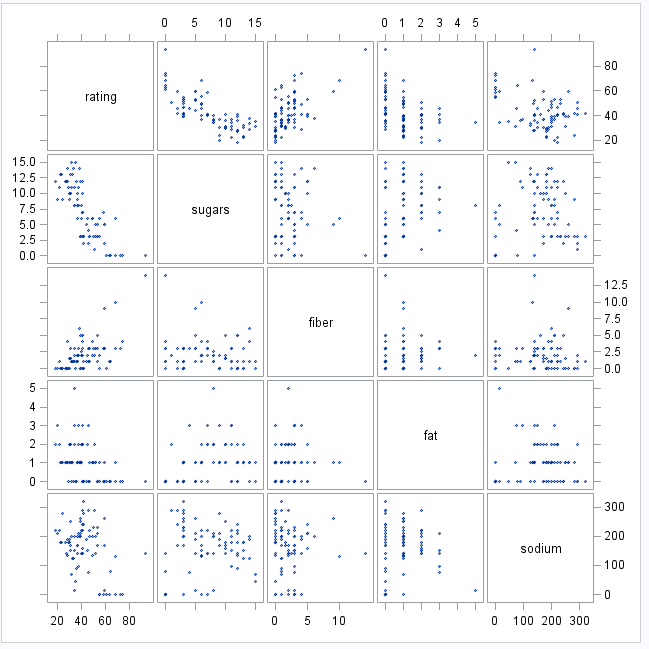
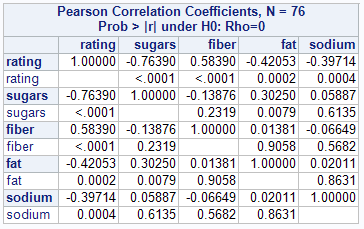
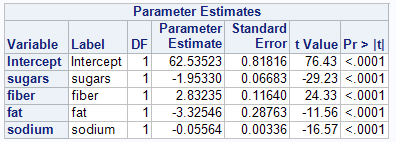
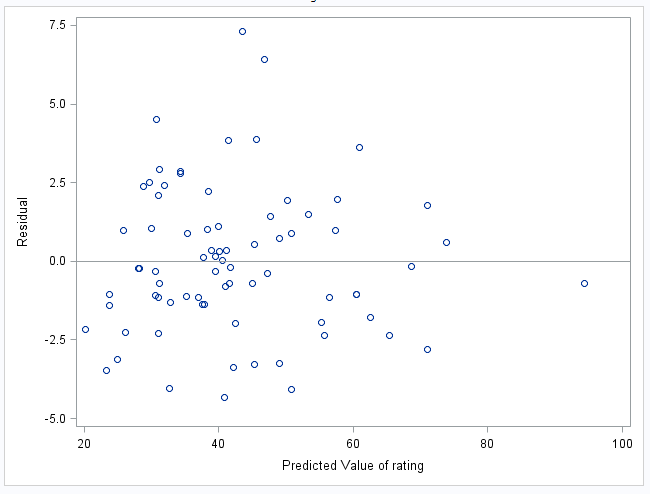
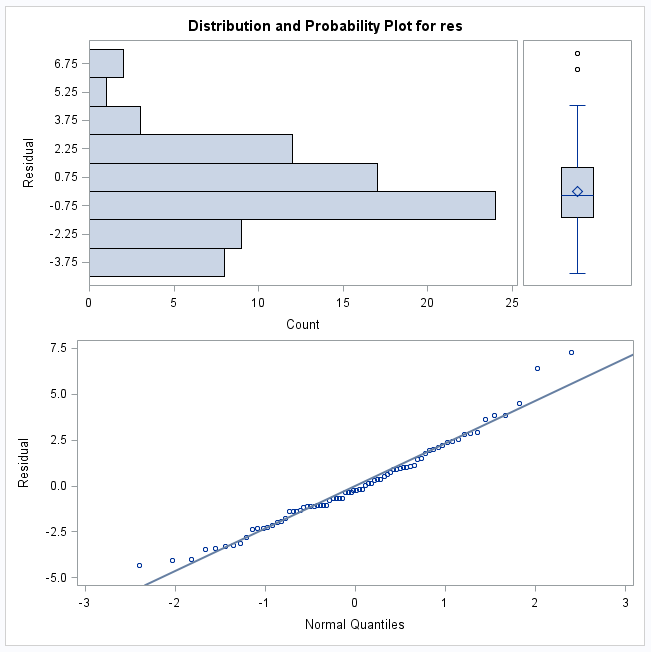
Lab 6

1. 
   1. 
   2. Sugars- negative. Fiber- positive. Fat- Negative. Sodium- Negative
   3. This is because the Fat value is either a 0, 1, 2, or 3. There is no variation for it.
   4. 
   5. Sugars: -0.76. Fiber: 0.58. Fat: -0.42. Sodium: -0.40. The strongest correlation is Sugars because it is the closest number to 1 or -1
   6. It shows that Sodium, Sugars, and Fat are a negative correlation and Fiber is a positive correlation.
   7. The strongest correlation outside of the ratings variable would be between sugars and fat. The Correlation value is 0.30 which is a pretty high.
   8. Y(hat) = 62.54 -1.95\*sugars + 2.83\*fiber – 3.33\*fat – 0.06\*sodium
   9. 
   10. 4 assumptions
       1. Linearity. Equivalently
       2. Constant variance.
       3. Normality
       4. Independence
       5. 
       6. This plot indicates that the linearity is satisficed because the residuals are centered around the zero line for all ranges of predicted values. This plot also indicates that the constant variance is satisfied because the vertical spread of the residuals is roughly the same for all ranges of predicted values.
   11. 
       1. The normality and independence is checked because the histogram is representing a bell-shaped curve
   12. 23.29 predicted rating
   13. Yhat = 62.54 -1.95\*9 + 2.83\*0 – 3.33\*3 – 0.06\*210
   14. -3.47
   15. Rating – yhat value
   16. 62.53. Yhat = 62.54 -1.95\*0 + 2.83\*0 – 3.33\*0 – 0.06\*0
   17. This is the estimated value because the estimated intercept is 62.54 and the other values are all 0
   18. 65.36. Yhat = 62.54 -1.95\*0 + 2.83\*1 – 3.33\*0 – 0.06\*0
   19. 0.116. When you add 0.116 fiber, the rating goes up 1.
   20. When you add one fiber to it, the rating increases by 2.83.
       1. 71.03. Yhat = 62.54 -1.95\*0 + 2.83\*3 – 3.33\*0 – 0.06\*0
       2. When you add 2 more fibers, the value goes up 2.83(2) because the rating went up 2.83(2)
   21. 57.38. Y(hat) = 62.54 -1.95\*7 + 2.83\*3 – 3.33\*0 – 0.06\*0
   22. 0.06. When you add 0.06 sugars, the rating goes down by one.
   23. The difference between the predicted ratings for Shredded Wheat and Frosted-Mini-Wheats is that you are doing 7\*1.95 so the rating goes down from the amount of sugar being added (1 rating point down for each 0.06 sugars added)
   24. R-Squared: 0.973 and Root MSE: 2.385
   25. R-Squared is interpretated as: 97.3% of total variation in rating is explained by the model with sugars, fat, sodium, and fiber.
   26. The Root MSE value is the average distance from actual rating to the predicted rating is 2.385.
   27. I think that it predicts it very well. If the average distance from the actual rating and the predicted rating is only 2.385, that means the prediction is not far off at all.
   28. 
   29. 
   30. I think if you add protein, your correlation rate will be higher and adding it will help reduce the Root MSE value to help strengthen the relationship between the variables.
2. Using all of the variables will give you the strongest correlation.