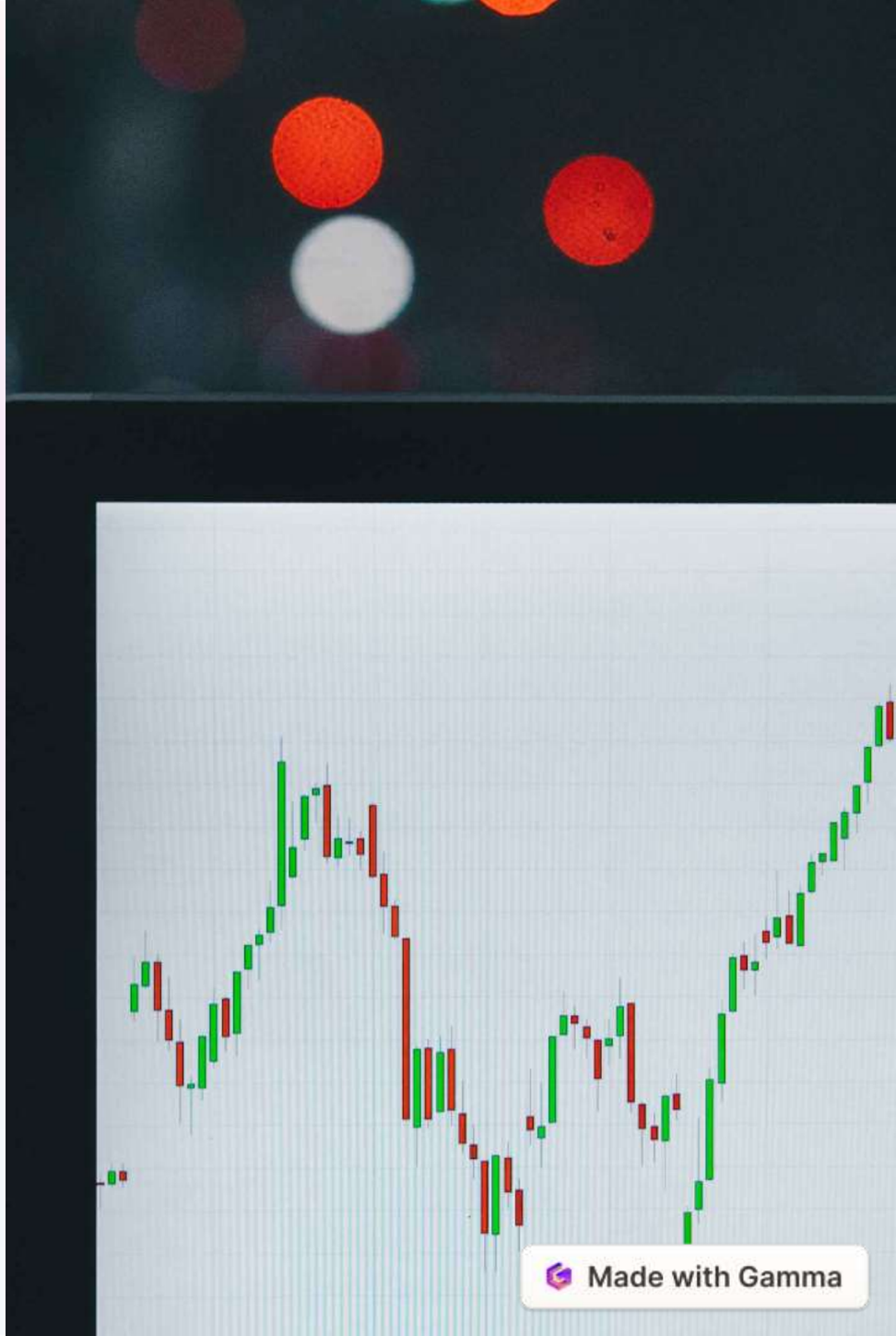
Project on

Data Analyzer

Submitted By : Lata (8721126)

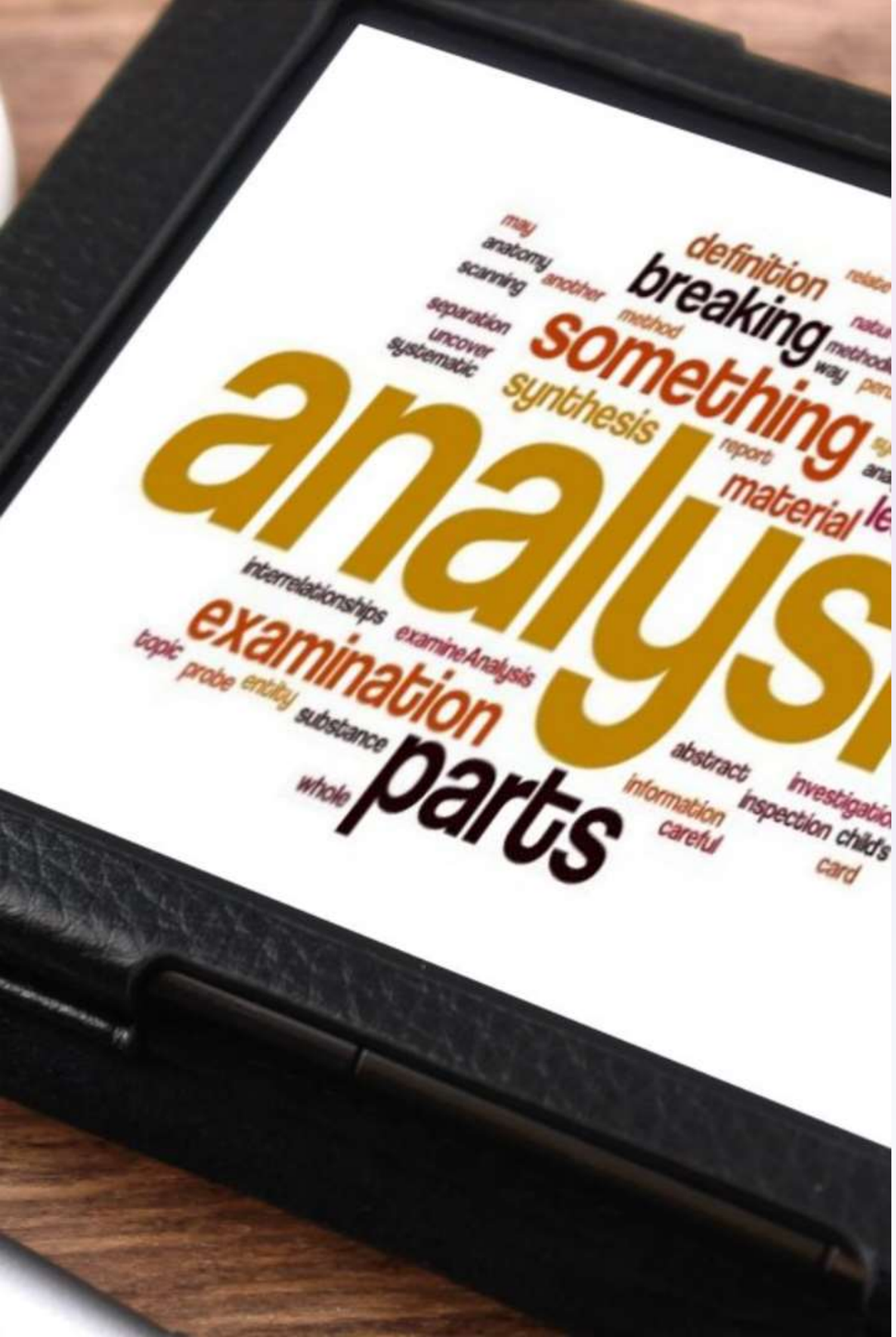
Computer Engineering (3rd year)

Submitted To : Dr. Rama Chawla (Assistant Prof.)



Content of the Presentation

1. Introduction about project
2. Objective
3. Problem Statement
4. Methodology
5. Modules
6. Result
7. Conclusion
8. References



Introduction about project

Data Analysis is the technique of collecting, transforming, and organizing data to make future predictions and informed data-driven decisions. It also helps to find possible solutions for a business problem.

Objective

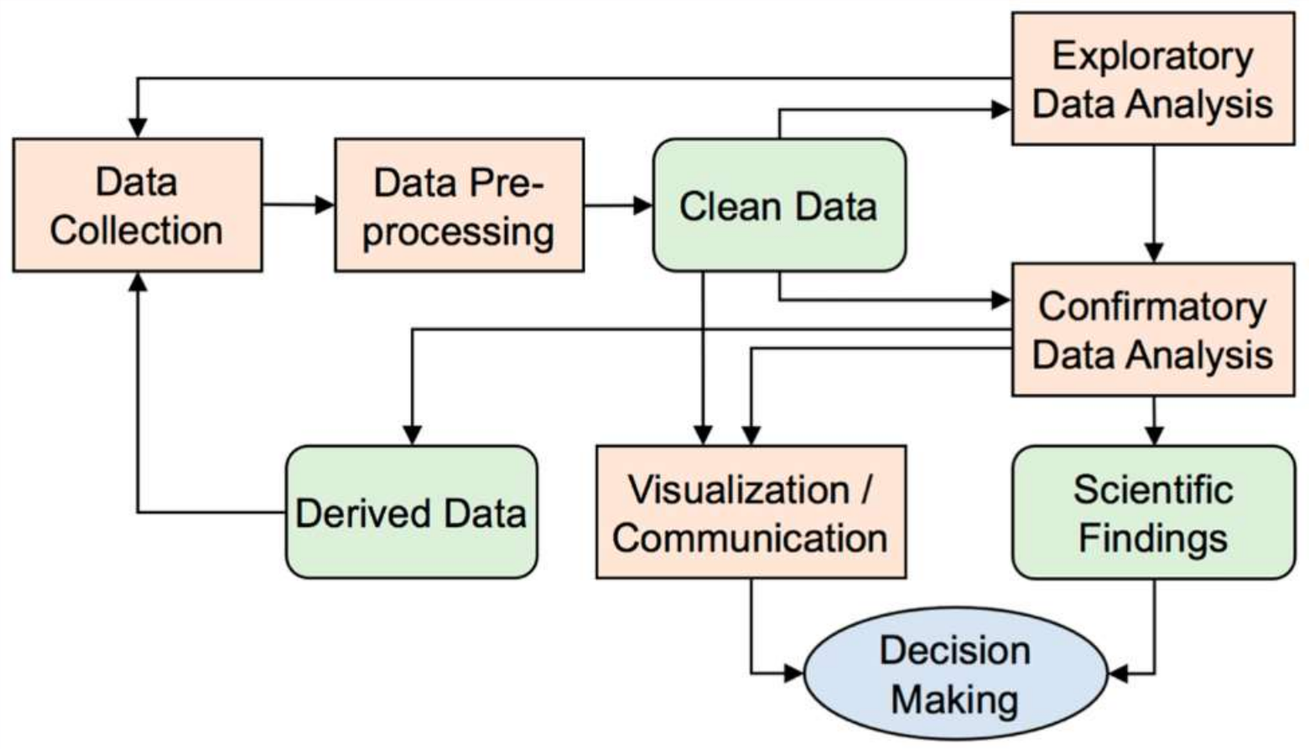
- Overview
- Managing data
- Business profit

Problem Statement

- Timeline
- Data Quality
- End-user
- Cost

Methodology

- Data Collection
- Data Cleaning
- Data Analysis
- Data Visualization



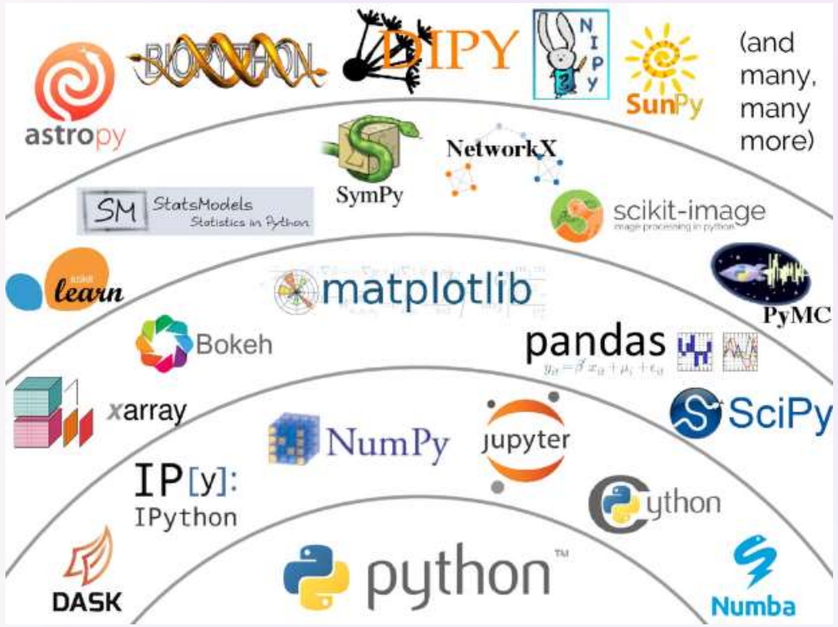
Working

1. Step 1: Acquiring the Data
2. Step 2: Preprocessing and Cleaning the Data
3. Analyzing the Data
4. Step 4:Exploratory Data Analysis
5. Create Interactive Dashboards and Visualizations in Power BI

Modules

- Import Libraries and Csv file
- Preprocessing and Cleaning the Data
  - Shape
  - head
  - info
  - drop
  - isnull
  - dropna
  - rename
  - describe
- Exploratory Data Analysis
  - Gender
  - Age
  - State
  - Marital Status
  - Occupation
  - Product Category

Tools



- **Pandas**: for data manipulation and analysis
- **Matplotlib**: for data visualization
- **Seaborn**: for enhanced statistical data visualization
- **NumPy**: for mathematical operations on arrays
- **Power BI**: enables users to visualize and share insights from their data.

These libraries provide a wide range of functionalities that are essential for working with data and conducting data analysis tasks efficiently.

Result

Married women age group 26-35 yrs from UP, Maharashtra and Karnataka working in IT, Healthcare and Aviation are more likely to buy products from Food, Clothing and Electronics category*

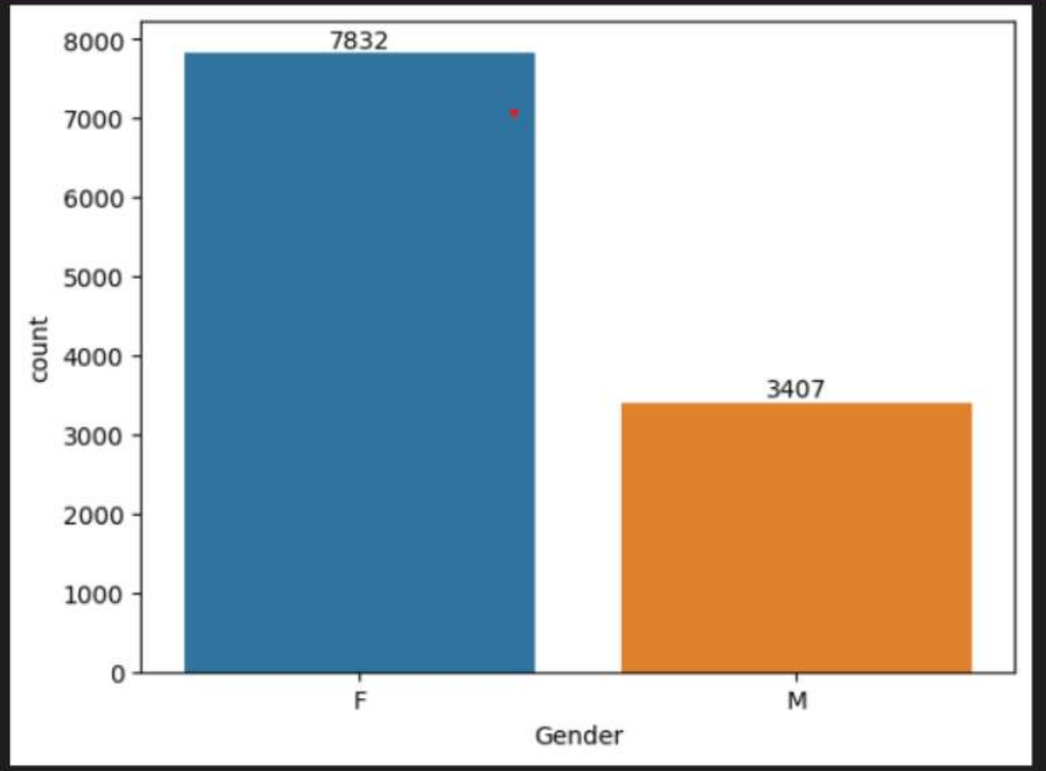
Gender

```
# plotting a bar chart for Gender and it's count

ax = sns.countplot(x = 'Gender',data = df)

for bars in ax.containers:
    ax.bar_label(bars)
```

[15]



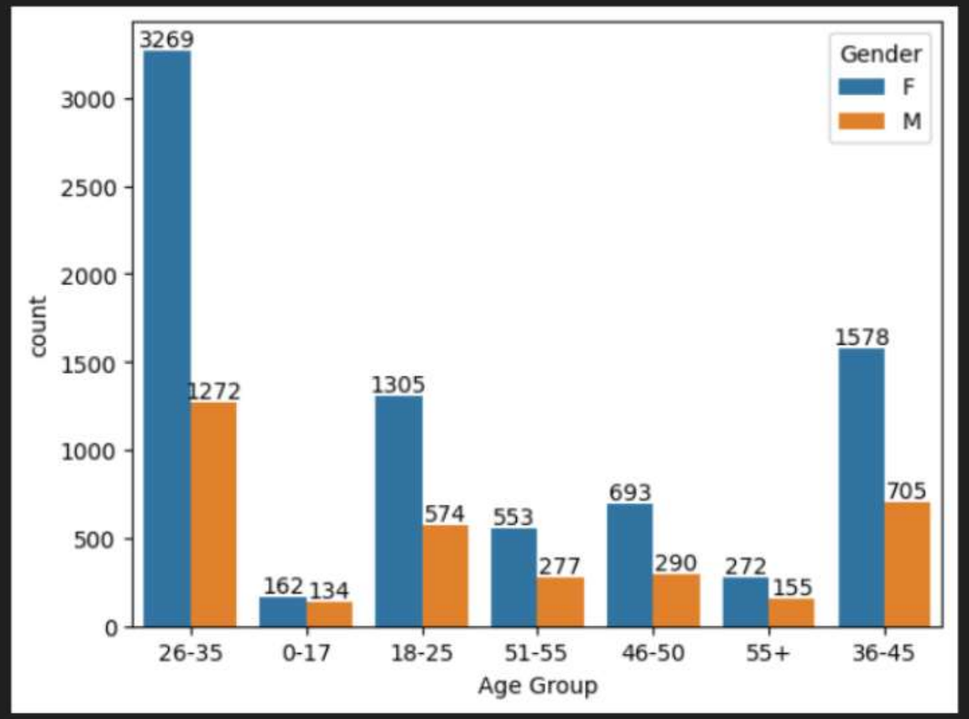
Age

```
ax = sns.countplot(data = df, x = 'Age Group', hue = 'Gender')

for bars in ax.containers:
    ax.bar_label(bars)
```

[17]

...

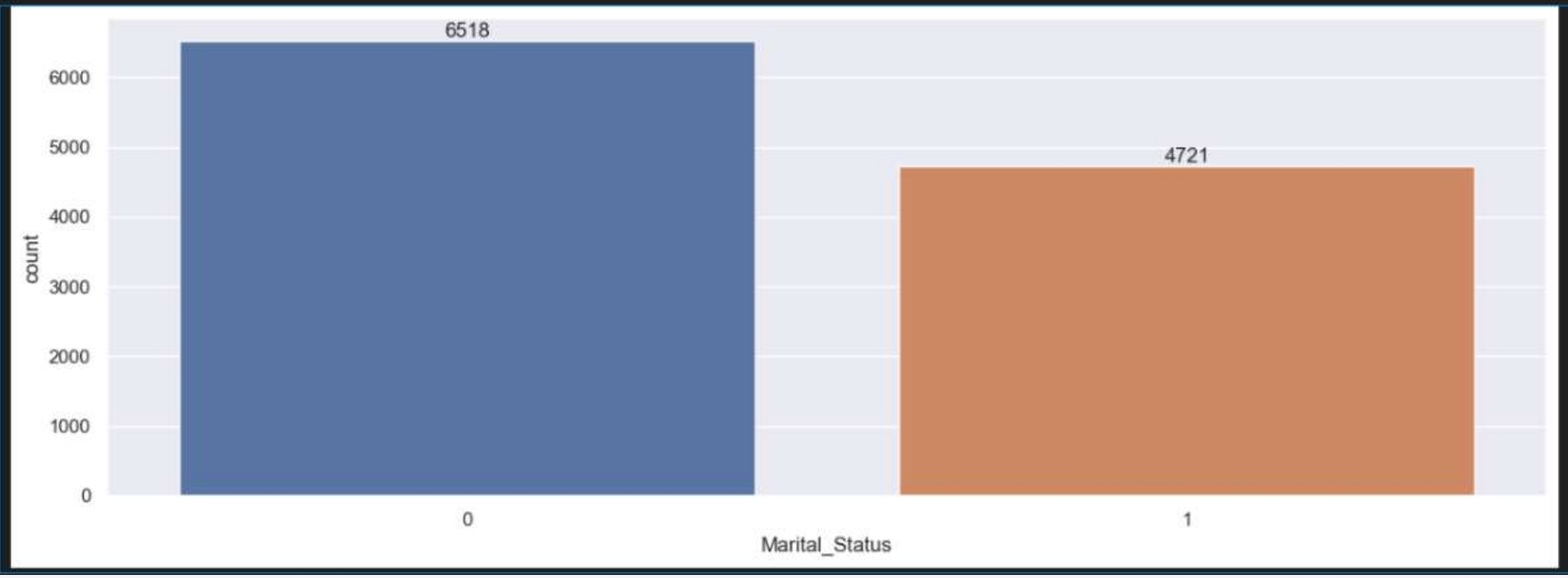


Marital Status

```
ax = sns.countplot(data = df, x = 'Marital_Status')

sns.set(rc={'figure.figsize':(7,5)})
for bars in ax.containers:
    ax.bar_label(bars)
```

[21]

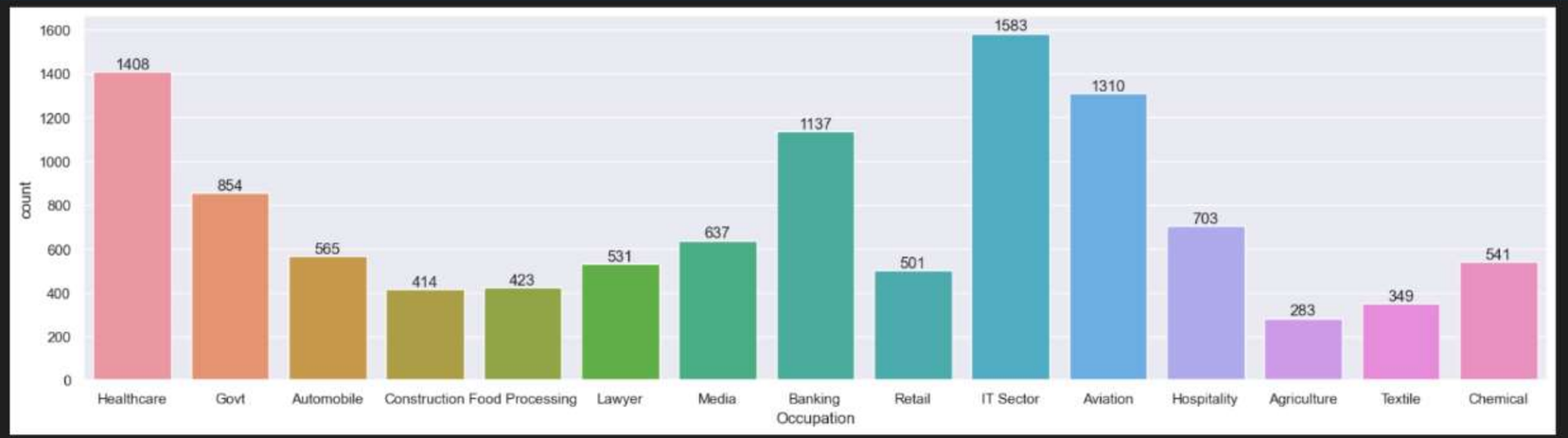


Occupation

```
sns.set(rc={'figure.figsize':(20,5)})
ax = sns.countplot(data = df, x = 'Occupation')

for bars in ax.containers:
    ax.bar_label(bars)
```

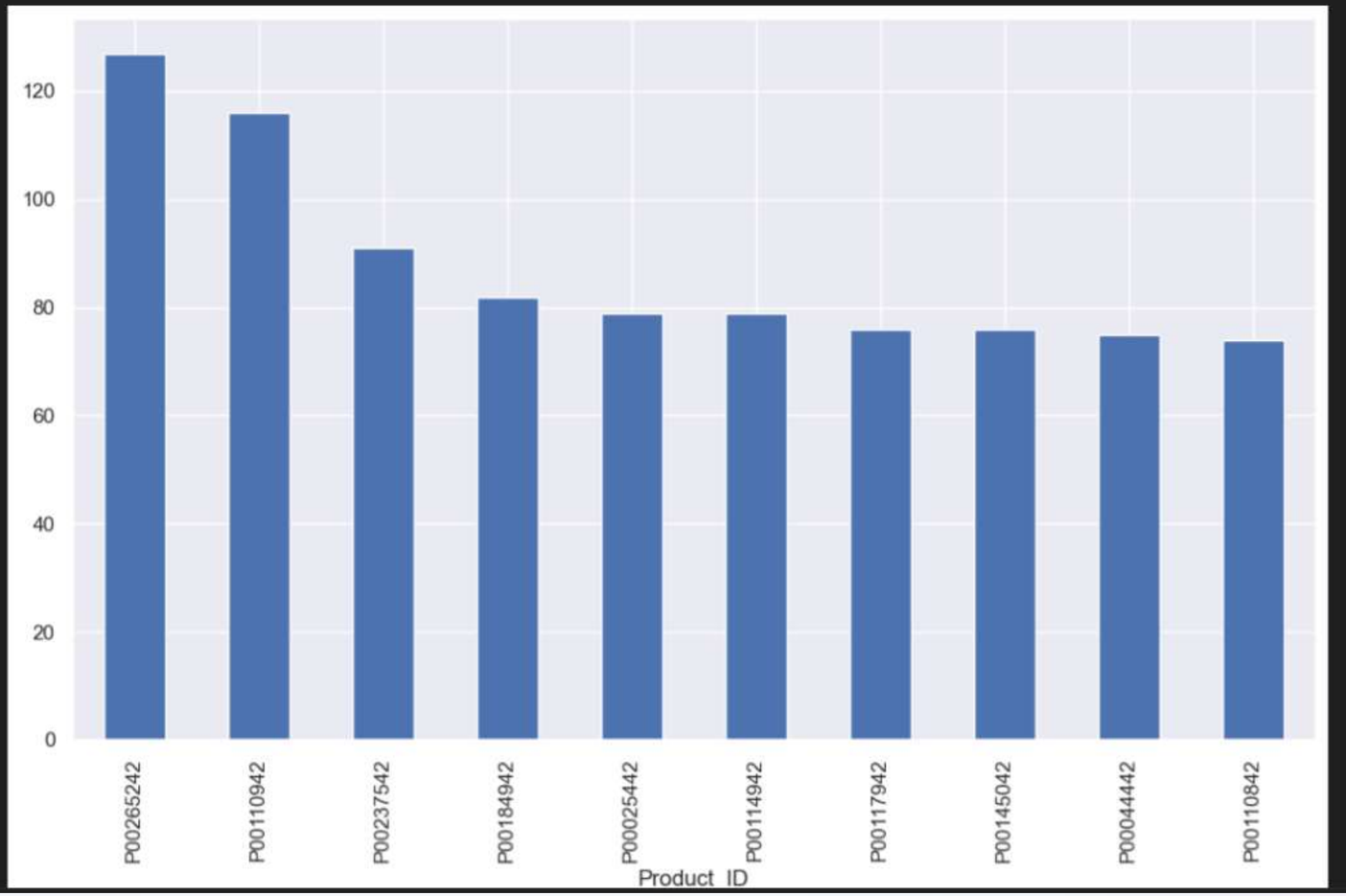
[23]

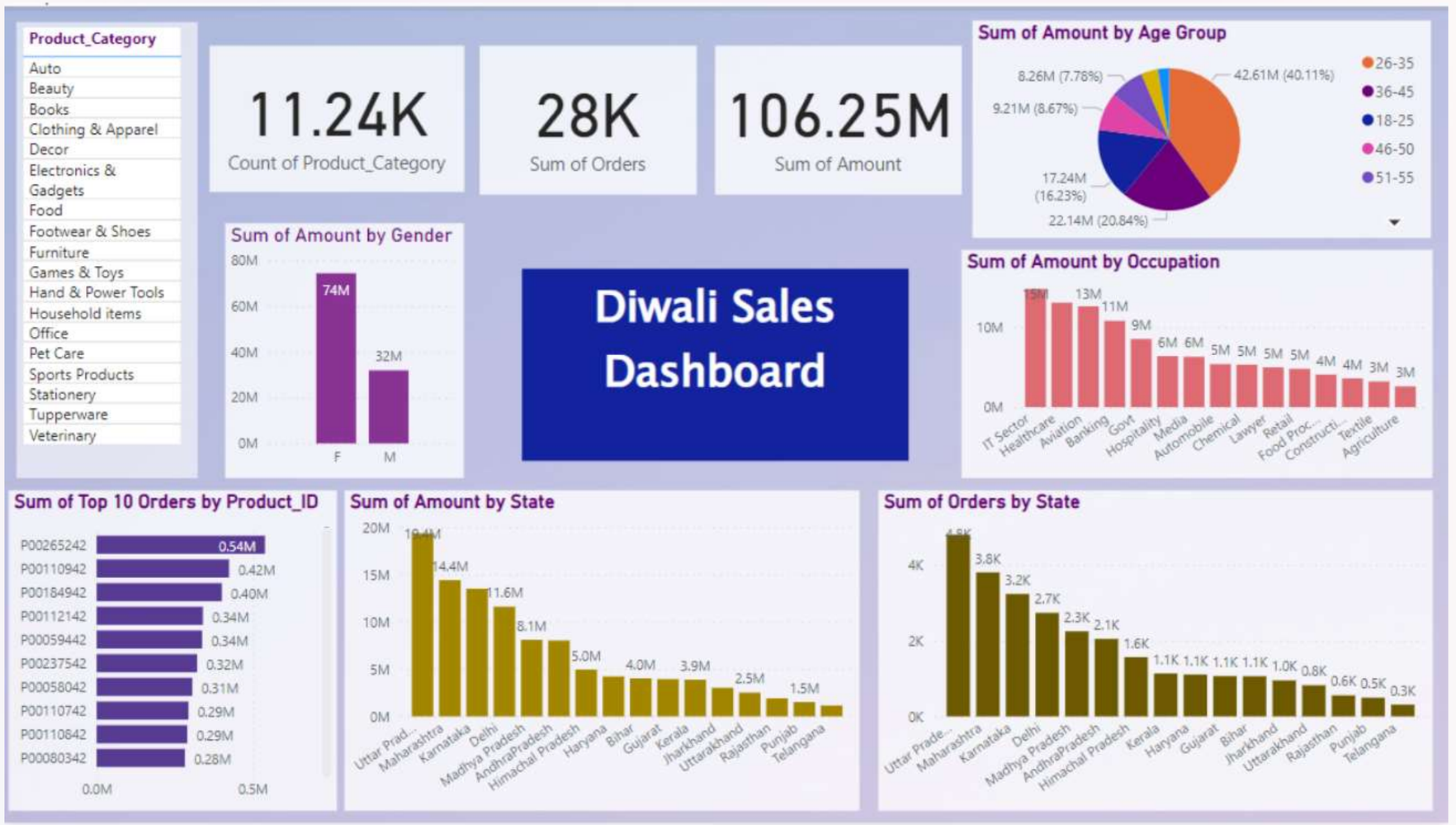


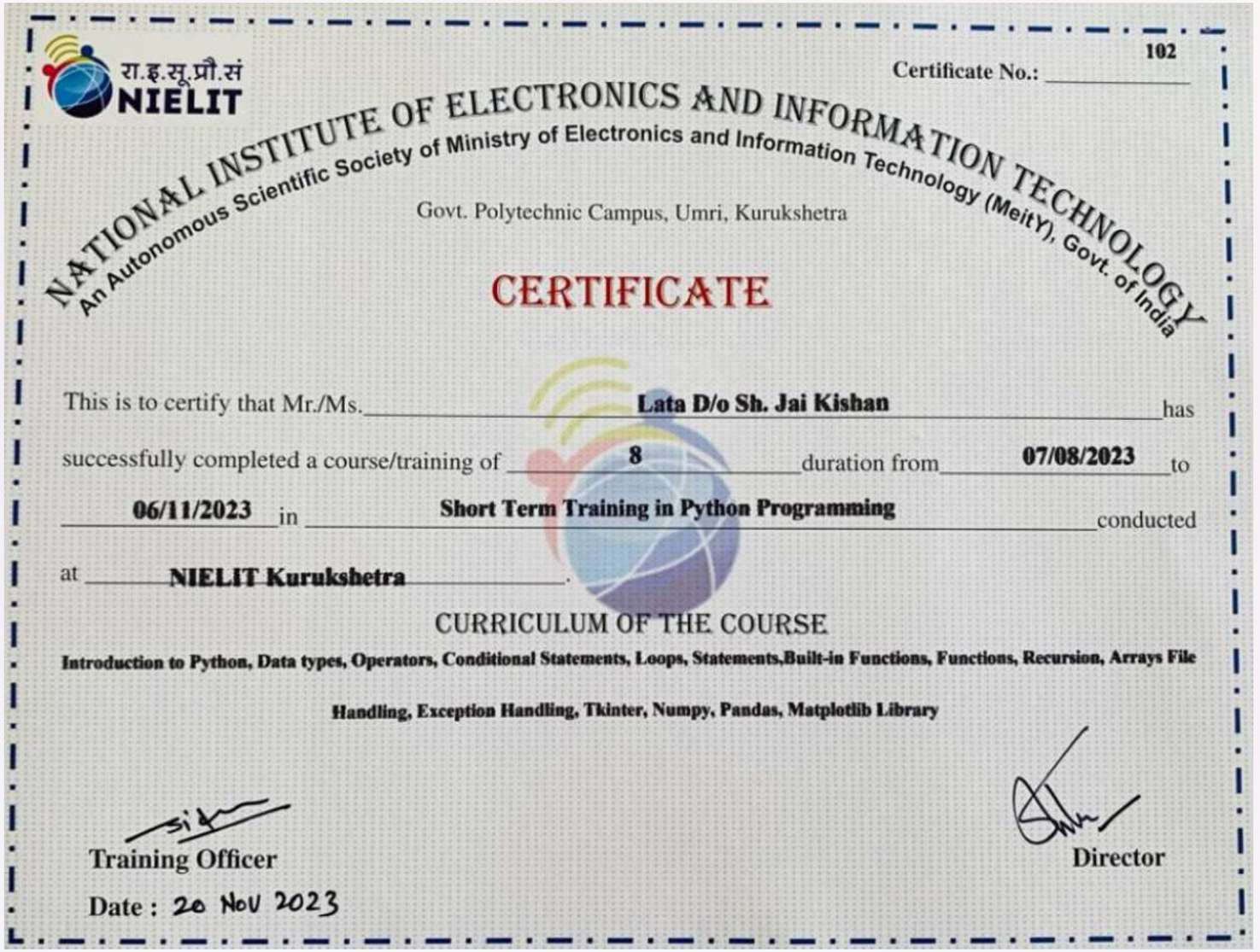
```
# top 10 most sold products (same thing as above)

fig1, ax1 = plt.subplots(figsize=(12,7))
df.groupby('Product_ID')['Orders'].sum().nlargest(10).sort_values(ascending=False).plot(kind='bar')
```
[28]

```
<Axes: xlabel='Product_ID'>
```







रा.इ.सू.प्रौ.सं
NIELIT

Certificate No.: 102

NATIONAL INSTITUTE OF ELECTRONICS AND INFORMATION TECHNOLOGY

An Autonomous Scientific Society of Ministry of Electronics and Information Technology (MeitY), Govt. of India

Govt. Polytechnic Campus, Umri, Kurukshetra

CERTIFICATE

This is to certify that Mr./Ms. **Lata D/o Sh. Jai Kishan** has successfully completed a course/training of **8** duration from **07/08/2023** to **06/11/2023** in **Short Term Training in Python Programming** conducted at **NIELIT Kurukshetra**.

CURRICULUM OF THE COURSE

Introduction to Python, Data types, Operators, Conditional Statements, Loops, Statements,Built-in Functions, Functions, Recursion, Arrays File Handling, Exception Handling, Tkinter, Numpy, Pandas, Matplotlib Library

Training Officer

Director

Date : 20 Nov 2023

References

[1] https://www.python.org/

[2] https://pandas.pydata.org/

[3] https://matplotlib.org/

[4] https://www.w3schools.com/python/

[5] https://powerbi.microsoft.com/en-gb/

Thank You

Email ID: vanshusharma4045@gmail.com