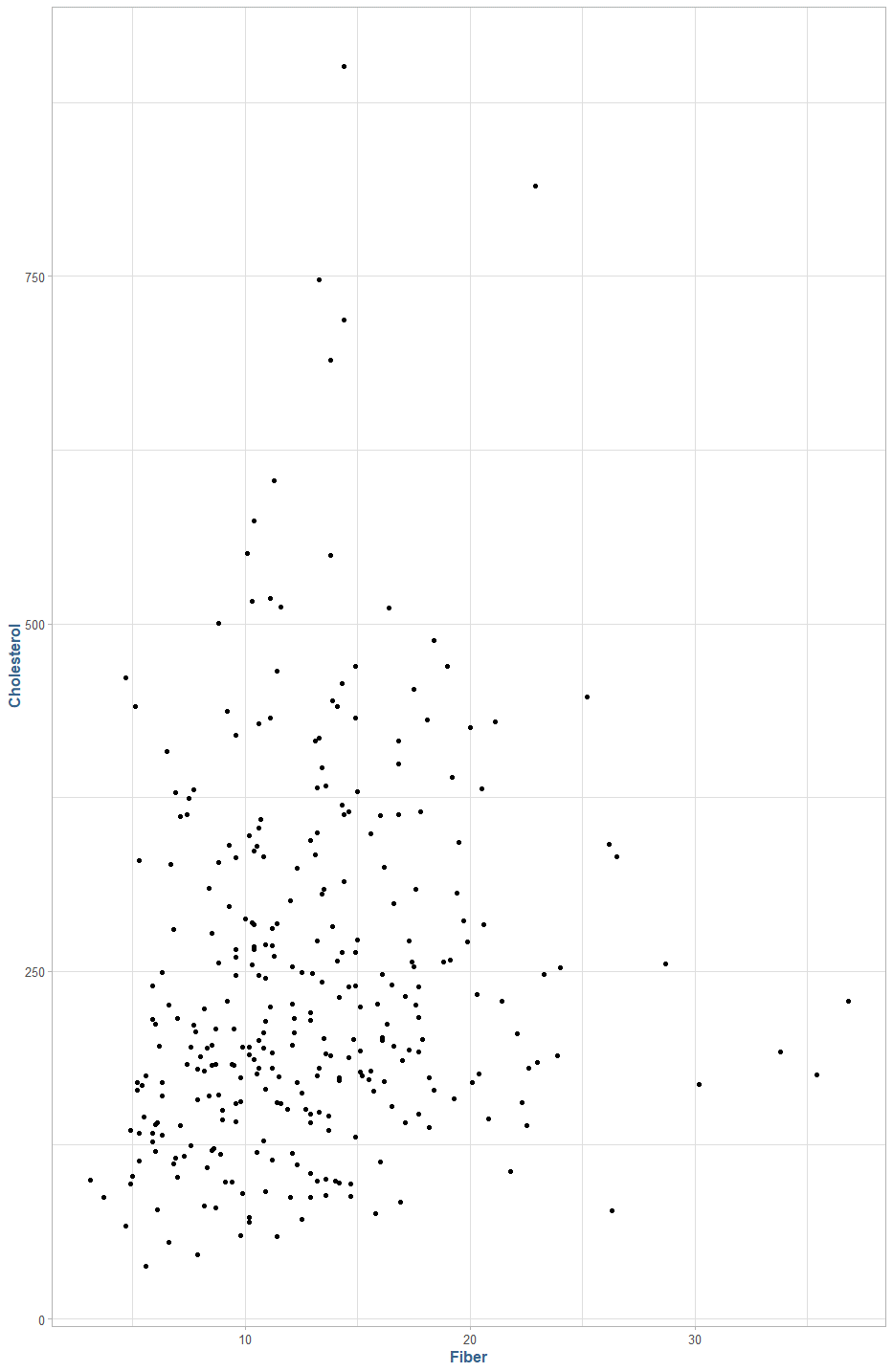
Assigment #

Brandon Moretz

### Introduction

1. Consider the continuous variable, FIBER. Is this variable correlated with Cholesterol? Obtain a scatterplot and appropriate statistics to address this question.

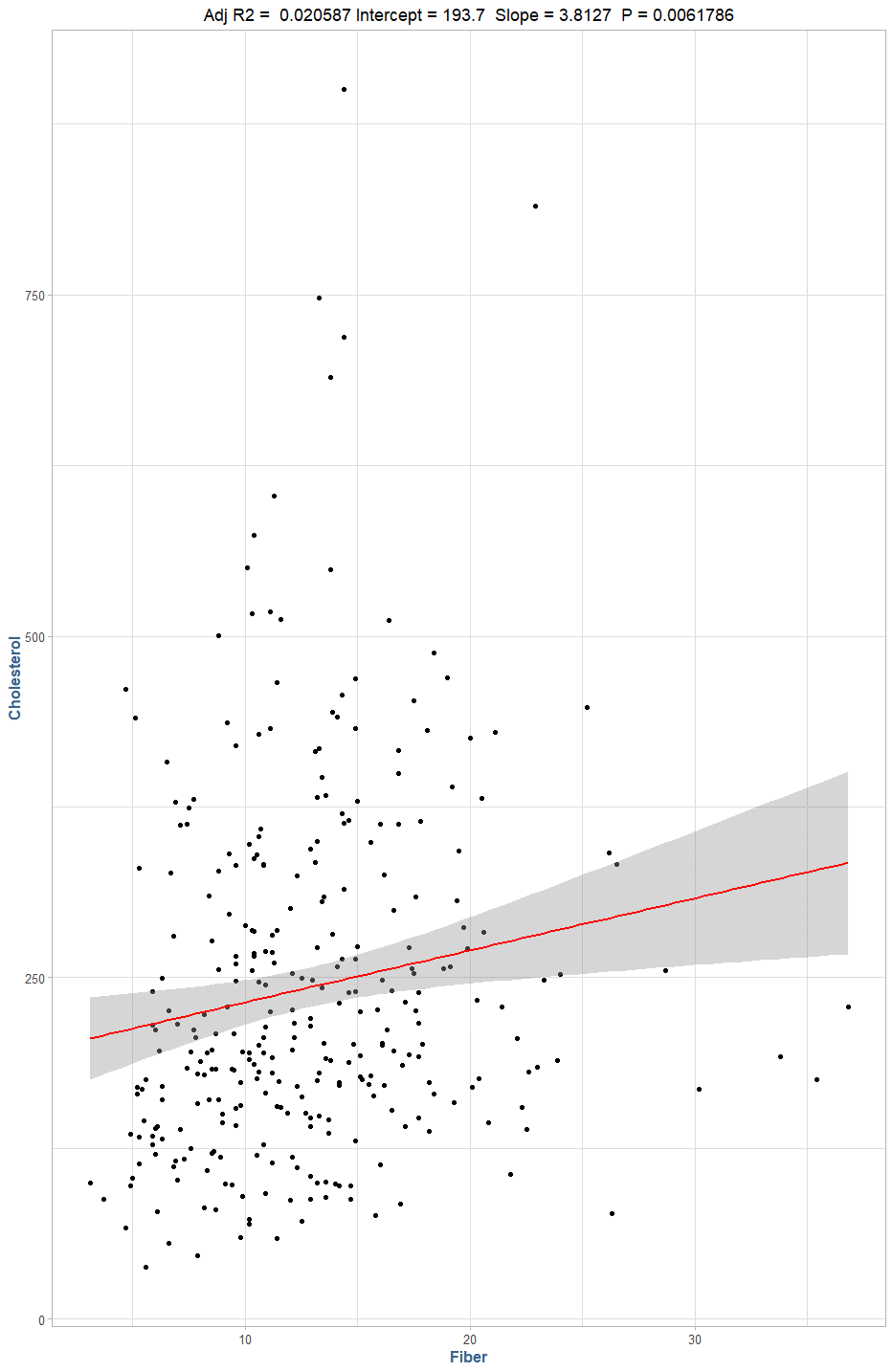
In the following diagram we see a lot of variance in the following scatterplot of Cholesterol ~ Fiber.



There is a liner correlation of .1539 between Cholesterol and Fiber.

1. Fit a simple linear regression model that uses FIBER to predict CHOLESTEROL(Y). Report the model, interpret the coefficients, discuss the goodness of fit.

The Fiber variable appears to be a relatively poor indicator of cholesterol. In the graphic below, we can see an R2 value of 0.0237, indicating that approximately 2% of the variance in the data is explained by the Fiber variable. Additionally,



Model: 193.7014 + 3.8127 β1, where β1 is fiber.

3.) For the ALCOHOL categorical variable, create a set of dummy coded (0/1) indicator variables. Fit a multiple linear model that uses the FIBER continuous variable and the ALCOHOL dummy coded variables to predict the response variable Y=CHOLESTEROL. Remember to leave one of the dummy coded variables out of the model so that you have a basis of interpretation for the constant term. Report the model, interpret the coefficients, discuss hypothesis test results, goodness of fit statistics, diagnostic graphs, and leverage, influence and Outlier statistics. This is called an Analysis of Covariance Model (ANCOVA)

### Research

### Conclusion