Diwali Sales Analysis - Project Documentation

Title

Diwali Sales Analysis using Python (Pandas, Matplotlib, Seaborn)

Abstract

The Diwali Sales Analysis project aims to extract actionable insights from retail sales data collected during the Diwali festive season. By analyzing customer demographics, purchase behavior, and product categories, this project helps businesses understand trends and boost future festive sales through targeted marketing strategies.

Objectives

- To analyze and visualize Diwali season customer sales data.
- To identify the most profitable age groups, states, product categories, and customer segments.
- To understand customer behavior based on gender, age, occupation, and marital status.
- To guide businesses in making data-driven decisions for future festive campaigns.

Tools & Technologies Used

- Python
- Libraries: Pandas, NumPy, Matplotlib, Seaborn
- Jupyter Notebook
- CSV Dataset: Diwali Sales Data.csv

Advantages

- Helps businesses tailor marketing strategies to key customer groups.
- Identifies high-performing product categories and customer segments.
- Enhances sales forecasting and planning for future Diwali seasons.
- Encourages data-driven decision-making in retail management.

Data Preprocessing Steps

- Load dataset using pandas.read_csv().
- Handle missing values by using dropna().

- Convert data types (e.g., Amount to int).
- Drop unnecessary columns like 'Status', 'unnamed1'.
- Basic EDA (count, sum, groupby, etc.) for insights.

Key Analyses Performed

- Gender-wise total sales and purchases.
- State-wise sales distribution.
- Age group and occupation-wise purchase patterns.
- Marital status vs amount analysis.
- Product category performance.
- Top performing products (by sales).

Visualizations Used

- Bar plots (using Seaborn) for categorical comparisons.
- Pie charts and line plots for state and category distribution.
- Histograms for age, amount, and quantity analysis.

Sample Insights

- Female customers contribute significantly to total purchase amount.
- Married customers tend to spend more during Diwali.
- States like Uttar Pradesh and Maharashtra generate the highest sales.
- Professionals and working adults (age 26–35) spend the most.
- Clothing and food products are the most purchased categories.

Dataset Summary

- Rows: ~11251
- Columns: Gender, Age Group, State, Occupation, Product Category, Amount, etc.
- Source: Mock retail sales data during the Diwali festival.

Conclusion

The Diwali Sales Analysis provides valuable insights into customer behavior and product trends during the festive season. This helps e-commerce businesses optimize their marketing campaigns, improve customer targeting, and maximize revenue.