**UNIVERSITY OF ASIA PACIFIC**

Department of Computer Science and Engineering



**Course Title :**

**Artificial Intelligence and Expert Systems Lab**

**Course Code : CSE 404**

**Project : 02**

|  |  |
| --- | --- |
| **Submitted By:**  **Shawon Barman**  Section: A  ID: 18201043 | **Submitted to:**  **Dr. Nasima Begum**  Assistant Professor  Department of CSE, University Of Asia Pacific |

**Project Name: Implementation of Multivariable Linear Regression**

**Introduction:**

Multiple linear regression is simply the extension of simple linear regression, that predicts the value of a dependent variable (sometimes it is called as the outcome or target variable) on the basis of two or more independent variables (or sometimes, the predictor or explanatory variables). It is a method of statistical analysis that provides the statistical significance to explanatory variables, or which potential explanatory variables are crucial predictors for a given response (target) variable.

**Problem statement:**

Implement the Multivariable Linear Regression Using Open-Source Dataset with and without SK-Learn.

**Dataset:**

Dataset name : USA Housing

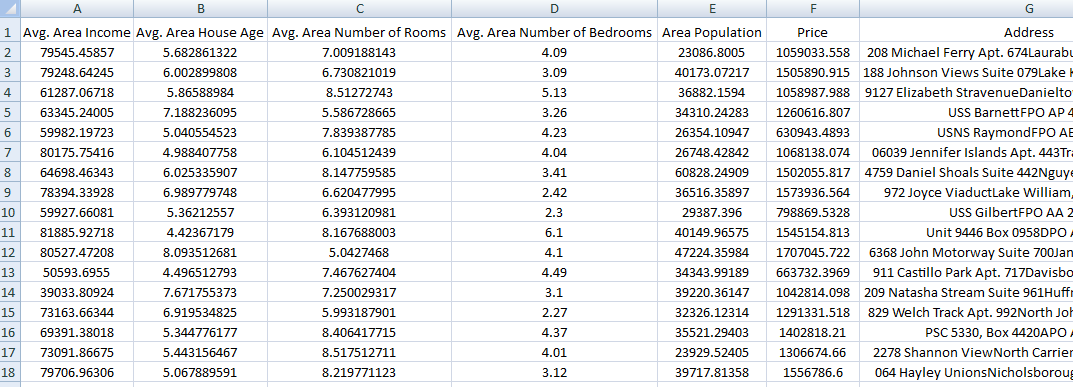
Dataset URL : <https://github.com/huzaifsayed/Linear-Regression-Model-for-House-Price-Prediction/blob/master/USA_Housing.csv>

The dataset has 7 columns:

* Avg. Area Income,
* Avg. Area House Age,
* Avg. Area Number of Rooms,
* Avg. Area Number of Bedrooms,
* Area Population,
* Price,
* Address

The variables are as follows Response:

* Variable (x1) : Avg. Area Income,
* Variable (x2) : Avg. Area House Age,
* Variable (x3) : Avg. Area Number of Rooms,
* Variable (x4) : Avg. Area Number of Bedrooms,
* Variable (x5) : Area Population,
* Variable (y) : Price



Here in the dataset the Price column is a dependent feature and the rest of the columns are independent features plotting the independent variable vs department variable to see how independent variables affect the independent variable.

**Tools And Languages:**

* **Programing Language:** Python
* **IDE**: Jupyter Notebook

**Parameter Initialization:**

For the weights of the features and bias we use 1,1.5, 3, 8, and 9 values.

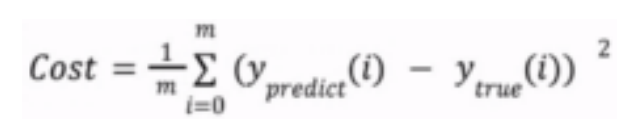
**Hypothesis Function:**

For Multiple Linear Regression the hypothesis function will be:

hypo.PNG

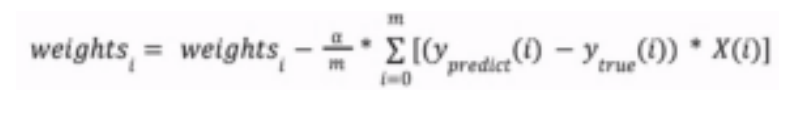
**Cost Function:**

For cost function, we are using Means Squared Error. The formula for the MSE



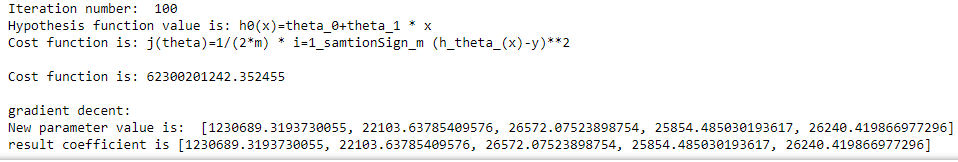
**Gradient Function:**

Updating the weights and bias value we are using Gradient Descent(GD)

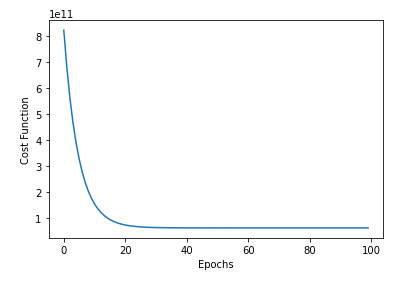


**Linear Regression Without SK-Learn: Output**

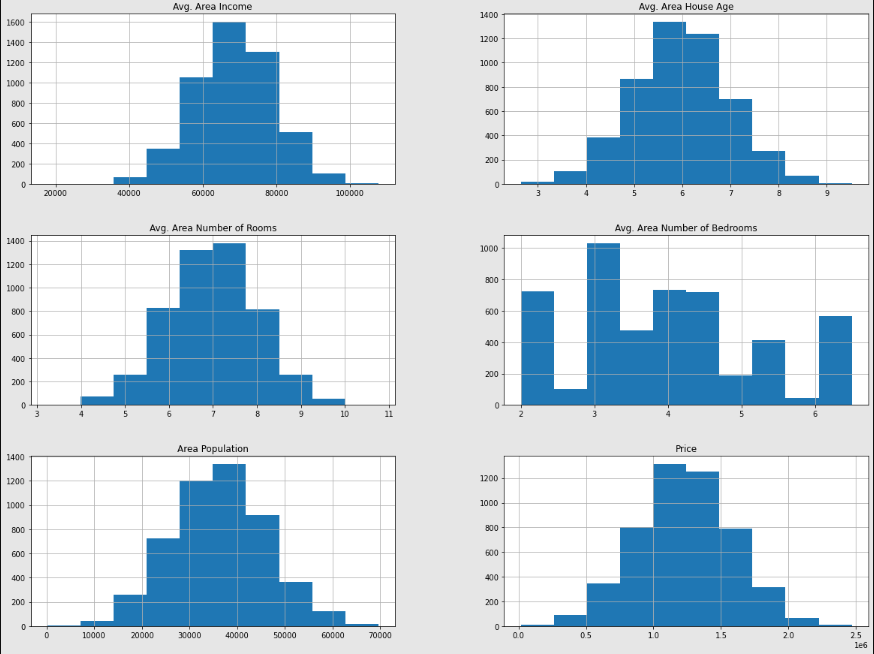
***Cost Function & Gradient Decent:***



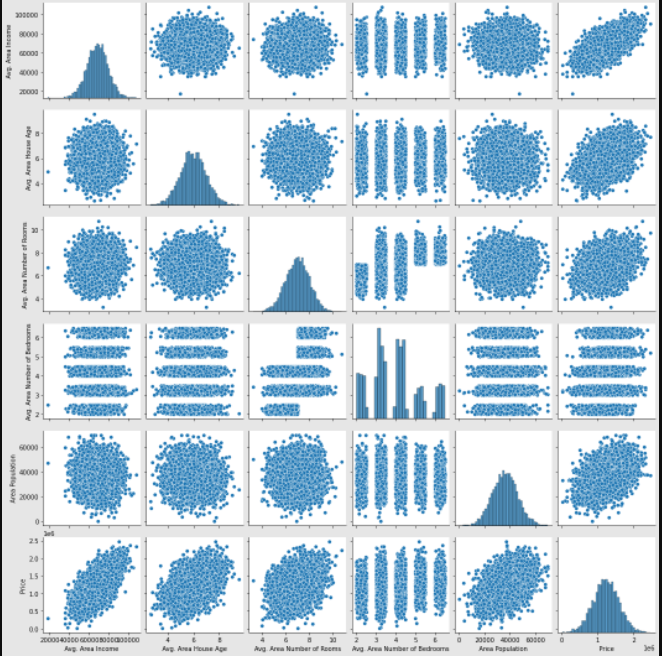
***Graph Plotting:***



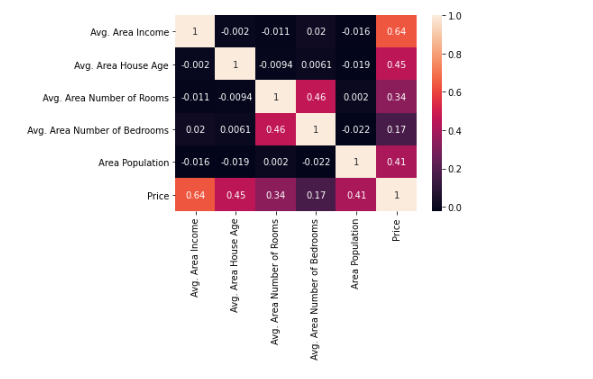
***Figure Plotting:***



***Pairwise Figure:***

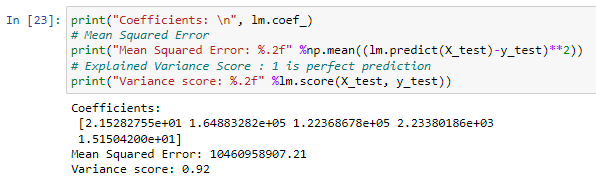


***HeatMap:***

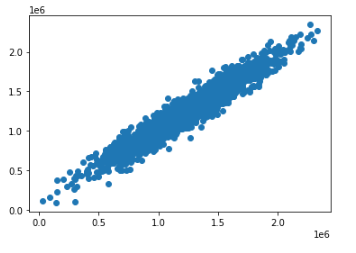


**Linear Regression With SK-Learn: Output:**

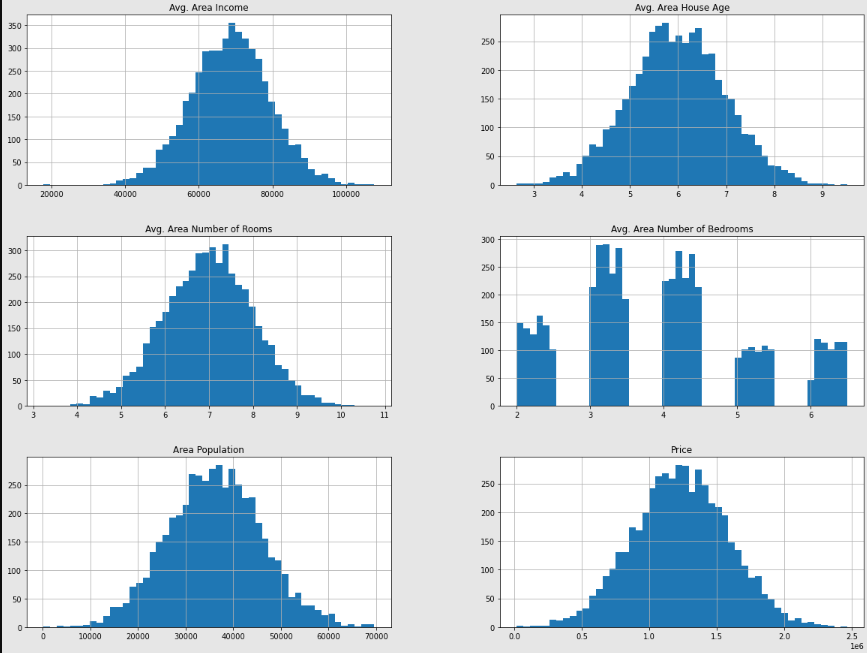
***Coefficients , Mean Error & Variance Score:***



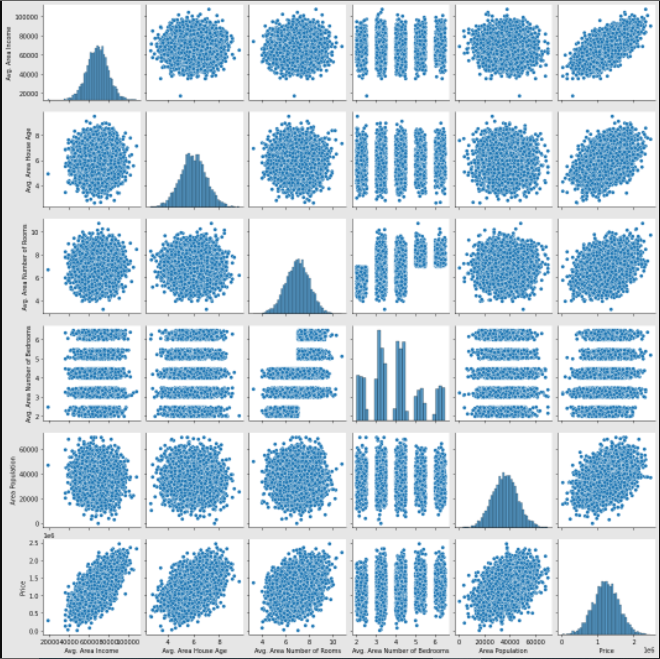
***Graph Plotting:***



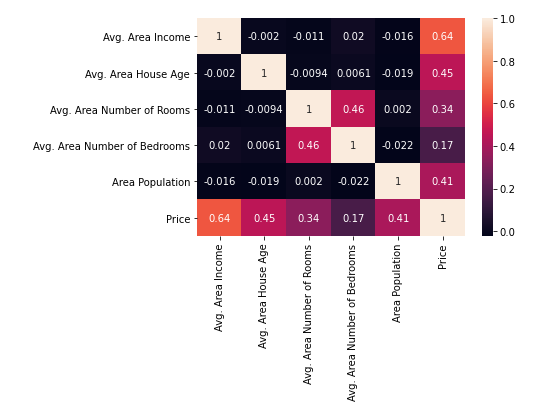
***Figure Plotting:***



***Pairwise Figure:***



***HeatMap:***



**Conclusion:**

Multiple regression is an extension of simple linear regression. In this project, after successful implementation, I’ve a brief knowledge about linear regression. Multiple linear regression models are useful in helping an enterprise to consider the impact of multiple independent predictors and variables on a dependent variable and can be beneficial for forecasting and predicting results. So, the outcome knowledge of this project will help me a lot to real-life problem predicting and solving on various field.