

Presented by Taufiq Hidayat HL



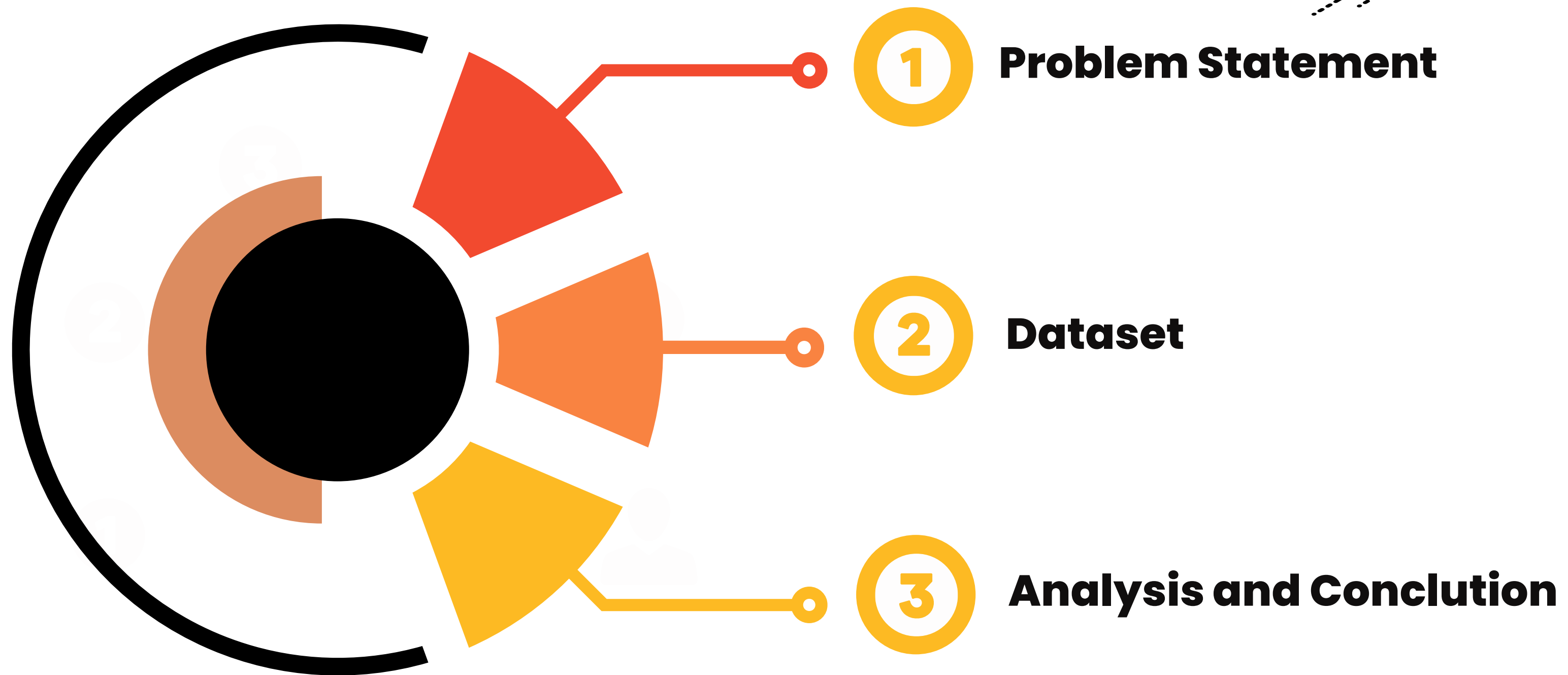
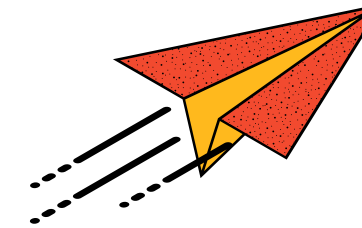
FINAL PROJECT

SQL

Bootcamp Data Analysis by @myskill.id



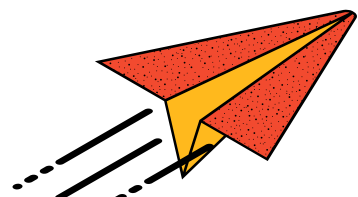
Table Of Content



PROBLEM STATEMENT



- Selama transaksi yang terjadi selama 2021, pada bulan apa total nilai transaksi (after_discount) paling besar? Gunakan is_valid = 1 untuk memfilter data transaksi. (Source table: order_detail)
- Selama transaksi pada tahun 2022, kategori apa yang menghasilkan nilai transaksi paling besar? Gunakan is_valid = 1 untuk memfilter data transaksi. (Source table: order_detail, sku_detail).
- 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. Gunakan is_valid = 1 untuk memfilter data transaksi. (Source table: order_detail, sku_detail).
- Tampilkan top 5 metode pembayaran yang paling populer digunakan selama 2022 (berdasarkan total unique order). Gunakan is_valid = 1 untuk memfilter data transaksi. (Source table: order_detail, payment_method).
- Urut dari ke-5 produk ini berdasarkan nilai transaksinya. (Samsung, Apple, Sony, Huawei, Lenovo) Gunakan is_valid = 1 untuk memfilter data transaksi. (Source table: order_detail, sku_detail).



Dataset



Terdapat 4 dataset yang digunakan dalam menganalisis dan menjawab pertanyaan di atas yaitu order_detail, sku_detail, customer_detail, payment_detail. Data tersebut diambil dari salah satu Perusahaan e-commerce yang memungkinkannya mengumpulkan transaksi dari pelanggannya, melakukan pembelian selama 2 tahun terakhir (2021 dan 2022). Mengenaipenjelasan dataset adalah sebagai berikut:

order_detail:

- | | |
|--------------------|--|
| 1. id | → angka unik dari order / <i>id_order</i> |
| 2. customer_id | → angka unik dari pelanggan |
| 3. order_date | → tanggal saat dilakukan transaksi |
| 4. sku_id | → angka unik dari produk (sku adalah <i>stock keeping unit</i>) |
| 5. price | → harga yang tertera pada tagging harga |
| 6. qty_ordered | → jumlah barang yang dibeli oleh pelanggan |
| 7. before_discount | → nilai harga total dari produk ($price * qty_ordered$) |
| 8. discount_amount | → nilai diskon product total |
| 9. after_discount | → nilai harga total produk ketika sudah dikurangi dengan diskon |
| 10. is_gross | → menunjukkan pelanggan belum membayar pesanan |
| 11. is_valid | → menunjukkan pelanggan sudah melakukan pembayaran |
| 12. is_net | → menunjukkan transaksi sudah selesai |
| 13. payment_id | → angka unik dari metode pembayaran |



Dataset



sku_detail:

- | | |
|---------------|---|
| 1. id | → angka unik dari produk (dapat digunakan untuk <i>key</i> saat join) |
| 2. sku_name | → nama dari produk |
| 3. base_price | → harga barang yang tertera pada tagging harga / <i>price</i> |
| 4. cogs | → <i>cost of goods sold</i> / total biaya untuk menjual 1 produk |
| 5. category | → kategori produk |

customer_detail:

- | | |
|--------------------|---|
| 1. id | → angka unik dari pelanggan |
| 2. registered_date | → tanggal pelanggan mulai mendaftarkan diri sebagai anggota |

Payment_detail:

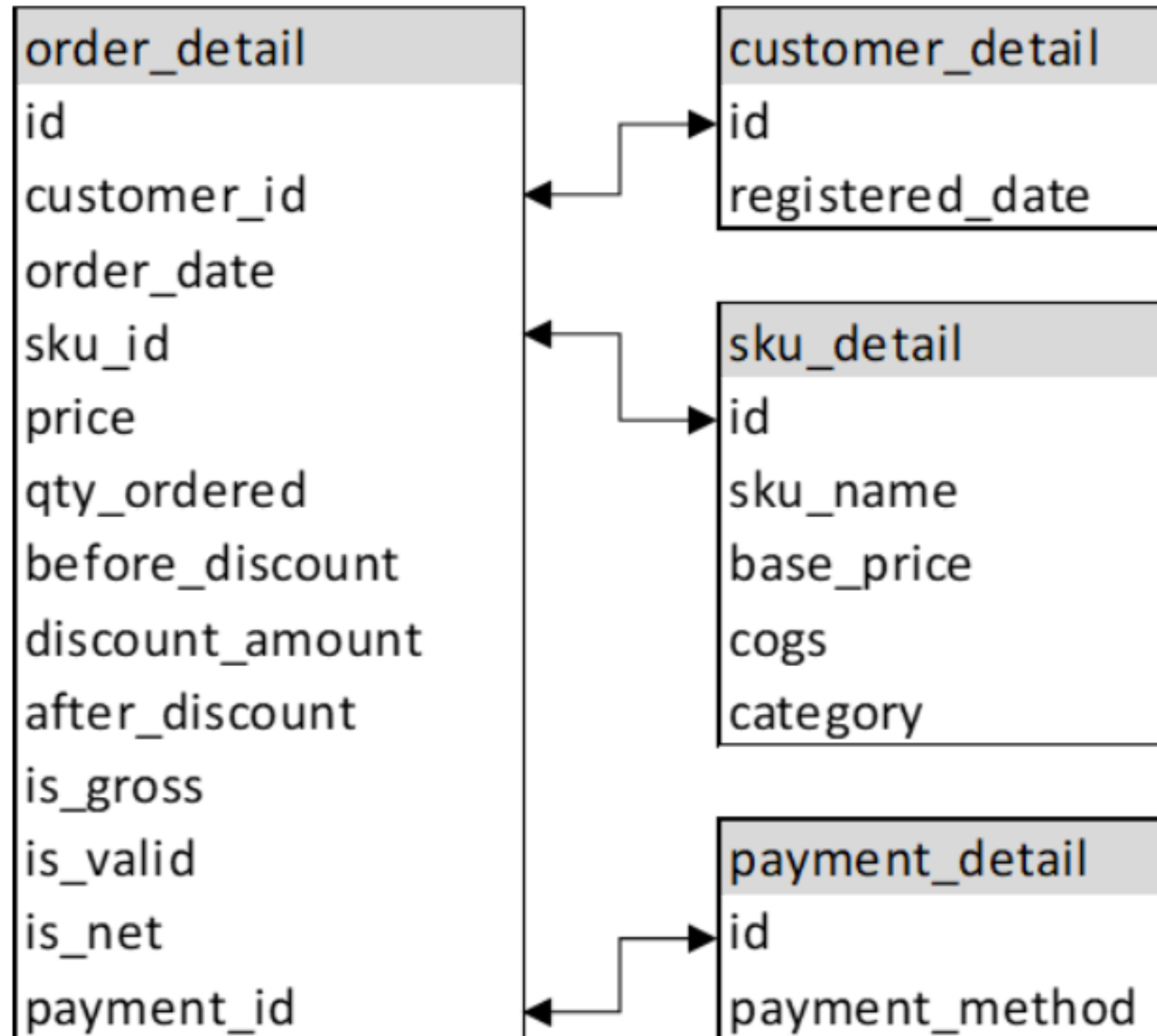
- | | |
|-------------------|-------------------------------------|
| 1. id | → angka unik dari metode pembayaran |
| 2. payment_method | → metode pembayaran yang digunakan |



Tabel Flow



Skema Tabel



Dataset



Tabel Order_detail

1 SELECT * FROM order_detail

id	customer_id	order_date	sku_id	price	qty_ordered	before_discount	discount_amount	after_discount	is_gross	is_val
ODR9939707760w	C713589L	2021-11-19	P858068	26100	200	5220000	2610000	2610000	1	1
ODR7448356649d	C551551L	2021-11-19	P886455	1971942	5	9859710	2464927.5	7394782.5	1	0
ODR4011281866z	C685596L	2021-11-25	P678648	7482000	1	7482000	2065344.62	5416655.38	1	0
ODR3378927994s	C830683L	2021-11-22	P540013	3593680	1	3593680	1455440.4	2138239.6	1	1
ODR4904430099k	C191766L	2021-11-21	P491032	4413220	1	4413220	1059172.8	3354047.2	1	1
ODR7618778722h	C299859L	2021-11-20	P886455	1971942	2	3943884	985971	2957913	1	0
ODR7610732813d	C313534L	2022-12-01	P849301	2697000	1	2697000	809100	1887900	1	1
ODR4415476736l	C215074L	2022-12-01	P918122	2533672	1	2533672	760101.6	1773570.4	1	1
ODR6474755771h	C368335L	2021-11-22	P240605	2479790	1	2479790	743937	1735853	1	0
ODR9699658949w	C246762L	2022-05-21	P168052	6487648	1	6487648	638000	5849648	1	0
ODR6859592351r	C822265L	2021-01-10	P910456	4092886	5	20464430	613932.9	19850497.1	1	0
ODR2741972772y	C483469L	2021-11-21	P239399	3683000	1	3683000	552450	3130550	1	1
ODR5064165483d	C329053L	2021-11-23	P239399	3683000	1	3683000	552450	3130550	1	1
ODR8655835343e	C545190L	2021-11-24	P239399	3683000	1	3683000	552450	3130550	1	1
ODR6859592351r	C822265L	2021-01-10	P240605	2479790	1	2479790	743937	1735853	1	0

Dataset



Tabel sku_detail

```
1 SELECT * FROM sku_detail
```

id	sku_name	base_price	cogs	category
P798444	AT-FSM-35	57631.7	46052	Kids & Baby
P938347	AYS_Haier-18HNF	3931789.26	3499256	Appliances
P826364	Atalian_DV206A-Brown-41	324597	243426	Men Fashion
P467533	Darul_Sakoon_Food_Bundle	2870.42	2378	Superstore
P229955	HP_15AY-15-Ay072NIA-ci3	2265625	1631250	Computing
P985828	UnzeLondon_GS4666-9	202855	172376	Men Fashion
P381119	UnzeLondon_GS5348-10	202855	139954	Men Fashion
P499423	UnzeLondon_GS5348-9	202855	158224	Men Fashion
P672852	Paramount_9781107677364	182856.6	138968	School & Education
P447157	Paramount_9781107697690	182856.6	153584	School & Education
P346967	AKL_A131128712_SS-48_Yellow	173275	103936	Women Fashion
P129115	Jaza_2Premium5kg+FOC	99302.38	75458	Superstore
P176799	Paramount_9781107657823	96552.6	71398	School & Education
P834446	emo_Emo-VB-06 Navy Blue Blazer	93745.4	84332	Women Fashion
P948478	Jaza_2Sela5kg+FOC	93502.38	61654	Superstore

Dataset



Tabel customer_detail

```
1 SELECT * FROM customer_detail
```

id	registered_date
C996508L	2021-07-10
C180415L	2021-07-18
C535451L	2021-07-23
C177843L	2021-07-12
C951682L	2021-07-27
C483469L	2021-07-26
C346553L	2021-07-11
C631361L	2021-07-15
C534453L	2021-07-18
C404192L	2021-07-06
C720927L	2021-07-17

Tabel payment_detail

```
1 SELECT * FROM payment_detail
```

id	payment_method
1	cod
2	jazzvoucher
3	customercredit
4	Payaxis
5	jazzwallet
6	easypay_voucher
7	Easypay
8	ublcreditcard
9	mygateway
10	mcblite
11	cashatdoorstep
12	internetbanking
13	Easypay_MA
14	productcredit
15	marketingexpense



Analysis and Conclution



Q1 : Selama transaksi yang terjadi selama 2021, pada bulan apa total nilai transaksi (after_discount) paling besar? Gunakan is_valid = 1 untuk memfilter data transaksi. (Source table: order_detail)

Pada pertanyaan di atas meminta kita untuk menampilkan bulan ketika total transaksi mencapai maksimum, artinya kita harus mendapatkan transaksi yang mana customer telah membayar (difilter menggunakan is_valid = 1), selain itu untuk mendapatkan nilai total transaksi kita memerlukan table **order_detail**.

berikut adalah query untuk menjawab pertanyaan diatas:



```
14 SELECT DISTINCT sku_detail.category,  
15    SUM(after_discount) AS Total_2022_after_discount  
16 FROM order_detail  
17 LEFT JOIN sku_detail  
18 ON order_detail.sku_id = sku_detail.id  
19 WHERE  
20    order_date BETWEEN '2022-01-01' AND '2022-12-31'  
21    AND is_valid=1  
22 GROUP BY sku_detail.category  
23 ORDER BY SUM(after_discount) DESC  
24
```



Analysis and Conclusion



Dari hasil query sebelumnya dapat dilihat bahwa Bulan Agustus merupakan bulan dengan total transaksi mencapai nilai maksimum di tahun 2021 yaitu sekitar ±227 juta

month	total_2021_after_discount
august	227862744
december	217309963
october	207603259.9
november	180396009.74
july	148007735
september	145943335.2
june	43154552
january	36822126.5
february	35611797
may	34163856



Analysis and Conclution



Q2 : Selama transaksi pada tahun 2022, kategori apa yang menghasilkan nilai transaksi paling besar? Gunakan `is_valid = 1` untuk memfilter data transaksi. (Source table: `order_detail`, `sku_detail`)

Pada pertanyaan di atas meminta kita untuk menampilkan kategory apa saja ketika total transaksi mencapai maksimum, artinya kita harus mendapatkan transaksi yang mana customer telah membayar (difilter meggunakan `is_valid = 1`),

Karena total nilai transaksi tidak berada dalam tabel yang sama dengan **kategori produk**, jadi kita akan menjawab pertanyaan diatas dengan menggabungkan (join) dua tabel, yaitu table **order_detail** (di mana terdapat kolom `after_discount`) dan tabel **sku_detail** (di mana terdapat kolom `category`)



```
14 SELECT DISTINCT sku_detail.category,  
15    SUM(after_discount) AS Total_2022_after_discount  
16 FROM order_detail  
17 LEFT JOIN sku_detail  
18 ON order_detail.sku_id = sku_detail.id  
19 WHERE  
20    order_date BETWEEN '2022-01-01' AND '2022-12-31'  
21    AND is_valid=1  
22 GROUP BY sku_detail.category  
23 ORDER BY SUM(after_discount) DESC  
24
```


Analysis and Conclution



Dari hasil tabel dapat dilihat bahwa **Mobiles & Tablets** adalah category yang mencapai nilai transaksi terbesar selama tahun 2022 yaitu sekitar 918.451.576 juta

category	total_2022_after_discount
Mobiles & Tablets	918451576
Entertainment	365344148.9
Appliances	316358100
Computing	214028543.4
Men Fashion	135588253
Women Fashion	93014970.62
Home & Living	79483716.2
Health & Sports	54235579.599999994
Beauty & Grooming	46211019.18
Superstore	32643266.519999996
Kids & Baby	25931276.84
Others	21744646.02
Soghaat	17658332
School & Education	17362465.3
Books	6792519.2



Analysis and Conclution



Q3: 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. Gunakan `is_valid = 1` untuk memfilter data transaksi.(Source table: `order_detail`, `sku_detail`)

Kita dapat menyelesaikan masalah di atas dengan membandingkan nilai penjualan setiap kategori produk pada tahun 2021 dan tahun 2022. Oleh karena itu, kita perlu membuat tabel sementara yang berisi nilai penjualan dari tabel **Order_detail** dan **sku_detail** serta menghitung perbedaan antara dua periode yang berbeda. Jika perbedaannya positif, maka kategori tersebut mengalami peningkatan nilai transaksi, sebaliknya Jika perbedaannya Negatif, maka kategori tersebut mengalami penurunan nilai transaksi.

Selain itu, karena kolom kategori berada di tabel `sku_detail` sedangkan nilai penjualan (`after_discount`) ada di tabel `order_detail`, kita perlu melakukan join antara kedua tabel tersebut.



Analysis and Conclution



Q3: 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. Gunakan is_valid = 1 untuk memfilter data transaksi.(Source table: order_detail, sku_detail)

Berikut adalah query untuk menyelesaikan pertanyaan diatas:

```
7 WITH sales_revenue AS
8 (SELECT
9     sd.category,
10     SUM(CASE WHEN EXTRACT(YEAR FROM order_date) = 2021 THEN after_discount END) AS revenue_2021,
11     SUM(CASE WHEN EXTRACT(YEAR FROM order_date) = 2022 THEN after_discount END) AS revenue_2022
12 FROM order_detail od
13 LEFT JOIN sku_detail sd
14 ON od.sku_id = sd.id
15 WHERE
16     EXTRACT(YEAR FROM order_date) IN (2021,2022)
17     AND is_valid = 1
18 GROUP BY sd.category
19 ORDER BY 2,3 DESC)
20 SELECT
21     *,
22     (revenue_2022 - revenue_2021) AS revenue_diff
23 FROM sales_revenue
24 ORDER BY revenue_diff DESC
```



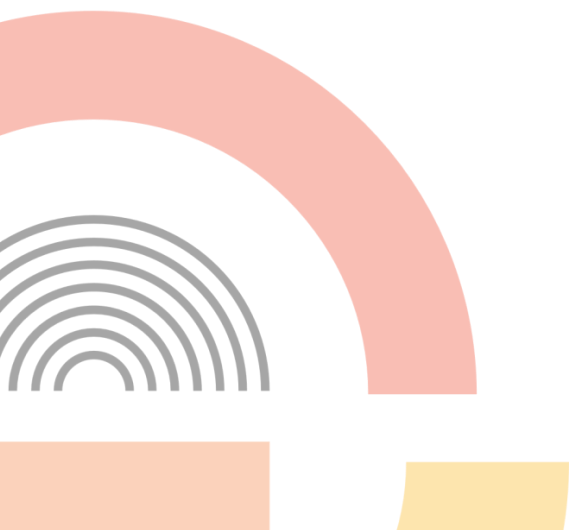
Analysis and Conclution



Berdasarkan hasil query diatas dapat dilihat terdapat 2 kategori yang mengalami penurunan pendapatan di tahun 2022 yaitu kategori **Books dan other**

Selain itu 13 category lainnya mengalami kenaikan pendapatan di tahun 2022. salah satunya kategori **Mobile and Tablets** mencapai peningkatan pendapatan tertinggi yaitu sebanyak 547.844.858 juta

category	revenue_2021	revenue_2022	revenue_diff
Mobiles & Tablets	370606718	918451576	547844858
Entertainment	162326357.4	365344148.9	203017791.49999997
Appliances	218550177	316358100	97807923
Men Fashion	58628198	135588253	76960055
Computing	172878860	214028543.4	41149683.400000006
Home & Living	45797873	79483716.2	33685843.2
Health & Sports	33837965.6	54235579.6	20397614
Women Fashion	84045961.4	93014970.62	8969009.219999999
School & Education	11558982.4	17362465.3	5803482.9
Superstore	28828088	32643266.519999996	3815178.519999996
Soghaat	15056202.6	17658332	2602129.4000000004
Kids & Baby	23971057.8	25931276.84	1960219.039999999
Beauty & Grooming	46047360	46211019.18	163659.1799999997
Books	10124596	6792519.2	-3332076.8
Others	40468515.74	21744646.02	-18723869.720000003



Analysis and Conclusion



Q4: Tampilkan top 5 metode pembayaran yang paling populer digunakan selama 2022 (berdasarkan total unique order). Gunakan `is_valid = 1` untuk memfilter data transaksi. (Source table: `order_detail`, `payment_method`)

Pertanyaan di atas meminta kita untuk menampilkan metode pembayaran beserta jumlah total pesanan unik yang sesuai. Kita mendefinisikan jumlah total pesanan unik dengan `nou_order` yang dihitung dengan menghitung secara unik jumlah id dalam tabel `order_detail`.

Karena `payment_method` ada di tabel `payment_detail` dan berada di tabel yang berbeda dengan `nou_order`, kita perlu melakukan join (dalam hal ini LEFT JOIN) antara tabel `payment_detail` (didefinisikan sebagai `pd`) dan tabel `order_detail` (didefinisikan sebagai `od`). Setelah itu, kita akan mengambil nilai dalam periode tahun 2022 di mana pesanan telah dibayar (`is_valid=1`). Terakhir, kita akan mengurutkan tabel berdasarkan `payment_method` dan `nou_order` secara menurun, sebelum membatasi tabel hingga 5 baris teratas sesuai dengan persyaratan di atas.



Analysis and Conclution



Q4: Tampilkan top 5 metode pembayaran yang paling populerdigunakan selama 2022(berdasarkan total unique order). Gunakanis_valid = 1 untuk memfilter data transaksi. (Source table: order_detail, payment_method)



Berikut adalah query untuk menyelesaikan pertanyaan diatas:

```
5 SELECT
6     pd.payment_method,
7     COUNT (DISTINCT od.id) AS nou_order
8 FROM order_detail od
9 LEFT JOIN payment_detail pd
10 ON od.payment_id = pd.id
11 WHERE
12     order_date BETWEEN '2022-01-01' AND '2022-12-31'
13     AND is_valid = 1
14 GROUP BY payment_method
15 ORDER BY nou_order DESC
16 LIMIT 5
17
```



Analysis and Conclution



Dari hasil tabel menunjukkan bahwa **COD, Payaxis, Customer Credit, Easypay, dan Jazzwallet** merupakan Top 5 metode pembayaran yang terpopuler digunakan berdasarkan total unik order pada tahun 2022.

payment_method	nou_order
cod	1809
Payaxis	181
customercredit	75
Easypay	69
jazzwallet	26



Analysis and Conclution



Q5: Urutkan dari ke-5 produk ini berdasarkan nilai transaksinya. (Samsung, Apple, Sony, Huawei, Lenovo)Gunakanis_valid = 1 untuk memfilter data transaksi. (Source table: order_detail, sku_detail)

Pertanyaan di atas meminta kita untuk mengurutkan antara Samsung, Apple, Sony, Huawei, dan Lenovo. Diantara kelima produk tersebut yang mana memiliki nilai transaksi terbesar. Selain itu kita akan mengambil nilai total transaksi di mana pesanan telah dibayar(Is_Valid=1).

untuk menyortir nama nama produk di kolom Sku_name kita dapat menggunakan fungsi case then else sesuai dengan kelima nama produk yang di tanyakan.

karena kolom nama produk berada di tabel sku_detail sedangkan nilai penjualan (after_discount) ada di tabel order_detail, kita perlu melakukan join antara kedua tabel tersebut



Analysis and Conclusion



Q5: Urutkan dari ke-5 produk ini berdasarkan nilai transaksinya. (Samsung, Apple, Sony, Huawei, Lenovo)Gunakanis_valid = 1 untuk memfilter data transaksi. (Source table: order_detail, sku_detail)

Berikut adalah query untuk menyelesaikan pertanyaan diatas:

```
9
10 SELECT
11     CASE
12         WHEN LOWER(sku_name) LIKE '%samsung%' THEN 'Samsung'
13         WHEN LOWER(sku_name) LIKE '%macbook%'
14         OR LOWER(sku_name) LIKE '%iphone%'
15         OR LOWER(sku_name) LIKE '%apple%' THEN 'Apple'
16         WHEN LOWER(sku_name) LIKE '%sony%' THEN 'Sony'
17         WHEN LOWER(sku_name) LIKE '%huawei%' THEN 'Huawei'
18         WHEN LOWER(sku_name) LIKE '%lenovo%' THEN 'Lenovo'
19     END AS brand,
20     SUM(after_discount) AS total_revenue
21 FROM order_detail od
22 LEFT JOIN sku_detail sd
23 ON od.sku_id = sd.id
24 WHERE
25     is_valid = 1 AND (LOWER(sku_name) LIKE '%samsung%'
26     OR LOWER(sku_name) LIKE '%macbook%' OR LOWER(sku_name) LIKE '%iphone%'
27     OR LOWER(sku_name) LIKE '%apple%' OR LOWER(sku_name) LIKE '%sony%'
28     OR LOWER(sku_name) LIKE '%huawei%' OR LOWER(sku_name) LIKE '%lenovo%')
29 GROUP BY brand
30 ORDER BY total_revenue DESC
31
```



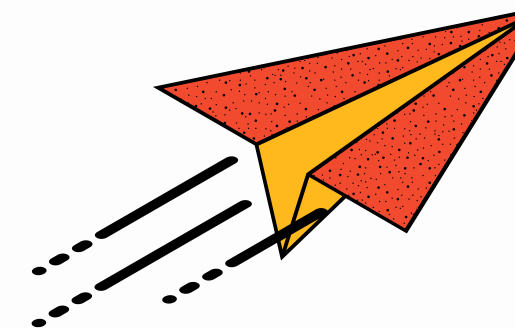
Analysis and Conclution



Berdasarkan hasil query diatas dapat dilihat bahwa brand yang memiliki Nilai transaksi tertinggi yaitu **Samsung** sebanyak ±588 juta, di posisi kedua ada brand **Appel** dengan nilai transaksi sebanyak ±444 juta, lanjut ada brand **Sony** menempati posisi ketiga dengan nilai transaksi sebanyak ±63,9 juta, brand **Huawei** berada di posisi keempat dengan nilai transaksi sebanyak ±63 juta Sedangkan **Lenovo** yang menduduki nilai transaksi terendah yaitu sebanyak ±62 juta

brand	total_revenue
Samsung	588764148
Apple	444855360
Sony	63960718
Huawei	63160260
Lenovo	62379800.4





THANK YOU



[linkedin.com/in/tfqhidayat](https://www.linkedin.com/in/tfqhidayat)



082251404245



taufiqhidayat1807@gmail.com



@fiq.hl

