

An Exploratory Data Analysis of Festive Sales for Consumer Insights

Introduction

The objective of this project is to perform an exploratory data analysis (EDA) on a dataset of festive sales, specifically from Diwali sales, to identify trends and gain insights into consumer behavior. By analyzing various demographic and transactional attributes, we aim to understand the key factors driving sales and to pinpoint the most valuable customer segments.

Data Cleaning and Preparation

The initial dataset consisted of 11,251 rows and 15 columns. Before analysis, several data cleaning and preparation steps were undertaken to ensure the quality and integrity of the data:

- 1. Initial Data Inspection:** The data was loaded and its structure was examined.
- 2. Dropping Irrelevant Columns:** The 'Status' and 'unnamed1' columns were identified as being empty or irrelevant to the analysis and were subsequently dropped.
- 3. Handling Null Values:** The dataset was checked for null values. The 'Amount' column was found to have 12 null entries. These rows were removed to maintain the accuracy of sales-related calculations.
- 4. Data Type Conversion:** The 'Amount' column was converted from a float to an integer data type for consistency.

After these cleaning steps, the dataset was ready for exploratory analysis.

Exploratory Data Analysis and Visualizations

1. Gender Analysis

- **Gender Distribution of Buyers:**

A bar chart was created to visualize the number of buyers by gender. The analysis revealed that female buyers significantly outnumbered male buyers.

- **Purchasing Power by Gender:**

The total sales amount was analyzed by gender. The results showed that the purchasing power of females was substantially greater than that of males.

Conclusion on Gender: The majority of buyers during the festive season are female, and they also contribute more to the total sales amount.

2. Age Analysis

- **Age Group Distribution:**

A count plot was generated to show the distribution of buyers across different age groups, segmented by gender.

- **Sales by Age Group:**

The total sales amount was analyzed by age group.

Conclusion on Age: The age group of 26-35 years represents the largest segment of buyers, with females in this category being the most prominent. This age group also contributes the most to the total sales amount.

3. Geographical Analysis (State)

- **Top 10 States by Orders:**

The top 10 states with the highest number of orders were identified and visualized.

- **Top 10 States by Sales Amount:**

The top 10 states with the highest total sales amount were also analyzed.

Conclusion on State: Uttar Pradesh, Maharashtra, and Karnataka are the top three states, both in terms of the number of orders placed

and the total sales amount generated.

4. Marital Status Analysis

- **Distribution by Marital Status:**

A count plot was used to show the number of married versus unmarried buyers.

- **Sales by Marital Status and Gender:**

The total sales amount was analyzed based on marital status and gender.

Conclusion on Marital Status: The data indicates that married individuals, particularly women, are the most significant group of buyers and have the highest purchasing power.

5. Occupation Analysis

- **Distribution by Occupation:**

The number of buyers from different occupational sectors was visualized.

- **Sales by Occupation:**

The total sales amount was analyzed by occupation.

Conclusion on Occupation: The highest number of buyers come from the IT, Healthcare, and Aviation sectors, and these sectors also contribute the most to the total sales amount.

6. Product Category Analysis

- **Top Selling Product Categories by Count:**

The most frequently purchased product categories were identified.

- **Top Selling Product Categories by Sales Amount:**

The product categories that generated the highest sales amount were analyzed.

Conclusion on Product Category: The Food, Clothing & Apparel, and Electronics & Gadgets categories are the most popular among buyers

and generate the highest revenue.

7. Top Selling Products

- **Top 10 Products by Orders:**

The top 10 most ordered products were identified to understand specific product demand.

Overall Conclusion

The exploratory data analysis provides a clear profile of the key consumer segment during the festive sales period. The most significant insights are:

- **Primary Customer Profile:** Married women between the ages of 26-35.
- **Key Geographic Markets:** Uttar Pradesh, Maharashtra, and Karnataka.
- **Top Professional Segments:** Individuals working in the IT, Healthcare, and Aviation sectors.
- **Most Popular Product Categories:** Food, Clothing & Apparel, and Electronics & Gadgets.

This detailed analysis suggests that marketing and sales strategies should be targeted towards this specific demographic to maximize revenue during the festive season.