

Final Task

Virtual Internship Program Kimia Farma

by Alfa Eridani

Petunjuk

Silahkan merujuk pada Data Source Task 5 yang telah disediakan untuk mengerjakan soal soal di bawah ini

Pada bagian data analytics, terdiri dari 4 soal dengan use case & tabel yang sama. Bayangkan kamu memiliki database erp yang terdiri dari 3 tabel: penjualan, pelanggan, barang. Tabel tersebut akan dibuat menjadi sebuah datamart yang nantinya digunakan untuk visualisasi.

Query

Soal 1 *:

Dari 2 query ini, mana yang bekerja lebih baik? Jelaskan mengapa.

(a) `SELECT * FROM pelanggan WHERE SUBSTR(alamat, 1, 3) = Mat;`

(b) `SELECT * FROM pelanggan WHERE alamat LIKE 'Mat%'`

**disclaimer: soal ini tidak terkait dengan data source*

Jawaban : (b) `SELECT * FROM pelanggan WHERE alamat LIKE 'Mat%'`

Alasan : *Query* (b) bekerja lebih baik karena dapat memanfaatkan indeks sehingga pencariannya dapat lebih cepat. *Query* (a) sendiri kurang efisien karena fungsi SUBSTR harus dipanggil untuk setiap baris dalam tabel yang menyebabkan kinerja menjadi lebih lambat. Selain itu, *query* (b) juga lebih mudah dibaca dibandingkan *query* (a).

Query

Soal 2 *:

Anggap kita memiliki tabel pelanggan dengan kolom: id, nama, tanggal_lahir, alamat. Bagaimana cara yang lebih tepat dalam menulis query untuk mendapatkan data pelanggan yang tanggal_lahir nya ada di antara 2000-01-01 sampai 2008-12-31? Pilihlah salah satu jawaban dan berikan alasannya.

- (a) `SELECT * FROM pelanggan WHERE tanggal_lahir >= '2000-01-01' AND tanggal_lahir <= '2008-12-31'`
- (b) `SELECT * FROM pelanggan WHERE tanggal_lahir BETWEEN '2000-01-01' AND '2008-12-31'`

**disclaimer: soal ini tidak terkait dengan data source*

Jawaban : (b) `SELECT * FROM pelanggan WHERE tanggal_lahir BETWEEN '2000-01-01' AND '2008-12-31'`

Alasan : Query (b) lebih efisien karena penggunaan BETWEEN lebih mudah dipahami dan dibaca dibandingkan menggunakan dua operator sekaligus seperti pada query (a). Operator BETWEEN sendiri sudah dirancang khusus untuk mengambil data yang berada di antara dua nilai.

Soal 3: Menentukan Primary Key

A. Tugas

Tentukan primary key dari table penjualan. jelaskan alasannya

B. Jawaban & Penjelasan :

Primary key dari tabel penjualan adalah kolom id_invoice dan id_customer, karena keduanya memiliki nilai yang unik (tidak ada nilai duplikat) dan tidak ada data NULL.

Soal 4: Design Datamart

A. Tugas

Buatlah design datamart (Terdiri dari tabel base, dan tabel aggregate). Upload file query dalam gdrive mu (pastikan dapat diakses public). Lalu masukkan linknya di tabel di bawah, dan cantumkan juga screenshoot query nya (jika lebih dari 1 file, maka masing masing file di-screenshoot)

Silahkan tambah halaman jika dibutuhkan

B. Jawaban :

No	Nama	Link
1	base_table.sql	https://drive.google.com/file/d/1Y7yB3Zgi46Obw2F5iPzjaF_G4v-a41W9/view?usp=share_link
2	aggregate_table.sql	https://drive.google.com/file/d/1PWllhT0GSyv2zOyMPR9CeQ1iFjYYnfiX/view?usp=share_link
3	Data csv	https://drive.google.com/drive/folders/1Z0rSHmLQJfvFpnL4ehZCQ72Y6kgIPhLa?usp=share_link

Table Base “base_table”

```
1 CREATE TABLE base_table (  
2     WITH penjualan_baru AS (  
3         SELECT id_invoice, id_customer, id_barang, tanggal, jumlah_barang  
4         FROM penjualan  
5         UNION DISTINCT  
6         SELECT id_invoice, id_customer, id_barang, tanggal, jumlah_barang  
7         FROM penjualan_ds  
8     )  
9  
10    SELECT  
11        pen.id_invoice,  
12        pen.id_customer,  
13        pen.id_barang,  
14        pel.id_cabang_sales,  
15        pel.nama,  
16        pel.cabang_sales,  
17        pel.group,  
18        bar.kode_lini,  
19        bar.lini,  
20        pel_ds.id_distributor,  
21        bar.nama_barang,
```

Table Base “base_table”

```
22     str_to_date(pen.tanggal, '%m/%d/%Y') as tanggal,  
23     pen.jumlah_barang,  
24     bar.kemasan,  
25     bar_ds.harga  
26 FROM penjualan_baru AS pen  
27 LEFT JOIN pelanggan AS pel  
28     ON pen.id_customer = pel.id_customer  
29 LEFT JOIN barang AS bar  
30     ON pen.id_barang = bar.kode_barang  
31 LEFT JOIN pelanggan_ds AS pel_ds  
32     ON pen.id_customer = pel_ds.id_customer  
33 LEFT JOIN barang_ds AS bar_ds  
34     ON pen.id_barang = bar_ds.kode_barang  
35 ORDER BY id_invoice  
36 )
```


Table Base “base_table”

column	data type	description	transformation
id_invoice	text	ID untuk identifikasi penjualan	
id_customer	mediumtext	ID untuk identifikasi pelanggan	
id_barang	mediumtext	ID untuk identifikasi barang	
id_cabang_sales	text	ID untuk identifikasi cabang penjualan	
nama	text	Nama cabang	
cabang_sales	text	Kota cabang	
group	text	Tipe cabang (apotek/klinik)	
kode_lini	int	Kode lini	

Table Base “base_table”

column	data type	description	transformation
lini	text	Lini barang	
id_distributor	text	ID untuk identifikasi distributor	
nama_barang	text	Nama barang	
tanggal	date	Tanggal pembelian	Mengubah text ke date dengan "str_to_date"
jumlah_barang	integer	Jumlah barang yang dibeli	
kemasan	text	Tipe kemasan barang	
harga	integer	Harga barang	

Table Aggregate “aggregate_table”

```
1 • ○ CREATE TABLE aggregate_table (  
2     SELECT *, (jumlah_barang * harga) AS total_harga  
3     FROM base_table  
4 )
```

Table Aggregate “aggregate_table”

column	data type	description	transformation
id_invoice	text	ID untuk identifikasi penjualan	
id_customer	mediumtext	ID untuk identifikasi pelanggan	
id_barang	mediumtext	ID untuk identifikasi barang	
id_cabang_sales	text	ID untuk identifikasi cabang penjualan	
nama	text	Nama cabang	
cabang_sales	text	Kota cabang	
group	text	Tipe cabang (apotek/klinik)	
kode_lini	int	Kode lini	

Table Aggregate “aggregate_table”

column	data type	description	transformation
lini	text	Lini barang	
id_distributor	text	ID untuk identifikasi distributor	
nama_barang	text	Nama barang	
tanggal	date	Tanggal pembelian	Mengubah text ke date dengan "str_to_date"
jumlah_barang	integer	Jumlah barang yang dibeli	
kemasan	text	Tipe kemasan barang	
harga	integer	Harga barang	
total_harga	integer	Total harga pembelian	Mengalikan jumlah_barang dengan harga

Soal 5 : Data Visualization

A. Tugas

buatlah data visualiasasi nya, dan cantumkan linknya di bawah (pastikan bisa diakses publik). Lalu cantumkan juga screenshot visualisasinya

Silahkan tambah halaman jika dibutuhkan

A. Jawaban :

Link visualisasi (ex link Google Data Studio) :

<https://lookerstudio.google.com/reporting/913ea634-d2dc-453b-9255-b42c6f81ca5e>

Brand Salicyl Sales Dashboard

by Alfa Eridani

Select date range

group

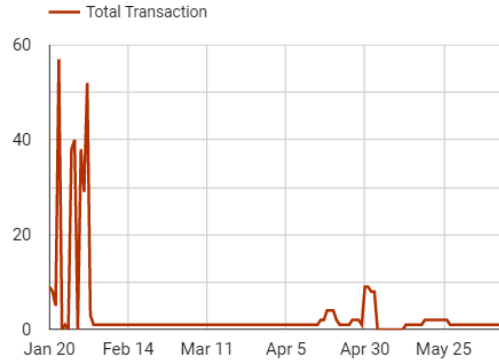
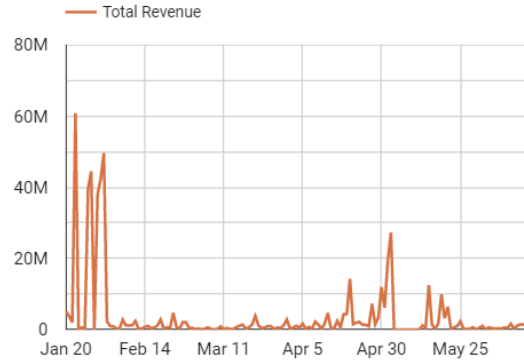


Total Transaction

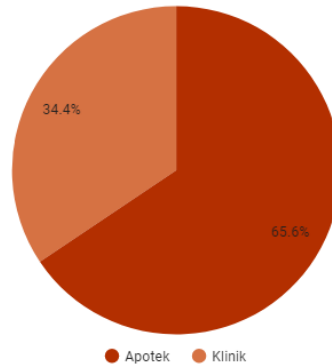
465

Total Revenue

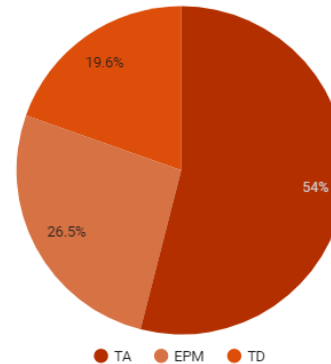
Rp523,747,500.00



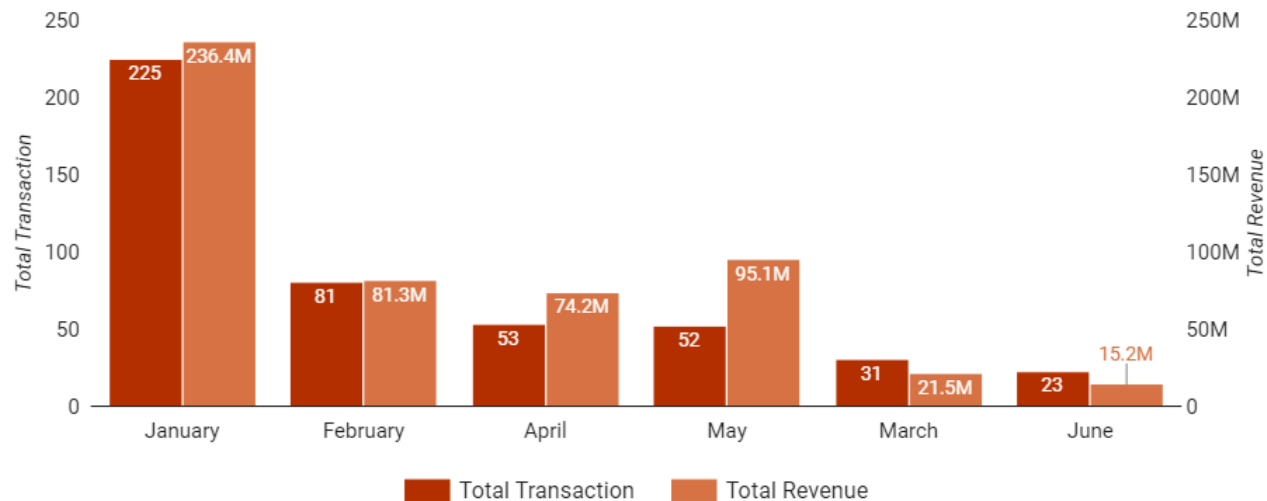
Total Transaction by Group



Total Transaction by Distributor

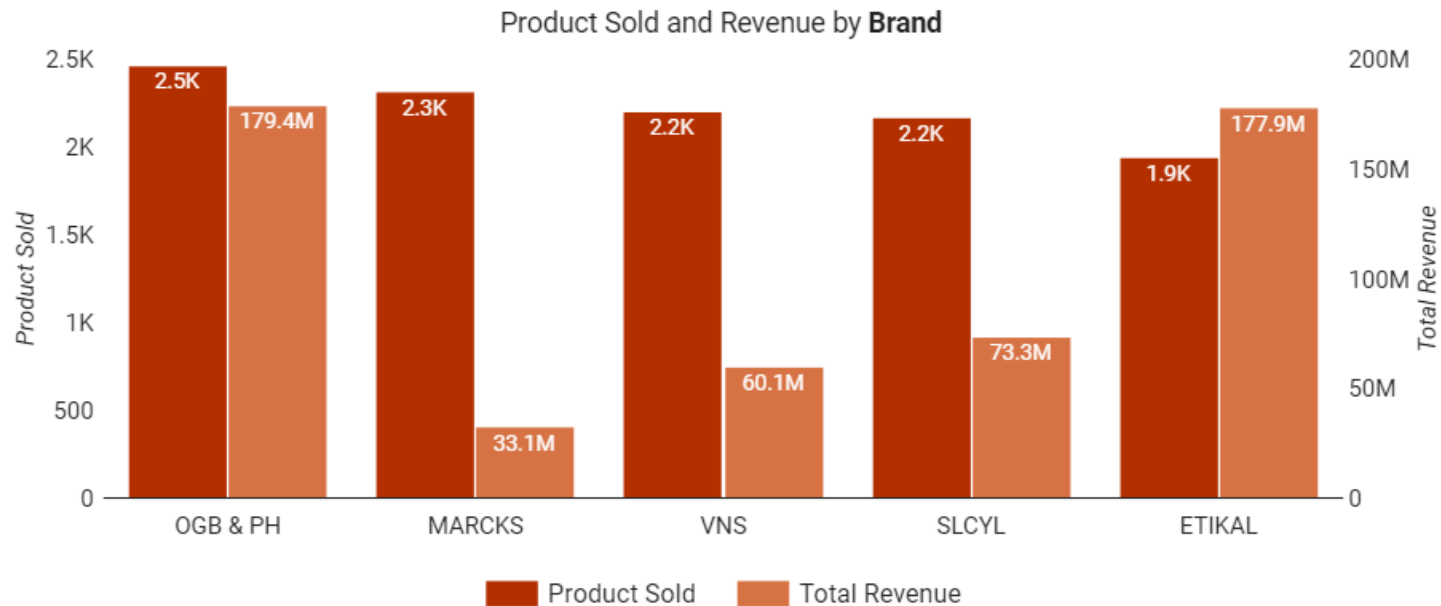


Total Transaction and Revenue by Month



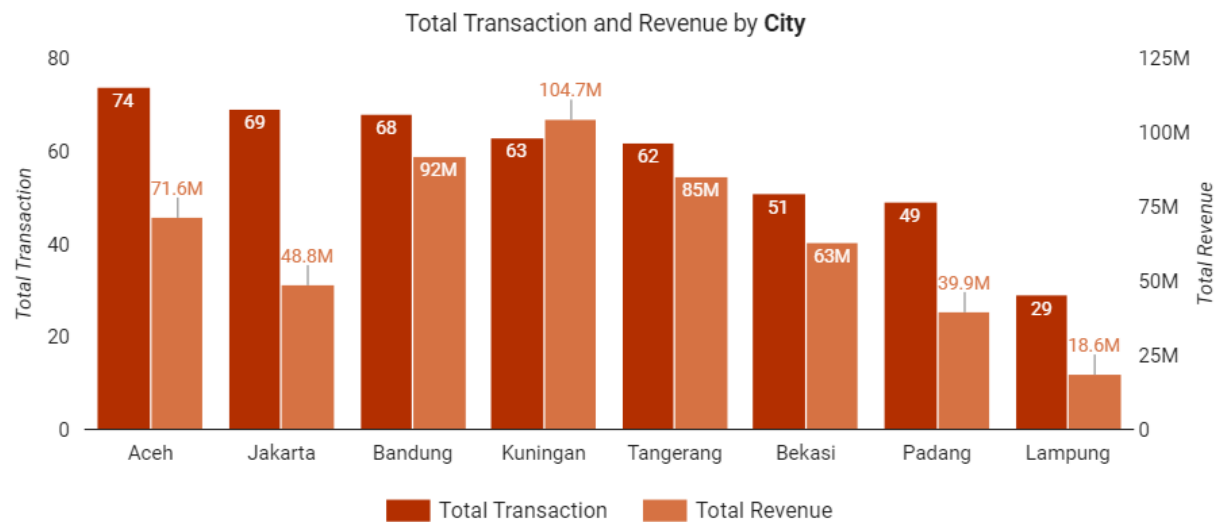
Transaction Data by Month

Bulan		Total Transactio...	Product Sold	Total Revenue
1.	January	225	5,057	Rp236,433,800
2.	February	81	2,060	Rp81,328,000
3.	April	53	1,704	Rp74,189,500
4.	May	52	1,427	Rp95,102,700
5.	March	31	444	Rp21,519,400
6.	June	23	410	Rp15,174,100



Transaction Data by Brand

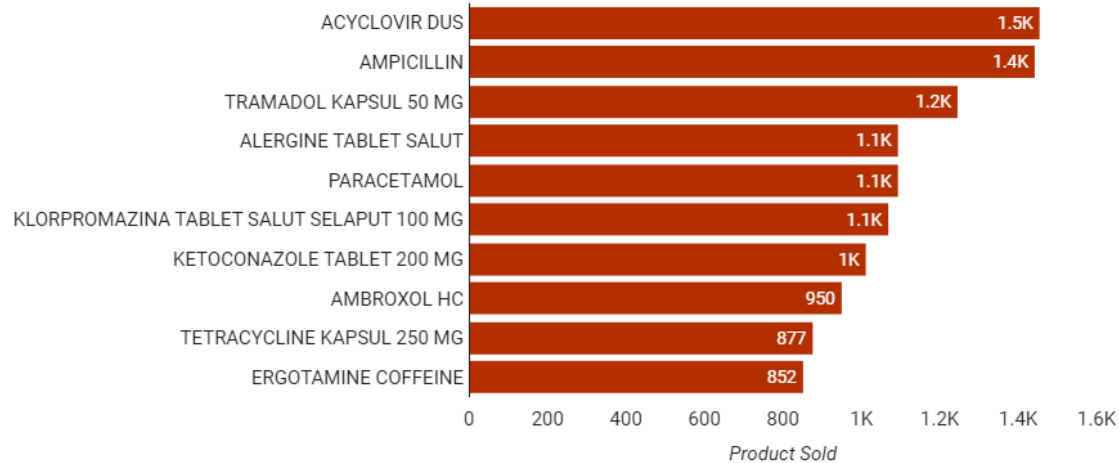
lini		Total Transaction	Product Sold...	Total Revenue
1.	OGB & PH	103	2,470	Rp179,436,000
2.	MARCKS	88	2,320	Rp33,125,600
3.	VNS	98	2,199	Rp60,050,500
4.	SLCYL	84	2,165	Rp73,259,000
5.	ETIKAL	92	1,948	Rp177,876,400



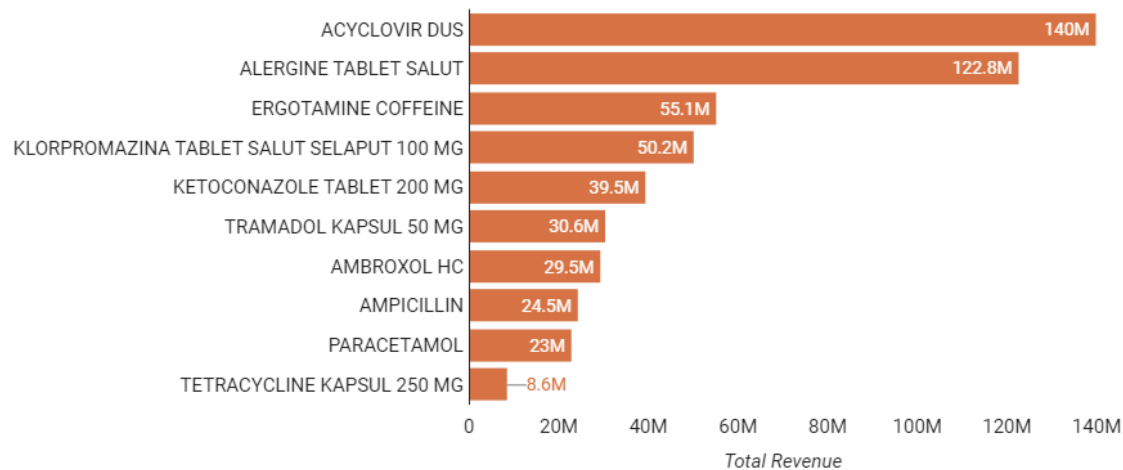
Transaction Data by City

City		Total Transactio...	Product Sold	Total Revenue
1.	Aceh	74	1,564	Rp71,589,900
2.	Jakarta	69	2,002	Rp48,810,300
3.	Bandung	68	1,840	Rp92,031,400
4.	Kuningan	63	1,822	Rp104,736,100
5.	Tangerang	62	1,345	Rp85,027,000
6.	Bekasi	51	1,175	Rp63,048,300
7.	Padang	49	829	Rp39,895,300
8.	Lampung	29	525	Rp18,609,200

Product Sold by **Product Name**



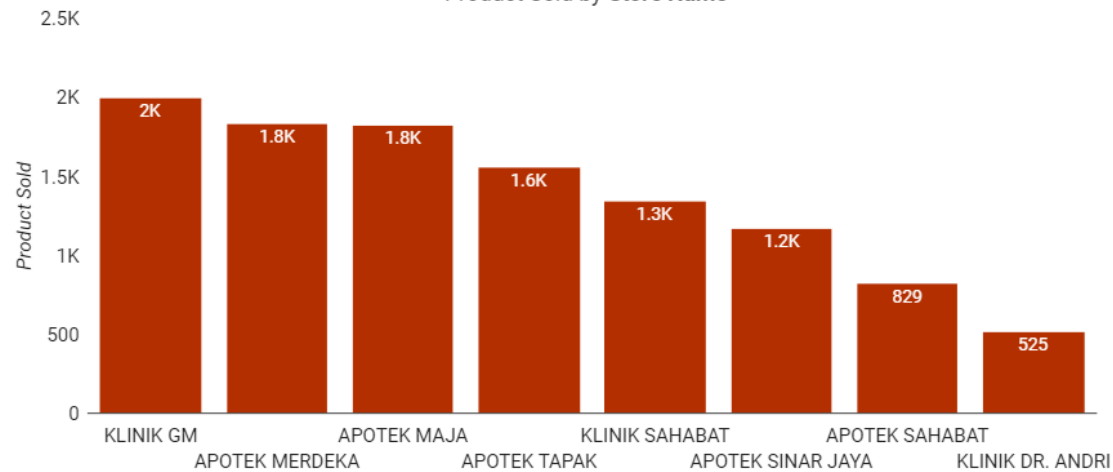
Total Revenue by **Product Name**



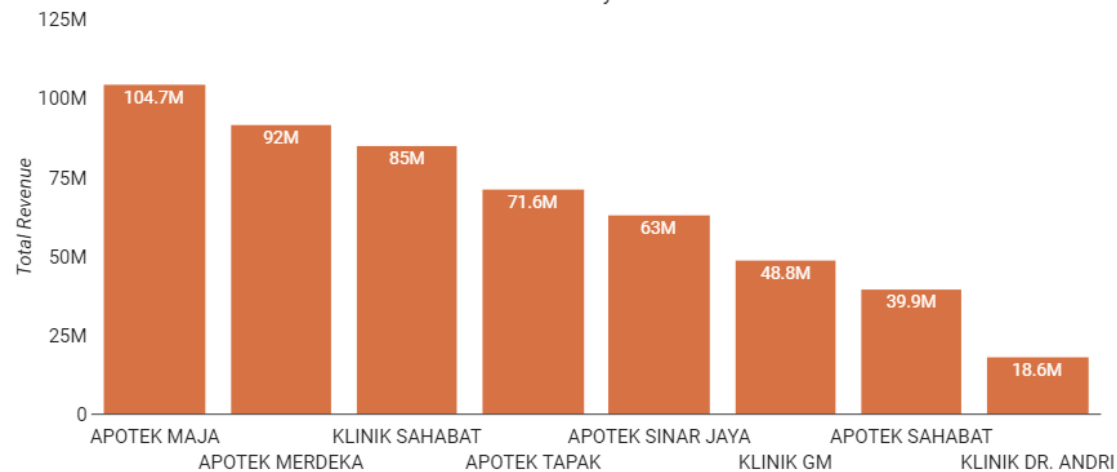
Transaction Data by **Product**

	nama_barang	Total Transaction	Product Sold	Total Revenue ▾
1.	ACYCLOVIR DUS	56	1,458	Rp139,968,000
2.	ALERGINE TABLET SALUT	55	1,096	Rp122,752,000
3.	ERGOTAMINE COFFEINE	37	852	Rp55,124,400
4.	KLORPROMAZINA TABLET SALUT SELAPUT 100 ...	49	1,069	Rp50,243,000
5.	KETOCONAZOLE TABLET 200 MG	47	1,012	Rp39,468,000
6.	TRAMADOL KAPSUL 50 MG	62	1,249	Rp30,600,500
7.	AMBROXOL HC	36	950	Rp29,450,000
8.	AMPICILLIN	55	1,443	Rp24,531,000
9.	PARACETAMOL	35	1,096	Rp23,016,000
10.	TETRACYCLINE KAPSUL 250 MG	33	877	Rp8,594,600

Product Sold by **Store Name**



Total Revenue by **Store Name**



Transaction Data by **Store Name**

	nama	Total Transaction	Product Sold	Total Revenue ▾
1.	APOTEK MAJA	63	1,822	Rp104,736,100
2.	APOTEK MERDEKA	68	1,840	Rp92,031,400
3.	KLINIK SAHABAT	62	1,345	Rp85,027,000
4.	APOTEK TAPAK	74	1,564	Rp71,589,900
5.	APOTEK SINAR JAYA	51	1,175	Rp63,048,300
6.	KLINIK GM	69	2,002	Rp48,810,300
7.	APOTEK SAHABAT	49	829	Rp39,895,300
8.	KLINIK DR. ANDRI	29	525	Rp18,609,200

Soal 6 : Additional Complementary Data

A. Tugas :

Dari data yang tersedia, menurut kamu untuk melengkapi analisis nya apakah diperlukan data lain juga? jika iya, sebutkan data apa yang kamu maksud dan mengapa memerlukan data tersebut

Jawaban :

Data pelanggan dapat lebih didetailkan lagi, karena data yang diberikan masih dalam level company. Dengan data yang lebih detail, analisis dari perspektif konsumen dapat dimaksimalkan lagi.

Selain itu, data penjualan dari perusahaan lain dapat membantu untuk melakukan analisis kompetitor. Data ini mungkin bisa didapatkan dari distributor.