Diwali Sales Analysis – Python EDA Project

Project Overview

Project Title: Diwali Sales Analysis
Tools: Python, Pandas, NumPy, Matplotlib, Seaborn

This project focuses on analyzing Diwali sales data using **Python** to understand customer purchasing behavior during the festive season. The project demonstrates essential data analysis skills such as data cleaning, exploratory data analysis (EDA), and deriving business insights through visualizations.

Objectives

1. Load and understand the dataset
2. Clean the data by handling missing and unnecessary columns
3. Perform Exploratory Data Analysis (EDA)
4. Analyze customer demographics and purchasing patterns
5. Extract meaningful business insights from Diwali sales data

---

Project Structure

```
Diwali-Sales-Analysis/
│
├── Diwali_Sales_Analysis.ipynb
├── Diwali Sales Data.csv
└── README.md
```

Dataset Description

- **Dataset Name**: Diwali Sales Data
- **File Format**: CSV
- **Encoding Used**: `unicode_escape`

Key Columns:

- User_ID
- Gender
- Age Group
- State
- Marital_Status
- Occupation
- Product_Category
- Amount

Tools & Libraries Used

- **Python**
- **Pandas** – Data manipulation and cleaning
- **NumPy** – Numerical operations
- **Matplotlib** – Data visualization
- **Seaborn** – Statistical visualization

Data Cleaning & Preparation

The following data cleaning steps were performed:

- Removed unnecessary columns
- Dropped rows with null values
- Converted data types where required
- Ensured data consistency for analysis

Exploratory Data Analysis (EDA)

The analysis was performed to answer key business questions such as:

1. **Gender-wise purchasing behavior**
2. **Age group contributing the most to sales**
3. **State-wise sales and order distribution**
4. **Impact of marital status on spending**
5. **Occupation-wise sales trends**
6. **Most popular product categories during Diwali**

Example Analysis Code:

```
sns.barplot(x='Gender', y='Amount', data=df)
sns.barplot(x='Age Group', y='Amount', data=df)
```

Key Findings

- **Female customers** contributed more to total sales compared to male customers.
- The **26–35 age group** showed the highest purchasing power.
- Certain states recorded significantly higher sales during Diwali.
- **Married customers** spent more on average.
- A few **product categories dominated festive purchases**.
- Specific occupations showed higher spending trends.

Reports Generated

- **Customer Demographics Analysis**
- **State-wise Sales Performance**
- **Product Category Demand Analysis**
- **Festive Season Spending Patterns**

Conclusion

This project demonstrates how **Python-based exploratory data analysis** can uncover valuable insights from sales data. The findings can help businesses improve customer targeting, inventory planning, and marketing strategies during festive seasons like Diwali.

How to Use

1. **Clone the repository**
2. Open the Jupyter Notebook:
3. `jupyter notebook Diwali_Sales_Analysis.ipynb`
4. Run the cells sequentially to view the analysis and visualizations
5. Modify or extend the analysis to explore additional insights