

SQL PROJECT

Walmart Sales Analysis

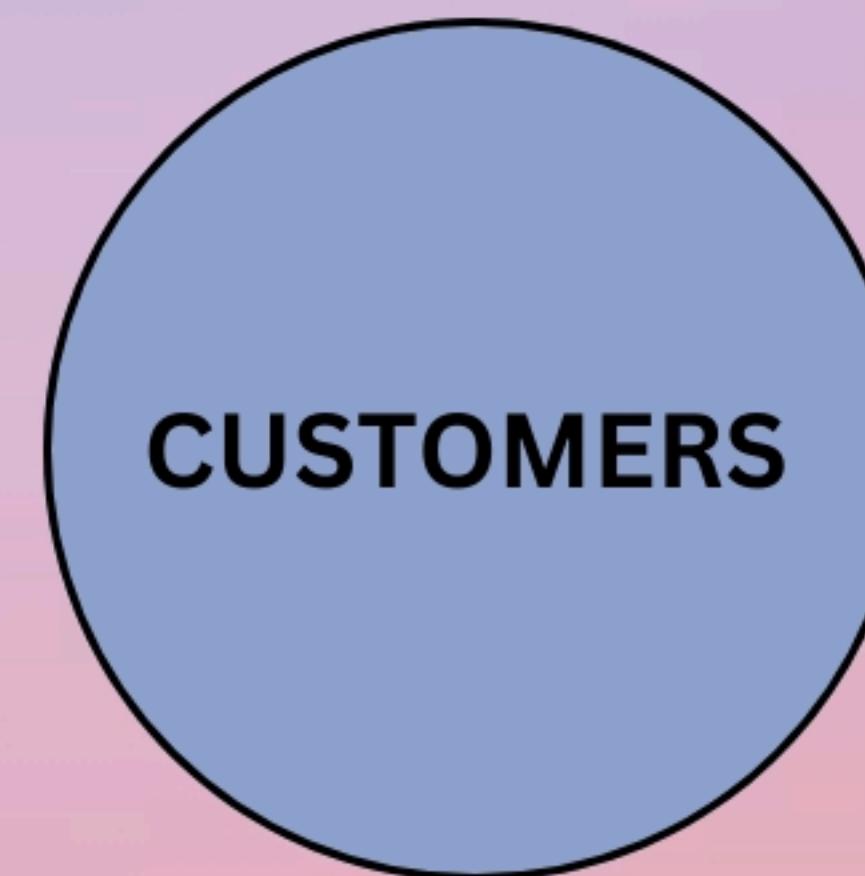
- Viraz Anand Gupta

ABOUT PROJECT

This project aims to explore the "Walmart Sales Data" to understand top performing branches and products, Sales trend of different products and Customer Behavior. The main aim is to study how sales strategies can be Improved and Optimized.

Business Question to Answer

The Questions are divided into Three Ways:



QUESTIONS ABOUT PRODUCT

Q1 - How many unique product line does the data have?

Q2 - What is the most common payment method?

Q3 - What is the most selling product line?

Q4 - What is the total revenue by month?

Q5 - What month had the largest GOGS?

Q6 - What product line the largest revenue?

Q7 - What is the city with the largest revenue?

Q8 - What product line had the largest VAT?

Q9 - Which branch sold more product than average product sold?

Q10 - What is the most common product line by gender?

Q11 - What is the average rating of each product line?

QUESTIONS ABOUT SALES

- Q1 - Number of Sales made in each time of the day per weekday?**
- Q2 - Which of the Customer types brings the most revenue?**
- Q3 - Which city has the largest tax/VAT precent?**
- Q4 - Which customer types pays the most in VAT?**

QUESTIONS ABOUT CUSTOMER

Q1 - How many unique customer types does the data have?

Q2 - How many unique payment methods does the data have?

Q3 - What is the most common customer type?

Q4 - Which customer type buys the most?

Q5 - What is the Gender of the most of the customer?

Q6 - What is the Gender distribution per branch?

Q7 - Which time of the day do customers give the most rating?

Q8 - Which time of the day do Customer give the most rating per day?

Q9 - Which day of the week has the best Average rating?

Q10 - Which day of the week has the best Average rating per branch?

INSIGHTS AFTER ANALYSIS

-- Q1 - How many unique product lines does the data have?

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

COUNT(DISTINCT
product_line)

6

-- Q2 - What is the most common payment method?

```
SELECT payment, COUNT(payment) FROM sales  
GROUP BY payment  
ORDER BY COUNT(payment) DESC;
```

	payment	COUNT(payment)
▶	Cash	344
	Ewallet	342
	Credit card	309

-- Q3 - What is the most selling product line?

```
SELECT SUM(quantity),product_line FROM sales  
GROUP BY product_line  
ORDER BY SUM(quantity) DESC;
```

	SUM(quantity)	product_line
▶	961	Electronic accessories
	952	Food and beverages
	911	Home and lifestyle
	902	Sports and travel
	902	Fashion accessories
	844	Health and beauty

```
-- Q4 - What is the total revenue by month?  
SELECT month_name AS month,  
SUM(total) AS total_revenue FROM sales  
GROUP BY month_name  
ORDER BY total_revenue;
```

	month	total_revenue
▶	February	95727.3765
	March	108867.1500
	January	116291.8680

-- Q5 - What month had the largest COGS?

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

	month	cogs
▶	February	91168.93
	March	102692.00
	January	103683.00

-- Q6 - What product line had the largest revenue?

```
SELECT product_line,  
       SUM(total) as total_revenue FROM sales  
    GROUP BY product_line  
ORDER BY total_revenue DESC;
```

product_line	total_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

-- Q7 - What is the city with the largest revenue?

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

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

-- Q8 - What product line had the largest VAT?

```
SELECT product_line,  
       AVG(tax_pct) as avg_tax FROM sales  
    GROUP BY product_line  
ORDER BY avg_tax DESC;
```

product_line	avg_tax
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

```
-- Q9 - Which branch sold more products than average product sold?  
  
SELECT product_line,  
CASE  
WHEN AVG(quantity) > 6 THEN "Good"  
ELSE "Bad"  
END AS remark  
FROM sales  
GROUP BY product_line;
```

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

```
-- Q10 - What is the most common product line by gender?  
SELECT gender,product_line,  
|COUNT(gender) AS total_cnt FROM sales  
GROUP BY gender, product_line  
ORDER BY total_cnt DESC;
```

gender	product_line	total_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

```
-- Q11 - What is the average rating of each product line?  
SELECT  
    ROUND(AVG(rating), 2) as avg_rating, product_line  
FROM sales  
GROUP BY product_line  
ORDER BY avg_rating DESC;
```

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

INSIGHTS AFTER ANALYSIS

```
-- Q1 - 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

```
-- Q2 - 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

```
-- Q3 - Which city has the largest tax/VAT percent?-  
SELECT city,  
ROUND(AVG(tax_pct), 2) AS avg_tax_pct FROM sales  
GROUP BY city  
ORDER BY avg_tax_pct DESC;
```

city	avg_tax_pct
Naypyitaw	16.09
Mandalay	15.13
Yangon	14.87

```
-- Q4 - Which customer type pays the most in VAT?
```

```
SELECT customer_type,  
AVG(tax_pct) AS total_tax FROM sales  
GROUP BY customer_type  
ORDER BY total_tax;
```

customer_type	total_tax
Normal	15.09805040
Member	15.61457214

INSIGHTS AFTER ANALYSIS

-- Q1 - How many unique customer types does the data have?-

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

customer_type
Normal
Member

-- Q2 - How many unique payment methods does the data have?

```
SELECT DISTINCT payment FROM sales;
```

payment

Credit card

Ewallet

Cash

```
-- Q3 - What is the most common customer type?  
SELECT customer_type,  
count(*) as count FROM sales  
GROUP BY customer_type  
ORDER BY count DESC;
```

customer_type	count
Member	499
Normal	496

-- Q4 - Which customer type buys the most?

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

customer_type	COUNT(*)
Normal	496
Member	499

-- Q5 - What is the gender of most of the customers?-

```
SELECT gender,COUNT(*) as gender_cnt FROM sales  
GROUP BY gender  
ORDER BY gender_cnt DESC;
```

gender	gender_cnt
Male	498
Female	497

```
-- Q6 - What is the gender distribution per branch?  
SELECT gender, COUNT(*) as gender_cnt FROM sales  
WHERE branch = "C"  
GROUP BY gender  
ORDER BY gender_cnt DESC;
```

gender	gender_cnt
Female	177
Male	150

```
-- Q7 - 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;
```

time_of_day	avg_rating
Afternoon	7.02340
Morning	6.94474
Evening	6.90536

```
-- Q8 - Which time of the day do customers give most ratings per branch?  
SELECT time_of_day, AVG(rating) AS avg_rating FROM sales  
WHERE branch = "A"  
GROUP BY time_of_day  
ORDER BY avg_rating DESC;
```

time_of_day	avg_rating
Afternoon	7.18889
Morning	7.00548
Evening	6.87143

```
-- Q9 - Which day of the week has the best avg ratings?  
SELECT day_name,AVG(rating) AS avg_rating FROM sales  
GROUP BY day_name  
ORDER BY avg_rating DESC;
```

day_name	avg_rating
Monday	7.13065
Friday	7.05507
Tuesday	7.00316
Sunday	6.98864
Saturday	6.90183
Thursday	6.88986
Wednesday	6.76028

```
-- Q10 - Which day of the week has the best average ratings per branch?  
SELECT day_name,COUNT(day_name) total_sales FROM sales  
WHERE branch = "C"  
GROUP BY day_name  
ORDER BY total_sales DESC;
```

day_name	total_sales
Tuesday	54
Saturday	54
Wednesday	50
Thursday	48
Sunday	46
Monday	38
Friday	37

Thank You!