**📝 Project Report: Diwali Sales Data Analysis**

**📄 Cover Page**

**Project Title:** Diwali Sales Data Analysis  
**Submitted by:** Vikas Kumar  
**Tools Used:** Python, Jupyter Notebook, Pandas, NumPy, Matplotlib, Seaborn  
**Date:** June 2025

**📌 Executive Summary**

This report presents an in-depth analysis of customer purchase behavior during the Diwali festive season using a real-world retail dataset. The goal is to identify sales trends, understand customer demographics, and uncover actionable insights to improve marketing and sales strategies for future festive campaigns. Python and its data science libraries were used to clean, explore, and visualize the dataset.

**📖 Introduction**

Diwali, also known as the festival of lights, marks one of the most significant shopping periods in India. Retailers across the country witness a major spike in customer purchases. This project focuses on understanding who is shopping, what they are buying, and how much they are spending during this period.

The dataset provides customer demographic information along with product categories and purchase amounts. By exploring this data, we aim to draw insights that can help businesses better plan for future festive seasons.

**🔍 Methodology**

The following steps were followed during this project:

1. **Data Collection:**  
   The dataset used was a cleaned CSV file titled Diwali\_Sales\_Data.csv, containing customer demographics and sales transactions.
2. **Data Preprocessing:**
   * Removed null values
   * Converted data types for accurate analysis
   * Filtered irrelevant or redundant columns (e.g., Status, Unnamed: 10)
3. **Exploratory Data Analysis (EDA):**
   * Used visualizations (bar plots, pie charts, histograms) to examine:
     + Age distribution
     + Gender-based purchase behavior
     + City category impact
     + Product category and amount trends
4. **Tools & Libraries Used:**
   * **Python:** Main programming language
   * **Pandas & NumPy:** Data manipulation
   * **Matplotlib & Seaborn:** Data visualization
   * **Jupyter Notebook:** Code execution and presentation

**📊 Data Analysis & Insights**

**1. Gender-Based Insights:**

* Male customers contributed to more sales than female customers.
* Marketing efforts can be focused more towards male audiences during Diwali.

**2. Age Group Analysis:**

* The **26–35** age group was the most active in terms of purchases.
* They can be considered the most responsive demographic for Diwali promotions.

**3. Marital Status:**

* Married customers had slightly higher average purchases.
* Loyalty or couple-based discounts could be effective.

**4. City Category:**

* Cities in **Tier A** and **Tier B** recorded higher sales.
* These urban segments can be targeted for region-specific advertising.

**5. Product Categories:**

* **Product Category 1** had the highest number of sales and revenue.
* Popular categories should be featured more prominently during festive marketing campaigns.

**📌 Key Findings**

| **Attribute** | **Insight** |
| --- | --- |
| Gender | Males shop more than females |
| Age | 26–35 age group is the most active |
| Marital Status | Married users have higher average spend |
| Cities | Tier A & B cities contribute to major sales |
| Product Category | Category 1 generates the highest revenue |

**💼 Business Recommendations**

1. **Personalized Marketing:**  
   Use customer segmentation to send customized promotions to 26–35 age group males in Tier A/B cities.
2. **Product Positioning:**  
   Highlight Category 1 products on websites, apps, and physical displays during festive seasons.
3. **City-Wise Discounts:**  
   Offer region-specific discount codes based on city tiers.
4. **Married Segment Offers:**  
   Promote couple/family bundles to leverage higher spending from married customers.

**✅ Conclusion**

The Diwali Sales Data Analysis successfully highlights key consumer behaviors and product preferences during the festive season. These insights can be used by retail businesses to tailor their marketing strategies, manage inventory more effectively, and create more impactful customer engagement plans.

This analysis not only helps improve Diwali sales strategies but also serves as a reusable framework for analyzing seasonal retail patterns.

**📎 Appendix**

* Dataset file: Diwali\_Sales\_Data.csv
* Jupyter Notebook: Data\_Analysis.ipynb
* Tools: Python, Pandas, Matplotlib, Seaborn
* Visualization snapshots (available in .ipynb)