**Titanic Dataset Analysis**

**1. Dataset Loading**

* Imported and loaded the Titanic dataset using Jupyter Notebook.
* Checked the basic structure

**2. Missing Values Analysis**

* Visualized missing values using a heatmap to identify columns with null data.

**3. Distribution Analysis**

* Performed distribution analysis on key numerical and categorical columns:
  + **Age:** Overall distribution and its effect on survival.
  + **Age vs Fare by Survival:** Compared age and fare with survival outcomes.
  + **Gender vs Survival:** Analyzed survival rates by gender.
  + **Survival rate by Parch:** Checked the relationship between number of parents/children aboard and survival.

**4. Relationship Analysis**

* Identified correlations between numerical columns using a heatmap to detect relationships and multicollinearity.

**5. Grouping and Aggregation**

* Calculated survival rates based on categorical measures:
  + Survival rate by **gender**
  + Survival rate by **passenger class (Pclass)**
  + Survival rate by **number of siblings/spouses aboard (SibSp)**

**6. Outlier Detection**

* Identified outliers in numerical columns using boxplots and statistical methods.

**7. Skewness and Transformation**

* Analyzed skewness of numerical variables.
* Applied visualizations and log transformations where necessary to normalize skewed distributions.