



WALMART SALES ANALYSIS


INTRODUCTION

Walmart, one of the largest retail corporations in the world, operates with an extensive product catalog and a vast customer base. Analyzing Walmart's sales data is crucial for understanding consumer behavior, optimizing inventory, and improving business strategies. This project focuses on leveraging SQL to analyze Walmart's sales data, with the goal of extracting valuable insights that can drive strategic decision-making.



SCHEMA



WalmartSalesData ...	
Branch	
City	
Σ cogs	
Customer type	
 Date	
Gender	
Σ gross income	
Σ gross margin percentage	
Invoice ID	
Payment	
Product line	
Σ Quantity	
Σ Rating	
Σ Tax 5%	
Time	
Σ Total	
Σ Unit price	
Collapse ^	



List of Contents



Generic

- How many unique cities does the data have?
- In which city is each branch?

Product

- How many unique product lines does the data have?
- What is the most selling product line
- What is the most selling product line?
- What is the total revenue by month?
- What month had the largest COGS?
- What product line had the largest revenue?
- What is the city with the largest revenue?
- What product line had the largest VAT?
- Fetch each product line and add a column to those product line showing "Good", "Bad".
- Which branch sold more products than average product sold?
- What is the most common product line by gender?
- What is the average rating of each product line?



Customers

- How many unique customer types does the data have?
- How many unique payment methods does the data have?
- What is the most common customer type?
- Which customer type buys the most?
- What is the gender of most of the customers?
- What is the gender distribution per branch?
- Which time of the day do customers give most ratings?
- Which time of the day do customers give most ratings per branch?
- Which day of the week has the best avg ratings?
- Which day of the week has the best average ratings per branch?

Sales

- Number of sales made in each time of the day per weekday.
- Which of the customer types brings the most revenue?

Generic



- How many unique cities does the data have?
- In which city is each branch?

Product



- How many unique product lines does the data have?
- What is the most selling product line
- What is the most selling product line?
- What is the total revenue by month?
- What month had the largest COGS?
- What product line had the largest revenue?
- What is the city with the largest revenue?
- What product line had the largest VAT?
- Fetch each product line and add a column to those product line showing "Good", "Bad".
- Which branch sold more products than average product sold?
- What is the most common product line by gender?
- What is the average rating of each product line?

CUSTOMERS



- How many unique customer types does the data have?
- How many unique payment methods does the data have?
- What is the most common customer type?
- Which customer type buys the most?
- What is the gender of most of the customers?
- What is the gender distribution per branch?
- Which time of the day do customers give most ratings?
- Which time of the day do customers give most ratings per branch?
- Which day of the week has the best avg ratings?
- Which day of the week has the best average ratings per branch?

SALES



- Number of sales made in each time of the day per weekday.
- Which of the customer types brings the most revenue?

How many unique cities does the data have?



```
SELECT DISTINCT  
  (city)  
FROM  
  sales;
```

Result Grid	
	city
▶	Yangon
	Naypyitaw
	Mandalay

In which city is each branch?



```
SELECT  
    DISTINCT city,  
    branch  
FROM sales;
```

city	branch
Yangon	A
Naypyitaw	C
Mandalay	B

How many unique product lines does the data have?



```
SELECT  
    COUNT(DISTINCT product_line)  
FROM  
    sales;
```

COUNT(DISTINCT product_line)
6

What is the most common payment method?



```
SELECT
    payment_method, COUNT(payment_method) AS CNT
FROM
    sales
GROUP BY payment_method
ORDER BY CNT DESC;
```

payment_method	CNT
Cash	344
Ewallet	342
Credit card	309

What is the most selling product line?



```
SELECT
    product_line, COUNT(product_line) AS CNT
FROM
    sales
GROUP BY product_line
ORDER BY CNT DESC;
```

product_line	CNT
Fashion accessories	178
Food and beverages	174
Electronic accessories	169
Sports and travel	163
Home and lifestyle	160
Health and beauty	151

What is the total revenue by month?



```
SELECT
    MONTH_NAME AS MONTH, SUM(total) AS total_revenue
FROM
    sales
GROUP BY MONTH
ORDER BY total_revenue DESC;
```

MONTH	total_revenue
January	116291.8680
March	108867.1500
February	95727.3765

What month had the largest COGS?



```
SELECT
    month_name, SUM(cogs) AS COGS
FROM
    sales
GROUP BY month_name
ORDER BY COGS DESC;
```

month_name	COGS
January	110754.16
March	103683.00
February	91168.93

What product line had the largest revenue?



```
SELECT
    product_line, SUM(total) AS revenue
FROM
    sales
GROUP BY product_line
ORDER BY revenue DESC;
```

product_line	revenue
Food and beverages	56144.8440
Fashion accessories	54305.8950
Sports and travel	53936.1270
Home and lifestyle	53861.9130
Electronic accessories	53783.2365
Health and beauty	48854.3790

What is the city with the largest revenue?



```
SELECT  
    branch, city, SUM(total) AS revenue  
FROM  
    sales  
GROUP BY branch , city  
ORDER BY revenue DESC;
```

branch	city	revenue
C	Naypyitaw	110490.7755
A	Yangon	105861.0105
B	Mandalay	104534.6085

What product line had the largest VAT?



```
SELECT
    product_line, AVG(VAT) AS Vat
FROM
    sales
GROUP BY product_line
ORDER BY Vat DESC;
```

product_line	Vat
Home and lifestyle	16.03033124
Sports and travel	15.75697549
Health and beauty	15.40661591
Food and beverages	15.36531029
Electronic accessories	15.15447632
Fashion accessories	14.52806181

Which branch sold more products than average product sold?



```
SELECT
    branch, SUM(quantity) AS qty
FROM
    sales
GROUP BY branch
HAVING SUM(quantity) > (SELECT
    AVG(quantity)
FROM
    sales);
```

branch	qty
A	1849
C	1828
B	1795

What is the most common product line by gender?



```
SELECT
    gender, product_line, COUNT(gender) AS CNT
FROM
    sales
GROUP BY gender , product_line
ORDER BY CNT DESC;
```

gender	product_line	CNT
Female	Fashion accessories	96
Female	Food and beverages	90
Male	Health and beauty	88
Female	Sports and travel	86
Male	Electronic accessories	86
Male	Food and beverages	84
Female	Electronic accessories	83
Male	Fashion accessories	82
Male	Home and lifestyle	81
Female	Home and lifestyle	79
Male	Sports and travel	77
Female	Health and beauty	63

What is the average rating of each product line?



```
SELECT
    product_line, ROUND(AVG(rating), 2) AS RATING
FROM
    sales
GROUP BY product_line
ORDER BY RATING DESC;
```

product_line	RATING
Food and beverages	7.11
Fashion accessories	7.03
Health and beauty	6.98
Electronic accessories	6.91
Sports and travel	6.86
Home and lifestyle	6.84

Number of sales made in each time of the day per weekday



```
SELECT
    TIME_OF_DAY, COUNT(*) AS Total_Sales
FROM
    sales
WHERE
    DAY_NAME = 'Wednesday'
GROUP BY TIME_OF_DAY
ORDER BY Total_Sales DESC;
```

TIME_OF_DAY	Total_Sales
Afternoon	61
Evening	58
Morning	22

Which of the customer types brings the most revenue?



```
SELECT
    customer_type, SUM(total) AS revenue
FROM
    sales
GROUP BY customer_type
ORDER BY revenue DESC;
```

customer_type	revenue
Member	163625.1015
Normal	157261.2930

Which city has the largest tax percent/ VAT (Value Added Tax)?



```
SELECT
    city, AVG(VAT) AS max_vat
FROM
    sales
GROUP BY city
ORDER BY max_vat DESC;
```

city	max_vat
Naypyitaw	16.09010850
Mandalay	15.13020824
Yangon	14.87020798

Which customer type pays the most in VAT?



```
SELECT
    customer_type, AVG(VAT) AS max_vat
FROM
    sales
GROUP BY customer_type
ORDER BY max_vat DESC;
```

customer_type	max_vat
Member	15.61457214
Normal	15.09805040

How many unique customer types does the data have?



```
SELECT DISTINCT  
  customer_type  
FROM  
  sales;
```

customer_type
Normal
Member

How many unique payment methods does the data have?



```
SELECT DISTINCT  
    payment_method  
FROM  
    sales;
```

payment_method
Credit card
Ewallet
Cash

What is the most common customer type?



```
SELECT
    customer_type, COUNT(customer_type) AS customer
FROM
    sales
GROUP BY customer_type
ORDER BY customer DESC
LIMIT 1;
```

customer_type	customer
Member	499

Which customer type buys the most?



```
SELECT
    customer_type,
    COUNT(*)
FROM sales
GROUP BY customer_type;
```

customer_type	COUNT(*)
Normal	496
Member	499

What is the gender of most of the customers?



```
SELECT
    gender, COUNT(customer_type) AS customer
FROM
    sales
GROUP BY gender
ORDER BY customer DESC
LIMIT 1;
```

gender	customer
Male	498

What is the gender distribution per branch?



```
SELECT  
    branch, gender, COUNT(*) AS distribution  
FROM  
    sales  
GROUP BY branch , gender;
```

branch	gender	distribution
A	Male	179
C	Female	177
C	Male	150
B	Female	160
B	Male	169
A	Female	160

Which time of the day do customers give most ratings?



```
SELECT
    TIME_OF_DAY, AVG(rating) AS avg_rating
FROM
    sales
GROUP BY TIME_OF_DAY
ORDER BY avg_rating DESC
LIMIT 1;
```

	TIME_OF_DAY	avg_rating
►	Afternoon	7.02340

Which time of the day do customers give most ratings per branch?



```
SELECT
    branch, TIME_OF_DAY, AVG(rating) AS avg_rating
FROM
    sales
WHERE
    branch = 'B'
GROUP BY branch , TIME_OF_DAY
ORDER BY avg_rating DESC
LIMIT 1;
```

branch	TIME_OF_DAY	avg_rating
B	Morning	6.83793

Which day of the week has the best avg ratings?



```
SELECT
    DAY_NAME, AVG(rating) AS avg_ratings
FROM
    sales
GROUP BY DAY_NAME
ORDER BY avg_ratings DESC
LIMIT 1;
```

DAY_NAME	avg_ratings
Monday	7.13065

Which day of the week has the best average ratings per branch?



```
SELECT
    DAY_NAME, branch, AVG(rating) AS avg_ratings
FROM
    sales
GROUP BY DAY_NAME , branch
ORDER BY avg_ratings DESC
LIMIT 1
```

DAY_NAME	branch	avg_ratings
Friday	A	7.31200

Number of sales made in each time of the day per weekday.



```
SELECT
    time_of_day,
    COUNT(*) AS total_sales
FROM sales
WHERE day_name = "Sunday"
GROUP BY time_of_day
ORDER BY total_sales DESC;
```

time_of_day	total_sales
Evening	58
Afternoon	52
Morning	22

Which of the customer types brings the most revenue?



```
SELECT
    customer_type,
    SUM(total) AS total_revenue
FROM sales
GROUP BY customer_type
ORDER BY total_revenue;
```

customer_type	total_revenue
Normal	157261.2930
Member	163625.1015

THANK YOU

