**Comments on Presentation**

**Group 5 (Bruins)**

Good description of the data.

The goals of the analysis are a little unclear, eg, whether it is focused on inferences about associations or on prediction. Try to make this clearer in the written report. In some places you focused on the associations but in others discussed results in terms of prediction, e.g., low R-squared.

Linear regression is a reasonable approach. However, the diagnostic plots suggest some extreme outliers and the statement “satisfactory normality” (slide 9) is incorrect.

Good to mention the influential point – also point out that this point has approximately 0 residual.

It wasn’t clear why a different set of predictors was used for analysis of review scores than for analysis of price. Choice of predictors could use better justification in your written report.

Best not to use raw R output (slide 10).

Logistic regression model – show as formula rather than R code

It wasn’t clear if 5,302 is the number of unique hosts or if there are multiple records for the same host.

Be careful about ranking importance of factors based on the OR – the importance also depends on the range of the predictor variable as well as the OR.