

FINAL PROJECT

E-Commerce Sales Analysis

Presented By Faridz Salman Al Parissy

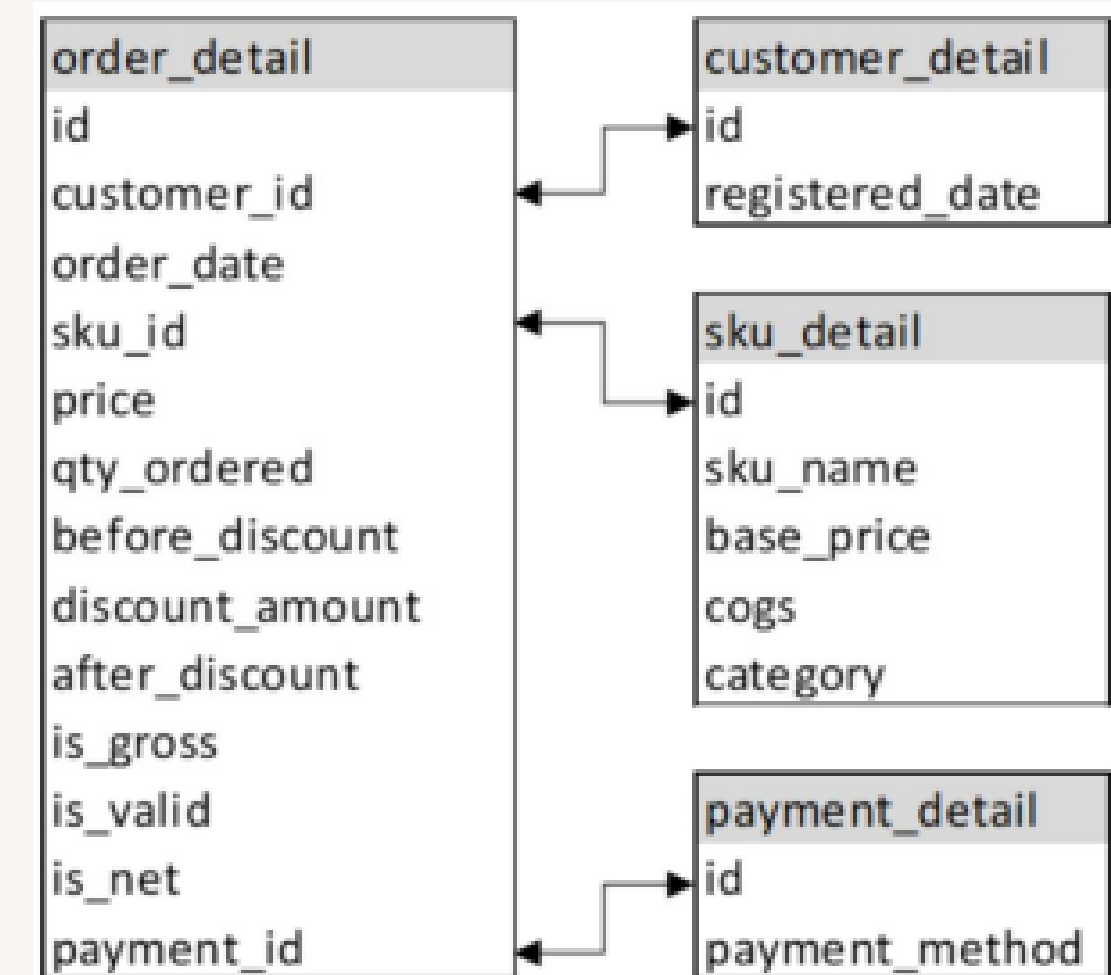
About the dataset

The dataset contains 4 tables

1. **Order_detail** table stores comprehensive information about each item within a customer's order
2. **Customer_detail** provides information about customers and when they registered as a customer.
3. **Payment_detail** table provides information about the payment method of each customer
4. **Sku_detail** table provides detailed information about each product offered, going beyond basic identification to include production cost and selling price data as well as categories.

The ecommerce dataset make up 14 relevant features

1. **Order_id** : Unique identifier for each order placed by a table
2. **Customer_id** : Identifier for each customer
3. **Order_date** : Date the order was placed by the customer
4. **Category** : Product Category
5. **Sku_name** : Name of product
6. **Price** : Price of each product
7. **Cogs** : Cost of production before marketing
8. **Qty_ordered** : Number of orders
9. **Before_discount** : (price*qty_ordered) Product price of each order
10. **Discount_amount** : Discounted price every order
11. **After_discount** : The total transaction price of each order
12. **Is_gross** : Customers who have not made payments
13. **Is_valid** : Customers who have made payments
14. **Is_net** : Customers who have completed transactions



Data Overview

All data use after_discount and is_valid = 1



Total Customer
2864



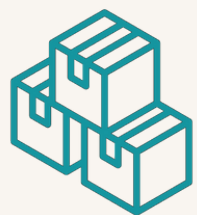
Total Revenue
Rp 3,67 Miliar



Total Orders
3955



Net Profit
Rp 820 Juta



Total Quantity
8307



Average Order Value
Rp 927,326.25

Based Query



Question/Task

01

During the transactions that occurred during 2021, in which month did the total transaction value (after_discount) be the largest?
Use is_valid = 1 to filter transaction data.

02

During transactions in 2022, which category generated the largest transaction value?
Use is_valid = 1 to filter transaction data.

03

Compare the transaction value of each category in 2021 with 2022. State which categories experienced an increase and which categories experienced a decrease in transaction value from 2021 to 2022.
Use is_valid = 1 to filter transaction data.

04

Show the top 5 most popular payment methods used in 2022 (based on total unique orders).
Use is_valid = 1 to filter transaction data.

05

Sort these 5 products by their transaction value.
1. Samsung 3. Sony 5. Lenovo
2. Apple 4. Huawei
Use is_valid = 1 to filter transaction data

01

Showing the month with the largest transactions after_discount in 2021

Logical Thinking

Create aggregation for **total_transaction_value**

Extract **order_date** to month

Filtering data **year** = 2021 and **valid** by 1

Grouping data by **month_2021_sales**

Sort data by **total_transaction_value**

Query

```
SELECT
    monthname(order_date) month_2021_sales,
    round(sum(after_discount),2) total_transaction_value
FROM order_detail
WHERE
    extract(YEAR FROM order_date) = 2021 AND
    is_valid = 1
GROUP BY month_2021_sales
ORDER BY total_transaction_value DESC;
```

Query Postgre

```
to_char(order_date, 'Month') AS month_2021_sales,
```

	A-Z month_2021_sales	123 total_transaction_value
1	August	227,862,744
2	December	217,309,963
3	October	207,603,259.88
4	November	180,396,009.7
5	July	148,007,735
6	September	145,943,335.24
7	June	43,154,552
8	January	36,822,126.5
9	February	35,611,797
10	May	34,163,856
11	March	23,643,062
12	April	22,208,472.59

Insight

The **biggest sales** in 2021 occurred in **August** IDR 227,862,744, possibly due to promos or campaigns related to Independence Day. **December** also showed high transaction numbers IDR 217,309,963, which can be attributed to Christmas and New Year, when people tend to increase their spending. In addition, **October** and **November** also recorded sizable numbers, likely due to promos such as Harbolnas.

April was the month with the **lowest transactions** IDR 22,208,472.59, which could be a concern for future sales improvement strategies.

Recommendation

- August** Maximize promotions across multiple channels, offer special discounts/bundles, and use data-driven marketing. **December** Provide early Christmas promos, year-end product discounts, and flash sales.
- April** and **March** are the **lowest** Analyze the cause and create "Spring/Mid-Year Sale" campaigns, exclusive discounts, and retargeting.
- Long-term Strategy**, Harness historical data to predict future sales trends, conduct A/B testing to identify the most effective promotional strategies, and develop membership or subscription programs to foster consistent sales throughout the year.

02

Showing the categories that generated the largest transaction value in 2022

Logical Thinking

Create aggregation for **total_transaction_2022**

Joining table **sku_detail** and **id**

Filtering data **year** = 2022 and **valid** by 1

Grouping data by **category**

Sort data by **total_transaction_2022**

Query

```
SELECT
    sd.category,
    round(sum(after_discount),2) total_transaction_2022
FROM order_detail od
    LEFT JOIN sku_detail sd
        ON od.sku_id = sd.id
WHERE
    extract(YEAR FROM order_date) = 2022 AND
    od.is_valid = 1
GROUP BY sd.category
ORDER BY total_transaction_2022 DESC;
```

	A-Z category	123 total_transaction_2022
1	Mobiles & Tablets	918,451,576
2	Entertainment	365,344,151
3	Appliances	316,358,100
4	Computing	214,028,543.38
5	Men Fashion	135,588,253
6	Women Fashion	93,014,970.62
7	Home & Living	79,483,716.2
8	Health & Sports	54,235,579.59
9	Beauty & Grooming	46,211,019.19
10	Superstore	32,643,266.53
11	Kids & Baby	25,931,276.84
12	Others	21,744,646.02
13	Soghaat	17,658,332
14	School & Education	17,362,465.3
15	Books	6,792,519.2

Insight

The **Mobiles & Tablets** category recorded the **highest total transaction** of IDR 918,451,576, indicating that electronic products, especially mobile phones and tablets, have very high demand compared to other categories. This is likely influenced by technology trends, communication needs, and the increasing use of mobile devices in various aspects of life, such as work and entertainment.

The category with the **lowest transactions** was **Books**, which only recorded IDR 6,792,519.2, far below the other categories. This could be due to the shift in consumer interest to digital media such as e-books or audiobooks, or the lack of attractive promotions and campaigns for this category.

Recommendation

1. The **Mobiles & Tablets** category to maintain or improve marketing strategies, such as exclusive promotions or bundling with related accessories. **Entertainment** and **Appliances** also have high value. Upselling and cross-selling strategies can be explored.
2. The **Books** and **School & Education** , Specific promotions, such as student discounts or digital education campaigns, evaluate market demand for this category, product diversification, such as e-books or online courses.
3. **Long-term Strategy**, segment your customers to see shopping patterns, whether these trends are consistent every year or whether there are changes in consumption patterns. Use historical data and customer feedback to understand why transactions are low in certain categories.

03

Shows a comparison of transaction values for each category from 2021 and 2022, as well as categories that experienced increases and decreases.

Logical Thinking

Create aggregation for **total_transaction_value**, **transaction_growth**, and **growth_percentage**

Extracting **order_date** to year

Make CTE **yearly_transaction** and **growth** for virtual table that exists only for the duration of the query. **It helps for breaking down** into logical blocs, improving readability, and maintainability.

Joining table **sku_detail**, **order_detail**, and **Self-join** for comparison between 2021 and 2022

Filtering data **year** = 2021, 2022 and **valid** by 1

Grouping data by **category** and **year**

Sort data by **transaction_growth**

y0 is prior year (2021) and **y1** is curret year (2022)

Query

```
WITH
  yearly_transaction AS (
    SELECT
      sd.category,
      extract(YEAR FROM order_date) tahun,
      round(sum(od.after_discount),2) total_transaction_value
    FROM order_detail od
    LEFT JOIN sku_detail sd
      ON od.sku_id = sd.id
    WHERE
      od.is_valid = 1 AND
      extract(YEAR FROM order_date) IN (2021,2022)
    GROUP BY
      sd.category,
      tahun
  ),
  growth AS (
    SELECT
      y1.category,
      y0.total_transaction_value transaction_2021,
      y1.total_transaction_value transaction_2022,
      round((y1.total_transaction_value - y0.total_transaction_value),2) transaction_growth,
      round(((y1.total_transaction_value -
      y0.total_transaction_value)/y0.total_transaction_value)*100,2) growth_percentage
    FROM yearly_transaction y1
    LEFT JOIN yearly_transaction y0
      ON y1.category = y0.category
      AND y1.tahun = y0.tahun + 1
    WHERE
      y0.total_transaction_value IS NOT NULL AND
      y1.total_transaction_value IS NOT NULL
  )
SELECT *
FROM growth
ORDER BY transaction_growth DESC;
```


	A-Z category	123 transaction_2021	123 transaction_2022	123 transaction_growth	123 growth_percentage
1	Mobiles & Tablets	370,606,718	918,451,576	547,844,858	147.82
2	Entertainment	162,326,357.38	365,344,151	203,017,793.62	125.07
3	Appliances	218,550,177	316,358,100	97,807,923	44.75
4	Men Fashion	58,628,198	135,588,253	76,960,055	131.27
5	Computing	172,878,860	214,028,543.38	41,149,683.38	23.8
6	Home & Living	45,797,873	79,483,716.2	33,685,843.2	73.55
7	Health & Sports	33,837,965.6	54,235,579.59	20,397,613.99	60.28
8	Women Fashion	84,045,961.4	93,014,970.62	8,969,009.22	10.67
9	School & Education	11,558,982.39	17,362,465.3	5,803,482.91	50.21
10	Superstore	28,828,088	32,643,266.53	3,815,178.53	13.23
11	Soghaat	15,056,202.6	17,658,332	2,602,129.4	17.28
12	Kids & Baby	23,971,057.8	25,931,276.84	1,960,219.04	8.18
13	Beauty & Grooming	46,047,360	46,211,019.19	163,659.19	0.36
14	Books	10,124,596	6,792,519.2	-3,332,076.8	-32.91
15	Others	40,468,515.75	21,744,646.02	-18,723,869.73	-46.27

Insight

The **Mobile & Tablets** category saw the most significant **transaction growth**, a remarkable **147.82%** increase from **IDR 370,606,718** in 2021 to **IDR 918,451,576** in 2022. This substantial growth suggests rising demand for electronic products, potentially fueled by technological advancements, the increasing prevalence of digital devices, the continued rise of online shopping, and trends in gaming (game consoles) and entertainment (streaming subscriptions).

The **Others** category saw a significant **-46.27%**, dropping from **IDR 40,468,515.75** in 2021 to **IDR 21,744,646.02** in 2022. This substantial decrease suggests that some product categories within this group may be losing traction or are being impacted by shifting consumer shopping habits. Notably, the **Books** experienced the steepest drop, falling **-32.91%** from **IDR 10,124,396** in 2021 to **IDR 6,792,519.2** in 2022. This decline likely reflects a preference shift towards e-books or audiobooks, coupled with potentially insufficient promotion or appeal in the physical book category.

Recommendation

- Maximizing High-Growth Categories** such as "Mobiles & Tablets", "Entertainment", and "Men's Fashion" have **significant growth**, so companies must continue to maintain momentum by adding product variety, increasing exclusive promotions, and optimizing digital marketing strategies.
- Revitalizing Declining Categories** such as "Books" and "Others" have **significant decline**, companies need to explore new strategies, such as product innovation, collaboration with communities, and more engaging event-based marketing campaigns to re-energize customers.
- Improving the Use of Data and Market Trends**, Analyzing customer purchasing patterns and market trends can help in developing more effective business strategies. Having analyzed customer purchasing patterns and market trends, the company has developed more effective business strategies. This analysis allows for the design of targeted promotions and adjustments to product stock based on customer demand.

04

Show the top 5 most popular payment methods used in 2022

Logical Thinking

Create aggregation unique **order_id**

Joining table **payment_detail** and **order_detail**

Filtering data **year** = 2022 and **valid** by 1

Grouping data by **payment_method**

Sort data by most **payment_2022**

Limiting data by 5

Query

```
SELECT
    pd.payment_method,
    count(DISTINCT od.id) payment_2022
FROM order_detail od
LEFT JOIN payment_detail pd
    ON od.payment_id = pd.id
WHERE
    extract(YEAR FROM order_date) = 2022 AND
    od.is_valid = 1
GROUP BY pd.payment_method
ORDER BY payment_2022 DESC
LIMIT 5;
```

	A-Z payment_method	123 payment_2022
1	cod	1,809
2	Payaxis	181
3	customercredit	75
4	Easypay	69
5	jazzwallet	26

Insight

Cash on Delivery (COD) remains the dominant payment method with 1,809 transactions, significantly outpacing all other options. Several key factors contribute to its continued popularity. Trust is a major driver, as many customers feel more secure paying only after receiving their goods, ensuring both quality and authenticity. Limited access to banking services and digital wallets, particularly in certain regions, makes digital payments less practical for a segment of the population. Established habits and a preference for cash transactions also play a role. Furthermore, concerns about data security and perceived complexities in digital payment refund processes lead some customers to avoid them altogether. While **Payaxis** is the most popular digital method, its transaction volume is considerably **lower than COD**. Other digital payment options, including **Customer Credit**, **Easypay**, and **Jazzwallet** very low transaction numbers.

Recommendation

1. **Encourage digital payment** adoption by offering incentives such as discounts or cashback for non-COD transactions.
2. **Increase customer education** on the security and convenience of digital payments to reduce reliance on COD.

05

Showing brands with the highest transaction value between Samsung, Apple, Sony, Huawei, and Lenovo

Logical Thinking

Identification brand with letter

Joining table **sku_detail** and **order_detail**

Filtering data **valid** by 1 and not null

Grouping data by **product_brand**

Sort data by **total_transaction**

Query

```
SELECT
CASE
    WHEN LOWER(sd.sku_name) LIKE '%samsung%' THEN 'Samsung'
    WHEN LOWER(sd.sku_name) LIKE '%apple%' OR
        LOWER(sd.sku_name) LIKE '%macbook%' OR
        LOWER(sd.sku_name) LIKE '%iphone%' THEN 'Apple'
    WHEN LOWER(sd.sku_name) LIKE '%sony%' THEN 'Sony'
    WHEN LOWER(sd.sku_name) LIKE '%huawei%' THEN 'Huawei'
    WHEN LOWER(sd.sku_name) LIKE '%lenovo%' THEN 'Lenovo'
    ELSE NULL
END AS product_brand,
ROUND(SUM(od.after_discount), 2) AS total_transaction
FROM order_detail od
LEFT JOIN sku_detail sd
    ON od.sku_id = sd.id
WHERE is_valid = 1
GROUP BY product_brand
HAVING product_brand IS NOT NULL
ORDER BY 2 DESC;
```

	A-Z product_brand	123 total_transaction
1	Samsung	588,764,150
2	Apple	444,855,360
3	Sony	63,960,718
4	Huawei	63,160,260
5	Lenovo	62,379,800.38

Insight

Samsung dominates the market with the **highest total** transactions of IDR 588,764,150, far ahead of **Apple** in second place with IDR 403,443,360. Meanwhile, **Sony, Huawei,** and **Lenovo** have relatively **small** and almost equal total transactions, indicating that they compete in a smaller segment. The significant difference between Samsung and Apple compared to other brands shows that the market is dominated by two major brand.

Any Further Discussion?

Let's Connect !

Faridz Salman Al Parissy

Data Analyst, Data Visualization, Front-End Developer



+62 813-8828-2752



faridzsalman09@gmail.com



Faridz Salman Al Parissy



Explore more project **here!**