**Computer Sales**

1. Loading the Data set
2. Removed unwanted Sequence number
3. Created Dummy variables for CD, Multi, Premium and deleting the categorical column

4) No NaN entry in the dataset

5) Used Box plot,Histogram, QQ plot to understand the features and how other features influence the price

6) Correlation Plot to understand multiple variable

7) Correlation Coefficient to find the strength and direction

8) library(corpcor) to Estimation of Co variance

Model 1 : All Variables

Result : RSq value is 0.7756

Influence index plot used to get the influencing data

Model 2 : without entry 1441,1701, 3784, 4478

Result : RSq Value is 0.7774

Model 3 : Square Root Model

Result : RSq Value is 0.7872

Model 4 : Log Model

Result : RSq Value is 0.5081

Model 5 : Two Degree model

Result : RSq Value is 0.6849

Square Root Model is best comparing with other model , but still R Sq value is less than .85.

9) library(car) used to identify the influencing plot and VIF

10) Added Variable plot to check correlation between variables and output variable

11VIF shows that the model is not col-linearity in variables

Variance Inflation Factors ( VIF) greater than 10 means col-linearity in the variables. All the values are less than 10 in this dataset

1. Model Building

80% Train Data

20% Test Data

RSq Value: .7869

RMSE is 266.4478