WALLMART SALES ANALSIS

BY Mariyum Baig

Introduction

This project explores Walmart sales data from the Kaggle Walmart Sales Forecasting Competition to uncover key insights into product performance, sales trends, and customer behavior.
The dataset includes sales transactions from three branches containing 17 columns such as invoice details,product information, pricing, customer demographics, and sales metrics

Objectives:

- Identify top-performing branches & products
- Analyze sales trends
- Understand customer behavior
- Optimize sales strategies

Data Analysis Approach

Steps followed:

1.Data Wrangling: Handling missing/null value.

2.Feature Engineering: Adding new insights

(e.g., time of day, day name,month name)

3.Exploratory Data Analysis (EDA):

Answering business questions you need it

Feature Engineering

- ✓ Time of Day (Morning, Afternoon, Evening)

  to analyze peak sales hours

- ✓ Day Name (Mon-Sun)

  to find the busiest day ofthe week

- ✓ Month Name (Jan-Dec)

  to determine the highest sales month

ADD Column : DAY NAME

```sql
/*add column day name*/
SELECT DATENAME(weekday, Date) AS day_name
FROM WalmartData;
/*ADD COLUMN*/
ALTER TABLE WalmartData ADD day_name VARCHAR(10);
/*UPDATE COLUMN*/
UPDATE WalmartData
SET day_name = DATENAME(weekday, [Date]);
/* VIEW TABLE*/
Select * FROM WalmartData;
```

gross_margin_percentage	gross_income	Rating	day_name
4.7619047164917	26.1415004730225	9.10000038146973	Saturday
4.7619047164917	3.8199999332428	9.60000038146973	Friday
4.7619047164917	16.2154998779297	7.40000009536743	Sunday
4.7619047164917	23.2880001068115	8.39999961853027	Sunday
4.7619047164917	30.2084999084473	5.30000019073486	Friday
4.7619047164917	29.8864994049072	4.09999990463257	Monday
4.7619047164917	20.6520004272461	5.80000019073486	Monday
4.7619047164917	36.7799987792969	8	Sunday

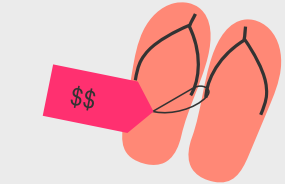
ADD COLUMN :MONTH NAME

```sql
/*add column MONTH name*/
SELECT DATENAME(MONTH, Date) AS month_name
FROM WalmartData;
/*ADD COLUMN*/
ALTER TABLE WalmartData ADD month_name VARCHAR(10);
/*UPDATE COLUMN*/
UPDATE WalmartData
SET month_name= DATENAME(MONTH, [Date]);
/* VIEW TABLE*/
Select * FROM WalmartData;
```

gross_income	Rating	day_name	month_name
26.1415004730225	9.10000038146973	Saturday	January
3.8199999332428	9.60000038146973	Friday	March
16.2154998779297	7.40000009536743	Sunday	March
23.2880001068115	8.39999961853027	Sunday	January
30.2084999084473	5.30000019073486	Friday	February
29.8864994049072	4.09999990463257	Monday	March
20.6520004272461	5.80000019073486	Monday	February
36.7799987792969	8	Sunday	February

Business Questions & Analysis

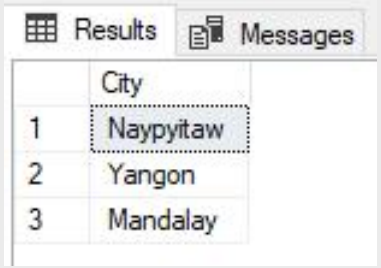
1. How many unique cities does the data have?
2. In which city is each branch?
3. How many unique product line does the data have?
4. What product_line had the largest revenue?
5. What city had the largest revenue?
6. What product line had the largest VAT?
7. What is most common payment method?
8. What is most selling product?
9. What is revenue by month?
10. What month had the largest cogs?
11. Which branch sold more project than average product sold
12. What is the most common product line by gender?
13. What is the average ratings for each product line



How many unique cities does the data have?

```
USE [Wallmart_Sales_Analysis];
GO

  Select * FROM WalmartData;
 /*How many unique cities does the data have?*/
  Select Distinct(City) From WalmartData;
```



	City
1	Naypyitaw
2	Yangon
3	Mandalay

Count no of branches in each city?

```sql
/*Count no of branches in each city?*/
  Select City,Count(Branch) AS noofbrancheseachcity
  FROM WalmartData
  GROUP BY City;
```

Results | Messages

	City	noofbrancheseachcity
1	Naypyitaw	328
2	Yangon	340
3	Mandalay	332

How many unique product line does the data have?

```sql
/*How many unique product line does the data have?*/
Select DISTINCT Product_line
FROM WalmartData;
```

Results | Messages

	Product_line
1	Fashion accessories
2	Health and beauty
3	Electronic accessories
4	Food and beverages
5	Sports and travel
6	Home and lifestyle

What product_line had the largest revenue??

$$

```sql
/*What city had the largest revenue?*/
  Select City,ROUND(SUM(Total),2) AS Revenue
  FROM WalmartData
  GROUP BY City
  ORDER BY Revenue DESC;
```

100 %

Results | Messages

	City	Revenue
1	Naypyitaw	110568.71
2	Yangon	106200.37
3	Mandalay	106197.67

What city had the largest revenue?

$$

```sql
/*What city had the largest revenue?*/
  Select City,ROUND(COUNT(Total),2) AS Revenue
  FROM WalmartData
  GROUP BY City
  ORDER BY Revenue DESC;
```

100 %

Results | Messages

	City	Revenue
1	Yangon	340
2	Mandalay	332
3	Naypyitaw	328

What product line had the largest VAT?

$$

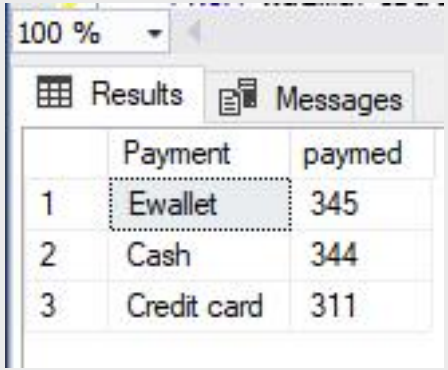
```sql
/*What product line had the largest VAT?*/
 Select product_line,ROUND(SUM(Tax_5),2) AS total_vat
 FROM WalmartData
 GROUP BY product_line
 ORDER BY total_vat DESC;
```

Results | Messages

	product_line	total_vat
1	Food and beverages	2673.56
2	Sports and travel	2624.9
3	Electronic accessories	2587.5
4	Fashion accessories	2586
5	Home and lifestyle	2564.85
6	Health and beauty	2342.56

What is most common payment method?

```sql
/*What is most common payment method?*/
Select Payment,Count(Payment) as paymed
FROM WalmartData
GROUP BY Payment
Order BY paymed desc;
```



	Payment	paymed
1	Ewallet	345
2	Cash	344
3	Credit card	311

What is most selling product?

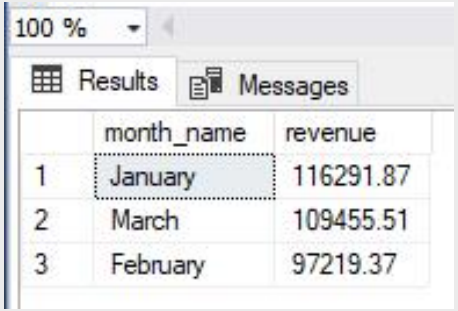
```sql
/*What is most selling product?*/
 Select Product_line,Round(count(total ),2) as total
 FROM WalmartData
 GROUP BY Product_line
 ORDER BY total DESC;
```

Results | Messages

	Product_line	total
1	Fashion accessories	178
2	Food and beverages	174
3	Electronic accessories	170
4	Sports and travel	166
5	Home and lifestyle	160
6	Health and beauty	152

What is revenue by month?

```sql
/*What is revenue by each month?*/
 Select month_name,ROUND(SUM(total),2) AS revenue
 FROM WalmartData
 GROUP BY month_name
 ORDER BY revenue DESC;
```



	month_name	revenue
1	January	116291.87
2	March	109455.51
3	February	97219.37

What month had the largest cogs?

$$$

```sql
/*What month had the largest cogs?*/
  Select month_name,ROUND(SUM(cogs),2) AS cogs
  FROM WalmartData
  GROUP BY month_name
  ORDER BY cogs DESC;
```

Results | Messages

	month_name	cogs
1	January	110754.16
2	March	104243.34
3	February	92589.88

Which branch sold more project than average product sold?

$$$

```
/*Which branch sold more project than average product sold*/
 Select Branch, SUM(Quantity) AS total_quantity
 FROM WalmartData
 GROUP BY Branch
 ORDER BY total_quantity DESC;
```

100 %

Results | Messages

	Branch	total_quantity
1	A	1859
2	C	1831
3	B	1820

What is the most common product line by gender?

$$$

Results | Messages

	Gender	Product_line	Count
1	Female	Fashion accessories	96
2	Female	Food and beverages	90
3	Female	Sports and travel	88
4	Female	Electronic accessories	84
5	Female	Home and lifestyle	79
6	Female	Health and beauty	64
7	Male	Health and beauty	88
8	Male	Electronic accessories	86
9	Male	Food and beverages	84
10	Male	Fashion accessories	82
11	Male	Home and lifestyle	81
12	Male	Sports and travel	78

```sql
/* What is the most common product line by gender?*/
  SELECT Gender, Product_line, COUNT(*) AS Count
FROM WalmartData
GROUP BY Gender, Product_line
ORDER BY Gender, Count DESC;
```

What is the average ratings for each product line?

```sql
/*What is the average ratings for each product line*/
Select product_line,ROUND(AVG(Rating),2) AS avg_rat
FROM WalmartData
GROUP BY product_line
ORDER BY avg_rat DESC;
```

100 %

Results | Messages

	product_line	avg_rat
1	Food and beverages	7.11
2	Fashion accessories	7.03
3	Health and beauty	7
4	Electronic accesso...	6.92
5	Sports and travel	6.92
6	Home and lifestyle	6.84

Thanks

www.linkedin.com/in/mariyum-baig-706128217