POLS 671: Introduction to Data Analysis

Monday 5:30–8:15 p.m.
407 Ford Hall

Professor: Adam Enders
Office: 105 Ford Hall
Office Hours: Monday 4:00–5:15 p.m. (or by appointment)
Email: amende01@louisville.edu

Course Description

> *"The best thing about being a statistician is that you get to play in everybody's backyard."* –John W. Tukey

The goal of this course is to introduce graduate students to basic statistics and statistical models that are frequently employed to answer questions of importance to social scientists. As the only course in the graduate program dedicated to statistical inference, this course will cover a great deal of material, from basic descriptive statistics up to ordinary least squares regression, the workhorse of multivariate statistics upon which nearly all more "sophisticated" models are built. By the end of the semester, students should be able to manipulate and describe data using statistical software, produce powerful statistical graphics for visualizing salient features of data and presentation to readers, and use statistical methods to test theories about political phenomena.

Unlike with courses designed around particular substantive topics, class meetings will be heavily lecture-based, punctuated with demonstrations of the calculation and interpretation of particular quantities of interest using statistical software. Of course, students are encouraged to use class time to ask questions about course material. However, students should attempt to troubleshoot specific software issues beforehand. Furthermore, students should expect homework assignments every week, which will be due in class the following week.

Class Meetings

This course is designated as "hybrid," meaning that at least 25% of instruction is to be delivered face-to-face. For the time being, we will be meeting in person at the designated meeting time. Minimally, this will allow us to build some rapport with one another in the event that all courses are forced online. If and when we meet virtually, we will continue to meet at the designating meeting times using Blackboard Collaborate. Students can access online meetings using the following steps:

1. Open the Blackboard course
2. Select **Tools** in the left hand menu
3. Select either of the **Blackboard Collaborate Ultra** links

4. Select the session to open
5. Select **Join Session**

Course Materials

We will primarily be using the book listed below. This is the only text that students are required to have regular access to.

> Agresti, Alan. 2017. *Statistical Methods for the Social Sciences (5^{th} edition).* Pearson. ISBN: 978-0134507101.

Note that some weeks the textbook information will be supplemented by additional journal articles or book chapters. All of this supplementary is available via the course Blackboard page.

Other texts that might prove useful to students who are looking to supplement the presentation of material in the Agresti text with other presentations of the same material might consider the following (both of which should be available at the library):

> Monogan, James E. III. 2015. *Political Analysis Using R.* Springer. ISBN: 978-3319234458.

> Fox, John, and Sanford Weisberg. 2010. *An R Companion to Applied Regression (2^{nd} edition).* SAGE. ISBN: 978-1412975148.

Software

We will be using the R statistical computing environment to carry out statistical analyses and produce graphical depictions of data. Though not the most user friendly option for data analysis, R is increasingly being employed in both academia and the private sector because it is A) extremely powerful and flexible, and B) open source (or, free!). Exposure to some basic programming can never hurt. Thus, learning to do some basic data analysis in R is a transferable skill that will almost certainly benefit students in some way or another regardless of career track.

To install R on Windows or Mac, follow these instructions:

1. Go to `https://cran.r-project.org/`
2. At the top of the page, click the "Download R" link that corresponds with the operating system of the machine you're working on (i.e., Mac or Windows)
    - For Windows: click "install R for the first time" and then "Download R 3.6.2 for Windows" on the following page

- For Mac: click "R-3.6.2.pkg" if you're running OS X 10.11 (El Capitan) or later; if not, scroll down and click "R-3.3.3.pkg"

3. After binaries have downloaded, install as you would any other application

RStudio is a slightly more user friendly "integrated development environment" for R. Basically, it makes interacting with R a little easier for those new to programming. Follow these instructions to install RStudio (note that you must already have R installed – per instructions above – in order to install RStudio):

1. Go to `http://www.rstudio.com/ide/download/`
2. Scroll down and click "Download RStudio Desktop"
3. Scroll down to the "All Installers" list
4. Click "RStudio-1.2.5033.exe" if you're running Windows, or "RStudio-1.2.5033.dmg" if you're running Mac
5. After RStudio has downloaded, install as you would any other application

Finally, students should make use of excellent online resources for troubleshooting any issues they may encounter with R:

- `Stack Exchange`
- `UCLA Institute for Digital Research and Education`
- `CRAN Task Views`
- `Quick-R`
- To understand R functions one is already aware of, simply submit "?`FUNCTION`" to R
  - This `website` explains getting help within R at greater length

Course Requirements

Participation: While attendance will not be formally recorded, systematic absences and lack of participation will be obvious in such a small class, and those absences will be reflected in the participation component of the student's final grade.

Assignments: Assignments are designed to test students' comprehension of course material. They will be regularly assigned over the course of the semester. Students should expect one assignment per week.

Final Report: Students must complete a final report where they test a theory they are interested in using their own dataset by employing the statistical and data analytic techniques they learned about in class. Students will present preliminary results during the last class meeting. More details will be revealed in class.

Grades

Distribution

Participation	10%
Assignments	50%
Final Report	40%

Scale

94-100 = A	75-78 = C+
90-93 = A-	71-74 = C
86-89 = B+	68-70 = C-
82-85 = B	60-67 = D
79-81 = B-	60 and below = F

Class Schedule

All of the following reading assignments are to be completed *before* attending class on the associated date/week. The material not included in the main text (journal articles and book chapters) are available for download on the course Blackboard page.

Week 1: Introductory Material (January 11)

Agresti, Chapters 1-2

Week 2: No class (January 18)

MLK day

Agresti, Chapters 1-2

Week 3: Distributions, Descriptive Statistics, & R(Studio) (January 25)

Agresti, Chapters 1-3

"Introduction to R" manual

"R Reference Sheet"

Week 4: Visualizing Data & More R(Studio) (February 1)

Cleveland, William S. 1994. *The Elements of Graphing Data*. Chapter 4.

Week 5: Probability & Sampling Distributions, & the CLT (February 8)

Agresti, Chapter 4

Week 6: Interval Estimation (February 15)

Agresti, Chapters 5 (less emphasis on pages 121-130)

Week 7: Hypothesis Testing (February 22)

Agresti, Chapter 6

Gill, Jeff. 1999. "The Insignificance of Null Hypothesis Significance Testing." *Political Research Quarterly* 52(3): 647-674

Ioannidis, John P. A. 2005. "Why Most Published Research Findings Are False." *PLoS Medicine* 2(8): e124.

Paper proposal due

Week 8: Comparison of Groups (March 1)

Agresti, Chapter 7 (less emphasis on pages 193-203)

Week 9: Cross-Tabulations & Correlation (March 8)

Agresti, Chapter 8 (pages 215-225 only) & Chapter 9 (pages 247-262 only)

Week 10: Bivariate Regression (March 15)

Agresti, Chapters 9 (finish) & 10 (very brief)

(Optional) For more help with regression:

Lewis-Beck, Colin, and Michael S. Lewis-Beck. 2015. *Applied Regression: An Introduction*. SAGE. ISBN: 978-1483381473

Berry, William D., and Stanley Feldman. 1985. *Multiple Regression in Practice*. SAGE. ISBN: 978-0803920545.

Week 11: Multiple Regression (March 22)

Agresti, Chapters 10 & 11

Week 12: Categorical Variables & Modeling (March 29)

Agresti, 13 (up to page 404) & 14

Week 13: Interactions & Practice (April 5)

Agresti, 13 (up to page 404) & 14

Paper draft due

Week 14: Student Presentations (April 12)

No readings

Week 15: Free time (April 19)

Work on final paper!

Final Report due Friday, April 23 @ 5:00 PM

Course Policies

Title IX/Clery Act Notification: Sexual misconduct (including sexual harassment, sexual assault, and any other nonconsensual behavior of a sexual nature) and sex discrimination violate University policies. Students experiencing such behavior may obtain **confidential** support from the PEACC Program (852-2663), Counseling Center (852-6585), and Campus Health Services (852-6479). To report sexual misconduct or sex discrimination, contact the Dean of Students (852-5787) or University of Louisville Police (852-6111).

Disclosure to **University faculty or instructors** of sexual misconduct, domestic violence, dating violence, or sex discrimination occurring on campus, in a University-sponsored program, or involving a campus visitor or University student or employee (whether current or former) is **not confidential** under Title IX. Faculty and instructors must forward such reports, including names and circumstances, to the UniversityÕs Title IX officer.

For more information, see the `Sexual Misconduct Resource Guide`.

Academic Integrity: Academic dishonesty is prohibited at the University of Louisville. It is a serious offense because it diminishes the quality of scholarship, makes accurate evaluation of student progress impossible, and defrauds those in society who must ultimately depend upon the knowledge and integrity of the institution and its students and faculty. For more information, see the `Code of Student Rights and Responsibilities`.

Disabilities Accommodation: The University of Louisville is committed to providing access to programs and services for qualified students with disabilities. If you are a student with a disability and require accommodation to participate and complete requirements for this class, notify me immediately and contact the Disability Resource Center (Stevenson Hall, 502-852-6938) for verification of eligibility and determination of specific accommodations. For more information, visit the `Disability Resource Center`.

Religious Observation Policy: Federal law and university policy prohibit discrimination on the basis of religious belief. It is the policy of the University of Louisville to accommodate students, faculty, and staff who observe religious work-restricted holy days. Students who need to alter a quiz/exam deadline due to religious observations must, however, bring the

matter to the attention of the instructor during the first week of the semester.

Final Caveat: While this course has been devised carefully, the instructor does reserve the right to amend the assignments or schedule as presented above.