**Titanic Dataset - Exploratory Data Analysis (EDA) Report**

**Introduction**

**The goal of this report is to explore the Titanic dataset using statistical summaries and visualizations to uncover trends, patterns, and anomalies related to passenger survival during the Titanic disaster.**

**Dataset: Titanic training dataset  
Target Variable: Survived (0 = No, 1 = Yes)**

**Dataset Overview**

* **Rows**: 891
* **Columns**: 12
* **Key Features**: Pclass, Sex, Age, SibSp, Parch, Fare, Embarked, Survived

**Missing Values**

| **Column** | **Missing Count** |
| --- | --- |
| Age | 177 |
| Cabin | 687 |
| Embarked | 2 |

**Note**: The Cabin column is mostly missing and might be dropped or imputed with caution.

**Descriptive Statistics**

Key numerical columns:

* Age: Mean ~29.7, Range: 0.42–80
* Fare: Mean ~32.2, Wide range up to 512
* SibSp and Parch: Most values are low, indicating smaller family groups

**Univariate Visualizations**

* **Age**: Bell-shaped distribution, mostly between 20–40 years
* **Fare**: Right-skewed, many passengers paid under $50
* **Survival Count**:
  + Survived: ~38%
  + Did not survive: ~62%

**Categorical Insights**

| **Feature** | **Observation** |
| --- | --- |
| **Sex** | Females had significantly higher survival rates |
| **Pclass** | 1st class passengers had the highest survival |
| **Embarked** | Port C passengers had better odds than S |

**Boxplots & Violinplots**

* **Age vs Survival**: Younger passengers more likely to survive
* **Fare vs Survival**: Higher fare positively correlates with survival

**Correlation Heatmap**

Key insights from correlation matrix:

* Sex shows strong correlation with survival (female = more likely)
* Pclass and Fare are moderately correlated with survival
* Family size (SibSp, Parch) has less effect

**Pairplot**

Clear clusters observed:

* Survived females mostly from higher classes
* Non-survivors concentrated in lower-class, lower-fare zones

**Summary of Findings**

* **Survival is influenced by gender, class, and fare.**
* **Females and 1st-class passengers had the highest survival rates.**
* **Younger age and higher fare showed positive associations with survival.**
* **Port of embarkation slightly affects survival.**