



Task 5

Nama : Maulana Kavaldo

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 query a tidak efisien, menggunakan fungsi pada column yang difilter. Saat kolom alamat diindeks, indeks menjadi tidak berguna untuk mengoptimalkan kueri karena kolom tersebut memiliki fungsi. Sehingga query b lebih baik.

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 operator AND, OR, & NOT sebaiknya dihindari. Hal tersebut nantinya juga akan berpengaruh terhadap optimalisasi query. Sehingga pada kasus ini dengan query b lebih efisien dan mengurangi kesalahan dalam penulisan query dibandingkan dengan operator AND, OR, & NOT `>=` / `<=` meskipun diperoleh hasil data yang sama.

Soal 3: Menentukan Primary Key

Tugas

Tentukan primary key dari table penjualan. Jelaskan alasannya?

Jawaban & Penjelasan:

Dengan membuat kolom baru untuk primary key-nya.

Karena pada table penjualan tidak ada kolom yang unik sehingga membuat kolom baru dengan menggabungkan kolom id_invoice dan id_barang sehingga diperoleh kolom baru dengan nama id_penjualan.

Soal 4: Design Datamart

Tugas

Buatlah design datamart (Terdiri dari tabel base dan tabel aggregate).

Table Base “Penjualan”

```
SELECT
CONCAT(id_invoice, '_', id_barang) id_penjualan,
pjl.id_invoice,
pjl.tanggal,
pjl.id_barang,
brg.nama_barang, pjl.harga,
pjl.unit,
pjl.jumlah_barang,
(jumlah_barang * pjl.harga) AS total_harga_per_barang,
pjl.mata_uang,
brg.kode_brand, brg.brand,
pjl.id_customer,
plg.nama nama_customer,
plg.cabang_sales,
plg.id_distributor,
plg.group AS group_category
FROM `virtual-internship-rakamin.kimia_farma.penjualan` pjl
LEFT JOIN `virtual-internship-rakamin.kimia_farma.barang` brg
ON (pjl.id_barang = brg.kode_barang)
LEFT JOIN `virtual-internship-rakamin.kimia_farma..pelanggan` plg
ON (pjl.id_customer = plg.id_customer)
```

Table Base: Penjualan

Column	Data Type	Description	Transformation
Id_penjualan	string	Promary key tabel penjualan	Gabungan Id_invoice dengan id_barang
id_invoice	varchar (255)	Menyimpan id invoice	-
tanggal	date	Tanggal transaksi	-
id_barang	varchar (255)	Id barang	-
jumlah_barang	Numeric (int)	Jumlah barang	-
harga	double	Harga barang	-
total	Numeric (int)	Kolom harga x jumlah barang	harga x jumlah barang
kode_brand	varchar (255)	Kode brand	filtering

Table Aggregate "Total_Penjualan"

```
SELECT
id_invoice, tanggal,
id_customer, id_distributor, category,
COUNT (DISTINCT id_barang) total_barang,
SUM(total_harga_per_barang) total_penjualan
FROM `virtual-internship-rakamin.kimia_farma.penjualan`
GROUP BY 1,2,3,4,5,6,7
ORDER BY 1
```


Table Aggregate: "Penjualan Per-Hari"

Column	Data Type	Description	Transformation
Id_invoice	String	Primary ke	-
Tanggal	Date	Tanggal transaksi	-
Id_customer	String	Id untuk customer	-
Id_distributor	String	Id untuk distributor	-
kategori	String	Nama-nama untuk kategori produk	-
total_barang	Numeric (int)	Jumlah barang	Count, Distinct
total_penjualan	Numeric (int)	Menyimpan total penjualan tiap invoice	SUM

Soal 5 : Data Visualization

Tugas

Buatlah data visualiasasi nya, dan cantumkan linknya di bawah (pastikan bisa diakses publik). Lalu cantumkan juga screenshot visualisasinya.

Jawaban:

<https://datastudio.google.com/reporting/489efaa7-34e8-4374-bfec-d4aaab012eb0>

TABEL PENJUALAN SLCYL

	ID INVOICE	ID BARANG	LINI BRAND	HARGA	JUMLAH...	TOTAL HAR...
1.	IN6181	BRG0010	SLCYL	6940.3	130	902239
2.	IN6203	BRG0005	SLCYL	5648.3	100	564830
3.	IN6180	BRG0005	SLCYL	5648.3	98	553533.4
4.	IN6114	BRG0010	SLCYL	6940.3	98	680149.4
5.	IN6303	BRG0010	SLCYL	6940.3	95	659328.5
6.	IN6273	BRG0005	SLCYL	5648.3	79	446215.7
7.	IN6133	BRG0010	SLCYL	6940.3	78	541343.4
8.	IN6252	BRG0004	SLCYL	6940.3	67	465000.100000...
9.	IN6262	BRG0010	SLCYL	6940.3	67	465000.100000...
10.	IN6159	BRG0010	SLCYL	6940.3	67	465000.100000...

1 - 70 / 70 < >

PILIH BRAND: SLCYL

(1) +

Pilih rentang tanggal

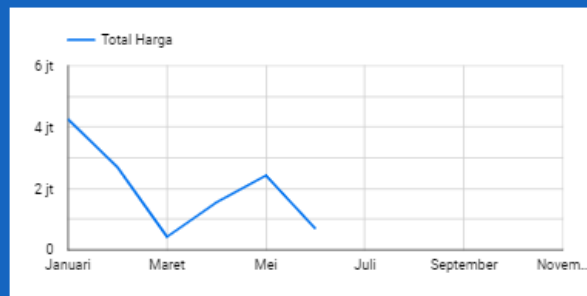
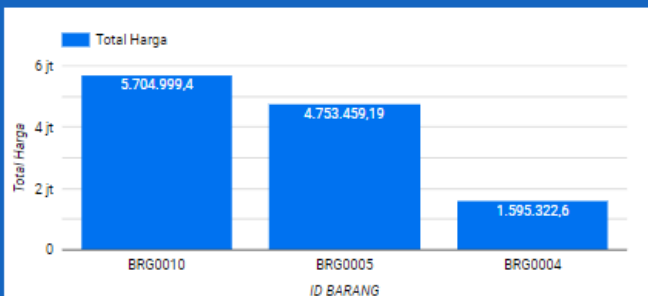
-

JUMLAH BARANG

1.893

TOTAL PENJUALAN

12.053.781



Soal 6 : Additional Complementary Data

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

Jawaban:

- Data koordinat dari suatu lokasi bisa dalam bentuk (latitude dan longitude) untuk mengetahui penyebaran disuatu tempat atau daerah.