DATA ANALYTICS

ASSIGNMENT 3

Linear And Multiple Regression

TEAM MEMBERS

NAME SRN  
VIMARSHA R PES1201701590

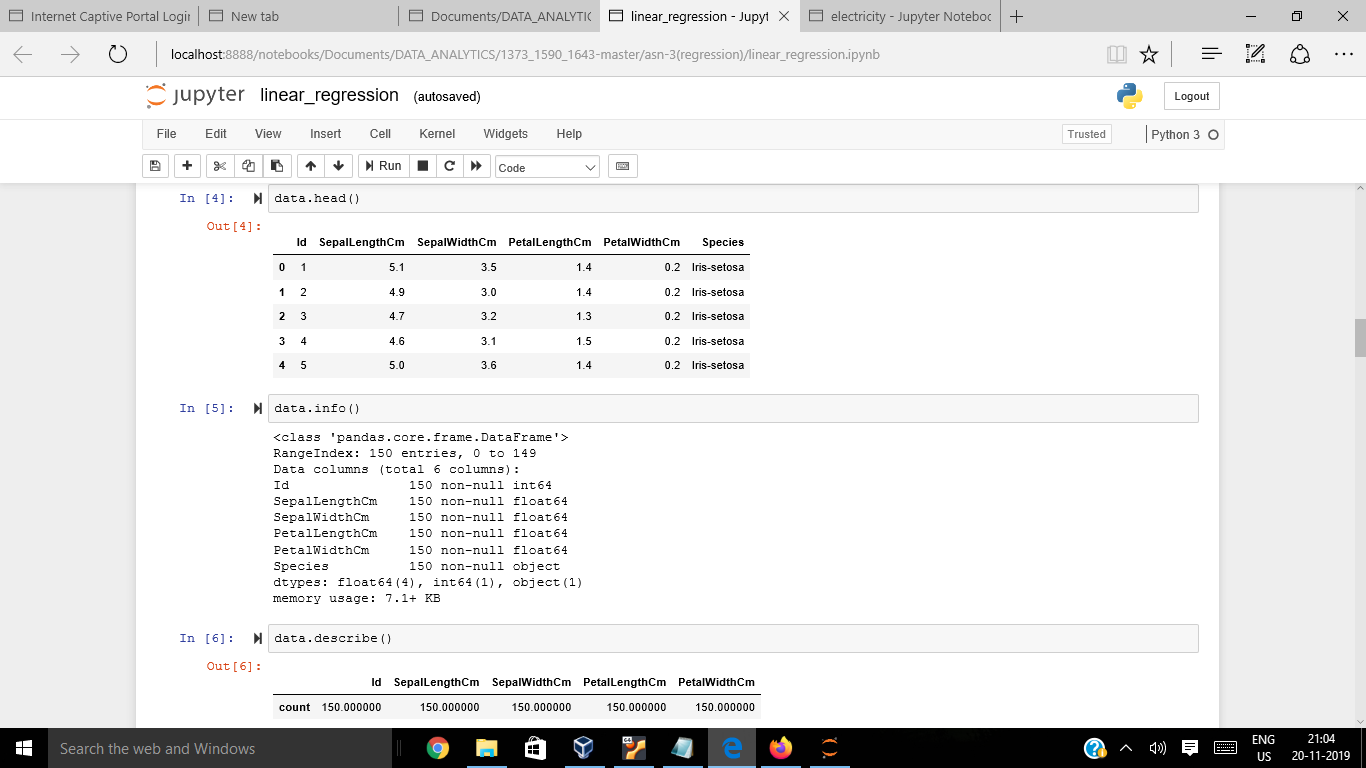
VINUTHA M N PES1201701373

KEERTHI PRIYA P PES1201701643

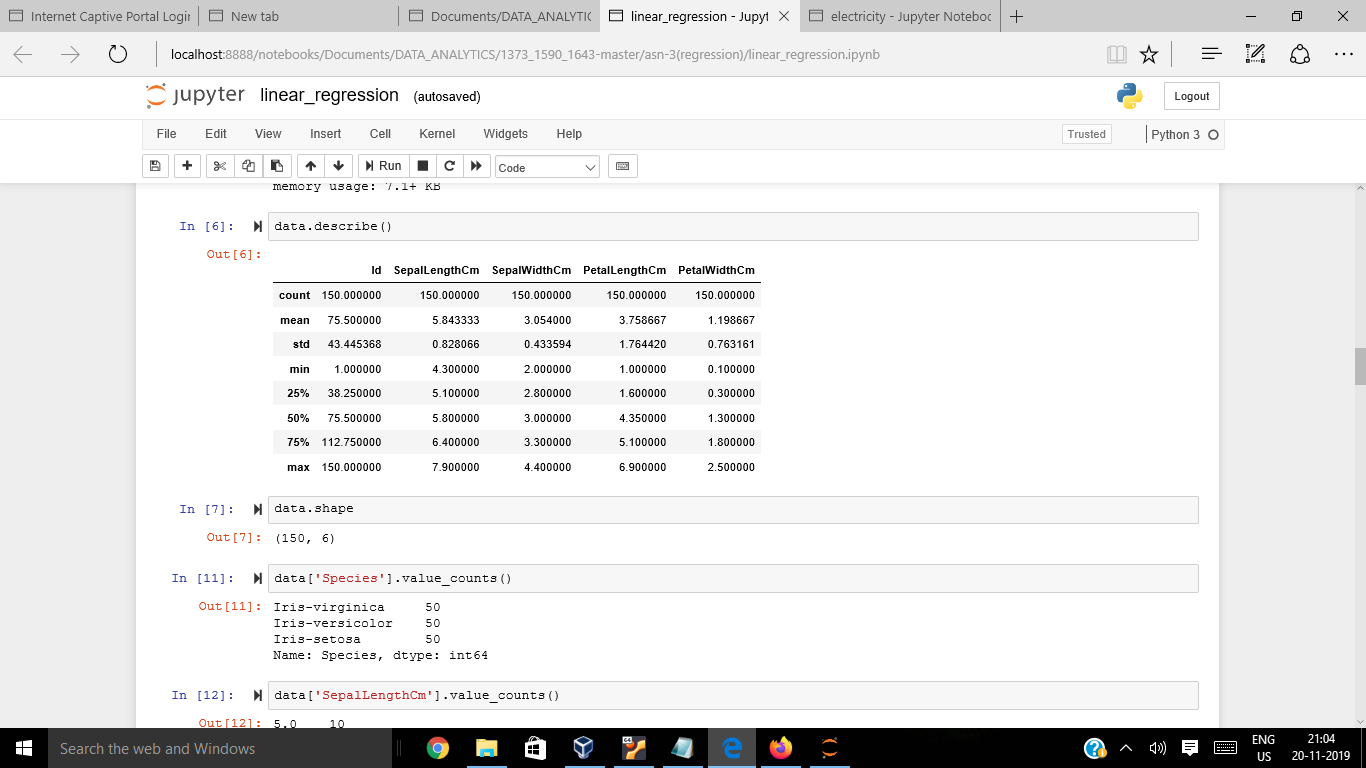
**Linear Regression:**

Linear regression is a basic and commonly used type of predictive analysis. The overall idea of regression is to examine two things: (1) does a set of ***predictor variables*** do a good job in predicting an outcome (dependent) variable? (2) Which variables in particular are significant predictors of the outcome variable, and in what way do they–indicated by the magnitude and sign of the beta estimates–impact the outcome variable? .

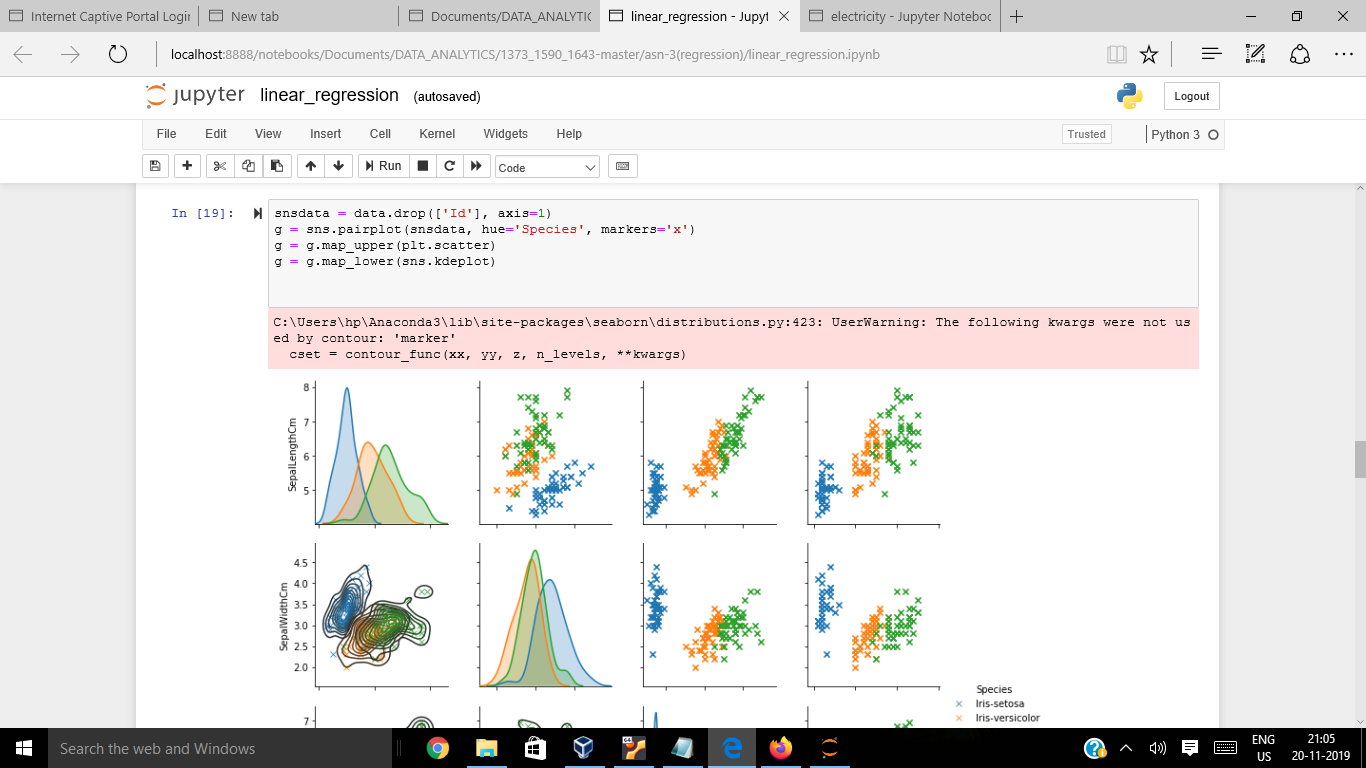
The simplest form of the regression equation with one dependent and one independent variable is defined by the formula ***y = c + b\*x***, where y = estimated dependent variable score, c = constant, b = regression coefficient, and x = score on the independent variable.

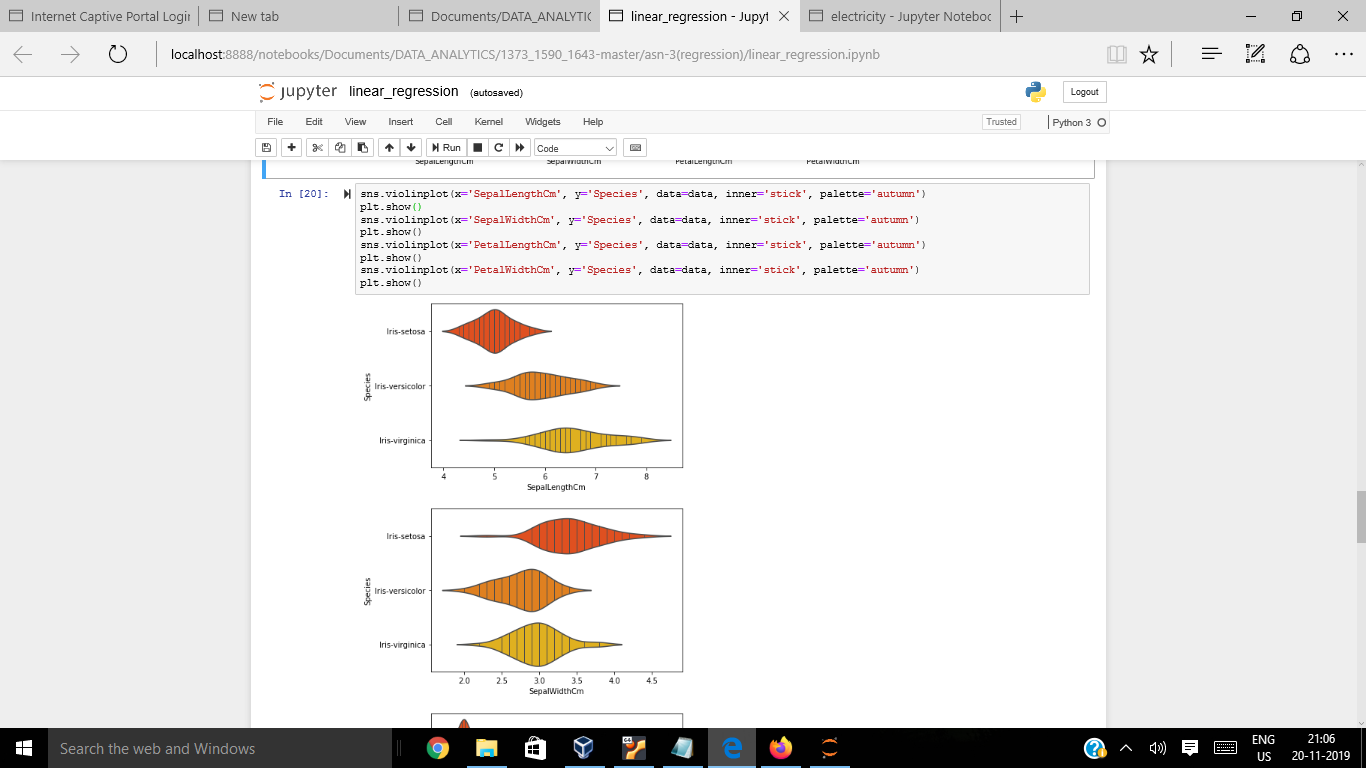


Firstly we are importing all the required modules and considering first 5 rows of the ***IRIS dataset***



In IRIS dataset we have three species of flowers, ***Verginica, Versicolor, Setosa.*** And for each species we have sepal\_length, ***sepal\_width, petal\_length, petal\_width***, for each species graphs are plotted considering one independent variable at a time.

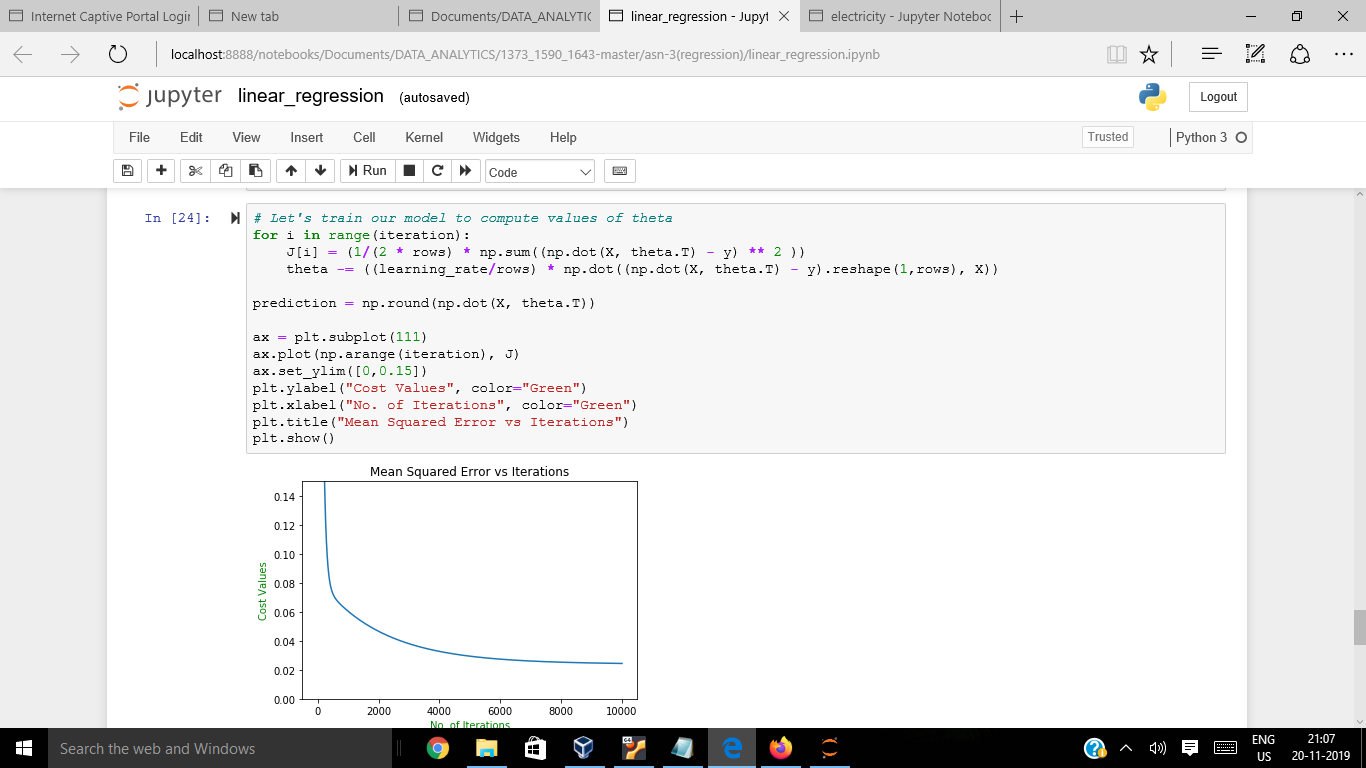




***Multilinear Regression:***

Multiple regression is the extension of ordinary least-squares (OLS) that involves more than one explanatory variable.

**The Formula for Multiple Linear Regression Is**

yi=β0+β1xi1+β2xi2+...+βpxip+ϵ

Finally we have to check the efficiency of the model by plotting a graph between original values and predicted values as shown in the graph below:

As per the below graph

***The Accuracy we are getting is 96%***

