**Experiment 03**

**Aim:** Multiple Linear Regression in Python.

**Tools:** Jupyter Notebook, Housing Dataset (Kaggle)

**Theory:**

What is multiple linear regression? Multiple linear regression is a regression model that estimates the relationship between a quantitative dependent variable and two or more independent variables using a straight line.

About dataset used in this practical

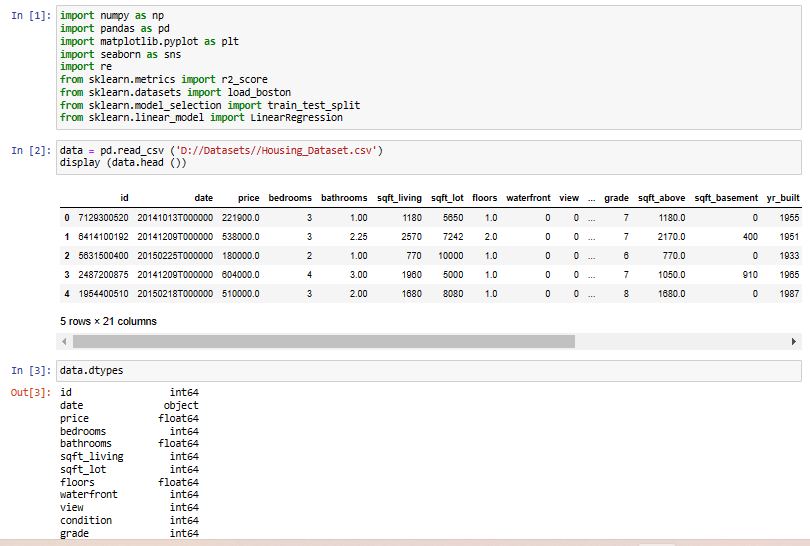
Online property companies offer valuations of houses using machine learning techniques. The aim of this report is to predict the house sales in King County, Washington State, USA using Multiple Linear Regression (MLR). The dataset consisted of historic data of houses sold between May 2014 to May 2015.  
We will predict the sales of houses in King County with an accuracy of at least 75-80% and understand which factors are responsible for higher property value - $650K and above.”

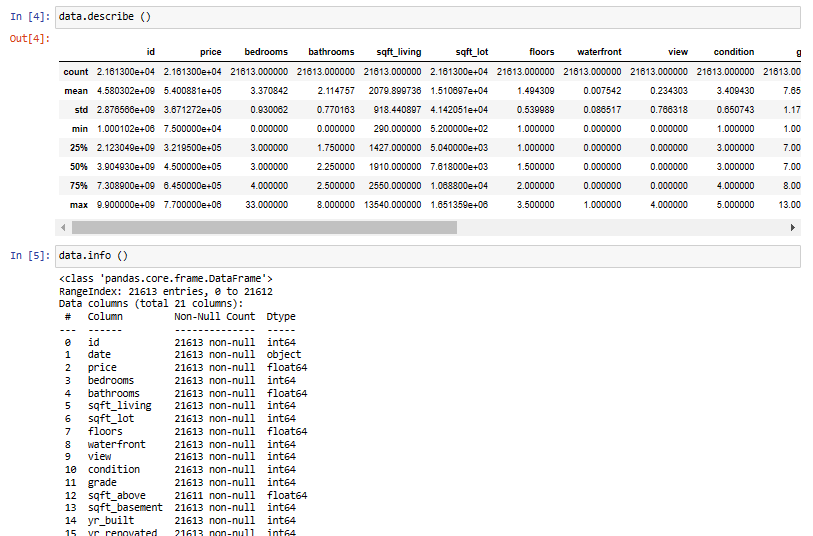
The dataset consists of house prices from King County an area in the US State of Washington, this data also covers Seattle. The dataset was obtained from Kaggle.  
The dataset consisted of 21 variables and 21613 observations.

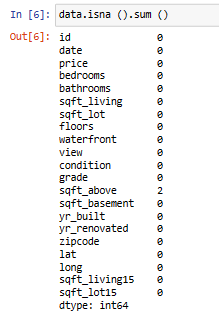
For creating a model we will be using SciKit Learn, we have seen this in previous practical as well.

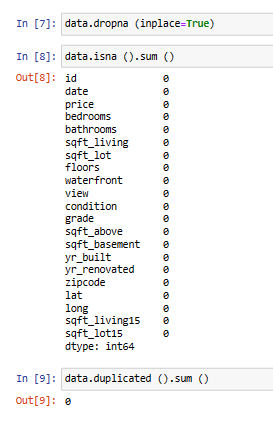
Linear Regression is a machine learning algorithm based on supervised learning. It performs a regression task. Regression models a target prediction value based on independent variables. It is mostly used for finding out the relationship between variables and forecasting. Different regression models differ based on – the kind of relationship between dependent and independent variables, they are considering and the number of independent variables being used.

**Code & Output:**

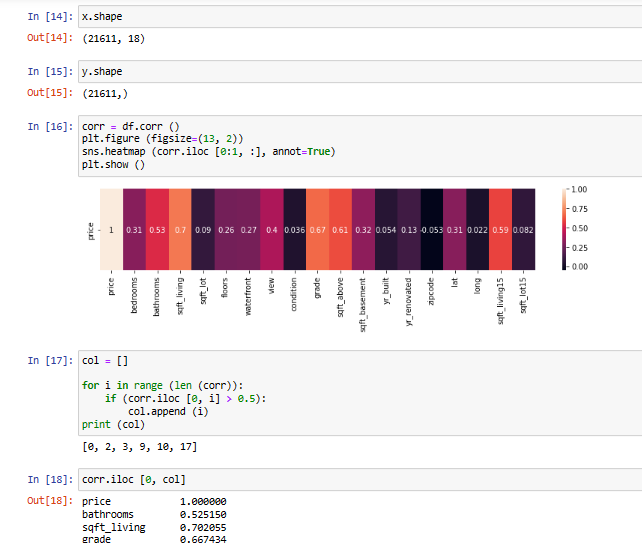
****

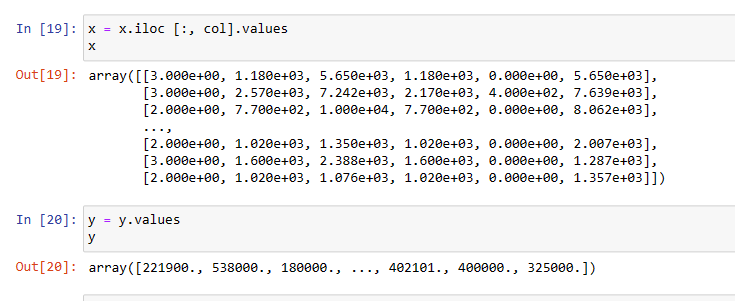
****

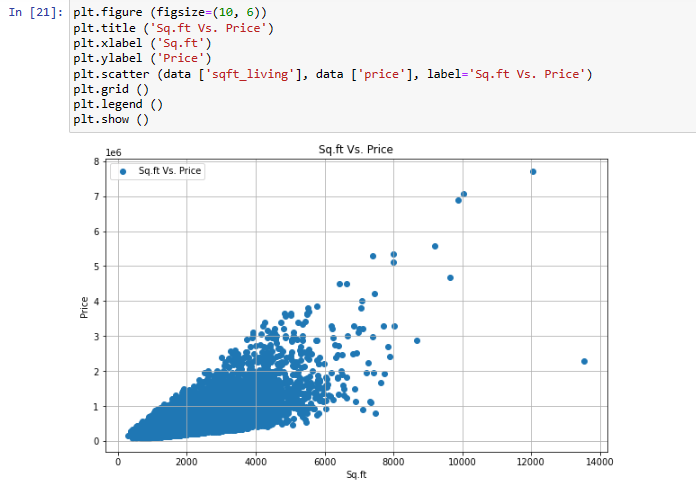
****

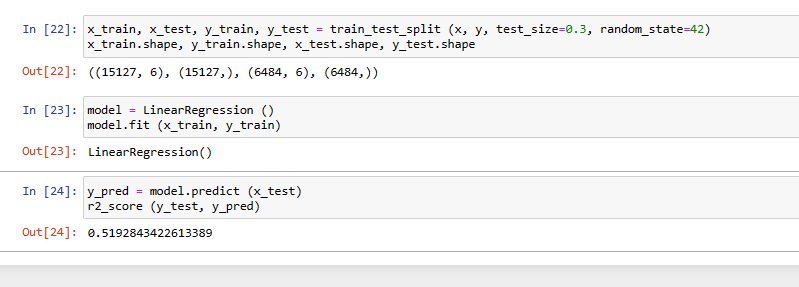
****

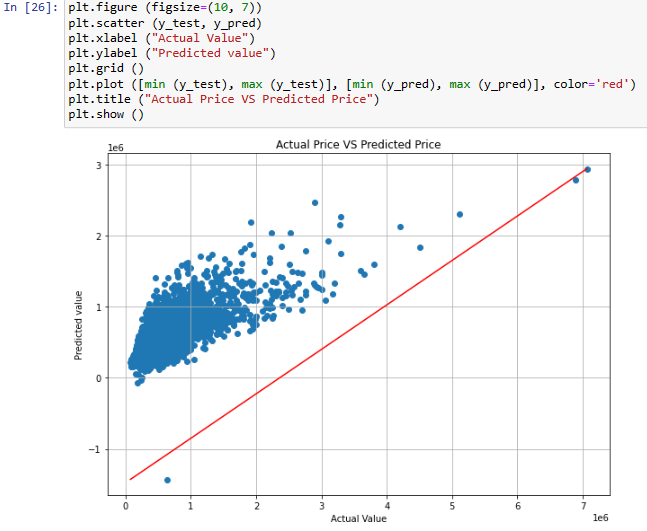
****

****

****

****

****

****

**Conclusion:** We have performed the given experiment successfully.