Data preprocessing

To apply various machine learning techniques to answer different questions, you should start by cleaning and preprocessing the data for each classification problem. For instance, with the Titanic dataset, you can prepare the data to be suitable for training and evaluating models for predicting or inferring demographic characteristics. Use this dataset to address eight distinct classification problems given below, ensuring that the data is properly preprocessed and that suitable features are selected for each model.

Instructions:

1. Data Preprocessing:

- Load the Titanic dataset and clean basic data.
- Handle missing values appropriately.
- Convert categorical variables into numerical representations.

2. Classification Tasks:

- For each of the following classification tasks, based on the target variable, select relevant features.

Classification Tasks:

1. Survival Prediction:

- Target Variable: `Survived`

2. Class Prediction:

- Target Variable: `Pclass`

3. Gender Prediction (Binary Classification):

- Target Variable: `Sex`

4. Embarkation Port Prediction:

- Target Variable: `Embarked`

5. Cabin Prediction (Binary: Has Cabin Information or Not):
   - Target Variable: Binary indicator for missing `Cabin` data.
6. Family Status Prediction (Binary Classification):
   - Target Variable: Binary indicator derived from `SibSp` and `Parch`.
7. High-Fare Prediction:
   - Target Variable: Binary indicator derived from `Fare`.
8. Age Group Prediction (Categorical Classification):
   - Target Variable: Age group categories.