**REPORT BY**

**NAME:KUSHAL SHAH**

**STUDENT ID:A20207420**

**EMAIL ID:kusshah@okstate.edu**

**Task 1)**

**B)** H0: The nightly price for the listing does not depend on combination of: 1)Number of beds guest can use

2)The number of bedroom available to guests

3)Number of bathroom available for guests

4)Number of guests that can be accommodates on

each stay factors.

Ha: H0 is not true.

(Technical Hypothesis:H0: MR model does not explain relationship between the Bathrooms,Bedrooms, Beds, accommodate

(independent) and Price(dependent) variable any better that base line model.

Ha:: H0 is not true.

H0: Values of regression coefficient for each of Bathroom, Bedroom, Beds and accommodates equals

to 0 in population.

Ha: H0 is not true)

**C)** MR test shows that price for the listing does depend upon combination of 4 variables listed in hypothesis(see Figure 1)

**D)** Assumptions: (Met or did not meet)

* 1)Errors are uncorrelated (Met, see Figure 2 in appendix )
* 2)Homoscedasticity (Met, see Figure 3 for reference)
* 3)Errors are is normally distributed(did not meet, See Figure 4 for reference)
* 4) X variables are linearly related to Y variable (Met, see Figure 5 for reference.)
* 5)variables are linearly independent: Multicolinearity ( Met, See Figure 6 for reference)

**Task 2:**

**Conclusion:**

* Price for the listing does depends upon number of beds, bedrooms bathrooms and Accommodates. (See Figure 8 for reference)
* Out of these four, Price is most sensitive to number of accommodates allowed on each stay(see Figure 7 for reference).
* Price would go up by $30.84 if we increase number of allowed accommodates by 1 and keeping other factors(number of beds, number of bedrooms and number of bathrooms) constant. (see Figure 9 for reference.)
* Number of Beds is the second most sensitive factor to Price. (see Figure 7 for reference).
* Price would go down by $10.85 if we include one more bed in our stay and keeping other factors(number of accommodates, number of bedrooms and number of bathrooms) constant. (see Figure 9 for reference.)
* Number of bedrooms and number of bathrooms does not impact Price significantly. (See Figure 8 for reference)
* Increasing Number of accommodates to stay would increase price where increasing number of beds would decrease the price. (see Figure 9 for reference.)

**Apendix:**

**Task 1: C)**

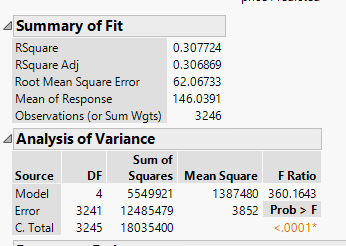


Figure 1

Here P value(<0.001) shows that we can reject technical null hypothesis. There is an association between price(dependent) and Beds, Bathrooms, Bedrooms, Accommodates (independent variable.) This model actually improvise error handling by 30.77% compare to baseline model.

**D-1)**

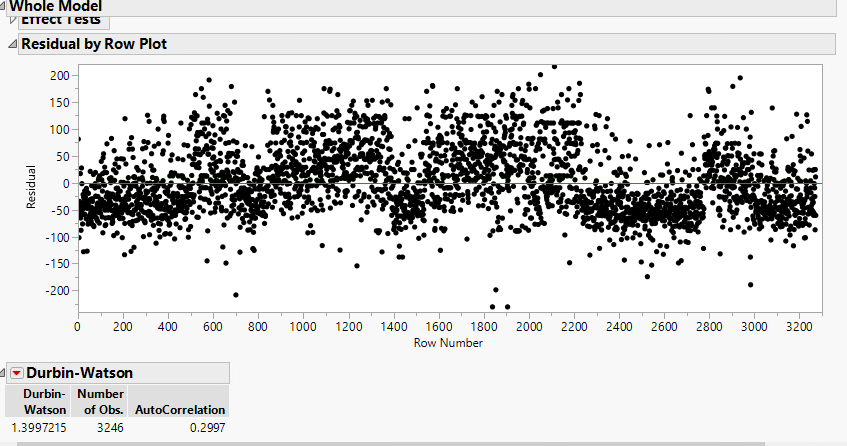


Figure 2

Residual plot by row(random up and down around 0 line on Y axis) and Durbin Watson number shows that

there is no correlation between errors.

**D-2)**

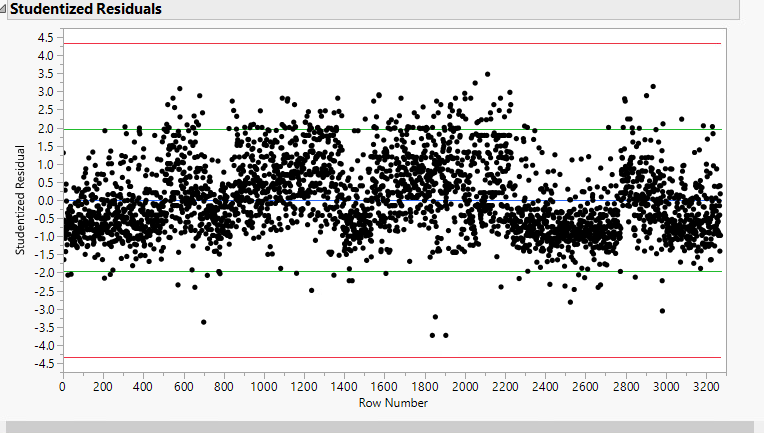


Figure 3

Studentized Residual plot shows that there is no particular patter of Residual along 0 line. That shows that variance of Error is constant(Homoscedanticity).

**D-3)**

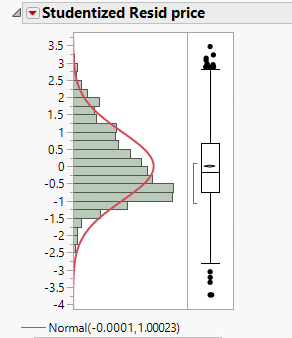
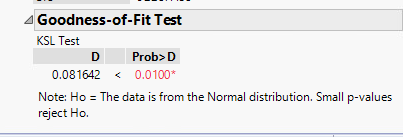
 

Figure 4

Above two figures about normal plot and goodness of fir test(P<0.01) shows that errors are not normally distributed.

**D-4)**

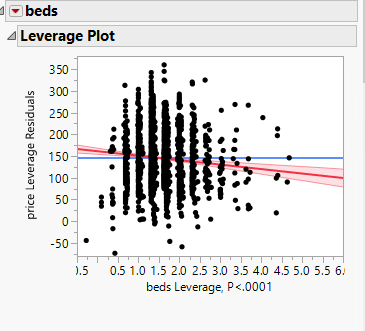
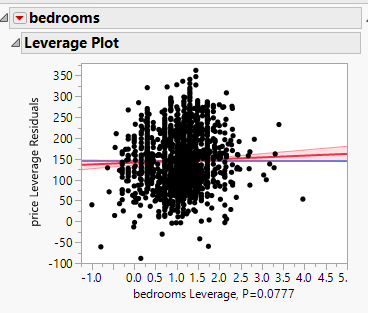
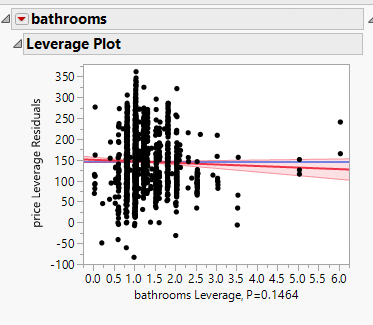
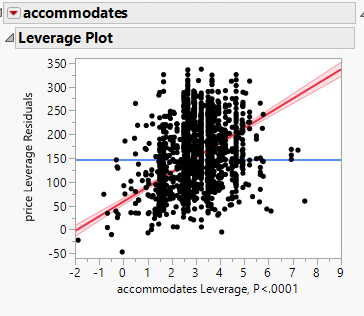


Figure 5.

Here, We can see that accommodation and beds are linearly related to Price. But bathrooms and bedrooms are not. Down the road, we can see in explanation of conclusion that bedrooms and bathrooms is taken as a control variable, so it’s not necessary to have this constraint satisfied for them (see Figure 8 for reference.)

**D-5)**

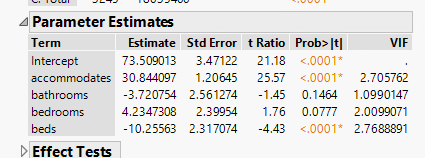


Figure 6

VIF values<10 shows that there is no serious multicollinearity issues associated.

**Task 2)** :

**Explanation of conclusion**.

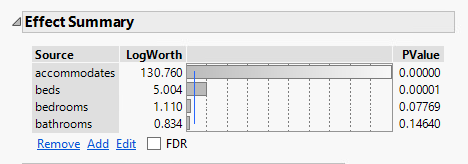


Figure 7

Value of log worth shows that standardized coefficient values is highest for Accommodates followed by Beds followed by Bedrooms followed by bathrooms.

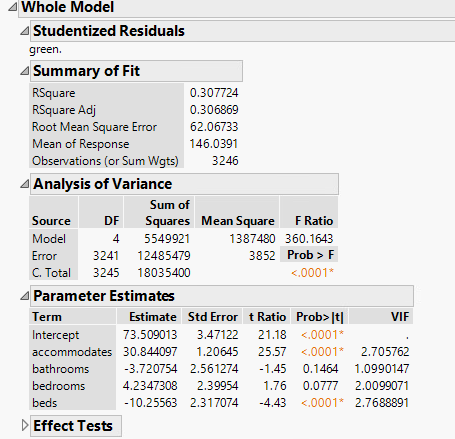


Figure 8

P value in analysis of Variance(<0.0001) show that we can reject null hypothesis. Rsquare adj. value shows that MR model handles error 30% better than baseline model. If we see P values for bathrooms and bedrooms in parameter estimate table, it is great than 0.05, which shows that these factors do not contribute significantly in the model. If we remove those variable and re-run MR test, Rsqaure value would approximately remain same but coefficients of Beds and accommodates drops down from -10.25 to

-9.8 and 30.84 to 31.30 respectively. Intercept value also gets changed from 73.51 to 71.55 Which ultimately shows concern. Instead we can let bedrooms and bathrooms remain as control variable to get accurate model.

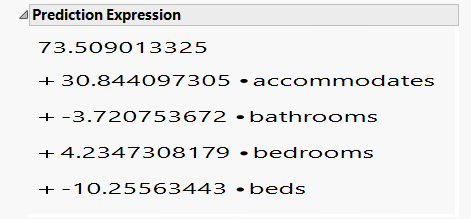


Figure 9

We can say from Figure 9 that ,Price would go up by $30.84 if we increase number of allowed accommodates by 1 and keep other variable constant. In same way, Price would go down by $10.85 if we increase one unit of beds while keeping other variables constant.