🛒 Walmart Sales Analysis (SQL Project)

A comprehensive SQL-driven data analysis project using Walmart's transactional data to uncover valuable business insights, optimize operations, and understand customer behaviors.

📁 Dataset Overview

The dataset contains the following fields:

- `invoice_id`, `Branch`, `City`, `category`, `unit_price_USD`, `quantity`, `date`, `time`, `payment_method`, `rating`, `profit_margin`, `Day`, `Month`, `Year`, `Total_price`

🛠️ Tools Used

- **PostgreSQL** – SQL queries & data analysis
- **Pandas (Python)** – Post-processing and aggregation
- **GitHub** – Version control and documentation

🔍 Business Questions, SQL Queries & Answers

1. Analyze Payment Methods and Sales

```sql
SELECT payment_method, COUNT(DISTINCT invoice_id) AS total_transactions,
       SUM(quantity) AS total_items_sold
FROM walmart_data
GROUP BY payment_method;
```

Answer:

Payment Method	Total Transactions	Total Items Sold
Cash	1832	4984
Credit card	4257	9569.35
Ewallet	3911	9002.60

---

2. Highest-Rated Category in Each Branch

```sql
SELECT Branch, category, AVG(rating) AS rating
FROM walmart_data
GROUP BY Branch, category
QUALIFY ROW_NUMBER() OVER(PARTITION BY Branch ORDER BY AVG(rating) DESC) = 1;
```

Answer (Top 5):

Branch	Category	Avg Rating
WALM001	Electronic accessories	7.45
WALM002	Food and beverages	8.25
WALM003	Sports and travel	7.50

---

3. Busiest Day for Each Branch

```
SELECT Branch, Day, COUNT(invoice_id) AS total_transactions
FROM walmart_data
GROUP BY Branch, Day
QUALIFY ROW_NUMBER() OVER(PARTITION BY Branch ORDER BY COUNT(invoice_id) DESC) =
1;
```

Answer (Top 5):

Branch	Busiest Day	Total Transactions
WALM001	Thursday	16
WALM002	Thursday	15
WALM003	Tuesday	33

---

4. Total Quantity Sold by Payment Method

```
SELECT payment_method, SUM(quantity) AS total_quantity_sold
FROM walmart_data
GROUP BY payment_method;
```

Payment Method	Total Quantity Sold
Cash	4984
Credit card	9569.35
Ewallet	9002.60

---

5. Category Ratings by City

```
SELECT City, category, AVG(rating), MIN(rating), MAX(rating)
FROM walmart_data
GROUP BY City, category;
```

Example (Abilene):

City	Category	Avg	Min	Max
Abilene	Health and beauty	9.7	9.7	9.7

---

6. Total Profit by Category

```sql
SELECT category, SUM(profit_margin) AS Total_Profit
FROM walmart_data
GROUP BY category
ORDER BY Total_Profit DESC;
```

Category	Total Profit
Home and lifestyle	1790.28
Fashion accessories	1789.17
Electronic accessories	164.73

---

7. Most Common Payment Method per Branch

```sql
SELECT Branch, payment_method, COUNT(*) AS count
FROM walmart_data
GROUP BY Branch, payment_method
QUALIFY ROW_NUMBER() OVER(PARTITION BY Branch ORDER BY COUNT(*) DESC) = 1;
```

Branch	Payment Method	Count
WALM001	Ewallet	46
WALM003	Credit card	115

---

8. Transactions by Time Shift

```sql
SELECT Branch,
       CASE
         WHEN time BETWEEN '00:00:00' AND '11:59:59' THEN 'Morning'
         WHEN time BETWEEN '12:00:00' AND '17:59:59' THEN 'Afternoon'
         ELSE 'Evening'
       END AS Shift,
       COUNT(invoice_id) AS Total_Transactions
FROM walmart_data
GROUP BY Branch, Shift;
```

Example (WALM001):

Shift	Transactions
Morning	8
Afternoon	37
Evening	30

---

9. Branches with Highest Revenue Decline YoY

```sql
WITH yearly_revenue AS (
  SELECT Branch, Year, SUM(Total_price) AS Total_Revenue
  FROM walmart_data
  GROUP BY Branch, Year
)
SELECT curr.Branch, curr.Year AS current_year, prev.Year AS previous_year,
       curr.Total_Revenue AS current_revenue,
       prev.Total_Revenue AS previous_revenue,
       (curr.Total_Revenue - prev.Total_Revenue) AS revenue_change
FROM yearly_revenue curr
JOIN yearly_revenue prev
ON curr.Branch = prev.Branch AND curr.Year = prev.Year + 1
ORDER BY revenue_change ASC;
```

Branch	Year	Revenue Drop
WALM095	2020	-4828.65
WALM057	2020	-4744.74
WALM049	2020	-4379.81

📊 Key Findings

- 💳 **Credit cards** dominate as the preferred payment method, but branches differ.
- 📅 **Fridays and Sundays** are often the busiest days — ideal for promotions.
- 🛒 **Fashion** and **Home & Lifestyle** are the most profitable categories.
- ⏰ **Afternoon and evening shifts** see the highest foot traffic.
- 📉 Certain branches showed a **notable revenue drop year-over-year**, highlighting performance issues.

📈 Reports & Insights

- **Branch-Level Revenue Trends:** Revealed underperforming branches needing strategic focus.
- **Category Profitability:** Helped in identifying high-margin product lines for optimization.

- **Payment Preferences:** Informed potential improvements to POS systems.
- **Customer Ratings:** Uncovered variation by city and category — guiding customer satisfaction efforts.

☑ Conclusion

This project leveraged SQL to explore Walmart's transactional data and delivered actionable business intelligence insights. From operational adjustments to strategic marketing and financial planning, this analysis empowers decision-makers to enhance efficiency, profitability, and customer experience.
