

Stat 6021: Project 2 Comments (Group 03: Lotane, Miller, Sery, Swain)

Score: 94

Section 1

- Were high-level results of analysis clearly written?
 - Yes.
- Was section written for the right audience?
 - Not exactly.
 - Remember the audience, so you need to avoid talking about statistical methods that do not mean anything to them. I would have removed the entirety of paragraph 2, the first line of paragraph 3, and the second to last sentence in the last paragraph.

Section 2

- Were description of data and variables provided?
 - Yes.
- Were variables group created clearly identified as such?
 - Yes.
- Were descriptions in an easy to read format?
 - Yes.

Section 3

- Were the two questions of interest stated clearly?
 - Yes.
- Were the two questions motivated and worth investigating?

- Yes.
- Was the response variable for each question clearly stated?
 - Yes.

Section 4

- Were appropriate visualizations presented for first question of interest?
 - Yes.
- Were contextual commentary provided for visualizations?
 - Yes.

Section 5

- Were clear reasons provided for initial model?
 - Yes.
- Were steps taken to improve upon the initial model with reasons provided?
 - Yes.
- Model diagnostics checked?
 - Yes.
- Were influential observations, outliers, and high leverage observations checked and how were they handled?
 - Yes.
 - You did explain the first observation but not the subsequent ones. Usually there is something about these. There is actually at least one more data point that is likely to be a data entry error.
- Was model assessed on test data?
 - Yes.
- Were interesting or relevant interpretations of models provided?
 - Yes.
 - Although I don't agree with your interpretation on page 14, first paragraph. Your response variable is raised to the power of 0.1, so your statement on how changing the size prop by one unit is unlikely to be correct, since such statements are impossible when you use a transformation other than a log transformation.

- Were relevant conclusions that addressed the first question of interest provided?
 - Yes
- Was relevant R output provided?
 - Not all. The R output for the final model was not provided so I cannot verify your interpretations.
- Other comments on Section 5:
 - Should have tried a log transform on price, as interpretations on model coefficients can still be done. I think the residual plot will look fine.

Section 6

- Were appropriate visualizations presented for second question of interest?
 - Yes.
- Were contextual commentary provided for visualizations?
 - Yes.

Section 7

- Were clear reasons provided for initial model?
 - Yes.
- Were steps taken to improve upon the initial model with reasons provided?
 - Yes.
- Was model assessed on test data?
 - Yes.
 - I wish there was some discussion about adjusting the threshold to lower the FNR.
- Were interesting or relevant interpretations of models provided?
 - Yes.
- Were relevant conclusions that addressed the second question of interest provided?
 - Yes.
- Was relevant R output provided?
 - Yes.
- Other comments on Section 7:
 - Plots on pages 18, 19 should have been in the previous section for visualizations.