Exploratory Data Analysis (EDA) – Titanic Dataset

Objective: To extract insights from the Titanic dataset using statistical summaries and visualizations, identifying patterns, trends, and anomalies.

Tools & Libraries:

1. Python (Pandas, Matplotlib, Seaborn)
2. Kaggle Notebook (for coding and visualization)
3. Dataset: Titanic Dataset from Kaggle

Steps Performed:

1. Data Loading & Overview: Imported dataset from Kaggle and explored basic info.
2. Data Cleaning: Handled missing values and converted categorical variables where necessary.
3. Statistical Summary: Used .describe() and .value_counts() for initial analysis.
4. Visualizations: Histograms, Countplots, Boxplots, Scatterplots, and Correlation Heatmap.
5. Insights: Survival rates higher for women, children, and first-class passengers.

Key Findings:

1. Women and children had significantly higher survival rates.
2. Passenger class strongly influenced survival chances.
3. Higher fares often correlated with survival.
4. Cabin data had too many missing values, reducing usefulness.

Outcome: Gained skills in data cleaning, statistical summarization, visual trend analysis, and deriving meaningful insights.