

**NAME: SABIKA ZAHRA**

**ROLL NUMBER: 074**

**CLASS: BSAI**

**SECTION: ALPHA**

**SEMESTER: 4**

**SUBMITTED TO: SIR RASIKH ALI**

**DATE: 2,20,2025**

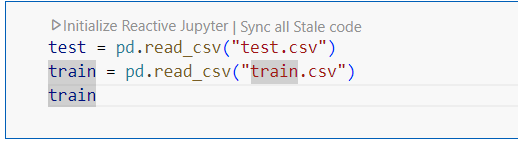
**1. Import Libraries**

The first step is to import the necessary libraries for data manipulation, preprocessing, and modeling.



### ****2. Load Data****

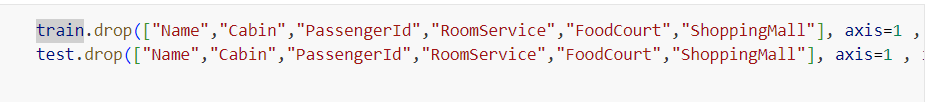
Load the training and testing datasets.





### ****4. Drop Unnecessary Columns****

Drop columns that are not needed for modeling. In this case, the column is dropped from the each dataset individually .



### ****5. Handle Missing Values****

Identify columns with a large number of missing values and drop them. Then, fill missing values in the remaining columns. Do it separately with both data sets.





### ****6. Convert Data Types****

Convert numerical columns with float data types to integers.



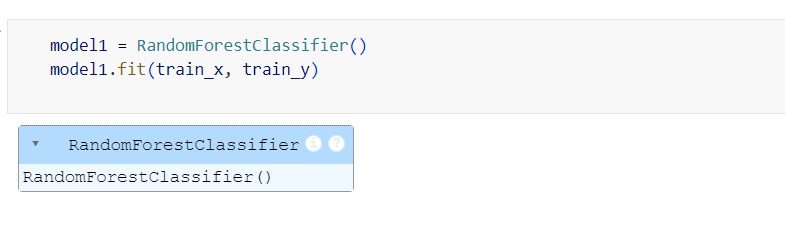


### ****7. Split into x and y:****



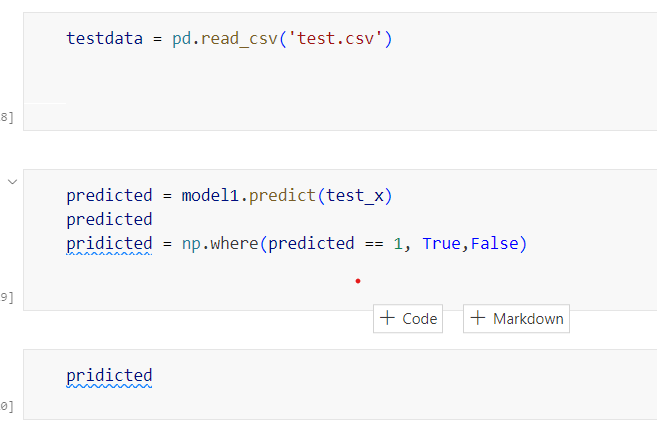
### ****8. Train the Model****

Train a RandomForestClassifier model on the training data.



### ****9. Make Predictions****

Use the trained model to make predictions on the test data.



### ****10. Create Submission File****

Create a submission file with the predicted values.

