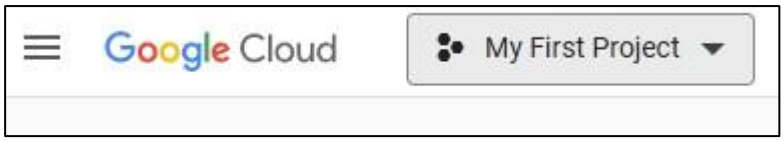
DATA WAREHOUSE

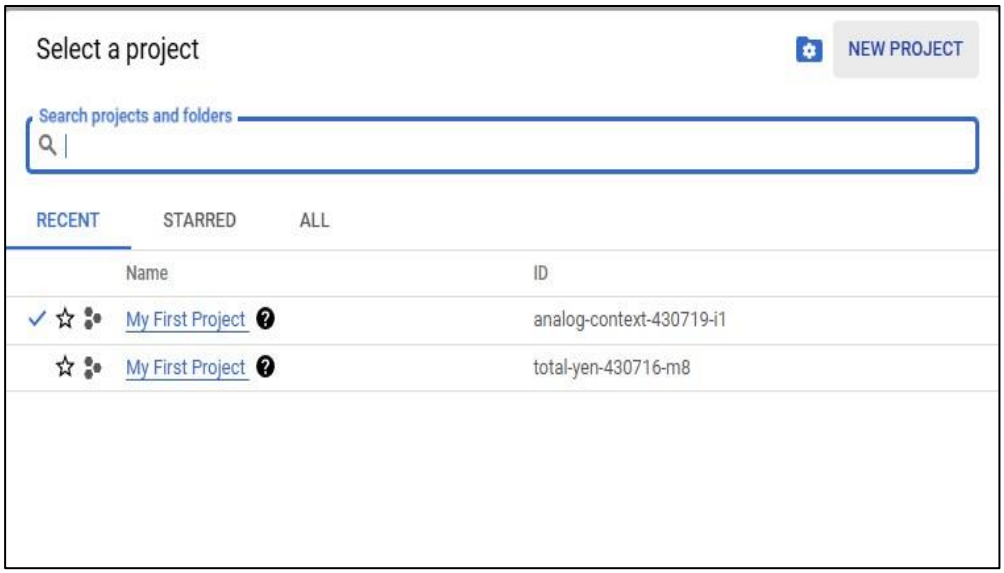
PART 4 – Big Query

TASK

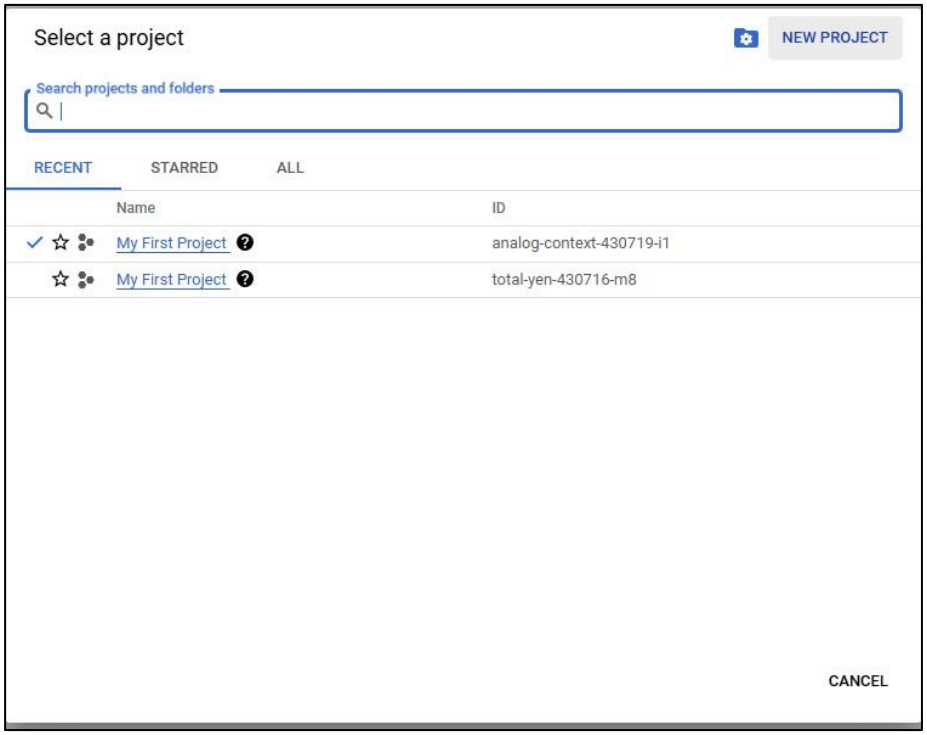
1. Buat sebuah bucket di GCP, upload beberapa file ke bucket tersebut!
   Jawab:
   a. Klik My First Project



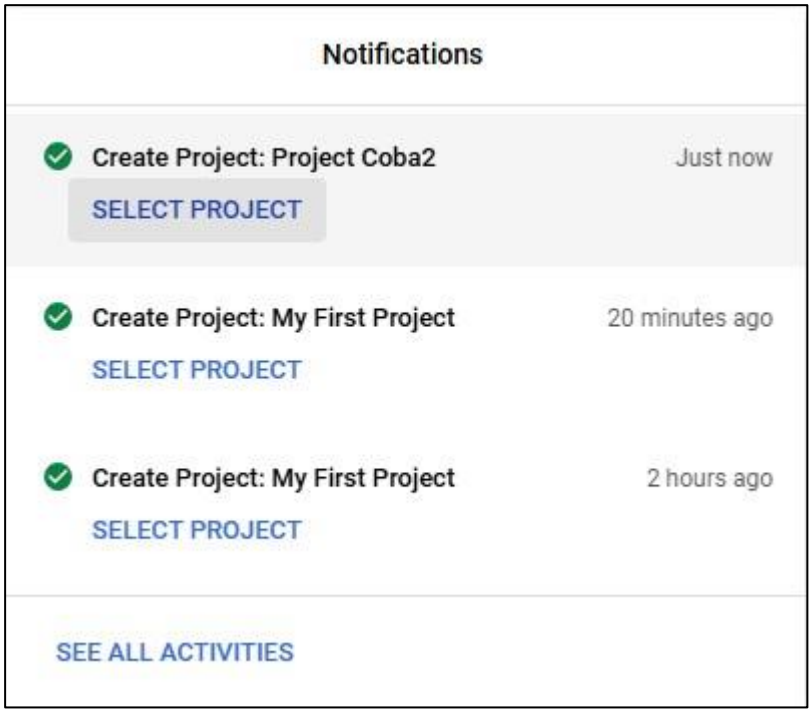
   b. Klik New Project



c. Ganti project name

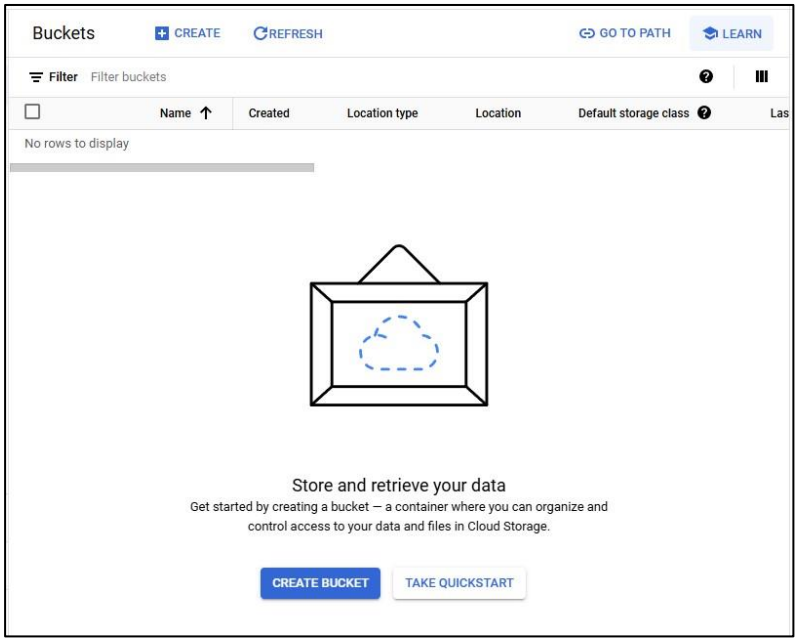


d. Select project yang telah dibuat

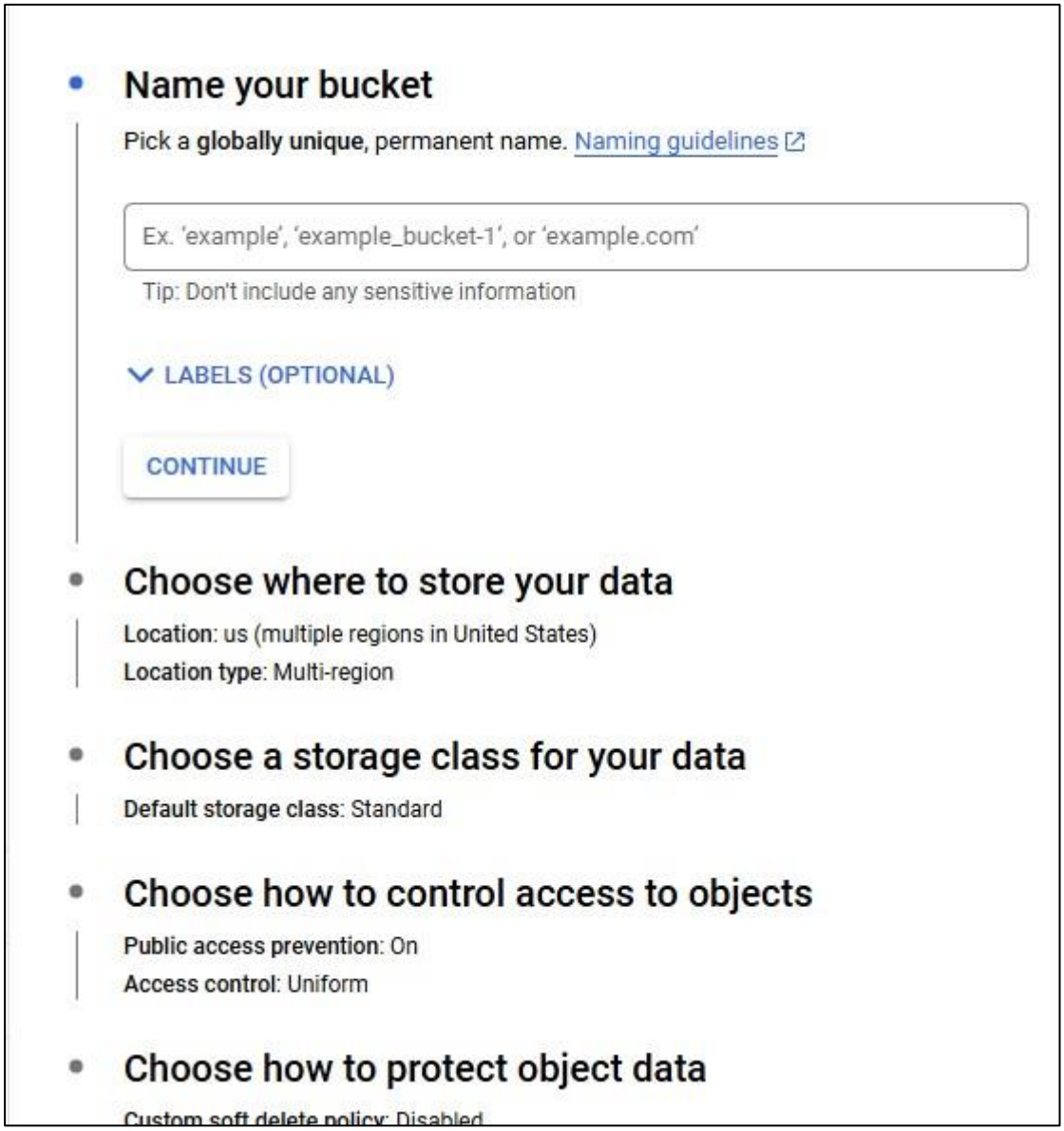


e. Search bucket di search bar dan pilih bucket

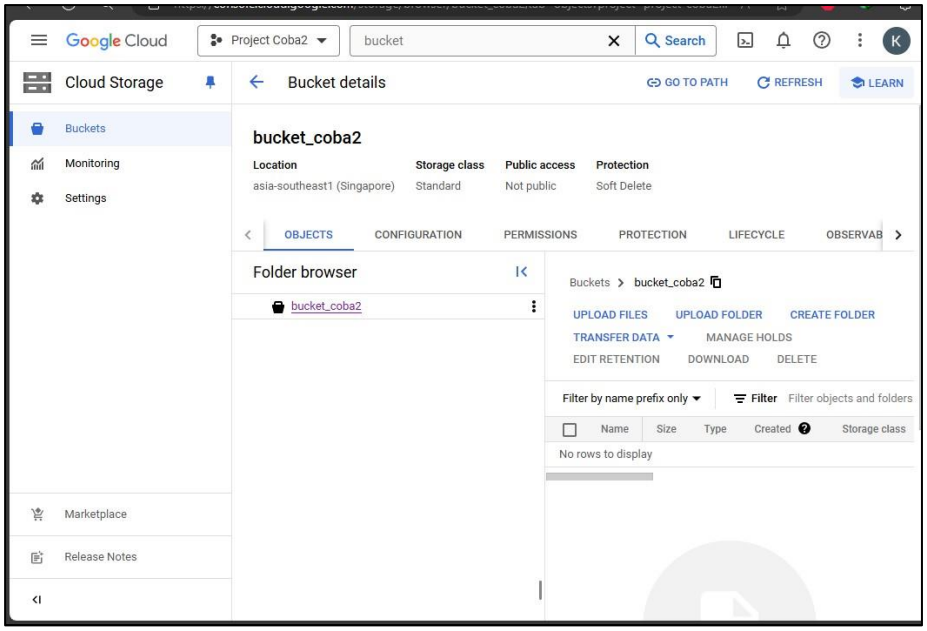
f. Pilih create bucket



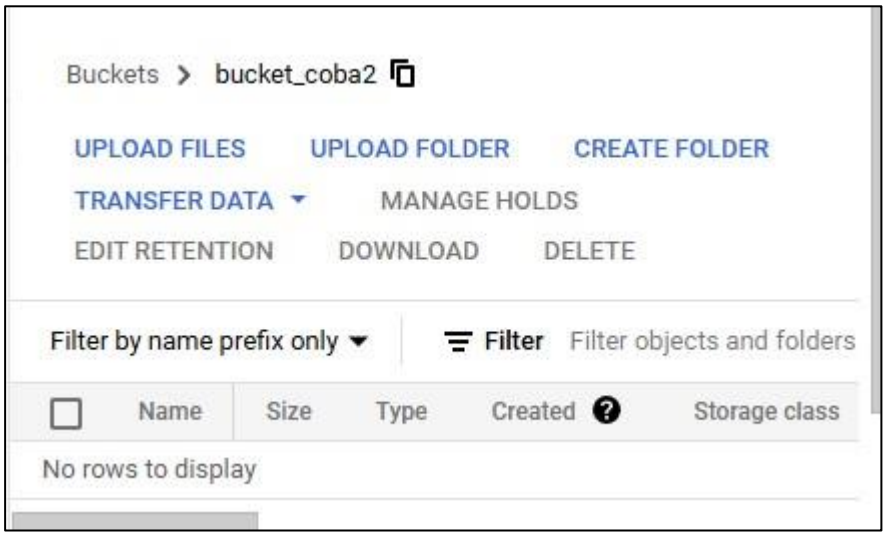
g. Buat nama bucket, pilih lokasi penyimpanan, dan atur sesuai kebutuhan



h. Seperti ini

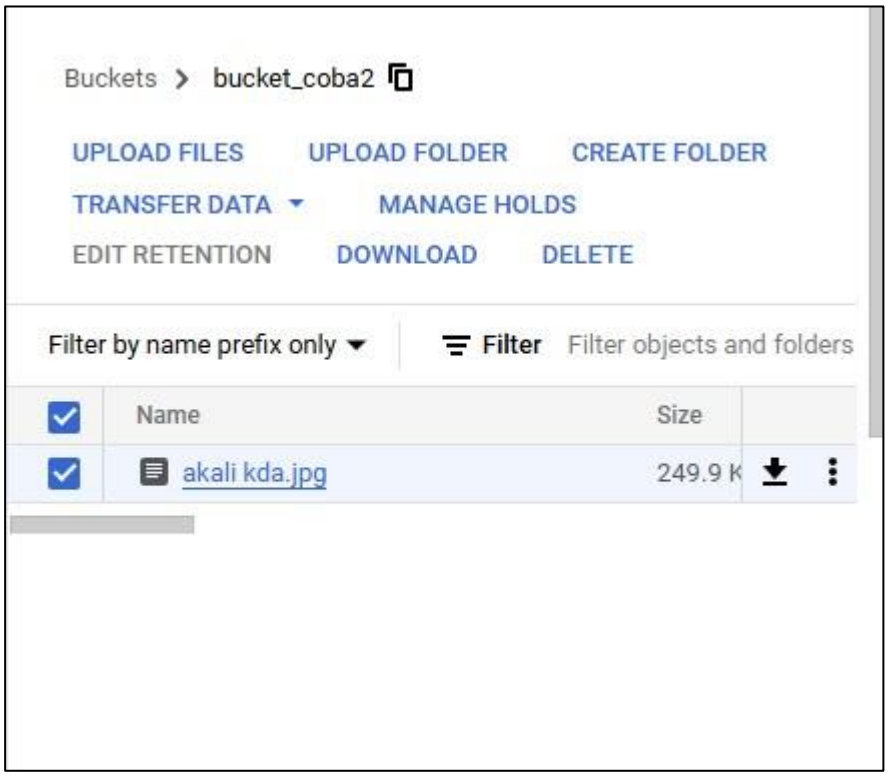


i. Pilih upload files

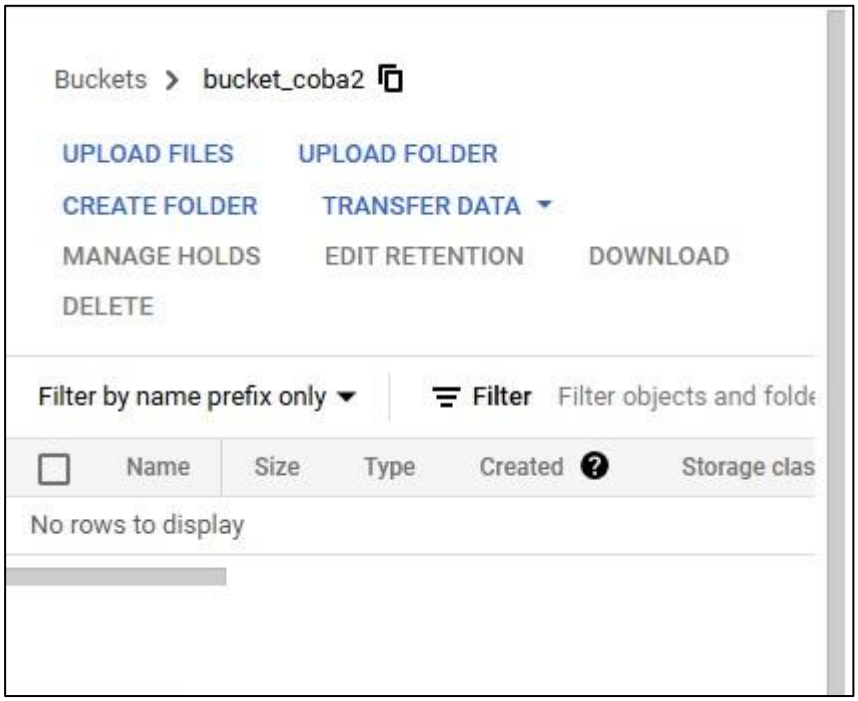


j. Sudah terupload

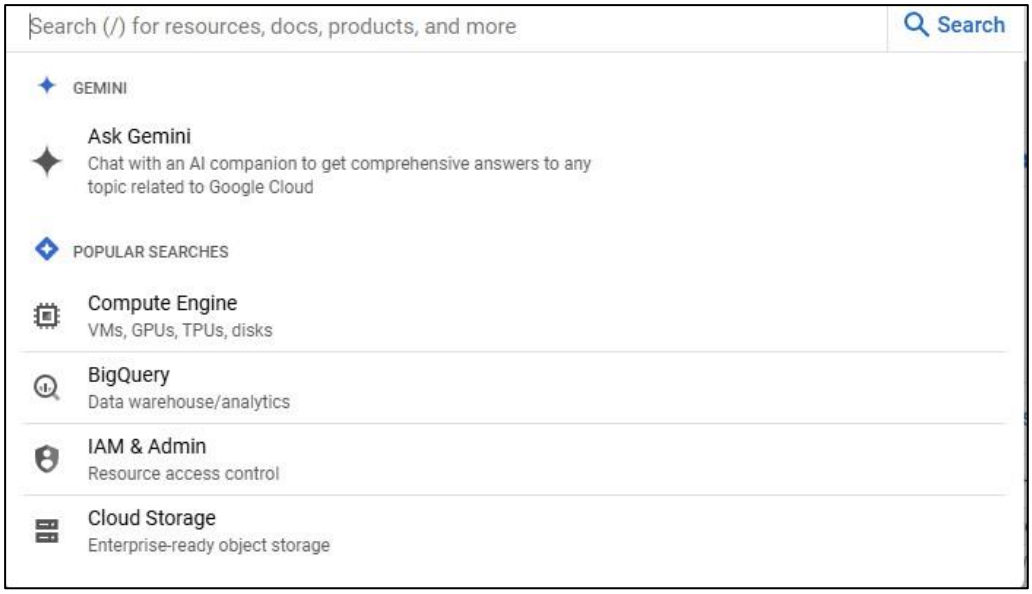
2. Hapus file yang sudah diupload!
   Jawab:
   a. Pilih file yang telah teruplod dan pilih delete



   b. Hasilnya file akan terhapus



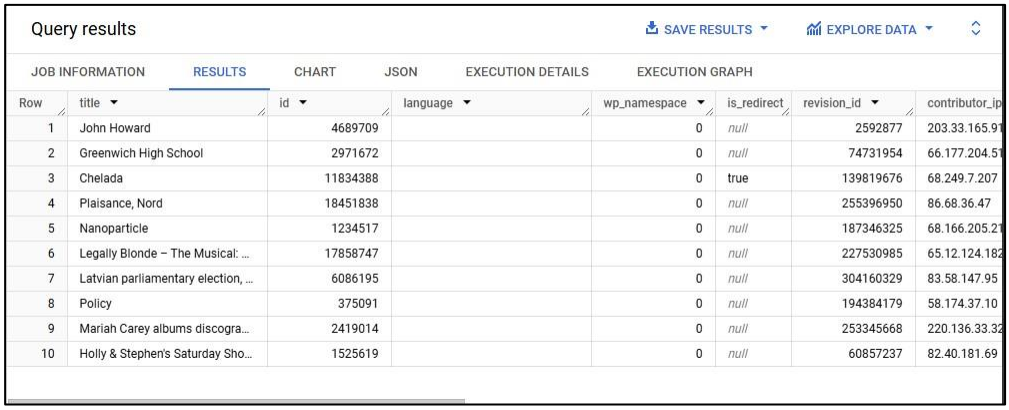
3. Lakukan eksplorasi sample data wikipedia (https://console.cloud.google.com/bigquery?p=bigquery-public-data&d=samples&t=wi kipedia&page=table) dengan menggunakan big query!
   Jawab:
   a. Pilih search bar dan pilih Big Query



   b. Jalankan query

   c. Jangan lupa di-save

d. Hasilnya



Query results

SAVE RESULTS | EXPLORE DATA

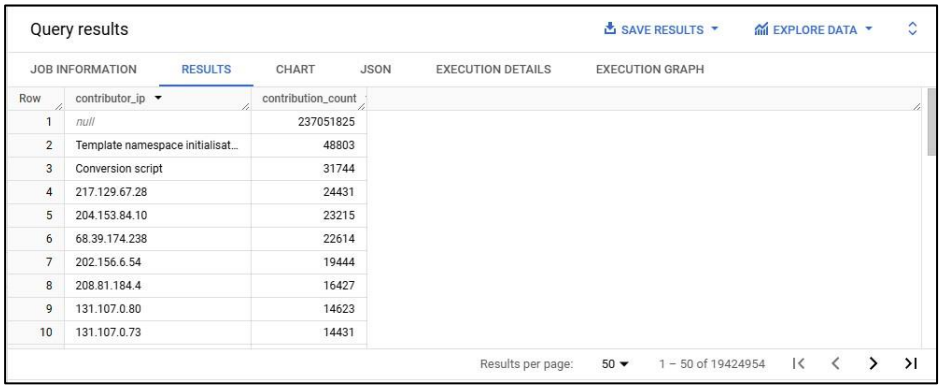
JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	title	id	language	wp_namespace	is_redirect	revision_id	contributor_ip
1	John Howard	4689709		0	null	2592877	203.33.165.9
2	Greenwich High School	2971672		0	null	74731954	66.177.204.5
3	Chelada	11834388		0	true	139819676	68.249.7.207
4	Plaisance, Nord	18451838		0	null	255396950	86.68.36.47
5	Nanoparticle	1234517		0	null	187346325	68.166.205.2
6	Legally Blonde – The Musical: ...	17858747		0	null	227530985	65.12.124.182
7	Latvian parliamentary election, ...	6086195		0	null	304160329	83.58.147.95
8	Policy	375091		0	null	194384179	58.174.37.10
9	Mariah Carey albums discogra...	2419014		0	null	253345668	220.136.33.32
10	Holly & Stephen's Saturday Sho...	1525619		0	null	60857237	82.40.181.69

4. Munculkan jumlah kontribusi dari masing-masing contributor_ip, urutkan dari kontribusi terbesar ke kontribusi terkecil!
   Jawab:
   a. Memunculkan kontributor IP

```
5  SELECT contributor_ip, COUNT(*) as contribution_count
6  FROM `bigquery-public-data.samples.wikipedia`
7  GROUP BY contributor_ip
8  ORDER BY contribution_count DESC;
9
```

   b. Hasilnya



Query results

SAVE RESULTS | EXPLORE DATA

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	contributor_ip	contribution_count
1	null	237051825
2	Template namespace initialisat...	48803
3	Conversion script	31744
4	217.129.67.28	24431
5	204.153.84.10	23215
6	68.39.174.238	22614
7	202.156.6.54	19444
8	208.81.184.4	16427
9	131.107.0.80	14623
10	131.107.0.73	14431

Results per page: 50 | 1 – 50 of 19424954