**House Price predictions using Linear Regression**

* Pandas, matplotlib, seaborn, NumPy and scikit-learn were used
* Divided the train.csv dataset into train(0.7) +test (0.3) data
* Selected the columns such that the Loss function (RMSE) is minimized
* Therefore, out of 15 columns only 5 were chosen that is - ['nitrogen\_o2','room per dwelling', 'distance', 'pupil teacher ratio', 'lower']

A screenshot of a graph

Description automatically generated

* Applied Linear Regression on the dataset taking ‘medv(median)’ as the value for testing

A screenshot of a computer

Description automatically generated

A screen shot of a graph

Description automatically generated

* The model was tested on the test data set and the Loss function came out to be as RMSE = 4.685090A screenshot of a computer code

  Description automatically generated
* Once the perfect pair of parameters were found which minimized the loss function, the model was then allowed to predict the house price in the test.csv file. And the output was saved as predictions.csv

A screenshot of a computer program

Description automatically generated

* The Output thus received is not accurate as the loss function is too high, Therefore we can conclude that Linear Regression is not the appropriate algorithm as the dataset is too scattered for a straight Hyperplane to predict it.
* Other Regression Technique such as Random Forest Regression or Polynomial Regression might give a better result.