

Soal & Template Jawaban

Task 5

Nama : Noviandi

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 : Karena query **(b) `SELECT * FROM pelanggan WHERE alamat LIKE 'Mat%'`** lebih sederhana/singkat dan dapat langsung mencari alamat yang mengandung kata 'Mat' diawal kalimat. Sedangkan pada query (a) perlu memastikan beberapa element **`SUBSTR(string, start, Length)`**. selain itu terdapat kesalahan pada query (a) dimana kata **Mat** harus diberi tanda petik '**Mat**' karena merupakan sebuah string.

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 : Karena query (b) dapat digunakan untuk memfilter kolom dengan batas atas dan batas bawah tertentu. Selain itu lebih singkat dari query (a) karena tidak harus menulis statement **tanggal_lahir** 2 kali.

Soal 3: Menentukan Primary Key

A. Tugas

Tentukan primary key dari table penjualan. jelaskan alasannya

B. Jawaban & Penjelasan : Primary Key dari table penjualan adalah **id_customer**. Karena id_customer memiliki nilai yang unik dan tidak terdapat nilai 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

A. Jawaban :

No	Nama File	Link
1.	base_tabel.sql	base table/noviandi
2.	aggregate_tabel.sql	aggregate table/noviandi

Table Base "base_table"

```
1 • USE kimia_farma;
2
3 -- create base table
4 • CREATE TABLE base_table AS(
5 SELECT
6     CONCAT(a.id_invoice,a.id_barang) AS id_penjualan,
7     c.id_customer,
8     a.id_invoice,
9     a.id_distributor,
10    str_to_date(a.tanggal, "%d/%m/%Y") tanggal,
11    a.brand_id,
12    b.lini,
13    a.id_barang,
14    b.nama_barang,
15    b.nama_tipe,
16    a.jumlah_barang,
17    b.kemasan,
18    a.harga,
19    a.unit,
20    a.id_cabang,
21    c.nama,
22    c.level,
23    c.cabang_sales,
24    c.group AS grup
25 FROM penjualan AS a
26 LEFT JOIN barang AS b ON a.id_barang = b.kode_barang
27 LEFT JOIN pelanggan AS c ON a.id_customer = c.id_customer
28 );
29
30 -- determine Primary Key
31 • ALTER TABLE base_table ADD PRIMARY KEY(id_penjualan);
```

Table Base “base_table”

column	data type	description	transformation
id_penjualan	varchar	primary key base table	concat(a.id_invoice,a.id_barang)
id_customer	varchar	id customer, key join ke table pelanggan	-
id_invoice	varchar	id invoice	-
id_distributor	varchar	distributor yang mensupply barang ke cabang	-
tanggal	date	tanggal transaksi	transfromation from text to date
brand_id	varchar	id brand	-
lini	varchar	jenis lini	-
id_barang	varchar	id barang, key join ke table barang	-
nama_barang	varchar	nama barang	-
nama_tipe	varchar	tipe produk	-

Table Base “base_table”

column	data type	description	transformation
jumlah_barang	integer	jumlah barang yang dibeli	-
kemasan	varchar	kemasan barang	-
harga	integer	harga satuan barang	-
unit	varchar	kemasan setiap penjualan	-
id_cabang	varchar	id cabang	-
nama	varchar	nama pelanggan	-
level	varchar	level pelanggan	-
cabang_sales	varchar	lokasi cabang	-
grup	varchar	kategori pelanggan	-

Table Aggregate “aggregate_table”

```
1      -- create aggregate table
2      • ○ CREATE TABLE aggregate_table AS(
3      SELECT
4          id_penjualan,
5          id_invoice,
6          tanggal,
7          id_customer,
8          id_barang,
9          nama_barang AS barang,
10         jumlah_barang,
11         harga,
12         kemasan,
13         nama_tipe,
14         id_cabang,
15         nama AS customer,
16         cabang_sales AS kota_cabang,
17         id_distributor,
18         grup AS group_customer,
19         brand_id,
20         SUM(jumlah_barang * harga) AS total_sales
21     FROM base_table
22     GROUP BY 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16
23 );
```

Table Aggregate “aggregate_table”

column	data type	description	Transformation
id_penjualan	varchar	primary key aggregate table	-
id_invoice	varchar	id invoice	-
tanggal	date	tanggal transaksi	-
id_customer	varchar	id customer	-
id_barang	varchar	id barang	-
barang	varchar	nama barang	nama_barang as barang
jumlah_barang	integer	jumlah barang yang dibeli	-
harga	interger	harga satuan barang	-
kemasan	varchar	kemasan barang	-

Table Aggregate “aggregate_table”

column	data type	description	transformation
nama_tipe	varchar	tipe produk	-
id_cabang	varchar	id cabang	-
customer	varchar	nama pelanggan	nama as customer
kota_cabang	varchar	lokasi cabang	sales_cabang as kota_cabang
id_distributor	varchar	id distributor	-
group_customer	varchar	ketegori pelanggan	grup as group_customer
brand_id	varchar	id brand	-
total_sales	integer	jumlah barang x harga	sum(jumlah_barang * harga) as total_sales

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) :

[link/dashboard/Noviandi](#)



SALES DASHBOARD KIMIA FARMA

JANUARY - JUNE 2022

Cari Barang?

Enter a value

Select date range

total sales
49.2M

Total Items Sold
9,176

Customers
8

Types of Products
10

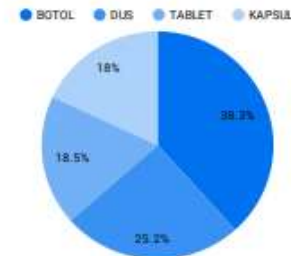
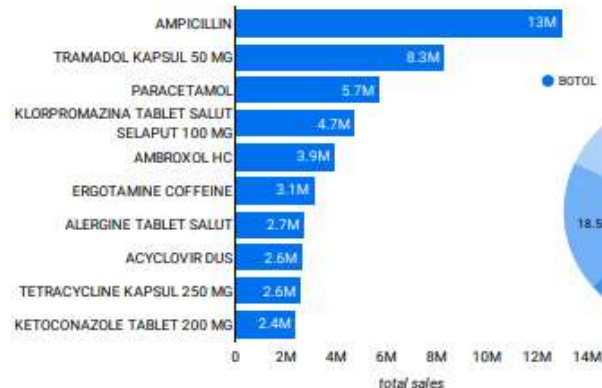
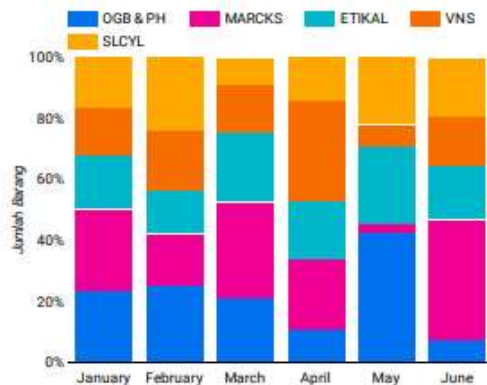
Customer

Cabang Sales



Kota Cabang	Nama Customer	Total Sales
1. Jakarta	KLINIK GM	12.2M
2. Kuningan	APOTEK MAJA	8M
3. Bekasi	APOTEK SINAR JAYA	6M
4. Tangerang	KLINIK SAHABAT	5.6M
5. Bandung	APOTEK MERDEKA	5.6M
6. Aceh	APOTEK TAPAK	4.7M
7. Padang	APOTEK SAHABAT	4.2M
8. Lampung	KLINIK DR. ANDRI	2.9M

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

A. Jawaban : Menurut saya data yang diperlukan adalah data kategori obat. Untuk mengetahui persentase jenis obat yang banyak dibeli oleh customer sehingga bisa digunakan untuk mengetahui penyakit apa yang sedang banyak diderita oleh masyarakat.