

**LAPORAN  
BUSINESS INTELLIGENCE**

**PROJECT AKHIR  
KELOMPOK 7**



**Disusun oleh:**

<b>Alvino Valerian D.R</b>	<b>2341720027</b>
<b>Anugerah Rachmadani</b>	<b>244107023012</b>
<b>Ilham Faturachman</b>	<b>244107023001</b>
<b>Muhammad Adityo Rahman</b>	<b>2341720177</b>

**PROGRAM STUDI D-IV TEKNIK INFORMATIKA**

**JURUSAN TEKNOLOGI INFORMASI**

**POLITEKNIK NEGERI MALANG**

**2025**

## Daftar isi

---

Daftar isi.....	2
<b>Link Github.....</b>	<b>3</b>
<b>Dataset.....</b>	<b>4</b>
Diagram Star Schema.....	5
Penjelasan Proses ETL Dimensi (Produk, Pelanggan, Pembayaran).....	6
<b>Penjelasan Proses ETL Dimensi Waktu.....</b>	<b>7</b>
<b>Penjelasan Proses ETL Tabel Fakta (Fact Sales).....</b>	<b>8</b>
<b>Orkestrasi Job (Main Job).....</b>	<b>9</b>
Perbandingan Query dan Visualisasi dari Looker Studio.....	9
1.Total Penjualan.....	9
2.Penjualan Terbanyak per Produk (Paling Sering Dipakai).....	10
3.Penjualan per Bulan (Trend Waktu).....	10
4.Penjualan per Kategori Produk.....	12
5.Metode Pembayaran yang Paling Sering Digunakan.....	12
6.TOP 5 Pelanggan Berdasarkan Nilai Belanja.....	13
7.Rata-Rata Rating Produk.....	13

## **Link Github**

---

[https://github.com/AlvinoValerian/PBL\\_BUSINESS-INTELLIGENCE](https://github.com/AlvinoValerian/PBL_BUSINESS-INTELLIGENCE)

## **Link Looker Studio**

---

<https://lookerstudio.google.com/reporting/7ebb3236-70ed-4138-927f-c2edd0bf5da2>

## Dataset

### Customer Purchase History Dataset

*Understand your shoppers, not just your sales. This 1 800-row dataset captures multi-product purchase behavior, payment preferences and review ratings—ideal for cohort analysis and Customer Lifetime Value calculations. Use it to reveal cross-sell opportunities, build “top customers” reports or explore rating-based product popularity.*

#### Columns Included:

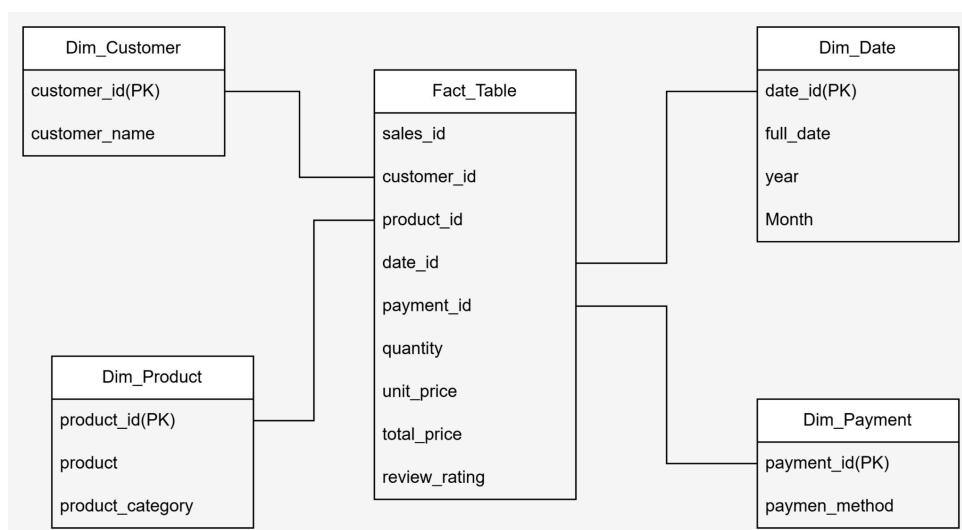
`CustomerID`, `Product`, `PurchaseDate`, `Quantity`, `UnitPrice`, `CustomerName`,  
`ProductCategory`, `PaymentMethod`, `ReviewRating`, `TotalPrice`

- Calculate Customer Lifetime Value (CLV).
- Rank top customers by spend or rating.
- Analyse product popularity within categories.
- Use PivotTables to find frequently bundled items.

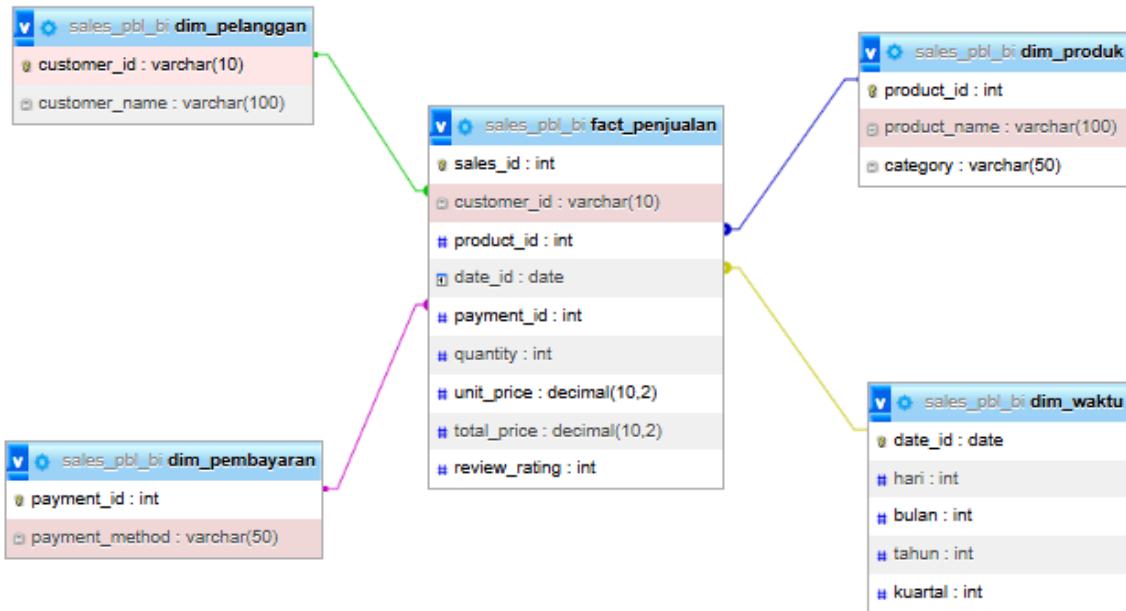
Link Web : <https://excelx.com/practice-data/sales-retail/>

Dataset :

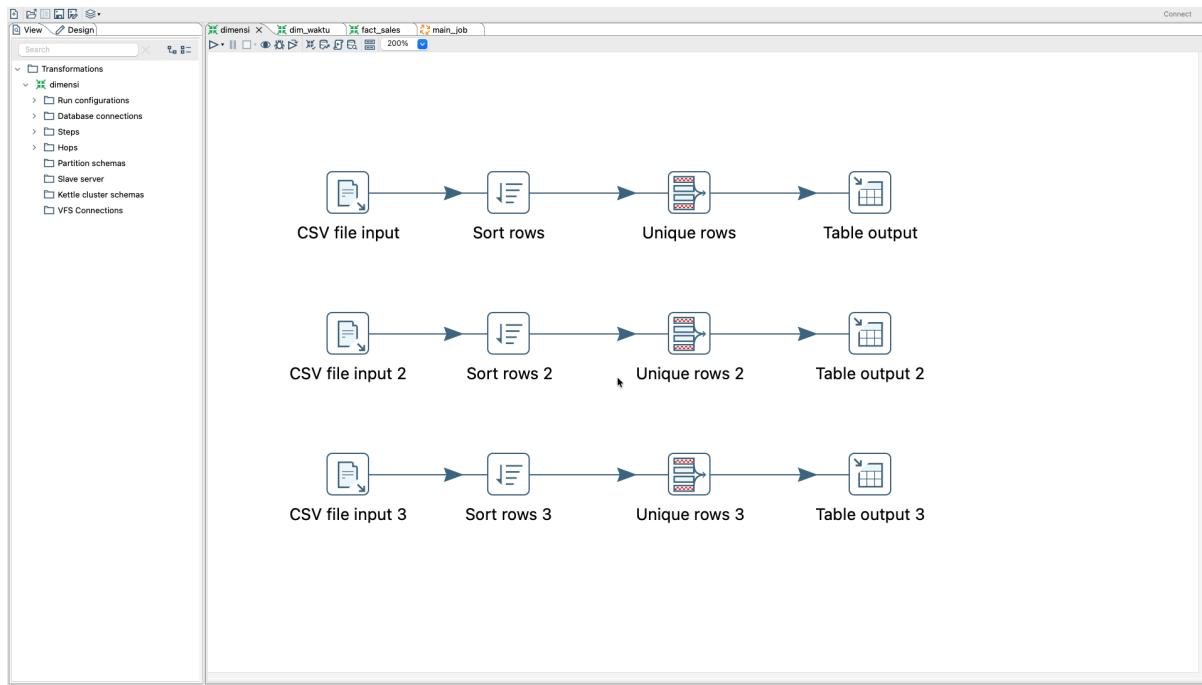
<https://excelx.com/wp-content/uploads/2025/06/Customer-Purchase-History.xlsx>



## Diagram Star Schema



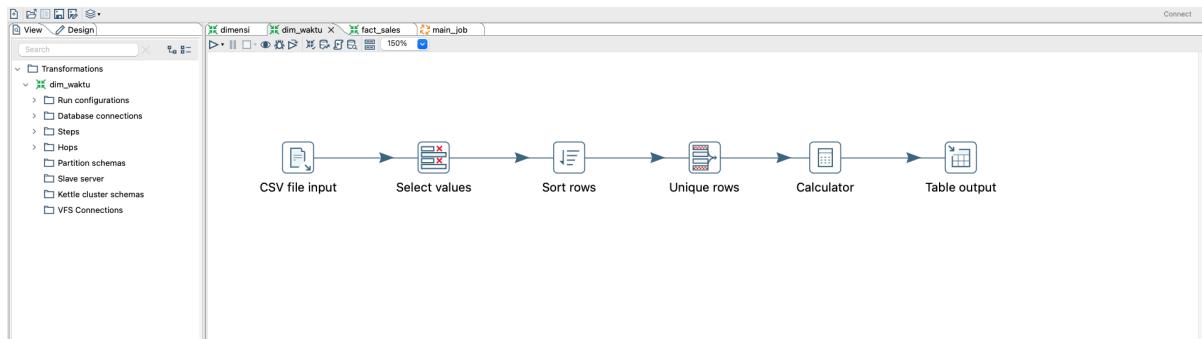
## Penjelasan Proses ETL Dimensi (Produk, Pelanggan, Pembayaran)



**Penjelasan Alur Kerja:** Proses ini bertujuan untuk menormalisasi data dengan memisahkan atribut deskriptif dari data transaksi utama untuk membentuk tabel dimensi dalam Skema Bintang.

- 1. CSV File Input (Extract):** Langkah ekstraksi data dari sumber mentah ([Customer-Purchase-History.csv](#)). Pada tahap ini, seluruh kolom dibaca dan tipe data disesuaikan.
- 2. Sort Rows (Transform):** Data diurutkan berdasarkan atribut kunci dimensi (misalnya: [Product](#) untuk Dimensi Produk atau [CustomerID](#) untuk Dimensi Pelanggan). Pengurutan ini **wajib** dilakukan sebagai prasyarat untuk langkah selanjutnya.
- 3. Unique Rows (Transform):** Melakukan proses *data cleaning* dengan menghapus duplikasi. Karena data sumber bersifat transaksional (berulang), langkah ini memastikan hanya ada satu baris unik untuk setiap entitas (misalnya: satu baris unik untuk setiap Nama Produk) agar sesuai dengan prinsip *Primary Key* di tabel dimensi.
- 4. Table Output (Load):** Memuat data yang sudah bersih dan unik ke dalam tabel dimensi target ([Dim\\_Produk](#), [Dim\\_Pelanggan](#), [Dim\\_Pembayaran](#)) di database [sales\\_dwh](#). Opsi *Truncate Table* diaktifkan untuk memastikan tabel bersih dari data lama sebelum proses muat dimulai.

## Penjelasan Proses ETL Dimensi Waktu



**Penjelasan Alur Kerja:** Proses ini bertujuan untuk membentuk hirarki waktu agar analisis penjualan dapat dilakukan berdasarkan periode (Tahun, Bulan, Kuartal).

1. Select Values (Transform - Meta-data):

Mengubah tipe data kolom PurchaseDate dari String menjadi format Date (yyyy-MM-dd) untuk memastikan kompatibilitas dengan fungsi kalkulasi tanggal.

2. Sort & Unique Rows (Transform):

Mengambil daftar tanggal unik dari data transaksi. Ini memastikan tidak ada tanggal ganda yang masuk ke tabel dimensi waktu.

3. Calculator (Transform - Derivation):

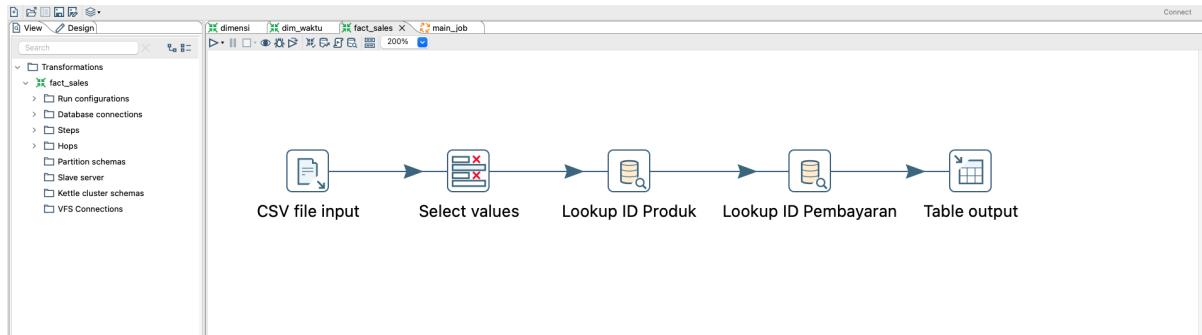
Mengekstrak atribut turunan dari PurchaseDate. Menggunakan fungsi kalkulasi internal Pentaho untuk menghasilkan kolom baru:

- **Year of date A => Kolom Tahun**
- **Month of date A => Kolom Bulan**
- **Quarter of date A => Kolom Kuartal**

4. Table Output (Load):

Menyimpan data tanggal beserta atribut turunannya ke tabel Dim\_Waktu. Mapping dilakukan secara selektif (kolom CustomerID diabaikan karena tidak relevan dengan dimensi waktu).

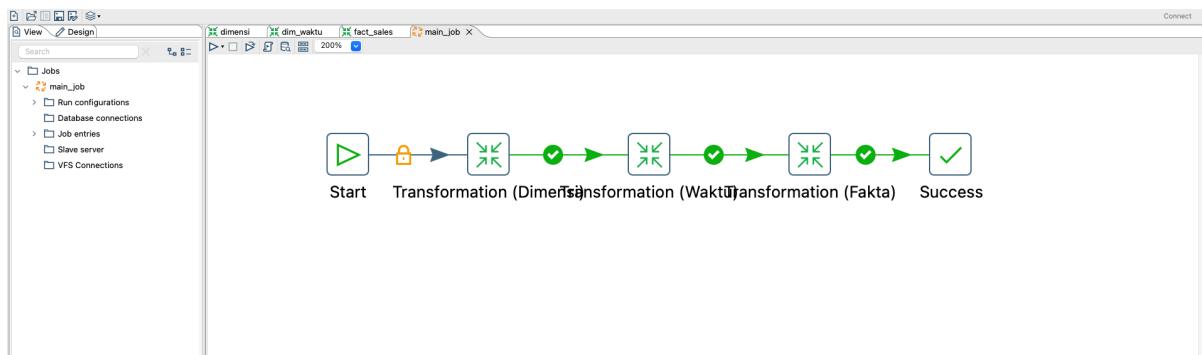
## Penjelasan Proses ETL Tabel Fakta (Fact Sales)



**Penjelasan Alur Kerja:** Ini adalah proses utama pembentukan *Data Warehouse*, menggabungkan metrik numerik dengan kunci referensi (*Foreign Keys*).

1. **CSV File Input (Extract):** Membaca kembali data transaksi lengkap.
2. **Database Lookup (Transform - Surrogate Key Replacement):** Langkah ini krusial untuk menjaga **Integritas Referensi** dalam *Star Schema*.
  - **Lookup Produk:** Sistem mencari `product_name` dari CSV ke tabel `Dim_Produk` di database, lalu mengambil `product_id`-nya.
  - **Lookup Pembayaran:** Sistem mencari `PaymentMethod` dari CSV ke tabel `Dim_Pembayaran`, lalu mengambil `payment_id`-nya.
  - **Tujuan:** Mengganti data teks di tabel fakta menjadi ID angka (*Surrogate Key*) agar kinerja *query* lebih cepat dan hemat penyimpanan.
3. **Table Output (Load):** Memuat data akhir ke tabel `Fact_Penjualan`. Kolom yang dimuat meliputi:
  - **Foreign Keys:** `customer_id`, `product_id`, `payment_id`, `date_id`.
  - **Measures (Fakta):** `quantity`, `total_price`, `unit_price`, `review_rating`.

## Orkestrasi Job (Main Job)



**Penjelasan:** Seluruh transformasi diatur dalam satu *Job* utama untuk memastikan urutan eksekusi yang benar (Dependency Management):

1. **Start:** Titik awal eksekusi.
2. **Transformation Dimensi & Waktu:** Dijalankan terlebih dahulu untuk mengisi tabel referensi.
3. **Transformation Fakta:** Dijalankan **terakhir**. Hal ini wajib dilakukan karena tabel fakta membutuhkan ID yang dihasilkan dari tabel dimensi. Jika urutan dibalik, proses akan gagal (*Error Foreign Key Constraint*).

## Perbandingan Query dan Visualisasi dar Looker Studio

### 1. Total Penjualan

```
Showing rows 0 - 0 (1 total, Query took 0.0278 seconds.)  
  
SELECT SUM(f.total_price) AS total_penjualan FROM sales_pbl.bi.fact_penjualan f;  
  
 Profiling [ Edit inline ] [ Edit ] [ Explain SQL ] [ Create PHP code ] [ Refresh ]  
  
 Show all | Number of rows: 25 Filter rows: Search this table  
  
  
  
total_penjualan  
3286656.80
```

Visualisasi:



KPI  
3.007.523,23

## 2.Penjualan Terbanyak per Produk (Paling Sering Dipakai)

Showing rows 0 - 6 (7 total, Query took 0.0341 seconds.)

```
SELECT p.product_name, SUM(f.quantity) AS total_qty, SUM(f.total_price) AS total_penjualan FROM sales_pbl.bi.fact_penjualan f JOIN sales_pbl.bi.dim_produk p ON f.product_id = p.product_id GROUP BY p.product_name ORDER BY total_penjualan DESC;
```

Profiling [ Edit inline ] [ Edit ] [ Explain SQL ] [ Create PHP code ] [ Refresh ]

Show all | Number of rows: 25 | Filter rows: Search this table

Extra options

product_name	total_qty	total_penjualan
Tablet	1316	533091,00
Laptop	1231	481721,40
Monitor	1180	468632,80
Desk	1156	454008,70
Phone	1117	444989,72
Chair	1147	443118,71
Printer	1113	441090,47

Visualisasi:



## 3.Penjualan per Bulan (Trend Waktu)

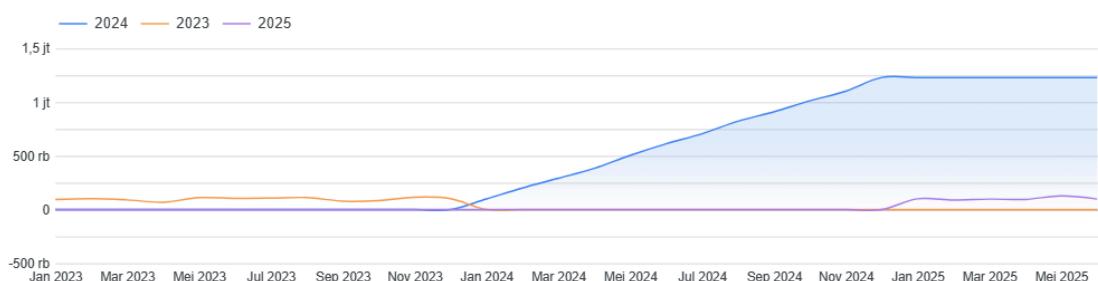
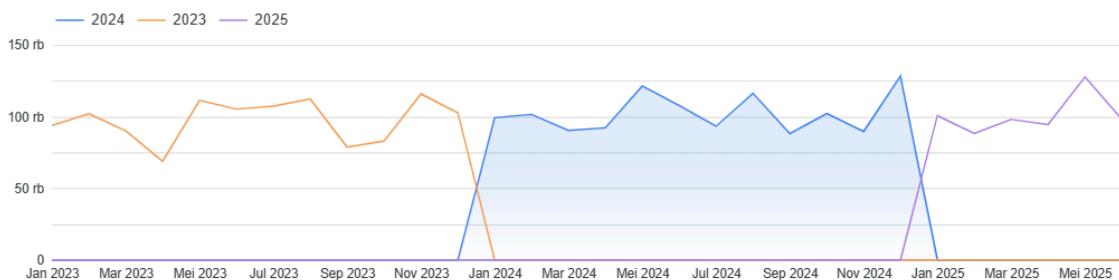
Showing rows 0 - 24 (30 total, Query took 0.0075 seconds.) [tahun: 2023... - 2025...] [bulan: 1... - 1...]

```
SELECT w.tahun, w.bulan, SUM(f.total_price) AS total_penjualan FROM sales_pbl.bi.fact_penjualan f JOIN sales_pbl.bi.dim_waktu w ON f.date_id = w.date_id GROUP BY w.tahun, w.bulan ORDER BY w.tahun, w.bulan;
```

Profiling [ Edit inline ] [ Edit ] [ Explain SQL ] [ Create PHP code ] [ Refresh ]

tahun	bulan	total_penjualan
2023	1	98390.67
2023	2	109746.60
2023	3	91665.47
2023	4	85517.05
2023	5	117663.53
2023	6	119837.23
2023	7	119476.63
2023	8	116736.14
2023	9	85040.67
2023	10	87857.58
2023	11	136455.33
2023	12	112883.28
2024	1	105158.94
2024	2	103280.28
2024	3	97450.02
2024	4	102020.38
2024	5	129387.77
2024	6	107748.12
2024	7	102536.36
2024	8	129835.63
2024	9	98900.14
2024	10	104198.67
2024	11	110434.39
2024	12	132860.98
2025	1	111583.86

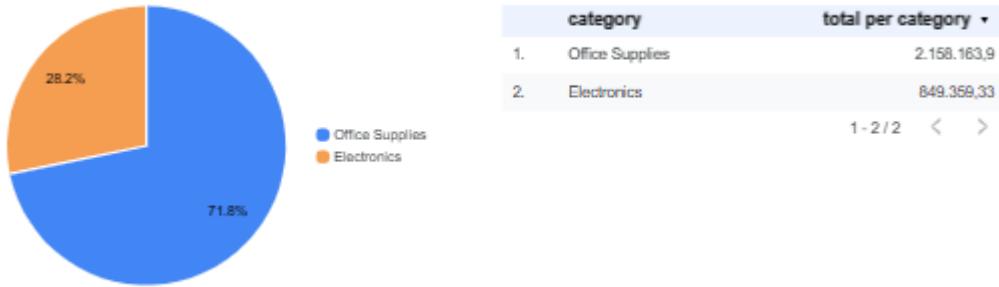
## Visualisasi:



## 4. Penjualan per Kategori Produk

Showing rows 0 - 1 (2 total). Query took 0.0081 seconds.		
<pre>SELECT p.category, SUM(f.total_price) AS total_penjualan FROM sales_pbl.bi.fact_penjualan f JOIN sales_pbl.bi.dim_produk p ON f.product_id = p.product_id GROUP BY p.category ORDER BY total_penjualan DESC;</pre>		
<input type="checkbox"/> Profiling [ Edit inline ] [ Explain SQL ] [ Create PHP code ] [ Refresh ]		
<input type="checkbox"/> Show all   Number of rows: 25 ▾ Filter rows: Search this table		
<a href="#">Extra options</a>		
category	total_penjualan	▼
Office Supplies	2316302,60	1
Electronics	950354,20	

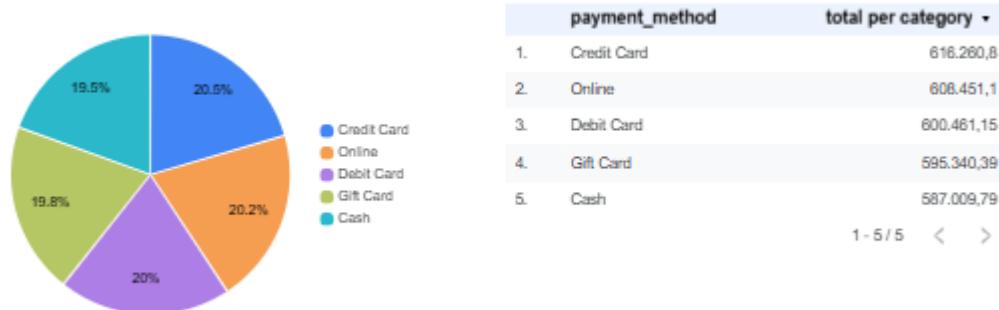
Visualisasi:



## 5. Metode Pembayaran yang Paling Sering Digunakan

Showing rows 0 - 4 (5 total). Query took 0.0140 seconds.		
<pre>SELECT pb.payment_method, COUNT(f.sales_id) AS total_transaksi, SUM(f.total_price) AS total_penjualan FROM sales_pbl.bi.fact_penjualan f JOIN sales_pbl.bi.dim_pembayaran pb ON f.payment_id = pb.payment_id GROUP BY pb.payment_method ORDER BY total_transaksi DESC;</pre>		
<input type="checkbox"/> Profiling [ Edit inline ] [ Explain SQL ] [ Create PHP code ] [ Refresh ]		
<input type="checkbox"/> Show all   Number of rows: 25 ▾ Filter rows: Search this table		
<a href="#">Extra options</a>		
payment_method	total_transaksi	▼
Debit Card	398	671304,23
Cash	356	615791,82
Gift Card	353	642235,87
Credit Card	349	675771,51
Online	348	661563,37

Visualisasi:



## 6. TOP 5 Pelanggan Berdasarkan Nilai Belanja

Showing rows 0 - 4 (5 total, Query took 0.0190 seconds.)

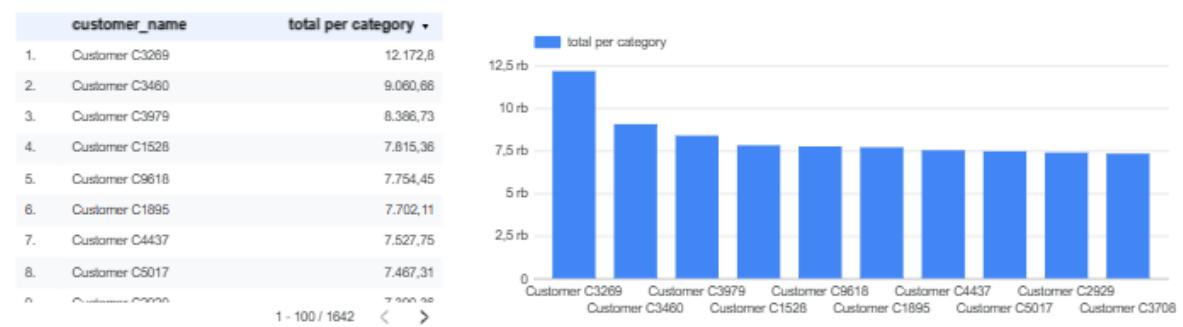
```
SELECT c.customer_name, SUM(f.total_price) AS total_belanja FROM sales_pbl.bi.fact_penjualan f JOIN sales_pbl.bi.dim_pelanggan c ON f.customer_id = c.customer_id GROUP BY c.customer_name ORDER BY total_belanja DESC LIMIT 5;
```

Profiling | [Edit inline](#) | [Edit](#) | [Explain SQL](#) | [Create PHP code](#) | [Refresh](#)

[Extra options](#)

customer_name	total_belanja
Customer C3269	12172,80
Customer C3460	9060,66
Customer C3979	8386,73
Customer C1528	7815,36
Customer C9618	7754,45

Visualisasi:



## 7. Rata-Rata Rating Produk

Showing rows 0 - 6 (7 total, Query took 0.7363 seconds.)

```
SELECT p.product_name, AVG(f.review_rating) AS rata_rating FROM sales_pbl.bi.fact_penjualan f JOIN sales_pbl.bi.dim_produk p ON f.product_id = p.product_id GROUP BY p.product_name ORDER BY rata_rating DESC;
```

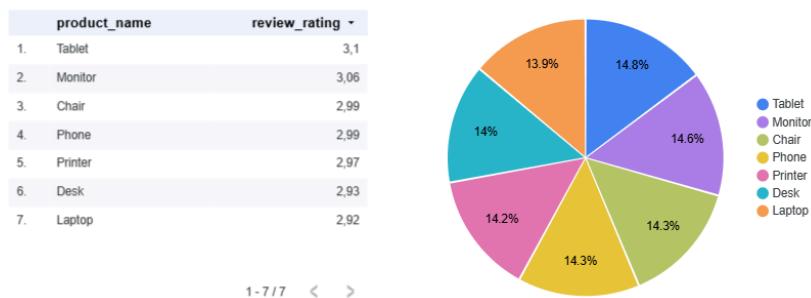
Profiling | [Edit inline](#) | [Edit](#) | [Explain SQL](#) | [Create PHP code](#) | [Refresh](#)

[Show all](#) | Number of rows: 25 | Filter rows:

[Extra options](#)

product_name	rata_rating
Tablet	3.1043
Monitor	3.0591
Chair	2.9921
Phone	2.9919
Printer	2.9700
Desk	2.9258
Laptop	2.9179

Visualisasinya:



## 8.Penjualan per Kuartal (Level Executive)

Showing rows 0 - 9 (10 total). Query took 0.0073 seconds.) [tahun: 2023... - 2025...] [kuartal: 1... - 2...]

```
SELECT w.tahun, w.kuartal, SUM(f.total_price) AS total_penjualan FROM sales_pbl.bi.fact_penjualan f JOIN sales_pbl.bi.dim_waktu w ON f.date_id = w.date_id GROUP BY w.tahun, w.kuartal ORDER BY w.tahun, w.kuartal;
```

Profiling  Edit inline  Explain SQL  Create PHP code  Refresh

Show all Number of rows: 25 Filter rows: Search this table

Extra options

tahun	kuartal	total_penjualan
2023	1	209802.74
2023	2	323017.81
2023	3	321253.44
2023	4	337196.19
2024	1	305889.24
2024	2	339156.27
2024	3	331272.13
2024	4	347294.04
2025	1	314363.29
2025	2	347411.65

Visualisasinya:

	tahun	kuartal	total_price
1.	2024	4	319.837,2
2.	2023	4	301.263,1
3.	2024	2	321.295,81
4.	2025	2	320.116,9
5.	2023	2	285.389,5
6.	2024	3	297.566,3
7.	2025	1	286.877,28
8.	2023	1	285.839,27
n.	2024	4	304.474,80

1 - 10 / 10 < >