**Ornery Outliers**

Introduction:

First, try to avoid rhetorical questions in academic writing. Second, the intro needs to be significantly expanded. Provide a little more background about what it is you are precisely modeling and why it is important.

Data:

As with paper two, we need to visualize the variables before we start visualizing the residuals of a linear model. Some basic scatterplots of explanatory variables vs price would be helpful. If those plots are too messy, then some boxplots of the quantitative x variables could be considered. We need something to help the reader understand the type of data you are working with.

Lines 35-45. It is not clear how you could reduce the number of dummy variables by translating the data to numerical values. The purpose of dummy variables is to handle data that cannot be appropriately mapped to numerical values. Please clarify what you mean here.

Figure 7 should be split be enlarged to be as big as it can without leaving the margins. It is currently very difficult to read.

90-100: Multicollinearity is only a problem when interpreting coefficients, but does not affect predictions. This should be clarified in the text. Also, the strategy of manually deciding which variable to keep and which to throw out is not appropriate. Variables should be removed using one of the selection techniques we have discussed in class.

135-140: Remember that interactions and multicollinearity are different things. If two things seem related, that is not evidence of an interaction. Comments of this nature should be clarified in the text and/or removed.

140: In addition, interaction terms should be tested one at a time, not all at once. Please confirm that the p-value results are the same when you look at interaction terms individually.

160-170: You cannot exponentiate coefficients to interpret them on the log-scale. You must provide interpretations on the original scale. In addition, you can never get a negative number when you exponentiate as you do with the yr\_built transformation.

230: What is the accuracy of the null model (a model with no explanatory variables)? Please add this.

240: Avoid rhetorical questions. Also, the conclusion should better summarized what was accomplished in the paper.