A

PROJECT REPORT

On

“BLACK FRIDAY DATA ANALYSIS”

USING A PYTHON

A report submitted to

Seven Mentors Institute, Pune

2024-25

Date :

10 NOV 2024

Submitted by :

Kanak Manoj Garg

CERTIFICATE

This is to certify that the project called **"BLACK FRIDAY DATA ANALYSIS"** is the bonafide work carried out by **Kanak Manoj Garg** at **SEVEN MENTOR INSTITUTE, PUNE** under the guidance Course tutor **Mr. Sagar Gade Sir** and this project report leads the career path by exploring Data Analysis using Python.

DATE: 10 NOV 2024

PLACE: Pune

SIGNATURE OF GUIDE

DECLARATION

I hereby declare that, this project entitled is an outcome of my own efforts Under the guidance of Mr. Sagar Gade Sir. The project is submitted to the **Seven Mentor Institute Pune**. For the partial fulfilment of the "**Python Course"** 2024-2025.

I also declare that this project report has not been previously submitted to any other institute.

DATE: 10 NOV 2024

PLACE: Pune **Kanak Manoj Garg**

INDEX

ACKNOWLEDGEMENT

The successful completion of this data analysis project is the result of contributions and support from various individuals and entities. It is only fitting to acknowledge their invaluable assistance and guidance.

I am sincerely grateful to my mentor and colleagues who provided valuable insights and feedback throughout this project. Their support was instrumental in navigating the complexities of the data analysis process.

I would like to extend my heartfelt thanks to mentor for their unwavering guidance and encouragement. Their expertise and constructive criticism were crucial in shaping the direction and quality of this work.

Additionally, I appreciate the resources and tools provided by various platforms and institutions that facilitated the successful execution of this project. Their support was indispensable in achieving the project's objectives. Thank you all for your contributions and support. This project would not have been possible without your help.

INTRODUCTION

Black Friday is one of the most significant shopping events globally, characterized by massive sales and consumer spending. Understanding the trends and patterns during this period is crucial for retailers and analysts. This project aims to analyze a dataset of Black Friday sales to uncover key insights related to consumer behaviour, product preferences, and sales performance.

Items in Quotation:

- **TransactionID**: Unique identifier for each transaction.
- **ProductID**: Identifier for the product sold.
- **UserID**: Identifier for the customer.
- **Gender**: Gender of the customer.
- **Age**: Age group of the customer.
- **Occupation**: Occupation category of the customer.
- **City_Category**: Category of the city where the customer resides.
- **Stay_In_Current_City_Years**: Number of years the customer has stayed in the current city.
- **Marital_Status**: Marital status of the customer.
- **Product_Category_1, 2, 3**: Categories of the products sold.
- **Purchase**: Purchase amount.

TECHNOLOGY USED IN PROJECT

1. Python

Python is the primary programming language used for data analysis due to its simplicity and powerful libraries. Python provides a versatile environment for data manipulation, statistical analysis, and visualization.

- **Pandas**: Essential for data manipulation and analysis. It provides data structures such as Data Frames to handle and preprocess the dataset efficiently.
- **NumPy**: Used for numerical operations and handling large arrays and matrices of data.
- **Seaborn**: A statistical data visualization library built on Matplotlib, used for creating informative and attractive visualizations. It simplifies the creation of complex plots like bar charts and heatmaps.
- **Matplotlib**: A fundamental plotting library used for creating static, animated, and interactive visualizations. It provides extensive customization options for plots.

2. Data Analysis Libraries

- **Pandas**: Handles data cleaning, transformation, and aggregation. Key functionalities include data filtering, grouping, and summarization.
- **NumPy**: Supports numerical calculations and array operations, essential for data preprocessing and analysis.

3. Data Visualization Tools

- **Seaborn**: Utilized for creating advanced statistical plots such as bar charts and pair plots, which help in understanding distributions and relationships in the data.
- **Matplotlib**: Used for general plotting needs and customizing visualizations. It allows for detailed control over plot appearance and formatting.
- **Jupyter Notebooks**: An interactive computing environment that enables the creation of documents containing live code, equations, visualizations, and narrative text. It facilitates iterative analysis and documentation.

4. Development Environment

- **Jupyter Notebook / JupyterLab:** Provides an interactive platform for writing and executing Python code. It supports the integration of code with visualizations and explanatory text.
- **Integrated Development Environments (IDEs):** IDEs like PyCharm or VS Code are used for coding and debugging, providing a comprehensive development environment with support for Python.

5. Data Storage and Management

- **CSV Files:** The dataset is typically stored in CSV (Comma-Separated Values) format, which is easily read and processed using Pandas.
- **Excel Files:** For datasets available in Excel format, Pandas can also handle Excel files for data extraction and analysis.

EXPLAINATION

Making of the Project

SAMPLE CODES AND DATA VISUALIZATION

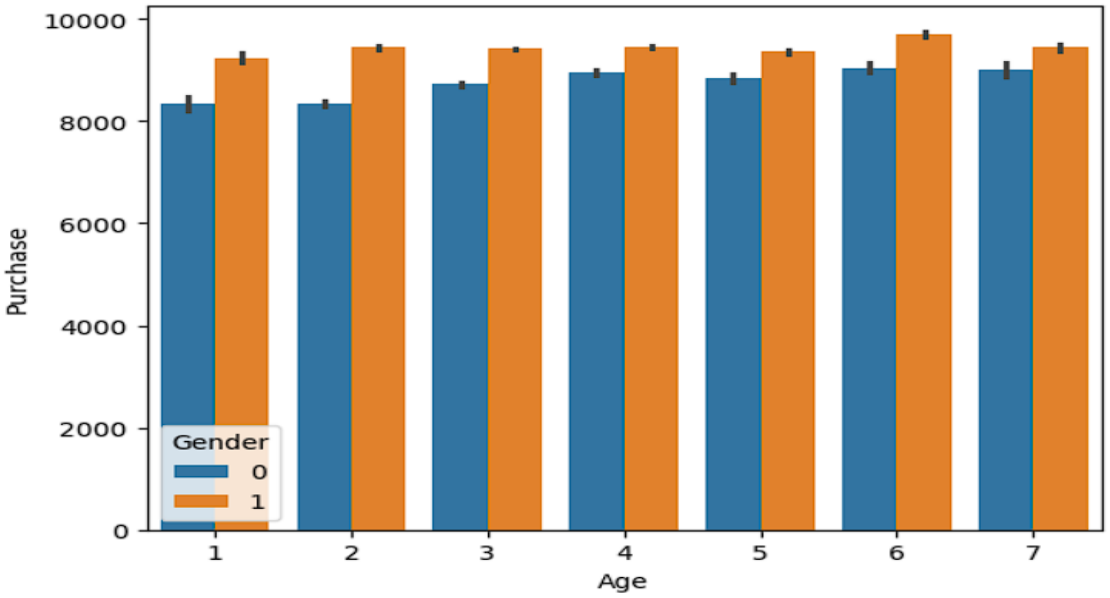
```
# Importing Libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
from warnings import filterwarnings
filterwarnings("ignore")

# Load the dataset
df = pd.read_csv("train.csv", encoding = 'latin')

# Number of columns
len(df.columns)

# Name of columns in the data
df.columns

# Information of the data
df.info()

# Summary statistics
df.describe()

# Display the first few rows of the dataset
df.head()

# Display the last few rows of the dataset
df.tail()
```

```
# Check for missing values
df.isnull().sum()

# Drop valueless columns
df.drop(['User_ID'], axis = 1, inplace =True)
df.drop('City_Category',axis=1,inplace=True)

# Handling Categorical Feature Gender, Age
df["Gender"] = df["Gender"].map({'F':0,"M":1})
df['Age'].unique()
df['Age']=df['Age'].map({'0-17':1,'18-25':2,'26-35':3,'36-45':4,'46-50':5,
'51-55':6,'55+':7})
df['Age']

# Replacing missing values
df["Product_Category_1"].mode()[0]
df['Product_Category_1']=
df['Product_Category_1'].fillna(df['Product_Category_1'].mode()[0])
df["Product_Category_2"].mode()[0]
df['Product_Category_2']=
df['Product_Category_2'].fillna(df['Product_Category_2'].mode()[0])
df["Product_Category_3"].mode()[0]
df['Product_Category_3’] =
df['Product_Category_3'].fillna(df['Product_Category_3'].mode()[0])
df["Purchase"].mean()
df['Purchase']= df['Purchase'].fillna(value)

# Showing/ Printing the dataset
df
print(df)
```

 # Visualization Age Vs Purchased

sns.barplot(data=df, x = "Age", y = "Purchase", hue="Gender")

```
<Axes: xlabel='Age', ylabel='Purchase'>
```



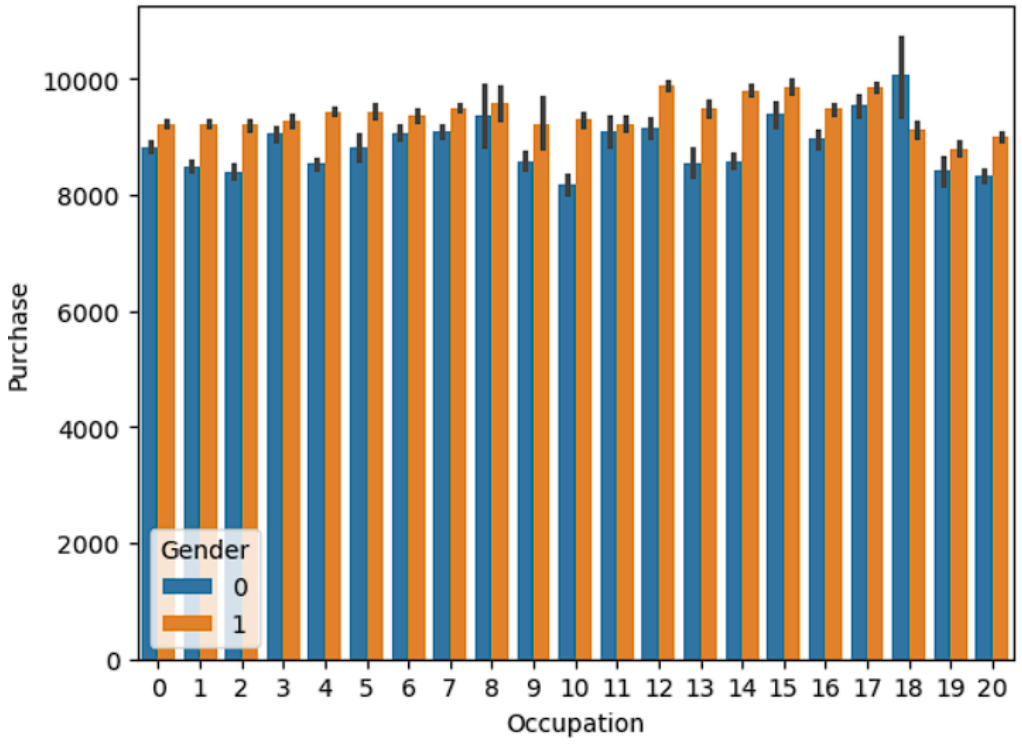
Observation

purchasing of man is high then women

 # Visualization Occupation Vs Purchased

sns.barplot(x = "Occupation", y= "Purchase", hue="Gender", data = df)

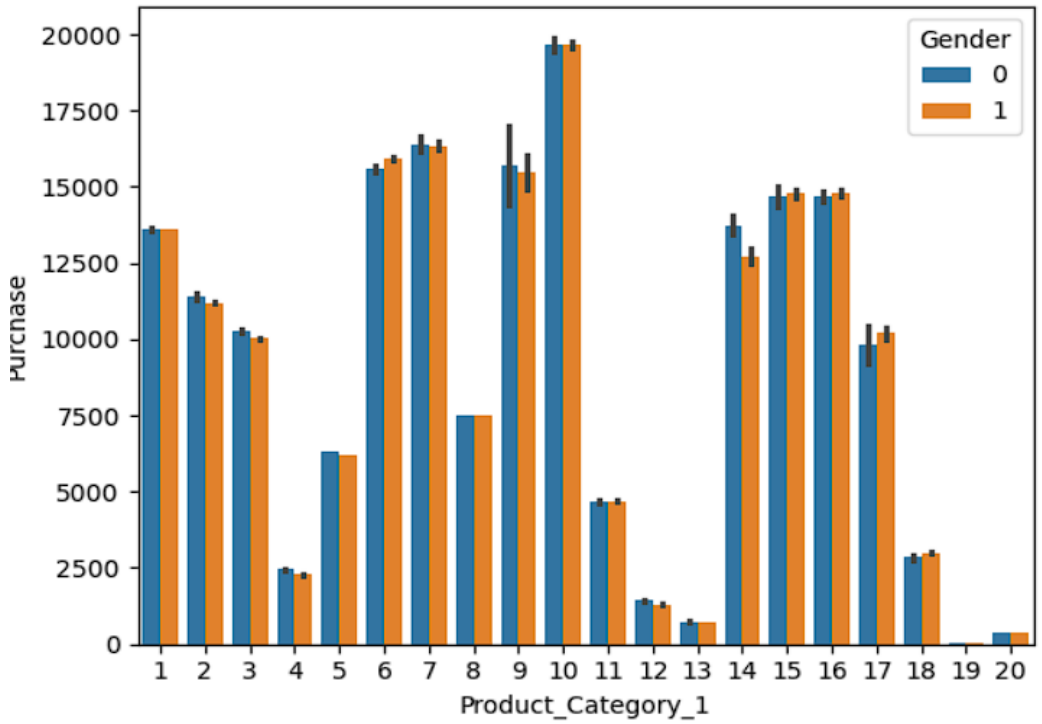
```
<Axes: xlabel='Occupation', ylabel='Purchase'>
```



 # Visualization Product_Category_1 Vs Purchased

sns.barplot(x = "Product_Category_1", y = "Purchase", hue="Gender", data = df)

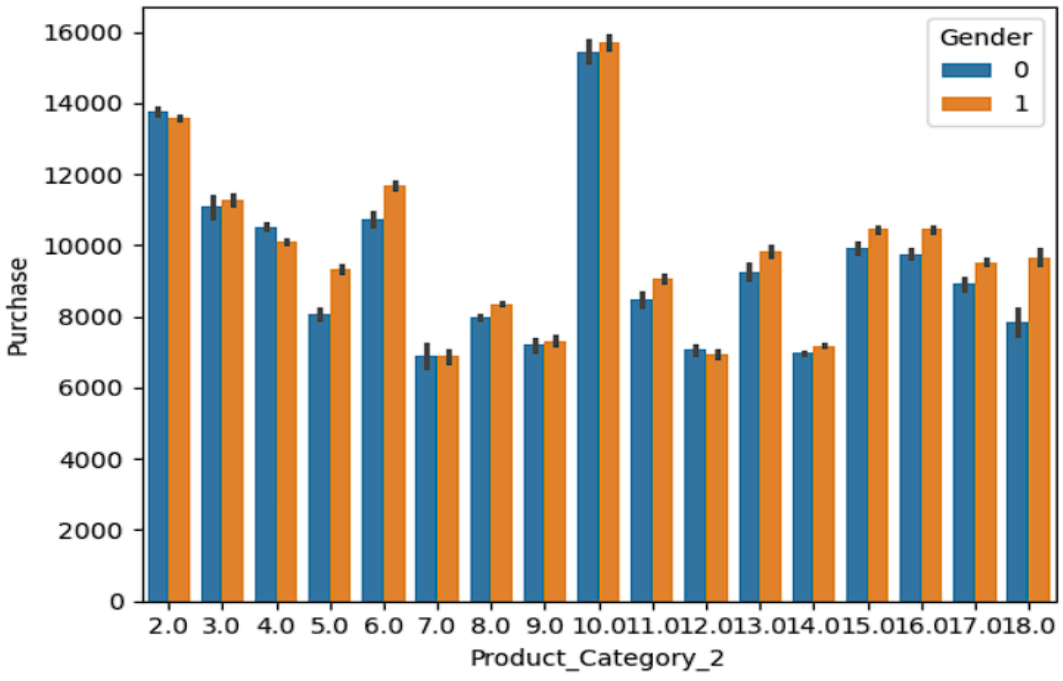
```
<Axes: xlabel='Product_Category_1', ylabel='Purchase'>
```



 # Visualization Product_Category_2 Vs Purchased

sns.barplot(x = "Product_Category_2", y = "Purchase", hue="Gender", data = df)

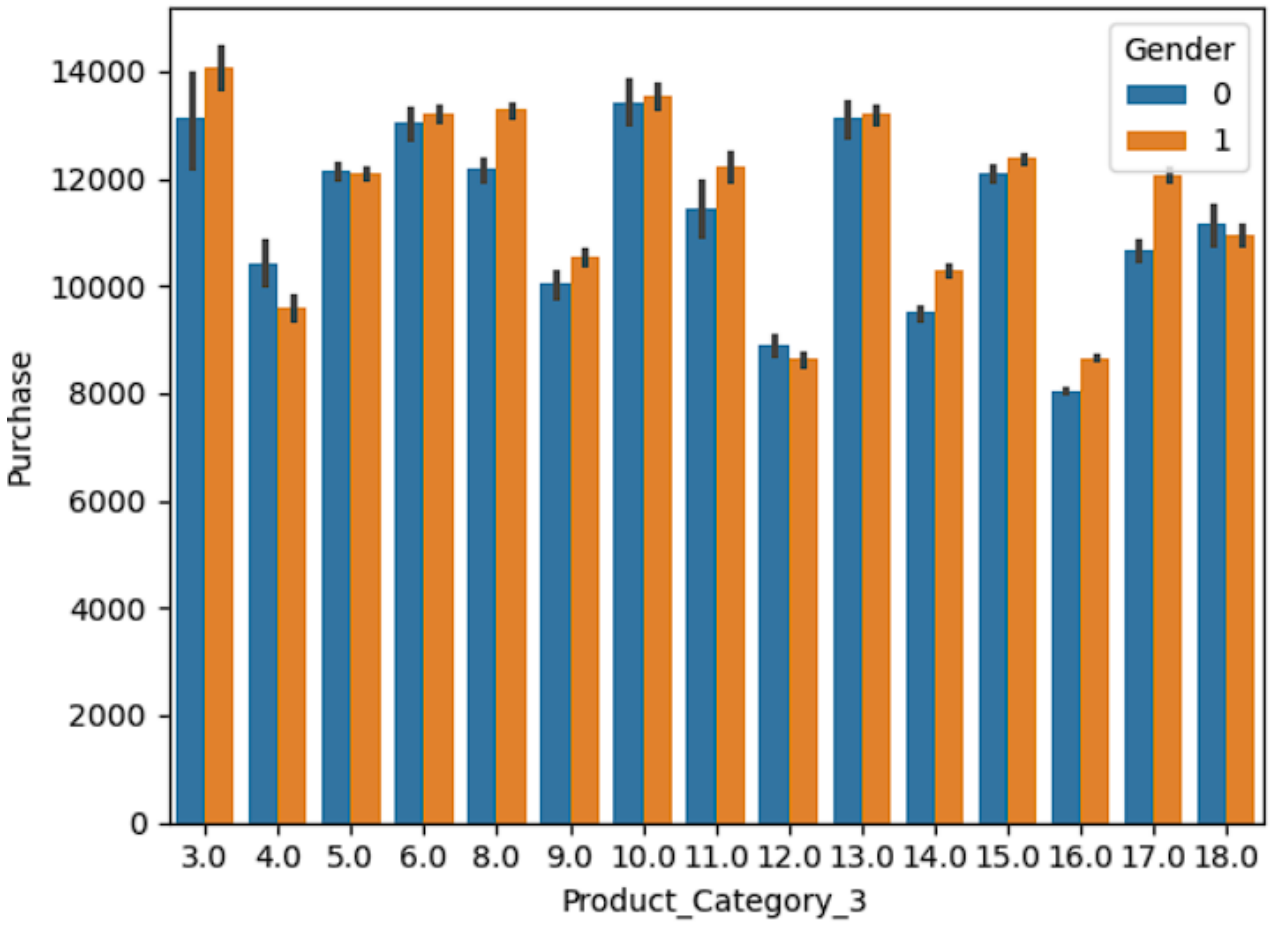
```
<Axes: xlabel='Product_Category_2', ylabel='Purchase'>
```



 # Visualization Product_Category_3 Vs Purchased

sns.barplot(x = "Product_Category_3", y = "Purchase", hue="Gender", data = df)

```
<Axes: xlabel='Product_Category_3', ylabel='Purchase'>
```



PROBLEM STATEMENT

Objective:

The goal of this analysis is to explore and understand various aspects of Black Friday sales data. The focus is on identifying key trends, customer behaviors, and sales patterns during the Black Friday period.

Data:

The dataset used for this analysis includes information on Black Friday sales with the following relevant attributes:

- **Customer ID:** Unique identifier for each customer.
- **Product ID:** Unique identifier for each product.
- **Purchase Amount:** The amount spent by the customer on a particular product.
- **Product Category:** The category to which the product belongs (e.g., Electronics, Clothing).
- **Purchase Date:** The date on which the purchase was made.

Analysis Tasks:

- **Identify Key Sales Trends:** Analyze the overall sales trends during the Black Friday period.
- **Examine Customer Behavior:** Study the purchasing behavior of different customer segments.
- **Analyze Product Performance:** Evaluate the performance of various product categories.
- **Visualize Sales Patterns:** Create visualizations to illustrate sales patterns And customer behaviors.

CONCLUSION

The analysis of the Black Friday Dataset has provided valuable insights into various aspects of consumer behavior, purchasing patterns, and sales trends.

Overall, the analysis has enhanced our understanding of Black Friday shopping habits and preferences. By identifying popular product categories, examining demographic influences, and comparing spending patterns, we gain insights that can inform marketing strategies, inventory management, and sales forecasting.

This comprehensive analysis supports informed decision-making and highlights areas for further exploration in retail analytics.

FUTURE SCOPE

E-Commerce Growth:

Continued Shift to Online Shopping: The COVID-19 pandemic accelerated the shift toward online shopping, and this trend is likely to continue. Retailers will invest more in their e-commerce platforms, enhancing user experience, and providing personalized shopping experiences.
Omnichannel Strategies: Retailers will increasingly adopt omnichannel strategies, integrating online and offline shopping experiences. This could include options like buy online, pick up in-store (BOPIS), and enhanced return policies.
Globalization of Black Friday:

International Adoption: Black Friday is already gaining traction in countries outside the U.S., such as Canada, the UK, and various Asian markets. As global e-commerce continues to grow, Black Friday sales may expand further into new international markets.
Localization of Promotions: Retailers may tailor Black Friday promotions to fit local cultures and shopping habits, making the event more relevant to diverse consumer bases.
Sustainability and Ethical Shopping:

Focus on Sustainability: There is a growing consumer demand for sustainable and ethically produced products. Retailers may respond by promoting eco-friendly products during Black Friday, appealing to environmentally conscious shoppers.
Circular Economy Initiatives: Some retailers may introduce initiatives that promote recycling, upcycling, or reselling of products, encouraging consumers to make more sustainable choices during the shopping event.
Technological Innovations:

Augmented Reality (AR) and Virtual Reality (VR): Retailers may use AR and VR to enhance the shopping experience, allowing customers to visualize products in their homes or try on clothing virtually.
Artificial Intelligence: AI can be leveraged for personalized marketing, inventory management, and customer service, enhancing the overall shopping experience and efficiency during Black Friday sales.
Social Media and Influencer Marketing:

Increased Role of social media: Social media platforms will play a significant role in promoting Black Friday deals, with retailers leveraging targeted ads and influencer partnerships to reach potential customers.
Live Shopping Events: The rise of live shopping events on platforms like Instagram, Facebook, and TikTok may become more prevalent, allowing consumers to shop in real-time while engaging with influencers and brands.
Consumer Behavior Changes:

Shift to Experience Over Products: As consumers increasingly value experiences over material goods, retailers may adapt by offering unique experiences or services during Black Friday, such as exclusive events or workshops.
Early Shopping Trends: The trend of starting Black Friday promotions earlier in the month may continue, with retailers encouraging consumers to shop earlier to avoid crowds and spread out sales.
Mobile Shopping:

Growth of Mobile Commerce: As mobile shopping continues to rise, retailers will focus on optimizing their mobile platforms and apps for a seamless shopping experience, including mobile-exclusive deals and promotions.
Data Privacy and Security:

Increased Focus on Data Protection: With the rise of online shopping, consumers will be more concerned about data privacy and security. Retailers will need to invest in robust security measures and transparent data practices to build trust with customers.

REFERENCES

www.google.com
www.python.org/doc
www.youtube.com
www.microsoft.com
www.copilotmicrosoft.com