

Dataset Analysis – Titanic Dataset The Titanic dataset contains information about 891 passengers and 12 attributes such as age, gender, ticket class, and fare. The target variable in this dataset is Survived. The dataset contains both numerical and categorical data. Numerical features include Age, Fare, SibSp, and Parch. Categorical features include Sex, Ticket, Cabin, and Embarked. The passenger class (Pclass) is an ordinal variable. Exploratory analysis shows that the average age of passengers is around 29 years and ticket fares vary widely, indicating different economic classes. There are missing values in the Age and Cabin columns, which means data cleaning will be required before applying machine learning models. The target variable is slightly imbalanced, as more passengers died than survived. Overall, the Titanic dataset is suitable for classification tasks.