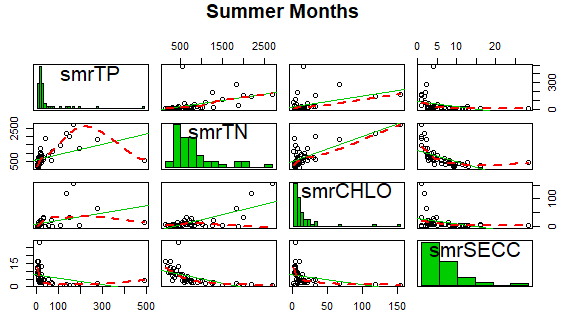
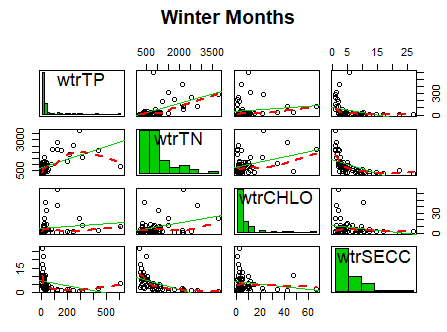
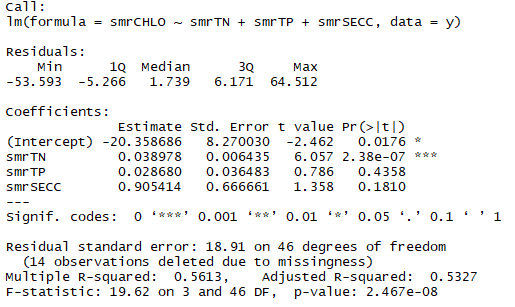
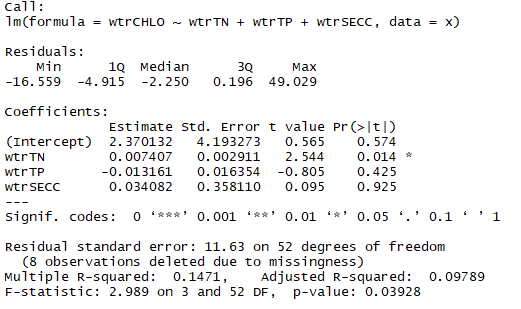
James Young Homework 7

**1. Show side-by-side scatterplots of summer and winter variables and comment on your initial assessment of linear relationships for summer and winter**



There seems to be no strong correlations between the independent variables in the summer or winter months. There seems to be a week correlation between total phosphorus and total nitrogen, however.

**2. Fit the BOTH regressions for summer and winter using R and provide the snip of the coefficients side by side and label both. Example below**



Winter Summer

**3. Fit BOTH the estimated regression function (use the ‘hat’ function) using your results from step 2**

Summer :

Winter :

**5. Test if all βn’ s is equal to zero. State the hypothesis (once), use the R output and use the p-value from the F test to perform the test, and provide the decision rule, and conclusion for each Summer test: Winter test: (α = 0.05)**

For βinterceptSummer , P value = 0.0232, therefore we reject the null hypothesis and conclude that βinterceptSummer  0. For βsmrTN, P value = 2.59\*10-7, therefore we reject the null hypothesis and conclude that βsmrTN 0. For βsmrTP, P value = 0.4421, therefore we fail to reject the null hypothesis and conclude that βsmrTP may = 0. For βsmrSECC, P value = 0.2265, therefore we fail to reject the null hypothesis and conclude that βsmrSECC may = 0.

For βinterceptWinter, P value = 0.574, therefore we fail to reject the null hypothesis and conclude that βinterceptWinter may = 0. For βwtrTN, P value = 0.014, therefore we reject the null hypothesis and conclude that βwtrTN 0. For βwtrTP, P value = 0.425, therefore we fail to reject the null hypothesis and conclude that βwtrTP may = 0. For βwtrSECC, P value = 0.925, therefore we fail to reject the null hypothesis and conclude that βwtrSECC may = 0.

**6. Interpret the coefficient(s), by this I mean give the following**

βinterceptSummer = -, there is no unit x increase that would effect y for the intercept, it is not interpretable as it is less than 0

βsmrTN = , for every unit increase in x, y increases by 0.039878, this variable is significant.

βsmrTP = , for every unit increase in x, y increases by 0.02868, this variable is not significant.

βsmrSECC = , for every unit increase in x, y increases by 0.905414, this variable is not significant.

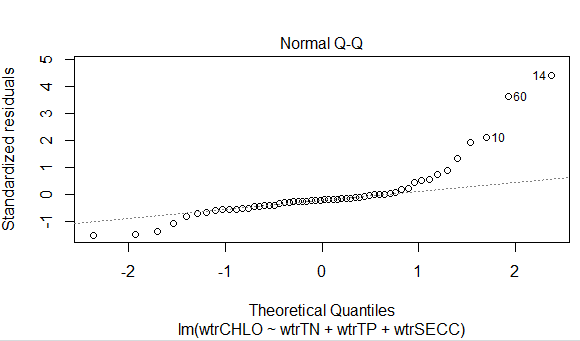
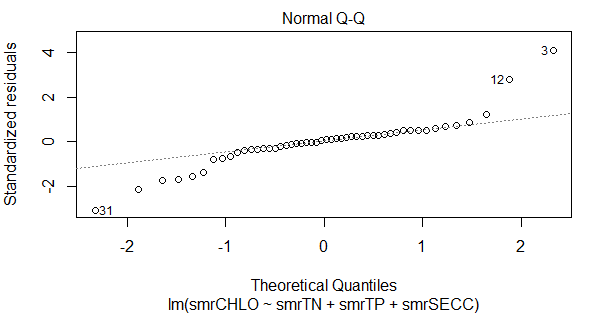
βinterceptWinter = 2.37, there is no unit x increase that would effect y for the intercept, it is interpretable as it is greater than 0

βwtrTN = 0.007407, for every unit increase in x, y increases by 0.007407, this variable is significant.

βwtrTP = -0.013161, for every unit increase in x, y decreases by 0.013161, this variable is not significant.

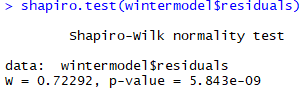
βwtrSECC = 0.034082, for every unit increase in x, y increases by 0.034082, this variable is not significant.

**7. Provide the only the qqplot with the qqline for summer and winter side by side. Examine the residuals distribution from the qqplot and comment on the distribution FOR EACH**



Winter Summer

Winter Q-Q plot shows large deviation in residuals at tails and so does Summer Q-Q. The distribution for both graphs does not look normal visually.

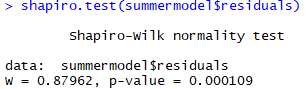
**8. Perform the Shapiro-Wilks Test to test the normality FOR EACH. You need hypothesis test, R snip with the result and the decision rule and conclusion**

H0 : All residuals came from normal distribution

H1: Any residuals did not come from a normal distribution

Decision: Calculated p-value = 5.8e-09. 5.8e-09 < 0.05 therefore

We reject H0 and conclude the residuals are not from a normal distribution.

H0 : All residuals came from normal distribution

H1: Any residuals did not come from a normal distribution

Decision: Calculated p-value = 0.0001. 0.8372 > 0.05 therefore

we reject H0 and conclude the residuals are not from a normal distribution.

**9. Comment and compare each model fit and express any evidence that whether summer or winter models are adequate use R squared value and any other diagnostic tools.**

The winter model had a multiple R-Squared value of 0.1471 while the summer model had a multiple R-Squared value of 0.5631. Neither of these multiple R-Squared values accurately or adequately represent the data, but the summer model is a better model choice (for summer conditions) than the winter model is for winter conditions.

**10. Describe if any of the variables that are limiting factors for y.**

Total nitrogen appears to be the limiting factor of y and is the only significant coefficient of the independent variables.

**11. Are answers the same for summer and for winter? why?**

Yes, nitrogen is the only significant independent variable for both winter and summer so it is the limiting factor for both.

R Code

