# Exploratory Data Analysis Summary: Titanic Dataset

1. Overview

- Objective: To uncover insights, trends, and anomalies using statistical and visual methods.
- Dataset Used: Titanic Dataset
- Tools: Python (Pandas, Matplotlib, Seaborn)

2. Data Understanding

- Dataset Shape: (number of rows, number of columns)
- Basic Info: Summary of data types, missing values, etc.
- Statistical Summary: Using .describe() for numerical columns

3. Univariate Analysis

- Distribution of individual columns:
- Categorical: Countplots of Sex, Embarked, Pclass
- Numerical: Histograms and boxplots of Age, Fare

4. Bivariate/Multivariate Analysis

- Relationships between variables:
- Survival vs. Sex, Pclass, Age, Fare
- Heatmap of correlations between numerical variables
- Scatterplots or violin plots for deeper insights

5. Key Insights

- Example:

Females had a higher survival rate.

Passengers in 1st class survived more than those in 3rd.

Younger passengers had slightly better chances of survival.

6. Anomalies or Outliers

- Noted in Fare or Age distributions (e.g., extremely high fares)

7. Conclusion

- Patterns Found: (Summarize trends)
- Anomalies Noted: (Mention any strange or unexpected values)
- Next Steps (Optional): Consider feature engineering or preparing for modeling.