

Nama : Aida Alya Rahmadani  
NIM : 12030123130121  
Kelas : E  
Mata Kuliah : Pengkodean dan Pemrograman  
Dosen Pengampu : Dr. Totok Dewayanto, S.E., M.Si., Akt.

## PERTEMUAN KE-9

### PERINTAH:

Sebutkan urutan data analitik secara bertahap, dan buat tahapan tersebut untuk kepentingan pembuatan kueri sql berbasis google big query!

### JAWABAN:

Untuk keperluan pembuatan kueri SQL berbasis Google BigQuery, urutan data analitik secara bertahap dapat diuraikan sebagai berikut. Tahapan ini mencerminkan proses analisis data yang sistematis untuk menghasilkan kueri SQL yang efektif:

#### 1) Definisi Tujuan Analitik

- a) **Deskripsi:** Tentukan pertanyaan bisnis atau tujuan analisis, seperti "Berapa total penjualan per kategori produk dalam 3 bulan terakhir?" atau "Bagaimana tren pengguna aktif harian?"
- b) **Dalam BigQuery:** Terjemahkan tujuan ini ke dalam kebutuhan data spesifik, seperti tabel, kolom, dan metrik yang relevan.
- c) **Contoh Kueri:** Identifikasi tabel (misalnya, sales\_data) dan kolom (misalnya, product\_category, sale\_amount, sale\_date).

#### 2) Eksplorasi dan Pemahaman Data

- a) **Deskripsi:** Pahami struktur dataset, termasuk nama tabel, kolom, tipe data, dan hubungan antar tabel (schema). Identifikasi potensi masalah seperti data yang hilang atau inkonsisten.
- b) **Dalam BigQuery:** Gunakan perintah seperti `SELECT * FROM project.dataset.table LIMIT 10` untuk melihat sampel data atau `INFORMATION_SCHEMA` untuk memeriksa metadata.

### 3) Pembersihan dan Persiapan Data

- a) **Deskripsi:** Tangani data yang hilang, duplikat, atau format yang tidak konsisten. Transformasi data seperti mengubah tipe data atau menggabungkan tabel mungkin diperlukan.
- b) **Dalam BigQuery:** Gunakan fungsi seperti CAST, COALESCE, atau CASE untuk menangani data, serta JOIN untuk menggabungkan tabel.

### 4) Pemodelan dan Agregasi Data

- a) **Deskripsi:** Tentukan metrik atau dimensi yang akan dianalisis, seperti total, rata-rata, atau grup berdasarkan kategori. Pilih operasi agregasi seperti SUM, COUNT, atau AVG.
- b) **Dalam BigQuery:** Gunakan GROUP BY, HAVING, atau fungsi jendela (window functions) untuk analisis kompleks.

### 5) Analisis dan Visualisasi

- a) **Deskripsi:** Ekstrak wawasan dari hasil kueri, seperti tren, pola, atau anomali. Hasil kueri dapat dihubungkan ke alat visualisasi seperti Google Data Studio.
- b) **Dalam BigQuery:** Pastikan kueri menghasilkan data dalam format yang siap untuk visualisasi, seperti kolom yang terstruktur untuk sumbu X/Y.

### 6) Optimasi dan Validasi

- a) **Deskripsi:** Periksa performa kueri (waktu eksekusi, biaya) dan validasi hasilnya untuk memastikan keakuratan. Optimalkan dengan indeks, partisi, atau clustering jika memungkinkan.
- b) **Dalam BigQuery:** Gunakan EXPLAIN untuk menganalisis rencana eksekusi kueri, atau manfaatkan tabel yang dipartisi berdasarkan tanggal untuk efisiensi.

### 7) Iterasi dan Implementasi

- a) **Deskripsi:** Berdasarkan hasil analisis, ulangi proses jika diperlukan untuk menyempurnakan kueri atau menjawab pertanyaan baru. Implementasikan kueri dalam pipeline atau laporan otomatis.
- b) **Dalam BigQuery:** Jadwalkan kueri menggunakan fitur BigQuery Scheduled Queries atau integrasikan dengan alat seperti Cloud Composer untuk otomatisasi.

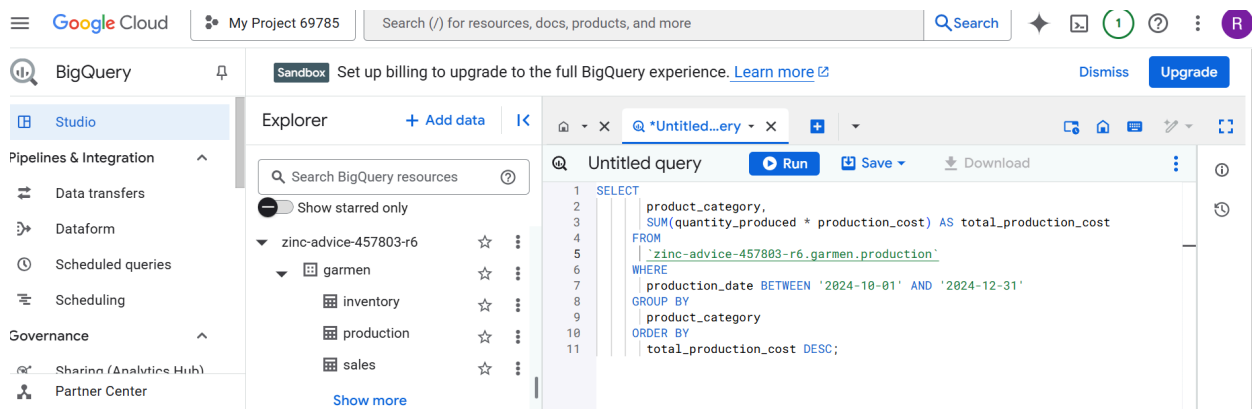
## PERINTAH:

Sesuai dengan urutan masing-masing 'task' diatas buat soal untuk kepentingan bisnis pabrikasi khususnya pabrik garment. Buat tabel datasetnya sesuai dengan kondisi urutan diatas. Tabel dibuat dengan 2 versi yaitu excel dan csv.

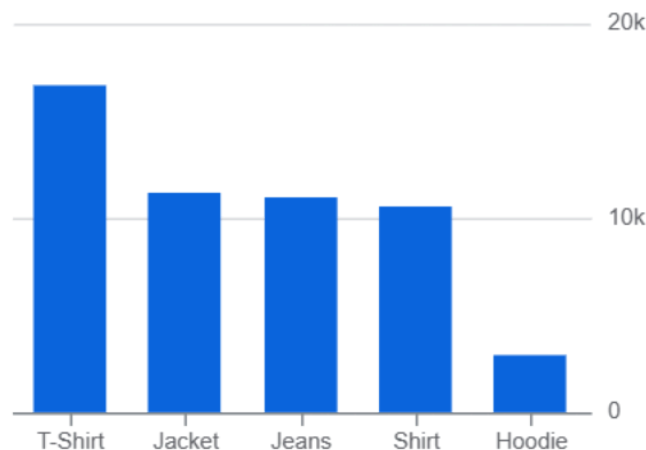
## JAWABAN:

### Identifikasi Tujuan Analitik

- **Soal:** Seorang manajer pabrik garment ingin mengetahui performa produksi dan penjualan untuk mengoptimalkan alokasi sumber daya di Q4 2024. Tentukan tujuan analitik untuk menghitung total biaya produksi per kategori produk dan membandingkannya dengan pendapatan penjualan.
- **Kode SQL:** Menghitung total biaya produksi per kategori produk di Q4 2024.



total\_production\_cost by product\_category



	Job information	Results	Chart	JSON	Execution details	Export
Row	product_category	total_production...				
1	T-Shirt	16780.0				
2	Jacket	11250.0				
3	Jeans	11000.0				
4	Shirt	10540.0				
5	Hoodie	3000.0				

## Pemahaman Struktur Data

- **Soal:** Jelaskan struktur tabel sales dan identifikasi data penjualan awal untuk memahami distribusi penjualan di berbagai wilayah. Bagaimana cara mengeksplorasi 10 transaksi penjualan pertama?
- **Kode SQL:** Mengeksplorasi data awal dari tabel sales.

The screenshot shows the Google Cloud BigQuery Studio interface. On the left, the 'Explorer' pane displays a project named 'zinc-advice-457803-r6' with a dataset named 'garmen'. Under 'garmen', there are four tables: 'inventory', 'production', 'sales', and 'sales'. The 'Query editor' pane on the right shows a SQL query:

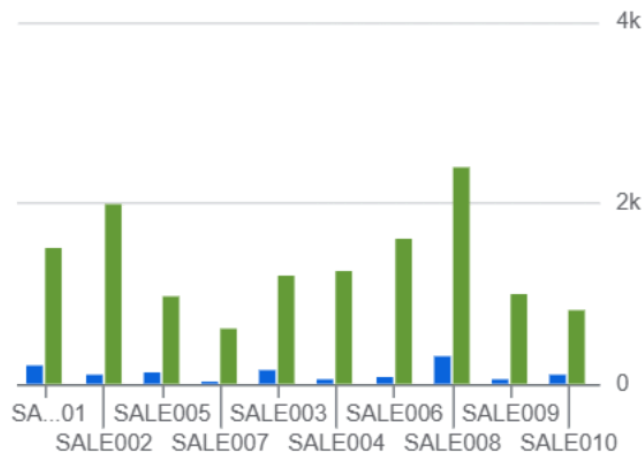
```

1 SELECT
2   sale_id,
3   product_id,
4   sale_date,
5   quantity_sold,
6   sale_amount,
7   region
8 FROM
9   `zinc-advice-457803-r6.garmen.sales`
10 ORDER BY
11   sale_date

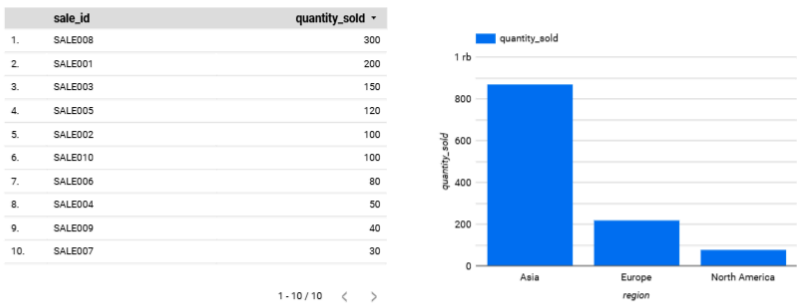
```

The query is completed, as indicated by the 'Query completed' message at the bottom of the editor.

quantity\_sold, sale\_amount by sale\_id

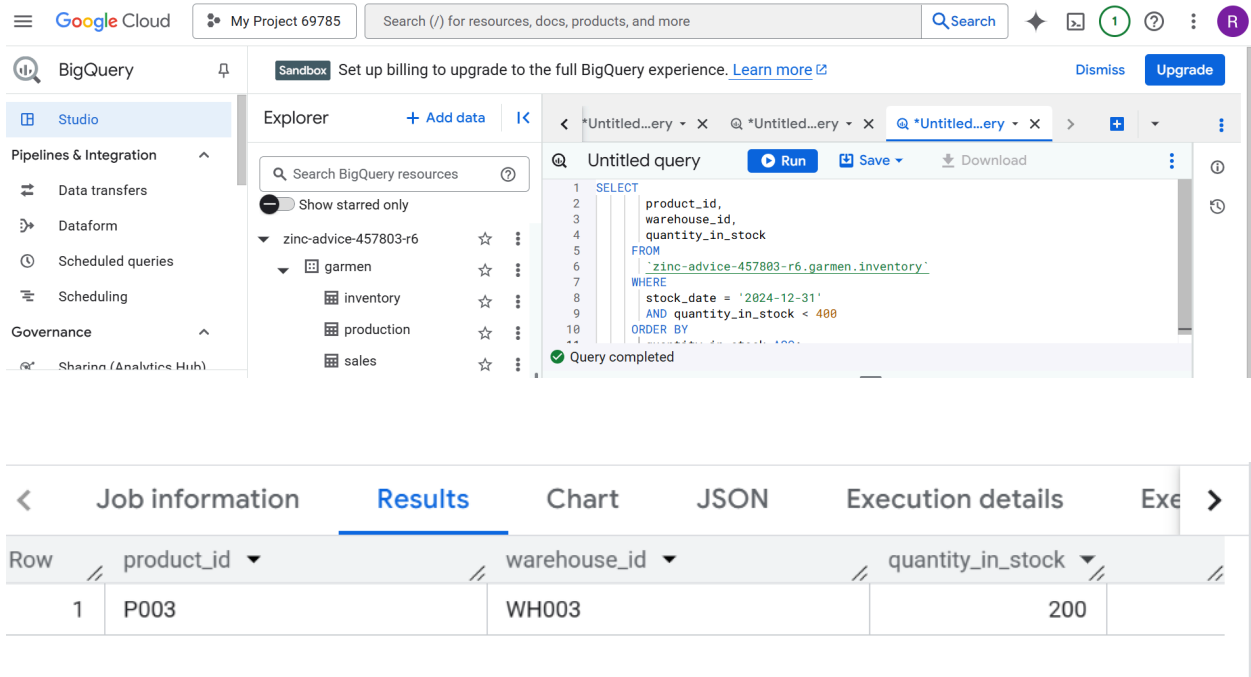


# BigQuery SQL kustom

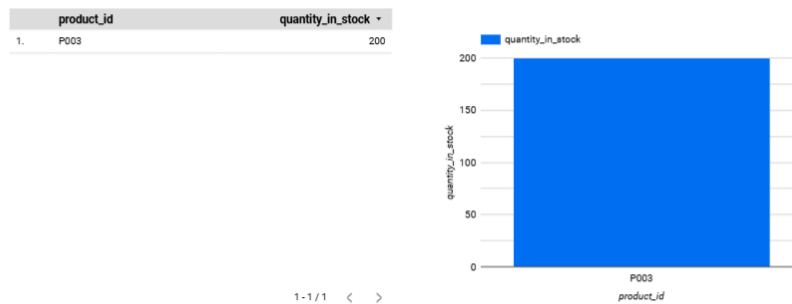


## Perumusan Kebutuhan Data

- **Soal:** Untuk mencegah kekurangan stok, identifikasi produk dengan stok di bawah 400 unit di gudang pada tanggal pencatatan stok terbaru (31 Desember 2024). Tentukan dimensi (product\_id, warehouse\_id) dan metrik (quantity\_in\_stock).
- **Kode SQL:** Mencari produk dengan stok rendah pada 31 Desember 2024.



## BigQuery SQL kustom



### Desain Kueri SQL

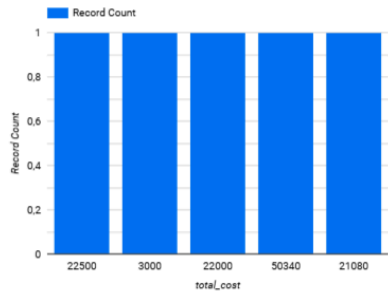
- **Soal:** Buat kueri untuk menghitung margin keuntungan per produk berdasarkan biaya produksi dan penjualan di Q4 2024.
- **Kueri SQL:**

The screenshot shows the Google Cloud BigQuery Studio interface. The top navigation bar includes the Google Cloud logo, the project name 'My Project 69785', a search bar, and a 'Search' button. The left sidebar contains the 'BigQuery' logo and a list of navigation items: 'Studio', 'Pipelines & Integration', 'Data transfers', 'Dataform', 'Scheduled queries', 'Scheduling', 'Governance', 'Sharon (Analytics Hub)', 'Partner Center', 'Settings', and 'Release Notes'. The main area is divided into three sections: 'Explorer', 'Query Editor', and 'Results'. The 'Explorer' section shows a tree view of the project's data assets, including 'zinc-advice-457803-r6', 'garmen', 'inventory', 'production', and 'sales'. The 'Query Editor' section displays a custom SQL query for calculating profit margins. The query is as follows:

```
1 SELECT
2   p.product_category,
3   SUM(s.sale_amount) AS total_sales,
4   SUM(p.quantity_produced * p.production_cost) AS total_cost,
5   SUM(s.sale_amount) - SUM(p.quantity_produced * p.production_cost) AS
6   profit_margin
7 FROM
8   `zinc-advice-457803-r6.garmen.production` p
9 INNER JOIN
10  `zinc-advice-457803-r6.garmen.sales` s
11 ON
12  p.product_id = s.product_id
13 WHERE
14  s.sale_date BETWEEN '2024-10-01' AND '2024-12-31'
15  AND p.production_date BETWEEN '2024-10-01' AND '2024-12-31'
16 GROUP BY
17  p.product_category
18 ORDER BY
19  profit_margin DESC;
```

# BigQuery SQL kustom

product_category		Record Count
1.	Jacket	1
2.	Hoodie	1
3.	Jeans	1
4.	T-Shirt	1
5.	Shirt	1



1 - 5 / 5 < >

Row	product_category	total_sales	total_cost	profit_margin
1	Hoodie	600.0	3000.0	-2400.0
2	Jeans	7200.0	22000.0	-14800.0
3	Shirt	3560.0	21080.0	-17520.0
4	Jacket	4500.0	22500.0	-18000.0
5	T-Shirt	15300.0	50340.0	-35040.0

## Optimasi Kueri

- **Soal:** Optimalkan kueri untuk analisis penjualan harian agar lebih efisien, dengan memanfaatkan partisi tabel sales berdasarkan sale\_date.
- **Kueri SQL** (diasumsikan tabel sales dipartisi berdasarkan sale\_date):

BigQuery

Sandbox Set up billing to upgrade to the full BigQuery experience. [Learn more](#)

Dismiss Upgrade

Studio

Explorer + Add data

Search BigQuery resources

Show starred only

zinc-advice-457803-r6

- garmen
  - inventory
  - production
  - sales

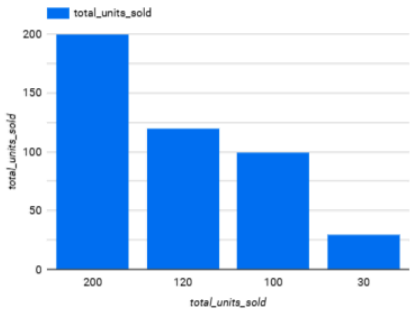
Untitled query

Run Save Download

```
1 SELECT
2   DATE(sale_date) AS sale_day,
3   SUM(quantity_sold) AS total_units_sold
4 FROM
5   `zinc-advice-457803-r6.garmen.sales`
6 WHERE
7   sale_date BETWEEN '2024-10-01' AND '2024-10-31'
8 GROUP BY
9   sale_day
10 ORDER BY
11   sale_day;
```

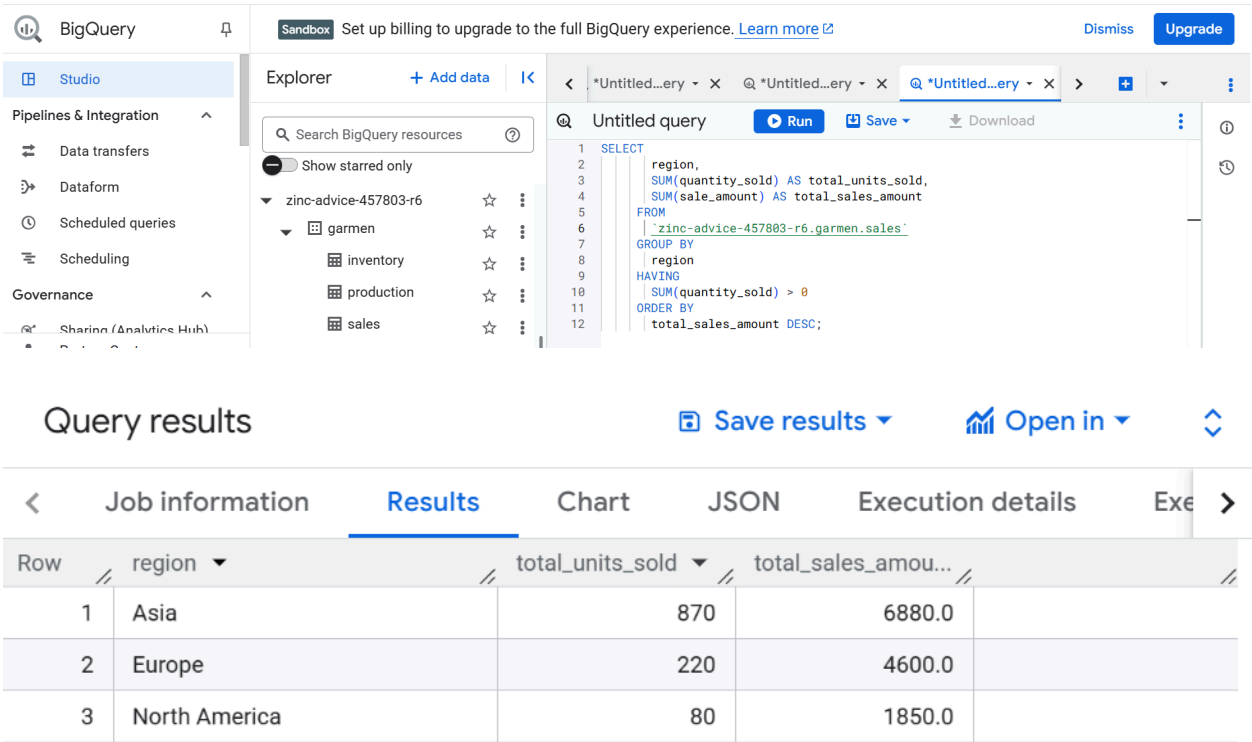
# BigQuery SQL kustom

	sale_day	total_units_sold
1.	5 Okt 2024	200
2.	15 Okt 2024	120
3.	6 Okt 2024	100
4.	20 Okt 2024	30



## Validasi Hasil

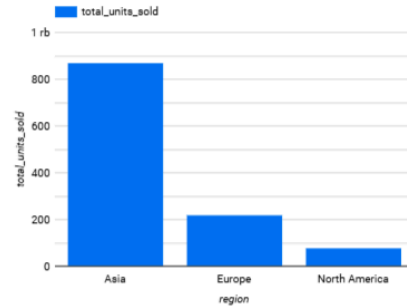
- **Soal:** Bagaimana cara memvalidasi bahwa total penjualan per wilayah sesuai dengan data mentah? Buat kueri untuk membandingkan total penjualan dengan jumlah unit terjual.





# BigQuery SQL kustom

	region	total_units_sold
1.	Asia	870
2.	Europe	220
3.	North America	80



1 - 3 / 3 < >

## Iterasi dan Penyempurnaan

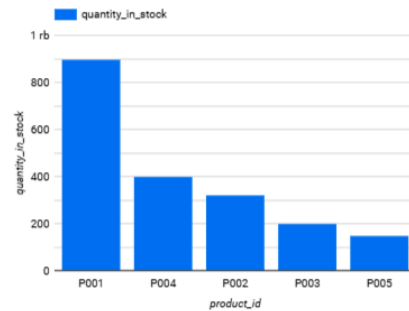
- **Soal:** Jika analisis stok menunjukkan data yang tidak konsisten, bagaimana cara memperbaiki kueri untuk memastikan hanya data stok terbaru yang diambil?
- **Kueri SQL** (mengambil stok terbaru per produk):

The screenshot shows the Google Cloud BigQuery Studio interface. The top bar includes the Google Cloud logo, the project name 'My Project 69785', and a search bar. The left sidebar contains navigation options: Studio, Pipelines & Integration, Data transfers, Dataform, Scheduled queries, and Governance. The main area is divided into an Explorer pane on the left and a query editor on the right. The Explorer pane shows a tree view of the project's resources, including a dataset named 'garmen' with tables 'inventory', 'production', and 'sales'. The query editor displays a custom SQL query titled 'Untitled query'.

```
1 WITH LatestStock AS (  
2   SELECT  
3     product_id,  
4     warehouse_id,  
5     quantity_in_stock,  
6     ROW_NUMBER() OVER (PARTITION BY product_id, warehouse_id ORDER BY  
7       stock_date DESC) AS rn  
8   FROM  
9     `zinc-advice-457803-r6.garmen.inventory`  
10 )  
11 SELECT  
12   product_id,  
13   warehouse_id,  
14   quantity_in_stock  
15 FROM  
16   LatestStock  
17 WHERE  
18   rn = 1  
19 ORDER BY  
20   product_id, warehouse_id;
```

## BigQuery SQL kustom

	product_id	quantity_in_stock
1.	P001	900
2.	P004	400
3.	P002	320
4.	P003	200
5.	P005	150



1 - 5 / 5 < >

### Pelaporan dan Implementasi

- **Soal:** Buat laporan otomatis untuk menampilkan tren penjualan bulanan per kategori produk, dan simpan hasilnya sebagai view di Big Query.