Data Analysis of Diwali Sales

1. Project Objective

The goal of this project is to analyze Diwali sales data to identify key customer demographics, purchasing trends, and popular product categories. These insights can support businesses in tailoring their marketing strategies, improving customer satisfaction, and optimizing inventory management.

2. Data Collection

- **Dataset**: The data for this project was sourced from a CSV file containing customer transaction records for Diwali sales.
- **Initial Structure**:
    - Rows: 11,251
    - Columns: 15
    - Information includes details like User_ID, Cust_name, Product_ID, Gender, Age Group, State, Occupation, Product_Category, and Amount.

3. Data Preprocessing

- **Step 1: Data Import**
    - Loaded the dataset using pandas.
    - Verified data import by checking the top rows with df.head() and overall structure with df.info().
- **Step 2: Data Cleaning**
    - **Identified Missing Values**:
        - Columns Status and Unnamed1 were found to have null values throughout and were irrelevant to our analysis.
        - Dropped these columns using df.drop().
    - **Handling Remaining Missing Data**:
        - The Amount column had 12 missing values. These rows were removed to maintain data consistency.
    - **Data Type Conversion**:
        - Converted Amount column from float to int for easier analysis and visualization.
    - **Renaming Columns**:
        - Renamed Marital_Status to Shaadi to make it more descriptive for this context.
    - **Final Dataset Structure**:
        - After cleaning, the dataset consisted of 11,239 rows and 13 columns.

4. Exploratory Data Analysis (EDA)

- Used matplotlib and seaborn for visualizing trends, distributions, and comparisons.
- **Analysis by Gender**:
  - **Objective**: Understand the distribution of sales between male and female customers.
  - **Findings**: Female customers made the majority of purchases, contributing to higher sales.
- **Analysis by Age Group**:
  - **Objective**: Identify which age groups are most active in purchasing.
  - **Findings**: The 26-35 age group, especially women, showed the highest purchasing power.
- **Analysis by State**:
  - **Objective**: Analyze regional sales patterns to understand the geographic distribution of demand.
  - **Findings**: Uttar Pradesh, Maharashtra, and Karnataka had the highest number of orders, indicating strong sales from these states.
- **Analysis by Marital Status**:
  - **Objective**: Determine if marital status impacts purchasing behavior.
  - **Findings**: Married individuals, particularly women, showed higher purchasing tendencies.
- **Analysis by Occupation**:
  - **Objective**: Explore which occupations are associated with higher sales.
  - **Findings**: Customers in IT, Healthcare, and Aviation had significant spending, suggesting these sectors have more purchasing power.
- **Analysis by Product Category**:
  - **Objective**: Identify the most popular product categories.
  - **Findings**: Food, Clothing, and Electronics were the top-selling categories.

5. Key Insights and Conclusions

- **Customer Demographics**:
  - Married women aged 26-35, particularly from states like Uttar Pradesh, Maharashtra, and Karnataka, working in IT, Healthcare, and Aviation, are the primary customer segment.
- **Popular Products**:
  - High demand for products in Food, Clothing, and Electronics categories.
- **Business Impact**:
  - The analysis helps in understanding key customer demographics, allowing businesses to personalize their marketing efforts and optimize inventory for high-demand products.

6. Tools and Libraries Used

- **Python Libraries**: pandas for data manipulation, matplotlib and seaborn for visualizations.
- **Techniques**: Grouping, sorting, and plotting were employed to gain insights.

7. Future Recommendations

- **Marketing**: Focus on targeted marketing for married women in the 26-35 age group from high-sales states.
- **Inventory Management**: Prioritize stocking high-demand items in Food, Clothing, and Electronics to meet customer demand effectively.

Findings from **Diwali Sales Analysis Project**:

1. **Gender Insights**:
    - The analysis reveals that female customers made more purchases than male customers and had a higher total spending. This indicates that females have a stronger purchasing power during the Diwali season.
2. **Age Group Analysis**:
    - The primary age group of buyers was between 26-35 years, with females in this age range showing the highest purchasing behaviour. This age group contributed significantly to the total sales amount, highlighting it as a valuable customer segment.
3. **State-Based Findings**:
    - The states with the highest sales and number of orders were Uttar Pradesh, Maharashtra, and Karnataka. These states represent a high concentration of demand, suggesting a regional focus for marketing campaigns.
4. **Marital Status**:
    - Married individuals, particularly women, showed a higher purchasing tendency and spending. This suggests that marital status may influence purchase decisions, possibly linked to family needs during festivals.
5. **Occupation Trends**:
    - The most active buyers worked in the IT, Healthcare, and Aviation sectors. These occupations showed higher spending, possibly due to higher disposable income, making them valuable for targeted marketing
6. **Product Category Preferences**:
    - The top-selling categories were Food, Clothing, and Electronics, indicating a strong demand for these items during Diwali. This insight can help with inventory planning and promotions focused on these categories.
7. **Top Products**:
    - Analysis of product data showed the top 10 most purchased products, providing insights for inventory optimization and product bundling strategies.
8. **Conclusion**:
    - Married women aged 26-35, primarily from Uttar Pradesh, Maharashtra, and Karnataka, and working in IT, Healthcare, and Aviation, are likely to spend more on Food, Clothing, and Electronics during Diwali. Businesses can improve customer experience and optimize sales by targeting these demographics and focusing on the popular product categories.

These insights guide effective decision-making in marketing, inventory management, and customer targeting. Let me know if you need further details on any specific aspect!