

Soal & Template Jawaban

Task 5

Nama : Athiya Fathinati Anindya

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 query (b) lebih simple dan singkat. Selain itu, penulisan alamat di query (a) harusnya 'Alamat'.

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 : Dengan operator BETWEEN, tanggal 2000-01-01 dan 2008-12-31 sudah termasuk dalam pencarian. Hal ini memudahkan dan mempersingkat kita mengetik query. Sedangkan dengan query (a), kita perlu mengetik lebih banyak dan kemungkinan kesalahan syntax lebih besar karena kita harus memastikan dengan benar penggunaan operator =, >, < dan AND.

Soal 3: Menentukan Primary Key

A. Tugas

Tentukan primary key dari table penjualan. jelaskan alasannya

B. Jawaban & Penjelasan : id_invoice. Karena kolom ini semuanya bernilai unique value. Hanya ada 1 row untuk 1 id_invoice, karena ini menunjukkan 1 pesanan.

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	kimiafarma.sql	https://drive.google.com/drive/folders/17Vm9xIU8dnDyK5zrcMV2dDKYHK8hFYZ2?usp=drive_link
2	kimiafarma_agg.sql	

Table Base “table_base”

```
1  -- delete null values
2  delete from public.pelanggan
3  where id_customer isnull;
4
5  -- rename 'nama'
6  ALTER TABLE public.pelanggan
7  RENAME nama TO nama_customer;
8
9  -- rename 'lini'
10 ALTER TABLE public.penjualan
11 RENAME lini TO brand;
12
13 -- join column 'nama_customer', 'cabang_sales' & 'group' from pelanggan and 'nama_barang' from barang
14 SELECT DISTINCT
15     pn.id_invoice,
16     pn.id_distributor,
17     pn.tanggal,
18     pl.nama_customer,
19     b.nama_barang,
20     b.kemasan,
21     pn.jumlah_barang,
22     pn.harga,
23     pn.brand,
24     pl.cabang_sales,
25     pl.group
26     INTO tabel_base
27 FROM public.penjualan AS pn
28 JOIN public.pelanggan AS pl ON pn.id_customer = pl.id_customer
29 JOIN public.barang AS b ON pn.id_barang = b.kode_barang
30 ORDER BY tanggal
```

Table Base “table_base”

[illegible]

Table Aggregate “agg_cabang”

```
4  -- sales by cabang_sales & group
5  select cabang_sales, "group",
6         avg(jumlah_barang * harga) as avg_sales,
7         sum(jumlah_barang * harga) as total_sales
8         into agg_cabang
9  from public.tabel_base
10 group by 1, 2
11 order by 4 desc
```

Table Aggregate “agg_cabang”





	cabang_sales character varying 	group character varying 	avg_sales double precision 	total_sales double precision 
1	Jakarta	Klinik	258541.58936170206	12151454.699999997
2	Kuningan	Apotek	170614.914893617	8018901
3	Bekasi	Apotek	130111.86347826089	5985145.720000001
4	Tangerang	Klinik	119236.46595744681	5604113.9
5	Bandung	Apotek	118839.40106382976	5585451.849999999
6	Aceh	Apotek	100646.64191489361	4730392.17
7	Padang	Apotek	92173.46739130435	4239979.5
8	Lampung	Klinik	126131.57826086957	2901026.3000000003

Table Aggregate “agg_brand”

```
22  --sales by brand
23  select brand,
24         avg(jumlah_barang * harga) as avg_sales,
25         sum(jumlah_barang * harga) as total_sales
26         into agg_brand
27  from public.tabel_base
28  group by 1
29  order by 3 desc
```

Table Aggregate “agg_brand”




	brand character varying 	avg_sales double precision 	total_sales double precision 
1	MARCKS	213010.66428571427	14910746.499999998
2	VNS	174286.04714285722	12200023.300000004
3	SLCYL	172082.39428571422	12045767.599999996
4	ETIKAL	93374.81142857144	6536236.800000001
5	OGB & PH	50338.44199999999	3523690.939999999

Table Aggregate “agg_sales_bulan”

```
40 --sales by month
41 with bybulan as(
42 select
43     extract(month from tanggal) as bulan,
44     tanggal, nama_customer, cabang_sales, "group", nama_barang, brand, jumlah_barang, harga
45 from public.tabel_base
46 order by bulan
47 )
48 select bulan,
49     to_char(to_timestamp (bulan::text, 'MM'), 'TMmon') as "month",
50     avg(jumlah_barang * harga) as avg_sales,
51     sum(jumlah_barang * harga) as total_sales
52 into agg_sales_bulan
53 from bybulan
54 group by bulan
55 order by bulan
```

Table Aggregate “agg_sales_bulan”





	bulan numeric 	month text 	avg_sales double precision 	total_sales double precision 
1	1	jan	145385.63028985503	20063216.979999993
2	2	feb	154760.9733962264	8202331.59
3	3	mar	82133.2790322581	2546131.650000001
4	4	apr	190592.44905660386	10101399.800000004
5	5	may	109998.7736538462	5719936.230000002
6	6	jun	112323.8647826087	2583448.89

Table Aggregate “agg_barang”

```
57 --sales by nama_barang
58 select nama_barang, brand,
59         avg(jumlah_barang * harga) as avg_sales,
60         sum(jumlah_barang * harga) as total_sales
61         into agg_barang
62 from public.tabel_base
63 group by 1, 2
64 order by 3 desc
```

Table Aggregate “agg_barang”

	nama_barang character varying	brand character varying	avg_sales double precision	total_sales double precision
1	AMPICILLIN	MARCKS	328048.6971428571	11481704.399999999
2	PARACETAMOL	SLCYL	237705.27500000002	5704926.600000001
3	TRAMADOL KAPSUL 50 MG	VNS	192658.35714285713	6743042.499999999
4	AMBROXOL HC	VNS	164027.36249999996	3936656.6999999993
5	TRAMADOL KAPSUL 50 MG	SLCYL	145115.36363636365	1596269
6	AMPICILLIN	VNS	138211.2818181818	1520324.0999999999
7	ACYCLOVIR DUS	ETIKAL	138180.46363636365	1519985.1
8	KLORPROMAZINA TABLET SALUT SELAPUT 100 MG	SLCYL	135559.2	4744572
9	ERGOTAMINE COFFEINE	ETIKAL	129535.4875	3108851.7
10	TETRACYCLINE KAPSUL 250 MG	MARCKS	108113.95833333333	2594735
11	ALERGINE TABLET SALUT	MARCKS	75846.09999999999	834307.0999999999
12	KETOCONAZOLE TABLET 200 MG	OGB & PH	68788.48000000001	2407596.8000000003
13	ALERGINE TABLET SALUT	ETIKAL	54497.142857142855	1907400
14	ACYCLOVIR DUS	OGB & PH	31888.404	1116094.14

Table Aggregate “agg_jenisbarang”

```
66 --sales by jenis_barang
67 select split_part(nama_barang, ' ', 1) as jenis_barang,
68        avg(jumlah_barang * harga) as avg_sales,
69        sum(jumlah_barang * harga) as total_sales
70        into agg_jenisbarang
71 from public.tabel_base
72 group by 1
73 order by 3 desc
```

Table Aggregate “agg_jenisbarang”

	jenis_barang text	avg_sales double precision	total_sales double precision
1	AMPICILLIN	282652.7934782608	13002028.499999998
2	TRAMADOL	181289.3804347826	8339311.499999999
3	PARACETAMOL	237705.27500000002	5704926.600000001
4	KLORPROMAZINA	135559.2	4744572
5	AMBROXOL	164027.36249999996	3936656.6999999993
6	ERGOTAMINE	129535.4875	3108851.7
7	ALERGINE	59602.328260869566	2741707.1
8	ACYCLOVIR	57306.0704347826	2636079.2399999998
9	TETRACYCLINE	108113.95833333333	2594735
10	KETOCONAZOLE	68788.48000000001	2407596.8000000003

Table Aggregate “agg_jenisbarang_bulan”

```
75 --top 5 jenis_barang per bulan
76 select split_part(nama_barang, ' ', 1) as jenis_barang,
77        avg(jumlah_barang * harga) as avg_sales,
78        sum(jumlah_barang * harga) as total_sales
79 from public.tabel_base
80 group by 1
81 order by 3 desc
82 limit 5;
83
84 with b as(
85 select bulan, "month",
86        split_part(nama_barang, ' ', 1) as jenis_barang,
87        avg(jumlah_barang * harga) as avg_sales,
88        sum(jumlah_barang * harga) as total_sales
89 from public.agg_bulan
90 group by 1, 2, 3
91 order by 1
92 )
93 select *
94 into agg_jenisbarang_bulan
95 from b
96 where jenis_barang in ('AMPICILLIN', 'TRAMADOL', 'PARACETAMOL', 'KLORPROMAZINA', 'AMBROXOL')
```

Table Aggregate “agg_jenisbarang_bulan”

	bulan numeric	month text	jenis_barang text	avg_sales double precision	total_sales double precision
1	1	January	AMBROXOL	129802.9909090909	1427832.9
2	1	January	AMPICILLIN	320263.1727272727	7045789.8
3	1	January	KLORPROMAZINA	141560.51875	2264968.3
4	1	January	PARACETAMOL	247537.3666666667	1485224.2
5	1	January	TRAMADOL	167613.5562500000	2681816.9000000004
6	2	February	AMBROXOL	222556.9499999999	445113.8999999999
7	2	February	AMPICILLIN	329111.1000000000	987333.3
8	2	February	KLORPROMAZINA	50834.7000000000	152504.1
9	2	February	PARACETAMOL	341066.1714285714	2387463.1999999997
10	2	February	TRAMADOL	218048.0222222222	1962432.2000000004
11	3	March	AMBROXOL	54916.6499999999	109833.2999999999
12	3	March	AMPICILLIN	279261.8000000000	1117047.2000000002
13	3	March	KLORPROMAZINA	28241.5	84724.5
14	3	March	PARACETAMOL	83283.6	166567.2

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/8ac07da2-91a1-4ce8-8e82-580f26c3da8d>

Date ▾

City ▾

Prod... ▾

Brand ▾

Sales Performance Dashboard

Product Name ▾



Product name:
ACYCLOVIR DUS
Brand:
OGB & PH
Package:
DUS
Price:
1,169.91

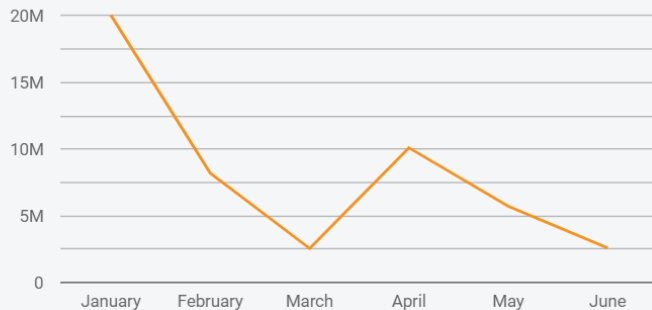
Sales
49.2M

Avg Order Value
140.6K

Orders
350.0

Quantity
9.2K

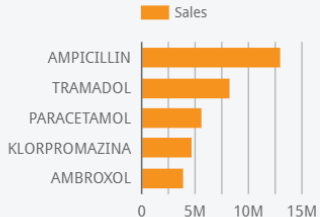
Sales Trend



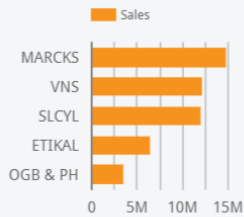
Sales by City



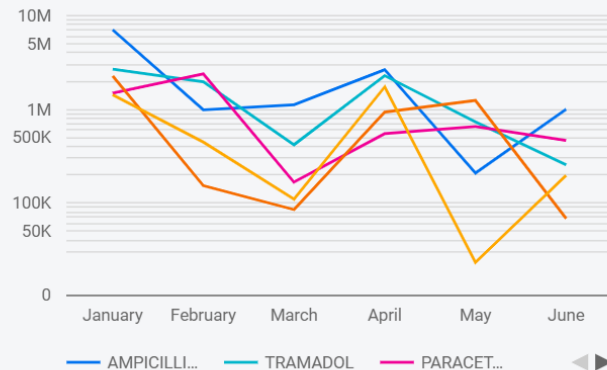
Sales by Product



Sales by Brand



Sales Trend by Product



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. Date rangenya diperbanyak hingga include tahun sebelumnya juga, agar bisa dibandingkan Sales tahun ini dengan tahun lalu.
2. Tambah kolom Cost atau/dan Profit tiap produk sehingga audience bisa mengurangi cost produk dan meningkatkan profit.