

Soal & Template Jawaban

Task 5

Nama : Ahmad Iman Waworuntu

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 lebih fleksibel, mudah dibaca, dan cenderung lebih efisien dalam situasi seperti ini.

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 '200001-01' 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 : Alasan penggunaan BETWEEN lebih disarankan adalah karena membuat query menjadi lebih mudah dibaca dan meminimalkan peluang kesalahan. Selain itu, BETWEEN secara implisit mencakup kedua tanggal batas yang dijelaskan, sehingga lebih jelas dan meminimalkan risiko kehilangan data yang tepat pada tanggal batas.

Soal 3: Menentukan Primary Key

A. Tugas

Tentukan primary key dari table penjualan. jelaskan alasannya

B. Jawaban & Penjelasan :

“id_invoice ”

Karena kolom id_invoice menjelaskan kode faktur yang dimana masing-masing kode memiliki nilai unik pada tiap transaksinya serta tidak dipengaruhi oleh factor dari kolom / tabel lain.

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 Table	https://tinyurl.com/4cpxac8d
2	Aggregate Table	https://tinyurl.com/2mycbrsp

Table Base “base_table”

```
1 • ○ CREATE TABLE base_table (  
2     SELECT  
3         pj.id_distributor, pj.id_cabang, pj.id_invoice,  
4         pj.tanggal, pj.id_customer, pj.id_barang,  
5         pj.jumlah_barang, pj.unit, pj.harga,  
6         pj.mata_uang, pj.brand_id, pj.lini,  
7         pl.level, pl.nama, pl.id_cabang_sales,  
8         pl.cabang_sales, pl.id_group, pl.group,  
9         b.sektor, b.nama_barang, b.tipe,  
10        b.nama_tipe, b.kode_lini, b.kemasan  
11    FROM penjualan pj  
12        LEFT JOIN barang b  
13            ON pj.id_barang = b.kode_barang  
14        LEFT JOIN pelanggan pl  
15            ON pj.id_customer = pl.id_customer  
16    );
```

Table Base “base_table”

column	data type	description	transformation
id_distributor	Text	id unik distributor	
Id_cabang	Text	Id unik cabang	
Id_invoice	Text	Id unik faktur	
Tanggal	Text	Tanggal transaksi	Datetime
Id_customer	Text	Id unik pelanggan	
Id_barang	Text	Id unik barang	
Jumlah_barang	Int	Jumlah produk terjual	
Unit	text	Unit ukuran produk terjual	

Table Base “base_table”

column	data type	description	transformation
Harga	double	Harga per unit terjual	
Mata_uang	Text	Mata uang transaksi	
Brand_id	Text	Id unik dari merek produk terjual	
Lini	Text	Nama merek product terjual	
Level	Text	Tingkat pelanggan	
Nama	Text	Nama pelanggan	
Id_cabang sales	Text	Id unik cabang sales	
Cabang sales	Text	Lokasi cabang sales	

Table Base “base_table”

column	data type	description	transformation
Id_group	Text	Id unik grup pelanggan	
Group	Text	Grup pelanggan	
Sektor	Text	Sektor barang	
Nama_barang	Text	Nama produk terjual	
Tipe	Text	Tipe produk terjual	
Nama_tipe	Text	Nama tipe produk terjual	
Kode_lini	Text	Kode nama barang	
kemasan	Text	Kemasan produk terjual	

Table Aggregate aggregate_table”

```
1 • ○ CREATE TABLE aggregate_table (  
2     SELECT  
3         EXTRACT(MONTH FROM tanggal) bulan_ke,  
4         MONTHNAME(tanggal) bulan,  
5         cabang_sales,  
6         bt.group,  
7         nama nama_customer,  
8         lini,  
9         SUM(jumlah_barang * harga) total_harga  
10    FROM base_table bt  
11    GROUP BY 1,2,3,4,5,6  
12    ORDER BY 1)
```

Table Aggregate “aggregate_table”

column	data type	Description	transformation
Bulan_ke	Int	Urutan bulan	Mengambil nomor dari tanggal
Bulan	Varchar(9)	Nama bulan	Mengambil nama bulan dari tanggal
Cabang_sales	Text	Lokasi cabang sales	
Group	Text	Nama grup pelanggan	
Nama_cutomer	Text	Nama pelanggan	
Lini	Text	Merek produk terjual	
Total_harga	double	Total pendapatan dari transaksi	Menjumlahkan harga dan jumlah barang terjual

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://public.tableau.com/app/profile/iman.waworuntu/viz/SalicylSalesDashoard/SalesDashboard>



Salicyl Sales Dashboard

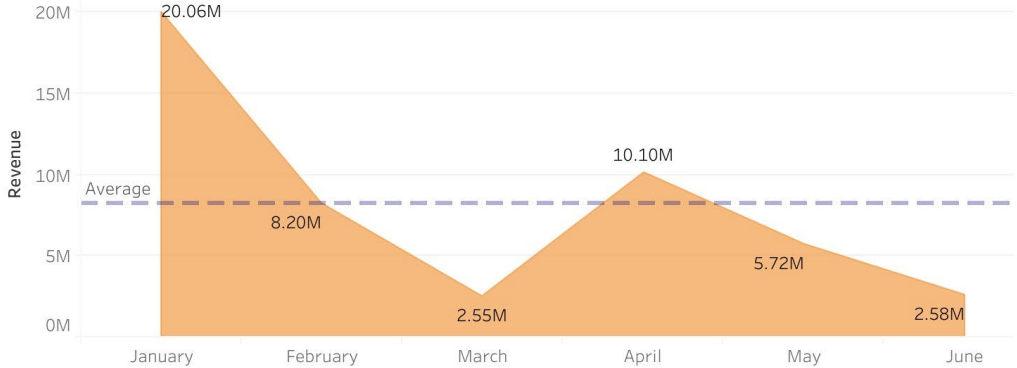
Jan 2022 - Jun 2022

Total Revenue
Rp49.22M

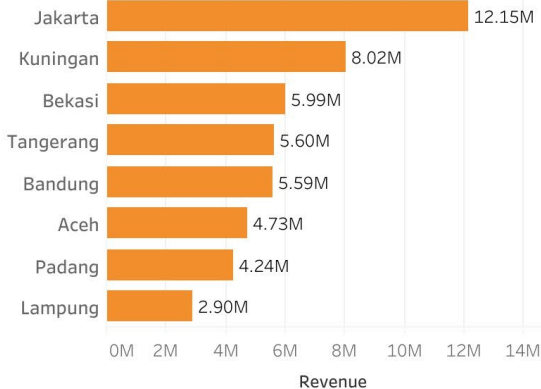
Total Products Sold
9,176

Total Transactions
350

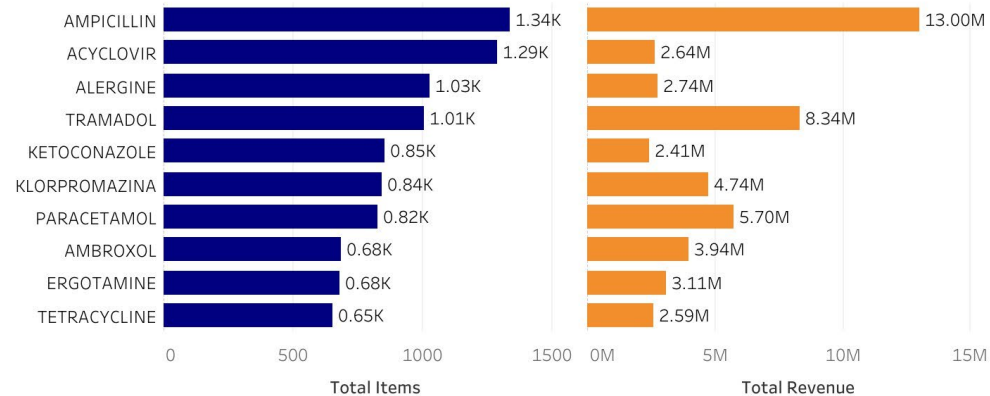
Monthly Revenue



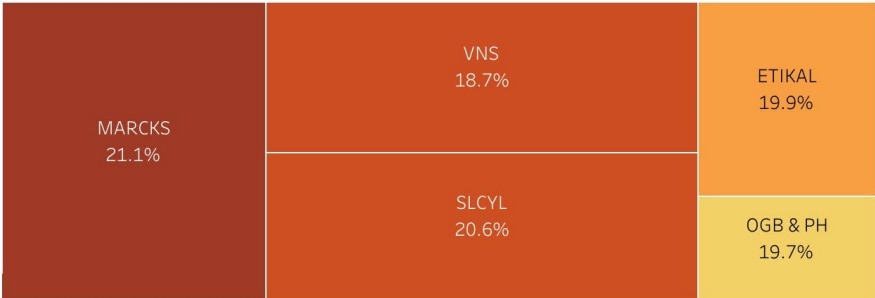
Revenue by Location



Revenue per Product



Revenue per Brand



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 : Data yang dibutuhkan untuk menemukan pola spesifik tertentu :

- Data promosi : untuk mengetahui apakah pelanggan tersebut membeli melalui promosi yang diterima atau tidak untuk mengetahui efektifitas berbagai promosi yang dilakukan oleh perusahaan (PT Kimia Farma Tbk.)
- Data demografi pelanggan : untuk mengetahui karakteristik (usia, jenis kelamin, pendapatan, dsb) serta alasan pembelian dari pelanggan yang membeli produk spesifik yang telah disalurkan oleh sales melalui gerai pelanggan.
- Data kepuasan pelanggan : untuk mengetahui seberapa puas pelanggan terhadap produk yang dijual. Dengan mengetahui kepuasan pelanggan, perusahaan dapat mengevaluasi serta merancang strategi perbaikan atau inovasi untuk kedepannya.