R Workshop
Synopsis
Instructors: Gustavo Robles and Jonathan Mummolo

Day 1
Instructor: Gustavo Robles

Morning Session: Introduction to R

Introduction to R
R Basics: Workspace, working directory, and scripts.
Creating and storing objects.
Vectors, matrices, strings & lists.
Functions.
R-Packages (code, functions, and datasets).

Afternoon Session: Reading and Manipulating Data

Reading data in different formats.
Merging and appending datasets.
Selecting and subsampling data.
Creating and transforming variables.

Day 2
Instructor: Gustavo Robles

Morning Session: Analyzing Data

Frequency and cross tables.
Summary statistics by groups.
Statistical tests.
Regression models.
Regression output, coefficients, and standard errors.
Predictions and marginal effects.

Afternoon Session: Reporting Results and Advanced Topics

Exporting tables.
Reporting regression output from R to different formats.
Basic programming skills: conditionals, loops, and recursion.
User-defined functions.
Running R in BATCH mode.

Day 3
Instructor: Jonathan Mummolo

Data Visualization in R:
The following lessons will be taught using a pre-supplied data set, most likely a random sample of the 2012 Cooperative Congressional Election Study. The goal is to use a real social science data set to explore methods of visual summarization. Particular emphasis will be placed on managing and displaying the results of statistical models after estimation.

Morning Session: Intro to "plot"

Basic inputs to the plot function
Scatter plots
Bar plots
Box Plots
Density Plots / Histograms
Using the par() function
    Combining multiple plots
    Using loops to generate large plot matrices
    Plotting more than one y-axis for a time series
Three-dimensional scatter plots (two predictors and an outcome)
Specifics covered throughout: labeling; custom axes; point types; color options; legends; custom text placement and fonts, including Greek symbols
Saving plots as .pdf, .jpg, etc.

Afternoon Session: Plotting After Estimation (and why you should pretty much never include a regression table in the main text of your paper ever again...)

Saving model parameters and organizing estimates for plotting
    Store output after lm() and predict() functions
    Store output after estimating new var-covar matrix (e.g. HC standard errors)
Plotting predicted values (from OLS, loess); overlaying onto raw data
Residual Plots (with brief statistical explanation)
Plotting coefficients with confidence intervals (tree plots)
Plotting continuous interaction effects (with brief statistical explanation)

*Please come prepared to follow along on your laptop.

R program download links:

R Software
http://www.r-project.org

R Studio
https://www.rstudio.com

Some useful books:

Maindonald, John, and W. John Braun (2010). "Data Analysis and Graphics Using R". Cambridge University Press.

Fox, John, and Sanford Weisberg (2011). "An R Companion to Applied Regression". SAGE Publications, Inc.