Amazon Sales Data Analysis Report

Peiris K.N.W | 21020736 | 2021/IS/073

Code & dataset - https://github.com/NisalWishwajith/amazon_sales_data_analysis.git

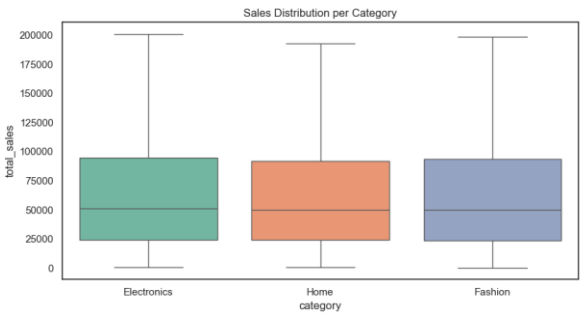
01. Introduction

In this assignment, we analyze an Amazon sales dataset to understand business performance and identify important patterns. The main goal is to explore sales trends, customer payment behavior, and price distribution. This analysis helps the company understand how products perform, which payment methods customers prefer, and how pricing affects total sales. The findings can support better business decisions and improve sales strategies.

02. Data Analytics Methodology

The data set used for this analysis contains 10,000 sales records and 21 features such as product category, price, quantity, discount, shipping cost, total sales, and payment method. Before analysis, the dataset was checked for missing values and data consistency. Basic data cleaning and preprocessing were performed using Python. Then exploratory data analysis and visualization techniques were applied to identify patterns and relationships in the data.

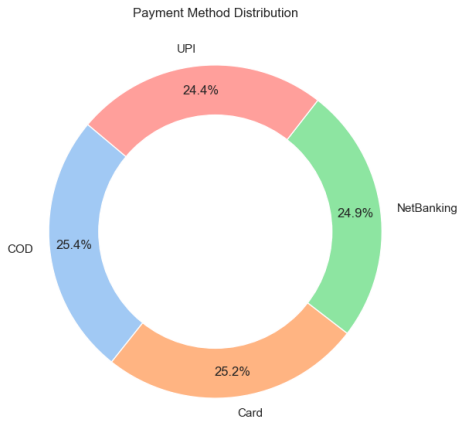
03. Exploratory Data Analysis (EDA)



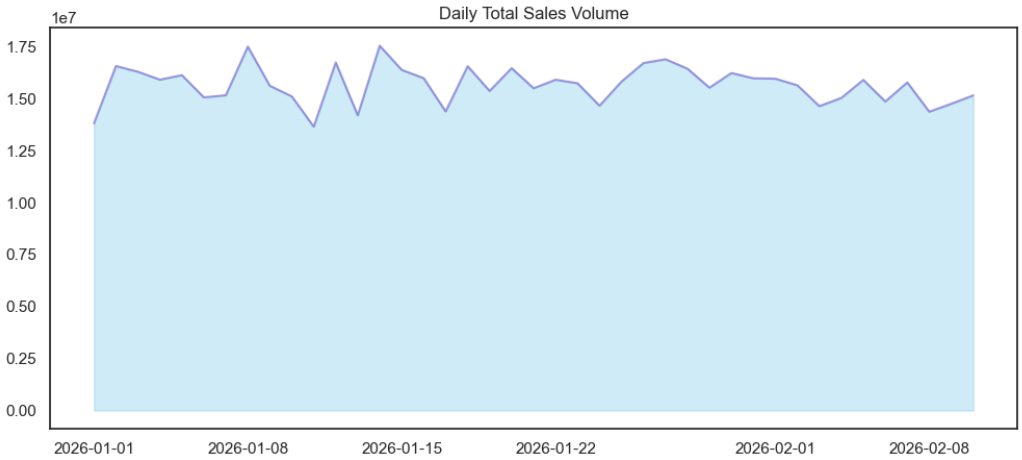
The boxplot shows that sales across Electronics, Home, and Fashion range from near 0 up to about **200,000**, with median values around **50,000 - 60,000**. Fashion and Electronics show slightly higher upper ranges, indicating higher potential revenue,

while all categories show wide variation in sales values.

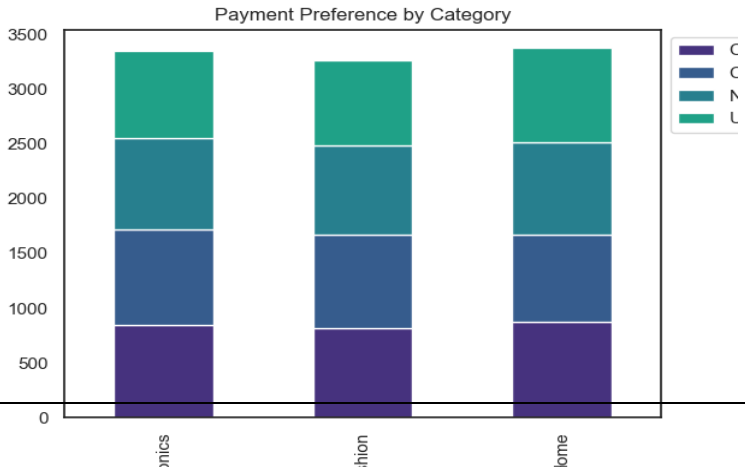
The chart shows an almost equal distribution of payment methods, with **COD (25.4%) being the most used**, followed by Card (25.2%), Net Banking (24.9%), and UPI (24.4%). This indicates customers use all payment methods similarly, but cash on delivery has a slightly higher preference.

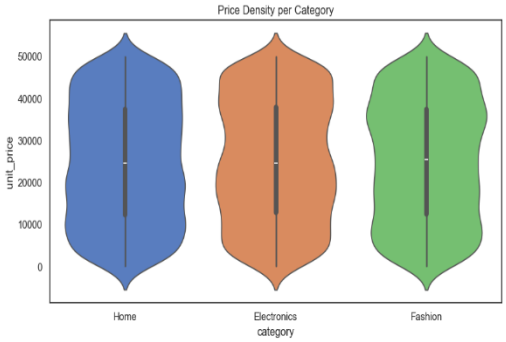


The area chart shows daily total sales fluctuating between approximately **1.3 million and 1.5 million**. Sales remain relatively stable with small variations, indicating consistent business performance over time.



The stacked chart shows similar payment usage across Electronics, Fashion, and Home categories, with total transactions around **3300 - 3400 per category**. Each payment method contributes nearly equal proportions, showing consistent customer payment behavior across product types.





The violin plot shows product prices ranging roughly from **5,000 to 45,000**, with median prices around **25,000 - 30,000** across categories. Fashion shows slightly wider price distribution, indicating greater price variation compared to Electronics and Home.

03. Key Findings

The analysis shows that sales values vary widely, indicating different product price ranges. Some products generate very high sales compared to others. Customers use multiple payment methods, but some methods are more popular than others. Understanding customer payment preference can help improve customer experience. Sales trends show changes over time, which can help businesses plan future marketing and inventory strategies. Price distribution analysis shows that different product categories have different pricing patterns.

04. Business Recommendations

1. Focus on high performing products to increase revenue.

2. Improve support for popular payment methods to enhance customer satisfaction.

3. Monitor sales trends regularly to plan inventory and promotions.

4. Review pricing strategies across categories to remain competitive in the market.

05. Conclusion

This analysis provided useful insights into sales performance, customer behavior, and pricing patterns. By using data-driven decisions, the company can improve sales performance, customer satisfaction, and business growth.