

Task 2: Exploratory Data Analysis (EDA) - Titanic Dataset

✓ Objective:

To understand the structure and relationships within the Titanic dataset using descriptive statistics and visualizations.

✨ Steps Performed:

1. Generated summary statistics using `.describe()`.
2. Created histograms to observe distributions of numerical features.
3. Plotted boxplots to detect outliers and spread of data.
4. Built a correlation matrix to evaluate linear relationships.
5. Used pairplots to explore multivariate interactions, especially with the `'Survived'` column.

☐ Tools Used:

- Python
- Pandas
- Matplotlib
- Seaborn

📁 Files in this Repo:

- `'titanic_eda.py'`: Python script with EDA steps and visualizations.
- `'histograms.png'`: Histogram plot for numeric columns.
- `'boxplot_<column>.png'`: Boxplots for individual numeric columns.
- `'correlation_matrix.png'`: Heatmap of feature correlations.
- `'pairplot.png'`: Pairwise plots of key features.

📥 Dataset Source:

[Kaggle Titanic Dataset] (<https://www.kaggle.com/datasets/yasserh/titanic-dataset>)

📄 Submission:

Submit this GitHub repository link using the form provided in the internship task.