

# Soal & Template Jawaban

Task 5

Nama : Mohammad Sulaeman

# 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 : pilihan b lebih fleksibel karena dapat menemukan baris dengan kata-kata lain yang dimulai dengan 'Mat' setelah karakter pertama sedangkan pilihan a hanya cocok untuk kasus saat ingin mencari baris dengan tiga karakter pertama alamat yang tepat 'Mat' dan tidak bisa mencari kata lain yang dimulai dengan 'Mat' setelah tiga karakter pertama. Selain itu pilihan b juga lebih hemat dan cepat dalam proses run.

# 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 : pilihan b lebih fleksibel dan aman dalam menangani format tanggal yang berbeda dan memastikan bahwa semua data yang relevan diambil dengan benar, karena pilihan a memiliki potensi kesalahan jika format tanggal di kolom tanggal\_lahir bukan dalam format yang diharapkan (misalnya, jika formatnya adalah DD-MM-YYYY atau MM-DD-YYYY)

## Soal 3: Menentukan Primary Key

### A. Tugas

Tentukan primary key dari table penjualan. jelaskan alasannya

### B. Jawaban & Penjelasan : primary key dari table penjualan adalah id\_invoice karena salah satu syarat dari primary ke adalah unik dan not null, dan pada table penjualan id\_invoice lah yang memenuhi persyaratan ini. Selain itu juga ada id\_costumer yang menjadi foreign key pada table ini.

## 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	hamud	<a href="https://lookerstudio.google.com/u/0/reporting/ba123c51-5d79-478f-9161-6abef3e6133f/page/p_1ha9wgkl7c/edit">https://lookerstudio.google.com/u/0/reporting/ba123c51-5d79-478f-9161-6abef3e6133f/page/p_1ha9wgkl7c/edit</a>

# Table Base “<<penjualan>>”

	id_invoice	id_barang	id_cabang	id_customer	id_distributo...	lini	tanggal	unit	mata_uang	jumlah_b...	brand_id	h...
1...	IN6287	BRG0002	CAB01	CUST55391	TD	ETIKAL	Jan 30, 2022	DUS	IDR	17	BRND002	2...
2...	IN6280	BRG0005	CAB04	CUST55384	TD	SLCYL	Jan 23, 2022	DUS	IDR	1	BRND005	5...
3...	IN6208	BRG0006	CAB09	CUST55475	TD	OGB & ...	Jan 30, 2022	DUS	IDR	8	BRND006	2...
4...	IN6054	BRG0006	CAB03	CUST55435	TD	OGB & ...	Feb 1, 2022	DUS	IDR	34	BRND006	2...
5...	IN6235	BRG0002	CAB01	CUST55451	TD	ETIKAL	Jan 30, 2022	DUS	IDR	10	BRND002	2...
6...	IN6002	BRG0007	CAB05	CUST55546	TD	ETIKAL	Feb 4, 2022	DUS	IDR	13.667832...	BRND007	4...
7...	IN6141	BRG0002	CAB10	CUST55401	TD	ETIKAL	Jan 28, 2022	DUS	IDR	25	BRND002	2...
8...	IN6055	BRG0004	CAB01	CUST55583	TD	VNS	Mar 13, 2022	DUS	IDR	4	BRND004	8...
9	IN6253	BRG0002	CAB10	CUST55431	TD	ETIKAL	Jan 28, 2022	DUS	IDR	34	BRND002	2...

## Table Base “<<Nama Tabel>>”

column	data type	description	transformation
id_invoice	string	primary key table penjualan	
id_barang	string	foreign key ke table barang	
id_costumer	string	foreign key ke table cstumer	
unit	string	nama barang	
harga	int	harga barang	
jumlah_barang	int	jumlah barang	



## Table Aggregate “<<penjualan>>”

	id_invoice ▾	id_customer	unit	harga	jumlah_barang
1.	IN6308	CUST55397	TABLET	39,000	10
2.	IN6308	CUST55397	DUS	96,000	8
3.	IN6307	CUST55396	TABLET	110,500	41
4.	IN6307	CUST55396	BOTOL	81,700	46
5.	IN6307	CUST55396	DUS	208,000	8
6.	IN6306	CUST55430	DUS	96,000	45
7.	IN6306	CUST55430	TABLET	71,500	79
8.	IN6301	CUST55393	BOTOL	31,000	10
9.	IN6301	CUST55393	TABLET	79,300	47 91
					1 - 62 / 62 < >

# Table Aggregate “<<Nama Tabel>>”

column	data type	description	transformation
persen barang	percent	column berisi data persenan masing masing row terhadap jumlah total barang	data jumlah barang berubah menjadi persenan terhadap jumlah total barang
id_invoice	string	primary key table penjualan	
id_barang	string	foreign key ke table barang	
id_costumer	string	foreign key ke table cstumer	
unit	string	nama barang	
harga	int	harga barang	
jumlah_barang	int	jumlah barang	

# Soal 5 : Data Visualization

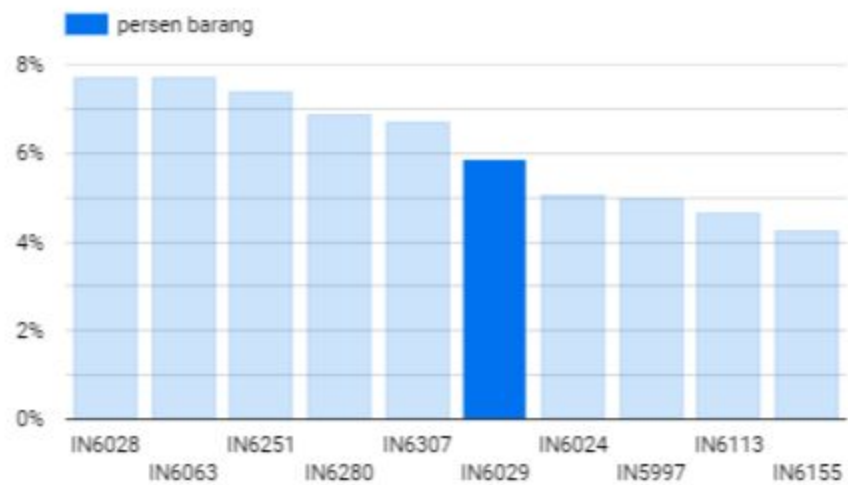
## A. Tugas

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) :



## 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 : memerlukan data transaksi karena data transaksi akan memudahkan untuk track barang yang terjual dan detailnya.