

Exploratory Data Analysis

Retail Purchases

Submitted by
Jashan Grover
22119027
II Year, Production and Industrial Engineering



Introduction

This report presents an exploratory analysis of a retail purchases dataset sourced from Kaggle. The dataset includes over 300,000 entries spread across 30 columns, providing insights into consumer behavior, product preferences, and geographical trends in retail purchases.

Retail businesses face challenges in understanding dynamic consumer preferences and market trends. Exploratory Data Analysis (EDA) plays a crucial role in extracting meaningful insights from data, guiding decisions to enhance customer satisfaction, optimize inventory management, and drive growth.

Using statistical methods and data visualization techniques, this analysis explores the demographic distribution of buyers, regional variations in purchasing behavior, seasonal trends, and transactional preferences. It also examines correlations between income levels and spending habits, as well as the performance of different product categories and brands.

The findings aim to provide actionable insights for improving marketing strategies, operational efficiency, and customer engagement within the retail sector. This report will guide stakeholders in making informed decisions that align with market dynamics and consumer expectations.

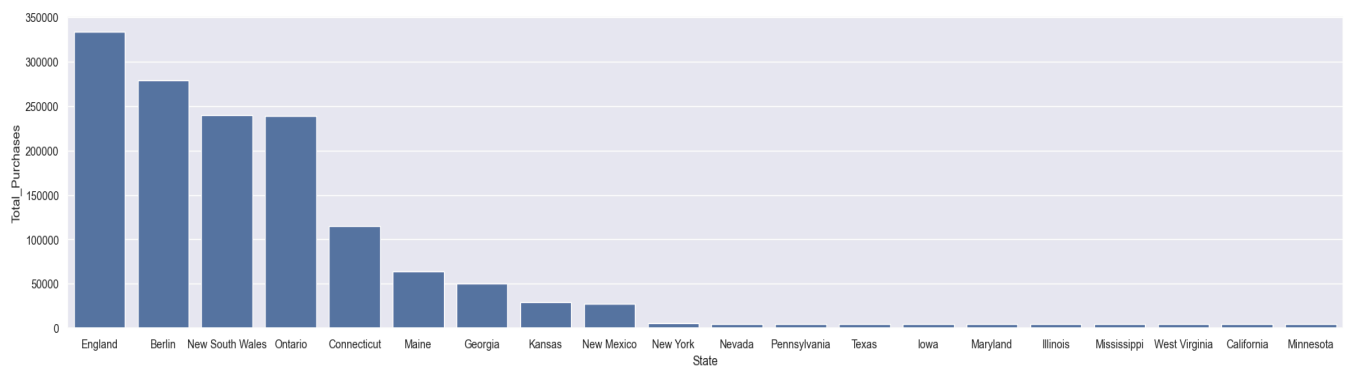
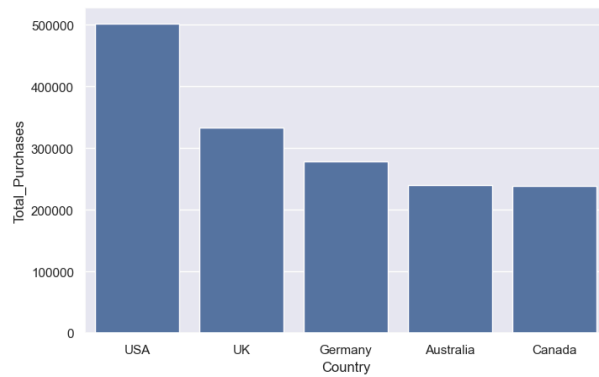
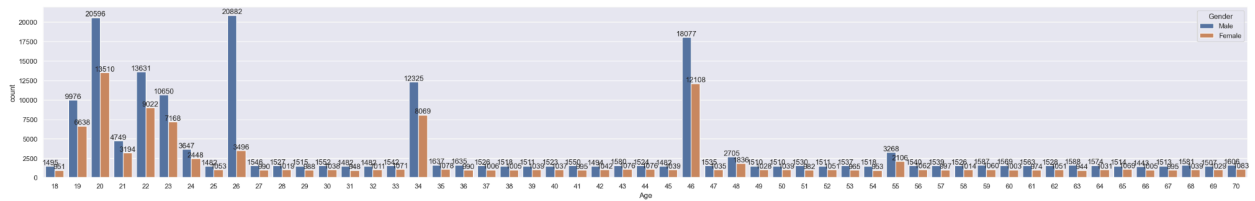
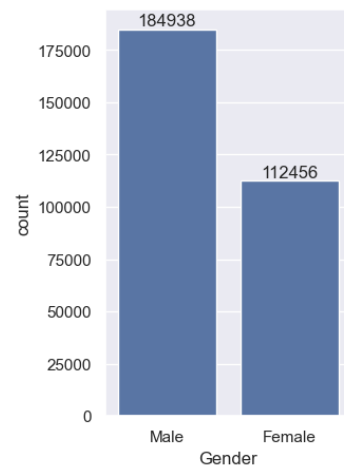
Methodology

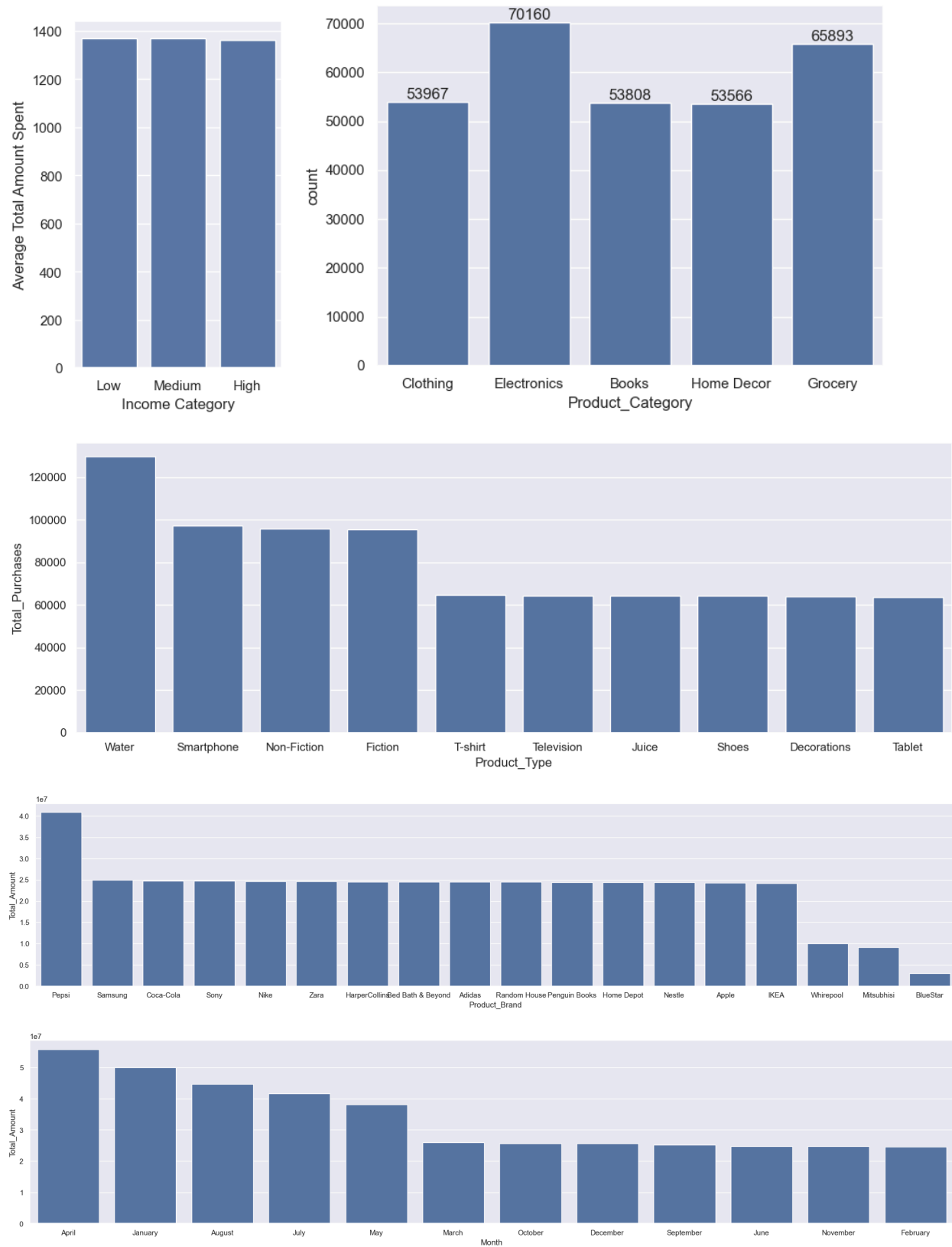
This exploratory analysis of the retail transactions dataset was performed within Jupyter Notebook, utilizing essential Python libraries such as NumPy, Pandas, Matplotlib, and Seaborn. The methodology employed a systematic approach aimed at extracting valuable insights from the dataset following comprehensive data preprocessing and meticulous cleaning. Some unnecessary columns that were not relevant to the main analysis were dropped.

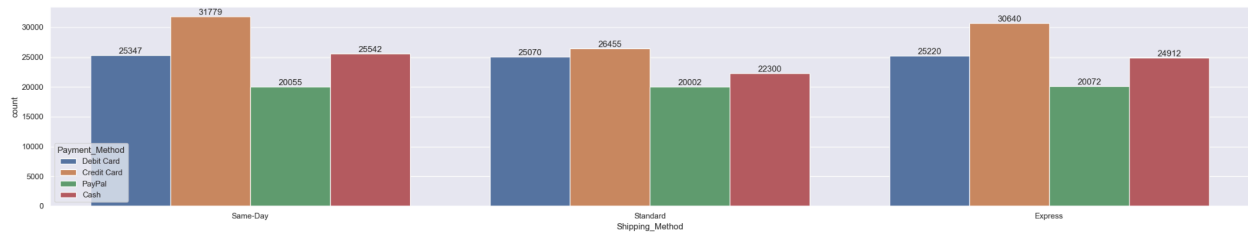
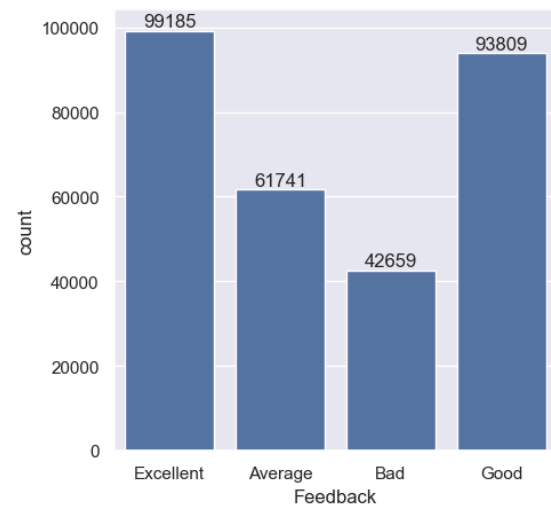


Key Findings/Insights

1. There are more male buyers, and they spend more money than female buyers.
2. Buying is highest among young people aged 19-26, with peaks also seen at ages 34 and 46. Male buyers outnumber female buyers in every age group.
3. Total purchases and total amount spent follow similar patterns across countries, with the USA having the highest figures.
4. States show similar trends in total purchases and spending, with England (a state of the UK) leading among states. The first USA state appears fifth despite the USA being the largest buyer overall.
5. The average total amount spent is nearly the same across all types of buyers and income categories.
6. People spend the most on electronics, followed by groceries, clothing, books, and home decor.
7. Water is the top selling and revenue generating product, with Pepsi being the highest revenue generating brand.
8. April experiences the highest shopping activity, indicating it as the peak season for purchases.
9. A significant portion of buyers rate products as average or below average in satisfaction.
10. All shipping methods are equally popular among buyers, with credit cards being the most preferred payment method in every scenario.









Conclusion

1. Demographic Insights:

- Male buyers outnumber female buyers and tend to spend more overall.
- Younger demographics, particularly aged 19-26, show the highest buying activity, with notable peaks at ages 34 and 46, possibly suggesting specific life stages impacting buying behavior.

2. Geographical Trends:

- The USA leads in both total purchases and total amount spent, across the countries.
- England (considered as a state of the UK in the dataset) is at the topmost position among all the states, with the first USA state at fifth position, despite the USA being the largest buyer overall.

3. Income and Spending Patterns:

- Average total spending remains consistent across all income categories, indicating that income level does not significantly influence spending behavior.

4. Product Preferences:

- Electronics are the most purchased category, followed by grocery, clothing, books, and home decor in descending order of spending.

5. Product and Brand Performance:

- Water is the highest selling product in terms of units moved, while Pepsi stands out as the top revenue generating brand.

6. Seasonal and Behavioral Insights:

- April emerges as the peak shopping season, likely due to various factors such as seasonal discounts, tax refunds, or holiday-related spending.
- A significant proportion of buyers rate products as average or below average, indicating potential areas for product improvement or customer satisfaction enhancements.

7. Transactional Behavior:

- All types of shipping methods are equally popular among buyers, with credit cards being the most preferred payment method across various scenarios.



Recommendations

1. Targeted Marketing and Product Development:

- Capitalize on the strong preference for electronics by focusing marketing efforts and product development in this category. Consider introducing new and innovative electronics products to maintain consumer interest.
- Explore strategies to increase sales in other categories such as clothing, books, and home decor, potentially through bundled promotions or targeted marketing campaigns aimed at specific demographic segments.

2. Enhanced Customer Experience:

- Address the feedback regarding average or below average product satisfaction by investing in product quality improvements and customer service enhancements.
- Implement user friendly policies for returns and exchanges to improve overall customer satisfaction and loyalty.

3. Geographical Expansion Strategies:

- Given the prominence of the USA as a country and certain states including England (UK), consider expanding operations or intensifying marketing efforts in these regions to further capitalize on high transaction volumes and spending levels.

4. Seasonal Promotions:

- Leverage the peak shopping season in April by launching seasonal promotions, special offers, and exclusive deals to maximize sales during this period.
- Analyze historical data to anticipate consumer preferences and buying patterns during peak seasons, enabling proactive planning and targeted marketing strategies.

5. Payment and Shipping Optimization:

- Continue to support all popular payment methods, especially credit cards, while ensuring seamless checkout experiences to minimize cart abandonment rates.
- Maintain high standards of shipping efficiency and reliability across all shipping methods to uphold customer satisfaction and loyalty.