Stat 384/484   Final Group Project

This is a group project that contributes 10% to your final grade. You work in a group with 3-5 people. Each group is to submit one report. The submission includes both a regression model and a written report to present the results of your analysis, for each of the following two data sets.

1. **Condo sales**

   Data: condo.xlsx
   Variables:
   Price100: condo price in hundreds of dollars
   Floor: floor location of the unit
   Distance: distance from the elevator
   View: indicator whether it has a view of ocean
   End: indicator whether it is an end unit
   Furnish: indicator whether it is furnished

   Objective: the goal is to build a multiple regression model to predict the sale price of Oceanside condominium units

   The body of your report should indicate how you developed your conclusion. Begin with a concise statement of the regression objective. Next, use your knowledge of the dependent variable and predictor variables to formulate a model. You may need to identify several possible models before finalizing a fitted model which best serves your objective. You may follow the procedure given below when you explore the data.

   1. Begin your study with a graphical investigation of the nature of the relationship between the dependent variable and each predictor variable. You can use scatterplots or boxplots whichever deemed appropriate. Comment on your graphs

   2. The initial regression model to consider is a first-order model which includes all five predictor variables. Use statistical techniques learned to analyze the model.

   3. Next you may want to try models that include the second-order terms of the two quantitative predictor variables without interaction.

   4. To investigate the effect of interaction, add a two-way interaction term, one at a time, between two predictor variables, to the model. Try all the possible two-way interaction terms among the five predictor variables to see if any two-way interaction is significant.

   In your report, you do not need to include full details of all the regression work that you tried. But you should list things that were attempted. Your report should provide enough details to justify your final selection of the best model and to show the major steps that lead to your conclusion. Make sure to include the residual plot for the final best fit model to verify the model assumptions and explain the effect of the predictor variables on the price of the condo.

2. Admission to graduate school

Data: Admission.xlsx
Variables:
Admission: indicator of admission to graduate school
GRE: GRE score
GPA: GPA
Rank: prestige of the undergraduate institution in 4 levels. 1 means the highest prestige, 4 means the lowest prestige

Objective: the goal is to investigate what factors can have effect on admission into graduate school.

The body of your report should indicate how you developed your conclusion. Begin with a concise statement of the regression objective. Next, use your knowledge of the dependent variable and predictor variables to formulate a model. You may need to identify several possible models before finalizing a fitted model which best serves your objective. You may follow the procedure given below when you explore the data.

1. Begin your study with a contingency table of admission by prestige level. Test whether the two classifications are independent.

2. You can then add additional continuous predictor variables into the model. Use statistical techniques learned to analyze the model.

3. You don't need to add higher order terms or interactions, but you may want to explain whether any modifications are needed to lead into your final model.

In your report, you do not need to include full details of all the regression work that you tried. But you should list things that were attempted. Your report should identify factors that contribute to the likelihood of admission to graduate school and explain how they affect the likelihood of admission to graduate school.