

A summary of EDA of D_Sales

Sales Data Overview

1. Dataset Information:

- The dataset contains various columns, including gender, marital status, age group, occupation, product category, amount spent, and state.
- Several irrelevant columns (like **Status** and **unnamed1**) were dropped, and missing values in the **Amount** column were filled with the mean value.

Gender Analysis

- **Gender Distribution:**
 - A count plot for gender shows the distribution of male and female customers.
- **Average Spending by Gender:**
 - The bar plot reveals the average amount spent by each gender, with a noticeable difference in spending patterns between genders.

Age Group and Gender

- **Purchases by Age and Gender:**
 - The count plot for age group, segmented by gender, shows which age group has the highest purchases for each gender. This provides insights into the most active customer demographic in terms of age and gender.

State-Level Insights

- **Top 10 States by Total Orders:**
 - A bar plot shows the states with the highest number of orders.
 - The top contributing states include Uttar Pradesh (UP), Maharashtra, and Karnataka.

- **Top 10 States by Total Amount:**
 - Another bar plot shows the states with the highest total spending. Similar states, such as **UP, Maharashtra, and Karnataka**, appear here as well.

Marital Status and Gender

- **Purchases by Marital Status and Gender:**
 - A bar plot reveals that married women tend to spend more compared to others, indicating that marital status impacts purchasing behavior.

Occupation Analysis

- **Occupation Distribution:**
 - A count plot shows the distribution of customers across various occupations.
- **Spending by Occupation:**
 - A bar plot reveals that people working in sectors like **IT, Healthcare, and Aviation** tend to spend more, particularly on products from categories like **Food, Clothing, and Electronics**.

Product Category Insights

- **Product Categories:**
 - The most popular product categories, based on total amount spent, are **Food, Clothing & Apparel, and Electronics & Gadgets**.

Key Recommendations for Targeting

1. **Focus on Married Women Aged 26-35:**
 - This group represents the largest customer base. Marketing campaigns targeted at **married women** in **UP, Maharashtra, and Karnataka** could be especially effective, particularly for categories like **Food, Clothing, and Electronics**.
2. **Leverage Occupation Insights:**
 - Customers in the **IT, Healthcare, and Aviation** sectors are key spenders, so tailored campaigns in these sectors could improve conversion rates.

- **IT professionals** could be targeted with ads for **Electronics & Gadgets**, while **Healthcare workers** and **Aviation employees** might be more responsive to ads for **Clothing** or **Food** products.
- 3. **Product Focus:**
 - Marketing for **Food, Clothing & Apparel**, and **Electronics** should be prioritized, as these categories have the highest sales figures. Special deals or product bundles in these categories could drive higher engagement and sales.
- 4. **State-Based Targeting:**
 - Focus on regions like **UP, Maharashtra**, and **Karnataka** where high spending and purchase volumes are prevalent. These states could benefit from localized promotions and events.
- 5. **Age Group Segmentation:**
 - Consider developing age-specific campaigns, especially for the **26-35** age group, which is the most active purchasing demographic.

By focusing on these key demographic and product insights, businesses can enhance their marketing strategies and optimize inventory planning.