Exploratory Data Analysis (EDA) – Python Diwali Sales Analysis

■ Objective

To perform Exploratory Data Analysis (EDA) on a sales dataset to uncover insights, trends, and patterns in Diwali sales using Python. The goal is to clean, visualize, and interpret the data to support better business decisions.

■ Tools & Libraries Used

- Python
- Pandas – Data manipulation and cleaning
- NumPy – Numerical computations
- Matplotlib – Data visualization
- Seaborn – Statistical and aesthetic plotting
- Jupyter Notebook / Google Colab – Interactive code and output execution

■ Dataset Information

Dataset Name: Diwali Sales Dataset
File: Python_Diwali_Sales_Analysis.csv (contained inside the ZIP file)

- User_ID – Unique customer ID
- Cust_name – Customer name
- Gender – Male/Female
- Age Group – Customer's age category
- State – State of purchase
- Occupation – Customer's job type
- Marital_Status – 0 = Unmarried, 1 = Married
- Product_Category – Type of product purchased
- Orders – Number of items purchased
- Amount – Total purchase amount

■ Steps Performed

1 Data Loading: Imported CSV using Pandas and checked structure using .head(), .info(), and .describe().

2 Data Cleaning: Removed null values, corrected column names and data types, and filtered unnecessary columns.

3. Exploratory Data Analysis: Performed univariate, bivariate, and category-wise analyses.
4. Visualizations: Created graphs using Matplotlib & Seaborn such as bar plots, pie charts, and heatmaps.
5. Insights: Found top contributing states, age groups, and occupations affecting Diwali sales.

■ Key Insights

- Married women aged 26–35 years from states like Uttar Pradesh, Maharashtra, and Karnataka contributed the most to sales.
- IT, Healthcare, and Aviation professionals showed higher purchasing power.
- Food, Clothing, and Electronics were the most purchased categories.

■ Deliverables

- Jupyter Notebook: Python_Diwali_Sales_Analysis.ipynb
- PDF Report: Python_Diwali_Sales_Analysis_Report.pdf
- README File: Project summary and usage guide

■ How to Run

1. Extract the ZIP file.
2. Open the .ipynb file in Jupyter Notebook or Google Colab.
3. Run all cells sequentially.
4. Check the generated graphs and insights.

■ Conclusion

The EDA process provided useful business insights on consumer behavior during Diwali sales. Such analyses can help marketing teams target the right customer groups and optimize product offerings.

Author: Jharna Verma
Tool Used: Python (Pandas, Matplotlib, Seaborn)
Date: October 2025