Final Project Instructions

0. Find a dataset suitable for regression. It should have one designated response variable y and multiple predictor variables x. Describe the data and its variables and what is your research goal.
1. Perform a multiple regression with all predictors.
2. Perform model selection to decide if some predictors should be removed. Report appropriate p-values and conclusions based on them.
3. Try to include quadratic and interaction effects and again decide if they should be kept.
4. Perform model diagnostics and if needed remove some observations.
5. Report your final model together with R^2 and interpret the coefficients. Include CIs for each slope.
6. Depending on previous steps, you might have to fit other models: nonlinear, autocorrelation or dealing with heteroskedasticity.
7. Depending on your y variable you might have to fit a GLM model.
8. Bonus points: try shrinkage methods to see if there is an improvement.
9. If any changes are made after step 5 you need to report again the final model.
10. Summarize all your findings in a short paragraph written for the broad scientific community.