PS218: Understanding Political Numbers

Fall 2015

Mr. Patrick Kearney
North Hall  411
University of Wisconsin, Madison
pmkearney@wisc.edu

Office Hours: Tuesdays 8:30-10:30am and by appointment.

TA: Michael Decrescenzo (North Hall 121)
Office Hours: Tuesday 2-3pm; Wednesday 10am-12pm

About the Course:

Formally, this course will teach you how to understand quantitative analysis, how to develop a question that you can answer, how to collect and analyze the data and how to present your conclusions to others.

There are essentially two components to the course. The first is the theoretical component. Why are numbers and data important to political analysis? What is the proper way to ask a research question? How do we find and use political data to tell a story and support an argument? The second component is technical. How do we employ statistical computing and graphics software to clearly present our data and findings to others?

Yet, in essence, this course is about more than just numbers. It is a skills and tools course. It is about learning how to ask (and answer) the right questions about politics. It is about presenting those answers clearly to an audience. And, it is about employing the computing tools to help get to this end goal. In so doing, it is about getting and “A” on every term paper you write from here on out.

Computing Requirements:

You need to have a computer you can use and on which you can install R and R Studio. Both are free and can be run on both Windows and Mac operating systems. They are available for download at www.r-project.org and www.rstudio.com respectively.

R is a free, open source statistical computing environment. R is extremely powerful and flexible. It is at the cutting edge of statistical software and it is the best producer of statistical graphics available. RStudio is an Integrated Development Environment (IDE) for R that combines a text editor, R, and other tools in a single program.

We recognize that R and RStudio take time to master and can be frustrating. That is why we’ll spend so much time during the course learning it. We DO NOT assume any prior knowledge of statistics or R. While we do expect some knowledge of basic computing tasks (creating folders, etc.), this course is designed to provide you all the
necessary instruction in the programs you will be using.

**Text:**

There is no *required* text for this course. In general, the lectures are self-contained and will allow students to succeed in the course.

That said, there are a number of “additional resources” that I will recommend here for students who wish to expand their understanding of the course material.

*R Graphs Cookbook* by Hrishi V. Mittal.

The best basic text on creating basic graphics in **R**.

*Visualize This* by Nathan Yau

This course draws significantly on Yau's ideas about data visualization.

*Getting Started with R Studio* by John Verzani

*Learning R Studio for R Statistical Computing* by van der Loo and de Jong

*A Beginner's Guide to R* by Zuur, Ieno and Meesters

Additionally, websites such as *flowingdata.com*, *fivethirtyeight.com*, and *nytimes.com/upshot* are excellent examples of data analysis and visualization.

Any substantive reading for the course will come in the form of short articles, book chapters and online readings. These will be posted on **Learn@UW** and/or sent to you via e-mail. Lectures, R Script files and other course related material will also be posted on **Learn@UW** throughout the semester.

There are a number of other texts and websites that I can recommend for the more advanced user. If you find yourself in this category, just ask.

**Components of the Grade:**

First, we will have a number of exercises designed to demonstrate you've learned the basic skills of programming in **R**. These will be brief, with the goal that you learn how to read data, or make a chart, or conduct an analysis. You will submit your R code and output but generally won't have to write much beyond that. These will be graded as 1 (satisfactory), .5 (turned in but incorrect or incomplete) and 0 (not turned in). Because we will be discussing solutions to these exercises in class the day they are due, late work **WILL NOT** be accepted.

Second, you will submit two more in depth analyses of topics that we cover in class. This will be in the form of short essays and several accompanying graphics that illustrate your analysis, created using **R**. The goal here is to develop your ability to “tell a story with data”, the fundamental objective of the class. These will be graded on the conventional A-F scale.

At the end of the semester, you will write a final paper, similar to the above but longer (12 pages, excluding graphics) and with more detailed analysis. For this
project you will choose the subject and be responsible for collecting the data. This essay will also be graded on an A-F scale.

Since this is the most important component of your grade for the course, we will ask you to submit 2 “drafts” in the final weeks of the semester. The first will be a short (2-3 pages) paper proposal in which you identify your research question and plan for your paper. The second will be the submission of the data that you will use in the analysis section of your final paper.

In addition to developing this paper you will also present your memo/proposal to the members of your discussion section in no more than 7 minutes. You will develop a presentation that relies on quantitative information and graphics. Each section will choose the best project from their section to advance to the finals. Finalists will present their projects to the entire class at the last class meeting. Section presentations begin the week of December 1 2015 (the week after Thanksgiving).

The winning presentation, as determined by a ballot of the class and the instructors, will earn an A for the course. All of the finalists are guaranteed at least an AB in the course. All members of the section from which the winning presentation is taken get 2% added to their final course grade.

***All assignments must be submitted online via the Dropbox section of the Learn@UW site. All documents must be submitted in .pdf format. All assignments involving computer coding should also include the .R script that generates all your output as a separate file.

All sources for data and other information used in your written work must be appropriately documented and sourced such that we can go and confirm every statement and number in the document.

There will be NO final exam.

**Grading:**

Your final grade will be computed based on

- Exercises 20%
- 2 short essays 20%
- 2 paper “drafts” 20%
- Final paper 40%

The standard UW Grading Scale will be used.
A Note on Attendance:

A number of sessions of this course will be conducted in more of a workshop format as opposed to a traditional lecture. Learning R is like learning a new language; it takes lots of hands on practice. As such, attendance is extraordinarily important. While I won't often take formal attendance, it is my experience that it is extraordinarily difficult to pass this course without regularly showing up. Additionally, since a large component of this course involves computing, it is perfectly acceptable to use laptops during lecture. That said, if your laptop is open I expect you to be working in R Studio not your social media accounts.

***Students affiliated with the McBurney Center should see me about any necessary accommodations.
Course Outline

Below you will find the tentative outline for the semester. Be aware that these topics and dates are subject to change. This is especially important since we will take this course at your pace (i.e. if we need to spend more time on certain skills, we will). As such, consult Learn@UW, your e-mail and/or in class announcements for up to date changes.

Thursday, September 3—NO CLASS; APSA Conference

Week 1
Tuesday, September 8—Intro and Syllabus
Thursday, September 10—Thinking about Political Numbers
   Readings--Intro from The Signal and the Noise by Silver
   Ch. 1&2 from Visualize This by Yau

Week 2
Tuesday, September 15—Thinking About Data Visualization
Thursday, September 17—Introduction to R and R Studio

Week 3
Tuesday, September 22—Elections and Numbers
   Online readings regarding 2012 Presidential Election
Thursday, September 24—Histograms in R
   **Exercise 1 distributed**

Week 4
Tuesday, September 29—More on Histograms, Bar and Box Plots in R
Thursday, October 1—World Health Data
   Reading---The Healing of America by Reid, Ch. 1&3

Week 5
Tuesday, October 6—Scatter Plots in R
   **Exercise 1 due**
Thursday, October 8—More on Scatter Plots
   **Exercise 2 distributed**

Week 6
Tuesday, October 13—Basic linear trend lines in R
Thursday, October 15—Alabama Election and Plotting Linear Models
   Readings—Articles on Alabama and Wisconsin Gubernatorial Elections
   **Exercise 3 distributed**
Week 7
Tuesday, October 20 — Data Collection and Management in R and Excel
**Exercise 2 due**
Thursday, October 22 — Linear Models and p-values

Week 8
Tuesday, October 27 — Non-Linear Relationships (Quadratic and Logarithmic)
**Essay 1 distributed**
Thursday, October 29 — More on Non-Linear Models
**Exercise 3 due**

Week 9
Tuesday, November 3 — Plotting Data Over Time in R
Thursday, November 5 — Black Death and Presidential Approval (More Time Series)
**Essay 1 due**

Week 10
Tuesday, November 10 — Basic Multivariate Linear Models in R
Thursday, November 12 — More with Basic Multivariate Models and Forecasting
Readings — Articles on Forecasting Elections
**Essay 2 distributed**

Week 11
Tuesday, November 17 — What are Good Research Questions??
Thursday, November 19 — Where do I find data??
**Essay 2 due**

Week 12
Tuesday, November 24 — Discuss Paper Topics
**Paper Proposal due**
Thursday, November 26 — NO CLASS, THANKSGIVING

Week 13
Tuesday, December 1 — Writing your paper...what we expect
Thursday, December 3 — Revisiting the Baseline
**Data Set due**

Week 15
Tuesday, December 8 — Advanced topics in R and Visualization
Thursday, December 10 — Advanced Topics in R and Visualization

Tuesday, December 15 — Final Paper Presentations
Wednesday, December 23 — **Final Paper Due**