

Soal & Template Jawaban

Task 5

Nama : Arini Arumsari

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 : Query yang lebih baik tergantung pada kebutuhan. Jika ingin mencari baris yang alamatnya tepat "Mat", gunakan query 1. Jika ingin mencari baris yang alamatnya berisi "Mat" di mana pun dalam string, gunakan query 2.

Alesan saya milih b karena kita ingin mencari Alamat pelanggan dengan isi mat dimanapun nama mat berada

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 : a

Alasan :

Kedua query tersebut menghasilkan hasil yang sama. Operator `>=` dan `<=`: Lebih mudah dibaca dan dipahami.

Lebih mudah digunakan jika hanya ingin memilih baris yang tepat pada tanggal tertentu

Soal 3: Menentukan Primary Key

A. Tugas

Tentukan primary key dari table penjualan. jelaskan alasannya

B. Jawaban & Penjelasan :

Jawaban

```
ALTER TABLE penjualan ADD PRIMARY KEY (id_invoice);
```

```
ALTER TABLE barang ADD PRIMARY KEY (kode_barang);
```

Penjelasan:

Primary key dari tabel penjualan adalah id_invoice karena invoice memiliki nilai unique begitupun dengan kolom kode_barang memiliki nilai yang unique

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 screenshot 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.	postgresql	https://github.com/ariniamsr/-Kimia-Farma-Big-Data-Analytics-Project-Based-Internship-Program/tree/main/postgresql

Table Base “<<Nama Tabel>>”

	tanggal date	id_invoice character var	id_customer character var	id_barang character var	nama_cabang character varying	cabang_sales character var	kode_barang character var	produk character var	lini character var	jumlah_unit integer	unit character var	harga_per_unit numeric
1	2022-01-20	IN5997	CUST553...	BRG0001	APOTEK TAPAK	Aceh	BRG0001	ACYCLOV...	OGB & PH	1	DUS	1169.91
2	2022-01-20	IN6297	CUST553...	BRG0002	APOTEK MAJA	Kuningan	BRG0002	ALERGIN...	ETIKAL	5	DUS	2337.5
3	2022-01-21	IN6155	CUST553...	BRG0003	KLINIK GM	Jakarta	BRG0003	AMPICILL...	MARCKS	9	DUS	10690.6
4	2022-01-22	IN6144	CUST553...	BRG0004	APOTEK MERDEKA	Bandung	BRG0004	TRAMAD...	VNS	13	DUS	8700.7
5	2022-01-23	IN6280	CUST553...	BRG0005	KLINIK SAHABAT	Tangerang	BRG0005	KLORPRO...	SLCYL	1	DUS	5648.3
6	2022-01-23	IN6052	CUST553...	BRG0006	APOTEK SINAR JAYA	Bekasi	BRG0006	KETOCO...	OGB & PH	5	DUS	2819.2
7	2022-01-25	IN6089	CUST553...	BRG0007	APOTEK SAHABAT	Padang	BRG0007	ERGOTA...	ETIKAL	9	DUS	4592.1
8	2022-01-23	IN6251	CUST553...	BRG0008	APOTEK TAPAK	Aceh	BRG0008	TETRACY...	MARCKS	1	DUS	3991.9

Table Base “<<Nama Tabel>>”

column	data type	description	transformation
Tanggal	Date	Tanggal transaksi	
Id_invoice	varchar	Nilai unique tiap transaksi	
Id_customer	varchar	Identitas customer tiap transaksi	
Id_barang	Varchar	Identitas barang	
Nama_cabang	varchar	Nama Apotek tiap cabang	nama as nama_cabang
Cabang_sales	varchar	Penjualan tiap kota	
Kode_barang	varchar	Kode disetiap barang	
Produk	varchar	Nama Produk	nama_barang as produk
lini	varchar	Merk	
Jumlah_unit	Integer	Jumlah unit	jumlah_barang as jumlah_unit
Unit	Varchar	Unit	
Harga_per_unit	Numeric	Harga tiap unit	harga as harga_per_unit

Table Aggregate “<<Nama Tabel>>”

	tanggal date	bulan double precis	id_invoice character var	cabang_sales character var	pelanggan character varying	produk character varying	merek character vary	jumlah_produk_terjual integer	harga_satuan numeric	pendapatan numeric
1	2022-01-20	1	IN5997	Aceh	APOTEK TAPAK	ACYCLOVIR DUS	OGB & PH	1	1169.91	1169.91
2	2022-01-20	1	IN6297	Kuningan	APOTEK MAJA	ALERGINE TABLET SALUT	ETIKAL	5	2337.5	11687.5
3	2022-01-21	1	IN6155	Jakarta	KLINIK GM	AMPICILLIN	MARCKS	9	10690.6	96215.4
4	2022-01-22	1	IN6144	Bandung	APOTEK MERDEKA	TRAMADOL KAPSUL 50 MG	VNS	13	8700.7	113109.1
5	2022-01-23	1	IN6317	Aceh	APOTEK TAPAK	AMPICILLIN	MARCKS	10	10690.6	106906.0
6	2022-01-23	1	IN6125	Aceh	APOTEK TAPAK	AMPICILLIN	MARCKS	14	10690.6	149668.4
7	2022-01-23	1	IN5980	Aceh	APOTEK TAPAK	AMPICILLIN	MARCKS	20	10690.6	213812.0

Table Aggregate “<<Nama Tabel>>”

column	data type	description	transformation
tanggal	date	Tanggal transaksi	
bulan	Double precision	Bulan transaksi	date_part('month', tanggal) AS bulan
Id_invoice	Varchar	Identitas dokumen	
Cabang_sales	varchar	Kota penjualan	
pelanggan	varchar	Nama pelanggan	nama_cabang AS pelanggan
Produk	varchar	Nama produk	
Merk	varchar	Nama merk	lini AS merek
Jumlah produk terjual	Integer	Jumlah yang terjual	jumlah_unit AS jumlah_produk_terjual
Harga satuan	numeric	Harga tiap barang	harga_per_unit AS harga_satuan
pendapatan	numeric	revenue	(jumlah_unit * harga_per_unit) AS pendapatan

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://lookerstudio.google.com/reporting/f91064ea-a92b-4e4a-a20b-57ffb9c96d8d>

Bulan ▼

Merek ▼

Kota ▼

Produk ▼

Total Pendapatan

49.2M

Jumlah Transaksi

350

Total Produk

10

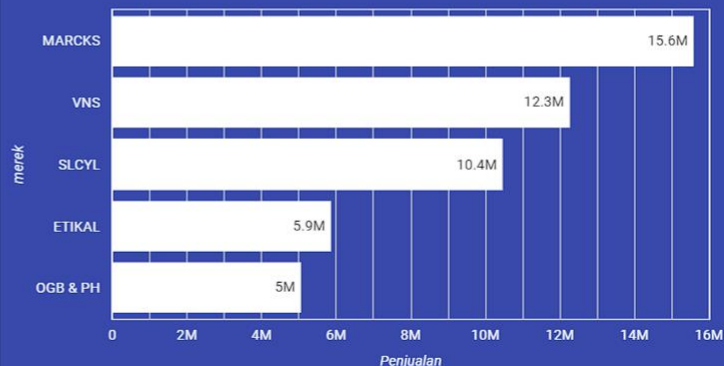
Total Unit Terjual

9,176

Unit Terjual Berdasarkan Kota



Total Penjualan berdasarkan Merek



pelanggan	pendapatan ▼	produk	harga satuan	produk terjual	pendapatan ▼
KLINIK GM	12,151,454.7	AMPICILLIN	437,758.7	1,337	13,002,028.5
APOTEK MAJA	8,018,901	TRAMADOL KAPSUL 50 MG	380,867.8	1,005	8,339,311.5
APOTEK SINAR JAYA	5,985,145.72	PARACETAMOL	166,567.2	822	5,704,926.6
KLINIK SAHABAT	5,604,113.9	KLORPROMAZINA TABLET SALUT SELAPUT 10...	197,690.5	840	4,744,572
APOTEK MERDEKA	5,585,451.85	AMBROXOL HC	138,736.8	681	3,936,656.7

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 :

Perlu adanya Data persaingan

Data persaingan dapat memberikan informasi tentang produk dan layanan yang ditawarkan oleh perusahaan pesaing. Data ini dapat digunakan untuk memahami posisi perusahaan di pasar. Misalnya, jika perusahaan pesaing menawarkan produk yang serupa dengan produk perusahaan, maka perusahaan dapat membuat strategi untuk bersaing dengan perusahaan pesaing tersebut.