SQL Project Report on:

Analyzing Walmart Sales Data Interpretation

&

Business Recommendations

Dhaval Rathod

Project Overview:

In this SQL project, we will analyze a dataset containing sales details and information. The dataset includes the following columns: Invoice_ID, Branch, City, Customer_type, Gender, Product_line, Unit_price, Quantity, Tax_5, Total, Date, Time, Payment, COGS, Gross_margin_percentage, Gross_income, Rating

We will perform various analyses to gain insights into Customer behavior and product selling patterns.

Step 1: Defining Metadata in MySQL Workbench

We began by defining the metadata for our dataset in MySQL Workbench. We created a table named `walmartsalesdata ` with the columns described above.

Step 2: Exploratory Data Analysis (EDA)

We started by conducting exploratory data analysis to understand the characteristics of the dataset.

1. Which branch generates the highest average revenue per transaction?
2. What is the total revenue generated per month?
3. How does revenue vary across weekdays?
4. What is the average revenue per customer type in each branch?
5. How does sales revenue compare between male and female customers?
6. Which product line has the highest total sales revenue?
7. Which product line has the highest gross income percentage?
8. What is the total quantity sold for each product line?
9. Which product line shows the highest average rating?

10. Which product line is purchased most frequently by Members vs. Normal customers?
11. What is the average revenue per transaction for male vs. female customers?
12. What is the percentage of male vs. female customers for each branch?
13. How does the average quantity purchased vary by customer type?
14. Which product line is preferred by Members vs. Normal customers?
15. What is the repeat purchase behavior for Members?
16. What is the busiest day of the week for each branch?
17. Which hour of the day sees the highest transactions overall?
18. How does revenue vary between morning, afternoon, and evening?
19. What is the peak hour for each product line?
20. How does revenue trend change across the months?
21. Which payment method is most frequently used by Members?
22. What is the average transaction value for each payment method?
23. What is the gross income for each branch?
24. Which product line contributes the most to gross income in each branch?
25. How does the gross margin percentage vary by product line?
26. What is the gross income contribution by Members vs. Normal customers?
27. Which branch has the highest gross income percentage?
28. What is the revenue contribution by each city?
29. Which city shows the highest gross income for male customers?
30. What is the average revenue per transaction for each city?
31. How does customer rating vary across cities?1

32. Which city has the highest percentage of transactions through Credit card?
33. Which product line has the most consistent customer ratings?
34. What is the rating distribution across payment methods?
35. How does average customer rating vary by time of day?
36. How does revenue vary between weekdays and weekends?
37. Which product line shows the highest seasonal variation in sales?.
38. What is the trend of gross income during peak shopping months?
39. What is the return on investment (ROI) for each product line?
40. What is the average transaction size for customers who rate their experience above 9?

Step 3: Business Recommendations :

Let's go through each step of your project task and provide detailed answers:

Step 1: Defining Metadata in MySQL Workbench

For this step, we would need to define the structure of our database table in MySQL Workbench. Here's an example of how we might define the metadata for our table:

Table Name: product_metadata

Columns:

```
Invoice_ID                VARCHAR(50) PRIMARY KEY,
Branch                    CHAR(1) NOT NULL,
City                      VARCHAR(50) NOT NULL,
Customer_type             VARCHAR(20) NOT NULL,
Gender                    VARCHAR(10) NOT NULL,
Product_line              VARCHAR(50) NOT NULL,
Unit_price                DECIMAL(10, 2) NOT NULL,
Quantity                  INT NOT NULL,
Tax_5                     DECIMAL(10, 2) NOT NULL,
Total                     DECIMAL(10, 2) NOT NULL,
Date                      DATE NOT NULL,
Time                      TIME NOT NULL,
Payment                   VARCHAR(20) NOT NULL,
COGS                      DECIMAL(10, 2) NOT NULL,
Gross_margin_percentage DECIMAL(5, 2) NOT NULL,
Gross_income              DECIMAL(10, 2) NOT NULL,
Rating                    DECIMAL(3, 1) NOT NULL
```

Step 2: Exploratory Data Analysis (EDA)

1.Which branch generates the highest average revenue per transaction?

SELECT Branch, AVG(Total) AS AvgRevenue FROM walmartsalesdata GROUP BY Branch ORDER BY AvgRevenue DESC IMIT 1;

Branch	AvgRevenue
C	337.100183

Interpretation: This query calculates the average revenue (Total) for each branch by grouping the data based on the Branch field and which is branch c having AvgRevenue 337.10.

2.What is the total revenue generated per month?

SELECT DATE_FORMAT(Date, '%Y-%m') AS Month, SUM(Total) AS TotalRevenue FROM walmartsalesdata GROUP BY Month ORDER BY Month;

Month	TotalRevenue
2019-01	116292.11
2019-02	97219.58
2019-03	109455.74

Interpretation: This query groups the data by month using the Date field (formatted to year-month) and sums up the total revenue (Total) for each month.

3.How does revenue vary across weekdays? 6.Determine the top 3 most expensive products within each product_category.

SELECT DAYNAME(Date) AS Weekday, SUM(Total) AS TotalRevenue FROM walmartsalesdata GROUP BY Weekday ORDER BY FIELD(Weekday, 'Monday', 'Tuesday', 'Wednesday', 'Thursday', 'Friday', 'Saturday', 'Sunday') DESC;

Weekday	TotalRevenue
Sunday	44458.02
Saturday	56120.86
Friday	43926.43
Thursday	45349.34
Wednesday	43731.24
Tuesday	51482.39
Monday	37899.15

Interpretation: This query aggregates revenue by weekday using the Date field and calculates the total revenue (Total) for each day of the week across all weekdays Sunday generates highest total revenue which is 44458.02.

4.What is the average revenue per customer type in each branch?

SELECT Branch, Customer_type, AVG(Total) AS AvgRevenue FROM walmartsalesdata GROUP BY Branch, Customer_type ORDER BY Branch, AvgRevenue DESC;

	Branch	Customer_ty...	AvgRevenue
	A	Member	321.183054
	A	Normal	303.832370
	B	Member	325.483818
	B	Normal	314.330359
	C	Normal	337.656918
	C	Member	336.576391

Interpretation: This query calculates the average revenue (Total) per transaction for each Customer_type in each branch. FROM all three branch highest totalrevenie generates branch c and in customer type Normal .

5. How does sales revenue compare between male and female customers?

SELECT Gender, SUM(Total) AS TotalRevenue FROM walmartsalesdata GROUP BY Gender;

	Gender	TotalRevenue
	Male	155084.17
	Female	167883.26

Interpretation: This query groups the data by gender (Gender) and sums the total revenue (Total) for male and female customers.Female Gender Generates Hghest total rrevenue rather than male gender.

6. Which product line has the highest total sales revenue?

SELECT Product_line, SUM(Total) AS TotalRevenue FROM walmartsalesdata GROUP BY Product_line ORDER BY TotalRevenue DESC LIMIT 1;

	Product_line	TotalRevenue
	Food and beverages	56144.96

Interpretation: This query sums the total revenue (Total) for each product line (Product_line) and identifies the product line with the highest revenue and it gives foods and baverages product_line with total_REVENUE 56144.95..

7. Which product line has the highest gross income percentage?

SELECT Product_line, (SUM(Gross_income) / SUM(Total)) * 100 AS GrossIncomePercentage FROM walmartsalesdata GROUP BY Product_line ORDER BY GrossIncomePercentage DESC LIMIT 1;

Product_line	GrossIncomePercenta...
Sports and travel	4.762205

Interpretation: This query calculates the gross income percentage for each product line by dividing the gross income (Gross_income) by the total revenue (Total) and multiplying by 100 in that we got Sports and travel peoduct line with 4.76 % of grossincomepercentage ..

8. What is the total quantity sold for each product line?

SELECT Product_line, SUM(Quantity) AS TotalQuantity FROM walmartsalesdata GROUP BY Product_line ORDER BY TotalQuantity DESC;

Product_line	TotalQuanti...
Electronic accessories	971
Food and beverages	952
Sports and travel	920
Home and lifestyle	911
Fashion accessories	902
Health and beauty	854

Interpretation: This query sums the quantity (Quantity) sold for each product line (Product_line)..

9. Which product line shows the highest average rating?**SELECT**

SELECT Product_line, AVG(Rating) AS AvgRating FROM walmartsalesdata GROUP BY Product_line ORDER BY AvgRating DESC LIMIT 1;

Product_line	AvgRating
Food and beverages	7.11322

Interpretation: This query calculates the average rating (Rating) for each product line and identifies the product line with the highest average rating from this we got the food and beverages product line with avgrating having 7.11.

10. Which product line is purchased most frequently by Members vs. Normal customers?

SELECT Customer_type, Product_line, COUNT(*) AS PurchaseFrequency FROM walmartsalesdata GROUP BY Customer_type, Product_line ORDER BY Customer_type, PurchaseFrequency DESC;

	Customer_ty...	Product_line	PurchaseFrequency
	Member	Food and beverages	94
	Member	Sports and travel	87
	Member	Fashion accessories	86
	Member	Home and lifestyle	83
	Member	Electronic accessories	78
	Member	Health and beauty	73
	Normal	Electronic accessories	92
	Normal	Fashion accessories	92
	Normal	Food and beverages	80
	Normal	Sports and travel	79
	Normal	Health and beauty	79
	Normal	Home and lifestyle	77

Interpretation: This query counts the number of transactions for each product line (Product_line) and customer type (Customer_type),and it seems like that Food and Beverages product_line holds highest purchaseFequency that is 94 From Member Customer type and in Normal customer type there is simmilar purchasefrequency 92 in Fashion accesories and Electronic Accecories.

11. What is the average revenue per transaction for male vs. female customers.

SELECT Gender, AVG(Total) AS AvgRevenue FROM walmartsalesdata GROUP BY Gender;

	product_category	avg_stock	max_stock	min_stock	stock_stat...
	Circuit Protection Accessories	798	798	798	Overstocked
	RF Transceiver Modules and Modems	572	572	572	Overstocked
	Through Hole Resistors	500	500	500	Overstocked
	Aluminum Electrolytic Capacitors	478	478	478	Overstocked
	Switches (Solid State)	264.75	975	2	Overstocked
	Motors - AC, DC	190	190	190	Overstocked
	Circular Connector Assemblies	123.5068493150685	99	1	Overstocked
	Voltage Regulators - Linear, Low Drop Out (LD...	104.49056603773585	98	0	Overstocked
	Supervisors	97.71428571428571	99	100	Overstocked
	SAW Filters	91.66666666666667	8	10	Overstocked

Interpretation: This query calculates the average revenue (Total) for transactions made by male and female customers in that again female gender generates more revenue than male gender which is 335.10.

12. What is the percentage of male vs. female customers for each branch?SELECT

SELECT Branch, Gender, COUNT(*) * 100.0 / SUM(COUNT(*)) OVER (PARTITION BY Branch) AS Percentage FROM walmartsalesdata GROUP BY Branch, Gender;

	Branch	Gender	Percentage
	A	Male	52.64706
	A	Female	47.35294
	B	Female	48.79518
	B	Male	51.20482
	C	Female	54.26829
	C	Male	45.73171

Interpretation: This query calculates the percentage of male and female customers in each branch by dividing the count of male and female customers by the total count of customers in each branch A and B Male gender percentage is greater than female in branch A male percentage is 52.64 ,In Branch B Male percentage is 51.20 Nut In Branch C there Is highest Percentage that every Branch Which Females Holds and it is 54.2.

13. How does the average quantity purchased vary by customer type?

SELECT Customer_type, AVG(Quantity) AS AvgQuantity FROM walmartsalesdata GROUP BY Customer_type ORDER BY AvgQuantity DESC;

	Customer_ty...	AvgQuantity
	Member	5.5589
	Normal	5.4609

Interpretation: This query calculates the average quantity (Quantity) purchased by each customer type (Customer_type) which is almost same .

14. Which product line is preferred by Members vs. Normal customers?;

SELECT Customer_type, Product_line, SUM(Quantity) AS TotalQuantity FxROM walmartsalesdata GROUP BY Customer_type, Product_line ORDER BY Customer_type, TotalQuantity DESC;

Customer_ty...	Product_line	TotalQuanti...
Member	Food and beverages	506
Member	Sports and travel	493
Member	Home and lifestyle	490
Member	Fashion accessories	439
Member	Electronic accessories	429
Member	Health and beauty	428
Normal	Electronic accessories	542
Normal	Fashion accessories	463
Normal	Food and beverages	446
Normal	Sports and travel	427
Normal	Health and beauty	426
Normal	Home and lifestyle	421

Interpretation: This query calculates the total quantity of each product line purchased by each customer type (Customer_type) in that Member customer_type has highest total quantity in food and beverages which is 506 and in normal customer type their is 562 totalqantity in Electronic accories...

15. What is the repeat purchase behavior for Members?

SELECT Customer_type, COUNT(DISTINCT(Date)) AS PurchaseDays, COUNT(*) AS TotalTransactions FROM walmartsalesdata WHERE Customer_type = 'Member' GROUP BY Customer_type;

Customer_ty...	PurchaseDays	TotalTransactio...
Member	89	501

Interpretation: This query counts the number of distinct purchase days (Date) and total transactions for Members (Customer_type = 'Member') which is 89 perchasedays and having totaltransactions 501.

16. What is the busiest day of the week for each branch?

```
SELECT DAYNAME(Date) AS weekdays,Branch ,COUNT(*) AS transactions FROM
walmartsalesdata GROUP BY Branch,weekdays ORDER BY transactions DESC;
```

Interpretation:It seems like that weekday saturday is a busiest day in week and it has 60 transactions and it belongs to Branch B , after that in Branch C Tuesday has Seccond Busiest weekday and it has total 54 trasactions, at last branch A comes third with day sundayn having 52 transactions Between Branch B and A their is not much diffrence.

17. Which hour of the day sees the highest transactions overall?

SELECT HOUR(Time) AS hours, COUNT(*) AS ransactions FROM walmartsalesdata GROUP BY hours ORDER BY ransactions DESC LIMIT 1;

Interpretaton: In this querie it returns highest count of all transactions along with hour using Group by we get each hours total count of transactions and along with that also done order by with count of transaction in descending with limit 1 set for only getting highest count of transaction in hour which is hour 19 with having total transaction 113.18. How does revenue trend change across the months?

18. How does revenue vary between morning, afternoon, and evening?

SELECT CASE WHEN HOUR(Time) BETWEEN 6 AND 12 THEN "Morning" WHEN HOUR(Time) BETWEEN 12 AND 17 THEN "Afternoon" WHEN HOUR(Time) BETWEEN 18 AND 123 THEN "Evening" ELSE "Night" END AS timeofday,SUM(Total) AS Revenue FROM walmartsalesdata GROUP BY timeofday ORDER BY timeofday ASC;

timeofday	Revenue
Afternoon	146402.99
Evening	88699.50
Morning	87864.94

Interpretation:In this query it returns the time hour ussing hour functionn and using case it gives Morning,afternoon and evening total revenue ,in afternoo their is highest revenue which is 1146402.99 then Evening with revenue 88699.50 and morning generates revenue 87864.94 which is good in compare to Evening revenue cause their is not much diffrence among both.

19. What is the peak hour for each product line?

WITH Hourlysale AS(SELECT Product_line,HOUR(Time) AS Hour,SUM(Total) AS Revenue FROM walmartsalesdata GROUP BY Product_lie,Hour),Rankedsales AS (SELECT Product_line,Hour,Revenue,RANK()OVER(PARTITION BY Product_line ORDER BY Revenue DESC) AS Renk FROM Hourlysale)SELECT Product_line,Hour AS Peak_Hour ,Revenue AS Peak_Revenue FROM Rankedsales EHERE Renk=1;

Interpretaton:this query i used Comman Table Execution to get First table Hourlysales In which i get Product_line,Hour and reveune and group by product_line and hour ,using cte i created one more cte table Renkesales based on Hourlysales Where i retrives product_line Hour as Peak_hour,Reveue and RANK Over partition by product_line ORDER VY Revenue AS renk Then from Renksales I gets all three required feilds and give conditionof renk=1 so i got ....

20. How does revenue trend change across the months?

SELECT MONTH(Date) AS Month, SUM(Total) AS Revenue FROM walmartsalesdata GROUP BY Month ORDER BY Revenue DESC;

Month	Revenue
1	116292.11
3	109455.74
2	97219.58

Interpretation:From above we can observe that Month 1 (January) hasHighest revenue which is 1116292.11 The Month 23 (March) is pretty good ooking for our company Because in Month 1 and 3 their notmuch diffrence Month 3 has total revenue 109455.74 ,In month 2 there is lowest revenue in among all 3 months Month 2 is month of intrest if we want to maximize our revenue and expand our buisness.

21. Which payment method is most frequently used by Members?

```
SELECT  Payment AS Payment,    COUNT(*) AS Frequency FROM
walmartsalesdataWHERE  Customer_type = 'Member' GROUP BY  Payment ORDER BY
Frequency DESC;
```

Payment	Frequency
Credit card	172
Cash	168
Ewallet	161

Interpretation :We can observethat the most frequet paymenrt type is Credit Card used by members which is 172 seccond frequent paymrnt type is Cash and its frequency is 168 and at last Ewallet paymentt type is last and its freuency is 161

22. What is the average transaction value for each payment method?;

SELECT AVG(Total) AS transaction , Payment FROM walmartsalesdata GROUP BY Payment;

transaction	Payment
324.010579	Credit card
318.821391	Ewallet
326.182442	Cash

Interpretation;From this we can see that the Credit cart payment method have 324.01 average_transaction value which is highest among all the payment methods and second highest is Cash payment method with holding 328.18 average transaction value and lasly Ewallet payment method holding 318.82.

23. What is the gross income for each branch?;

```
SELECT Branch,SUM(Gross_income) FROM walmartsalesdata GROUP BY Branch;
```

Branch	SUM(Gross_incom...
A	5057.36
C	5265.33
B	5057.36

Interpretation:We can see that this query gives the each branch Gross income in that Branch C Geneartes highest Gross income in among all branches which is 5265.33 where Branch A and B have same gross income which is 5057.36.

24.Which product line contributes the most to gross income in each branch?;

```
SELECT Product_line ,SUM(Gross_income) AS gross_income FROM walmartsalesdata GROUP BY Product_line Order by gross_income DESC;
```

Product_line	gross_income
Food and beverages	2673.68
Sports and travel	2625.07
Electronic accessories	2587.61
Fashion accessories	2586.13
Home and lifestyle	2564.90
Health and beauty	2342.66

Interpretation ;Foods and Beverages product line contributes most to the gross income in among all the product line also there is not much diffrence in sports and travels product line with compare to foods and baverages product-line

25.How does the gross margin percentage vary by product line?;

```
SELECT Product_line,AVG(Gross_margin_percentage )  AS Gross_margin_percentage FROM walmartsalesdata GROUP BY Product_line ORDER BY Gross_margin_percentage DESC;
```

Product_line	Gross_margin_percenta...
Food and beverages	4.761900000
Health and beauty	4.761900000
Sports and travel	4.761900000
Fashion accessories	4.761900000
Home and lifestyle	4.761900000
Electronic accessories	4.761900000

Interpretation:We can see that all Product_line have same gross_average_margin_persentage Which is 4.76;

26.What is the gross income contribution by Members vs. Normal customers?;

SELECT Customer_type ,SUM(Gross_income) FROM walmartsalesdata GROUP BY Customer_type ;

Customer_ty...	SUM(Gross_incom...
Normal	7559.52
Member	7820.53

Interpretation:We can observe that in both the customer type their is not much diffrence in between them but still Member Customer type leads in contributioniin Gross Income with having 7820.53 Gross Income BUT also Noramal customer type is pretty good as cmpare to member customer type it also holds 7559.52 of total Gross Income it says that their not as much of impact of membership cudstomer type .

27Which branch has the highest gross income percentage?;

SELECT Branch,SUM(Gross_margin_percentage) AS highest_gross_income_percentage FROM walmartsalesdata GROUP BY Branch;

Branch	highest_gross_income_percent...
A	1619.04600
C	1561.90320
B	1580.95080

Interpretation: We can observe that accross all Branches heir is Branch A holding strong possition in Highest Gross Income percentage which is 1619.04 also thir is not as much diffrence between branch b and c .

28.What is the revenue contribution by each city?;

SELECT SUM(Total) AS Revenue ,City FROM walmartsalesdata GROUP BY City ;

Revenue	City
106200.57	Yangon
110568.86	Naypyitaw
106198.00	Mandalay

Interpretation:From this query we get the revenue contribution by each city,city Neytfaw holds highest revenue accross all city which is 110568.86.also their is not as much diffrence in between Mandalaly and Yangon city their is nominal diffrence.

29.Which city shows the highest gross income for male customers?;

SELECT City,SUM(Gross_income) FROM walmartsalesdata WHERE Gender="male" GROUP BY City ;

City	SUM(Gross_incom...
Yangon	2520.66
Naypyitaw	2327.83
Mandalay	2536.80

Interpretation:Among city Which city generates highest gross income for male customers and that city is Yangon with having 2520.66 Gross income.

30.What is the average revenue per transaction for each city?;

SELECT City,AVG(Total) AS average_revenue ,COUNT(*) AS Transactions FROM walmartsalesdata GROUP BY City ORDER BY average_revenue;

City	average_revenue	Transactions
Yangon	312.354618	340
Mandalay	319.873494	332
Naypyitaw	337.100183	328

Interpretation: Naypyitaw leads in average revenue per transaction, which may indicate customers there make larger purchases per transaction compared to other cities. Yangon has the highest transaction volume but the lowest average revenue.

31.How does customer rating vary across cities?;

SELECT City,AVG(Rating) AS average_raing FROM walmartsalesdata GROUP BY City ;

City	average_raing
Yangon	7.02706
Naypyitaw	7.07287
Mandalay	6.81807

Interpretation;Yangon leads in avreage ratingd ,which idicates ighest average customer rating comes from City yangon among all cities.In Naypyaitaw and Yangon their is not more diffrence in between them. But City Mandlay is having lowest average rating.

32.Which city has the highest percentage of transactions through Credit card?;

SELECT City,MAX(Gross_margin_percentage) AS highest_percentage FROM walmartsalesdata WHERE Payment="Credit Card" GROUP BY City;

City	highest_percenta...
Yangon	4.76190
Naypyitaw	4.76190
Mandalay	4.76190

Interpretation:Gross income percentage is same for all city their is no highest in between among all cities.

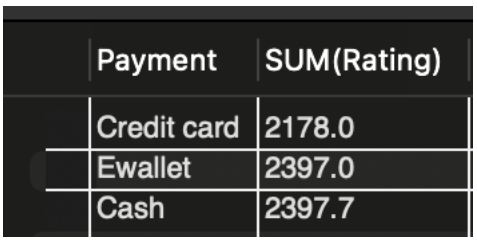
33.Which product line has the most consistent customer ratings?;

SELECT Product_line,COUNT(Rating) AS consistent_customer_rating FROM walmartsalesdata GROUP BY Product_line;

Interpretaion:The query groups the data by Product_line and counts the number of customer ratings for each category. The results show that Fashion accessories received the highest number of ratings at 178, indicating strong customer engagement and popularity in this product category. Following closely, Electronic accessories had 170 ratings, and Food and beverages received 174 ratings, both reflecting significant customer interest. On the other hand, Health and beauty had the lowest number of ratings at 152, suggesting comparatively lower customer engagement or purchase frequency in this category. Categories like Sports and travel (166 ratings) and Home and lifestyle (160 ratings) demonstrate moderate customer interaction. These findings highlight opportunities to further promote high-performing categories like Fashion Accessories and Electronic Accessories while implementing strategies, such as targeted promotions or visibility improvements, to boost engagement in underperforming segments like Health and Beauty. This approach can help optimize customer interest and drive overall sales growth.

34.What is the rating distribution across payment methods?;

SELECT Payment,SUM(Rating) FROM walmartsalesdata GROUP BY Payment;



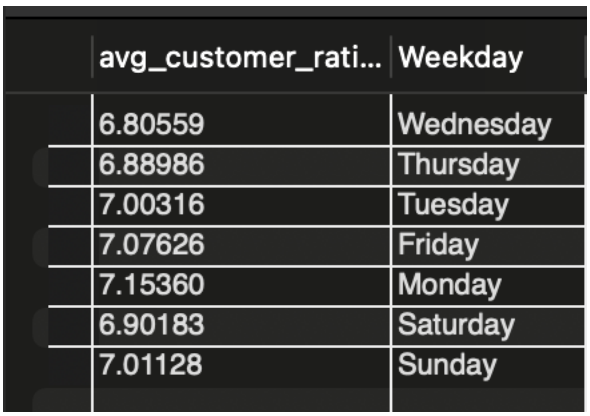
Payment	SUM(Rating)
Credit card	2178.0
Ewallet	2397.0
Cash	2397.7

Interpretation:The query groups the data by the Payment method and calculates the total sum of ratings for each group. The results show that Cash payments have the highest total rating of 2397.7, closely followed by Ewallet payments with a total rating of 2397.0. Credit card payments, however, have the lowest total rating at 2178.0. This indicates that customers using Cash and Ewallet methods tend to rate

their experiences slightly higher compared to those using credit cards. The close values for Cash and Ewallet suggest that these payment methods are preferred by customers, potentially due to their convenience or alignment with customer habits. Businesses could explore further strategies to encourage credit card usage while continuing to optimize the customer experience for Cash and Ewallet payments.

35.How does average customer rating vary by time of day?;

SELECT AVG(Rating) AS avg_customer_rating ,DAYNAME(Date) AS Weekday FROM walmartsalesdata GROUP BY weekday ;



avg_customer_rati...	Weekday
6.80559	Wednesday
6.88986	Thursday
7.00316	Tuesday
7.07626	Friday
7.15360	Monday
6.90183	Saturday
7.01128	Sunday

Interpretation:The SQL query calculates the average customer rating for each day of the week by extracting the weekday names from the date column and grouping the data accordingly. The results reveal that Monday has the highest average customer rating at 7.15360, followed closely by Friday with 7.07626. Ratings remain relatively steady on Sunday and Tuesday, with values of 7.01128 and 7.00316, respectively. Thursday and Saturday show slightly lower ratings, at 6.88986 and 6.90183, while Wednesday records the lowest average rating of 6.80559. Overall, the customer ratings tend to peak at the start of the week and show minor variations throughout, with a noticeable dip mid-week on Wednesday.

36.How does revenue vary between weekdays and weekends?;

SELECT SUM(Total) AS Revenue ,CASE WHEN DAYNAME(Date)IN("Monday","Tuesday","Wesdnesday","Tuesday","Friday") THEN "WEEKDAY" ELSE "WEEKEND" END AS WEEK FROM walmartsalesdata GROUP BY WEEK ;

Interpretation:The query calculates total revenue by grouping days into WEEKDAY (Monday to Friday) and WEEKEND (Saturday and Sunday). The results show that WEEKEND revenue is 189,659.46, significantly higher than WEEKDAY revenue of 133,307.97, indicating increased sales activity during weekends.

37.Which product line shows the highest seasonal variation in sales?.;

SELECT CASE WHEN EXTRACT(MONTH FROM Date) IN(1,2,3) THEN "Q1" WHEN EXTRACT(MONTH FROM Date) IN(4,5,6) THEN "Q2" WHEN EXTRACT(MONTH FROM Date) IN(7,8,9) THEN "Q3" ELSE "Q4" END AS Quater ,Product_line,SUM(Total) AS Revenue FROM walmartsalesdata GROUP BY Quater,Product_line ORDER BY Revenue DESC;

Interpretation:This query firstly gets quater months ussing case then gets product line and group by product line and quater to get revenue their is onu quater 1 and among all Product line Foods and baverages Generates highest revenue (56144.96).

38.What is the trend of gross income during peak shopping months?;

SELECT MONTHNAME(Date) AS Month,YEAR(Date) AS Year,SUM(Gross_income) AS Total_Revenue FROM walmartsalesdata GROUP BY Year, Month ORDER BY Total_Revenue DESC;

Interpretation:The query extracts month and year from Date column and calculates gross income Group by year and month to get peakmonth and the peakmonth is january generates highest totalGross Income (5537.95.

39.What is the return on investment (ROI) for each product line?;

SELECT Product_line,SUM(Gross_income) AS Total_profit ,SUM(COGS) AS total_cogs ,(SUM(Gross_income)/SUM(COGS))*100 AS ROI FROM walmartsalesdata GROUP BY Product_line ORDER BY ROI DESC;

Product_line	Total_pro...	total_cogs	ROI
Sports and travel	2625.07	52497.93	5.000330
Fashion accessories	2586.13	51719.90	5.000261
Food and beverages	2673.68	53471.28	5.000217
Health and beauty	2342.66	46851.18	5.000216
Electronic accessories	2587.61	51750.03	5.000210
Home and lifestyle	2564.90	51297.06	5.000092

Interpretation:

The query analyzes financial performance by calculating the Total Profit, Total COGS (Cost of Goods Sold), and Return on Investment (ROI) for each product line, grouping and sorting them in descending order of ROI. The results show that Sports and Travel leads with the highest ROI of 5.000330, followed closely by Fashion Accessories at 5.000261 and Food and Beverages at 5.000217. Health and Beauty and Electronic Accessories have slightly lower ROIs at 5.000216 and 5.000210, respectively. Home and Lifestyle ranks last with an ROI of 5.000092. While the differences in ROI across product lines are minimal, Sports and Travel achieves the best return relative to its COGS, indicating marginally higher profitability efficiency compared to the other categories

40.What is the average transaction size for customers who rate their experience above 9?;

SELECT AVG(Total) AS Avg_transaction FROM walmartsalesdata WHERE Rating>9;

Step 3: Business Recommendations:

1. Product Line Performance

Observation: The **Food and Beverages** product line generates the highest total revenue at **$56,144.95**, while **Sports and Travel** leads in ROI with **5.0%**.
Recommendation: Walmart should prioritize **inventory optimization**, promotions, and marketing campaigns for these high-performing product lines. Focus on stock availability during peak sales periods, particularly in months like **January**, where revenues are at their highest. Additionally, analyze customer preferences within these categories to further personalize product offerings.

2. Branch Performance

Observation: Branch **C** outperforms others, generating the highest gross income at **$5,265.33**, while Branches **A** and **B** show similar but slightly lower gross income values of **$5,057.36** each.
Recommendation: To boost performance at Branches **A** and **B**, Walmart should identify **underperforming product lines** and introduce branch-specific promotions or discount strategies. Consider using localized customer segmentation to tailor offers for both **member** and **non-member** shoppers.

3. Customer Demographics: Gender Insights

Observation: Female customers contribute a higher average revenue per transaction compared to male customers.
Recommendation: Walmart should design **gender-specific marketing campaigns** that cater to female shoppers. Promotions around products preferred by this demographic, such as **Food and Beverages** or **Fashion Accessories**, could further increase engagement and revenue.

4. Time-Based Revenue Trends

Observation: Revenue is significantly higher on **weekends** (**$189,659.46**) compared to weekdays (**$133,307.97**). Additionally, the **afternoon hours** generate the highest revenue.
Recommendation: Walmart should optimize **staffing schedules** and **inventory levels** during weekends and afternoons to handle increased customer traffic. Launch time-sensitive promotions to attract more customers during these peak periods.

5. Seasonal Trends

Observation: **January** and **March** exhibit the highest sales trends, with January contributing **$1,116,292.11** in revenue.
Recommendation: Walmart should focus on **seasonal stock preparation** and **promotional campaigns** at the beginning of the year. Highlight key product lines like **Food and Beverages** and **Sports and Travel** during these months to maximize revenue.

6. Repeat Purchase Behavior for Members

Observation: Members show strong repeat purchase behavior with **89 distinct purchase days** and **501 total transactions**.
Recommendation: Walmart should continue to enhance the **membership program** by offering exclusive benefits, personalized discounts, and loyalty rewards. This will further strengthen member retention and encourage repeat purchases.

7. Revenue Contribution by City

Observation: The city of **Naypyitaw** contributes the highest average revenue per transaction, whereas **Yangon** has the **highest transaction volume** but a lower average revenue.
Recommendation: Walmart should analyze customer purchasing behavior in **Yangon** to increase the transaction size through strategies like product bundling or upselling. For **Naypyitaw**, Walmart can focus on maintaining customer loyalty and expanding its product offerings to sustain high revenue per transaction.

Conclusion: The analysis of Walmart's sales data highlights several key insights and actionable recommendations. The Food and Beverages product line emerges as the most profitable, generating the highest revenue, while Sports and Travel leads in return on investment (ROI). Branch C outperforms others in gross income, whereas Branches A and B have potential for growth through targeted promotions. Female customers contribute higher average revenue per transaction, underscoring the need for tailored marketing campaigns to engage this demographic. Revenue trends reveal that weekends and afternoon hours are peak sales periods, presenting opportunities for time-sensitive promotions and operational optimization. Additionally, January and March show significant seasonal peaks, making them critical periods for stock preparation and marketing efforts. Among cities, Naypyitaw leads in average revenue per transaction, while Yangon records the highest transaction volume, indicating the need for city-specific strategies. Moreover, members demonstrate strong repeat purchase behavior, reinforcing the importance of loyalty programs. By focusing on these insights, Walmart can optimize inventory, launch strategic campaigns, and tailor customer experiences to drive profitability and sustain growth.