

bantu saya Saya sedang menggunakan BigQuery untuk menganalisis data penjualan dan produksi produk dari dua tabel yang telah saya unggah, yaitu tabel_produksi dan tabel_penjualan, yang ada dalam dataset perusahaan di project saya bernama margaretha-daraayu-damai. Saya ingin mencari id_produk dengan jumlah penjualan tertinggi selama 6 bulan pertama tahun 2024, yaitu dari tanggal 1 Januari hingga 15 Desember 2023. Kriteria lainnya adalah id_produksi nya bernomor genap. Catatan untuk membantu anda mengerjakan: Kolom id_produk, id_produksi, dan tanggal produksi ada di tabel_produksi, Kolom jumlah_terjual ada di tabel_penjualan Untuk cari penjualan terbesar: jumlah_terjual nya bisa diurutkan saja dari yang terbesar ke terkecil Kolom yang dipakai hanya id_produk, id_produksi, tanggal_produksi, dan jumlah terjual saja Untuk menghubungkan 2 tabel tersebut menggunakan foreign key, foreign key nya yaitu kolom id_produk Dapatkah Anda memberikan query SQL khusus BigQuery untuk menjalankan analisis ini?

ChatGPT said:

ChatGPT

Berikut adalah query SQL untuk memenuhi kebutuhan Anda menggunakan BigQuery:

```
sql
Copy code
-- Query untuk mencari id_produk dengan jumlah penjualan tertinggi
WITH penjualan_6_bulan_pertama AS (
  SELECT
    p.id_produk,
    tp.id_produksi,
    tp.tanggal_produksi,
    SUM(tp.jumlah_terjual) AS total_penjualan
  FROM
    `margaretha-daraayu-damai`.`table_penjualansel`
)
```

4o

You said:

tabel_penjualan.csv

Spreadsheet

tabel_produksi.csv

Spreadsheet

bantu saya Saya sedang menggunakan BigQuery untuk menganalisis data penjualan dan produksi produk dari dua tabel yang telah saya unggah, yaitu tabel_produksi dan tabel_penjualan, yang ada dalam dataset perusahaan di project saya bernama margaretha-daraayu-damai. Saya ingin mencari id_produk dengan jumlah penjualan tertinggi selama 6 bulan pertama tahun 2024, yaitu dari tanggal 1 Januari hingga 15 Desember 2023. Kriteria lainnya adalah id_produksi nya bernomor genap. Catatan untuk membantu anda mengerjakan: Kolom id_produk, id_produksi, dan tanggal produksi ada di tabel_produksi, Kolom jumlah_terjual ada di tabel_penjualan Untuk cari penjualan terbesar: jumlah_terjual nya bisa diurutkan saja dari yang terbesar ke terkecil Kolom yang dipakai hanya id_produk, id_produksi, tanggal_produksi, dan jumlah terjual saja Untuk menghubungkan 2 tabel tersebut menggunakan foreign key, foreign key nya yaitu kolom id_produk Dapatkah Anda memberikan query SQL khusus BigQuery untuk menjalankan analisis ini?

ChatGPT said:

ChatGPT

Berikut adalah query SQL untuk kebutuhan analisis Anda di BigQuery. Query ini akan mencari produk dengan jumlah penjualan tertinggi selama periode 6 bulan pertama 2024, dengan kriteria tambahan bahwa `id_produk` harus bernomor genap.

Namun sebelum itu, saya akan memeriksa isi file `tabel_produk.csv` dan `tabel_penjualan.csv` untuk menyesuaikan query dengan struktur kolomnya. Mari kita lihat terlebih dahulu.

Dari data yang diunggah, struktur kolomnya adalah sebagai berikut:

Tabel Penjualan (`tabel_penjualan.csv`):

- `penjualan_id`
- `tanggal_penjualan`
- `barang_id` (kolom ini merupakan `id_produk`)
- `jumlah_terjual`
- `harga_penjualan`
- `pelanggan_id`

Tabel Produksi (`tabel_produk.csv`):

- `produksi_id` (kolom ini adalah `id_produk`)
- `tanggal_produk`
- `barang_id` (kolom ini merupakan `id_produk`)
- `jumlah_diproduksi`
- `biaya_produk`

Query SQL berikut dapat digunakan untuk melakukan analisis sesuai kebutuhan Anda:

```
sql
Copy code
SELECT
  tp.barang_id AS id_produk,
  tp.produksi_id AS id_produk,
  tp.tanggal_produk,
  SUM(tj.jumlah_terjual) AS total_penjualan
FROM
  `margaretha-daraayu-damai.perusahaan.tabel_produk` tp
JOIN
  `margaretha-daraayu-damai.perusahaan.tabel_penjualan` tj
ON
  tp.barang_id = tj.barang_id
WHERE
  DATE(tj.tanggal_penjualan) BETWEEN '2024-01-01' AND '2024-06-30'
  AND MOD(tp.produksi_id, 2) = 0
GROUP BY
  tp.barang_id, tp.produksi_id, tp.tanggal_produk
ORDER BY
  total_penjualan DESC
LIMIT 1;
```

Penjelasan:

1. **Join Tabel:** Tabel `tabel_produksi` (tp) dan `tabel_penjualan` (tj) dihubungkan berdasarkan `barang_id` (yang setara dengan `id_produk`).
2. **Filter Waktu:** Hanya data dari 1 Januari hingga 30 Juni 2024 yang digunakan.
3. **ID Produksi Genap:** Menggunakan `MOD(tp.produksi_id, 2) = 0` untuk memfilter `id_produksi` bernomor genap.
4. **Pengelompokan:** Data dikelompokkan berdasarkan `barang_id`, `produksi_id`, dan `tanggal_produksi`.
5. **Urutan Penjualan:** Hasil diurutkan dari penjualan terbesar ke terkecil.
6. **Pembatasan Hasil:** Menggunakan `LIMIT 1` untuk mendapatkan produk dengan penjualan tertinggi.

Jika ada kebutuhan tambahan atau perubahan, beri tahu saya!

4o

You said:

coba selesaikan masalah ini No matching signature for function MOD for argument types: FLOAT64, INT64. Supported signatures: MOD(INT64, INT64); MOD(NUMERIC, NUMERIC); MOD(BIGNUMERIC, BIGNUMERIC) at [14:7]

ChatGPT said:

ChatGPT

Masalah terjadi karena kolom `produksi_id` dalam tabel memiliki tipe data `FLOAT64`. Fungsi `MOD` tidak mendukung operasi dengan tipe data `FLOAT64`, tetapi hanya bekerja dengan tipe data `INT64`, `NUMERIC`, atau `BIGNUMERIC`.

Untuk memperbaikinya, kita perlu mengonversi tipe data `produksi_id` menjadi bilangan bulat (`INT64`) sebelum menggunakan fungsi `MOD`. Berikut adalah perbaikan query Anda:

```
sql
Copy code
SELECT
    tp.barang_id AS id_produk,
    tp.produksi_id AS id_produksi,
    tp.tanggal_produksi,
    SUM(tj.jumlah_terjual) AS total_penjualan
FROM
    `margaretha-daraayu-damai.perusahaan.tabel_produksi` tp
JOIN
    `margaretha-daraayu-damai.perusahaan.tabel_penjualan` tj
ON
    tp.barang_id = tj.barang_id
WHERE
    DATE(tj.tanggal_penjualan) BETWEEN '2024-01-01' AND '2024-06-30'
    AND MOD(CAST(tp.produksi_id AS INT64), 2) = 0
GROUP BY
    tp.barang_id, tp.produksi_id, tp.tanggal_produksi
ORDER BY
    total_penjualan DESC
```

```
LIMIT 1;
```

Perubahan:

1. **Konversi Tipe Data:** Menggunakan `CAST(tp.produksi_id AS INT64)` untuk mengonversi `produksi_id` dari `FLOAT64` menjadi `INT64`.
2. **Penerapan MOD:** Setelah konversi, fungsi `MOD` dapat digunakan tanpa masalah.

Coba jalankan query ini, dan beri tahu saya jika ada kendala lain!

4o

You said:

coba selesaikan masalah ini Syntax error: Unexpected keyword DATE at [20:3]

ChatGPT said:

ChatGPT

Masalah terjadi karena fungsi `DATE()` di BigQuery digunakan untuk mengonversi tipe data menjadi `DATE`, tetapi jika kolom Anda sudah berupa tipe data `DATE` atau `DATETIME`, penggunaannya tidak diperlukan.

Jika kolom `tanggal_penjualan` sudah berupa string atau `DATE`, berikut adalah perbaikan query Anda:

```
sql
Copy code
SELECT
  tp.barang_id AS id_produk,
  tp.produksi_id AS id_produksi,
  tp.tanggal_produksi,
  SUM(tj.jumlah_terjual) AS total_penjualan
FROM
  `margaretha-daraayu-damai.perusahaan.tabel_produksi` tp
JOIN
  `margaretha-daraayu-damai.perusahaan.tabel_penjualan` tj
ON
  tp.barang_id = tj.barang_id
WHERE
  PARSE_DATE('%Y-%m-%d', tj.tanggal_penjualan) BETWEEN DATE('2024-01-01')
AND DATE('2024-06-30')
  AND MOD(CAST(tp.produksi_id AS INT64), 2) = 0
GROUP BY
  tp.barang_id, tp.produksi_id, tp.tanggal_produksi
ORDER BY
  total_penjualan DESC
LIMIT 1;
```

Perubahan:

1. **Menggunakan PARSE_DATE:** Jika kolom `tanggal_penjualan` berupa string dalam format `YYYY-MM-DD`, fungsi `PARSE_DATE` mengonversinya menjadi tipe `DATE`.

- o Format %Y-%m-%d sesuai dengan format tanggal dalam data Anda.
2. **Menghapus DATE ()** : Fungsi DATE () hanya digunakan jika Anda ingin mendeklarasikan literal tanggal (seperti DATE ('2024-01-01')) atau konversi eksplisit.

Coba query ini, dan beri tahu saya jika masih ada masalah!

The image shows two screenshots of a Google Cloud environment. The top screenshot is the BigQuery console, displaying a SQL query and its results. The bottom screenshot is the Pelaporan Looker Studio dashboard, showing a report titled 'BigQuery SQL kustom' with a bar chart and a data panel.

BigQuery Query and Results:

```
1 SELECT
2   tp.barang_id AS id_produk,
3   tp.produksi_id AS id_produksi,
4   tp.tanggal_produksi,
5   SUM(tj.jumlah_terjual) AS total_penjualan
6 FROM
7   `margaretha-daraayu-damai.Company.tabel_produksi` tp
8 JOIN
9   `margaretha-daraayu-damai.Company.tabel_penjualan` tj
10  ON
11   tp.barang_id = tj.barang_id
12 WHERE
```

Row	id_produk	id_produksi	tanggal_produksi	total_penjualan
1	1004.0	3.0	2023-10-14	248.0
2	1004.0	15.0	2023-02-02	248.0
3	1002.0	16.0	2023-11-21	239.0
4	1002.0	17.0	2023-03-26	239.0

Pelaporan Looker Studio Dashboard:

The dashboard is titled 'BigQuery SQL kustom' and shows a bar chart with the following data:

id_produk	Record Count
3	1
15	1
7	1
11	1
12	1
1	1
9	1
12	1

The data panel on the right lists the following fields:

- id_produk
- id_produksi
- tanggal_produksi
- total_penjualan
- Record Count