Exploratory Data Analysis (EDA) Report

1. Tasks Performed

- Library Imports: numpy, pandas, matplotlib, seaborn.
- Dataset Loading: Loaded using pandas read_csv().
- Initial Data Inspection: head(), info(), describe().
- Checking Missing Values: isnull().sum().
- Basic Data Cleaning: Removed/handled missing values.
- Descriptive Statistics: Summarized data characteristics.
- Data Visualization: Histograms, heatmaps, scatterplots.
- Correlation Analysis: Identified feature relationships.

2. Learnings

- Efficient use of Python data science libraries.
- Importance of early data inspection.
- Techniques to handle missing values.
- Skill to summarize data with statistics.
- Use of visualizations to extract insights.
- Understanding of feature correlations.
- Data cleaning for better model building.
- Realization of EDA's role in projects.

3. Graph Insights

1. Histograms:
   - Showed distributions of features.
   - Identified skewness and outliers.
2. Correlation Heatmaps:
   - Visualized feature relationships.

Exploratory Data Analysis (EDA) Report

- Highlighted important feature pairs.

3. Pairplots:

- Revealed clustering/grouping patterns.

4. Scatter Plot (Age vs Fare):

- Younger passengers paid less.

- Outliers: Older passengers paying high fares.

5. Age Distribution Histogram:

- Majority are young adults.

- Few older individuals.