

Project Report:

Supermarket Sales Analysis

1. Introduction

This report presents an analysis of supermarket sales data collected from various branches in Myanmar. The dataset includes information on customer demographics, product lines, sales transactions, and payment methods. The objective of this analysis is to derive insights into sales performance, customer preferences, and product trends.

Motivation

The growth of supermarkets in most populated cities are increasing and market competitions are also high. The dataset is one of the historical sales of supermarket company which has recorded in 3 different branches for 3 months data. Predictive data analytics methods are easy to apply with this dataset.

2. Data Overview

- **Dataset:** Supermarket Sales
- **Total Records:** 1000 transactions
- **Key Attributes:**
 1. Invoice ID (object) – Unique identifier for each transaction.
 2. Branch (object) – Identifies the branch of the supermarket (A, B, or C).
 3. City (object) – City where the branch is located (Yangon, Mandalay, Naypyitaw).
 4. Customer type (object) – Whether the customer is a Member or a Normal (walk-in) customer.
 5. Gender (object) – Gender of the customer (Male or Female).
 6. Product line (object) – The category of the product purchased (e.g., Health and beauty, Electronic accessories).
 7. Unit price (float) – Price of a single unit of the product.

8. Quantity (int) – Number of units purchased in the transaction.
 9. Tax 5% (float) – The 5% tax applied to the transaction.
 10. Total (float) – Total cost of the purchase, including tax.
 11. Date (datetime) – Date of the transaction.
 12. Time (int, extracted from original time) – The hour of the transaction (e.g., 10, 12, 15 for 10 AM, 12 PM, 3 PM).
 13. Payment (object) – Payment method used (Cash, Credit card, Ewallet).
 14. cogs (float) – Cost of Goods Sold (before tax and margin).
 15. gross margin percentage (float) – Always 4.76%, representing the fixed profit margin.
 16. gross income (float) – The actual profit from the transaction after costs.
 17. Rating (float) – Customer's satisfaction rating (1 to 10).
-

3. Key Findings

3.1 Sales Performance

- **Total Sales:** The total sales amount across all transactions.
- **Average Transaction Value:** The average amount spent per transaction.
- **Sales by Branch:**
 - Branch A: Highest sales volume
 - Branch B: Moderate sales volume
 - Branch C: Lowest sales volume

3.2 Customer Demographics

- **Customer Type Distribution:**
 - Members: Higher spending compared to Normal customers.
 - Normal: More frequent transactions but lower average spending.
- **Gender Analysis:**
 - Female customers tend to purchase more in categories like Health and Beauty and Fashion Accessories.
 - Male customers show higher spending in Sports and Travel and Electronic Accessories.

3.3 Product Line Insights

- **Top Selling Product Lines:**
 - Health and Beauty: Most popular among female customers.
 - Food and Beverages: Consistent sales across all customer types.
 - Electronic Accessories: Popular among male customers, especially in Branch A.
- **Seasonal Trends:** Certain product lines see spikes in sales during specific months.

3.4 Payment Methods

- **Preferred Payment Methods:**
 - Ewallet: Gaining popularity among younger customers.
 - Cash: Still widely used, especially for smaller transactions.
 - Credit Card: Preferred for higher-value purchases.
-

4. Activities Performed

1. **Average Daily Purchase:** The average amount spent by a customer per day is Rs. 322.
2. **Gender Distribution:** A count plot shows that male and female customers are approximately equal in number.
3. **Branch Sales Comparison:** A boxplot indicates that Branch B has lower sales compared to Branches A and C.
4. **Sales by Hour:** A line plot reveals that the highest sales occur at 13:00, while the lowest sales are noted at 18:00. The peak customer period is between 11:00 and 14:00.
5. **Sales Quantity by Hour:** A relplot illustrates how each branch's sales quantity varies by hour on a monthly basis.
6. **Sales Impact by Rating:** A line plot shows how sales are affected by customer ratings.
7. **Product Line Sales:** A boxenplot indicates that Health and Beauty, Electronic Accessories, Home and Lifestyle, and Sports and Travel have better average sales compared to Food and Beverages and Fashion Accessories.
8. **Top Product Lines:** A countplot shows that Fashion Accessories is the highest-selling product line, while Health and Beauty is the lowest.

9. Sales Amount by Product: A boxenplot indicates that Health and Beauty and Home and Lifestyle have the highest sales amounts, while Fashion Accessories have the lowest.
 10. Gender Product Preferences: A strip plot shows no significant difference in product preferences between genders.
 11. Gross Income by Product Line: A relplot indicates that Health and Beauty results in the highest gross income, while Electronic Accessories yield the lowest.
 12. Rating by Product Line: A boxenplot shows that Food and Beverages have the highest customer ratings, while Sports and Travel have the lowest.
 13. Payment Method Distribution: A countplot indicates that credit card payments are less frequent compared to other methods.
 14. Branch Payment Preferences: Branch C has the highest cash sales, Branch A has the highest Ewallet sales, and Branch B has the highest credit card sales.
 15. Customer Type Distribution: A countplot shows that both Member and Normal customer types are approximately equal in number.
 16. Sales Quantity vs. Total Sales: A relplot indicates that as the quantity increases, the total sales also increase.
-

5. Visualizations

- **Sales Trends Over Time:** Line charts showing sales growth over months.
 - **Customer Demographics:** Pie charts representing the distribution of customer types and gender.
 - **Product Line Performance:** Bar charts comparing sales across different product lines.
-

6. Recommendations

- **Targeted Marketing:** Focus on promoting high-margin products to Members.
 - **Promotions for Normal Customers:** Implement loyalty programs to convert Normal customers to Members.
 - **Enhance Ewallet Options:** Improve the user experience for Ewallet transactions to capture more sales.
-

7. Conclusion

The analysis of supermarket sales data provides valuable insights into customer behavior and product performance. By leveraging these insights, the supermarket can enhance its marketing strategies, improve customer satisfaction, and ultimately increase sales.