#importing libraries

import numpy as np

import pandas as pd

from sklearn import linear\_model

from sklearn.cross\_validation import train\_test\_split

#loading California dataset from dataset library

from sklearn.datasets import fetch\_california\_housing

california=fetch\_california\_housing()

print(california)

#we are transferring the data into two different dataframes

df\_x=pd.DataFrame(california.data,columns=california.feature\_names)

df\_y=pd.DataFrame(california.target)

#Describing the dataset california

df\_x.describe()

#training the regression model

reg=linear\_model.LinearRegression()

# we are splitting the data into test and train

x\_train,x\_test,y\_train,y\_test=train\_test\_split(df\_x,df\_y,test\_size=0.2,random\_state=4)

#fitting the data into the model

reg.fit(x\_train,y\_train)

#calculating the coefficients

reg.coef\_

#we are predicting the prices

a=reg.predict(x\_test)

a[6] # for checking the data to see hoew much difference from the orginalprice

y\_test #data

#finding the Mean Square Error(MSE)

np.mean((a-y\_test)\*\*2) #to see how much error we are getting