Practice assignment

Linear Regression

All data files referenced in this assignment are in the Jigsaw lab at:

C\Data Science with R\Assignments\Non Graded Assignments\ Topic 9 Linear Regression

(Please note that you will need to save all your work in C drive in your user folder)

Objective
To predict the Housing prices in the suburbs of Boston as a function of the different attributes

Please refer to the data dictionary to get a better understanding of the variables.

Details
Location: **C/Data Science R/Assignments/Non Graded Assignments/Topic 9 Linear Regression**

Data Dictionary : *Boston Housing Prices Data Dictionary.csv*
Data set : *boston_prices.csv*

Following are some steps to be followed

- **Data Preparation :**Understand summaries of the data, clean the data- missing value treatment, outlier treatment, create new variables, dummy variables ,etc.

- **Data Profiling:** Understand the relationship between every IDV and DV (Using Correlations, visualisations, etc.)

- **Build a Linear Regression model using IDV's that have good correlation with the DV.** Your goal will be to create a model that best explains or predicts price

- **Evaluate the model performance and validate the model.** Generate residual plots and actual Vs. predicted plots in Excel