Diwali Sales Analysis – Project Report

1. Introduction

The **Diwali Sales Analysis** project focuses on understanding customer purchasing behavior during the Diwali festival season. Diwali is one of the biggest shopping periods in India, and analyzing sales data during this time helps businesses improve marketing strategies, inventory planning, and customer targeting.

This project applies **Exploratory Data Analysis (EDA)** techniques using Python to extract meaningful insights from sales data.

---

2. Objectives

- Analyze customer demographics and purchasing patterns
- Identify top-performing product categories and states
- Understand the impact of gender, age group, and occupation on sales
- Derive actionable insights to support business decision-making

---

3. Dataset Description

- **Dataset Name:** Diwali Sales Data
- **File Format:** CSV
- **Key Columns:**
- User_ID
- Gender
- Age Group
- State
- Marital_Status
- Occupation
- Product_Category
- Amount (Sales Value)

The dataset contains customer-level transactional data collected during the Diwali season.

---

4. Tools & Technologies Used

- **Programming Language:** Python
- **Libraries:**
- Pandas – data manipulation
- NumPy – numerical operations
- Matplotlib & Seaborn – data visualization
- **Environment:** Jupyter Notebook

---

5. Methodology

1. **Data Importing** – Loaded the CSV file into a Pandas DataFrame.
2. **Data Understanding** – Checked shape, structure, and summary of the dataset.
3. **Data Cleaning** –
4. Removed null and irrelevant columns
5. Corrected data types
6. **Exploratory Data Analysis (EDA)** –
7. Analyzed sales by gender, age group, state, occupation, and product category
8. Used bar plots and count plots for visualization
9. **Insight Generation** – Identified trends and key findings from visual analysis.

---

6. Exploratory Data Analysis & Insights

6.1 Gender-wise Analysis

- Female customers contributed more to overall sales compared to male customers.
- Average purchase value was higher among female buyers.

6.2 Age Group Analysis

- The **26–35 age group** showed the highest purchasing power.
- Younger working professionals dominated Diwali shopping.

6.3 State-wise Sales

- Top contributing states included **Uttar Pradesh, Maharashtra, and Karnataka**.
- These states recorded the highest order count and sales value.

6.4 Occupation Analysis

- Customers from **IT, Healthcare, and Aviation sectors** spent more during Diwali.
- Occupation plays a significant role in spending behavior.

6.5 Product Category Analysis

- Highest sales were observed in **Food, Clothing, and Electronics** categories.
- These categories are major revenue drivers during festive seasons.

---

7. Key Findings

- Married women aged 26–35 from metro and tier-1 states are the most valuable customers.
- Festive sales are strongly influenced by demographics and profession.
- Businesses should focus promotions on high-performing product categories and customer segments.

---

8. Conclusion

The Diwali Sales Analysis project successfully demonstrates how data analysis can uncover valuable business insights. By leveraging Python and visualization techniques, the project highlights customer trends that can help organizations optimize marketing strategies, improve customer engagement, and maximize festive-season revenue.

---

9. Future Scope

- Apply machine learning models for sales prediction
- Perform customer segmentation using clustering techniques
- Integrate real-time sales dashboards using Power BI or Tableau

---

Project By: Akash Rajendra Sagar