****

**House price prediction model**

**Using linear regression**

**DATA & WEB Mining**

**Department of Computer Science and Engineering**

**SRM University, AP**

**April - 2021**

By

M Pawan Aditya AP18110010085

B Akshay AP18110010641

M Kartheek AP18110010043

Krishna Santosh AP18110010458

D Narasimha AP18110010648

**Abstract:**

The house price prediction is a model which will be predicting the house price based on attributes like average area, average no. of bedrooms, area population etc. so as this data model is the basic case of prediction we will be using the linear regression algorithm for the training of the model.

**INTRODUCTION:**

In the modern world everyone would be interested to buy house at any time in life so we have come with a application of predicting the price of house so the one in need can use this application so he can have a estimation of price of the house to buy. The problem statement is like a real estate agent want help to predict the house price in USA region and so the data set will be containing 7 attributes columns namely average area income, average area house age, average area number of rooms, average area number of bedrooms, area population, Address and price. The price the outcome for the testing data set.

**DATASET Description:**

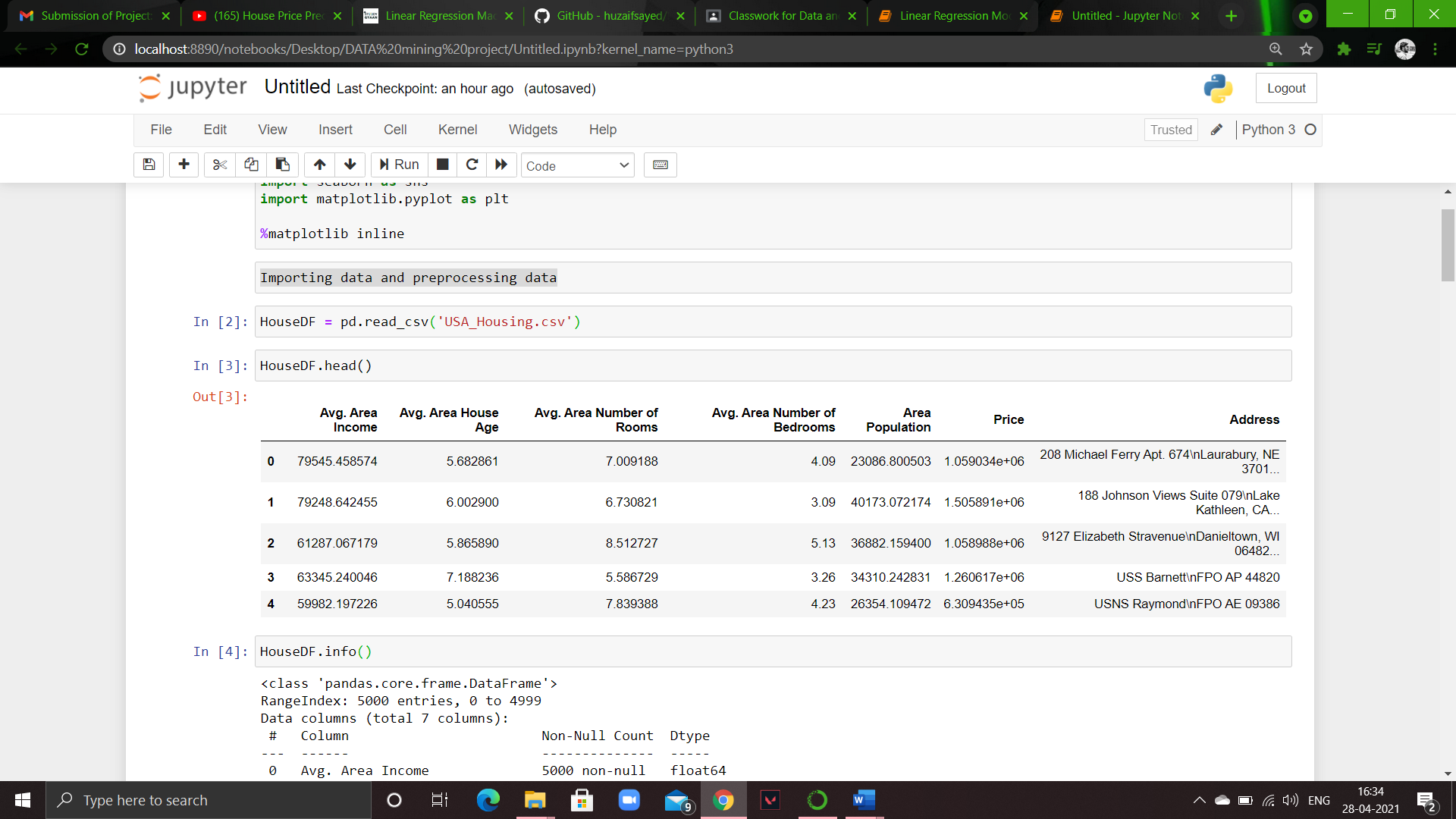
We have taken the dataset from Kaggle site one of most famous site for dataset. The data set has 5000 house prices based on the attributes. the data set will be containing 7 attributes columns namely average area income, average area house age, average area number of rooms, average area number of bedrooms, area population, Address and price .We will be using python language for the implementation of the model and Jupiter notebook as platform . we have used various libraries like pandas, numpy, seaborn, matplotlib for various functions.

**Preprocessing:**

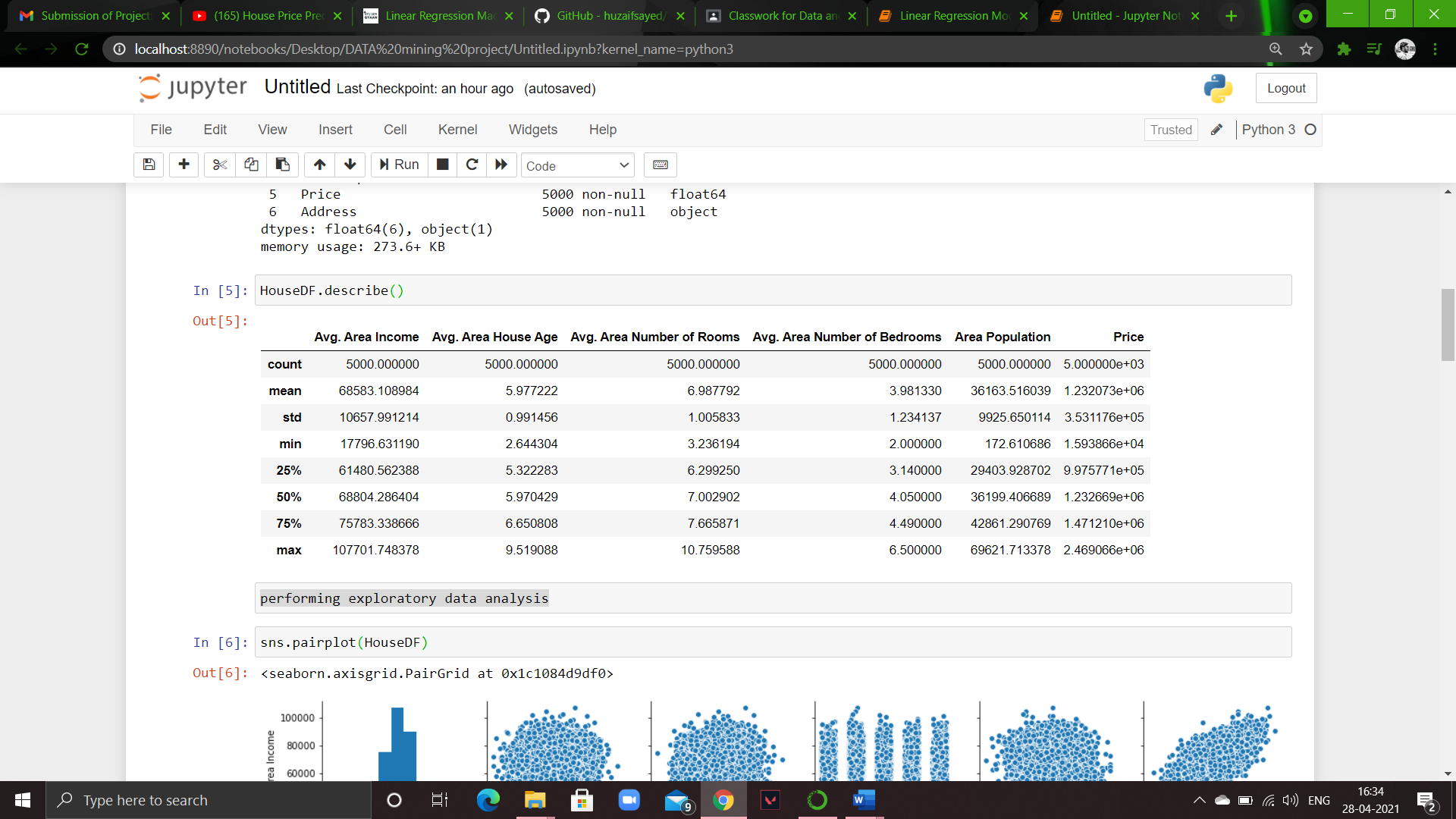
First we have check the head of the dataset and description of the data set so we can know what type of data is present in each column and to know the mean, standard deviation, maximum and minimum values of the column so we can have an idea of the data set.

Then we have checked for (isnull) to check if and column or row values is null that is empty so we can fill those cells to avoid the error.

This is the head of dataset



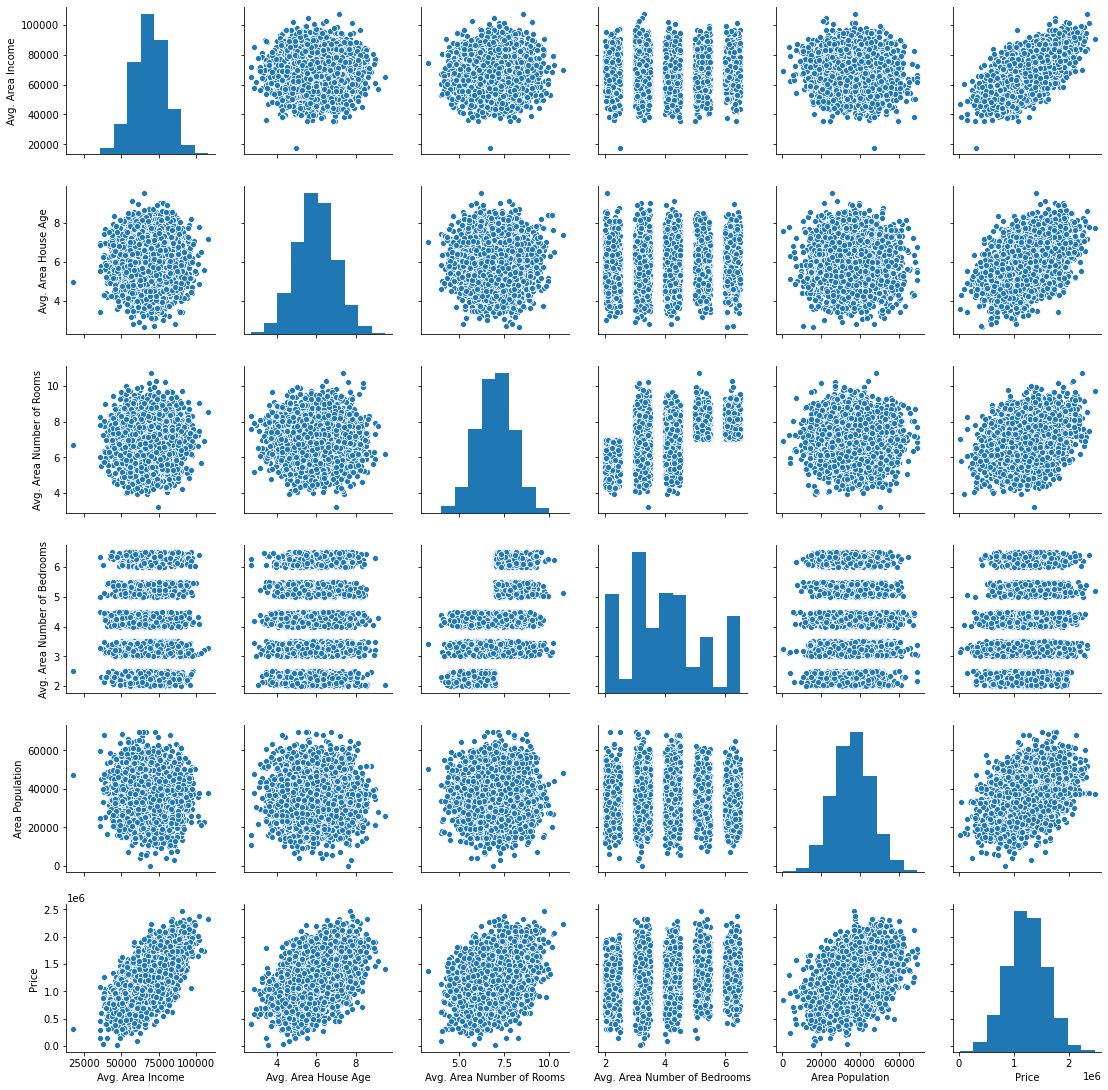
This the description of dataset



**Exploratory Data analysis:**

we have performed the exploratory data analysis to study the dataset to have and overview of the dataset and columns in it and to know the relationship between the columns so we can reduce or do something with it.

At first we have pair plot of the columns to know the relationship between them.

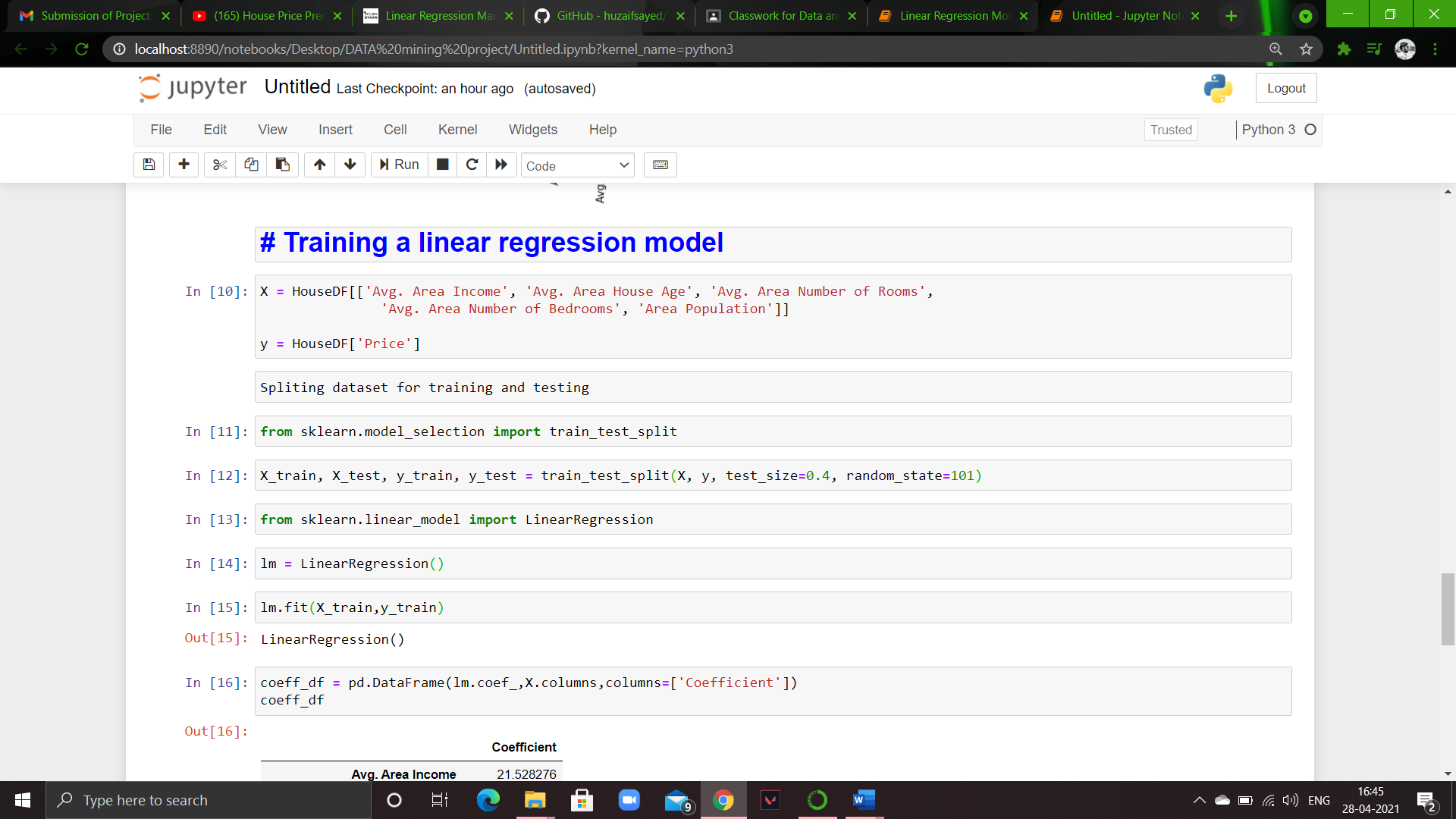


Then we have plot a heat map to see the corelation of the dataframe

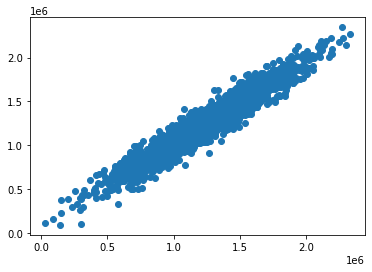


**Code implementation:**

The we have trained the model by splitting the dataset for training and testing.



**Prediction of our model:**

****

**Conclusion:**

As we can see the prediction graph have a shape of a line the model has predicted well

These are the evaluation of the model

