Project Report

on

Walmart Sales Data

Table of Contents

Introduction of the Project

Walmart, one of the largest retail corporations globally, operates numerous stores across various regions. Understanding the sales patterns and the factors influencing them is crucial for maintaining a competitive edge and optimizing operations. The ability to analyze historical sales data provides insights that can lead to better inventory management, targeted marketing strategies, and overall improved decision-making. This project focuses on analysing Walmart sales data to derive insights into sales performance, customer demographics, product line performance, and other key metrics. The goal is to identify patterns and trends that can help Walmart make data-driven decisions to enhance their business strategies.

The primary goal of this project is to perform a comprehensive analysis of Walmart Sales Data to uncover trends, patterns, and insights that can inform business decisions. The analysis spans from basic Exploratory Data Analysis (EDA) to more advanced techniques, including time series forecasting and predictive modeling. The project aims to shed light on branch performance, product lines, sales trends, and customer behaviour. Through meticulous data preparation, in-depth exploratory data analysis (EDA), and the creation of interactive visualizations using Power BI. I have extracted valuable insights that can inform business strategies and decision-making.

The insights derived from this analysis can significantly impact Walmart's operational efficiency and strategic planning. By understanding sales dynamics and customer behaviour, Walmart can improve its inventory management, optimize supply chain logistics, and tailor marketing strategies to enhance customer satisfaction and sales performance.

Objective of the Project

The project aims to conduct a comprehensive analysis of Walmart's sales data to extract actionable insights and support data-driven decision-making. The objectives are broken down into several key analytical areas:

1. **Initial Exploration of the Walmart Sales Dataset:**
   - **Objective:** To gain a preliminary understanding of the dataset and its structure.
   - **Key Focus Areas:**
     - Understanding the dataset's dimensions and basic characteristics.
     - Identifying key variables and their data types.
     - Checking for missing values, duplicates, and inconsistencies in the data.
2. **Exploratory Data Analysis:**
   - **Objective:** To explore the data in-depth and uncover hidden patterns and relationships.
   - **Key Focus Areas:**
     - Descriptive statistics and summary metrics for key variables.
     - Distribution analysis of numerical and categorical variables.
     - Correlation analysis to understand relationships between variables.
3. **Visualizations Analysis:**
   - **Objective:** To visualize data patterns, trends, and distributions for better interpretation.
   - **Key Focus Areas:**
     - Creating visual representations such as histograms, bar charts, line graphs, and scatter plots.
     - Visualizing sales trends over time, across different stores, and departments.
     - Using heatmaps and other advanced visualization techniques for deeper insights.
4. **Feature Engineering:**
   - **Objective:** To create new features that enhance the predictive power of the data.
   - **Key Focus Areas:**
     - Creating time-based features like day of the week, month, quarter, and year.
     - Generating lag features to capture historical sales trends.
     - Creating interaction features to capture combined effects of variables (e.g., holidays and promotions).
5. **Other Analysis and Insights:**
   A. **Sales Performance by Branch and City:**
      - **Objective:** To evaluate and compare sales performance across different locations.
      - **Key Focus Areas:**
        - Analyzing total sales, average sales per transaction, and growth rates by branch and city.

- Identifying high-performing and underperforming locations.

B. **Customer Demographics (Sales Distribution by Customer Type and Gender):**

- **Objective:** To understand the demographic distribution of customers and its impact on sales.
- **Key Focus Areas:**
  - Analyzing sales distribution by customer type (e.g., new vs. returning customers) and gender.
  - Identifying demographic segments contributing most to revenue.

C. **Product Line Performance (Based on Revenue and Sales Volume):**

- **Objective:** To assess the performance of different product lines.
- **Key Focus Areas:**
  - Ranking product lines based on total revenue and sales volume.
  - Analyzing product line profitability and customer preferences.

D. **Pricing and Quantity (Unit Price vs. Quantity Sold, and Total Sales Based on Unit Price and Quantity):**

- **Objective:** To analyze the relationship between pricing strategies and sales volume.
- **Key Focus Areas:**
  - Examining the correlation between unit price and quantity sold.
  - Analyzing total sales performance in relation to pricing strategies.

E. **Payment Methods (Based on Popular Payment Methods, and Payment Method Impact on Sales)**

- **Objective:** To analyze the popularity and impact of different payment methods on sales.
- **Key Focus Areas:**
  - Analyzing sales distribution by payment method (e.g., cash, credit, mobile payments).
  - Assessing the impact of payment method preferences on overall sales.

F. **Sales Trends Over Time (Based on Daily Sales, Monthly Sales, and Hourly Sales)**

- **Objective:** To analyze temporal sales trends at various granularities.
- **Key Focus Areas:**
  - Analyzing daily, weekly, monthly, and hourly sales patterns.
  - Identifying peak sales periods and potential seasonality in sales.

G. **Customer Feedback (Based on Rating Distribution and Area of Improvements on Low Ratings)**

- **Objective:** To analyze customer feedback and its impact on sales.
- **Key Focus Areas:**
  - Analyzing customer ratings and reviews to gauge satisfaction levels.
  - Identifying trends in customer feedback and correlating with sales performance.

Methodology of the Project

The methodology outlines the systematic approach taken to analyze Walmart's sales data. It encompasses data preparation, exploratory analysis, feature engineering, and the application of various analytical techniques to derive actionable insights. The methodology is structured as follows:

1. **Data Collection and Understanding:**
    - **Objective:** To gather and understand the dataset, including its structure, features, and potential data issues.
    - **Steps:**
        - Load the dataset and explore its structure, including the number of records and features.
        - Review the data dictionary, if available, to understand the meaning of each feature.
        - Identify key variables related to sales, customer demographics, product lines, and external factors.
2. **Data Cleaning and Preprocessing:**
    - **Objective:** To ensure data quality and prepare the dataset for analysis.
    - **Steps:**
        - Handle missing values through appropriate strategies such as imputation or removal.
        - Correct any data inconsistencies or errors, such as duplicate records or incorrect data entries.
        - Normalize and standardize numerical features if necessary.
        - Encode categorical variables using techniques like one-hot encoding or label encoding.
3. **Exploratory Data Analysis (EDA):**
    - **Objective:** To explore the data, identify patterns, and understand the relationships between variables.
    - **Steps:**
        - Conduct descriptive statistical analysis to summarize key features (e.g., mean, median, mode, standard deviation).
        - Create visualizations such as histograms, bar charts, line graphs, and scatter plots to visualize distributions and relationships.
        - Perform correlation analysis to identify relationships between numerical features.
        - Analyze trends and patterns in sales data, including seasonal and cyclical variations.
4. **Feature Engineering:**
    - **Objective:** To create new features that enhance the predictive power of the dataset.

- **Steps:**
    - Generate time-based features such as day of the week, month, quarter, and year from the date variable.
    - Create lag features to capture historical sales data and identify trends.
    - Develop interaction features that combine multiple variables to capture complex relationships.
    - Calculate aggregate features such as total sales per customer, average sales per transaction, and customer lifetime value.

5. **Data Analysis:**
    - **Objective:** To analyze the data for understanding sales trends, Customer Feedback and Customer Behaviour.
    - **Steps:**
        - Perform segmentation analysis to identify different customer segments based on purchasing behaviour.
        - Conduct sales performance analysis by branch, city, and department to identify high-performing and underperforming areas.
        - Analyze the impact of pricing, promotions, and external factors on sales performance.

6. **Visualization and Communication of Insights:**
    - **Objective:** To effectively communicate the findings and insights derived from the analysis.
    - **Steps:**
        - Create comprehensive dashboards and visualizations to present key insights.
        - Use storytelling techniques to explain the data insights, trends, and recommendations.
        - Highlight actionable insights and provide strategic recommendations based on the analysis.

Dataset Overview of the Project

The dataset used in this analysis was sourced from the Kaggle Walmart Sales Forecasting Competition. It consists of 17 columns and 1000 rows, offering a rich source of sales transitions data. The dataset includes the following columns:

- Invoice ID: Unique identifier for each transaction.
- Branch: Store branch where the transactions took place.
- City: City where the store branch is located.
- Customer type: Member or normal customer.
- Gender: Gender of the customer.
- Product line: Category of the product purchased.
- Unit Price: Price per unit of the product.
- Quantity: Number of units purchased.
- Tax 5%: Tax applied to the transaction.
- Total: Total amount for the transaction including tax.
- Date: Date of the transaction.
- Time: Time of the transaction.
- Payment: Payment method used.
- cogs: Cost of goods sold.
- Gross margin percentage: Percentage margin on the sales.
- Gross income: Gross income from the transaction.
- Rating: Customer rating for the transaction.

Data Types:

A. object: 9 columns (Invoice ID, Branch, City, Customer type, Gender, Product line, Date, Time, Payment)
B. float64: 7 columns (Unit price, Tax 5%, Total, cogs, gross margin percentage, gross income, Rating)
C. int64: 1 column (Quantity)

Components of the Project

The Walmart Sales Data project is structured around three main components that facilitate the extraction, analysis, and visualization of the data:

1) **SQL for Data Extraction, Cleaning, and Feature Engineering**:

   SQL (Structured Query Language) is a programming language designed for managing and manipulating databases. In the context of this project, SQL is used to extract relevant data from the Walmart sales database, clean the data by identifying and correcting errors or inconsistencies, and perform feature engineering. Feature engineering involves creating new features or modifying existing ones to improve the performance of machine learning models or to make the data more suitable for analysis.

2) **Python for Basic to Advanced Analysis**:

   Python is a versatile programming language that is widely used in data analysis and scientific computing. For the Walmart Sales Data project, Python is employed for a range of analytical tasks, from basic exploratory data analysis (EDA) to advanced techniques such as time series forecasting and predictive modeling. Python libraries like Pandas, NumPy, Matplotlib, and Seaborn are likely used for data manipulation, statistical analysis, and visualization. Additionally, Python's machine learning libraries like Scikit-learn may be utilized for more complex analyses like feature engineering and predictive modeling.

3) **Power BI for Interactive Visualizations**:

   Power BI is a business analytics service by Microsoft that provides interactive visualizations with self-service business intelligence capabilities. It is used in this project to create interactive dashboards and reports that allow stakeholders to explore the results of the analysis in a user-friendly manner. Power BI visualizations can help in identifying trends, patterns, and outliers in the data, making it easier for decision-makers to understand the implications of the analysis and take informed actions.

Analysis and Insights the Project

Based on the comprehensive analysis of the Walmart sales data, including visualizations, feature engineering, and insights across various dimensions, here are detailed analysis and insights of the project for optimizing business strategies and operations:

1. **Exploratory Data Analysis (EDA) Analysis:**

   A. Numerical Columns:

   - Unit price: Mean = 55.67, Std = 26.49, Min = 10.08, Max = 99.96
   - Quantity: Mean = 5.51, Std = 2.92, Min = 1, Max = 10
   - Tax 5%: Mean = 15.38, Std = 11.71, Min = 0.51, Max = 49.65
   - Total: Mean = 322.97, Std = 245.89, Min = 10.68, Max = 1042.65
   - cogs: Mean = 307.59, Std = 234.18, Min = 10.17, Max = 993.00
   - gross margin percentage: Constant value of 4.76%
   - gross income: Mean = 15.38, Std = 11.71, Min = 0.51, Max = 49.65
   - Rating: Mean = 6.97, Std = 1.72, Min = 4, Max = 10

   B. Categorical Columns:

   - Invoice ID: 1000 unique values
   - Branch: 3 unique values (A, B, C)
   - City: 3 unique values (Yangon, Naypyitaw, Mandalay)
   - Customer type: 2 unique values (Member, Normal)
   - Gender: 2 unique values (Male, Female)
   - Product Line: 6 unique values (fashion accessories, Food and beverage, Health and Beauty, Home and Lifestyle, Sports and Travel, Electronic accessories)
   - Date: 89 unique values
   - Time: 506 unique values
   - Payment: 3 unique values (Ewallet, Cash, Credit card)

   C. Observations:

   - There's a balanced distribution of male and female customers, with a slight preference for membership among customers.
   - Fashion accessories are the most popular product line, while Sports and travel are the least popular.
   - The average unit price is approximately 55.67, with a standard deviation of 26.49.
   - The average quantity per transaction is 5.51, with a standard deviation of 2.92.
   - The total amount spent per transaction varies significantly, from 10.68 to 1,042.65, with an average of 322.97.
   - Ewallets are the most popular payment method, suggesting a trend towards digital transactions.
   - Gross income per transaction aligns with the 5% tax and gross margin, ranging from 0.51 to 49.65, with an average of 15.38.
   - The average rating is around 7, with a significant number of ratings above 7, indicating overall customer satisfaction.

- Gross Margin Percentage is constant at 4.76% across all transactions, which might indicate a fixed margin across all products or branches.
- Total and cogs have a wide range and high variability, indicating varying transactions sizes and costs.

2. Visualizations Analysis:

A. **Distribution Plots for Numerical Variables**:

- This will show histogram with KDE plots for the variables: 'Unit price', 'Quantity', 'Total', 'Rating'.
- It helps me understand the distribution and destiny of these numerical variables.

B. **Count Plots for Categorical Variables:**

- This will show count plots for the variables: Branch, City, Customer Type, Gender, Product Line, and payments.
- It helps you see the frequency of each category within these variables.

C. **Correlation Heatmap**:

- This will show a heatmap of the correlation matrix for numerical variables.
- It helps me identify relationships between numerical variables, such as how strongly they are correlated with each other.

3. Feature Engineering Analysis:

- **Sales Patterns**:

  Understanding day, month, and hour patterns helps in forecasting demand and optimizing operations. For instance, if certain days or months see higher sales, you can plan for increased inventory and targeted marketing.

- **Customer Segmentation**:

  CLV (Customer Lifetime Value) segmentation allows for personalized marketing and improved customer retention strategies. High-value customers might appreciate exclusive offers, while low-value customers might benefit from incentives to boost their spending.

- **Product Strategy**:

  Analyzing product category diversity can guide product bundling and cross-selling strategies. If customers buying more categories tend to spend more, promoting product bundles or offering discounts on multiple categories can be effective.

4. Other Analysis and Insights:

A. Sales Performance by Branch and City:

a) **Sales by Branch:**

- **Branch Performance**: Identify which branch has the highest and lowest total sales. This can indicate which branches are performing better or worse.
- **Operational Adjustments**: A branch with significantly higher sales may require additional resources or inventory, while branches with lower sales might need targeted promotions or operational improvements.

- **Market Potential**: Analyze if the high-performing branches are located in areas with higher customer density or economic activity.

b) Sales by City:

- **City Performance**: Determine which cities generate the most and least revenue. This helps in understanding regional market potential and customer preferences.
- **Marketing Focus**: Cities with higher sales could be targeted for increased marketing efforts or expansion of product lines, while cities with lower sales might benefit from local promotions or market research.
- **Sales Trends**: Look for geographic trends. For example, if certain cities have consistently higher sales, they might have unique market characteristics or customer preferences that can be leveraged.

B. **Customer Demographics (Sales Distribution by Customer Type and Gender):**

a) Sales by Customer Type:

- **Customer Type Contribution**: Determine whether 'Members' or 'Normal' customers contribute more to overall sales. For instance, if 'Members' contribute significantly more, it suggests that loyalty programs or memberships are effective in driving sales.
- **Retention Strategies**: If 'Members' generate more sales, it reinforces the importance of maintaining and expanding membership programs. Conversely, if 'Normal' customers generate a substantial number of sales, there might be opportunities to convert them into members.
- **Marketing Focus**: Tailor marketing strategies based on the contribution of each customer type. For example, targeted promotions could be designed to increase sales among 'Normal' customers.

b) Gender-Based Sales Analysis:

- **Gender-Based Preferences**: Identify if one gender contributes more to sales than the other. For example, if 'Female' customers contribute more, it might indicate that marketing strategies or product lines are more appealing to women.
- **Targeted Marketing**: Use this information to develop gender-targeted marketing campaigns. If one gender shows higher spending, consider tailoring promotions and product recommendations accordingly.
- **Product Preferences**: Investigate if there are specific product lines or categories that are more popular among a particular gender, which can help in product placement and promotions.

C. **Product Line Performance (Based on Revenue and Sales Volume):**

a) **Revenue by Product Line:**

- **Top Performers**: Identify which product lines are the top revenue generators. This information can highlight which product categories are most profitable.

- **Strategic Focus**: Focus marketing and inventory efforts on high-revenue product lines to maximize returns. Consider expanding these lines or introducing similar products to capitalize on their success.
- **Resource Allocation**: Allocate more resources, such as marketing spend or shelf space, to high-revenue product lines to further boost their performance.

b) **Sales Volume by Product Line:**

- **Popular Products**: Determine which product lines are sold in the highest quantities. This indicates customer demand and product popularity.
- **Inventory Management**: Ensure that high-volume products are well-stocked to meet customer demand. Consider promotions or discounts to increase sales of lower-volume items.
- **Cross-Selling Opportunities**: Use popular product lines to drive cross-selling strategies. For example, if a particular line sells well, consider bundling it with other products to increase overall sales.

D. **Pricing and Quantity (Unit Price vs. Quantity Sold, and Total Sales Based on Unit Price and Quantity):**

a) Scatter Plot of Unit Price vs. Quantity Sold:

- **Price Sensitivity**: Analyze if there is a trend where higher prices lead to lower quantities sold, or vice versa. This can indicate price sensitivity among customers.
- **Sales Strategy**: If high unit prices correlate with lower quantities sold, consider exploring price elasticity and adjusting pricing strategies to optimize sales.
- **Outliers**: Identify any outliers or anomalies in the data. For example, very high-priced items with unexpectedly high quantities sold could indicate special promotions or discounts.

b) Heatmap of Total Sales Based on Unit Price and Quantity:

- **Sales Patterns**: Determine which combinations of unit price and quantity result in the highest total sales. This helps in understanding which pricing strategies and quantities yield the most revenue.
- **Optimization Opportunities**: Identify price and quantity combinations that lead to higher sales. For instance, finding an optimal price point that maximizes sales volume can guide pricing decisions.
- **Product Pricing Strategy**: Use this information to set prices that balance between unit price and quantity sold, aiming for combinations that boost overall sales.

E. **Payment Methods (Based on Popular Payment Methods, and Payment Method Impact on Sales)**

a) Popular Payment Methods:

- **Customer Preferences**: Identify which payment methods are most popular among customers. This insight helps in understanding customer preferences and tailoring services accordingly.

- **Payment Options**: If certain payment methods dominate, it might indicate customer convenience or trust issues with other methods. It's essential to ensure that the most popular methods are always available and functioning smoothly.
- **Promotion and Loyalty Programs**: Popular payment methods could be linked to specific promotions or loyalty programs. Understanding customer preferences can help in designing targeted promotions or offers.

b) Payment Method Impact on Sales:

- **Revenue Contribution**: Determine which payment methods contribute the most to overall sales. This can indicate which payment methods are favoured for higher-value transactions.
- **Transaction Trends**: Analyze if there's a trend where certain payment methods are associated with higher average transaction values. For instance, credit cards might be used for larger purchases compared to cash.
- **Payment Method Fees**: Consider the cost associated with different payment methods. High transaction fees on certain methods can impact net revenue, and if these methods are popular, strategies to minimize costs should be considered.

F. **Sales Trends Over Time (Based on Daily Sales, Monthly Sales, and Hourly Sales)**

a) Daily Sales Trend:

- **Seasonal Patterns**: Identify patterns or fluctuations in daily sales, which can indicate seasonal trends, holidays, or specific events that influence sales.
- **Peak and Low Periods**: Recognize the days with the highest and lowest sales, helping in planning inventory, staffing, and promotional activities accordingly.
- **Anomalies**: Spot any unusual spikes or drops in sales, which could be linked to external factors like promotions, weather, or market events.

b) Monthly Sales Trend:

- **Growth Analysis**: Analyze the overall growth or decline in sales over the months. This can indicate the effectiveness of business strategies or external economic conditions.
- **Seasonal Impact**: Understand the impact of seasonal factors on sales, such as holiday seasons or back-to-school periods, which can drive higher sales in certain months.
- **Planning**: Use monthly sales trends to plan for future inventory, marketing campaigns, and resource allocation.

c) Hourly Sales Trend:

- **Peak Hours**: Identify the hours with the highest sales, helping in optimizing store operations, staffing, and promotional efforts during peak times.
- **Customer Behaviour**: Understand customer shopping behaviour, such as whether customers prefer shopping in the morning, afternoon, or evening.
- **Operational Efficiency**: Plan store hours and staffing levels based on peak sales times to maximize efficiency and customer service.

G. **Customer Feedback (Based on Rating Distribution and Area of Improvements on Low Ratings):**

a) Rating Distribution:

- **Customer Satisfaction**: Understanding the overall distribution of ratings provides a snapshot of customer satisfaction. A higher number of positive ratings indicates good customer experiences, while a significant number of low ratings might point to areas needing improvement.
- **Quality Assessment**: Analyzing the spread of ratings helps in assessing the quality and consistency of products or services offered.
- **Customer Expectations**: The distribution can reveal whether customer expectations are being met, exceeded, or unmet.

b) Areas for Improvement Based on Low Ratings:

- **Identifying Problem Areas**: Highlighting product lines with the lowest ratings helps in identifying specific areas or products that may be underperforming or causing customer dissatisfaction.
- **Quality Control**: Focus on product lines that receive consistent low ratings for quality checks, product redesigns, or discontinuation if necessary.
- **Customer Feedback Utilization**: Leverage customer feedback to make targeted improvements in product lines or services, aiming to address the issues highlighted by the low ratings.

Findings of the Project

Based on the comprehensive analysis of the Walmart sales data, including visualizations, feature engineering, and insights across various dimensions, here are some key findings for optimizing business strategies and operations:

1. **Branch and City Performance:**
    - Branch A and the city of Yangon demonstrated strong sales performance, suggesting successful strategies that can be replicated in other locations.
    - Branch C and cities like Naypyitaw showed lower sales, indicating potential areas for strategic improvement.
2. **Customer Demographics:**
    - A nearly equal split between Member and Normal customers was observed, with no significant difference in gender distribution. This indicates an opportunity for targeted marketing to convert more Normal customers into Members and to tailor promotions based on gender-specific preferences.
3. **Product Line Performance:**
    - Fashion accessories and Food and beverages emerged as top-performing product lines, while Sports and travel lagged. This suggests a need for targeted efforts to boost sales in underperforming categories.
4. **Pricing and Quantity:**
    - The analysis of unit price and quantity sold revealed that competitive pricing strategies are essential for maintaining sales volume. The potential for volume discounts and bulk purchase promotions was identified.
5. **Payment Methods:**
    - Ewallets were the most popular payment method, highlighting a trend towards digital transactions. This underscores the importance of optimizing digital payment options and possibly introducing exclusive offers for digital payments.
6. **Sales Trends Over Time:**
    - Sales trends indicated peak shopping periods and hours, providing valuable data for inventory management and promotional timing.
7. **Customer Feedback:**
    - While overall ratings were positive, some product lines received lower ratings, indicating areas for improvement. Highlighting positive reviews and addressing issues in low-rated product lines can enhance customer satisfaction and trust.

Recommendations of the Project

Based on the comprehensive analysis of the Walmart sales data, including visualizations, feature engineering, and insights across various dimensions, here are some detailed recommendations for optimizing business strategies and operations:

1) **Product Line Optimization:**

- **Focus on Best-Selling Lines**: The Fashion accessories and Food and beverages product lines are the top performers. Increase inventory and marketing efforts for these categories to maximize sales.
- **Revitalize Low-Performing Categories**: The Sports and travel line has lower sales. Conduct market research to understand customer preferences and consider introducing new or complementary products within this category.
- **Cross-Selling Opportunities**: Encourage customers to purchase from multiple product lines by bundling popular items with underperforming ones, potentially increasing overall sales.

2) **Customer Segmentation and Targeting:**

- **Member vs. Normal Customers**: With nearly equal splits, consider exclusive promotions or loyalty programs to convert more Normal customers into Members. Analyze purchase behaviour differences to tailor marketing campaigns accordingly.
- **Gender-Specific Campaigns**: The balanced gender distribution allows for targeted campaigns for both male and female customers. Consider developing gender-specific promotions, especially around holidays or special events.

3) **Branch and City Strategies:**

- **Branch A and Yangon Dominance**: Invest in these high-performing locations with more inventory, staff, and localized marketing. Understand what drives their success and replicate those strategies in other branches.
- **Underperforming Branches**: Analyze the challenges in branches B and C, as well as cities like Naypyitaw and Mandalay. Implement targeted campaigns, community engagement, and possibly store renovations to boost sales.

4) **Payment Methods:**

- **Emphasize Digital Payments**: The preference for Ewallet suggests a shift towards digital transactions. Ensure smooth, secure, and varied digital payment options. Consider partnerships with popular e-wallet providers for exclusive discounts or cashback offers.
- **Promote Lesser-Used Methods**: Encourage the use of Credit cards through promotional offers, such as discounts for first-time users or rewards for frequent purchases.

5) **Pricing and Quantity Strategies:**

- **Competitive Pricing**: The distribution of unit prices suggests a competitive pricing strategy. Continue monitoring competitor prices and adjust accordingly to remain attractive to price-sensitive customers.

- **Volume Discounts**: Introduce discounts for bulk purchases, particularly in product lines with higher price points. This can encourage customers to buy in larger quantities.

6) **Customer Feedback and Ratings:**

- **Address Low Ratings**: Focus on product lines receiving lower ratings. Implement quality improvements, gather more detailed customer feedback, and address specific concerns.
- **Highlight Positive Reviews**: Use high customer ratings in marketing materials to build trust and attract new customers. Showcase testimonials and case studies.

7) **Sales Trends and Inventory Management:**

- **Seasonal Trends**: Use the insights from the monthly and daily sales trends to prepare for peak shopping periods. Adjust inventory levels and marketing efforts accordingly to meet demand.
- **Time-Based Promotions**: The hourly sales trends indicate peak shopping hours. Consider time-limited promotions to capitalize on high-traffic periods.

8) **Marketing and Promotions:**

- **Targeted Marketing Campaigns**: Use customer demographic data (such as age, gender, and customer type) to craft targeted campaigns. Personalized marketing can significantly enhance customer engagement and conversion rates.
- **Utilize Digital Channels**: With a significant portion of transactions happening through Ewallets, leverage digital marketing channels like social media, email campaigns, and mobile app notifications for promotions.

9) **Data-Driven Decision Making:**

- **Continuous Data Monitoring**: Regularly update and analyze sales data to identify emerging trends and shifts in customer behaviour. Use dashboards and real-time analytics tools for continuous monitoring.
- **A/B Testing**: Implement A/B testing for promotions, pricing strategies, and product placements to identify the most effective tactics.

10) **Customer Experience Enhancement:**

- **Streamline Checkout Process**: Ensure a seamless checkout process, both online and offline, to reduce cart abandonment rates.
- **Customer Service Excellence**: Train staff to handle customer inquiries efficiently and politely, both in-store and online. Address any issues promptly to maintain high customer satisfaction levels.

These recommendations aim to enhance overall business performance, increase customer satisfaction, and drive growth. Implementing these strategies can help Walmart optimize its operations and capitalize on market opportunities.

Conclusion of the Project

This project involved an in-depth analysis of Walmart's sales data, encompassing various dimensions such as sales performance by branch and city, customer demographics, product line performance, pricing and quantity analysis, payment methods, sales trends over time, and customer feedback. The comprehensive exploratory data analysis (EDA) and feature engineering provided valuable insights that can drive strategic business decisions. This analysis has provided Walmart with a comprehensive understanding of its sales dynamics, customer preferences, and potential areas for growth. By implementing the recommended strategies, Walmart can optimize its operations, improve customer satisfaction, and drive overall business growth.

The project underscores the value of data-driven decision-making in retail, demonstrating how detailed data analysis can uncover hidden opportunities and provide actionable insights. Moving forward, continuous data monitoring and iterative improvements will be crucial for sustaining and enhancing Walmart's market position. This project serves as a foundation for further analyses and strategic planning, enabling Walmart to adapt to changing market conditions and customer needs effectively. Based on the findings, several actionable recommendations were provided, including optimizing product lines, enhancing customer targeting, focusing on branch-specific strategies, leveraging digital payment methods, refining pricing strategies, and improving customer service.