будет исключен из рассмотрения. Если набор хранилищ отмечен как **GARBAGE** и вы возвращаетесь во время, когда он был зафиксирован, он будет восстановлен в контексте этого запроса.

На рис. 6.30 показана таблица с тремя наборами хранилищ, зафиксированными в моменты времени T1, T2 и T3 соответственно. Если вы захотите посмотреть, какой была таблица во время между T2 и T3, вам понадобятся только первые два набора хранилищ.

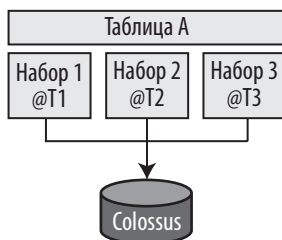


Рис. 6.30. Таблица с тремя наборами хранилищ, зафиксированными в моменты времени T1, T2 и T3

Оптимизация хранения

Со временем, когда вы записываете или изменяете данные, хранилище может стать фрагментированным. Например, предположим, что вы загружаете 100 Кбайт данных каждые две минуты. Каждые из этих 100 Кбайт получают свой набор хранилищ и свой файл. Через месяц у вас будет 2 Тбайт данных, что не очень много, но 21 000 файлов и наборов хранилищ могут ухудшить неэффективность запросов, потому что BigQuery будет тратить много времени на открытие файлов и чтение метаданных о наборах хранилищ.

Оптимизатор хранилища помогает упорядочить данные и придать им оптимальную для запросов форму. Это обеспечивается путем периодической перезаписи файлов. Сначала файлы могут записываться в формате, обеспечивающем быструю запись (хранилище, оптимизированное для записи), а затем в формате, обеспечивающем быстрое чтение (хранилище, оптимизированное для чтения). Можно сказать, что система хранения поддерживает понятие *поколений*, когда

данные записываются в несколько поколений, каждое из которых чем старше, тем более оптимизировано.

На рис. 6.31 показана таблица с данными первого поколения, которые оптимизируются и переписываются в поколение 1.

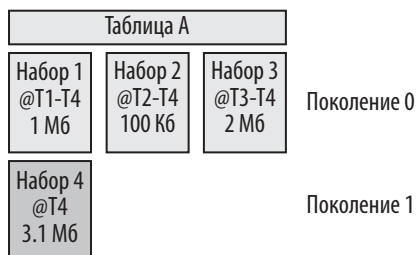


Рис. 6.31. Данные из поколения 0 оптимизируются и переписываются в поколение 1

Оптимизированный набор хранилищ (набор 4) содержит те же данные из наборов 1, 2 и 3. Когда произойдет фиксация набора 4, первые три набора хранилищ будут отмечены как подлежащие утилизации, но они не будут удалены немедленно. Этот факт важен для путешествий во времени, потому что пользователям может потребоваться прочитать таблицу, соответствующую моментам времени, когда существовали только более ранние наборы хранилищ, поэтому BigQuery должна хранить эти метаданные.

Секционирование

Поддержка секционирования в BigQuery позволяет разбивать большие логические таблицы на более мелкие и выполнять запросы только к нужным частям. Представьте, что вам требуется проанализировать только данные, полученные после 3 мая 2019 года; если таблица секционирована по дате, запрос может прочитать только нужные данные.

Секция, с точки зрения внутренней структуры, — это фактически облегченная таблица. Данные для одной секции хранятся физически отдельно от других секций, и для каждой секции имеется полный набор метаданных. Во многих случаях это позволяет рассматривать секции как таблицы. Например, можно обратиться к конечной точке REST API и применить метод `tables.delete()` к секции, чтобы удалить ее. Если настроить продолжительность хранения секции, можно организовать автоматическое удаление секций, если секционирование выполнено по дате, через определенный промежуток, как если бы они были таблицами.

Преимущество секционирования перед использованием нескольких таблиц — в более высокой эффективности запросов. Например, запрашивая диапазон дат,

можно использовать обычный фильтр по столбцу с датой и сканировать только те данные, которые соответствуют фильтру. Кроме того, секционированные таблицы проще в управлении, чем множество логических таблиц, и обеспечивают более эффективный доступ. Фильтры секций часто можно перенести на уровень базы данных с метаданными (Spanner), чтобы сэкономить на чтении не только ненужных данных, но и ненужных метаданных.

Секционирование задумывалось для полей с низкой мощностью (то есть с небольшим числом уникальных значений), обычно меньше нескольких тысяч. При злоупотреблении секционированием таблиц есть риск создать большой объем метаданных. Это не повредит в случаях, когда в результате фильтрации для сканирования остается небольшое количество секций, но если вам когда-нибудь понадобится просканировать всю таблицу, такая операция окажется очень неэффективной, потому что для этого придется прочитать все метаданные. Если таблица имеет большую мощность, используйте кластеризацию. Один из способов уменьшения числа секций заключается в создании более крупных секций, например ежемесячных. Это можно сделать, создав отдельное поле, усекающее дату события до уровня месяца, и выполнив секционирование по этому полю.

Для представления секций в метаданных BigQuery использует наборы хранилищ, отмеченные идентификатором секции. Это упрощает фильтрацию на основе секций и позволяет читать только секции, соответствующие определенной дате. BigQuery может применить фильтр на уровне метаданных, даже не открывая физические данные. Наборы хранилищ также содержат информацию о размере поля, которая позволяет при пробном запуске определить, сколько данных потребуется просканировать, без фактического выполнения запроса.

На рис. 6.32 показана таблица с тремя наборами хранилищ. Каждый набор в этом случае представляет отдельную секцию.

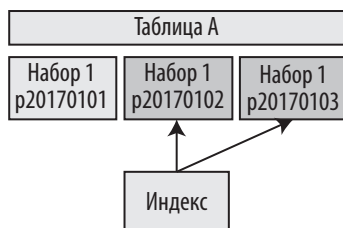


Рис. 6.32. Три набора хранилищ, представляющих три разные секции

Допустим, что секционирование выполнено по столбцу `eventDate` и выполняется запрос, включающий фильтр `WHERE eventDate >= '20170102'`. Этому фильтру соответствуют только два набора хранилищ: `20170102` и `20170103`. В базе данных Spanner существует индекс (см. рис. 6.32), который помогает найти наборы, соответствующие этому диапазону. То есть согласно условию

в запросе можно считать, что таблица содержит только эти два набора хранилищ, а значит, для его выполнения придется сканировать меньше данных и скорость будет выше.

Кластеризация

Кластеризация — это способ хранения данных в полусортированном формате по ключу, сконструированному из столбцов в данных. В файлы данных записываются непересекающиеся диапазоны пространства ключей. Это помогает увеличить эффективность поиска и сканирования диапазонов, потому что механизм запросов должен будет открывать только файлы, соответствующие ключу, и сможет выполнять двоичный поиск, чтобы найти файлы, содержащие начало и конец диапазона.

На рис. 6.33 показано, как можно использовать кластеризацию.

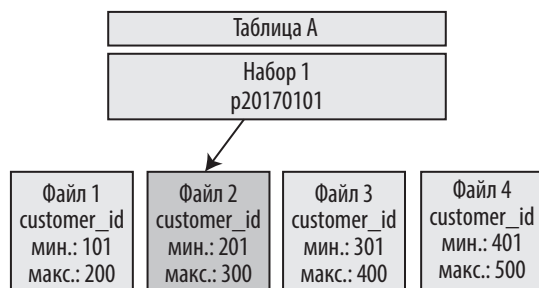


Рис. 6.33. Набор хранилищ в случае кластеризации по столбцу `customer_id`

Таблица кластеризована по столбцу `customer_id`, а данные отсортированы так, что файлы хранят непересекающиеся диапазоны значений в столбце `customer_id`. Файл 1 содержит идентификаторы клиентов (`customer_id`) от 101 по 200, файл 2 содержит идентификаторы от 201 по 300, и т. д.

Предположим, что выполняется запрос `SELECT ... WHERE customer_id = 275`. Мы знаем, что файлы расположены в порядке возрастания значений `customer_id`, и можем просто просмотреть заголовки файлов, чтобы понять, что `customer_id` 275 находится в файле 2. Мы можем выполнить быстрый двоичный поиск, чтобы найти начальный файл, и после этого нам не нужно будет просматривать другие файлы, потому что весь искомый диапазон находится в одном файле. Поскольку таблица была кластеризована, достаточно прочитать только один файл, а не все.

Файлы с данными имеют заголовки, содержащие минимальные и максимальные значения всех полей. Практически все колоночные хранилища поддерживают такую информацию (включая хранилища с открытым исходным кодом, такие как

Parquet и ORC). Это дает дополнительное преимущество, позволяя проверить присутствие искомого значения в таблице простым анализом заголовков. Более того, заголовки файлов кешируются, поэтому часто необходимые файлы можно определить, вообще не выполняя дисковых операций ввода/вывода.

Когда данные сортируются по файлам, каждый файл имеет узкий диапазон ключей. Например, файл 1 хранит значения **aa-ac**, файл 2 хранит значения **ad-ag**, файл 3 хранит значения **ah-ap**, и т. д. Если вам понадобится найти значение **ae**, вы можете открыть заголовок файла 1, увидеть, что искомого значения там нет, затем открыть заголовок файла 2 и обнаружить, что искомое значение попадает в диапазон значений, хранящихся в этом файле. После этого можно использовать только этот файл и не заглядывать в другие. Это означает, что механизму запросов придется просканировать только один файл, что намного быстрее и дешевле, чем сканирование всех файлов.

Обратите внимание, что внутри файла данные не сортируются, сортировка производится только по файлам. Данные в файлах *Сараситор* (подробнее об этом чуть ниже) переупорядочиваются для увеличения коэффициента сжатия, а сортировка данных замедляет чтение. Кроме того, сортировка данных внутри файла мало чем может помочь, потому что вам придется прочитать с диска то же самое количество блоков.

Повторная кластеризация. Одна из сложностей, связанных с кластеризацией, связана с ее поддержанием при изменении данных. Представьте, например, что в предыдущем примере кто-то добавит две записи с ключами **ab** и **ao**. Эти ключи соответствуют диапазонам в файлах 1 и 3. Но файлы не могут изменяться, и вы, скорее всего, не захотите создавать файлы, содержащие по одной записи. BigQuery запишет эти новые записи в один файл в новом наборе хранилищ. Теперь, при обращении к таблице, придется проверить два набора хранилищ и выполнять лишнюю работу. Если данные постоянно меняются со временем, появится много наборов хранилищ, которые могут иметь перекрывающиеся диапазоны. В какой-то момент вам придется просканировать все файлы, чтобы найти единственное значение.

Эта проблема фрагментирования данных решается методом повторной кластеризации. Иногда BigQuery будет выполнять повторную кластеризацию таблиц в фоновом режиме. BigQuery поддерживает *коэффициент кластеризации*, который описывает долю полностью кластеризованных данных. Если эта доля падает ниже определенного предела, производится перезапись данных в отсортированном формате. При этом будет создан новый набор хранилищ, чтобы обеспечить возможность возврата к предыдущим состояниям таблицы. Повторная кластеризация происходит автоматически, без вмешательства пользователя и с использованием «системных ресурсов», а не ресурсов, за которые пользователи должны платить.

Диаграмма на рис. 6.34 иллюстрирует выполнение повторной кластеризации.

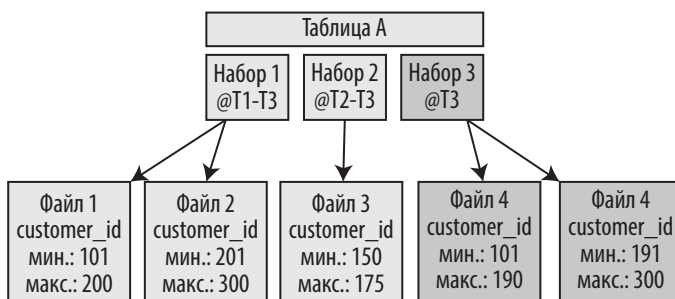


Рис. 6.34. Повторная кластеризация создает новый набор хранилищ

Здесь есть два набора хранилищ с перекрывающимися диапазонами значений ключа; первый набор включает два файла с диапазоном значений ключа от 101 до 300. Второй набор, зафиксированный позднее (T2), имеет меньший диапазон ключей, от 150 до 175, но он перекрывается с диапазонами в файлах из набора 1.

В момент T3 наступает событие повторной кластеризации. В результате создается копия всех данных из файлов 1, 2 и 3, а затем выполняется их повторная кластеризация. В ходе этого процесса хранилища, созданные ранее, отмечаются как «мусор», а в наборе 3 фиксируется новая копия, соответствующая моменту T3. После этого момента все запросы будут видеть хорошо кластеризованные данные в наборе 3.

СЕКЦИОНИРОВАНИЕ И КЛАСТЕРИЗАЦИЯ

Секционирование можно рассматривать как деление таблицы на множество подтаблиц на основе данных в определенном столбце. Кластеризация, напротив, больше похожа на сортировку таблиц по определенному набору столбцов. Разница может показаться незначительной, но кластеризация работает лучше при множестве уникальных значений. Например, если у вас миллион клиентов и вы часто выполняете запросы, чтобы отыскать одного из них, кластеризация по идентификатору клиента обеспечит высокую эффективность такого поиска. Если выполнить секционирование по столбцу `customer_id`, поиск тоже будет выполняться быстро, но из-за большого объема метаданных, необходимого для отслеживания всех секций, запросы, выбирающие всех пользователей, будут выполняться медленнее.

Секционирование часто используется в сочетании с кластеризацией; есть возможность выполнить секционирование по столбцу с низкой мощностью (например, по дате события) и кластеризовать по столбцу с большим количеством уникальных значений (например, по идентификатору клиента). Это позволит работать со срезом диапазона дат в таблице, как если бы она была отдельной таблицей, а также находить записи с конкретными клиентами без сканирования всех данных в секции.

Оптимизация производительности путем кластеризации таблиц. Кластеризация также обеспечивает ряд оптимизаций, применяемых к запросам. Например, одна из оптимизаций перемещает ограничения с одной стороны соединения в другую. Предположим, что вы выполняете следующий запрос:

```
SELECT orders.order_id
FROM retail.orders AS orders JOIN retail.customers
ON orders.customer_id = customers.customer_id
WHERE customers.customer_name = 'Jordan Tigani'
```

Он найдет все заказы клиента по имени Джордан Тайджани (Jordan Tigani). Предположим, что таблица заказов кластеризована по столбцу `customer_id`. В простейшем случае можно было бы применить фильтр к таблице `customers` и разослать оставшиеся записи всем рабочим серверам сегментов, а затем просканировать всю таблицу заказов (`orders`), чтобы найти заказы, имеющие соответствующий идентификатор клиента. Однако поскольку таблица заказов кластеризована по полю `customer_id`, достаточно просмотреть только файлы, в которых может храниться искомый `customer_id`, а значит, нет необходимости сканировать таблицу целиком. Это снижает стоимость и значительно повышает производительность.

Следует также отметить, что преимущества кластеризации проявляются не только при фильтрации по столбцу, использованному для кластеризации; они также проявляются при фильтрации по столбцам, значения которых коррелируют со значениями в столбце кластеризации. Представьте, что у вас есть таблица заказов, кластеризованная по столбцу `order_id`, и заказы следуют почти по порядку. Если выполнить запрос, фильтрующий сделки по узкому диапазону дат, эти даты окажутся в небольшом количестве файлов; BigQuery придется просканировать только файлы с указанными датами сделок, даже если столбец `order_id` не указан в фильтре. Эта оптимизация не только увеличивает производительность, но также снижает стоимость запроса при работе в режиме до востребования. Вообще все, что BigQuery делает для уменьшения сканируемого объема данных, ведет к снижению затрат при работе с кластерными таблицами.

DML

DML — это набор специальных операторов SQL, позволяющих изменять таблицы (см. главу 8). Всего существует четыре таких оператора: `INSERT`, `DELETE`, `UPDATE` и `MERGE`.

`INSERT` добавляет записи в таблицу. Это очень простая операция, потому что в своей основе она подобна загрузке дополнительных данных в таблицу. Когда выполняется операция `INSERT`, файлы, представляющие новые данные, записываются в Colossus, и в метаданные добавляется новый набор хранилищ. Новый набор хранилищ имеет время фиксации, совпадающее со временем загрузки данных.

`DELETE` удаляет записи из таблицы. Это более сложная операция, потому что файлы и метаданные (наборы хранилищ) в BigQuery не изменяются. Пред-

положим, что вы решили удалить одну запись (`DELETE ... WHERE customer_id = 1234`) и эта запись хранится в файле C, в наборе хранилищ 3.

Поскольку файлы в BigQuery не изменяются, чтобы удалить одну запись, нельзя просто удалить строку из середины файла C. Вместо этого BigQuery создаст копию файла без записи (назовем его C2), затем отметит старый набор хранилищ меткой **GARBAGE**, потому что он больше не используется. Новому файлу потребуются новый набор хранилищ, набор 4. Однако этого недостаточно, потому что новый набор хранилищ должен содержать все остальное, что имелось в наборе 3. Новый набор хранилищ будет ссылаться на старые файлы из набора 3, кроме файла C, из которого была удалена запись.

На рис. 6.35 показано, как это происходит.

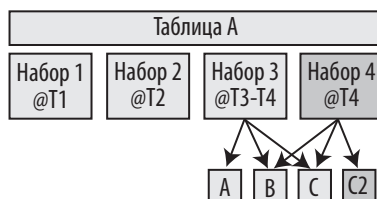


Рис. 6.35. Операция удаления включает создание нового набора хранилищ, ссылающегося на комбинацию файлов из старого набора и новых файлов, без удаленных записей

Если вам показалось, что такая простая операция, как удаление, требует выполнить слишком много действий, то знайте, что вам не показалось. Именно поэтому лучше объединять обновления в пакеты и применять их одновременно.

UPDATE обычно реализуется как атомарная комбинация операций **INSERT** и **DELETE**. То есть вместо фактического изменения данных происходит удаление старых записей и добавление новых. Это позволяет основной массе файлов оставаться неизменными и применять изменения к оперативным записям или файлам. **MERGE**, по сути, является всего лишь очень необычной формой оператора **UPDATE**. Этот оператор позволяет одновременно выполнять чтение, изменение и запись. С точки зрения архитектуры операции **UPDATE** и **MERGE** очень похожи на операцию **DELETE**; **UPDATE** просто записывает новые значения в существующие записи. Аналогично **MERGE** записывает дополнительные данные.

Метафайл

Метафайл — это один из механизмов, помогающих BigQuery добиться высокой производительности за счет исключения ненужных секций. Как мы уже упоминали в этой главе, этот файл хранит метаданные обо всех файлах данных, составляющих таблицу. Он содержит минимальные и максимальные значения

для всех полей, а также местоположение каждого файла данных. Этот файл имеет тот же формат, что и файлы данных в BigQuery, а значит, его можно запрашивать как любой другой файл. Это важно, потому что, выбирая файлы для сканирования, обработчику запросов не нужно просматривать заголовки всех файлов. Достаточно прочитать этот файл и затем выбрать только нужные файлы.

На рис. 6.36 показан пример метафайла.

Столбец	Мин.	Макс.	Файл
Столбец 1	0	25	Файл 0001
Столбец 2	aaa	ccc	Файл 0001
Столбец 1	26	50	Файл 0002
Столбец 2	bbb	fff	Файл 0002
Столбец 1	50	99	Файл 0003
Столбец 1	eee	zzz	Файл 0003

Рис. 6.36. Метафайл с метаданными обо всех файлах данных, составляющих таблицу

Если в запросе используется предикат `WHERE field1 = 30`, можно сначала отправить запрос в этот файл, а затем отправить обратно результат с единственным файлом, который нам нужен, в данном случае `file0002`. Имея эту информацию, мы можем не открывать любые другие файлы, чтобы выполнить запрос.

Выводы

В этой главе мы познакомились с внутренней архитектурой BigQuery, чтобы разобраться и понять процесс выполнения запросов.

Начав с общего обзора архитектуры, мы проследили путь запроса от момента получения сервером GFE и передачи соответствующему серверу заданий BigQuery до его обработки сервером запросов Dremel. Мы исследовали этапы выполнения различных типов запросов: от простейшего запроса сканирования-фильтрации-подсчета до более сложных запросов сканирования-фильтрация-агрегирования, которые имеют большую мощность и могут потребовать переупорядочения. Мы также рассмотрели два способа реализации соединений: всенаправленные соединения для небольших таблиц и соединения хешированием для более крупных — и выяснили, как по плану запроса узнать, какой механизм соединения используется.

Мы также обсудили способ хранения данных, преимущества колоночного формата BigQuery, словарного кодирования и то, как использование наборов хранилищ делает возможным возврат к более старой версии. Наконец, мы рассмотрели, как реализованы секционирование и кластеризация и почему они делают производительность запросов лучше.

Оптимизация производительности и затрат

Настройка производительности BigQuery обычно выполняется с целью уменьшить время выполнения запроса или его стоимость либо и то и другое. В этой главе мы рассмотрим ряд оптимизаций производительности, которые могут пригодиться вам на практике.

Принципы производительности

Дональд Кнут (Donald Knuth), легендарный ученый в области информатики, сделал важное наблюдение о том, что преждевременная оптимизация является корнем всех зол. Впрочем, полное высказывание Кнута выглядит более сбалансированным:¹

Мы *должны* забыть об оптимизации, скажем, в 97% случаев: преждевременная оптимизация — корень всех зол. Напротив, все свое внимание мы должны уделить оставшимся 3%. Хороший программист не успокоится таким рассуждением, он внимательно исследует критически важный код, но только после того, как идентифицирует его.

Вслед за Кнутом мы хотели бы предостеречь: настройку производительности следует выполнять только в конце разработки и только если обнаружится, что типичные запросы выполняются слишком долго. Гораздо лучше иметь таблицы с гибкими схемами и элегантные, удобочитаемые и легко сопровождаемые запросы, чем запутывать макеты таблиц и запросы ради крошечной прибавки производительности. Тем не менее иногда вам действительно будет необходимо повысить производительность запросов, возможно, потому, что они выполняют

¹ Эта цитата взята из его статьи 1974 года «Structured Programming with go to Statements». Получить PDF-файл с текстом статьи можно по адресу <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.103.6084>.

ся так часто, что даже небольшие улучшения будут иметь смысл. Другой аспект, который следует учитывать, — знание компромиссов производительности может помочь вам в выборе альтернативного дизайна.

Ключевые составляющие производительности

В этой главе мы исследуем два аспекта. Сначала мы посмотрим, как измерить производительность запросов, чтобы можно было идентифицировать критически важные части вашей программы, которые действительно необходимо оптимизировать. Затем, опираясь на опыт пользователей и наши знания архитектуры BigQuery, мы выясним, что относится к критическим 3% по Кнуту. Это позволит нам разрабатывать схемы таблиц и запросы с учетом вероятности появления узких мест в производительности и поможет сделать оптимальный выбор на этапе проектирования.

Для оптимизации производительности запросов в BigQuery важно понимать ключевые составляющие производительности запросов, о которых рассказывается во второй части этой главы. Время, затрачиваемое на выполнение запроса, зависит от объема данных, извлекаемых из хранилища, организации этих данных, количества этапов, необходимых для обработки запроса, возможности распараллеливания этих этапов, объема данных, обрабатываемых каждым этапом, и вычислительной дороговизны каждого из этапов.

В целом простой запрос, который читает три столбца, будет выполняться на 50% дольше запроса, читающего только два столбца, потому что запрос с тремя столбцами должен прочитать на 50% больше данных.¹ Запрос, включающий группировку, обычно выполняется медленнее, чем запрос без группировки, потому что операция группировки добавляет дополнительный этап обработки запроса.

Управление затратами

Стоимость запроса зависит от вашего тарифного плана. В BigQuery существует два типа тарифных планов. Первый тип — тарифный план до востребования, когда ваш работодатель платит нам (Google) в зависимости от объема данных, обработанных вашими запросами. Если вы используете единый тарифный план, ваша компания получает определенное количество слотов² (например, 500), и вы можете выполнять сколько угодно запросов без дополнительных затрат.

¹ Различие может быть трудно измерить точно, потому что BigQuery — это сетевая служба, и время, необходимое для ее достижения по сети, и нагрузка на саму службу могут быстро меняться. Вам может потребоваться выполнить запрос много раз, чтобы получить более или менее точную оценку скорости выполнения.

² Слот — это единица вычислительной мощности, выделяемой для обработки запросов SQL. BigQuery автоматически вычисляет, сколько слотов потребуется для каждого запроса. Более подробно об этом рассказывается в главе 6.

В тарифном плане до востребования (на запрос) стоимость запроса пропорциональна объему обработанных данных. Чтобы снизить затраты при использовании тарифного плана до востребования, ваши запросы должны обрабатывать как можно меньше данных. Кроме того, уменьшение количества читаемых данных также увеличивает скорость выполнения запросов. Третья часть этой главы посвящена оптимизации хранения данных и доступа к ним.

Если вы используете резервирование с фиксированной ставкой, чистая стоимость вашего запроса вполне соответствует времени, затраченному на его выполнение.¹ В модели с фиксированной ставкой можно косвенно сократить расходы, если занимать забронированные слоты меньшее время, то есть путем увеличения скорости запросов, как описано во второй части этой главы.

Оценка затрат на запрос

Пользующиеся тарифным планом до востребования могут получить оценку стоимости запроса, прежде чем отправить его в сервис. Веб-интерфейс BigQuery проверяет запросы и предоставляет оценку объема данных, которые будут обработаны. Вы можете увидеть, сколько байтов обработает запрос до его выполнения, кликнув на **Validator Query** (Проверить запрос). Пользователи командной строки `bq` могут передать ключ `--dry_run`, чтобы получить план запроса и оценку объема данных для обработки, прежде чем фактически запустить запрос. Пробные прогоны выполняются бесплатно. Зная, какой объем данных будет обработан запросом, можно использовать калькулятор цен Google Cloud Platform (GCP) (<https://cloud.google.com/products/calculator/>), чтобы оценить стоимость запроса в долларах. Такие инструменты, как BigQuery Mate (<https://oreil.ly/dGMCK>) и superQuery (<https://web.superquery.io>), тоже предлагают оценку стоимости, но могут не иметь доступа к информации о возможных скидках. На момент написания этой книги обработка одного терабайта данных стоила пять долларов США (для объединенных регионов US и EU) после превышения бесплатного уровня в один терабайт в месяц. Обратите внимание, что BigQuery должна читать только столбцы, на которые ссылаются ваши запросы, а секционирование и кластеризация могут еще больше уменьшить объем сканируемых данных (и, соответственно, стоимость).



Для экспериментов с BigQuery можно использовать песочницу BigQuery. На нее распространяются те же ограничения, что и на уровень бесплатного пользования (10 Гбайт активного хранилища и 1 Тбайт обработанных данных в месяц), но доступ к ней можно получить без кредитной карты.²

¹ Операции, не способные выполняться параллельно, могут не увеличивать стоимость, потому что оставшиеся слоты могут быть задействованы другими рабочими нагрузками.

² Дополнительную информацию можно найти по ссылке <https://cloud.google.com/bigquery/docs/sandbox>.

Запуская запрос, можно указать параметр `--maximum_bytes_billed`, чтобы ограничить объем данных, обрабатываемых запросом. Если количество обрабатываемых байтов превысит указанный предел, запрос не будет выполнен без оплаты. Также управлять расходами можно, устанавливая пользовательскую квоту в облачной консоли GCP Cloud Console (<https://console.cloud.google.com/iam-admin/quotas>), чтобы ограничить количество данных, обрабатываемых за день. Этот лимит можно задать на уровне проекта или пользователя.

Поиск наиболее дорогостоящих запросов

Для контроля расходов полезно создать краткий список запросов, на которые следует обратить внимание. Сделать это можно, обратившись к схеме `INFORMATION_SCHEMA`, связанной с проектом, и выбрать самые дорогие запросы:

```
SELECT
  job_id
  , query
  , user_email
  , total_bytes_processed
  , total_slot_ms
FROM `some-project`.INFORMATION_SCHEMA.JOBS_BY_PROJECT
WHERE EXTRACT(YEAR FROM creation_time) = 2019
ORDER BY total_bytes_processed DESC
LIMIT 5
```

Этот запрос вернет пять самых дорогих запросов, выполненных в 2019 году в рамках некоторых проектов, если судить по значению `total_bytes_processed`. Если у вас фиксированный тариф, вы можете получить запросы с наибольшим значением `total_slot_ms`.

Измерение производительности и поиск проблем

Для настройки производительности запроса важно ответить на следующие вопросы, чтобы знать, на чем сосредоточить внимание:

- Какой объем данных читается из хранилища и как эти данные организованы.
- Сколько этапов требуется для выполнения запроса и как распараллелить эти этапы.
- Какой объем данных обрабатывается на каждом этапе и насколько они дорогие с вычислительной точки зрения.

Как видите, в каждом вопросе есть аспект, который можно измерить (например, объем читаемых данных), и аспект, который необходимо понять (например, влияние организации данных на производительность запроса).

В этом разделе мы посмотрим, как измерить производительность запроса, проанализировать журналы BigQuery и исследовать объяснения запросов. Затем, опираясь на наше понимание архитектуры BigQuery (глава 6) и знание характеристик производительности (речь о них пойдет в последующих разделах этой главы), мы сможем увеличить производительность. Далее в этой главе мы представим запросы и их характеристики производительности, не углубляясь в описание подходов к измерениям, которые могут привести к улучшению их производительности.

Определение скорости выполнения запроса с помощью REST API

Поскольку BigQuery имеет REST API, для оценки времени выполнения запросов можно использовать любой инструмент для работы с веб-службами. Если в вашей организации уже применяется один из таких инструментов, используйте его.

Иногда нужно измерить скорость выполнения с сервера, на котором эти инструменты отсутствуют. В таких случаях можно прибегнуть к Unix-утилитам `time` и `curl`. Как говорилось в главе 5, код SQL и JSON можно поместить в переменные Bash:¹

```
read -d '' QUERY_TEXT << EOF
SELECT
    start_station_name
    , AVG(duration) as duration
    , COUNT(duration) as num_trips
FROM `bigquery-public-data`.london_bicycles.cycle_hire
GROUP BY start_station_name
ORDER BY num_trips DESC
LIMIT 5
EOF

read -d '' request << EOF
{
    "useLegacySql": false,
    "useQueryCache": false,
    "query": "\${QUERY_TEXT}"
}
EOF
request=$(echo "$request" | tr '\n' ' ')
```

Обратите внимание на необходимость отключить кэширование запросов, чтобы при каждом запросе на стороне сервера выполнялись все необходимые операции.

¹ См. сценарий `07_perf/time_query.sh` в репозитории GitHub с примерами для этой книги (<https://github.com/GoogleCloudPlatform/bigquery-oreilly-book>).

В главе 5 также говорилось о том, как использовать инструмент командной строки `gcloud` для получения токена доступа и идентификатора проекта, необходимых для обращения к REST API:

```
access_token=$(gcloud auth application-default print-access-token)
PROJECT=$(gcloud config get-value project)
```

Теперь, когда все готово, можно выполнить запрос несколько раз и получить общее время, чтобы затем рассчитать среднюю производительность запроса и уменьшить влияние на оценку случайных сетевых сбоев:

```
NUM_TIMES=10
time for i in $(seq 1 $NUM_TIMES); do
echo -en "\r ... $i / $NUM_TIMES ..."
curl --silent \
  -H "Authorization: Bearer $access_token" \
  -H "Content-Type: application/json" \
  -X POST \
  -d "$request" \
  "https://www.googleapis.com/bigquery/v2/projects/$PROJECT/queries" > /dev/null
done
```

Мы получили следующие результаты:

```
Real    0m16.875s
User    0m0.265s
Sys     0m0.109s
```

Общее время выполнения запроса 10 раз составило 16.875 секунды, то есть в среднем на выполнение одного запроса уходило 1.7 секунды. Обратите внимание, что сюда входит время обращения к серверу и время, потраченное на выборку результатов, то есть это не чистое время обработки запроса.

Оценить время, затраченное на прием-передачу, можно, включив кеширование запросов:

```
read -d '' request << EOF
{
  "useLegacySql": false,
  "useQueryCache": true,
  "query": "${QUERY_TEXT}"
}
EOF
```

Повторив эксперимент, мы получили следующие результаты:

```
Real    0m6.760s
User    0m0.264s
Sys     0m0.114s
```

Поскольку теперь запросы кешируются, полученные результаты отражают почти чистое время, обусловленное задержками в сети. Полученные значения

показывают, что фактическое время обработки запроса составляет (16.875 – 6.760) / 10, или около 1 секунды.

Определение скорости выполнения запроса с помощью BigQuery Workload Tester

Использование инструментов измерения производительности веб-служб или низкоуровневых инструментов Unix возможно и желательно в обычных системах, тем не менее мы рекомендуем использовать BigQuery Workload Tester (<https://github.com/GoogleCloudPlatform/pontem/tree/dev/BigQueryWorkloadTester>) для измерения скорости запросов BigQuery в вашей среде разработки. В отличие от обычных инструментов измерения производительности веб-служб, Workload Tester может вычислять время, затрачиваемое на передачу данных по сети туда и обратно (управлять которым практически невозможно), и сообщать только время обработки запроса (которое можно и нужно оптимизировать) без многократного его выполнения. Этот инструмент может измерять время, затрачиваемое на отдельные запросы и рабочие нагрузки (последовательности запросов), и может вызывать запросы параллельно, если вы захотите проверить параллелизм.

Для Workload Tester требуется Gradle (<https://gradle.org/>), инструмент с открытым исходным кодом, предназначенный для сборки проектов. То есть перед установкой Workload Tester нужно установить Gradle. Cloud Shell предлагает возможность быстро опробовать Workload Tester. В ней и других системах Linux на основе Debian можно установить Gradle с помощью следующей команды:

```
sudo apt-get -y install gradle
```

В macOS то же самое можно сделать командой

```
brew install gradle
```

Описание порядка установки Gradle в других операционных системах можно найти в документации (<https://gradle.org/install/>).

Затем следует скопировать репозиторий GitHub с исходными текстами Workload Tester и собрать инструмент:¹

```
git clone https://github.com/GoogleCloudPlatform/pontem.git
cd pontem/BigQueryWorkloadTester
gradle clean :BigQueryWorkloadTester:build
```

Давайте измерим скорость запроса, определяющего среднюю продолжительность велосипедных поездок в Лондоне. По аналогии с примером в предыдущем

¹ См. сценарий 07_perf/install_workload_tester.sh в репозитории GitHub с примерами для этой книги.

разделе, мы могли бы просто сохранить текст запроса в переменной Bash, но было бы полезнее хранить эталонные запросы в другом месте, поэтому запишем текст запроса в файл:¹

```
cat <<EOF| tr '\n' ' ' > queries/busystations.sql
SELECT
  start_station_name
  , AVG(duration) as duration
  , COUNT(duration) as num_trips
FROM `bigquery-public-data`.london_bicycles.cycle_hire
GROUP BY start_station_name
ORDER BY num_trips DESC
LIMIT 5
EOF
```

Затем создадим конфигурационный файл для каждой рабочей нагрузки, состоящей из набора запросов в файлах:

```
cat <<EOF>./config.yaml
concurrencyLevel: 1
isRatioBasedBenchmark: true
benchmarkRatios: [1.0, 2.0]
outputFileFolder: $OUTDIR
workloads:
- name: "Busy stations"
  projectId: $PROJECT
  queryFiles:
  - queries/busystations.sql
  outputFileName: busystations.json
EOF
```

В этой конфигурации мы устанавливаем базовый уровень конкурентности равным 1, то есть запросы будут выполняться строго последовательно. Мы также указываем набор коэффициентов для измерения времени выполнения запросов на уровнях конкурентности 1.0 и 2.0 по отношению к базовому (то есть 1 и 2 конкурентно выполняемых запросов). Чтобы протестировать уровни конкурентности 1, 2, 5, 10, 15 и 20, используйте следующие настройки:

```
concurrencyLevel: 10
isRatioBasedBenchmark: true
benchmarkRatios: [0.1, 0.25, 0.5, 1.0, 1.5, 2.0]
```

И наконец, запустим инструмент измерения:

```
gradle clean :BigQueryWorkloadTester:run
```

Результатом является файл, содержащий общее истекшее время (включающее время на передачу данных о сети туда и обратно) и фактическое время обработки каждого из запросов. В нашем случае первый запрос (с уровнем

¹ См. сценарий 07_perf/time_bqwt.sh в репозитории GitHub с примерами для этой книги.

конкурентности 1) обрабатывался 1111 миллисекунд (1.111 секунды), а второй и третий запросы (выполнявшиеся одновременно из-за выбранного нами уровня параллелизма 2) обрабатывались 1.108 секунды и 1.026 секунды. Иначе говоря, BigQuery обеспечивает почти одинаковую производительность, независимо от количества одновременно обрабатываемых запросов.

Выявление проблем в рабочих нагрузках с помощью Stackdriver

Помимо измерения скорости отдельных запросов, иногда полезно оценить производительность всего рабочего процесса, прибегнув к помощи журналов BigQuery. Это можно сделать в веб-консоли GCP, в разделе **Stackdriver Logging** (Stackdriver. Журналы). Например, можно посмотреть предупреждения (и более серьезные ошибки), касающиеся запросов и операций над набором данных, которые мы выполняли в главе 5 с набором данных `ch05`, выбрав пункт **Warning** (Предупреждения) в раскрывающемся меню, как показано на рис. 7.1.

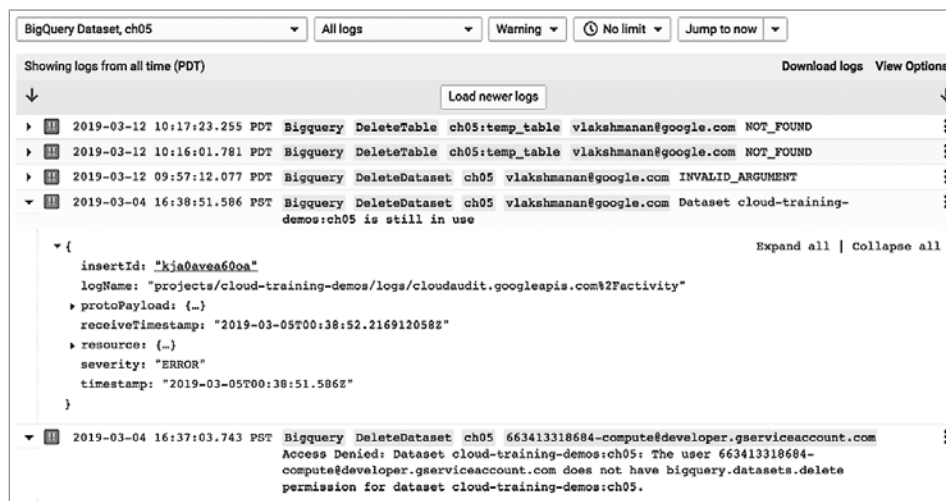


Рис. 7.1. Обзор сообщений в Stackdriver, которые отправила BigQuery

Эта возможность просмотра всех сообщений об ошибках, возникших в проекте, в одном месте может очень пригодиться, особенно если запросы отправляются сценариями и панелями мониторинга, которые не выводят сообщения об ошибках, возвращаемых BigQuery. Фактически можно (при наличии необходимых разрешений) просмотреть журналы и получить набор операций, выполняемых рабочей нагрузкой, и по этому набору определить, выполняются ли ненужные операции. Посмотрите на список операций с набором данных `ch05eu`, показанный на рис. 7.2, и прочитайте его снизу вверх.

2019-03-12 09:59:11.079 PDT	Bigquery	DeleteDataset	ch05eu	vlakshmanan@google
2019-03-12 09:58:44.440 PDT	Bigquery	DeleteTable	ch05eu:cycle_stations_copy	
2019-03-12 09:58:36.189 PDT	Bigquery	DeleteTable	ch05eu:bad_bikes	vlakshmana
2019-03-04 18:05:04.403 PST	Bigquery	InsertJob	ch05eu:bad_bikes	vlakshmanan@g
2019-03-03 21:55:54.661 PST	Bigquery	DeleteDataset	ch05eu	663413318684-compu
2019-03-03 16:41:05.594 PST	Bigquery	InsertJob	ch05eu:cycle_stations_copy	vl
2019-03-03 16:26:49.334 PST	Bigquery	InsertDataset	ch05eu	vlakshmanan@google

Рис. 7.2. С помощью Stackdriver можно получить набор операций, выполняемых рабочей нагрузкой

В этом случае видно, что был создан набор данных с именем ch05eu и в него была добавлена таблица с именем cycle_stations_copy. Затем была предпринята неудачная попытка удалить набор данных ch05eu, потому что набор данных не был пустым. Затем в него была добавлена новая таблица с именем bad_bikes. После этого обе таблицы, bad_bikes и cycle_stations_copy, были удалены. Наконец, был удален и сам набор данных.

Мы можем изучить детали каждого задания. Например, давайте рассмотрим первое задание, создавшее таблицу cycle_stations_copy. Здесь можно увидеть схему созданной таблицы, как показано на рис. 7.3.

2019-03-03 16:41:05.594 PST	Bigquery	InsertJob	ch05eu:cycle_stations_copy
<pre> { insertId: "hrcug1e2g4g5" logName: "projects/cloud-training-demos/logs/cloudaudit.googleapis.com%2Fa protoPayload: { @type: "type.googleapis.com/google.cloud.audit.AuditLog" authenticationInfo: {...} authorizationInfo: [1] metadata: { @type: "type.googleapis.com/google.cloud.audit.BigQueryAuditMetadata" tableCreation: { jobName: "projects/cloud-training-demos/jobs/3051c867-ad0c-413d-a6de reason: "JOB" table: { schemaJson: "{ \"fields\": [{ \"name\": \"id\", \"type\": \"INTEGER\", \"mode\": \"NULLABLE\" }, { \"name\": \"install_date\", </pre>			

Рис. 7.3. Исследование деталей задания с целью выяснить схему созданной таблицы

Учитывая эту информацию и знание контекста, получается, что таблица `cycle_stations_copy` не использовала никаких полей из `bad_bikes`. Возможно, весь набор операций с `bad_bikes` был ненужным и их можно убрать из рабочего процесса.

Чтение плана запроса

Кроме измерения скорости запросов и изучения журналов, выявлять проблемы с производительностью можно, просматривая информацию о плане запроса. В плане запроса перечисляются этапы обработки запроса и предоставляется информация о данных, обрабатываемых на каждом шаге каждого этапа. Информация о плане запроса доступна в формате JSON в информации о задании, а также отображается в веб-интерфейсе BigQuery.

В BigQuery граф выполнения оператора SQL разбивается на этапы, каждый из которых состоит из единиц работы, которые выполняются параллельно рабочими серверами. Взаимодействия между этапами осуществляются посредством распределенной архитектуры переупорядочения (см. главу 6), поэтому большинство этапов начинается с чтения выходных данных предыдущих этапов и заканчивается передачей результатов на вход следующих этапов. Имейте в виду, что следующий этап может начаться до полного окончания предыдущего этапа — этапы могут начинать обработку уже имеющихся данных. То есть этапы не обязательно выполняются последовательно.



Помните, что план запроса может меняться динамически: точный объем данных и вычислительные затраты на промежуточных этапах неизвестны до начала выполнения рассматриваемого этапа. Если фактический объем данных сильно отличается от ожидаемого, могут быть добавлены новые этапы с целью перераспределения данных между рабочими серверами. Из-за динамического характера плана запроса информацию о нем следует изучать после завершения запроса.

Получение плана запроса из информации о задании

Информация о каждом этапе завершеного запроса в информации о задании включает время начала и конца каждого этапа выполненного запроса, общее количество прочитанных и записанных записей, а также количество байтов, записанных всеми рабочими серверами в ходе обработки запроса. Например, попробуйте выполнить следующий запрос:

```
SELECT
  start_station_name,
  AVG(duration) AS duration,
  COUNT(duration) AS num_trips
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire
GROUP BY
```

```

    start_station_name
ORDER BY
    num_trips DESC
LIMIT
5

```

Получить информацию о соответствующем задании можно обращением к REST API:¹

```

JOBID=8adb3fd-e310-44bb-9c6e-88254958ccac # CHANGE
access_token=$(gcloud auth application-default print-access-token)
PROJECT=$(gcloud config get-value project)
curl --silent \
    -H "Authorization: Bearer $access_token" \
    -X GET \
    "https://www.googleapis.com/bigquery/v2/projects/$PROJECT/jobs/$JOBID"

```

Например, в сведениях о первом этапе приводится следующая информация о времени, затраченном на ожидание данных, отношении операций ввода/вывода к вычислительным операциям, а также об объеме прочитанных, переупорядоченных и записанных данных:

```

"waitRatioAvg": 0.058558558558558557,
"readRatioAvg": 0.070270270270270274,
"computeRatioAvg": 0.86036036036036034
...
"shuffleOutputBytes": "356596",
"shuffleOutputBytesSpilled": "0",
"recordsRead": "24369201",
"recordsWritten": "6138",
"parallelInputs": "7",

```

Поскольку это этап ввода, нас больше всего интересует производительность операций чтения и переупорядочения, потому что на выполнение этого этапа ушло 86% процессорного времени. Так как данные не нужно было вытеснять на диск, можно сделать вывод, что узкие места (если они есть) в этом запросе не связаны с вводом/выводом. Если запрос выполняется слишком медленно, причины нужно искать в другом месте, возможно, одна из них связана с относительно небольшим уровнем параллелизма (7). Однако в данном конкретном случае дело не в этом, потому что входных данных относительно немного — всего 24 миллиона записей — и разбиения такого объема на семь фрагментов вполне достаточно.

¹ Этот код находится в сценарии `07_perf/get_job_details.sh` в репозитории GitHub с примерами для этой книги. Получить требуемый идентификатор задания можно в разделе Query history (История запросов) в веб-интерфейсе BigQuery. Если ваш запрос был выполнен за пределами США и ЕС, вам также потребуется указать местоположение задания в объекте ресурса (неудобно, но что поделаешь!). Также подробную информацию о задании можно получить командой `bq ls -j`.


Визуализация информации о плане запроса


Большую часть информации о плане запроса из сведений о задании можно получить в визуальном представлении в веб-интерфейсе BigQuery. Откройте запрос и перейдите на вкладку **Execution details** (Сведения о выполнении), чтобы увидеть описание плана запроса. Оно включает общее время, а также время выполнения всех шагов, составляющих этапы выполнения запроса.



Вкладка **Execution details** (Сведения о выполнении) содержит важную информацию о производительности запросов (рис. 7.4).

Query results

 SAVE RESULTS

 EXPLORE IN DATA STUDIO


Query complete (1.1 sec elapsed, 862.1 MB processed)

Job information

Results

JSON

Execution details

 For help debugging or optimizing your query, check our documentation. Learn more

Elapsed time

1.1 sec

Slot time consumed

4.048 sec

Bytes shuffled

348.49 KB

Bytes spilled to

0 B

Рис. 7.4. Ключевые показатели на вкладке со сведениями о выполнении запроса; целью оптимизации производительности, как правило, является сокращение времени слотов и/или объема переупорядочиваемых байтов

График общего распределения времени на рис. 7.5 показывает, что семь параллельных входов действовали в первые 0.6 секунды, а затем были добавлены еще два входа.

В веб-интерфейсе также доступна подробная информация о шагах и этапах, как показано на рис. 7.6.

На рис. 7.6 видно, что запрос обрабатывался в три этапа. Первый этап (S00) — это этап ввода, второй этап (S01) выполнил сортировку, а третий (S02) осуществил вывод. На рис. 7.6 также визуальное показано распределение времени на каждом этапе. Соотношения, доступные в численном виде в ответе JSON, полученном на запрос сведений о задании, изображены в виде цветных полосок, как показано на рис. 7.7.¹

¹ Если в книге вы видите черно-белые иллюстрации, попробуйте выполнить запрос и посмотрите, как выглядит информация о запросе в веб-интерфейсе BigQuery.

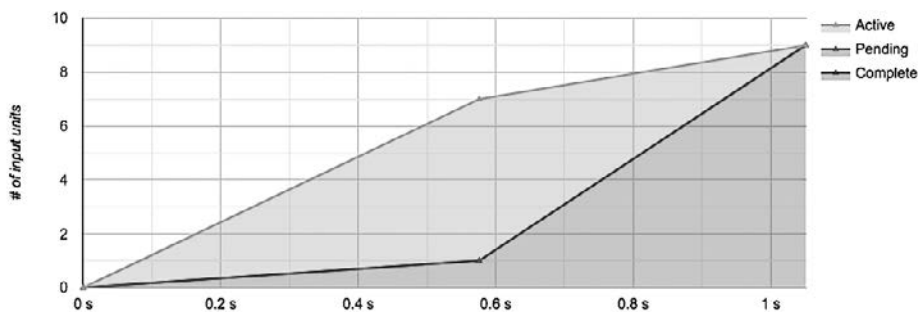


Рис. 7.5. График распределения времени на основе информации из плана запроса

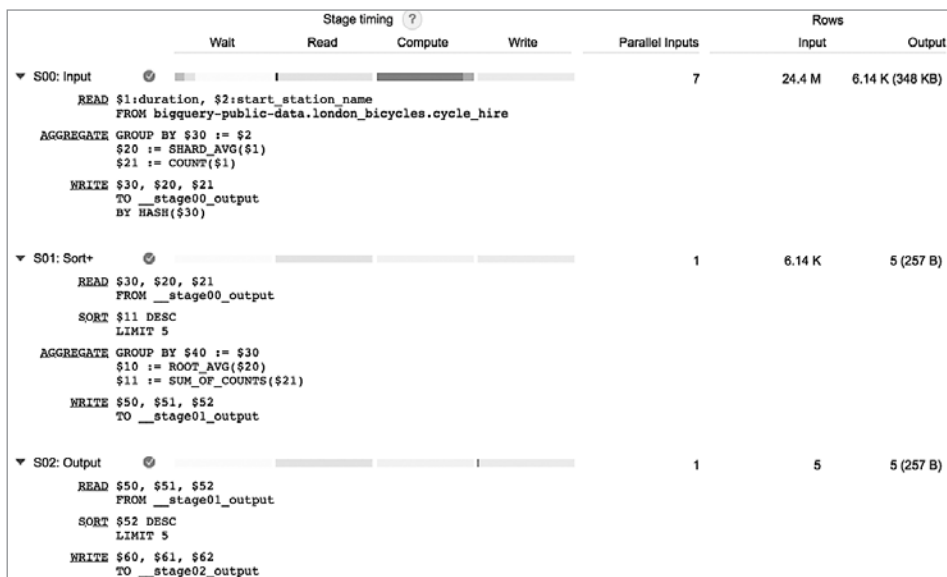


Рис. 7.6. Информация о шагах и этапах из плана запроса



Рис. 7.7. Распределение времени на каждом этапе

Чем темнее цвет, тем больше времени было потрачено на ожидание, чтение, вычисления или запись. В данном случае большая часть времени тратится на вычисления и совсем немного уходит на ожидание. Этап ожидания весьма изменчив (промежуточный цвет показывает максимальное ожидание, а разницу

между средним и максимальным можно рассматривать как показатель изменчивости) — здесь максимальное время ожидания почти вдвое превышает среднее время ожидания. Этап чтения, с другой стороны, сохраняет стабильность, а расходы на запись вообще незначительны.

Рассмотрев этап ввода, показанный на рис. 7.8, можно заметить, что он состоит из трех шагов: чтения двух столбцов (*duration* и *start_station_name*, обозначенные ссылками \$1 и \$2) из таблицы BigQuery, агрегирования и записи выходных данных.



Рис. 7.8. Шаги, составляющие первый этап

Шаг агрегирования состоит из трех операций: группировки по *start_station_name* (напомним, что этому входному столбцу соответствует ссылка \$2), поиска средней продолжительности (\$1) в каждом сегменте и подсчета записей с ненулевой продолжительностью поездки. Этап записи передает группы (\$30), среднюю продолжительность (\$20) и количество поездок (\$21) в промежуточный вывод, выполняя распределение по хешу названия пункта проката. Если на следующем этапе потребуется задействовать несколько рабочих серверов, хеш названия пункта проката определит, какие части вывода *stage_00* и каким рабочим серверам передать для обработки.

Снова вернемся к рис. 7.6. Здесь можно заметить, что первый этап выполняется параллельно семью рабочими серверами, а остальные два этапа — одним. Этап ввода читает 24.4 миллиона записей и выводит примерно 6140 записей, что составляет 348 Кбайт. Эти записи сортируются на втором этапе, и пять из них передаются на третий этап. Обнаружив этап с единственным рабочим сервером, мы должны гарантировать, что нагрузка на память этого сервера не превысит его возможностей (как это сделать, мы покажем далее в этой главе), — 348 Кбайт это немного, поэтому данный запрос не должен создавать проблем с производительностью из-за ограниченности ресурсов, доступных одному рабочему серверу.

Другой вариант визуализации плана запроса в BigQuery — использовать BigQuery Visualizer (https://oreil.ly/Unw_c), доступный по адресу <https://bqvisualiser.appspot.com/> (рис. 7.9).

BigQuery Visualizer особенно удобен при исследовании сложных запросов с десятками этапов, которые бывает трудно понять, просматривая лишь краткую сводку в веб-интерфейсе BigQuery.

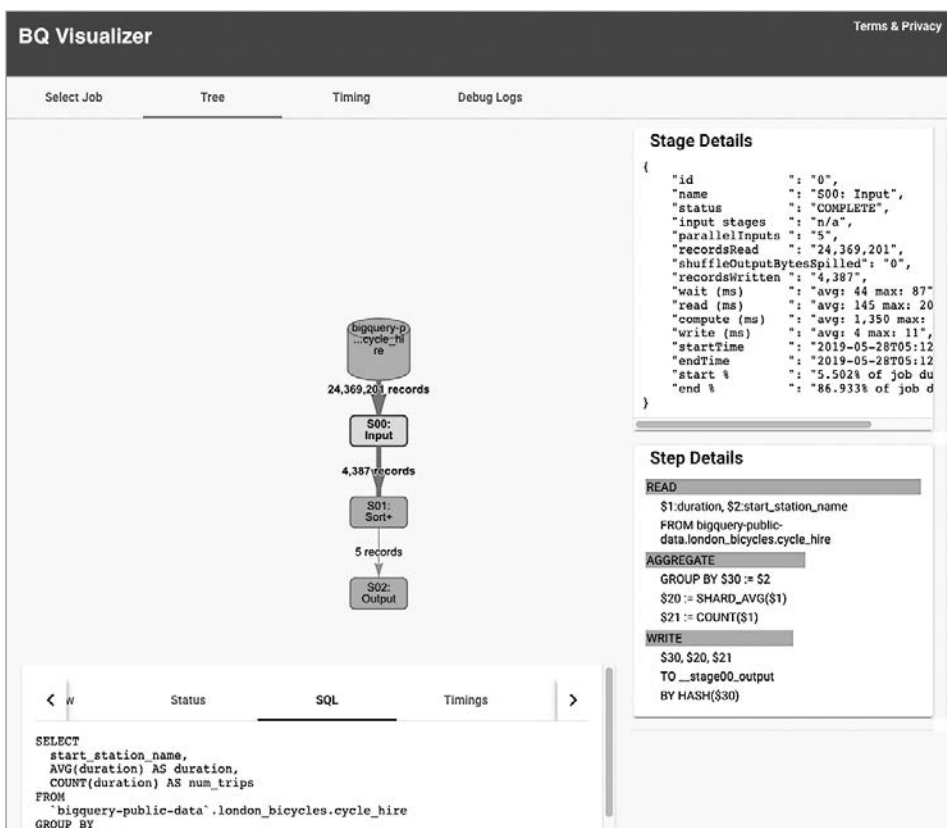


Рис. 7.9. Визуализация заданий с помощью BigQuery Visualizer

Увеличение скорости выполнения запросов

Мы уже говорили в предыдущем разделе о том, что для оценки скорости запросов и выявления потенциальных проблем вы должны выполнить следующие шаги:

1. Определить общее время выполнения рабочей нагрузки с помощью BigQuery Workload Tester.
2. Изучить журналы, чтобы убедиться, что рабочая нагрузка не выполняет никаких непредвиденных операций.
3. Изучить план запросов, составляющих рабочую нагрузку, чтобы выявить узкие места или ненужные этапы.

Когда вы определите, что проблема существует и в рабочем процессе нет явных ошибок, можно начинать думать, как повысить скорость выполнения критических частей рабочей нагрузки. В этом разделе мы предлагаем несколько возможных способов повышения производительности запросов, включая:

- минимизацию ввода/вывода;
- кеширование результатов предыдущих запросов;
- эффективное выполнение соединений;
- исключение перегрузки рабочих серверов;
- использование приближенных функций агрегирования.

Минимизация ввода/вывода

Как отмечалось выше, запрос, вычисляющий сумму трех столбцов, будет выполняться медленнее запроса, вычисляющего сумму двух столбцов, причем основная разница в производительности будет обусловлена чтением большего количества данных, а не дополнительным сложением. То есть запрос, вычисляющий среднее значение столбца, будет выполняться так же быстро, как запрос, вычисляющий дисперсию данных (даже при том, что для вычисления дисперсии требуется, чтобы BigQuery вычислила сумму и сумму квадратов), потому что большая часть накладных расходов в простых запросах связана с вводом/выводом, а не с вычислениями.

Ограничения числа столбцов в SELECT

BigQuery использует колоночные форматы файлов, поэтому чем меньше столбцов перечислено в операторе SELECT, тем меньше данных ему придется прочитать. Например, оператор SELECT * читает все столбцы во всех записях в таблице, что делает его довольно медленным и затратным. Исключения составляют случаи, когда SELECT * используется в подзапросе и ссылается лишь на несколько полей, извлекаемых внешним запросом; оптимизатор BigQuery сумеет понять, что прочитать нужно только необходимые столбцы.

Явно перечислите столбцы, которые должны присутствовать в конечном результате. Например, следующий запрос найдет bike_id, соответствующий самой долгой поездке в наборе данных:

```
SELECT
  bike_id
  , duration
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire
ORDER BY duration DESC
LIMIT 1
```

И сделает это гораздо эффективнее, чем запрос:

```
SELECT
  *
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire
ORDER BY duration DESC
LIMIT 1
```

Первый запрос выполнялся 1.8 секунды и прочитал 372 Мбайт данных, а второй — 5.5 секунды (в три раза дольше) и прочитал 2.59 Гбайт (в семь раз больше).



Если таблица не кластеризована, применение предложения `LIMIT` не влияет на объем читаемых данных. Кластеризация таблиц позволяет уменьшить количество читаемых байтов (хотя величину экономии трудно спрогнозировать). Для ознакомления с содержимым таблицы вместо запроса `SELECT *` с предложением `LIMIT` используйте кнопку предварительного просмотра в веб-интерфейсе. Плата за предварительный просмотр не взимается, тогда как запрос `SELECT *` будет стоить тех же денег, что и сканирование таблицы.

Если вам нужны почти все столбцы в таблице, используйте запросы вида `SELECT * EXCEPT`, чтобы не читать лишние данные (см. главу 2).

Уменьшение объема читаемых данных

Настройку запросов важно начинать с исследования читаемых данных, чтобы определить, можно ли уменьшить их объем. Предположим, вы решили найти типичную продолжительность самой обычной поездки в один конец. Это можно сделать, например, так:

```
SELECT
  MIN(start_station_name) AS start_station_name
  , MIN(end_station_name) AS end_station_name
  , APPROX_QUANTILES(duration, 10)[OFFSET(5)] AS typical_duration
  , COUNT(duration) AS num_trips
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire
WHERE
  start_station_id != end_station_id
GROUP BY
  start_station_id, end_station_id
ORDER BY num_trips DESC
LIMIT 10
```

Этот запрос выполнялся 14.7 секунды и вернул следующие результаты:

Row	start_station_name	end_station_name	typical_duration	num_trips
1	Black Lion Gate, Kensington Gardens	Hyde Park Corner, Hyde Park	1500	12 000
2	Black Lion Gate, Kensington Gardens	Palace Gate, Kensington Gardens	780	11 833
3	Hyde Park Corner, Hyde Park	Albert Gate, Hyde Park	1920	11 745
4	Hyde Park Corner, Hyde Park	Triangle Car Park, Hyde Park	1380	10 923
5	Hyde Park Corner, Hyde Park	Black Lion Gate, Kensington Gardens	1680	10 652

В информации о выполнении запроса видно, что сортировка (для вычисления приблизительных квантилей для каждой пары пунктов проката) требует переупорядочения выходных данных на этапе ввода, но основное время расходуется на вычисления, как можно видеть на рис. 7.10.

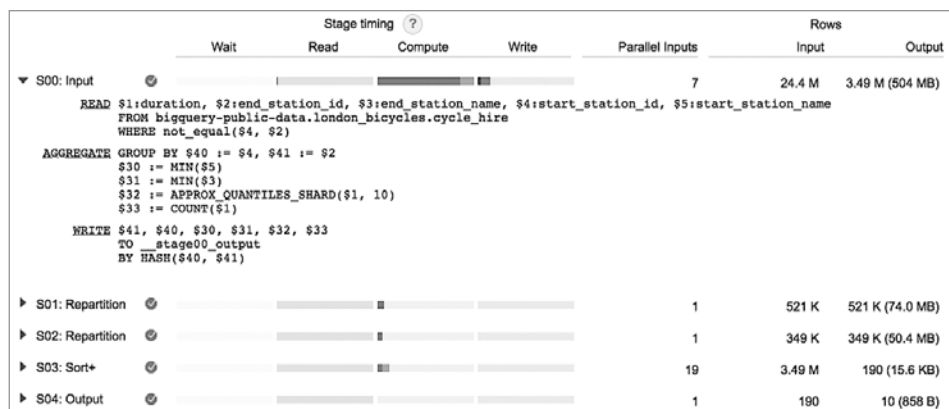


Рис. 7.10. Этот запрос требует выполнения двух этапов переупорядочения, но основное время тратится на вычисления

Тем не менее мы можем уменьшить издержки на ввод/вывод, если будем выполнять фильтрацию и группировку по названиям пунктов проката, а не по их идентификаторам, потому что при этом потребуется прочитать меньше столбцов:

```

SELECT
  start_station_name
  , end_station_name
  , APPROX_QUANTILES(duration, 10)[OFFSET(5)] AS typical_duration
  , COUNT(duration) AS num_trips
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire
WHERE
  start_station_name != end_station_name
GROUP BY
  start_station_name, end_station_name
ORDER BY num_trips DESC
LIMIT 10

```

Этот запрос исключает из чтения два столбца с идентификаторами и выполняется за 9.6 секунды, то есть на 30% быстрее. Это увеличение обусловлено накапливающимся эффектом чтения меньшего количества данных: запросу требуется на одно переупорядочение меньше и меньшее количество рабочих серверов (10 вместо 19) для сортировки, как показано на рис. 7.11.

Этот запрос возвращает тот же результат, потому что между именем и идентификатором пункта проката существует однозначное соответствие.

		Stage timing ?				Rows	
		Wait	Read	Compute	Write	Parallel Inputs	Input Output
▼ S00: Input	⊗					7	24.4 M 2.85 M (380 MB)
	<pre> READ \$1:duration, \$2:end_station_name, \$3:start_station_name FROM bigquery-public-data.london_bicycles.cycle_hire WHERE not_equal(\$3, \$2) AGGREGATE GROUP BY \$40 := \$3, \$41 := \$2 \$30 := APPROX_QUANTILES_SHARD(\$1, 10) \$31 := COUNT(\$1) WRITE \$41, \$40, \$30, \$31 TO _stage00_output BY HASH(\$40, \$41) </pre>						
▶ S01: Repartition	⊗					1	509 K 509 K (70.0 MB)
▶ S02: Sort+	⊗					10	2.85 M 100 (8.36 KB)
▶ S03: Output	⊗					1	100 10 (858 B)

Рис. 7.11. Используя преимущество однозначного соответствия между `station_id` и `station_name`, можно прочитать меньше столбцов, убрать один этап и использовать меньше рабочих серверов для сортировки

Уменьшение объема затратных вычислений

Допустим, вам нужно определить, какое общее расстояние преодолел каждый велосипед в наборе данных. Самый простой способ — найти все поездки с участием каждого велосипеда и суммировать пройденные расстояния:

```

WITH trip_distance AS (
  SELECT
    bike_id
    , ST_Distance(ST_GeogPoint(s.longitude, s.latitude),
                  ST_GeogPoint(e.longitude, e.latitude)) AS distance
  FROM
    `bigquery-public-data`.london_bicycles.cycle_hire,
    `bigquery-public-data`.london_bicycles.cycle_stations s,
    `bigquery-public-data`.london_bicycles.cycle_stations e
  WHERE
    start_station_id = s.id
    AND end_station_id = e.id
)

SELECT
  bike_id
  , SUM(distance)/1000 AS total_distance
FROM trip_distance
GROUP BY bike_id
ORDER BY total_distance DESC
LIMIT 5

```

У нас этот запрос выполнялся 7.1 секунды (44 секунды суммарного времени слотов) и перераспределит 1.69 Мбайт. Как выяснилось, некоторые велосипеды проехали почти 6000 километров:

Row	bike_id	total_distance
1	12925	5990.988493972133
2	12757	5919.736998793672
3	12496	5883.1268196056335
4	12841	5870.757769474104
5	13071	5853.763514457338

Вычисление расстояния — довольно затратная операция, а кроме того, есть возможность отказаться от соединения таблиц `cycle_stations` и `cycle_hire`, если предварительно вычислить расстояния между всеми парами пунктов проката:

```
WITH stations AS (
  SELECT
    s.id AS start_id
    , e.id AS end_id
    , ST_Distance(ST_GeogPoint(s.longitude, s.latitude),
                  ST_GeogPoint(e.longitude, e.latitude)) AS distance
  FROM
    `bigquery-public-data`.london_bicycles.cycle_stations s,
    `bigquery-public-data`.london_bicycles.cycle_stations e
),
```

В остальном усовершенствованный запрос не изменился, кроме того, что он выполняет соединение с предварительно собранной таблицей расстояний:

```
trip_distance AS (
  SELECT
    bike_id
    , distance
  FROM
    `bigquery-public-data`.london_bicycles.cycle_hire,
    stations
  WHERE
    start_station_id = start_id
    AND end_station_id = end_id
)

SELECT
  bike_id
  , SUM(distance)/1000 AS total_distance
FROM trip_distance
GROUP BY bike_id
ORDER BY total_distance DESC
LIMIT 5
```

Теперь суммарное время слотов составляет 31.5 секунды (меньше на 30%), несмотря на увеличение объема данных (33.44 Мбайт), перераспределяемых между узлами.

Кеширование результатов предыдущих запросов

Сервис BigQuery автоматически кеширует результаты запросов во временные таблицы. Если в течение 24 часов будет отправлен идентичный запрос, результаты для него будут извлечены из этой временной таблицы. Кешированные результаты возвращаются очень быстро, и плата за них не взимается.

Однако есть несколько важных аспектов, о которых следует помнить. Кеширование запросов основано на точном сравнении строк. Поэтому даже появление лишнего пробела в тексте запроса может привести к промаху кеша. Запросы никогда не кешируются, если демонстрируют недетерминированное поведение (например, используют `CURRENT_TIMESTAMP` или `RAND`), если запрашиваемая таблица или представление изменились (даже если столбцы/записи, представляющие интерес, не изменились), если таблица связана с потоковым буфером (даже если в нем не появилось новых записей), если в запросе используются операторы языка манипулирования данными (Data Manipulation Language, DML) или если запрашиваются внешние источники данных.



Мы не рекомендуем читать данные непосредственно из временных таблиц кеша, потому что срок хранения таких таблиц может истечь. Если результаты запроса могут служить исходными данными для других запросов, используйте обычные таблицы или материализованные представления, как описано в следующем разделе.

Кеширование промежуточных результатов

Общую производительность можно улучшить за счет увеличения объема ввода/вывода, используя преимущества временных таблиц и материализованных представлений. Допустим, у вас есть ряд запросов, которые начинаются с поиска типичной продолжительности поездок между парами пунктов проката:

```
WITH typical_trip AS (  
  SELECT  
    start_station_name  
    , end_station_name  
    , APPROX_QUANTILES(duration, 10)[OFFSET(5)] AS typical_duration  
    , COUNT(duration) AS num_trips  
  FROM  
    `bigquery-public-data`.london_bicycles.cycle_hire  
  GROUP BY  
    start_station_name, end_station_name  
)
```


ВСЕГДА ЛИ МАТЕРИАЛИЗОВАННЫЕ ТАБЛИЦЫ ЭФФЕКТИВНЕЕ?

Материализованные таблицы не всегда уменьшают стоимость вычислений. Например, вы решили использовать предложение `WITH`, чтобы абстрагировать дорогостоящую функцию, использующую регулярное выражение, но основное выражение всегда имеет ограничительный фильтр (в данном случае для поездок, продолжающихся дольше 84 000 секунд):

```
WITH trip AS (  
  SELECT  
    REGEXP_REPLACE(start_station_name,  
      r"^# ([a-zA-Z0-9\s]+)$", "FROM: \\1") AS start_station_name  
    , REGEXP_REPLACE(end_station_name,  
      r"^# ([a-zA-Z0-9\s]+)$", "TO: \\1") AS end_station_name  
    , duration  
  FROM  
    `bigquery-public-data`.london_bicycles.cycle_hire  
)  
  
SELECT * FROM trip  
WHERE duration > 84000
```

Оптимизатор BigQuery, просмотрев запрос, мог бы ограничить количество обращений к регулярному выражению. Однако если предложение `WITH` материализовать в таблицу, функция, использующая регулярное выражение, будет вызвана для каждой записи.

Предложение `WITH` (также называемое табличным выражением) улучшает читаемость, но не скорость и стоимость запроса, потому что результаты не кешируются. То же относится к представлениям и подзапросам. Если вы часто используете предложение `WITH`, представления или подзапросы, сохраняйте результаты в таблице (или материализованном представлении) — это поможет увеличить производительность:¹

```
CREATE OR REPLACE TABLE ch07eu.typical_trip AS  
SELECT  
  start_station_name  
  , end_station_name  
  , APPROX_QUANTILES(duration, 10)[OFFSET(5)] AS typical_duration  
  , COUNT(duration) AS num_trips  
FROM  
  `bigquery-public-data`.london_bicycles.cycle_hire  
GROUP BY  
  start_station_name, end_station_name
```

¹ Конечно, вы должны будете измерить это увеличение. В некоторых случаях дополнительные издержки, связанные с чтением промежуточных результатов из таблицы, наоборот, делают запрос дороже по сравнению с простым пересчетом результатов с использованием предложения `WITH`.

Используем предложение **WITH**, чтобы определить, в какие дни поездки на велосипедах длятся дольше обычного:

```
SELECT
  EXTRACT (DATE FROM start_date) AS trip_date
  , APPROX_QUANTILES(duration / typical_duration, 10)[OFFSET(5)] AS ratio
  , COUNT(*) AS num_trips_on_day
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire AS hire
JOIN typical_trip AS trip
ON
  hire.start_station_name = trip.start_station_name
  AND hire.end_station_name = trip.end_station_name
  AND num_trips > 10
GROUP BY trip_date
HAVING num_trips_on_day > 10
ORDER BY ratio DESC
LIMIT 10
```

Этот запрос выполнялся 19.1 секунды и обработал 1.68 Гбайт. Теперь задействуем таблицу с промежуточными результатами:

```
SELECT
  EXTRACT (DATE FROM start_date) AS trip_date
  , APPROX_QUANTILES(duration / typical_duration, 10)[OFFSET(5)] AS ratio
  , COUNT(*) AS num_trips_on_day
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire AS hire
JOIN ch07eu.typical_trip AS trip
ON
  hire.start_station_name = trip.start_station_name
  AND hire.end_station_name = trip.end_station_name
  AND num_trips > 10
GROUP BY trip_date
HAVING num_trips_on_day > 10
ORDER BY ratio DESC
LIMIT 10
```

Этот запрос выполнялся 10.3 секунды (меньше на 50% из-за исключения затратных вычислений) и обработал 1.72 Гбайт (небольшой прирост из-за необходимости читать данные из новой таблицы). Оба вопроса вернули один и тот же результат — поездки в Рождество длятся дольше обычного:

Row	trip_date	ratio	num_trips_on_day
1	2016-12-25	1.6	34477
2	2015-12-25	1.5263157894736843	20871
3	2015-08-01	1.25	41200
4	2016-07-30	1.2272727272727273	43524
5	2015-08-02	1.2222222222222223	41243

Таблица `ch07eu.typical_trip` не обновляется при добавлении новых данных в таблицу `cycle_hire`. Один из способов решить эту проблему устаревших данных заключается в использовании материализованного представления или периодическом выполнении обновляющих запросов по расписанию. Вы должны определить стоимость таких обновлений, чтобы увидеть, компенсирует ли увеличение производительности запросов дополнительные затраты на поддержание промежуточной таблицы в актуальном состоянии.

Ускорение выполнения запросов с помощью BI Engine

Если есть таблицы, к которым вы часто обращаетесь из инструментов бизнес-аналитики (Business Intelligence, BI), такие как информационные панели с агрегатами и фильтрами, для ускорения запросов можно воспользоваться движком BI Engine. Он автоматически сохраняет соответствующие фрагменты данных в памяти (фактические столбцы таблицы или полученные результаты) и использует специализированный процессор запросов, настроенный для работы с данными, большей частью расположенными в памяти. С помощью консоли администратора BigQuery Admin Console можно зарезервировать объем памяти (до 10 Гбайт), который BigQuery должна использовать для своего кеша, как показано на рис. 7.12.

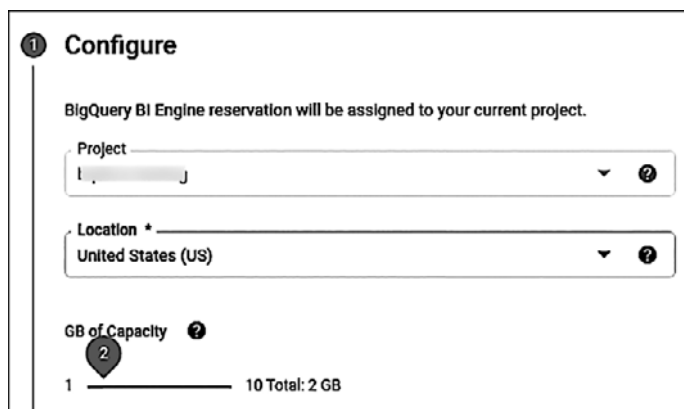


Рис. 7.12. Резервирование памяти для кеширования данных из таблиц в настройках BI Engine

Зарезервируйте память в том же регионе, в котором находится используемый набор данных. После этого BigQuery начнет кешировать таблицы, их части и агрегаты в памяти и быстрее возвращать результаты.

В основном BI Engine используется для кеширования таблиц, доступ к которым осуществляется из инструментов панели мониторинга, таких как Google Data Studio. Резервируя память для BI Engine, можно существенно улучшить отклик информационных панелей, основанных на данных в BigQuery.

Эффективное выполнение соединений

Соединение двух таблиц требует координации данных и подвержено ограничениям пропускной способности между слотами. Если есть возможность избежать соединения или уменьшить объем соединяемых данных, сделайте это.

Денормализация

Один из способов повысить производительность чтения и избежать соединений — отказаться от эффективного хранения данных и добавить избыточные копии. Это называется *денормализацией*. То есть вместо хранения широты и долготы пункта прокатов велосипедов отдельно от информации о прокате можно создать денормализованную таблицу:

```
CREATE OR REPLACE TABLE ch07eu.london_bicycles_denorm AS
SELECT
    start_station_id
  , s.latitude AS start_latitude
  , s.longitude AS start_longitude
  , end_station_id
  , e.latitude AS end_latitude
  , e.longitude AS end_longitude
FROM
    `bigquery-public-data`.london_bicycles.cycle_hire as h
JOIN
    `bigquery-public-data`.london_bicycles.cycle_stations as s
ON
    h.start_station_id = s.id
JOIN
    `bigquery-public-data`.london_bicycles.cycle_stations as e
ON
    h.end_station_id = e.id
```

В этом случае всем последующим запросам не придется выполнять соединение, потому что таблица будет содержать всю необходимую информацию о местоположении для всех поездок:

Row	start_station_id	start_latitude	start_longitude	end_station_id	end_latitude	end_longitude
1	439	51.5338	−0.118677	680	51.47768469	−0.170329317
2	597	51.473471	−0.20782	622	51.50748124	−0.205535908
3	187	51.49247977	−0.178433004	187	51.49247977	−0.178433004
4	15	51.51772703	−0.127854211	358	51.516226	−0.124826
5	638	51.46663393	−0.169821175	151	51.51213691	−0.201554966

В этом случае вы меняете вычислительную дороговизну соединения на затраты, связанные с хранением большого объема данных. Вполне возможно, что стоимость чтения большого количества данных с диска перевесит стоимость соединения, поэтому вы должны оценить, останетесь ли вы в плюсе от денормализации за счет увеличения производительности.

Предотвращение самосоединений больших таблиц

Самосоединения происходят, когда таблица соединяется сама с собой. BigQuery поддерживает самосоединения таблиц, однако они могут ухудшить производительность, если таблица, для которой выполняется самосоединение, очень большого размера. Часто самосоединения можно избежать, если использовать такие возможности SQL, как оконные функции и функции агрегирования.

Рассмотрим пример. В BigQuery есть общедоступный набор данных с именами детей, опубликованный Управлением социального обеспечения США. Попробуем найти в этом наборе данных мужские имена, наиболее распространенные в 2015 году в штате Массачусетс:

```
SELECT
  name
  , number AS num_babies
FROM `bigquery-public-data`.usa_names.usa_1910_current
WHERE gender = 'M' AND year = 2015 AND state = 'MA'
ORDER BY num_babies DESC
LIMIT 5
```

Вот результаты этого запроса:

Row	name	num_babies
1	Benjamin	456
2	William	445
3	Noah	403
4	Mason	365
5	James	354

Точно так же найдем женские имена (`gender = 'F'`), наиболее распространенные в 2015 году в штате Массачусетс:

Row	name	num_babies
1	Olivia	430
2	Emma	402
3	Sophia	373
4	Isabella	350
5	Charlotte	344

А теперь представьте, что нам понадобилось выяснить наиболее распространенные имена мальчиков и девочек за все годы. Простейшее решение этой задачи — дважды прочитать исходную таблицу и выполнить самосоединение:

```
WITH male_babies AS (
SELECT
    name
    , number AS num_babies
FROM `bigquery-public-data`.usa_names.usa_1910_current
WHERE gender = 'M'
),
female_babies AS (
SELECT
    name
    , number AS num_babies
FROM `bigquery-public-data`.usa_names.usa_1910_current
WHERE gender = 'F'
),
both_genders AS (
SELECT
    , SUM(m.num_babies) + SUM(f.num_babies) AS num_babies
    , SUM(m.num_babies) / (SUM(m.num_babies) + SUM(f.num_babies)) AS frac_male
FROM male_babies AS m
JOIN female_babies AS f
USING (name)
GROUP BY name
)

SELECT * FROM both_genders
WHERE frac_male BETWEEN 0.3 and 0.7
ORDER BY num_babies DESC
LIMIT 5
```

У нас этот запрос выполнялся 74 секунды и вернул следующие результаты:

Row	name	num_babies	frac_male
1	Jordan	982149616	0.6705115608373867
2	Willie	940460442	0.5722103705452823
3	Lee	820214744	0.689061146650151
4	Jessie	759150003	0.5139710590240227
5	Marion	592706454	0.32969114589732473

Заметим, однако, что этот ответ неправильный: как бы нам ни нравилось имя Jordan (Джордан), но все население США составляет всего около 330 миллионов

человек, поэтому не может быть 982 миллиона детей с таким именем. Самосоединение, к сожалению, не признает границ годов и штатов.¹

Более быстрое, красивое и правильное (!) решение состоит в том, чтобы прочитать исходные данные только один раз и вообще избежать самосоединения. Следующий запрос выполнялся всего 2.4 секунды — быстрее в 30 раз:

```
WITH all_babies AS (  
  SELECT  
    name  
    , SUM(IF(gender = 'M', number, 0)) AS male_babies  
    , SUM(IF(gender = 'F', number, 0)) AS female_babies  
  FROM `bigquery-public-data.usa_names.usa_1910_current`  
  GROUP BY name  
) ,  
  
both_genders AS (  
  SELECT  
    name  
    , (male_babies + female_babies) AS num_babies  
    , SAFE_DIVIDE(male_babies, male_babies + female_babies) AS frac_male  
  FROM all_babies  
  WHERE male_babies > 0 AND female_babies > 0  
)  
  
SELECT * FROM both_genders  
WHERE frac_male BETWEEN 0.3 and 0.7  
ORDER BY num_babies desc  
limit 5
```

Вот результаты этого запроса, если вам интересно:

Row	name	num_babies	frac_male
1	Jessie	229263	0.4327213723976394
2	Riley	187762	0.46760792918694943
3	Casey	181176	0.5916456925862145
4	Jackie	161428	0.4624042916966078
5	Johnnie	136208	0.6842549629977681

¹ Вы можете убедиться, что именно это является основной причиной получения неправильного значения для `num_babies`, добавив штат и год в предложение `USING` (и не забудьте добавить эти два поля в два первых оператора `SELECT`). Тогда число детей скорректируется в верном направлении (например, вы получите 2 018 162 для имени Джесси (Jessie), тогда как правильный ответ 229 263). Ответ все еще неверен, потому что соединение игнорирует записи с `NULL` в этих полях (`NULL` никогда не равно чему-либо еще).

Уменьшение объема данных, участвующих в соединении

Предыдущий запрос можно ускорить, добавив эффективное соединение и уменьшив объем соединяемых данных за счет группировки по имени и полу на ранних этапах:

```
with all_names AS (  
    SELECT name, gender, SUM(number) AS num_babies  
    FROM `bigquery-public-data`.usa_names.usa_1910_current  
    GROUP BY name, gender  
)  
  
male_names AS (  
    SELECT name, num_babies  
    FROM all_names  
    WHERE gender = 'M'  
)  
  
female_names AS (  
    SELECT name, num_babies  
    FROM all_names  
    WHERE gender = 'F'  
)  
  
ratio AS (  
    SELECT  
        name  
        , (f.num_babies + m.num_babies) AS num_babies  
        , m.num_babies / (f.num_babies + m.num_babies) AS frac_male  
    FROM male_names AS m  
    JOIN female_names AS f  
    USING (name)  
)  
  
SELECT * from ratio  
WHERE frac_male BETWEEN 0.3 and 0.7  
ORDER BY num_babies DESC  
LIMIT 5
```

Ранняя группировка уменьшает объем данных, достигающих оператора JOIN. В результате переупорядочение и другие сложные операции применяются к гораздо меньшему объему данных и выполняются весьма эффективно. Этот запрос выполнялся за две секунды и вернул правильный результат.

Применение оконных функций вместо самосоединения

Предположим, что вы решили определить продолжительность простоя велосипеда между поездками, то есть время, прошедшее от конца предыдущей поездки до начала следующей. Это пример зависимых отношений между записями. Может показаться, что единственный способ решить эту задачу — выполнить

самосоединение таблицы, сопоставив дату окончания одной поездки с датой начала следующей.

Однако и в этой задаче можно избежать самосоединения, используя оконные функции (подробнее об оконных функциях рассказывается в главе 8):

```
SELECT
  bike_id
, start_date
, end_date
, TIMESTAMP_DIFF(
  start_date,
  LAG(end_date) OVER (PARTITION BY bike_id ORDER BY start_date),
  SECOND) AS time_at_station
FROM `bigquery-public-data`.london_bicycles.cycle_hire
LIMIT 5
```

Вот результаты этого запроса:

Row	bike_id	start_date	end_date	time_at_station
1	2	2015-01-05 15:59:00 UTC	2015-01-05 16:17:00 UTC	<i>null</i>
2	2	2015-01-07 01:31:00 UTC	2015-01-07 01:50:00 UTC	119640
3	2	2015-01-21 07:56:00 UTC	2015-01-21 08:12:00 UTC	1231560
4	2	2015-01-21 16:15:00 UTC	2015-01-21 16:31:00 UTC	28980
5	2	2015-01-21 16:57:00 UTC	2015-01-21 17:23:00 UTC	1560

Обратите внимание, что первая запись в результатах имеет значение *null* в столбце *time_at_station*, потому что в наборе данных нет даты окончания предыдущей поездки. Затем по столбцу *time_at_station* можно проследить продолжительность простоя велосипеда.

Используя эти результаты, можно найти среднее время простоя велосипедов для каждого пункта проката и ранжировать станции по этому показателю:

```
WITH unused AS (
SELECT
  bike_id
, start_station_name
, start_date
, end_date
, TIMESTAMP_DIFF(start_date, LAG(end_date) OVER (PARTITION BY bike_id ORDER BY
start_date), SECOND) AS time_at_station
FROM `bigquery-public-data`.london_bicycles.cycle_hire
)

SELECT
  start_station_name
```

```

, AVG(time_at_station) AS unused_seconds
FROM unused
GROUP BY start_station_name
ORDER BY unused_seconds ASC
LIMIT 5

```

Теперь мы можем назвать пункты проката, где велосипеды простаивают меньше всего:

Row	start_station_name	unused_seconds
1	LSP1	1500.0
2	Wormwood Street, Liverpool Street	4605.427372968633
3	Hyde Park Corner, Hyde Park	5369.884926322234
4	Speakers' Corner 1, Hyde Park	6203.571977906734
5	Albert Gate, Hyde Park	6258.720194303267

Выполнение соединения с предварительно вычисленными значениями

Иногда полезно предварительно применить функции к небольшим таблицам, а затем выполнить соединение с предварительно вычисленными значениями, не повторяя дорогостоящие вычисления каждый раз.

Например, представьте, что вам понадобилось найти пару пунктов проката станций, между которыми клиенты ездят на велосипедах в наиболее быстром темпе. Чтобы вычислить темп¹ (минут на километр) езды, нужно разделить продолжительность поездки на расстояние между пунктами проката.

Мы можем создать денормализованную таблицу с расстояниями между пунктами проката, а затем найти средний темп:

```

with denormalized_table AS (
  SELECT
    start_station_name
    , end_station_name
    , ST_DISTANCE(ST_GeogPoint(s1.longitude, s1.latitude),
                  ST_GeogPoint(s2.longitude, s2.latitude)) AS distance
    , duration
  FROM
    `bigquery-public-data`.london_bicycles.cycle_hire AS h
  JOIN
    `bigquery-public-data`.london_bicycles.cycle_stations AS s1
  ON h.start_station_id = s1.id
  JOIN

```

¹ В велоспорте и беге темп — это величина, обратная скорости.

```

        `bigquery-public-data`.london_bicycles.cycle_stations AS s2
    ON h.end_station_id = s2.id
),

durations AS (
    SELECT
        start_station_name
        , end_station_name
        , MIN(distance) AS distance
        , AVG(duration) AS duration
        , COUNT(*) AS num_rides
    FROM
        denormalized_table
    WHERE
        duration > 0 AND distance > 0
    GROUP BY start_station_name, end_station_name
    HAVING num_rides > 100
)

SELECT
    start_station_name
    , end_station_name
    , distance
    , duration
    , duration/distance AS pace
FROM durations
ORDER BY pace ASC
LIMIT 5

```

Этот запрос вызывает геопространственную функцию ST_DISTANCE один раз для каждой записи в таблице cycle_hire (24 миллиона раз), выполняется 16.1 секунды и обрабатывает 1.86 Гбайт данных.

Как вариант, можно использовать таблицу cycle_stations, чтобы заранее вычислить расстояние между каждой парой станций (это самосоединение), а затем выполнить соединение полученного результата с таблицей, хранящей средние продолжительности поездок между пунктами проката:

```

with distances AS (
    SELECT
        a.id AS start_station_id
        , a.name AS start_station_name
        , b.id AS end_station_id
        , b.name AS end_station_name
        , ST_DISTANCE(ST_GeogPoint(a.longitude, a.latitude),
            ST_GeogPoint(b.longitude, b.latitude)) AS distance
    FROM
        `bigquery-public-data`.london_bicycles.cycle_stations a
    CROSS JOIN
        `bigquery-public-data`.london_bicycles.cycle_stations b
    WHERE a.id != b.id
),

```

```

durations AS (
  SELECT
    start_station_id
    , end_station_id
    , AVG(duration) AS duration
    , COUNT(*) AS num_rides
  FROM
    `bigquery-public-data`.london_bicycles.cycle_hire
  WHERE
    duration > 0
  GROUP BY start_station_id, end_station_id
  HAVING num_rides > 100
)

```

```

SELECT
  start_station_name
  , end_station_name
  , distance
  , duration
  , duration/distance AS pace
FROM distances
JOIN durations
USING (start_station_id, end_station_id)
ORDER BY pace ASC
LIMIT 5

```

Для выполнения этого запроса с более эффективными соединениями потребовалось всего 5.4 секунды, скорость выполнения увеличилась в три раза, а объем обработанных данных уменьшился до 554 Мбайт, благодаря чему стоимость запроса уменьшилась почти в четыре раза.

JOIN и денормализация

А что, если хранить протяженность каждой поездки в денормализованной таблице?

```

CREATE OR REPLACE TABLE ch07eu.cycle_hire AS
SELECT
  start_station_name
  , end_station_name
  , ST_DISTANCE(ST_GeogPoint(s1.longitude, s1.latitude),
    ST_GeogPoint(s2.longitude, s2.latitude)) AS distance
  , duration
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire AS h
JOIN
  `bigquery-public-data`.london_bicycles.cycle_stations AS s1
ON h.start_station_id = s1.id
JOIN
  `bigquery-public-data`.london_bicycles.cycle_stations AS s2
ON h.end_station_id = s2.id

```

Этот запрос возвращает результаты за 8.7 секунды и обрабатывает 1.6 Гбайт. Иначе говоря, он на 60% медленнее и примерно в три раза дороже, чем предыдущий запрос. То есть в данном случае соединение с меньшей таблицей оказалось более эффективным решением, чем обращение к большой денормализованной таблице. Однако вы обязательно должны проверять это в каждом конкретном случае. Далее вы увидите, как эффективно хранить данные с разными уровнями детализации в одной денормализованной таблице с вложенными и повторяющимися полями.

Исключение перегрузки рабочих серверов

Некоторые операции (например, оформление заказа) должны выполняться на одном рабочем сервере. Однако сортировка слишком большого объема данных может привести к переполнению памяти рабочего сервера и к ошибке «исчерпания ресурсов». Старайтесь не перегружать рабочие серверы слишком большим количеством данных. По мере обновления оборудования в вычислительных центрах Google понятие «слишком много» со временем расширяется. В настоящее время оно составляет порядка одного гигабайта.

Ограничение сортировки больших объемов данных

Предположим, вы хотите узнать номера договоров аренды и велосипедов — 1, 2, 3 и т. д. — в порядке завершения поездок. Это можно реализовать с помощью функции `ROW_NUMBER()` (мы рассмотрим оконные функции в главе 8):

```
SELECT
  rental_id
  , ROW_NUMBER() OVER(ORDER BY end_date) AS rental_number
FROM `bigquery-public-data`.london_bicycles.cycle_hire
ORDER BY rental_number ASC
LIMIT 5
```

Вот полученные результаты:

Row	rental_id	rental_number
1	40346512	1
2	40346508	2
3	40346519	3
4	40346510	4
5	40346520	5

Этот запрос выполняется 29.9 секунды, но обрабатывает всего 372 Мбайт. Такая низкая скорость обусловлена необходимостью сортировки всего набора данных

london_bicycles на одном рабочем сервере. Если бы мы попытались обработать набор данных большего размера, то исчерпали бы ресурсы рабочего сервера.

В таких случаях желательно подумать, как можно ограничить объем сортируемых данных и распределить их. На самом деле можно извлечь дату проката, а затем отсортировать поездки в пределах каждого дня:

```
WITH rentals_on_day AS (  
  SELECT  
    rental_id  
    , end_date  
    , EXTRACT(DATE FROM end_date) AS rental_date  
  FROM `bigquery-public-data.london_bicycles.cycle_hire`  
)  
  
SELECT  
  rental_id  
  , rental_date  
  , ROW_NUMBER() OVER(PARTITION BY rental_date ORDER BY end_date) AS  
  rental_number_on_day  
FROM rentals_on_day  
ORDER BY rental_date ASC, rental_number_on_day ASC  
LIMIT 5
```

Этот запрос выполняется 8.9 секунды (скорость увеличилась в три раза), потому что теперь сортировка выполняется только по данным за один день. Он возвращает номера договоров аренды ежедневно:

Row	rental_id	rental_date	rental_number_on_day
1	40346512	2015-01-04	1
2	40346508	2015-01-04	2
3	40346519	2015-01-04	3
4	40346510	2015-01-04	4
5	40346520	2015-01-04	5

Асимметричные данные

Та же проблема исчерпания ресурсов рабочего сервера (в данном случае — проблема переполнения памяти) может возникнуть при выполнении `ARRAY_AGG` с `GROUP BY`, если один из ключей встречается гораздо чаще, чем другие.¹

¹ Это текущее ограничение динамического выполнения; в будущем оно может быть ослаблено.

Поскольку операции фиксации среди более трех миллионов репозиторий в GitHub распределены достаточно равномерно, этот запрос успешно выполняется:

```
SELECT
  repo_name
  , ARRAY_AGG(STRUCT(author, committer, subject, message, trailer, difference,
encoding) ORDER BY author.date.seconds)
FROM `bigquery-public-data.github_repos.commits`, UNNEST(repo_name) AS repo_name
GROUP BY repo_name
```

Однако большинство людей, использующих GitHub, живут лишь в нескольких часовых поясах, поэтому мы получаем неудачную группировку по часовому поясу — в этом случае мы запросили, чтобы один рабочий сервер отсортировал значительную долю в 750 Гбайт:

```
SELECT
  author.tz_offset, ARRAY_AGG(STRUCT(author, committer, subject, message,
trailer, difference, encoding) ORDER BY author.date.seconds)
FROM `bigquery-public-data.github_repos.commits`
GROUP BY author.tz_offset
```

Одно из решений проблемы: добавить `LIMIT` в `ORDER BY`:

```
SELECT
  author.tz_offset, ARRAY_AGG(STRUCT(author, committer, subject, message,
trailer, difference, encoding) ORDER BY author.date.seconds LIMIT 1000)
FROM `bigquery-public-data.github_repos.commits`
GROUP BY author.tz_offset
```

Если потребуется отсортировать все данные, используйте более детальные ключи (то есть распределите обработку данных между большим количеством рабочих серверов), а затем агрегируйте результаты, соответствующие желаемому ключу. Например, вместо группировки только по часовому поясу можно группировать по часовому поясу *и* `repo_name`, а затем агрегировать по репозиториям, чтобы получить фактический ответ для каждого часового пояса:

```
SELECT
  repo_name, author.tz_offset
  , ARRAY_AGG(STRUCT(author, committer, subject, message, trailer, difference,
encoding) ORDER BY author.date.seconds)
FROM `bigquery-public-data.github_repos.commits`, UNNEST(repo_name) AS repo_name
GROUP BY repo_name, author.tz_offset
```

Оптимизация пользовательских функций

Для вызова пользовательской функции (User-Defined Function, UDF) на JavaScript требуется запустить подпроцесс V8, что снижает производительность. Пользовательские функции на JavaScript являются вычислительно затратными

и имеют доступ к ограниченной памяти, поэтому сокращение объема данных, обрабатываемых пользовательской функцией, может повысить производительность.

Хотя BigQuery поддерживает UDF на JavaScript, по возможности пишите свои функции на SQL — оптимизированном языке, встроенном в BigQuery, и поддерживающем распределенные вычисления. Все UDF на SQL — встроенные в запрос, временные или постоянные — имеют одинаковую производительность. Кроме того, пользовательские функции на SQL расширяют возможности повторного использования и композиции и улучшают читабельность.

Использование приближенных функций агрегирования

BigQuery предлагает быстрые, аппроксимирующие версии агрегатных функций, отличающиеся низким потреблением памяти. Вместо `COUNT(DISTINCT...)` к большим потокам данных можно применить `APPROX_COUNT_DISTINCT`, когда допустима небольшая статистическая неопределенность в результате.

Примерный подсчет

Рассмотрим для примера запрос, определяющий количество уникальных репозиториях GitHub:

```
SELECT
COUNT(DISTINCT repo_name) AS num_repos
FROM `bigquery-public-data`.github_repos.commits, UNNEST(repo_name) AS repo_name
```

Этот запрос выполняется 7.1 секунды и вычисляет правильный результат 3 348 576. А следующий запрос выполняется 3.2 секунды (быстрее в два раза) и возвращает приблизительный результат 3 400 927, больше правильного на 1.5%:

```
SELECT
  APPROX_COUNT_DISTINCT(repo_name) AS num_repos
FROM `bigquery-public-data`.github_repos.commits, UNNEST(repo_name) AS repo_name
```

Однако применение аппроксимирующих функций может оказаться бесполезным. Рассмотрим поиск общего количества уникальных велосипедов в наборе данных `london_bicycles`:

```
SELECT
  COUNT(DISTINCT bike_id) AS num_bikes
FROM `bigquery-public-data`.london_bicycles.cycle_hire
```

Этот запрос выполняется 0.9 секунды и возвращает точный результат 13 705. Аппроксимирующая версия запроса выполняется 1.6 секунды (даже медленнее точного запроса) и возвращает примерный результат 13 699:


```
SELECT
  APPROX_COUNT_DISTINCT(bike_id) AS num_bikes
FROM `bigquery-public-data`.london_bicycles.cycle_hire
```



Приближенные алгоритмы намного эффективнее точных только на больших наборах данных, и их рекомендуется использовать, когда допустимы отклонения примерно в 1%. Прежде чем использовать аппроксимирующую функцию, всегда оценивайте ее пригодность для своего случая!

Примерные наибольшие значения

Среди других аппроксимирующих функций можно назвать: `APPROX_QUANTILES` для вычисления перцентилей, `APPROX_TOP_COUNT` для поиска наибольших элементов и `APPROX_TOP_SUM` для вычисления суммы наибольших элементов.

Вот пример применения `APPROX_TOP_COUNT` для поиска пяти наиболее часто арендуемых велосипедов:

```
SELECT
  APPROX_TOP_COUNT(bike_id, 5) AS num_bikes
FROM `bigquery-public-data`.london_bicycles.cycle_hire
```

Он возвращает следующие результаты:

Row	num_bikes.value	num_bikes.count
1	12925	2922
	12841	2489
	13071	2474
	12926	2467
	12991	2444

Обратите внимание, что все результаты возвращаются в одной записи, в виде массива значений, что обеспечивает сохранение порядка.



Если ваши запросы выполняются слишком долго, попробуйте использовать `APPROX_TOP_COUNT` для проверки — возможно, причина заключается в асимметрии данных. Если это так, попробуйте применить советы, приведенные выше в этой главе и относящиеся к работе с асимметричными данными, и выполнить операцию на более детальном уровне или используйте `LIMIT` для сокращения обрабатываемых данных.

Чтобы найти пять пунктов проката с наибольшей длительностью поездок на взятых там велосипедах, можно использовать `APPROX_TOP_SUM`:

```
SELECT
  APPROX_TOP_SUM(start_station_name, duration, 5) AS num_bikes
FROM `bigquery-public-data`.london_bicycles.cycle_hire
WHERE duration > 0
```

Вот результаты этого запроса:

Row	num_bikes.value	num_bikes.sum
1	Hyde Park Corner, Hyde Park	600037440
	Black Lion Gate, Kensington Gardens	581085720
	Albert Gate, Hyde Park	367235700
	Speakers' Corner 1, Hyde Park	318485820
	Speakers' Corner 2, Hyde Park	268442640

Функции HLL

В дополнение к описанным функциям `APPROX_*` (которые реализуют аппроксимирующие алгоритмы агрегирования), BigQuery также поддерживает алгоритм HyperLogLog++ (HLL++) (<https://research.google.com/pubs/pub40671.html>), который позволяет разбить задачу подсчета уникальных значений на три отдельные операции:

1. Инициализация набора, который называют HLL-скетчем, добавлением в него новых элементов с помощью `HLL_COUNT.INIT`.
2. Определение мощности (количества элементов) HLL-скетча с помощью `HLL_COUNT.EXTRACT`.
3. Слияние двух HLL-скетчей в один с использованием `HLL_COUNT.MERGE_PARTIAL`.

Кроме того, `HLL_COUNT.MERGE` объединяет шаги 2 и 3, вычисляя количество по набору HLL-скетчей.

Например, вот запрос, который находит количество разных пунктов проката по таблице велосипедных поездок в Лондоне, независимо от того, началась или закончилась поездка в этом пункте проката:

```
WITH sketch AS (
SELECT
  HLL_COUNT.INIT(start_station_name) AS hll_start
  , HLL_COUNT.INIT(end_station_name) AS hll_end
FROM `bigquery-public-data`.london_bicycles.cycle_hire
)

SELECT
  HLL_COUNT.MERGE(hll_start) AS distinct_start
  , HLL_COUNT.MERGE(hll_end) AS distinct_end
  , HLL_COUNT.MERGE(hll_both) AS distinct_station
FROM sketch, UNNEST([hll_start, hll_end]) AS hll_both
```

Вот результаты этого запроса:

Row	distinct_start	distinct_end	distinct_station
1	880	882	882

Конечно, тот же подсчет можно выполнить с использованием `APPROX_COUNT_DISTINCT`:

```
SELECT
  APPROX_COUNT_DISTINCT(start_station_name) AS distinct_start
, APPROX_COUNT_DISTINCT(end_station_name) AS distinct_end
, APPROX_COUNT_DISTINCT(both_stations) AS distinct_station
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire
, UNNEST([start_station_name, end_station_name]) AS both_stations
```

Этот запрос вернет тот же результат, но он выглядит проще и понятнее. Поэтому мы обычно предпочитаем использовать версии `APPROX_`.

Одной из причин использования функций `HLL` может быть необходимость выполнять агрегирование вручную или предотвращение сохранения определенных столбцов. Предположим, что ваши данные имеют схему `user_id`, `date`, `product`, `country` и вам нужно определить количество уникальных пользователей. Столбец `user_id` идентифицирует личность, поэтому лучше не хранить его бесконечно. В таком случае можно вычислить агрегирование вручную, используя `HLL_COUNT.INIT`:

```
INSERT INTO approx_distinct_users_agg AS
SELECT date, product, country, HLL_COUNT.INIT(user_id) AS sketch
GROUP BY date, product, country, sketch
```

Теперь вам не нужно хранить `user_id`; достаточно сохранить скетч. После этого вы сможете использовать его всякий раз, когда потребуется вычислить агрегированное значение более высокого уровня, например:

```
SELECT date, HLL_COUNT.MERGE(sketch)
FROM approx_distinct_users_agg
GROUP BY date
```

Оптимизация хранения данных и доступа к ним

В предыдущем разделе мы обсудили способы увеличения производительности запросов, но ограничились методами, не меняющими компоновки таблицы, ее местоположения или способов обращения к ней. В этом разделе мы покажем, насколько существенное влияние на производительность может оказать учет этих факторов. Вспоминать эти советы следует на этапе разработки таблиц,

потому что изменение схемы существующей таблицы повлечет нарушение работоспособности существующих запросов.

Минимизация сетевых издержек

BigQuery — это региональный сервис, доступный по всему миру. Например, если вы запрашиваете набор данных, хранящийся в регионе EU, запрос будет выполняться на серверах, расположенных в вычислительном центре в Европе. Чтобы вы могли сохранить результаты запроса в таблице, она должна находиться в наборе данных, который также находится в регионе EU. Однако BigQuery REST API можно вызывать (то есть запустить запрос) из любой точки мира, даже с компьютеров за пределами GCP.

При работе с другими ресурсами GCP, такими как Google Cloud Storage или Cloud Pub/Sub, наилучшая производительность достигается, если они находятся в том же регионе, что и набор данных. Поэтому если запрос выполняется из экземпляра Compute Engine или кластера Cloud Dataproc, сетевые издержки окажутся минимальными, если экземпляр или кластер также находятся в том же регионе, что и запрашиваемый набор данных.

Обращаясь к BigQuery из-за пределов GCP, учитывайте топологию сети и постарайтесь свести к минимуму число переходов между клиентским компьютером и вычислительным центром GCP, в котором находится набор данных.

Сжатые, неполные ответы

При непосредственном обращении к REST API сетевые издержки можно уменьшить, принимая сжатые, неполные ответы. Для приема сжатых ответов можете указать в HTTP-заголовке, что вы готовы принять gzip-архив, и обеспечить наличие строки «gzip» в заголовке User-Agent, например:

```
Accept-Encoding: gzip
User-Agent: programName (gzip)
```

В этом случае все ответы будут сжиматься с помощью gzip.

По умолчанию ответы BigQuery содержат все поля, перечисленные в документации. Однако если известно, какая часть ответа нам интересна, мы можем попросить BigQuery послать только эту часть, уменьшив тем самым сетевые издержки. Например, в этой главе мы видели, как получить полную информацию о задании с помощью Jobs API. Если вас интересует только подмножество полного ответа (например, только шаги в плане запроса), можно указать интересующие поля, чтобы ограничить размер ответа:¹

¹ Это сценарий `07_perf/get_job_details_compressed.sh` в репозитории GitHub с примерами для этой книги.

```
JOBSURL="https://www.googleapis.com/bigquery/v2/projects/$PROJECT/jobs"
FIELDS="statistics(query(queryPlan(steps)))"
curl --silent \
  -H "Authorization: Bearer $access_token" \
  -H "Accept-Encoding: gzip" \
  -H "User-Agent: get_job_details (gzip)" \
  -X GET \
  "${JOBSURL}/${JOBID}?fields=${FIELDS}" \
| zcat
```

Обратите внимание: здесь также указано, что мы принимаем сжатые данные gzip.

Объединение нескольких запросов в пакеты

При использовании REST API есть возможность объединить несколько вызовов BigQuery API, используя тип содержимого `multipart/mixed` и вложенные HTTP-запросы в каждой из частей. В теле каждой части указывается HTTP-операция (GET, PUT и т. д.), путь в URL, заголовки и тело. В ответ сервер отправит единственный HTTP-ответ с типом содержимого `multipart/mixed`, каждая часть которого будет содержать ответ (по порядку) на соответствующий запрос в пакетном запросе. Несмотря на то что ответы возвращаются в определенном порядке, сервер может обрабатывать вызовы в любом порядке. Поэтому пакетный запрос можно рассматривать как группу запросов, выполняемых параллельно.

Вот пример отправки пакетного запроса для получения некоторых деталей из планов выполнения последних пяти запросов в нашем проекте. Сначала мы используем инструмент командной строки BigQuery, чтобы получить пять последних успешных заданий:¹

```
# 5 последних успешных заданий
JOBS=$(bq ls -j -n 50 | grep SUCCESS | head -5 | awk '{print $1}')
```

Запрос отправляется в конечную точку BigQuery, предназначенную для обработки пакетов:

```
BATCHURL="https://www.googleapis.com/batch/bigquery/v2"
JOBSPATH="/projects/$PROJECT/jobs"
FIELDS="statistics(query(queryPlan(steps)))"
```

В пути URL можно определить отдельные запросы:

```
request=""
for JOBID in $JOBS; do
  read -d '' part << EOF

  --batch_part_starts_here
```

¹ Это сценарий `07_perf/get_recent_jobs.sh` в репозитории GitHub с примерами для этой книги.

```
GET ${JOBSPATH}/${JOBID}?fields=${FIELDS}

EOF
request=$(echo "$request"; echo "$part")
done
```

Затем можно отправить запрос в виде составного запроса:

```
curl --silent \
  -H "Authorization: Bearer $access_token" \
  -H "Content-Type: multipart/mixed; boundary=batch_part_starts_here" \
  -X POST \
  -d "$request" \
  "${BATCHURL}"
```

Массовое считывание с использованием BigQuery Storage API

В главе 5 мы обсудили использование BigQuery REST API и клиентских библиотек для перечисления таблиц и получения результатов запросов. REST API возвращает данные в виде записей с разбивкой по страницам, которые лучше подходят для относительно небольших наборов результатов. Однако с появлением машинного обучения и распределенных инструментов извлечения, преобразования и загрузки (Extract, Transform, Load — ETL) теперь внешним инструментам требуется быстрый и эффективный массовый доступ к управляемому хранилищу BigQuery. Такой доступ к массовому чтению обеспечивается в BigQuery Storage API через протокол вызова удаленных процедур (Remote Procedure Call, RPC). С помощью BigQuery Storage API структурированные данные передаются по сети в двоичном формате сериализации, который точнее соответствует колоночному формату хранения данных. Это обеспечивает дополнительное распараллеливание набора результатов между несколькими потребителями.

Конечные пользователи не используют BigQuery Storage API напрямую.¹ Вместо этого они применяют Cloud Dataflow, Cloud Dataproc, TensorFlow, AutoML, и другие инструменты, использующие Storage API для чтения данных напрямую из управляемого хранилища, а не через BigQuery API.

Поскольку Storage API напрямую обращается к хранимым данным, разрешение на доступ к BigQuery Storage API отличается от существующего BigQuery API. Разрешения BigQuery Storage API должны настраиваться независимо от разрешений BigQuery.

BigQuery Storage API предоставляет несколько преимуществ инструментам, читающим данные непосредственно из управляемого хранилища BigQuery. Например, потребители могут читать непересекающиеся наборы записей из таблицы,

¹ Хотя это возможно; см. <https://cloud.google.com/bigquery/docs/reference/storage/samples>. Конечная точка BigQuery Storage API отличается от конечной точки BigQuery REST API — это `bigquerystorage.googleapis.com`.

используя несколько потоков (например, разрешив распределенное чтение данных от разных рабочих серверов в Cloud Dataproc), динамически сегментировать эти потоки (таким способом уменьшая хвостовую задержку, которая может быть серьезной проблемой для заданий MapReduce), выбирать подмножество столбцов для чтения (для передачи в структуры машинного обучения только признаков, используемых моделью), фильтровать значения столбцов (уменьшая объем данных, передаваемых по сети) и при этом гарантировать согласованность мгновенных снимков (то есть читая данные с определенного момента времени).

В главе 5 мы рассмотрели использование расширения `%bigquery` в Jupyter Notebook для загрузки результатов запросов в объекты `DataFrame`. Однако в примерах использовались относительно небольшие наборы данных — от десятка до нескольких сотен записей. А можно ли загрузить весь набор данных `london_bicycles` (24 миллиона записей) в `DataFrame`? Да, можно, но в этом случае для загрузки данных в `DataFrame` следует использовать `Storage API`, а не `BigQuery API`. Сначала нужно установить клиентскую библиотеку `Storage API` для Python с поддержкой `Avro` и `pandas`. Сделать это можно командой

```
%pip install google-cloud-bigquery-storage[fastavro,pandas]
```

Затем остается только использовать расширение `%bigquery`, как и прежде, но добавить параметр, требующий использовать `Storage API`:

```
%bigquery df --use_bqstorage_api --project $PROJECT
SELECT
  start_station_name
  , end_station_name
  , start_date
  , duration
FROM `bigquery-public-data`.london_bicycles.cycle_hire
```

Обратите внимание, что здесь мы используем способность `Storage API` предоставлять прямой доступ к отдельным столбцам; необязательно читать всю таблицу `BigQuery` в объект `DataFrame`. Если запрос вернет небольшой объем данных, расширение автоматически будет использовать `BigQuery API`. Поэтому не страшно, если вы всегда будете указывать этот флаг в ячейках блокнота. Чтобы включить флаг `--use_bqstorage_api` во всех ячейках блокнота, можно установить флаг контекста:

```
import google.cloud.bigquery.magics
google.cloud.bigquery.magics.context.use_bqstorage_api = True
```

Выбор эффективного формата хранения

Производительность запроса зависит от того, где и в каком формате хранятся данные, составляющие таблицу. В общем случае производительность тем выше, чем меньше запросу требуется выполнять поиск или преобразование типов.

Внутренние и внешние источники данных

BigQuery поддерживает запросы к внешним источникам, таким как Google Cloud Storage, Cloud Bigtable и Google Sheets, однако максимальная производительность запросов возможна только при использовании собственных таблиц.

В качестве хранилища аналитических данных для всех ваших структурированных и полуструктурированных данных мы советуем использовать BigQuery. Внешние источники данных лучше использовать только для промежуточного хранения (Google Cloud Storage), загрузки в режиме реального времени (Cloud Pub/Sub, Cloud Bigtable) или периодического обновления (Cloud SQL, Cloud Spanner). Далее настройте конвейер данных для загрузки данных по расписанию из этих внешних источников в BigQuery (см. главу 4).

Если вам понадобится запросить данные из Google Cloud Storage, по возможности сохраняйте их в сжатом колоночном формате (например, Parquet). Используйте форматы на основе записей, такие как JSON или CSV, только в крайнем случае.

Управление жизненным циклом промежуточных корзин

Если вы загружаете данные в BigQuery, предварительно помещая в облачное хранилище Google Cloud Storage, не забудьте удалить их из облачного хранилища после загрузки. Если для загрузки данных в BigQuery используется конвейер ETL (чтобы попутно значительно преобразовать их или оставить только часть данных), у вас может появиться желание сохранить исходные данные в Google Cloud Storage. В таких случаях снизить затраты вам поможет определение правил управления жизненным циклом корзин, понижающих класс хранения в Google Cloud Storage.

Вот как можно включить управление жизненным циклом корзины и настроить автоматическое перемещение данных из объединенных регионов или стандартных классов, возраст которых превысил 30 дней, в хранилище Nearline Storage, а данных, хранящихся в Nearline Storage дольше 90 дней, — в хранилище Coldline Storage:

```
gsutil lifecycle set lifecycle.yaml gs://some_bucket/
```

В этом примере файл *lifecycle.yaml* содержит следующий код:

```
{
"lifecycle": {
  "rule": [
    {
      "action": {
        "type": "SetStorageClass",
        "storageClass": "NEARLINE"
      },
      "condition": {
```



```

    "age": 30,
    "matchesStorageClass": ["MULTI_REGIONAL", "STANDARD"]
  }
},
{
  "action": {
    "type": "SetStorageClass",
    "storageClass": "COLDLINE"
  },
  "condition": {
    "age": 90,
    "matchesStorageClass": ["NEARLINE"]
  }
}
]}}

```

Вы можете использовать управление жизненным циклом не только для изменения класса объекта, но и для удаления объектов старше определенного порога.¹

Хранение данных в виде массивов и структур

Кроме прочих общедоступных наборов данных, в BigQuery имеется набор данных с информацией о циклонических штормах (ураганах, тайфунах, циклонах и т. д.), полученной метеорологическими службами по всему миру. Циклонические штормы могут длиться до нескольких недель, и замеры их метеорологических показателей производятся примерно каждые три часа. Предположим, что вы решили найти в этом наборе данных все штормы, случившиеся в 2018 году, максимальную скорость ветра, достигнутую каждым штормом, а также время и местоположение шторма, когда эта максимальная скорость была достигнута. Все эти сведения из общедоступного набора данных извлекает следующий запрос:

```

SELECT
  sid, number, basin, name,
  ARRAY_AGG(STRUCT(iso_time, usa_latitude, usa_longitude, usa_wind) ORDER BY
  usa_wind DESC LIMIT 1)[OFFSET(0)].*
FROM
  `bigquery-public-data`.noaa_hurricanes.hurricanes
WHERE
  season = '2018'
GROUP BY
  sid, number, basin, name
ORDER BY number ASC

```

Запрос извлекает идентификатор шторма (*sid*), его порядковый номер в сезоне, бассейн и имя шторма (если оно было присвоено), а затем находит массив на-

¹ Дополнительную информацию можно найти по ссылке https://cloud.google.com/storage/docs/managing-lifecycles#change_an_objects_storage_class.

блюдений, выполненных для этого шторма, ранжируя наблюдения в порядке убывания скорости ветра и выбирая максимальную скорость для каждого шторма. Сами штормы упорядочены по порядковому номеру. Результат включает 88 записей и выглядит примерно так:

Row	sid	number	basin	name	iso_time	usa_latitude	usa_longitude	usa_wind
1	2018002N09123	1	WP	BOLAVEN	2018-01-02 18:00:00 UTC	9.7	117.2	29
2	2018003S15053	2	SI	AVA	2018-01-05 06:00:00 UTC	-17.9	50.0	93
3	2018006S13092	3	SI	IRVING	2018-01-07 18:00:00 UTC	-15.8	83.0	89
4	2018010S18123	4	SI	JOYCE	2018-01-11 18:00:00 UTC	-18.7	121.6	54

Запрос выполнялся 1.4 секунды и обработал 41.7 Мбайт. Первая запись описывает шторм Болавен (Bolaven), достигший максимальной скорости 29 м/с 2 января 2018 года в 18:00 UTC.

Поскольку наблюдения проводятся несколькими метеорологическими службами, эти данные можно стандартизировать с использованием вложенных полей и сохранить структуры в BigQuery, как показано ниже:¹

```
CREATE OR REPLACE TABLE ch07.hurricanes_nested AS

SELECT sid, season, number, basin, name, iso_time, nature, usa_sshs,
       STRUCT(usa_latitude AS latitude, usa_longitude AS longitude, usa_wind AS
wind, usa_pressure AS pressure) AS usa,
       STRUCT(tokyo_latitude AS latitude, tokyo_longitude AS longitude,
tokyo_wind AS wind, tokyo_pressure AS pressure) AS tokyo,
       ... AS cma,
       ... AS hko,
       ... AS newdelhi,
       ... AS reunion,
       ... AS bom,
       ... AS wellington,
       ... AS nadi
FROM `bigquery-public-data`.noaa_hurricanes.hurricanes
```

¹ Все запросы из этого раздела можно найти в файле 07_perf/hurricanes.sql в репозитории GitHub с примерами для этой книги.

Запросы к этой таблице выглядят так же, как запросы к исходной таблице, но с незначительным изменением имен столбцов (`usa.latitude` вместо `usa_latitude`):

```
SELECT
    sid, number, basin, name,
    ARRAY_AGG(STRUCT(iso_time, usa.latitude, usa.longitude, usa.wind) ORDER BY
    usa.wind DESC LIMIT 1)[OFFSET(0)].*
FROM
    ch07.hurricanes_nested
WHERE
    season = '2018'
GROUP BY
    sid, number, basin, name
ORDER BY number ASC
```

Этот запрос обрабатывает тот же объем данных и выполняется в течение того же времени, что и исходный, использующий общедоступный набор данных. Применение вложенных полей (структур) не меняет скорости или стоимости запроса, но может сделать запрос более читабельным.

Поскольку существует множество наблюдений одного и того же шторма в течение его продолжительности, мы можем изменить хранилище так, чтобы уместить в одну запись весь массив наблюдений для каждого шторма:

```
CREATE OR REPLACE TABLE ch07.hurricanes_nested_track AS
```

```
SELECT sid, season, number, basin, name,
    ARRAY_AGG(
        STRUCT(
            iso_time,
            nature,
            usa_sshs,
            STRUCT(usa_latitude AS latitude, usa_longitude AS longitude, usa_wind AS
wind, usa_pressure AS pressure) AS usa,
            STRUCT(tokyo_latitude AS latitude, tokyo_longitude AS longitude,
                tokyo_wind AS wind, tokyo_pressure AS pressure) AS tokyo,
            ... AS cma,
            ... AS hko,
            ... AS newdelhi,
            ... AS reunion,
            ... AS bom,
            ... AS wellington,
            ... AS nadi
        ) ORDER BY iso_time ASC ) AS obs
FROM `bigquery-public-data`.noaa_hurricanes.hurricanes
GROUP BY sid, season, number, basin, name
```

Обратите внимание, что теперь мы храним `sid`, `season` и другие характеристики шторма в виде скалярных столбцов, потому что они не меняются в зависимости от его продолжительности. Остальные данные, изменяющиеся с каждым на-

блюдением, хранятся в виде массива структур. Вот как выглядит запрос к новой таблице:¹

```
SELECT
  number, name, basin,
  (SELECT AS STRUCT iso_time, usa.latitude, usa.longitude, usa.wind
   FROM UNNEST(obs) ORDER BY usa.wind DESC LIMIT 1).*
FROM ch07.hurricanes_nested_track
WHERE season = '2018'
ORDER BY number ASC
```

Этот запрос вернет тот же результат, но на этот раз обработает только 14.7 Мбайт (снижение стоимости в три раза) и завершится за одну секунду (увеличение скорости на 30%). Чем обусловлено это улучшение производительности? Когда данные хранятся в виде массива, количество записей в таблице резко сокращается (с 682 000 до 14 000),² потому что теперь на один шторм приходится только одна запись, а не много записей — по одной для каждого наблюдения. Затем, когда мы фильтруем строки по сезону, BigQuery может одновременно отбрасывать множество связанных наблюдений, как показано на рис. 7.13.

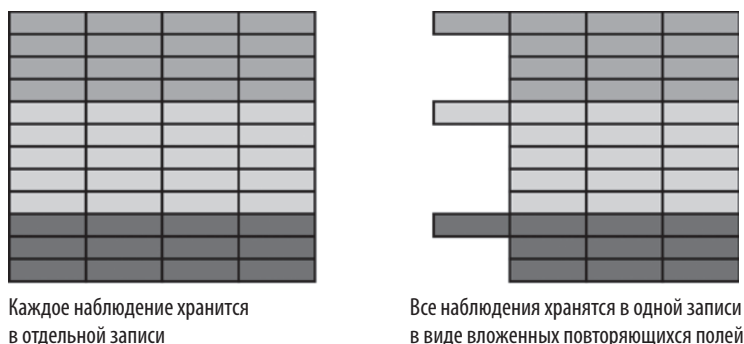


Рис. 7.13. Использование вложенных и повторяющихся полей может повысить производительность запросов, потому что BigQuery может отбрасывать сразу множество связанных наблюдений

Еще одно преимущество — отсутствие необходимости дублировать записи с данными, когда в одной таблице хранятся наблюдения с разными уровнями детализации. В одной таблице можно хранить как данные об изменении широты и долготы штормов, так и высокоуровневые данные, такие как название штормов и сезон. А поскольку BigQuery хранит табличные данные по столбцам с использованием сжатия, запрашивать и обрабатывать высокоуровневые дан-

¹ Синтаксис инструкций SQL для работы с массивами описывается в главе 2.

² Поскольку набор данных о штормах постоянно обновляется, число записей может увеличиться к тому моменту, когда вы будете читать эту книгу.

ные можно, не опасаясь затрат за работу с детальными данными, — теперь они хранятся в виде отдельного массива значений для каждого шторма.

Например, чтобы узнать число штормов по годам, можно запросить только нужные столбцы:

```
WITH hurricane_detail AS (
SELECT sid, season, number, basin, name,
  ARRAY_AGG(
    STRUCT(
      iso_time,
      nature,
      usa_sshs,
      STRUCT(usa_latitude AS latitude, usa_longitude AS longitude, usa_wind AS
wind, usa_pressure AS pressure) AS usa,
      STRUCT(tokyo_latitude AS latitude, tokyo_longitude AS longitude,
        tokyo_wind
AS wind, tokyo_pressure AS pressure) AS tokyo
    ) ORDER BY iso_time ASC ) AS obs
FROM `bigquery-public-data`.noaa_hurricanes.hurricanes
GROUP BY sid, season, number, basin, name
)

SELECT
  COUNT(sid) AS count_of_storms,
  season
FROM hurricane_detail
GROUP BY season
ORDER BY season DESC
```

Предыдущий запрос обработал 27 Мбайт, что в два раза меньше 56 Мбайт, которые пришлось бы обработать, если не использовать вложенные повторяющиеся поля.

Вложенные поля сами по себе не повышают производительность, хотя могут улучшить удобочитаемость, фактически выполняя соединение с другими связанными таблицами. Кроме того, вложенные повторяющиеся поля чрезвычайно полезны с точки зрения производительности. Подумайте над возможностью использовать вложенные повторяющиеся поля в своей схеме, потому что они могут значительно повысить скорость и снизить стоимость запросов, фильтрующих по столбцу, не являющемуся вложенным или повторяющимся (в нашем случае `season`).

Ключевой недостаток вложенных повторяющихся полей — сложность реализации потоковой передачи в такую таблицу, если потоковые обновления включают добавление элементов в существующие массивы. Реализовать это намного сложнее, чем добавление новых записей: вам необходимо будет изменить существующую запись, — для таблицы с информацией о штормах это существенный недостаток, поскольку в нее постоянно добавляются новые наблюдения, и это объясняет, почему в этом общедоступном наборе данных не используются вложенные повторяющиеся поля.

ПРАКТИКА ПРИМЕНЕНИЯ МАССИВОВ

Как показывает опыт, для успешного применения вложенных повторяющихся полей требуется некоторая практика. Образцовый набор данных Google Analytics¹ в BigQuery идеально подходит для этой цели. Самый простой способ идентифицировать вложенные данные в схеме — найти слово **RECORD** в *столбце Type*, которое соответствует типу данных **STRUCT**, и слово **REPEATED** в *столбце Mode*, как показано ниже:

Field name	Type	Mode
visitorId	INTEGER	NULLABLE
visitStartTime	INTEGER	NULLABLE
date	STRING	NULLABLE
totals	RECORD	NULLABLE
totals. visits	INTEGER	NULLABLE
totals. hits	INTEGER	NULLABLE
hits	RECORD	REPEATED
hits. hitNumber	INTEGER	NULLABLE
hits. time	INTEGER	NULLABLE

В этом примере поле **TOTALS** имеет тип **STRUCT** (но не повторяется), а **HITS** — имеет тип **STRUCT** и повторяется. В этом есть определенный смысл, потому что Google Analytics отслеживает данные сеанса посетителя (**visitor**) на уровне агрегирования (одно значение сеанса для **totals.hits**) и на уровне детализации (отдельные значения **hit.time** для каждой страницы и изображения, полученные с вашего сайта). Хранение данных на этих разных уровнях детализации без дублирования **visitorId** в записях возможно только в случае применения массивов.

После сохранения данных в повторяющемся формате с массивами вам нужно предусмотреть развертывание этих данных в своих запросах с помощью **UNNEST**, например:²

```
SELECT DISTINCT
  visitId
  , totals.pageviews
  , totals.timeOnsite
  , trafficSource.source
  , device.browser
```

¹ Набор данных `bigquery-public-data.google_analytics_sample.ga_sessions_20170801`.

² Этот код можно найти в файле `07_perf/google_analytics.sql` в репозитории GitHub с примерами для этой книги.

```

, device.isMobile
, h.page.pageTitle
FROM
  `bigquery-public-data`.google_analytics_sample.ga_sessions_20170801,
  UNNEST(hits) AS h
WHERE
  totals.timeOnSite IS NOT NULL AND h.page.pageTitle =
  'Shopping Cart'
ORDER BY pageviews DESC
LIMIT 10

```

Операция развертывания разбивает массив на части, например превращает [1,2,3,4,5] в отдельные записи:

```

[1,
2
3
4
5]

```

После этого можно выполнять обычные операции SQL, такие как WHERE, чтобы отфильтровать попадания на страницы с такими заголовками, как «Shopping Cart». Попробуйте!

С другой стороны, общедоступный набор данных с информацией об операциях фиксации в GitHub (`bigquery-publicdata.github_repos.commits`) использует вложенное повторяющееся поле (`repo_name`) для хранения списка репозитория, затронутых операциями фиксации. Оно не меняется с течением времени и обеспечивает ускорение запросов, которые выполняют фильтрацию по любому другому полю.

Хранение данных в виде географических типов

В общедоступном наборе со вспомогательными данными в BigQuery имеется таблица границ зон действия почтовых индексов США (`bigquery-public-data.utility_us.zipcode_area`) и еще одна таблица с многоугольниками, описывающими границы городов США (`bigquery-publicdata.utility_us.us_cities_area`). Столбец с границами зон действия почтовых индексов (`zipcode_geom`) представляет собой строку,¹ тогда как столбец с границами городов (`city_geom`) представлен географическим типом.

¹ На самом деле это устаревший набор данных, сохранившийся по причинам, которые будут описаны далее в этом разделе. В более современном наборе `bigquery-public-data.geo_us_boundaries.us_zip_codes` для представления тех же данных используются географические типы.

Из этих двух таблиц можно получить список всех почтовых индексов для Санта-Фе (Santa Fe) в Нью-Мексико (New Mexico), как показано ниже:¹

```
SELECT name, zipcode
FROM `bigquery-public-data`.utility_us.zipcode_area
JOIN `bigquery-public-data`.utility_us.us_cities_area
ON ST_INTERSECTS(ST_GeogFromText(zipcode_geom), city_geom)
WHERE name LIKE '%Santa Fe%'
```

Этот запрос выполняется 51.9 секунды, обрабатывает 305.5 Мбайт данных и возвращает следующие результаты:

Row	name	zipcode
1	Santa Fe, NM	87505
2	Santa Fe, NM	87501
3	Santa Fe, NM	87507
4	Eldorado at Santa Fe, NM	87508
5	Santa Fe, NM	87508
6	Santa Fe, NM	87506

Почему этот запрос выполняется так долго? Вовсе не из-за операции ST_INTERSECTS, а главным образом потому, что функция ST_GeogFromText должна вычислить ячейки S2² и построить тип GEOGRAPHY, соответствующий каждому почтовому индексу.

Мы можем попробовать изменить таблицу почтовых индексов, выполнив эту операцию заранее, и сохранить геометрию в виде значения типа GEOGRAPHY:

```
CREATE OR REPLACE TABLE ch07.zipcode_area AS
SELECT
  * REPLACE(ST_GeogFromText(zipcode_geom) AS zipcode_geom)
FROM
  `bigquery-public-data`.utility_us.zipcode_area
```



SELECT * REPLACE (см. предыдущий запрос) — это удобный способ заменить столбец из выражения SELECT *.

Новый набор данных имеет размер 131.8 Мбайт, что значительно больше 116.5 Мбайт в исходной таблице. Однако запросы к этой таблице могут использовать охват S2 и выполняться намного быстрее. Например, следующий

¹ Географические типы и геопространственные функции описываются в главе 4.

² См. <https://oreil.ly/PkIsx>.

запрос выполняется 5.3 секунды (увеличение скорости в 10 раз) и обрабатывает 320.8 Мбайт (небольшое увеличение стоимости при использовании тарифного плана «до востребования»):

```
SELECT name, zipcode
FROM ch07.zipcode_area
JOIN `bigquery-public-data`.utility_us.us_cities_area
ON ST_INTERSECTS(zipcode_geom, city_geom)
WHERE name LIKE '%Santa Fe%'
```

Преимущества в производительности, которые дает хранение географических данных в столбце типа `GEOGRAPHY`, более чем убедительны. Вот почему набор данных `utility_us` устарел (он все еще доступен для сохранения работоспособности уже написанных запросов). Мы советуем использовать таблицу `bigquery-public-data.geo_us_boundaries.us_zip_codes`, в которой географическая информация хранится в столбце типа `GEOGRAPHY` и постоянно обновляется.

Секционирование таблиц для уменьшения объема сканирования

Представьте, что вы часто запрашиваете набор данных `london_bicycles`, выполняя фильтрацию по годам:

```
SELECT
  start_station_name
  , AVG(duration) AS avg_duration
FROM `bigquery-public-data`.london_bicycles.cycle_hire
WHERE EXTRACT(YEAR from start_date) = 2015
GROUP BY start_station_name
ORDER BY avg_duration DESC
LIMIT 5
```

Этот запрос выполняется 2.8 секунды, обрабатывает 1 Гбайт данных и возвращает пункты проката с самыми долгими поездками в 2015 году:

Row	start_station_name	avg_duration
1	Mechanical Workshop Penton	105420.0
2	Contact Centre, Southbury House	5303.75
3	Stewart's Road, Nine Elms	4836.380090497735
4	Black Lion Gate, Kensington Gardens	4788.747908066496
5	Speakers' Corner 2, Hyde Park	4610.192911183014

Однако чтобы найти записи, соответствующие 2015 году, этот запрос должен прочитать всю таблицу. В этом разделе мы рассмотрим разные способы уменьшения объема обрабатываемых данных.

Антипаттерн: окончания в именах таблиц и подстановочные знаки

Если в запросах часто приходится использовать фильтрацию по годам, сократить количество читаемых данных можно, сохранив данные в нескольких таблицах путем добавления в конец каждой номер года. При такой организации для получения данных за 2015 год не потребуется извлекать записи, соответствующие другим годам, достаточно просто прочитать таблицу `cycle_hire_2015`.

Давайте попробуем этот способ, создав таблицу, как показано ниже:

```
CREATE OR REPLACE TABLE ch07eu.cycle_hire_2015 AS (  
  SELECT * FROM `bigquery-public-data`.london_bicycles.cycle_hire  
  WHERE EXTRACT(YEAR from start_date) = 2015  
)
```

Теперь перепишем запрос, чтобы он использовал эту таблицу (см. ниже). Этот запрос выполняется за одну секунду (увеличение скорости в три раза) и обрабатывает всего 345 Мбайт (экономия в три раза):

```
SELECT  
  start_station_name  
  , AVG(duration) AS avg_duration  
FROM ch07eu.cycle_hire_2015  
GROUP BY start_station_name  
ORDER BY avg_duration DESC  
LIMIT 5
```



Вместо разбиения данных вручную по нескольким таблицам используйте секционированные и шаблонные (описываются ниже) таблицы.

Для организации поиска по нескольким годам можно использовать подстановочные знаки в окончаниях имен таблиц:

```
SELECT  
  start_station_name  
  , AVG(duration) AS avg_duration  
FROM `ch07eu.cycle_hire_*`  
WHERE _TABLE_SUFFIX BETWEEN '2015' AND '2016'  
GROUP BY start_station_name  
ORDER BY avg_duration DESC  
LIMIT 5
```

Секционированные таблицы

Секционированные таблицы позволяют хранить все связанные данные в одной логической таблице и при этом эффективно запрашивать подмножество этих данных. Например, если вы храните данные за весь год, но обычно запрашиваете данные только за последнюю неделю, секционирование по времени позволит

удешевить такие запросы, потому что они будут сканировать только секции, соответствующие последним семи дням. Такой подход может значительно ускорить запросы и сделать их дешевле.

Создание таблиц с именами по годам, как было показано в предыдущем разделе, неэффективно: для этого BigQuery должна будет поддерживать копию схемы и метаданных для каждой из таких таблиц и проверять их разрешения. Кроме того, в одном запросе нельзя использовать более 1000 таблиц, и это может затруднить выполнение запросов ко всему набору данных, если таблицы будут разделены по дате, а не по году (в таком случае 1000 таблиц не охватит даже трех лет). Кроме того, потоковая передача данных в таблицы, разделенные по датам, может потребовать синхронизации часов и часовых поясов между несколькими клиентами. Поэтому мы советуем использовать секционированные таблицы.



Секционирование и кластеризация являются наиболее эффективными способами снижения стоимости и повышения производительности запросов. Не сомневайтесь, благодаря этим способам вы сможете использовать оптимизации, недоступные для несекционированных или некластеризованных таблиц.

Секционированная таблица — это специальная таблица, разделенная на секции (или разделы), управляемые сервисом BigQuery. Вот как можно создать секционированную версию `cycle_hire`:

```
CREATE OR REPLACE TABLE ch07eu.cycle_hire_partitioned
  PARTITION BY DATE(start_date) AS
SELECT * FROM `bigquery-public-data`.london_bicycles.cycle_hire
```



Контролировать расходы на хранение можно, указав предельное время хранения секций и отправив BigQuery запрос, обеспечивающий условия, при которых пользователи всегда будут использовать фильтр по секциям (и не запрашивать всю таблицу по ошибке):

```
CREATE OR REPLACE TABLE ch07eu.cycle_hire_partitioned
  PARTITION BY DATE(start_date)
  OPTIONS(partition_expiration_days=1000,
          require_partition_filter=true) AS
SELECT * FROM `bigquery-publicdata`.
london_bicycles.cycle_hire
```

Если вы забудете настроить этот параметр при создании таблицы, его всегда можно добавить позже:

```
ALTER TABLE ch07eu.cycle_hire_partitioned
SET OPTIONS(require_partition_filter=true)
```

Теперь, чтобы найти пункты проката, в которых аренда велосипедов в 2015 году осуществлялась в среднем на более длительный срок, можно выполнить запрос к секционированной таблице, не забыв добавить фильтр по столбцу секционирования (`start_date`):

```
SELECT
  start_station_name
    , AVG(duration) AS avg_duration
FROM ch07eu.cycle_hire_partitioned
WHERE start_date BETWEEN '2015-01-01' AND '2015-12-31'
GROUP BY start_station_name
ORDER BY avg_duration DESC
LIMIT 5
```

Этот запрос выполняется одну секунду и обрабатывает всего 419.4 Мбайт, что немного больше, чем в случае использования отдельных таблиц по годам (из-за необходимости читать столбец `start_date`), однако он все равно экономит время из-за отсутствия необходимости читать полный набор данных. Обратите внимание, что при такой формулировке запроса проявляются некоторые недостатки:

```
SELECT
  start_station_name
    , AVG(duration) AS avg_duration
FROM ch07eu.cycle_hire_partitioned
WHERE EXTRACT(YEAR FROM start_date) = 2015
GROUP BY start_station_name
ORDER BY avg_duration DESC
LIMIT 5
```

Этот запрос обрабатывает 1 Гбайт данных и не дает никакой экономии. Чтобы секционирование приносило пользу, среда выполнения BigQuery должна иметь возможность статически определять фильтры по секциям.



Можно отправить BigQuery запрос автоматически секционировать таблицу по времени приема, а не по столбцу с датой/временем. Для этого в роли столбца секционирования можно использовать `_PARTITIONTIME` или `_PARTITIONDATE`. Это псевдостолбцы, определяющие время приема и в действительности не существующие в наборе данных. Однако их можно использовать в запросах, чтобы ограничить число проверяемых записей.

При потоковой передаче в таблицу с секционированием по времени загрузки данные в потоковом буфере хранятся в разделе `__UNPARTITIONED__`. Чтобы получить данные из этого раздела, ищите значение `NULL` в псевдостолбце `_PARTITIONTIME`.

Кластеризация таблиц на основе ключей с большой мощностью множества

Кластеризация, как и секционирование, — это способ хранения информации в BigQuery, позволяющий запросам читать меньше данных. В отличие от секционированной таблицы, которая ведет себя аналогично группе независимых таблиц

(по одной на секцию), кластеризованные таблицы хранятся в отсортированном формате как единое целое. Такой порядок позволяет хранить неограниченное количество уникальных значений без снижения производительности, а также, при применении фильтра, позволяет BigQuery пропустить любые файлы, не содержащие запрашиваемого диапазона значений.

Кластеризация может выполняться по любым простым неповторяющимся столбцам (INT64, BOOL, NUMERIC, STRING, DATE, GEOGRAPHY и TIMESTAMP). Обычно кластеризация выполняется по столбцам с очень большим количеством уникальных значений, например `customerId`, при наличии миллионов клиентов. Если есть столбцы с меньшим количеством элементов, но часто используемые вместе, можно кластеризовать таблицу сразу по нескольким столбцам. В таком случае фильтрацию можно осуществлять по любой последовательности столбцов кластеризации, начиная с первого, и пользоваться всеми преимуществами кластеризации.

Если большинство запросов к набору данных проката велосипедов будет использовать столбцы `start_station_name` и `end_station_name`, мы могли бы оптимизировать хранилище, чтобы дать возможность использовать эту общность наших запросов, создав следующую таблицу:

```
CREATE OR REPLACE TABLE ch07eu.cycle_hire_clustered
PARTITION BY DATE(start_date)
CLUSTER BY start_station_name, end_station_name
AS (
  SELECT * FROM `bigquery-public-data`.london_bicycles.cycle_hire
)
```

После этого запросы, использующие столбцы кластеризации по порядку, могли бы принести значительную пользу, например:

```
SELECT
  start_station_name
  , end_station_name
  , AVG(duration) AS duration
FROM ch07eu.cycle_hire_clustered
WHERE
  start_station_name LIKE '%Kennington%'
  AND end_station_name LIKE '%Hyde%'
GROUP BY start_station_name, end_station_name
```

Но в данном случае вся таблица занимает всего 1.5 Гбайт и уместается в один блок, поэтому улучшения не наблюдается.

Преимущества кластеризации заметны только на больших таблицах. Наш коллега Фелипе Хоффа (Felipe Hoffa) создал таблицу объемом 2.20 Тбайт, описывающую просмотры статей в Википедии и кластеризованную следующим образом:

```
CLUSTER BY wiki, title
```

Пока наши запросы будут выполнять фильтрацию по столбцу `wiki` (и, возможно, по столбцу `title`), они будут получать преимущества кластеризации. Например, подсчитаем количество просмотров страниц в англоязычной Википедии в июне 2017 года, содержащих слово «Liberia»:

```
SELECT title, SUM(views) AS views
FROM `fh-bigquery.wikipedia_v3.pageviews_2017`
WHERE DATE(datehour) BETWEEN '2017-06-01' AND '2017-06-30'
AND wiki = 'en'
AND title LIKE '%Liberia%'
GROUP BY title
```

Этот запрос был выполнен за 4.8 секунды и обработал 38.6 Гбайт. Если бы таблица не была кластеризована (а только секционирована),¹ на выполнение запроса потребовалось бы 25.9 секунды (в пять раз медленнее) и он обработал бы 180.2 Гбайт (в пять раз дороже). С другой стороны, запрос, который не фильтрует по столбцу `wiki`, не будет иметь никаких преимуществ.

Кластеризация по столбцу секционирования

Секционирование таблиц осуществляется по дате (по фактическому столбцу или по времени приема). Чтобы организовать секционирование на уровне часов, можно использовать разбиение на основе даты, а затем кластеризовать по часам вместе с другими подходящими атрибутами.

То есть кластеризация по столбцу секционирования довольно часто используется на практике. Например, если секционировать таблицу по столбцу `event_time` — отметке времени события, — это позволит выполнять очень быстрые и эффективные запросы, охватывающие произвольные периоды времени, которые меньше границ дней, использованных для секционирования. Например, можно запросить данные только за последние 10 минут, и при этом запросу не придется сканировать более старые записи.

В секционированной таблице каждая секция состоит из данных за одни сутки, и BigQuery поддерживает метаданные, обеспечивающие использование и поддержание целостности секций всеми запросами, заданиями загрузки и операторами языка определения данных (Data Definition Language, DDL) или DML.

Повторная кластеризация

BigQuery сортирует данные в кластеризованной таблице по значениям в столбцах кластеризации и организует их в хранимые блоки оптимального размера для эффективного сканирования и удаления ненужных данных. Однако, в отличие

¹ Попробуйте использовать `fh-bigquery.wikipedia_v2.pageviews_2017`.

от секционирования, BigQuery не поддерживает сортировку данных внутри кластеров. BigQuery периодически повторно кластеризует данные, чтобы обеспечить эффективное удаление ненужных данных и скорость сканирования. Увидеть, насколько эффективно кластеризована таблица, можно, взглянув на показатель `clustering_ratio` таблицы (значение 1.0 соответствует максимально оптимальной кластеризации).

Изменение таблицы с использованием инструкций DML приводит к повторной кластеризации секции. Предположим, что вы периодически получаете таблицу исправлений (которая содержит поездки, пропущенные в предыдущих обновлениях). Использование оператора `MERGE`, как показано ниже, вызовет повторную кластеризацию обновленных секций:

```
MERGE ch07eu.cycle_hire_clustered all_hires
USING ch07eu.cycle_hire_corrections some_month
ON all_hires.start_station_name = some_month.start_station_name
WHEN MATCHED
  AND all_hires._PARTITIONTIME = DATE(some_month.start_date) THEN
  INSERT (rental_id, duration, ...)
  VALUES (rental_id, duration, ...)
```



Если вы не хотите ждать, пока BigQuery наконец решит выполнить повторную кластеризацию таблицы, в которую вы записали изменения, можете воспользоваться возможностью операторов DML принудительно запускать кластеризацию. Например, можно применить пустую инструкцию `UPDATE` к конкретным секциям (например, записанным за последние 24 часа):

```
UPDATE ch07eu.cycle_hire_clustered
SET start_station_id = 300
WHERE start_station_id = 300
AND start_date > TIMESTAMP_SUB(CURRENT_TIMESTAMP(),
INTERVAL 1 DAY)
```

В табл. 7.1 перечислены различия между секционированием и кластеризацией, и вы можете использовать ее, выбирая между двумя данными методами.

Таблица 7.1. Секционирование и кластеризация

	Секционирование	Кластеризация
Уникальные значения	Меньше 10 000	Не ограничено
Управление данными	Подобно таблицам (можно определять сроки действия, удалять и т. д.)	Только средствами DML
Определение стоимости по пробным прогонам	Точно	Верхняя граница

Таблица 7.1 (окончание)

	Секционирование	Кластеризация
Определение окончательной стоимости	Точно	На уровне блока (трудно предсказать)
Поддержка	Не требуется (секционирование производится сразу же)	В фоновом режиме (могут возникать задержки, вызванные выполнением повторной кластеризации в фоновом режиме)

Дополнительные преимущества кластеризации

Как вы наверняка помните, мы говорили, что запросы вида `SELECT * ... LIMIT 10` в BigQuery являются антишаблоном, потому что стоят ровно столько же, сколько и полное сканирование таблицы. С кластеризованными таблицами ситуация иная. При чтении из кластеризованной таблицы BigQuery будет использовать любые оптимизации, которые помогут предотвратить чтение ненужных данных. Поэтому если вы выполните запрос `SELECT * ... LIMIT 10` для кластеризованной таблицы, механизм выполнения сможет остановить чтение данных, как только подготовит 10 записей. Механизм запросов использует несколько рабочих серверов, действующих параллельно, и любой из них может завершить выполнение задания первым, поэтому объем сканирования нельзя предсказать. С другой стороны, запросы к большим таблицам обходятся намного дешевле.

Удивительным побочным эффектом «ранней остановки» для сокращения затрат является возможность получить выигрыш в производительности, даже не выполняя фильтрацию по столбцам кластеризации: если выполнять фильтрацию по столбцам, коррелирующим со столбцами кластеризации, BigQuery сможет читать меньше данных!

Предположим, что у вас есть таблица с двумя столбцами: `zip_code` (почтовый индекс) и `state` (штат). Она кластеризована по столбцу `zip_code`, то есть на диске данные сортируются по `zip_code`. В Соединенных Штатах существует прямая связь между штатами и почтовыми индексами, поскольку диапазоны почтовых индексов назначаются по географическому местоположению (от 00000 на северо-востоке до 99999 на северо-западе). Если выполнить запрос, фильтрующий записи по названию штата и не использующий в фильтре столбец кластеризации, BigQuery пропустит все блоки данных, не включающие искомый штат, и в итоге вы заплатите за запрос меньше.

При работе с кластеризованными таблицами BigQuery может применять ряд оптимизаций производительности, которые невозможны для некластеризованных таблиц. Одна из этих оптимизаций предназначена для *схем типа «звезда»*, которые позволяют выполнять фильтрацию на основе ограничений в размерной таблице. Например, у вас есть таблица фактов `orders`, содержащая заказы,

сгруппированные по столбцу `customer_id`, и размерная таблица, содержащая столбец `customers`, и вы выполняете следующий запрос:

```
SELECT o.*
FROM orders o
JOIN customers c USING (customer_id)
WHERE c.name = "Changying Bao"
```

Обычно подобные запросы сканируют всю таблицу, чтобы выполнить соединение. Но поскольку таблица `orders` кластеризована по столбцу, который используется для соединения, правая сторона соединения — запрос клиентов и поиск соответствующих идентификаторов — выполняется первой. После этого второй части запроса остается только найти в столбце кластеризации соответствие `customer_id`. С точки зрения низкоуровневого сканирования BigQuery параллельно выполняет следующие запросы:

```
// Сначала выполняется поиск идентификатора клиента.
// При этом сканируется лишь малая часть размерной таблицы
SET id = SELECT customer_id FROM customers
WHERE c.name = "Changying Bao"

// Затем выполняется поиск клиента в таблице orders.
// Фильтрация выполняется по столбцу кластеризации,
// и поэтому читается лишь малая часть данных.
SELECT * FROM orders WHERE customer_id=$id ;
```

Проще говоря, мы настоятельно советуем применять кластеризацию, если вы хотите снизить затраты на выполнение запросов и повысить производительность.

Случаи использования, нечувствительные ко времени

BigQuery стремится минимизировать время, необходимое для получения информации из данных. Как результат, мы можем проводить интерактивный анализ больших наборов данных и передавать в них обновления практически в режиме реального времени. Иногда, однако, требования к скорости выполнения могут быть не особенно жесткими. Примером могут служить ночные отчеты. В таких случаях иногда желательно ставить запросы в очередь и выполнять их, когда это возможно, и часто вполне достаточно того, чтобы отчеты отражали данные по состоянию на час назад.

Пакетные запросы

Вы можете отправить в сервис наборы запросов, которые называют пакетными запросами. Они будут поставлены в очередь от вашего имени и выполняться,

когда будут доступны свободные ресурсы. В 2012 году, когда впервые появилась их поддержка (<https://developers.googleblog.com/2012/08/now-in-bigquery-batch-queries-and.html>), преимущество пакетных запросов заключалось в их стоимости. Однако на момент написания этой книги интерактивные и пакетные запросы стоили одинаково. Если вы используете тариф с фиксированной оплатой, для выполнения пакетных и интерактивных запросов будут использоваться выделенные вам слоты.

Поскольку пакетные запросы не менее дороги и используют те же забронированные ресурсы, что и интерактивные запросы, основная причина, по которой может возникнуть желание использовать пакетные запросы, заключается в том, что они не учитывают ограничения параллелизма и могут упростить планирование сотен запросов.

Существует ряд ограничений, влияющих на параллельное выполнение интерактивных (не пакетных) запросов. Например, параллельно может выполняться не более 50 запросов,¹ и вместе с тем ограничивается количество параллельно извлекаемых байтов и количество «больших запросов». Если эти пределы достигнуты, запрос тут же завершится ошибкой. BigQuery предполагает, что интерактивный запрос — это то, что необходимо выполнить сразу. В случае пакетных запросов, как только будет достигнуто ограничение скорости, BigQuery поставит эти запросы в очередь и позже повторит попытку выполнить их. Для пакетных запросов существуют свои, похожие ограничения, но они устанавливаются отдельно от ограничений для интерактивных запросов, поэтому пакетные запросы не будут влиять на интерактивные.

Одним из примеров могут служить периодические запросы, запускаемые ежедневно или ежечасно с целью сбора информации для панелей мониторинга. Допустим, у вас есть 500 запросов, которые нужно выполнить. Если попытаться запустить их одновременно в виде интерактивных запросов, часть из них выдаст ошибку из-за ограничений на число одновременно выполняемых запросов. Кроме того, совершенно необязательно, чтобы эти запросы конкурировали с другими запросами, которые выполняются вручную из веб-интерфейса BigQuery. Поэтому такие запросы можно запустить с пакетным приоритетом, а остальные запросы выполнять как интерактивные.

Для запуска запросов в пакетном режиме добавьте флаг `--batch` при вызове инструмента командной строки `bq` или укажите приоритет `BATCH` задания в консоли или REST API вместо `INTERACTIVE`. Если BigQuery не запустит запрос в течение 24 часов, она изменит приоритет задания на интерактивный. Однако обычно время ожидания перед запуском запроса составляет не более нескольких минут, если вы не отправляете запросов больше, чем определяет ваша квота на число

¹ Это значение по умолчанию, но вы можете запросить увеличение квоты. См. <https://cloud.google.com/bigquery/quotas>.

одновременно выполняемых запросов. Последующие запросы будут выполняться после завершения предыдущих.

Загрузка файлов

Если вы хотите свести к минимуму «время анализа» или получить максимально простой конвейер данных, мы настоятельно рекомендуем использовать потоковые вставки в BigQuery через Cloud Pub/Sub и Cloud Dataflow. Эта архитектура обеспечивает скорость загрузки порядка 100 000 записей в секунду даже в небольших кластерах Cloud Dataflow. Кластер Dataflow можно масштабировать по горизонтали, добавлять больше компьютеров и достигать скорости загрузки в несколько миллионов записей в секунду без каких-либо дополнительных настроек.¹

Потоковая загрузка оплачивается отдельно, тогда как задания загрузки выполняются бесплатно. В некоторых сценариях задержка в несколько минут является вполне приемлемой ценой за снижение затрат на передачу данных. Поэтому подумайте о возможности использовать загрузку файлов вместо потоковых вставок. Это можно сделать с помощью BigQueryIO из Apache Beam, например:

```
BigQueryIO.writeTableRows()  
    .to("project-id:dataset-id.table-id")  
    .withCreateDisposition(  
        BigQueryIO.Write.CreateDisposition.CREATE_IF_NEEDED)  
    .withMethod(Method.FILE_LOADS)  
    .withTriggeringFrequency(Duration.standardSeconds(600))  
    .withNumFileShards(10)  
    .withSchema(new TableSchema(...))  
    .withoutValidation()
```

Этот фрагмент записывает новые данные в BigQuery каждые 10 минут, используя механизм загрузки файлов, что позволяет избежать затрат на загрузку потокового контента. В кластерах Dataflow среднего размера загрузка файлов может достигать скорости 300 000 записей в секунду. Однако вы должны знать, что оконные вычисления требуют времени, поэтому задержка может составить порядка нескольких минут. Из-за квот загрузки для каждой таблицы и проекта, а также из-за того, что в квотах учитываются ошибки и повторные попытки, мы советуем выполнять загрузку файлов не чаще чем раз в пять минут.

¹ Поскольку потоковая передача — платная, вам может потребоваться увеличить квоту. Проверьте квоты по умолчанию на странице https://cloud.google.com/bigquery/quotas#streaming_inserts и запросите дополнительную квоту на потоковую передачу в консоли GCP.

Выводы

В этой главе мы рассмотрели способы управления затратами с использованием механизма пробных прогонов `dry_run` и с установкой ограничений на количество оплачиваемых байтов. Затем мы рассмотрели способы измерения скорости запросов с помощью REST API и специализированного инструмента.

Мы также рассмотрели несколько способов увеличения скорости выполнения запросов. Чтобы свести к минимуму накладные расходы на ввод/вывод, мы предложили способы уменьшения объема читаемых данных. Мы также рассмотрели приемы кеширования, от результатов предыдущих запросов до промежуточных результатов и целых таблиц. Исследовали способы эффективного соединения таблиц, исключения самосоединения, сокращения объемов соединяемых данных, в том числе за счет использования предварительно вычисленных значений. Ограничивая большие сортировки, защищаясь от асимметрии данных и оптимизируя пользовательские функции, можно минимизировать вероятность перегрузки слотов. Наконец, мы рекомендовали использовать аппроксимирующие функции агрегирования, включая функции подсчета и определения наибольших значений.

В заключение мы рассмотрели способы оптимизации хранения и ускорения доступа к данным. Мы предложили реорганизовать приложения для приема сжатых и неполных ответов, отправлять запросы в пакетном режиме и выполнять массовое считывание данных с использованием Storage API. Мы обнаружили, что хранение данных в виде массивов структур и географических типов помогает повысить производительность. Мы также познакомились с некоторыми способами уменьшения объема сканируемых данных, в частности, путем сегментирования и кластеризации.

Контрольный список

Мы хотели бы завершить эту главу той же рекомендацией, с которой начали: описываемые здесь приемы могут давать существенное повышение производительности, но вы обязательно должны убедиться, что они применимы к вашему рабочему процессу. Следуйте этому контрольному списку, если вам покажется, что ваш запрос работает медленно:

Проблема	Возможные решения
Выполняется самосоединение	Используйте функции агрегирования, чтобы исключить возможность самосоединения больших таблиц. Используйте оконные (аналитические) функции для вычисления самозависимых отношений
Используются инструкции DML	Объедините в пакеты инструкции DML (INSERT, UPDATE, DELETE)

Проблема	Возможные решения
Соединение выполняется слишком медленно	<p>Уменьшите объем данных, участвующих в соединении.</p> <p>Возможно, поможет денормализация данных.</p> <p>Используйте вложенные повторяющиеся поля</p>
Запросы вызываются повторно	<p>Воспользуйтесь возможностью кеширования результатов запросов.</p> <p>Материализуйте результаты предыдущих запросов в таблицы</p>
Происходит исчерпание ресурсов рабочих серверов	<p>Ограничьте объем сортировки с применением оконных функций.</p> <p>Проверьте асимметричность данных.</p> <p>Оптимизируйте пользовательские функции</p>
Используются функции COUNT, TOP, DISTINCT	<p>Подумайте о возможности использования аппроксимирующих версий этих функций</p>
Медленный ввод/вывод	<p>Уменьшите нагрузку на сеть.</p> <p>Попробуйте использовать соединение для уменьшения размера таблицы.</p> <p>Выберите более эффективный формат хранения.</p> <p>Секционировать таблицы.</p> <p>Кластеризуйте таблицы</p>

Продвинутые запросы

В главах 2 и 3 мы рассмотрели основы запросов на Standard SQL и типов данных, поддерживаемых в BigQuery. Исходный код парсера и анализатора диалекта Standard SQL, поддерживаемого в BigQuery, был открыт как ZetaSQL (<https://github.com/google/zetasql>). Парсер и анализатор ZetaSQL используются для обеспечения постоянного функционирования, проверки и неявного приведения типов, разрешения имен и многого другого во всех продуктах Google Cloud Platform (GCP), которые поддерживают SQL (например, Cloud Spanner и Cloud Dataflow). Однако эти механизмы запросов могут поддерживать не все возможности языка ZetaSQL. Например, на момент написания этой книги сервис BigQuery не поддерживал транзакции с несколькими состояниями. Также на момент написания этой книги Cloud Dataflow не поддерживал географические запросы, но когда такая поддержка появится, запросы GIS SQL и географические типы в Cloud Dataflow будут работать так же, как в BigQuery.

В этой главе мы рассмотрим возможности, типы данных и функции ZetaSQL, поддерживаемые в BigQuery, которые выходят за рамки Standard SQL или могут быть незнакомы многим аналитикам. Для начала мы рассмотрим синтаксис параметризованных запросов и пользовательских функций, поддерживающих возможность повторного использования. Затем углубимся в синтаксис SQL, включая массивы, окна, метаданные таблиц, а также инструкции определения данных и работы с ними. Мы расскажем о поддержке сценариев и хранимых процедур в BigQuery, а в конце главы рассмотрим геоинформационные системы, статистические функции и функции шифрования.

Многократные запросы

BigQuery поддерживает возможность повторного использования запросов и их частей. Мы можем параметризовать запросы и выделять часто используемый код в функции, подзапросы или предложения WITH. Рассмотрим каждую из этих возможностей по порядку.

Параметризованные запросы

Поддержка параметризованных запросов позволяет задать запрос с учетом параметров, которые определяются во время выполнения. Благодаря этому один и тот же запрос можно использовать в разных контекстах без необходимости форматировать строки для создания запроса во время выполнения.¹

На момент написания этой книги веб-интерфейс BigQuery не поддерживал параметризованные запросы. В главе 4 мы рассмотрели приемы выполнения параметризованных запросов с использованием REST API и запросы из Jupyter Magics. В этой главе мы покажем, как выполнять параметризованные запросы с использованием Python Cloud Client API.

Именованные параметры

Вернемся к нашему набору данных с информацией о прокате велосипедов в Лондоне и предположим, что нам часто требуется вычислять количество поездок, начинающихся в отдельных пунктах проката, продолжительность которых находится в пределах определенного диапазона, но при этом название пункта проката и значения продолжительности могут меняться. Для решения этой задачи можно определить параметризованный запрос, принимающий название пункта проката и пороговые значения продолжительности в виде параметров:²

```
query = """
SELECT
    start_station_name
    , AVG(duration) as avg_duration
FROM
    `bigquery-public-data`.london_bicycles.cycle_hire
WHERE
    start_station_name LIKE CONCAT('%', @STATION, '%')
    AND duration BETWEEN @MIN_DURATION AND @MAX_DURATION
GROUP BY start_station_name
"""
```

Запрос имеет именованные параметры, начинающиеся с символа @, но обратите внимание, как параметр @STATION используется в LIKE. Если бы мы просто написали '%@STATION%', BigQuery интерпретировала бы символ @ буквально из-за одинарных кавычек. Поэтому параметр указывается отдельно и объединяется со строками, содержащими подстановочный знак %.

Во время выполнения именованные параметры заменяются фактическими значениями, каждое из которых определяется с помощью соответствующего имени

¹ Конструирование запросов с помощью операции форматирования строк открывает возможность для атак вида «инъекция SQL» (см. главу 5).

² См. сценарий 08_advqueries/param_named.py в репозитории GitHub с примерами для этой книги (<https://github.com/GoogleCloudPlatform/bigquery-oreilly-book>).

параметра и типа данных SQL (здесь мы используем Python Cloud Client API для BigQuery, поэтому это код на Python):

```
query_params = [
    bigquery.ScalarQueryParameter(
        "STATION", "STRING", station_name),
    bigquery.ScalarQueryParameter(
        "MIN_DURATION", "FLOAT64", min_duration),
    bigquery.ScalarQueryParameter(
        "MAX_DURATION", "FLOAT64", max_duration),
]
```

Чтобы выполнить запрос, нужно создать задание, передать параметры запроса, запустить запрос и проанализировать результаты:

```
job_config = bigquery.QueryJobConfig()
job_config.query_parameters = query_params
query_job = client.query(
    query,
    location="EU",
    job_config=job_config,
)
for row in query_job:
    print("{}: \t{}".format(
        row.start_station_name, row.avg_duration))
```

Предыдущий код можно завернуть в функцию на Python, которая принимает параметры запроса в аргументах:

```
def print_query_results(client,
                        station_name,
                        min_duration=0,
                        max_duration=84000):
```

Эту функцию можно вызывать многократно с разными значениями параметров:

```
client = bigquery.Client()
print_query_results(client, 'Kennington', 300)
print_query_results(client, 'Hyde Park', 600, 6000)
```

Этот код выведет следующее:

```
Kennington between 300 and 84000
Kennington Oval, Oval: 1269.0798128928543
Doddington Grove, Kennington: 1243.7377963737788
Kennington Road Post Office, Oval: 1360.2854550952536
Kennington Lane Rail Bridge, Vauxhall: 991.4344845855808
Cleaver Street, Kennington: 1075.6050140700947
Kennington Cross, Kennington: 996.2538654101008
Kennington Road, Vauxhall: 1228.6673653660118
Cotton Garden Estate, Kennington: 996.7003600110778
Kennington Lane Tesco, Vauxhall: 929.6523615439942
```


Kennington Station, Kennington: 1238.4088412072647

Hyde Park between 600 and 6000

Bayswater Road, Hyde Park: 1614.2670577732417

Wellington Arch, Hyde Park: 1828.9651324965134

Hyde Park Corner, Hyde Park: 2120.4145144213744

Cumberland Gate, Hyde Park: 1899.3282223532708

Speakers' Corner 1, Hyde Park: 2070.2458069837776

Triangle Car Park, Hyde Park: 1815.661582196573

Albert Gate, Hyde Park: 1897.9349474341027

Knightsbridge, Hyde Park: 1963.0815096317635

Serpentine Car Park, Hyde Park: 1688.0595490490423

Park Lane, Hyde Park: 2055.451932776309

Speakers' Corner 2, Hyde Park: 2093.6202531645563

Во многих движках SQL параметризованные запросы предварительно компилируются и за счет этого способны обеспечить улучшенную производительность. В BigQuery это не так. Как показывает предыдущий фрагмент кода, запрос и настройки задания с параметрами запроса передаются в сервис BigQuery одновременно (в вызове `client.query`). Соответственно, производительность параметризованного запроса ничем не отличается от производительности статического запроса с жестко заданными значениями параметров.

И все же использование параметров в BigQuery дает определенные преимущества. Одним из таких преимуществ является безопасность и предотвращение атак вида «инъекция SQL». Генерируя запрос и добавляя в него фильтр по значению, указанному пользователем, вы должны убедиться, что в пользовательском значении отсутствуют специальные символы. Иногда очень трудно обеспечить «безопасность» сгенерированных запросов, когда они включают идентификаторы, указанные пользователем. Параметризация дает простой способ защиты запросов; этот прием гарантирует невозможность атак «инъекция SQL», независимо от значений, указанных в параметрах.

Именованные параметры с отметками времени

В предыдущем запросе использовались строковые и числовые параметры. Однако точно так же в запросе можно использовать параметры с отметками времени, при условии, что в коде на Python (или Go, или любом другом языке) будет использован соответствующий тип данных. В Python мы должны использовать тип `datetime.datetime`. То есть чтобы найти среднюю продолжительность поездок в течение определенного часа, нужно сначала импортировать необходимые библиотеки для Python:

```
from google.cloud import bigquery
from datetime import datetime
from datetime import timedelta
import pytz
```

Затем вычислить значения параметров:

```
def print_query_results(client, mid_time):
    start_time = mid_time - timedelta(minutes=30)
    end_time = mid_time + timedelta(minutes=30)
```

Настроить параметры запроса с этими значениями:

```
query = """
    SELECT
        AVG(duration) as avg_duration
    FROM
        `bigquery-public-data`.london_bicycles.cycle_hire
    WHERE
        start_date BETWEEN @START_TIME AND @END_TIME
    """
query_params = [
    bigquery.ScalarQueryParameter(
        "START_TIME", "TIMESTAMP", start_time),
    bigquery.ScalarQueryParameter(
        "END_TIME", "TIMESTAMP", end_time),
]
job_config = bigquery.QueryJobConfig()
job_config.query_parameters = query_params
query_job = client.query(
    query,
    location="EU",
    job_config=job_config,
)
for row in query_job:
    print(row.avg_duration)
print("_____")
```

И наконец, вызвать функцию, передав ей объект `datetime`, например, с датой Рождества в 2015 году:

```
client = bigquery.Client()
print_query_results(client,
    datetime(2015, 12, 25, 15, 0, tzinfo=pytz.UTC))
```

Этот вызов вернет среднюю продолжительность поездок между 14:30 и 15:30 25 декабря 2015 года, которая составляет 3658.5000000000005 секунды.



Запросы, выполняемые по расписанию, автоматически получают два параметра при вызове: `@run_time`, типа `TIMESTAMP`; и `@run_date`, типа `DATA`. То есть любой запрос, выполняемый по расписанию, можно параметризовать этими двумя параметрами.

Позиционные параметры

Несмотря на то что BigQuery поддерживает позиционные параметры, мы настоятельно рекомендуем использовать в запросах только именованные параметры,

потому что они улучшают читаемость как самих запросов, так и вызывающего кода.¹ Позиционные параметры указываются с помощью ? и обязательно должны передаваться по порядку.²

```
def print_query_results(client, params):
    query = """
        SELECT
            start_station_name
            , AVG(duration) as avg_duration
        FROM
            `bigquery-public-data`.london_bicycles.cycle_hire
        WHERE
            start_station_name LIKE CONCAT('%', ?, '%')
            AND duration BETWEEN ? AND ?
        GROUP BY start_station_name
    """
    query_params = [
        bigquery.ScalarQueryParameter(
            None, "STRING", params[0]),
        bigquery.ScalarQueryParameter(
            None, "FLOAT64", params[1]),
        bigquery.ScalarQueryParameter(
            None, "FLOAT64", params[2]),
    ]
```

Параметры с массивами и структурами

До сих пор мы демонстрировали примеры *скалярных параметров запроса*. Однако BigQuery поддерживает также *параметры с массивами*. Представьте, что у нас есть пользовательский интерфейс, отображающий количество поездок от пунктов проката, которые пользователь может выбрать интерактивно. Для этого нужно, чтобы набор идентификаторов пунктов проката, выбранных пользователем, был предоставлен в запросе в виде массива.

Вот как можно узнать количество поездок, начинающихся в пунктах проката, выбранных пользователем и перечисленных в параметре-массиве @STATIONS, используя функции IN и UNNEST:³

```
query = """
    SELECT
        start_station_id
```

¹ Наличие поддержки позиционных параметров обусловлено тем, что драйверы Open Database Connectivity и Java Database Connectivity ожидают, что механизмы запросов будут обозначать параметры знаком вопроса.

² См. сценарий 08_advqueries/param_positional.py в репозитории GitHub с примерами для этой книги.

³ См. сценарий 08_advqueries/param_array.py в репозитории GitHub с примерами для этой книги.

```

        , COUNT(*) as num_trips
FROM
`bigquery-public-data`.london_bicycles.cycle_hire
WHERE
    start_station_id IN UNNEST(@STATIONS)
    AND duration BETWEEN @MIN_DURATION AND @MAX_DURATION
GROUP BY start_station_id
"""
query_params = [
    bigquery.ArrayQueryParameter(
        'STATIONS', "INT64", ids),
    bigquery.ScalarQueryParameter(
        'MIN_DURATION', "FLOAT64", min_duration),
    bigquery.ScalarQueryParameter(
        'MAX_DURATION', "FLOAT64", max_duration),
]

```

Теперь передадим массив идентификаторов пунктов проката в параметре `stations`:

```
print_query_results(client, [270, 235, 62, 149], 300, 600)
```

И получим в результате число поездок, начавшихся в каждом из указанных пунктов проката:

```

270:    26400
149:    4143
235:    8337
62:     5954

```

Точно так же в запрос можно передать структуру, создав параметр структуры:

```

bigquery.StructQueryParameter(
    "bicycle_trip",
    bigquery.ScalarQueryParameter("start_station_id", "INT64", 62),
    bigquery.ScalarQueryParameter("end_station_id", "INT64", 421),
)

```

Обратите внимание, что нельзя параметризовать имена таблиц и столбцов и другие части самого запроса. Если запустить параметризованный запрос с параметром `--dry_run` и не передать обязательные параметры, ответ будет включать типы всех параметров, определяемые автоматически.

Пользовательские функции SQL

Кроме повторного использования целых запросов путем их параметризации, можно организовать повторное использование более мелких частей запросов. Например, можно повторно использовать набор операций, определив его в виде пользовательской функции (User-Defined Function, UDF).

Предположим, что вы решили узнать количество ночных поездок по дням недели (здесь имеются в виду поездки, начало и конец которых приходятся на разные дни). Такой запрос требует манипуляций с датами. Такие манипуляции удобно определить как временные пользовательские функции SQL:

```
CREATE TEMPORARY FUNCTION dayOfWeek(x TIMESTAMP) AS
(
  ['Sun', 'Mon', 'Tue', 'Wed', 'Thu', 'Fri', 'Sat']
  [ORDINAL(EXTRACT(DAYOFWEEK from x))])
);
CREATE TEMPORARY FUNCTION getDate(x TIMESTAMP) AS
(
  EXTRACT(DATE FROM x)
);
```

Определив эти функции, вы сможете использовать их в предложении **WITH** для поиска ночных поездок:

```
WITH overnight_trips AS (
  SELECT
    duration
    , dayOfWeek(start_date) AS start_day
  FROM
    `bigquery-public-data`.london_bicycles.cycle_hire
  WHERE
    getDate(start_date) != getDate(end_date)
)
```

Чтобы узнать количество ночных поездок по дням недели, можно сгруппировать результаты предыдущего табличного выражения по дням недели:

```
SELECT
  start_day
  , COUNT(*) AS num_overnight_rentals
  , AVG(duration)/3600 AS avg_duration_hours
FROM
  overnight_trips
GROUP BY
  start_day
ORDER BY num_overnight_rentals DESC
```

Как показывают результаты, наибольшее количество ночных поездок приходится на субботу и пятницу, а наименьшее — на начало рабочей недели:

Row	start_day	num_overnight_rentals	avg_duration_hours
1	Sat	28095	9.13462063237824
2	Fri	23746	8.772040203262295
3	Thu	18153	9.792348372169885
4	Sun	16648	13.834484622777499

Row	start_day	num_overnight_rentals	avg_duration_hours
5	Wed	15571	10.848297047930972
6	Mon	12507	10.729399536259686
7	Tue	12461	9.430337319102266

Сохраняемые пользовательские функции

Пользовательские функции в предыдущих разделах были определены как временные, поэтому их можно применять только в рамках одного запроса. Чтобы использовать их в других запросах, вам придется скопировать и вставить определения этих функций. Естественно, это довольно неудобно и чревато ошибками.

Если у вас есть функция, которую вы хотите использовать в разных запросах, ее лучше сохранить в наборе данных, а затем ссылаться на нее по мере необходимости (создайте свой набор данных с именем `ch08eu` в регионе EU, прежде чем опробовать следующий запрос):

```
CREATE OR REPLACE FUNCTION ch08eu.dayOfWeek(x TIMESTAMP) AS
(
  ['Sun', 'Mon', 'Tue', 'Wed', 'Thu', 'Fri', 'Sat']
  [ORDINAL(EXTRACT(DAYOFWEEK from x))])
);
```

Так же как в случае с таблицами, выбор между `CREATE FUNCTION`, `CREATE OR REPLACE FUNCTION` и `CREATE FUNCTION IF NOT EXISTS` зависит от желаемого функционирования операции сохранения функции, если эта функция уже существует: ошибка, замена существующей функции или по-ор¹ соответственно. Так же как в случае с таблицами, функцию можно удалить из набора данных с помощью `DROP FUNCTION`. Эта возможность доступна не только в SQL, но также в REST API, при использовании инструмента командной строки `bq` или клиентских библиотек.

После сохранения функции в наборе данных ее можно использовать в запросах, сославшись на проект и набор данных, в котором находится функция (так же, как в случае с именами таблиц в запросах, по умолчанию подразумевается текущий активный проект):

```
WITH overnight_trips AS (
  SELECT
    duration
    , ch08eu.dayOfWeek(start_date) AS start_day
  FROM
    `bigquery-public-data`.london_bicycles.cycle_hire
  ...
```

¹ по-ор, сокращенно от англ. «no operation» (ничего не делать) — так обозначаются операции, которые ничего не делают. См. [https://en.wikipedia.org/wiki/NOP_\(code\)](https://en.wikipedia.org/wiki/NOP_(code)). На русском языке: <https://ru.wikipedia.org/wiki/NOP>.

Определение функции можно увидеть в веб-интерфейсе BigQuery, как показано на рис. 8.1.

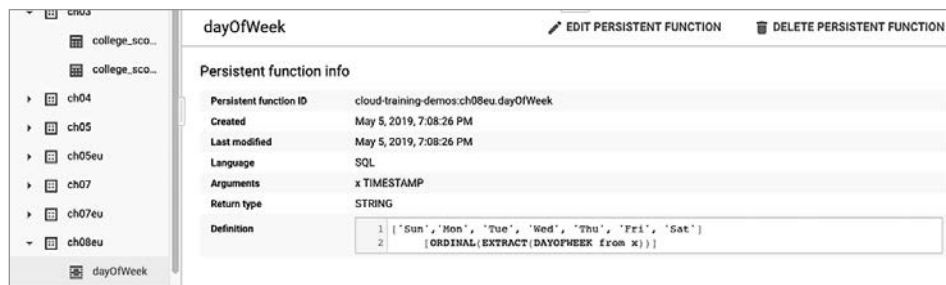


Рис. 8.1. Выберите хранимую функцию в веб-интерфейсе BigQuery, чтобы увидеть/отредактировать ее определение

Разрешения на доступ к пользовательским функциям, так же как и для таблиц, хранятся на уровне набора данных и ведут себя аналогично. Например, разрешение `bigquery.routines.list` позволяет владельцам получить список функций в наборе данных, а разрешение `bigquery.routines.[create/get/update/delete]` — создавать, вызывать, изменять или удалять функции.

Общедоступные пользовательские функции

Кроме всего прочего, вы можете определить пользовательские функции в наборе данных с разрешениями `allAuthenticatedUsers`, чтобы расширить возможности BigQuery. Например, в BigQuery есть встроенная функция `AVG`, но отсутствует функция `MEDIAN`. Наши коллеги Эллиот Броссар (Elliott Brossard) и Фелипе Хоффа (Felipe Hoffa) определили UDF SQL для вычисления медианы, позаботившись о правильном определении медианы для массивов с четным и нечетным числом элементов:¹

```
CREATE OR REPLACE FUNCTION fhoffa.x.median (arr ANY TYPE) AS ((
  SELECT IF (MOD(ARRAY_LENGTH(arr), 2) = 0,
    ( arr[OFFSET(DIV(ARRAY_LENGTH(arr), 2) - 1)] +
      arr[OFFSET(DIV(ARRAY_LENGTH(arr), 2))] ) / 2,
    arr[OFFSET(DIV(ARRAY_LENGTH(arr), 2))])
  )
  FROM (SELECT ARRAY_AGG(x ORDER BY x) AS arr FROM UNNEST(arr) AS x)
));
```

¹ Массивы с нечетным числом элементов имеют один элемент посередине и равное количество элементов с каждой стороны. Если массив имеет четное количество элементов, в середине находятся два элемента, соответственно, медиана определяется как среднее этих двух средних элементов.

Поскольку набор данных `x` в проекте `fhoffa` является общедоступным, вы можете воспользоваться этой функцией, чтобы найти пункты проката с наибольшей средней продолжительностью поездок:

```
SELECT
  start_station_name
  , COUNT(*) AS num_trips
  , fhoffa.x.median(ARRAY_AGG(tripduration)) AS typical_duration
FROM `bigquery-public-data`.new_york_citibike.citibike_trips
GROUP BY start_station_name
HAVING num_trips > 1000
ORDER BY typical_duration DESC
LIMIT 5
```

Вот какие результаты мы получили:

Row	start_station_name	num_trips	typical_duration
1	Mobile 01	1039	1697.0
2	Soissons Landing	18484	1525.0
3	Yankee Ferry Terminal	18013	1496.0
4	Central Park North & Adam Clayton Powell Blvd	54465	1419.0
5	Broadway & Moylan Pl	6121	1413.0

Почему для иллюстрации использования функции вычисления медианы мы использовали набор данных с информацией о прокате велосипедов для Нью-Йорка, а не для Лондона? Дело в том, что все наборы данных, используемые в запросе, должны находиться в одном месте; поскольку набор данных `fhoffa.x` находится в Соединенных Штатах, мы были вынуждены использовать набор данных с информацией о прокате велосипедов, также хранящийся в США. Мы не смогли бы использовать лондонский набор данных.

Коллекция таких пользовательских функций с открытым исходным кодом, разработанных сообществом, доступна по адресу <https://github.com/GoogleCloudPlatform/bigquery-utils> и синхронизирована с общедоступным набором данных `bqutil.fn`. Функция вычисления медианы Эллиота и Фелипе является частью этой коллекции и набора данных, поэтому ее можно использовать, как показано ниже (попробуйте сами):

```
SELECT
  start_station_name
  , COUNT(*) AS num_trips
  , bqutil.fn.median(ARRAY_AGG(tripduration)) AS typical_duration
FROM `bigquery-public-data`.new_york_citibike.citibike_trips
GROUP BY start_station_name
HAVING num_trips > 1000
ORDER BY typical_duration DESC
LIMIT 5
```


Даже если у вас нет возможности использовать общедоступные пользовательские функции, вы все равно можете пользоваться преимуществами повторного использования кода, поместив нужные вам пользовательские функции в набор данных своей компании.

Повторное использование частей запросов

Представьте, что вам нужно найти дни с наибольшим числом особенно длинных поездок в оба конца. Под «особенно длинными» в данном случае подразумеваются поездки, продолжающиеся дольше обычных более чем в два раза. Эта задача решается в три шага:

1. Выбрать поездки в оба конца.
2. Вычислить среднюю продолжительность таких поездок для каждого пункта проката.
3. Определить дни с наибольшим числом поездок, длительность которых превышает среднюю продолжительность более чем в два раза.

Коррелированные подзапросы

Написать необходимый запрос можно в порядке «наоборот»:

```
SELECT
  start_date,
  COUNT(*) AS num_long_trips
FROM -- "первое предложение FROM"
  (SELECT
    start_station_name
    , duration
    , EXTRACT(DATE from start_date) AS start_date
  FROM
    `bigquery-public-data`.london_bicycles.cycle_hire
  WHERE
    start_station_name = end_station_name) AS roundtrips
WHERE -- "внешнее предложение WHERE"
  duration > 2*(
    SELECT
      AVG(duration) as avg_duration
    FROM
      `bigquery-public-data`.london_bicycles.cycle_hire
    WHERE
      start_station_name = end_station_name
      AND roundtrips.start_station_name = start_station_name
  )
GROUP BY start_date
ORDER BY num_long_trips DESC
LIMIT 5
```

Первое предложение **FROM** состоит из подзапроса, который выбирает необходимые столбцы из набора данных и извлекает дату из столбца с отметкой времени. Внешнее предложение **WHERE** сравнивает текущую продолжительность поездки с удвоенной средней продолжительностью, где сама средняя продолжительность вычисляется с использованием подзапроса.

Второй подзапрос является примером коррелированного подзапроса — подзапроса, который использует значения (в данном случае **start_station_name**, **duration** и **end_station_name**) из внешнего запроса. Здесь второй подзапрос должен вычислить среднюю продолжительность поездок, начинающихся и заканчивающихся в одном и том же пункте проката.

Предложение **WITH**

Предложение **WITH** позволяет повторно использовать табличные выражения и делать запросы более читабельными. Запрос из предыдущего раздела можно переписать с использованием двух предложений **WITH**:

```
WITH roundtrips AS (  
  SELECT  
    start_station_name  
    , duration  
    , EXTRACT(DATE from start_date) AS start_date  
  FROM  
    `bigquery-public-data`.london_bicycles.cycle_hire  
  WHERE  
    start_station_name = end_station_name  
)  
,  
station_avg AS (  
  SELECT  
    start_station_name  
    , AVG(duration) as avg_duration  
  FROM  
    roundtrips  
  GROUP BY start_station_name  
)
```

Эти два запроса упрощают определение третьего, производящего вычисления:

```
SELECT  
  start_date,  
  COUNT(*) AS num_long_trips  
FROM  
  roundtrips  
JOIN station_avg USING(start_station_name)  
WHERE duration > 2*avg_duration  
GROUP BY start_date  
ORDER BY num_long_trips DESC  
LIMIT 5
```

Stages		Wait		Read	
✓	S00: Input ^	Avg:	<div><div></div></div> 51 ms	<div><div></div></div> 144 ms	
		Max:	<div><div></div></div> 113 ms	<div><div></div></div> 262 ms	
	READ	\$10:duration, \$12:start_station_name, \$11:end_station_name FROM bigquery-public-data.london_bicycles.cycle_hire WHERE equal(\$12, \$11)			
	AGGREGATE	GROUP BY \$70 := \$12 \$60 := SHARD_AVG(\$10)			
	WRITE	\$70, \$60 TO __stage00_output BY HASH(\$70)			
✓	S01: Aggregate v	Avg:	<div><div></div></div> 1 ms	<div><div></div></div> 0 ms	
		Max:	<div><div></div></div> 1 ms	<div><div></div></div> 0 ms	
✓	S03: Coalesce v	Avg:	<div><div></div></div> 0 ms	<div><div></div></div> 0 ms	
		Max:	<div><div></div></div> 2 ms	<div><div></div></div> 0 ms	
✓	S04: Join+ ^	Avg:	<div><div></div></div> 0 ms	<div><div></div></div> 236 ms	
		Max:	<div><div></div></div> 1 ms	<div><div></div></div> 318 ms	
	READ	\$1:duration, \$3:start_date, \$4:start_station_name, \$2:end_station_name FROM bigquery-public-data.london_bicycles.cycle_hire WHERE equal(\$4, \$2)			
	READ	\$50, \$80 FROM __stage03_output			

Рис. 8.2. План выполнения запроса наглядно показывает, что чтение данных происходит дважды

Обратите внимание, как значение `roundtrips` повторно используется в предложении `WITH station_avg`, а также в основном запросе. Однако мы должны предупредить вас, что повторное использование предложений `WITH` улучшает только удобочитаемость — фактические данные не обязательно будут кешироваться и использоваться повторно. Действительно, заглянув в план выполнения запроса, можно увидеть, что этапы 0 и 4 начинаются с инструкции чтения (`READ`) из набора данных `bigquery-public-data` и в первом случае извлекаются только три столбца, а во втором — четыре, как показано на рис. 8.2.

Результаты включают две даты, приходящиеся на Рождество:

Row	start_date	num_long_trips	a)
1	2016-12-25	740	b)
2	2016-05-08	714	c)
3	2017-04-09	667	d)
4	2015-08-01	663	e)
5	2015-12-25	660	f)

Предложение `WITH` обеспечивает возможность повторного использования только внутри запроса. Чтобы повторно использовать набор результатов в нескольких запросах, можно создать промежуточную таблицу или материализованное представление. Дополнительные затраты на хранение в этом случае могут компенсироваться более быстрыми и менее дорогостоящими вычислениями, но имейте в виду, что вы должны проверить и убедиться, что это действительно так. Предложение `WITH` будет работать быстрее, если промежуточная таблица больше исходной. Подробнее об этом рассказывается в главе 7.

Определение констант

Предложение `WITH` можно использовать для хранения констант и как единственное место, где их можно изменить. Например, представьте, что нужно найти пункты проката, для которых наибольшее количество поездок превышает определенную продолжительность:

```
WITH params AS (  
  SELECT 600 AS DURATION_THRESH  
)  
SELECT  
  start_station_name  
  , COUNT(duration) as num_trips  
FROM  
  `bigquery-public-data`.london_bicycles.cycle_hire  
  , params  
WHERE duration >= DURATION_THRESH  
GROUP BY start_station_name  
ORDER BY num_trips DESC  
LIMIT 5
```

Определяя порог продолжительности в предложении `WITH`, можно организовать удобное место для сбора всех «магических» чисел, используемых в запросе, и определения удобочитаемых имен для них. Обратите внимание, что предложение `FROM` в предыдущем запросе включает параметры, благодаря чему появляется возможность использовать в запросе константу `DURATION_THRESH`.

Другой способ выразить константы — использовать синтаксис сценариев и объявить их как переменные. Мы рассмотрим этот способ далее в этой главе.

НЕОДНОЗНАЧНОСТИ В STANDARD SQL

Темы, рассматриваемые в этой главе, относятся к области, не охваченной или оставленной неопределенной в Standard SQL. Взяв, к примеру, массивы. В PostgreSQL поддерживается синтаксис массивов в следующей форме:

```
SELECT pay_by_quarter[3] FROM sal_emp;
```

Тогда как BigQuery в этом случае потребует, чтобы мы уточнили свое намерение с помощью UNNEST, указав:

```
SELECT pay_by_quarter[ORDINAL(3)] FROM sal_emp, UNNEST(pay_by_quarter)
```

Подход, реализованный в PostgreSQL, выглядит понятнее, но в более сложных ситуациях возникает возможность неоднозначной интерпретации. Допустим, что запрос `SELECT author, book.title FROM ds.all_books` в BigQuery выполняет коррелированное перекрестное соединение вложенных повторяющихся полей и для каждой записи возвращает `author: STRING, book.title: STRING`. Этот подход кажется очень удобным, но при наличии нескольких вложенных повторяющихся полей возникает проблема. Предположим, что каждая глава имеет заголовок. Каким образом движок должен обработать следующий запрос:

```
SELECT author, book.title, book.chapter.title  
FROM ds.all_books
```

Возможно, пользователь хотел, чтобы под видом `book.chapter.title` запрос вернул массив, а может быть, он хотел получить простую строку со всеми заголовками глав. Если вы сохраните ограничения на все выражения, кроме JOIN, которое изменяет число записей на выходе и требует объединения через запятую с UNNEST для доступа к массивам, то эта двусмысленность исчезнет.

Продвинутый SQL

В этом разделе мы рассмотрим синтаксис SQL, который, как правило, несколько отличается в разных движках SQL (см. «Неоднозначности в стандартном SQL» ниже), но, тем не менее, достоин того, чтобы держать его в своем арсенале. Массивы и оконные функции могут значительно повысить производительность запросов. Табличные метаданные обеспечивают возможность рефлексии¹ и решения некоторых типичных задач. Язык определения данных (Data Definition Language, DDL) и язык манипулирования данными (Data Manipulation Language, DML)

¹ В данном контексте под рефлексией понимается возможность запроса SQL изменять свою структуру и поведение во время выполнения в зависимости от схемы таблицы.

поддерживают многие операции по обслуживанию базы данных в самом SQL, что позволяет использовать те же среды разработки, проверки, тестирования или непрерывной интеграции/непрерывного развертывания (Continuous Integration/Continuous Deployment, CI/CD), которые используются для SQL-запросов обслуживания базы данных.

Работа с массивами

Массивы в BigQuery — это упорядоченные списки, содержащие значения одного типа данных. Массивы полезны везде, где есть необходимость упорядочить или сохранить повторяющиеся значения в одной записи либо и то и другое.



Хранение данных в виде массивов может сократить потребление памяти и, в зависимости от количества элементов, значительно ускорить запросы. Из-за этого следует разобраться с массивами и познакомиться с ними поближе.

Использование массивов для сохранения порядка

В качестве примера, когда может понадобиться сохранить порядок следования данных, представим, что нам нужно выявить 100 наиболее часто арендуемых велосипедов:

```
SELECT
  bike_id,
  COUNT(*) AS num_trips
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire
GROUP BY
  bike_id
ORDER BY
  num_trips DESC
LIMIT
  100
```

Этот запрос вернет 100 записей. Вот первые три из них:

Row	bike_id	num_trips
1	12925	2922
2	12841	2871
3	13071	2860

Мы могли бы сохранить эти 100 записей в таблицу для последующего анализа с использованием других запросов, может быть, даже за пределами BigQuery,

но проблема в том, что считанные из таблицы записи не обязательно будут упорядочены.¹ Как быть, если для последующего анализа необходимо, чтобы данные были упорядочены? Одно из решений — использовать материализованное представление, другое — сохранить результат в одной записи, объединив результаты в массив и переместив `ORDER BY` и `LIMIT` в инструкцию `ARRAY_AGG`:

```
WITH numtrips AS (  
  SELECT  
    bike_id AS id,  
    COUNT(*) AS num_trips  
  FROM  
    `bigquery-public-data`.london_bicycles.cycle_hire  
  GROUP BY  
    bike_id  
)  
  
SELECT  
  ARRAY_AGG(STRUCT(id,num_trips)  
            ORDER BY num_trips DESC LIMIT 100)  
  AS bike  
FROM  
  numtrips
```

Этот запрос вернет единственную запись:

Row	bike_id	bike.num_trips
1	12925	2922
	12841	2871
	13071	2860

Если сохранить эту запись в таблице, список наиболее часто арендуемых велосипедов останется упорядоченным и все приложения, обращающиеся к этой таблице, будут получать упорядоченный список. Таким образом, ответственность за работу с массивами и умение разворачивать их целиком теперь полностью ложится на аналитика.

Как видно на примере, создавая массив структур, можно поддерживать отношение «один к одному» между столбцами. Теперь представьте себе массив структур как своего рода мини-таблицу, хранящуюся в каждой записи. BigQuery не поддерживает массивы массивов, но всегда можно создать массив структур, каждая из которых, в свою очередь, может содержать массивы.

¹ На самом деле 100 записей с двумя столбцами — это достаточно мало, чтобы данные поместились в один сегмент, и, вероятнее всего, они бы возвращались по порядку. Однако обычно нельзя рассчитывать, что записи будут возвращаться в каком-то определенном порядке.

Использование массивов для хранения повторяющихся полей

Еще одна причина использовать массивы — необходимость хранения повторяющихся полей, даже если порядок их следования неважен. Например, в BigQuery существует набор данных, содержащий налоговые декларации США благотворительных организаций. Подавляющее большинство таких организаций выплачивает налоги только один раз в год, но некоторые подают несколько деклараций в год. Эти декларации удобно хранить в одной записи:

```
SELECT
  ein
  , ARRAY_AGG(STRUCT(elf, tax_pd, subsecdd)) AS filing
FROM `bigquery-public-data`.irs_990.irs_990_2015
WHERE ein BETWEEN '390' AND '399'
GROUP BY ein
LIMIT 3
```

Этот запрос вернет следующие результаты:

Row	ein	filing.elf	filing.tax_pd	filing.subsecdd
1	390123480	E	201412	8
		E	201312	8
2	390233059	E	201412	12
3	390201015	E	201412	8

Обратите внимание, что первая организация подала две декларации в 2015 году (за 2014 и 2013 годы), тогда как другие — только по одной.

МАССИВЫ ГАРАНТИРУЮТ ЦЕЛОСТНОСТЬ ДАННЫХ

Изменение схемы таблицы для использования массивов позволяет применять лучшие методы и препятствует появлению неправильной логики в запросах. Исходная схема, в которой каждой декларации соответствует отдельная запись, может приводить к логическим ошибкам, когда аналитик забывает, что организация может подать несколько отчетов в течение года. Например, вы решили найти благотворительные организации, которые не подают декларации в электронном виде (если декларация подавалась в электронном виде, в столбце `elf` будет храниться значение 'E'). Вы можете поддасться соблазну и выбрать самое простое решение:

```
SELECT
  ein
FROM `bigquery-public-data`.irs_990.irs_990_2015
WHERE elf != 'E'
```

Однако поскольку некоторые благотворительные организации сдают декларации несколько раз в год, возможно, что часть из них будет подана в электронном

виде, а часть — нет. И это действительно так, в чем можно убедиться, выполнив следующий запрос:

```
SELECT
  ein
  , COUNTIF(elf = 'E', 1, 0) AS num_elf
  , COUNTIF(elf = 'E', 0, 1) AS num_not_elf
FROM `bigquery-public-data`.irs_990.irs_990_2015
GROUP BY ein
HAVING num_elf > 0 AND num_not_elf > 0
ORDER BY num_elf DESC
LIMIT 3
```

Судя по результатам этого запроса, действительно встречаются благотворительные организации, подающие декларации и в электронном виде, и на бумаге:

Row	ein	num_elf	num_not_elf
1	271157077	3	1
2	030555726	3	1
3	363977636	3	3

Если бы схема таблицы предусматривала хранение деклараций в виде массива, аналитикам было бы сложнее допустить эту ошибку — аналитик, который ищет благотворительные организации, подающие декларации не в электронном виде, знал бы, что декларации хранятся в массиве, и поэтому написал бы такой запрос:

```
SELECT
  ein
FROM
  [TABLENAME]
WHERE
  'E' NOT IN (SELECT elf FROM UNNEST(filing))
LIMIT 5
```



Поскольку декларации хранятся в таблице не в виде массива, попробуйте выполнить запрос, предварительно преобразовав его в требуемую схему с помощью предложения `WITH` и использовав `filings` в роли имени таблицы в основном запросе:

```
WITH filings AS (
  SELECT
    ein
    , ARRAY_AGG(STRUCT(elf, tax_pd, subsecdd)) AS filing
  FROM `bigquery-public-data`.irs_990.irs_990_2015
  GROUP BY ein
)
```

Функция `UNNEST` разворачивает массив и должна находиться в предложении `FROM`. Обратите внимание на использование коррелированного подзапроса для

извлечения поля `elf` из массива структур `filing` и `IN` для проверки наличия значения в массиве.

Другой способ проверить наличие значения 'E' в массиве структур — использовать `EXISTS`:

```
SELECT
  ein
FROM
  [TABLENAME]
WHERE
  EXISTS (SELECT elf FROM UNNEST(filing) WHERE elf != 'E')
LIMIT 5
```



Проверить правильность результатов, полученных этим способом ('520910763', '237048405', '410519270', '592515700' и '420655796'), можно так:

```
SELECT
  ein
  , SUM(IF(elf = 'E', 1, 0)) AS num_elf
  , SUM(IF(elf = 'E', 0, 1)) AS num_not_elf
FROM `bigquery-public-data`.irs_990.irs_990_2015
WHERE ein in UNNEST(
  ['520910763', '237048405', '410519270', '592515700',
   '420655796'])
GROUP BY ein
ORDER BY num_elf DESC
```

Как и ожидалось, значение `num_elf` для всех `ein` оказалось равным нулю.

Использование массивов для генерирования данных

Еще одна причина использовать массивы — генерирование данных. Представьте, что летом вам нужно подстригать газон каждые 10 дней и у вас есть три добровольца. Вы хотите, чтобы они выполняли эту работу по очереди.

Для начала можно сгенерировать список дней стрижки газона:

```
SELECT
  GENERATE_DATE_ARRAY('2019-06-23', '2019-08-22', INTERVAL 10 DAY) AS summer
```

Вот результаты этого запроса:

Row	summer
1	2019-06-23
	2019-07-03
	2019-07-13

Row	summer
	2019-07-23
	2019-08-02
	2019-08-12
	2019-08-22

Все они находятся в одной записи. Создать таблицу с несколькими записями, по одной для каждого дня, можно с помощью **UNNEST** (обратите внимание, что таблицу **days** и массив **summer** мы используем в предложении **FROM** — функцию **UNNEST** можно использовать в BigQuery только в предложении **FROM**):

```
WITH days AS (
  SELECT
    GENERATE_DATE_ARRAY('2019-06-23', '2019-08-22', INTERVAL 10 DAY) AS summer
)
SELECT summer_day
FROM days, UNNEST(summer) AS summer_day
```

Этот запрос вернет несколько записей — по одной для каждого дня:

Row	summer_day
1	2019-06-23
2	2019-07-03
3	2019-07-13
4	2019-07-23
5	2019-08-02
6	2019-08-12
7	2019-08-22



Запятая в предыдущем запросе выполняет коррелированное перекрестное соединение **CROSS JOIN**¹ и поэтому исключает записи с пустыми массивами или значениями **NULL** вместо массивов. Чтобы включить их, замените запятую на **LEFT JOIN**:

```
FROM days LEFT JOIN UNNEST(summer) AS summer_day
```

¹ Это коррелированное перекрестное соединение: оператор **UNNEST** ссылается на столбец **ARRAY** в каждой записи в исходной таблице, который раньше присутствовал в предложении **FROM**. Для каждой записи в исходной таблице **UNNEST** разворачивает массив **ARRAY** в набор записей, содержащих элементы **ARRAY**, а затем **CROSS JOIN** соединяет этот новый набор записей с одной записью из исходной таблицы. Поскольку эти перекрестные соединения в **UNNEST** коррелируются только с элементами в массиве, их влияние на производительность минимально.

Мы можем жестко закодировать список добровольцев, например, так:

```
SELECT ['Lak', 'Jordan', 'Graham'] AS minions
```

Теперь нужно распределить летние дни между добровольцами, чтобы они работали по очереди. Для этого сгенерируем массив `dayno` и используем деление по модулю (получение остатка от деления) для получения индекса в массиве `minions`:

```
WITH days AS (  
  SELECT  
    GENERATE_DATE_ARRAY('2019-06-23', '2019-08-22',  
      INTERVAL 10 DAY) AS summer,  
    ['Lak', 'Jordan', 'Graham'] AS minions  
)  
  
SELECT  
  summer[ORDINAL(dayno)] AS summer_day  
  , minions[OFFSET(MOD(dayno,  
    ARRAY_LENGTH(minions)))]  
    AS minion  
FROM  
  days, UNNEST(GENERATE_ARRAY(1,ARRAY_LENGTH(summer),1)) dayno  
ORDER BY summer_day ASC
```

В результате получаем список летних дней для стрижки газона и тех, кто будет выполнять эту работу в каждый из дней:

Row	summer_day	minion
1	2019-06-23	Jordan
2	2019-07-03	Graham
3	2019-07-13	Lak
4	2019-07-23	Jordan
5	2019-08-02	Graham
6	2019-08-12	Lak
7	2019-08-22	Jordan

Обратите внимание, что `ORDINAL` индексирует массив, начиная с единицы, а `OFFSET` — с нуля. Мы также использовали функцию `ARRAY_LENGTH`, чтобы получить длину массива `summer`.

Функции для работы с массивами

Два массива можно объединить с помощью функции `ARRAY_CONCAT`:

```
SELECT  
  ARRAY_CONCAT(  
    GENERATE_DATE_ARRAY('2019-03-23', '2019-06-22', INTERVAL 20 DAY)  
    , GENERATE_DATE_ARRAY('2019-08-23', '2019-11-22', INTERVAL 20 DAY)  
  ) AS shoulder_season
```

Для отладки иногда полезно вывести содержимое массива строк в виде строки `STRING`. Сделать это можно с помощью `ARRAY_TO_STRING`:

```
SELECT
  ARRAY_TO_STRING(['A', 'B', NULL, 'D'], '*', 'na') AS arr
```

Этот запрос вернет следующий результат:

Row	arr
1	A*B*na*D

Во втором параметре функции `ARRAY_TO_STRING` передается разделитель, а в третьем — строка для замены значений `NULL` (по умолчанию они просто пропускаются).

Во время отладки часто бывает полезно вывести содержимое массива в виде строки. Функция `ARRAY_TO_STRING` работает только с массивами строк, зато есть функция `TO_JSON_STRING`, которая работает с массивами любого типа. Например, вот как можно вывести содержимое массива дат:

```
SELECT
  TO_JSON_STRING(
    GENERATE_DATE_ARRAY('2019-06-23', '2019-08-22',
      INTERVAL 10 DAY)) AS json
```

Этот запрос вернет содержимое массива, преобразовав даты в формат JSON:

Row	json
1	["2019-06-23","2019-07-03","2019-07-13","2019-07-23", "2019-08-02","2019-08-12","2019-08-22"]

Формат JSON позволяет получить особенно наглядные результаты, если массив хранит структуры:

```
SELECT
  TO_JSON_STRING([
    STRUCT(1 AS a, 'bbb' AS b),
    STRUCT(2 AS a, 'ccc' AS b)
  ]) AS json
```

Вот получившийся результат:

Row	json
1	[{"a":1,"b":"bbb"},{"a":2,"b":"ccc"}]

В табл. 8.1 приводится краткая характеристика функций для работы с массивами. Здесь предполагается, что в столбце `minions` хранится массив строк.

Таблица 8.1. Функции для работы с массивами

Функция	Что делает	Пример использования
<code>GENERATE_ARRAY</code> <code>GENERATE_DATE_ARRAY</code>	Генерирует массив	<code>SELECT GENERATE_ARRAY(10, 20, 3)</code>
<code>OFFSET</code> <code>ORDINAL</code>	Обращается к элементам массива	<code>SELECT minions[OFFSET(0)] FROM ... SELECT minions[ORDINAL(1)] FROM ...</code>
<code>ARRAY_LENGTH</code>	Возвращает длину массива	<code>SELECT ARRAY_LENGTH(minions)</code>
<code>UNNEST</code>	Разворачивает массив; может использоваться только в предложении <code>FROM</code>	<code>WITH workers AS (SELECT ['Lak', 'Jordan', 'Graham'] AS minions) SELECT m FROM workers, UNNEST(minions) AS m</code>
<code>IN</code>	Проверяет наличие значения в массиве	<code>WITH workers AS (SELECT ['Lak', 'Jordan', 'Graham'] AS minions) SELECT 'Lak' IN UNNEST(minions) FROM workers</code>
<code>EXISTS</code>	Проверяет наличие хотя бы одного элемента в массиве	<code>WITH workers AS (SELECT ['Lak', 'Jordan', 'Graham'] AS minions UNION ALL SELECT [] AS minions) SELECT EXISTS (SELECT * FROM UNNEST(minions)) FROM workers</code>
<code>ARRAY_AGG</code>	Агрегирует сгруппированные элементы в массив	<code>SELECT ein , ARRAY_AGG(elf) AS elf FROM `bigquery-public-data`.irs_990. irs_990_2015 GROUP BY ein LIMIT 3</code>
<code>ARRAY_CONCAT</code>	Объединяет два массива одного типа	<code>SELECT ARRAY_CONCAT(['A', 'B', 'C'], ['D', 'E', 'F'])</code>
<code>ARRAY_TO_STRING</code> <code>TO_JSON_STRING</code>	Преобразует массив в строку	<code>SELECT TO_JSON_STRING([STRUCT(1 AS a, 'bbb' AS b), STRUCT(2 AS a, 'ccc' AS b))</code>

Оконные функции

Аналитические оконные функции (иногда их сокращенно называют *аналитическими функциями*, или *оконными функциями*) вычисляют совокупную аналитику (например, SUM, COUNT и т. д.) по группе строк (называемой окном). Есть три типа аналитических функций: агрегатные аналитические функции, функции навигации и функции нумерации.

Агрегатные аналитические функции

Агрегатные функции работают со всей таблицей и возвращают одно агрегированное значение. Например, следующий запрос возвращает наибольшую продолжительность поездки и две другие агрегатные статистики, вычисленные по полному набору данных `london_bicycles`:

```
SELECT
  MAX(duration) AS longest_duration
, COUNT(*) AS num_trips
, AVG(duration) AS average_duration
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire
```

В результате возвращается одна запись:

Row	longest_duration	num_trips	average_duration
1	2674020	24369201	1332.2942381245884

Также агрегатные аналитические функции могут возвращать агрегатные статистики для каждой записи, вычисляя их по «окружающим» записям. Например, следующий запрос вычисляет скользящее среднее длительности поездок по предшествующим 100 записям:

```
SELECT
  AVG(duration)
  OVER(ORDER BY start_date ASC
        ROWS BETWEEN 100 PRECEDING AND 1 PRECEDING)
  AS average_duration
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire
LIMIT 5
```

В отличие от агрегатной функции `AVG(duration)` в предыдущем примере, здесь среднее значение вычисляется не по всей таблице, а по окну, которое определено с помощью предложения `OVER`. Само окно включает 100 предшествующих записей, упорядоченных по столбцу `start_date`. Этот запрос возвращает `average_duration` для каждой записи, причем для первой возвращается значение `null`, потому что в предшествующем ей окне нет ни одной записи:

Row	average_duration
1	null
2	360.0
3	510.0
4	380.0
5	390.0

Если бы нам понадобилось вычислить среднюю продолжительность по окну, включающему 50 предшествующих и 50 последующих поездок, мы могли бы использовать такой запрос:

```
ROWS BETWEEN 50 PRECEDING AND 50 FOLLOWING
```

Чтобы вычислить среднее значение по всем предшествующим записям, от начала набора данных до текущей записи (включительно), мы могли бы использовать такой запрос:

```
ROWS BETWEEN UNBOUNDED PRECEDING AND CURRENT ROW
```

Если отбросить оператор **FOLLOWING**, то роль границы по умолчанию будет играть данная запись, то есть выражение **ROWS 50 PRECEDING** будет равноценно выражению **ROWS BETWEEN 50 PRECEDING AND CURRENT ROW**.

А как подсчитать среднюю продолжительность по предыдущим 100 поездкам, но ограничить окно поездками, которые начинаются с того же пункта проката, который указан в данной записи? В этом случае в предложение **OVER()** нужно добавить **PARTITION BY**:

```
SELECT
  AVG(duration)
  OVER(PARTITION BY start_station_id
        ORDER BY start_date ASC
        ROWS BETWEEN 100 PRECEDING AND 1 PRECEDING)
  AS average_duration
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire
LIMIT 5
```

Вычислить скользящее среднее за последний час можно с помощью **RANGE**, указав целочисленное смещение в единицах упорядочения:

```
SELECT
  AVG(duration)
  OVER(PARTITION BY start_station_id
        ORDER BY UNIX_SECONDS(start_date) ASC
        RANGE BETWEEN 3600 PRECEDING AND CURRENT ROW)
```



```

        AS average_duration
FROM
    `bigquery-public-data`.london_bicycles.cycle_hire
LIMIT 5

```

Обратите внимание, что здесь упорядочивание выполняется по `UNIX_SECONDS(start_date)`, а не просто по `start_date`, чтобы можно было найти записи, для которых `UNIX_SECONDS(start_date)` находится между 3600 секундами назад и настоящим временем.

Функции навигации

Агрегатные аналитические функции вычисляют агрегированные статистики по всем записям в окне. То есть `AVG(duration)` вычисляет среднюю продолжительность, а `MAX(duration)` — максимальную продолжительность в окне. А что, если понадобится найти единственное значение, определяемое местоположением записи? В этом случае вам пригодятся функции навигации.

Допустим, вам нужно найти «следующую» поездку на данном велосипеде. В предложении `OVER()` вам потребуется выполнить сегментирование по `bike_id`, упорядочить по `start_date` и определить окно между текущей и следующей записью. Чтобы найти `start_date` последней записи в этом окне, можно применить функцию `LAST_VALUE` и получить время начала следующей поездки на данном велосипеде:

```

SELECT
    start_date
    , end_date
    , LAST_VALUE(start_date)
      OVER(PARTITION BY bike_id
           ORDER BY start_date ASC
           ROWS BETWEEN CURRENT ROW AND 1 FOLLOWING)
      AS next_rental_start
FROM
    `bigquery-public-data`.london_bicycles.cycle_hire
LIMIT 5

```

Этот запрос вернет следующие результаты:

Row	start_date	end_date	next_rental_start
1	2015-01-06 20:01:00 UTC	2015-01-06 20:13:00 UTC	2015-01-07 08:03:00 UTC
2	2015-01-07 08:03:00 UTC	2015-01-07 08:22:00 UTC	2015-01-07 10:06:00 UTC
3	2015-01-07 10:06:00 UTC	2015-01-07 10:14:00 UTC	2015-01-07 13:03:00 UTC
4	2015-01-07 13:03:00 UTC	2015-01-07 13:09:00 UTC	2015-01-07 14:46:00 UTC
5	2015-01-07 14:46:00 UTC	2015-01-07 15:00:00 UTC	2015-01-07 17:35:00 UTC

Можно не указывать явно границы окна, а использовать функцию LEAD:

```
SELECT
  start_date
  , end_date
  , LEAD(start_date, 1)
    OVER(PARTITION BY bike_id
         ORDER BY start_date ASC )
    AS next_rental_start
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire
LIMIT 5
```

Во втором параметре функции LEAD передается количество последующих записей. Значение по умолчанию равно 1, поэтому здесь мы могли бы опустить этот параметр.

Родственными для LAST_VALUE и LEAD являются функции FIRST_VALUE и LAG. Также можно использовать более общую функцию NTH_VALUE, чтобы получить значение в любой позиции окна.

Функции нумерации

Функции нумерации позволяют определить позицию текущей записи в окне, если эти записи должны быть упорядочены определенным образом. Например, найти в наборе данных london_bicycles пять самых долгих поездок, начавшихся в каждом пункте проката, можно, создав окна, сегментированные по start_station_id, упорядочив эти окна по продолжительности, а затем вычислив ранг каждой поездки:

```
SELECT
  start_station_id
  , duration
  , RANK()
    OVER(PARTITION BY start_station_id ORDER BY duration DESC)
    AS nth_longest
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire
LIMIT 5
```

RANK — это аналитическая оконная функция. Ее применение в предыдущем запросе позволяет получить ранг каждой поездки среди всех поездок, начавшихся в том же пункте проката:

Row	start_station_id	duration	nth_longest
1	13	1448640	1
2	13	346080	2

Row	start_station_id	duration	nth_longest
3	13	225420	3
4	13	165000	4
5	13	92700	5

Поместив предыдущий SELECT в предложение WITH, сгруппировав по `start_station_id` и использовав `ARRAY_AGG`, можно получить три самые долгие поездки для каждого пункта проката:

```
WITH longest_trips AS (
  SELECT
    start_station_id
    , duration
    , RANK()
      OVER(PARTITION BY start_station_id ORDER BY duration DESC)
      AS nth_longest
  FROM
    `bigquery-public-data`.london_bicycles.cycle_hire
)

SELECT
  start_station_id
  , ARRAY_AGG(duration ORDER BY nth_longest LIMIT 3) AS durations
FROM
  longest_trips
GROUP BY start_station_id
LIMIT 5
```

Вот полученные результаты:

Row	start_station_id	durations
1	10	243300
		101700
		93840
2	17	737280
		247920
		197460
3	34	1017180
		588900
		560700
4	43	244740

Row	start_station_id	durations
		193980
		176580
4	60	1303260
		596520
		319140

Помимо `RANK()`, в BigQuery также поддерживаются функции `DENSE_RANK()` и `ROW_NUMBER()`. Они отличаются только способом обработки привязки. Следующий короткий пример иллюстрирует эту особенность:

```
WITH example AS (
  SELECT 'A' AS name, 32 AS age
  UNION ALL SELECT 'B', 32
  UNION ALL SELECT 'C', 33
  UNION ALL SELECT 'D', 33
  UNION ALL SELECT 'E', 34
)

SELECT
  name
  , age
  , RANK() OVER(ORDER BY age) AS rank
  , DENSE_RANK() OVER(ORDER BY age) AS dense_rank
  , ROW_NUMBER() OVER(ORDER BY age) AS row_number
FROM example
```

Как показывает результат, `RANK()` пропускает номера, если есть связь, `DENSE_RANK()` присваивает ранг хотя бы один раз, а `ROW_NUMBER()` — уникальный номер каждой записи:

Row	name	age	rank	dense_rank	row_number
1	A	32	1	1	1
2	B	32	1	1	2
3	C	33	3	2	3
4	D	33	3	2	4
5	E	34	5	3	5

Метаданные таблиц

До сих пор мы рассматривали примеры, демонстрирующие способы запроса данных из таблицы. Но кроме самих данных есть еще и информация об этих данных, которую называют *метаданными*. Сюда относится, например, перечень

имеющихся столбцов, время создания таблицы, кому она принадлежит, кто имеет право доступа к ней и т. д. В этом разделе мы расскажем, как можно получить такие метаданные таблицы.

Динамическое конструирование запросов

В этой главе мы уже говорили о том, что лучший способ организовать информацию о налоговых декларациях, помогающий уменьшить число логических ошибок, — хранить перечень деклараций, поданных в течение года, в виде массива:

```
SELECT
  ein
  , ARRAY_AGG(STRUCT(elf, tax_pd, subseccd)) AS filing
FROM `bigquery-public-data`.irs_990.irs_990_2015
GROUP BY ein
```

В нашем примере показаны только три столбца — `elf`, `tax_pd` и `subseccd`, потому что набирать имена всех столбцов вручную достаточно непросто. Однако проблему создания требуемого оператора с именами всех столбцов в таблице легко решить с помощью `INFORMATION_SCHEMA` (<https://cloud.google.com/bigquery/docs/information-schema-tables>) — специальной таблицы, входящей в состав любого набора данных и содержащей метаданные обо всех таблицах в этом наборе.

Например, вот как можно получить имена столбцов в таблице `irs_990_2015` в наборе данных `irs_990`:

```
SELECT column_name
FROM `bigquery-public-data`.irs_990.INFORMATION_SCHEMA.COLUMNS
WHERE table_name = 'irs_990_2015'
```

Этот запрос вернет набор результатов с 246 столбцами, первые из которых показаны ниже:

Row	column_name
1	ein
2	elf
3	tax_pd
4	subseccd
5	s501c3or4947a1cd

На основании этих результатов можно с помощью `CONCAT` сгенерировать текст требуемого запроса:

```

WITH columns AS (
  SELECT column_name
  FROM `bigquery-public-data`.irs_990.INFORMATION_SCHEMA.COLUMNS
  WHERE table_name = 'irs_990_2015' AND column_name != 'ein'
)

SELECT CONCAT(
  'SELECT ein, ARRAY_AGG(STRUCT(',
  ARRAY_TO_STRING(ARRAY(SELECT column_name FROM columns), ',\n '),
  '\n) FROM `bigquery-public-data`.irs_990.irs_990_2015\n',
  'GROUP BY ein')

```

В результате получится такой запрос:

```

SELECT ein, ARRAY_AGG(STRUCT(ein,
  elf,
  tax_pd,
  subseccd,
  ...
  othrinc509,
  totsupp509
) FROM `bigquery-public-data`.irs_990.irs_990_2015
GROUP BY ein

```

Далее мы рассмотрим скрипты. На момент написания этой книги динамические SQL-запросы уже были включены в планы развития BigQuery, но еще не были реализованы. Если к тому моменту, когда вы будете читать эти строки, поддержка динамических запросов в BigQuery уже будет реализована, вы сможете не только вывести текст запроса, но и выполнить его!

Метки и теги

Таблица с информацией о схеме, описанная в предыдущем разделе, содержит сведения об именах и типах столбцов в таблицах, время создания таблиц и многое другое. Кроме того, можно добавлять свои метаданные о ресурсах (наборах данных, таблицах, представлениях и т. д.) в виде меток, таких как: название приложения, создавшего ресурс (например, `component:salesportal`), группа-владелец (например, `team:emeasales`), в каком окружении находится и используется (например, `environment:production`) и на какой стадии жизненного цикла находится (например, `state:validated`). Каждая метка определяется как пара ключ/значение (например, `environment` — это ключ, а его значениями могут быть `development`, `staging`, `test` или `production`). Метка с пустым значением называется тегом, но его можно использовать аналогично.

Ключи и значения в BigQuery можно выбирать абсолютно произвольно; ваша команда управления данными должна определить ключи и значения, которые могут присутствовать в таблицах, принадлежащих вашей организации. Присвоение меток ресурсам позволяет искать их по меткам и систематически применять одинаковые политики ко всем ресурсам с одной меткой.

Например, добавим метку к набору данных `ch08eu`, используя инструмент командной строки `bq` (то же самое можно сделать с помощью консоли GCP, выполнив инструкцию `ALTER TABLE SET OPTIONS`, и различных клиентских библиотек):

```
bq update --set_label costcenter:abc342 ch08eu
```

Изменить значение метки можно с помощью ключа `--set_label` с требуемым значением:

```
bq update --set_label costcenter:def456 ch08eu
```

Также есть возможность присвоить метку самому запросу:

```
bq query --label costcenter:def456 --nouse_legacy_sql 'SELECT ...'
```



Не указывайте конфиденциальные данные в метках — метки доступны даже тем, у кого нет разрешений на чтение ресурсов, которым эти метки присвоены.

После присвоения меток наборам данных или таблицам их можно искать по меткам. Например, можно выполнить такой поиск (аналогичный поиск можно выполнить с помощью REST API или различных клиентских библиотек):

```
bq ls --filter 'labels.costcenter:def456'
```

Эта команда вернет следующее:

```
datasetId
-----
ch08eu
```

Путешествия во времени

Как вы уже знаете, в BigQuery есть возможность запросить предыдущее состояние таблицы сроком давности до семи дней. Например, вот как с помощью `SYSTEM_TIME` можно запросить данные, хранившиеся в таблице шесть часов назад:

```
SELECT
*
FROM `bigquery-public-data`.london_bicycles.cycle_stations
FOR SYSTEM_TIME AS OF
    TIMESTAMP_SUB(CURRENT_TIMESTAMP(), INTERVAL 6 HOUR)
```

Обратите внимание, что в одном запросе нельзя сослаться на состояние таблицы более чем в один момент времени. Если вам понадобится найти отличия между разными временными версиями таблицы, скопируйте одну из версий (используя язык определения данных (Data Definition Language, DDL); см. следующий раздел) в новую таблицу, а затем выполните соединение.



Используйте возможность перемещения во времени, если вам понадобится несколько раз выполнить один и тот же запрос к таблице, которая передает данные через поток.

Язык определения данных и язык манипулирования данными

Инструкции языка определения данных (Data Definition Language, DDL) — это те же инструкции SQL, позволяющие создавать, изменять и удалять таблицы и представления. Язык манипулирования данными (Data Manipulation Language, DML) позволяет изменять, добавлять и удалять данные из таблиц.

DDL

Мы уже разбирали способы создания таблиц или представлений с использованием запросов. Например, вот как можно создать таблицу, включающую только пункты проката, имеющие слово «Hyde» в названии:

```
CREATE OR REPLACE TABLE ch08eu.hydepark_stations AS
SELECT
  * EXCEPT(longitude, latitude)
  , ST_GeogPoint(longitude, latitude) AS location
FROM `bigquery-public-data`.london_bicycles.cycle_stations
WHERE name LIKE '%Hyde%'
```

Список параметров. В главе 7 мы видели, как добавлять спецификации, управляющие секционированием и кластеризацией таблиц во время их создания. Вы также можете указать список параметров, определяющих отметку времени, когда таблица станет недействительной, а также различные метки и т. д.:

```
CREATE OR REPLACE TABLE ch08eu.hydepark_stations
OPTIONS(
  expiration_timestamp=TIMESTAMP "2020-01-01 00:00:00 UTC",
  description="Stations with Hyde Park in the name",
  labels=[("cost_center", "abc123")]
) AS
SELECT
  * EXCEPT(longitude, latitude)
  , ST_GeogPoint(longitude, latitude) AS location
FROM `bigquery-public-data.london_bicycles.cycle_stations`
WHERE name LIKE '%Hyde%'
```

Пустая таблица. Чтобы создать пустую таблицу, достаточно указать имена столбцов и их типы:


```
CREATE OR REPLACE TABLE ch08eu.hydepark_rides
(
  start_time TIMESTAMP,
  duration INT64,
  start_station_id INT64,
  start_station_name STRING,
  end_station_id INT64,
  end_station_name STRING
)
PARTITION BY DATE(start_time)
CLUSTER BY start_station_id
```

Изменение параметров. Некоторые параметры можно изменять после создания таблицы, например:

```
ALTER TABLE ch08eu.hydepark_rides
SET OPTIONS(
  expiration_timestamp=TIMESTAMP "2021-01-01 00:00:00 UTC",
  require_partition_filter=True,
  labels=[("cost_center", "def456")]
)
```

DML

BigQuery — это прежде всего хранилище данных, куда обычно загружаются или передаются потоковые данные. Поддержка возможности изменения данных не является основной задачей. Однако такая возможность есть: DML позволяет вставлять, изменять, удалять и объединять данные в таблицы.



BigQuery — это аналитическая база данных, а не база данных оперативной обработки транзакций (OnLine Transaction Processing, OLTP). Она не предназначена для частого изменения данных с помощью инструкций DML. Если вам понадобится изменить много данных, попробуйте сделать это в пакетном режиме. Например, можно создать промежуточную таблицу с необходимыми изменениями, а затем запустить один оператор UPDATE или MERGE, чтобы применить все эти изменения в одной операции.

Вставка выбранных записей в SELECT. Следующий запрос извлечет данные из таблицы `cycle_hires` и вставит их в таблицу `hydepark_rides` (которую мы создали в предыдущем разделе как секционированную таблицу):

```
INSERT ch08eu.hydepark_rides
SELECT
  start_date AS start_time
  , duration
  , start_station_id
  , start_station_name
  , end_station_id
```

```

    , end_station_name
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire
WHERE
  start_station_name LIKE '%Hyde%'

```

Эта инструкция добавит в таблицу 1.01 миллиона записей, которые можно получить так:

```

WITH rides_in_year AS (
SELECT
  EXTRACT(MONTH from start_time) AS month
  , duration
FROM ch08eu.hydepark_rides
WHERE
  DATE(start_time) BETWEEN '2016-01-01' AND '2016-12-31'
  AND start_station_id = 300
  AND end_station_id = 303
)

SELECT
  month
  , AVG(duration)/60 as avg_duration_minutes
FROM rides_in_year
GROUP BY month
ORDER BY avg_duration_minutes DESC
LIMIT 5

```

Обратите внимание на наличие в предложении `WHERE` в этом запросе столбца секционирования `start_time` и столбца кластеризации `start_station_id`. Наличие столбца секционирования является обязательным, потому что мы изменили таблицу, применив параметр `require_partition_filter = True`, а создание таблицы с `CLUSTER BY start_station_id` помогло добиться высокой эффективности запросов. Этот запрос обработал 12 Мбайт данных (вместо 740 Мбайт, если бы мы использовали некластеризованную таблицу). Из этого можно сделать вывод, что использование секционированной таблицы снижает стоимость запроса независимо от используемого тарифа — фиксированного с резервированием или до востребования.

Вставка значений VALUES. С помощью оператора `INSERT` можно вставить в таблицу пару поездок уже после загрузки данных:

```

INSERT ch08eu.hydepark_rides (
  start_time
  , duration
  , start_station_id
  , start_station_name
  , end_station_id
  , end_station_name
)
VALUES

```

```
('2016-02-18 17:21:00 UTC', 720, 300,
'Serpentine Car Park, Hyde Park', 303, 'Albert Gate, Hyde Park'),
('2016-02-18 16:30:00 UTC', 1320, 300,
'Serpentine Car Park, Hyde Park', 303, 'Albert Gate, Hyde Park')
```

Хотя это необязательно, имена столбцов все же лучше указать в операторе INSERT, как это сделано в предыдущем примере.



Если вас не интересует читабельность запроса, можете записать оператор INSERT, как показано ниже:

```
INSERT ch08eu.hydepark_rides
VALUES
('2016-02-18 17:21:00 UTC', 720, 300,
'Serpentine Car Park, Hyde Park', 303, 'Albert Gate, Hyde Park'),
('2016-02-18 16:30:00 UTC', 1320, 300,
'Serpentine Car Park, Hyde Park', 303, 'Albert Gate, Hyde Park')
```

Вставка значений VALUES с подзапросом SELECT. Мы не рекомендуем делать вставку способом, описанным выше, поскольку высока вероятность неправильного ввода названий пунктов проката (например, вместо 'Albert Gate, Hyde Park' можно по ошибке ввести 'Hyde Park: Albert Gate'). Гораздо безопаснее использовать подзапрос SELECT и выбирать названия из таблицы пунктов проката:

```
...
VALUES
('2016-02-18 17:21:00 UTC', 720,
 300, (SELECT name FROM `bigquery-public-data`.london_bicycles.cycle_stations
WHERE id = 300),
 303, (SELECT name FROM `bigquery-public-data`.london_bicycles.cycle_stations
WHERE id = 303)),
...
```

Имена таблиц нельзя использовать в функциях SQL, поэтому у вас не получится избавиться от повторяющегося кода перемещением подзапроса в функцию:

```
CREATE TEMPORARY FUNCTION stationName(stationId INT64) AS(
  (SELECT name FROM
    `bigquery-public-data`.london_bicycles.cycle_stations
    WHERE id = stationId)
);
```

На момент написания этой книги BigQuery не поддерживала хранимые процедуры, но они были включены в план развития сервиса. Возможно, когда вы будете читать эти строки, такая опция уже будет доступна, и вы сможете определить хранимую процедуру, чтобы вызывать ее в INSERT.

Удаление записей. Когда в нашем конвейере извлечения, преобразования и загрузки (ETL) обнаружилась ошибка, внесенная в декабре 2016 года, из-за

которой конвейер по ошибке загружал поездки с нулевой продолжительностью, вы смогли удалить эти записи с помощью DML:

```
DELETE ch08eu.hydepark_rides
WHERE
    start_time > '2016-12-01' AND
    (duration IS NULL OR duration = 0)
```

Помимо таких ошибок, приводящих к вставке ошибочных записей, есть и другие причины, по которым вам может потребоваться удалить записи из хранилища. Например, «право на забвение» требует, чтобы компании удаляли всю пользовательскую информацию при некоторых обстоятельствах. Это легко реализовать с помощью оператора DELETE:¹

```
DELETE ch08eu.hydepark_rides
WHERE
    userId = 3452123
```

Вместо выполнения подобных запросов для каждого случая в отдельности лучше объединить изменения в операторе MERGE (который мы рассмотрим далее).

Изменение значений в записях. Представьте, что продолжительности поездок от одного из пунктов проката пришли в минутах и мы забыли преобразовать эти значения в секунды перед сохранением. Проблему можно исправить, изменив значения в таблице:

```
UPDATE ch08eu.hydepark_rides
SET duration = duration * 60
WHERE
    start_time > '2016-12-01' AND
    start_station_id = 303
```

Если один из столбцов является массивом и в массив необходимо добавить новый элемент, это тоже можно сделать с помощью UPDATE. Предположим, что в таблице пунктов проката имеется столбец maintenance (график технического обслуживания). Вы сможете получить существующий график технического обслуживания и добавить в него новое событие:²

```
UPDATE ch08eu.stations_table
SET maintenance = ARRAY_CONCAT(maintenance,
    ARRAY_STRUCT<time TIMESTAMP, employeeID STRING>[
        (CURRENT_TIME(), emp303)
    ])
WHERE id = 303
```

¹ В таблице нет столбца userId, поэтому этот запрос не будет работать.

² В таблице stations_table нет такого столбца, поэтому этот запрос не будет работать.

Прежде чем использовать операторы DML, подумайте, возможно, есть более подходящий способ решить ту же задачу. Например, если график обслуживания требуется менять при каждом обслуживании, лучше хранить события в отдельной таблице и при необходимости соединять ее с таблицей пунктов проката в запросе.

Оператор MERGE. Оператор MERGE представляет собой атомарную комбинацию из INSERT, UPDATE и DELETE, которая выполняется (успешно или не выполняется вовсе) как один оператор. Записи из исходной таблицы (или из подзапроса) вставляются в целевую таблицу, и для каждой выполняется набор операций. Можно определить разные наборы операций для трех сценариев: MATCHED, NOT MATCHED BY TARGET или NOT MATCHED BY SOURCE.

Например, следующий запрос объединит записи из общедоступного набора данных london_bicycles с таблицей в нашем наборе ch08eu и выполнит разные операции в разных сценариях:

```
MERGE ch08eu.hydepark_stations T
USING
  (SELECT *
   FROM `bigquery-public-data`.london_bicycles.cycle_stations
   WHERE name LIKE '%Hyde%') S
ON T.id = S.id
WHEN MATCHED THEN
  UPDATE
  SET bikes_count = S.bikes_count
WHEN NOT MATCHED BY TARGET THEN
  INSERT(id, installed, locked, name, bikes_count)
  VALUES(id, installed, locked, name, bikes_count)
WHEN NOT MATCHED BY SOURCE THEN
  DELETE
```

Этот запрос объединит записи из таблицы в общедоступном наборе данных, полученные подзапросом (исходная таблица), с записями в ch08eu.hydepark_stations (целевая таблица), имеющими то же значение в столбце id. Если записи совпадают по столбцу id, столбец bikes_count в целевой таблице получит значение столбца bikes_count из исходной таблицы (другие столбцы остаются без изменений). Если значение id присутствует в исходной, но отсутствует в целевой таблице, запись из исходной таблицы вставляется в целевую. Если значение id присутствует в целевой таблице, но отсутствует в исходной, соответствующая запись удаляется из целевой таблицы. Оператор MERGE выполняется атомарно.

За пределами SQL

До сих пор мы рассматривали только SQL-запросы. Несмотря на широту возможностей SQL, некоторые операции проще выполнить на каком-нибудь другом процедурном языке. Для этого можно использовать пользовательские функции (UDF) на JavaScript.

Многие системы управления базами данных поддерживают возможность включения операторов SQL в код на другом процедурном языке. Конечно, можно организовать управление наборами операторов SQL с помощью клиентских библиотек BigQuery, но иногда удобнее выполнять весь набор операций на сервере, например, чтобы упростить обработку сбоев, откат изменений и передачу данных. Именно такую возможность дают сценарии и хранимые процедуры.

Пользовательские функции на JavaScript

По возможности пользовательские функции (UDF) лучше писать на SQL. BigQuery способна оптимизировать код на SQL, эффективно распределять его и выполнять. Но это не относится к пользовательским функциям на JavaScript. Кроме того, на JavaScript есть ограничения на размер кода, объем выходных данных и количество функций в запросе. Пользовательские функции на JavaScript более универсальны, чем пользовательские функции на SQL, но имейте в виду, что их выполнение связано с дополнительными накладными расходами, потому что они запускаются под управлением движка V8 в изолированном процессе.

Тем не менее иногда требуется выполнить сложные вычисления, которые тяжело реализовать на SQL.¹ В таких случаях пользовательские функции на JavaScript, даже со всеми их ограничениями, оказываются очень удобной альтернативой. Например, представьте, что ваша функция определения цены оказалась очень сложной, но у вас есть код для `computePrice()`, реализованный на JavaScript для вашего веб-сайта. Вы можете определить эту функцию и сохранить ее в своем наборе данных (также поддерживаются временные функции на JavaScript):

```
CREATE OR REPLACE FUNCTION ch08eu.computePrice(dur INT64)
RETURNS INT64
LANGUAGE js AS """
  function factorial(n) {
    return (n > 1) ? n * factorial(n - 1) : 1;
  }
  var nhours = 1 + Math.floor(dur/3600.0);
  var f = factorial(nhours);
  var discount = 0.8/(1+Math.pow(Math.E, -f));
  return 3 + Math.floor(dur * (1-discount) * 0.0023)
""";

SELECT
  duration, ch08eu.computePrice(duration) AS price
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire
LIMIT 5
```

¹ Сложно, но *не* невозможно. Пожалуйста, не присылайте нам свои реализации этой функции на SQL.

Обратите внимание на ключевые различия между пользовательскими функциями на JavaScript и SQL: нужно указать язык `js` и заключить всю функцию в тройные кавычки. Тело функции может включать другие функции на JavaScript (например, наша функция включает реализацию функции вычисления факториала). Вот результат этого запроса:

Row	duration	price
1	3180	6
2	7380	6
3	2040	4
4	2280	5
5	2340	5

Бывает и так, что в JavaScript уже есть библиотечная функция, которую вы захотите использовать повторно. Вы сможете вызвать функцию из внешнего пакета JavaScript, если загрузите его и сохраните в Google Cloud Storage. Например, если представить, что предыдущая функция вычисления факториала определена в файле с именем *mathfn.js*, вы можете сохранить этот файл в Google Cloud Storage и сослаться на него в списке параметров `OPTIONS` функции:

```
CREATE TEMPORARY FUNCTION computePrice(dur INT64)
RETURNS INT64
LANGUAGE js AS """
  var nhours = 1 + Math.floor(dur/3600.0);
  var f = factorial(nhours);
  var discount = 0.8/(1+Math.pow(Math.E, -f));
  return 3 + Math.floor(dur * (1-discount) * 0.0023)
""";
OPTIONS (
  library=["gs://somebucket/path/to/mathfn.js",
          "gs://somebucket/path/to/someother.js"]
);
```

Пользовательские функции на JavaScript выполняются на одном рабочем сервере и ограничены тем, что доступно этому серверу.¹ На момент написания этой книги асинхронные функции JavaScript не поддерживались.

Скрипты

BigQuery позволяет писать скрипты, состоящие из множества инструкций, и отправлять их в одном запросе. Скрипт может сохранять результаты в пере-

¹ Дополнительные рекомендации по использованию пользовательских функций на JavaScript можно найти по адресу <https://cloud.google.com/bigquery/docs/reference/standard-sql/user-defined-functions#best-practices-for-javascript-udfs>.

менных и использовать циклы для выполнения одного и того же запроса несколько раз.

Последовательность инструкций

Одна из самых простых причин, по которым может понадобиться писать скрипты, — поочередный запуск нескольких запросов. Допустим, вы решили создать промежуточную таблицу, чтобы выполнить соединение с ней, а затем удалить ее. Это можно сделать с помощью скрипта, просто написав последовательность инструкций:¹

```
CREATE OR REPLACE TABLE ch08eu.typical_trip AS
SELECT
    start_station_name
  , end_station_name
  , APPROX_QUANTILES(duration, 10)[OFFSET(5)] AS typical_duration
  , COUNT(*) AS num_trips
FROM
    `bigquery-public-data`.london_bicycles.cycle_hire
GROUP BY
    start_station_name, end_station_name
;

CREATE OR REPLACE TABLE ch08eu.unusual_days AS
SELECT
    EXTRACT (DATE FROM start_date) AS trip_date
  , APPROX_QUANTILES(duration / typical_duration, 10)[OFFSET(5)] AS ratio
  , COUNT(*) AS num_trips_on_day
FROM
    `bigquery-public-data`.london_bicycles.cycle_hire AS hire
  , ch08eu.typical_trip AS trip
WHERE
    hire.start_station_name = trip.start_station_name
    AND hire.end_station_name = trip.end_station_name
    AND num_trips > 10
GROUP BY trip_date
HAVING num_trips_on_day > 10
ORDER BY ratio DESC
;

DROP TABLE ch08eu.typical_trip;
```

Одна из отличительных черт скриптов заключается в необходимости ставить точку с запятой после каждой отдельной инструкции. Этот сценарий можно отправить в виде одного запроса.

¹ Полный скрипт можно найти в файле https://github.com/GoogleCloudPlatform/bigquery-oreilly-book/blob/master/08_advqueries/script_seq.sql.



Во многих случаях скрипты можно заменить предложениями `WITH`, соединениями, коррелированными подзапросами или `GROUP BY`. Прежде чем писать свой скрипт, посмотрите, нельзя ли решить стоящую перед вами задачу с помощью одного запроса. Один запрос, вероятно, будет выполнен более эффективно.

Например, предыдущий скрипт довольно легко заменить предложением `WITH`. К тому же предложения `WITH` выполняются быстрее. Даже очень сложные задачи можно решить с помощью одного запроса. Представьте, что у вас есть запрос, который ищет поездки длительностью более 30 минут для каждого `bike_id` в наборе данных `london_bicycles`. Если понадобится повторить этот анализ для порогов 60, 120, ..., 300 минут, может показаться, что для решения этой задачи нужен цикл, выполняющий запрос несколько раз. Однако результат можно получить с помощью одного запроса, выполнив соединение с массивом порогов и сгруппировав результат по `bike_id` и `threshold`.

Не злоупотребляйте скриптами.

Временные таблицы

Предыдущий скрипт можно упростить, если вместо создания и удаления таблицы использовать временную таблицу:¹

```
CREATE TEMPORARY TABLE typical_trip AS
SELECT
    start_station_name
    , end_station_name
    , APPROX_QUANTILES(duration, 10)[OFFSET(5)] AS typical_duration
    , COUNT(*) AS num_trips
FROM
    `bigquery-public-data`.london_bicycles.cycle_hire
GROUP BY
    start_station_name, end_station_name
;

CREATE OR REPLACE TABLE ch08eu.unusual_days AS
SELECT
    EXTRACT (DATE FROM start_date) AS trip_date
    , APPROX_QUANTILES(duration / typical_duration, 10)[OFFSET(5)] AS ratio
    , COUNT(*) AS num_trips_on_day
FROM
    `bigquery-public-data`.london_bicycles.cycle_hire AS hire
    , typical_trip AS trip
WHERE
    hire.start_station_name = trip.start_station_name
    AND hire.end_station_name = trip.end_station_name
```

¹ Полный скрипт можно найти в файле https://github.com/GoogleCloudPlatform/bigquery-oreilly-book/blob/master/08_advqueries/script_seq.sql.

```

    AND num_trips > 10
  GROUP BY trip_date
  HAVING num_trips_on_day > 10
  ORDER BY ratio DESC
;

```

Временные таблицы существуют, только пока выполняется скрипт, и автоматически удаляются по его завершении. Также обратите внимание, что, в отличие от постоянных таблиц, временные таблицы не связаны с набором данных.

Устройство простого скрипта

Чтобы получить более полное представление о скриптах, напомним скрипт поиска пунктов проката, в которых оканчиваются наиболее продолжительные поездки, начавшиеся в Гайд-парке. Большинство скриптов начинаются с раздела объявлений, в котором определяются переменные, используемые скриптом:¹

```

-- Переменные
DECLARE PATTERN STRING DEFAULT '%Hyde%';
DECLARE stations ARRAY<STRING>;
DECLARE MIN_TRIPS_THRESH INT64 DEFAULT 100;

```

Здесь мы описали шаблон названий искомых пунктов проката и массив для их хранения. Третья переменная — это минимальная продолжительность поездки, которую мы будем использовать в скрипте. Обратите внимание на типы переменных. Шаблон — это строка, а порог — целое число. Присваивая им значения по умолчанию, мы получаем возможность использовать их в роли именованных констант.

Переменные могут быть любого типа, поддерживаемого в BigQuery. Например, переменная `stations` — это массив строк, который можно использовать для хранения результата запроса (также мы могли бы использовать временную таблицу):

```

SET stations = (
  SELECT
    ARRAY_AGG(name)
  FROM
    `bigquery-public-data`.london_bicycles.cycle_stations
  WHERE
    name LIKE PATTERN
);

```

Обратите внимание, что в предыдущем коде оператор `SET` присваивает результат запроса переменной `stations` (массив) и использует переменную `PATTERN`

¹ Полный скрипт можно найти в файле https://github.com/GoogleCloudPlatform/bigquery-oreilly-book/blob/master/08_advqueries/script_temptbl.sql.

в запросе. Поскольку `stations` — это массив, запрос `SELECT` вызывает `ARRAY_AGG` для агрегирования столбца `name` по всем записям. В остальной части сценария массива `stations` будет использоваться подобно любому другому массиву (в `UNNEST` и т. д.).

Теперь можно найти конечные пункты наиболее продолжительных поездок:

```
SELECT
  start_station_name
  , end_station_name
  , AVG(duration) AS avg_duration
  , COUNT(duration) AS num_trips
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire
  , UNNEST(stations) AS station
WHERE
  start_station_name = station
GROUP BY start_station_name, end_station_name
HAVING num_trips > MIN_TRIPS_THRESH
ORDER BY avg_duration DESC
LIMIT 5
```

Если число поездок невелико, среднее значение может оказаться сильно смещенным из-за отдельных выбросов. Поэтому мы выбираем пункты проката с наибольшим числом поездок, чтобы получить более достоверное среднее значение. Предложение `HAVING` в этом запросе отбрасывает пары пунктов проката, насчитывающих меньше 100 поездок:

Row	start_station_name	end_station_name	avg_duration	num_trips
1	Hyde Park Corner, Hyde Park	Abbey Orchard Street, Westminster	10718.507462686563	268
2	Wellington Arch, Hyde Park	Imperial College, Knightsbridge	10062.04724409449	127
3	Hyde Park Corner, Hyde Park	Westminster University, Marylebone	9726.05042016807	119
4	Park Lane, Hyde Park	Westbourne Grove, Bayswater	9712.5	104
5	Albert Gate, Hyde Park	Paddington Green Police Station, Paddington	9182.72727272727	132

Циклы

В скриптах можно также использовать условный оператор `IF` и различные примитивы для организации циклов. Порог 100 был выбран совершенно произвольно. Что получится, если изменить его? Иногда лучше попробовать несколько по-

рогов и посмотреть на результаты. Для этого можно изменить `MIN_TRIPS_THRESH` в цикле и поместить инструкцию `SELECT` в этот цикл.¹

```
WHILE MIN_TRIPS_THRESH < 1000 DO
  SELECT ...;
  SET MIN_TRIPS_THRESH = MIN_TRIPS_THRESH * 2;
END WHILE
```

Теперь, запустив сценарий, вы получите три набора результатов, последний из которых выглядит так:

Row	start_station_name	end_station_name	avg_duration	num_trips
1	Bayswater Road, Hyde Park	Bayswater Road, Hyde Park	4289.155172413791	3480
2	Hyde Park Corner, Hyde Park	Wellington Arch, Hyde Park	3817.0747740345105	2434
3	Knightsbridge, Hyde Park	Hyde Park Corner, Hyde Park	3582.595834591005	3313
4	Park Lane, Hyde Park	Park Lane, Hyde Park	3524.174474450833	12701
5	Hyde Park Corner, Hyde Park	Knightsbridge, Hyde Park	3479.189736664417	1481

Также поддерживается конструкция простого безусловного цикла. Например, цикл `WHILE` в предыдущем сценарии может быть реализован следующим образом:

```
LOOP
  IF MIN_TRIPS_THRESH >= 1000 THEN
    BREAK;
  END IF;

  SELECT MIN_TRIPS_THRESH;
  SET MIN_TRIPS_THRESH = MIN_TRIPS_THRESH * 2;
END LOOP;
```

Обратите внимание, как используется цикл `LOOP`, — он действует подобно циклу `WHILE` со всегда истинным условием. Выйти из цикла можно с помощью инструкции `BREAK`, как показано выше, а пропустить оставшуюся часть тела цикла итерации — с помощью инструкции `CONTINUE`.²

¹ Полный скрипт можно найти в файле https://github.com/GoogleCloudPlatform/bigquery-oreilly-book/blob/master/08_advqueries/script_loop.sql.

² Ключевые слова `BREAK` и `CONTINUE` действуют так же, как в языках, подобных языкам C (таких, как Python, C# или Java). В документации Matlab довольно хорошо объясняется применение инструкций `BREAK` (<https://www.mathworks.com/help/matlab/ref/break.html>) и `CONTINUE` (<https://www.mathworks.com/help/matlab/ref/continue.html>).

Хранимые процедуры

Готовый сценарий можно сохранить в наборе данных подобно пользовательским функциям (UDF). Такие сохраненные сценарии называют *хранимыми процедурами*. Вот пример определения процедуры:¹

```
CREATE OR REPLACE PROCEDURE ch08eu.sp_unusual_trips()
BEGIN

-- Начало сценария
CREATE TEMPORARY TABLE typical_trip AS
  SELECT
    start_station_name
    , end_station_name
    , APPROX_QUANTILES(duration, 10)[OFFSET(5)] AS typical_duration
    , COUNT(*) AS num_trips
  FROM
    `bigquery-public-data`.london_bicycles.cycle_hire
  GROUP BY
    start_station_name, end_station_name
;

CREATE OR REPLACE TABLE ch08eu.unusual_days AS
  SELECT
    EXTRACT (DATE FROM start_date) AS trip_date
    , APPROX_QUANTILES(duration / typical_duration, 10)[OFFSET(5)] AS ratio
    , COUNT(*) AS num_trips_on_day
  FROM
    `bigquery-public-data`.london_bicycles.cycle_hire AS hire
    , typical_trip AS trip
  WHERE
    hire.start_station_name = trip.start_station_name
    AND hire.end_station_name = trip.end_station_name
    AND num_trips > 10
  GROUP BY trip_date
  HAVING num_trips_on_day > 10
  ORDER BY ratio DESC
;

-- Конец сценария

END;
```

После определения хранимую процедуру можно вызвать, как показано ниже (обратите внимание на круглые скобки):

```
CALL ch08eu.sp_unusual_trips();
```

¹ См. 08_advqueries/stored_procedure_def.sql в репозитории GitHub с примерами для этой книги.

Параметры хранимых процедур

Хранимые процедуры похожи на пользовательские функции. Они могут принимать входные параметры и возвращать выходные параметры.¹ Например, давайте изменим скрипт, добавив параметр `MIN_TRIPS_THRESH` и возвращаемое значение с результатами, вместо сохранения в таблицу:²

```
CREATE OR REPLACE PROCEDURE ch08eu.sp_most_unusual(  
  IN MIN_TRIPS_THRESH INT64,  
  OUT result ARRAY<STRUCT<trip_date DATE, ratio FLOAT64, num_trips_on_day INT64>>)  
  
BEGIN  
  CREATE TEMPORARY TABLE typical_trip AS  
  SELECT  
    start_station_name  
  , end_station_name  
  , APPROX_QUANTILES(duration, 10)[OFFSET(5)] AS typical_duration  
  , COUNT(*) AS num_trips  
FROM  
  `bigquery-public-data`.london_bicycles.cycle_hire  
GROUP BY  
  start_station_name, end_station_name  
;  
  
SET result = (  
  WITH unusual_trips AS (  
    SELECT  
      EXTRACT (DATE FROM start_date) AS trip_date  
    , APPROX_QUANTILES(duration / typical_duration, 10)[OFFSET(5)] AS ratio  
    , COUNT(*) AS num_trips_on_day  
  FROM  
    `bigquery-public-data`.london_bicycles.cycle_hire AS hire  
    , typical_trip AS trip  
  WHERE  
    hire.start_station_name = trip.start_station_name  
    AND hire.end_station_name = trip.end_station_name  
    AND num_trips > MIN_TRIPS_THRESH  
  GROUP BY trip_date  
  HAVING num_trips_on_day > MIN_TRIPS_THRESH  
  )  
  SELECT  
    ARRAY_AGG(STRUCT(trip_date, ratio, num_trips_on_day)  
      ORDER BY ratio DESC LIMIT 3)  
  FROM unusual_trips  
);  
  
END;
```

¹ Параметры могут быть входными и выходными; они отмечаются как `IN` и `OUT` соответственно.

² См. `08_advqueries/stored_procedure_inout.sql` в репозитории GitHub с примерами для этой книги.

При вызове процедуры нужно передать необходимые параметры, объявив все необходимые для этого переменные:

```
DECLARE y ARRAY<STRUCT<trip_date DATE, ratio FLOAT64, num_trips_on_day INT64>>;  
CALL ch08eu.sp_most_unusual(10, y);  
SELECT y;
```

Вот что вернул этот вызов:

Row	y.trip_date	y.ratio	y.num_trips_on_day
1	2016-12-25	1.6111111111111112	34477
	2015-12-25	1.5161290322580645	20871
	2015-08-01	1.25	41200

Продвинутые функции

В этом разделе мы рассмотрим некоторые продвинутые функции, поддерживаемые в BigQuery: функции для анализа географических данных, вычисления статистик, а также для хеширования и генерации уникальных чисел. А в следующей главе мы познакомимся с машинным обучением.

Геоинформационная система BigQuery

Сколько таблиц в вашем хранилище данных содержит информацию о местоположении? Возможно, в течение продолжительного времени вы записываете широту и долготу места, где находятся транспортные средства или груз, или, может быть, у вас есть адреса ваших клиентов и поставщиков. Если вы фиксируете сделки с клиентами, то могли бы выполнить соединение с другой таблицей, хранящей местоположения магазинов.

Информация о местоположении не только очень распространена, но от нее зависят многие деловые решения. Выявление районов и городов, на которые следует обратить внимание при проведении рекламной кампании, является обычной задачей анализа в отделах маркетинга. Логистическим и производственным компаниям, занимающимся доставкой грузов и продукции, часто необходимо отслеживать местонахождение миллионов упаковок в своих цепочках поставок. Поэтому геопространственный анализ прочно укрепился в аналитике данных. Наш коллега Чед Дженнингс (Chad Jennings) недавно произвел поиск по GitHub¹ и обнаружил, что 9% всех *.sql*-файлов в нем выполняют геопространственные запросы, используя в своих исследованиях очень консервативное определение

¹ В BigQuery есть общедоступный набор данных всех файлов в GitHub.

понятия «геопространственный запрос». BigQuery обеспечивает простоту и эффективность анализа и визуализации геопространственных данных.

Традиционно системы, поддерживающие анализ географических данных, были довольно уникальными и назывались геоинформационными системами (Geographic Information Systems, GIS). В настоящее время многие универсальные аналитические инструменты, такие как BigQuery, поддерживают виды анализа, ранее требовавшие использования специализированных инструментов, а функции геопространственного анализа в них часто называют *GIS-функциями*.

Географические типы

GIS-функции оперируют географическими типами. BigQuery поддерживает точки, линии и многоугольники на поверхности Земли, используя в качестве основы эталонный эллипсоид WGS84.¹ Например, создать географический тип с долготой и широтой определенного пункта проката велосипедов в Лондоне можно с помощью `ST_GeogPoint` (обратите внимание, что параметр `longitude` с долготой стоит на первом месте):

```
SELECT
  name
  , ST_GeogPoint(longitude, latitude) AS location
FROM
  `bigquery-public-data`.london_bicycles.cycle_stations
WHERE
  id BETWEEN 300 and 305
```

В наборе с результатами столбец местоположения имеет тип `Point`, и в текстовом представлении он отображается в хорошо известном формате Well Known Text — WKT (<https://www.opengespatial.org/standards/wkt-crs>) — общепринятом стандарте представления географических точек в виде строк:

Row	name	location
1	Marylebone Lane, Marylebone	POINT(-0.148105415 51.51475963)
2	Serpentine Car Park, Hyde Park	POINT(-0.17306032 51.505014)
3	Albert Gate, Hyde Park	POINT(-0.158456089 51.50295379)
4	Kennington Lane Tesco, Vauxhall	POINT(-0.115853961 51.48677988)
5	Putney Pier, Wandsworth	POINT(-0.216573 51.466907)

¹ Поскольку Земля не имеет форму идеальной сферы, существует множество возможных ее приближений, и разные эллипсоидальные аппроксимации могут быть более точными в разных частях мира. Глобальная система позиционирования (Global Positioning System, GPS) использует WGS84; поэтому почти все координаты местоположений, с которыми вы столкнетесь, будут относиться к WGS84.

При загрузке геопространственных данных из других систем можно заметить, что географические данные вводятся либо в формате WKT, либо в формате GeoJSON — еще одном стандартном строковом формате представления географических координат. Если система-источник дает вам возможность выбора, примите решение в пользу GeoJSON вместо WKT, потому что GeoJSON точно определяет, является ли интересующий многоугольник «внутренним» или «внешним»¹ (и те и другие имеют одинаковые координаты), а также определяет необходимость тесселяции ребер.² Если геопространственные данные имеют какой-то другой формат, например Shapefiles, воспользуйтесь инструментом с открытым исходным кодом `ogr2ogr`, чтобы преобразовать данные в GeoJSON перед загрузкой в BigQuery.³



Как мы уже говорили в главе 7, GIS-запросы в BigQuery работают гораздо эффективнее, если географические данные хранятся в виде географических типов, а не в виде примитивов (например, долготы и широты) или строк.

Если в вашем наборе данных географические координаты хранятся в виде примитивов или строк, преобразуйте их в географические типы в конвейере извлечения, загрузки и преобразования (ELT) или ETL:

```
CREATE OR REPLACE TABLE ch08eu.cycle_stations AS
SELECT
  *, ST_GeogPoint(longitude, latitude) AS location
FROM
  `bigquery-public-data`.london_bicycles.cycle_stations
```

Чтобы преобразовать строки в формате WKT или GeoJSON в географические типы, используйте функцию `ST_GeogFromText` или `ST_GeogFromGeoJSON` соответственно.

Чтобы преобразовать географические объекты в строки в формате WKT или GeoJSON, используйте функцию `ST_AsText` или `ST_AsGeoJSON` соответственно:

¹ При парсинге WKT имеет место неоднозначность интерпретации многоугольника. По умолчанию BigQuery предполагает, что подразумевается многоугольник меньшего размера. `ST_GeogFromText` поддерживает параметр `oriented`, в котором можно передать значение `TRUE` и тем самым подсказать, что вершины многоугольника перечислены в направлении против часовой стрелки.

² Под *тесселяцией* понимается требование, согласно которому многоугольники полностью должны заполнять плоскость без перекрытий и пропусков — например, все многоугольники, описывающие границы районов в штате, должны заполнять всю площадь штата. Эта проблема отсутствует в GeoJSON, потому что края района уже находятся на плоскости, но при анализе WKT BigQuery должна тесселировать полученные координаты, из-за чего точки могут смещаться на расстояние до 10 метров.

³ Дополнительную информацию можно найти по ссылке <https://oreil.ly/jZVpC>.

```

SELECT
    name
    , ST_AsGeoJSON(location) AS location_string
FROM
    ch08eu.cycle_stations
WHERE
    id BETWEEN 300 and 305

```

Этот запрос вернет следующие результаты:

Row	name	location_string
1	Marylebone Lane, Marylebone	{ "type": "Point", "coordinates": [-0.148105415, 51.51475963] }
2	Serpentine Car Park, Hyde Park	{ "type": "Point", "coordinates": [-0.17306032, 51.505014] }
3	Albert Gate, Hyde Park	{ "type": "Point", "coordinates": [-0.158456089, 51.50295379] }
4	Kennington Lane Tesco, Vauxhall	{ "type": "Point", "coordinates": [-0.115853961, 51.48677988] }
5	Putney Pier, Wandsworth	{ "type": "Point", "coordinates": [-0.216573, 51.466907] }

Создание многоугольников

Предположим, что у вас есть поездка от пункта проката 300 до пункта проката 305, затем до пункта проката 302 и, наконец, обратно до пункта проката 300. Вы можете использовать `ST_MakeLine` и `ST_MakePolygon` для представления одного сегмента поездки или пути туда и обратно соответственно:

```

WITH stations AS (
SELECT
    (SELECT location FROM ch08eu.cycle_stations WHERE id = 300)
      AS loc300,
    (SELECT location FROM ch08eu.cycle_stations WHERE id = 302)
      AS loc302,
    (SELECT location FROM ch08eu.cycle_stations WHERE id = 305)
      AS loc305
)
SELECT
    ST_MakeLine(loc300, loc305) AS seg1
    , ST_MakePolygon(ST_MakeLine(
        [loc300, loc305, loc302])) AS poly
FROM
    stations

```

Этот запрос вернет следующие результаты:

Row	seg1	poly
1	LINESTRING(-0.17306032 51.505014, -0.115853961 51.48677988)	POLYGON((-0.216573 51.466907, -0.115853961 51.48677988, -0.17306032 51.505014, -0.216573 51.466907))

ГЕОГРАФИЧЕСКИЕ ДАННЫЕ В МАШИННОМ ОБУЧЕНИИ

Большинство систем машинного обучения, работающих со структурированными данными, могут оперировать только числовыми или категориальными переменными. Использование названий штатов или стран может не дать нужной детализации, поэтому предпочтительнее использовать точное местоположение. Не преобразуйте географические местоположения в формат WKT или GeoJSON для предоставления географического местоположения в категориальных признаках, лучше используйте геохеш, потому что символы в хеше передают географическую близость.

Для примера взгляните на первые несколько букв геохешей центральных точек близлежащих почтовых индексов в штате Аляска:

```
SELECT
  state_code
  , zip_code
  , ST_GeoHash(internal_point, 2) AS ziphash_2
  , ST_GeoHash(internal_point, 5) AS ziphash_5
  , ST_GeoHash(internal_point, 10) AS ziphash_10
FROM
  `bigquery-public-data`.geo_us_boundaries.us_zip_codes
WHERE
  state_code = 'AK'
ORDER BY ziphash_10 ASC
LIMIT 5
```

Этот запрос вернет следующие результаты:

Row	state_code	zip_code	ziphash_2	ziphash_5	ziphash_10
1	AK	99546	b1	b14qu	b14queqr8k
2	AK	99547	b1	b1k15	b1k158vcqn
3	AK	99638	b1	b1rug	b1rugtepv7
4	AK	99685	b3	b39d7	b39d7x4cgz
5	AK	99692	b3	b39dd	b39dd3d7xf

Оба индекса, 99546 (Адак) и 99638 (Никольский), принадлежат местностям в цепочке Алеутских островов, и пространственная близость фиксируется совпадением двух первых букв геохеша (b1). Еще более близкими являются четвертая и пятая области в примере с результатами: область 99692 (Датч Харбор) фактически окружена областью 99685 (<https://www.unitedstateszipcodes.org/99685>, Уналашка), и факт их пространственной близости подтверждается совпадением первых четырех букв геохеша (b39d). Обратите внимание, что в обоих случаях пространственная близость не определяется по числовым значениям почтовых индексов.

Проще говоря, лучшим способом включения географических точек из BigQuery в модели машинного обучения является передача первых нескольких символов их геохешей. Выбор количества символов зависит от требуемой точности представления местоположений. Для организации обучения с несколькими разрешениями можно передать три отдельные категориальные переменные: в первой — один первый символ геохеша, во второй — два первых символа, в третьей — три первых символа.

Обратите внимание на то, как конструируется многоугольник, и на то, что линия может иметь несколько сегментов — многоугольник всегда замкнут, а линия может оставаться открытой.

GIS-предикативные функции

BigQuery поддерживает ряд пространственных *предикативных функций*. Эти функции можно использовать в предложениях `WHERE` и `JOIN`. Например, можно найти почтовые индексы, в границах которых имеется больше всего пунктов проката сети New York Citibike, находящихся на расстоянии до одного километра (1000 метров) от центральной точки почтового индекса:

```
SELECT
  z.zip_code
  , COUNT(*) AS num_stations
FROM
  `bigquery-public-data`.new_york.citibike_stations AS s,
  `bigquery-public-data`.geo_us_boundaries.us_zip_codes AS z
WHERE
  ST_DWithin(
    z.zcta_geom,
    ST_GeogPoint(s.longitude, s.latitude),
    1000) -- 1 км
GROUP BY z.zip_code
ORDER BY num_stations DESC
LIMIT 5
```

Основой этого запроса является функция `ST_DWithin` (или `distance within` — расстояние в пределах). Вот как выглядит результат:

Row	zip_code	num_stations
1	11201	116
2	11217	112
3	10003	112
4	11238	103
5	10011	95

Точно так же, используя `ST_Intersects`, `ST_Contains` или `ST_CoveredBy`, можно проверить пересечение двух геометрий, включение одной в другую или охват одной геометрии другой. Есть и другие пространственные предикативные функции (https://cloud.google.com/bigquery/docs/reference/standard-sql/geography_functions).

GIS-метрики

Одной из самых полезных GIS-функций является `ST_Distance`, которая возвращает расстояние между двумя географическими точками. Например, вот как можно найти расстояние между Сиэтлом и Майами:

```
WITH seattle AS (  
  SELECT ANY_VALUE(internal_point) as loc  
  FROM `bigquery-public-data`.geo_us_boundaries.us_zip_codes  
  WHERE city = 'Seattle' and state_code = 'WA'  
)  
,  
miami AS (  
  SELECT ANY_VALUE(internal_point) as loc  
  FROM `bigquery-public-data`.geo_us_boundaries.us_zip_codes  
  WHERE city = 'Miami city' and state_code = 'FL'  
)  
  
SELECT  
  ST_Distance(seattle.loc, miami.loc)/1000 AS dist  
FROM seattle, miami
```

В результате у нас получилось 4364 километра. Ваш результат может немного отличаться из-за использования `ANY_VALUE` для выбора любого из почтовых индексов в двух городах.



Чтобы не выходить за рамки политики конфиденциальности, действующей в вашей организации, вам часто придется увеличивать или уменьшать разрешение данных о местоположении. `ST_SnapToGrid` позволяет округлять координаты местоположения. Например, `ST_SnapToGrid(pt, 0.01)` округлит широту и долготу `pt` до второго знака после запятой. Эта функция также может применяться к многоугольникам и линиям и производит правильные вычисления (например, отрезает сегменты линий, теряющие смысл после округления координат вершин).

Преобразование и агрегирование геометрии

В BigQuery также есть множество функций для вычисления объединений, пересечений и выполнения других операций с географическими областями.

Запрос в предыдущем разделе выбирал произвольный почтовый индекс в городе и использовал в расчетах только центр многоугольника. Иногда лучше объ-

единить все многоугольники почтовых индексов в обоих городах с помощью `ST_UNION` и вычислить расстояние между получившимися объединениями:

```
WITH seattle AS (  
  SELECT ST_UNION(ARRAY_AGG(zcta_geom)) as loc  
  FROM `bigquery-public-data`.geo_us_boundaries.us_zip_codes  
  WHERE city = 'Seattle' and state_code = 'WA'  
)  
,  
miami AS (  
  SELECT ST_UNION(ARRAY_AGG(zcta_geom)) as loc  
  FROM `bigquery-public-data`.geo_us_boundaries.us_zip_codes  
  WHERE city = 'Miami city' and state_code = 'FL'  
)  
  
SELECT  
  ST_Distance(seattle.loc, miami.loc)/1000 AS dist  
FROM seattle, miami
```

Мы получили результат 4356, и на этот раз он более детерминированный: это расстояние от самого юго-восточного угла Сиэтла до самого северо-западного угла Майами.

На практике часто возникает потребность объединения географических областей, поэтому в BigQuery была добавлена функция `ST_UNION_AGG`:

```
WITH seattle AS (  
  SELECT ST_UNION_AGG(zcta_geom) as loc  
  FROM `bigquery-public-data`.geo_us_boundaries.us_zip_codes  
  WHERE city = 'Seattle' and state_code = 'WA'  
)
```

Вычислить центральную точку совокупности географических областей можно с помощью `ST_CENTROID_AGG`. То есть вот так можно найти расстояние между геометрическими центрами Сиэтла и Майами:

```
WITH seattle AS (  
  SELECT ST_CENTROID_AGG(zcta_geom) as loc  
  FROM `bigquery-public-data`.geo_us_boundaries.us_zip_codes  
  WHERE city = 'Seattle' and state_code = 'WA'  
)  
,  
miami AS (  
  SELECT ST_CENTROID_AGG(zcta_geom) as loc  
  FROM `bigquery-public-data`.geo_us_boundaries.us_zip_codes  
  WHERE city = 'Miami city' and state_code = 'FL'  
)  
  
SELECT  
  ST_Distance(seattle.loc, miami.loc)/1000 AS dist  
FROM seattle, miami
```

Этот запрос вернет в результате 4363 километра.

Для визуализации геопространственных данных можно использовать BigQuery Geo Viz или Jupyter Notebook (описывается в главе 5).

Полезные статистические функции

BigQuery поддерживает вычисление статистик по петабайтным наборам данных. Например, можно вычислить среднее значение, стандартное отклонение и процентиля по столбцу, а также корреляцию Пирсона между парой столбцов.

Статистики

Вот пример запроса, вычисляющего статистики по столбцу `duration` в наборе данных `london_bicycles`:

```
SELECT
  MIN(duration) AS min_dur
  , MAX(duration) AS max_dur
  , COUNT(duration) AS num_dur
  , AVG(duration) AS avg_dur
  , SUM(duration) AS total_dur
  , STDDEV(duration) AS stddev_dur
  , VARIANCE(duration) AS variance_dur
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire
```

Он возвращает следующие результаты:

Row	min_dur	max_dur	num_dur	avg_dur	total_dur	stddev_dur	variance_dur
1	-3540	2674020	24369201	1332.29	32466946080	9827.99	9.66E7

Минимальное (-3540?) и максимальное (2 674 020 секунд = 31 день!) значения длительности почти наверняка представляют ошибочные наблюдения. Соответственно, средняя продолжительность не является доверительной, потому что включает в себя эти ошибочные значения.

Квантили

Более надежной статистической оценкой для столбцов со значительными выбросами является медиана. В BigQuery ее можно получить с помощью `APPROX_QUANTILES`. Эта функция позволяет указать количество квантилей, поэтому используем 3:

```
SELECT
  APPROX_QUANTILES(duration, 3)
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire
```

Этот запрос вернет следующие результаты:

Row	f0_
1	-3540
	600
	1080
	2674020

Почему в результате получилось четыре числа? Потому, что мы запросили три квантиля, а для описания границ трех квантилей требуется четыре числа: (–3540, 600), (600, 1080) и (1080, 2 674 020). То есть число –3540 — это минимальная продолжительность, а 2 674 020 — максимальная. Одна треть поездок длилась 600 секунд или меньше. Верхняя треть поездок продолжалась 1080 секунд или дольше. Средняя треть — от 600 до 1080 секунд.

Чтобы найти медиану, или среднюю точку распределения, можно запросить два квантиля:

```
SELECT
  APPROX_QUANTILES(duration, 2)
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire
```

Этот запрос вернет:

Row	f0_
1	-3540
	840
	2674020

Согласно этим результатам медиана приблизительно равна 840 секундам.¹ Конечно, можно выполнить более дробное квантование и выбрать, скажем, 95-й процентиль² длительностей:

```
SELECT
  APPROX_QUANTILES(duration, 100)[OFFSET(95)]
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire
```

¹ Обратите внимание, что полученная нами медиана является приблизительным квантилем. Ожидаемая ошибка составляет $\pm 1\%$.

² 95-й и 99-й процентиля часто используются для моделирования «наихудших» сценариев.

Этот запрос вернет:

Row	f0_
1	3000

Согласно этому результату, 95% поездок длится меньше 3000 секунд.

Корреляция

Есть ли корреляция между близостью пунктов проката к центру Лондона и продолжительностью поездок, начинающихся в этих пунктах проката? Чтобы ответить на этот вопрос, создадим столбец, отражающий расстояние от пункта проката до центра города, и еще один — отражающий среднюю продолжительность поездок, а затем вычислим корреляцию:

```
WITH distances AS (  
    SELECT  
        id  
        , ST_Distance(location, ST_GeogPoint(-0.12574, 51.50853)) AS distance  
    FROM  
        ch08eu.cycle_stations  
) ,  
  
durations AS (  
    SELECT  
        start_station_id AS id  
        , APPROX_QUANTILES(duration, 2)[OFFSET(1)] AS median_duration  
    FROM  
        `bigquery-public-data`.london_bicycles.cycle_hire  
    GROUP BY start_station_id  
)  
  
SELECT CORR(distance, median_duration) AS pearson  
FROM distances  
JOIN durations  
USING(id)
```

В результате получится число 0.14, сообщающее об отсутствии значимой корреляции между расстоянием от центра города и продолжительностью поездок. Коэффициент корреляции 1.0 соответствует прямо пропорциональной связи, а -1.0 — обратно пропорциональной. Если две переменные являются линейно-независимыми, коэффициент корреляции между ними будет равен нулю.

Алгоритмы хеширования

Во многих задачах анализа часто бывает необходимо сгенерировать короткую строку, уникально представляющую запись. Например, представьте, что вам нужно

быстро выявить повторяющиеся записи, не выполняя дорогостоящее сравнение каждого столбца. Сделать это можно двумя способами: вычислить уникальный «отпечаток» значений в записи или присвоить записи универсально-уникальный идентификатор (Universally Unique Identifier, UUID) в момент ее создания. Конечно, нет никакой гарантии, что один и тот же отпечаток не будет присвоен двум совершенно разным наборам значений, но вероятность этого очень мала. Точно так же нет гарантии, что один и тот же UUID не будет сгенерирован какой-либо другой системой где-то еще. Тем не менее вероятность этого события очень мала.

То есть даже если вы не собираетесь использовать отпечатки или UUID для криптографии или управления денежными средствами, вы можете использовать их, чтобы не посчитать один и тот же велосипед дважды. Есть небольшая теоретическая вероятность, что некий клиент получит бесплатную поездку, но никому из клиентов не придется платить дважды.¹ Впрочем, для тех, кому это необходимо, BigQuery поддерживает популярные алгоритмы шифрования (MD5 и SHA).

Функции создания отпечатков

Для получения отпечатка BigQuery использует алгоритм вычисления хеш-кода FARM (<https://github.com/google/farmhash>) с открытым исходным кодом на нескольких языках программирования (что позволяет в разных системах получить один и тот же хеш-код для тех же данных). Функция принимает аргумент типа STRING и возвращает INT64.

Мы знаем, что поездку можно однозначно идентифицировать по `bike_id`, времени начала и пункту проката, где был арендован велосипед. Вместо сравнения всех трех значений, когда потребуется проверить соответствие двух записей одной и той же поездке, можно просто сравнить отпечатки двух записей (конечно, при условии, что вы уже их вычислили). Вот как можно вычислить эти отпечатки:

```
WITH identifier AS (
  SELECT
    CONCAT(
      CAST(bike_id AS STRING), '***',
      CAST(start_date AS STRING), '***',
      CAST(start_station_id AS STRING)
    ) AS rowid
  FROM `bigquery-public-data.london_bicycles.cycle_hire`
  LIMIT 10
)

SELECT
  rowid, FARM_FINGERPRINT(rowid) AS fingerprint
FROM identifier
```

¹ Деньги на вычисления, которые вы сэкономите благодаря тому, что вам не придется сравнивать каждое поле, или за счет создания по-настоящему уникальной и глобально согласованной инфраструктуры (например, Cloud Spanner!), компенсируют стоимость этой случайной бесплатной поездки на велосипеде.

ПОВТОРЯЕМАЯ ВЫБОРКА ДЛЯ МАШИННОГО ОБУЧЕНИЯ

Алгоритм получения отпечатка часто применяется для создания повторяемых хешей записей в наборах данных, предназначенных для машинного обучения. Отпечатки можно использовать для деления набора данных на контрольную и обучающую выборки. Например, чтобы разделить поездки по пунктам проката и дням, можно вычислить хеши для этих двух полей и использовать их для деления набора данных:

```
WITH datasets AS (  
  SELECT  
    CONCAT(  
      CAST(start_station_id AS STRING),  
      CAST(EXTRACT(DATE FROM start_date) AS STRING))  
  AS key,  
    *  
  FROM `bigquery-public  
data.london_bicycles.cycle_hire`  
  LIMIT 100  
)  
  
SELECT  
  IF(MOD(ABS(FARM_FINGERPRINT(key)), 10) < 8, 'train', 'test') AS ds,  
  * EXCEPT(key)  
FROM datasets  
LIMIT 10
```

Вот что возвращает этот запрос (в реальности столбцов больше — здесь показана только часть из них; но обратите внимание, что первый столбец идентифицирует записи как обучающие (train) или контрольные (test)):

Row	ds	rental_id	duration	bike_id	end_station_id	end_station_name
1	train	63022577	600	8930	55	Finsbury Circus, Liverpool Street
2	test	41340757	540	9260	58	New Inn Yard, Shoreditch
3	train	47466877	2040	11137	210	Hinde Street, Marylebone
4	test	53171329	5760	7033	733	Park Lane, Mayfair
5	train	42268703	900	7005	248	Triangle Car Park, Hyde Park

Этот запрос вернет следующие результаты:

Row	rowed	fingerprint
1	8168***2016-09-15 10:09:00+00***176	6524654244988303787
2	7218***2016-06-08 18:49:00+00***114	-4994312061947208007
3	3648***2015-07-23 13:21:00+00***304	3924490378672877823
4	9403***2017-03-14 18:43:00+00***574	-1509385442790305242

Row	rowed	fingerprint
5	10704***2016-08-16 20:58:00+00***223	–8271736518219415928
6	6048***2017-03-31 08:25:00+00***632	6083880842645302266
7	14039***2017-03-06 19:29:00+00***529	–5809138520111495006
8	7956***2015-07-27 09:55:00+00***14	–4999466933100478693
9	5744***2015-01-27 09:40:00+00***341	–8567341349676429749
10	13088***2016-08-06 22:50:00+00***29	6415473001431984902

MD5 и SHA

BigQuery также поддерживает популярные алгоритмы хеширования. Эти односторонние алгоритмы могут пригодиться для представления информации идентификации личности (Personally Identifiable Information, ПИ), правда, следует отметить, что не все алгоритмы хеширования обеспечивают одинаковый уровень безопасности и вы должны учитывать это, выбирая алгоритм для хеширования своих данных (о шифровании мы поговорим в главе 10):

```
SELECT
  name
  , MD5(name) AS md5_
  , SHA256(name) AS sha256_
  , SHA512(name) AS sha512_
FROM UNNEST(['Joe Customer', 'Jane Employee']) AS name
```

Этот запрос вернет следующие результаты:

Row	name	md5_	sha256_	sha512_
1	Joe Customer	9JFfot7XXNa9 IFXrZYpklQ==	ITPGdZjjJNgvrYfvHRP2HX ofhTntHalPMAn5tdA4AY8=	ysAXoRHTb+ENWL9jB2pCD1 arBasmuush7KJVasKWMbz1v zyUKHUS5CDI9jBNR3yxBDwRFL SQbHwPLklBuLptQ==
2	Jane Employee	g6HbGfBF02V JLdJoXs8tXQ==	wXJxfwK/hP4dgjQuz lcPOLVZEryACurXmL7qM cnC3tE=	N9tGIXX6AibvHpDNaZciAMHSYK/9/ nA9886fVkcPwykLONRlpiIM 7zE25yUZy6RSEpVKM+sdM +lcsG682qtj2Q==

UUID

Чтобы сгенерировать универсально-уникальную строку из 32 шестнадцатеричных цифр, которая вряд ли будет сгенерирована в другой системе, используйте функцию `GENERATE_UUID`:

```
SELECT GENERATE_UUID() AS uuid;
```

Запустив этот запрос, мы получили следующий результат (у вас, конечно же, получится совсем другая строка):

Row	uuid
1	5ae248e9-5872-410f-862f-8a27bb527b53

Генератор случайных чисел

В BigQuery также есть свой генератор случайных чисел, который можно с успехом использовать в видах анализа, требующих перемешивания или добавления шума. Вот как сгенерировать равномерно распределенное случайное число в диапазоне от 0 до 1:

```
SELECT RAND()
```

Выводы

В этой главе мы рассмотрели способы повторного использования кода на SQL. Запросы могут иметь именованные или позиционные параметры, и эти параметры можно передавать во время выполнения. Часто используемый код SQL можно выделить в функции, временные (доступные только текущему запросу) или хранимые в наборе данных и доступные любому, у кого есть доступ к набору данных для чтения. Другой подход, способствующий повторному использованию и читабельности, основан на применении предложения WITH и подзапросов.

Мы также поближе познакомились с массивами, обсудили ситуации, когда можно использовать массивы: для сохранения порядка, для хранения повторяющихся полей или для генерации данных. Мы также рассмотрели оконные функции, вычисляющиеся подмножеством таблицы, например, для расчета скользящих средних, для навигации («следующие три строки») и для нумерации (чтобы найти первую, последнюю и т. д.).

Мы представили примеры использования табличных метаданных для динамического конструирования запросов и подробно описали инструкции DDL и DML. Обсудили поддержку скриптов в BigQuery, обеспечивающую как простое выполнение линейной последовательности инструкций, так и позволяющую реализовать более сложные алгоритмы, например, с применением циклов.

Наконец, мы обсудили GIS-функции BigQuery, показали, как создавать данные географических типов, вычислять GIS-предикативные функции, преобразовывать существующие географические объекты с помощью ST_Union и многое другое. В заключение мы представили обзор статистических функций и алгоритмов хеширования.

Машинное обучение в BigQuery

Искусственный интеллект (ИИ) — это область computer science, основная цель которой заключается в создании вычислительных систем, способных действовать автономно. За прошедшие годы в сфере ИИ возникло много разных подразделов, но наибольший успех был достигнут в подходе, основанном на использовании больших наборов данных для обучения универсальных моделей (таких, как деревья решений и нейронные сети), способных решать сложные задачи с большой точностью.

Обучение компьютера на наборах примеров называется *машинным обучением с учителем*, и этот способ обучения на хранимых данных поддерживается в BigQuery. В этой главе мы посмотрим, как решать широкий спектр задач машинного обучения с помощью BigQuery ML. Несмотря на возможность машинного обучения в BigQuery, использование эффективных фреймворков машинного обучения, таких как TensorFlow, и данных из BigQuery может открыть доступ к гораздо более широкому разнообразию моделей и компонентов машинного обучения. Поэтому в этой главе мы также рассмотрим связи между BigQuery и полноценными фреймворками машинного обучения.

Что такое машинное обучение?

Если собрать архивные данные (а для чего еще нужно хранилище данных, если не для этого?) и эти данные будут содержать правильные ответы (называемые *метками*), мы сможем обучить модели машинного обучения на этих данных прогнозировать результаты для случаев, когда метки еще неизвестны. Например, при наличии набора данных о фактических продажах можно обучить модели машинного обучения прогнозировать продажи в будущем. Как и в случае анализа данных, машинное обучение в BigQuery также выполняется в SQL.

Формулировка задачи машинного обучения

Допустим, вы управляете несколькими сотнями кинотеатров по всей стране и хотите спрогнозировать, сколько билетов будет продано за определенное время показа в конкретном кинотеатре, — это может очень пригодиться при планировании графика показа фильмов. При наличии статистических данных о показе фильмов задачу машинного обучения можно сформулировать следующим образом: на основе информации о прокате, имеющейся в наборе данных, узнать, сколько билетов было продано на каждый сеанс в каждом кинотеатре, затем применить полученную модель к фильму-кандидату и определите, каким будет спрос на этот фильм в конкретное время.

Атрибуты фильма, которые мы используем в качестве входных данных в модели машинного обучения, называются *признаками* модели. Метка — это прогнозное значение, которое требуется узнать, и в данном случае метка — это количество проданных билетов. Ниже приводятся некоторые примеры признаков, которые вы можете включить в свою модель:

- Оценка содержимого кинофильма¹ (например, PG-13 означает, что фильм рекомендуется для детей младше 13 лет).
- Фильм будет показан в рабочие или выходные дни?
- В какое время дня будет демонстрироваться фильм (днем, вечером, ночью)?
- Жанр фильма (комедия, триллер и т. д.).
- Как давно снят фильм (в днях).
- Средний рейтинг фильма по оценкам критиков (по шкале от 1 до 10).
- Общая сумма кассовых сборов за предыдущий фильм этого же режиссера, если применимо.
- Общая сумма кассовых сборов за предыдущий фильм с этим же актером в главной роли, если применимо.
- Местоположение кинотеатра.
- Тип кинотеатра (например, большой кинотеатр с несколькими залами, кинотеатр под открытым небом для автомобилистов, кинозал в торговом центре и т. д.).

Обратите внимание, что название фильма само по себе не является хорошим обучающим признаком.² Даже если фильм «Шпион, выйди вон!» («Tinker Tailor

¹ См. https://en.wikipedia.org/wiki/Motion_picture_content_rating_system.

² Отдельные слова в названии фильма могут оказаться неплохими признаками, если применить к названиям фильмов методы обработки естественного языка, такие как лексемизация, морфологический поиск и получение векторных представлений слов. Также могут пригодиться признаки, вычисленные из названий фильмов; например,

Soldier Spy») 2011 года будет присутствовать в наборе обучающих данных, нам будет неинтересно прогнозировать прокат именно этого фильма (потому что он уже был показан в наших кинотеатрах). Гораздо интереснее спрогнозировать сборы, скажем, от проката фильма «Глубоководный горизонт» («Deep Water Horizon»), еще одного триллера с похожими отзывами критиков, выпущенного в 2016 году.

То есть модель машинного обучения должна опираться на особенности фильма (характеризующие фильм), а не на признаки, однозначно его идентифицирующие. В этом случае модель сможет догадаться, что при одинаковых условиях показа фильм «Глубоководный горизонт» принесет такую же прибыль, как и «Шпион, выйди вон!», потому что они относятся к одному жанру и имеют похожий рейтинг.

Первые четыре признака (рейтинг, дни показа, время показа, жанр) являются категориальными, то есть принимают одно из конечного числа возможных значений. В BigQuery категориальным считается любой признак, являющийся строкой. Если в базе данных категориальные признаки представлены значениями другого типа (например, признак времени показа может быть числом, например 1430, или отметкой времени), в запросе они должны преобразовываться в строки. Следующие четыре признака (время с момента выпуска, рейтинги критиков, кассовые сборы за предыдущий фильм этого же режиссера и за предыдущий фильм с этим же актером в главной роли) — числовые, то есть имеют значимое числовое выражение. Последние два признака (тип и местоположение кинотеатра) должны быть представлены особым образом; возможные варианты мы обсудим далее в этой главе.

Меткой, или правильным прогнозом, является количество проданных билетов. Во время обучения модели машинного обучения BigQuery передаются входные признаки и соответствующие метки, на основе которых создается модель, обобщающая эту информацию (рис. 9.1). Затем, во время прогнозирования, обученную модель можно применить к новому набору входных признаков и получить оценку количества билетов, которое можно продать, если запланировать показ фильма в определенное время и в определенном месте.

Типы задач машинного обучения

Мы часто используем разные модели и методы машинного обучения в зависимости от характера входных признаков и меток. В этом подразделе мы кратко перечислим поддерживаемые типы задач машинного обучения, а в оставшейся части главы подробно рассмотрим способы их решения.

хорошим прогнозирующим элементом может послужить длина названия или наличие в нем слова «шпион».

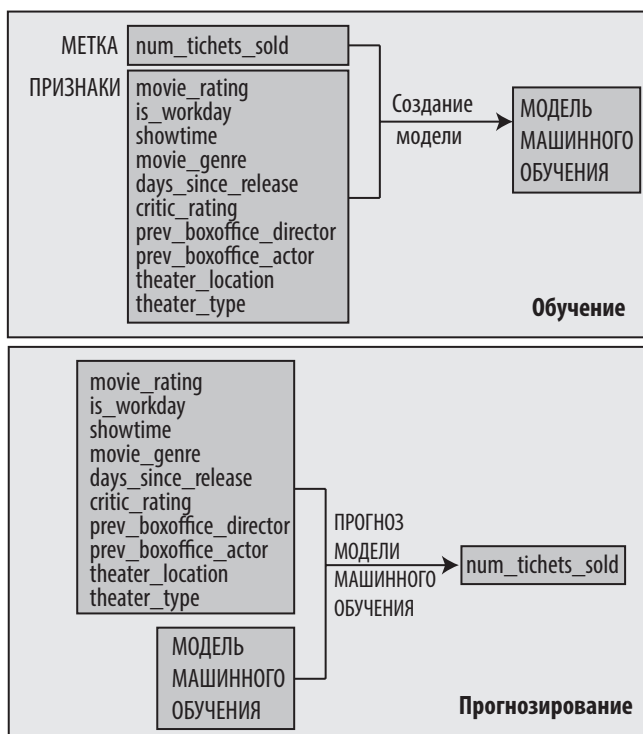


Рис. 9.1. Во время обучения модели машинного обучения передаются входные признаки и соответствующие метки. Затем обученную модель можно использовать для прогнозирования. На основе набора входных признаков модель способна спрогнозировать метку

Регрессия

В примере, описанном в предыдущем разделе, требовалось спрогнозировать количество проданных билетов на конкретный фильм. В данном случае метка является числом, и задачи машинного обучения подобного типа называют *регрессией*.

Классификация

Если результатом (меткой) задачи машинного обучения является категориальная переменная, такие задачи относятся к разряду задач *классификации*. Модели классификации возвращают вероятность принадлежности записи к категории (классу), определяемой меткой. Например, чтобы получить модель машинного обучения, прогнозирующую успех шоу, вы могли бы создать модель классификации, возвращающую вероятность успеха.

Многие задачи классификации имеют два класса, например: успех или провал шоу, высокий или низкий спрос на товар, прибытие рейса вовремя или с опозданием. Это задачи так называемой *бинарной классификации*. В таких случаях столбец метки должен иметь значение True или False либо 1 или 0. Прогнозом такой модели является вероятность, что метка будет иметь значение True. Чтобы определить наиболее вероятный класс, мы обычно устанавливаем пороговую вероятность равной 0.5.

Задача классификации может иметь несколько классов. Например, возвращаясь к нашему сценарию проката велосипедов, может понадобиться предсказать пункт проката, куда будет возвращен велосипед, и поскольку для этой категориальной метки существуют сотни возможных значений, эта задача относится к разряду задач *мультиклассовой (множественной) классификации*. Результатом такой модели машинного обучения является набор вероятностей, по одной для каждого пункта проката, и сумма этих вероятностей будет равна 1.0. В задачах множественной классификации, как правило, наибольший интерес представляют три или пять наибольших предсказаний, а не фактические значения вероятностей.

Рекомендательная система

Особый случай многоклассовой классификации, когда цель состоит в том, чтобы рекомендовать «следующий» продукт на основании рейтингов или предыдущих покупок, называется *рекомендательной системой*. Как и любую другую задачу множественной классификации, задачу подбора рекомендаций можно решить с помощью стандартного подхода. Тем не менее для таких задач были созданы специальные типы моделей машинного обучения, и для реализации рекомендаций лучше использовать именно их. Рекомендательные системы также лучше подходят для решения задач таргетинга — поиска клиентов, которым понравится продукт или рекламное предложение.

Кластеризация

Если в наборе исходных данных вообще нет меток, к ним не получится применить методы машинного обучения с учителем. Зато можно попробовать выявить сложившиеся группы в данных; задачи этого вида называются *кластеризацией (группировкой)*. Например, клиентов можно разделить на кластеры (группы) в зависимости от характеризующих их признаков. Также можно воспользоваться службой Cloud Data Labeling Service и прибегнуть к помощи людей, осуществляющих маркировку данных, перед проведением обучения с учителем.

Неструктурированные данные

До сих пор мы исходили из того, что данные могут быть структурированными или полуструктурированными. Если какие-то входные признаки не структу-

рированы (например, являются изображениями или текстом на естественном языке), их можно обработать с помощью готовых моделей, таких как Cloud Vision API или Cloud Natural Language, и использовать полученные результаты в виде числовых или категориальных входных признаков. Например, с помощью Cloud Natural Language API можно определить ключевые сущности в электронных письмах клиентов или эмоциональную окраску их отзывов и использовать полученные сущности в качестве категориальных переменных, а оценки эмоциональной окраски — в качестве числовых признаков.

Преобразовать неструктурированные данные в структурированные можно с помощью строковых функций или машинного обучения. На практике часто используется прием под названием «*мешок слов*», суть которого состоит в том, чтобы разделить текстовое поле на отдельные слова и интерпретировать наличие/отсутствие отдельных слов как признаки. Например, из названия фильма «Шпион, который меня любил» («The Spy Who Loved Me») можно выделить два признака со значением True: `has_spy` и `has_love`. Все остальные признаки будут иметь значение False (слова «the», «Who» и «Me» лучше отбросить, потому что они слишком часто встречаются в тексте и не имеют большого значения для прогноза). Также в качестве признака можно использовать количество слов в названии (фильмы с большим количеством слов в названии, скорее всего, окажутся авторским кино и могут быть привлекательными для разных аудиторий).

Если не структурирована сама метка (например, требуется, чтобы модель генерировала идеальные ответы на вопросы клиентов, опираясь на набор архивных ответов), такие задачи относят к разряду задач генерирования естественного языка — они не поддерживаются в BigQuery.

Краткая сводка по типам моделей

В табл. 9.1 перечисляются типы задач машинного обучения. В следующем разделе мы обсудим типы моделей, поддерживаемые в BigQuery.

Таблица 9.1. Типы моделей машинного обучения и как они реализуются в BigQuery

Описание задачи	Тип задачи машинного обучения	Тип модели в BigQuery
Метки отсутствуют, данные нельзя обеспечить метками	Кластеризация	kmeans
Числовые метки	Регрессия	linear_reg dnn_regressor boosted_tree_regressor
Рекомендация продуктов пользователям	Рекомендации	matrix_factorization

Таблица 9.1 (окончание)

Описание задачи	Тип задачи машинного обучения	Тип модели в BigQuery
Выбор целевой аудитории для предложения продукта	Таргетинг клиентов	<code>matrix_factorization</code>
Метки имеют значения 1/0, True/False (две категории)	Бинарная классификация	<code>logistic_reg</code> <code>dnn_classifier</code> <code>boosted_tree_classifier</code>
Метки представлены фиксированным набором строк	Многоклассовая классификация	<code>logistic_reg</code> <code>dnn_classifier</code> <code>boosted_tree_classifier</code>
Входные признаки не структурированы	Классификация изображений, классификация текста, анализ эмоциональной окраски, извлечение сущностей	Используйте выходные данные Cloud Vision API или Cloud Natural Language API в качестве входных данных для любой из стандартных моделей BigQuery, перечисленных выше
Метки не структурированы	Ответы на вопросы, аннотирование текста, генерирование подписей к изображениям	Используйте продукты Cloud AutoML

Построение регрессионной модели

В примере построения регрессионной модели мы используем набор данных `london_bicycles`. Предположим, что у нас есть два типа велосипедов: тяжелые и надежные городские велосипеды и быстрые, но менее надежные шоссейные велосипеды. Для клиентов, совершающих длительные поездки, у нас должны быть в наличии шоссейные велосипеды, а для коротких поездок — обычные городские. Чтобы организовать правильное распределение велосипедов, необходимо спрогнозировать продолжительность поездок.

Выбор метки

Первым шагом к решению задачи машинного обучения является ее формулирование — определение признаков и меток для модели. Поскольку задача первой модели состоит в том, чтобы на основе набора архивных данных предсказать продолжительность поездки, меткой будет служить продолжительность поездки.

Но насколько верно мы определили цель задачи? Должна ли модель прогнозировать продолжительность каждой поездки или общую продолжительность

всех поездок для каждого пункта проката, например, в течение часа? Если мы выберем последний вариант, тогда меткой должна быть сумма продолжительностей всех поездок за определенный час. Допустим, по опыту ведения бизнеса мы знаем, что пункт проката, выдающий 1000 велосипедов на 20 минут каждый, должен выдавать городские велосипеды, тогда как пункт проката, выдающий ежедневно 100 велосипедов на 200 минут, должен выдавать шоссейные велосипеды. То есть прогнозирование общей продолжительности не поможет принять правильное решение, а вот предсказание продолжительности каждой отдельной поездки будет весьма кстати.

Другой вариант — оценка вероятности, что поездка продлится меньше 30 минут. В этом случае метка будет иметь два значения — True/False — в зависимости от длительности поездки (больше или меньше 30 минут). Это еще больше поможет бизнесу, потому что вероятность может определять относительную пропорцию городских и шоссейных велосипедов, имеющихся в каждом пункте проката.

На практике довольно часто приходится выбирать между несколькими целями. Иногда можно создать метку в виде взвешенной комбинации из целей и обучить единственную модель. Иногда полезнее обучить несколько моделей, по одной для каждой цели, и использовать разные модели в разных сценариях. А иногда лучше представить конечному пользователю результаты всех моделей и дать ему возможность выбора. Все зависит от того, чем вы занимаетесь.

В этом примере мы построим две модели: одну для прогнозирования продолжительности поездки, а другую для прогнозирования вероятности того, что поездка займет больше 30 минут. Затем мы дадим конечному пользователю возможность принять решение на основе двух прогнозов.

Выбор признаков в наборе данных

Если предположить, что продолжительность поездок зависит от пункта проката, дня недели и времени суток, эти параметры могут послужить нам входными признаками. Прежде чем продолжить и создать модель с этими тремя признаками, желательно убедиться, что эти факторы действительно влияют на метку.

Выбор и формирование признаков для модели машинного обучения называется *конструированием признаков* (feature engineering). Конструирование признаков часто является наиболее важным условием создания точных моделей машинного обучения, и этот шаг может гораздо сильнее повлиять на точность прогнозирования, чем выбор алгоритма или настройка гиперпараметров. Чтобы получить хороший набор признаков, необходимо глубоко понимать данные и предметную область. Часто на этом этапе проверяется множество гипотез; у вас есть идея относительно признака, вы проверяете ее обоснованность (влияние признака на метку), а затем добавляете этот признак в модель. Если эта идея не подтверждается, вы проверяете следующую.

Влияние пункта проката

Чтобы проверить зависимость продолжительности поездок от пункта проката, можно визуализировать результат следующего запроса в Data Studio, используя `start_station_name` в качестве области определения и `duration` в качестве конкретных значений:¹

```
SELECT
  start_station_name
  , AVG(duration) AS duration
FROM `bigquery-public-data`.london_bicycles.cycle_hire
GROUP BY start_station_name
```

Результат визуализации показан на рис. 9.2.

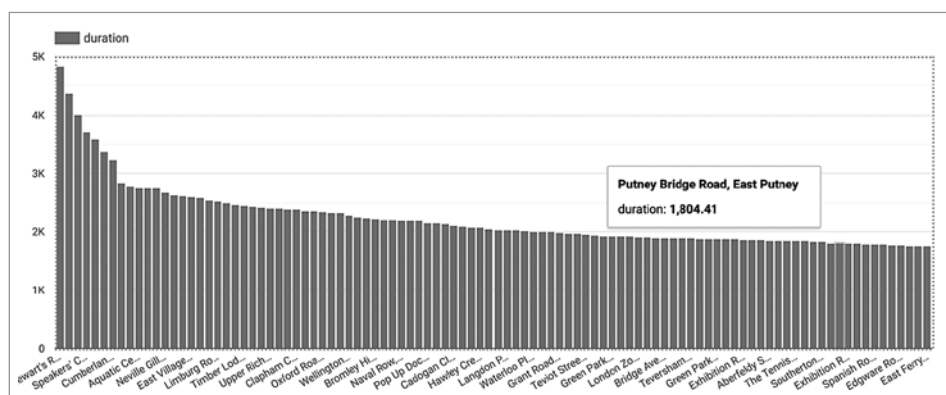


Рис. 9.2. Как видите, длительные поездки в основном совершаются лишь из нескольких пунктов проката

На основании результатов, показанных на рис. 9.2, можно сделать вывод, что длительные поездки (дольше 3000 секунд) в основном совершаются лишь из нескольких пунктов проката, и для большинства пунктов проката длительности поездок находятся в относительно узком диапазоне. Если бы они находились в узком диапазоне для всех пунктов проката в Лондоне, тогда пункт проката, где выдан велосипед, был бы плохим прогнозным признаком. Но в нашем случае, как показывает график на рис. 9.2, признак `start_station_name` имеет значение.

Обратите внимание, что `end_station_name` не может использоваться в роли признака, потому что он неизвестен, пока велосипед не вернется на начальный пункт проката. Поскольку мы создаем модель машинного обучения для прогнозирования

¹ В веб-интерфейсе BigQuery кликните на Explore in Data Studio (Исследовать в Data Studio).

ния событий в будущем, не следует использовать столбцы, которые неизвестны на момент составления прогноза. Этот критерий причинно-следственной связи накладывает ограничения на выбор признаков.

День недели

Оценка следующего кандидата в признаки выполняется аналогично: проверим значимость `dayofweek` — дня недели (точно так же можно проверить значимость `hourofday` — часа дня):

```
SELECT
  EXTRACT(dayofweek FROM start_date) AS dayofweek
, AVG(duration) AS duration
FROM `bigquery-public-data`.london_bicycles.cycle_hire
GROUP BY dayofweek
```

Результат визуализации показан на рис. 9.3.

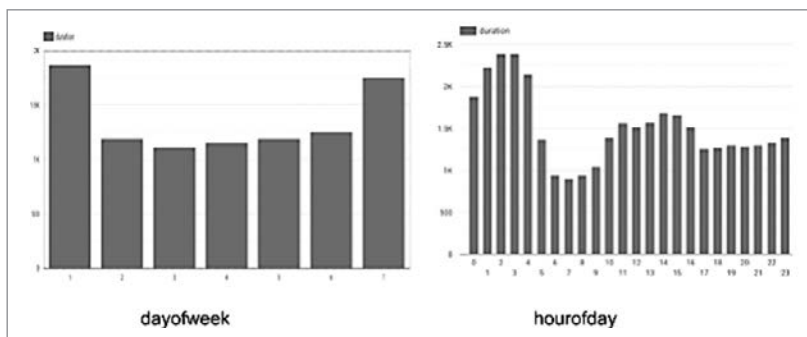


Рис. 9.3. Длительные поездки чаще совершаются в выходные, а также утром и в начале дня

Как показано на рис. 9.3, продолжительность поездок зависит от дня недели и от времени суток. Поездки, совершаемые в выходные дни (дни 1 и 7), длятся дольше, чем в будние дни. Аналогично дольше длятся поездки, совершаемые рано утром и в первой половине дня. То есть обе переменные — `dayofweek` и `hourofday` — являются хорошими признаками.

Количество велосипедов

Еще один потенциальный признак — количество велосипедов на пункте проката. Можно предположить, что люди совершают более длительные поездки, если на пункте проката, куда они обратились, доступно небольшое число велосипедов. Проверить это можно с помощью следующего запроса:

```

SELECT
    bikes_count
    , AVG(duration) AS duration
FROM `bigquery-public-data`.london_bicycles.cycle_hire
JOIN `bigquery-public-data`.london_bicycles.cycle_stations
ON cycle_hire.start_station_name = cycle_stations.name
GROUP BY bikes_count

```

Результат визуализации в Data Studio показан на рис. 9.4.

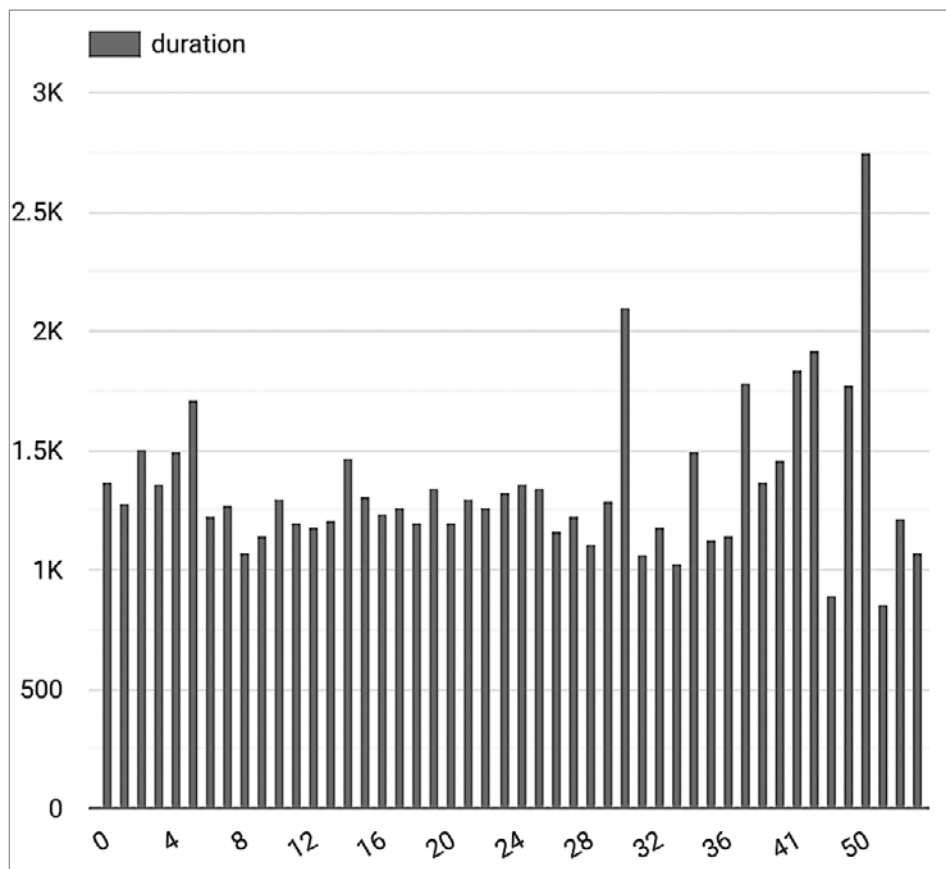


Рис. 9.4. Связь между продолжительностью поездки и количеством велосипедов на пункте проката, где арендован велосипед

Как видно на рис. 9.4, между продолжительностью поездки и количеством велосипедов на пункте проката нет очевидной связи (в отличие от времени суток). Это подсказывает нам, что количество велосипедов не является хорошим признаком. Подтвердим это количественно, вычислив коэффициент корреляции Пирсона:


```
SELECT
  CORR(bikes_count, duration) AS corr
FROM `bigquery-public-data`.london_bicycles.cycle_hire
JOIN `bigquery-public-data`.london_bicycles.cycle_stations
ON cycle_hire.start_station_name = cycle_stations.name
```

Результат -0.0039 показывает, что переменные `bikes_count` и `duration` практически независимы, потому что значение 0.0 коэффициента Пирсона соответствует полному отсутствию, а абсолютное значение 1.0 — абсолютной линейной зависимости.

Коэффициент корреляции Пирсона не является идеальным показателем полезности того или иного признака, потому что учитывает только линейную зависимость. Иногда признаки могут иметь нелинейную зависимость от метки. Но в любом случае коэффициент Пирсона является хорошей первичной оценкой. Специалисты в области машинного обучения часто используют более сложные статистические проверки, такие как взаимная информация, которая вычисляет случайность изменения признака с изменением метки.

Создание обучающего набора данных

На основании результатов исследований набора данных `london_bicycles` и выявленных взаимосвязей между различными столбцами со столбцом меток можно подготовить обучающий набор данных, включающий выбранные признаки и метки:

```
SELECT
  duration
  , start_station_name
  , CAST(EXTRACT(dayofweek FROM start_date) AS STRING) as dayofweek
  , CAST(EXTRACT(hour FROM start_date) AS STRING) AS hourofday
FROM `bigquery-public-data`.london_bicycles.cycle_hire
```

Столбцы признаков должны быть числовыми (`INT64`, `FLOAT64` и т. д.) или категориальными (`STRING`). Если признак выражается числом, но должен рассматриваться как категориальный, его следует привести к типу `STRING` — это объясняет преобразование в строки столбцов `dayofweek` и `hourofday` в этом запросе, которые хранят целые числа (в диапазонах от 1 до 7 и от 0 до 23 соответственно).¹

¹ Мы могли бы рассматривать эти переменные как значения из непрерывного диапазона, но тогда мы столкнулись бы с необходимостью бороться с тем фактом, что `dayofweek = 7` ближе к `dayofweek = 1`, чем к `dayofweek = 5`. Для справки, нам понадобилось бы: (а) сохранить `dayofweek` дважды, один раз в текущей форме, а другой как `MOD (dayofweek + 3, 7)` и (б) заменить `dayofweek` на $\sin(2\pi * \text{dayofweek} / 7.0)$. Эти сложности трудно объяснить заинтересованным сторонам. Если это не проблема и вы решаете похожую задачу, поэкспериментируйте со всеми тремя представлениями, чтобы определить, какое из них позволяет получить более точные результаты.



Если подготовка данных включает выполнение дорогостоящих в вычислительном отношении преобразований или соединений, желательно сохранить подготовленные обучающие данные в таблице, чтобы не повторять эту работу во время экспериментов. Если преобразования несложные, но сам запрос получается слишком громоздким, тогда может оказаться удобнее сохранить его в виде представления.

В данном случае запрос простой и короткий, поэтому (для большей ясности) мы просто будем повторять его в следующих разделах.

Обучение и оценка модели

Чтобы обучить модель и сохранить ее в наборе данных `ch09eu`,¹ нужно вызвать инструкцию `CREATE MODEL`, которая работает аналогично `CREATE TABLE`:

```
CREATE OR REPLACE MODEL ch09eu.bicycle_model
OPTIONS(input_label_cols=['duration'], model_type='linear_reg')
AS

SELECT
  duration
  , start_station_name
  , CAST(EXTRACT(dayofweek FROM start_date) AS STRING) as dayofweek
  , CAST(EXTRACT(hour FROM start_date) AS STRING) AS hourofday
FROM `bigquery-public-data`.london_bicycles.cycle_hire
```

Обратите внимание, что столбец меток и тип модели определяются в `OPTIONS`. Поскольку метка является числом, эта задача относится к разряду задач регрессии. Вот почему мы выбрали тип модели `linear_reg` (другие поддерживаемые типы моделей мы обсудим далее в этой главе). Как мы уже говорили в предыдущем разделе, оператор `SELECT` в запросе выше подготавливает обучающий набор данных и извлекает столбцы меток и элементов.

Оценка модели

Этот запрос был выполнен за 2.5 минуты и произвел единственную итерацию обучения,² в чем можно убедиться, заглянув на вкладку **Training** (Обучение) в разделе **BigQuery** в консоли GCP Cloud Console. Средняя абсолютная ошибка (доступная на вкладке оценки) составляет 1026 секунд, или около 17 ми-

¹ Создать его при необходимости; он должен находиться в регионе ЕУ, потому что обучающие данные находятся в ЕУ.

² Это объясняется тем, что BigQuery способен найти аналитическое решение этой задачи линейной регрессии. Узнать больше можно по ссылке <https://oreil.ly/0svPQ>.

нут.¹ То есть модель способна прогнозировать продолжительность поездок со средней ошибкой около 17 минут.

Результаты оценки также можно получить, выполнив следующий SQL-запрос:

```
SELECT * FROM ML.EVALUATE(MODEL ch09eu.bicycle_model)
```

Обратите внимание, что инструкция `OPTIONS` также определяет тип модели. Здесь мы выбрали простейшую модель регрессии, которую поддерживает BigQuery. Мы настоятельно рекомендуем выбирать простые модели и больше времени тратить на обдумывание и выбор комбинаций признаков, потому что отдача от нового/улучшенного набора входных признаков значительно превышает отдачу от использования более совершенной модели. Только достигнув максимальных показателей в ходе экспериментов с признаками, можно попробовать более сложные модели.

Комбинирование дней недели

Выше мы уже говорили, что есть несколько способов представления признаков. Например, исследуя взаимосвязь между днем недели и продолжительностью поездки, мы обнаружили, что в выходные дни поездки длятся дольше, чем в будни. Зная это, в роли признака вместо исходного значения `dayofweek` можно использовать категории, объединяющие несколько значений `dayofweek`:

```
CREATE OR REPLACE MODEL ch09eu.bicycle_model_weekday
OPTIONS(input_label_cols=['duration'], model_type='linear_reg')
AS
```

```
SELECT
  duration
  , start_station_name
  , IF(EXTRACT(dayofweek FROM start_date) BETWEEN 2 and 6,
    'weekday', 'weekend') as dayofweek
  , CAST(EXTRACT(hour FROM start_date) AS STRING) AS hourofday
FROM `bigquery-public-data`.london_bicycles.cycle_hire
```

Эта модель имеет среднюю абсолютную ошибку 967 секунд, то есть меньше 1026 секунд в предыдущей модели. Это повод перейти к модели `weekend/weekday`.

¹ Также в отчет с результатами обучения включаются другие меры ошибки (средне-квадратичная ошибка, среднеквадратичная ошибка логарифма, абсолютная средняя ошибка и т. д.). Для большинства задач регрессии абсолютная средняя ошибка обеспечивает хороший баланс между нечувствительностью к выбросам и чувствительностью к итеративным улучшениям. Используйте абсолютную среднюю ошибку, если нет веских причин не делать этого.

Группировка по времени суток

Опять же, исходя из установленной зависимости между временем суток и продолжительностью поездок, можно поэкспериментировать с делением времени на четыре периода: $(-\infty, 5)$, $[5, 10)$,¹ $[10, 17)$ и $[17, \infty)$:

```
CREATE OR REPLACE MODEL ch09eu.bicycle_model_bucketized
OPTIONS(input_label_cols=['duration'], model_type='linear_reg')
AS

SELECT
  duration
  , start_station_name
  , IF(EXTRACT(dayofweek FROM start_date) BETWEEN 2 and 6, 'weekday', 'weekend')
  as dayofweek
  , ML.BUCKETIZE(EXTRACT(hour FROM start_date), [5, 10, 17]) AS hourofday
FROM `bigquery-public-data`.london_bicycles.cycle_hire
```

ML.BUCKETIZE может служить примером функций предварительной обработки, поддерживаемых в BigQuery, — мы передаем ей число для группировки и устанавливаем границы групп, где предполагается, что первая группа ограничена слева значением $-\infty$, а последняя ограничена справа значением $+\infty$. Эта модель дает среднюю абсолютную ошибку в 901 секунду, что меньше, чем 967 секунд в модели weekend/weekday. Поэтому используем модель с группировкой по времени суток.

Получение прогнозов с помощью модели

Обученную модель можно использовать для прогнозирования, передавая ей наборы записей. Например, вот как можно получить прогнозируемую продолжительность поездок от пункта проката в Гайд-парке в 17:00 во вторник:

```
-- НЕПРАВИЛЬНЫЙ СПОСОБ! (см. следующий раздел)
SELECT * FROM ML.PREDICT(MODEL ch09eu.bicycle_model_bucketized,
  (SELECT 'Park Lane', Hyde Park' AS start_station_name
    , 'weekday' AS dayofweek, '17' AS hourofday)
)
```

Этот запрос вернет прогнозируемую продолжительность, равную 2225 секундам, но это неверный ответ. Заметили проблему?

Необходимость преобразований с помощью TRANSFORM

В предыдущем запросе мы должны были передать в поле dayofweek значение 'weekday', а не '3', потому что при обучении модель получала в dayofweek

¹ Запись интервала $[a, b)$ означает, что интервал включает a , но не включает b ; иначе говоря, это интервал $a \leq x < b$.

значения 'weekday' и 'weekend'. Также не следовало передавать непосредственное значение '17' в поле `hourofday` — мы должны передать имя группы, представляющей 17 часов. Проще говоря, чтобы получить правильные значения, в запросе прогноза необходимо выполнить те же преобразования, что и в обучающем коде.

Было бы неплохо, если бы BigQuery могла запоминать, какие преобразования применялись на этапе обучения, и автоматически применяла их на этапе предсказания. Самое замечательное, что это возможно, — именно это делает предложение **TRANSFORM**!

Более того, извлечение времени суток и дня недели можно перенести в предложение **TRANSFORM**, чтобы клиентский код передавал только отметку времени, соответствующую началу поездки:

```
CREATE OR REPLACE MODEL ch09eu.bicycle_model_bucketized
TRANSFORM(* EXCEPT(start_date)
    , IF(EXTRACT(dayofweek FROM start_date) BETWEEN 2 and 6,
'weekday', 'weekend') as dayofweek
    , ML.BUCKETIZE(EXTRACT(HOUR FROM start_date), [5, 10, 17]) AS hourofday
)
OPTIONS(input_label_cols=['duration'], model_type='linear_reg')
AS

SELECT
    duration
    , start_station_name
    , start_date
FROM `bigquery-public-data`.london_bicycles.cycle_hire
```

Используйте предложение **TRANSFORM** и сформулируйте задачу машинного обучения так, чтобы любому, кому потребуется прогноз, достаточно было передать только исходные данные.¹

Если в запросе есть предложение **TRANSFORM**, модель обучается на его выходных данных. В таком случае предложению **TRANSFORM** передаются все признаки и метки, возвращаемые исходным запросом **SELECT**, кроме `start_date`, а затем добавляется пара признаков (`dayofweek` и `hourofday`), извлеченных из `start_date`.

Для прогнозирования продолжительности поездки получившаяся модель требует передать ей только `start_station_name` и `start_date`. Преобразования сохраняются и применяются к переданным данным для получения входных признаков.

¹ В действительности именно так и действует BigQuery по умолчанию, если входной признак имеет тип **TIMESTAMP**. Точно так же, как по умолчанию BigQuery применяет к строковым значениям прием прямого кодирования (one-hot encode), из значений типа **TIMESTAMP** извлекается такая информация, как день недели. Предложения **TRANSFORM** обеспечивают возможность более точного управления.



Размещение всех функций предварительной обработки внутри предложения **TRANSFORM** дает важное преимущество: для использования модели клиентам не нужно знать, какая предварительная обработка была выполнена, — BigQuery позаботится о применении необходимых преобразований к исходным данным во время прогнозирования. Поэтому мы советуем использовать в обучающем запросе инструкцию **SELECT**, возвращающую только исходные данные, и выполнять все преобразования в предложении **TRANSFORM**.

После добавления предложения **TRANSFORM** запрос на получение прогноза приобретает следующий вид:

```
SELECT * FROM ML.PREDICT(MODEL ch09eu.bicycle_model_bucketized,
  (SELECT 'Park Lane , Hyde Park' AS start_station_name
    , CURRENT_TIMESTAMP() AS start_date)
)
```

Результат выглядит как показано ниже (ваш результат может отличаться, если у вас другое время суток или день недели):

Row	predicted_duration	start_station_name	start_date
1	3498.804224263982	Park Lane, Hyde Park	2019-05-19 04:24:03.376064 UTC

Получение пакетных прогнозов

Также есть возможность создать таблицу прогнозов для каждого часа и каждого пункта проката начиная с 3 часов утра следующего дня, если использовать функцию генерирования массива:

```
DECLARE tomorrow_3am TIMESTAMP;
SET tomorrow_3am = TIMESTAMP_ADD(
  TIMESTAMP(DATE_ADD(CURRENT_DATE(), INTERVAL 1 DAY)),
  INTERVAL 3 HOUR);

WITH generated AS (
  SELECT
    name AS start_station_name
    , GENERATE_TIMESTAMP_ARRAY(
      tomorrow_3am,
      TIMESTAMP_ADD(tomorrow_3am, INTERVAL 24 HOUR),
      INTERVAL 1 HOUR) AS dates
  FROM
    `bigquery-public-data`.london_bicycles.cycle_stations
),

features AS (
  SELECT
    start_station_name
    , start_date
```

```

FROM
  generated
  , UNNEST(dates) AS start_date
)

SELECT * FROM ML.PREDICT(MODEL ch09eu.bicycle_model_bucketized,
  (SELECT * FROM features)
)

```

Этот запрос вернет около 20 000 прогнозов. Вот часть из них:

6	2707.621807505363	Palace Gate, Kensington Gardens	2019-05-19 15:00:00 UTC
7	2707.621807505363	Palace Gate, Kensington Gardens	2019-05-19 16:00:00 UTC
8	2571.887817969073	Palace Gate, Kensington Gardens	2019-05-19 17:00:00 UTC
9	2571.887817969073	Palace Gate, Kensington Gardens	2019-05-19 18:00:00 UTC

Как видите, весь процесс машинного обучения, от создания обучающего набора данных до обучения и прогнозирования, можно выполнить без переноса данных из BigQuery.

Исследование весов модели

Модель линейной регрессии вычисляет прогноз как взвешенную сумму входных данных. Используемые веса можно получить (или экспортировать) с помощью следующей команды:

```
SELECT * FROM ML.WEIGHTS(MODEL ch09eu.bicycle_model_bucketized)
```

Каждому числовому признаку соответствует единственный вес, а для каждого категориального признака число весов равно числу возможных значений. Например, признак `dayofweek` имеет следующие веса:

Row	processed_input	weight	category_weights.category	category_weights.weight
2	dayofweek	<i>null</i>	weekday	1709.4363890323655
			weekend	2084.400311228229

Это означает, что если день недели является рабочим, этот признак вносит в общую прогнозируемую продолжительность поездки 1709 секунд (оптимальные веса не уникальны, поэтому вы можете получить другое значение). Веса, применяемые к разным признакам, не имеют особого физического смысла — практически единственная причина, по которой может понадобиться исследовать весовые коэффициенты, как было показано выше, — это если вы решите организовать прогнозирование за пределами BigQuery.



Не пытайтесь строить предположения о работе модели, опираясь на величины или знаки весов. Если входные признаки являются линейно зависимыми (что не редкость в реальных наборах данных), величины и знаки весов не имеют смысла. Для объяснения модели можно использовать инструмент What-If Tool (<https://ai.googleblog.com/2018/09/the-what-if-tool-code-free-probing-of.html>) или пакет объяснения моделей, такой как LIME (<https://www.oreilly.com/learning/introduction-to-local-interpretable-model-agnostic-explanations-lime>).

Поскольку линейная модель очень проста (она вычисляет средневзвешенное значение входных данных), можно извлечь веса модели и реализовать математические вычисления для получения прогноза, например, на Python (<https://towardsdatascience.com/how-to-do-online-prediction-with-bigquery-ml-db2248c0ae5>):

```
def compute_regression(rowdict,
                       numeric_weights, scaling_df, categorical_weights):
    input_values = rowdict
    # числовые входы
    pred = 0
    for column_name in numeric_weights['input'].unique():
        wt = numeric_weights[ numeric_weights['input'] == column_name
                               ][ 'input_weight' ].values[0]
        if column_name != '__INTERCEPT__':
            meanv = (scaling_df[ scaling_df['input'] ==
                                column_name ][ 'mean' ].values[0])
            stddev = (scaling_df[ scaling_df['input'] ==
                                  column_name ][ 'stddev' ].values[0])
            scaled_value = (input_values[column_name] - meanv)/stddev
        else:
            scaled_value = 1.0
        contrib = wt * scaled_value
        pred = pred + contrib
    # категориальные входы
    for column_name in categorical_weights['input'].unique():
        category_weights = categorical_weights[ categorical_weights['input'] ==
                                                column_name ]
        wt = category_weights[ category_weights['category_name'] ==
                               input_values[column_name] ][ 'category_weight' ].values[0]
        pred = pred + wt
    return pred
```

Здесь значения для массива `numeric_weights` можно получить с помощью такого запроса:

```
SELECT
    processed_input AS input,
    model.weight AS input_weight
FROM
    ml.WEIGHTS(MODEL dataset.model) AS model
```


Данные для масштабированного объекта `DataFrame` — `scaling_df` можно получить с помощью следующего запроса:

```
SELECT
    input, min, max, mean, stddev
FROM
    ml.FEATURE_INFO(MODEL dataset.model) AS model
```

Значения для массива `categorical_weights` можно получить, выполнив этот запрос:

```
SELECT
    processed_input AS input,
    model.weight AS input_weight,
    category.category AS category_name,
    category.weight AS category_weight
FROM
    ml.WEIGHTS(MODEL dataset.model) AS model,
    UNNEST(category_weights) AS category
```

Модель `logistic_reg` на этапе прогнозирования возвращает результат применения сигмоидной функции к средневзвешенному значению. То есть на Python прогноз можно вычислить так:

```
def compute_classifier(rowdict,
    numeric_weights, scaling_df, categorical_weights):
    pred=compute_regression(rowdict, numeric_weights, scaling_df,
    categorical_weights)
    return (1.0/(1 + np.exp(-pred)) if (-500 < pred) else 0)
```

Более сложные регрессионные модели

Модель линейной регрессии является самой простой формой регрессионных моделей — каждому входному признаку присваивается вес, а выходным значением является взвешенная сумма входных данных плюс константа, называемая поправкой. BigQuery также поддерживает модели `dnn_regressor` и `xgboost`.

Глубокие нейронные сети

Глубокую нейронную сеть (Deep Neural Network, DNN) можно рассматривать как расширение линейных моделей, в которых каждый узел в первом слое вычисляет взвешенную сумму входных признаков, преобразованных с помощью функции, как правило, нелинейной. Второй слой состоит из узлов, каждый из которых вычисляет взвешенную сумму выходных данных первого слоя, преобразованных с помощью нелинейной функции и т. д., как показано на рис. 9.5.

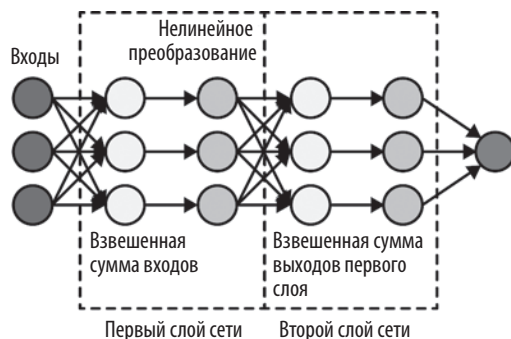


Рис. 9.5. Глубокая нейронная сеть состоит из слоев с «узлами». Здесь показаны два слоя, расположенные между входами и выходами. В каждом слое имеется по три узла, однако число слоев и узлов в них может быть каким угодно

Чтобы обучить модель DNN с 64 узлами в первом слое и 32 узлами во втором слое, можно использовать такой запрос:

```
CREATE OR REPLACE MODEL ch09eu.bicycle_model_dnn
TRANSFORM(* EXCEPT(start_date)
            , IF(EXTRACT(dayofweek FROM start_date) BETWEEN 2 and 6, 'weekday',
'weekend') as dayofweek
            , ML.BUCKETIZE(EXTRACT(HOUR FROM start_date), [5, 10, 17]) AS
hourofday
)
OPTIONS(input_label_cols=['duration'],
         model_type='dnn_regressor',
         hidden_units=[64, 32])
AS

SELECT
    duration
    , start_station_name
    , start_date
FROM `bigquery-public-data`.london_bicycles.cycle_hire
```

Для обучения этой модели потребовалось около 20 минут. В результате средняя абсолютная ошибка модели составила 1016 секунд. Это, конечно, хуже 901 секунды, достигнутой линейной моделью. К сожалению, это нормальное явление — глубокие нейронные сети, как известно, весьма неустойчивы в обучении.



Мы настоятельно советуем начинать с линейных моделей и только после выбора эффективных признаков и преобразований переходить к экспериментам с более сложными моделями. Это связано с тем, что при создании модели `dnn_regressor` вам наверняка придется поэкспериментировать, меняя количество слоев и узлов (то есть с `hidden_units`) и настройки регуляризации (то есть `l2_reg`), чтобы добиться высокой точности прогнозирования. Учитывая требовательный характер сетей глубокого обучения, изменение представлений признаков — это верный путь к противоречивым результатам.

Один из способов преодолеть этот недостаток — выполнить настройку гиперпараметров, чтобы найти их оптимальные значения. Эта возможность поддерживается полноценными инфраструктурами машинного обучения, такими как Cloud AI Platform (CAIP).¹ Иногда желательно использовать этот механизм обучения или AutoML (мы рассмотрим оба варианта далее в этой главе), а пока попробуем использовать меньшую сеть:

```
CREATE OR REPLACE MODEL ch09eu.bicycle_model_dnn
TRANSFORM(* EXCEPT(start_date)
           , IF(EXTRACT(dayofweek FROM start_date) BETWEEN 2 and 6, 'weekday',
'weekend')) as dayofweek
           , ML.BUCKETIZE(EXTRACT(HOUR FROM start_date), [5, 10, 17]) AS
hourofday
)
OPTIONS(input_label_cols=['duration'],
        model_type='dnn_regressor',
        hidden_units=[10, 5])
AS

SELECT
    duration
    , start_station_name
    , start_date
FROM `bigquery-public-data`.london_bicycles.cycle_hire
```

Точность этой модели выше (981 секунда), но все же она не так хороша, как линейная модель. Необходима дополнительная настройка гиперпараметров, чтобы получить глубокую нейронную сеть, имеющую более высокую точность, чем линейная модель, с которой мы начинали. Кроме того, в общем случае глубокие нейронные сети обеспечивают высокую точность, только если имеется много признаков с непрерывными диапазонами значений.

Деревья решений с градиентным бустингом

Деревья решений — это метод машинного обучения, пользующийся большой популярностью из-за простоты интерпретации результатов (по сути, они представляют собой простую комбинацию правил «if-then»). Однако, как правило, деревья решений имеют невысокую точность из-за ограниченности диапазона функций, которые они способны аппроксимировать, и склонности к переобучению. Одним из способов увеличить точность деревьев решений (в ущерб интерпретируемости²) является обучение ансамбля деревьев решений, каждое из

¹ См. <https://cloud.google.com/ml-engine/docs/tensorflow/hyperparameter-tuning-overview>. Cloud AI Platform Predictions позволяет посылать задания машинного обучения с диапазонами значений для поиска.

² Многие пакеты, реализующие деревья решений, поддерживают метрику «важность признака», которая, если говорить в общих чертах, определяется частотой использования признака в ансамбле деревьев. Однако если у вас есть два взаимосвязанных признака, их важность будет разделена между ними, из-за чего может пострадать интерпретируемость результатов.

которых является слабым предиктором, но вместе они дают достаточно высокую точность. Бустинг — это метод выбора деревьев в ансамбле, а XGBoost¹ — это масштабируемый, распределенный способ построения деревьев решений на очень больших и разреженных наборах данных. До появления сетей глубокого обучения примерно в 2015 году XGBoost считался самым современным методом машинного обучения. Он продолжает пользоваться популярностью при решении задач с использованием структурированных данных.

Обучить модель XGBoost в BigQuery можно, выбрав тип модели `boosted_tree_regressor`:

```
CREATE OR REPLACE MODEL ch09eu.bicycle_model_xgboost
TRANSFORM(* EXCEPT(start_date)
           , IF(EXTRACT(dayofweek FROM start_date) BETWEEN 2 and 6,
'weekday', 'weekend') as dayofweek
           , ML.BUCKETIZE(EXTRACT(HOUR FROM start_date), [5, 10, 17])
                                     AS hourofday
)
OPTIONS(input_label_cols=['duration'],
        model_type='boosted_tree_regressor',
        max_tree_depth=4)
AS

SELECT
    duration
    , start_station_name
    , start_date
FROM `bigquery-public-data`.london_bicycles.cycle_hire
```

Полученная модель в результате имеет более низкую точность прогнозирования (1363 секунды), чем линейная модель. Важность входных признаков можно получить с помощью такой команды:

```
SELECT * FROM ML.FEATURE_INFO(MODEL ch09eu.bicycle_model_xgboost)
```

Использование человеческих знаний и вспомогательных данных

Помимо использования моделей с разными архитектурами и настройки их параметров, можно подумать о добавлении новых входных признаков, включая вспомогательные данные и человеческие знания.

Например, в предыдущей модели мы использовали функцию `ML.BUCKETIZE`, чтобы разбить непрерывную переменную (час, извлеченный из отметки времени) на четыре диапазона. Еще одна чрезвычайно полезная функция — `ML.FEATURE_`

¹ Название XGBoost расшифровывается как eXtreme Gradient Boost, где Gradient Boost (бустинг градиента) — это название метода, предложенного Джеромом Фридманом (Jerome H. Friedman) в статье «Greedy Function Approximation: A Gradient Boosting Machine» (<https://www.kdd.org/kdd2016/papers/files/rfp0697-chenAemb.pdf>).

CROSS — позволяет объединить отдельные категориальные признаки в условие по AND (подобные взаимозависимости между признаками с трудом определяются моделями машинного обучения). Как подсказывает интуиция, сочетание дня недели и утренних часов в нашей задаче является хорошим прогнозирующим признаком продолжительности поездок на велосипедах, гораздо более значимым, чем день недели или утренние часы сами по себе. Возможно, есть смысл определить пересечение этих двух признаков вместо того, чтобы рассматривать день недели и время по отдельности:

```
ML.FEATURE_CROSS(STRUCT(
  IF(EXTRACT(dayofweek FROM start_date) BETWEEN 2 and 6,
    'weekday', 'weekend') as dayofweek,
  ML.BUCKETIZE(EXTRACT(HOUR FROM start_date),
    [5, 10, 17]) AS hr
)) AS dayhr
```

До сих пор в качестве входных данных для моделей мы использовали `start_station_name`. При этом пункты проката считаются независимыми друг от друга. В главе 8 мы рассмотрели преимущества ST_GeoHash как способа сбора данных о пространственной близости. А что, если добавить в модель вспомогательный признак с информацией о местонахождении пунктов проката?

Следующий обучающий запрос объединяет эти две идеи:

```
CREATE OR REPLACE MODEL ch09eu.bicycle_model_fc_geo
  TRANSFORM(duration
    , ML.FEATURE_CROSS(STRUCT(
      IF(EXTRACT(dayofweek FROM start_date) BETWEEN 2 and 6,
        'weekday', 'weekend') as dayofweek,
      ML.BUCKETIZE(EXTRACT(HOUR FROM start_date),
        [5, 10, 17]) AS hr
    )) AS dayhr
    , ST_GeoHash(ST_GeogPoint(latitude, longitude), 4) AS start_station_loc4
    , ST_GeoHash(ST_GeogPoint(latitude, longitude), 6) AS start_station_loc6
    , ST_GeoHash(ST_GeogPoint(latitude, longitude), 8) AS start_station_loc8
  )
  OPTIONS(input_label_cols=['duration'], model_type='linear_reg')
AS

SELECT
  duration
  , latitude
  , longitude
  , start_date
FROM `bigquery-public-data`.london_bicycles.cycle_hire
JOIN `bigquery-public-data`.london_bicycles.cycle_stations
ON cycle_hire.start_station_id = cycle_stations.id
```

Эта модель имеет среднюю абсолютную ошибку, равную 898 секундам, то есть показывает более высокую точность по сравнению с 901 секундой, которую мы наблюдали выше. Однако улучшение получилось весьма незначительным.

Так как каждое следующее улучшение становится все меньше, пришло время двигаться дальше.

Создание модели классификации

В предыдущем разделе мы построили модели машинного обучения, прогнозирующие продолжительность поездок. Однако в течение одного часа в прокат будет сдаваться большое число велосипедов и их сроки аренды будут разными. Рассмотрим распределение продолжительностей поездок на велосипедах, взятых в прокат на пункте Роял-авеню 1 в Челси по будням в 14:00:

```
SELECT
  APPROX_QUANTILES(duration, 10) AS q
FROM `bigquery-public-data`.london_bicycles.cycle_hire
WHERE
  EXTRACT(dayofweek FROM start_date) BETWEEN 2 and 6
  AND EXTRACT(hour FROM start_date) = 14
  AND start_station_name = 'Royal Avenue 1, Chelsea'
```

Вот полученные результаты:

Row	q
1	0
	240
	420
	540
	660
	840
	1020
	1260
	1500
	2040
	386460

Восемьдесят процентов поездок, совершенных в будни от этого пункта проката, длилось меньше 1500 секунд. Если бы вы ограничились только этим прогнозом, то могли бы решить разместить на этом пункте проката в такие дни только городские велосипеды. Однако зная, что примерно от 10 до 20% поездок длится больше 1800 секунд, вы могли дополнительно разместить 15% шоссейных велосипедов. Спрогнозировать вероятность продолжительности поездки дольше 1800 секунд вам поможет модель классификации.

Обучение

Для простоты возьмем тот же набор признаков, который мы использовали в регрессионной модели, и обучим модель прогнозировать вероятность того, что поездка продлится больше 30 минут:

```
CREATE OR REPLACE MODEL ch09eu.bicycle_model_longrental
TRANSFORM(* EXCEPT(start_date)
, IF(EXTRACT(dayofweek FROM start_date) BETWEEN 2 and 6,
'weekday', 'weekend') as dayofweek
, ML.BUCKETIZE(EXTRACT(HOUR FROM start_date), [5, 10, 17])) AS
hourofday
)
OPTIONS(input_label_cols=['biketype'], model_type='logistic_reg')
AS

SELECT
  IF(duration > 1800, 'roadbike', 'commuter') AS biketype
  , start_station_name
  , start_date
FROM `bigquery-public-data`.london_bicycles.cycle_hire
```

Обратите внимание, что на этот раз выбран тип модели `logistic_reg` (логистическая регрессия) — это самый простой тип моделей для задач классификации. Для классификации с использованием глубокой нейронной сети или дерева решений с градиентным бустингом используйте `dnn_classifier` или `boosted_tree_classifier` соответственно.

Мы создали метки, определив порог продолжительности поездок, равный 1800 секундам, и дали двум категориям имена `roadbike` (шоссейный велосипед) и `commuter` (городской велосипед). Примерно так же мы определили категории `weekend/weekday` на основе числовой переменной `dayofweek`. Мы могли бы применить логическое значение (`True/False`), но использование фактических имен категорий делает наши намерения более понятными.

В конце обучения вы увидите, что за семь итераций обучения ошибка уменьшилась и достигла минимального уровня, как показано на рис. 9.6 (из-за случайного характера начальных значений весов ваши результаты могут немного отличаться).

На самом деле на рис. 9.6 показаны две кривые потерь: одна для обучающих данных и другая для контрольных данных (BigQuery автоматически делит набор данных на обучающие и контрольные данные). Здесь кривые очень похожи. Если бы кривая потерь на контрольных данных располагалась намного выше кривой потерь на обучающих данных, можно было бы говорить о переобучении модели. Переключившись на табличное представление, можно убедиться, что обе потери действительно очень близки на протяжении всего обучения:

Итерация	Потери на обучающих данных	Потери на контрольных данных	Скорость обучения	Продолжительность (с)
6	0.3072	0.3024	3.2000	41.59
5	0.3078	0.3029	6.4000	39.66
4	0.3119	0.3069	3.2000	40.54
3	0.3240	0.3195	1.6000	42.15
2	0.3576	0.3543	0.8000	37.96
1	0.4502	0.4483	0.4000	38.01
0	0.5812	0.5805	0.2000	22.10

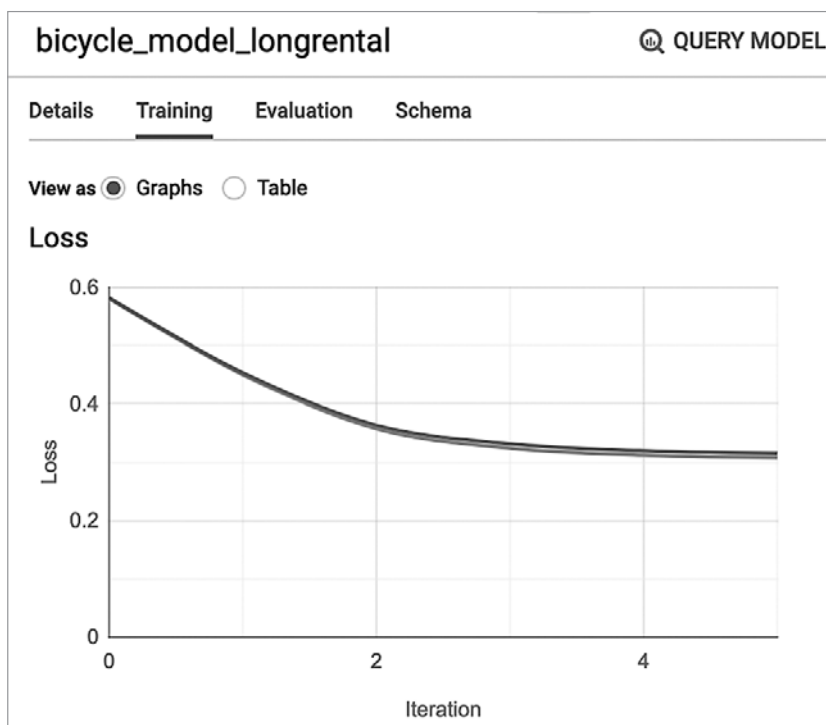


Рис. 9.6. Кривая потерь приближается к минимуму в процессе обучения

Оценка

В качестве меры в классификации используется перекрестная энтропия, поэтому на графике изображены кривые обучения. Если перейти на вкладку **Evaluation** (Оценка) в веб-интерфейсе BigQuery, как показано на рис. 9.7, можно посмотреть более знакомые метрики оценки, такие как точность.

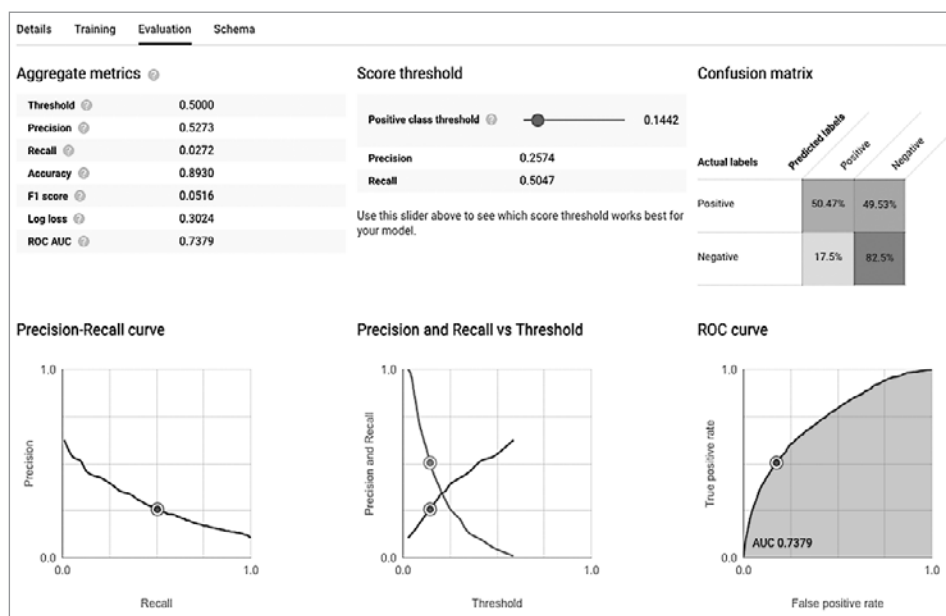


Рис. 9.7. Вкладка Evaluation (Оценка) в веб-интерфейсе BigQuery с характеристиками модели классификации

Прогнозирование

Прогнозирование с использованием этой модели производится точно так же, как в случае с моделью линейной регрессии, только на этот раз возвращается вероятность каждого класса:

```
SELECT * FROM ML.PREDICT(MODEL ch09eu.bicycle_model_longrental,
  (SELECT 'Park Lane , Hyde Park' AS start_station_name
    , TIMESTAMP('2019-05-09 16:16:00 UTC') AS start_date)
)
```

Этот запрос возвращает следующие результаты:

Row	predicted_biketype	predicted_biketype_probs.label	predicted_biketype_probs.prob	start_station_name	start_date
1	commuter	roadbike	0.4419...	Park Lane, Hyde Park	2019-05-10 16:16:00 UTC
		commuter	0.5580...		

Как видите, вероятность того, что для поездки, начавшейся в 16 часов в будний день в пункте проката Гайд-парка потребуется шоссейный велосипед, равна 0.44, или 44%. То есть в идеале в этом пункте проката 44% велосипедов должны быть шоссейными.

Выбор порога

В нашем случае интерес представляет прогнозируемая вероятность. Однако во многих задачах классификации предпочтительнее получить не только вероятность, но и прогнозируемый класс. То есть результат прогноза (см. предыдущий раздел) должен включать не только вероятность, но также класс с наибольшей вероятностью. В задаче бинарной классификации это равносильно установке порогового значения вероятности, равного 0.5, и выбору класса с вероятностью больше этого порога.

Полнотой (recall) в машинном обучении называют процент фактических истинных результатов (отношение истинно-положительных результатов к общему числу положительных результатов) в определенной пороговой точке. Хороший порог должен иметь высокую полноту. Однако выбор пороговой точки с высоким значением полноты таит опасность получения большого количества ложноположительных результатов. Если выбрать порог, равный нулю, вы получите весь набор данных и, соответственно, идеальную полноту.

Другим важным показателем является *точность* (precision) — процент истинно-положительных результатов по всему набору данных. Это можно выразить так: «Я спрогнозировал, что этот результат истинный, но какова вероятность, что я прав?» Если установить порог равным нулю, истинными будут далеко не все полученные результаты. (То есть если вы спрогнозировали, что все результаты верны, но в действительности верными будут только 10%, точность составит 10%. Это плохой результат для классификатора.)

Совокупные показатели на вкладке **Evaluation** (Оценка), например `accuracy=0.89`, рассчитываются на основе порога 0.5.

Если вы хотите, чтобы в 50% случаев, когда требуется шоссейный велосипед, он был в наличии, тогда вам нужна полнота 0.5, потому что достаточно охватить половину продолжительных поездок. Если хотите, можете воспользоваться ползунком на вкладке **Evaluation** (Оценка) и установить пороговое значение равным 0.144, как показано на рис. 9.8, чтобы получить желаемую величину метрики полноты. Обратите внимание, что это происходит за счет точности; при этом пороговом значении модель обеспечит точность на уровне 0.26 — только в 26% случаев спрогнозированные поездки будут длиться дольше 30 минут и для них потребуется выдать шоссейный велосипед.¹

¹ Точность (или доля истинно-положительных результатов) — это доля прогнозов, когда модель правильно предсказала положительный класс. Другими словами, если модель

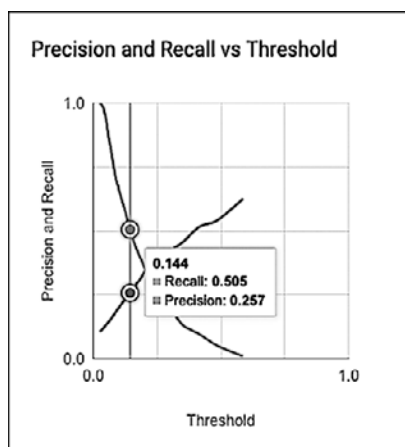


Рис. 9.8. Изменение порога для получения желаемой полноты или точности

При использовании моделей бинарной классификации желаемый порог можно передать в `ML.PREDICT`:

```
SELECT * FROM ML.PREDICT(MODEL ch09eu.bicycle_model_longrental,
  (SELECT 'Park Lane , Hyde Park' AS start_station_name
    , TIMESTAMP('2019-05-09 16:16:00 UTC') AS start_date),
  STRUCT(0.144 AS threshold)
)
```

Этот запрос возвращает следующие результаты:

Row	predicted_biketype	predicted_biketype_probs.label	predicted_biketype_probs.prob	start_station_name	start_date
1	roadbike	roadbike	0.4419...	Park Lane, Hyde Park	2019-05-09 16:16:00 UTC

Обратите внимание, что теперь прогнозируемый тип велосипеда (`predicted_biketype`) — шоссейный велосипед (`roadbike`), хотя вероятность, соответствующая классу шоссейных велосипедов, меньше порогового значения — по умолчанию 0.5.

100 раз спрогнозировала выбор шоссейного велосипеда (`roadbike`), ее прогноз будет правильным в 25.7 случая. Полнота — это доля положительных результатов, правильно спрогнозированных моделью, то есть доля предсказаний выбора шоссейного велосипеда, когда такой велосипед действительно необходим. Для задач множественной классификации точность (или полнота) соответствует средней точности при интерпретации каждой категории как задачи бинарной классификации.

Настройка механизма машинного обучения в BigQuery

По умолчанию механизм машинного обучения BigQuery ML разумно подходит к выбору скорости обучения,¹ масштабированию входных признаков,² делению данных на обучающую и контрольную выборки³ и т. д. При создании модели в предложении `OPTIONS` можно определить дополнительные настройки (https://cloud.google.com/bigquery-ml/docs/reference/standard-sql/bigqueryml-syntax-create#model_option_list). В этом разделе мы обсудим некоторые из них.

Управление делением данных

По умолчанию для наборов данных среднего размера BigQuery случайным образом отбирает 20% данных и сохраняет их для оценки. Обучение проводится только на 80% предоставленных данных. Для небольших наборов данных (с числом записей меньше 500) для обучения используются все данные, а для больших наборов данных (с числом записей больше 50 000) для оценки используется только 10 000 записей. Управлять выбором данных для оценки можно с помощью трех параметров: `data_split_method`, `data_split_eval_fraction` и `data_split_col`, как указано в табл. 9.2.

Таблица 9.2. Управление делением данных на обучающую и контрольную выборки

Сценарий	<code>data_split_method</code>	<code>data_split_eval_fraction</code>	<code>data_split_col</code>
По умолчанию	<code>auto_split</code>	0.2	–
Обучение на всех данных	<code>no_split</code>	–	–
В контрольную выборку для оценки случайным образом отбирается 10% данных	<code>random</code>	0.1	–
Конкретно указывается, какие записи должны отбираться в контрольную выборку	<code>custom</code>	–	<code>colname</code> В данном столбце записи с логическим значением <code>True</code> / <code>NULL</code> выбираются для оценки
В контрольную выборку отбираются 10% последних записей	<code>seq</code>	0.1 (по умолчанию 0.2)	<code>colname</code> В данном столбце записи упорядочиваются по возрастанию значений

¹ BigQuery выбирает наиболее подходящее значение путем поиска строк в данных в начале каждой итерации.

² Все числовые входные признаки масштабируются так, чтобы они имели нулевое среднее значение и единичную дисперсию.

³ По умолчанию в контрольную выборку случайным образом отбираются 20% записей.

Лучшие результаты достигаются при обучении модели на первых 80% (упорядоченных по времени) записей с информацией о поездках и тестировании на оставшихся 20%.¹ То есть вместо случайного деления данных на обучающую и контрольную выборки обучение будет производиться на старых поездках, а тестироваться на новых:

```
CREATE OR REPLACE MODEL ch09eu.bicycle_model_bucketized_seq
TRANSFORM(* EXCEPT(start_date)
           , IF(EXTRACT(dayofweek FROM start_date) BETWEEN 2 and 6, 'weekday',
'weekend') as dayofweek
           , ML.BUCKETIZE(EXTRACT(HOUR FROM start_date), [5, 10, 17])
                                     AS hourofday
           , start_date – used to split the data
)
OPTIONS(input_label_cols=['duration'], model_type='linear_reg',
        data_split_method='seq',
        data_split_eval_fraction=0.2,
        data_split_col='start_date')
AS

SELECT
  duration
  , start_station_name
  , start_date
FROM `bigquery-public-data`.london_bicycles.cycle_hire
```

Обратите внимание, что в обоих предложениях — `SELECT` и `TRANSFORM` — присутствует столбец, используемый для деления данных, а предложение `OPTIONS` включает три параметра, управляющих делением данных.

Средняя абсолютная ошибка этой модели составляет 860 секунд, но мы не можем сравнить это число с результатом, полученным при случайном разделении данных, — оценочные показатели сильно зависят от выбора данных для оценки, и поскольку сейчас используется другой набор контрольных данных, мы не можем сравнить эти результаты с результатами, полученными ранее. Кроме того, предыдущие результаты были искажены утечкой информации, как описывалось на примере рождественских дней.

¹ Лучшие, потому что может произойти так, что дни большой нагрузки на пункт проката А совпадут с днями большой нагрузки на пункт проката В. Когда выборки создаются случайным образом, может произойти утечка этой информации, если данные по пункту проката А, соответствующие Рождеству 2009 года, попадут в обучающую выборку, а данные по пункту проката В, соответствующие Рождеству того же 2009 года, попадут в контрольную выборку. Управляя делением данных так, чтобы последние несколько дней в наборе данных не попали в обучающую выборку, мы сможем точнее смоделировать обучение модели на исторических данных, а затем развернуть ее.

Балансировка классов

В задаче классификации менее 12% поездок длится дольше 1800 секунд. Это пример несбалансированного набора данных. Возможно, будет полезно придать более редкому классу более высокий вес. Это можно сделать, явно передав массив весов классов или отправив запрос BigQuery установить веса классов, опираясь на обратную частоту.

Вот пример использования метода автоматического балансирования:

```
CREATE OR REPLACE MODEL ch09eu.bicycle_model_longrental_balanced
TRANSFORM(* EXCEPT(start_date)
           , IF(EXTRACT(dayofweek FROM start_date) BETWEEN 2 and 6, 'weekday',
'weekend') as dayofweek
           , ML.BUCKETIZE(EXTRACT(HOUR FROM start_date), [5, 10, 17])
                                   AS hourofday
           , start_date
)
OPTIONS(input_label_cols=['biketype'], model_type='logistic_reg',
        data_split_method='seq',
        data_split_eval_fraction=0.2,
        data_split_col='start_date',
        auto_class_weights=True)
AS

SELECT
  IF(duration > 1800, 'roadbike', 'commuter') AS biketype
  , start_station_name
  , start_date
FROM `bigquery-public-data`.london_bicycles.cycle_hire
```

Обратите внимание, что после балансировки весов вероятность, возвращаемая моделью, больше не является оценкой фактической частоты событий. Это связано с тем, что оценка вероятности, полученная в результате логистической регрессии, основана на частоте появления событий в данных, наблюдаемых моделью, и мы искусственно увеличили частоту появления редких событий.

Регуляризация

Как уже отмечалось выше, большинство пунктов проката имеют почти одинаковую продолжительность поездок, за исключением нескольких случаев с необычно продолжительными поездками. Многие из этих пунктов проката насчитывают очень мало поездок. Категориальные признаки, имеющие такое распределение с длинным хвостом, могут вызвать *переобучение*. Под переобучением понимается ситуация, при которой модель запоминает шум (случайные флуктуации) в данных, а не изучает сигнал. Другими словами, модель может стать настолько сложной, что будет представлять сам набор данных, а не его базовые качества.

Регуляризация помогает избежать переобучения, снижая сложность, в частности, за счет введения штрафа за большие значения весов. Большие значения весов часто являются признаком переобучения, потому что могут внезапно включаться, когда имеется ровно одна точка данных.

BigQuery ML поддерживает два типа регуляризации: L1 и L2. Регуляризация L1 подталкивает отдельные веса к нулю и лучше поддается интерпретации, а регуляризация L2 стремится сохранить все веса относительно одинаковыми и лучше справляется с предотвращением переобучения.¹ Управлять величиной регуляризации L1 или L2 можно при создании модели:

```
CREATE OR REPLACE MODEL ch09eu.bicycle_model_bucketized_seq_l2
TRANSFORM(* EXCEPT(start_date)
    , IF(EXTRACT(dayofweek FROM start_date) BETWEEN 2 and 6,
        'weekday', 'weekend') as dayofweek
    , ML.BUCKETIZE(EXTRACT(HOUR FROM start_date), [5, 10, 17]) AS
        hourofday
    , start_date — used to split the data
)
OPTIONS(input_label_cols=['duration'], model_type='linear_reg',
    data_split_method='seq',
    data_split_eval_fraction=0.2,
    data_split_col='start_date',
    l2_reg=0.1)
AS

SELECT
    duration
    , start_station_name
    , start_date
FROM `bigquery-public-data`.london_bicycles.cycle_hire
```

В этом случае средняя абсолютная ошибка получилась равной 857 секундам, то есть она почти идентична полученной без регуляризации L2. Причина, скорее всего, в том, что у нас есть достаточно большой набор данных и определено всего несколько настроек, чего оказалось достаточно, чтобы переобучения не произошло. Регуляризация L2 обычно считается наилучшим выбором, особенно для небольших наборов данных или при использовании достаточно сложной модели (например, глубокой нейронной сети) с большим числом параметров.

Кластеризация методом k-средних

Рассматриваемые до настоящего времени алгоритмы машинного обучения относились к категории методов обучения с учителем — мы должны были предо-

¹ За дополнительной информацией о регуляризации L1 и L2 обращайтесь по ссылке www.robotics.stanford.edu/~ang/papers/icml04-l1l2.ps.

ставить столбец меток. Однако BigQuery поддерживает также методы обучения без учителя. Например, мы можем применить к данным алгоритм кластеризации методом k -средних (https://en.wikipedia.org/wiki/K-means_clustering)¹ для группировки их в кластеры на основе сходства. Такое название алгоритм получил, потому что определяет k кластеров, каждый из которых описывается средним значением его членов. В отличие от машинного обучения с учителем, помогающим предсказать значение столбца метки для новых данных, которые модель еще не видела, обучение без учителя носит описательный характер. Используйте `model_type = kmeans` в BigQuery, чтобы получить описание данных с точки зрения центроидов k кластеров, выявленных в данных, и принять решение о принадлежности к тому или иному кластеру на основе атрибутов его центроида.

Выбор признаков для кластеризации

Первым шагом в использовании алгоритма кластеризации методом k -средних является выбор признаков для кластеризации. Поскольку таблицы в BigQuery склонны к сглаживанию и описывают несколько аспектов, выбор способа помогает понять, какие общие характеристики имеют члены кластера.

Предположим, у вас есть данные, каждая запись в которых представляет розничную покупку. Кластеризовать такую таблицу можно несколькими способами, и выбор способа зависит от того, что предполагается делать с кластерами:

- Можно попробовать выявить естественные группы среди клиентов. Этот прием называется *сегментацией клиентов*. Данные, используемые для сегментации клиентов, будут играть роль атрибутов, описывающих клиентов, совершающих покупки, — к ним можно отнести, например, магазин, где совершена покупка, наименование купленных товаров, стоимость покупки и т. д. Цель кластеризации по клиентам — желание понять, какие отличительные особенности имеют группы клиентов (они называются *персонами*), чтобы на их основе можно было создать элементы, представляющие членов этих групп, — так называемые клиенты-центроиды.
- Можно попробовать найти естественные группы среди купленных товаров — так называемые *товарные группы*. Данные, используемые для выявления товарных групп, будут играть роль атрибутов, описывающих товары, приобретенные в рамках покупки, — к ним можно отнести, например, личность покупателя, время покупки, магазин, где были куплены товары и т. д. Цель кластеризации по товарам — желание понять, какие отличительные особенности имеют группы товаров, и уменьшить *каннибализацию*² или повысить *перекрестные продажи*.

¹ Аналогичная статья в русскоязычном сегменте Википедии: https://ru.wikipedia.org/wiki/Метод_k-средних. — *Примеч. пер.*

² Каннибализация — явление, при котором новый товар сокращает рыночную долю другого товара того же производителя. — *Примеч. ред.*

В обоих случаях мы используем кластеризацию в качестве эвристики для принятия решений — слишком сложно было бы проектировать маркетинговые акции для товаров или клиентов по отдельности или понять их взаимовлияние, проделать то же самое для групп товаров или клиентов намного проще.

Обратите внимание, что в случае выработки рекомендаций (рекомендация товаров клиентам или ориентация клиентов на товар) лучше обучать модель `matrix_factorization`, как описано далее в этой главе. Но в других ситуациях, для которых не существует готовых методов прогнозирующей аналитики, кластеризация методом k -средних может дать вам возможность принимать решения на основе данных.

Кластеризация пунктов проката велосипедов

Предположим, что вам часто приходится принимать решение — на какие пункты проката поставить новые типы велосипедов, какие нуждаются в ремонте или расширении и т. д. Такие решения лучше принимать на основе данных. Это означает, что вам нужно сгруппировать пункты проката, схожие по таким атрибутам, как продолжительность поездок от пункта проката, количество поездок в день, количество велосипедных стоек на пункте проката и расстояние от пункта проката до центра города. Поскольку первые два атрибута различаются в зависимости от дня недели (рабочий или выходной), вычислим для каждого из них два значения.

Поскольку запрос получается довольно длинным и громоздким, сохраним полученные значения в таблице:

```
CREATE OR REPLACE TABLE ch09eu.stationstats AS

WITH hires AS (
  SELECT
    h.start_station_name as station_name,
    IF(EXTRACT(DAYOFWEEK FROM h.start_date) BETWEEN 2 and 6,
       "weekday", "weekend") as isweekday,
    h.duration,
    s.bikes_count,
    ST_DISTANCE(ST_GEOGPOINT(s.longitude, s.latitude),
                ST_GEOGPOINT(-0.1, 51.5))/1000 as distance_from_city_center
  FROM `bigquery-public-data.london_bicycles.cycle_hire` as h
  JOIN `bigquery-public-data.london_bicycles.cycle_stations` as s
  ON h.start_station_id = s.id
  WHERE EXTRACT(YEAR from start_date) = 2015
),

stationstats AS (
  SELECT
    station_name,
    AVG(IF(isweekday = 'weekday', duration, NULL)) AS duration_weekdays,
    AVG(IF(isweekday = 'weekend', duration, NULL)) AS duration_weekends,
```

```

COUNT(IF(isweekday = 'weekday', duration, NULL)) AS numtrips_weekdays,
COUNT(IF(isweekday = 'weekend', duration, NULL)) AS numtrips_weekends,
MAX(bikes_count) as bikes_count,
MAX(distance_from_city_center) as distance_from_city_center
FROM hires
GROUP BY station_name
)

SELECT *
FROM stationstats

```

В результате у нас получилась таблица, насчитывающая 802 записи, по одной для каждого пункта проката, действовавшего в 2015 году, которая выглядит примерно так:

Row	station_name	duration_weekdays	duration_weekends	numtrips_weekdays	numtrips_weekends	bikes_count	distance_from_city_center
1	Borough Road, Elephant & Castle	1109.932...	2125.095...	5749	1774	29	0.126...
2	Webber Street, Southwark	795.439...	938.357...	6517	1619	34	0.164...
3	Great Suffolk Street, The Borough	802.530...	1018.310...	8418	2024	18	0.193...

Кластеризация

Так же как при обучении с учителем, кластеризация выполняется с помощью инструкции `CREATE MODEL` для таблицы, созданной в предыдущем разделе, но при этом необходимо удалить поле `station_name`, потому что оно однозначно идентифицирует каждый пункт проката:

```

CREATE OR REPLACE MODEL ch09eu.london_station_clusters
OPTIONS(model_type='kmeans',
        num_clusters=4,
        standardize_features = true) AS
SELECT * EXCEPT(station_name)
FROM ch09eu.stationstats

```

В параметре `model_type` указано значение `kmeans`. Если параметр `num_clusters` отсутствует, BigQuery автоматически выберет разумное значение, основываясь на количестве записей в таблице. Также для этого набора данных необходимо указать еще один параметр, `standardize_features`, потому что разные столбцы имеют очень разные диапазоны. Расстояние до центра города составляет не-

сколько километров, а количество и продолжительность поездок измеряются тысячами. Поэтому было бы неплохо, чтобы BigQuery масштабировала эти значения в диапазоны с нулевым средним и единичной дисперсией.

Исследование кластеров

Чтобы определить, к какому кластеру принадлежит конкретный пункт проката, можно использовать `ML.PREDICT`. Вот запрос, который ищет кластер для каждого пункта проката со словом «Kennington» в названии:

```
SELECT * except(nearest_centroids_distance)
FROM ML.PREDICT(MODEL ch09eu.london_station_clusters,
(SELECT * FROM ch09eu.stationstats
 WHERE REGEXP_CONTAINS(station_name, 'Kennington')))
```

Он возвращает следующие результаты:

Row	CENTROID_ ID	station_ name	duration_ weekdays	duration_ weekends	numtrips_ weekdays	numtrips_ weekends	bikes_ count	distance_ from_city_ center
1	2	Kennington Road, Vauxhall	1209.433...	1720.598...	8135	2975	26	0.891...
2	2	Kennington Lane Rail Bridge, Vauxhall	979.391...	1812.217...	20263	5014	28	2.175...
3	2	Cotton Garden Estate, Kennington	1572.919...	997.949...	5313	1600	14	1.117...
4	3	Kennington Station, Kennington	1689.587...	3579.285...	4875	1848	15	1.298...

Несколько пунктов проката на Kennington принадлежит центроиду 2, а остальные — центроиду 3.¹ Для анализа этих групп можно исследовать атрибуты центроидов:

```
SELECT *
FROM ML.CENTROIDS(MODEL ch09eu.london_station_clusters)
ORDER BY centroid_id
```

¹ Алгоритм *k*-средних чувствителен к выбору начальных точек, а поскольку они выбираются случайным образом, вы можете получить немного отличающиеся результаты.

Этот запрос возвращает таблицу, содержащую по одной записи для каждого атрибута кластера:

Row	centroid_id	feature	numerical_value	categorical_value.category	categorical_value.value
1	1	distance_from_city_center	2.978...		
2	1	bikes_count	10.013...		
3	1	numtrips_weekends	8273.849...		

Эту таблицу можно свести, выполнив следующий запрос:

```
CREATE TEMP FUNCTION cvalue(x ANY TYPE, col STRING) AS (
  (SELECT value from unnest(x) WHERE name = col)
);

WITH T AS (
  SELECT
    centroid_id,
    ARRAY_AGG(STRUCT(feature AS name,
                      ROUND(numerical_value,1) AS value)
              ORDER BY centroid_id) AS cluster
  FROM ML.CENTROIDS(MODEL ch09eu.london_station_clusters)
  GROUP BY centroid_id
)

SELECT
  CONCAT('Cluster#', CAST(centroid_id AS STRING)) AS centroid,
  cvalue(cluster, 'duration_weekdays') AS duration_weekdays,
  cvalue(cluster, 'duration_weekends') AS duration_weekends,
  cvalue(cluster, 'numtrips_weekdays') AS numtrips_weekdays,
  cvalue(cluster, 'numtrips_weekends') AS numtrips_weekends,
  cvalue(cluster, 'bikes_count') AS bikes_count,
  cvalue(cluster, 'distance_from_city_center') AS distance_from_city_center
FROM T
ORDER BY centroid_id ASC
```

И получить такой результат:

Row	centroid	duration_weekdays	duration_weekends	numtrips_weekdays	numtrips_weekends	bikes_count	distance_from_city_center
1	Cluster#1	1362.6	1968.4	25427.3	8273.8	10.0	3.0
2	Cluster#2	1193.5	1738.1	8457.4	2584.3	21.0	3.0
3	Cluster#3	1675.0	2460.5	4702.4	2136.8	14.9	6.7
4	Cluster#4	1124.0	1543.1	8519.0	2342.1	5.7	4.1

Для визуализации этой таблицы в веб-интерфейсе BigQuery щелкните на **Explore in Data Studio** (Исследовать в Data Studio), а затем выберите **Table with bars**. (Таблица со столбиками). Сделайте столбец **centroid** областью определения, а оставшиеся столбцы — значениями. Результат показан на рис. 9.9.



Рис. 9.9. Атрибуты кластеров

На рис. 9.9 видно, что кластер 1 объединяет чрезвычайно загруженные пункты проката (см. столбец с количеством поездок (**num_trips**)), находящиеся недалеко от центра города. Кластер 2 объединяет менее загруженные пункты проката недалеко от центра города. Кластер 3 включает удаленные от центра города пункты проката, которые, скорее всего, пользуются большим спросом в выходные для длительных поездок (это единственные пункты проката, где число поездок в выходные дни превышает число поездок в будние дни). Кластер 4 включает совсем маленькие пункты проката (см. столбец **bikes_count**), находящиеся вне центра города, возможно, в жилых районах. Основываясь на этих характеристиках и немного зная Лондон, можно придумать описательные названия для этих кластеров. Кластеру 1, вероятно, подойдет название «Туристическая зона», кластеру 2 — «Деловой район», кластеру 3 — «Дневные поездки», а кластеру 4 — «Пригородные пункты проката».

Принятие решений на основе данных

Теперь полученные кластеры можно использовать для принятия различных решений. Предположим, что вы только что получили финансирование и можете увеличить число стоек для парковки велосипедов. На каких пунктах проката вы увеличили бы число парковочных мест? Не имея результатов кластеризации, можно было бы предположить, что наиболее подходящими кандидатами на увеличение числа мест являются пункты проката с большим количеством поездок и недостаточным количеством велосипедов — пункты в кластере 1. Но выполнив кластеризацию, вы теперь знаете, что эта группа пунктов проката обслуживает в основном туристов. Они не голосуют на выборах, поэтому добавим велосипеды в кластер 4 (пригородные пункты проката).

Рассмотрим еще один пример: допустим, вы решили поэкспериментировать с новым типом замков. В каком кластере пунктов проката лучше провести этот эксперимент? Наиболее логичным, на первый взгляд, это было бы в деловом районе, то есть пункты проката с большим количеством велосипедов и достаточно большим числом поездок, чтобы обеспечить представительность

результатов А/В-тестирования. С другой стороны, если у вас появилась возможность приобрести шоссейные (гоночные) велосипеды, на каких пунктах проката их лучше разместить? Кластер 3, включающий пункты проката, от которых люди отправляются в поездки за город на весь день, кажется удачным выбором.

Очевидно, вы могли бы прийти к аналогичным решениям, каждый раз заново анализируя данные. Однако кластеризация, придумывание описательных имен и их использование для принятия решений выглядят намного проще и понятнее.

Рекомендательные системы

Коллаборативная (совместная) фильтрация позволяет подбирать рекомендуемые товары для пользователей или выбирать целевые группы. Отправной точкой является таблица с тремя столбцами: идентификатор пользователя, идентификатор товара и рейтинг, присвоенный пользователем данному товару. Эта таблица может быть разреженной — от пользователей не требуется, чтобы они оценили все имеющиеся товары. Опираясь только на имеющиеся оценки, метод коллаборативной фильтрации позволяет находить похожих пользователей или похожие товары и определять оценку, которую пользователь дал бы товару, которого он никогда не видел. Затем, исходя из полученных результатов, мы можем рекомендовать товары с самыми высокими прогнозируемыми оценками или выбирать для товаров целевые группы пользователей с самыми высокими прогнозируемыми оценками.

Набор данных MovieLens

Чтобы наглядно показать способы реализации рекомендательных систем, давайте используем набор данных MovieLens. Это набор отзывов к фильмам, выпущенный исследовательской лабораторией GroupLens (<https://grouplens.org/about/what-is-grouplens/>) на кафедре информатики Университета штата Миннесота при финансовой поддержке Национального научного фонда США.

В Cloud Shell загрузите данные в таблицу BigQuery, как показано ниже:

```
curl -O 'http://files.grouplens.org/datasets/movielens/ml-20m.zip'
unzip ml-20m.zip
bq --location=EU load --source_format=CSV \
  --autodetect ch09eu.movielens_ratings ml-20m/ratings.csv
bq --location=EU load --source_format=CSV \
  --autodetect ch09eu.movielens_movies_raw ml-20m/movies.csv
```

Получившаяся таблица с рейтингами включает следующие столбцы:

Row	userId	movieId	rating	timestamp
1	70141	6219	2.0	1070338674
2	70159	2657	2.0	1427155558

Вот короткий и быстрый запрос, помогающий получить некоторые характеристики о наборе данных:

```
SELECT
  COUNT(DISTINCT userId) numUsers,
  COUNT(DISTINCT movieId) numMovies,
  COUNT(*) totalRatings
FROM ch09eu.movielens_ratings
```

Согласно результатам, возвращаемым этим запросом, набор данных включает отзывы более чем 138 000 пользователей на почти 27 000 фильмов и чуть более 20 миллионов оценок (рейтингов), что подтверждает успешную загрузку данных.

Рассмотрим несколько первых фильмов, выполнив такой запрос:

```
SELECT *
FROM ch09eu.movielens_movies_raw
WHERE movieId < 5
```

Как видите, жанр фильма определяется строкой в специальном формате:

Row	movieId	title	genres
1	3	Grumpier Old Men (1995)	Comedy Romance
2	4	Waiting to Exhale (1995)	Comedy Drama Romance
3	2	Jumanji (1995)	Adventure Children Fantasy

Строку с перечнем жанров можно преобразовать в массив и сохранить в другой таблице:

```
CREATE OR REPLACE TABLE ch09eu.movielens_movies AS
SELECT
  * REPLACE(SPLIT(genres, "|") AS genres)
FROM
  ch09eu.movielens_movies_raw
```

Теперь таблица выглядит так:

Row	movieId	title	genres
1	4	Waiting to Exhale (1995)	Comedy
			Drama

Row	movieId	title	genres
			Romance
2	3	Grumpier Old Men (1995)	Comedy
			Romance
3	2	Jumanji (1995)	Adventure
			Children
			Fantasy

Загрузив набор данных MovieLens, можно приступить к коллаборативной фильтрации.

Разложение матрицы

Разложение матрицы — это метод коллаборативной фильтрации, основанный на разложении матрицы рейтингов на два вектора, которые называют *пользовательскими факторами* и *факторами товаров*. Вектор пользовательских факторов является представлением `user_col` с небольшим числом измерений, а вектор факторов товаров — представлением `item_col`.

Создать рекомендательную модель можно с помощью следующего запроса:

```
-- это не окончательная модель; см. movie_recommender_16
CREATE OR REPLACE MODEL ch09eu.movie_recommender
options(model_type='matrix_factorization',
        user_col='userId', item_col='movieId', rating_col='rating')
AS

SELECT
userId, movieId, rating
FROM ch09eu.movielens_ratings
```

Обратите внимание, что модель создается как обычно, за исключением того, что в параметре `model_type` указывается тип модели `matrix_factorization` и требуется определить, какую роль играют столбцы в настройках коллаборативной фильтрации.

Получившейся модели потребовался час для обучения, при этом потери на обучающих данных сначала выглядят крайне плохо, но затем сходятся к нулю в течение следующих четырех итераций:¹

¹ Причина столь существенного изменения продолжительности итераций обусловлена тем, что основной алгоритм оптимизации обрабатывает пользователей в одной итера-

Итерация	Потери на обучающих данных	Потери на контрольных данных	Продолжительность (с)
4	0.5734	172.4057	180.99
3	0.5826	187.2103	1,040.06
2	0.6531	4758.2944	219.46
1	1.9776	6297.2573	1093.76
0	63 287 833 220.5795	168 995 333.0464	1091.21

Однако потери на контрольных данных остаются довольно высокими — намного больше потерь на обучающих данных. Это говорит о переобучении, поэтому понадобится регуляризация.

Попробуем так:

```
-- это не окончательная модель; см. movie_recommender_16
CREATE OR REPLACE MODEL ch09eu.movie_recommender_l2
options(model_type='matrix_factorization',
        user_col='userId', item_col='movieId',
        rating_col='rating', l2_reg=0.2)
```

AS

```
SELECT
userId, movieId, rating
FROM ch09eu.movielens_ratings
```

Теперь мы имеем более быструю сходимость (три итерации вместо пяти) и намного меньший эффект переобучения:

Итерация	Потери на обучающих данных	Потери на контрольных данных	Продолжительность (с)
2	0.6509	1.4596	198.17
1	1.9829	33 814.3017	1 066.06
0	481 434 346 060.7928	2 156 993 687.7928	1 024.59

По умолчанию BigQuery выбирает количество факторов, равное \log_2 от количества записей. В данном случае мы имеем 20 миллионов записей, поэтому число факторов было выбрано равным 24. Как и в случае с алгоритмом выбора числа кластеров в методе кластеризации k -средних, это вполне разумное значение по

ции, а фильмы — в следующей, а также тем, что число пользователей намного больше, чем фильмов.

умолчанию, но нередко имеет смысл попробовать также числа примерно на 50% больше (36) и на треть меньше (16):¹

```
CREATE OR REPLACE MODEL ch09eu.movie_recommender_16
options(model_type='matrix_factorization',
        user_col='userId', item_col='movieId',
        rating_col='rating', l2_reg=0.2, num_factors=16)
AS

SELECT
userId, movieId, rating
FROM ch09eu.movielens_ratings
```

В итоге мы обнаружили, что потери на контрольных данных при `num_factors=16` оказались ниже (0.97), чем при `num_factors=36` (1.67) и `num_factors=24` (1.45). Мы могли бы продолжить эксперименты, но, скорее всего, столкнулись бы с эффектом уменьшения отдачи. Поэтому остановимся на числе факторов, равном 16, и продолжим.

Получение рекомендаций

Имея обученную модель, вы можете генерировать рекомендации. Например, найдем лучшие комедийные фильмы, которые можно рекомендовать пользователю с идентификатором (`userId`) 903:

```
SELECT * FROM
ML.PREDICT(MODEL ch09eu.movie_recommender_16, (
  SELECT
    movieId, title, 903 AS userId
  FROM ch09eu.movielens_movies, UNNEST(genres) g
  WHERE g = 'Comedy'
))
ORDER BY predicted_rating DESC
LIMIT 5
```

В этом запросе мы вызываем функцию `ML.PREDICT`, передаем ей обученную модель рекомендаций и параметры `movieId` и `userId`, для которых требуется получить прогноз. В данном случае мы хотим получить рекомендации для единственного пользователя (с `userId = 903`), ограничившись фильмами в жанре комедии. Вот результат:

¹ Это может показаться странным. Почему на треть, а не на половину? Фактически идея состоит в том, что если за основу взять число 16, тогда 24 будет больше него на 50%. Наша цель — попробовать геометрическую прогрессию значений-кандидатов для `num_factors`, чтобы быстро охватить пространство кандидатов. Если вы решите испытать больше трех значений `num_factors`, попробуйте использовать последовательность значений `num_factors`, каждое из которых примерно в $\sqrt{2}$ раз превышает предыдущее. Например: 4, 6, 8, 12, 16, 24, 32, 48, 64 и т. д.

Row	predicted_rating	movieId	title	userId
1	4.747231361947591	107434	Diplomatic Immunity (2009–)	903
2	4.372639637398302	62206	Supermarket Woman (Sûpâ no onna) (1996)	903
3	4.325021974040314	122441	Tales That Witness Madness (1973)	903
4	4.296062517241643	120313	Otakus in Love (2004)	903
5	4.277251207896746	130347	Bill Hicks: Sane Man (1989)	903

Фильтрация фильмов, оцененных пользователем

В этот список могут входить также фильмы, которые пользователь уже видел и оценивал в прошлом. Давайте удалим их:

```
SELECT * FROM
ML.PREDICT(MODEL ch09eu.movie_recommender_16, (
  WITH seen AS (
    SELECT ARRAY_AGG(movieId) AS movies
    FROM ch09eu.movielens_ratings
    WHERE userId = 903
  )
  SELECT
    movieId, title, 903 AS userId
  FROM ch09eu.movielens_movies, UNNEST(genres) g, seen
  WHERE g = 'Comedy' AND movieId NOT IN UNNEST(seen.movies)
))
ORDER BY predicted_rating DESC
LIMIT 5
```

Так получилось, что для этого пользователя данный запрос вернул тот же набор фильмов — фильмов с самыми высокими прогнозируемыми рейтингами, которые пользователь еще не видел.

Выбор целевой аудитории

В предыдущем разделе мы показали, как определить фильмы, которые указанный пользователь мог бы оценить наиболее высоко. Но иногда требуется решить обратную задачу — найти клиентов, которые оценят указанный товар. К примеру, вы хотите получить больше отзывов для фильма с `movieId=96481`, имеющего только одну оценку, и с этой целью вам нужно разослать предложения 100 пользователям, которые с большой долей вероятности высоко его оценят. Найти таких пользователей можно так:

```
SELECT * FROM
ML.PREDICT(MODEL ch09eu.movie_recommender_16, (
  WITH allUsers AS (
    SELECT DISTINCT userId
```

```

    FROM ch09eu.movielens_ratings
  )
  SELECT
    96481 AS movieId,
    (SELECT title FROM ch09eu.movielens_movies WHERE movieId=96481) title,
    userId
  FROM
    allUsers
))
ORDER BY predicted_rating DESC
LIMIT 100

```

В результате он возвращает целевую аудиторию из 100 пользователей, первые пять из которых перечислены здесь:

Row	predicted_rating	movieId	title	userId
1	4.8586009640376915	96481	American Mullet (2001)	54192
2	4.670093338552966	96481	American Mullet (2001)	84240
3	4.544395037073204	96481	American Mullet (2001)	109638
4	4.422718574118088	96481	American Mullet (2001)	26606
5	4.410969328468145	96481	American Mullet (2001)	138139

Пакетные прогнозы для всех пользователей и фильмов

А что, если вы захотите спрогнозировать оценку для каждой комбинации пользователя и фильма? Чтобы не извлекать отдельных пользователей и фильмы, как в предыдущем запросе, можно воспользоваться удобной функцией пакетного прогнозирования для всех `movieId` и `userId`, имевшихся на этапе обучения:

```

SELECT *
FROM ML.RECOMMEND(MODEL ch09eu.movie_recommender_16)

```

Как было показано в предыдущем разделе, можно отфильтровать фильмы, которые пользователь уже видел и оценивал в прошлом. Причина, по которой ранее просмотренные фильмы не фильтруются по умолчанию, заключается в том, что иногда (например, представьте, что вы подбираете рекомендуемые рестораны) требуется рекомендовать что-то (например, рестораны), что пользователю уже понравилось в прошлом.

Включение информации о пользователях и фильмах

Подход на основе разложения матрицы не использует информацию о пользователях или фильмах, которая недоступна в матрице рейтингов. Однако на практике у исследователя часто имеется дополнительная информация о поль-

зователях (например, город проживания, годовой доход и расход и т. д.) и почти всегда есть масса дополнительной информации о товарах. Как включить эту информацию в модель рекомендаций?

Прежде всего важно понимать, что пользовательские факторы и факторы товаров, получающиеся в результате разложения матрицы, в итоге являются краткими представлениями информации о пользователях и товарах, доступных в матрице рейтингов. Мы можем объединить эти данные с другой имеющейся у нас информацией и создать регрессионную модель для прогнозирования рейтинга.

Получение факторов

Получить пользовательские факторы или факторы товаров можно с помощью `ML.WEIGHTS`. Например, ниже показано, как получить факторы товаров для `movieId=96481` и пользовательские факторы для `userId=54192`:

```
SELECT
  processed_input
  , feature
  , TO_JSON_STRING(factor_weights)
  , intercept
FROM ML.WEIGHTS(MODEL ch09eu.movie_recommender_16)
WHERE
  (processed_input = 'movieId' AND feature = '96481')
  OR
  (processed_input = 'userId' AND feature = '54192')
```

Этот запрос вернет следующие результаты:

Row	processed_input	feature	f0_	intercept
1	movieId	96481	[{"factor":16,"weight":0.01274324364248563}, {"factor":15,"weight":-0.026002830400362179}, {"factor":14,"weight":-0.0088894978851240675}, {"factor":13,"weight":0.010309411637259363}, {"factor":12,"weight":-0.025990228913849212}, {"factor":11,"weight":0.0037023423385396021}, {"factor":10,"weight":-0.0016743710047063861}, {"factor":9,"weight":0.018434530705228803}, {"factor":8,"weight":-0.0016500835388799462}, {"factor":7,"weight":-0.021652088589080184}, {"factor":6,"weight":-0.00097969747732716637}, {"factor":5,"weight":-0.056352201014532581}, {"factor":4,"weight":-0.025090456181039382},	-1.1915305828542884

Row	processed_input	feature	f0_	intercept
			{“factor”:3,“weight”:0.015317626028966519}, {“factor”:2,“weight”:-0.00046084151232374118}, {“factor”:1,“weight”:-0.0009461271544545048}}	
2	userId	54192	[{“factor”:16,“weight”:-0.66257902781387934}, {“factor”:15,“weight”:-0.089502881890795027}, {“factor”:14,“weight”:-0.14498342867805328}, {“factor”:13,“weight”:0.57708118940369757}, {“factor”:12,“weight”:-0.25409266698347688}, {“factor”:11,“weight”:0.243523510689305}, {“factor”:10,“weight”:0.48314159427498959}, {“factor”:9,“weight”:0.21335694312220596}, {“factor”:8,“weight”:0.34206958377350211}, {“factor”:7,“weight”:-0.076313491055098021}, {“factor”:6,“weight”:0.21214183741037482}, {“factor”:5,“weight”:0.19387028511697624}, {“factor”:4,“weight”:-0.42699681695332414}, {“factor”:3,“weight”:0.046570444717220438}, {“factor”:2,“weight”:0.25934273163373722}, {“factor”:1,“weight”:-0.18839802656522864}]	2.511409230366029

Умножая эти веса и прибавляя поправку **intercept**, можно получить прогнозируемый рейтинг для данной комбинации **movieId** и **userId**.

Эти веса также играют роль низкоразмерного представления фильма и поведения пользователя. Вы можете создать регрессионную модель для прогнозирования рейтинга с учетом пользовательских факторов, факторов товаров и любой другой известной информации о пользователях и товарах.

Создание входных признаков

Набор данных MovieLens не содержит никакой информации о пользователях и очень мало информации о самих фильмах. Чтобы это проиллюстрировать, добавим некоторую синтетическую информацию о пользователях:

```
CREATE OR REPLACE TABLE ch09eu.movielens_users AS
SELECT
  userId
  , RAND() * COUNT(rating) AS loyalty
  , CONCAT(SUBSTR(CAST(userId AS STRING), 0, 2)) AS postcode
FROM
```

```
ch09eu.movielens_ratings
GROUP BY userId
```

Входные признаки, относящиеся к пользователям, можно получить путем соединения пользовательской таблицы с весами модели машинного обучения и выбора информации о пользователях и пользовательских факторов из массива весов:

```
WITH userFeatures AS (
  SELECT
    u.*
    , (SELECT ARRAY_AGG(weight) FROM UNNEST(factor_weights)) AS user_factors
  FROM
    ch09eu.movielens_users u
  JOIN
    ML.WEIGHTS(MODEL ch09eu.movie_recommender_16) w
  ON
    processed_input = 'userId' AND feature = CAST(u.userId AS STRING)
)

SELECT * FROM userFeatures
LIMIT 5
```

Этот запрос вернет следующие признаки пользователей (перед передачей этих данных в регрессионную модель нужно удалить userId):

Row	userId	loyalty	postcode	user_factors
1	65536	72.51794801197904	65	0.038901538776462
				0.0019075355240976716
				0.011537776936285278
				-0.0322503841197857
				0.046464397209825425
				-0.015348467879503527
				0.05865111283285229
				0.04859058815259179
				0.017664456774125117
				0.006847553039523945
				0.012585216564478762
				-0.06506297976701378
				-0.005041156227839918
				-0.04187860699038322
				0.006216526560890197
				0.02711744261644579

Точно так же можно получить признаки товаров на основе данных о фильмах, только при этом вам придется решить, как обрабатывать жанр, потому что фильм относится к нескольким жанрам. Если вы решите для каждого жанра создать отдельную запись в обучающих данных, сконструировать признаки товаров можно так:

```
WITH productFeatures AS (
  SELECT
    p.* EXCEPT(genres)
    , g
    , (SELECT ARRAY_AGG(weight) FROM UNNEST(factor_weights)) AS product_factors
  FROM
    ch09eu.movielens_movies p, UNNEST(genres) g
  JOIN
    ML.WEIGHTS(MODEL ch09eu.movie_recommender_16) w
  ON
    processed_input = 'movieId' AND feature = CAST(p.movieId AS STRING)
)

SELECT * FROM productFeatures
LIMIT 5
```

Этот запрос сгенерирует записи в следующей форме:

Row	movieId	title	g	product_factors
1	1450	Prisoner of the Mountains (Kavkazsky plennik) (1996)	War	0.9883690055578206
				1.3052751077485096
				-1.4000285383517228
				1.3901032474256991
				-0.32863748198986686
				-0.7688057246956399
				-1.1853591273232054
				-0.4553668299329251
				-0.14564591302024543
				-0.18609388556738163
				-0.3547198526732644
				0.06067380147330148
				-0.2733324088164271
				1.8302213060412562
				0.4753820155626278
				1.559946725190114

Комбинируя эти два предложения WITH и извлекая рейтинг, соответствующий комбинации movieId-userId (если она существует в таблице рейтингов), можно создать обучающий набор данных:¹

```
CREATE OR REPLACE TABLE ch09eu.movielens_hybrid_dataset AS

WITH userFeatures AS (
  SELECT
    u.*,
    (SELECT ARRAY_AGG(weight) FROM UNNEST(factor_weights)) AS user_factors
  FROM
    ch09eu.movielens_users u
  JOIN
    ML.WEIGHTS(MODEL ch09eu.movie_recommender_16) w
  ON
    processed_input = 'userId' AND feature = CAST(u.userId AS STRING)
),
productFeatures AS (
  SELECT
    p.* EXCEPT(genres)
    , g
    , (SELECT ARRAY_AGG(weight) FROM UNNEST(factor_weights)) AS product_factors
  FROM
    ch09eu.movielens_movies p, UNNEST(genres) g
  JOIN
    ML.WEIGHTS(MODEL ch09eu.movie_recommender_16) w
  ON
    processed_input = 'movieId' AND feature = CAST(p.movieId AS STRING)
)

SELECT p.* EXCEPT(movieId), u.* EXCEPT(userId), rating
FROM productFeatures p, userFeatures u
JOIN
  ch09eu.movielens_ratings r
ON
  r.movieId = p.movieId AND r.userId = u.userId
```

Вот как выглядит одна из записей в этой таблице:

1	Hunted, The (2003)	Action	2.6029616190628015	692.7156232519949	70	0.026523240535672774	2.0
			0.33485455845698525			0.0019319939217823622	
			0.31628840722516194			-0.0020145595411925534	
			-0.3075233831543138			-0.002646563034985453	
			-0.4473419662482839			-0.01594551937825673	

¹ См. 09_bqml/hybrid.sql в репозитории GitHub с примерами для этой книги (<https://github.com/GoogleCloudPlatform/bigquery-oreilly-book>).

1	Hunted, The (2003)	Action	2.6029616190628015	692.7156232519949	70	0.026523240535672774	2.0
			−1.0222758233057185			−0.010801066706191506	
			−0.42418301494313826			4.772572135005211E-4	
			−1.2447809221572947			0.014766024570817101	
			−0.20242685993451942			−0.007500869241538576	
			1.330350771422776			−0.020383420117709883	
			−0.3354935275410769			−0.007863867111381763	
			0.32404375319192513			0.019901597021923123	
			1.402657314320568			−0.003178194776711233	
			0.4728896971092763			0.013146874239054253	
			−0.5743444547904143			−0.0017117741950437	
			0.35632448579921905			−0.030130776462043048	

Фактически мы получили пару атрибутов, характеризующих фильм, массив факторов товара, соответствующих фильму, пару атрибутов, характеризующих пользователя, и массив пользовательских факторов, соответствующих этому пользователю. Они образуют входные признаки для «гибридной» модели рекомендаций, которая строится на основе модели разложения матрицы и добавляет метаданные о пользователях и фильмах.

Обучение гибридной модели рекомендаций

На момент написания этой книги механизм машинного обучения BigQuery ML не мог обрабатывать входные признаки регрессионной модели в виде массивов. Поэтому определим функцию, преобразующую массив в структуру, в которой элементы массива являются полями:

```
CREATE OR REPLACE FUNCTION ch09eu.arr_to_input_3(a ARRAY<FLOAT64>)
RETURNS STRUCT<a1 FLOAT64, a2 FLOAT64, a3 FLOAT64> AS (
STRUCT(
  a[OFFSET(0)]
  , a[OFFSET(1)]
  , a[OFFSET(2)]
));
```

Если теперь выполнить такой запрос:

```
SELECT
  ch09eu.arr_to_input_3(a).*
FROM
  (SELECT [34.23, 43.21, 63.21] AS a)
```

Он вернет:

Row	a1	a2	a3
1	34.23	43.21	63.21

При желании можно определить аналогичную функцию с именем `ch09eu.arr_to_input_16_users` для преобразования массива пользовательских факторов в именованные столбцы и функцию для преобразования массива факторов товаров.¹ После этого можно связать метаданные о пользователях и товарах с пользовательскими факторами и факторами товаров, полученными из разложения матрицы, и создать регрессионную модель для прогнозирования рейтинга:

```
CREATE OR REPLACE MODEL ch09eu.movielens_recommender_hybrid
OPTIONS(model_type='linear_reg', input_label_cols=['rating'])
AS
SELECT
  * EXCEPT(user_factors, product_factors)
  , ch09eu.arr_to_input_16_users(user_factors).*
  , ch09eu.arr_to_input_16_products(product_factors).*
FROM
  ch09eu.movielens_hybrid_dataset
```

Нет смысла рассматривать результаты оценки этой модели, потому что при создании обучающего набора мы использовали фиктивную информацию о пользователях (обратите внимание на вызов `RAND()` в создании столбца `loyalty`). Мы представили этот пример, только чтобы продемонстрировать, как можно реализовать подобные рекомендации. Конечно, при желании получить нечто более сложное мы могли бы обучить модель `dnn_regressor` и оптимизировать ее гиперпараметры. Но прежде чем заходить так далеко, желательно рассмотреть возможность использования таблиц AutoML, о которых рассказывается в следующем разделе.

Нестандартные модели машинного обучения в GCP

Механизм машинного обучения BigQuery ML предлагает широкий выбор моделей,² которые можно быстро создать и обучить, но, кроме того, есть также механизм AutoML, предлагающий современную высококачественную модель для задач, ради которых можно пожертвовать часами или даже днями обучения. Библиотеки Keras и TensorFlow, обеспечивающие низкоуровневое управление процессом машинного обучения, позволяют проектировать, разрабатывать

¹ См. `09_bqml/arr_to_input16.sql` в репозитории GitHub с примерами для этой книги.

² К тому времени, когда вы будете читать эти строки, `automl` вполне может стать одним из типов моделей, поддерживаемых в BigQuery.

и развертывать нестандартные модели машинного обучения. Мы советуем для начала воспользоваться средствами BigQuery ML для обучения моделей на структурированных или полуструктурированных данных и, в зависимости от имеющихся у вас навыков и важности решаемой задачи, использовать AutoML или Keras для тонкой настройки машинного обучения.

Настройка гиперпараметров

В машинном обучении используется много параметров, которые выбираются довольно произвольно. К ним относятся, например, скорость обучения, уровень регуляризации L2, количество слоев и узлов в нейронной сети, максимальная глубина дерева решений и количество факторов в матричном разложении. Часто бывает так, что выбор других значений может дать в результате более точную модель (в смысле величины ошибки на контрольном наборе данных). Выбор значений для этих параметров называется *настройкой гиперпараметров*.

Настройка гиперпараметров с использованием скриптов

Возьмем для примера модель кластеризации методом k -средних. На вкладке Evaluation (Оценка) в веб-интерфейсе BigQuery (а также выполнив запрос `SELECT * FROM ML.EVALUATE`) можно увидеть индекс Дэвиса — Болдина (Davies — Bouldin), который помогает определить оптимальное число кластеров в данных (чем меньше это значение, тем оптимальнее выполнена кластеризация).

Например, вот скрипт, пытающийся применить разное число кластеров:

```
DECLARE NUM_CLUSTERS INT64 DEFAULT 3;
DECLARE MIN_ERROR FLOAT64 DEFAULT 1000.0;
DECLARE BEST_NUM_CLUSTERS INT64 DEFAULT -1;
DECLARE MODEL_NAME STRING;

WHILE NUM_CLUSTERS < 8 DO

  SET MODEL_NAME = CONCAT('ch09eu.london_station_clusters_',
                           CAST(NUM_CLUSTERS AS STRING));

  CREATE OR REPLACE MODEL MODEL_NAME
  OPTIONS(model_type='kmeans',
          num_clusters=NUM_CLUSTERS,
          standardize_features = true) AS
  SELECT * except(station_name)
  from ch09eu.stationstats;

  SET error = (SELECT davies_bouldin_index FROM ML.EVALUATE(MODEL MODEL_NAME));
  IF error < MIN_ERROR THEN
    SET MIN_ERROR = error;
    SET BEST_NUM_CLUSTERS = NUM_CLUSTERS;
  END IF;

  SET NUM_CLUSTERS = NUM_CLUSTERS + 1;

END WHILE
```

Настройка гиперпараметров на Python

При желании то же самое можно реализовать на Python, используя его многопоточность для ограничения количества одновременно выполняемых запросов:¹

```
def train_and_evaluate(num_clusters: Range, max_concurrent=3):
    # поиск по сетке предполагает опробование всех возможных значений
    # в диапазоне
    params = []
    for k in num_clusters.values():
        params.append(Params(k))

    # Запустить все задания
    print('Grid search of {} possible parameters'.format(len(params)))
    pool = ThreadPool(max_concurrent)
    results = pool.map(lambda p: p.run(), params)

    # Отсортировать в порядке возрастания
    return sorted(results, key=lambda p: p._error)
```

Метод `run()` класса `Params` в этом примере выполняет соответствующие запросы обучения и оценки:

```
class Params:
    def __init__(self, num_clusters):
        self._num_clusters = num_clusters
        self._model_name = (
            'ch09eu.london_station_clusters_{}'.format(num_clusters))
        self._train_query = """
            CREATE OR REPLACE MODEL {}
            OPTIONS(model_type='kmeans',
                    num_clusters={},
                    standardize_features = true) AS
            SELECT * except(station_name)
            from ch09eu.stationstats
            """.format(self._model_name, self._num_clusters)
        self._eval_query = """
            SELECT davies_bouldin_index AS error
            FROM ML.EVALUATE(MODEL {});
            """.format(self._model_name)
        self._error = None

    def run(self):
        bq = bigquery.Client(project=PROJECT)
        job = bq.query(self._train_query, location='EU')
        job.result() # ждать завершения задания
        evaldf = bq.query(self._eval_query, location='EU').to_dataframe()
        self._error = evaldf['error'][0]
        return self
```

¹ Полный код см. в файле `09_bqml/hyperparam.ipynb` в репозитории GitHub с примерами для этой книги.

Выполнив поиск в диапазоне [3,9], мы обнаружили, что минимальная ошибка соответствует числу кластеров, равному 7:

ch09eu.london_station_clusters_7	1.551265	7
ch09eu.london_station_clusters_9	1.571020	9
ch09eu.london_station_clusters_6	1.571398	6
ch09eu.london_station_clusters_4	1.596398	4
ch09eu.london_station_clusters_8	1.621974	8
ch09eu.london_station_clusters_5	1.660766	5
ch09eu.london_station_clusters_3	1.681441	3

Настройка гиперпараметров с использованием AI Platform

В обоих методах настройки гиперпараметров, представленных выше, мы испытали каждое возможное значение параметра, попадающее в заданный диапазон. По мере роста числа возможных параметров поиск по сетке становится все более ресурсоемким. Хорошо было бы иметь более эффективный алгоритм поиска, и такой алгоритм предлагает Cloud AI Platform. Эту службу можно использовать для настройки гиперпараметров любых моделей (не только на основе TensorFlow). Давайте попробуем для примера применить ее для выбора признаков и количества узлов в модели DNN.¹

Сначала создадим конфигурационный файл, определяющий диапазоны значений параметров, количество одновременно выполняемых запросов и общее число попыток:

```
trainingInput:
  scaleTier: CUSTOM
  masterType: standard # See: https://cloud.google.com/ml-engine/docs/tensorflow/machine-types
hyperparameters:
  goal: MINIMIZE
  maxTrials: 50
  maxParallelTrials: 2
  hyperparameterMetricTag: mean_absolute_error
  params:
    - parameterName: afternoon_start
      type: INTEGER
      minValue: 9
      maxValue: 12
      scaleType: UNIT_LINEAR_SCALE
    - parameterName: afternoon_end
      type: INTEGER
      minValue: 15
      maxValue: 19
      scaleType: UNIT_LINEAR_SCALE
    - parameterName: num_nodes_0
      type: INTEGER
```

¹ Полный код примера доступен в блокноте https://github.com/GoogleCloudPlatform/bigquery-oreilly-book/blob/master/09_bqml/hyperparam.ipynb.

```

minValue: 10
maxValue: 100
scaleType: UNIT_LOG_SCALE
- parameterName: num_nodes_1
  type: INTEGER
  minValue: 3
  maxValue: 10
  scaleType: UNIT_LINEAR_SCALE

```

Обратите внимание: здесь мы определили минимальное и максимальное значение для каждого из параметров и метрику (средняя абсолютная ошибка) для минимизации. Для оптимизации мы также установили максимальное число попыток равным 50 (если бы выполнялся поиск по сетке, потребовалось бы проверить $4 \times 4 \times 90 \times 7$ — более 10 000 вариантов). То есть использование службы настройки гиперпараметров AI Platform дает нам 200-кратную экономию!

Теперь напишем программу на Python, которая посылает запросы BigQuery для обучения и оценки модели с набором параметров:

```

def train_and_evaluate(args):
    model_name = "ch09eu.bicycle_model_dnn_{}_{_}_{_}".format(
        args.afternoon_start, args.afternoon_end, args.num_nodes_0,
args.num_nodes_1
    )
    train_query = """
        CREATE OR REPLACE MODEL {}
        TRANSFORM(* EXCEPT(start_date)
            , IF(EXTRACT(dayofweek FROM start_date) BETWEEN 2 and 6,
'weekday', 'weekend')) as dayofweek
            , ML.BUCKETIZE(EXTRACT(HOUR FROM start_date), [5, {}, {}]) AS
hourofday
        )
        OPTIONS(input_label_cols=['duration'],
            model_type='dnn_regressor',
            hidden_units=[{}, {}])
        AS

        SELECT
            duration
            , start_station_name
            , start_date
        FROM `bigquery-public-data`.london_bicycles.cycle_hire
    """.format(model_name,
        args.afternoon_start,
        args.afternoon_end,
        args.num_nodes_0,
        args.num_nodes_1)
    logging.info(train_query)
    bq = bigquery.Client(project=args.project,
        location=args.location,
        credentials=get_credentials())
    job = bq.query(train_query)
    job.result() # ждать завершения задания

```

```
eval_query = """
    SELECT mean_absolute_error
    FROM ML.EVALUATE(MODEL {})
    """.format(model_name)
logging.info(eval_info)
evaldf = bq.query(eval_query).to_dataframe()
return evaldf['mean_absolute_error'][0]
```

Обратите внимание, что этот код использует определенное значение для каждого из настраиваемых параметров и возвращает среднюю абсолютную ошибку — метрику для минимизации.

Затем это значение ошибки записывается вызовом:

```
hpt.report_hyperparameter_tuning_metric(
    hyperparameter_metric_tag='mean_absolute_error',
    metric_value=error,
    global_step=1)
```

Программа обучения пересылается в службу AI Platform Training:

```
gcloud ai-platform jobs submit training $JOBNAME \
    --runtime-version=1.13 \
    --python-version=3.5 \
    --region=$REGION \
    --module-name=trainer.train_and_eval \
    --package-path=$(pwd)/trainer \
    --job-dir=gs://$BUCKET/hparam/ \
    --config=hyperparam.yaml \
    - \
    --project=$PROJECT --location=EU
```

Результаты с лучшими значениями параметров выводятся в консоль AI Platform.

AutoML

AutoML включает семейство продуктов, которые дают возможность автоматически создавать и развертывать современные модели машинного обучения, не написав ни строчки кода. В большинстве своем они опираются на применение разнообразных методов проектирования признаков, настройки гиперпараметров, поиска нейронной архитектуры, переноса обучения и методов ансамблирования для создания моделей, сопоставимых по качеству с моделями, созданными вручную ведущими специалистами по машинному обучению.



Используйте BigQuery ML, чтобы сформулировать задачу машинного обучения, — для определения признаков и меток, быстрой диагностики увеличения точности за счет использования какого-либо нового набора данных, выявления ошибок в предположениях о зависимости от времени и определения наилучшего способа представления знаний предметной области. Возможность быстро выполнять итерации, которую предлагает BigQuery

ML, бесценна, так же как и способность обучать модели без перемещения данных за пределы хранилища. После определения задачи машинного обучения вы сможете использовать AutoML, чтобы получить модель, показывающую очень высокую точность на конкретном наборе данных (с учетом заданных признаков и меток). От себя можем сказать, что AutoML вобрала в себя богатый опыт, накопленный экспертами в данной области, превзойти который будет очень непросто как с точки зрения точности, так и с точки зрения времени развертывания.

AutoML Vision, например, предлагает веб-интерфейс для загрузки изображений (или для передачи ссылок на изображения в облачном хранилище Google), определения их меток и запуска обучения модели классификации изображений или обнаружения объектов на них.

Поскольку в BigQuery обычно сохраняются структурированные или полуструктурированные данные, к ним можно применять такие модели, как AutoML Natural Language (для классификации текста и обнаружения сущностей), AutoML Tables (для регрессии, классификации и прогнозирования временных рядов по структурированным данным) и AutoML Recommendations (для создания современных моделей рекомендаций).

Чтобы использовать AutoML Tables (рис. 9.10), просто перейдите в консоль GCP, выберите таблицу в BigQuery, столбцы признаков и столбец меток, а затем кликните на **Train** (Обучить). Конечно, обучение в этом случае займет гораздо больше времени (примерно от 12 до 24 часов), но полученная в результате модель будет гораздо точнее, чем та, которую вы могли бы создать и обучить на том же наборе данных с помощью BigQuery ML.

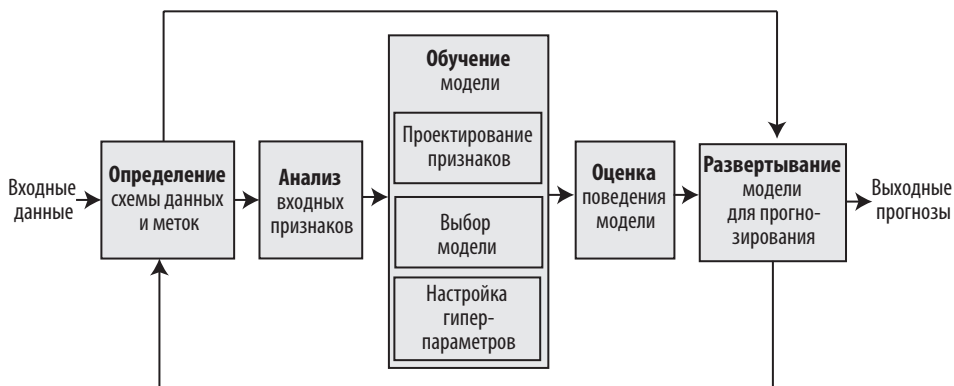


Рис. 9.10. Создание модели в AutoML Tables можно начать с выбора таблицы в BigQuery, содержащей набор обучающих данных, который был создан в ходе исследований и экспериментов в BigQuery ML. Как показывает наш опыт, применение AutoML Tables к тщательно отобранному набору обучающих данных обеспечивает особенно высокую точность прогнозирования

Поддержка TensorFlow

Даже при том, что BigQuery ML — удобная и масштабируемая служба, а AutoML предлагает широкие возможности и высокую точность, иногда требуется создать свою модель с использованием Keras или TensorFlow. Также может пригодиться возможность обучать модели с применением TensorFlow и прогнозировать с помощью BigQuery или обучать модели в BigQuery, но развертывать их в TensorFlow Serving.

Доступ к BigQuery можно получить непосредственно из кода TensorFlow и экспортировать таблицы BigQuery в записи TensorFlow, попутно преобразуя данные. Также можно выгрузить модель TensorFlow в BigQuery и экспортировать модель BigQuery в формате TensorFlow SavedModel. Мы рассмотрим эти возможности в данном разделе.

С помощью BigQueryReader из библиотеки TensorFlow

Конвейер ввода в TensorFlow способен читать данные из таблицы BigQuery с помощью объекта `BigQueryReader`. Для этого нужно сначала создать словарь признаков, перечисляющий интересующие нас столбцы:

```
features = dict(
    start_station_name=tf.FixedLenFeature([1], tf.string),
    duration=tf.FixedLenFeature([1], tf.int32))
```

Затем создайте объект `Reader`, указав отметку времени, начиная с которой должны читаться данные (потому что таблица BigQuery может получать потоковые данные в процессе чтения) и количество потоков (разделов) чтения:

```
reader = tf.contrib.cloud.BigQueryReader(project_id=PROJECT,
    dataset_id=DATASET,
    table_id=TABLE,
    timestamp_millis=TIME,
    num_partitions=NUM_PARTITIONS,
    features=features)
```

Наконец, заполните очередь разделами из таблицы BigQuery и используйте ее для чтения образцов в TensorFlow:

```
queue = tf.train.string_input_producer(reader.partitions())
row_id, examples_serialized = reader.read(queue)
examples = tf.parse_example(examples_serialized, features=features)
```

Однако у этого подхода есть некоторые недостатки. При обучении модели вам потребуется читать сразу по записям `batch_size`, перетасовывать порядок чтения между рабочими потоками, предварительно выбирать записи и т. д. Поэтому мы не рекомендуем пользоваться этим приемом на практике.

С помощью библиотеки pandas

Если таблица BigQuery не слишком большая, ее можно прочитать прямо в память, в объект `DataFrame`:

```
query = """
SELECT
    start_station_name
    , duration
FROM `bigquery-public-data`.london_bicycles.cycle_hire
GROUP BY start_station_name
"""
df = bq.query(query, location='EU').to_dataframe()
```

и затем использовать `tf.data` для чтения из pandas:

```
tf.estimator.inputs.pandas_input_fn(
    df,
    batch_size=128,
    num_epochs=10,
    shuffle=True,
    num_threads=8,
    target_column='duration'
)
```

С помощью Apache Beam/Cloud Dataflow

Если таблица слишком большая и не помещается в памяти, экспортируйте данные из BigQuery в записи TensorFlow в Google Cloud Storage, используя Cloud Dataflow (см. главу 5):

```
_ = (
    examples
    | 'get_tfrecords' >> beam.Map(lambda x: x['tfrecord'])
    | 'writetfr' >> beam.io.tfrecordio.WriteToTFRecord(
        os.path.join(options['outdir'], 'tfrecord', step)))
```

Все предыдущие образцы созданы путем извлечения записей из BigQuery:

```
tfexample = tf.train.Example(
    features=tf.train.Features(
        feature={
            'start_station_name': _bytes_feature(row['start_station_name']),
            'duration': _int64_feature(row['duration']),
        })
    )))
```

При необходимости можно также преобразовать записи с помощью `tf.transform` (https://www.tensorflow.org/tfx/transform/get_started). Затем можно использовать высокопроизводительные методы TensorFlow, предоставляемые `tf.data`. `tfrecorddataset` для чтения данных.

Экспорт в TensorFlow

Экосистема TensorFlow Serving позволяет реализовать прогнозирование с использованием моделей TensorFlow в веб-браузерах с помощью JavaScript и *tensorflow.js* на встроенных устройствах или в мобильных приложениях с помощью TensorFlow Lite, в кластерах Kubernetes с помощью KubeFlow, посредством REST API с помощью AI Platform Predictions и т. д. Поэтому вам может пригодиться возможность экспортировать модель из BigQuery ML в формате TensorFlow SavedModel. Экспортированную модель можно использовать в любой среде, поддерживающей модели TensorFlow.

Прогнозирование с использованием моделей TensorFlow

Обученную модель TensorFlow, экспортированную в формат SavedModel, можно импортировать в BigQuery и использовать для прогнозирования с помощью функции SQL `ML.PREDICT`. Эта опция может быть полезна для организации пакетного прогнозирования (например, на всех данных, собранных за последний час), учитывая, что BigQuery поддерживает возможность запланировать любой запрос SQL.

Чтобы импортировать модель в BigQuery, достаточно указать тип `tensorflow` модели в `model_type` и путь `model_path`, откуда была экспортирована модель SavedModel (обратите внимание на подстановочный знак в конце — он нужен для выбора ресурсов, словаря и т. д.):

```
CREATE OR REPLACE MODEL ch09eu.txtclass_tf
OPTIONS (model_type='tensorflow',
        model_path='gs://bucket/some/dir/1549825580/*')
```

Этот запрос создаст модель в BigQuery, которая действует подобно любой другой встроенной модели, как показано на рис. 9.11. Здесь представлена схема, указывающая, что входные данные для модели извлекаются из строкового столбца `input`.

На основании этой схемы мы можем получить прогноз:

```
SELECT
  input,
  (SELECT AS STRUCT(p, ['github', 'nytimes', 'techcrunch'][ORDINAL(s)])
   prediction)
FROM
  (SELECT p, ROW_NUMBER() OVER() AS s FROM
   (SELECT * FROM UNNEST(dense_1) AS p))
  ORDER BY p DESC LIMIT 1).*
FROM ML.PREDICT(MODEL advdata.txtclass_tf,
 (
  SELECT 'Unlikely Partnership in House Gives Lawmakers Hope for Border Deal' AS
  input
```

```
UNION ALL SELECT "Fitbit\'s newest fitness tracker is just for employees and
health insurance members"
UNION ALL SELECT "Show HN: Hello, a CLI tool for managing social media"
))
```

txtclass_tf			
Details	Training	Evaluation	Schema
Labels			
Field name	Type	Mode	Description
dense_1	FLOAT	NULLABLE	
Features			
Field name	Type	Mode	Description
input	STRING	NULLABLE	

Рис. 9.11. Схема данных для импортированной модели TensorFlow

Это очень эффективное решение, позволяющее обучить модель, сохранить ее в Google Cloud Storage, импортировать в BigQuery и периодически выполнять прогнозы без необходимости переносить нужные данные из хранилища.

Выводы

В этой главе мы кратко рассказали о машинном обучении в BigQuery. Сначала мы рассмотрели разные типы задач машинного обучения с использованием структурированных и полуструктурированных данных, а затем показали, как обучать модели в BigQuery и выполнять прогнозирование с их помощью.

Для обучения регрессионной модели в BigQuery мы создали набор данных, состоящий из признаков и меток. Затем создали обученную модель, оценили ее и использовали для прогнозов. Мы также применили ряд улучшений к базовой модели и показали, как можно извлечь веса модели. Наконец, мы показали, как обучать не только линейные модели, но и глубокие нейронные сети и деревья решений с градиентным бустингом.

Аналогично была обучена модель классификации, только для оценки использовались более сложные метрики, в частности, мы обсудили, как выбрать порог

в задаче бинарной классификации, чтобы получить желаемое значение точности или полноты.

Мы также описали различные настройки, которые могут оказаться важными для конкретных задач. Например, мы рассмотрели изменение способа распределения данных между обучающей и контрольной выборками, балансировку классов, когда один класс встречается реже другого, и регуляризацию для ослабления эффекта переобучения.

Кроме того, мы показали, как находить кластеры в структурированных данных с помощью алгоритма кластеризации методом k -средних, как визуализировать атрибуты кластеров с помощью Data Studio и как принимать решения на основании данных.

Последний тип моделей машинного обучения, который мы рассмотрели в этой главе, — системы рекомендаций. Мы создали модель матричного разложения для выработки рекомендаций по выбору товаров и для определения целевой аудитории. Мы также рассказали, как использовать пользовательские факторы и факторы товаров, являющиеся результатом разложения матрицы, для обучения более сложной модели, включающей данные о пользователях и товарах, находящихся вне матрицы рейтингов.

Наконец, мы кратко рассмотрели остальную часть экосистемы GCP, обеспечивающую поддержку нестандартных моделей, инструменты настройки гиперпараметров, AutoML и TensorFlow, обсудили взаимосвязь между разными способами конструирования моделей машинного обучения, а также немного остановились на том, когда и какие модели использовать.

Администрирование и безопасность BigQuery

Одной из привлекательных особенностей управляемых бессерверных продуктов, таких как BigQuery, является наличие инфраструктуры поддержки безопасности общедоступных облачных сервисов. В Google Cloud Platform (GCP) данные хранятся и передаются в зашифрованном виде и доступ к служебным API осуществляется только по зашифрованным каналам. Чтобы получить доступ к ресурсам BigQuery, пользователи и приложения должны пройти аутентификацию и авторизацию с использованием механизма управления идентификацией и доступом (Identity and Access Management, IAM). Для администрирования (пользователей, таблиц, заданий, представлений и т. д.) можно использовать веб-интерфейс BigQuery, инструмент командной строки bq или REST API.

В этой главе мы обсудим особенности защиты инфраструктуры BigQuery и порядок настройки Cloud IAM и других инструментов администрирования, которые можно использовать для мониторинга заданий и авторизации пользователей. В конце главы мы поговорим о поддержке различных инструментов в BigQuery, которые вы сможете использовать для удовлетворения своих требований, опираясь на прочный фундамент инфраструктуры безопасности, IAM и административных инструментов. Обязательно проконсультируйтесь у своего юриста и узнайте, насколько реализация любого из этих инструментов и механизмов соответствует вашим нормативно-правовым требованиям.

Защищенность инфраструктуры

Инфраструктура безопасности, которая лежит в основе BigQuery, является сквозной — начиная с людей и заканчивая вычислительными центрами, серверным оборудованием, стеком программного обеспечения, журналированием, шифрованием, обнаружением вторжений и, наконец, самой облачной платформой.

Команда разработчиков из подразделения информационной безопасности Google разрабатывает процессы проверки безопасности, определяет инфраструктуру защиты и внедряет процессы, обеспечивающие безопасную работу продуктов Google. В этом подразделении работают ведущие эксперты в области информационной безопасности, отвечающие за обнаружение и исправление проблем, таких как уязвимость *Heartbleed*¹ (https://www.owasp.org/index.php/Heartbleed_Bug) и *эксплойт SSL 3.0* (<https://www.us-cert.gov/ncas/alerts/TA14-290A>). Вычислительные центры Google используют многоуровневую модель физической безопасности с индивидуально разработанными средствами защиты, видеокамерами высокого разрешения, помогающими отслеживать злонамеренные действия, а также журналами доступа и обычным просмотром.

Высокая безопасность серверов обеспечивается десятками тысяч идентичных специализированных серверов. Эта однородность, наряду с организацией полного стека, включающего оборудование, сети и специализированное программное обеспечение Linux, снижает риски, связанные с цифровыми следами, и способствует быстрому реагированию на угрозы безопасности. Само серверное оборудование включает в себя специальную микросхему *Titan* для проверки встроенного и внешнего программного обеспечения, что обеспечивает надежную аппаратную идентификацию системы.

Безопасность информации о клиентах обеспечивается различными мерами и методами. Каждый слой приложения Google и стека хранения проверяет аутентификацию и авторизацию запросов, поступающих от других компонентов. Централизованная система управления группами и ролями определяет и контролирует доступ инженеров к промышленным службам и производственному окружению. К методам обеспечения безопасности относится использование протокола безопасности для аутентификации инженеров с помощью кратковременных сертификатов личных ключей, выдача которых защищена двухфакторной проверкой подлинности. Чтобы защитить информацию клиентов, все данные, которые хранятся на жестких дисках, извлекаемых из систем Google, подлежат удалению перед тем, как покинуть экосистему Google. Данные стираются с жестких дисков, которые проходят через несколько этапов контроля и проверки перед выпуском.

Управление BigQuery, как и другими службами Google, осуществляется через защищенную глобальную инфраструктуру шлюза API, доступную только по зашифрованным каналам Secure Sockets Layer (SSL)/Transport Layer Security (TLS). Каждый запрос должен включать токен аутентификации с ограниченным временем действия, генерируемый на основе логина пользователя или секретных ключей. Все запросы к API регистрируются в журнале, и администратор проекта может с помощью инструментов GCP читать журналы с операциями и попытками доступа к BigQuery.

¹ Статья в Википедии на русском языке: <https://ru.wikipedia.org/wiki/Heartbleed>. — *Примеч. пер.*

Все новые данные, записываемые на диски хранилища, шифруются в соответствии с 256-битным стандартом расширенного шифрования (Advanced Encryption Standard, AES-256), и каждый ключ сам шифруется с помощью регулярно сменяемого набора мастер-ключей. Для этого используются те же методы шифрования и управления ключами, криптографические библиотеки и корень доверия, что и во многих сервисах Google, включая Gmail. Такая общность распространяется также на сетевую инфраструктуру. Глобальная сеть Google помогает повысить безопасность передаваемых данных за счет ограничения количества переходов через общедоступный интернет. С помощью Cloud Interconnect и управляемой виртуальной частной сети (Virtual Private Network, VPN) можно создавать зашифрованные каналы между локальной частной IP-средой и сетью Google.

BigQuery использует все эти возможности. Однако это не освобождает вас от ответственности за обеспечение надлежащего доступа к данным и анализ журналов запросов. Вы также несете полную ответственность за передачу важной информации конечным пользователям, находящимся за пределами вашей корпоративной сети и общедоступной облачной инфраструктуры (то есть за предотвращение утечки данных) и обеспечение безопасности любых сведений, которые могут идентифицировать конкретное лицо, то есть личной информации. В оставшейся части этой главы мы рассмотрим инструменты, которые GCP и BigQuery предоставляют для достижения этих целей.

Управление идентификацией и доступом

Механизм управления идентификацией и доступом (Identity and Access Management, IAM) дает пользователям BigQuery возможность управлять доступом, определяя три параметра: идентичность, роль и ресурс. Фактически нужно указать, кто (идентичность) и какие привилегии (роль) имеет при обращении к каждому ресурсу.

Идентификация

Идентичность определяет пользователя, обращающегося к ресурсу. Это может быть конечный пользователь, идентифицируемый по учетной записи Google (например, по учетной записи @gmail.com или @example.com, где example.com — это домен G Suite), или приложение, идентифицируемое учетной записью службы. Учетные записи служб, по сути, являются адресами электронной почты, назначенными GCP, и могут создаваться (например, с помощью облачной консоли)¹ для определения подмножества разрешений, которыми обладает

¹ Дополнительную информацию можно найти по ссылкам <https://console.cloud.google.com/apis/credentials/serviceaccountkey> и https://cloud.google.com/iam/docs/creating-managing-service-accounts#creating_a_service_account.

создатель учетной записи службы в этом проекте. Как правило, такие учетные записи создаются для представления (ограниченного) набора разрешений, необходимых приложениям, запускаемым от нашего имени.

Доступ также может быть предоставлен виртуальным группам учетных записей Google, таким как группы Google, домены G Suite или домены Cloud Identity.¹ Лучше отдавать предпочтение группам Google, а не отдельным пользователям, потому что проще добавлять и удалять участников из группы Google, чем обновлять несколько процедур Cloud IAM для включения или удаления пользователей. Даже предоставляя доступ виртуальной группе, вы не теряете возможность аудита: при журналировании и аудите будут распознаваться действительные учетные записи Google или учетные записи служб, обращающихся к BigQuery.

Также возможно предоставить доступ `allAuthenticatedUsers` (специальный идентификатор для любого прошедшего аутентификацию с использованием учетной записи Google или учетной записи службы).² Обычно эта возможность используется для публикации общедоступного набора данных — именно так был опубликован набор данных `london_bicycles`, который мы использовали в этой книге. Обратите внимание, что `allAuthenticatedUsers` открывает доступ любому аутентифицированному пользователю, а не только пользователям в вашем домене.

Роль

Роль определяет, какие операции может выполнять конкретный пользователь. Роль состоит из набора разрешений. Можно создать роль для предоставления ограниченного доступа, определяемого настраиваемым списком разрешений. Однако на практике чаще используются предопределенные роли.

Предопределенные роли

В BigQuery есть ряд предопределенных ролей, таких как `dataViewer`, которые состоят из комбинации часто требуемых разрешений. Например, роль `dataViewer`, кроме всего прочего, предоставляет разрешение `bigquery.datasets.get`, позво-

¹ Они похожи на домены G Suite, но не имеют доступа к приложениям G Suite. Для управления пользователями, которым не нужны приложения G Suite или дополнительные возможности, такие как управление мобильными устройствами, можно создать бесплатные учетные записи Cloud Identity. Узнать больше можно по адресу <https://support.google.com/cloudidentity/answer/7319251>.

² Обратите внимание, что идентификатор `allUsers`, хотя и поддерживается в GCP, не имеет никакого эффекта в BigQuery, потому что все пользователи BigQuery должны проходить аутентификацию.

ляющее получать метаданные о наборе данных, и `bigquery.tables.getData` для получения данных из таблиц, но не `bigquery.datasets.delete`, которое позволяет любой личности с этим разрешением удалить набор данных.

На момент, когда мы писали эту книгу, было доступно восемь предопределенных ролей, в том числе четыре роли, определяющие доступ к наборам данных, а также таблицам и представлениям в них. Ниже представлены эти роли, в порядке увеличения прав:

1. `metadataViewer` (полное имя `roles/bigquery.metadataViewer`) предоставляет доступ только к метаданным наборов данных, таблиц и представлений.
2. `dataViewer` предоставляет право читать данные и метаданные.
3. `dataEditor` предоставляет право читать наборы данных, а также перечислять, создавать, изменять, читать и удалять таблицы в наборе данных.
4. `dataOwner` добавляет возможность удалить набор данных.
5. `readSessionUser` предоставляет доступ к BigQuery Storage API, оплачиваемый за счет проекта.
6. `jobUser` позволяет запускать задания (и выполнять запросы), оплачиваемые за счет проекта.
7. `user` позволяет запускать задания и создавать наборы данных, хранение которых оплачивается за счет проекта.
8. `admin` позволяет управлять всеми данными в проекте и отменять задания, запущенные другими пользователями.

Эти два набора ролей никак не зависят друг от друга. У пользователей с ролью `bigquery.readSessionUser` нет доступа к данным в таблицах — это могут быть пользователи, которым требуется читать данные из наборов, принадлежащих другим проектам! Чтобы они могли читать данные, вы должны дать им разрешение `bigquery.tables.getData`. Аналогично, роль `jobUser` не дает возможности создавать, изменять или удалять таблицы (она позволяет выполнять только запросы `SELECT`); чтобы пользователь смог выполнять инструкции языка определения данных (DDL) или языка манипулирования данными (DML), вы должны отдельно присвоить ему роль `dataEditor` кроме роли `jobUser`.

Также вполне возможно присвоить пользователю только роль `dataViewer`, без предоставления роли `user`. В этом случае пользователи сами будут оплачивать свои запросы (то есть они будут создавать задания запросов в своих проектах, но смогут запрашивать наборы данных, принадлежащие вам). Информацию о конкретных возможностях каждой роли, полном наборе разрешений и о том, какие разрешения необходимы для обращения к тем или иным методам REST API, вы найдете в документации BigQuery (<https://cloud.google.com/bigquery/docs/access-control>).

Простейшие роли

Кроме предопределенных и настраиваемых ролей BigQuery также поддерживает *простейшие роли*, появившиеся еще до того, как GCP получила поддержку Cloud IAM. Чаще всего вы будете использовать вышеупомянутые предопределенные роли, но иногда удобнее присвоить пользователям определенные роли в проекте (просмотр, правка или владение) и дать разрешение на работу со всеми наборами данных и заданиями в проекте BigQuery согласно этим ролям.

Идентифицированные личности с правом на просмотр проекта получают роль `dataViewer` для просмотра всех наборов данных в проекте, а также возможность создавать задания (то есть выполнять запросы), которые оплачиваются за счет проекта. Те, у кого есть право на изменение проекта, в дополнение к роли `dataViewer` получают роль `dataEditor`, а владельцы проектов, кроме ролей `dataViewer` и `dataEditor`, получают еще и роль `dataOwner`. Единственное исключение — только пользователь, выполнивший запрос, имеет доступ к кешированной таблице с результатами (потому что результаты могут включать соединения с наборами данных, доступ к которым может быть закрыт для владельцев других проектов). Присвоить или отозвать простейшую роль можно в консоли GCP.

Простейшие роли, предоставляющие доступ к наборам данных для чтения, записи или владения, однозначно соответствуют ролям `dataViewer`, `dataEditor` и `dataOwner` соответственно, поэтому могут присваиваться любыми способами, которыми могут назначаться предопределенные роли, но самый простой способ — щелкнуть на ссылке в веб-интерфейсе BigQuery.

Настраиваемые роли

Если по каким-то причинам предопределенные роли вас не устраивают, можно создать настраиваемую роль, но учтите, что необходимость предоставления нескольких ролей (таких, как `jobUser` и `dataViewer`) группам людей не может служить причиной для создания настраиваемой роли.



Если вам потребуется присвоить несколько ролей, чтобы дать возможность выполнять конкретную задачу, создайте группу Google, присвойте требуемые роли этой группе, а затем добавьте в эту группу пользователей или другие группы. Возможно, вы сочтете полезным создавать группы Google для решения разных задач и присваивать всем членам этих групп наборы предопределенных ролей. Например, всем членам команды, занимающейся исследованием данных, можно дать разрешения `dataViewer` и `jobUser` на доступ к хранимым наборам данных. При таком подходе, если сотрудники поменяют сферу деятельности, вам достаточно будет только изменить их членство в соответствующих группах вместо настройки их прав доступа к каждому набору данных и проекту.

Одной из причин создания настраиваемой роли является отказ от разрешений в предопределенных стандартных ролях. Например, предопределенная роль `dataEditor` позволяет владельцу создавать, изменять и удалять таблицы. Предположим, вы хотите разрешить своим поставщикам данных создавать таблицы, но у них не должно быть возможности изменять или удалять существующие таблицы. В этом случае можно создать новую роль с именем `dataSupplier` и определить для нее список соответствующих разрешений. Для этого создайте файл YAML (например, *dataSupplier.yaml*) со следующим кодом:

```
title: "Data Supplier"
description: "Can create, but not delete tables"
stage: "ALPHA"
includedPermissions:
- bigquery.datasets.get
- bigquery.tables.list
- bigquery.tables.get
- bigquery.tables.getData
- bigquery.tables.export
- bigquery.datasets.create
- bigquery.tables.create
- bigquery.tables.updateData
```

Затем выполните следующую команду `gcloud`, которая создаст требуемую роль:

```
PROJECT=$(gcloud config get-value project)
gcloud iam roles create dataSupplier --project $PROJECT \
  --file dataSupplier.yaml
```

Проверить разрешения, связанные с присвоенными ролями, можно командой

```
gcloud iam roles describe dataSupplier --project $PROJECT
```

Ее также можно применить к предопределенной роли.

Также удобно создавать роли в первоначальном состоянии (**stage**) **ALPHA**, проверить их на небольшом количестве пользователей и только потом переводить в состояние **BETA** или **GA**. Так можно точно настроить набор разрешений (начав с наиболее ограниченного), прежде чем применять его в более широком масштабе.

Ресурсы

Доступ регулируется отдельно для каждого конкретного ресурса. Роль `dataViewer` или разрешение `bigquery.tables.getData` не дает права доступа ко всем ресурсам в BigQuery; только к определенным наборам данных или таблицам.

Поскольку роль **dataViewer** предоставляет доступ только к таблицам или наборам данных, ее обладатель не сможет получить информацию о заданиях; задания — это отдельный ресурс и для доступа к ним необходим другой набор разрешений. Конечно, личности могут быть присвоены обе роли, **dataViewer** и **jobUser**, что даст ей возможность не только читать данные из таблиц, но и создавать задания (и выполнять запросы) и отменять свои задания.

Старайтесь избегать чрезмерного расширения разрешений/ролей; лучше перестраховаться и предоставить минимально необходимый объем привилегий, включая ограниченный набор ролей и круг доступных ресурсов. Не забывайте о необходимости обновлять разрешения для новых ресурсов по мере их создания. Вполне разумно будет установить доверительные границы, отображающие проекты в вашу организационную структуру, определить роли на уровне проекта — в этом случае политики IAM могут распространяться вниз от проектов к ресурсам внутри проекта и автоматически применяться к новым наборам данных в проекте.

Администрирование BigQuery

Администрировать BigQuery можно с помощью веб-интерфейса, REST API или инструмента командной строки **bq**. В этом разделе мы будем исходить из предположения, что вы являетесь, например, ответственным специалистом по BigQuery в компании или администратором проекта с соответствующими ресурсами (заданиями или наборами данных) и вам присвоена роль **admin** в BigQuery. Здесь мы рассмотрим наиболее типичные задачи, которые решают администраторы, сосредоточив внимание на применении инструмента командной строки **bq**.

Управление заданиями

Когда задание отправляется в BigQuery, оно последовательно проходит три стадии: **PENDING** — запланировано, но еще не запущено; **RUNNING** — запущено; и либо **SUCCESS**, либо **FAILURE**, в зависимости от результата выполнения.

Ниже показано, как можно получить список заданий, созданных в рамках проекта в течение последних 24 часов:

```
NOW=$(date +%s)
START_TIME=$(echo "($NOW - 24*60*60)*1000" | bc)
bq --location=US ls -j -all --min_creation_time $START_TIME
```

bq требует, чтобы отметка времени была выражена в миллисекундах, поэтому мы получаем значение для параметра **min_creation_time**, вычитая один день (24*60*60 секунд) из текущего времени и преобразуя секунды в миллисекунды с помощью калькулятора командной строки **bc**.

Если известен идентификатор выполняющегося задания, его можно отменить:¹

```
bq --location=US cancel bqjob_180ae24c_16b04a8d28d
```

Обратите внимание, что в некоторых случаях известен полный идентификатор задания, включающий не только имя проекта, но и местоположение,² например, из записей в журналах (или в веб-интерфейсе BigQuery). В этом случае параметр местоположения `location` можно опустить:

```
bq cancel someproject:US.bqjob_180ae24c_16b04a8d28d
```

Любой, кому присвоена роль `jobUser` или `user`, сможет запускать и отменять свои задания; ему не нужна роль `admin`, если он не собирается просматривать или отменять задания, запущенные другими пользователями.

Авторизация пользователей

Мы рекомендуем создавать группы Google и добавлять в них пользователей, вместо того чтобы предоставлять им индивидуальный доступ к ресурсам. Следуя этой рекомендации, можно организовать доступ к ресурсам, используя только группы Google. Есть несколько удобных способов добавления и удаления нескольких пользователей в группу Google. Подробнее об этом рассказывается на странице со справкой о G Suite (<https://support.google.com/a/answer/6191469?hl=ru>).

Чтобы предоставить однократный доступ к BigQuery отдельным пользователям, учетным записям служб или группам Google, используйте страницу Cloud Console IAM (<https://console.cloud.google.com/iam-admin/iam>). Чтобы поделиться определенными ресурсами, в веб-интерфейсе BigQuery выберите набор данных, а затем кликните на `Share dataset` (Поделиться набором данных).

Восстановление удаленных записей и таблиц

Если пользователь повредил содержимое таблицы, загрузив повторяющиеся данные или удалив важные записи, вы сможете исправить проблему в течение семи дней. Удаленные таблицы (в отличие от удаленных записей в существующих таблицах) можно восстановить, только если прошло не больше двух дней.

Например, чтобы восстановить состояние таблицы, в котором она была 24 часа назад, можно использовать функцию `SYSTEM_TIME AS OF` и инструкции DDL:³

¹ Отмененные задания тоже оплачиваются.

² Напомним, что данные хранятся в BigQuery в определенном регионе или объединении регионов. Запросы должны выполняться там, где расположены данные, и метаданные заданий также хранятся с учетом региона.

³ Прежде чем испытать этот пример, создайте выходной набор данных `ch10eu` в местоположении EU.

```
CREATE OR REPLACE TABLE ch10eu.restored_cycle_stations AS
SELECT
  *
FROM `bigquery-public-data`.london_bicycles.cycle_stations
FOR SYSTEM_TIME AS OF
  TIMESTAMP_SUB(CURRENT_TIMESTAMP(), INTERVAL 24 HOUR)
```

Также в течение двух ближайших дней можно восстановить удаленную таблицу. Например, давайте удалим таблицу, созданную минуту назад:

```
bq rm ch10eu.restored_cycle_stations
```

А теперь восстановим ее в состоянии, в котором она была 120 секунд назад:

```
NOW=$(date +%s)
SNAPSHOT=$(echo "($NOW - 120)*1000" | bc)
bq --location=EU cp \
  ch10eu.restored_cycle_stations@$SNAPSHOT \
  ch10eu.restored_table
```



Восстановить удаленную таблицу можно, только если за это время в наборе данных не была создана другая таблица с тем же идентификатором.¹ Это значит, что нельзя восстановить удаленную таблицу, если она получает потоковые данные и требуется создать таблицу с помощью `create-disposition`, если она не существует. В этом случае потоковый конвейер, скорее всего, создаст пустую таблицу и начнет вставлять в нее новые записи. Вот почему следует с осторожностью использовать `CREATE OR REPLACE TABLE`: это делает таблицу невозможной.

Непрерывная интеграция/непрерывное развертывание

SQL-запросы желательно хранить в системе управления версиями, чтобы иметь возможность получить версию скрипта по состоянию на определенный момент времени и видеть, как он изменялся. Для выполнения таких запросов можно использовать Cloud Source Repositories и Cloud Functions (или Cloud Run, если у вас есть более сложные зависимости).

Вызов BigQuery из Cloud Function

В облачном репозитории Cloud Source Repository создайте файл `.sql` с текстом SQL-запроса для BigQuery и файл на Python с кодом функции для Cloud

¹ При этом вмещающий набор данных не должен удаляться или создаваться заново. См. <https://cloud.google.com/bigquery/docs/managing-tables#undeletetable>.

Function. После этого Cloud Function сможет использовать клиентскую библиотеку BigQuery для отправки запроса в BigQuery и экспорта результатов в Google Cloud Storage, если время выполнения запроса будет меньше времени тайм-аута облачной функции.¹

Создать облачную функцию для Cloud Function можно в консоли GCP Cloud Console. В текстовое поле введите следующий код:

```
from google.cloud import bigquery
def query_to_gcs():
    client = bigquery.Client()

    # Запустить запрос и ждать его завершения
    query_job = client.query("""
        ...
    """)
    query_job.result()

# Запустить извлечение результатов в GCS и ждать завершения процедуры
extract_job = client.extract_table(
    query_job.destination, "gs://bucket/file.csv")
extract_job.result()
```

Теперь с помощью Cloud Scheduler (<https://cloud.google.com/scheduler/>) вместо запроса можно запланировать запуск облачной функции.



Обратите внимание, что предыдущий код вызывает функцию `extract_table` и передает ей временную таблицу, созданную в результате выполнения запроса. Это — быстрый способ экспортировать результат запроса в файл в формате CSV.

Сохранение запросов создания таблиц, представлений и функций в системе управления версиями

Возможность управления версиями и повторяемость важны не только для запросов, извлекающих данные, но и для создающих таблицы, представления, модели, хранимые процедуры и функции. Поэтому желательно поместить весь код создания в скрипт, который затем можно будет вызывать, когда понадобится воссоздать нужную таблицу, представление, модель или функцию.

Создать таблицу из результата запроса можно с помощью клиентской библиотеки BigQuery, установив нужную таблицу целью задания:

¹ Величину тайм-аута для облачных функций можно настраивать, но на момент написания этой книги максимальное значение составляло не более девяти минут.

```

from google.cloud import bigquery
client = bigquery.Client()
sql = """
WITH stations AS (
    SELECT [300, 314, 287] AS closed
)
SELECT
    station_id
    , (SELECT name FROM `bigquery-public-data`.london_bicycles.cycle_stations
        WHERE id=station_id) AS name
FROM
    stations, UNNEST(closed) AS station_id
"""

job_config = bigquery.QueryJobConfig()
job_config.destination = (
    client.dataset('ch10eu').table('stations_under_construction'))
query_job = client.query(sql, location='EU', job_config=job_config)
query_job.result() # Ждать завершения запроса

```

Вот эквивалентный запрос на языке DDL:

```

CREATE OR REPLACE TABLE -- или TABLE/MODEL/FUNCTION
ch10eu.stations_under_construction
(
    station_id INT64 OPTIONS(description = 'Station ID'),
    name string OPTIONS(description = 'Official station name')
)
OPTIONS(
    description = 'Stations in London.',
    labels=[("pii", "none")] -- обязательно в нижнем регистре.
)
AS

WITH stations AS (
    SELECT [300, 314, 287] AS closed
)
SELECT
    station_id
    , (SELECT name FROM `bigquery-public-data`.london_bicycles.cycle_stations
        WHERE
id=station_id) AS name
FROM
    stations, UNNEST(closed) AS station_id

```

Обратите внимание, что описания таблиц и столбцов в этом примере находятся непосредственно в инструкциях CREATE и сохраняются в репозитории. Если у вас уже есть таблицы BigQuery, вы можете запросить список таблиц и столбцов из ресурса метаданных и программно создать операторы SQL DDL с помощью функций (https://cloud.google.com/bigquery/docs/information-schema-tables#advanced_example).



Будьте внимательны, планируя запускать по расписанию облачные функции, которые создают или заменяют таблицы: если у вас есть облачная функция, которая выполняет по расписанию или повторно запускает предыдущий оператор для ЗАМЕНЫ таблицы (обратите внимание на REPLACE), любые изменения в этой таблице (включая обновленные записи и описания схем) будут затерты.

Экспорт биллинга — получение информации о расходах

Вы можете получить посуточный график использования служб GCP, а также расчеты стоимости посуточного обслуживания набора данных в BigQuery. Но имейте в виду, что получение этой информации тоже будет включено в счет и за нее нужно будет заплатить!

Чтобы включить экспорт биллинга, в облачной консоли GCP Cloud Console откройте раздел **Billing** (Оплата), выберите свою платежную учетную запись и проект, выберите набор данных BigQuery,¹ куда следует экспортировать биллинг, а затем включите экспорт.

Данные биллинга загружаются в BigQuery через равные промежутки времени, поэтому может пройти несколько часов, прежде чем появится первая информация. Частота обновления в BigQuery зависит от используемых служб GCP.

Как и любую другую таблицу в BigQuery, вы можете изучить схему экспортированных данных биллинга, чтобы выяснить, какие запросы можно выполнять и какие данные имеются для мониторинга. Далее перечисляются некоторые интересные возможности, с которых можно начать.

Месячная стоимость продукта

Чтобы узнать ежемесячную сумму, в которую обходится продукт, используйте следующий запрос:

```
SELECT
  invoice.month
  , product
  , ROUND(SUM(cost)
    + SUM(IFNULL((SELECT SUM(c.amount) FROM UNNEST(credits) c),
      0))
    , 2) AS monthly_cost
FROM ch10eu.gcp_billing_export_v1_XXXXXX_XXXXXX_XXXXXX
GROUP BY 1, 2
ORDER BY 1 ASC, 2 ASC
```

¹ Создайте набор данных, если необходимо.

Ежемесячная стоимость — это сумма затрат на продукт, скорректированная на сумму долга.

Визуализация данных биллинга

Как показано на рис. 10.1, для визуализации данных биллинга в Data Studio доступна начальная панель мониторинга.

Подробнее о том, как создать ее копию и начать работать в Data Studio, можно узнать по ссылке <https://cloud.google.com/billing/docs/how-to/visualize-data>.

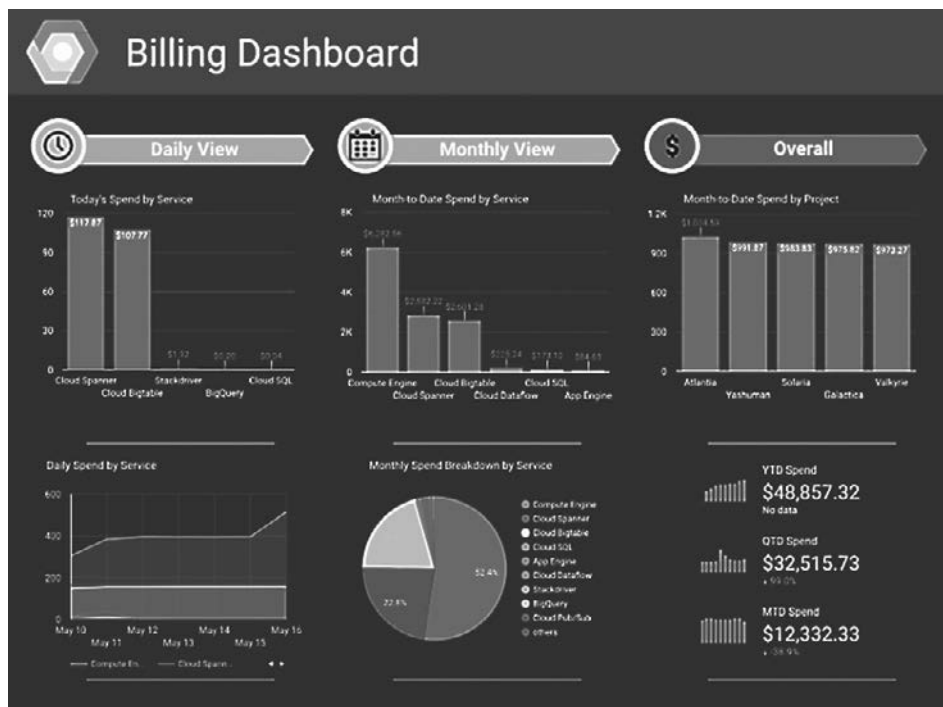


Рис. 10.1. Пример панели мониторинга в Data Studio, отображающей данные биллинга, экспортированные в BigQuery

Метки

Конечно, интересно иметь информацию о затратах по продуктам, но часто намного интереснее узнать о затратах по подразделениям в вашей организации. Чтобы получить такой уровень детализации в отчетах, необходимо добавить метки (пары ключ/значение) к вашим ресурсам GCP. В этом случае каждая запись в экспорте биллинга будет содержать в столбцах `label.key` и `label.value`

значения, соответствующие метке ресурса GCP, за использование которого начислена плата.

Если метки присваивать по названиям подразделений, ключом может служить `team`, а значением — `marketing` или `research`. Метки также можно определять по окружению (например, `key=environment, value=production` или `value=test`), приложению или компоненту.

Также можно присваивать метки ресурсам GCP, таким как виртуальные машины Compute Engine, кластеры Dataproc или задания Dataflow. Разумеется, метки можно присваивать наборам данных BigQuery, таблицам, моделям и даже запросам.

Например, вот как можно применить метку «`environment:learning`» к набору данных `ch10eu`:

```
bq update --set_label environment:learning ch10eu
```

Присваивание меток таблицам и представлениям осуществляется аналогично, но (на момент написания книги) метки таблиц/представлений не отображаются в данных биллинга:

```
bq update --set_label environment:learning ch10eu.restored_table
```

Кроме того, можно присвоить метку заданию перед его отправкой из командной строки:

```
bq query --label environment:learning --nouse_legacy_sql 'SELECT 17'
```

Чтобы добавить метку при отправке задания через REST API, нужно заполнить свойство `labels` ресурса задания.

В этом случае данные биллинга будут отражать затраты на запросы, и, агрегируя затраты по меткам, вы сможете детализировать расходы по окружениям, подразделениям или любым другим ключам меток:

```
SELECT
  invoice.month
  , label.value
  , ROUND(SUM(cost)
    + SUM(IFNULL((SELECT SUM(c.amount) FROM UNNEST(credits) c),
      0))
    , 2) AS monthly_cost
FROM
  ch10eu.gcp_billing_export_v1_XXXXXX_XXXXXX_XXXXXX
  , UNNEST(labels) AS label
WHERE
  label.key = 'environment'
GROUP BY 1, 2
ORDER BY 1 ASC, 2 ASC
```

Оперативные панели, мониторинг и журналы аудита

Ключевым аспектом безопасности является возможность проверки эффективности мер защиты. Доступность всех развернутых ресурсов для наблюдения очень важна.

Cloud Security Command Center

Командный центр безопасности (Cloud Security Command Center, SCC) предлагает комплексную платформу управления безопасностью для GCP. Предлагая информацию об имеющихся активах и состоянии безопасности, Cloud SCC упрощает обнаружение, предотвращение и реагирование на угрозы. Набор встроенных детекторов угроз предупредит вас о подозрительной активности.

Со страницы GCP Cloud Console Security Command Center Marketplace можно получить доступ к Cloud SCC и запустить поиск активов. После сканирования проектов вы сможете использовать оперативную панель Cloud SCC для поиска распространенных проблем, таких как открытый порт 22 (Secure Shell [SSH]). После этого поиск активов будет производиться не реже одного раза в день.

Мониторинг и анализ журналов аудита с помощью Stackdriver

Для мониторинга ресурсов BigQuery можно использовать инструмент Stackdriver. К числу его возможностей относятся: визуализация таких показателей, как общее время, потраченное на выполнение запросов, количество доступных слотов, и многих других. BigQuery также автоматически отправляет журналы аудита в Stackdriver Logging. Stackdriver Logging позволяет пользователям фильтровать и экспортировать сообщения в другие службы, включая Cloud Pub/Sub, Cloud Storage и BigQuery.

Кроме долговременного хранения журналов, имеется также возможность экспорта журналов в BigQuery, благодаря которой можно реализовать обзорный анализ журналов. Вот, например, запрос, оценивающий затраты (до применения скидок) по идентификатору пользователя (<https://cloud.google.com/bigquery/docs/reference/auditlogs/>):

```
WITH data as
(
  SELECT
    protopayload_auditlog.authenticationInfo.principalEmail as principalEmail,
    protopayload_auditlog.servicedata_v1_bigquery.jobCompletedEvent AS
jobCompletedEvent
  FROM
    ch10.cloudaudit_googleapis_com_data_access_2019*
)
```

```

SELECT
  principalEmail,
  SUM(jobCompletedEvent.job.jobStatistics.totalBilledBytes)/POWER(2, 40)) AS
Estimated_USD_Cost
FROM
  data
WHERE
  jobCompletedEvent.eventName = 'query_job_completed'
GROUP BY principalEmail
ORDER BY Estimated_USD_Cost DESC

```

Информация о журналах аудита BigQuery доступна в `protoPayload.metadata` в объекте `LogEntry`. Они организованы в три потока: действия администратора, системные события и доступ к данным. Журнал с действиями администратора включает такие события, как запуск и завершение заданий. Системные события — это такие события, как `TableDeletion`, возникающие по истечении срока действия таблицы или раздела. Поток обращений к данным содержит информацию о новых заданиях, заданиях, изменивших состояние, изменениях в табличных данных и операциях чтения данных из таблиц.

Доступность, восстановление после отказа и шифрование

Архитектура BigQuery обеспечивает высочайшую надежность. Например, благодаря бессерверному характеру службы, выход из строя практически любого аппаратного компонента не повлияет на способность BigQuery выполнять запросы. В отличие от многих систем, которые привязаны к конкретным виртуальным машинам или узлам, BigQuery действует в гигантском общем пуле серверов и может почти мгновенно перенаправлять трафик из одного места в другое.

Зоны, регионы и объединения регионов

В GCP есть три типа местоположений. *Зоны* — это вычислительные кластеры, обычно расположенные в одном здании. Зоны имеют очень высокую доступность, но в случае серьезного сбоя оборудования (например, из-за пожара или выхода из строя трансформатора) зона может отключиться. Иногда в зонах возникают проблемы, которые не связаны с оборудованием. Высокая частота запросов может привести к сбою некоторых служб и вызвать проблемы с зависимыми службами. Отказоустойчивые сервисы, подобные BigQuery, обладают всем необходимым для преодоления любых проблем, связанных с зонами.

Регионы, в свою очередь, представляют более широкие области, охватывающие несколько независимых друг от друга зон. Обычно регионы распределены по

нескольким зданиям в большом комплексе. В целом очень редко случаются ситуации, когда весь регион отключается от сети. Однако такое возможно из-за стихийных бедствий. В преддверии предсказуемых стихийных бедствий, таких как ураганы, во избежание потери любых данных возможно штатное отключение региона. Непредсказуемые стихийные бедствия, такие как землетрясения, могут привести к потере данных службами в регионе, а также к невозможности повторного запуска региона после катастрофы.

Объединения регионов обладают наибольшей устойчивостью; они обычно предполагают некоторую гибкость в выборе местоположения, будучи распределенными по нескольким вычислительным центрам, удаленным друг от друга на сотни километров. Например, объединение EU (Европейский союз) состоит только из физических вычислительных центров, расположенных на территории Европейского союза. Какие конкретные вычислительные центры относятся к их числу? Относится ли вычислительный центр во Франкфурте к объединению EU? А вычислительный центр в Финляндии? Некоторые службы, такие как Google Cloud Storage, четко указывают, какие регионы входят в состав объединения. Другие, такие как BigQuery, определяют менее четкие критерии размещения данных в объединении регионов, чтобы сохранить гибкость как в вычислениях, так и в размещении данных.

BigQuery и обработка отказов

Один из способов предсказать, насколько хорошо служба противостоит отказам определенного типа, — подсчитать их частоту; если отказы случаются очень редко, это может говорить о том, что служба не проверяется и может иметь ошибки или другие проблемы. Службы Google разрабатываются так, чтобы справляться практически с любыми типами аппаратных и даже программных отказов и продолжать работать. Они не только проектируются таким образом, но и проходят строгие испытания, чтобы убедиться, что отказы обрабатываются регулярно. Давайте обсудим некоторые типы отказов и то, как на них реагирует BigQuery.

Отказы дисков

Вращающиеся диски имеют движущиеся детали и, подобно многим механизмам с подобными деталями, довольно часто выходят из строя. Благодаря алгоритму стирающего кодирования (erasure encoding) для кодирования данных в Colossus (инфраструктура хранения, лежащая в основе BigQuery; см. главу 6), даже выход из строя большого числа дисков может не приводить к потере данных. При наличии 100 000 дисков можно ожидать, что десятки из них будут выходить из строя каждый день. Кроме того, в процессе текущего профилактического обслуживания число отключаемых дисков может исчисляться сотнями. В случае отключения электроэнергии есть риск потерять намного больше дисков при по-

вторном запуске. Даже с учетом всех этих факторов вероятность потери данных из-за аппаратного сбоя в BigQuery близка к нулю.¹

Когда диск выходит из строя, Colossus обнаруживает это и реплицирует данные на другой диск. Сотрудники, обслуживающие оборудование вычислительного центра Google, извлекают диск и уничтожат его. Программное обеспечение и службы, использующие этот диск, даже не заметят сбоя, разве что некоторые запросы будут выполняться на несколько миллисекунд дольше.

Отказы серверов

Несмотря на многочисленные усилия поддерживать работоспособность серверов, они тоже иногда выходят из строя. В операционных системах встречаются ошибки, космическое излучение повреждает память, перегорают процессоры и блоки питания, возникают утечки памяти в программном обеспечении и многое другое. Для противостояния таким сбоям в мире серверного программного обеспечения используется общий подход, основанный на избыточности и усилении защиты. Дорогие серверы имеют резервные источники питания, допускают замену памяти и процессоров непосредственно во время работы, чтобы они могли работать безостановочно.

В Google используется иная философия и иной подход. Вычислительные центры Google организованы так, что готовы к выходу из строя любой машины и в любое время. Программное обеспечение должно быть готово к подобным событиям, и поэтому в Google было создано несколько распределенных систем с горизонтальным масштабированием, таких как MapReduce, Google File System (GFS, предшественница Colossus) и Dremel.

Серверы BigQuery постоянно выходят из строя. Когда у вас сотни тысяч или даже миллионы независимых серверов, подобное будет случаться постоянно. Практически любой сервер BigQuery может выйти из строя в любой момент, но пользователи ничего не заметят, кроме небольшого замедления своих запросов. Программное обеспечение управления кластерами Borg, на котором работают вычислительные центры Google, может автоматически перезапустить задачи на других машинах, если какая-то машина перестанет откликаться на запросы проверки работоспособности даже на несколько секунд. Программное обеспечение движка запросов даже не замечает этих проблем; оно просто повторяет задачу и продолжит работу.

Даже более крупные проблемы, такие как выход из строя целой стойки или сетевого коммутатора, обрабатываются просто и понятно. Благодаря большому

¹ Конечно, все эти уважительные причины, ошибки в коде, стихийные бедствия и т. д. могут привести к потере данных. Но вероятность такой потери очень мала. Если у вас есть 1 Пбайт данных и каждый день в одной зоне из строя выходят 1000 дисков, то вероятность потерять данные в течение миллиона лет будет ниже 0.01%.

масштабу кластеров в вычислительных центрах, если стойка или коммутатор выйдут из строя, это повлияет только на некоторые выполняемые сервисом задачи, поэтому он сможет обойти проблему. Никто не заметит неполадок, кроме обслуживающего персонала в вычислительном центре, который должен будет ее устранить. И даже персонал, отвечающий за работу сервиса, может спать спокойно.

Зональные отказы

Но что произойдет при значительном сбое, после которого встроенные механизмы автоматического самовосстановления не помогут? Для ясности сразу отметим, что подобное происходит довольно редко. При проектировании зон предусматривается максимально высокая устойчивость к аппаратным и сетевым сбоям. Для всего, что может привести к отключению зоны, например сетевых коммутаторов или трансформаторов, обычно есть дополнительное резервное оборудование. Но при проведении земляных работ с экскаватором бывали случаи обрыва оптоволоконного кабеля, что приводило к пожарам на трансформаторах и отключению всей зоны.

Некоторые службы, такие как Google Compute Engine, на момент написания этой книги привязаны к одной зоне доступности. Если зона отключается, все экземпляры виртуальных машин в этой зоне становятся недоступными. BigQuery, напротив, предусматривает обработку практически всех зональных отказов. Все облачные проекты в BigQuery имеют первичное и вторичное местоположение. Если в первичном местоположении происходит сбой, BigQuery переключается на вторичное местоположение.

Зональные сбои делятся на две категории: *частичные* (soft) и *полные* (hard). Под частичными сбоями подразумеваются проблемы, которые не нарушают общей работоспособности, но могут снизить пропускную способность. Частичные отказы часто являются результатом программного сбоя, а не отказа оборудования. Например, может остановиться сервер квот, зависнуть Bigtable или планировщик BigQuery будет тратить слишком много времени для планирования. Под полным отказом подразумевается отключение зоны. Полный отказ может быть обусловлен нарушением снабжения вычислительного центра электроэнергией или появлением какой-то неисправимой аппаратной проблемы.

В ответ на частичные сбои BigQuery активно *осушает* зону, то есть новые запросы отправляются куда-то еще, но при этом обработка уже запущенных запросов не прекращается. Новые запросы направляются во вторичную зону. Выполнение уже запущенных запросов может продолжаться в этой зоне, но если проблема достаточно серьезная, они будут перезапущены во вторичной зоне.

Частичные сбои вовсе не редкость. BigQuery действует в десятках зон доступности, разбросанных по всему миру, и вероятность того, что какая-то служба столкнется с проблемами в какой-то зоне, довольно высока. Однако благодаря

механизмам отказоустойчивости в BigQuery пользователи практически никогда не замечают никаких сбоев.

Полные сбои случаются гораздо реже; обычно они означают серьезные проблемы для всей зоны. В случае полного зонального сбоя пользователи могут заметить нарушение работы по таким признакам, как прерывание запущенных запросов и их повторный запуск в новой зоне, из-за чего на их выполнение может потребоваться вдвое больше времени. При особенно серьезных сбоях последние данные могут оказаться нереплицированными во вторичную зону и могут оставаться недоступными до восстановления работоспособности зоны. В таких случаях запросы к затронутым таблицам потерпят неудачу. BigQuery предпочитает прервать выполнение запроса с сообщением об ошибке, чем вернуть противоречивые данные.

Региональные отказы

На следующем уровне в классификации сбоев находятся сбои, при которых отключаются целые регионы. Вероятность отключения региона еще ниже, чем вероятность отключения зоны. Как и в зонах, в регионах могут случаться частичные отказы, при появлении которых можно аккуратно отключить весь регион. Регионы, находящиеся на пути урагана, можно полностью отключить до того, как ураган на них обрушится. Выполнив отключение до урагана, Google может минимизировать вероятность потери данных при перезапуске региона или в редких случаях, когда он окажется поврежден. Перебои в подаче электроэнергии не должны приводить к отключению региона, потому что имеются местные резервные источники питания, но если перебои в подаче электроэнергии продолжаются слишком долго, емкость резервных источников может быть исчерпана и тогда будет инициировано упорядоченное отключение региона.

В регионах также могут происходить полные отказы, но они встречаются еще реже. На практике трудно оценить, насколько часто происходят катастрофические события, которые называют «черными лебедями».¹ Катастрофическое землетрясение может совершенно неожиданно вызвать отключение региона. Так же неожиданно могут происходить другие экстремальные погодные явления или стихийные бедствия. Полный отказ региона может быть вызван повреждением оборудования и, соответственно, сопровождаться потерей данных, не реплицированных за границы региона.

На момент написания этой книги местоположения BigQuery, охватывающие один регион (такие, как `asia-east1` или `europa-north1`), не предполагали воз-

¹ Неожиданные и очень редкие крупномасштабные события с катастрофическими последствиями; свое название «черный лебедь» они получили в теории, выдвинутой Нассимом Николасом Талебом (Nassim Nicholas Taleb) в книге «The Black Swan: The Impact of the Highly Improbable» (Random House). (Талеб Нассим Николас. Черный лебедь. Под знаком непредсказуемости. КоЛибри, 2019. — Примеч. пер.)

возможности хранения данных за их пределами, потому что невозможно сохранить резервную копию в другом месте, не нарушив требований клиентов к местонахождению их данных. Например, Сингапур — это остров длиной всего 42 километра; если клиент требует, чтобы его данные не покидали пределов Сингапура, согласно местному законодательству, эти данные нельзя будет сохранить где-то еще. Однако прежде чем делать какие-либо выводы относительно сохранности ваших данных, обязательно загляните в документацию, актуальность которой постоянно поддерживается компанией Google.

Местоположения, охватывающие несколько регионов, такие как US и EU, хранят резервные копии в нескольких регионах. В такой ситуации катастрофические отказы регионов не будут угрожать целостности данных. Но может пройти некоторое время, прежде чем эта резервная копия станет доступной.

Сохранность, резервное копирование и восстановление после аварий

В общих чертах механизм репликации в BigQuery выглядит так:

- Для данных, которые можно хранить в нескольких регионах, создаются копии как минимум в двух регионах (данные, которые можно хранить только в одном регионе, в другие регионы не копируются).
- Все данные копируются в две зоны доступности.
- Внутри зоны данные кодируются с использованием алгоритма стирающего кодирования.

У внешних резервных копий также есть защита в виде вторичных механизмов предотвращения случайного удаления; для этого внутренняя прошивка дисков изменяется так, чтобы сделать невозможным удаление до истечения определенного периода. То есть если в BigQuery попадет программная ошибка, вызывающая преждевременное удаление данных, низкоуровневые системы встроенного ПО на дисках не допустят этого.

Если клиент случайно удалит данные, он сможет воспользоваться функцией поддержки путешествий во времени в BigQuery. Пользователь сможет запросить таблицу в том виде, в каком она была до удаления данных, используя синтаксис `SYSTEM_TIME AS OF` (см. главу 8). Также есть возможность скопировать таблицу, какой она была в определенный момент времени, используя `tablename@timestamp` в задании копирования.

Этот прием копирования устаревших данных может пригодиться для восстановления таблицы. Чтобы выполнить такое восстановление, скопируйте данные из таблицы со старым именем, имевшимся до удаления, в таблицу с новым именем. Имейте в виду, что если удалить таблицу, а затем создать новую с тем

же именем, данные в старой таблице будут потеряны навсегда, поэтому будьте осторожны. На момент написания этой книги восстановить удаленную таблицу можно было только в течение 48 часов после удаления, что меньше обычного семидневного периода.

Конфиденциальность и шифрование

В Google очень серьезно относятся к безопасности и конфиденциальности. Все данные в BigQuery хранятся и передаются только в зашифрованном виде. Шифрование при записи на диск осуществляется прозрачно, с использованием механизма шифрования файлов в файловой системе Colossus. Поточковые данные шифруются в Bigtable и в файлах журналов. Метаданные шифруются в Spanner. Сетевой трафик прозрачно шифруется с помощью внутренних протоколов вызова удаленных процедур (Remote Procedure Call, RPC). Даже если кто-то и имеет физический доступ к дискам или к сетевым подключениям, то он не сможет получить данные в открытом виде.

Прозрачность доступа

Google предпринимает множество мер для защиты доступа к данным. Доступ к данным пользователей могут получить только небольшое число инженеров, обеспечивающих безопасность и надежную работу системы. Прозрачность доступа в GCP предполагает, что всякий раз, когда кто-то в Google получает доступ к вашим данным, вам отправляется уведомление через записи в журнале аудита. Как правило, все очень просто; если кто-то прочитает ваши данные, вы сможете об этом узнать.

Управление службами в виртуальном частном облаке

Управление службами в виртуальном частном облаке (Virtual Private Cloud Service Controls, VPC-SC) — за этим громким названием скрывается механизм, обеспечивающий полный контроль над доступом к службам и передачей данных в GCP. Например, доступ к BigQuery можно ограничить узким диапазоном IP-адресов, находящихся в сети вашей компании. Также можно обеспечить контроль над передачей данных между службами и предотвратить экспорт данных из BigQuery в Google Cloud Storage. Как вариант, можно разрешить экспорт в Google Cloud Storage, но только в корзины Cloud Storage, принадлежащие вашей организации.

Механизм VPC-SC используется не только в BigQuery, но и во многих продуктах GCP. Это позволяет определить одну общую политику, описывающую методы экспортирования и перемещения данных. С его помощью вы можете вообще запретить доступ к BigQuery (хотя это кажется странным). За дополнительной

информацией о VPC обращайтесь к документации по Google Cloud (<https://cloud.google.com/vpc/docs/>).

Ключи шифрования, управляемые клиентом

Все данные хранятся в BigQuery в зашифрованном виде, но как организовать шифрование данных с использованием ваших собственных ключей? Для этого можно использовать механизм поддержки ключей шифрования, управляемых клиентом (Customer-Managed Encryption Keys, СМЕК). Добавить свои ключи можно в Cloud KMS — центральной службе управления ключами GCP, а затем определить, какие наборы данных или таблицы должны шифроваться с помощью этих ключей.

BigQuery использует несколько слоев для обертывания ключей, то есть мастер-ключи не видны за пределами KMS. Каждая таблица, защищенная с помощью СМЕК, хранит в метаданных обернутый ключ. Когда BigQuery обращается к таблице, она посылает запрос в Cloud KMS, чтобы развернуть ключ. Затем развернутый ключ используется для развертывания отдельных ключей для каждого файла. Этот протокол обертывания ключей имеет ряд преимуществ, которые снижают риск утечки развернутых ключей. Получив развернутый ключ для файла, вы не сможете читать другие файлы. Получив развернутый ключ для таблицы, вы сможете развернуть ключи файлов, только пройдя проверку контроля доступа. И Cloud KMS никогда не развертывает мастер-ключ. Если удалить ключ из KMS, вы никогда не сможете развернуть другие ключи (поэтому будьте осторожны со своими ключами!).

Соответствие требованиям законодательств

Многие организации руководствуются теми или иными положениями национального законодательства, и ваша организация, вероятно, тоже устанавливает требования к программному обеспечению и выбору местоположения хранимых данных, следуя законодательным актам. В этом разделе мы посмотрим, как BigQuery может помочь вам обеспечить поддержку такого соответствия нормативным требованиям. Но имейте в виду: вы обязательно должны проконсультироваться с юристом, чтобы определить, насколько реализация того или иного инструмента соответствует вашим нормативным требованиям.

Местоположение данных

Многие правительства во всем мире выдвигают определенные требования к местоположению хранимых данных, и база BigQuery в полной мере им соответствует и готова обеспечить возможность выполнения запросов к любым наборам данных только в вычислительном центре, где этот набор данных

хранится. Ограничения на местоположение данных накладываются во время создания набора данных. Например, вот как создать набор данных в регионе `asia-east2` (Гонконг):

```
bq --location=asia-east2 mk --dataset ch10hk
```

BigQuery поддерживает местоположения двух типов: региональные и многорегиональные (объединения регионов). Примером регионального местоположения может служить регион `asia-east2`, соответствующий Гонконгу. Он представляет конкретную географическую область. Другой тип местоположения — многорегиональный (например, `US` или `EU`¹) — включает два или более региональных местоположения. Актуальный список поддерживаемых местоположений можно найти в документации BigQuery (<https://cloud.google.com/bigquery/docs/locations>).

Как мы уже говорили в главе 6, BigQuery определяет местоположение для запуска задания, опираясь на параметры по умолчанию проекта, резервирования и наборов данных, указанных в запросе. Также можно явно указать местоположение, где должно выполняться задание, — в веб-интерфейсе BigQuery (указав местоположение в настройках запроса), при обращении к REST API (определив свойство `location` в разделе `jobReference`) или при использовании инструмента командной строки `bq` (параметр `--location`). Если запрос не может быть выполнен в указанном местоположении (например, если выбрано местоположение `US`, а данные находятся в `EU`), BigQuery вернет сообщение об ошибке.

Непосредственное перемещение данных между регионами в настоящее время невозможно — только с помощью службы передачи данных BigQuery Data Transfer Service. Но есть одно исключение: данные из многорегиональной корзины Cloud Storage `US` можно переместить в набор данных BigQuery, хранящийся в любом регионе или объединении регионов. Если вам потребуется загрузить данные в BigQuery из региональной корзины Cloud Storage, эта корзина должна находиться там же, где и набор данных BigQuery (например, оба должны находиться в регионе `asia-east2`), если только корзина не находится в объединенном регионе `US`.

Если у вас нет возможности использовать службу BigQuery Data Transfer Service, есть более сложная процедура перемещения данных, состоящая из нескольких этапов: экспортировать данные из BigQuery в Google Cloud Storage в том же регионе, переместить данные в целевой регион в Cloud Storage и затем загрузить их в набор данных BigQuery. Но имейте в виду, что в этом случае вы понесете дополнительные расходы на хранение данных, пока они находятся в Cloud Storage, а также расходы на передачу данных между регионами в Cloud Storage.

¹ На момент написания книги данные, находящиеся в объединении `EU`, не могли храниться в Цюрихе или Лондоне.

Ограничение доступа к подмножествам данных

Чтобы ограничить доступ ко всему набору данных, можно использовать IAM. Но во многих случаях таблицы могут содержать конфиденциальные данные, и вам может понадобиться ограничить доступ к соответствующим столбцам таблицы. Это можно сделать с помощью авторизованных представлений, динамических фильтров или средств управления доступом.

Авторизованные представления

Авторизованные представления позволяют ограничить доступность определенных столбцов или записей для запросов SQL пользователей. Допустим, у вас есть множество пользователей, которые должны видеть только определенное подмножество столбцов и записей из набора данных `london_bicycles`. Для этого им можно предоставить доступ не к исходному набору данных, а к набору `ch10eu`, содержащему такое представление:

```
CREATE OR REPLACE VIEW ch10eu.authorized_view_300 AS
SELECT
  * EXCEPT (bike_id, end_station_priority_id)
FROM
  [PROJECTID].ch07eu.cycle_hire_clustered
WHERE
  start_station_id = 300 OR end_station_id = 300
```

Пользователи, наделенные правом доступа к этому представлению, не смогут получить доступ к столбцу `bike_id` или к информации с других пунктов проката, кроме пункта с идентификатором 300. Дайте пользователям доступ к этому представлению, открыв набор данных, в состав которого входит это представление. Для этого в веб-интерфейсе BigQuery выберите целевой набор данных (`ch10eu`) и кликните на **Share Dataset** (Поделиться набором данных), а затем откройте к нему доступ нужным пользователям или группе Google. Попробуйте для примера открыть этот набор данных для своей второй учетной записи Google с ролью `BigQuery User`, как показано на рис. 10.2.

Однако авторизованное представление само должно иметь доступ к исходному набору данных. Для этого выберите исходный набор данных (`ch07eu`) в веб-интерфейсе BigQuery, кликните на **Share Dataset** (Поделиться набором данных), а затем в панели **Dataset permissions** (Разрешения для набора данных) выберите авторизованное представление, которому вы решили предоставить доступ, как показано на рис. 10.3.

Если теперь открыть в браузере страницу со следующим URL (замените `[PROJECT]` именем своего проекта):

```
https://console.cloud.google.com/bigquery?p=[PROJECT]&d=ch10eu&page=dataset
```


Dataset permissions

To grant access to this dataset, add members and assign Identity and Access Management (IAM) roles to specify their level of access. Multiple roles allowed.

You can no longer set ACLs in the console to manage access. To learn how IAM and ACLs are related, see the documentation.

DATASET PERMISSIONS **AUTHORIZED VIEWS**

Add members ?

BigQuery Data Viewer
Access to view datasets and all of their contents

Search members

Filter by

Owner (3)
Full access

Editor (9)
Edit access

BigQuery Admin
BigQuery Data Editor
BigQuery Data Owner
✓ BigQuery Data Viewer
BigQuery Metadata Viewer
BigQuery User

BigQuery Data Vi...

Selected
✓ BigQuery Data Viewer

Project
BigQuery
Custom
Data Catalog
IAM

Рис. 10.2. Предоставление доступа к набору данных (ch10eu) с авторизованным представлением

Dataset permissions

To grant access to this dataset, add members and assign Identity and Access Management (IAM) roles to specify their level of access. Multiple roles allowed.

You can no longer set ACLs in the console to manage access. To learn how IAM and ACLs are related, see the documentation.

DATASET PERMISSIONS **AUTHORIZED VIEWS**

Currently authorized views

None.

Share authorized view

Select project Select dataset Select view

Рис. 10.3. Предоставление авторизованному представлению из набора ch10eu права доступа к исходному набору данных ch07eu

зарегистрировавшись со своей второй учетной записью Google, вы сможете просматривать содержимое набора данных `ch10eu`, но набор данных `ch07eu` останется для вас недоступным. Также вы сможете выполнять запросы к представлению:

```
SELECT AVG(duration)
FROM [PROJECT].ch10eu.authorized_view_300
```

Обратите внимание, что из-за ограничений, накладываемых представлением, этот запрос вычислит среднюю продолжительность поездок, которые начались или закончились в пункте проката с идентификатором 300.



Управление доступом с использованием авторизованных представлений не представляет сложностей, пока в игру не вступают многоуровневые представления (представления, вызывающие представления, которые, в свою очередь, вызывают другие представления...). При использовании многоуровневых представлений в запросе SQL появляются ссылки на многие таблицы, и эти таблицы могут находиться в разных наборах данных. Это может усложнить задачу администрирования, потому что придется объединить списки управления доступом (Access Control List, ACL) ко всем представлениям/наборам данных.

Динамическая фильтрация по пользователю

В предыдущем разделе мы создали представление, которое фильтрует набор данных, оставляя доступным для всех имеющих доступ к этому представлению только подмножество столбцов и записей. А как быть, если потребуется открыть доступ к разным записям для разных пользователей? Используйте для этого встроенную функцию `SESSION_USER`, как показано в следующем примере.

Предположим, что вам понадобилось создать представление, возвращающее топ-10 сделок на сумму более 1 000 000 долларов США, но вы хотите, чтобы доступ к нему имели только пользователи из той же компании, которым принадлежит информация о сделках:¹

```
CREATE OR REPLACE VIEW ecommerce.vw_large_transactions
OPTIONS(
  description="large transactions for review",
  labels=[('org_unit', 'loss_prevention')],
  expiration_timestamp=TIMESTAMP_ADD(CURRENT_TIMESTAMP(), INTERVAL 90 DAY)
)
AS

SELECT
```

¹ Указанный набор данных является общедоступным, но по причинам защиты личной информации он не включает адреса электронной почты посетителей. Поэтому данный запрос не будет работать. Мы приводим его только в иллюстративных целях.

```

    visitorId,
    REGEXP_EXTRACT(SESSION_USER(), r'@(.+)') AS user_domain,
    REGEXP_EXTRACT(visitorEmailAddress, r'@(.+)') AS customer_domain,
    date,
    totals.transactions,
    totals.transactionRevenue,
    totals.totalTransactionRevenue,
    totals.timeOnScreen

FROM `bigquery-public-data`.google_analytics_sample.ga_sessions_20170801

WHERE
    (totals.totalTransactionRevenue / 1000000) > 1000
    AND REGEXP_EXTRACT(visitorEmailAddress, r'@(.+)') =
        REGEXP_EXTRACT(SESSION_USER(), r'@(.+)')
ORDER BY totals.totalTransactionRevenue DESC
LIMIT 10

```

Обратите внимание, что представление фильтрует сделки, оставляя только те, которые относятся к тому же домену, откуда пришел пользователь, поэтому топ-10 сделок для @example.com смогут видеть пользователи из @example.com, но они не будут доступны пользователям из @acme.com.

Удаление всех сделок, связанных с конкретным физическим лицом

Предположим, что вы храните в BigQuery данные о сделках пользователя и получаете запрос на удаление из хранилища всех записей с информацией о сделках для `userId=xyz`. Это можно сделать двумя способами (всегда уточняйте у юриста, является ли тот или иной способ подходящим в вашем случае, если удаление выполняется в соответствии с требованиями закона): с использованием DML и методом шифрования.

DML

Удалить данные о пользователе можно с помощью DML-инструкции `DELETE`:

```

DELETE someds.user_transactions
WHERE
    userId = 'xyz'

```

Не забудьте также удалить записи из всех резервных копий и временных таблиц. Удаленные записи можно восстановить в течение семи дней, поэтому если существуют определенные требования к сроку, когда данные должны быть удалены, удалите их как минимум на семь дней раньше этого срока. Как отмечалось в главе 8, эффективнее объединять подобные операции и удалять записи для группы пользователей с помощью `MERGE`.

Уничтожение шифрованием

Второй способ удаления записей состоит в том, чтобы назначить уникальный ключ шифрования каждому `userID` и зашифровать все конфиденциальные данные, принадлежащие пользователю, с помощью этого ключа. Этот способ дает дополнительное преимущество, обеспечивая защиту конфиденциальности пользователей. Чтобы удалить записи с информацией о пользователе, достаточно просто уничтожить ключ. Преимущество этого подхода в том, что он делает удаленные записи невозможными (если нет возможности восстановить уничтоженный ключ) сразу во всех таблицах хранилища данных, включая резервные копии и временные таблицы.

Предположим, что каждый велосипед в наборе данных `london_bicycles` — это «персона», все данные о которой нужно удалить. В частности, мы решили зашифровать информацию о начальном и конечном пунктах проката, между которыми осуществлялись поездки на данном велосипеде.

Сначала создадим таблицу с ключами для всех `bike_id` в наборе данных:

```
CREATE OR REPLACE TABLE ch10eu.encrypted_bike_keys AS
WITH bikes AS (
  SELECT
    DISTINCT bike_id
  FROM
    `bigquery-public-data`.london_bicycles.cycle_hire
)
SELECT
  bike_id, KEYS.NEW_KEYSET('AEAD_AES_GCM_256') AS keyset
FROM
  bikes
```

Для каждого `bike_id` в оригинальном наборе данных в таблице `encrypted_bike_keys` есть свой ключ, например:

Row	bike_id	keyset
1	3792	CJ/erfElEmQKWAowdHlwZS5nb29nbGVhcGlzMnbS9nb29nbGUuY3J5cHRvLnRpbmsuQWVzR2NtS2V5EilalB0bdZJ8sfaEzaN8RyShvuCZL5r0OXf7EztsBLB9V1tMGAEQARif3q3xCCAB
2	5331	CKSmz/ILEmQKWAowdHlwZS5nb29nbGVhcGlzMnbS9nb29nbGUuY3J5cHRvLnRpbmsuQWVzR2NtS2V5EilalKtZqxsil9t415v6ITKAWEi/wxLS0cizdbCDWVPpdl8JGAEQARikps/yCyAB

Мы можем использовать этот набор ключей для шифрования любых данных в `cycle_hire` (в нашем случае к этим данным относится информация о начальном и конечном пунктах поездки, а также конкретные отметки времени). Сначала создадим пару вспомогательных функций для шифрования данных об отдельном велосипеде, потому что некоторые из конфиденциальных столбцов содержат целочисленные значения, а другие — строковые:

```
CREATE TEMPORARY FUNCTION encrypt_int(keyset BYTES, data INT64, trip_start
TIMESTAMP) AS (
  AEAD.ENCRYPT(keyset, CAST(data AS STRING), CAST(trip_start AS STRING))
);
CREATE TEMPORARY FUNCTION encrypt_str(keyset BYTES, data STRING, trip_start
TIMESTAMP) AS (
  AEAD.ENCRYPT(keyset, data, CAST(trip_start AS STRING))
);
```

Затем выполним соединение с `encrypted_bike_keys` по `bike_id`, чтобы получить ключ шифрования для каждой записи, и вызовем вспомогательные функции:

```
CREATE OR REPLACE TABLE ch10eu.encrypted_cycle_hire AS
SELECT
  cycle_hire.* EXCEPT(start_station_id, end_station_id,
                        start_station_name, end_station_name)
  , encrypt_int(keyset, start_station_id, start_date) AS start_station_id
  , encrypt_int(keyset, end_station_id, start_date) AS end_station_id
  , encrypt_str(keyset, start_station_name, start_date) AS start_station_name
  , encrypt_str(keyset, end_station_name, start_date) AS end_station_name
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire
JOIN
  ch10eu.encrypted_bike_keys
USING (bike_id)
```

Этот запрос создаст таблицу, в которой каждый столбец с конфиденциальными данными (здесь `start_station_id`, `end_station_id`, `start_station_name` и `end_station_name`) будет зашифрован с помощью ключей, уникальных для `bike_id` (напомним, что `bike_id` в этом примере играет роль «персоны», конфиденциальность которой мы обеспечиваем). Третий аргумент функции `ENCRYPT` — это «дополнительные» данные, которые можно использовать, чтобы гарантировать возможность расшифровывания только в определенном контексте. Здесь в качестве контекста используется `start_date`. Если этого не сделать, любой имеющий доступ к таблице мог бы поменять местами данные с другой записью, соответствующей этому велосипеду, и получить открытый текст с информацией о местонахождении велосипеда. Но в данном случае, поскольку контекст (`start_date`) отличается, такая атака со стороны инсайдера становится невозможной.¹

¹ Представьте, что программист использует авторизованное приложение, которое отображает местонахождение велосипеда в текущий момент времени. Если у него будет возможность изменить приложение, он сможет подставить ранее зашифрованное местонахождение на место текущего и получить приложение для его отображения. Таким способом он сможет получить историю местонахождений велосипеда, даже не зная ключей шифрования. Однако если использовать `start_date` в качестве дополнительного поля, подобная атака становится невозможной, потому что более раннее местонахождение зашифровано с использованием дополнительных данных, состоящих из зашифрованной даты начала поездки, и его невозможно расшифровать, используя самую последнюю дату. В качестве таких дополнительных данных в функции `ADEAD` можно использовать любые метаданные, которые могут служить таким контекстом.

Чтобы запросить данные из зашифрованной таблицы, их нужно предварительно расшифровать. Например, вот как можно найти пункты проката с самыми продолжительными поездками:

```
CREATE TEMPORARY FUNCTION
decrypt(keyset BYTES, encrypted BYTES, trip_start TIMESTAMP) AS (
    AEAD.DECRYPT_STRING(keyset, encrypted, CAST(trip_start AS STRING))
);

WITH duration_by_station AS (
    SELECT
        duration
        , decrypt(keyset, start_station_name, start_date) AS start_station_name
    FROM
        ch10eu.encrypted_cycle_hire
    JOIN
        ch10eu.encrypted_bike_keys
    USING (bike_id)
)

SELECT
    start_station_name
    , AVG(duration) AS duration
FROM
    duration_by_station
GROUP BY
    start_station_name
ORDER BY duration DESC
LIMIT 5
```

Этот запрос вернет ожидаемые результаты:

Row	start_station_name	duration
1	Stewart's Road, Nine Elms	4836.380090497737
2	Contact Centre, Southbury House	4364.000000000001
3	Speakers' Corner 2, Hyde Park	4006.0086554245627
4	Speakers' Corner 1, Hyde Park	3710.4661268713203
5	Black Lion Gate, Kensington Gardens	3588.0120035566083

Обратите внимание, что мы не можем использовать группировку по зашифрованному значению, потому что шифрование не детерминировано (из-за использования дополнительного контекста) — для одного и того же пункта проката с разными значениями `start_date` получаются разные зашифрованные названия, в результате их оказывается гораздо больше, чем в исходной таблице:

```
SELECT COUNT (DISTINCT start_station_name)
FROM ch10eu.encrypted_cycle_hire
```

Этот запрос вернет число 24369201, тогда как

```
SELECT COUNT (DISTINCT start_station_name)
FROM `bigquery-public-data`.london_bicycles.cycle_hire
```

вернет только 880.

Запрос поиска пунктов проката с самыми продолжительными поездками в зашифрованной таблице выполняется намного медленнее (более чем в два раза), чем аналогичный запрос поиска в незашифрованной таблице. Зато теперь, чтобы уничтожить все сведения, относящиеся `bike_id`, достаточно просто удалить соответствующий ключ:

```
DELETE ch10eu.encrypted_bike_keys
WHERE bike_id = 300
```

Выбирая метод удаления всех данных, относящихся к отдельной личности, — с помощью инструкций DML или путем шифрования, — вы должны решить, что для вас лучше: технические издержки на дополнительное обслуживание или вычислительная сложность.

Предотвращение потери данных

Во многих случаях вы можете даже не знать о существовании конфиденциальных данных. Поэтому иногда полезно сканировать таблицы BigQuery в поисках известных шаблонов, таких как номера кредитных карт, конфиденциальные коды проектов компании и медицинская информация. Результаты сканирования можно использовать как первый шаг на пути к обеспечению надлежащей защиты таких конфиденциальных данных и снизить риск их раскрытия. Также иногда важно периодически проводить такое сканирование, чтобы вовремя заметить изменение характера их использования.

Для сканирования таблиц в BigQuery и защиты конфиденциальных данных можно использовать Cloud Data Loss Prevention (Cloud DLP).¹ Cloud DLP — это управляемая служба, использующая более 90 встроенных детекторов для выявления шаблонов, форматов и контрольных сумм. Она также позволяет определять свои детекторы с использованием словарей, регулярных выражений и контекстных элементов. Служба Cloud DLP включает набор инструментов, позволяющих выполнить деидентификацию данных множеством методов, в том числе маскировкой, токенизацией, псевдонимизацией, сдвигом даты и т. д., причем без репликации данных клиента. Cloud DLP можно применять не только к таблицам в BigQuery, но также к потокам данных и файлам в Google Cloud Storage и к изображениям. Наконец, с ее помощью можно организовать анализ структурированных данных,

¹ См. <https://cloud.google.com/dlp/docs/>. На момент написания этой книги Cloud DLP была глобальной службой. Если у вас есть ограничения на местоположение данных, загляните в документацию и проверьте, отвечает ли служба вашим ограничениям на текущий момент.

чтобы оценить риск реидентификации,¹ включая вычисление таких показателей, как *k*-анонимность (<https://en.wikipedia.org/wiki/K-anonymity>).

Чтобы просканировать таблицу BigQuery, выберите ее в облачной консоли Cloud Console, а затем выберите **Export** ▶ **Scan with DLP** (Экспорт ▶ Сканировать в DLP), параллельно настроив поиск определенных форм данных, как показано на рис. 10.4.

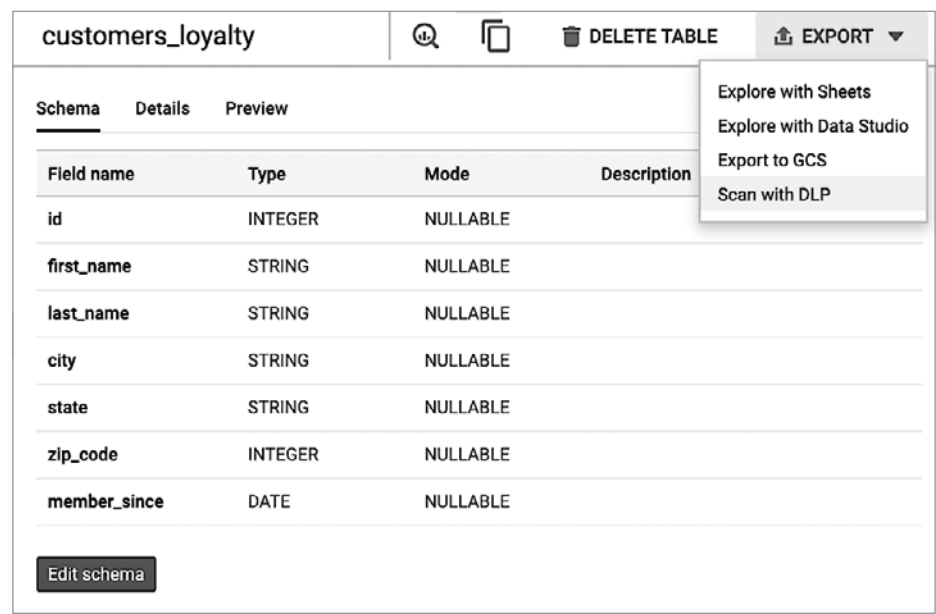


Рис. 10.4. Сканирование таблицы BigQuery с использованием Cloud DLP

Чтобы отредактировать или иным образом деидентифицировать конфиденциальные данные, обнаруженные при сканировании с помощью Cloud DLP, защитите их с помощью ключей Cloud KMS, которые мы обсудим в следующем разделе.

СМЕК

BigQuery использует *двойное шифрование* (envelope encryption) данных в таблице без всякого вмешательства с вашей стороны. При двойном шифровании данные в таблице BigQuery сначала шифруются с использованием ключа шифрования данных (Data Encryption Key, DEK), а затем ключи DEK шифруются

¹ См. https://en.wikipedia.org/wiki/Data_re-identification. Это риск, связанный с тем, что анонимные данные удастся сопоставить со вспомогательными данными и идентифицировать лица, с которыми эти данные связаны.

ключом шифрования ключей. Данные каждого клиента GCP, доступ к которым он не открывал для других, делятся на блоки и шифруются другими ключами, которые не используются для других клиентов. Эти ключи DEK отличаются даже от ключей, которыми шифруются другие части тех же данных, принадлежащих тому же клиенту. Ключи шифрования ключей используются для шифрования ключей DEK, которые, в свою очередь, используются для шифрования ваших данных. Эти ключи шифрования ключей (Key Encryption Keys, KEK) управляются централизованной службой управления ключами Google Key Management Service (KMS), как показано на рис. 10.5.



Рис. 10.5. Двойное шифрование с помощью ключей DEK и KEK; ключи KEK управляются централизованной службой KMS, которая производит циклическую смену ключей

Благодаря шифрованию злоумышленник не сможет получить доступ к данным, случайно завладев ими, не имея также доступа к ключам шифрования. Даже если злоумышленник каким-то образом завладеет устройствами хранения с вашими данными, он не сможет их расшифровать. Шифрование также действует как «контрольно-пропускной пункт» — централизованная служба управления ключами шифрования создает единую точку, обеспечивающую и контролирующую доступ к данным. Наконец, шифрование увеличивает защищенность данных клиента, позволяя системам манипулировать данными, например выполнять резервное копирование, а инженерам — поддерживать инфраструктуру, не имея прямого доступа к самим данным.

Если правила требуют, чтобы вы контролировали ключи, используемые для шифрования данных, вам может пригодиться механизм поддержки ключей шифрования, управляемых клиентом (СМЕК). Напомним, что ключи DEK, используемые для шифрования данных, сами шифруются с помощью ключей КМК, которые хранятся и управляются централизованно. Если вам понадобится самим управлять шифрованием ключей, можно запустить службу KMS в центральном проекте и использовать свои ключи для шифрования табличных данных во всех проектах вашей организации.

В Cloud KMS ключи находятся в связке, которая хранится в определенном месте. Если вы решите запустить службу KMS в центральном проекте, создайте связку ключей, указав период циклической смены:

```
gcloud kms keyrings create acmecorp --location US
gcloud kms keys create xyz --location US \
  --keyring acmecorp --purpose encryption \
  --rotation-period 30d \
  --next-rotation-time 2019-07-01T12:00:00Z
```

Связка ключей должна создаваться в том же местоположении, где хранятся ваши наборы данных BigQuery. Например, связка ключей для защиты набора данных в регионе US должна храниться в том же регионе US, а связка ключей для защиты набора данных в регионе `asia-northeast1` должна храниться в регионе `asia-northeast1`.

Ключи из Cloud KMS используются для шифрования ключей DEK, с помощью которых шифруются данные в BigQuery. Итак, создав ключ, вы должны разрешить учетной записи службы BigQuery в каждом проекте (не в проекте KMS) использовать его для шифрования и дешифрования данных:

```
SVC=$(bq show --encryption_service_account)
gcloud kms keys add-iam-policy-binding \
  --project=[KMS_PROJECT_ID] \
  --member serviceAccount:$SVC \
  --role roles/cloudkms.cryptoKeyEncrypterDecrypter \
  --location=US \
  --keyring=acmecorp \
  xyz
```

Затем, при создании таблицы, укажите используемый ключ:

```
bq mk ... --destination_kms_key \
projects/[PROJECT_ID]/locations/US/keyRings/acmecorp/cryptoKeys/xyz \
mydataset.transactions
```

Помимо этой настройки таблиц на этапе создания не требуется никаких специальных мер для запроса данных из таких таблиц, защищенных службой Cloud KMS. BigQuery хранит имя ключа, с помощью которого зашифровано содержимое таблицы, и будет использовать его при обработке запроса к такой таблице. Все существующие инструменты, веб-интерфейс BigQuery и утилита командной строки `bq` по умолчанию прозрачно обрабатывают зашифрованные таблицы, при условии, что BigQuery имеет доступ к ключу Cloud KMS, использованному для шифрования содержимого таблицы.

Защита от утечки данных

Управление службами в виртуальном частном облаке (Virtual Private Cloud Service Controls, VPC-SC) позволяет пользователям выстроить периметр безопасности вокруг ресурсов Google Cloud Platform, таких как корзины в Cloud Storage и наборы данных в BigQuery, и снизить риск утечки данных за счет

предотвращения их эксфильтрации за периметр VPC. Дополнительно используя Private Google Access, можно настроить гибридную облачную среду, обеспечивающую защиту конфиденциальных данных.

Механизм VPC-SC обеспечивает дополнительную защиту периметра на основе контекста, помимо контроля доступа на основе идентификации личности, предлагаемого Cloud IAM. С помощью VPC-SC можно минимизировать риски безопасности, связанные с доступом из неавторизованных сетей с использованием украденных учетных данных, несанкционированным экспортом данных недовольными инсайдерами и непреднамеренной утечкой личных данных из-за неправильно настроенных политик IAM. Механизм VPC-SC позволяет также предотвратить чтение данных из ресурса или их копирование в ресурс за пределами периметра с использованием таких инструментов, как `gsutil` и `bq`.

Для настройки VPC Service Controls перейдите в раздел **VPC Service Controls** в консоли GCP и добавьте новый периметр. Затем укажите проекты и службы в этих проектах, которым разрешено связываться друг с другом в пределах периметра. Допустим, что вы выбрали только проект Project A и две службы: BigQuery и Cloud Storage. В этом случае вы сможете загружать данные из GCS в BigQuery и экспортировать данные из BigQuery в GCS, но только в корзины, принадлежащие тому же проекту. Вам не удастся загрузить данные в BigQuery из корзин или экспортировать данные из BigQuery в корзины, принадлежащие другим проектам. То же касается копирования данных между корзинами или обращений к наборам данных, которые являются частью других проектов. Конечно, вы можете включить в периметр два проекта и тем самым обеспечить возможность взаимодействий между проектами (но только между этими двумя проектами).

Выводы

В этой главе мы описали инфраструктуру безопасности, лежащую в основе BigQuery, и обсудили, как пользователи и приложения могут осуществлять аутентификацию и авторизацию с использованием IAM. Затем мы рассмотрели различные инструменты, которые помогают обеспечить соответствие нормативным и правовым требованиям. Однако повторим еще раз, что вы всегда должны консультироваться с юристом, чтобы понять, насколько реализация любого из этих инструментов удовлетворяет вашим нормативным требованиям или требованиям соответствия.

Благодарим вас за участие в нашем совместном путешествии по BigQuery. В главе 1 мы познакомились с основами сервиса, в главе 2 немного углубились в синтаксис SQL, в главе 3 описали типы данных, в главе 4 рассмотрели загрузку данных, а в главе 5 исследовали среду разработки. В главе 6 мы присту-

пили к изучению более продвинутых аспектов, начав с описания архитектуры BigQuery. В главе 7 мы дали различные советы по повышению производительности, а в главе 8 рассмотрели некоторые тонкости BigQuery. Глава 9 была посвящена машинному обучению в BigQuery, а эта — вопросам безопасности. BigQuery является бессерверным хранилищем корпоративных данных, требующим минимальных настроек, поэтому мы смогли в этой книге уделить большое внимание анализу данных и продемонстрировать важные идеи. Мы надеемся, что вам понравится работать с BigQuery и вы преуспеете в этом!

Об авторах

Валиappa (Лак) Лакшманан (Valliappa (Lak) Lakshmanan) — руководитель отдела аналитики данных и решений искусственного интеллекта в Google Cloud. Его команда разрабатывает программные решения для бизнес-задач, используя BigQuery и другие продукты Google Cloud для анализа данных и машинного обучения. Он также является автором книги «Data Science on the Google Cloud Platform», выпущенной издательством O'Reilly.

Джордан Тайджани (Jordan Tigani) — директор по управлению продуктами для BigQuery. Был одним из основателей базы BigQuery и помог сделать ее одним из самых успешных продуктов в Google Cloud. Написал первую книгу о BigQuery, а также занимался его популяризацией. Джордан имеет 20-летний опыт разработки программного обеспечения, работал в Microsoft Research и во многих стартапах, занимающихся машинным обучением.

Об обложке

На обложке изображен масайский страус (лат. *Struthio camelus massaicus*), подвид страуса — крупнейшей птицы в мире. Масайский страус обитает на открытых равнинах и в травянистых саваннах Восточной Африки.

Масайский страус достигает 2–2,7 метра в высоту и, несмотря на размах крыльев до 2 метров, не умеет летать. Страусы прекрасно обходятся без этой способности: бородки их перьев не сцеплены друг с другом и не позволяют образовать опахало (как у летающих птиц), тем не менее страусиные крылья обеспечивают птице некоторую подъемную силу и устойчивость при побеге от хищников. У страуса длинные сильные ноги, позволяющие ему развивать скорость до 70 км/час, что делает его самой быстрой птицей на земле, а также самым быстрым двуногим животным.

Самцы имеют характерное черное оперение с небольшой белой каймой вокруг крыльев и хвоста, контрастирующее с красноватой шеей и ногами (которые становятся еще ярче в период спаривания). Самки же имеют коричнево-серый окрас. Кроме того, у масайского страуса два пальца, тогда как у большинства птиц их четыре, причем один из них напоминает копыто. Страусы кочуют стадами, насчитывающими до 50 особей, часто совместно с другими травоядными животными, такими как антилопы или зебры.

Существует распространенное мнение, что в случае опасности страус прячет голову в песок. Считается, что этот миф исходит из писания Плиния Старшего, который действительно мог видеть, как страусы глотают песок и гальку (это помогает птицам переваривать пищу, потому что у них нет зубов). Также вполне возможно, что он мог видеть, как во время высиживания они переворачивают яйца, которые откладывают в песок. Как бы то ни было, когда масайский страус чувствует угрозу, он или убегает, или пытается спрятаться, прижавшись к земле. В экстремальных ситуациях они будут давать отпор и даже способны убить льва.

Многие животные, изображенные на обложках книг издательства O'Reilly, находятся под угрозой вымирания; все они очень важны для биосферы.

Иллюстрацию для обложки нарисовал Хосе Марзан (Jose Marzan) по черно-белой гравюре из книги «Meyers Kleines Lexicon».

Валиappa Лакшманан, Джордан Тайджани
**Google BigQuery. Всё о хранилищах данных,
аналитике и машинном обучении**

Перевел с английского *А. Киселев*

Заведующая редакцией	<i>Ю. Сергиенко</i>
Ведущий редактор	<i>К. Тульцева</i>
Литературный редактор	<i>А. Попова</i>
Художественный редактор	<i>В. Мостипан</i>
Корректоры	<i>С. Беляева, Н. Викторова</i>
Верстка	<i>Л. Егорова</i>

Изготовлено в России. Изготовитель: ООО «Прогресс книга».

Место нахождения и фактический адрес: 194044, Россия, г. Санкт-Петербург,
Б. Сапсониевский пр., д. 29А, пом. 52. Тел.: +78127037373.

Дата изготовления: 08.2020. Наименование: книжная продукция. Срок годности: не ограничен.

Налоговая льгота — общероссийский классификатор продукции ОК 034-2014, 58.11.12 — Книги печатные
профессиональные, технические и научные.

Импортер в Беларусь: ООО «ПИТЕР М», 220020, РБ, г. Минск, ул. Тимирязева, д. 121/3, к. 214, тел./факс: 208 80 01.

Подписано в печать 22.07.20. Формат 70×100/16. Бумага офсетная. Усл. п. л. 39,990. Тираж 700. Заказ 0000.

Отпечатано в ОАО «Первая Образцовая типография». Филиал «Чеховский Печатный Двор».

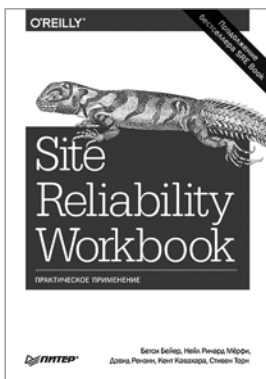
142300, Московская область, г. Чехов, ул. Полиграфистов, 1.

Сайт: www.chpk.ru. E-mail: marketing@chpk.ru

Факс: 8(496) 726-54-10, телефон: (495) 988-63-87

*Бетси Бейер, Нейл Ричард Мёрфи, Дэвид Рензин,
Кент Кавахара, Стивен Торн*

SITE RELIABILITY WORKBOOK: ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ



Книга Site Reliability Engineering спровоцировала бурную дискуссию. Что сегодня понимается под эксплуатацией и почему столь фундаментальную важность имеют вопросы надежности? Теперь инженеры Google, участвовавшие в создании этого бестселлера, предлагают перейти от теории к практике — Site Reliability Workbook покажет, как принципы и практика SRE воплощаются в вашем продакшене. Опыт специалистов Google дополнен кейсами пользователей Google Cloud Platform. Представители Evernote, The Home Depot, The New York Times и других компаний описывают свой боевой опыт, рассказывают, какие практики у них прижились, а какие — нет. Эта книга поможет адаптировать SRE к реалиям вашей собственной практики, независимо от размеров вашей компании.

Вы научитесь: обеспечивать надежность сервисов в облаках и средах, которые вы не полностью контролируете; применять различные методы создания, запуска и мониторинга сервисов, ориентируясь на SLO; трансформировать команды админов в SRE-инженеров; внедрять методы запуска SRE с чистого листа и на базе существующих систем. Бетси Бейер, Нейл Ричард Мёрфи, Дэвид Рензин, Кент Кавахара и Стивен Торн занимаются обеспечением надежности систем Google.

КУПИТЬ