Diwali Sales Analysis Report

Overview

The Diwali Sales dataset consists of 11,251 entries with 15 columns. The dataset contains various features related to customers, including demographic information, product details, order information, and sales data. The analysis will focus on identifying trends and insights based on different factors like gender, age group, marital status, state, and product category.

1. Data Preprocessing

Missing Data: Initially, there were 12 missing values in the 'Amount' column, which were removed. The columns `Status` and `unnamed1` had no meaningful data, so they were dropped.

Data Types: The 'Amount' column, which was initially of type float64, was converted to integer type for better readability and precision in the analysis.

Renaming Columns: The 'Marital_Status' column was renamed to 'Shaadi' for clarity, as it indicates whether the customer is married or not.

2. Data Summary

The cleaned data now consists of the following columns:

- User_ID: Unique identifier for each user
- Cust_name: Name of the customer
- Product_ID: Unique product identifier
- Gender: Gender of the customer (F/M)
- Age Group: Age group of the customer
- Age: Exact age of the customer
- Shaadi (Marital Status): 0 for unmarried, 1 for married
- State: State in India where the customer is located
- Zone: Geographical zone of the customer
- Occupation: Customer's occupation

- Product_Category: Category of the purchased product

- Orders: Number of orders made by the customer

- Amount: Total amount spent by the customer

3. Exploratory Data Analysis (EDA)

The analysis was performed on the cleaned dataset using various plots and summaries to identify key trends.

Gender Distribution: The dataset shows a higher number of female customers than male customers.

Purchasing Power by Gender: Female customers have spent more on average than male customers. A bar chart depicting total sales for each gender reveals that females contributed significantly more to the total sales compared to males.

Age Distribution: The age group 26-35 is the most active group, with the highest number of customers. This is followed by the 36-45 age group.

Total Sales by Age Group: The highest sales were recorded from the 26-35 age group, particularly among female customers, showing that this group is not only the largest but also the most engaged in purchasing.

State Distribution: Uttar Pradesh, Maharashtra, and Karnataka were the top three states with the most orders.

Total Sales by State: Uttar Pradesh leads in terms of the total number of orders, while Maharashtra and Karnataka are the top two states in terms of total sales.

Marital Status Distribution: Most customers in the dataset are married, with a smaller portion being unmarried.

Sales by Marital Status: Married customers tend to spend more on average than unmarried ones. A bar plot shows the total amount spent by customers based on their marital status.

Occupation Distribution: The most common occupation among customers is in the 'Office' category, followed by sectors like healthcare, government, and construction.

4. Insights and Observations

Gender Influence: Female customers are more likely to purchase products and spend higher amounts compared to their male counterparts.

Age Trends: The 26-35 age group is the most active, both in terms of the number of purchases and the total spending. Younger and older age groups show less engagement, which may suggest that marketing efforts should target the 26-35 age group.

State Preferences: The sales data shows regional preferences, with Uttar Pradesh, Maharashtra, and Karnataka driving the highest sales. This insight can be useful for targeted marketing in these regions.

Marital Status: Married individuals tend to spend more during Diwali, possibly due to larger family purchases or gifts.

Customize Offers for Office-Goers: As office workers represent a large portion of the consumer base, promotional campaigns offering products suited for office use (e.g., work-related gifts or accessories) could be more effective.

5. Recommendations

Target Female Customers: Since women are responsible for a larger portion of the sales, marketing campaigns should focus on products that appeal to this demographic.

Focus on the 26-35 Age Group: As this group shows the highest spending behavior, personalized offers or discounts tailored for this age group could yield high returns.

Regional Targeting: Given the high sales in Uttar Pradesh, Maharashtra, and Karnataka, targeted promotions in these regions could help maximize sales.

Promote Family Products: Since married individuals tend to spend more, it would be beneficial to promote family-oriented products during Diwali sales.

Customize Offers for Office-Goers: As office workers represent a large portion of the consumer base, promotional campaigns offering products suited for office use (e.g., work-related gifts or accessories) could be more effective.

6. Conclusion

The Diwali sales dataset reveals significant insights about customer behavior, preferences, and trends. The 26-35 age group, female customers, and people from specific states like Uttar Pradesh, Maharashtra, and Karnataka are crucial to driving sales during Diwali. By understanding these patterns, businesses can tailor their marketing strategies to maximize their reach and sales during the festive season.

This analysis provides a foundation for more targeted campaigns and offers for future Diwali sales events, leading to more efficient and effective marketing strategies.

7. Visualizations

Gender Distribution: Bar chart displaying the count of male and female customers.

Total Sales by Gender: Bar plot showing total sales by gender.

Age Group Analysis: Bar chart showing distribution of customers by age group and gender.

Total Sales by Age Group: Total amount spent by different age groups.

State Distribution: Top 10 States by Orders: Bar plot of the number of orders per state.

Top 10 States by Sales: Bar plot of the total sales per state.

Marital Status and Sales: Bar chart comparing total amount spent by married and unmarried customers.