PS218: Understanding Political Numbers

Fall 2016

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

Office Hours: Tuesdays 9:30-11:30am and by appointment.

TA: Michael Decrescenzo
Office Hours: North Hall 101A, Tuesday 1-3 and Thursday 10-12

About the Course:
Formally, this course will teach you how to understand quantitative analysis, how to develop a data-based research 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 a 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 you will develop the ability to write a data-focused research paper, an essential skill in your academic career and beyond.

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 either 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:**

In the past I have not assigned a required text for this course. In general, the lectures are self-contained and will allow students to succeed in the course. At the same time, some of your predecessors have asked for the additional resources that a required text provides. With that in mind, I suggest two *highly recommended* texts. Readings from these will be listed in the syllabus and will supplement the lectures. These are:

*R Graphs Cookbook* by Hrishi V. Mittal. The best basic text on creating basic graphics in R.
*R in a Nutshell: A Desktop Quick Reference* by Joseph Adler.

Specific Adler chapters are not currently listed in the syllabus. I will post these online as they become relevant.

Beyond these two texts, I will be assigning selected chapters and articles for you to read over the course of the semester. These are listed in the syllabus and will be posted on Learn@UW. There is no *required* text for this course. I understand that for some the lack of a text to follow along with and for those students there are a number of “additional resources” that I can recommend here for anyone who prefers a text. The following are “manuals” of sorts to help you learn R and R Studio.

*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.

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 and codebook 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 November 30, with sign up online the week before. 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 percentage points added to their final course grade (you must be present for voting to receive this extra credit).

***All assignments must be submitted online via the Dropbox section of the Learn@UW site. All documents should 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 15%, 2 short essays 15%, 2 paper
“drafts” 20%, Presentation 10% and 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.

Week 1
Wednesday, September 7—Intro and Syllabus

Week 2
Monday, September 12—Thinking about Political Numbers
   Readings--Intro from The Signal and the Noise by Silver and Ch. 1&2 from Visualize This by Yau
   Wednesday, September 14—Thinking About Data Visualization

Week 3
Monday, September 19—Introduction to R and R Studio
   Readings from Adler (posted online)
   Wednesday, September 21—Elections and Numbers
      Online readings regarding 2012 Presidential Election

Week 4
Monday, September 26—Histograms in R
   Mital Chapters 6 and 2
   **Exercise 1 distributed**
   Wednesday, September 28—More on Histograms, Bar and Box Plots in R
      Mital Chapter 7

Week 5
Monday, October 3—World Health Data
   Reading---The Healing of America by Reid, Ch. 1&3
   Wednesday, October 5—Scatter Plots in R
      **Exercise 1 due**
      Mital Chapter 3
Week 6
Monday, October 10 — More on Scatter Plots
   **Exercise 2 distributed**
Wednesday, October 12 — Basic linear trend lines in R
   Readings from Adler (posted online)

Week 7
Monday, October 17 — Alabama Election and Plotting Linear Models
   Readings—Articles on Alabama and Wisconsin Gubernatorial Elections
   **Exercise 3 distributed**
Wednesday, October 19 — Data Collection and Management in R and Excel
   **Exercise 2 due**

Week 8
Monday, October 24 — Interpreting Linear Models and p-values
   Readings from Ellenberg (posted online)
Wednesday, October 26 — Non-Linear Relationships (Logarithmic)
   **Exercise 3 due**
   **Essay 1 distributed**

Week 9
Monday, October 31 — What are Good Research Questions??
Wednesday, November 2 — Where do I find data??
   **Essay 1 Due**

Week 10
Monday, November 7 — Basic Multivariate Linear Models in R
Wednesday, November 9 — More with Basic Multivariate Models and Forecasting
   Readings—Articles on Forecasting Elections
   **Essay 2 distributed**
Week 11
Monday, November 14—Plotting Data Over Time in R
Wednesday, November 16—Black Death and Presidential Approval (More Time Series)
  **Essay 2 due**

Week 12
Monday, November 21—Discuss Paper Topics
Wednesday, November 23—NO CLASS DUE TO THANKSGIVING
  **Paper Proposal due Wednesday 11/23 by 5pm**

Week 13
Monday, November 28—Writing your paper and presentations...a full outline of what we expect
  **Data Set and Codebook due Sunday 12/4 by 5pm**
Wednesday, November 30—Advanced topics in R and Visualization—the Challenger

  **Presentations in Section Begin Wednesday 11/30**

Week 14
Monday, December 5—Advanced topics in R and Visualization Wednesday, December 7—Political Numbers in the News

Week 15
Monday, December 12—The Future of Political Numbers
Wednesday, December 14—Final Paper Presentations

Wednesday, December 21—**Final Paper Due** 11:59pm