Pengantar SQL untuk BigQuery dan Cloud SQL

BigQuery adalah data warehouse berskala petabyte terkelola sepenuhnya yang berjalan di Google Cloud. Analis data dan data scientist dapat dengan cepat mengkueri dan memfilter set data besar, menggabungkan hasil, dan melakukan operasi kompleks tanpa harus khawatir tentang penyiapan dan pengelolaan server. BigQuery hadir dalam bentuk alat command line (sudah terinstal dalam Cloud Shell) atau konsol web. Keduanya siap digunakan untuk mengelola dan mengkueri data yang disimpan dalam project Google Cloud.

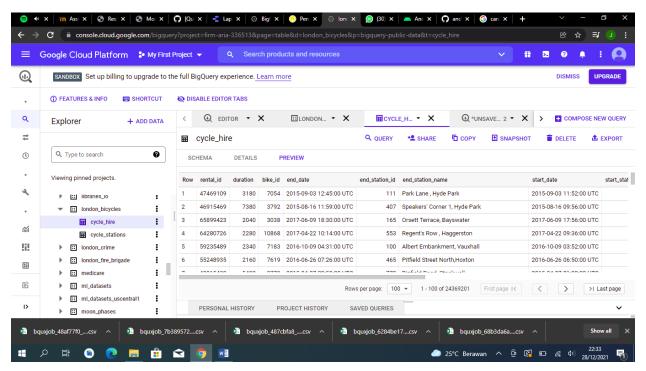
Dengan menarik beberapa data publik ke project sehingga bisa berlatih menjalankan perintah SQL di BigQuery.

Klik ADD DATA > Explore Public Datasets

Search "london", kemudian pilih kotak London Bicycle Hires,

Project baru bernama bigquery-public-data

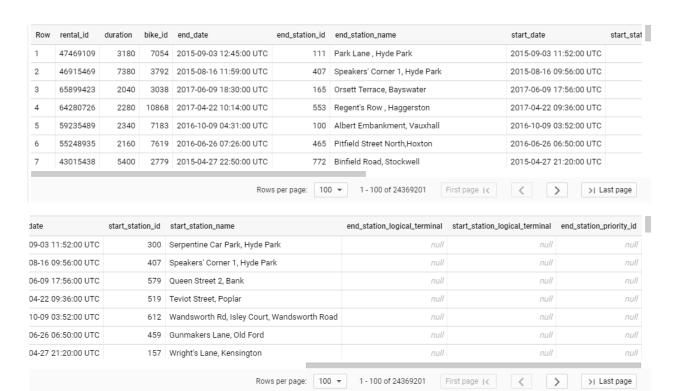
Luaskan project **bigquery-public-data** > cari data sets **london_bicycles**, lalu pilih tabel **cycle_hire**



Tabel cycle hire memiliki 24369201 baris dan memiliki 12 kolom sebagai berikut :

- 1. Row
- rental id

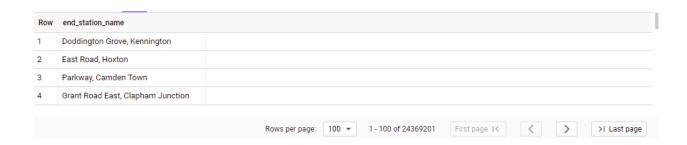
- 3. duration
- 4. bike idend date
- 5. end station id
- 6. end station name
- 7. start_date
- 8. start_station_id
- 9. start_station_name
- 10. end_station_logical_terminal
- 11. start_station_logical_terminal
- 12. end_station_priority_id



1. SELECT, FROM, dan WHERE

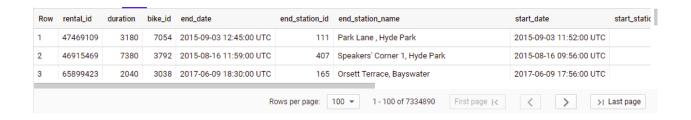
SELECT kolom end_station_name

SELECT end station name FROM 'bigquery-public-data.london bicycles.cycle hire';



SELECT * menampilkan semua nilai kolom dari tabel. Durasi diukur dalam detik, itulah sebabnya Anda menggunakan nilai 1200 (60 * 20).

SELECT * FROM `bigquery-public-data.london_bicycles.cycle_hire` WHERE duration>=1200;



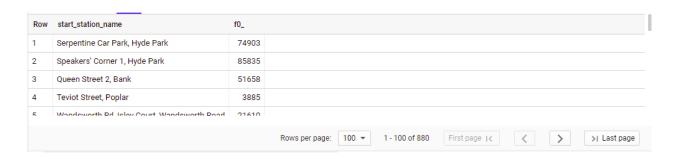
2. GROUP BY

SELECT start_station_name FROM `bigquery-public-data.london_bicycles.cycle_hire` GROUP BY start_station_name;



3. COUNT

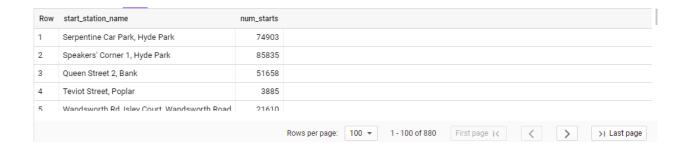
SELECT start_station_name, COUNT(*) FROM `bigquery-public-data.london_bicycles.cycle_hire` GROUP BY start_station_name;



4. AS

Membuat alias dari tabel atau kolom

SELECT start_station_name, COUNT(*) AS num_starts FROM `bigquery-public-data.london bicycles.cycle hire` GROUP BY start station name;



ORDER BY

mengurutkan data yang dihasilkan dari kueri dalam urutan naik atau turun berdasarkan kriteria atau nilai kolom yang ditentukan.

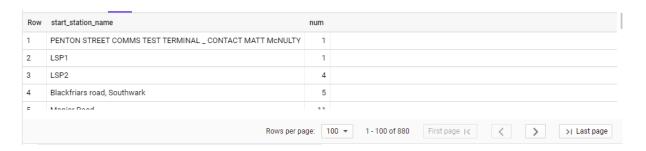
 Menampilkan tabel berisi jumlah perjalanan bikeshare yang dimulai di setiap stasiun awal, yang disusun secara alfabetis berdasarkan stasiun awal.

SELECT start_station_name, COUNT(*) AS num FROM `bigquery-public-data.london_bicycles.cycle_hire` GROUP BY start_station_name ORDER BY start_station_name;



 Menampilkan tabel berisi jumlah perjalanan bikeshare yang dimulai di setiap stasiun awal, yang disusun secara numerik dari terendah ke tertinggi.

SELECT start_station_name, COUNT(*) AS num FROM `bigquery-public-data.london_bicycles.cycle_hire` GROUP BY start_station_name ORDER BY num;



 Menampilkan tabel berisi jumlah perjalanan bikeshare yang dimulai di setiap stasiun awal, yang disusun secara numerik dari tertinggi ke terendah.

SELECT start_station_name, COUNT(*) AS num FROM `bigquery-public-data.london_bicycles.cycle_hire` GROUP BY start_station_name ORDER BY num DESC;



 Menampilkan tabel berisi jumlah tabel perjalanan bikeshare yang berakhir di setiap stasiun akhir, yang disusun secara numerik dari jumlah perjalanan tertinggi ke terendah

SELECT end_station_name, COUNT(*) AS num FROM `bigquery-public-data.london_bicycles.cycle_hire` GROUP BY end_station_name ORDER BY num DESC;

Row	end_station_name	num				
1	Belgrove Street , King's Cross	231802				
2	Hyde Park Corner, Hyde Park	215038				
3	Waterloo Station 3, Waterloo	193200				
4	Albert Gate, Hyde Park	157943				