Stat 021 Final Project

Suzanne Thornton

Due: Friday, Dec. 13, 11:59pm

Resources

- http://r-pkgs.had.co.nz/ This is a free online textbook all about creating packages in R.

- https://hilaryparker.com/2014/04/29/writing-an-r-package-from-scratch/ This is an easy-to-follow step-by-step guide on writing your own R package.

- https://teachdatascience.com/datapackage/ This is another step-by-step guide that includes pictures of the process of creating an R package within R studio.

- https://www.dataquest.io/blog/career-guide-data-science-projects-portfolio/ While I don't agree with every opinion in this article, it does give you a sense of why I have assigned this final project.

- https://kbroman.org/pkg_primer/pages/data.html This is a blog post about how to include a dataset in a user-defined R package.

Your package must include:

- A unique data set (no two students should use the same data set, let me know if you need assistance finding a data set) with one numerical response variable and at least one numerical predictor variable and one categorical predictor variable;

- Functions that produce the following output and take your data set as the input

  1. a summary of the goodness of fit of the model that incorporates both the omnibus F-test and the adjusted R squared value;

  2. a summary of which predictor variable has the largest effect size ~~(in the same units as the data)~~;

  3. a plot of the residuals versus the predicted response values ~~where the size of the plotted points corresponds to the size of the residual~~;

  4. the model summary for a regression model fit to the standardized data;

  5. ~~the results for a test of heterscedasticity~~.