

# Soal & Template Jawaban

Task 5

Nama : Bayu Bagus Bagaswara

# 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

Alasan : Untuk mempercepat proses query biasanya menggunakan index. Pada kasus diatas index nya adalah kolom alamat. Jawaban A akan menghasilkan pencarian data lebih lambat karena menggunakan fungsi pada kolom yang terindeks, yakni fungsi SUBSTR. Hal tersebut akan membatalkan proses index pada kolom alamat. Sedangkan Jawaban B akan lebih cepat query nya karena langsung membuat index pada kolom alamat.

# 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

Alasan : Operator BETWEEN digunakan dalam kondisi WHERE untuk memfilter record dalam rentang yang ditentukan. Rentang nilai dapat berupa string, angka, atau tanggal. Rentang nilai harus ditentukan dengan operator AND.

## Soal 3: Menentukan Primary Key

### A. Tugas

Tentukan primary key dari table penjualan. jelaskan alasannya

- B. Jawaban & Penjelasan : id\_invoice dan id\_barang. Invoice memuat rincian transaksi dan pembayaran dari penjual kepada pembeli, sedangkan id\_barang adalah barang yang telah dijual. Hasil penggabungan kolom id\_invoice dan id\_barang akan menciptakan data yang unik untuk memberikan informasi penjualan.

## Soal 4: Design Datamart

### A. Tugas

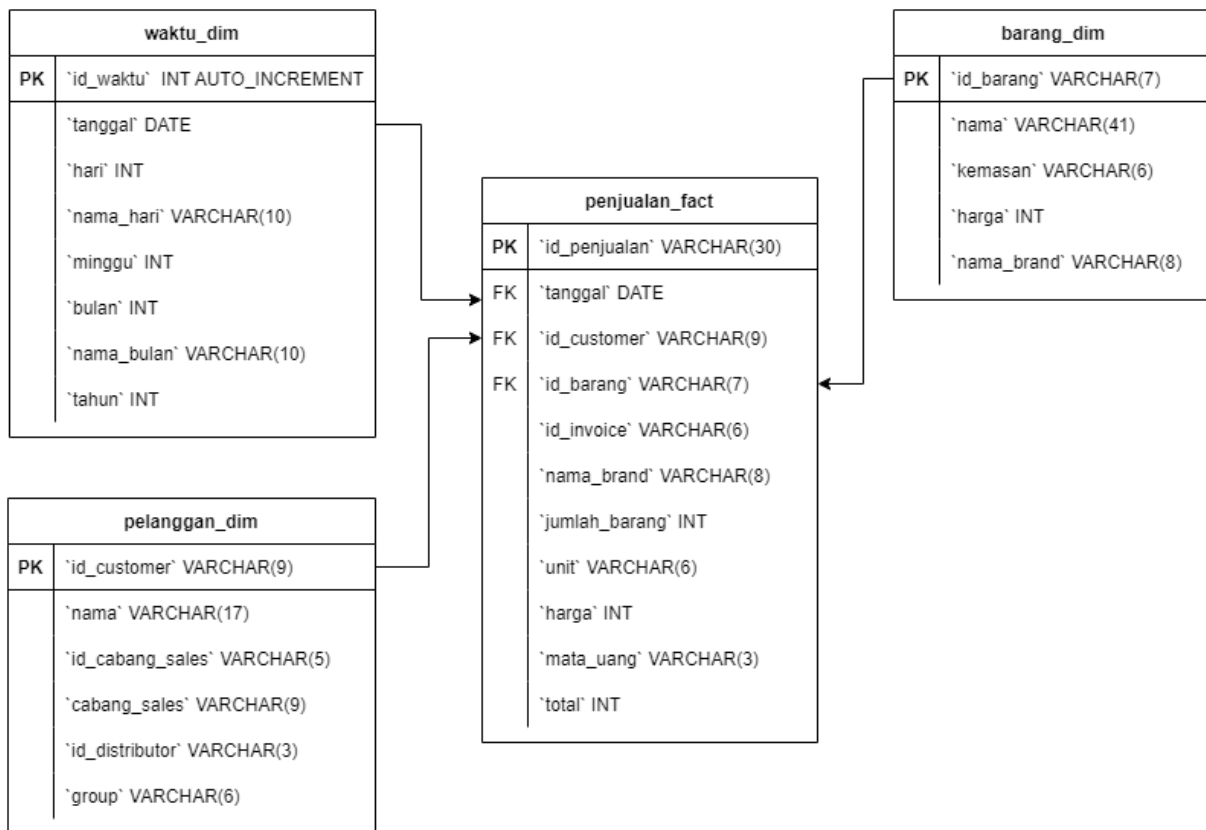
Buatlah design datamart (Terdiri dari tabel base, dan tabel aggregate). Upload file query dalam gdrive mu (pastikan dapat diakses public). Lalu masukkan linknya di tabel di bawah, dan cantumkan juga screenshoot query nya (jika lebih dari 1 file, maka masing masing file di-screenshoot)

Silahkan tambah halaman jika dibutuhkan


### A. Jawaban :

No	Nama File	Link
1	waktu_dim	<a href="https://drive.google.com/file/d/1nGfGJ0mNUVCeixbi-EHK8JCCsvRSR6xs/view?usp=sharing">https://drive.google.com/file/d/1nGfGJ0mNUVCeixbi-EHK8JCCsvRSR6xs/view?usp=sharing</a>
2	barang_dim	<a href="https://drive.google.com/file/d/1mL9VVhHcUeD1RMuq9fIB_SermQbT3XGw/view?usp=sharing">https://drive.google.com/file/d/1mL9VVhHcUeD1RMuq9fIB_SermQbT3XGw/view?usp=sharing</a>
3	pelanggan_dim	<a href="https://drive.google.com/file/d/1e-zBRcWoERET4pVbSQf0v3l2HZQSLHEm/view?usp=sharing">https://drive.google.com/file/d/1e-zBRcWoERET4pVbSQf0v3l2HZQSLHEm/view?usp=sharing</a>
4.	penjualan_fact	<a href="https://drive.google.com/file/d/1cujiXgEe0yHuFeUp-rQRLE83nRCbXeEY/view?usp=sharing">https://drive.google.com/file/d/1cujiXgEe0yHuFeUp-rQRLE83nRCbXeEY/view?usp=sharing</a>

# Rancangan ER Diagram



# Table Base “<<waktu\_dim>>”



```
1  create table `waktu_dim` (  
2      `id_waktu` int auto_increment primary key,  
3      `tanggal` date,  
4      `hari` int,  
5      `nama_hari` varchar(10),  
6      `minggu` int,  
7      `bulan` int,  
8      `nama_bulan` varchar(10),  
9      `tahun` year  
10 );
```




## Table Base “<<barang\_dim>>”



```
1 create table `barang_dim`(  
2     `id_barang` varchar(7),  
3     `nama` varchar(41),  
4     `kemasan` varchar(6),  
5     `harga` int,  
6     `nama_brand` varchar(8),  
7     primary key (id_barang)  
8 );
```

## Table Base “<<pelanggan\_dim>>”



```
1 create table `pelanggan_dim`(  
2     `id_customer` varchar(9),  
3     `nama` varchar(17),  
4     `id_cabang_sales` varchar(5),  
5     `cabang_sales` varchar(9),  
6     `id_distributor` varchar(3),  
7     `group` varchar(6),  
8     primary key(id_customer)  
9 );
```

## Table Base “<<waktu\_dim>>”

column	data type	description	transformation
id_waktu	int	primary key	-
tanggal	date	data tanggal	Mengubah datetime menjadi date
hari	int	hari dalam minggu	Pengenaan fungsi DAYOFWEEK
nama_hari	varchar	hari dalam minggu	Pengenaan fungsi DAYNAME
minggu	int	minggu dalam tahun	Pengenaan fungsi WEEK
bulan	int	bulan	Pengenaan fungsi MONTH
nama_bulan	varchar	nama bulan	Pengenaan fungsi MONTHNAME
tahun	year	tahun	Pengenaan fungsi YEAR

## Table Base “<<barang\_dim>>”

column	data type	description	transformation
id_barang	varchar	primary key	Diambil dari kode_barang
nama	varchar	nama barang	Diambil dari nama barang
kemasan	varchar	kemasan barang	Diambil dari kemasan atau unit
harga	int	harga per kemasan	Diambil dari harga barang
nama_brand	varchar	nama brand per barang	Diambil dari nama brand

## Table Base “<<pelanggan\_dim>>”

column	data type	description	transformation
id_customer	varchar	primary key	Diambil dari id_customer
nama	varchar	nama customer	Diambil dari nama di table pelanggan
id_cabang_sales	varchar	id cabang sales (fk)	Diambil dari id_cabang_sales di table pelanggan
cabang_sales	varchar	nama cabang sales	Diambil dari nama cabang sales di table pelanggan
id_distributor	varchar	id distributor (fk)	Diambil dari id_distributor di table pelanggan
group	varchar	nama group per customer	Diambil dari kolom group

# Table Aggregate “<<penjualan\_aggregation>>”

```
1 create table `penjualan_fact`(  
2     `id_penjualan` varchar(30),  
3     `id_invoice` varchar(6),  
4     `tanggal` date,  
5     `id_customer` varchar(9),  
6     `id_barang` varchar(7),  
7     `nama_brand` varchar(8),  
8     `jumlah_barang` int,  
9     `unit` varchar(6),  
10    `harga` int,  
11    `mata_uang` varchar(3),  
12    `total` int,  
13    primary key (id_penjualan)  
14 );
```

## Table Aggregate “<<penjualan\_fact>>”

column	data type	description	transformation
id_penjualan	varchar	primary key	-
id_invoice	varchar	id untuk invoice	-
tanggal	date	tanggal penjualan	-
id_customer	varchar	fk dari id customer	-
id_barang	varchar	fk dari id barang	-
nama_barang	varchar	nama barang yang terjual	-
jumlah_barang	int	jumlah barang yang terjual	-
unit	varchar	satuan barang yang terjual	-

## Table Aggregate “<<Nama Tabel>>”

column	data type	description	transformation
harga	int	harga barang per unit/kemasan	-
mata_uang	varchar	mata uang penjualan	-
total	int	total penjualan	perkalian jumlah barang dan 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://datastudio.google.com/reporting/767d6416-0837-4d10-a36c-8dde528e39b9>



## SALICYL SALES DASHBOARD

berdasarkan data penjualan brand Salicyl bulan Januari sampai Juni tahun 2022.

### Overview

Total Transaksi

59

↑ 293.3%

Total Penjualan (Rupiah)

56.742.000

↑ 297.2%

Jumlah Customer

3

0.0%

Jumlah Barang

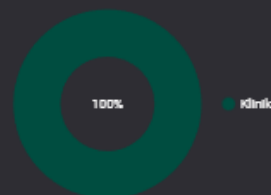
2

0.0%

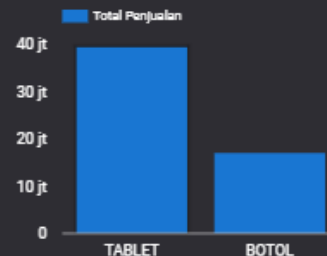
Total Penjualan tiap Bulan



Group Customer



Penjualan per Unit



### Barang

Deskripsi Barang

Jumlah Barang Keseluruhan

Nama Barang	Unit	Harga
1. KLORPROMAZINA TABLET SALUT SELAPUT 100 MG	TABLET	1.645.000

1 rb

Jumlah Barang

## 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

A. Jawaban : Iya, seperti data pelanggan dan distributor. Data pelanggan yang diberikan kurang detail, karena memiliki id\_customer berbeda tetapi detail nama dan cabang\_salesnya sama. Sedangkan data seperti distributor dan cabang\_sales hanya bisa diambil nilai id nya, tidak bisa mendapatkan datanya lebih detail, karena dengan data tersebut kita bisa menganalisa kinerja penjualan cabang\_sales dan distributor .