**Jelly Beans Review**

35-36: Can you clarify what you mean by an “extremely significant decrease in the number and diversity of the dataset”? I assume you mean in comparison to the national dataset?

38: This section should include some visualizations of the data beyond a table of variable names. We need to visualize the actual variables we will be using in our models before we start modeling. At least some of these visualizations should be reported to the reader.

Page 5: Please place Figures 6 and 8 side by side, do the same for Figures 5 and 7. It will be much easier to see the demonstrated improvement if the comparable plots are closer to each other.

94-95: If the tree only has two branches, please include it in the report. More generally, the discussion regarding the regression tree needs to be heavily expanded. Which variables did you consider in the regression tree? Did you transform the variables like in linear regression or use the raw forms?

98-101: Multiplying a variable by itself is not an interaction, but rather a higher ordered term. This should be clarified. In addition, you should test the significance of the interaction/higher ordered terms one at a time, rather than including them in a model all at the same time.

126-130: Be sure to specify that the increases and decreases in the response variable occur “on average”.