SQL

Data Analysis Project

- Sujal Parmar | Data Analyst

Introduction:

This project aims to analyze sales data from three Amazon branches located in Mandalay, Yangon, and Naypyitaw. By examining 1,000 sales transactions across various product lines and customer types, we seek to understand what factors influence sales performance.

Business Problem:

Amazon wants to enhance its sales strategies and customer targeting across its branches in Mandalay, Yangon, and Naypyitaw. The challenge is to identify which product lines and customer segments are most profitable, understand sales trends, and uncover patterns that can drive more effective business decisions.

Goal:

The goal is to analyze sales data to identify top-performing product lines, understand customer behavior, and reveal trends and patterns that can help improve sales strategies and decision-making across Amazon’s branches.

Question- 1

```sql
-- Which product line has the highest sales?
SELECT
    product_line, SUM(total) AS total_sales
FROM
    amazon
GROUP BY product_line
ORDER BY total_sales DESC
LIMIT 1;
```

	product_line	total_sales
▶	Food and beverages	56144.844000000005

Insights:

- The **"Food and Beverages"** product line generates the highest sales, indicating strong customer demand.
- This category significantly impacts overall revenue.

Recommendations:

- Focus on targeted **promotions and seasonal offers** for "Food and Beverages" to boost sales further.
- Maintain adequate stock levels to meet high demand and avoid stockouts.
- Consider adding new products to the "Food and Beverages" line to attract more customers.

Question- 2

```sql
-- How much revenue is generated each month?
SELECT
    monthname, SUM(total)
FROM
    amazon
GROUP BY 1;
```

	monthname	total_revenue
▶	January	116291.86800000005
	March	109455.50700000004
	February	97219.37399999997

Insights:

- **January** generated the highest revenue, followed by March and February. This indicates that sales peak in January, possibly due to seasonal factors such as holiday shopping.

Recommendations:

- Increase **marketing** efforts and **promotions** before and during January to capitalize on higher customer spending.

Question- 3

```sql
-- Which product line generated the highest revenue?
SELECT
    Product_line, SUM(total) AS total_revenue
FROM
    amazon
GROUP BY Product_line
ORDER BY total_revenue DESC
LIMIT 1;
```

Product_line	total_revenue
Food and beverages	56144.844000000005

Question- 4

```sql
-- In which city was the highest revenue recorded?
SELECT
    city, SUM(total) AS total_revenue
FROM
    amazon
GROUP BY 1
ORDER BY 2 DESC
LIMIT 1;
```

city	total_revenue
Naypyitaw	110568.70649999994

Insights:

- The "Food and Beverages" product line generated the highest revenue.
- **Naypyitaw** recorded the highest revenue. This indicates that Naypyitaw is the most profitable city among the branches.

Recommendations:

- **Allocate** more **resources** and marketing efforts to **Naypyitaw** to sustain and further boost its revenue.
- Analyze successful strategies and practices in Naypyitaw and consider applying them to other branches to improve their performance.

Question- 5

```sql
-- Count the sales occurrences for each time of day on every weekday.
SELECT
    CASE
        WHEN HOUR(Time) >= 6 AND HOUR(Time) < 12 THEN 'Morning'
        WHEN HOUR(Time) >= 12 AND HOUR(Time) < 18 THEN 'Afternoon'
        ELSE 'Evening'
    END AS Time_of_Day,
    DAYNAME(Date) AS Weekday,
    COUNT(*) AS Sales_Occurrences
FROM
    amazon
GROUP BY
    Time_of_Day, Weekday
ORDER BY
    Weekday, FIELD(Time_of_Day, 'Morning', 'Afternoon', 'Evening');
```

Time_of_Day	Weekday	Sales_Occurrences
Morning	Friday	29
Afternoon	Friday	74
Evening	Friday	36
Morning	Monday	21
Afternoon	Monday	75
Evening	Monday	29
Morning	Saturday	28
Afternoon	Saturday	81
Evening	Saturday	55
Morning	Sunday	22
Afternoon	Sunday	70
Evening	Sunday	41
Morning	Thursday	33
Afternoon	Thursday	76
Evening	Thursday	29
Morning	Tuesday	36
Afternoon	Tuesday	71
Evening	Tuesday	51
Morning	Wednes...	22
Afternoon	Wednes...	81
Evening	Wednes...	40

Insights:

- The "**Afternoon**" has the highest number of sales occurrences across most weekdays. This suggests that customers are more active in the afternoon.

Recommendations:

- Schedule more **staff** during the afternoon hours to **handle** the **higher volume** of sales and improve customer service during peak times.

Question- 6

```sql
--Identify the customer type contributing the highest revenue.
SELECT
    Customer_type, SUM(Total) AS Total_Revenue
FROM
    amazon
GROUP BY Customer_type
ORDER BY Total_Revenue DESC
LIMIT 1;
```

	Customer_type	Total_Revenue
▶	Member	164223.44400000002

Question- 7

```sql
-- Which customer type occurs most frequently?
SELECT
    Customer_type, COUNT(*) AS Frequency
FROM
    amazon
GROUP BY Customer_type
ORDER BY Frequency DESC
LIMIT 1;
```

	Customer_type	Frequency
▶	Member	501

Insights:

- The "**Member**" customer type generates the highest revenue. This indicates that members are the most valuable customer segment in terms of revenue contribution.
- A large portion of the **customer base consists of members,** suggesting that membership offers are appealing and widely adopted.

Recommendations:

- Strengthen and expand benefits for members to further encourage their loyalty and increase their spending.
- **Enhance membership** benefits and loyalty programs to maintain and further grow this key customer segment.

Question- 8

```sql
-- Determine the city with the highest VAT percentage.
SELECT
    City, (SUM(VAT) / SUM(Total)) * 100 AS VAT_Percentage
FROM
    amazon
GROUP BY City
ORDER BY VAT_Percentage DESC
LIMIT 1;
```

City	VAT_Percentage
Mandalay	4.761904761904766

Insights:

- **Mandalay has the highest VAT percentage. This indicates that Mandalay's sales incur the highest tax rate compared to other cities.**

Recommendations:

- **Consider adjusting pricing or promotional strategies in Mandalay to account for the higher VAT and maintain competitiveness.**

Question- 9

```sql
-- Identify the day of the week with the highest average ratings.
SELECT
    DAYNAME(Date) AS Day_of_Week, AVG(Rating) AS Average_Rating
FROM
    amazon
WHERE
    Rating IS NOT NULL
GROUP BY Day_of_Week
ORDER BY Average_Rating DESC
LIMIT 1;
```

	Day_of_Week	Average_Rating
▶	Monday	7.153599999999999

Insights:

- Monday has the highest average customer rating at 7.15. The high ratings on Mondays suggest that customers are generally more satisfied with their shopping experiences at the start of the week.

Recommendations:

- Use Monday's positive ratings to promote special offers or campaigns, potentially boosting sales and engagement on this day.
- Analyze what contributes to higher ratings on Mondays and implement similar practices or enhancements throughout the week to improve overall customer satisfaction.

Thank You!