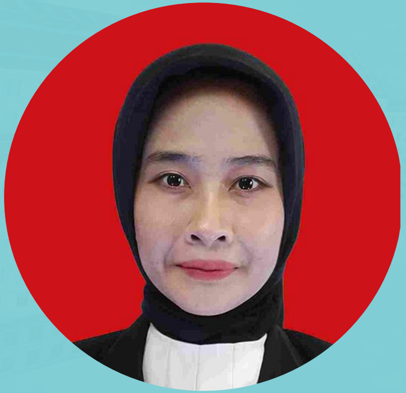
Rakamin Academy X kimia farma

Big Data Analytics Project

Salicyl Datamart and Sales Dashboard Using PostgreSQL & Looker Studio

Kimia Farma Big Data Analytics Project Based Internship Program

Presented by
Nicken Shidqia Nurahman



Nicken Shidqia Nurahman

About Me

Civil engineer graduate with some experience in administration and project management, who is interested in data science.
Detail oriented, and time management person, and familiar with Microsoft Office, Python, SQL and Jupyter.
Motivated to continue to learn and grow as a professional.

My Experience

Rakamin Academy

- Data Science Bootcamp Student – RAKAMIN ACADEMY
  Oct 2023 - Now
- Project Management Masters Degree Student – UNIVERSITAS INDONESIA
  Sep 2021 - Sep 2023
- Engineering Administration and Project Control Staff - PT. ISTAKA KARYA
  Aug 2019 - Sep 2021
- Project Control Intern - PT. ISTAKA KARYA
  Feb 2019 - Jul 2019
- Surveying Laboratory Assistant – UNIVERSITAS TRISAKTI
  Jul 2017- Agust 2019

Query

Soal 1

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

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

Jawaban :
B

Alasan :

- **Query B** akan bekerja lebih baik dan **lebih umum digunakan** di SQL. Query B menggunakan operator "**LIKE**" , Dimana dia akan memfilter kolom "alamat" pada table "pelanggan" dengan character yang diawali "Mat".
- Sementara query a menggunakan "**SUBSTR**" untuk meng-ekstrak 3 character pertama dari kolom "address" dan membandingkannya dengan string " Mat".
- Meskipun mungkin menghasilkan output yang sama, tetapi **query B lebih mudah dibaca** dan **lebih umum digunakan** jika dibandingkan dengan 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.

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

Jawaban :
B

Alasan :

- Menggunakan operator **BETWEEN** pada umumnya lebih mudah dibaca, serta secara **jelas** mengekspresikan rentang tanggal **tanpa perlu** menggunakan >= dan <=

Query

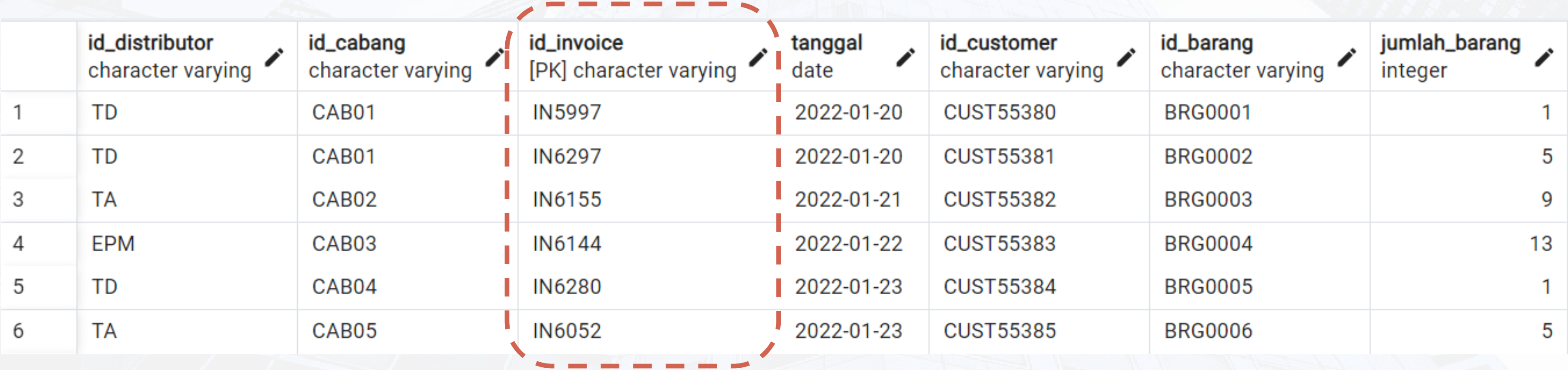
Soal 3

Tentukan Primary Key (PK) dari table penjualan. jelaskan alasannya

Jawaban :
id_invoice merupakan PK dari table penjualan.

Alasan :

- Primary Key merupakan unique identifier pada suatu table. Untuk table penjualan, dipilih id_invoice sebagai primary key karena kolom tersebut merupakan **unique identifier** yang secara unik mengidentifikasi setiap transaksi penjualan.



	id_distributor character varying	id_cabang character varying	id_invoice [PK] character varying	tanggal date	id_customer character varying	id_barang character varying	jumlah_barang integer
1	TD	CAB01	IN5997	2022-01-20	CUST55380	BRG0001	1
2	TD	CAB01	IN6297	2022-01-20	CUST55381	BRG0002	5
3	TA	CAB02	IN6155	2022-01-21	CUST55382	BRG0003	9
4	EPM	CAB03	IN6144	2022-01-22	CUST55383	BRG0004	13
5	TD	CAB04	IN6280	2022-01-23	CUST55384	BRG0005	1
6	TA	CAB05	IN6052	2022-01-23	CUST55385	BRG0006	5

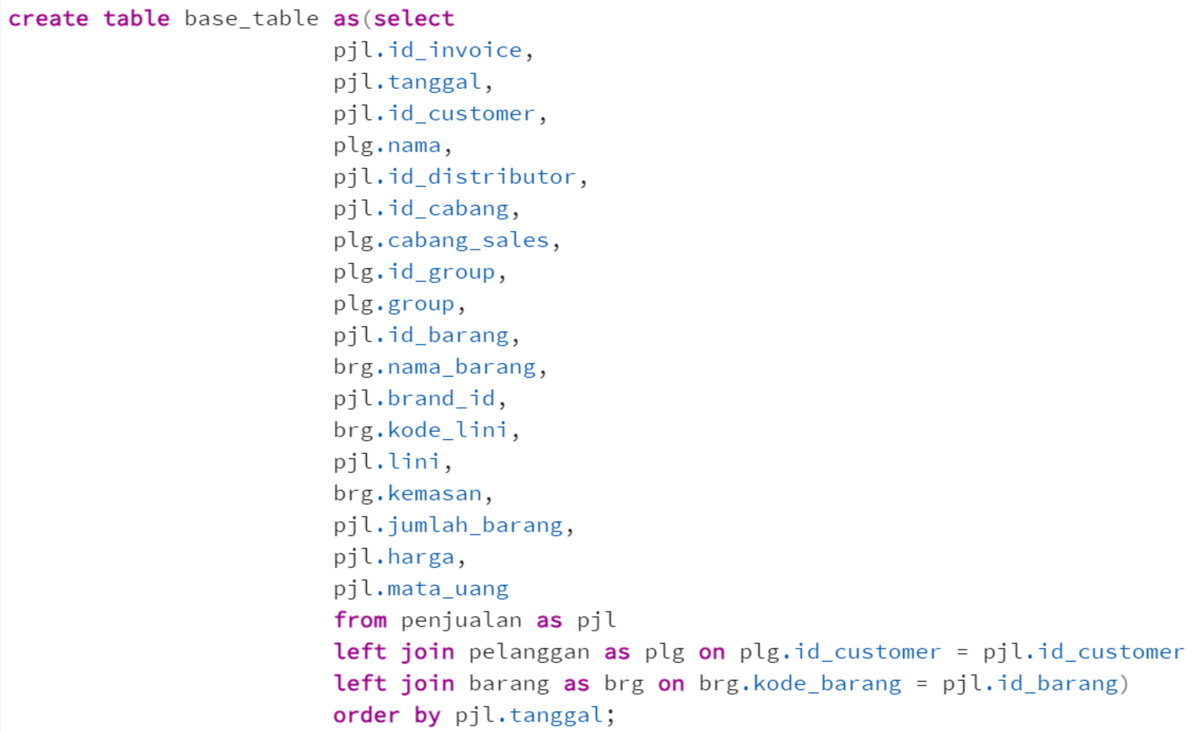
Design Datamart

Soal 4

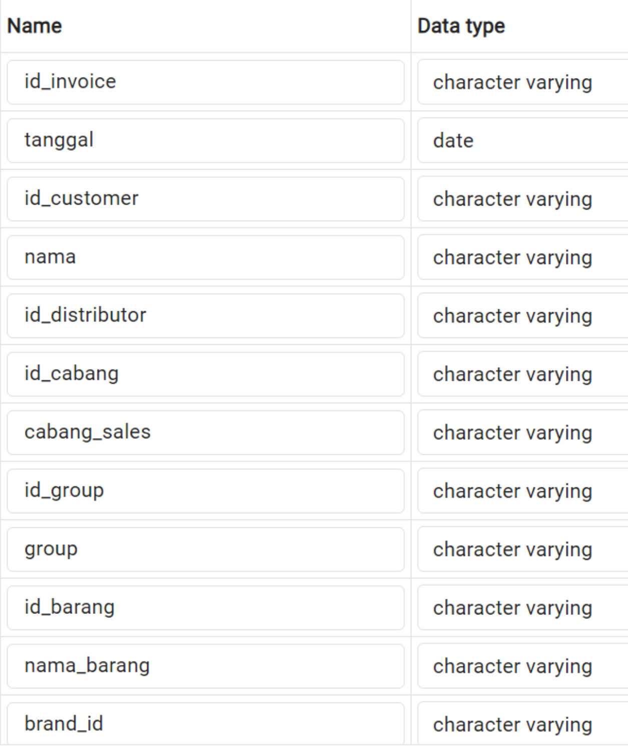
Buatlah design datamart (Terdiri dari tabel base, dan tabel aggregate).

No.	Nama File	Link
1.	create_database.sql • penjualan • pelanggan • barang	https://github.com/nickenshidqia/Big_Data_Analytics_Kimia_Farma/blob/9df3e5deb7ddd779cb3ff7bb21f05313496010a6/create_database.sql
2.	base_table.sql • base_table	https://github.com/nickenshidqia/Big_Data_Analytics_Kimia_Farma/blob/9df3e5deb7ddd779cb3ff7bb21f05313496010a6/base_table.sql
3.	aggregate_table.sql • sales_per_date • sales_per_customer • sales_per_product • sales_per_branch	https://github.com/nickenshidqia/Big_Data_Analytics_Kimia_Farma/blob/9df3e5deb7ddd779cb3ff7bb21f05313496010a6/aggregate_table.sql

Base Table



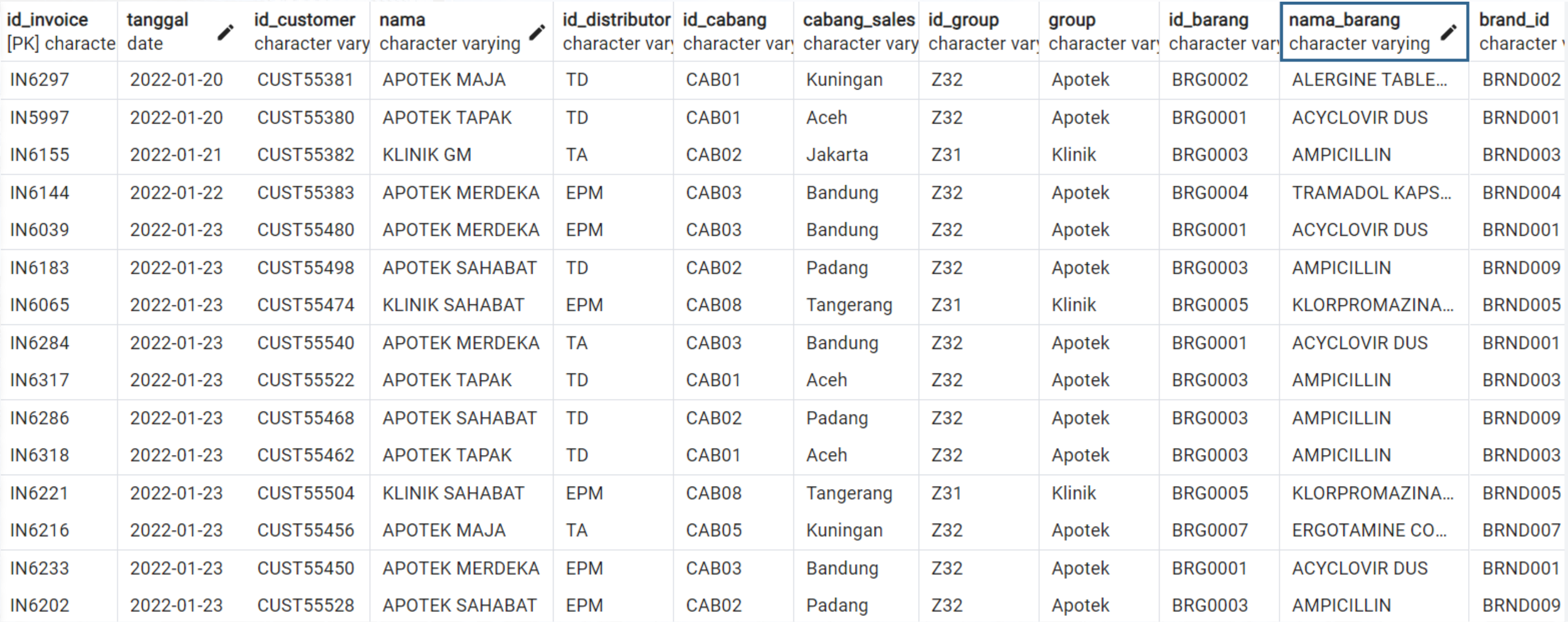
```sql
create table base_table as(select
                        pjl.id_invoice,
                        pjl.tanggal,
                        pjl.id_customer,
                        plg.nama,
                        pjl.id_distributor,
                        pjl.id_cabang,
                        plg.cabang_sales,
                        plg.id_group,
                        plg.group,
                        pjl.id_barang,
                        brg.nama_barang,
                        pjl.brand_id,
                        brg.kode_lini,
                        pjl.lini,
                        brg.kemasan,
                        pjl.jumlah_barang,
                        pjl.harga,
                        pjl.mata_uang
                        from penjualan as pjl
                        left join pelanggan as plg on plg.id_customer = pjl.id_customer
                        left join barang as brg on brg.kode_barang = pjl.id_barang)
                        order by pjl.tanggal;
```



Name	Data type
id_invoice	character varying
tanggal	date
id_customer	character varying
nama	character varying
id_distributor	character varying
id_cabang	character varying
cabang_sales	character varying
id_group	character varying
group	character varying
id_barang	character varying
nama_barang	character varying
brand_id	character varying

Rakamin Academy

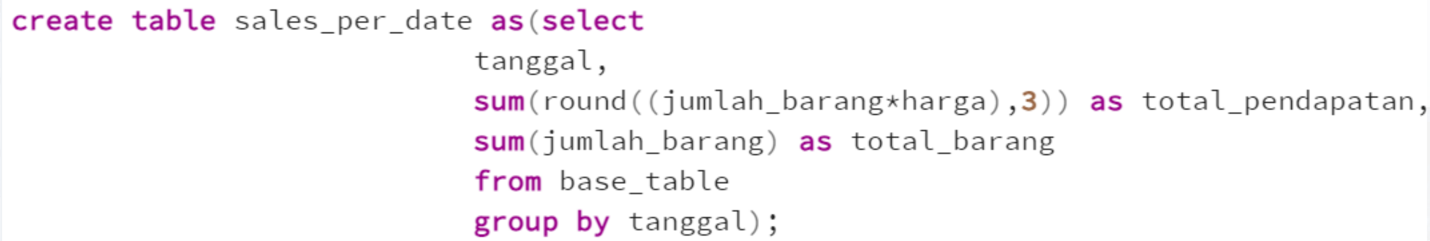
Base Table



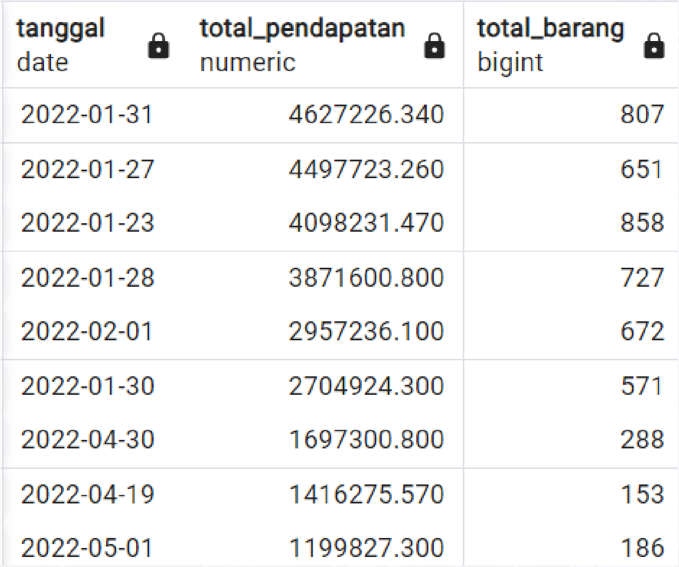
id_invoice	tanggal	id_customer	nama	id_distributor	id_cabang	cabang_sales	id_group	group	id_barang	nama_barang	brand_id
[PK] characte	date	character vary	character varying	character var	character var	character vary	character var	character var	character var	character varying	character
IN6297	2022-01-20	CUST55381	APOTEK MAJA	TD	CAB01	Kuningan	Z32	Apotek	BRG0002	ALERGINE TABLE...	BRND002
IN5997	2022-01-20	CUST55380	APOTEK TAPAK	TD	CAB01	Aceh	Z32	Apotek	BRG0001	ACYCLOVIR DUS	BRND001
IN6155	2022-01-21	CUST55382	KLINIK GM	TA	CAB02	Jakarta	Z31	Klinik	BRG0003	AMPICILLIN	BRND003
IN6144	2022-01-22	CUST55383	APOTEK MERDEKA	EPM	CAB03	Bandung	Z32	Apotek	BRG0004	TRAMADOL KAPS...	BRND004
IN6039	2022-01-23	CUST55480	APOTEK MERDEKA	EPM	CAB03	Bandung	Z32	Apotek	BRG0001	ACYCLOVIR DUS	BRND001
IN6183	2022-01-23	CUST55498	APOTEK SAHABAT	TD	CAB02	Padang	Z32	Apotek	BRG0003	AMPICILLIN	BRND009
IN6065	2022-01-23	CUST55474	KLINIK SAHABAT	EPM	CAB08	Tangerang	Z31	Klinik	BRG0005	KLORPROMAZINA...	BRND005
IN6284	2022-01-23	CUST55540	APOTEK MERDEKA	TA	CAB03	Bandung	Z32	Apotek	BRG0001	ACYCLOVIR DUS	BRND001
IN6317	2022-01-23	CUST55522	APOTEK TAPAK	TD	CAB01	Aceh	Z32	Apotek	BRG0003	AMPICILLIN	BRND003
IN6286	2022-01-23	CUST55468	APOTEK SAHABAT	TD	CAB02	Padang	Z32	Apotek	BRG0003	AMPICILLIN	BRND009
IN6318	2022-01-23	CUST55462	APOTEK TAPAK	TD	CAB01	Aceh	Z32	Apotek	BRG0003	AMPICILLIN	BRND003
IN6221	2022-01-23	CUST55504	KLINIK SAHABAT	EPM	CAB08	Tangerang	Z31	Klinik	BRG0005	KLORPROMAZINA...	BRND005
IN6216	2022-01-23	CUST55456	APOTEK MAJA	TA	CAB05	Kuningan	Z32	Apotek	BRG0007	ERGOTAMINE CO...	BRND007
IN6233	2022-01-23	CUST55450	APOTEK MERDEKA	EPM	CAB03	Bandung	Z32	Apotek	BRG0001	ACYCLOVIR DUS	BRND001
IN6202	2022-01-23	CUST55528	APOTEK SAHABAT	EPM	CAB02	Padang	Z32	Apotek	BRG0003	AMPICILLIN	BRND009

Aggregate Table

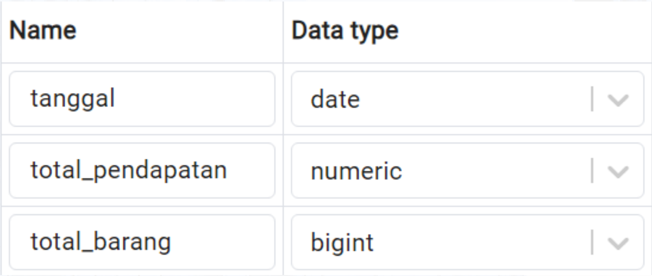
A. Sales per date



```
create table sales_per_date as(select
                           tanggal,
                           sum(round((jumlah_barang*harga),3)) as total_pendapatan,
                           sum(jumlah_barang) as total_barang
                           from base_table
                           group by tanggal);
```



tanggal date	total_pendapatan numeric	total_barang bigint
2022-01-31	4627226.340	807
2022-01-27	4497723.260	651
2022-01-23	4098231.470	858
2022-01-28	3871600.800	727
2022-02-01	2957236.100	672
2022-01-30	2704924.300	571
2022-04-30	1697300.800	288
2022-04-19	1416275.570	153
2022-05-01	1199827.300	186

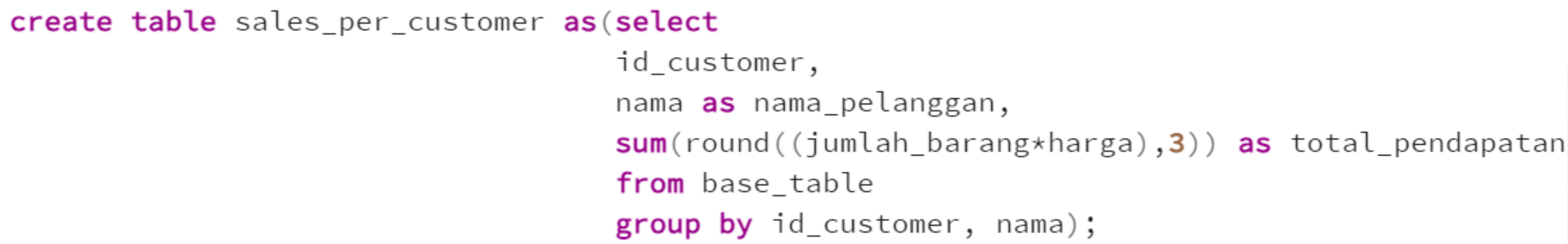


Name	Data type
tanggal	date
total_pendapatan	numeric
total_barang	bigint

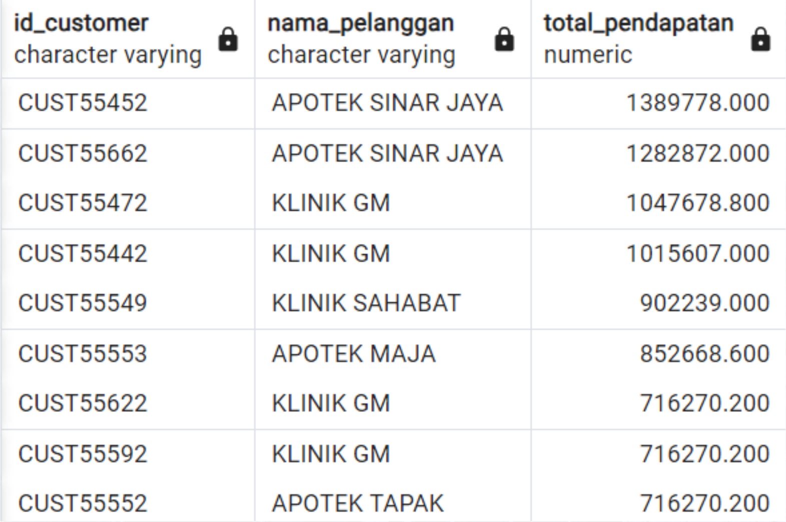
Insight :
Berdasarkan query sales per date diperoleh total **pendapatan tertinggi** adalah tanggal **31-01-2022** dengan hasil 4,627,226

Aggregate Table

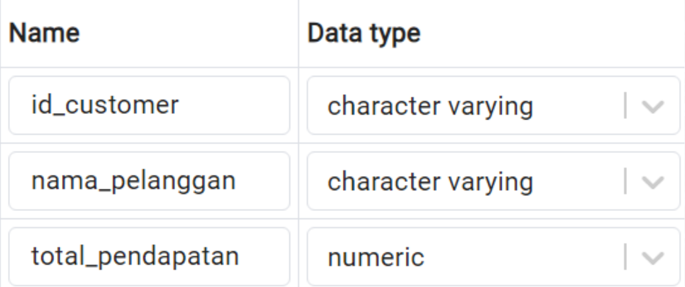
B. Sales per customer



```
create table sales_per_customer as(select
                                        id_customer,
                                        nama as nama_pelanggan,
                                        sum(round((jumlah_barang*harga),3)) as total_pendapatan
                                        from base_table
                                        group by id_customer, nama);
```



id_customer character varying	nama_pelanggan character varying	total_pendapatan numeric
CUST55452	APOTEK SINAR JAYA	1389778.000
CUST55662	APOTEK SINAR JAYA	1282872.000
CUST55472	KLINIK GM	1047678.800
CUST55442	KLINIK GM	1015607.000
CUST55549	KLINIK SAHABAT	902239.000
CUST55553	APOTEK MAJA	852668.600
CUST55622	KLINIK GM	716270.200
CUST55592	KLINIK GM	716270.200
CUST55552	APOTEK TAPAK	716270.200

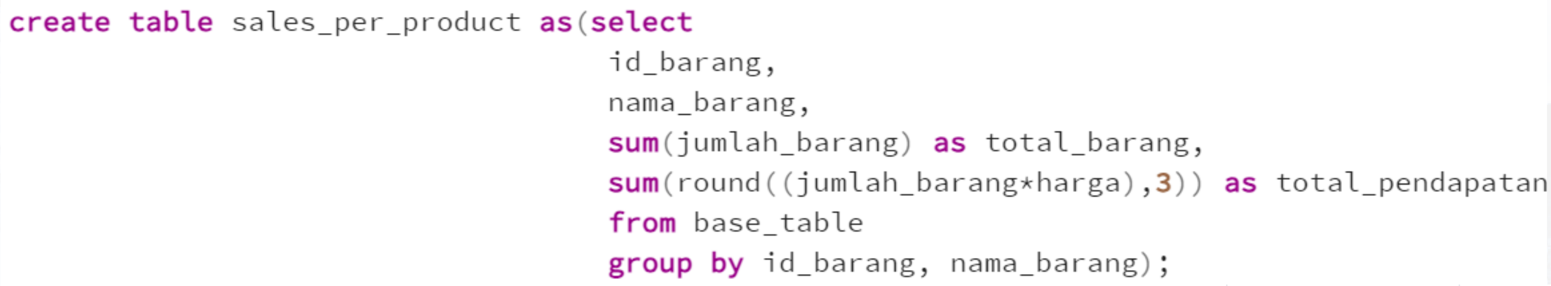


Name	Data type
id_customer	character varying
nama_pelanggan	character varying
total_pendapatan	numeric

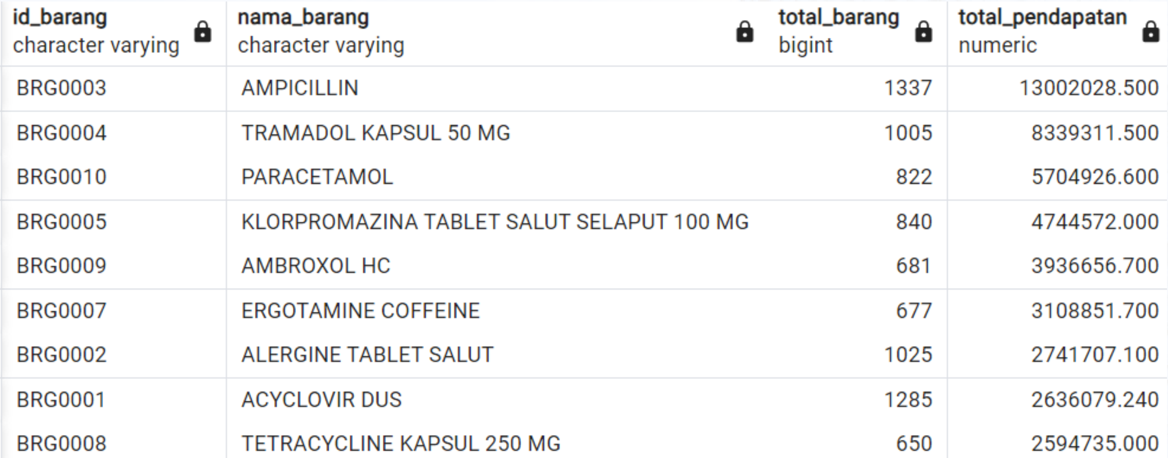
Insight :
Berdasarkan query sales per customer diperoleh total **pendapatan tertinggi** adalah **Apotek Sinar Jaya** dengan hasil 1,389,778

Aggregate Table

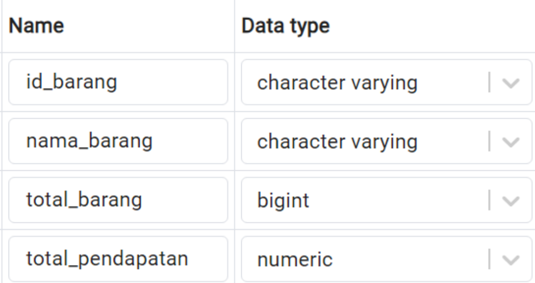
C. Sales per product



```
create table sales_per_product as(select
                                   id_barang,
                                   nama_barang,
                                   sum(jumlah_barang) as total_barang,
                                   sum(round((jumlah_barang*harga),3)) as total_pendapatan
                                   from base_table
                                   group by id_barang, nama_barang);
```



id_barang	nama_barang	total_barang	total_pendapatan
character varying	character varying	bigint	numeric
BRG0003	AMPICILLIN	1337	13002028.500
BRG0004	TRAMADOL KAPSUL 50 MG	1005	8339311.500
BRG0010	PARACETAMOL	822	5704926.600
BRG0005	KLORPROMAZINA TABLET SALUT SELAPUT 100 MG	840	4744572.000
BRG0009	AMBROXOL HC	681	3936656.700
BRG0007	ERGOTAMINE COFFEINE	677	3108851.700
BRG0002	ALERGINE TABLET SALUT	1025	2741707.100
BRG0001	ACYCLOVIR DUS	1285	2636079.240
BRG0008	TETRACYCLINE KAPSUL 250 MG	650	2594735.000

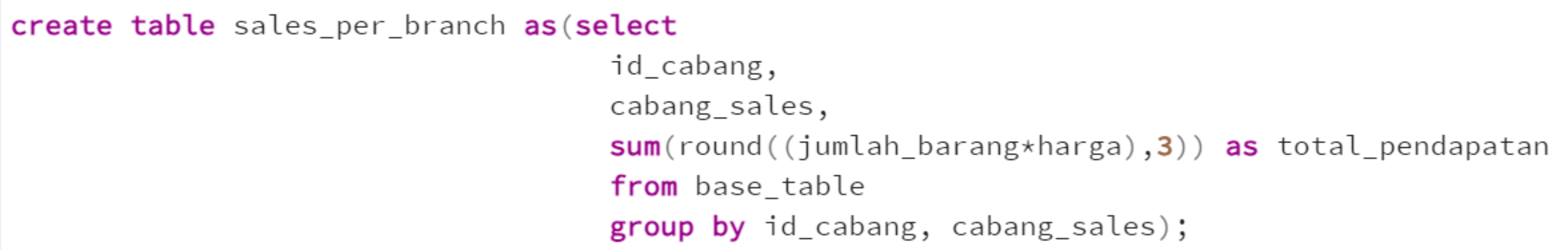


Name	Data type
id_barang	character varying
nama_barang	character varying
total_barang	bigint
total_pendapatan	numeric

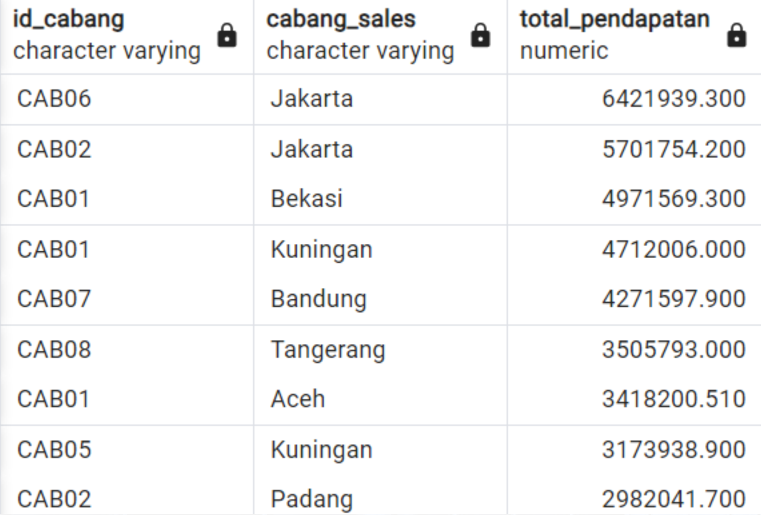
Insight :
Berdasarkan query sales per product diperoleh total **pendapatan tertinggi** adalah **Ampicillin** dengan hasil 13,002,028

Aggregate Table

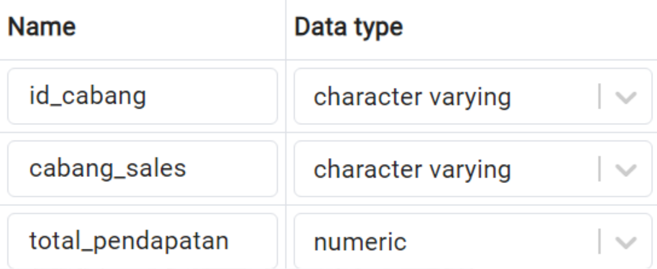
D. Sales per branch



```
create table sales_per_branch as(select
                                    id_cabang,
                                    cabang_sales,
                                    sum(round((jumlah_barang*harga),3)) as total_pendapatan
                                    from base_table
                                    group by id_cabang, cabang_sales);
```



id_cabang character varying	cabang_sales character varying	total_pendapatan numeric
CAB06	Jakarta	6421939.300
CAB02	Jakarta	5701754.200
CAB01	Bekasi	4971569.300
CAB01	Kuningan	4712006.000
CAB07	Bandung	4271597.900
CAB08	Tangerang	3505793.000
CAB01	Aceh	3418200.510
CAB05	Kuningan	3173938.900
CAB02	Padang	2982041.700



Name	Data type
id_cabang	character varying
cabang_sales	character varying
total_pendapatan	numeric

Insight :
Berdasarkan query sales per branch diperoleh total **pendapatan tertinggi** adalah **cabang Jakarta** dengan hasil 6,421,939

KIMIA FARMA SALES DASHBOARD

kimia farma

Author : Nicken Shidqia .N

Soal 5

Data Visualization

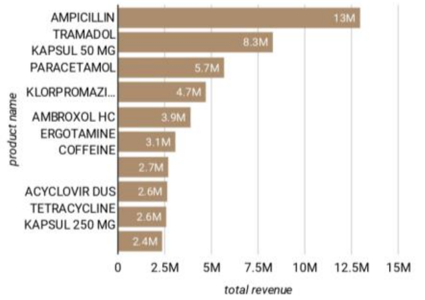
Month ▾

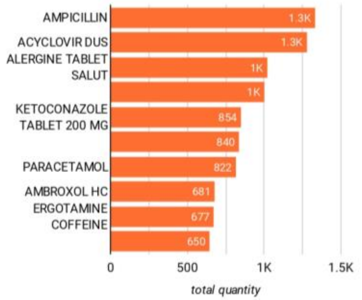
Brand ▾

Product quantity
9,176

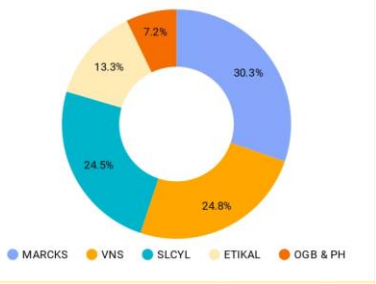
Total Revenue
49.2M

Revenue & Quantity Per Product

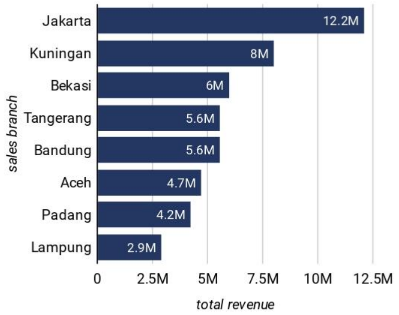




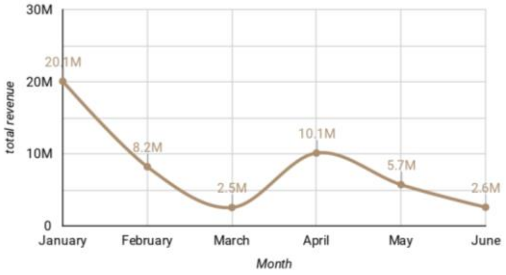
Revenue Per Brand

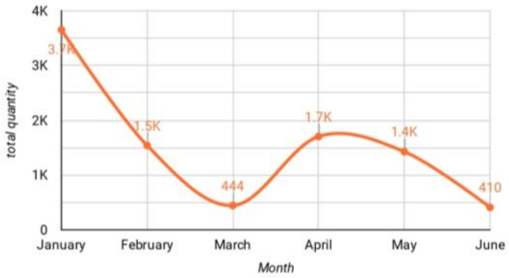


Revenue Per Sales Branch

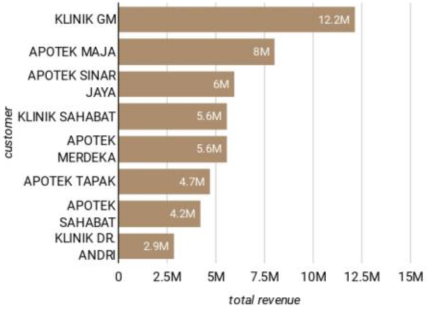


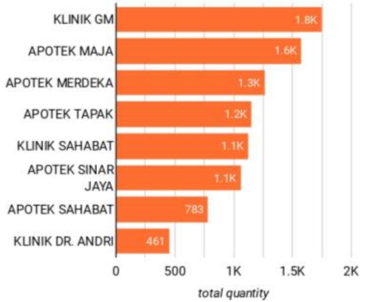
Revenue & Quantity Per Month





Revenue & Quantity Per Customer





Link : https://lookerstudio.google.com/reporting/8acda551-6e3e-4638-ae4f-3e9a20847100

Author : Nicken Shidqia .N

SALICYL SALES DASHBOARD

Month ▾

Bra... (1) ▾

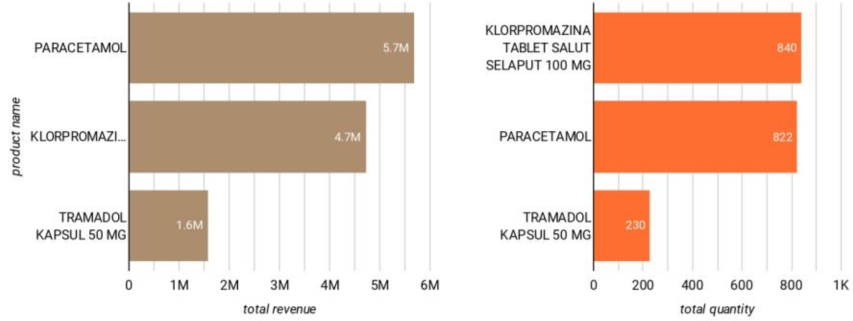
Product quantity
1,892

Total Revenue
12.0M

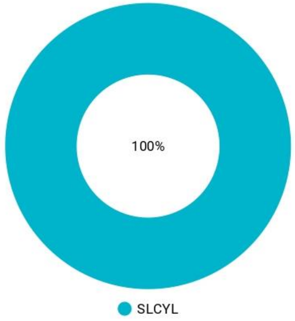
Soal 5

Data Visualization

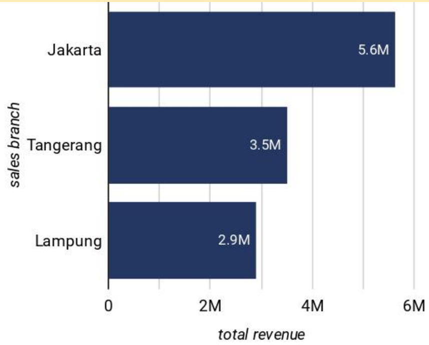
Revenue & Quantity Per Product



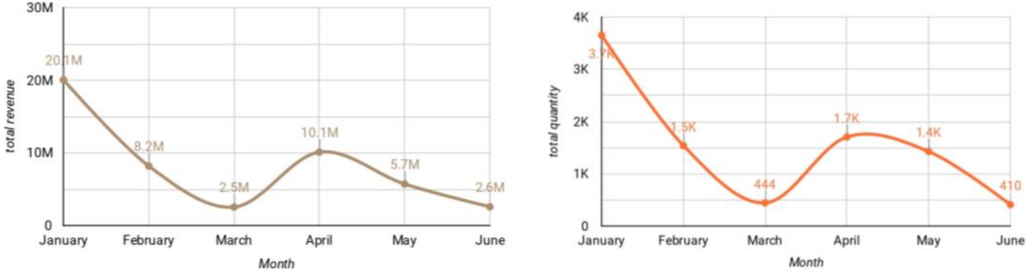
Revenue Per Brand



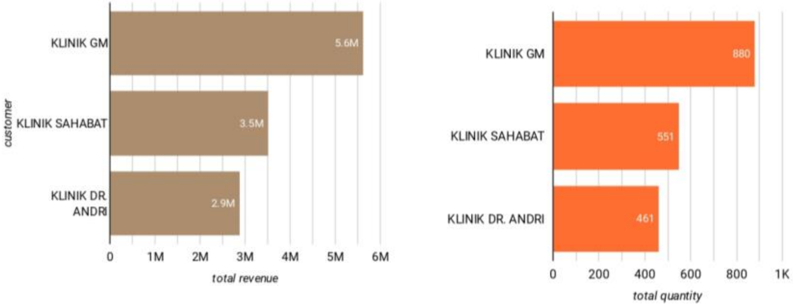
Revenue Per Sales Branch



Revenue & Quantity Per Month



Revenue & Quantity Per Customer



Link : https://lookerstudio.google.com/reporting/8acda551-6e3e-4638-ae4f-3e9a20847100

Insight

All Kimia Farma Brand Sales

- The highest revenue based on product category is Ampicillin with 13 M and total quantity 1.3K
- The highest revenue based on brand category is Marcks with 30.3%, followed by VNS 24.8%, and SLCYL 24.5%.
- The highest revenue based on sales branch category is Jakarta with 12.2 M.
- Sales of Kimia Farma are fluctuating with the highest revenue is happened on January 2022 with 20.1 M, while the lowest revenue is happened on March 2022 with 2.5 M
- The highest revenue based on customer category is Klinik GM 12.2 M with total quantity 1.8K.

Salicyl Brand Sales

- The highest revenue based on product category is Paracetamol with 5.7 M and total quantity 840.
- The highest revenue based on sales branch category is Jakarta with 5.6 M.
- Total revenue of sales Salicyl product is 12 M with total quantity 1,892.
- The highest revenue based on customer category is Klinik GM 5.6 M with total quantity 880.

Additional Complementary Data

Soal 6

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

Geographic Information:

- Latitude and longitude of each distributor’s and brach location.
- City, state, or region where distributors and brach are located.

Promotional Activities:

- Promotion Type, example discounts, bundle offers, seasonal promotions.
- Promotion Duration : Start and end dates for each promotional activity.
- Promotion Channels: Where the promotions are advertised or offered (in-store, online, specific platforms).

Competitor Data:

- Competitor Product Information
- Competitor Pricing
- Market Share
- Promotional Strategies
- Customer Reviews and Feedback

Link Portolio On Github :

https://github.com/nickenshidqia/Big_Data_Analytics_Kimia_Farma

LinkedIn:

https://www.linkedin.com/in/nickenshidqia/

Thank You

Rakamin Academy X kimia farma