ANALYZING SALES PERFORMANCE

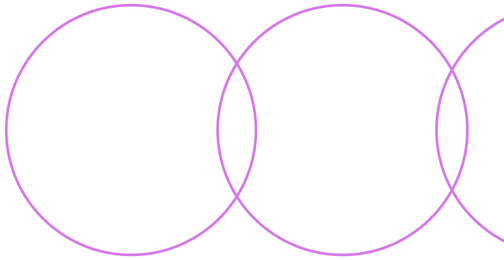
Amazon Sales Analysis

Devaraju Subramanyam

Amazon Sales Analysis Overview

Key Objectives and Presentation Agenda

This presentation provides a detailed analysis of Amazon's sales performance, focusing on factors influencing sales across branches. We'll explore key metrics and data insights, examining trends and offering sales strategy recommendations. The agenda covers an introduction, data insights, visual representations, and tailored recommendations for recruiters.



Amazon Sales Insights

Analyzing key metrics to understand sales performance and growth potential.

Product Analysis

To analyze data and understand product lines, start by examining sales volume, a key metric indicating performance. A high volume shows strong demand, while low volumes suggest a need for strategy improvements. Identifying trends helps refine marketing and optimize inventory, boosting underperforming lines.

Sales Analysis

This analysis examines sales trends to evaluate the effectiveness of business strategies. Insights will suggest modifications to enhance sales. Monitoring revenue is key to understanding financial performance and identifying growth opportunities. It also reveals how pricing and promotions affect profitability.

Customer Analysis

This analysis identifies customer segments and examines their purchasing patterns.
Understanding these elements helps tailor strategies to boost engagement.
Insights into trends allow optimization of product offerings.
This approach supports informed decisions and drives growth.

Database Overview

Column	Description	Data Type
invoice_id	Invoice of the sales made	VARCHAR(30)
branch	Branch at which sales were made	VARCHAR(5)
city	The location of the branch	VARCHAR(30)
customer_type	The type of the customer	VARCHAR(30)
gender	Gender of the customer making purchase	VARCHAR(10)
product_line	Product line of the product sold	VARCHAR(100)
unit_price	The price of each product	DECIMAL(10, 2)
quantity	The amount of the product sold	INT
VAT	The amount of tax on the purchase	FLOAT(6, 4)
total	The total cost of the purchase	DECIMAL(10, 2)
date	The date on which the purchase was made	DATE
time	The time at which the purchase was made	TIMESTAMP
payment_method	The total amount paid	DECIMAL(10, 2)
cogs	Cost Of Goods sold	DECIMAL(10, 2)
gross_margin_percentage	Gross margin percentage	FLOAT(11, 9)
gross_income	Gross Income	DECIMAL(10, 2)
rating	Rating	FLOAT(2, 1)

The Dataset has 1000 rows and 17 columns

Project Overview

Data Wrangling

Feature Engineering

Exploratory Data Analysis

Data Wrangling

Create Database and Table:
Start by creating a database and defining table structure using a CREATE TABLE statement with column names, data types, and NOT NULL constraints to ensure data quality.

Insert Data :
After creating the table, you can fill it with your dataset, such as by importing data from a CSV file. During this process, the database management system will automatically enforce the NOT NULL constraints, rejecting any rows that contain missing values in the specified columns.

Data integrity:
verify null values even with the NOT NULL constraint in place. Run a SQL query using IS NULL on any column. If set up correctly, a query like `SELECT * FROM your_table WHERE any_column IS NULL;` should return zero rows, indicating no nulls in the dataset and readiness for analysis.

```sql
CREATE TABLE amazon (
    invoice_id              VARCHAR(30)     NOT NULL,
    branch                  VARCHAR(5)      NOT NULL,
    city                    VARCHAR(30)     NOT NULL,
    customer_type           VARCHAR(30)     NOT NULL,
    gender                  VARCHAR(10)     NOT NULL,
    product_line            VARCHAR(100)    NOT NULL,
    unit_price              DECIMAL(10,2)   NOT NULL,
    quantity                INT             NOT NULL,
    vat                     FLOAT(6,4)      NOT NULL,
    total                   DECIMAL(10,2)   NOT NULL,
    date                    DATE            NOT NULL,
    time                    TIME            NOT NULL,
    payment_method          VARCHAR(20)     NOT NULL,
    cogs                    DECIMAL(10,2)   NOT NULL,
    gross_margin_percentage FLOAT(11,9)     NOT NULL,
    gross_income            DECIMAL(10,2)   NOT NULL,
    rating                  FLOAT(2,1),

    PRIMARY KEY (invoice_id)
);
```

Feature Engineering

Dayname

Extracted the day of the week (e.g., Monday, Tuesday) to identify which days have the highest sales activity.

Monthname

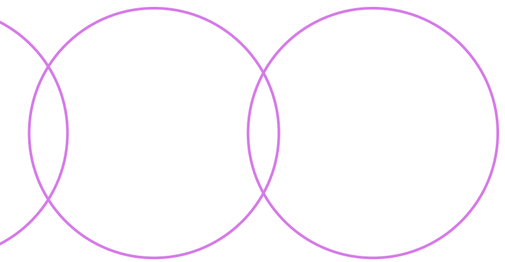
Extracted the month (e.g., January, February) to analyze monthly sales trends and performance.

Safe Update

Safe update mode is a safety feature that protects you from accidentally modifying or deleting a large number of records.

```sql
ALTER TABLE amazon
ADD COLUMN timeofday VARCHAR(10),
ADD COLUMN dayname VARCHAR(20),
ADD COLUMN MonthName VARCHAR(20);

-- Step 1: Disable safe update mode
SET SQL_SAFE_UPDATES = 0;

-- Step 2: Run your UPDATE statement
UPDATE amazon
SET
  timeofday = CASE
    WHEN Time BETWEEN '00:00:00' AND '11:59:59' THEN 'Morning'
    WHEN Time BETWEEN '12:00:00' AND '17:59:59' THEN 'Afternoon'
    ELSE 'Night'
  END,
  dayname = DAYNAME(Date),
  MonthName = MONTHNAME(Date);

-- Step 3 (Important!): Re-enable safe update mode
SET SQL_SAFE_UPDATES = 1;
```

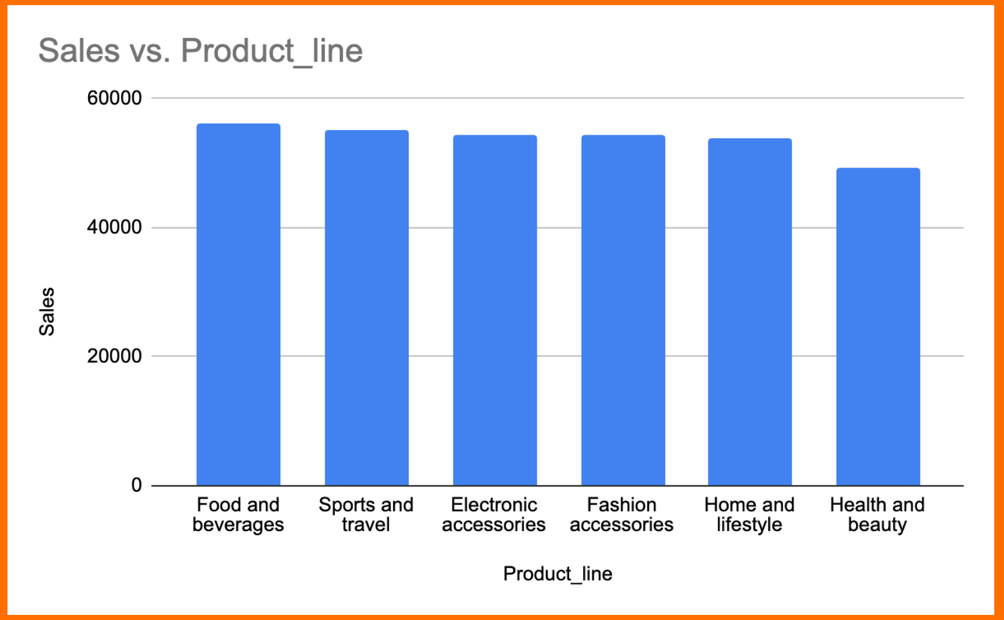
Exploratory Data Analysis



Product Analysis

```sql
-- Which product line has the highest sales?

Select Product_line,Round(sum(Total),2) as Sales
from amazon
group by Product_line
Order by sum(Total) desc;
```

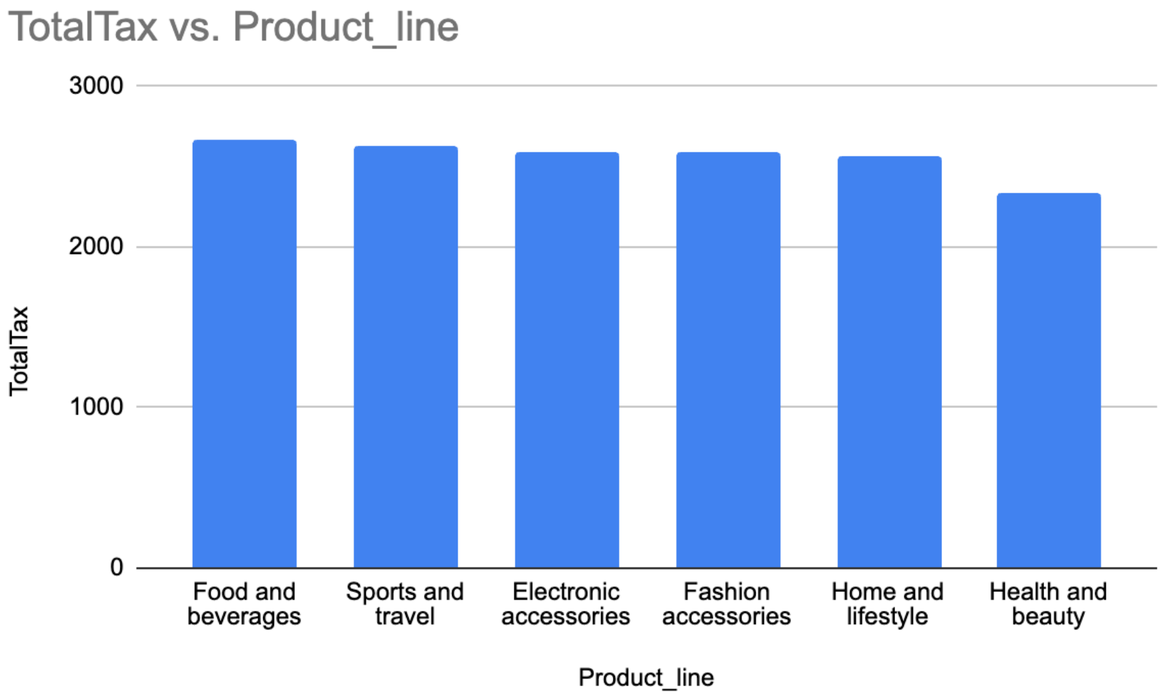
Food & berverages has highest Sales among all other ProductLines followed by sports & Travel



Product Analysis

```sql
-- Which product line incurred the highest Value Added Tax?
Select Product_line,sum(Tax5)
from amazon
group by Product_line
Order by sum(Tax5) desc
```

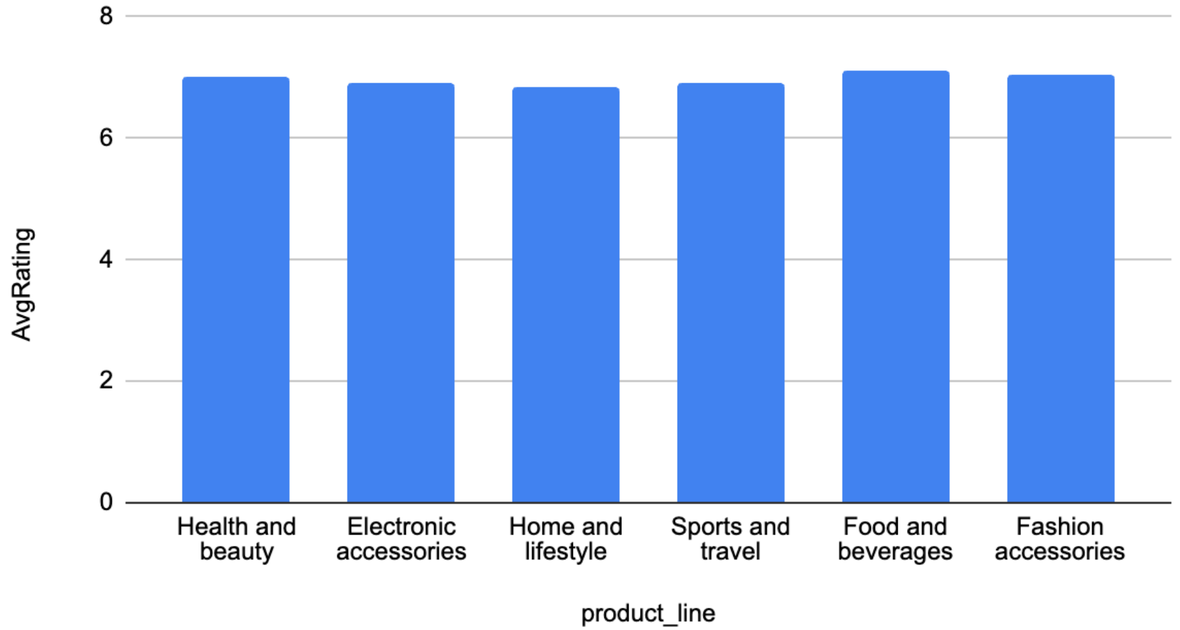
Food & berverrages has highest value added Tax folllowed by Sports & Beverages



Product Analysis

```sql
-- Calculate the average rating for each product line.
Select  product_line,Round(avg(rating),2) as AvgRating
from amazon
group by product_line;
```

Food & Berverages has highest Avg Rating followed by Fashion accessories



Product Analysis

```sql
WITH cte AS (
    SELECT
        Product_line,
        SUM(Total) AS TotalSales
    FROM amazon
    GROUP BY Product_line
),
overall AS (
    SELECT AVG(TotalSales) AS avgSales
    FROM cte
)
SELECT
    cte.Product_line,
    CASE
        WHEN cte.TotalSales > overall.avgSales THEN 'Good'
        ELSE 'Bad'
    END AS SalesCategory
FROM cte, overall;
```

Product_line	SalesCategory
Health and beauty	Bad
Electronic accessories	Good
Home and lifestyle	Good
Sports and travel	Good
Food and beverages	Good
Fashion accessories	Good

Sales Analysis

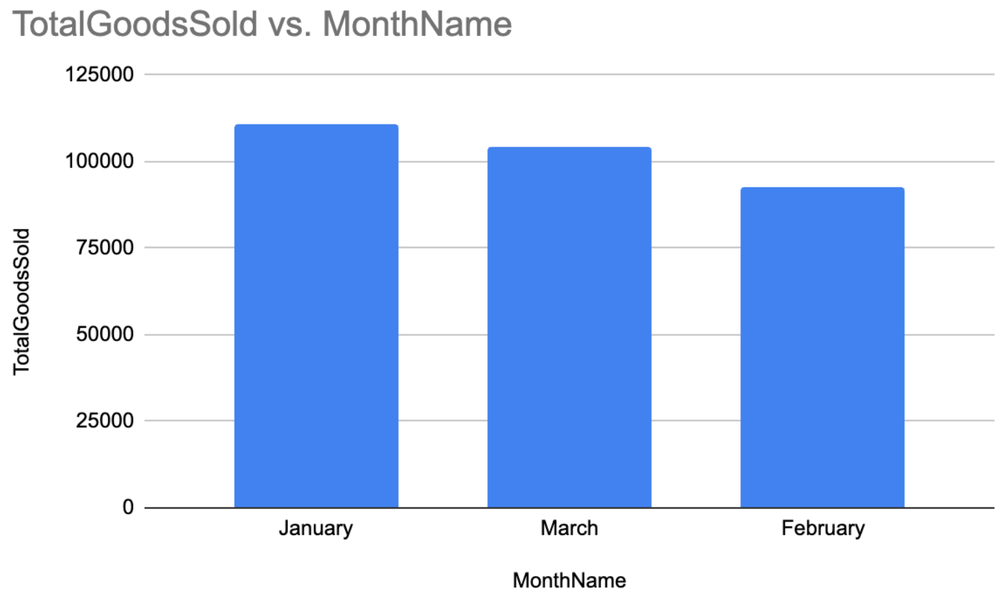
```sql
-- In which city was the highest revenue recorded?

Select city,Round(sum(cogs),2) as Revenue
from amazon
group by city
Order by sum(cogs) desc
limit 1;
```

Naypyitaw has highest revuenue among other cities

city	Revenue
Naypyitaw	105303.53
Yangon	101143.21
Mandalay	101140.64

Sales Analysis

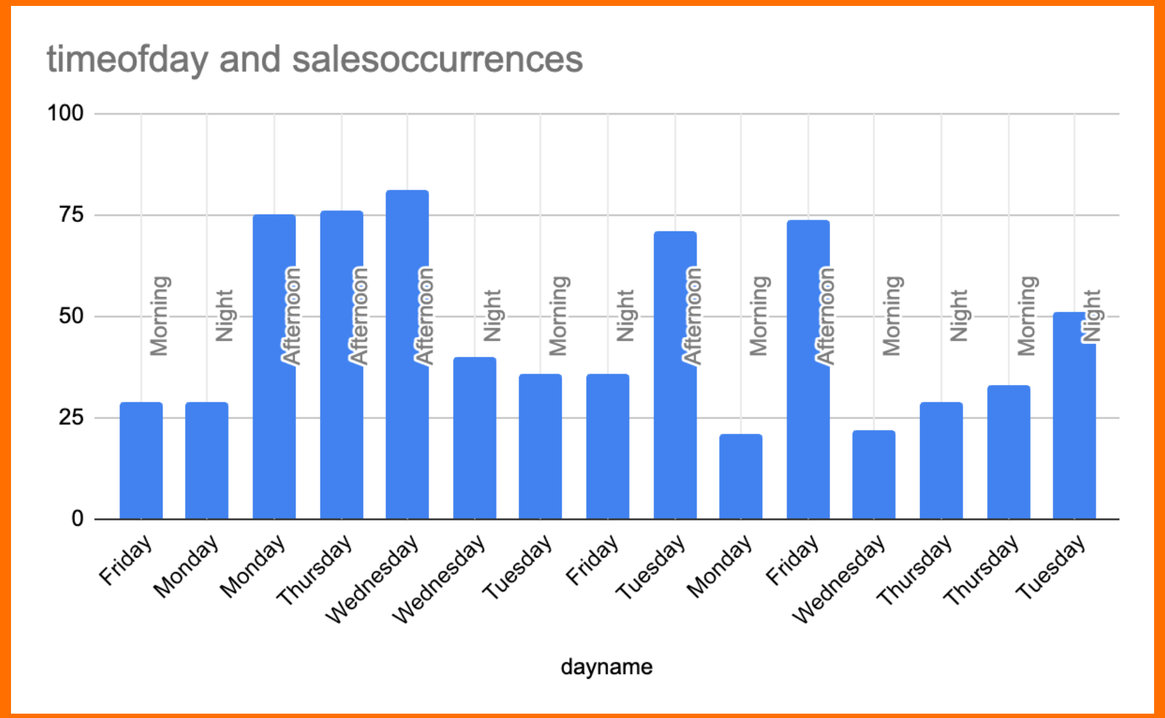
```sql
-- In which month did the cost of goods sold reach its peak?

Select MonthName ,Round(Sum(cogs),2) as TotalGoodsSold
from amazon
group by MonthName
Order by Sum(cogs) DESC
Limit 1;
```



Sales Analysis

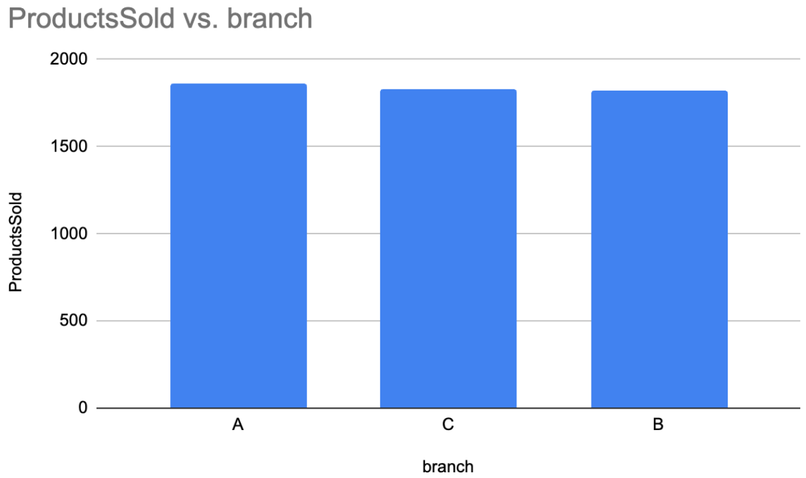
```sql
-- Count the sales occurrences for each time of day on every weekday.
Select dayname,timeofday,count(*) as salesoccurrences
from amazon
group by dayname,timeofday
having dayname NOT IN ('Saturday','Sunday');
```

It is observed the Sales are more in afternoon especially on Monday,Thursday & Wednesday . While Wednesday has highest Sales followed by Thursday & Monday



Sales Analysis

```sql
With cte as (
Select branch,Sum(quantity) as ProductsSold
from amazon
group by branch)
,OverallAvg as (
Select Avg(quantity) as AvgProdsSold
from amazon)

Select cte.branch,cte.ProductsSold
from cte,OverallAvg
where ProductsSold > AvgProdsSold;
```

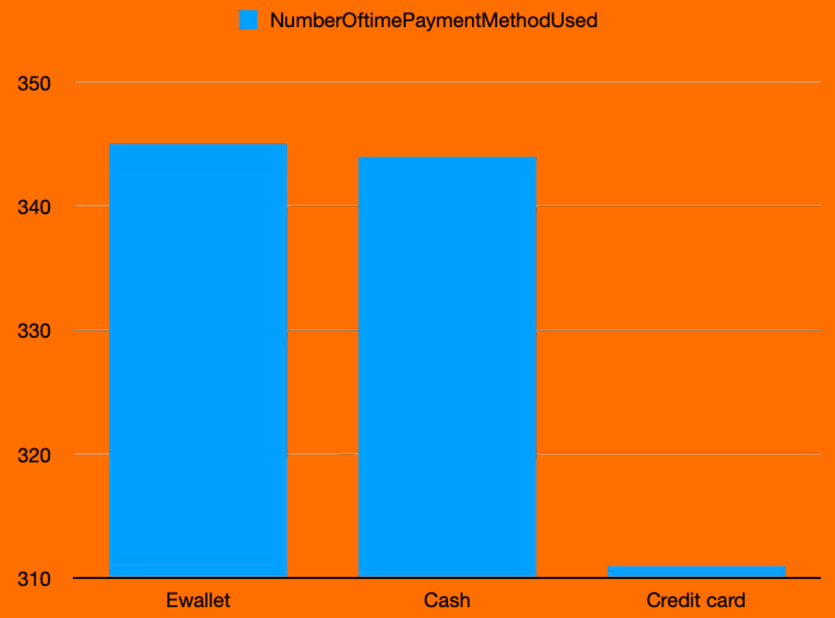


branch	ProductsSold
A	1859
C	1831
B	1820

Sales Analysis

```sql
-- Which payment method occurs most frequently?
Select payment,count(*) As NumberOftimePaymentMethodUsed
from amazon
group by payment
order by count(*) DESC
limit 1;
```

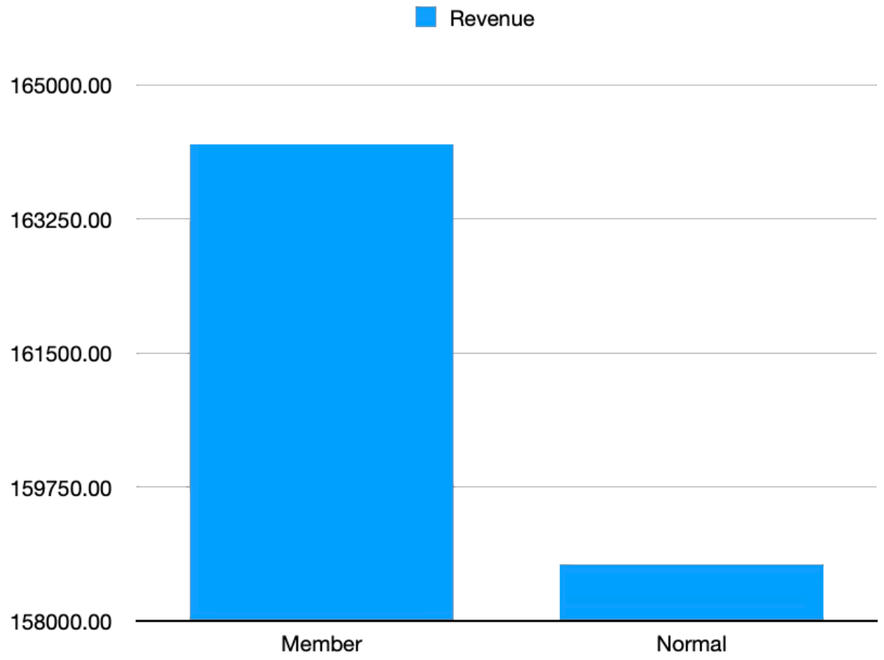
Most used paymentMode is Ewallet very closely followed by cash.



payment	NumberOftimePaymentMethodUse
Ewallet	345
Cash	344
Credit card	311

Customer Analysis

```sql
-- Identify the customer type contributing the highest revenue.
Select customer_type , Round(Sum(total),2) as Revenue
from amazon
group by customer_type
order by Revenue  DESC
Limit 1 ;
```

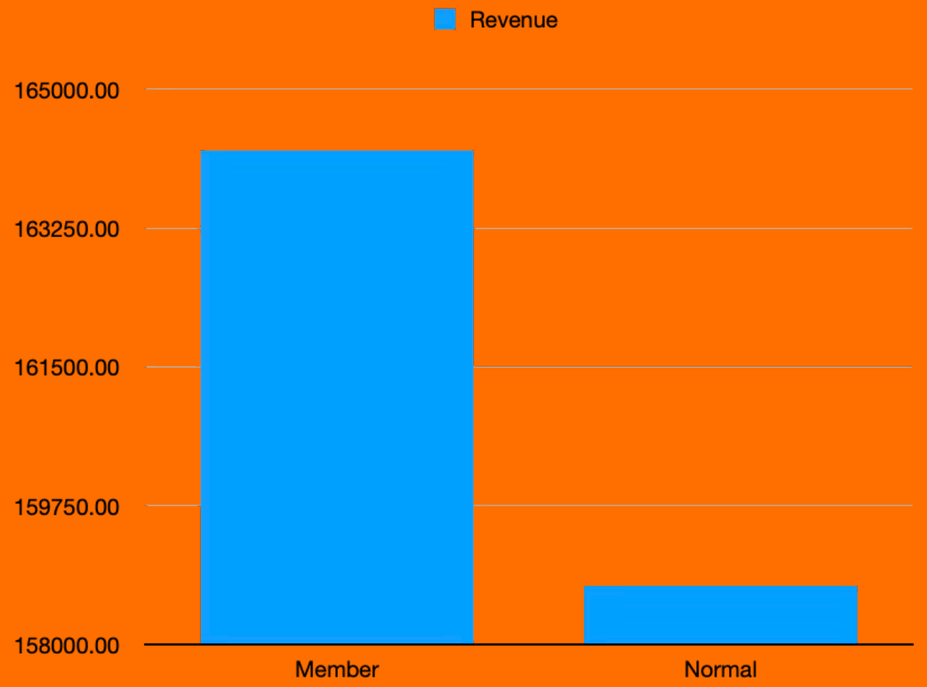


customer_type	Revenue
Member	164223.44
Normal	158743.31

Customer Analysis

```sql
-- Identify the customer type with the highest purchase frequency.
Select customer_type,count(*) As purchaseFrequency
from amazon
group by customer_type
Order by count(Invoice_ID) DESC
limit 1;
```

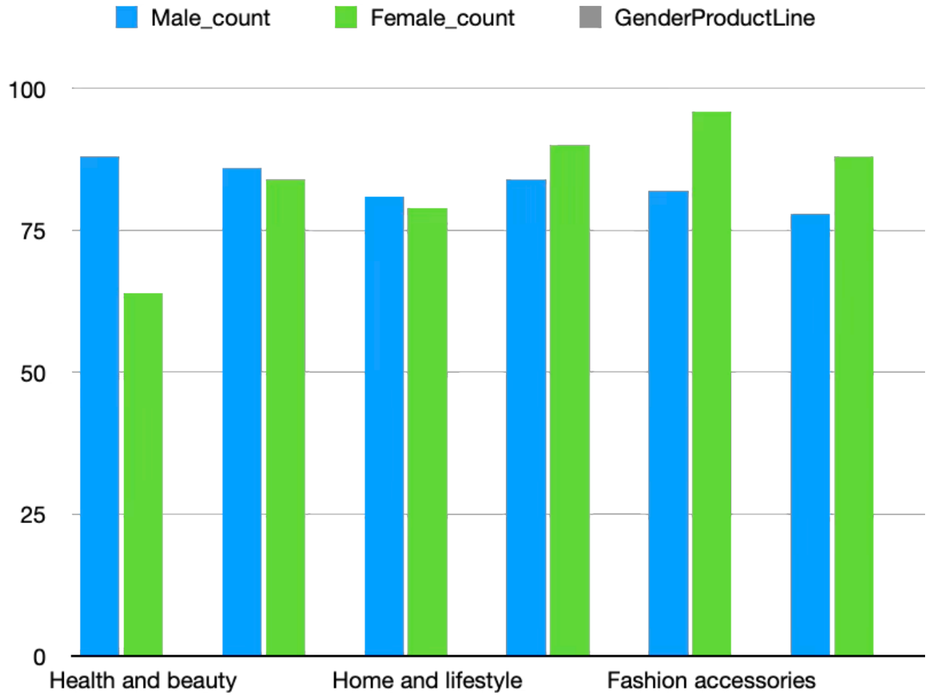
Members have the highest purchase Frequency .



customer_type	Revenue
Member	164223.44
Normal	158743.31

Customer Analysis

```sql
with MaleCte as(
Select product_line,Count(*) as Male_count,gender
from amazon
where gender = 'Male'
group by product_line,gender
),
FemaleCte as (
Select product_line,Count(*) as Female_count,gender
from amazon
where gender = 'Female'
group by product_line,gender
)
Select m.product_line ,Male_count,Female_count,
Case
    When Male_count > Female_count Then 'MaleProductLine' Else 'FemaleProductLine'
    End as GenderProductLine
from MaleCte m join FemaleCte f on
m.product_line = f.product_line;
```

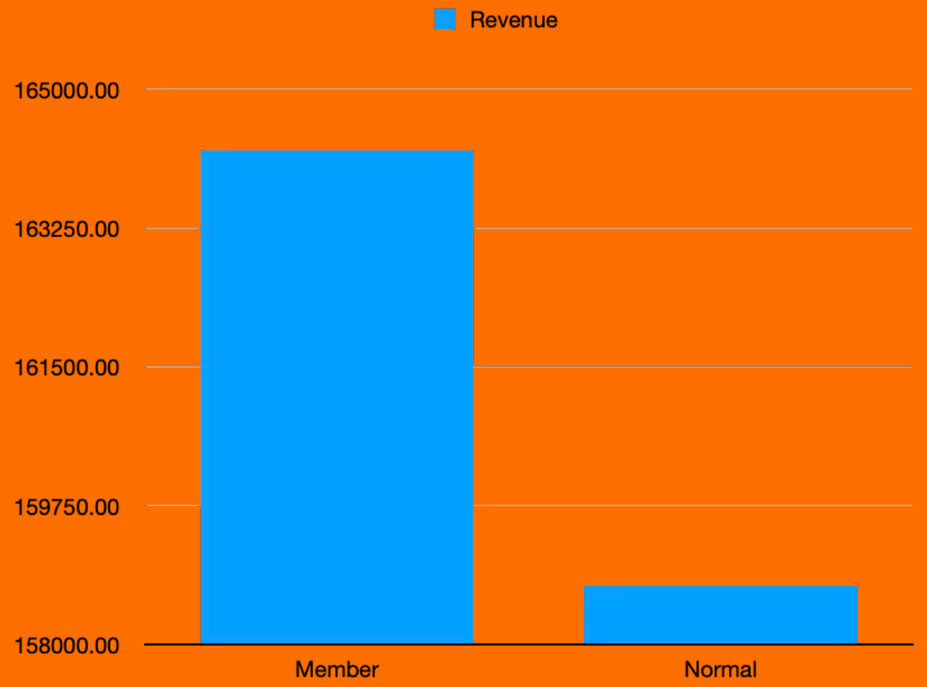


product_line	Male_count	Female_count	GenderProductLine
Health and beauty	88	64	MaleProductLine
Electronic accessories	86	84	MaleProductLine
Home and lifestyle	81	79	MaleProductLine
Food and beverages	84	90	FemaleProductLine
Fashion accessories	82	96	FemaleProductLine
Sports and travel	78	88	FemaleProductLine

Customer Analysis

```sql
-- Identify the customer type with the highest purchase frequency.
Select customer_type,count(*) As purchaseFrequency
from amazon
group by customer_type
Order by count(Invoice_ID) DESC
limit 1;
```

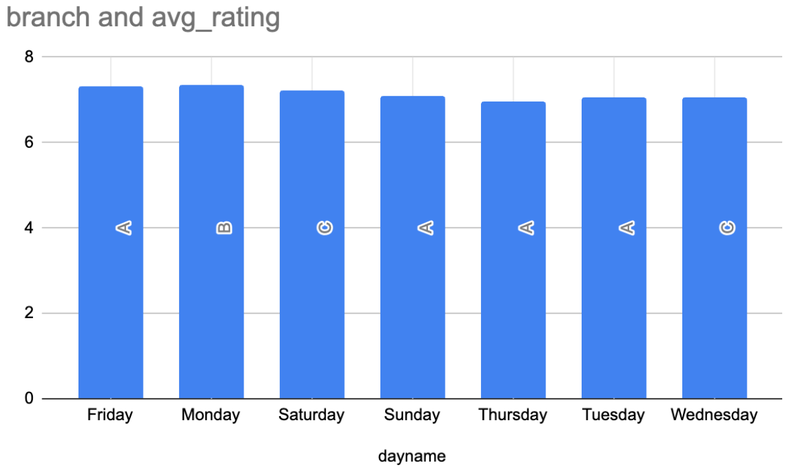
Members have the highest purchase Frequency .



customer_type	Revenue
Member	164223.44
Normal	158743.31

Customer Analysis

```sql
-- Determine the day of the week with the highest average ratings for each branch.

WITH rating_cte AS (
    SELECT
        dayname,
        branch,
        AVG(rating) AS avg_rating
    FROM amazon
    GROUP BY dayname, branch
)
,cte2 as (
SELECT
    dayname,
    branch,
    avg_rating,
    RANK() OVER (PARTITION BY dayname ORDER BY avg_rating DESC) AS rating_rank
FROM rating_cte)

Select dayname,branch,avg_rating
from cte2
where rating_rank = 1
ORDER BY dayname;
```



dayname	branch	avg_rating
Friday	A	7.31
Monday	B	7.34
Saturday	C	7.23
Sunday	A	7.08
Thursday	A	6.96
Tuesday	A	7.06
Wednesday	C	7.06

Key Bussiness Insights

Product Insights

- Top Performer: "Food and beverages" is the highest-selling product line.
- Area for Growth: "Health and beauty" is the only category with below-average sales.
- Customer Satisfaction: All product lines maintain consistently high average ratings of around 7 out of 10.

Sales Insights

- Top Market: Naypyitaw is the city with the highest recorded revenue.
- Peak Hours: Sales are highest in the afternoon, especially on Mondays, Wednesdays, and Thursdays.
- Payment Trends: Customers strongly prefer using Ewallet and Cash over credit cards.

Customer Insights

- Most Valuable Segment: "Member" customers are the most valuable, generating more revenue and purchasing more frequently than "Normal" customers.
- Gender Preferences: There are clear purchasing patterns based on gender. For instance, females purchase more "Fashion accessories," while males lead in "Health and beauty" sales.

Get in Touch !

dvssai.work@gmail.com
https://www.linkedin.com/in/dvs-sai/