Golf Demand in Seattle

Abstract:

Through this paper, we examine whether or not the demand for golf is seasonal, which factors contribute the most to the number of *ROUNDS* of Golf played, as well as whether or not the three city owned (MGS) courses should lower their rates in the winter months. To begin to answer these questions, we were given a data set from twenty 22 courses in the Seattle area: with eighteen explanatory variables for the number of rounds of golf played. We started by conducting various F-Wald tests, as well as an Auxiliary Regression on an unrestricted data set from twenty-two golf courses. We then were able to omit three variables from the data set: price of playing in the winter (*Winter\*Fee),* if the golf course had a driving range (*Range),* and length of the course (*Yard).* Then, by using a Breusch-Pagan test, heteroskedasticity was found, forcing us to use a robust regression with White Heterskedasticity-consistent Standard Errors, validating the t-statistics of the estimated coefficients on the explanatory variables. After the proper tests were executed, variables pertaining to the winter months: *Winter* (dummy variable equaling 1 during November-February)*, Rain* (average amount of rain for the month)*, Temp* (average temperature)*, MGS\*FEE* (the fact that MGS doesn’t lower their fee during the winter)*,* had a significant negative coefficient in the demand function for playing golf. This argues that there is a seasonal demand for golf, as there were drastically less rounds of golf played in the winter months from November to February, than in summer of spring. We would theoretically advise the MGS to lower their playing fees during the winter months to increase the number of rounds played, and to be able to better compete with other courses that lower their rates during winter. This being said, they should also not increase their rates in the summer or spring months.

Main Body:

To answer these questions, we have been given monthly data over the course of the year 1999 from the University of Washington’s Economics Department; containing information on 22 golf courses in the Seattle area. This dataset included an array of independent explanatory variables related to the dependent variable: (*Rounds*), the total number of rounds played at each course for every month. These explanatory variables were as follows: *Fee, FeeSub, Rain, Temp, Rating, Slope, Yard, Distance, Cart, Range, Winter, MGS).* For the first question of elasticity of demand for golf, we will examine the demand for playing a round of golf during both winter and non-winter months. However, before we can begin to determine whether it is seasonal, we must acquire the best possible model for the demand of *Rounds.*

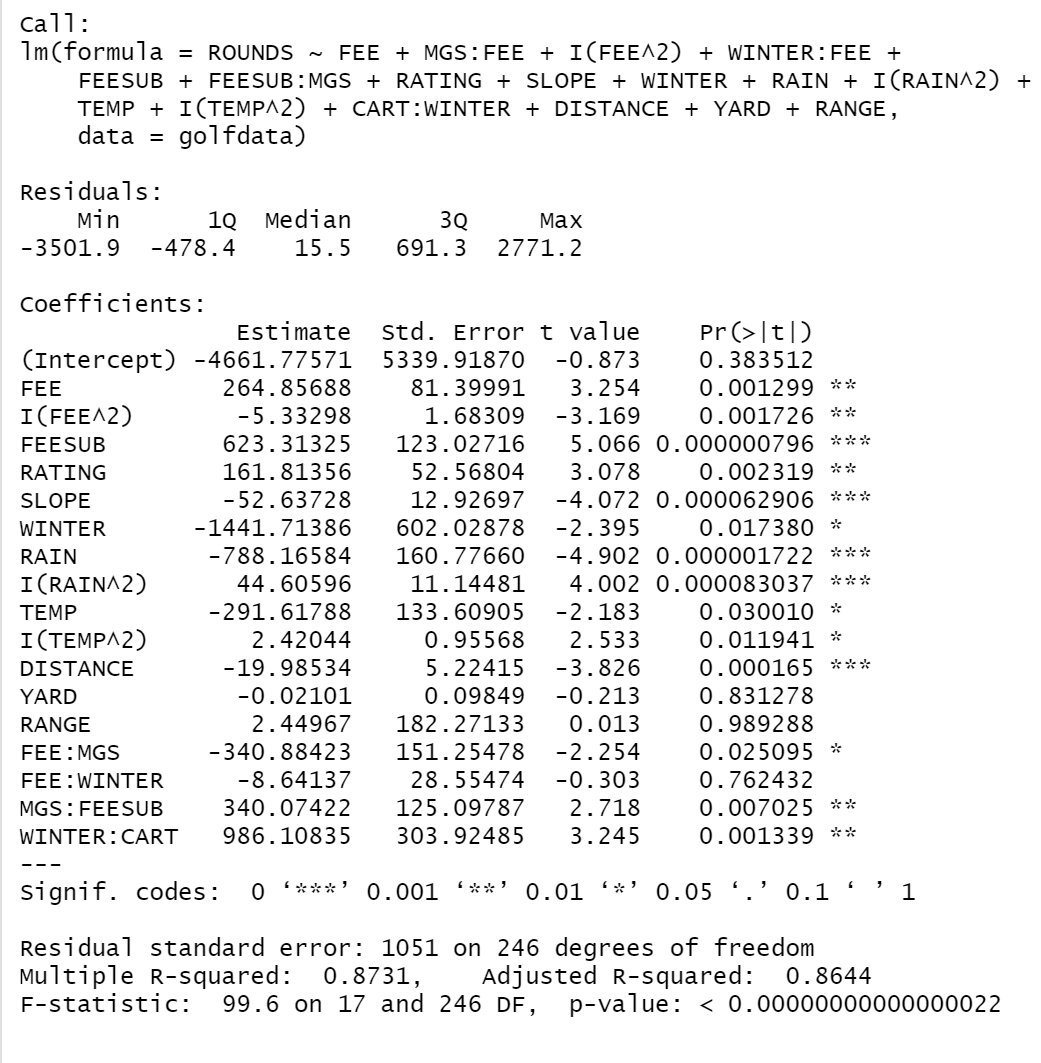
First, we began by using an ordinary least squares test (OLS), in order to provide unbiased coefficients on each of the explanatory variables. The unrestricted model (Figure A), is given in the dataset as an initial guess for these coefficients, however it is doubtful that it is the true best fit. We decided to include squared terms because we recognized that RAIN, TEMP, and FEE were not 100% linear in predicting the amount of ROUNDS played. Our reasoning was that as the amount of RAIN increases, most of the time it will lead to fewer people wanting to play Golf, however some don’t mind it as much. As well for TEMP, as the temperature increases, you can expect more people to play Golf, however it will eventually reach a point where it is too hot to play. Finally with the squared term of FEE, we figured that as the FEE increased, more and more people would cease to want to play a round of Golf at the course in question, however the course might have a high FEE because it is much more maintained, and has better playing conditions; leading to more rounds of Golf played at that course. At first glance, we suspect that the variables:  *Winter\*Fee, Range, Yard,* could be removed from the dataset, as they had significantly higher “p” values than the other variables. By conducting F-Wald tests for each of the variables in question (Figures X-Z), it is certain that these variables can be removed from the list, to create a new restricted model (Figure B). Going an extra step further, we calculated an auxiliary regression of the residuals of the new restricted model onto the explanatory variables of the unrestricted model. Given by Figure C, this shows that: the test statistic nR^2 < the chi-squared critical value of 7.81, at 3 degrees of freedom at the 5% level; meaning that we can omit these three variables.

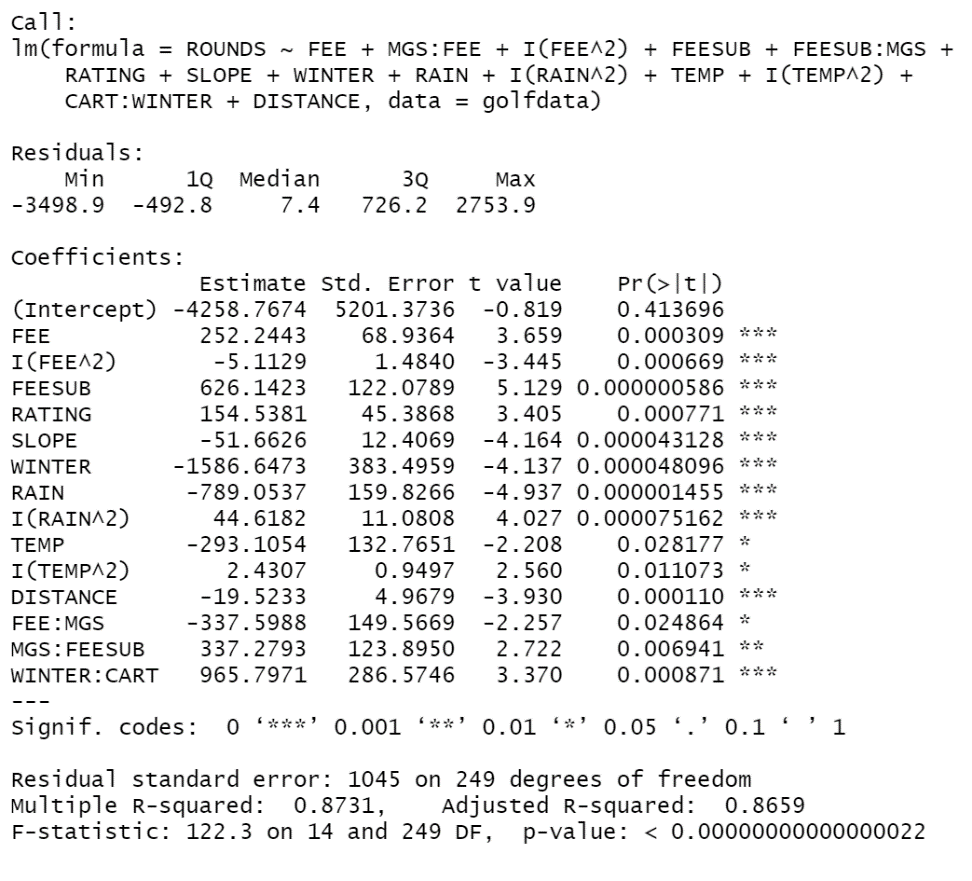
After the three variables are eliminated from the model, the next step is to determine if it contains heteroskedasticity. In order to do this, we used the Breusch-Pagan-Godfrey test on the restricted model, with null hypothesis of Homoskedasticity, and alternative hypothesis that at least one of the Epsilon error terms is equal to a different observation’s variance. This gave a statistic of: nR^2 = 70.40069, and a probability of failing to reject the null of: 0.00. This means that there is Heterskedasticity present in the restricted model, and we must eliminate it in order to validate the t-statistics on the coefficients. In order to do this, we will adjust the Standard Errors of each error term, and then estimated the restricted model using White Heteroskedastic consistent standard errors and Covariance (Figure D); ultimately adjusting for the heteroskedasticity in the model.

With the new Figure D, we can now be certain that this current model for *Rounds* contains Homoskedasticity. Next, we verified that each of the signs and magnitudes of the coefficients for each variable are logical. Looking at *MGS\*Fee* (-337.599), this coefficient is negative, which interprets to a high elasticity of the price (*FEE)* corresponding to the MGS courses. Next, for *FeeSub*, the model shows a positive correlation between the fee charged by each course’s competitors and the demand for a specific course, which follows the effect of the price of a substitute good. The coefficient for *Rating* is also positive, meaning that the courses with higher ratings in terms of difficulty will bring in adept players, but not necessarily discriminate against bad golfers. However, *Slope* has a negative coefficient, meaning that worse golfers will prefer a course with less inclines, but remain on the fence of deciding about the *Rating*. Moving on, the variable *Winter*, has a significantly large magnitude of -1586.647, meaning that for the months of winter, demand for golf will shrink dramatically. This is logical, because in the winter months from November to February, weather conditions are not favorable being cold and rainy; as well as having days be shorter with less time to actually play golf. When comparing the variables, *Rain* and *Temp*, it is certain that *Rain* has a larger magnitude of -789.054, as opposed to *Temp*’s -293.105. This makes sense because even if the temperature was suitable, people would still be less likely to golf if it was raining. With regards to *Cart\*Winter*, it is interpreted that golfers strongly prefer courses with a paved cart path over those without them; with a large magnitude of 965.797. Finally, with respect to *Distance*, its sign is negative, meaning that people would prefer courses within close vicinity to Seattle.

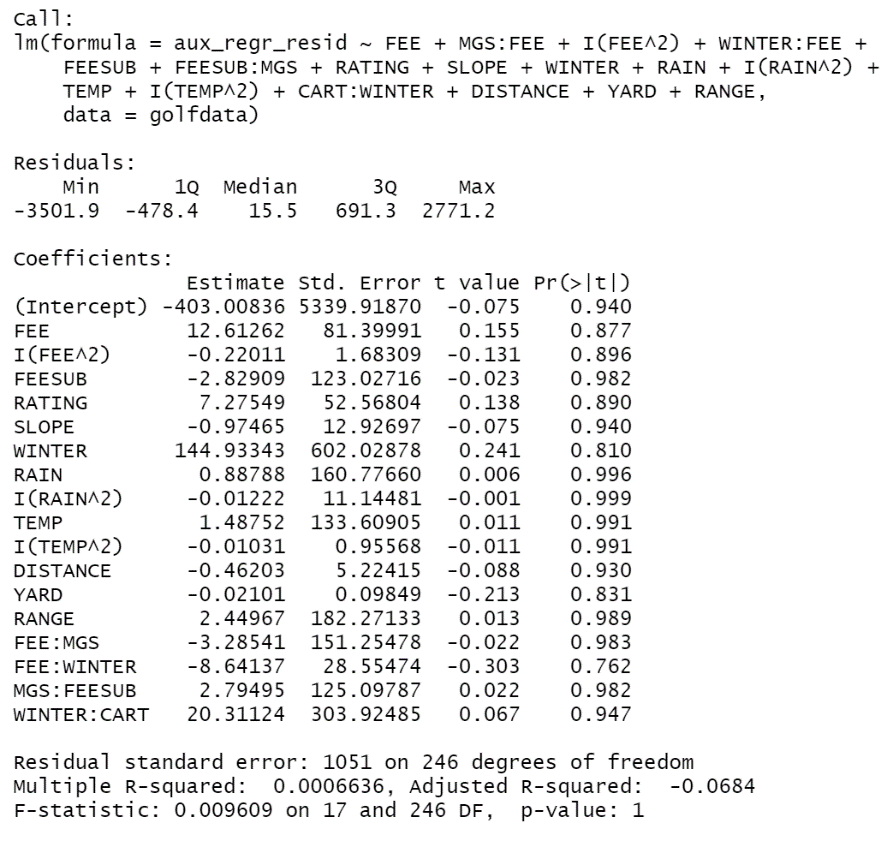
In conclusion, we believe that the demand of Golf is in fact seasonal, having a significant bigger proportion of *ROUNDS* in the summer and spring months than in the winter months of November-January. To answer our second question, the variables with the greatest significance on *ROUNDS* were: *FEESUB , WINTER:CART, WINTER, RAIN, TEMP,* and *FEE:MGS.* Furthermore, because the three MGS courses have the highest amount of *ROUNDS* played year round, we believe that they would benefit from lowering their prices in the winter. This being said, they should not increase their rates in the spring or summer months, because of the fact that *MGS:FEE* has a very large negative coefficient.

**APPENDIX**

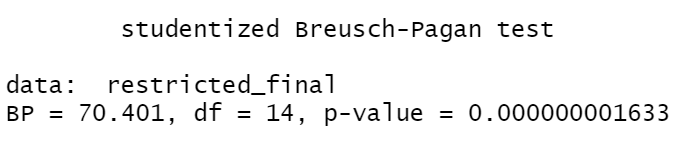
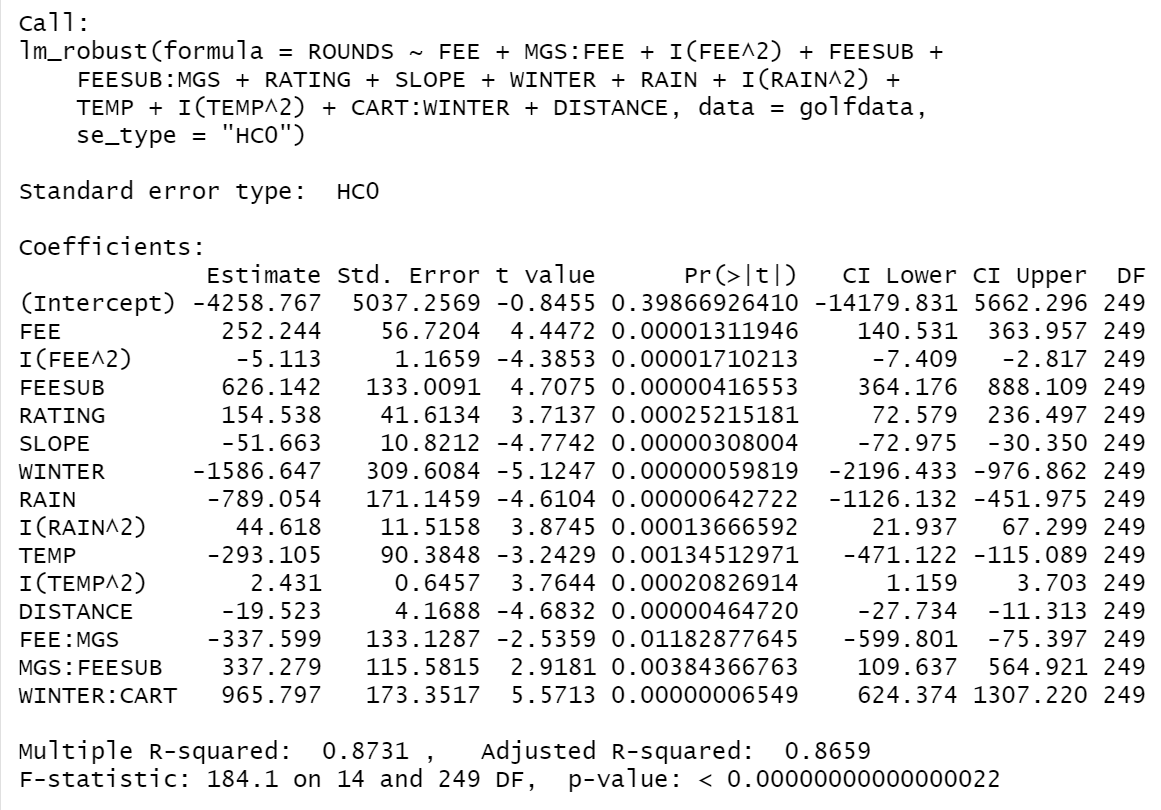
**Figure A:** Unrestricted Regression Model 

**Figure B:** Restricted Model Regression 

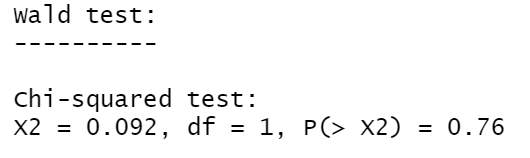
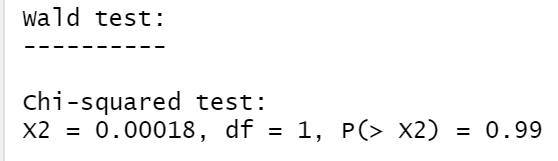
**Figure C:** Auxiliary Regression of Restricted Model residuals onto Unrestricted Model’s explanatory variables

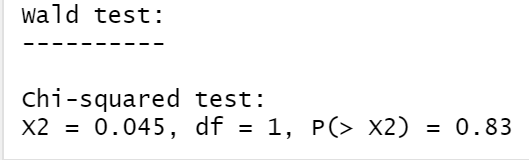


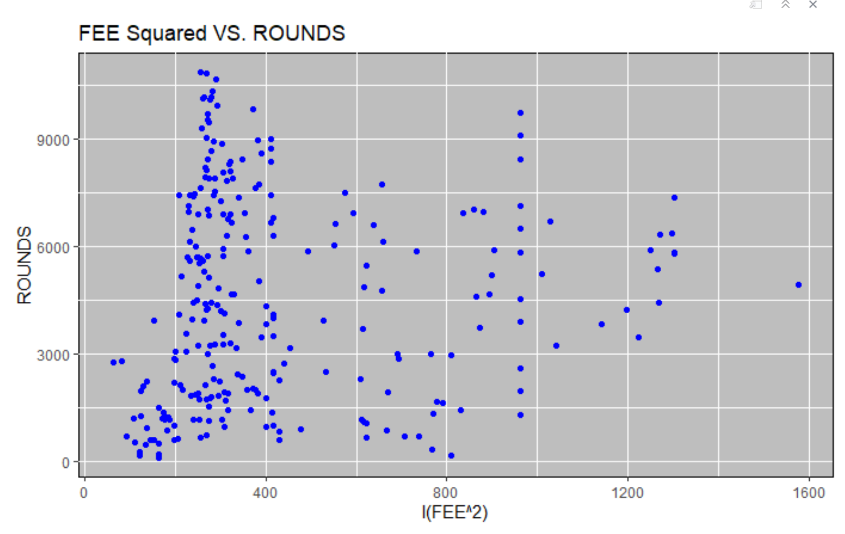
**Figure(s) D:** Breusch-Pagan-Godfrey Heteroskedasticity Test and New Robust Regression with White Heteroskedastic consistent standard errors and Covariance

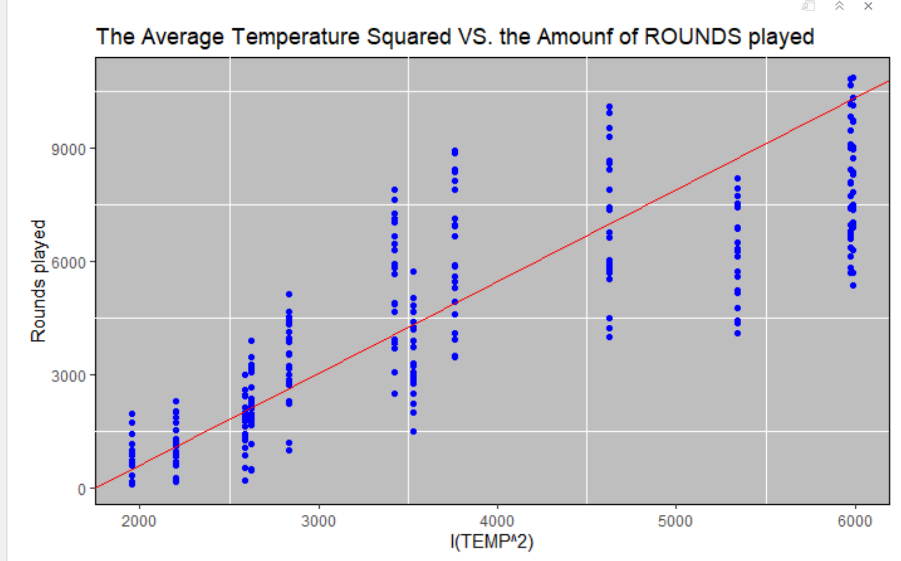


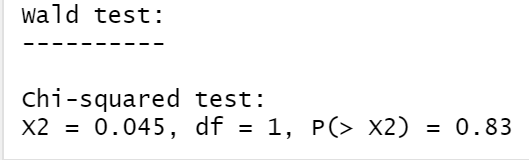
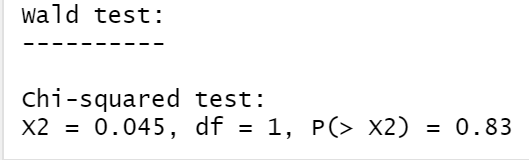
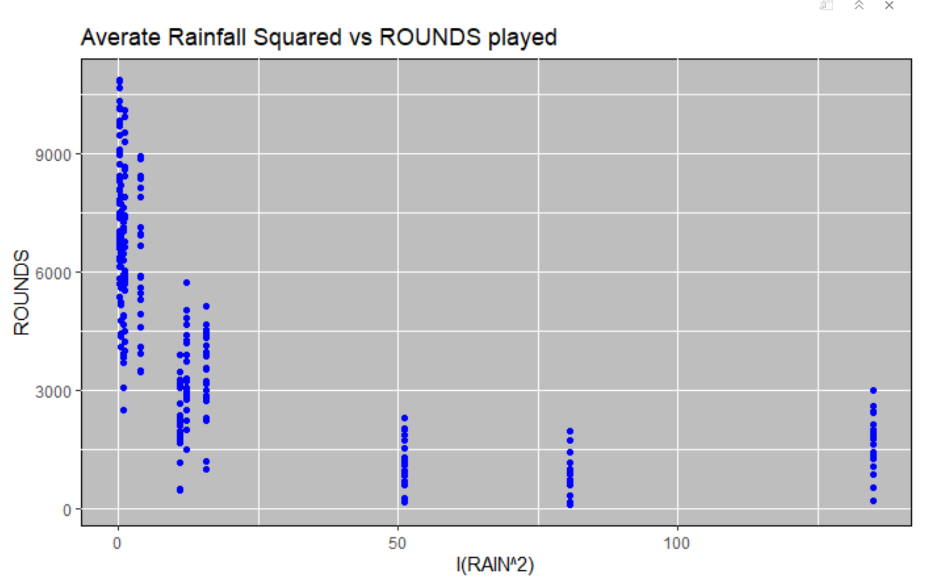
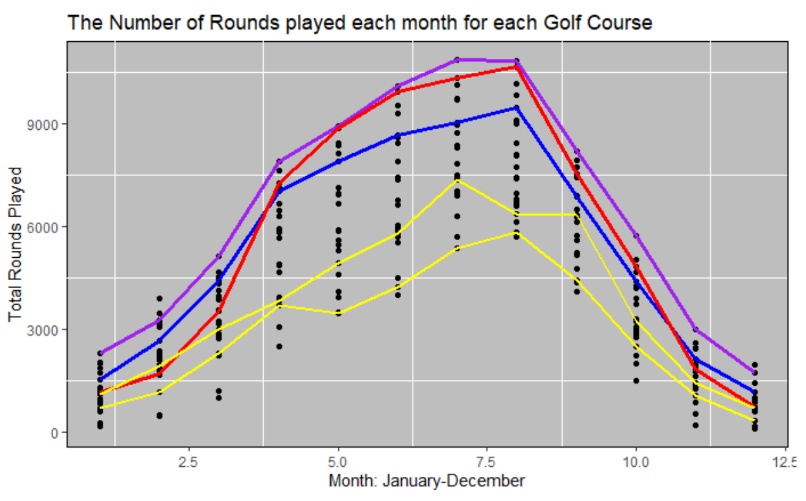
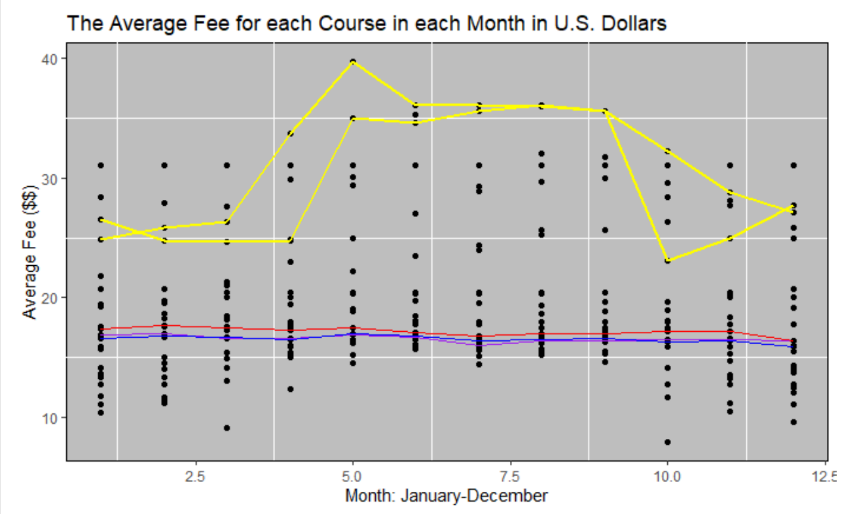
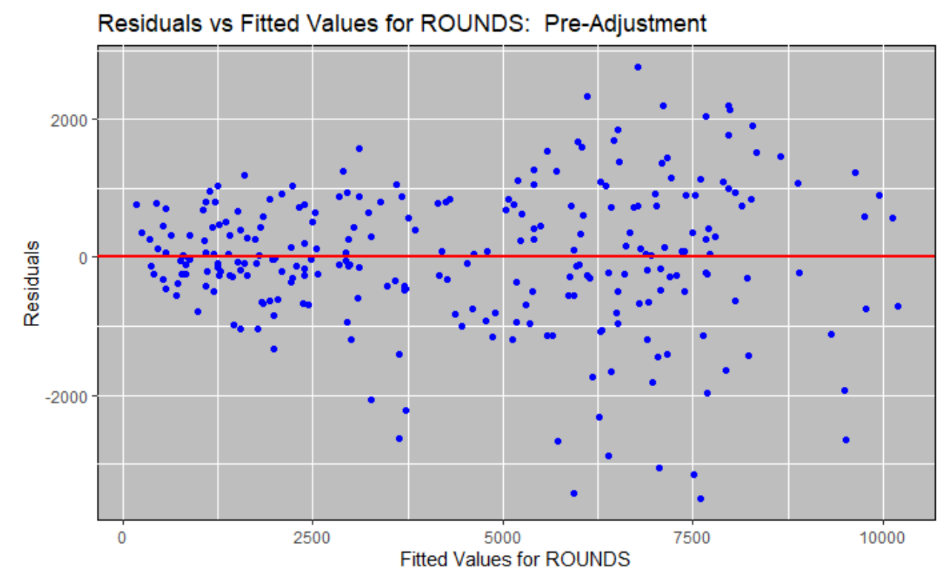
**Figures X, Y, Z:** F-Wald tests for *Winter\*Fee, Range, Yard* variables





**Figures cont’d:** Initial Plots





Citations:

1. Heteroskedasticity: <https://rpubs.com/cyobero/187387>
2. Auxillary Regression: <https://eeecon.uibk.ac.at/~zeileis/teaching/AER/Ch-Validation.pdf>
3. Regression Analysis: <http://web.utk.edu/~whwang/Regression.pdf>