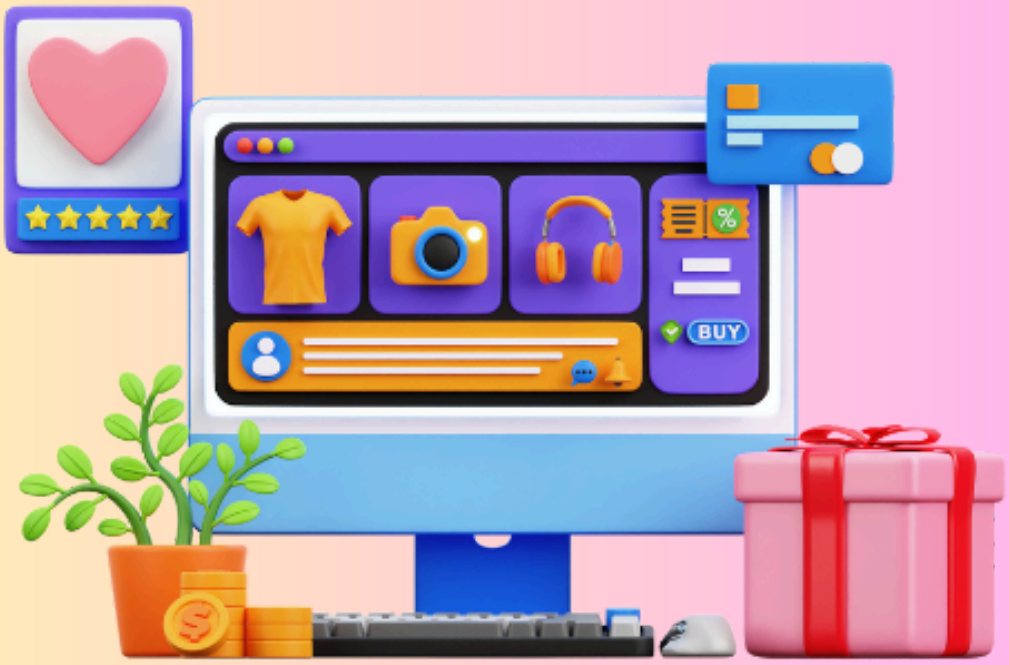
Walmart Sales Data Analysis Using SQL



in/tejassawaithul

BACKGROUND

Walmart is a leading global retailer with a significant online presence. In this project, Walmart's e-commerce data was analyzed to examine various aspects of sales performance. The data includes sales transactions, customer interactions, inventory levels, and product details. Using SQL, the data was queried and analyzed to uncover insights such as identifying top-selling products, understanding customer preferences, and tracking inventory trends. The goal of the project was to help Walmart make informed decisions on inventory management, pricing strategies, and product offerings, ultimately improving sales performance and customer satisfaction. SQL was applied to turn raw data into actionable insights that support business decisions.

OBJECTIVES

- To explore and retrieve all sales data, ensuring accessibility for further analysis and identifying specific transaction types based on conditions.
- To analyze sales performance across product lines, calculate total revenue, and assess contributions from different demographics and product lines.
- To evaluate branch and regional performance by analyzing customer ratings, sales, and income across different cities and branches.
- To gain insights into customer behavior by analyzing sales data based on customer type and item quantities sold.
- To examine tax contributions and identify the most used payment methods, analyzing their impact on total revenue.

DATASET OVERVIEW AND BASIC RETRIEVAL

Objective-1 :

To retrieve all data from the sales dataset for initial exploration, understanding its contents, and ensuring accessibility for further analysis.

1. Retrieve all records from the sales dataset.

```
SELECT * FROM Walmart.SalesData;
```

Invoice_ID	Branch	City	Customer_Type	Gender	Product_Line	Unit_Price	Quantity	Tax_5_Percent	Total	Date	Time	Payment	COGS	Gross_Margin_Percentage	Gross_Income	Rat
860-79-0874	C	San Francisco	Member	Female	Fashion accessories	99.30	10	49.6500	1042.6500	2014-09-27	14:53:00	Credit card	993.00	4.76	49.65	6.6
687-47-8271	A	San Francisco	Normal	Male	Fashion accessories	98.98	10	49.4900	1039.2900	2012-08-08	16:20:00	Credit card	989.80	4.76	49.49	8.7
283-26-5248	C	Broomfield	Member	Female	Food and beverages	98.52	10	49.2600	1034.4600	2013-07-09	20:23:00	Ewallet	985.20	4.76	49.26	4.5
751-41-9720	C	Los Angeles	Normal	Male	Home and lifestyle	97.50	10	48.7500	1023.7500	2012-12-08	16:18:00	Ewallet	975.00	4.76	48.75	8.0
303-96-2227	B	Moreno Valley	Normal	Female	Home and lifestyle	97.38	10	48.6900	1022.4900	2011-02-23	17:16:00	Ewallet	973.80	4.76	48.69	4.4
744-16-7898	B	Seattle	Normal	Female	Home and lifestyle	97.37	10	48.6850	1022.3850	2014-01-31	13:48:00	Credit card	973.70	4.76	48.69	4.9
271-88-8734	C	Seattle	Member	Female	Fashion accessories	97.21	10	48.6050	1020.7050	2013-08-24	13:00:00	Credit card	972.10	4.76	48.61	8.7
234-65-2137	C	Redondo Beach	Normal	Male	Home and lifestyle	95.58	10	47.7900	1003.5900	2011-04-12	13:32:00	Cash	955.80	4.76	47.79	4.8
554-42-2417	C	San Francisco	Normal	Female	Sports and travel	95.44	10	47.7200	1002.1200	2014-06-16	13:45:00	Cash	954.40	4.76	47.72	5.2
325-77-6186	A	Denver	Member	Female	Home and lifestyle	90.65	10	45.3250	951.8250	2014-12-02	10:53:00	Ewallet	906.50	4.76	45.33	7.3
280-17-4359	C	Phoenix	Member	Male	Health and beauty	90.50	10	45.2500	950.2500	2014-07-19	13:48:00	Cash	905.00	4.76	45.25	8.1
219-22-9386	B	San Diego	Member	Male	Sports and travel	99.96	9	44.9820	944.6220	2012-11-08	17:26:00	Credit card	899.64	4.76	44.98	4.2
702-83-5291	C	Seattle	Member	Male	Fashion accessories	99.82	9	44.9190	943.2990	2011-11-02	10:43:00	Cash	898.38	4.76	44.92	6.6
731-81-9469	C	Pomona	Member	Female	Sports and travel	89.80	10	44.9000	942.9000	2014-06-25	13:00:00	Credit card	898.00	4.76	44.90	5.4
817-69-8206	B	San Francisco	Normal	Female	Electronic accessories	99.73	9	44.8785	942.4485	2014-09-05	19:42:00	Credit card	897.57	4.76	44.88	6.5
393-65-2792	C	Los Angeles	Normal	Male	Food and beverages	89.48	10	44.7400	939.5400	2012-11-24	12:46:00	Credit card	894.80	4.76	44.74	9.6
751-69-0068	C	Los Angeles	Normal	Male	Sports and travel	99.24	9	44.6580	937.8180	2014-10-08	19:09:00	Ewallet	893.16	4.76	44.66	9.0
866-99-7614	C	Sunnyvale	Normal	Male	Food and beverages	89.20	10	44.6000	936.6000	2014-12-17	15:42:00	Credit card	892.00	4.76	44.60	4.4

2. Retrieve all records for sales where the total revenue is above $500 and the payment method is "Credit card".

```
SELECT * FROM Walmart.SalesData
WHERE Total > 500 AND
Payment = 'Credit card';
```

Invoice_ID	Branch	City	Customer_Type	Gender	Product_Line	Unit_Price	Quantity	Tax_5_Percent	Total	Date	Time	Payment	COGS	Gross_Margin_Percentage	Gross_Income	Rat
860-79-0874	C	San Francisco	Member	Female	Fashion accessories	99.30	10	49.6500	1042.6500	2014-09-27	14:53:00	Credit card	993.00	4.76	49.65	6.6
687-47-8271	A	San Francisco	Normal	Male	Fashion accessories	98.98	10	49.4900	1039.2900	2012-08-08	16:20:00	Credit card	989.80	4.76	49.49	8.7
744-16-7898	B	Seattle	Normal	Female	Home and lifestyle	97.37	10	48.6850	1022.3850	2014-01-31	13:48:00	Credit card	973.70	4.76	48.69	4.9
271-88-8734	C	Seattle	Member	Female	Fashion accessories	97.21	10	48.6050	1020.7050	2013-08-24	13:00:00	Credit card	972.10	4.76	48.61	8.7
219-22-9386	B	San Diego	Member	Male	Sports and travel	99.96	9	44.9820	944.6220	2012-11-08	17:26:00	Credit card	899.64	4.76	44.98	4.2
731-81-9469	C	Pomona	Member	Female	Sports and travel	89.80	10	44.9000	942.9000	2014-06-25	13:00:00	Credit card	898.00	4.76	44.90	5.4
817-69-8206	B	San Francisco	Normal	Female	Electronic accessories	99.73	9	44.8785	942.4485	2014-09-05	19:42:00	Credit card	897.57	4.76	44.88	6.5
393-65-2792	C	Los Angeles	Normal	Male	Food and beverages	89.48	10	44.7400	939.5400	2012-11-24	12:46:00	Credit card	894.80	4.76	44.74	9.6
866-99-7614	C	Sunnyvale	Normal	Male	Food and beverages	89.20	10	44.6000	936.6000	2014-12-17	15:42:00	Credit card	892.00	4.76	44.60	4.4
641-51-2661	C	Los Angeles	Member	Female	Food and beverages	87.10	10	43.5500	914.5500	2012-08-25	14:45:00	Credit card	871.00	4.76	43.55	9.9
746-04-1077	B	San Diego	Member	Female	Food and beverages	84.63	10	42.3150	888.6150	2012-04-26	11:36:00	Credit card	846.30	4.76	42.32	9.0
450-42-3339	C	Los Angeles	Normal	Male	Health and beauty	84.61	10	42.3050	888.4050	2012-11-24	18:58:00	Credit card	846.10	4.76	42.31	8.8
852-62-7105	B	Redmond	Normal	Female	Fashion accessories	83.25	10	41.6250	874.1250	2013-12-23	11:25:00	Credit card	832.50	4.76	41.63	4.4
607-65-2441	C	Los Angeles	Member	Male	Health and beauty	81.95	10	40.9750	860.4750	2013-12-07	12:39:00	Credit card	819.50	4.76	40.98	6.0
698-98-5964	A	San Francisco	Normal	Female	Food and beverages	81.21	10	40.6050	852.7050	2012-06-20	13:01:00	Credit card	812.10	4.76	40.61	6.3
138-17-5109	A	Los Angeles	Member	Female	Home and lifestyle	89.21	9	40.1445	843.0345	2014-12-30	15:42:00	Credit card	802.89	4.76	40.14	6.5
638-60-7125	A	Seattle	Normal	Female	Electronic accessories	99.56	8	39.8240	836.3040	2014-11-14	17:03:00	Credit card	796.48	4.76	39.82	5.2
431-66-2305	B	Seattle	Normal	Female	Electronic accessories	88.25	9	39.7125	833.9625	2014-03-21	20:51:00	Credit card	794.25	4.76	39.71	7.6

PRODUCT AND SALES PERFORMANCE

Objective-2 :

Analyze sales performance, revenue, and demographic contributions across product lines.

1. Calculate total revenue per product line, and include the highest and lowest revenue amounts for each product line.

```
SELECT Product_Line, SUM(Total) AS TotalRevenue,
MIN(Total) AS MinRevenue, MAX(Total) AS MaxRevenue
FROM Walmart.SalesData GROUP BY Product_Line
ORDER BY TotalRevenue DESC;
```

Product_Line	TotalRevenue	MinRevenue	MaxRevenue
Food and beverages	56144.8440	22.6590	1034.4600
Sports and travel	55122.8265	10.6785	1002.1200
Electronic accessories	54337.5315	26.7225	942.4485
Fashion accessories	54305.8950	12.6945	1042.6500
Home and lifestyle	53861.9130	14.6790	1023.7500
Health and beauty	49193.7390	18.6375	950.2500

2. Find the highest gross income for each product line.

```
SELECT Product_Line,
MAX(Gross_Income) AS HighestGrossIncome
FROM Walmart.SalesData GROUP BY Product_Line;
```

Product_Line	HighestGrossIncome
Fashion accessories	49.65
Food and beverages	49.26
Home and lifestyle	48.75
Sports and travel	47.72
Health and beauty	45.25
Electronic accessories	44.88

3. Calculate the total revenue for each product line by gender and find which gender has contributed the most to each product line's revenue.

```
SELECT Product_Line, Gender, SUM(Total) AS TotalRevenue
FROM Walmart.SalesData
GROUP BY Product_Line, Gender
ORDER BY Product_Line, TotalRevenue DESC;
```

Product_Line	Gender	TotalRevenue
Electronic accessories	Male	27235.5090
Electronic accessories	Female	27102.0225
Fashion accessories	Female	30437.4000
Fashion accessories	Male	23868.4950
Food and beverages	Female	33170.9175
Food and beverages	Male	22973.9265
Health and beauty	Male	30632.7525

4. Identify the top 5 product lines based on total revenue, and calculate the total quantity of items sold for each product line.

```
SELECT Product_Line, SUM(Total) AS TotalRevenue,
SUM(Quantity) AS TotalItemsSold
FROM Walmart.SalesData GROUP BY Product_Line
ORDER BY TotalRevenue DESC LIMIT 5;
```

Product_Line	TotalRevenue	TotalItemsSold
Food and beverages	56144.8440	952
Sports and travel	55122.8265	920
Electronic accessories	54337.5315	971
Fashion accessories	54305.8950	902
Home and lifestyle	53861.9130	911

Branch and Regional Analysis

Objective-3 :

To evaluate branch and regional performance by analyzing customer ratings, sales, and income across different cities and branches.

1. Find the average customer rating for each branch and order the branches based on the highest average rating.

```
SELECT Branch, AVG(Rating) AS AverageRating
FROM Walmart.SalesData GROUP BY Branch
ORDER BY AverageRating DESC;
```

Branch	AverageRating
C	7.07287
A	7.02706
B	6.81807

2. Calculate the total sales and average sales for each city.

```
SELECT City,  SUM(Total) AS TotalSales,
AVG(Total) AS AverageSales FROM Walmart.SalesData
GROUP BY City ORDER BY AverageSales DESC;
```

City	TotalSales	AverageSales
Ontario	829.7100	829.71000000
Covington	745.8360	745.83600000
Auburn	688.6215	688.62150000
San Mateo	634.3785	634.37850000
Gresham	1254.7080	627.35400000
Broomfield	1242.8850	621.44250000
Escondido	568.5120	568.51200000

3. Calculate the total gross income for each branch and rank the branches based on gross income in descending order.

```
SELECT Branch, SUM(Gross_Income) AS TotalGrossIncome
FROM Walmart.SalesData
GROUP BY Branch
ORDER BY TotalGrossIncome DESC;
```

Branch	TotalGrossIncome
C	5265.33
A	5057.36
B	5057.36

4. Calculate the total revenue and the average gross income for each branch.

```
SELECT Branch,
SUM(Total) AS TotalRevenue,
AVG(Gross_Income) AS AverageGrossIncome
FROM Walmart.SalesData
GROUP BY Branch ORDER BY TotalRevenue DESC;
```

Branch	TotalRevenue	AverageGrossIncome
C	110568.7065	16.052835
A	106200.3705	14.874588
B	106197.6720	15.233012

Customer and Transaction Insights

Objective-4 :

To gain insights into customer behavior by analyzing sales data based on customer type and item quantities sold.

1. Calculate the number of sales for each customer type and find the average sale value.

```
SELECT Customer_Type, COUNT(*) AS SalesCount,
AVG(Total) AS AverageSaleValue FROM Walmart.SalesData
GROUP BY Customer_Type;
```

Customer_Type	SalesCount	AverageSaleValue
Member	501	327.79130539
Normal	499	318.12285571

2. Retrieve all records where more than 9 items were sold, including key sales details.

```
SELECT Invoice_ID, Branch, City, Customer_Type,
Gender, Product_Line, Quantity FROM Walmart.SalesData
WHERE Quantity > 9;
```

Invoice_ID	Branch	City	Customer_Type	Gender	Product_Line	Quantity
860-79-0874	C	San Francisco	Member	Female	Fashion accessories	10
687-47-8271	A	San Francisco	Normal	Male	Fashion accessories	10
283-26-5248	C	Broomfield	Member	Female	Food and beverages	10
751-41-9720	C	Los Angeles	Normal	Male	Home and lifestyle	10
303-96-2227	B	Moreno Valley	Normal	Female	Home and lifestyle	10
744-16-7898	B	Seattle	Normal	Female	Home and lifestyle	10
271-88-8734	C	Seattle	Member	Female	Fashion accessories	10
234-65-2137	C	Redondo Beach	Normal	Male	Home and lifestyle	10
554-42-2417	C	San Francisco	Normal	Female	Sports and travel	10

Tax and Payment Method Trends

Objective-5 :

To examine tax contributions and identify the most used payment methods, analyzing their impact on total revenue.

1. Find the most used payment method, its count, and total revenue, sorted by count.

```
SELECT Payment, COUNT(*) AS PaymentCount,
SUM(Total) AS TotalRevenue FROM Walmart.SalesData
GROUP BY Payment
ORDER BY PaymentCount DESC;
```

Payment	PaymentCount	TotalRevenue
Ewallet	345	109993.1070
Cash	344	112206.5700
Credit card	311	100767.0720

2. Retrieve the total tax (5%) collected per city, and order by the highest total tax collected.

```
SELECT City
SUM(Tax_5_Percent) AS TotalTaxCollected
FROM Walmart.SalesData
GROUP BY City ORDER BY TotalTaxCollected DESC;
```

City	TotalTaxCollected
Los Angeles	3806.2855
San Francisco	2613.4730
Seattle	1951.2145
San Diego	694.6155
Phoenix	472.0775