

Soal & Template Jawaban

Task 5

Nama : **Ahmad Maulana** Malik Fattah

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%'`

Jawaban : a

Alasan :

Pada kasus ini, dibutuhkan untuk mencari data **alamat** berdasarkan **3** huruf awalnya, sehingga *query* pertama sudah cukup memenuhi kebutuhan.

Penggunaan *keyword* **LIKE** cocok digunakan apabila ingin melakukan pencarian yang lebih kompleks, misal untuk mencari potongan karakter di tengah kata/kalimat.

Untuk optimasi dan penyederhanaan lebih lanjut, *query* pertama juga dapat diganti dengan menggunakan *keyword* **LEFT**. Misal: `SELECT * FROM foo WHERE LEFT(foo, 3) = 'Bar';`

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'`

Jawaban : b

Alasan :

Keyword **BETWEEN** disediakan SQL untuk secara khusus melakukan filtrasi data rentang. Selain performa eksekusi yang lebih baik, *query* kedua juga dapat dituliskan secara lebih sederhana dan mudah dimengerti.

Soal 3: Menentukan Primary Key

A. Tugas

Tentukan primary key dari table penjualan. jelaskan alasannya

B. Jawaban & Penjelasan : Kolom **id_invoice**.

Salah satu syarat sebuah kolom menjadi *primary key* adalah nilainya harus unik untuk tiap *record*.

Table penjualan memiliki 350 *record* dengan dua kolom yang bersifat unik untuk tiap *record*, yaitu **id_invoice** dan **id_customer**.

Mengingat ini adalah *table* penjualan, maka **id_invoice** lebih representatif untuk digunakan sebagai *primary key*.

Soal 4: Design Datamart

A. Tugas

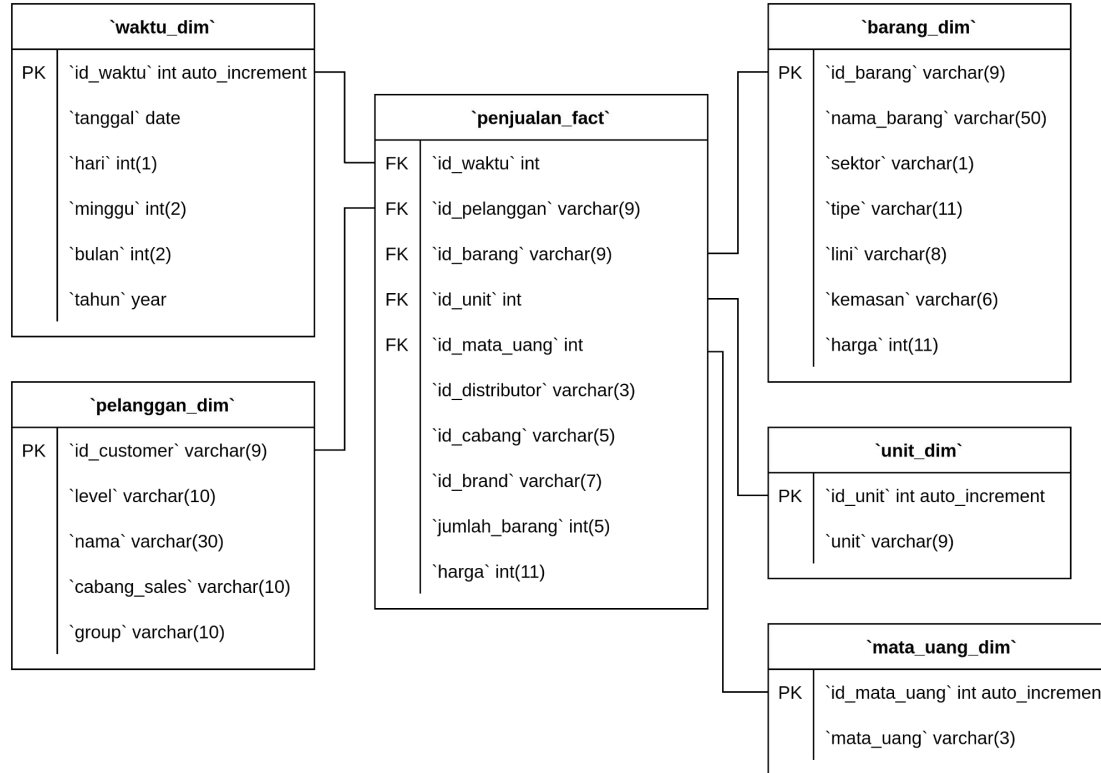
Buatlah design datamart (Terdiri dari tabel base, dan tabel aggregate). Upload file query dalam gdrive mu (pastikan dapat diakses public). Lalu masukkan linknya di tabel di bawah, dan cantumkan juga screenshoot query nya (jika lebih dari 1 file, maka masing masing file di-screenshoot)

Silahkan tambah halaman jika dibutuhkan

B. Jawaban : Pembuatan Datamart utamanya dilakukan pada file **03_datamart.sql**

No	Nama File	Link
1	01_create_tables.sql	https://s.id/18bad
2	02_data_exploration.sql	https://s.id/18baj
3	03_datamart.sql	https://s.id/18c8d

Rancangan ER Diagram



More detailed image [here](#).

Table Base waktu_dim

```
15● CREATE TABLE IF NOT EXISTS `waktu_dim` (  
16   `id_waktu` INT AUTO_INCREMENT PRIMARY KEY  
17   , `tanggal` DATE  
18   , `hari` INT(1)  
19   , `minggu` INT(2)  
20   , `bulan` INT(2)  
21   , `tahun` YEAR  
22 );
```


Table Base **waktu_dim**

waktu_dim mengacu pada data di tabel **penjualan** kolom **tanggal**.

column	data type	description	transformation
id_waktu	int(11)	Primary Key	-
tanggal	date	Data tanggal	Mengubah tipe data Datetime ke Date
hari	int(1)	Hari dalam minggu	Pengenaan fungsi DAYOFWEEK()
minggu	int(2)	Minggu dalam tahun	Pengenaan fungsi WEEK()
bulan	int(2)	Bulan	Pengenaan fungsi MONTH()
tahun	year(4)	Tahun	Pengenaan fungsi YEAR()

Table Base pelanggan_dim

```
24 CREATE TABLE IF NOT EXISTS `pelanggan_dim` (  
25   `id_pelanggan` VARCHAR(9) PRIMARY KEY  
26   , `level` VARCHAR(10)  
27   , `nama` VARCHAR(30)  
28   , `cabang_sales` VARCHAR(10)  
29   , `group` VARCHAR(10)  
30 );  
31
```

Table Base **pelanggan_dim**

pelanggan_dim mengacu pada data di tabel **pelanggan**.

column	data type	description	transformation
id_pelanggan	varchar(9)	Primary Key	Mengubah nama PK dari id_customer menjadi id_pelanggan
level	varchar(10)	Tingkat pelanggan	Memperluas cakupan VARCHAR ke 10
nama	varchar(30)	Nama pelanggan	Memperluas cakupan VARCHAR ke 30
cabang_sales	varchar(10)	Lokasi cabang sales	Memperluas cakupan VARCHAR ke 10
group	varchar(10)	Grup pelanggan	Memperluas cakupan VARCHAR ke 10

Table Base **barang_dim**

```
32 CREATE TABLE IF NOT EXISTS `barang_dim` (  
33   `id_barang` VARCHAR(9) PRIMARY KEY  
34   , `nama_barang` VARCHAR(50)  
35   , `sektor` VARCHAR(1)  
36   , `tipe` VARCHAR(11)  
37   , `brand` VARCHAR(10)  
38   , `kemasan` VARCHAR(6)  
39   , `harga` INT(11)  
40 );  
41
```

Table Base **barang_dim**

barang_dim mengacu pada data di tabel **barang** dan **barang_ds**.

column	data type	description	transformation
id_barang	varchar(9)	Primary Key	Mengubah nama PK dari barang_id menjadi id_barang
nama_barang	varchar(50)	Nama barang	Memperluas cakupan VARCHAR ke 50
sektor	varchar(1)	Sektor barang	-
tipe	varchar(11)	Tipe barang	Menggunakan isi dari kolom nama_tipe di tabel barang .
brand	varchar(10)		Memperluas cakupan VARCHAR ke 10
kemasan	varchar(6)		-
harga	int(11)		Mengubah tipe data dari decimal ke int .

Table Base **unit_dim**

```
42 CREATE TABLE IF NOT EXISTS `unit_dim` (  
43     `id_unit` INT AUTO_INCREMENT PRIMARY KEY  
44     , `unit` VARCHAR(9)  
45 );  
46
```

Table Base **unit_dim**

unit_dim mengacu pada data di tabel **penjualan** kolom **unit**.

column	data type	description	transformation
id_unit	int(11)	Primary Key	-
unit	varchar(9)	Nama satuan (<i>unit</i>)	-

Table Base mata_uang_dim

```
46  
47 CREATE TABLE IF NOT EXISTS `mata_uang_dim` (  
48     `id_mata_uang` INT AUTO_INCREMENT PRIMARY KEY  
49     , `mata_uang` VARCHAR(3)  
50 );  
51
```


Table Base **mata_uang_dim**

mata_uang_dim pada data di tabel **penjualan** kolom **mata_uang**.

column	data type	description	transformation
id_mata_uang	int(11)	Primary Key	-
mata_uang	varchar(3)	Nama mata uang	-

Table Aggregate **penjualan_fact**

```
52 CREATE TABLE IF NOT EXISTS `penjualan_fact` (  
53   `id_waktu` INT  
54   , `id_pelanggan` VARCHAR(9)  
55   , `id_barang` VARCHAR(9)  
56   , `id_unit` INT  
57   , `id_mata_uang` INT  
58   , `id_distributor` VARCHAR(3)  
59   , `id_cabang` VARCHAR(5)  
60   , `id_brand` VARCHAR(7)  
61   , `jumlah_barang` INT(5)  
62   , `harga` INT(11)  
63 );
```

Table Aggregate **penjualan_fact**

column	data type	description	transformation
id_waktu	int(11)	FK waktu_dim	-
id_pelanggan	varchar(9)	FK pelanggan_dim	-
id_barang	varchar(9)	FK barang_dim	-
id_unit	int(11)	FK unit_dim	-
id_mata_uang	int(11)	FK mata_uang_dim	-

Table Aggregate **penjualan_fact**

column	data type	description	transformation
id_distributor	varchar(3)	ID unik distributor	-
id_cabang	varchar(5)	ID unik cabang	-
id_brand	varchar(5)	ID unik <i>brand</i>	-
jumlah_barang	int(5)	Jumlah barang	Mengubah tipe data dari decimal ke int .
harga	int(11)	Harga barang	Mengubah tipe data dari decimal ke int .

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

B. Jawaban :

Link visualisasi (ex link Google Data Studio) :

<https://datastudio.google.com/reporting/c9816161-2235-476e-a9c5-060651b20109>

Kimia Farma Salicyl Brand Sales

1 Jan 2022 - 31 Mar 2022

total_penjualan
Rp12.860

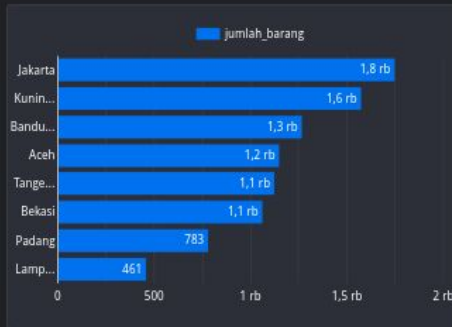
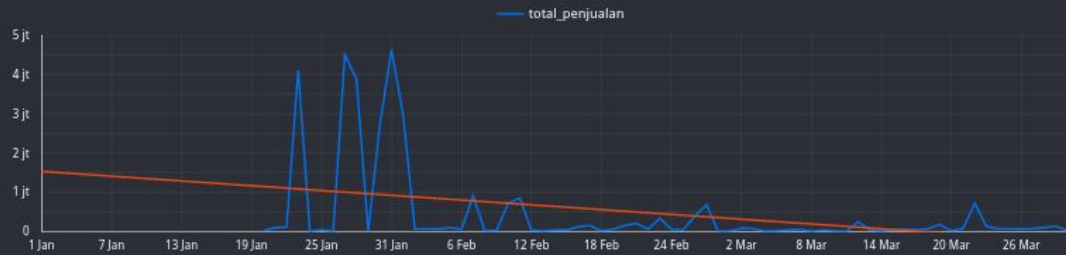
harga_penjualan
Rp1.162.292

jumlah_barang
5.635

Record Count
222

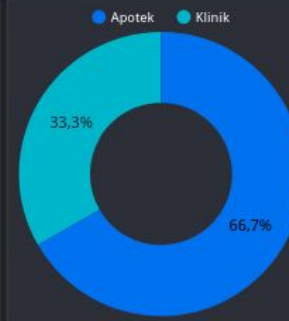
brand

kemasan



nama	Record Cou...
1. APOTEK TAPAK	30
2. APOTEK MAJA	30
3. KLINIK GM	30
4. APOTEK MERDEKA	30
5. KLINIK SAHABAT	30
6. APOTEK SINAR JAYA	29
7. APOTEK SAHABAT	29
8. KLINIK DR. ANDRI	14

1 - 8 / 8



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

B. Jawaban :

Analisis dapat diperdalam dengan diantaranya dengan melengkapi data dengan tabel *brand*, *distribusi*, dan *cabang*. Selain itu, analisis dari dimensi pelanggan juga dapat lebih diperkaya apabila tabel *pelanggan* kolom *level* memiliki data yang lebih variatif (tidak hanya level **Company**).