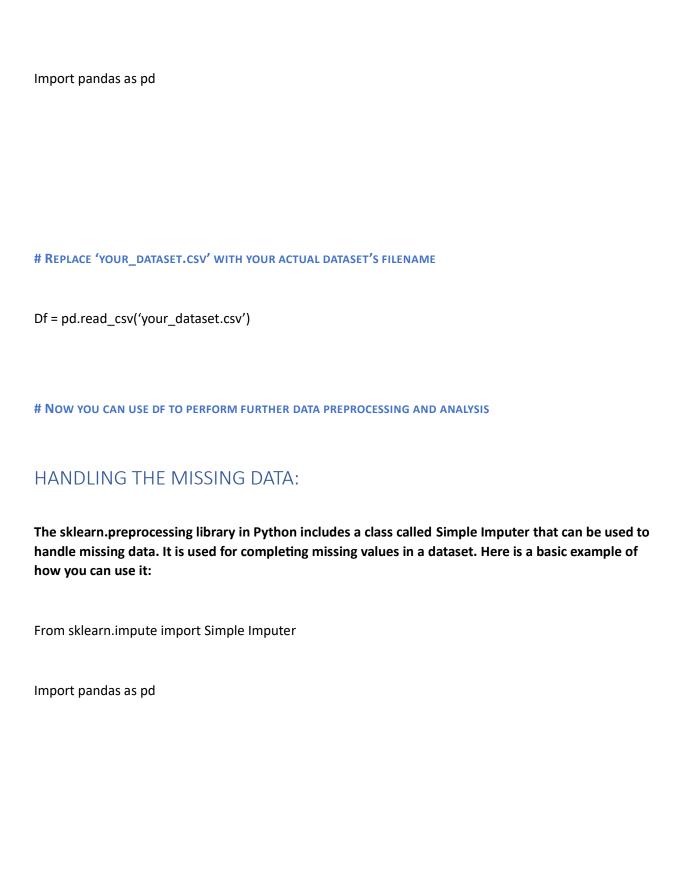
IMPORTING THE REQUIRED LIBRARIES:

| Additionally, if you plan to use advanced machine learning techniques, you might consider importing libraries like scikit-learn, TensorFlow, or Porch for building more complex models. |
|---|
| Import pandas as pd |
| Import numpy as np |
| Import matplotlib.pyplot as plt |
| Import seaborn as sns |
| From sklearn.model_selection import train_test_split |
| From sklearn.preprocessing import StandardScaler |
| From sklearn.linear_model import LinearRegression |
| From sklearn.metrics import mean_squared_error, r2_score |
| IMPORTING THE DATASET (READ DATA SET; CREATE MATRIX): |
| |

Make sure you have the file 'your_dataset.csv' in the same directory as your Python script or provide the full path if it's located elsewhere. This will load the data into a pandas DataFrame, which you can then use for data manipulation and analysis

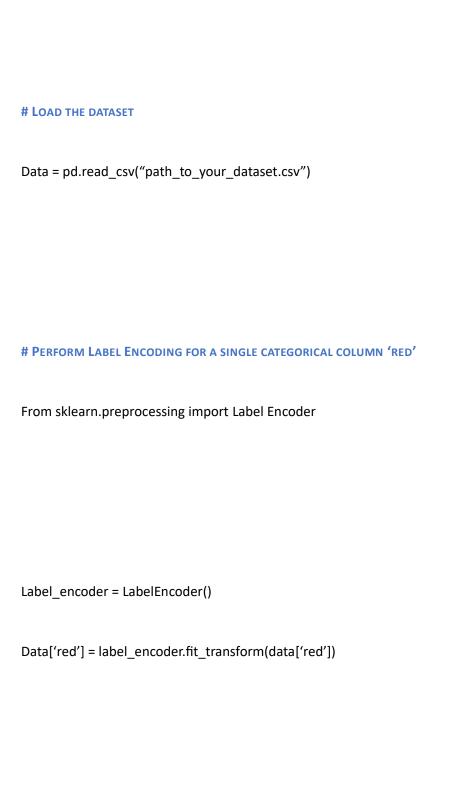


| # Load your dataset | |
|---|--|
| Data = pd.read_csv('path_to_your_dataset.csv') | |
| | |
| | |
| | |
| # | |
| # INITIALIZE THE IMPUTER | |
| Imputer = Simple Imputer (missing_values=np.nan, strategy='mean') # You can use 'median', 'most_frequent' or 'constant' as strategy | |
| | |
| | |
| | |
| # FIT AND TRANSFORM THE DATA | |
| THE AND TRANSFORM THE DATA | |
| Data_imputed = imputer.fit_transform(data) | |
| In this case, strategy can be set to 'mean', 'median', 'most_frequent', or 'constant' depending on how you want to impute the missing values. | |
| ENCODING THE CATEGORICAL DATA: | |

To encode categorical data in Python, you can use techniques like Label Encoding and One-Hot Encoding.

Here's an example using the pandas library:

Import pandas as pd

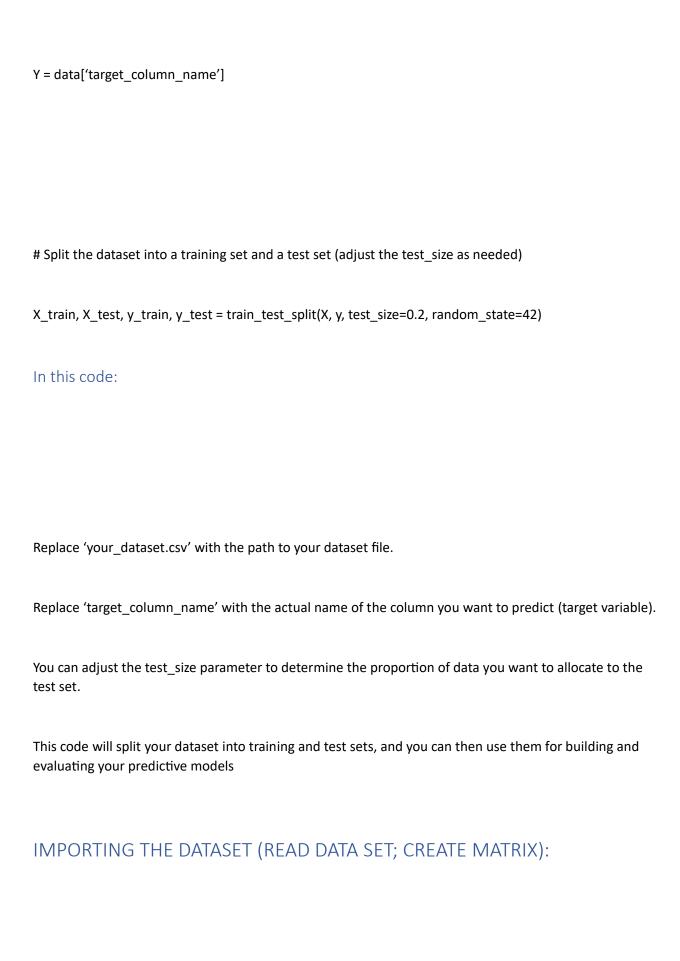


ALTERNATIVELY, FOR ONE-HOT ENCODING

| # data = pd.get_dummies(data, columns=['red']) |
|---|
| |
| |
| # Now 'RED' COLUMN HAS BEEN ENCODED. YOU CAN USE THE UPDATED DATAFRAME FOR FURTHER ANALYSIS OR MODELLING. |
| Make sure to replace "path_to_your_dataset.csv" with the actual path to your dataset file. |
| FEATURE SCALING AND IMPORT STANDARDSCALER: |
| To perform feature scaling and import the StandardScaler in Python, you can use the following code: |
| |
| |
| Python |
| Import pandas as pd |
| From sklearn.preprocessing import StandardScaler |
| |
| |









| Make sure to replace 'path_to_your_dataset.csv' with the actual path of your dataset. Also, ensure you have the necessary libraries installed. |
|--|
| |
| |
| |
| |
| |
| |
| |
| |
| |
| |