

Soal & Template Jawaban

Task 5

Nama : Bagus Ganjar Lugina

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

Alasan : Karena b akan menghasilkan output filter suatu kolom dari 3 huruf yang mengandung Mat, sedangkan a akan menghasilkan output error. Jika kolom alamat tersebut di index, index yang tujuannya untuk mengoptimalkan query jadi tidak berguna dikarenakan ada fungsi di dalam kolomnya.

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

Alasan : Karena b digunakan ketika ingin melakukan filter suatu kolom dengan batas bawah dan batas atas tertentu. Umumnya untuk tipe data angka atau tanggal.

Soal 3: Menentukan Primary Key

A. Tugas

Tentukan primary key dari table penjualan. jelaskan alasannya

B. Jawaban & Penjelasan :

Karena pada table penjualan tidak terdapat primary key (tidak ada unique values), oleh karena itu Primary key bisa dibuat dengan melakukan concat terhadap **id_invoice** dan **id_barang** sebagai **id_penjualan**.

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.	create_database.sql	https://drive.google.com/file/d/1bhGb6pkVQpzospmlt0zbJbcQ0ed1de0g/view?usp=sharing
2.	base_table	https://drive.google.com/file/d/1MkcHd6XNOheJrxB-ej2NY-KAGbmnM9LK/view?usp=sharing
3.	aggregat_table	https://drive.google.com/file/d/1Xd3owip77yuDodMH7b1tPiwbePViuO89/view?usp=sharing

Table Base “base_table”

```
-- Create Base Table
CREATE TABLE base_table AS(SELECT
    pjl.id_invoice ||'|| pjl.id_barang AS id_penjualan,
    pjl.id_invoice,
    pjl.tanggal,
    pjl.id_customer,
    pjl.id_barang,
    pjl.jumlah_barang,
    pjl.unit,
    pjl.harga,
    pjl.mata_uang,
    plg.level,
    plg.nama,
    plg.id_cabang_sales,
    plg.cabang_sales,
    plg.id_distributor,
    plg.grup,
    brg.kode_barang,
    brg.nama_barang,
    brg.kemasan,
    brg.nama_tipe,
    brg.kode_brand,
    brg.brand
FROM penjualan AS pjl
LEFT JOIN pelanggan AS plg ON plg.id_customer = pjl.id_customer
LEFT JOIN barang AS brg ON brg.kode_barang = pjl.id_barang
);

--Determine Primary Key
ALTER TABLE base_table ADD PRIMARY KEY(id_penjualan);\

--Export Queries to CSV
COPY(SELECT *
FROM base_table
) TO 'E:\Rakamin\JAP\Big Data Analytics - Kimia Farma\base_table.csv' DELIMITER ',' CSV HEADER;
```

Table Aggregate “base_table”

column	data type	description	transformation
id_penjualan	varchar	primary key base table	pjl.id_invoice " pjl.id_barang AS id_penjualan
id_invoice	varchar	id invoice	
tanggal	date	tanggal transaksi	
id_customer	varchar	id customer, key join ke tabel pelanggan	
id_barang	varchar	id barang, key join ke tabel barang	
jumlah_barang	integer	jumlah barang yang dibeli	
unit	varchar	kemasan setiap penjualan	
harga	integer	harga satuan barang	
mata_uang	varchar	mata uang penjualan IDR	
level	varchar	level pelanggan	

Table Aggregate “base_table”

column	data type	description	transformation
nama	varchar	nama pelanggan	
id_cabang_sales	varchar	id cabang sales	
cabang_sales	varchar	lokasi cabang	
id_distributor	varchar	distributor yang supply barang ke cabang	
grup	varchar	kategori pelanggan	
kode_barang	varchar	kode barang	
nama_barang	varchar	nama barang	
kemasan	varchar	kemasan barang	
nama_tipe	varchar	tipe produk	
kode_brand	varchar	kode brand	

Table Aggregate “base_table”

column	data type	description	transformation
brand	varchar	nama brand	

Table Aggregate “Figure”

```
-- Create Agregat Table
CREATE TABLE agregat_table AS(SELECT
    id_penjualan,
    id_invoice,
    tanggal,
    id_customer,
    id_barang,
    jumlah_barang,
    harga,
    nama AS customer,
    id_cabang_sales,
    cabang_sales,
    id_distributor,
    grup,
    nama_barang AS barang,
    kemasan,
    nama_tipe,
    kode_brand,
    brand,
    SUM(jumlah_barang * harga) AS total_sales
FROM base_table
GROUP BY 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17
);

-- Export Queries to CSV
COPY(SELECT *
FROM agregat_table
) TO 'E:\Rakamin\JAP\Big Data Analytics - Kimia Farma\agregat_table.csv' DELIMITER ',' CSV HEADER;
```

Table Aggregate “aggregate_table”

column	data type	description	transformation
id_penjualan	varchar	primary key aggregate tabel	
id_invoice	varchar	id invoice	
tanggal	date	tanggal transaksi	
id_customer	varchar	id customer	
id_barang	varchar	id barang	
jumlah_barang	integer	jumlah barang yang dibeli	
harga	integer	harga satuan barang	
customer	varchar	nama pelanggan	nama AS customer
id_cabang_sales	varchar	id cabang sales	
cabang_sales	varchar	cabang sales	

Table Aggregate “aggregate_table”

column	data type	description	transformation
id_distributor	varchar	distributor yang supply barang ke cabang	
grup	varchar	kategori pelanggan	
barang	varchar	nama barang	nama_barang AS barang
kemasan	varchar	kemasan barang	
nama_tipe	varchar	tipe produk	
kode_brand	varchar	kode brand	
brand	varchar	brand barang	
total_sales	integer	jumlah barang dikalikan dengan 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

B. Jawaban :

Link visualisasi (ex link Google Data Studio) :

<https://datastudio.google.com/reporting/7613d646-8f1b-479d-b69b-a60ef1236936>

Sales Report PT Kimia Farma



Cari barang

Enter a value

Select date range

Customer

Brand

Invoice

132

Customer

8

Jenis Barang

10

Brand

5

Kemasan

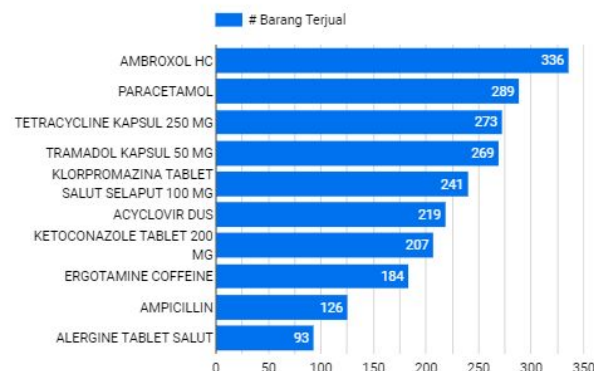
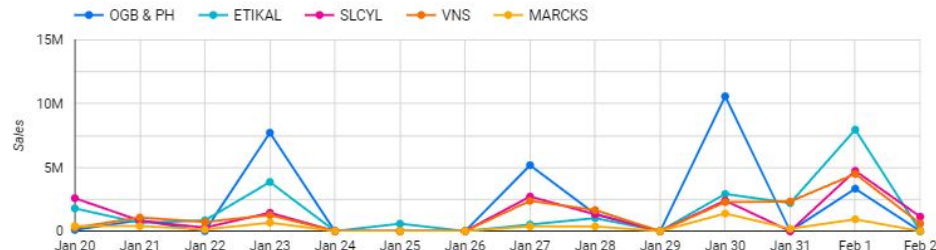
3

Total Barang Terjual

2,237

Total Sales

90,637,700



	Tanggal	Invoice ID	Customer	Barang	Jumlah Barang	Harga	Total Harga
1.	Feb 2, 2022	IN6064	APOTEK TAPAK	TRAMADOL KAPSUL 5...	24	24,500	588,000
2.	Feb 2, 2022	IN6064	APOTEK TAPAK	KLORPROMAZINA TAB...	24	47,000	1,128,000
3.	Feb 1, 2022	IN6301	APOTEK SAHABAT	TETRACYCLINE KAPS...	24	9,800	235,200
4.	Feb 1, 2022	IN6113	KLINIK GM	ACYCLOVIR DUS	2	96,000	192,000
5.	Feb 1, 2022	IN6023	KLINIK SAHABAT	AMBROXOL HC	67	31,000	2,077,000
6.	Feb 1, 2022	IN6023	KLINIK SAHABAT	KLORPROMAZINA TAB...	10	47,000	470,000
7.	Feb 1, 2022	IN6063	APOTEK TAPAK	PARACETAMOL	67	21,000	1,407,000
8.	Feb 1, 2022	IN6063	APOTEK TAPAK	TRAMADOL KAPSUL 5...	21	24,500	514,500
9.	Feb 1, 2022	IN6259	APOTEK MERDEKA	AMPICILLIN	14	17,000	238,000
10.	Feb 1, 2022	IN6023	KLINIK SAHABAT	KETOCONAZOLE TABL...	10	39,000	390,000
11.	Feb 1, 2022	IN6063	APOTEK TAPAK	ALERGINE TABLET SA	25	112,000	2,800,000

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 yang dibutuhkan adalah penggolongan pada barang, dimaksudkan untuk meningkatkan keamanan dan ketepatan penggunaan serta pengamanan distribusi obat. Obat dapat dibedakan menjadi berbagai kategori dan golongan. Contoh: Obat Bebas Terbatas, Obat Bebas, Obat Keras dll. Hal ini diperlukan untuk mengetahui distribusi penjualan golongan barang/obat di setiap pelanggan.