

Soal & Template Jawaban

Task 5

Nama : Rheza Paleva Uyanto

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 : Pada pilihan (a), Fungsi SUBSTR adalah untuk ekstraksi, sedangkan dalam konteks ini lebih ditekankan menggunakan filtering. Mat, tidak diberikan tanda petik, sehingga kemungkinan terjadi error..

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 : Walaupun secara logika, kedua syntax tersebut sama, pilihan (b) memiliki jumlah karakter lebih sedikit dibanding pilihan (a). Selain itu pilihan (a) memiliki proses yang lebih lama yaitu memfilter yang lebih dari 2000-01-01, dan memfilter kembali yang kurang dari 2008-12-31.

Soal 3: Menentukan Primary Key

A. Tugas

Tentukan primary key dari table penjualan. jelaskan alasannya

B. Jawaban & Penjelasan :

Primary Key dari tabel penjualan : id_invoice

Karena primary key harus unique (tidak memiliki nilai yang sama antar barisnya), sedangkan kolom lainnya memiliki nilai yang sama, sehingga tidak dapat dijadikan 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

A. Jawaban :

No	Nama File	Link
1	Data_Source_Task5.xlsx	https://docs.google.com/spreadsheets/d/1k_0L9GylEq6EuD6d_kdKu-Hj3GYW6p3T/edit?usp=drive_web&oid=112892435247988506765&rtpof=true
2	Aggregate Table, Query Base Table dan Aggrerate Table	https://drive.google.com/drive/u/0/folders/1kLoNQ_47ZgFXDzMCIdDyD0Q_0AJw1pWl

Table Base "<<barang>>"

```
CREATE TABLE IF NOT EXISTS public.barang
(
    kode_barang character varying COLLATE pg_catalog."default" NOT NULL,
    nama_barang character varying COLLATE pg_catalog."default",
    kemasan character varying COLLATE pg_catalog."default",
    harga integer,
    nama_tipe character varying COLLATE pg_catalog."default",
    kode_brand integer,
    brand character varying COLLATE pg_catalog."default",
    CONSTRAINT barang_pkey PRIMARY KEY (kode_barang)
);
```

Table Base “<<barang>>”

column	data type	description	transformation
kode_barang	character varying	kode barang	-
nama_barang	character varying	nama barang	-
kemasan	character varying	botol / tablet / dus	-
harga	integer	harga barang	-
nama_tipe	character varying	produk jadi	-
kode_brand	integer	kode barang	-
brand	character varying	kelompok barang	-

Table Base "<<pelanggan>>"

```
CREATE TABLE IF NOT EXISTS public.pelanggan
(
  id_customer character varying COLLATE pg_catalog."default" NOT NULL,
  level character varying COLLATE pg_catalog."default",
  nama character varying COLLATE pg_catalog."default",
  id_cabang_sales character varying COLLATE pg_catalog."default",
  cabang_sales character varying COLLATE pg_catalog."default",
  id_distributor character varying COLLATE pg_catalog."default",
  "group" character varying COLLATE pg_catalog."default",
  CONSTRAINT pelanggan_pkey PRIMARY KEY (id_customer)
);
```

Table Base “<<pelanggan>>”

column	data type	description	transformation
id_customer	character varying	kode customer	-
level	character varying	company	-
nama	character varying	nama customer	-
id_cabang_sales	character varying	kode nama daerah	-
cabang_sales	character varying	nama daerah	-
id_distributor	character varying	kode distributor	-
group	character varying	apotek / klinik	-

Table Base “<<penjualan>>”

```
CREATE TABLE IF NOT EXISTS public.penjualan
(
    id_invoice character varying COLLATE pg_catalog."default",
    tanggal character varying COLLATE pg_catalog."default",
    id_customer character varying COLLATE pg_catalog."default",
    id_barang character varying COLLATE pg_catalog."default",
    jumlah_barang integer,
    unit character varying COLLATE pg_catalog."default",
    harga integer,
    mata_uang character varying COLLATE pg_catalog."default"
);
```

Table Base “<<penjualan>>”

column	data type	description	transformation
id_invoice	character varying	kode invoice	-
tanggal	character varying	tanggal terbit invoice	-
id_customer	character varying	kode customer	-
id_barang	character varying	kode barang	-
jumlah_barang	integer	jumlah barang yang dibeli	-
unit	character varying	satuan barang	-
harga	integer	harga	-
mata_uang	character varying	idr	-

Table Aggregate “<<Tabel_Aggregate_task5>>”

Query Query History

```
1  with data1 as
2  (
3  select id_invoice, tanggal, a.id_customer, id_barang, jumlah_barang, unit, nama,
4         id_cabang_sales, cabang_sales, id_distributor, b.group
5  from penjualan as a left join pelanggan as b
6  on a.id_customer = b.id_customer
7  )
8  select id_invoice, tanggal, nama, cabang_sales, id_distributor, data1.group,
9         jumlah_barang, nama_barang, kemasan, harga, brand
10 from data1 left join barang as c
11 on data1.id_barang = c.kode_barang
12
```

Table Aggregate “<<Tabel_Agregate_task5>>”

column	data type	description	transformation
Id_invoice	character varying	nomor invoice	-
tanggal	character varying	tanggal	-
nama	character varying	nama customer	-
cabang_sales	character varying	nama daerah cabang	-
id_distributor	character varying	nama distributor	-
group	character varying	apotik / klinik	-
jumlah_barang	integer	jumlah barang yang dibeli	-
nama_barang	character varying	nama barang	-
kemasan	character varying	botol / tablet / dus	-
harga	integer	harga barang	-
brand	character varying	kelompok 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

A. Jawaban :

Link visualisasi (ex link Google Data Studio) :

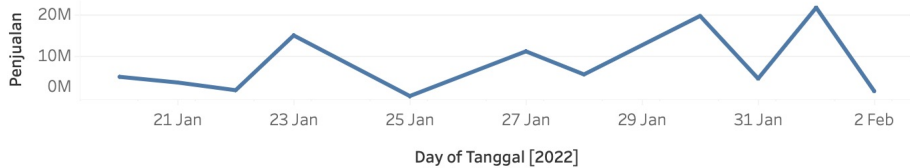
<https://public.tableau.com/app/profile/rheza.paleva.uyanto/viz/DashboardSalicylRakaminVIX-Rheza/Dashboard1>

Dashboard Penjualan Brand Salicyl - Kimia Farma
Virtual Internship Experience - Rheza Paleva Uyanto



Jumlah Produk	Jumlah Customer	Jumlah Barang	Penjualan	Mean Pembelian Per Customer
10	25	2,239	90,744,625	687,459

Penjualan per Waktu



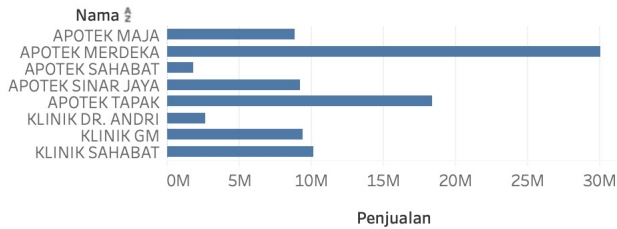
Penjualan per Area

Cabang Sales	Penjualan
Aceh	18,401,913
Bandung	30,035,339
Bekasi	9,220,709
Jakarta	9,469,900
Kuningan	8,851,487
Lampung	2,702,058
Padang	1,887,187
Tangerang	10,176,031

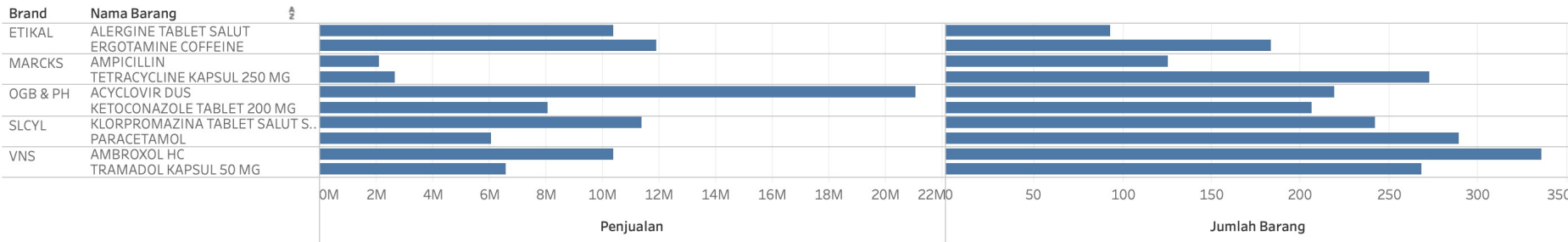
Penjualan



Penjualan per Customer



Penjualan 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 :

1. Perlu dijelaskan kembali, posisi data analist yang berada di industri farmasi, distributor, atau unit pelayanan farmasi. Karena subyek dari masing- masing user berbeda. Contoh : User dari Industri Farmasi adalah Distributor, User dari Distributor adalah Unit Pelayanan farmasi. Dalam data set ini masih tercampur, sehingga terjadi bias.
2. Data dari Industri Farmasi lain selain Kimia Farma. Mungkin data didapat dari sisi distributor, karena dari satu distributor dapat melayani dari lebih dari satu industri. Sehingga kita (jika sebagai data analis di distributor) melakukan perbandingan produk kita dengan kompetitor.