

Nama : Nissi Goklas Simamora
NIM : 40011423630231
Kelas : E
Mata Kuliah : Sistem Informasi Akuntansi

TUGAS DATA ANALISIS – Tahapan Membuat Project Dataset Perusahaan di BigQuery

1. Tahap pertama yang dilakukan yaitu meminta bantuan ChatGPT untuk membuat database perusahaan dengan tiga tabel(tabel produksi, tabel persediaan, dan tabel penjualan) yang masing- masing dilengkapi dengan struktur tabelnya dan tunjukkan satu primary key yang bisa digunakan untuk berinteraksi dengan tiga tabel tersebut dalam analitik.

Buatlah database perusahaan dengan tiga tabel(tabel produksi, tabel persediaan, dan tabel penjualan) masing - masing dilengkapi dengan struktur tabelnya dan tunjukkan satu primary key yang bisa digunakan untuk berinteraksi dengan tiga tabel tersebut dalam analitik!

Setelah memberikan perintah tersebut, maka ChatGPT akan memberikan output seperti berikut ini:

Berikut adalah contoh struktur database perusahaan dengan tiga tabel: tabel produksi, tabel persediaan, dan tabel penjualan. Setiap tabel akan memiliki struktur yang memungkinkan data terintegrasi dengan baik untuk analisis. Kita akan menggunakan kolom `product_id` sebagai *primary key* untuk berinteraksi dengan ketiga tabel tersebut.

Struktur Database Perusahaan

1. Tabel `produksi`

Tabel ini menyimpan informasi terkait produksi produk.

Nama Kolom	Tipe Data	Keterangan
production_id	INT	Primary Key, ID produksi
product_id	INT	Foreign Key ke <code>persediaan</code>
production_date	DATE	Tanggal produksi
produced_qty	INT	Jumlah barang yang diproduksi
production_cost	DECIMAL(10, 2)	Biaya produksi per unit
batch_number	VARCHAR(50)	Nomor batch produksi

2. Tabel `persediaan`

Tabel ini menyimpan informasi terkait stok dan persediaan produk.

Nama Kolom	Tipe Data	Keterangan
inventory_id	INT	Primary Key, ID persediaan
product_id	INT	Unique, ID produk
product_name	VARCHAR(100)	Nama produk
stock_qty	INT	Jumlah barang di persediaan
reorder_level	INT	Level stok minimal sebelum restock
purchase_price	DECIMAL(10, 2)	Harga beli per unit

3. Tabel penjualan

Tabel ini menyimpan informasi terkait penjualan produk.

Nama Kolom	Tipe Data	Keterangan
sales_id	INT	Primary Key, ID penjualan
product_id	INT	Foreign Key ke persediaan
sales_date	DATE	Tanggal penjualan
sold_qty	INT	Jumlah barang yang terjual
sales_price	DECIMAL(10, 2)	Harga jual per unit
customer_id	INT	ID pelanggan

2. Tahap kedua yaitu meminta ChatGPT untuk membuat tabel tersebut menjadi excel dimana masing-masing tabel terdiri dari 20 baris.

Buatlah tabel di atas ke dalam bentuk excel dengan masing-masing tabel terdiri dari 20 baris!

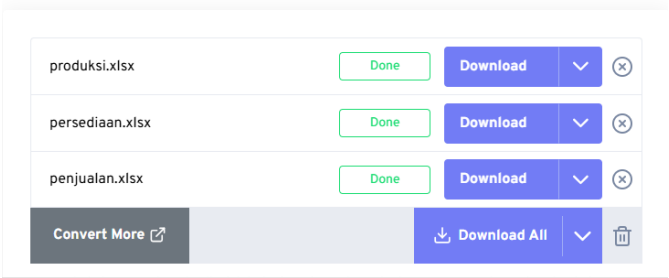


Berikut adalah tautan unduhan untuk masing-masing file Excel:

- 1. [Download Tabel Produksi](#)
- 2. [Download Tabel Persediaan](#)
- 3. [Download Tabel Penjualan](#)

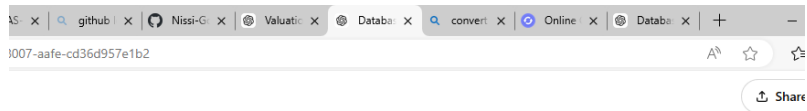
Setiap file Excel berisi tabel sesuai dengan data yang diminta. [↗]

3. Tahap ketiga yaitu meng-convert file excel yang telah di unduh tadi menjadi file csv melalui website convert file



4. Tahap keempat yaitu meminta bantuan ChatGPT untuk membuat pertanyaan dari tabel tersebut dengan beberapa kriteria.

Buatlah pertanyaan dari product_id yang tingkat penjualannya tertinggi tetapi production_date nya dimulai dari tanggal 1 Januari 2024 sampai 5 Februari 2024 untuk product_id yang bernomor genap!

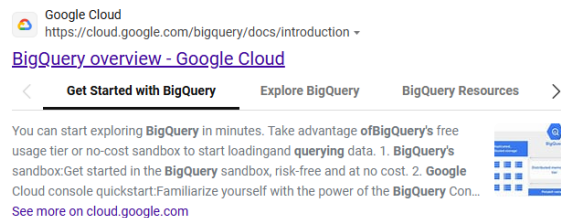


Berikut adalah pertanyaan analitik yang sesuai dengan kriteria yang diberikan:

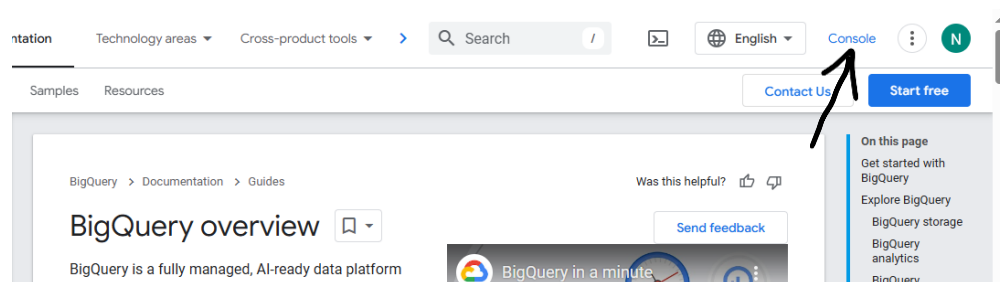
1. **Produk dengan `product_id` Bernomor Genap dan Tingkat Penjualan Tertinggi:**
 - Produk manakah (`product_id`) yang memiliki jumlah penjualan tertinggi (`sold_qty`)?
 - Tanggal produksi harus berada dalam rentang 1 Januari 2024 hingga 5 Februari 2024.
 - Hanya produk dengan `product_id` genap yang dipertimbangkan.
2. **Detail Produksi dan Penjualan Produk Tersebut:**
 - Apa saja detail produksi, seperti `production_date`, `produced_qty`, dan `production_cost` untuk produk tersebut?
 - Berapa jumlah produk yang dijual (`sold_qty`) dan pada harga berapa (`sales_price`)?
3. **Tanggal Produksi dan Penjualan:**
 - Kapan produk tersebut diproduksi dan dijual?

5. Tahap kelima yaitu membuat akun di Google Big Query. Adapun langkah-langkahnya adalah sebagai berikut:

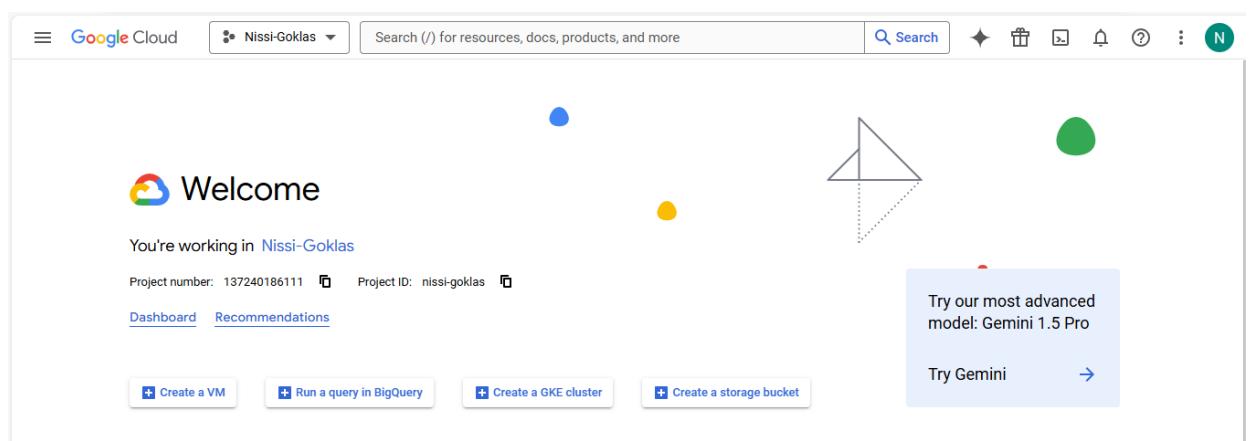
- Buka aplikasi google/chrome lalu ketik “google big query” kemudian klik BigQuery Overview



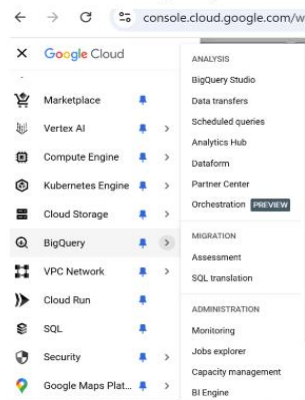
- Kemudian klik “console” yang ada di pojok kanan atas



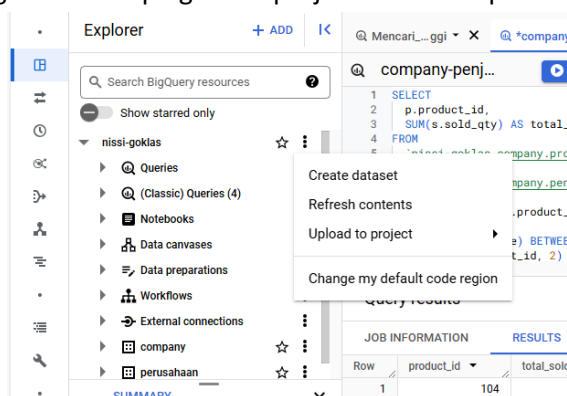
- Maka akan keluar tampilan seperti berikut:



- Langkah selanjutnya yaitu klik garis 3 yang ada di pojok kiri atas, pilih BigQuery, lalu klik BigQuery Studio



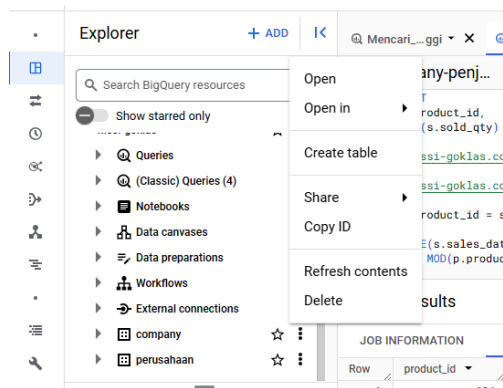
- Setelah langkah-langkah tersebut dilakukan, kemudian pilih create project dan isi nama project. Saya menamai project saya dengan nama **nissi-goklas**. Penggunaan tanda hubung dalam nama project sebagai pengganti spasi yang tidak bisa digunakan dalam dataset. Bagian location tidak perlu diisi atau dengan kata lain dikosongkan saja
- Setelah berhasil membuat nama project, maka nama project kita akan muncul. Klik titik tiga yang ada di samping nama project kemudian pilih create dataset.



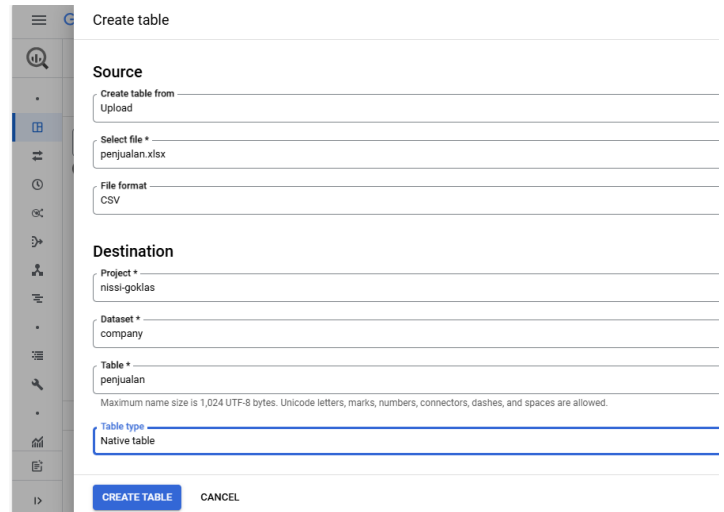
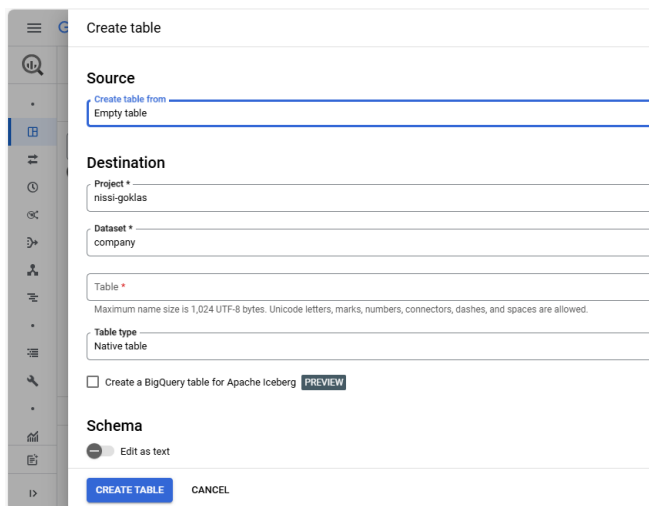
- Setelah itu, isi Dataset ID (sebaiknya dengan huruf kecil semua). Disini saya membuat 2 dataset yaitu perusahaan dan company. Untuk location type nya dipilih yang multi region saja. Kemudian klik create dataset.

 A screenshot of the 'Create dataset' dialog box in BigQuery. It shows the 'Project ID' as 'nissi-goklas'. The 'Dataset ID' field is empty, with a note that letters, numbers, and underscores are allowed. Under 'Location type', 'Multi-region' is selected. A dropdown shows 'US (multiple regions in United States)'. There are checkboxes for 'Link to an external dataset' and 'Enable table expiration'. At the bottom, there are 'CREATE DATASET' and 'CANCEL' buttons.

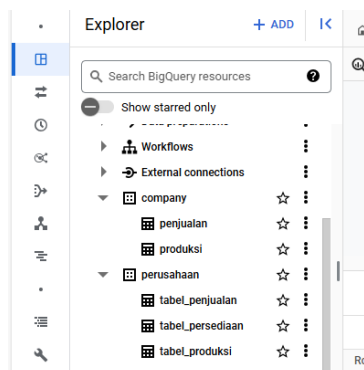
9. Setelah berhasil membuat dataset, maka dataset akan muncul sesuai dengan Dataset ID yang telah dibuat tadi. Kemudian klik titik tiga yang ada di samping dataset (company/perusahaan), klik create table.



10. Langkah selanjutnya yaitu mengisi source table dari create table yang telah di klik tadi. Source disini berasal dari tabel yang telah kita buat sebelumnya dengan bantuan ChatGPT yaitu tabel produksi, persediaan, dan penjualan. Namun yang diupload ke source ini bukan lah file excel melainkan file yang telah kita convert menjadi file csv. Di bagian bawah klik auto detect kemudian create table.



11. Setelah berhasil melakukan create table, maka akan muncul seperti gambar berikut:



12. Selanjutnya yaitu membuat Query SQL dengan bantuan ChatGPT. Adapun perintahnya adalah sebagai berikut: (file yang dilampirkan hanya file csv penjualan dan produksi)

produksi.csv
Spreadsheet

penjualan.csv
Spreadsheet

Baca detail 2 file yang saya lampirkan!

Setelah itu bantu saya,

Saya sedang menggunakan BigQuery untuk menganalisis data penjualan dan produksi produk dari dua tabel yang telah saya unggah, yaitu produksi dan penjualan, yang ada dalam dataset company di project saya bernama nissi-goklas. Saya ingin mencari product_id dengan jumlah penjualan tertinggi selama 6 bulan pertama tahun 2024, yaitu dari tanggal 1 Januari 2024 hingga 30 Juni 2024. Kriteria lainnya adalah product_id nya bernomor genap.

Catatan untuk membantu anda mengerjakan:

- Kolom product_id, production_id, dan production_date ada di produksi,
- Kolom sold_qty ada di penjualan
- Untuk cari penjualan terbesar: sold_qty nya bisa diurutkan saja dari yang terbesar ke terkecil
- Kolom yang dipakai hanya product_id, production_id, production_date, dan sold_qty saja
- Untuk menghubungkan 2 tabel tersebut menggunakan foreign key, foreign key nya yaitu kolom product_id

Dapatkan Anda memberikan query SQL khusus BigQuery untuk menjalankan analisis ini?

```
sql Copy code

WITH filtered_sales AS (
  SELECT
    p.product_id,
    SUM(s.sold_qty) AS total_sold_qty
  FROM
    `nissi-goklas.company.produksi` p
  JOIN
    `nissi-goklas.company.penjualan` s
  ON p.product_id = s.product_id
  WHERE
    DATE(s.sales_date) BETWEEN '2024-01-01' AND '2024-06-30'
    AND MOD(p.product_id, 2) = 0 -- Hanya product_id genap
  GROUP BY
    p.product_id
),
ranked_sales AS (
  SELECT
    product_id,
    total_sold_qty,
    RANK() OVER (ORDER BY total_sold_qty DESC) AS rank
  FROM
    filtered_sales
)
SELECT
  product_id,
  total_sold_qty
FROM
  ranked_sales
WHERE
  rank = 1;
```

13. Langkah selanjutnya yaitu copy SQL yang telah diberikan oleh ChatGPT kemudian paste ke BigQuery SQL. Setelah itu klik run, jika SQLnya sudah sesuai dengan isi data yang ada di tabel dataset kita, maka akan keluar result atau akan ada tanda cekli hijau di sudut kanan atas. Namun, jika SQL kita tidak sesuai dengan data yang ada di tabel dataset maka kan terjadi error.

Contoh jika SQL yang dimasukkan sudah benar dan sesuai

The screenshot shows the Google BigQuery interface. On the left is the Explorer panel with a search bar and a tree view of resources including 'nissi-goklas', 'Queries', 'Notebooks', 'Data canvases', 'Data preparations', 'Workflows', 'External connections', 'company', and 'perusahaan'. The main editor displays a SQL query named 'company-penjualan_tertinggi' with the following code:

```
1 SELECT
2   p.product_id,
3   SUM(s.sold_qty) AS total_sold
4 FROM
5   `nissi-goklas.company.produksi` p
6 JOIN
7   `nissi-goklas.company.penjualan` s
8 ON
9   p.product_id = s.product_id
10 WHERE
11   DATE(s.sales_date) BETWEEN '2024-01-01' AND '2024-02-28'
12   AND MOD(p.product_id, 2) = 0
```

Below the query editor, the 'Query results' section is visible, showing a table with columns 'product_id' and 'total_sold'. The first row shows '104' and '90'. A notification box at the bottom center states '"nissi simamora" deleted.' The bottom right corner indicates 'Results per page: 50' and '1 - 1 of 1'.

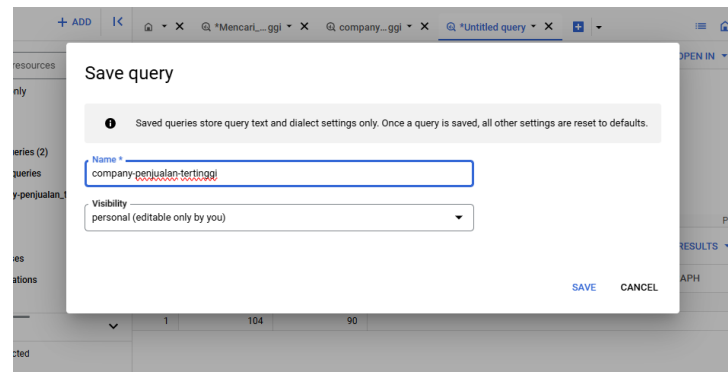
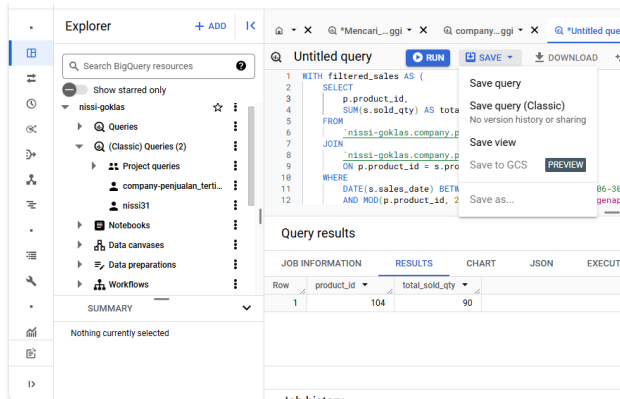
Contoh jika SQL yang dimasukkan salah atau tidak sesuai

The screenshot shows the Google BigQuery interface with a query named 'Untitled query'. The SQL code is as follows:

```
1 WITH filtered_sales AS (
2   SELECT
3     p.product_id,
4     SUM(s.sold_qty) AS total_sold_qty
5   FROM
6     `nissi-goklas.company.produksi` p
7   JOIN
8     `nissi-goklas.company.penjualan` s
9   ON p.product_id = s.product_id
10  WHERE
11    DATE(s.sales_date) BETWEEN '2024-01-01' AND '2024-06-30'
12    AND MOD(p.product_id, 2) = 0 -- Hanya product_id genap
```

The 'Query results' section shows an error message: 'Name product_id not found inside p at [3:11]'. The error is highlighted in a red box. The bottom right corner indicates 'Results per page: 50' and '1 - 1 of 1'.

14. Jika SQL yang dimasukkan sudah benar/sesuai, maka selanjutnya yaitu melakukan save. Klik save yang ada disamping tombol run, pilih save query classic. Kemudian isi nama sql (bisa disesuaikan tergantung pengerjaannya) lalu klik save dan sql kita akan tersimpan.



15. Jangan lupa untuk menyalin link console google cloud big query yang telah kita kerjakan tadi. Anda bisa menyimpan link tersebut di drive anda, agar nanti mudah untuk melanjutkan pekerjaan. Untuk melihat SQL yang telah kita kerja kan bisa di lihat di bagian (Classi) queries (2) seperti pada tampilan di bawah ini! Dan selesai.

