# import packages

import pandas as pd

import numpy as np

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

df = pd.read\_csv('StudentStudyHour.csv')

df

# dropping rows which have null values

df.dropna(inplace=True,axis=0)

df

y = df['Scores']

X = df.drop('Scores',axis=1)

# splitting the dataframe into train and test sets

X\_train,X\_test,y\_train,y\_test = train\_test\_split(X,y,test\_size=0.3,random\_state=101)

print(X\_train)

print(X\_test)

scaler = StandardScaler()

scaler.fit(X\_test)

X\_train = scaler.transform(X\_train)

X\_test = scaler.transform(X\_test)

print(X\_train)

print(X\_test)

lr = LinearRegression()

model = lr.fit(X\_train,y\_train)

y\_pred = model.predict(X\_test)

df = pd.DataFrame({'y\_test':y\_test,'y\_pred':y\_pred})

df

print(mean\_squared\_error(y\_test,y\_pred))