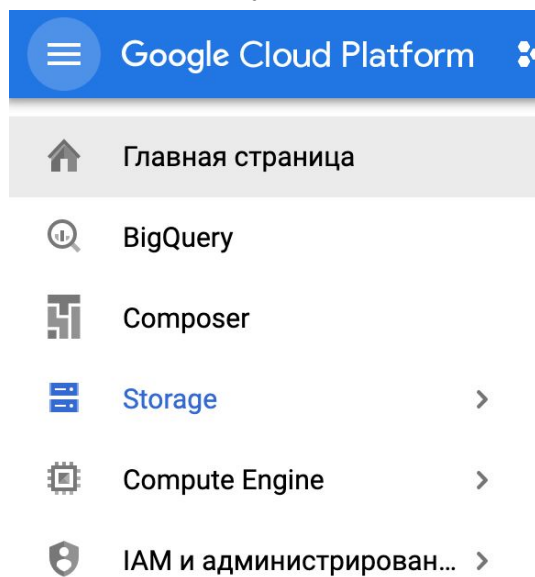


Инструкция к занятию “Облачные платформы”

1. [Зарегистрируйтесь в GCP](#)
2. Перейдите по адресу <https://console.cloud.google.com/> и создайте проект с произвольным названием (если у вас еще нет тестового проекта).
3. Создайте bucket (сегмент) в сервисе Cloud Storage.
 - а. Для этого найдите в меню навигации пункт “Storage > Обзор” и перейдите по нему



b. Нажмите кнопку “Создать сегмент” и установите следующие настройки:

Название ?
Названия сегментов Cloud Storage не должны повторяться. Если сегмент используется для [обслуживания веб-сайта](#), в качестве названия укажите доменное имя этого сайта.

Класс хранилища по умолчанию
Выбранный класс хранилища будет по умолчанию назначаться для всех объектов, добавляемых в этот сегмент. От класса хранилища объекта и местоположения сегмента зависят уровни георезервирования и доступности, а также стоимость хранения. Настроить класс хранения для отдельных объектов можно в gsutil. [Подробнее...](#)

i Данные в хранилищах Nearline и Coldline теперь резервируются в разных географических регионах. В новых местоположениях nat4 и eur4, доступных в бета-версии, используется георезервирование, при этом ресурсы для вычислений и хранения данных размещаются рядом для повышения производительности. [Подробнее...](#) Закреть

- Multi-Regional
- Regional
- Nearline
- Coldline

Местоположение

[Сравнить классы хранилища](#)

Стоимость хранения	Стоимость извлечения	Операции класса A ?
0,02 \$ за ГБ в месяц	Бесплатно	0,005 \$ за 1000 операций

Модель управления доступом
Выберите способ управления доступом к объектам в этом сегменте. [Подробнее...](#)

- Установить одинаковые права доступа на уровне сегмента (только политика сегмента)
Управление доступом будет осуществляться с помощью IAM-политики

4. Создайте виртуальную машину типа f1-micro в сервисе Compute Engine.
 - a. Для этого перейдите в меню навигации “Compute engine > Экземпляры VM”
 - b. Нажмите “Создать экземпляр” (этот пункт может скрываться за троеточием рядом с заголовком “Экземпляры VM”)

- c. Вбейте следующие настройки (поле “Название” заполните “gcp_intro_data_source_service”)

Название ?

Регион ? **Зона** ?

us-central1 (Айова) ▼ us-central1-a ▼

Тип машины
Чтобы выбрать количество ядер, объем памяти и количество графических процессоров, нажмите “Настроить”.


микромашина ... ▼ 0,6 ГБ памяти [Настроить](#)

Если вы [перейдете на платный аккаунт](#), то сможете создавать экземпляры с количеством ядер до 96.

Контейнер ?

Развернуть образ контейнера в этом экземпляре VM. [Подробнее...](#)

Загрузочный диск ?

 Новый стандартный постоянный диск объемом 10 ГБ
Образ
Debian GNU/Linux 9 (stretch) [Изменить](#)

Профиль и API-доступ ?

Сервисный аккаунт ?

Compute Engine default service account ▼

Область действия доступа ?

Доступ по умолчанию
 Полный доступ ко всем Cloud API
 Настроить доступ к каждому API

Брандмауэр ?

5. Откройте веб-консоль, нажав SSH

<input type="checkbox"/>	Название ^	Зона	Рекомендация	Используется	Внутренний IP-адрес	Внешний IP-адрес	Подключиться
<input type="checkbox"/>	<input checked="" type="checkbox"/> instance-1	us-central1-a			10.128.0.6 (nic0)	35.193.227.105	SSH ▼

6. Скачайте датасет

- a. В консоли SSH выполните команду
`wget -O winemag-data.json.tgz https://bit.ly/2R8z55k`
- b. Распакуйте файл командой
`tar -xzf winemag-data.json.tgz`

7. Создайте сервисную роль и авторизуйте её на машине
(<https://cloud.google.com/sdk/docs/authorizing>)

- a. Перейдите на страницу “IAM и администрирование > Сервисные аккаунты” и нажмите “Создать сервисный аккаунт”

Сведения о сервисном аккаунте

Название сервисного аккаунта
gcp_intro_account_1

Отображаемое название этого сервисного аккаунта

Идентификатор сервисн...
gcp-intro-account-1 @thermal-loop-237313.iam.gserviceaccount. X C

Описание сервисного аккаунта

Рекомендуем выбрать описательное название

СОЗДАТЬ ОТМЕНА

- b. На втором шаге создания добавьте права “Хранение данных > Создатель объектов в хранилище”

Права доступа для сервисного аккаунта (необязательно)

Предоставьте сервисному аккаунту доступ к проекту Showcase project, чтобы сервисный аккаунт мог выполнять определенные действия с ресурсами в этом проекте. [Подробнее...](#)

Выберите роль

Введите фильтр

Создание меток для ...	Администратор хранилища
Управление пользова...	Администратор объектов в х...
Управление сервиса...	Создатель объектов в хранил...
Утверждение доступа	Наблюдатель объектов в хра...
Хранение данных	Администратор переноса дан...
Android Management	Пользователь переноса данн...
App Engine	Наблюдатель переноса данн...
AutoML	

Создатель объектов в хранилище
Доступ для создания объектов в GCS.

УПРАВЛЕНИЕ РОЛЯМИ

- c. На третьем и четвертом шаге никаких изменений не требуется, просто нажимайте “Продолжить/Создать”
- d. Перейдите на страницу с сервисными аккаунтами и скачайте ключ

gcp-intro-account-2@thermal-loop-237313.iam.gserviceaccount.com gcp_intro_account_2

Изменить
Создать ключ
Удалить
Показать идентификатор клиента

- e. Откройте ключ в текстовом редакторе и скопируйте содержимое
- f. Создайте на машине в облаке (через SSH-консоль) файл key.json (для этого можно использовать текстовый редактор [vim](#)) и вставьте в него содержимое файла с предыдущего шага

- g. Откройте консоль и авторизуйтесь в качестве сервисной роли командой
`gcloud auth activate-service-account --key-file key.json`
- 8. Загрузите датасет командой
`gsutil cp winemag-data-130k-v2.json gs://gcp_intro/`
- 9. Загрузите данные в BigQuery
 - a. В меню навигации перейдите в “BigQuery”
 - b. Выделите свой проект в левом меню и создайте Dataset (Набор данных) с именем winemag (помните, что регион везде должен быть один и тот же - US)



- c. Установите права на бакет (сегмент), чтобы вы могли управлять объектами в нем (если этот пункт не будет выполнен, вы увидите ошибку “ResumableUploadAbortException: 403 Insufficient Permission”)

gcp_intro_bucket_1

Объекты Общая информация Права доступа Блокировка сегмента

Object-level permissions are disabled
 У вас осталось 89 дней, чтобы включить списки контроля доступа на уровне отдельных объектов . [Подробнее...](#) Enable

Добавить участников
Удалить
Введите имя или роль
Группировать по:
Участникам ▾

Тип	Участники ^	Роли
<input type="checkbox"/>	573182219267@cloudbuild.gserviceaccount.com	Сервисный аккаунт сборки Cloud унаследовано ✕ ✎
<input type="checkbox"/>	gcp-intro-1@thermal-loop-237313.iam.gserviceaccount.com	Создатель объектов в хранилище унаследовано ✕ ✎
<input type="checkbox"/>	gcp-intro-account-2@thermal-loop-237313.iam.gserviceaccount.com	Создатель объектов в хранилище унаследовано ✕ ✎
<input type="checkbox"/>	gcp-intro-role-1@thermal-loop-237313.iam.gserviceaccount.com	Несколько унаследовано ✕ ✎
<input checked="" type="checkbox"/>	Владельцы проекта thermal-loop-237313	Несколько ✕ ✎
<input type="checkbox"/>	Редакторы проекта thermal-loop-237313	Несколько ✕ ✎
<input type="checkbox"/>	Наблюдатели проекта thermal-loop-237313	Несколько ✕ ✎

Добавить участников
Удалить
Введите имя или роль
Группировать по:
Участникам ▾

Тип	Участники ^	Роли
<input type="checkbox"/>	573182219267@cloudbuild.gserviceaccount.com	Сервисный аккаунт с
<input type="checkbox"/>	gcp-intro-1@thermal-loop-237313.iam.gserviceaccount.com	Создатель объектов в
<input type="checkbox"/>	gcp-intro-account-2@thermal-loop-237313.iam.gserviceaccount.com	Создатель объектов в
<input type="checkbox"/>	gcp-intro-role-1@thermal-loop-237313.iam.gserviceaccount.com	
<input checked="" type="checkbox"/>	Владельцы проекта thermal-loop-237313	

Выбранные

- Администратор объектов в хранилище
- Владелец объектов устаревших хранилищ
- Владелец сегментов с устаревшими хранилищами

Cloud Build ▶

Cloud Composer ▶

IAM ▶

Устаревшие хранилища ▶

Хранение данных ▶

Сохранить
Отмена

d. Выберите набор данных в левом меню и создайте таблицу из загруженных данных

Предоставление



Планирование запроса



СОЗДАТЬ ТАБЛИЦУ



Создать таблицу

Источник

Создать таблицу на основе:

Выберите файл из сегмента GCS: [?](#)

Формат файла:

Google Cloud Storage



gcp_intro_bucket_1/winemag-data-130k-v2.json

Обзор

JSON (pa... [?](#)

Место назначения

Название проекта

Showcase project

Название набора данных

winemag

Тип таблицы [?](#)

Таблица в собстве... [?](#)

Название таблицы

winemag_reviews

Схема

Автоматическое определение

Схема и входные параметры

[i](#) Схема будет создана автоматически.

Секционирование и кластеризация

Секционирование [?](#)

Без секционирования

Порядок кластеризации (необязательно): [?](#)

Clustering order determines the sort order of the data. Clustering can only be used on a partitioned table, and works with tables partitioned either by column or ingestion time.

Не более 4 полей, определяющих порядок кластеризации, через запяту

Дополнительные параметры [v](#)



Создать таблицу



Отмена

10. Выполните запрос к таблице

```
Редактор запросов  
1 select country, avg(points) from winemag.winemag_reviews group by 1 order by 2 desc
```

11. * Дополнительное задание: постройте дашборд в DataStudio с парой графиков, сравнивающих цены и оценки для вин разных типов и стран.

авление |  Планирование запроса ▼ |  Ещё ▼

  **ПРОСМОТРЕТЬ В СТУДИИ ДАННЫХ**

ия о выполнении