University of Moratuwa
MBA in Information Technology
Department of Computer Science & Engineering

Cover Sheet for Assignment

Name with Index Numbers: K.M.Nethmini Dulanjalee 189103N

Title of Assignment: TEAM D Assignment 01–Report
Assignment No: 01 **Group** ☐ **Individual** ✓

Subject Code: CS5122
Subject: - Descriptive and Predictive Analytics
Lecturer: Dr. Uthayasanker Thayasivam

Student's Statement:

We certify that we have not plagiarized the work of others or participated in unauthorized collusion when preparing this assignment.

Date: 16/03/2019

Office use only:

On/ before deadline Submission ☐ Extension Given ☐ Late ☐

Signature:……

Marks Given:

In this survey results of "**Home Market Value**", there are 42 number of observations.

Based on the data we can raise following questions

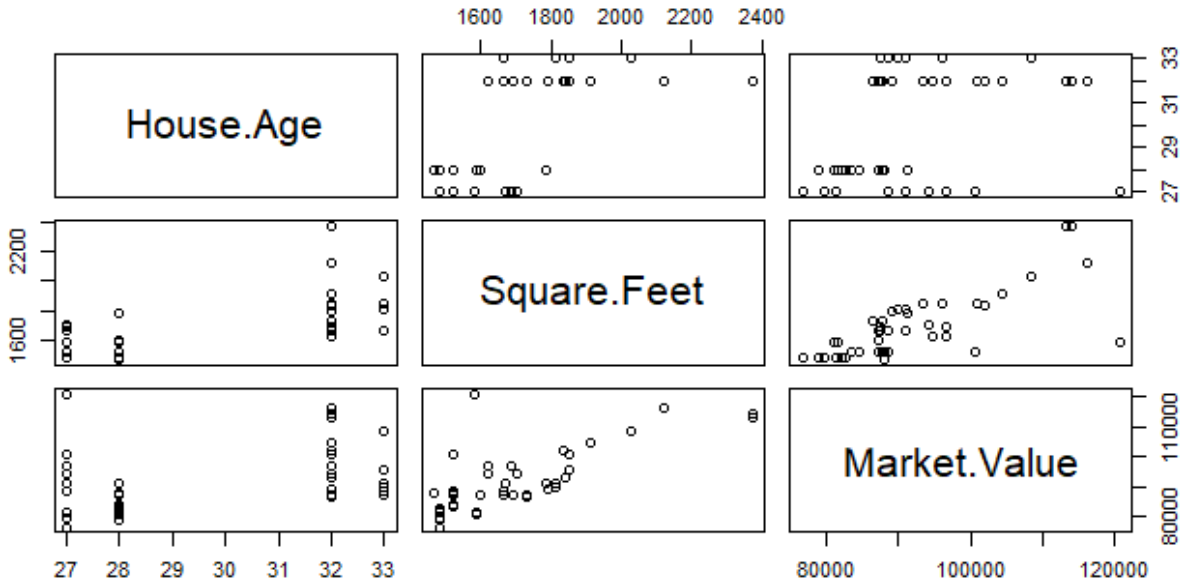
I. What is the relationship between the square feet of the house and the market value?
II. What is the relationship between the square feet of the house and the house age?
III. Does age of house has any impact on the market value?
IV. Does square feet of house has any impact on its market value?

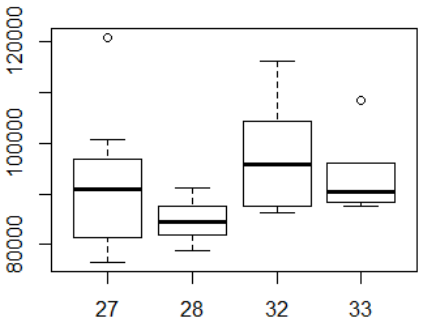
By analyzing statistical properties of data, we can come to following conclusions

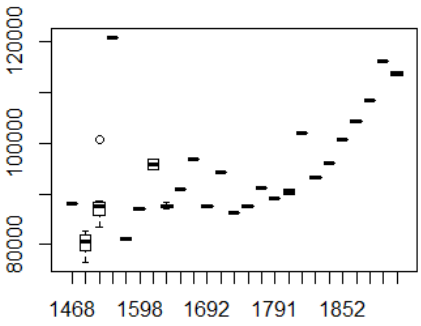
- *Executive summary of the data frame,*

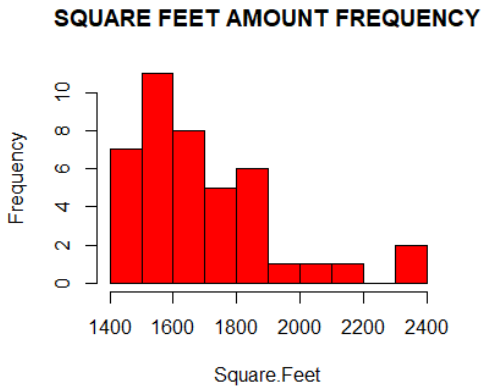
```
> summary(hmv)
   House.Age      Square.Feet    Market.Value
 Min.   :27.00   Min.   :1468   Min.   : 76600
 1st Qu.:28.00   1st Qu.:1520   1st Qu.: 86575
 Median :28.00   Median :1666   Median : 88500
 Mean   :29.83   Mean   :1695   Mean   : 92069
 3rd Qu.:32.00   3rd Qu.:1807   3rd Qu.: 96525
 Max.   :33.00   Max.   :2372   Max.   :120700
```

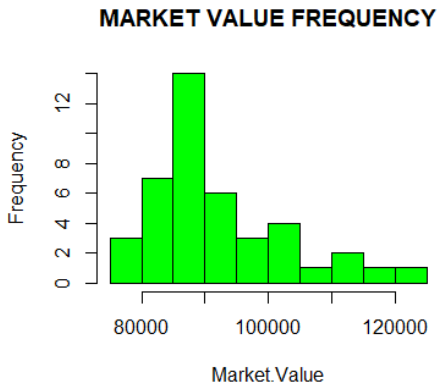
Below graph depicts the clustering of the given data set in to clusters and this represents a clear correlation between the house age, market value and square feet.

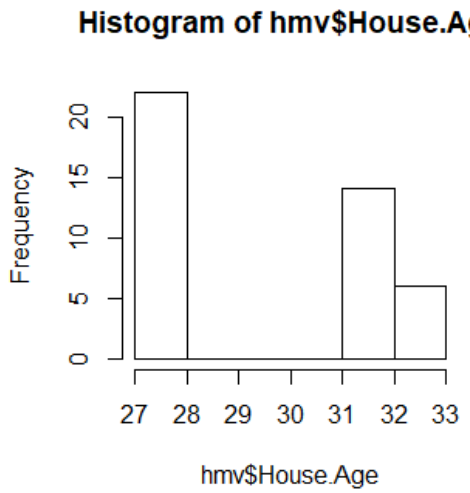


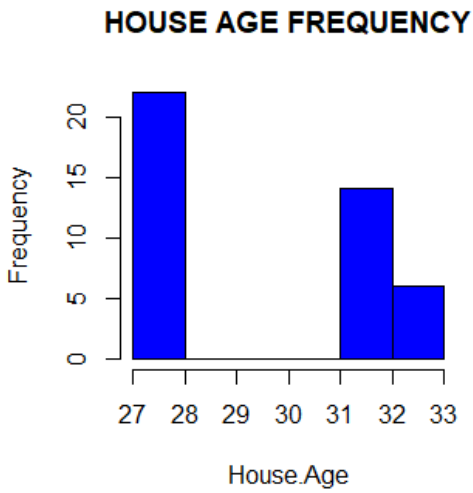






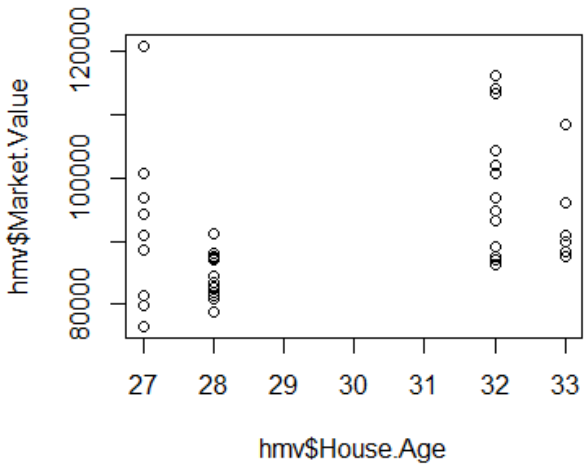


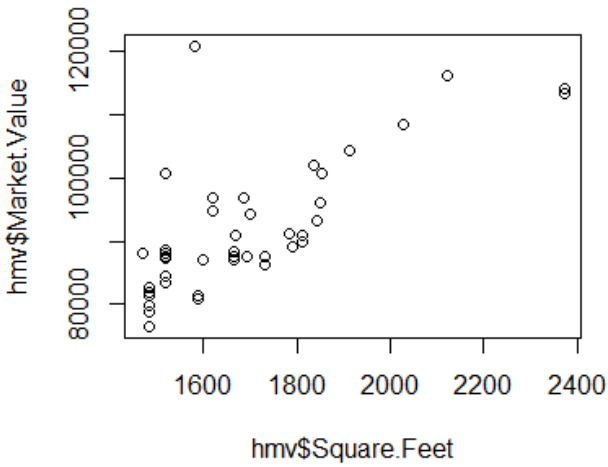




Suitable regression analytics technique

Linear Regression is one of the mostly known regression model used for predicting a quantitative response .Linear Regression creates a relationship between dependent variable (Y - Market Price) and independent variables (X - House age & Square feet) using a best fit straight line. The box plot illustrate there might be chance to perform regression lines by reading the pattern of points using naked eye.





Predicted Market Value of the following houses

Age	Square Feet
26	1650
28	1500
29	1800
30	2200
31	2400

```
> hmv.pred
        fit       lwr       upr
1  93380.45  88492.92  98267.99
2  85593.47  82520.73  88666.21
3  97041.63  93878.98 100204.28
4 112580.90 105506.93 119654.86
5 119937.95 110961.62 128914.27
>
```