

FINAL PROJECT

SQL



DYON SETIO UTOMO



Dataset

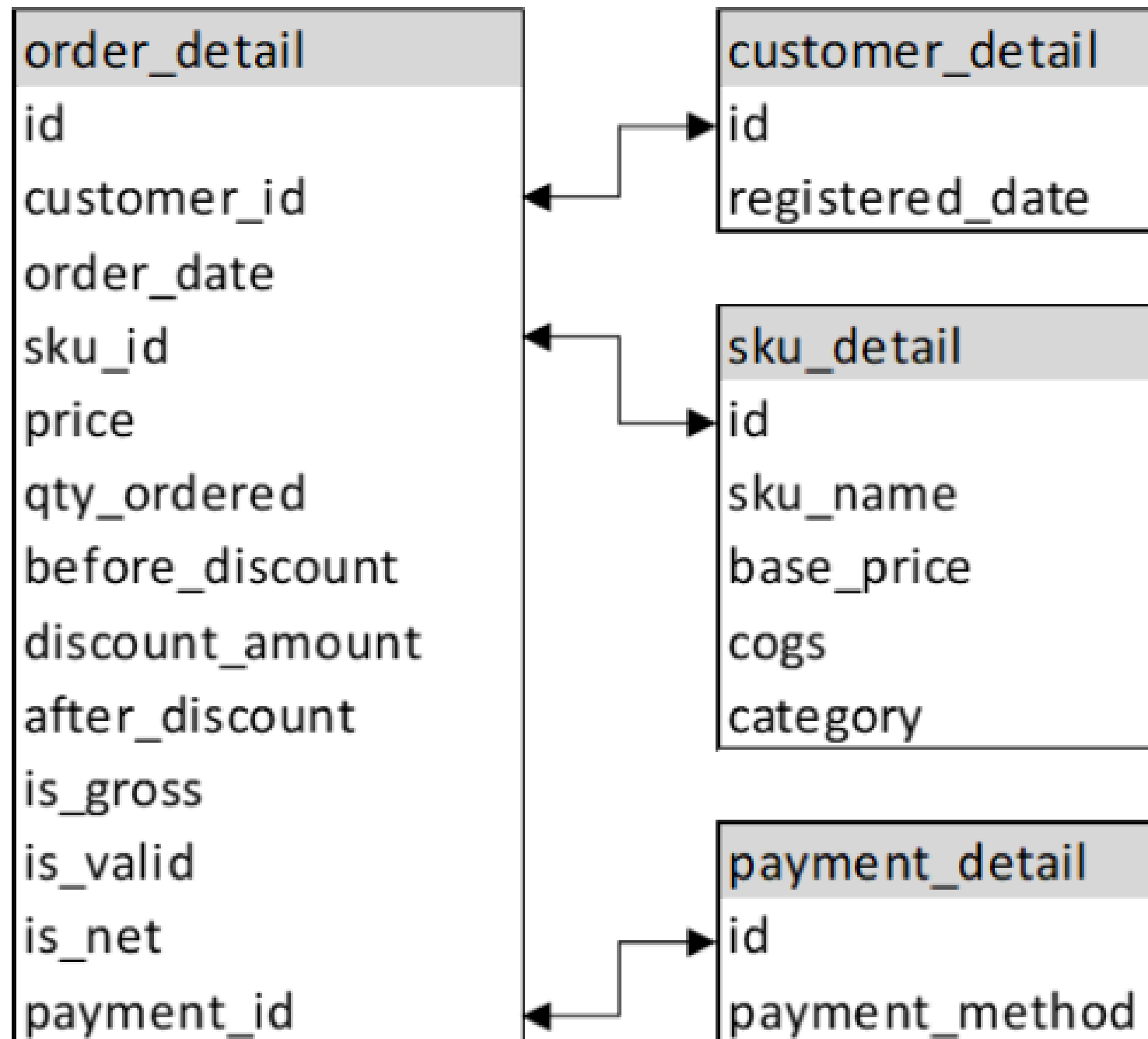
Dalam *project* ini, akan menggunakan dataset berikut:

1. **order_detail**
2. **sku_detail**
3. **customer_detail**
4. **payment_detail**

Data yang digunakan adalah data yang berasal dari Tokopedia (bukan data sesungguhnya).

Skema Tabel

Skema Tabel



Terdapat beberapa data yang saling terhubung. Hubungan ini dapat dijelaskan melalui skema tabel disamping.

Data **order_detail** menjadi data utama yang terhubung dengan data lain seperti **customer_detail**, **sku_detail**, dan **payment_detail**.

Pertanyaan 1

Selama transaksi yang terjadi selama 2021, pada bulan apa total nilai transaksi (after_discount) paling besar?

Tips:

Gunakan **is_valid = 1** untuk memfilter data transaksi.

Source table: order_detail

Query untuk Pertanyaan 1

```
2 SELECT
3     EXTRACT(MONTH FROM order_date) AS bulan,
4     SUM(after_discount) AS total_penjualan
5 FROM
6     `market_place.order_detail`
7 WHERE
8     EXTRACT(YEAR FROM order_date) = 2021
9     AND is_valid = 1
10 GROUP BY
11     bulan
12 ORDER BY
13     total_penjualan DESC;
```

EXTRACT(MONTH FROM order_date) AS bulan:

Menarik bulan dari tanggal pesanan.

SUM(after_discount) AS total_penjualan:

Menghitung total nilai transaksi setelah diskon.

EXTRACT(YEAR FROM order_date) = 2021:

Memfilter transaksi yang terjadi pada tahun 2021.

AND is_valid = 1:

Memfilter transaksi yang valid.

GROUP BY bulan:

Mengelompokkan data berdasarkan bulan.

ORDER BY total_penjualan DESC:

Mengurutkan hasil berdasarkan total nilai transaksi secara menurun.

Tabel Hasil untuk Pertanyaan 1

Row	bulan	total_penjualan
1	8	227862744.0
2	12	217309963.0
3	10	207603259.9
4	11	180396009.74
5	7	148007735.0
6	9	145943335.2000...
7	6	43154552.0
8	1	36822126.5
9	2	35611797.0
10	5	34163856.0
11	3	23643062.0
12	4	22208472.6

Berdasarkan data yang tersedia, kesimpulan dapat diambil bahwa **bulan Agustus** merupakan bulan dengan jumlah transaksi tertinggi pada tahun 2021, mencapai total penjualan sebesar **227.862.744**. Bulan Desember dan Oktober menyusul sebagai bulan dengan jumlah transaksi terbesar.

Perbedaan transaksi antar bulan cukup signifikan, dengan bulan bertransaksi paling tinggi memiliki jumlah transaksi sekitar lima kali lipat lebih besar daripada bulan dengan jumlah transaksi terendah.

Pertanyaan 2

Selama transaksi pada tahun 2022, kategori apa yang menghasilkan nilai transaksi paling besar?

Tips:

Gunakan **is_valid = 1** untuk memfilter data transaksi.

Source table: order_detail, sku_detail

Query untuk Pertanyaan 2

```
17 SELECT
18     sd.category AS kategori,
19     SUM(od.after_discount) AS total_penjualan
20 FROM
21     'market_place.order_detail' od
22 JOIN
23     'market_place.sku_detail' sd
24 ON
25     od.sku_id = sd.id
26 WHERE
27     EXTRACT(YEAR FROM od.order_date) = 2022
28     AND od.is_valid = 1
29 GROUP BY
30     kategori
31 ORDER BY
32     total_penjualan DESC;
```

sd.category AS kategori:

Memilih kategori produk dari tabel sku_detail.

SUM(od.after_discount) AS total_penjualan:

Menghitung total nilai after_discount dari tabel order_detail.

JOIN 'market_place.sku_detail sd ON od.sku_id = sd.id:

Menggabungkan tabel order_detail dan sku_detail berdasarkan sku_id.

EXTRACT(YEAR FROM od.order_date) = 2022:

Memfilter transaksi yang terjadi pada tahun 2022.

AND od.is_valid = 1:

Memfilter transaksi yang valid.

GROUP BY sd.kategori:

Mengelompokkan data berdasarkan kategori produk.

ORDER BY total_penjualan DESC:

Mengurutkan hasil berdasarkan total nilai transaksi secara menurun.

Tabel Hasil untuk Pertanyaan 2

Row	kategori	total_penjualan
1	Mobiles & Tablets	918451576.0
2	Entertainment	365344148.9
3	Appliances	316358100.0
4	Computing	214028543.4
5	Men Fashion	135588253.0
6	Women Fashion	93014970.62
7	Home & Living	79483716.2
8	Health & Sports	54235579.6
9	Beauty & Grooming	46211019.18

Dari data yang sudah diperoleh, dapat disimpulkan bahwa kategori "**Mobiles & Tablets**" merupakan kontributor utama terhadap total penjualan pada tahun 2022. Hal ini mengindikasikan tingginya permintaan pasar terhadap produk-produk elektronik, khususnya perangkat mobile.

Pertanyaan 3

Bandingkan nilai transaksi dari masing-masing kategori pada tahun 2021 dengan 2022. Sebutkan kategori apa saja yang mengalami peningkatan dan kategori apa yang mengalami penurunan nilai transaksi dari tahun 2021 ke 2022.

Tips:

Gunakan **is_valid = 1** untuk memfilter data transaksi

Source table: order_detail, sku_detail

Query untuk Pertanyaan 3

```
35 WITH kategori_pertahun AS (  
36     SELECT  
37         sd.category AS kategori,  
38         SUM(CASE WHEN EXTRACT(year FROM od.order_date) = 2021 THEN after_discount END) AS total_transaksi_2021,  
39         SUM(CASE WHEN EXTRACT(year FROM od.order_date) = 2022 THEN after_discount END) AS total_transaksi_2022  
40     FROM  
41         `market_place.order_detail` AS od  
42     LEFT JOIN  
43         `market_place.sku_detail` AS sd  
44     ON  
45         od.sku_id = sd.id  
46     WHERE  
47         EXTRACT(year FROM od.order_date) IN (2021, 2022)  
48         AND od.is_valid = 1  
49     GROUP BY  
50         1  
51 )  
52 SELECT  
53     kategori,  
54     total_transaksi_2021,  
55     total_transaksi_2022,  
56     CASE  
57         WHEN total_transaksi_2022 - total_transaksi_2021 < 0 THEN 'Penurunan'  
58         ELSE 'Kenaikan'  
59     END AS perubahan_transaksi  
60 FROM  
61     kategori_pertahun  
62 ORDER BY  
63     4
```

Common Table Expression (CTE)
'kategori_pertahun'

Query Utama

Query untuk Pertanyaan 3

(Common Table Expression (CTE) 'kategori_pertahun')

CTE (kategori_pertahun):

Bagian ini menyimpan hasil sementara dari query yang menghitung total nilai transaksi untuk setiap kategori (category) pada tahun 2021 dan 2022.

sd.category AS kategori:

Mengambil kategori produk dari tabel sku_detail dan memberinya alias kategori.

SUM(CASE WHEN EXTRACT(year FROM od.order_date) = 2021 THEN after_discount END) AS total_transaksi_2021:

Menghitung total nilai transaksi setelah diskon (after_discount) pada tahun 2021.

SUM(CASE WHEN EXTRACT(year FROM od.order_date) = 2022 THEN after_discount END) AS total_transaksi_2022:

Menghitung total nilai transaksi setelah diskon (after_discount) pada tahun 2022.

LEFT JOIN:

Menggabungkan tabel order_detail (od) dengan tabel sku_detail (sd) berdasarkan sku_id.

WHERE EXTRACT(year FROM od.order_date) IN (2021, 2022):

Memfilter transaksi yang terjadi pada tahun 2021 dan 2022.

AND od.is_valid = 1: Memfilter hanya transaksi yang valid.

GROUP BY 1: Mengelompokkan data berdasarkan kategori produk (sd.category).

Query untuk Pertanyaan 3

(Query Utama)

SELECT:

Mengambil data dari CTE kategori_pertahun.

kategori:

Menampilkan kategori produk.

total_transaksi_2021:

Menampilkan total nilai transaksi untuk tahun 2021.

total_transaksi_2022:

Menampilkan total nilai transaksi untuk tahun 2022.

CASE WHEN total_transaksi_2022 - total_transaksi_2021 < 0 THEN 'Penurunan' ELSE 'Kenaikan' END AS

perubahan_transaksi:

Menghitung perubahan nilai transaksi antara tahun 2021 dan 2022 untuk setiap kategori:

- Jika total transaksi di tahun 2022 lebih rendah daripada tahun 2021, maka akan dihasilkan nilai 'Penurunan'.
- Jika total transaksi di tahun 2022 lebih tinggi atau sama dengan tahun 2021, maka akan dihasilkan nilai 'Kenaikan'.

ORDER BY 4:

Mengurutkan hasil berdasarkan kolom keempat, yaitu perubahan_transaksi (menampilkan kategori dengan perubahan 'Penurunan' terlebih dahulu, kemudian 'Kenaikan').

Tabel Hasil untuk Pertanyaan 3

Row	kategori	total_transaksi_2021	total_transaksi_2022	perubahan_transaksi
1	School & Education	11558982.399999...	17362465.299999...	Kenaikan
2	Soghaat	15056202.6	17658332.0	Kenaikan
3	Kids & Baby	23971057.799999...	25931276.84	Kenaikan
4	Computing	172878860.0	214028543.4	Kenaikan
5	Men Fashion	58628198.0	135588253.0	Kenaikan
6	Home & Living	45797873.0	79483716.2	Kenaikan
7	Health & Sports	33837965.6	54235579.599999...	Kenaikan
8	Beauty & Grooming	46047360.0	46211019.18	Kenaikan
9	Women Fashion	84045961.4	93014970.62	Kenaikan
10	Superstore	28828088.0	32643266.519999...	Kenaikan
11	Appliances	218550176.99999...	316358100.0	Kenaikan
12	Mobiles & Tablets	370606718.0	918451576.0	Kenaikan
13	Entertainment	162326357.4	365344148.9	Kenaikan
14	Books	10124596.0	6792519.2	Penurunan
15	Others	40468515.74	21744646.02	Penurunan

Secara keseluruhan, data menunjukkan **tren positif** dalam nilai transaksi pada tahun 2022 dibandingkan dengan tahun 2021. Kategori **elektronik dan fashion** menjadi pendorong utama pertumbuhan ini.

Namun, perlu dilakukan analisis lebih lanjut untuk memahami faktor-faktor yang menyebabkan **penurunan** pada kategori **Books** dan **Others**.

Pertanyaan 4

Tampilkan top 5 metode pembayaran yang paling populer digunakan selama 2022 (berdasarkan total unique order).

Tips:

Gunakan **is_valid = 1** untuk memfilter data transaksi

Source table: order_detail, payment_method

Query untuk Pertanyaan 4

```
3  SELECT
4      pd.payment_method AS metode_pembayaran,
5      COUNT(DISTINCT od.id) AS total_unique_orders
6  FROM
7      `market_place.order_detail` od
8  JOIN
9      `market_place.payment_detail` pd
10 ON
11     od.payment_id = pd.id
12 WHERE
13     EXTRACT(YEAR FROM od.order_date) = 2022
14     AND od.is_valid = 1
15 GROUP BY
16     metode_pembayaran
17 ORDER BY
18     total_unique_orders DESC
19 LIMIT 5;
```


Query untuk Pertanyaan 4

pd.payment_method:

Memilih metode pembayaran dari tabel payment_detail.

COUNT(DISTINCT od.id) AS total_unique_orders:

Menghitung jumlah pesanan unik menggunakan COUNT(DISTINCT od.id).

JOIN market_place.payment_detail pd ON od.payment_id = pd.id:

Menggabungkan tabel order_detail dan payment_detail berdasarkan payment_id.

EXTRACT(YEAR FROM od.order_date) = 2022:

Memfilter transaksi yang terjadi pada tahun 2022.

AND od.is_valid = 1:

Memfilter transaksi yang valid.

GROUP BY pd.payment_method:

Mengelompokkan data berdasarkan metode pembayaran.

ORDER BY total_unique_orders DESC:

Mengurutkan hasil berdasarkan jumlah pesanan unik secara menurun.

LIMIT 5:

Membatasi hasil hanya pada 5 metode pembayaran teratas.

Tabel Hasil untuk Pertanyaan 4

Row	metode_pembayaran	total_unique_orders
1	cod	1809
2	Payaxis	181
3	customercredit	75
4	Easypay	69
5	jazzwallet	26

Meskipun tren pembayaran digital semakin meningkat, COD masih menjadi pilihan utama bagi sebagian besar konsumen. Platform seperti Payaxis, Easypay, dan Jazzwallet menunjukkan peningkatan penggunaan metode pembayaran digital. Metode pembayaran seperti customer credit memberikan fleksibilitas bagi konsumen.

Pertanyaan 5

Urutkan dari ke-5 produk ini berdasarkan nilai transaksinya.

- Samsung
- Apple
- Sony
- Huawei
- Lenovo

Tips:

Gunakan **is_valid = 1** untuk memfilter data transaksi

Source table: order_detail, payment_method

Query untuk Pertanyaan 5

```
3 WITH top_produk AS (  
4     SELECT  
5         CASE  
6             WHEN LOWER(sd.sku_name) LIKE '%samsung%' THEN 'Samsung'  
7             WHEN LOWER(sd.sku_name) LIKE '%apple%' THEN 'Apple'  
8             WHEN LOWER(sd.sku_name) LIKE '%iphone%' THEN 'Apple'  
9             WHEN LOWER(sd.sku_name) LIKE '%imac%' THEN 'Apple'  
10            WHEN LOWER(sd.sku_name) LIKE '%macbook%' THEN 'Apple'  
11            WHEN LOWER(sd.sku_name) LIKE '%sony%' THEN 'Sony'  
12            WHEN LOWER(sd.sku_name) LIKE '%huawe%' THEN 'Huawei'  
13            WHEN LOWER(sd.sku_name) LIKE '%lenovo%' THEN 'Lenovo'  
14            ELSE 'Other Brand'  
15        END AS nama_produk,  
16        od.after_discount AS nilai_transaksi  
17    FROM  
18        'market_place.order_detail' AS od  
19    LEFT JOIN  
20        'market_place.sku_detail' AS sd  
21    ON  
22        od.sku_id = sd.id  
23    WHERE  
24        od.is_valid = 1  
25        AND od.id IS NOT NULL  
26 )
```

Common Table Expression (CTE)
'top_produk'

Query Utama

```
28 SELECT  
29     nama_produk,  
30     SUM(nilai_transaksi) AS total_nilai_transaksi  
31 FROM  
32     top_produk  
33 WHERE  
34     nama_produk != 'Other Brand'  
35 GROUP BY  
36     nama_produk  
37 ORDER BY  
38     total_nilai_transaksi DESC;
```

Query untuk Pertanyaan 5

Common Table Expression (CTE) **top_produk**

WITH top_produk AS (...): Bagian ini mendefinisikan sebuah CTE bernama **top_produk** yang menyimpan data sementara mengenai produk-produk populer dan nilai transaksi terkait.

CASE: Struktur kondisional ini digunakan untuk mengelompokkan produk-produk berdasarkan nama merek mereka:

- Menggunakan fungsi **LOWER** untuk memastikan pencarian tidak sensitif terhadap huruf besar atau kecil.
- **LIKE '%samsung%':** Jika nama produk (**sku_name**) mengandung kata "samsung", maka akan digolongkan sebagai 'Samsung'.
- Pencarian ini juga dilakukan untuk produk lain seperti 'Apple', 'Sony', 'Huawei', dan 'Lenovo'.
- Jika nama produk tidak cocok dengan salah satu kondisi, maka produk tersebut akan digolongkan sebagai '**Other Brand**'.

od.after_discount AS nilai_transaksi: Mengambil nilai transaksi setelah diskon dari tabel **order_detail**.

LEFT JOIN market_place.sku_detail AS sd ON od.sku_id = sd.id: Melakukan join antara tabel **order_detail** (**od**) dan **sku_detail** (**sd**) menggunakan sku_id untuk mendapatkan detail produk.

WHERE od.is_valid = 1 AND od.id IS NOT NULL: Memfilter hanya transaksi yang valid (sudah dibayar) dan memiliki **id** yang tidak kosong.

Query untuk Pertanyaan 5

Query Utama

SELECT nama_produk, SUM(nilai_transaksi) AS total_nilai_transaksi: Mengambil **nama_produk** dan menghitung total nilai transaksi (**nilai_transaksi**) untuk setiap merek produk.

FROM top_produk: Mengambil data dari CTE **top_produk** yang telah dibuat sebelumnya.

WHERE nama_produk != 'Other Brand': Memfilter hanya produk yang termasuk dalam merek yang diinginkan dan mengecualikan produk yang dikategorikan sebagai 'Other Brand'.

GROUP BY nama_produk: Mengelompokkan data berdasarkan **nama_produk**.

ORDER BY total_nilai_transaksi DESC: Mengurutkan hasil berdasarkan total nilai transaksi dalam urutan menurun (dari yang terbesar ke terkecil).

Tabel Hasil untuk Pertanyaan 5

Row	nama_produk	total_nilai_transaksi
1	Samsung	588764148.0
2	Apple	444855360.0
3	Sony	63960718.0
4	Huawei	63160260.0
5	Lenovo	62379800.4

Dari data yang diberikan, dapat disimpulkan bahwa Samsung dan Apple merupakan pemimpin pasar dalam industri elektronik, diikuti oleh Sony, Huawei, dan Lenovo. Namun, persaingan di antara merek-merek ini sangat ketat dan terus berkembang.

TERIMA
KASIH