Here's a step-by-step guide to create a professional PDF report:



- a. Set up the environment
 - Use Jupyter Notebook .
 - Install required libraries:

pip install pandas matplotlib seaborn

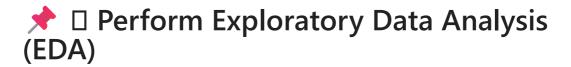
b. Import libraries

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



- Load the Titanic dataset or any dataset of your choice:

df = pd.read_csv('path_to_dataset.csv')



- a. Basic Information
 - Understand dataset structure:

df.shape df.columns df.info() df.describe() df.value_counts()

- b. Missing values analysis
 - Identify missing data:

df.isnull().sum()



- a. Univariate Analysis
 - Distribution of a single variable:

sns.histplot(df['column_name']) plt.show()

- b. Bivariate Analysis
 - Relationships between two variables:

sns.scatterplot(x='col1', y='col2', data=df) plt.show()

- c. Multivariate Analysis
 - Correlation heatmap:

sns.heatmap(df.corr(), annot=True, cmap='coolwarm') plt.show()

d. Other useful plots

Boxplot

Countplot

Pairplot



Interpretations and Observations

After each visualization, write your observations clearly:

What patterns do you see?

Any trends?

Anomalies?



↑ □ Summary of Findings

- Summarize key points at the end of the notebook:
- Trends observed
- Important variables
- Anomalies detected

- Business insights (if applicable)

In []: