

Soal & Template Jawaban

Task 5

Nama : Ferry Setefanus

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 : query b lebih baik

Alasan : Penggunaan operator LIKE untuk mencari pola yang diinginkan lebih efisien dibandingkan dengan query a yang harus memproses setiap baris pada kolom alamat yang membuat kinerja menjadi lebih berat. Selain itu, query b juga lebih efisien jika kita ingin mencari pola kata lainnya, hanya perlu mengubah pola pencarian pada operator LIKE.

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 : query b lebih tepat

Alasan : query b lebih singkat dalam penulisannya dan lebih mudah dipahami, sedangkan query a lebih sulit untuk dipahami terutama bagi developer yang tidak terbiasa dengan query SQL jika kondisi atau pernyataan yang digunakan sebagai pembanding menjadi lebih banyak.

Soal 3: Menentukan Primary Key

A. Tugas

Tentukan primary key dari table penjualan. jelaskan alasannya

B. Jawaban & Penjelasan :

Kolom id_invoice, alasannya adalah pada dasarnya nomor invoice merupakan nomor unik yang mewakili setiap transaksi yang terjadi. Nomor ini seharusnya tidak dapat muncul kembali pada transaksi yang berbeda, sehingga nomor yang digunakan harus bersifat unik dan dapat membedakan setiap transaksi yang terjadi.

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 :

No	Nama File	Link
1	table_base.sql	here
2	table_aggregate.sql	here

Table_base

```
CREATE TABLE table_base AS (  
  WITH data_penjualan AS (  
    SELECT CONCAT(id_invoice, '-', id_barang) AS id_penjualan,  
           id_invoice,  
           tanggal,  
           date_part('month', tanggal) AS month,  
           id_customer,  
           id_barang,  
           jumlah_barang  
    FROM penjualan  
  UNION  
    SELECT CONCAT(id_invoice, '-', id_barang) AS id_penjualan,  
           id_invoice,  
           tanggal,  
           date_part('month', tanggal) AS month,  
           id_customer,  
           id_barang,  
           jumlah_barang  
    FROM penjualan_ds  
  )  
  
  SELECT dp.id_penjualan, dp.id_invoice, dp.id_customer,  
         dp.id_barang, pds.id_distributor, dp.tanggal, dp.month,  
         pds.nama as customer, pds.cabang_sales, pds.group,  
         bds.nama_barang, bds.brand, bds.harga, dp.jumlah_barang  
  FROM data_penjualan as dp  
  LEFT JOIN pelanggan_ds as pds  
    ON dp.id_customer = pds.id_customer  
  LEFT JOIN barang_ds as bds  
    ON dp.id_barang = bds.kode_barang  
  ORDER BY dp.id_invoice  
)
```

Table_base

column	data type	description	transformation
id_penjualan	String	kode unik yang mewakili setiap transaksi penjualan	concat dari kolom id_invoice dengan id_barang
id_invoice	String	kode unik invoice	
id_customer	String	kode unik customer	
id_barang	String	kode unik barang	
tanggal	String	tanggal terjadinya transaksi	
month	String	bulan terjadinya transaksi	diextract dari kolom tanggal
id_distributor	String	kode unik distributor	

column	data type	description	transformation
customer	String	nama customer	
cabang_sales	String	tempat terjadinya transaksi	
group	String	jenis tempat terjadinya transaksi	
nama_barang	String	nama barang	
brand	String	merk barang	
harga	Integer	harga satuan barang	
jumlah_barang	Intger	jumlah barang yang dipesan	
total_harga	intger	total harga	hasil perkalian kolom harga dan jumlah_barang

Table_aggregate

```
CREATE TABLE table_aggregate AS (  
    SELECT *,  
           jumlah_barang * harga AS total_harga  
    FROM table_base  
)
```

Table_aggregate

column	data type	description	transformation
id_penjualan	String	kode unik yang mewakili setiap transaksi penjualan	concat dari kolom id_invoice dengan id_barang
id_invoice	String	kode unik invoice	
id_customer	String	kode unik customer	
id_barang	String	kode unik barang	
tanggal	String	tanggal terjadinya transaksi	
month	String	bulan terjadinya transaksi	diextract dari kolom tanggal
id_distributor	String	kode unik distributor	
customer	String	nama customer	

column	data type	description	transformation
cabang_sales	String	tempat terjadinya transaksi	
group	String	jenis tempat terjadinya transaksi	
id_customer	String	kode unik customer	
id_barang	String	kode unik barang	
nama_barang	String	nama barang	
brand	String	merk barang	
harga	Integer	harga satuan barang	
jumlah_barang	Intger	jumlah barang yang dipesan	
total_harga	intger	total harga	hasil perkalian kolom harga dan jumlah_barang

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://lookerstudio.google.com/reporting/c49ef8cb-9326-4043-96b7-bde951a07b73>

DASHBOARD DATA PENJUALAN KIMIA FARMA

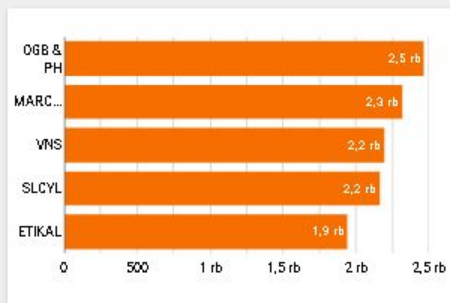
TOTAL DATA KESELURUHAN

Total produk terjual
11,1 rb

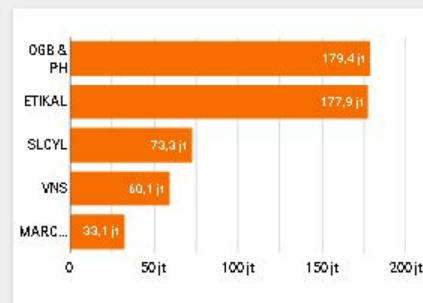
Total Customer
350

Total Revenue
523,7 jt

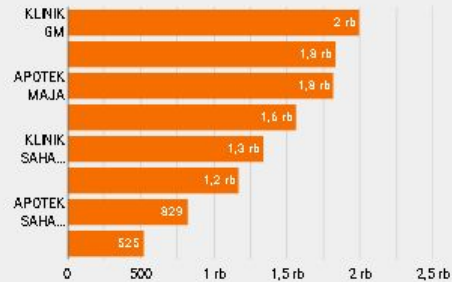
Total penjualan by brand



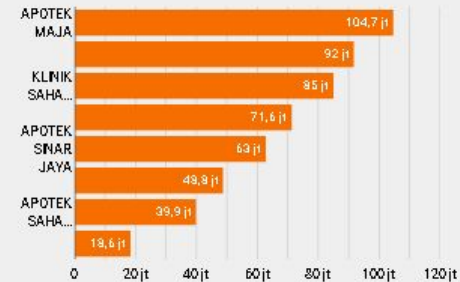
Total Revenue by brand



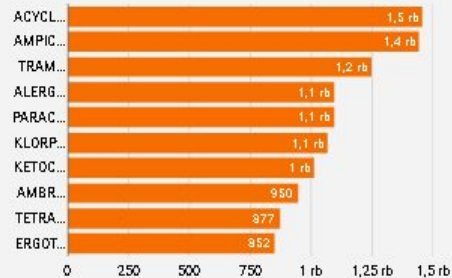
Total Penjualan by Customer



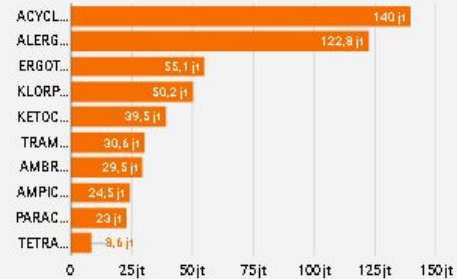
Total Revenue by Customer



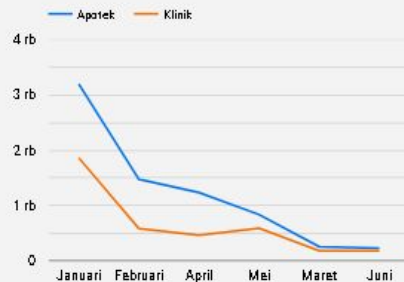
Total Penjualan by nama produk



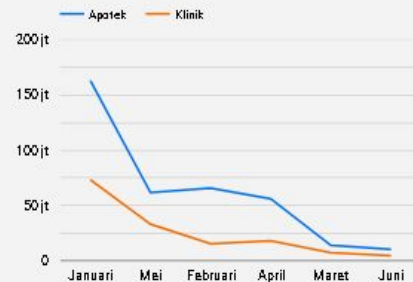
Total Revenue by nama produk



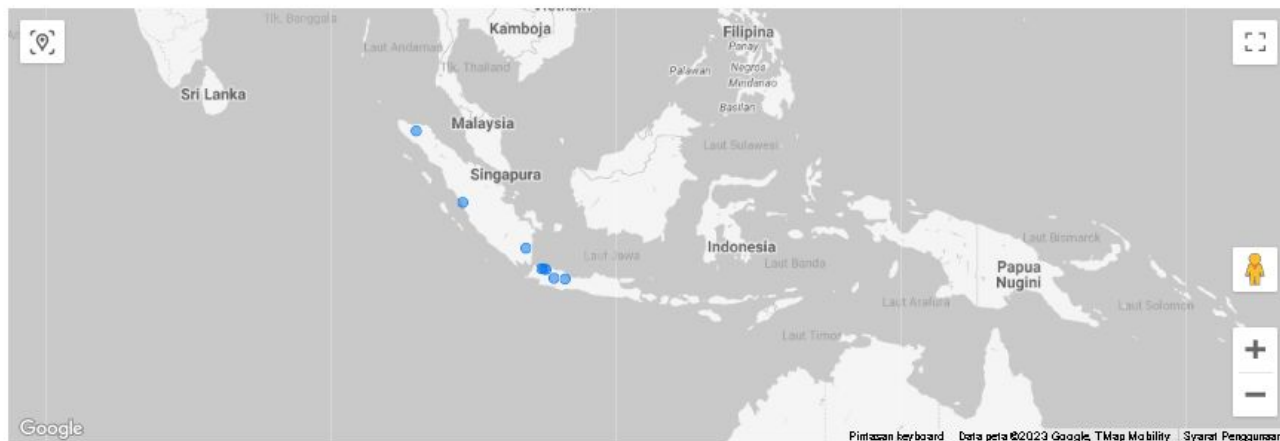
Total Penjualan perbulan by group



Total Revenue perbulan by group



City Distribution Map



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 :

- Data mengenai biaya untuk memperoleh obat - obatan yang dijual agar dapat mengetahui keuntungan bersih dari setiap transaksi yang terjadi.
- Data mengenai persaingan harga pasar agar dapat membandingkan harga obat yang dijual dengan kompetitor.
- Data mengenai promosi dan iklan yang dilakukan untuk menunjukkan apakah promosi dan iklan ini efektif untuk meningkatkan penjualan atau tidak.
- Data mengenai kondisi ekonomi dan pasar seperti pertumbuhan ekonomi, inflasi, dan kebijakan pemerintah yang bisa berdampak pada penjualan obat pada kimia farma.