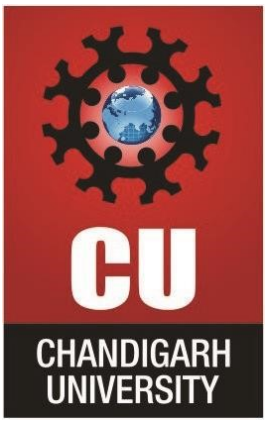
University Institute of Computing
Chandigarh University
Mohali, Punjab (140413)



Project Report
On
Diwali Sales Visualization Dashboard

Under Guidance of - Ms. Deepali Saini

Submitted To
Ms. Deepali Saini

Submitted By
Name – SAWAN KUMAR
UID - 24MCA20300
Section- 5A

S.No.	TABLE OF CONTENTS	PAGE No.
1	chapter 1: Introduction	3
2	Chapter 2: Literature Review	4
3	Chapter 3: Implementation	5
4	pie Chart	6
5	Histogram	7
6	References	11

Abstract:
Diwali is one of India's biggest festivals, marked by a surge in consumer spending. The Diwali sales dataset provides insights into purchasing patterns and spending habits during this period. By using data visualization techniques, we can extract

Chapter 1: Introduction

1.1 Overview

Diwali is one of India's biggest festivals, marked by a surge in consumer spending. The Diwali sales dataset provides insights into purchasing patterns and spending habits during this period. By using data visualization techniques, we can extract meaningful insights, such as peak sales times, top-performing products, and demographic trends.

1.2 Background

The dataset captures various features of Diwali sales, including product categories, purchase quantities, and customer demographics. Without visualization, it is challenging to recognize patterns in consumer behavior and sales distribution across regions and customer types.

1.3 Problem Statement

This project aims to analyze Diwali sales data using Python visualization tools to uncover trends and patterns that could help businesses understand consumer behavior and optimize future marketing strategies.

1.4 Objectives

The objectives are as follows:

- Analyze Diwali sales data by product categories, regions, and customer demographics.
- Visualize trends in sales volume, customer preferences, and spending patterns.
- Simplify data interpretation for better decision-making through informative visuals

1.5 Report Structure

The report is organized as follows:

- Chapter 1 introduces the project and outlines the problem, objectives, and structure.
- Chapter 2 reviews related literature and methods for data visualization.
- Chapter 3 describes the motivation, objectives, and scope of the project.
- Chapter 4 explains the tools, technologies, implementation details, and presents the analysis results and visualizations.

Chapter 2: Literature Review

2.1

Motivation:

The project aims to uncover patterns in Diwali sales that can help businesses tailor their strategies during this peak season. Visualizations can provide an intuitive understanding of consumer behavior, enabling companies to improve product offerings and marketing efforts..

2.2 Objectives:

The key objectives of the project include:

1. Explore sales distribution across product categories and regions.
2. Identify customer demographics contributing most to sales.
3. Generate visualizations that provide actionable insights for stakeholders.

2.3 Scope of the Project:

This project will focus on analyzing the dataset to identify and visualize sales trends. Visual aids such as bar plots, histograms, and pie charts will be used to

summarize findings on sales trends across demographics, product categories, and time.

Chapter 3: Implementation

3.1 Tools and Technologies Used

The following technologies were used:

- Python: Primary language for analysis and visualization.
- **Pandas**: Data handling and manipulation.
- **Matplotlib and Seaborn**: For creating visualizations to present findings.
-

3.2 Code Architecture

The data analysis followed these steps:

1. Data Loading: Importing the Diwali sales dataset.
2. Data Cleaning: Organizing the data for meaningful analysis.
3. Visualization: Using Matplotlib and Seaborn to generate insights.
4. Output: Displaying visualizations to interpret consumer behavior.

3.3 Data Validation and Preparation

Data validation was conducted to ensure accurate representation. Specific checks included ensuring proper data formatting, verifying key

columns (e.g., product categories, purchase quantities), and addressing any missing data.

3.4 Code Implementation and Output

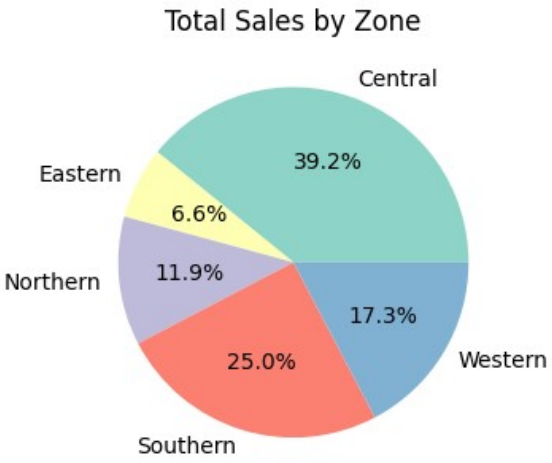
```
import pandas as pd import matplotlib.pyplot as plt import seaborn as sns
```

```
# Load the data
file_path = r"C:\Users\aarya\Desktop\AARYAN_PYTHON
data visualisation\AARYAN_PYTHON\Diwali Sales Data.csv"
df = pd.read_csv(file_path, encoding='ISO-8859-1')
```

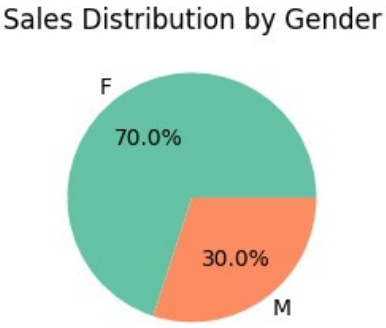
Pie Chart: Sales Distribution by Product Category

- Visualize sales across various product categories to identify top performers.

```
# Plot 1: Pie Chart - Total Sales by Zone

zone_sales = df.groupby('Zone')['Amount'].sum()

plt.figure(figsize=(8, 8))

plt.pie(zone_sales, labels=zone_sales.index, autopct='%1.1f%%',
colors=sns.color_palette("Set3"))

plt.title('Total Sales by Zone')

plt.show()
```



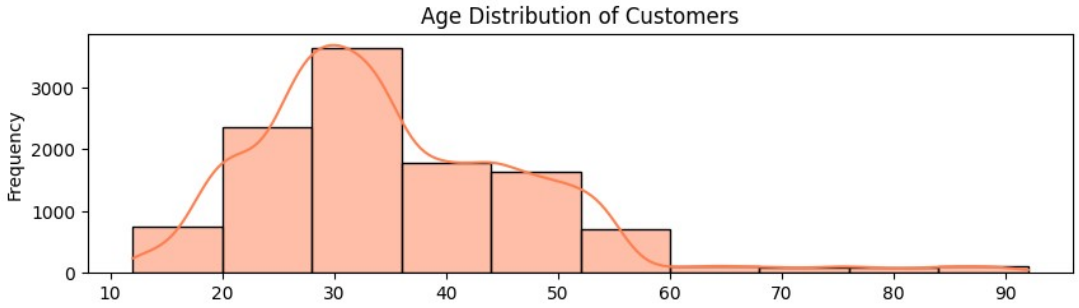
# Plot 2: Pie Chart - Gender Distribution in Sales
gender_sales = df.groupby('Gender')['Amount'].sum()
plt.figure(figsize=(8, 8))
plt.pie(gender_sales, labels=gender_sales.index,
autopct='%1.1f%%', colors=sns.color_palette("Set2")) plt.title('Sales Distribution by Gender')
plt.show()



Histogram: Customer Age Distribution

- Analyze the age distribution of customers to understand the primary age groups purchasing during Diwali.

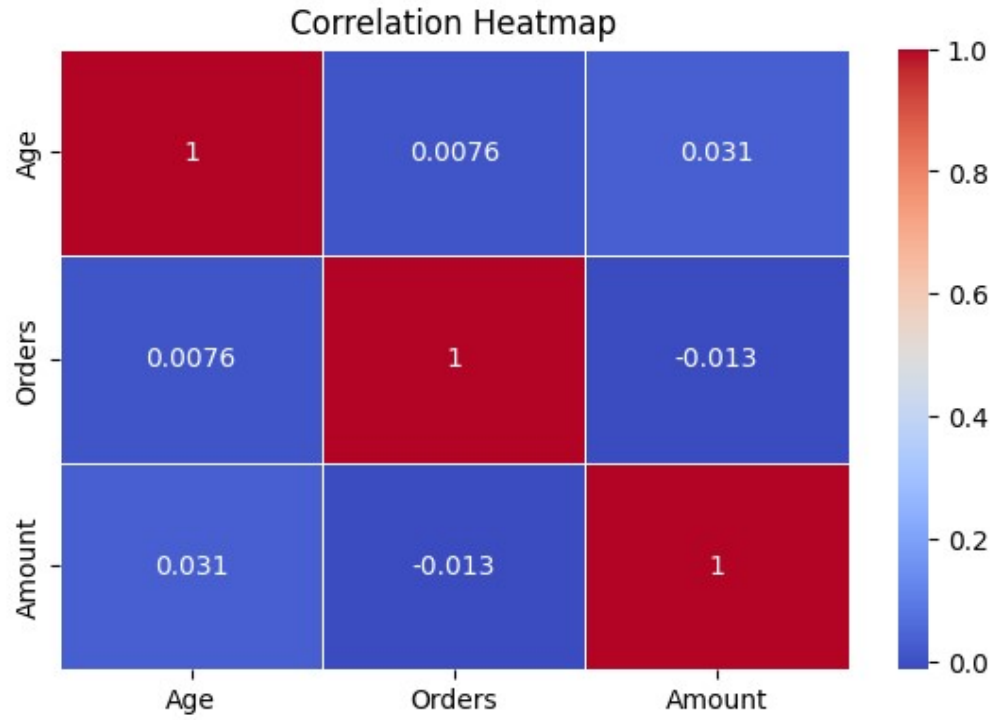
```
# Plot 3: Histogram - Age Distribution of Customers
plt.figure(figsize=(10, 6)) sns.histplot(df['Age'],
bins=10, kde=True, color='coral') plt.title('Age
Distribution of Customers') plt.xlabel('Age')
plt.ylabel('Frequency') plt.show()
```



Heatmap: Correlation Between Product Type and Sales Quantity

- Display correlations to understand which product types are in higher demand.

```
# Plot 4: Heatmap - Correlation Matrix plt.figure(figsize=(8, 6))
corr_matrix = df[['Age', 'Orders', 'Amount']].corr()
sns.heatmap(corr_matrix, annot=True, cmap='coolwarm',
linewidths=0.5) plt.title('Correlation Heatmap') plt.show()
```

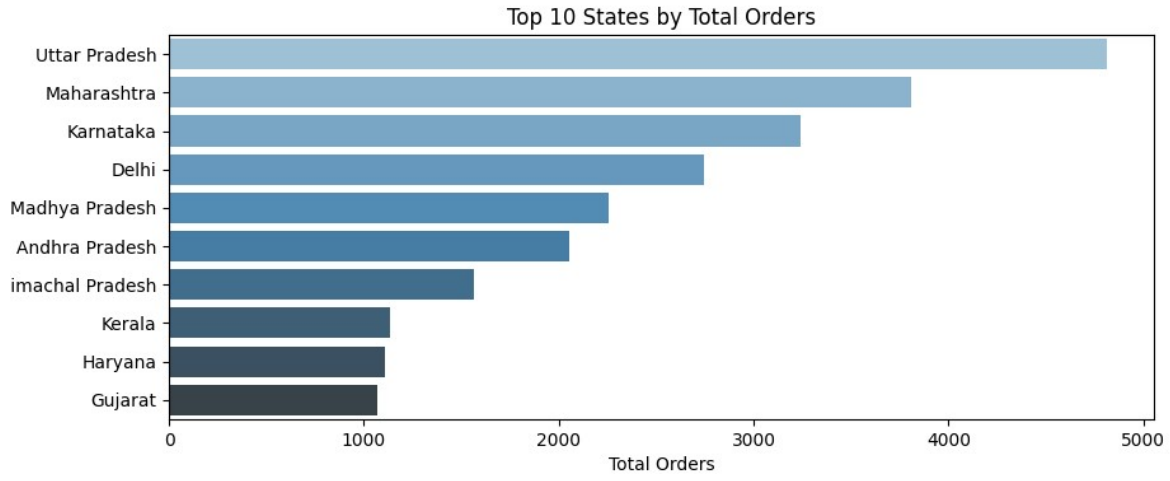


Plot 5: Bar Plot - Total Orders by State state_orders = df.groupby('State')['Orders'].sum().nlargest(10)

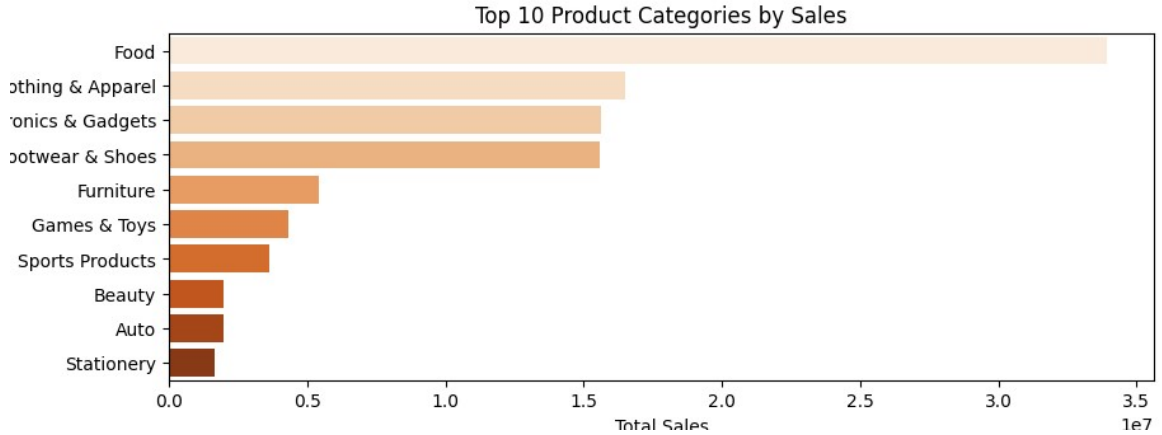
plt.figure(figsize=(10, 6)) sns.barplot(x=state_orders.values, y=state_orders.index,

palette='Blues_d')

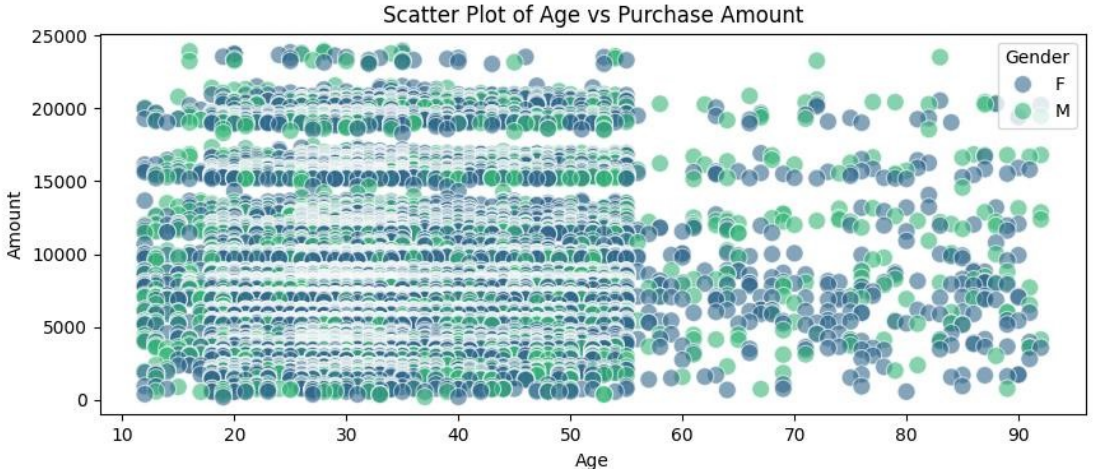
plt.title('Top 10 States by Total Orders') plt.xlabel('Total Orders') plt.ylabel('State') plt.show()



Plot 6: Bar Plot - Total Sales by Product Category
product_sales =
df.groupby('Product_Category')['Amount'].sum().nlar
g est(10) plt.figure(figsize=(10, 6))
sns.barplot(x=product_sales.values,
y=product_sales.index, palette='Oranges')
plt.title('Top 10 Product Categories by Sales')
plt.xlabel('Total Sales') plt.ylabel('Product Category')
plt.show()



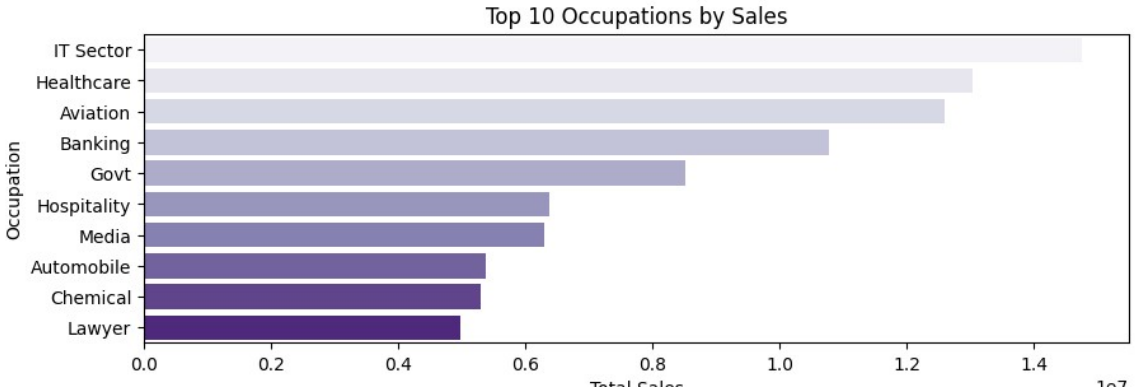
Plot 7: Scatter Plot - Age vs Amount plt.figure(figsize=(10, 6)) sns.scatterplot(x='Age', y='Amount', data=df, hue='Gender', palette='viridis', s=100, alpha=0.6) plt.title('Scatter Plot of Age vs Purchase Amount') plt.xlabel('Age') plt.ylabel('Amount') plt.show()



Plot 8: Bar Plot - Total Sales by Occupation occupation_sales = df.groupby('Occupation')['Amount'].sum().nlargest(10)

plt.figure(figsize=(10, 6)) sns.barplot(x=occupation_sales.values, y=occupation_sales.index, palette='Purples')

plt.title('Top 10 Occupations by Sales') plt.xlabel('Total Sales') plt.ylabel('Occupation') plt.show()



References

- Pandas Documentation: https://pandas.pydata.org/docs/
- Matplotlib Library - W3Schools: https://www.w3schools.com/python/matplotlib_intro.asp
- Seaborn Documentation: https://seaborn.pydata.org/
- Python Official Documentation: https://docs.python.org/3/
- Matplotlib Tutorial GeeksforGeeks: https://www.geeksforgeeks.org/matplotlib-tutorial/