WALMART ✳

SALES ANALYSIS

(using MySQL)

To analyze and derive insights from Walmart's sales data to improve business decision-making processes.

by Komal Nama
Batch Name: MIP-DA-13

OBJECTIVE

The objective of this analysis is to provide a comprehensive overview of Walmart's sales performance across various regions and product categories. The insights derived will guide strategic decision-making to optimize inventory management, enhance customer satisfaction, and drive revenue growth. Additionally, the analysis seeks to uncover opportunities for market expansion and product diversification, ultimately contributing to Walmart's overall business objectives and long-term profitability. The Problem statement includes 14 questions for analysis.

DATASET

The dataset consists of the following columns:

- **Invoice_ID:** Unique identifier for each invoice.
- **Branch:** Branch of the store where the sale took place.
- **City:** City where the store branch is located.
- **Customer_type:** Type of customer (e.g., Member, Normal).
- **Gender:** Gender of the customer.
- **Product_line:** Category of the product sold.
- **Unit_price:** Price per unit of the product.
- **Quantity:** Quantity of the product sold.
- **Tax_5%:** Tax applied on the sale.
- **Total:** Total amount of the sale.
- **Date:** Date of the transaction.
- **Time:** Time of the transaction.
- **Payment:** Payment method used (e.g., Cash, Credit Card).
- **cogs:** Cost of goods sold.
- **gross_margin_percentage:** Gross margin percentage.
- **gross_income:** Gross income from the sale.
- **Rating:** Customer rating for the purchase.

Problem Statements

1. Retrieve all columns for sales made in a specific branch (e.g., Branch 'A').

2. Find the total sales for each product line.

3. List all sales transactions where the payment method was 'Cash'.

4. Calculate the total gross income generated in each city.

5. Find the average rating given by customers in each branch.

6. Determine the total quantity of each product line sold.

7. List the top 5 products by unit price.

8. Find sales transactions with a gross margin percentage greater than 30%.

9. Retrieve sales transactions that occurred on weekends.

10. Calculate the total sales and gross income for each month.

11. Find the number of sales transactions that occurred after 6 PM.

12. List the sales transactions that have a higher total than the average total of all transactions.

13. Calculate the cumulative gross income for each branch by date.

14. Find the total cogs for each customer type in each city.

Q1. Retrieve all columns for sales made in a specific branch (e.g., Branch 'A').

```
SELECT *
FROM walmart_sales_data.walmartdata
WHERE Branch = 'A';
```

↓

Invoice_ID	Branch	City	Customer_type	Gender	Product_line	Unit_price	Quantity	Tax_5%	Total	Date	Time
750-67-8428	A	Yangon	Member	Female	Health and beauty	74.69	7	26.1415	548.9715	05-01-2019	13:08:00
631-41-3108	A	Yangon	Normal	Male	Home and lifestyle	46.33	7	16.2155	340.5255	03-03-2019	13:23:00
123-19-1176	A	Yangon	Member	Male	Health and beauty	58.22	8	23.288	489.048	27-01-2019	20:33:00
373-73-7910	A	Yangon	Normal	Male	Sports and travel	86.31	7	30.2085	634.3785	08-02-2019	10:37:00
355-53-5943	A	Yangon	Member	Female	Electronic accessories	68.84	6	20.652	433.692	25-02-2019	14:36:00
665-32-9167	A	Yangon	Member	Female	Health and beauty	36.26	2	3.626	76.146	10-01-2019	17:15:00
365-64-0515	A	Yangon	Normal	Female	Electronic accessories	46.95	5	11.7375	246.4875	12-02-2019	10:25:00
252-56-2699	A	Yangon	Normal	Male	Food and beverages	43.19	10	21.595	453.495	07-02-2019	16:48:00
829-34-3910	A	Yangon	Normal	Female	Health and beauty	71.38	10	35.69	749.49	29-03-2019	19:21:00
656-95-9349	A	Yangon	Member	Female	Health and beauty	68.93	7	24.1255	506.6355	11-03-2019	11:03:00
765-26-6951	A	Yangon	Normal	Male	Sports and travel	72.61	6	21.783	457.443	01-01-2019	10:39:00
329-62-1586	A	Yangon	Normal	Male	Food and beverages	54.67	3	8.2005	172.2105	21-01-2019	18:00:00

Q2. Find the total sales for each product line.

```
SELECT Product_line, SUM(Total) as Total_Sales
FROM walmartdata
GROUP BY Product_line;
```

Product_line	Total_Sales
Health and beauty	49193.73900000016
Electronic accessories	54337.531500000005
Home and lifestyle	53861.9130000001
Sports and travel	55122.826499999996
Food and beverages	56144.844000000005
Fashion accessories	54305.895

Q3. List all sales transactions where the payment method was 'Cash'.

```
SELECT *
FROM walmartdata
WHERE Payment = 'Cash';
```

Invoice_ID	Branch	City	Customer_type	Gender	Product_line	Unit_price	Quantity	Tax_5%	Total	Date	Time	Payment	cogs	gross_
226-31-3081	C	Naypyitaw	Normal	Female	Electronic accessories	15.28	5	3.82	80.22	08-03-2019	10:29:00	Cash	76.4	4.76190
529-56-3974	B	Mandalay	Member	Male	Electronic accessories	25.51	4	5.102	107.142	09-03-2019	17:03:00	Cash	102.04	4.76190
829-34-3910	A	Yangon	Normal	Female	Health and beauty	71.38	10	35.69	749.49	29-03-2019	19:21:00	Cash	713.8	4.76190
299-46-1805	B	Mandalay	Member	Female	Sports and travel	93.72	6	28.116	590.436	15-01-2019	16:19:00	Cash	562.32	4.76190
649-29-6775	B	Mandalay	Normal	Male	Fashion accessories	33.52	1	1.676	35.196	08-02-2019	15:31:00	Cash	33.52	4.76190
145-94-9061	B	Mandalay	Normal	Female	Food and beverages	88.36	5	22.09	463.89	25-01-2019	19:48:00	Cash	441.8	4.76190
848-62-7243	A	Yangon	Normal	Male	Health and beauty	24.89	9	11.2005	235.2105	15-03-2019	15:36:00	Cash	224.01	4.76190
149-71-6266	B	Mandalay	Member	Male	Sports and travel	78.07	9	35.1315	737.7615	28-01-2019	12:43:00	Cash	702.63	4.76190
640-49-2076	B	Mandalay	Normal	Male	Sports and travel	83.78	8	33.512	703.752	10-01-2019	14:49:00	Cash	670.24	4.76190
777-82-7220	B	Mandalay	Member	Male	Home and lifestyle	30.12	8	12.048	253.008	03-03-2019	13:01:00	Cash	240.96	4.76190
554-53-8700	C	Naypyitaw	Member	Male	Home and lifestyle	56.11	2	5.611	117.831	02-02-2019	10:11:00	Cash	112.22	4.76190
354-25-5821	B	Mandalay	Member	Female	Sports and travel	69.12	6	20.736	435.456	08-02-2019	13:03:00	Cash	414.72	4.76190

Q4. Calculate the total gross income generated in each city.

```
SELECT City, SUM(gross_income) as Total_Gross_Income
FROM walmartdata
GROUP BY City;
```

City	Total_Gross_Income
Yangon	5057.16050000002
Naypyitaw	5265.17650000002
Mandalay	5057.03200000003

Q5. Find the average rating given by customers in each branch.

SELECT Branch, AVG(Rating) as AVG_Rating
FROM walmartdata
GROUP BY Branch;

	Branch	AVG_Rating
▶	A	7.02705882352941
	C	7.07286585365853
	B	6.81807228915662

Q6. Determine the total quantity of each product line sold.

```
SELECT Product_line, SUM(Quantity) as Total_Quantity_Sold
FROM walmartdata
GROUP BY Product_line;
```

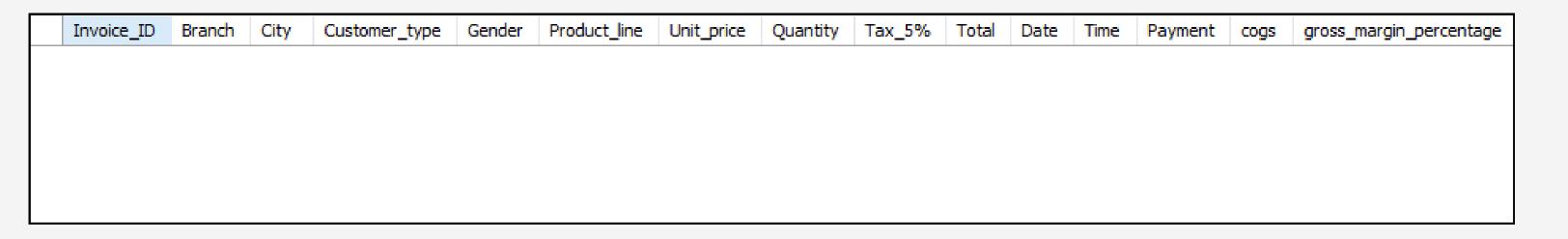| Product_line	Total_Quantity_Sold
Health and beauty	854
Electronic accessories	971
Home and lifestyle	911
Sports and travel	920
Food and beverages	952
Fashion accessories	902

Q7. List the top 5 products by unit price.

```
SELECT *
FROM walmartdata
ORDER BY UNIT_PRICE DESC
LIMIT 5;
```

Invoice_ID	Branch	City	Customer_type	Gender	Product_line	Unit_price	Quantity	Tax_5%	Total	Date	Time	Payment	cogs	gross_marg
148-41-7930	C	Naypyitaw	Normal	Male	Health and beauty	99.96	7	34.986	734.706	23-01-2019	10:33:00	Cash	699.72	4.76190476
219-22-9386	B	Mandalay	Member	Male	Sports and travel	99.96	9	44.982	944.622	09-03-2019	17:26:00	Credit card	899.64	4.76190476
641-62-7288	B	Mandalay	Normal	Male	Home and lifestyle	99.92	6	29.976	629.496	24-03-2019	13:33:00	Ewallet	599.52	4.76190476
437-53-3084	B	Mandalay	Normal	Male	Fashion accessories	99.89	2	9.989	209.769	26-02-2019	11:48:00	Ewallet	199.78	4.76190476
667-92-0055	A	Yangon	Member	Male	Health and beauty	99.83	6	29.949	628.929	04-03-2019	15:02:00	Ewallet	598.98	4.76190476

Q8. Find sales transactions with a gross margin percentage greater than 30%.

```
SELECT *
FROM walmartdata
WHERE gross_margin_percentage > 30;
```

Invoice_ID	Branch	City	Customer_type	Gender	Product_line	Unit_price	Quantity	Tax_5%	Total	Date	Time	Payment	cogs	gross_margin_percentage

**No such transactions where gross_margin_percentage is greater than 30

Q9. Retrieve sales transactions that occurred on weekends.

```
SELECT *
FROM walmartdata
WHERE DAYOFWEEK(STR_TO_DATE(Date, '%d-%m-%Y')) IN (1, 7);
```

	Branch	City	Customer_type	Gender	Product_line	Unit_price	Quantity	Tax_5%	Total	Date	Time	Payment	cogs	gros
750-67-8428	A	Yangon	Member	Female	Health and beauty	74.69	7	26.1415	548.9715	05-01-2019	13:08:00	Ewallet	522.83	4.761
631-41-3108	A	Yangon	Normal	Male	Home and lifestyle	46.33	7	16.2155	340.5255	03-03-2019	13:23:00	Credit card	324.31	4.761
123-19-1176	A	Yangon	Member	Male	Health and beauty	58.22	8	23.288	489.048	27-01-2019	20:33:00	Ewallet	465.76	4.761
315-22-5665	C	Naypyitaw	Normal	Female	Home and lifestyle	73.56	10	36.78	772.38	24-02-2019	11:38:00	Ewallet	735.6	4.761
529-56-3974	B	Mandalay	Member	Male	Electronic accessories	25.51	4	5.102	107.142	09-03-2019	17:03:00	Cash	102.04	4.761
636-48-8204	A	Yangon	Normal	Male	Electronic accessories	34.56	5	8.64	181.44	17-02-2019	11:15:00	Ewallet	172.8	4.761
549-59-1358	A	Yangon	Member	Male	Sports and travel	88.63	3	13.2945	279.1845	02-03-2019	17:36:00	Ewallet	265.89	4.761
189-17-4241	A	Yangon	Normal	Female	Fashion accessories	87.67	2	8.767	184.107	10-03-2019	12:17:00	Credit card	175.34	4.761
129-29-8530	A	Yangon	Member	Male	Sports and travel	62.62	5	15.655	328.755	10-03-2019	19:15:00	Ewallet	313.1	4.761
333-73-7901	C	Naypyitaw	Normal	Female	Health and beauty	54.92	8	21.968	461.328	23-03-2019	13:24:00	Ewallet	439.36	4.761
777-82-7220	B	Mandalay	Member	Male	Home and lifestyle	39.12	8	13.648	352.808	02-02-2019	17:01:00	Cash	249.96	4.761

Q10. Calculate the total sales and gross income for each month.

```sql
SELECT DATE_FORMAT(STR_TO_DATE(Date, '%d-%m-%Y'), '%Y-%m') AS Month,
    SUM(Total) AS Total_Sales,
    SUM(gross_income) AS Total_Gross_Income
FROM walmartdata
GROUP BY DATE_FORMAT(STR_TO_DATE(Date, '%d-%m-%Y'), '%Y-%m');
```

Month	Total_Sales	Total_Gross_Income
2019-01	116291.86800000005	5537.707999999999
2019-03	109455.50700000004	5212.167000000006
2019-02	97219.37399999997	4629.494000000001

Q11. Find the number of sales transactions that occurred after 6 PM.

```
SELECT COUNT(*) AS Transactions_After_6PM
FROM walmartdata
WHERE TIME(Time) > '18:00:00';
```

Transactions_After_6PM
▶ 281

Q12. List the sales transactions that have a higher total than the average total of all transactions.

```
SELECT *
FROM WalmartData
WHERE Total > (SELECT AVG(Total) FROM WalmartData);
```

Invoice_ID	Branch	City	Customer_type	Gender	Product_line	Unit_price	Quantity	Tax_5%	Total	Date	Time	Payment	cogs	gros:
750-67-8428	A	Yangon	Member	Female	Health and beauty	74.69	7	26.1415	548.9715	05-01-2019	13:08:00	Ewallet	522.83	4.761
631-41-3108	A	Yangon	Normal	Male	Home and lifestyle	46.33	7	16.2155	340.5255	03-03-2019	13:23:00	Credit card	324.31	4.761
123-19-1176	A	Yangon	Member	Male	Health and beauty	58.22	8	23.288	489.048	27-01-2019	20:33:00	Ewallet	465.76	4.761
373-73-7910	A	Yangon	Normal	Male	Sports and travel	86.31	7	30.2085	634.3785	08-02-2019	10:37:00	Ewallet	604.17	4.761
699-14-3026	C	Naypyitaw	Normal	Male	Electronic accessories	85.39	7	29.8865	627.6165	25-03-2019	18:30:00	Ewallet	597.73	4.761
355-53-5943	A	Yangon	Member	Female	Electronic accessories	68.84	6	20.652	433.692	25-02-2019	14:36:00	Ewallet	413.04	4.761
315-22-5665	C	Naypyitaw	Normal	Female	Home and lifestyle	73.56	10	36.78	772.38	24-02-2019	11:38:00	Ewallet	735.6	4.761
252-56-2699	A	Yangon	Normal	Male	Food and beverages	43.19	10	21.595	453.495	07-02-2019	16:48:00	Ewallet	431.9	4.761
829-34-3910	A	Yangon	Normal	Female	Health and beauty	71.38	10	35.69	749.49	29-03-2019	19:21:00	Cash	713.8	4.761
299-46-1805	B	Mandalay	Member	Female	Sports and travel	93.72	6	28.116	590.436	15-01-2019	16:19:00	Cash	562.32	4.761

Q13. Calculate the cumulative gross income for each branch by date.

```
SELECT Branch, Date,
     SUM(gross_income)
     OVER (PARTITION BY Branch
     ORDER BY STR_TO_DATE(Date, '%d-%m-%Y')) AS Cumulative_Gross_Income
FROM WalmartData;
```

Branch	Date	Cumulative_Gross_Income
A	01-01-2019	112.92
A	01-01-2019	112.92
A	01-01-2019	112.92
A	01-01-2019	112.92
A	01-01-2019	112.92
A	02-01-2019	127.54150000000001
A	02-01-2019	127.54150000000001
A	03-01-2019	172.18
A	03-01-2019	172.18
A	03-01-2019	172.18
A	04-01-2019	195.19250000000002
A	04-01-2019	195.19250000000002
A	05-01-2019	291.5975
A	05-01-2019	291.5975
A	05-01-2019	291.5975

Branch	Date	Cumulative_Gross_Income
B	26-01-2019	1378.2244999999996
B	26-01-2019	1378.2244999999996
B	26-01-2019	1378.2244999999996
B	26-01-2019	1378.2244999999996
B	26-01-2019	1378.2244999999996
B	27-01-2019	1421.9614999999997
B	27-01-2019	1421.9614999999997
B	28-01-2019	1571.0269999999994
B	28-01-2019	1571.0269999999994
B	28-01-2019	1571.0269999999994
B	28-01-2019	1571.0269999999994
B	28-01-2019	1571.0269999999994
B	28-01-2019	1571.0269999999994
B	29-01-2019	1616.1009999999992
B	29-01-2019	1616.1009999999992

Branch	Date	Cumulative_Gross_Income
C	14-01-2019	736.8265000000001
C	15-01-2019	797.5520000000001
C	15-01-2019	797.5520000000001
C	15-01-2019	797.5520000000001
C	16-01-2019	918.527
C	16-01-2019	918.527
C	16-01-2019	918.527
C	16-01-2019	918.527
C	16-01-2019	918.527
C	17-01-2019	949.3540000000002
C	17-01-2019	949.3540000000002
C	17-01-2019	949.3540000000002
C	18-01-2019	1013.3675000000002
C	18-01-2019	1013.3675000000002
C	18-01-2019	1013.3675000000002

Q14. Find the total cogs for each customer type in each city.

```
SELECT City, Customer_type, SUM(cogs) AS Total_COGS
FROM WalmartData
GROUP BY City, Customer_type;
```

City	Customer_type	Total_COGS
Yangon	Member	51083.30999999997
Naypyitaw	Normal	51130.87999999999
Yangon	Normal	50059.89999999994
Mandalay	Member	51147.32000000014
Mandalay	Normal	49993.32000000001
Naypyitaw	Member	54172.65000000016

FINDINGS

◆ **Branch Performance**

Branch A shows specific characteristics and sales patterns, which might indicate different customer preferences or market dynamics compared to other branches.

◆ **Product Line Sales**

Some product lines significantly contribute to total sales revenue, with varying performance across different product categories.

◆ **Payment Methods**

Some product lines significantly contribute to total sales revenue, with varying performance across different product categories.

◆ **Sales Timing**

A significant portion of sales occurs on weekends (Saturdays and Sundays), and there is a noticeable number of transactions occurring after 6 PM, indicating peak shopping times.

CONCLUSION

The Walmart sales analysis reveals several critical insights into the company's performance across various branches, product lines, and regions. The data indicates strong sales growth driven by certain high-performing product categories and specific regions, with notable peaks during weekends and after typical working hours. However, there are disparities in branch performance and customer satisfaction, suggesting areas for targeted improvement.