House price prediction

GOAL

prediction of house price in US using linear regression.

DATASET

In folder “USA\_Housing”.

DESCRIPTION

The main aim of the project is to create a model that can predict house prices in US.

WORK DONE

Importing the required packages into our python environment

Importing the house price data and do some EDA on it

Data Visualization on the house price data

Feature Selection & Data Split

Modeling the data using the algorithms

Evaluating the built model using the evaluation metrics

MODELS USED

We will need to first split up our data into an X list that contains the features to train on, and a y list with the target variable, in this case, the Price column. We will ignore the Address column because it only has text which is not useful for linear regression modeling.

X and y List

X = HouseDF[['Avg. Area Income', 'Avg. Area House Age', 'Avg. Area Number of Rooms',

'Avg. Area Number of Bedrooms', 'Area Population']]

y = HouseDF['Price']

Split Data into Train, Test

Now we will split our dataset into a training set and testing set using sklearn train\_test\_split(). the training set will be going to use for training the model and testing set for testing the model. We are creating a split of 40% training data and 60% of the training set.

from sklearn.model\_selection import train\_test\_split

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.4, random\_state=101)

X\_train and y\_train contain data for the training model. X\_test and y\_test contain data for the testing model. X and y are features and target variable names.

Creating and Training the LinearRegression Model

We will import and create sklearn linearmodel LinearRegression object and fit the training dataset in it.

from sklearn.linear\_model import LinearRegression

lm = LinearRegression()

lm.fit(X\_train,y\_train)

OUTPUT

LinearRegression(copy\_X=True, fit\_intercept=True, n\_jobs=None, normalize=False)

LinearRegression Model Evaluation

Now let’s evaluate the model by checking out its coefficients and how we can interpret them.

print(lm.intercept\_)

OUTPUT

-2640159.796851911

coeff\_df = pd.DataFrame(lm.coef\_,X.columns,columns=['Coefficient']) coeff\_df

LIBRARIES NEEDED

pandas

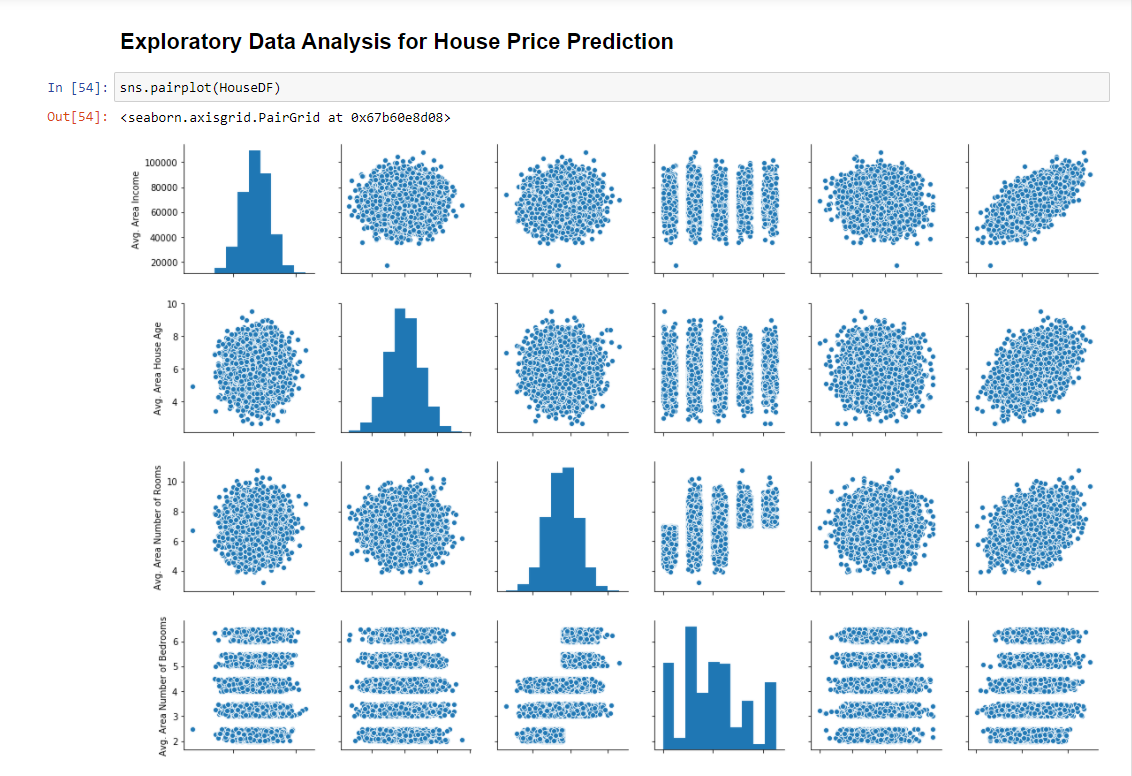
numpy

seaborn

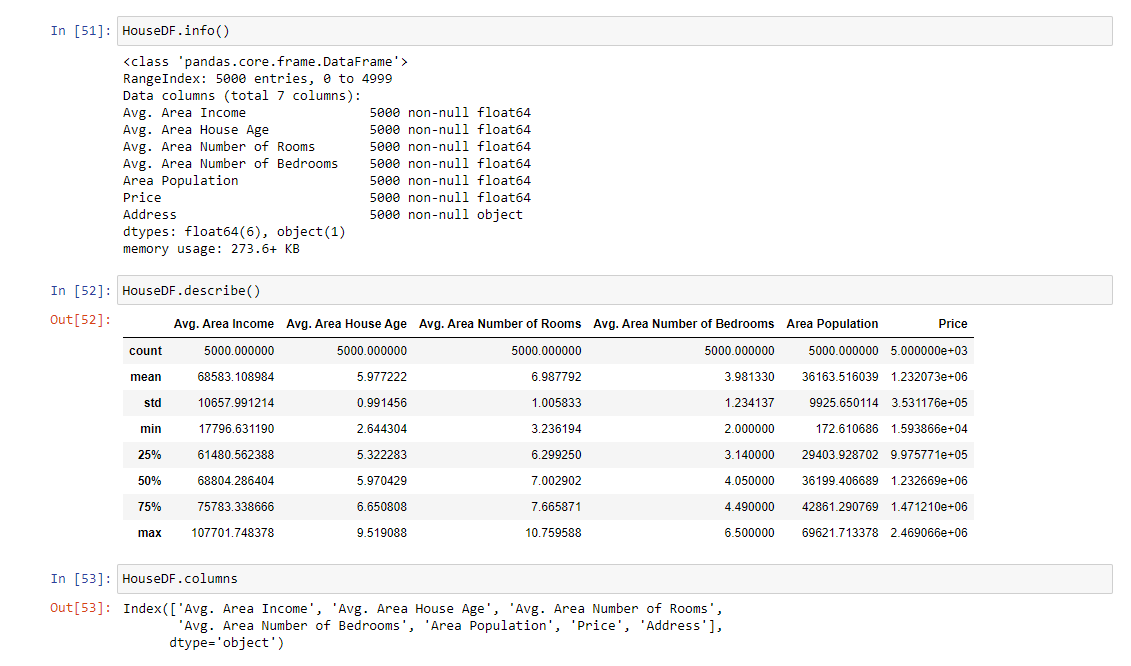
matplotlib.pyplot

EXAMPLES

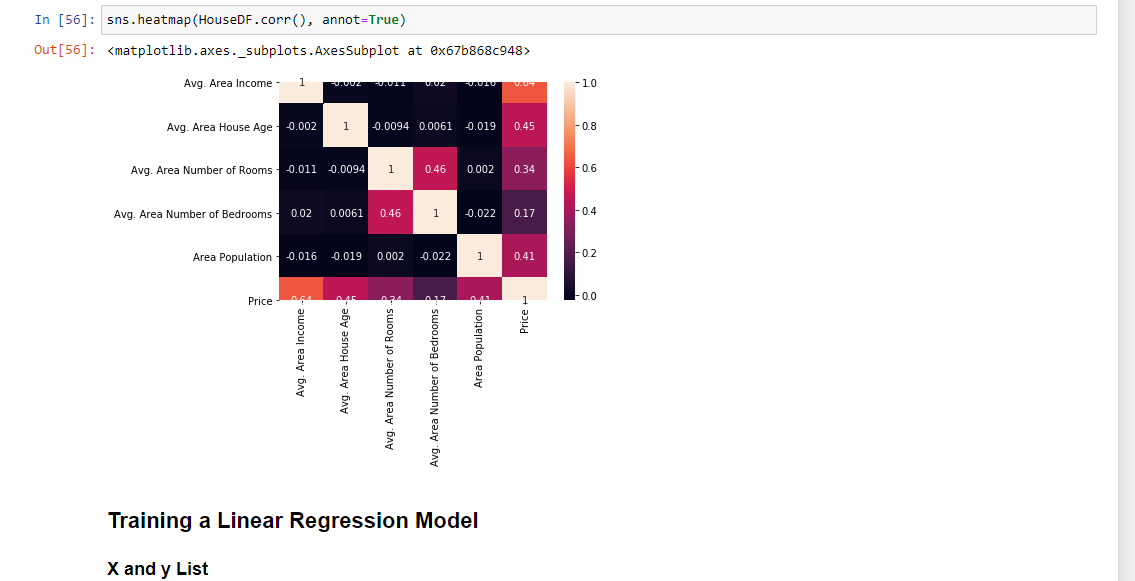
Example 1



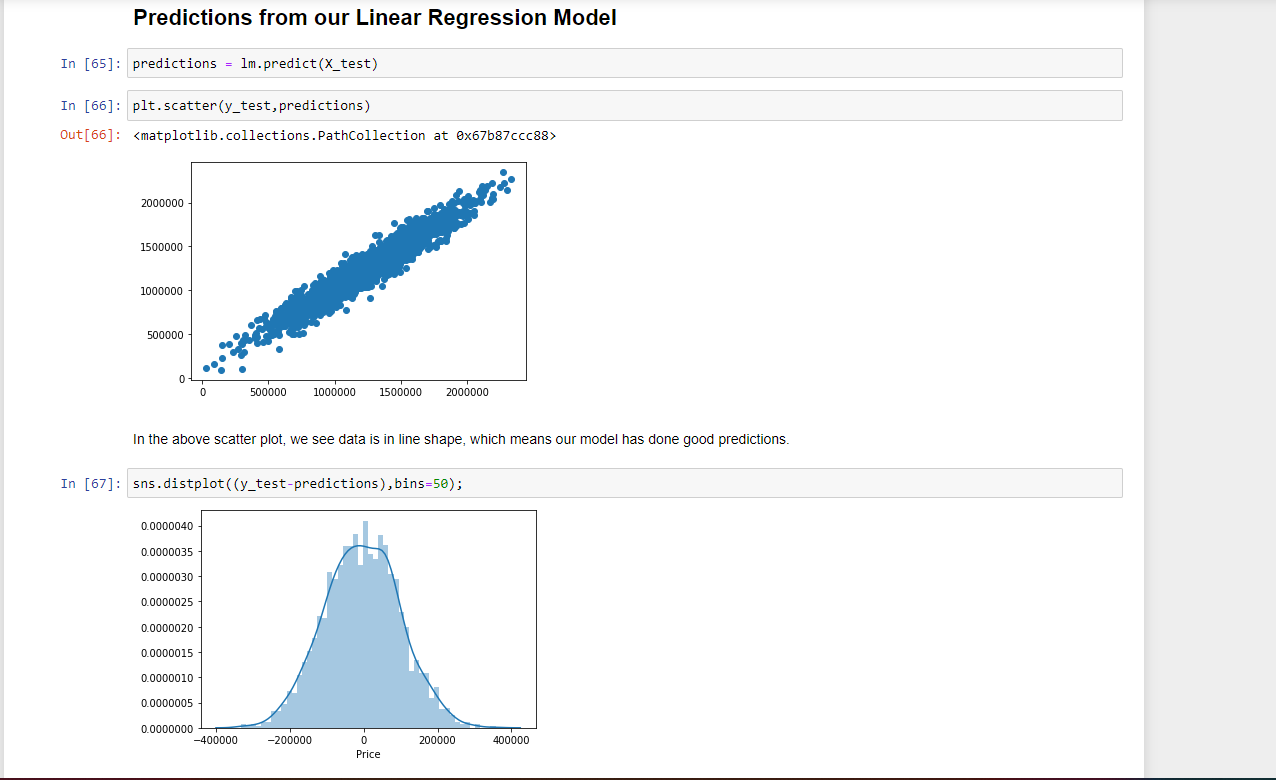
Example 2



Example 3



Example 4



CONCLUSION

We have created a Linear Regression Model which we help the real state agent for estimating the house price.

CONTRIBUTION BY

Vanshika goel

LinkedIn GitHub