```
#  IMPORTNING THE LIBRARY

import numpy as np      #Array

import matplotlib.pyplot as plt

import pandas as pd

#-----------------------------------------

# import the dataset & divided my dataset into independe & dependent

dataset = pd.read_csv(r"C:\Users\A3MAX SOFTWARE TECH\Desktop\WORK\1. KODI
WORK\1. NARESH\1. MORNING BATCH\N_Batch -- 10.30AM_ M25\3. Dec 24\24th- ML\5.
Data preprocessing\Data.csv")

X = dataset.iloc[:, :-1].values

y = dataset.iloc[:,3].values

#-----------------------------------------

from sklearn.impute import SimpleImputer # SPYDER 4


imputer = SimpleImputer()

#-------------------------------------------------------------------------

imputer = imputer.fit(X[:,1:3])

X[:, 1:3] = imputer.transform(X[:,1:3])


#  HOW TO ENCODE CATEGORICAL DATA & CREATE A DUMMY VARIABLE

from sklearn.preprocessing import LabelEncoder

labelencoder_X = LabelEncoder()

labelencoder_X.fit_transform(X[:,0])

X[:,0] = labelencoder_X.fit_transform(X[:,0])

#-------------------------------------------------------------------------
labelencoder_y = LabelEncoder()

y = labelencoder_y.fit_transform(y)

#-----------------------------------------------------------------

#SPLITING THE DATASET IN TRAINING SET & TESTING SET

from sklearn.model_selection import train_test_split

X_train,X_test,y_train,y_test = train_test_split(X, y, test_size= 0.2,
random_state=0)

# if you remove random_stat then your model not behave as accurate

#-----------------------------------------------------------------

#FEATURE SCALING
```

```
from sklearn.preprocessing import Normalizer

sc_X = Normalizer()

X_train = sc_X.fit_transform(X_train)

X_test = sc_X.transform(X_test)

#------------------------------------------------------------------------
```