Graded Project: Terro's Real Estate Agency

1) Using Descriptive Statistics we find the Mean (average), Standard error, Median, Mode, Standard Deviation, Sample variance, Kurtosis, Skewness, Range, Minimum, Maximum, Sum and Count of the given variable.

2) In this graph we came to know that the most number (133) of the flats are priced between $21k and $25k range and least number (6) of flats are priced between $37k and $41k and between $45k and $49k.

3) Using Covariance table we found that the relationship between each and every possible variable whether it is positive relationship or negative relationship.

4) Correlation used to find statistical relationship between two variables.
   - Highly positive correlated pairs.
     - Tax and Distance (0.910228).
     - Nox and Indus (0.763651).
     - Nox and Age (0.73147).
   - Highly Negative Correlated Pairs
     - Average price and Lstat (-0.73766).
     - Lstat and Average room (-0.61381).
     - Average Price and PTratio (-0.50779).

5) a) In this regression summary the value Rsquare is 0.544 so the model is good fit for Regression Model, if 1 unit increases in Lstat value the average price will be decreases 0.95 times, intercept value denotes that if value of x is 0 then the value of y is 34.55, all the plots are close to 0 and there is no pattern or trendline available in this plot so it is good fit model.

b) Yes, LSTAT variable significant for the analysis based on this because the Rsquare is greater than 0.5

6) a) The average price for the new house in this locality with 7 rooms is 21.458. The company is overcharging for the house.

b) Regression model including LSTAT and AVG_ROOM together as Independent variables and AVG_PRICE as dependent variable is better model when it is compared with AVG_PRICE as Dependent variable and LSTAT variable as Independent Variable because the Lstat and Avg_room with Avg_price is much close to 1.

7) If crime_rate increases by 1 unit the average price will be increased by 0.048 times, If age increase by 1 unit the average price will be increased by 0.032 times, if indus increased by 1 units the average price will be increased by 0.13, if nox is increased by 1 unit the average price will be decreased by 10.32 times, if distance increased by 1 unit the average is increased by

0.26 times, if tax increased by 1 unit the average price is decreased 0.014 times, if ptratio is increased by 1 unit the average is decreased by 1.07, if the average room is increased by 1 unit the average price is increased by 4.125 times, if the lstat is increased by 1 unit the average price decreased by 0.603.

8) a) R square value is closer 1 (0.7) so it is the best fit model.

b) Regression model excluding crime_rate as Independent variables and AVG_PRICE as dependent variable is better model when it is compared with AVG_PRICE as Dependent variable and all other variable as Independent Variable because the excluding crime_rate as independent variable with Avg_price is much close to 1.

c) When the Nitrous Oxide increases by 1 unit the average price will be decreased by 10.27 times

d) Regression equation for this model is

Predicted average price= 0.03*age1+0.130*indus1-10.2727*nox1+0.26*distance1-0.014*tax1-1.0717*ptratio+4.1254*avg_room-0.605*Lstat+29.4284.