PUBP 514: Introduction to Computation in R

Professor Waggoner
Department of Government
441 Tyler Hall

www.philipdwaggoner.com
pdwaggoner@wm.edu
757-221-1399

Location: Tyler 217 **Day/Time**: T TR 8:00–9:20 am

1 Overview & Course Description

Welcome to Introduction to Computation in R! The use of R is rapidly increasing in all corners of data science and empirical research. This is for good reason as R is not only a fast and efficient programming language and environment for doing statistics and data analysis, but it is also *free* and open source. As such, this course will introduce students to the statistical computing language of R. We will cover a range of topics in base R as well as fold in the "tidy" approach to coding in R at certain points. The goals of the course are to get students comfortable engaging in basic coding in R, troubleshooting errors on their own, estimating widely used models, and transforming numerical output into visually pleasing figures and plots. As the course is geared toward beginners, no prior coding experience (in or out of R) is assumed. We will start at the ground level to ensure that everyone is at the same place from Day 1.

2 Text & Materials

Required:

1. Download the *free* statistical computing program, R:
   https://www.r-project.org/ (mirror doesn't really matter)
2. Download the *free* R complementary platform, R Studio:
   https://www.rstudio.com/products/rstudio/download/

Recommended:

1. Dalgaard, "Introductory Statistics Using R"
2. Grolemund and Wickham, "R for Data Science" (*free*: http://r4ds.had.co.nz/)
3. Monogan, "Political Analysis Using R"

3 Evaluation & Assessment

Students will be assessed by pass/fail. The idea behind this approach is to maximize the amount of learning and retention. Learning R is hard (and *rewarding*) enough on its own, without the threat of a grade looming overhead. As this is a short course, the key methods of evaluation will be active participation in class discussion and working examples, and consistent attendance. Meeting these requirements will ensure your success, not just in the narrow confines of a short semester course, but in retaining the concepts we learn in service of the broader goal of efficient programming and application in substantive research.

4 Outline of Topics & Calendar

* ***This outline is subject to change.***

- *Week 1, Thursday, 1/17*: Introduction to `R` & Syllabus
    1. **Reading**: Interview with Hadley Wickham: `https://tinyurl.com/y8vjq2zp`

- *Week 2, Tuesday, 1/22*: Getting Started with the Basics of `R` and R Studio

- *Week 2, Thursday, 1/24*: Data Importation & Packages

- *Week 3, Tuesday, 1/29*: Data Wrangling & Cleaning

- *Week 3, Thursday, 1/31*: Programming Basics

- *Week 4, Tuesday, 2/5*: Statistics: Simple Numerical Relationships

- *Week 4, Thursday, 2/7*: Statistics: OLS Regression (Bivariate and Multiple)

- *Week 5, Tuesday, 2/12*: Statistics: Binary Choice Models

- *Week 5, Thursday, 2/14*: Statistics: Event Count Models

- *Week 6, Tuesday, 2/19*: Data Visualization & Graphical Output: Base `R`

- *Week 6, Thursday, 2/21*: Data Visualization & Graphical Output: `ggplot2`

- *Week 7, Tuesday, 2/26*: **Class Canceled**; Emailed Script on Basic Webscraping

- *Week 7, Thursday, 2/28*: Basic Text Mining & Quantitative Text Analysis (Natural Language Processing)