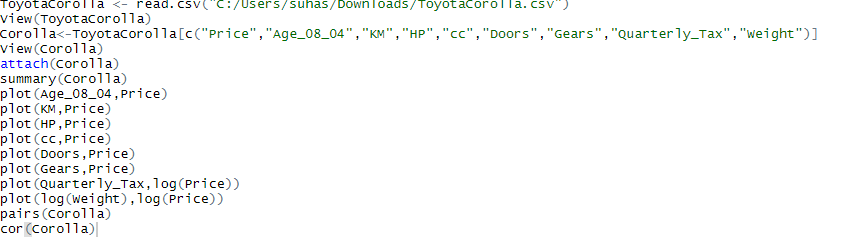
**Assignment on Multilinear Regression**

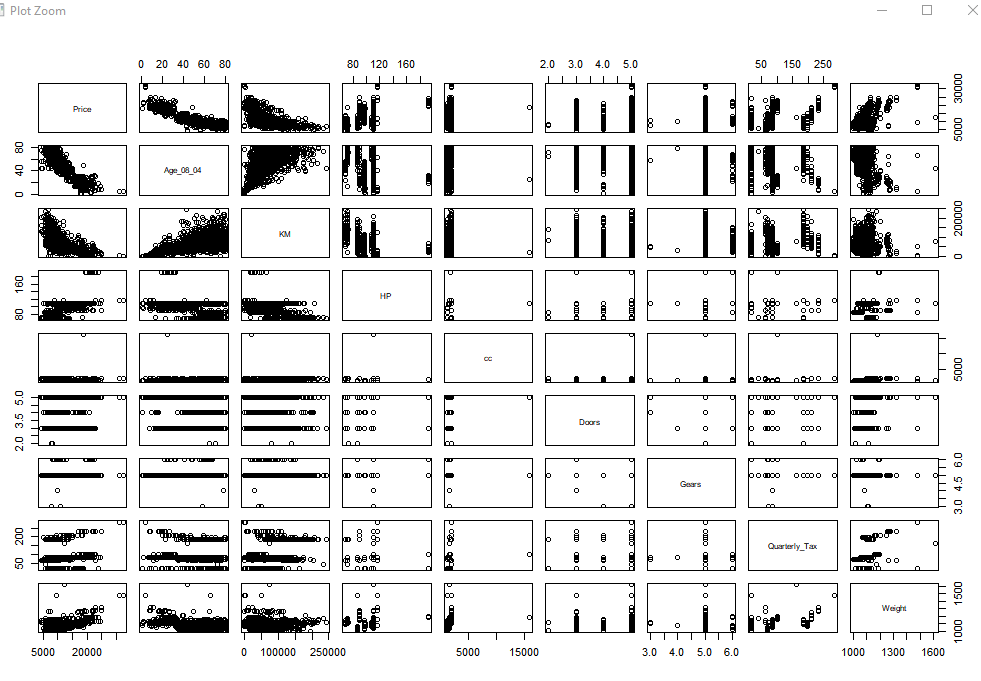
Q1: Consider only the below columns and prepare a prediction model for predicting Price.

Corolla<-Corolla[c("Price","Age\_08\_04","KM","HP","cc","Doors","Gears","Quarterly\_Tax","Weight")]

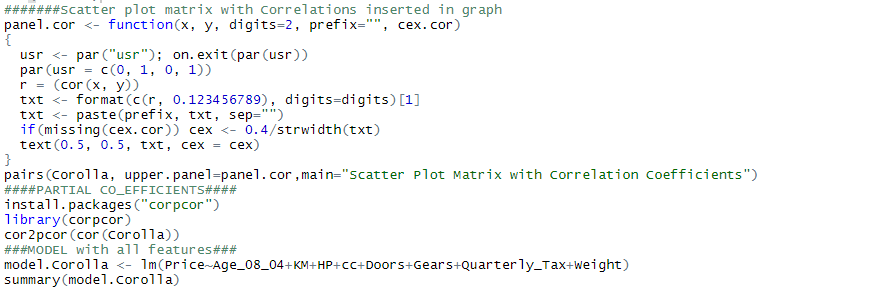
Ans: R Code :

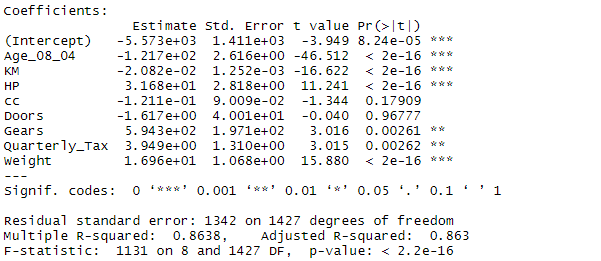
* Importing the given data set and then assigning it to an object. Performing scatterplot analysis to understand the correlation between different variables.





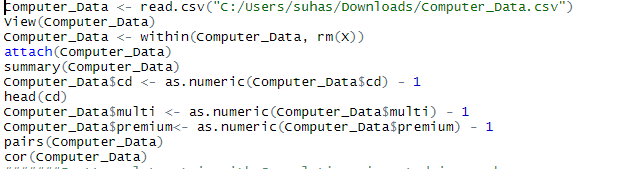
* From the above Pair Plot we can clearly identify the relationship between different features and also understand if there is a Multicollinearity problem
* As Part of the scatter diagram there seems to be a multicollinearity Problem between “**AGE\_08\_04”** and “**KM”** which are co-related
* Although there is a corerelation between Age and Kms it seems to be a Morderate coorelation of **0.505672180**.



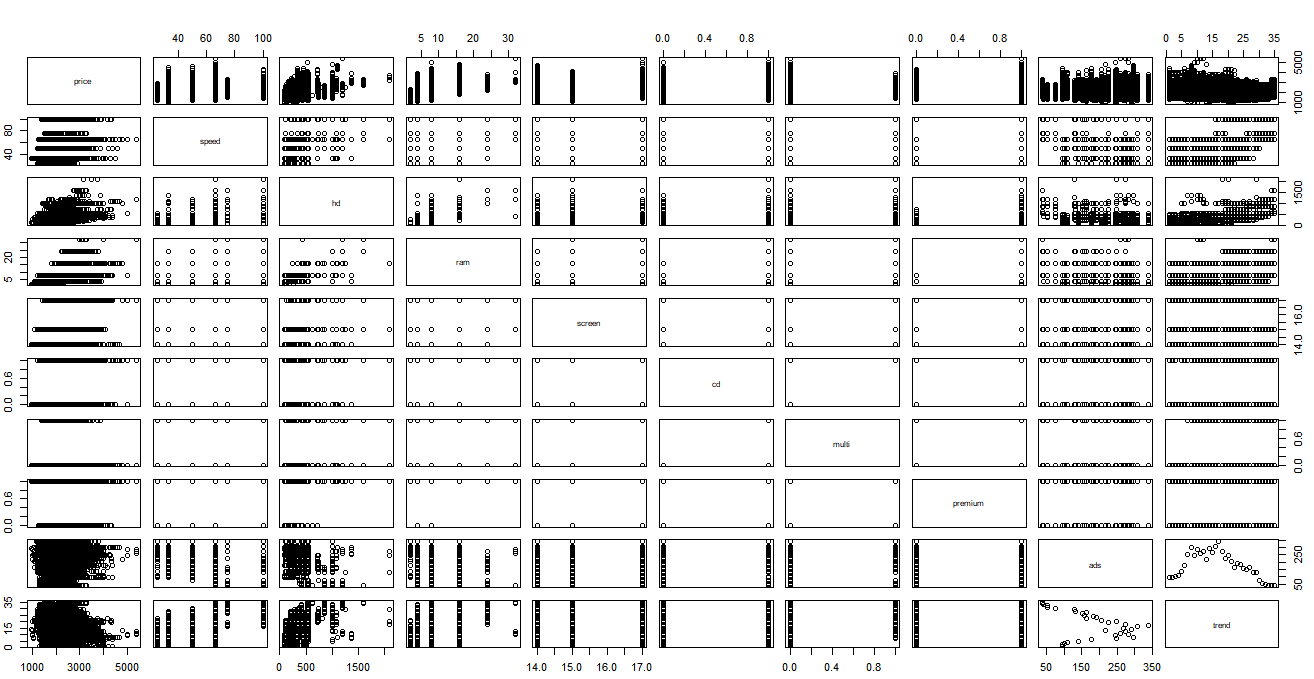
**Output:** 

**Question 2**) Predict Price of the computer A data frame containing:

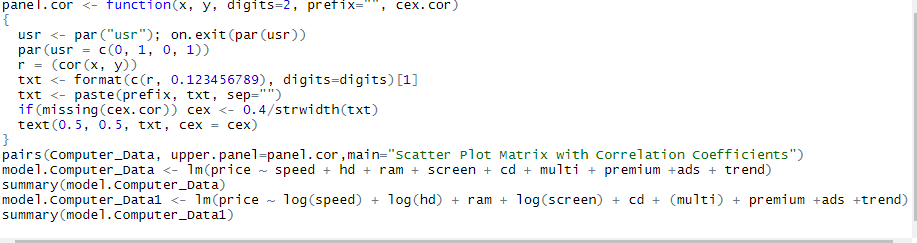
price : price in US dollars of 486 PCs, speed : clock speed in MHz, hd : size of hard drive in MB ram : size of Ram in in MB screen : size of screen in inches cd : is a CD-ROM present ? multi : is a multimedia kit (speakers, sound card) included ? premium : is the manufacturer was a "premium" firm (IBM, COMPAQ) ? ads : number of 486 price listings for each month trend : time trend indicating month starting from January of 1993 to November of 1995.

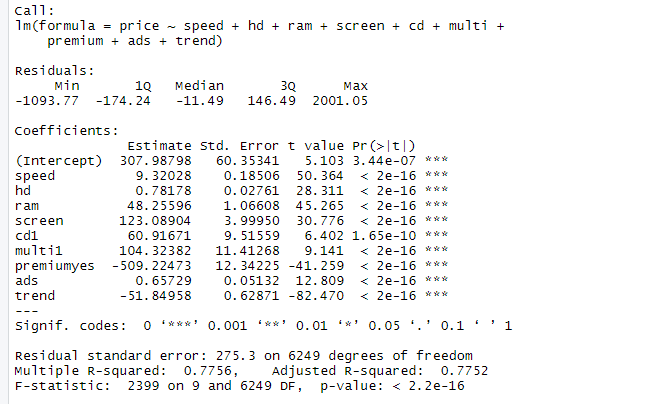
Ans : R code : 

* By running the above code we can see that the correlation coefficient as well as direction using the scatter plots.

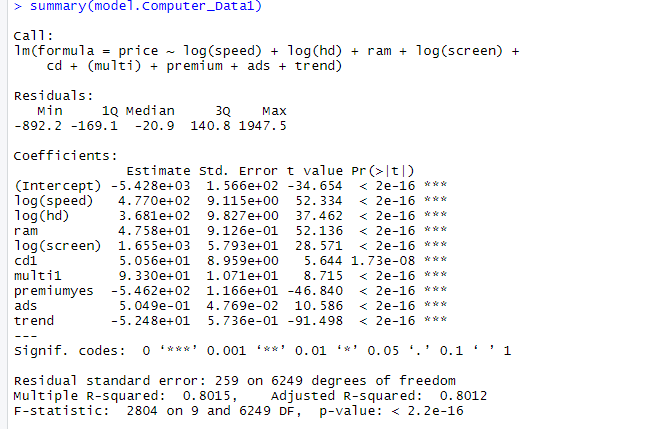


* Then we can build the model using the below code and we can see that the ouptput of the model that is the Rsquared values slightly increases when we take log values



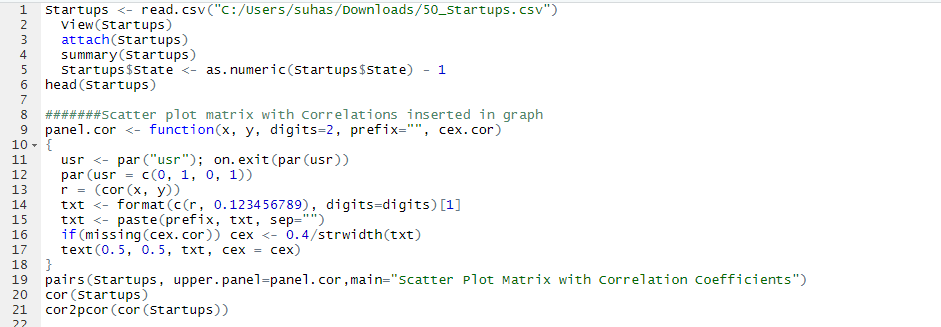
**Output**  without taking Log values

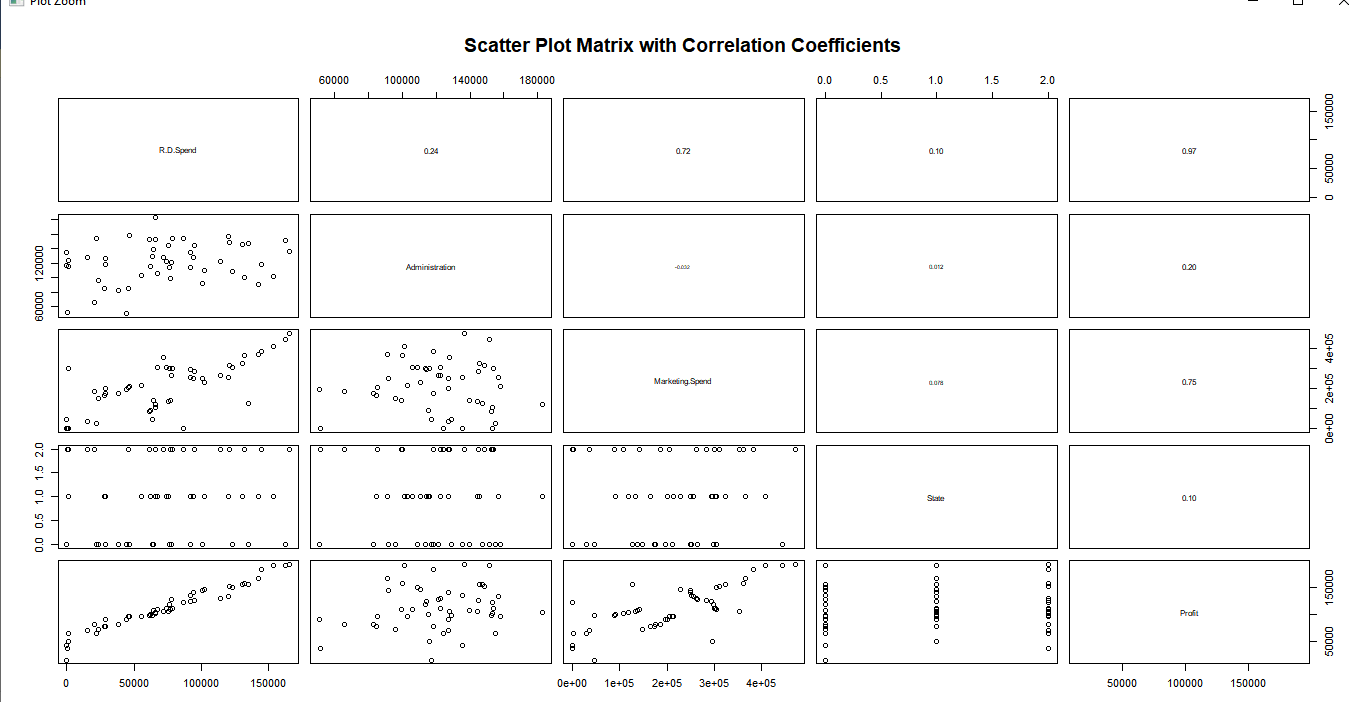
**Output taking Log on some of the input variables**

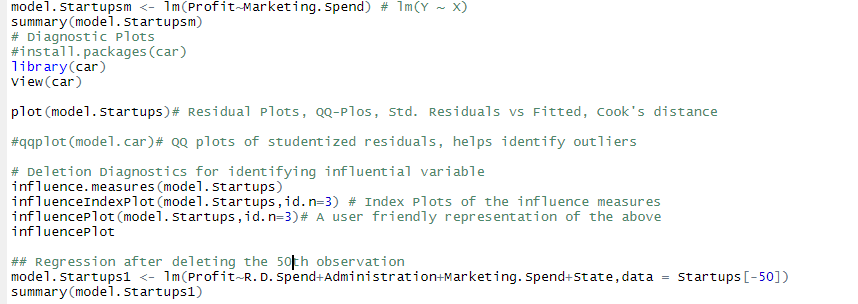


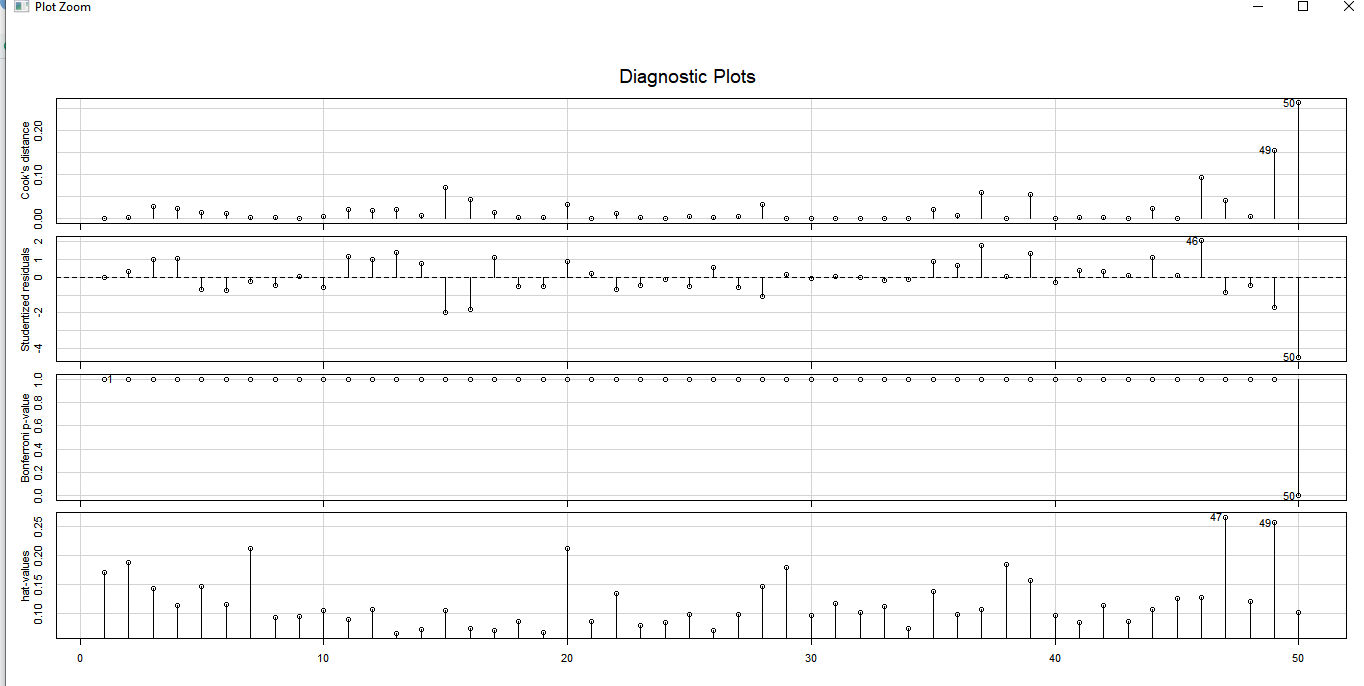
**Q3 )** Prepare a prediction model for profit of 50\_startups data. Do transformations for getting better predictions of profit and make a table containing R^2 value for each prepared model.

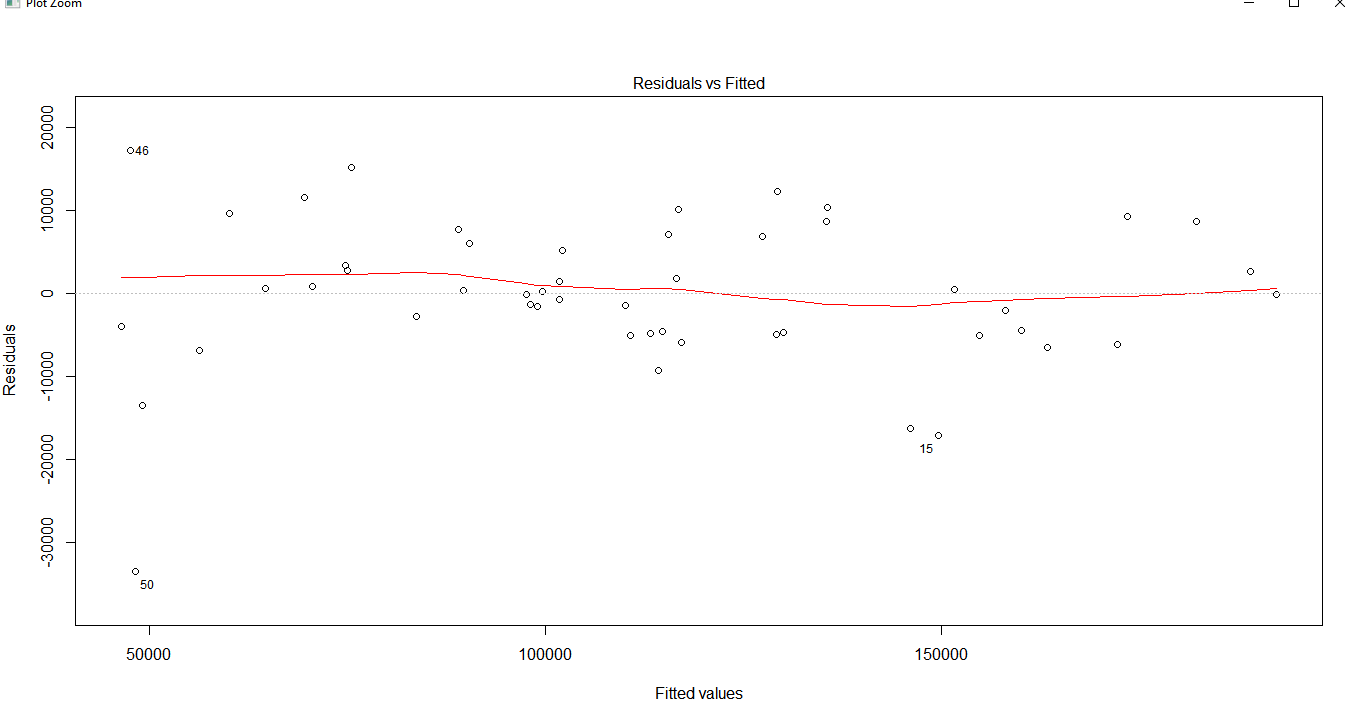
Ans: The discrete variable , i.e States is changed into dummy variables .0=California,1=Florida,2=New York.We can see the scatter plots and correlation values below











**Outputs**:

