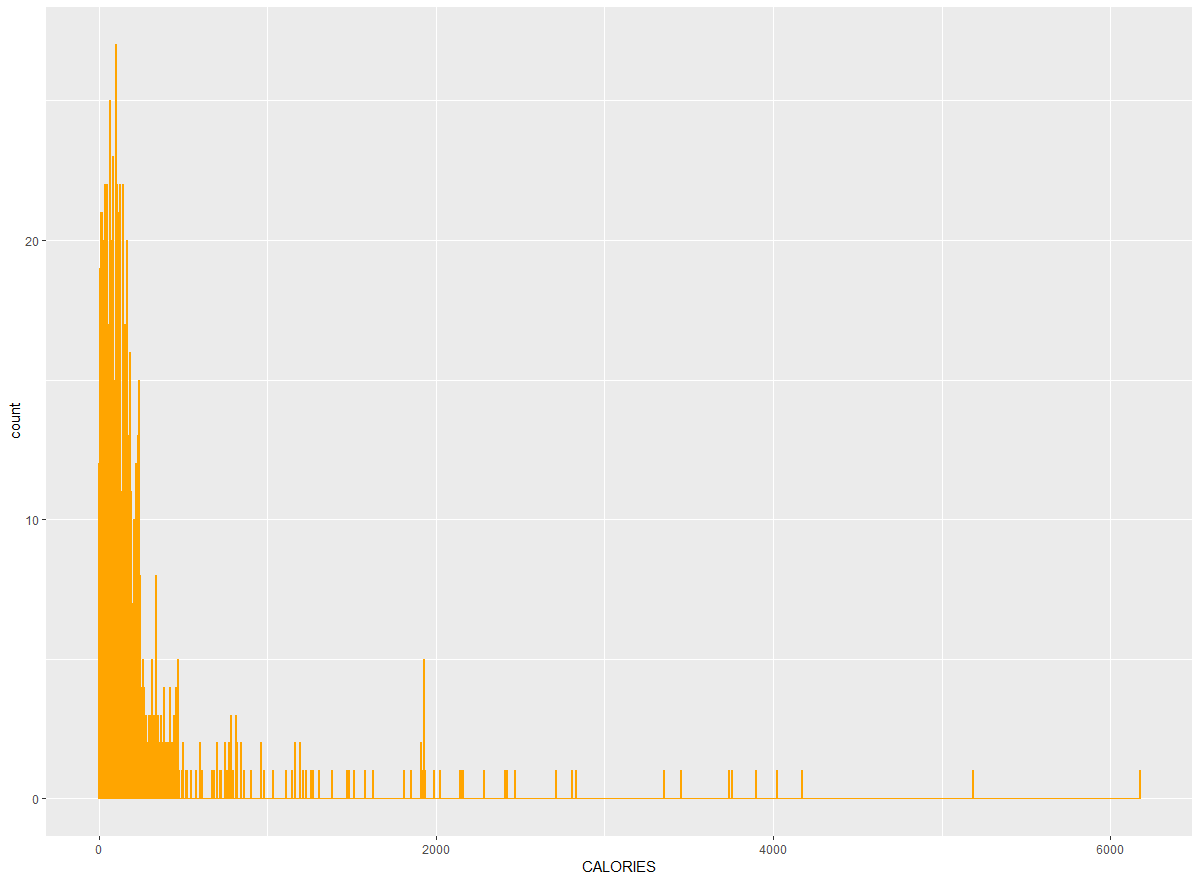
Mary Wishart

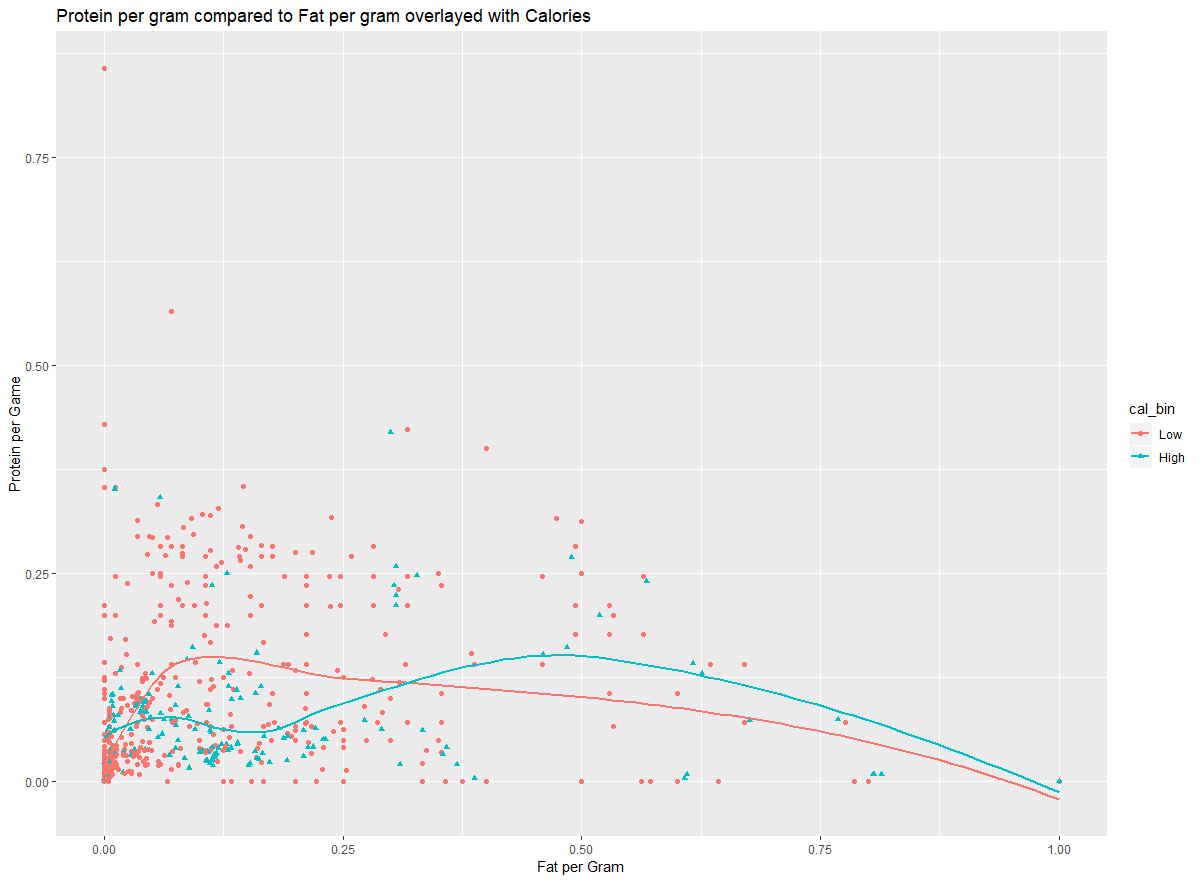
12/3/18

Final Project

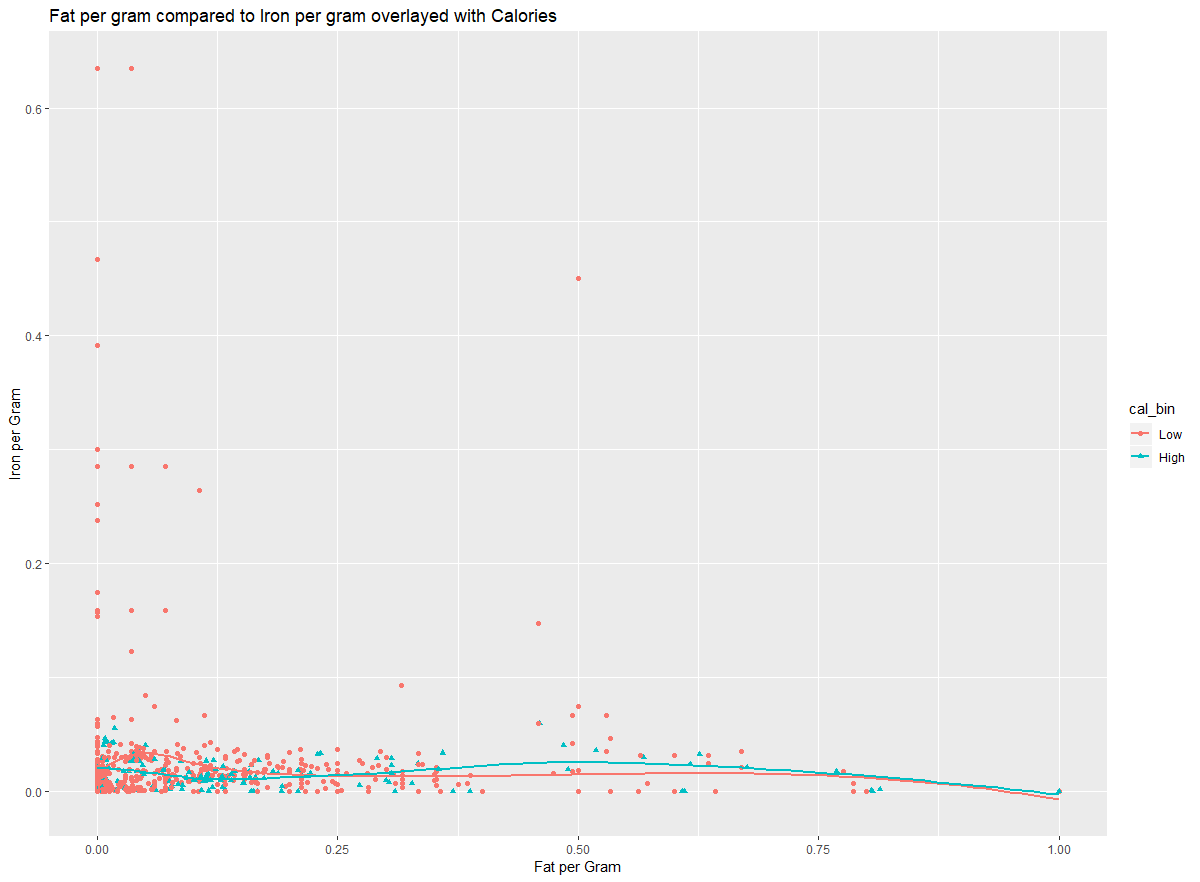
2. 

This histogram is skewed to the right, showing that most food is under 1,000 calories. However, there are outliers going all the way out to 6,000 calories. The histogram has a peak at around 100 calories.

3. I spilt the calories variable into two categories, one for foods less than or equal to 300 calories and the other for foods greater than 300 calories. Basically I spilt the data into low and high calorie foods. I spilt this based off the curve of the data. It seems that most food is low calorie and under 300, but there are quite a few high calorie foods that are greater than 300 calories.

4. 

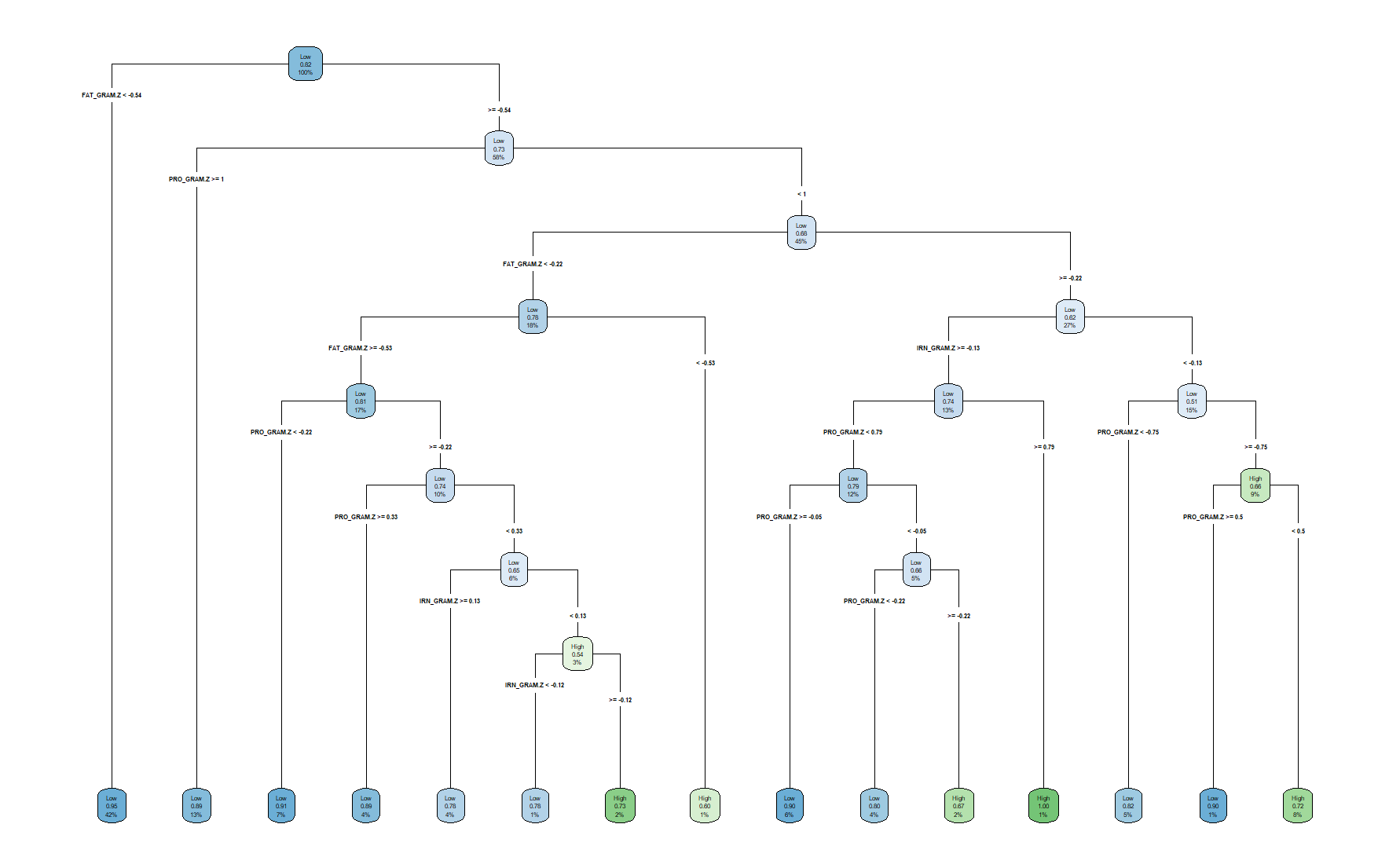
It appears that lower calorie foods have less protein and fat per gram than higher calorie foods. As the amount of fat increases, there is a general trend of protein decreasing and vice versa.

5. 

Iron and fat appear to also increase opposite to each other. Food high in fat are low in iron and vice versa.

6. Iron does not seem to impact the calories in food. All the high iron foods are low calorie. Fat however seems to impact the amount of calories. There are more high calorie foods as fat increases than low calorie foods. Protein is in the middle as it does impact the amount of calories in food, however it is not as much as fat.

8. A model that predicts that all foods are low calorie foods, the model will be correct 82.3% of the time. If the model predicts all the food is high calorie then the model will be correct 17.7% of the time.

9. 

The decision rule that applies to the highest amount of data is that a food that has less than 0.0086 fat per gram then that food is low calorie with 95% confidence. The decision that applies to the second highest amount of data, is that if the food has more or equal than 0.0086 fat per gram and then more or equal to 0.16 protein per gram then the food is also low calorie with a confidence percentage of 89%. The third decision rule states that is the food is more or equal to 0.0086 fat per gram, less than 0.16 protein per gram, then has more or equal to 0.074 fat per gram, less than 0.014 iron per gram, then more or equal to 0.0022 protein per gram but less than 0.11 protein per gram then the food is a high calorie food at 72% confidence. The fourth decision rule states the food is more or equal to 0.0086 fat per gram, less than 0.16 protein per gram, then less than 0.074 fat per gram but more than or equal to 0.011 fat per gram, and the food has less than 0.05 protein per gram then the food is low calorie with a confidence of 91%. The fifth highest decision rule is the food is more or equal to 0.0086 fat per gram, less than 0.16 protein per gram, then has more or equal to 0.074 fat per gram, has more or equal to 0.014 iron per gram, and has less than 0.14 protein per gram but more or equal to 0.065 protein per gram the food is low calorie with a 90% confidence.

10.