

Walmart



Walmart Sales Analysis

This analysis of Walmart's sales data provides critical insights that can be leveraged to increase revenue. By identifying top-selling products, peak purchasing times, customer preferences, and more, Walmart can implement targeted strategies to maximize sales and profitability.

- First, we created a database name Walmart_sales_data.
- Then we created a table to insert data from the excel sheet.

```
1 • create database walmart_sales_data;
2
3 • create table if not exists sales(
4     invoice_id varchar(30) not null primary key,
5     branch varchar(5) not null,
6     city varchar(30) not null,
7     customer_type varchar(10) not null,
8     gender varchar(10) not null,
9     product_line varchar(100) not null,
10    unit_price decimal(10, 2) not null,
11    quantitiy int not null,
12    VAT float(6, 4) not null,
13    total decimal(12, 4) not null,
14    date datetime not null,
15    time time not null,
16    payment_method varchar(15) not null,
17    cogs decimal(10, 2) not null,
18    gross_margin_percentage float(11, 9),
19    gross_income decimal(12, 4) not null,
20    rating float(2, 1)
21
22 );
```

Added a new column to give insights into the data

- The first column was about the time of the Morning, Afternoon and evening

```
25  -- -----feature engineering-----
26  -- -----
27  -- ----- Time of the day
28
29  -- 1. Adding a new column "time_of_day" to give insights of sales in the Morning, Afternoon and
30  -- Evening. This will help answer the question on which part of the day most sales are made.
31
32  ● SELECT time,
33      CASE
34          WHEN time BETWEEN '00:00:00' AND '11:59:59' THEN 'Morning'
35          WHEN time BETWEEN '12:00:00' AND '15:59:59' THEN 'Afternoon'
36          ELSE 'Evening'
37      END AS time_of_day
38  FROM sales;
39
40
41  ● alter table sales add column time_of_day varchar(20);
42
43  ● UPDATE sales
44      SET time_of_day =
45      CASE
46          WHEN time BETWEEN '00:00:00' AND '11:59:59' THEN 'Morning'
47          WHEN time BETWEEN '12:00:00' AND '15:59:59' THEN 'Afternoon'
48          ELSE 'Evening'
49      END;
```

- The second column was about day name of the week

```
-- Day name --  
-- 2. Add a new column "day_name" that contains the extracted days of the week on which  
-- the given transaction took place (Mon, Tue, Thur, Fri). This will help answer the question  
-- on which week of the day each branch is busiest.  
  
select date,  
dayname(date) as day_name  
from sales;  
  
alter table sales add column day_name varchar(10);  
  
update sales  
set day_name = dayname(date);
```

- The third column was about month name of the extracted year.

```
-- 3. Add a new column named month_name that contains the extracted months of the year  
-- on which the given transaction took place (Jan, Feb, Mar). Help determine which  
-- month of the year has the most sales and profit  
  
select date,  
monthname(date)  
from sales;  
  
alter table sales add column month_name varchar(10);  
  
update sales  
set month_name = monthname(date);
```

Objective: The primary objective of this analysis is to identify key operational insights that can drive revenue growth and improve customer satisfaction for Walmart. By examining sales data across various cities, product lines, payment methods, and customer demographics, the analysis aims to determine the most profitable cities, product lines, and customer segments. Additionally, understanding sales trends by time, day, and month, along with customer ratings and VAT contributions, will allow Walmart to optimize store operations, inventory management, and marketing strategies for maximum efficiency and profitability.

Questions:

- Q1. How many unique cities does the data have ?
- Q2. In which city is each branch ?
- Q3. How many unique products line does the data have ?
- Q4. What is the most common payment method ?
- Q5. What is the most selling product line ?
- Q6. What is the total revenue by month ?
- Q7. What month had the largest COGS (cost of goods sold) ?
- Q8. What product line had the largest revenue ?
- Q9. What is the city with the largest revenue ?
- Q10. What product line has the largest VAT ?
- Q12. What is the most common product line by gender ?
- Q13. What is the average rating of each product line ?
- Q14. Number of sales made in each time of the day per week. for ex. "Sunday"
- Q15. Which of the customer types brings the most revenue ?
- Q16. Which city has the largest tax percent / Value added tax (VAT) ?
- Q17. Which customer type pays the most VAT ?
- Q18. Which customer type buys the most ?
- Q19. What is the gender of the most customers ?
- Q20. Which time of the day customers give more ratings ?
- Q21. Which day of the week has the best average rating

Solutions:

Q1. How many unique cities does the data have?

```
4  -- Q1. How many unique cities does the data have ?
5
6  •  select
7     distinct city
8     from sales;
```

Result Grid | | Filter Rows: | Export: | Wrap Cell Conte

	city
▶	Yangon
	Naypyitaw
	Mandalay

Q2. In which city is each branch?

```
1  -- Q2. In which city is each branch ?
2
3  •  select
4     distinct city,
5     branch
6     from sales;
```

Result Grid | | Filter Rows: | Export: | Wrap Ce

	city	branch
▶	Yangon	A
	Naypyitaw	C
	Mandalay	B

Q3. How many unique products line does the data have?

```
1  -- Q3. How many unique product line does the data have ?
2
3  • select
4    distinct product_line
5  from sales;
```

Result Grid | | Filter Rows: | Export: | Wrap Cell Content:

	product_line
▶	Food and beverages
	Health and beauty
	Sports and travel
	Fashion accessories
	Home and lifestyle
	Electronic accessories

Result Grid | | Filter Rows:

	count(distinct product_line)
▶	6

Q4. What is the most common payment method?

```
1  -- Q4. What is the most common payment method ?
2  • select payment_method, count(payment_method) as Count
3  from sales
4  group by payment_method
5  order by Count desc;
6
```

Result Grid | | Filter Rows: | Export: | Wrap Cell Content:

	payment_method	Count
▶	Cash	344
	Ewallet	342
	Credit card	309

Q5. What is the most selling product line?

```
1  -- Q5. What is the most selling product line ?
2
3  •  select product_line,
4     count(product_line) as Count
5     from sales
6     group by product_line
7     order by Count desc;
```

Result Grid | | Filter Rows: | Export: | Wrap Cell C

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

Q6. What is the total revenue by month?

```
1  -- Q6. What is the total revenue by month ?
2
3  •  select month_name as month, round(sum(total), 2) as total_revenue
4     from sales
5     group by month
6     order by total_revenue;
```

Result Grid | | Filter Rows: | Export: | Wrap Cell Content:

	month	total_revenue
▶	February	95727.38
	March	108867.15
	January	116291.87

Q7. What month had the largest COGS (cost of goods sold) ?

```
1  -- Q7. What month had the largest COGS (cost of goods sold) ?
2
3  •  select month_name as month,
4     sum(cogs) as cogs
5  from sales
6  group by month
7  order by cogs desc;
```

Result Grid			Filter Rows:	Export:	Wrap Cell Content:
	month	cogs			
▶	January	110754.16			
	March	103683.00			
	February	91168.93			

Q8. What product line had the largest revenue ?



```
1  -- Q8. What product line had the largest revenue ?
2
3  •  select product_line,
4     round(sum(total), 2) as total_revenue
5  from sales
6  group by product_line
7  order by total_revenue desc;
```


Result Grid			Filter Rows:	Export:	Wrap Cell Cont
	product_line	total_revenue			
▶	Food and beverages	56144.84			
	Fashion accessories	54305.90			
	Sports and travel	53936.13			
	Home and lifestyle	53861.91			
	Electronic accessories	53783.24			
	Health and beauty	48854.38			

Q9. What is the city with the largest revenue ?

```
1      -- Q9. What is the city with the largest revenue ?
2
3      •  select branch, city,
4          round(sum(total), 2) as total_revenue
5      from sales
6      group by branch, city
7      order by total_revenue desc;
```

Result Grid



Filter Rows:

Export: 


Wrap Cell Con


	branch	city	total_revenue
▶	C	Naypyitaw	110490.78
	A	Yangon	105861.01
	B	Mandalay	104534.61

Q10. What product line has the largest VAT ?


```
1      -- Q10. What product line has the largest VAT ?
2
3      •  select product_line,
4          round(avg(VAT), 3) as VAT
5          from sales
6          group by product_line
7          order by VAT desc;
```

Result Grid



 Filter Rows:

Export:



Wrap Cell

	product_line	VAT
▶	Home and lifestyle	16.03
	Sports and travel	15.757
	Health and beauty	15.407
	Food and beverages	15.365
	Electronic accessories	15.154
	Fashion accessories	14.528

Q12. What is the most common product line by gender ?

```
1  -- Q12. What is the most common product line by gender ?
2
3  •  select gender, product_line,
4      count(gender) as total_count
5      from sales
6      group by gender, product_line
7      order by total_count desc;
```

gender	product_line	total_count
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

Q13. What is the average rating of each product line ?



```
1  -- Q13. What is the average rating of each product line ?
2
3  •  select product_line,
4      round(avg(rating), 2) as avg_rate
5      from sales
6      group by product_line
7      order by avg_rate desc;
```



product_line	avg_rate
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

Q14. Number of sales made in each time of the day per week. for ex. "Sunday"

```
1  -- Q14. Number of sale made in each time of the day per week. for ex. "Sunday"
2
3  •  select time_of_day,
4      count(time_of_day) as total_sale
5      from sales
6      where day_name = "Sunday"
7      group by time_of_day
8      order by total_sale desc; |
9
```

Result Grid

 Filter Rows:

Export:  Wrap Cell Content: 

	time_of_day	total_sale
▶	Evening	58
	Afternoon	52
	Morning	22

Q15. Which of the customer types brings the most revenue ?

```
1  -- Q15. Which of the customer types brings the most revenue ?
2
3  •  select customer_type,
4      round(sum(total), 2) as total_revenue
5      from sales
6      group by customer_type
7      order by total_revenue desc;
```

Result Grid

Filter Rows:

Export:

Wrap Cell Content:

	customer_type	total_revenue
▶	Member	163625.10
	Normal	157261.29

Q16. Which city has the largest tax percent / Value added tax (VAT) ?

```
1  -- Q16. Which city has the largest tax percent / Value added tax (VAT) ?
2
3  •  select city,
4     round(avg(VAT), 2) as tax_percent
5     from sales
6     group by city
7     order by tax_percent desc;
```

Result Grid			Filter Rows:	Export:	Wrap Cell Content:
	city	tax_percent			
▶	Naypyitaw	16.09			
	Mandalay	15.13			
	Yangon	14.87			

Q18. Which customer type buys the most ?

```
1  -- Q18. Which customer type buys the most ?
2
3  •  select customer_type,
4     count(customer_type) as customer_count
5     from sales
6     group by customer_type
7     order by customer_type desc;
```

Result Grid			Filter Rows:	Export:	Wrap
	customer_type	customer_count			
▶	Normal	496			
	Member	499			

Q19. What is the gender of the most customers ?

```
1  -- Q19. What is the gender of the most customers ?
2
3  •  select gender,
4     count(customer_type) as customers
5  from sales
6  group by gender
7  order by customers desc;
```

Result Grid			Filter Rows:	Export:	Wrap Cell Co
	gender	customers			
▶	Male	498			
	Female	497			

Q20. Which time of the day customers give more ratings ?

```
1  -- Q20. Which time of the day customers give more ratings ?
2
3  •  select time_of_day,
4     count(rating) as rating
5  from sales
6  group by time_of_day
7  order by rating desc;
```

Result Grid			Filter Rows:	Export:	Wrap Cell Content:
	time_of_day	rating			
▶	Evening	429			
	Afternoon	376			
	Morning	190			

Q21. Which day of the week has the best average rating ?

```
1  -- Q21. Which day of the week has the best avg rating ?
2
3  •  select day_name,
4      round(avg(rating), 2) as avg_customer_rating
5  from sales
6  group by day_name
7  order by avg_customer_rating desc;
8
```

Result Grid			Filter Rows:	Export:	Wrap Cell Content:
	day_name	avg_customer_rating			
▶	Monday	7.13			
	Friday	7.06			
	Tuesday	7			
	Sunday	6.99			
	Saturday	6.9			
	Thursday	6.89			
	Wednesday	6.76			

Walmart Revenue Enhancement Strategy Based on Data Insights:

1. Focus on Best-Selling Product Lines:

Based on the data, the product line with the largest revenue should be prioritized in promotions and stock availability. By ensuring adequate supply of this product line and promoting it through marketing campaigns, Walmart can capitalize on customer demand, leading to higher sales. Additionally, the most selling product line across different categories can be cross promoted with complementary products to encourage bulk buying.

2. Optimize Payment Methods:

The most common payment method should be streamlined for ease and convenience, with loyalty programs or rewards associated with that method. By promoting preferred payment options through discounts or reward points, Walmart can enhance customer satisfaction, potentially increasing repeat business.

3. City-Based Revenue Strategy:

The city with the largest revenue should receive special attention. Walmart can expand its store presence or offer city-specific promotions in this location to capture an even larger market share. Moreover, understanding the preferences of customers in this city, such as popular product lines or payment methods, can help in customizing the product offerings.

4. Maximize Sales by Time of Day:

The data indicates varying sales volumes at different times of the day. Walmart should tailor staffing, stock levels, and marketing campaigns to coincide with these peak times (e.g., mornings or evenings). By aligning store operations and promotions with the highest traffic periods, Walmart can increase sales efficiency and improve customer service.

5. Monthly and Seasonal Trends:

Walmart should analyze the total revenue generated by each month and focus on the month with the largest COGS (cost of goods sold). This could signify a high-demand period, which Walmart can capitalize on through promotions, bulk purchasing discounts, or targeted advertisements. By understanding these trends, Walmart can prepare in advance, ensuring stock levels are adequate to meet demand.

6. Gender and Customer Type Targeting:

Walmart can leverage insights about customer demographics, such as the most common gender of customers and which customer type brings in the most revenue. Tailored promotions and advertising campaigns focusing on this segment will likely increase conversion rates. Additionally, loyalty programs aimed at this customer type can incentivize repeat purchases.

7. Leverage Customer Reviews and Ratings:

Identifying the time of day when customers provide more ratings can help Walmart encourage reviews during these periods by offering small incentives (like discount coupons). Higher ratings can improve the store's reputation and attract new customers.

8. Tax and VAT Strategy:

Cities with the highest VAT percentages offer an opportunity for Walmart to explore partnerships or special VAT-inclusive pricing strategies to reduce the impact on customers. This can increase foot traffic and conversion rates in those regions.