Amazon Sales Data Analysis Project

amazon

Presented by: Rajvardhan Singh Parmar

# Agenda

- About The Company

- Project Overview

- Business Challenge

- AD-HOC Requests

- Conclusion

Amazon is a global e-commerce giant, primarily known for its vast online marketplace. In terms of sales, Amazon consistently ranks as one of the world's largest retailers.

The company generates significant revenue through its diverse product offerings, including electronics, apparel, and cloud services (AWS). Amazon's sales strategy is powered by a customer-centric approach, competitive pricing, and fast delivery services, contributing to its dominant market position.
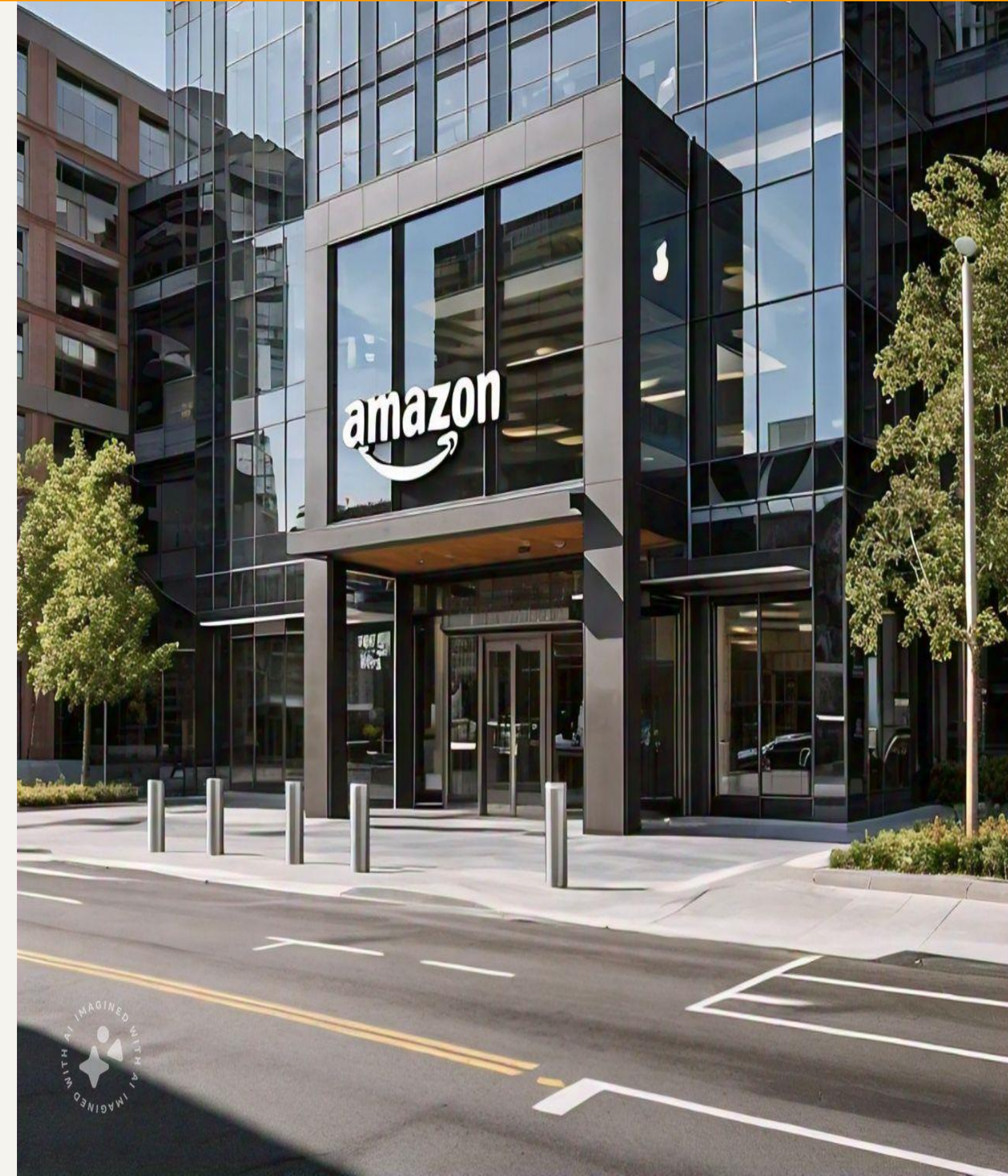
Project Overview

This project is dedicated to a comprehensive analysis of sales data from Amazon's branches in Mandalay, Yangon, and Naypyitaw. By scrutinizing 1,000 sales transactions, we will explore various product lines and customer demographics to identify trends and factors that influence sales performance. Our goal is to uncover actionable insights that can help optimize sales strategies, improve customer targeting, and enhance overall performance across these key locations.

Business Challenge

Amazon is seeking to revolutionize its sales strategies and customer targeting across its key branches in Mandalay, Yangon, and Naypyitaw. The objective is to delve deep into sales data to identify which product lines and customer segments are yielding the highest profitability. By thoroughly analyzing sales trends and uncovering insightful patterns, Amazon aims to develop more refined and effective business strategies. This effort will help tailor marketing approaches, optimize inventory management, and enhance overall decision-making, driving significant growth and improved customer satisfaction in these markets.

AD-HOC Requests

1) How much revenue is generated each month?

```sql
SELECT
    MONTHNAME(Date) AS Month_name,
    ROUND(SUM(Unit_price * Quantity), 2) AS Revenue
FROM
    sales_data
GROUP BY
    Month_name
ORDER BY
    Revenue DESC;
```

Rajvardhan Singh Parmar

OUTPUT

Month_name	Revenue
January	110754.16
March	104243.34
February	92589.88

2) In which month did the cost of goods sold reach its peak?

```sql
SELECT
    MONTHNAME(Date) AS Month_name,
    ROUND(SUM(COGS), 2) AS Maximum_COGS
FROM
    sales_data
GROUP BY
    Month_name
ORDER BY
    Maximum_COGS DESC
LIMIT 1;
```

Rajvardhan Singh Parmar

OUTPUT

Month_name	Maximum_COGS
January	110754.16

3) Determine the city with the highest VAT percentage.

```sql
SELECT
    City,
    ROUND(SUM(Tax) / SUM(Total)) * 100,2)
    AS vat_percentage
FROM
    sales_data
GROUP BY
    City
ORDER BY
    vat_percentage DESC
LIMIT
    1;
```

in Rajvardhan Singh Parmar

OUTPUT

City	vat_percentage
Yangon	4.76

4) For each product line, add a column indicating "Good" if its sales are above average, otherwise "Bad".

```sql
WITH cte1 AS (
    SELECT
        Product_line,
        ROUND(SUM(total), 2) AS Total_sales
    FROM
        sales_data
    GROUP BY
        Product_line
),
cte2 AS (
    SELECT
        ROUND(AVG(Total_sales), 2) AS AVG_sales
    FROM
        cte1
)
SELECT Product_line, Total_sales, AVG_sales,
CASE
    WHEN cte1.Total_sales > cte2.AVG_sales THEN 'GOOD'
    ELSE 'BAD'
    END AS Indicating
FROM
    cte1, cte2;
```

Rajvardhan Singh Parmar

OUTPUT

product_line	Total_sales	AVG_sales	Indicating
Food and beverages	56144.84	53827.79	GOOD
Health and beauty	49193.74	53827.79	BAD
Sports and travel	55122.83	53827.79	GOOD
Fashion accessories	54305.9	53827.79	GOOD
Home and lifestyle	53861.91	53827.79	GOOD
Electronic accessories	54337.53	53827.79	GOOD

5) Which product line is most frequently associated with each gender?

```sql
WITH cte1 AS (
    SELECT
        Product_line,
        Gender,
        COUNT(Gender) AS NO_of_count
    FROM sales_data
    GROUP BY Product_line, Gender
),
cte2 AS (
    SELECT
        MAX(NO_of_count) AS Max_count
    FROM cte1
)
SELECT
    cte1.Product_line,
    cte1.Gender,
    cte1.NO_of_count
FROM
    cte1, cte2
ORDER BY
    cte1.Product_line ASC,
    cte1.NO_of_count DESC;
```

Rajvardhan Singh Parmar

OUTPUT

Product_line	Gender	NO_of_count
Electronic accessories	Male	86
Electronic accessories	Female	84
Fashion accessories	Female	96
Fashion accessories	Male	82
Food and beverages	Female	90
Food and beverages	Male	84
Health and beauty	Male	88
Health and beauty	Female	64
Home and lifestyle	Male	81
Home and lifestyle	Female	79
Sports and travel	Female	88
Sports and travel	Male	78

6) Identify the customer type with the highest purchase frequency.

```sql
SELECT
    customer_type,
    COUNT(Invoice_ID) AS Purchase_frequency
FROM
    sales_data
GROUP BY
    customer_type
ORDER BY
    Purchase_frequency DESC
LIMIT
    1;
```

Rajvardhan Singh Parmar

OUTPUT

customer_type	Purchase_frequency
Member	501

7) Examine the distribution of genders within each branch.

```sql
SELECT
    Branch,
    Gender,
    COUNT(Gender) AS Gender_count
FROM
    sales_data
GROUP BY
    Branch,
    Gender
ORDER BY
    Branch,
    Gender_count DESC;
```

Rajvardhan Singh Parmar

OUTPUT

Branch	Gender	Gender_count
A	Male	179
A	Female	161
B	Male	170
B	Female	162
C	Female	178
C	Male	150

8) Determine the day of the week with the highest average ratings for each branch.

```sql
WITH cte1 AS (
    SELECT
        Branch,
        DAYNAME(Date) AS Day_name,
        ROUND(AVG(Rating), 2) AS Avg_rating
    FROM
        sales_data
    GROUP BY
        Branch, Day_name
),
cte2 AS (
    SELECT
        Branch,
        Day_name,
        Avg_rating,
        DENSE_RANK() OVER (
            PARTITION BY Branch
            ORDER BY Avg_rating DESC
        ) AS rn
    FROM
        cte1
)
SELECT
    Branch,
    Day_name,
    Avg_rating
FROM
    cte2
WHERE
    rn = 1;
```

Rajvardhan Singh Parmar

OUTPUT

Branch	Day_name	Avg_rating
A	Friday	7.31
B	Monday	7.34
C	Friday	7.28

9) Identify the branch that exceeded the average number of products sold.

```sql
WITH branch_sales AS (
    SELECT
        Branch,
        SUM(Quantity) AS Total_sales
    FROM
        sales_data
    GROUP BY
        Branch
),
branch_avg AS (
    SELECT
        AVG(Total_sales) AS Avg_sales
    FROM
        branch_sales
)
SELECT
    branch_sales.Branch,
    branch_avg.Avg_sales
FROM
    branch_sales
CROSS JOIN
    branch_avg
WHERE
    branch_sales.Total_sales > branch_avg.Avg_sales;
```

Rajvardhan Singh Parmar

OUTPUT

Branch	Avg_sales
A	1836.6667

10) Identify the customer type contributing the highest revenue.

```sql
WITH cte1 AS (
    SELECT
        Customer_type,
        (Unit_price * Quantity) AS Revenue
    FROM
        sales_data
),
cte2 AS (
    SELECT
        Customer_type,
        ROUND(SUM(Revenue), 2) AS Total_Revenue
    FROM
        cte1
    GROUP BY
        Customer_type
)
SELECT
    Customer_type,
    Total_Revenue
FROM
    cte2
ORDER BY
    Total_Revenue DESC
LIMIT 1;
```

Rajvardhan Singh Parmar

OUTPUT

Customer_type	Total_Revenue
Member	156403.28

Conclusion

In this analysis of Amazon's sales data from branches in Mandalay, Yangon, and Naypyitaw has provided valuable insights into the company's sales performance. By examining 1,000 transactions across various product lines and customer demographics, we identified key trends that can inform future business strategies. The findings emphasize the importance of targeted marketing, efficient inventory management, and a customer-centric approach to driving growth and enhancing customer satisfaction. As Amazon continues to innovate, these insights will be instrumental in refining strategies that align with both market demands and the company's goals.

Thank you!

Feel free to approach me
if you have any questions.