SUMMARY OUTPUT

	Regression Statistics
Multiple R	0.7991
R Square	0.6386
Adjusted R Square	0.6371
Standard Error	5540.2574
Observations	506

ANOVA

	df	SS	MS	F	Significance F
Regression	2	27276986214	1.3638E+10	444.3309	7.0085E-112
Residual	503	15439309201	30694451.7		
Total	505	42716295415			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	ower 95.0%	Ipper 95.0%
Intercept	-1358.27	3172.83	-0.43	0.67	-7591.90	4875.35	-7591.90	4875.35
AVG_ROOM	5094.79	444.47	11.46	3.47E-27	4221.55	5968.03	4221.55	5968.03
LSTAT	-642.36	43.73	-14.69	6.67E-41	-728.28	-556.44	-728.28	-556.44

RESIDUAL OUTPUT

Observation	Predicted AVG_PRICE	Residuals
1	28941.01368	-4941.013681
2	25484.20566	-3884.205661
3	32659.07477	2040.925231
4	32406.52	993.4800002
5	31630.40699	4569.593009
6	28054.52701	645.472994
7	21287.07846	1612.921545
8	17785.59653	9314.403473
9	8104.693384	8395.306616
10	18246.50673	653.4932695
11	17994.96223	-2994.962229
12	20732.21309	-1832.213091
13	18553.4842	3146.515803
14	23644.74107	-3244.741066
15	23108.95823	-4908.958231
16	22923.9452	-3023.945198
17	24652.57604	-1552.576036
18	19736.11045	-2236.110451
19	18929.7215	1270.278497
20	20573.77596	-2373.775964
21	13517.32408	82.67592493
22	20148.32175	-548.3217521
23	17908.96697	-2708.966971
24	15487.64606	-987.6460563
25	18352.81036	-2752.810359
26	16562.10901	-2662.109014
27	18744.40281	-2144.402811
28	18349.95811	-3549.958114
29	23510.18847	-5110.188468

AVG_ROOM Residual Plot

LSTAT Residual Plot
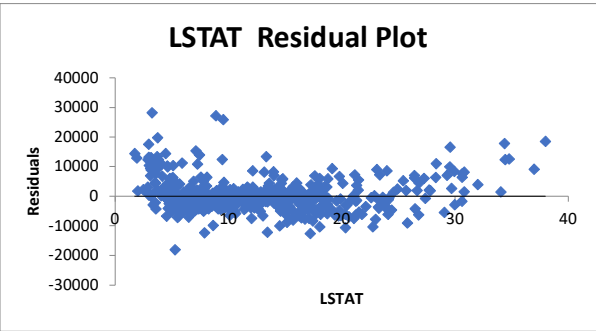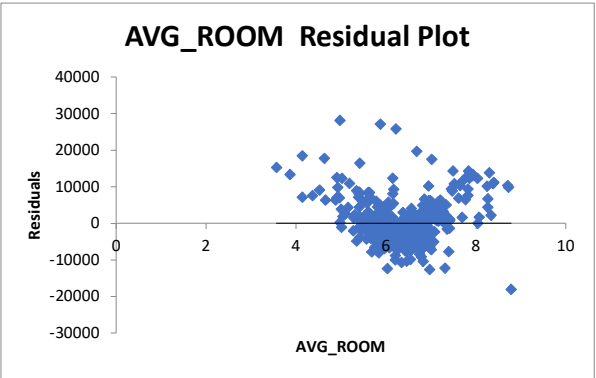
6(a). The regression equaition for the model is

$$Y = 5094.76X_1 - 642.36X_2 - 1358.27$$

where X_1 is the value of AVG_ROOM and X_2 is the value of LSTAT
If a new house in this locality has 7 rooms (on an average) and has a value of 20 for L-STAT,

AVG_PRICE = 5094.79(7) - 642.36(20) - 1358.27
= 35,663.53 - 12,847.2 - 1358.27
= **$ 21,458.06**

As per the comparison, the company is charging $21400 to the house with an average of 7 rooms with LSTAT of 12. The lower lower LSTAT value should increase the price of the house. The company is slightly over-charging.

6(b). Yes, the performance of this model is better than the previous model. Since we have included two significant variables in this model, this model is able to explain 63% variation of the Price where as the previous model was able to explain only 54% of the variation.