Walmart Sales Analysis Using MySQL

Project Overview

This project analyzes Walmart's transactional sales data using MySQL to evaluate overall business performance, customer behavior, product performance, and profitability. The objective is to transform raw sales records into meaningful insights that support data-driven business decisions in a retail environment.

Business Objective

The purpose of this analysis is to understand how Walmart performs across different branches, products, and customer segments. The project focuses on identifying revenue and profit drivers, customer purchasing patterns, peak sales periods, and areas where operational or strategic improvements can be made.

Dataset Description

The dataset consists of transactional sales records where each row represents a single customer purchase. It includes invoice details, branch and city information, customer demographics, product categories, pricing and cost metrics, payment methods, transaction timestamps, and customer ratings.

Tools and Technologies

MySQL was used as the primary database system, and SQL was used to perform data validation, aggregation, trend analysis, and business performance evaluation.

Database Design

A dedicated database named **walmart** was created to store sales analytics data. The analysis is performed on a single table named **walmartsalesdata**, which is structured to support reporting, grouping, and aggregation across multiple business dimensions such as time, location, product, and customer type.

Data Validation Approach

Basic validation checks were performed before analysis to ensure data quality. This included verifying the total number of records and confirming the transaction date range to ensure the dataset was complete and suitable for analysis.

Analysis Performed

Overall Business Performance

Total sales revenue and total profit were calculated to assess the financial health of the business and provide a high-level performance overview for management.

Branch Performance Analysis

Sales were analyzed at the branch level to compare performance across locations and identify high-performing and underperforming branches.

Product Performance Analysis

Product lines were evaluated based on quantity sold and total revenue generated. This analysis helps identify best-selling products and key revenue-driving categories.

Customer Behavior Analysis

Customer purchasing behavior was analyzed based on customer type and gender to understand how different customer segments contribute to overall sales.

Time-Based Sales Analysis

Sales trends were analyzed by month to identify seasonal patterns, and hourly sales analysis was performed to determine peak business hours and support operational planning.

Payment Method Analysis

Transaction counts and average transaction values were analyzed by payment method to understand customer payment preferences and spending behavior.

Profitability Analysis

Profit contribution was analyzed by branch, and profit margins were calculated for each product line to evaluate cost efficiency and pricing effectiveness.

Customer Satisfaction Analysis

Average customer ratings were analyzed by product line and branch to identify areas affecting customer experience and satisfaction.

Advanced Transaction Analysis

High-value transactions were identified at the branch level to support premium customer analysis and anomaly detection.

Key Insights

Certain branches and product lines consistently generate higher revenue and profit. A limited number of products contribute a significant share of total sales. Sales peak during specific hours, indicating opportunities for optimized staffing. Some payment methods are associated with higher average transaction values. Products and branches with higher customer ratings tend to perform better in terms of sales.

Business Impact

The insights from this analysis can support inventory optimization, pricing strategy refinement, targeted marketing campaigns, operational efficiency improvements, and customer experience enhancement.

Conclusion

This project demonstrates how MySQL can be used for real-world retail analytics by converting transactional sales data into actionable business insights. The structured queries and documented analysis reflect industry-standard practices used by data and analytics teams.

Future Enhancements

The analysis can be extended by integrating visualization tools such as Power BI or Tableau, automating recurring reports, applying predictive models for sales forecasting, and performing deeper customer segmentation analysis.