



COLLEGE OF SOCIAL AND BEHAVIORAL SCIENCES

School of Government and Public Policy

POL 596A: Programming for Statistical Analysis in R

Day/Time: Wed 2:00-4:00 PM

Room: Social Sciences Rm 311

Office Hours: Mon. 12-2 PM

(or by appointment, or anytime I am in my office)

Professor: Frank Gonzalez

Office: 337 Social Sciences Bldg.

Email: fgonzo@email.arizona.edu

Semester: Fall 2018

Course Description

This course is a companion workshop to POL 681, which you should already be enrolled in. As you embark on learning about more applied statistics in POL 681, this workshop will provide you with the basic tools needed to conduct statistical analyses using the magic of computing, which is generally preferred over the noble yet time-consuming task of doing analyses by hand (although yes, you will still need to do some analyses by hand before the year is over :-)). Specifically, you will be introduced to the R Statistical Computing Environment, a free, open-source implementation/extension of the S programming language, and shown how to use R to do statistical analyses such as those you are learning in POL 681.

Besides just being a companion workshop to POL 681, this course also introduces you to what may very well be a primary workhorse of your research over the coming years (at least for many of you). Someone else has probably already broken the news to you that there is no escaping math even in Political Science. What they might not have yet mentioned is that there is no escaping computer programming either. I know it is daunting, and to some degree that is well-justified. When I say “introduced,” it is literal, as there are many many things you can (and will need to) do with R beyond what is covered in this course. Many/most of the things you wind up doing in your own research (if you continue to do quantitative research) will only build on the fundamentals you learn here. You will spend time reading online discussion forums about how to do specific, more complex things. **But trust me**, as someone who was totally averse to such things early on, you can learn to have a lot of fun with it. Plus, it will take you a long way.

In this course, we will cover everything from answering “what is R?” (although I just sort of told you), to downloading R, loading data sets, manipulating data, doing basic statistical analyses, and even a bit of programming, as time permits. Why R, you ask? We will talk more about this in class, but basically, although other paid-for software is often more user-friendly and offers “canned” functions for specific analyses, R a) is free, b) is free, c) is programmable (meaning you are never limited by what the software developers have created - see Frank for “Matrix” metaphor), d) is possibly the most widely used statistical software in the world, with statisticians and social scientists constantly writing new programs to fit your specific needs! (roughly 12,000 add-on packages as of 2018), and e) forces you to learn “what is going on under the hood” of statistical analyses and data management in a more intimate way than other software does. Oh, and f) it’s FREE.

Course Objectives and Expected Learning Outcomes

1. Develop a broad understanding of the logic of programming and statistical analysis in R.
2. Learn how to load and manipulate different types of data in R.
3. Conduct basic statistical analyses in R.
4. Develop the basic skills necessary for learning more advanced programming and analysis in R.

Course Format and Expectations

This class will be similar to your statistics courses in that it will involve more lecturing by the instructor than discussion. However, questions, comments, and discussion during class are not only encouraged but also expected. Further, the course will often be interactive.

Assignments will be given on a weekly basis. They are not graded, but will be turned in each week on D2L. The instructor will provide feedback that should be used to revise assignments. Although assignments are not graded, the tasks you will do for them are fundamental for completing the final project (and likely the final project in POL 681 and your other stats courses). Collaboration with classmates is encouraged.

If there are any readings for a particular week, students are expected to have done all readings for the week *prior* to class, and to come to class prepared to reference the readings.

The final product for the course will be a large assignment (graded this time) that involves each of the tasks required in the previous assignments on a data set of your choosing (see instructor for data set recommendations).

Required Texts

- Monogan, James E., III (2015). *Political Analysis Using R*. Springer International Publishing.
- Available *for free* here (use JSTOR or purchase on campus):
<https://www.springer.com/us/book/9783319234458>

Other Recommended Texts

- Dalgaard, Peter. (2008). *Introductory Statistics with R*. Springer-Verlag New York.
- Fox, John, & Weisberg, Sanford. (2010). *An R Companion to Applied Regression* (2nd Edition). Sage Publications.
- Spector, Phil. (2008). *Data Manipulation with R*. Springer-Verlag New York.
- Also, www.rseek.org is Google (indeed power by Google) for R only.

Grading and Assignments

- Participation and Attendance - 10% each for a total of 20%: Attendance is required, and participation is as well when called for. Students are allowed **1** unexcused absence, but are penalized for each additional absence. Students who miss class due to illness or emergency are required to bring documentation from their health-care provider or

other relevant, professional third parties. Failure to submit third-party documentation will result in unexcused absences.

The UAs policy concerning Class Attendance, Participation, and Administrative Drops is available at: <http://catalog.arizona.edu/policy/class-attendance-participation-and-administrative-drop>. The UA policy regarding absences for any sincerely held religious belief, observance or practice will be accommodated where reasonable, <http://policy.arizona.edu/human-resources/religious-accommodation-policy>. Absences pre-approved by the UA Dean of Students (or Dean Designee) will be honored. See: <https://deanofstudents.arizona.edu/absences>.

- **Final Assignment - 80%:** The large majority of your grade will come directly from the final assignment, as this assignment will be the culmination of what you have done in the preceding (ungraded, but still required) assignments. Don't panic! You will essentially be working on this throughout the entire course, and so you should be working *with me* each week to make sure you understand and can perform each of the tasks necessary for the final assignment. If you stay on top of the assignments and material covered, it should be easy money. That said, it is your responsibility to make sure you stay up-to-date with the assignments and material. If you start to fall behind, let me know as soon as possible.

The final assignment will be made available midway through the course, but again, it is essentially just the culmination of all other assignments applied to a data set of your choosing.

The final assignment will be due by 2 PM on Wednesday of the last week of classes.

Grade Scale

Letter Grade	%
A	90-100
B	80-99.9
C	70-79.9
D	60-69.9
E	59.9 and below

- **Requests for incomplete (I) or withdrawal (W)** must be made in accordance with University policies, which are available at <http://catalog.arizona.edu/policy/grades-and-grading-system#incomplete> and <http://catalog.arizona.edu/policy/grades-and-grading-system#Withdrawal> respectively.
- Should you wish to discuss your grade, you must meet with me in person. I am not able to discuss grades by email.

Academic Integrity

Academic dishonesty is taken *extremely* seriously in this course. **Anyone found guilty of fabrication, falsification, or plagiarism will, at a minimum, receive an F for this course, and will be referred to university authorities.** Fabrication means altering information dishonestly, falsification means inventing information dishonestly, and plagiarism

means presenting someone else's work as your own, either by not giving proper acknowledgement of the source or by presenting in whole or in part any work that has been prepared by or copied from another person.

Students are encouraged to share intellectual views and discuss freely the principles and applications of course materials. However, graded work/exercises must be the product of independent effort unless otherwise instructed. Students are expected to adhere to the UA Code of Academic Integrity as described in the UA General Catalog. See: <http://deanofstudents.arizona.edu/academic-integrity/students/academic-integrity>.

The University Libraries have some excellent tips for avoiding plagiarism, available at <http://www.library.arizona.edu/help/tutorials/plagiarism/index.html>.

Selling class notes and/or other course materials to other students or to a third party for resale is not permitted without the instructors express written consent. Violations to this and other course rules are subject to the Code of Academic Integrity and may result in course sanctions. Additionally, students who use D2L or UA e-mail to sell or buy these copyrighted materials are subject to Code of Conduct Violations for misuse of student e-mail addresses. This conduct may also constitute copyright infringement.

Class Conduct

This course will consist of a substantial amount of discussion and debate, and I look forward to class discussions that are energized and thought-provoking, but this is not an excuse for incivility in the classroom. Your fellow classmates will undoubtedly often hold different opinions than your own, and I expect you to handle this disagreement with civility and respect for differing viewpoints. To foster a positive learning environment, students and instructors have a shared responsibility to do so. We want a safe, welcoming, and inclusive environment where all of us feel comfortable with each other and where we can challenge ourselves to succeed. Failure to show such respect will reflect in your final grade (specifically, points will be deducted from your participation grade).

It is also worth noting that laptops are acceptable in the classroom. However, they should be used strictly for note-taking.

Threatening Behavior Policy: The UA Threatening Behavior by Students Policy prohibits threats of physical harm to any member of the University community, including to oneself. See <http://policy.arizona.edu/education-and-student-affairs/threatening-behavior-students>.

UA Nondiscrimination and Anti-harassment Policy

The University is committed to creating and maintaining an environment free of discrimination; see <http://policy.arizona.edu/human-resources/nondiscrimination-and-anti-harassment-policy>.

Incompletes and Late Papers

Course incompletes will not be given except in extreme situations where all of the appropriate documentation is provided. Late assignments will be docked one letter grade if not handed in at the beginning of class on the due date (i.e., assignments handed in during the middle or at the end of class will already be docked one letter grade). Grades on the assignment will be lowered one additional letter grade for each additional day late.

Additional Resources for Students

UA Academic policies and procedures are available at <http://catalog.arizona.edu/policies>. Student Assistance and Advocacy information is available at <http://deanofstudents.arizona.edu/student-assistance/students/student-assistance>.

Confidentiality of Student Records

For more information on the confidentiality of student records at UA, please see: <http://www.registrar.arizona.edu/personal-information/family-educational-rights-and-privacy-act-1974-ferpa?topic=ferpa>.

Accessibility and Accommodations

Our goal in this classroom is that learning experiences be as accessible as possible. If you anticipate or experience physical or academic barriers based on disability, please let me know immediately so that we can discuss options. You are also welcome to contact the Disability Resource Center (520-621-3268) to establish reasonable accommodations. For additional information on the Disability Resource Center and reasonable accommodations, please visit <http://drc.arizona.edu>.

If you have reasonable accommodations, please plan to meet with me by appointment or during office hours to discuss accommodations and how my course requirements and activities may impact your ability to fully participate.

Please be aware that the accessible table and chairs in this room should remain available for students who find that standard classroom seating is not usable.

Subject to Change Statement

Information contained in the course syllabus, other than the grade and absence policy, may be subject to change with advance notice, as deemed appropriate by the instructor.

Course Schedule and Readings

All readings should be done *prior* to the date for which the readings are assigned.

<u>Date</u>	<u>Topics and Readings</u>
Oct. 11	What is R? Opening it and doing stuff <u>Reading: Ren Ch. 1 (on D2L); Monogan Ch. 1</u> Downloading, installing, running R Understanding R as a: calculator; programming language; computing environment Installing and loading “packages” Preview of what is possible <i>Frank Out of Town on Wednesday; Rescheduled for earlier - see email!</i>
Oct. 17	Loading Data/Basic Descriptives/Univariate Visualizations <u>Reading: Monogan Ch. 2, 3, & 4</u> Loading different types of data Common problems loading data “Looking” at your data/quality checks Basic descriptive analyses
Oct. 24	Vectors/Matrices/Arrays/Dataframes <u>Reading: Monogan Ch. 10</u> Creating and manipulating vectors/matrices/arrays/dataframes Some (light) matrix algebra
Oct. 31	Bivariate Analyses <u>Reading: Monogan Ch. 5</u> Some more matrix algebra Bivariate analyses (correlation, t-test, anova, etc.)
Nov. 7	Logical Expressions/Loops/Functions <u>Reading: Fox pp. 359-385 (on D2L); Ren Ch. 4 & 5 (on D2L); Monogan 11.2, 11.3, & 11.4</u> loops functions <i>Final Assignment Assigned</i>
Nov. 7	Logical Expressions/Loops/Functions (cont.) <u>Reading: see last week</u> Continuation of last week
Nov. 21	Linear Regression <u>Reading: Fox pp. 149-183 (on D2L); Monogan Ch. 6</u> Overview of linear regression Linear regression “by hand” (sort of) Linear regression using packages Obtaining/interpreting results
Nov. 28	Visualizing Results <u>Reading: no reading</u> Options: basic, lattice, ggplot Univariate

Date

Topics and Readings

Bi-variate
Multivariate
Interactions (continuous/categorical/interplot)
Combining plots
3D plots
4D plots
“Smell-o-vision” plots

Dec. 5

Last Week of Classes - Loose Ends

This week dedicated to tying loose ends

no assignments this week

Dec. 12

Final Assignments Due at 2 PM - NO EXCEPTIONS!
