

УРОК 1 🚩 Вступ у дата-аналітику. Мова SQL та ознайомлення з Google BigQuery.

Щодня мільйони людей по всьому світу гуглять інформацію. Гугл у відкритому доступі має таблиці із запитам. Тож ми з тобою уявимо себе дата-аналітиками й ці запити проаналізуємо.

Твоїм викладачем буде Нікіта Тимошенко, Data Analyst в Adante Consulting, Канада.
Просимо любити і шанувати 😊

▶ https://i.goit.global/da_lesson1
https://www.youtube.com/watch?v=dP8RjzfoU8w&ab_channel=GoIT

Тривалість: 27 хвилин.

PS: всі посилання для ДЗ знайдеш в описі під відео.

ПРАКТИЧНЕ ЗАВДАННЯ #1 (до 18:00 завтрашнього дня)

1. Додати до двох наявних колонок з бази `international_top_terms`, які ви створювали разом із викладачем на уроці, ще 5 колонок: `region_name`, `refresh_date`, `rank`, `score`, `week`.
2. Написати подібний код SQL, але вже до бази `top_terms`.

Формат здачі (приклад ДЗ
(https://drive.google.com/file/d/1rZ13Azzr_a4uzyx_kzRdeLZrClmuQV/view?usp=sharing)): 2 скриншоти кожного запиту.

📌 Практична підказка: роби одразу скриншот, де вміщається запит і відпрацьована таблиця.

УРОК 1 Конспект

SQL = Structured Query Language

-> Мова програмування для взаємодії користувача з реляційними базами даних

→

Використовується для створення, доступу та оновлення даних, збережених у реляційній БД

Для аналізу даних, ми маємо їх зібрати з якогось безкоштовного джерела. Таким джерелом є Google Cloud.

<https://console.cloud.google.com/welcome?pli=1>

Реєструємося і знаходимо продукт Google Big Query

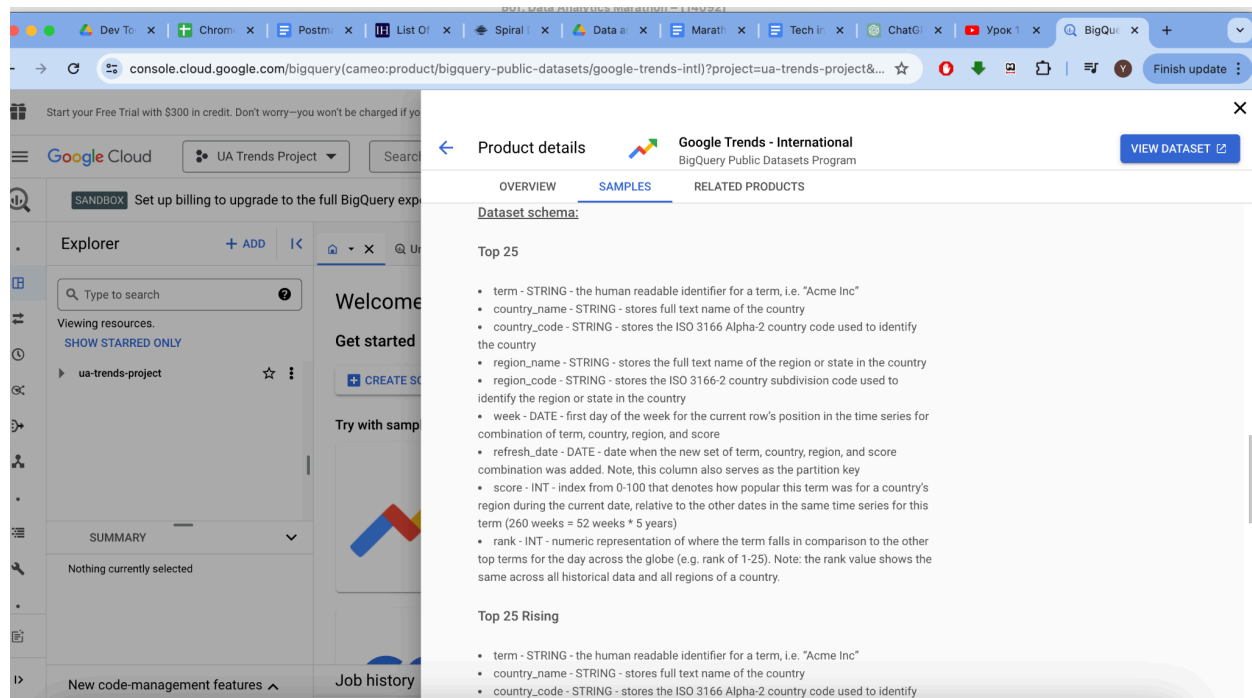
Create Project

Explore -> ADD

Public Datasets

Шукаємо "Trends"

Вибираємо Google Trends International



View dataset

Знаходимо ліворуч базу даних google_trends - тут є 4 таблиці

Таблиця International-top-terms - найпопулярніші запити до Гугла з багатьох країн світу.

Top-terms - ті ж дані, тільки на США

Закладки SCHEMA, OVERVIEW, DETAILS

Щоб для цієї сторінки створити запит, в меню "З вертикальні точки" знаходимо QUERY

Назву таблиці можна скопіювати з меню "З вертикальні точки" - Copy ID

Пишемо код, наприклад:

```
SELECT *
```

```
FROM bigquery-public-data.google_trends.international_top_terms
```

```
LIMIT 100;
```

Щоб запустити тільки цей код, виділяємо його і натискаємо RUN

Коментарі починаються з -- (можна закоментити за допомогою хоткейс: Cmd ?)

Розбір домашки №1 https://www.youtube.com/watch?v=4wbhnxxC8xQ&ab_channel=GoIT

УРОК 2

https://www.youtube.com/watch?v=mhpaY8a0CS0&ab_channel=GoIT

Розбір домашки: https://www.youtube.com/watch?v=N2Mc1vUPr1Q&ab_channel=GoIT

Лінк на квері:

<https://console.cloud.google.com/bigquery?sq=365695971000:d7170f46a4b247bb8fd26ff9eb58eabd>

```
-- SELECT *  
-- FROM `bigquery-public-data.google_trends.international_top_terms`  
-- WHERE refresh_date = "2024-05-12" LIMIT 1000;
```

-- 1. Мінімальна і максимальна дата/тиждень оновлення

```
SELECT  
MAX(week) AS MAX_week,  
MIN(week) AS min_week,  
MAX(refresh_date) AS MAXIMAL_DATE  
-- MIN(refresh_date) AS min_date  
-- SUM(rank) SUM_of_RANKS  
-- AVG()  
FROM bigquery-public-data.google_trends.international_top_terms;
```

-- 2. Список всіх країн

```
SELECT  
DISTINCT country_name AS COUNTRY_LIST  
FROM bigquery-public-data.google_trends.international_top_terms  
ORDER BY country_name ASC;
```

-- 3. Дані для певної країни

```
SELECT *
```

```
FROM bigquery-public-data.google_trends.international_top_terms
WHERE country_name = "Ukraine"
AND
-- refresh_date = "2024-04-10"
refresh_date = "2024-06-25"
AND
-- week = "2024-04-07"
week = "2024-06-23"
AND
region_name = "Vinnyts'ka oblast"
LIMIT 100;
```

-- 4. Список областей України

```
SELECT DISTINCT region_name, region_code
FROM bigquery-public-data.google_trends.international_top_terms
WHERE country_name = "Ukraine"
ORDER BY region_name ASC
;
```

-- 5. Список ранків для Вінницької області

```
SELECT DISTINCT region_name, region_code
FROM bigquery-public-data.google_trends.international_top_terms
WHERE country_name = "Ukraine"
ORDER BY region_name ASC
;
```

-- Д32, Вправа 1: надати для твого регіону топ-25 запитів за минулий тиждень

```
SELECT *
FROM bigquery-public-data.google_trends.international_top_terms
WHERE country_name = "Ukraine"
AND
week = "2024-04-07"
AND
refresh_date = "2024-04-10"
AND
```

```
region_name = "Vinnyts'ka oblast"
-- AND rank = 1
ORDER BY rank ASC
-- LIMIT 25
;
```

УРОК 3

https://www.youtube.com/watch?v=Uz05B1LE2us&ab_channel=GoIT

```
SELECT region_name, *
FROM bigquery-public-data.google_trends.international_top_terms
WHERE country_name = "Ukraine"
AND
week = "2024-04-07"
AND
refresh_date = "2024-04-10"
AND
region_name LIKE "Vinnyt%"
;
```

Щоб шукати термін зі схожою назвою, використовуємо команду LIKE і знак % (місце цього знаку залежить якого виду маска нам потрібна)

ПРАКТИЧНЕ ЗАВДАННЯ #3

Показати всі дані пошуків, де в запиті фігурує "2024". Можеш зробити запит по своїй області, по Україні або загалом для всіх країнах, що містяться у базі даних international_top_terms. Враховуй, що дані по останній актуальній даті можуть бути відсутні.

<https://console.cloud.google.com/bigquery?sq=365695971000:d7170f46a4b247bb8fd26ff9eb58eabd> - лінк на квері

```
SELECT
-- *
DISTINCT term
```

```
FROM bigquery-public-data.google_trends.international_top_terms
WHERE
region_name = "Vinnyts'ka oblast"
AND
term LIKE "%2024%"
ORDER BY term DESC
;
```

Бонусне відео. Візуалізація даних в Looker Studio.

https://www.youtube.com/watch?v=x7H35VQX7Ak&ab_channel=GoIT

<https://lookerstudio.google.com/u/0/navigation/reporting>

<https://cloud.google.com/bigquery/docs/> - більше дізнатись про можливості Big query

В Looker Studio дуже легко співпрацювати з даними з Big Query.
Інші унструменти для створення репортів: Power BI

Щоб створити новий звіт:

Looker Studio -> Blank Report

Тут можна вибрати різні джерела для формування звіту, такі як File Upload, YouTube, Big Query

Повертаємося в Big Query, знаходимо потрібну таблицю (в нашому випадку international_top_terms), пишемо квері, виконуємо

Кількість рядків для звіту не може бути більше ніж 80.000 для CSV (local file).

Результати зберігаємо: SAVE RESULTS -> (this time:) CSV (local file)

Мені не вдалося завантажити збережений файл, але я змогла завантажити з google drive:

https://docs.google.com/spreadsheets/d/1LqgKjq_rv_s5iOw8dKjQ66FEsK7QVIsN4hc3KonOUXQ/edit?gid=1429491937#gid=1429491937

Результат можемо відкрити в Google Sheets. Тут можна фільтрувати , сортувати і т.д.

Після завантаження файлу в Looker , перевіряємо типи даних, який Лукер визначив самостійно. (праворуч)

Є опція коригування назв колонок і їхніх типів даних (я так і не зрозуміла як я увійшла в режим редагування (через Add field? А потім Cancel?) : на назву таблиці в правій стороні екрану - з'явиться "олівчик" для редагування - відкриється вікно редагування.

По лівій стороні є табличка - за замовченням (її можна змінювати).

Theme and layout -> Layout -> Canvas size:

375*812 - для мобільних девайсів

Можна додати назву

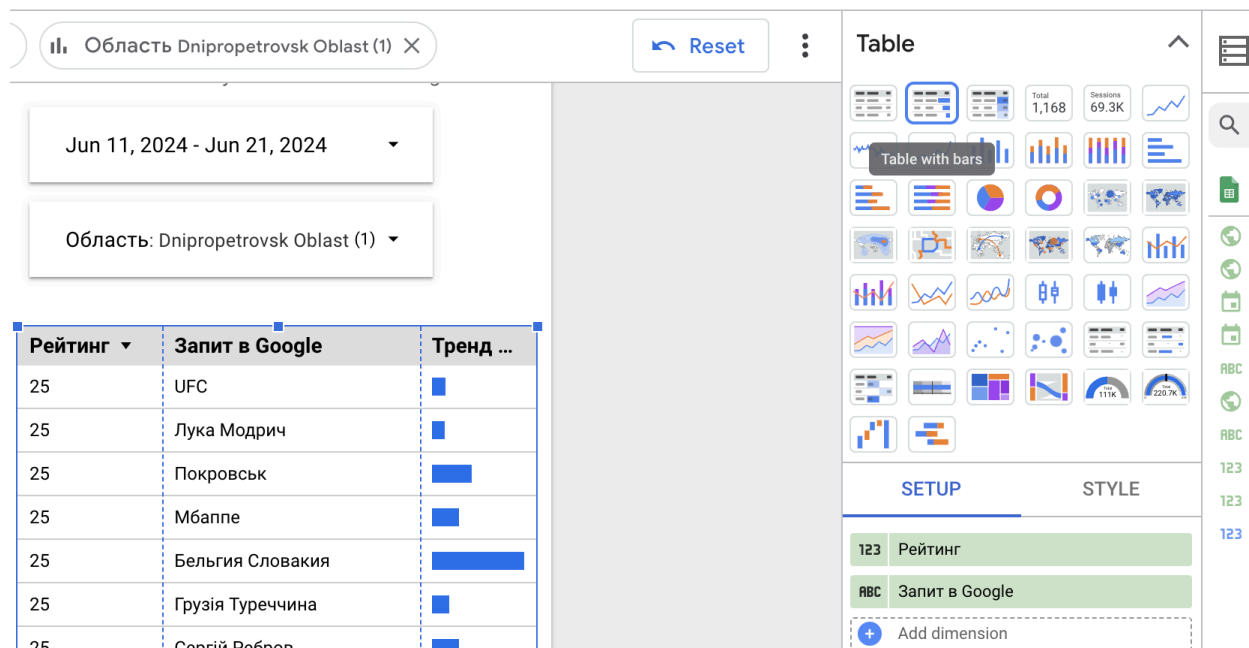
Можна додати фільтри над таблицею:

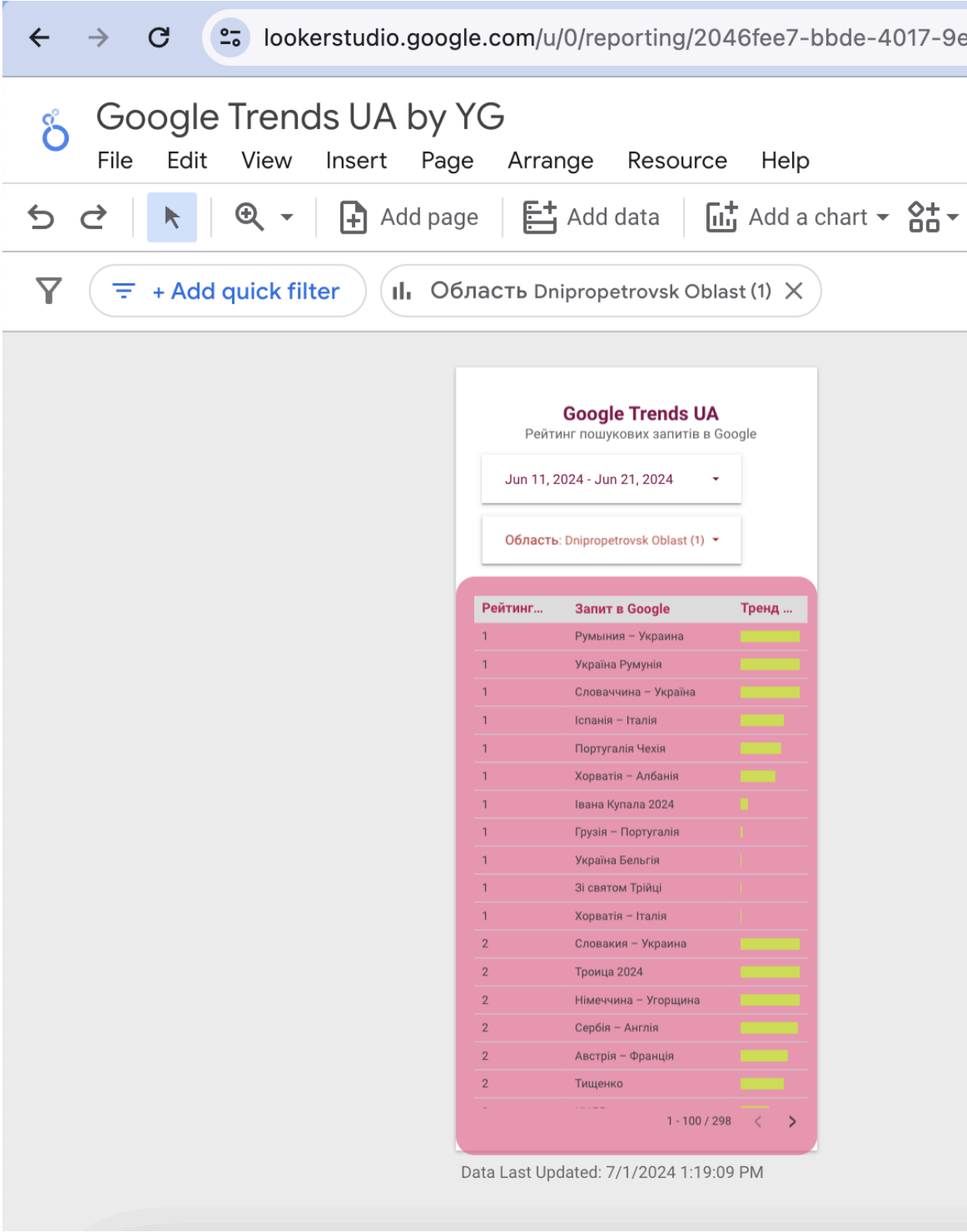
Add control -> Date Range Control

Add control -> Drop-down List

Змінити вигляд таблиці:

Стаємо на таблицю, вибираємо Chart і бажаний вигляд, наприклад:





<https://lookerstudio.google.com/u/0/reporting/2046fee7-bbde-4017-9ea5-874cf04f9061/page/v4i4D>