**Titanic Data Analysis - EDA Report**

**Objective**

Perform Exploratory Data Analysis (EDA) on the Titanic dataset to uncover patterns, trends, and insights using visual and statistical exploration techniques.

**Tools Used**

* Python
* Pandas
* Matplotlib
* Seaborn
* Jupyter Notebook

**Key Steps**

1. **Data Loading and Inspection**  
   Loaded the train.csv dataset and examined its structure using .info(), .describe(), and .isnull().sum().
2. **Univariate Analysis**  
   Analyzed individual variables such as:
   * Sex, Pclass, and Embarked (categorical)
   * Age, Fare, SibSp, and Parch (numerical)
3. **Bivariate Analysis**  
   Explored the relationship between each feature and the target variable (Survived) using:
   * Bar plots for categorical variables
   * Box plots for numerical variables
4. **Multivariate Analysis**  
   Used:
   * Heatmap to visualize correlation between numeric variables
   * Pairplot to understand relationships among multiple features colored by Survived
5. **Summary of Observations**  
   Documented key insights from visualizations and statistics.

**Key Observations**

* **Sex**: Females had a significantly higher survival rate than males.
* **Pclass**: First-class passengers had the highest survival rate; third-class had the lowest.
* **Age**: Younger passengers (especially children) had better chances of survival.
* **Fare**: Higher fares (indicative of upper class) correlated with higher survival.
* **SibSp & Parch**: Passengers with small families had slightly better survival than those alone or in large families.
* **Missing Values**: Notable missing data in Age, Cabin, and Embarked columns.
* **Cabin**: Many missing values; potential insight if properly imputed.

**Conclusion**

EDA helped identify crucial factors that influenced passenger survival on the Titanic. The most important variables were **Sex**, **Pclass**, and **Age**. These insights will guide future feature engineering and model development for prediction.