Public Policy 529.004: Statistics
Fall 2023 Syllabus

Instructor: Jonathan Hanson (he/him)
jkhanson@umich.edu
4223 Weill Hall, 734-615-1496
Office Hours: Tues. 1:00–2:30, Fri. 1:00–2:00, or by appointment

Graduate Student Instructor: Dominique Beata (she/they)
dbaeta@umich.edu
3202 Weill Hall, 734-615-4245
Office Hours: Tues. 10:00–11:30 or Thurs. 1:00–2:30, or by appointment

The objective of this course is to provide a foundation in statistics and their application to questions in public policy and social science research. Key topics include measurement, data collection, descriptive statistics, probability theory, statistical inference for different types of data, and significance testing. As the augmented version of Public Policy 529, this course is aimed toward students who foresee themselves as practitioners of data analysis in their future careers. Compared with the standard version, there are higher expectations for understanding statistical theory and a greater reliance on math for learning course concepts.

By nature of the material, this course is difficult. There will be times when you feel that you are just not getting it, and this is normal. Know this ahead of time, and keep working hard. There are no short cuts. Know also that I am on your side. My only goal is to help you learn, and I will push you because that is what I need to do in order for you to learn. The rewards will come if you persist!

Class Meeting Schedule

Unless otherwise noted, lectures are Mondays and Wednesdays from 1:00–2:20 pm in 1110 Weill Hall. Weekly section meetings are held on Fridays from 11:30–12:50 pm in 1110 Weill Hall.

Textbooks

Statistics textbooks have different strengths and weaknesses, and no single book will work well for everyone. It is useful, however, to have one text be a focal point for presentation of the material. The following book will serve in that role for this course:

- Alan Agresti, *Statistical Methods for the Social Sciences*, 5th edition, (Pearson: 2017).

This book is available for purchase, though I recommend that you look around for various options to access the book. Rentals on amazon.com, for example, are considerably cheaper than purchasing the text.

Although I strongly recommend you purchase or rent the Agresti text, alternatives issued under Creative Commons licenses have become increasingly viable in recent years, and I encourage you to utilize these resources as well. Two such books follow:

- David M. Diez, Christopher D. Barr, and Mine Çetinkaya-Rundel, *OpenIntro Statistics*, 4th edition, 2019. Available for download at https://www.openintro.org/stat/textbook.php.
- Barbara Illowsky and Susan Dean, *Introductory Statistics*, 2018. Available for download at https://tinyurl.com/y3s84pej.

Both of these books have good features. The relevant sections from these texts are listed on the syllabus along with those of the Agresti text so that you can refer to these books whenever you find it helpful. In general, the *Introductory Statistics* book organizes the material in a manner that is more similar to this course than *OpenIntro Statistics*.

Other reading selections will be made available on the Canvas site for the course. You can log into Canvas at http://canvas.umich.edu.

Assignments and Grading

Your grade for this course will be determined by the following:

Problem sets	30%
Quizzes	20%
Midterm exam	25%
Final exam	25%

By nature, this material is cumulative and you will become stronger with practice. Problem sets are thus assigned on a regular basis. You are encouraged to collaborate with other students to figure out how to answer questions on the problem sets. It is essential, however, that you write up all of your answers independently and in your own words. The ability to produce the answer yourself is a marker for your learning. Credit will not be given if it is determined that answers were not written independently.

The second portion of the course grade will come from two in-class quizzes worth 10% each. These quizzes, which will be given on September 27 and November 13, are designed to ensure that you are keeping up with the material between the two major exams for the course.

The midterm exam will be given on Wednesday, October 18, and the final exam will take place on Wednesday, December 13. Each of these exams is worth 25%.

The final course letter grade reflects the Ford School's guidelines. An A is awarded for work that is Excellent, an A- for work that is Very Good, a B+ for work that is Good, a B for work that is Acceptable, and a B- for work that is below expectations for graduate work. You should know I do not have a predetermined formula to convert numeric point totals into these categories. It would be a mistake, for instance, to assume that a grade of 75% on an exam translates into a C, since exams vary in their difficulty. The median grade for this course is typically in the A-/B+ range.

Software

Students will use either the Stata or R statistical packages for homework assignments. Both applications are very powerful, and they are available on computers in the Ford School computer lab and all of the larger computer labs on campus. Additionally, students can remotely log into the university's Virtual Sites (see information at https://its.umich.edu/computing/computers-software/virtual-sites) to access Stata or R when not on campus.

Of course, it may be more convenient to install one of these applications on a personal computer. R is an open-source program that is freely downloadable from https://cran.r-project.org. Students who use R are strongly encouraged to download the free RStudio Desktop companion application (https://rstudio.com/products/rstudio/download/) to serve as their interface with R. Stata is a commercial product. Time-limited licenses for Stata are available for purchase at student pricing from https://www.stata.com/order/new/edu/profplus/student-pricing/. The Stata/BE version is sufficient for PubPol 529 and PubPol 639.

The learning curve for R is steeper, but it offers a lot of flexibility and will remain available to you at no cost. Stata provides an easier path for getting started and has excellent documentation, but maintaining a license to the latest version can be costly.

The Friday section will include help with the statistical computing skills required to complete course assignments, and the Canvas site provides many resources. The Files section has folders with handouts for both software packages, the Media Gallery contains tutorial videos, and the Pages section has several links to free resources on the internet. Students who want additional training in using software for statistical work are encouraged to take one of the several courses offered in Ford that focus on learning software.

Additionally, Alton Worthington, a lecturer here at the Ford School who teaches several software-related courses, holds general office hours every Wednesday from 10:00-4:00. He is happy to help anyone with coding-related questions. You can sign up for these office hours at http://bit.ly/ABHW-OO. Meetings can be in his office (3216 Weill Hall) or via Zoom.

Academic Integrity

The Ford School academic community, like all communities, functions best when its members treat one another with honesty, fairness, respect, and trust. We hold all members of our community to high standards of scholarship and integrity. To accomplish its mission of providing an optimal educational environment and developing leaders of society, the Ford School promotes the assumption of personal responsibility and integrity and prohibits all forms of academic dishonesty, plagiarism and misconduct. Academic dishonesty may be understood as any action or attempted action that may result in creating an unfair academic advantage for oneself or an unfair academic advantage or disadvantage for any other member or members of the academic community. Plagiarism involves representing the words, ideas, or work of others as one's own in writing or presentations, and failing to give full and proper credit to the original source. Conduct, without regard to motive, that violates the academic integrity and ethical standards will result in serious consequences and disciplinary action. The Ford School's policy of academic integrity can be found in the MPP/MPA, BA, and PhD Program handbooks. Additional information regarding academic dishonesty, plagiarism and misconduct and their consequences is available here.

Inclusivity

Members of the Ford School community represent a rich variety of backgrounds and perspectives. We are committed to providing an atmosphere for learning that respects diversity. While working together to build this community we ask all members to:

- share their unique experiences, values and beliefs
- be open to the views of others
- honor the uniqueness of their colleagues
- appreciate the opportunity that we have to learn from each other in this community
- value one another's opinions and communicate in a respectful manner
- keep confidential discussions that the community has of a personal (or professional) nature
- use this opportunity together to discuss ways in which we can create an inclusive environment in Ford classes and across the UM community

Accommodations for Students with Disabilities

The University of Michigan recognizes disability as an integral part of diversity and is committed to creating an inclusive and equitable educational environment for students with disabilities. Students who are experiencing a disability-related barrier should contact Services for Students with Disabilities (SSD); 734-763-3000 or ssdoffice@umich.edu). For students who are connected with SSD, accommodation requests can be made in Accommodate. If you have any questions or concerns please contact your SSD Coordinator or visit SSD's Current Student webpage. SSD considers aspects of the course design, course learning objects and the individual academic and course barriers experienced by the student. Further conversation with SSD, instructors, and the student may be warranted to ensure an accessible course experience.

Student Mental Health and Wellbeing

The University of Michigan is committed to advancing the mental health and wellbeing of its students. We acknowledge that a variety of issues, both those relating to the pandemic and other issues such as strained relationships, increased anxiety, alcohol/drug problems, and depression, can directly impact students' academic performance and overall wellbeing. If you or someone you know is feeling overwhelmed, depressed, and/or in need of support, services are available.

You may access the Ford School's embedded counselor Paige Ziegler (pziegler@umich.edu) and/or counselors and urgent services at Counseling and Psychological Services (CAPS) and/or University Health Service (UHS). Students may also use the Crisis Text Line (text '4UMICH' to 741741) to be connected to a trained crisis volunteer. You can find additional resources both on and off campus through the University Health Service and through CAPS.

Ford School Public Health Protection Policy

In order to participate in any in-person aspects of this course–including meeting with other students to study or work on a team project–you must follow all the public health safety measures

and policies put in place by the State of Michigan, Washtenaw County, the University of Michigan, and the Ford School. Up to date information on U-M policies can be found here. It is expected that you will protect and enhance the health of everyone in the Ford School community by staying home and following self-isolation guidelines if you are experiencing any symptoms of COVID-19.

Use of Technology

Students should follow instructions from their instructor as to acceptable use of technology in the classroom, including laptops, in each course. All course materials (including slides, assignments, handouts, pre-recorded lectures or recordings of class) are to be considered confidential material and are not to be shared in full or part with anyone outside of the course participants. Likewise, your own personal recording (audio or video) of your classes or office hour sessions is allowed only with the express written permission of your instructor. If you wish to post course materials or photographs/videos of classmates or your instructor to third-party sites (e.g. social media), you must first have informed consent. **Without explicit permission from the instructor and in some cases your classmates, the public distribution or posting of any photos, audio/video recordings or pre-recordings from class, discussion section or office hours, even if you have permission to record, is not allowed and could be considered academic misconduct.**

Advances in generative artificial intelligence (AI) tools, such as ChatGPT, create a wide range of challenges and opportunities for higher education. While these tools can be extremely helpful in many respects, they do not change a fundamental principle of academic work: it is wrong to claim someone else's ideas or work as your own. For a student, the downsides and upsides of working with an AI tool can be much like working with a study group. Working with others can help you overcome roadblocks and help you learn things that you may have missed, but over-reliance on others to obtain answers likely will inhibit your learning. With statistics, we have many tools at our disposal. For example, ultimately we will use statistical software to do almost all the difficult work for us, but it is critical that we understand what the software is doing and that we know what to tell the software to do. This is why a key component of our course involves doing these problems by hand. When it comes time for quizzes and exams, this is what you will need to do, so practice it.

Here are a few guidelines for using generative AI tools. If you have any questions about whether a particular use is permissible, you should ask your instructors.

- It is okay to use these tools as a resource to look up information about theoretical concepts. Students should be aware that the information they receive may not be accurate or may not match what they are taught in lectures or their course textbook.
- It is okay to use these tools as a resource to help with learning the coding syntax of R or Stata.
- It is not permissible to use an AI tool to answer assignment questions for you. If you do use AI for assistance on how to do a particular kind of question, you should cite the query that you supplied. Please note that these tools make mistakes. You are responsible for any errors in the answer you submit.

Please review additional information and policies regarding academic expectations and resources at the Ford School of Public Policy at: https://intranet.fordschool.umich.edu/academic-expectations.

August 28: Introduction

- Agresti, chapter 1.
- Harford, Tim. "Statistics, Fast and Slow." *Financial Times*, April 13, 2018. See Canvas.

August 30: Measurement

- Agresti, section 2.1.
- Janet Buttolph Johnson and H.T. Reynolds 2008. "Measurement." *Political Science Research Methods*, chapter 4.
- Andrew Van Dam. 2023. "Are You Hispanic or Latino? A Census Mistake Reveals a Complicated Answer. " *Washington Post*, May 19, 2023. See Canvas.
- Additional resource: *OpenIntro Statistics*, sections 1.1–1.2; *Introductory Statistics*, chapter 1.
- Additional resource: for those interested in learning about survey question design, UM experts Frederick Conrad and Frauke Kreuter offer a course through Coursera entitled "Questionnaire Design for Social Surveys." Free to UM students. https://www.coursera.org/learn/questionnaire-design?.

September 6: Sampling

- Agresti, sections 2.2–2.5.
- Nate Cohn. 2023. "How Did We Do? A Review of 2022 Before Our First Poll of 2023. " *The New York Times*, July 28, 2023. See Canvas.
- Additional resource: *Intro Modern Statistics*, section 2.1; *Introductory Statistics*, chapter 1.
- Additional resource: for those interested in a deep dive on sampling, UM expert James Lepkowski offers a course through Coursera entitled "Sampling People, Networks, and Records." Free to UM students. https://www.coursera.org/learn/sampling-methods.
- See also a Coursera course by Frederick Conrad entitled "Data Collection: Online, Telephone and Face-to-face." https://www.coursera.org/learn/data-collection-methods.

September 11: Descriptive Statistics

▷ Problem Set 1 due Monday, September 11.

- Agresti, chapter 3.
- Philip Bump. 2023. "(White) America is getting older." *Washington Post*, June 22, 2023. See Canvas.
- Additional resource: *OpenIntro Statistics*, chapter 2; *Introductory Statistics*, chapter 2.

September 13 & 18: Probability

▷ Problem Set 2 due Tuesday, September 19.

- Agresti, section 4.1. Reading from the Additional resources is strongly recommended for the lectures on probability.
- Additional resources: *OpenIntro Statistics*, sections 3.1–3.3; *Introductory Statistics*, chapter 3.

September 20 & 25: Probability Distributions

▷ Problem Set 3 due Tuesday, September 26.

- Agresti, sections 4.2–4.3.
- Additional resource: *OpenIntro Statistics*, sections 3.4–3.5, 4.1, and 4.3; *Introductory Statistics*, sections 4.1–4.5, 5.1–5.3, and chapter 6.
- For more on the binomial distribution, see https://tinyurl.com/ydalpxke.

September 27: Sampling Distributions

⇒ Quiz on Wednesday, September 27. ⇐

- Agresti, sections 4.4–4.7 and pp. 118–123.
- Additional resource: *Introductory Statistics*, chapter 7.
- Helpful video about the Central Limit Theorem and normal distribution: https://youtu.be/zeJD6dqJ5lo

October 2: Statistical Inference (Estimation)

- Agresti, sections 5.1–5.4 and 5.6.
- Additional resource: *OpenIntro Statistics*, sections 5.1–5.2 and 7.1.1–7.1.4; *Introductory Statistics*, chapter 8.
- Glenn Kessler. "Did Exactly 4,645 People Die in Hurricaine Maria? Nope." *Washington Post*, June 1, 2018. See Canvas.
- Rink Hoekstra et al. (2014). "Robust Misinterpretation of Confidence Intervals." *Psychonomic Bulletin & Review* 21(5): 1157–1164. See Canvas.

October 4: Significance Tests

▷ Problem Set 4 due Wednesday, October 4.

- Agresti, sections 6.1–6.5.

- Additional resource: *OpenIntro Statistics*, sections 5.3, 6.1, 7.1.5; *Introductory Statistics*, chapter 9.
- A succinct summary of significance testing, with some simulation code in R, from Jose M Sallan: https://jmsallan.netlify.app/blog/introducing-hypothesis-testing/.

October 9 & 11: Significance Tests continued

▷ Problem Set 5 due Wednesday, October 11.

- Agresti, sections 6.6–6.8.
- Aaron E. Carroll (2018). "Congratulations. Your Study Went Nowhere." *The New York Times*, September 24, 2018. See Canvas.

October 16: Fall Study Break

October 18: Midterm Exam

October 23: Experimental Design and Causality

- "Research Design," in Kellstedt and Whitten, *The Fundamentals of Political Science Research*, 2009, chapter 4. See Canvas.
- Additional resource: *OpenIntro Statistics*, section 1.4.
- Jean Drèze (2018). "Evidence, Policy, and Politics." *Ideas for India*. http://www.ideasforindia.in/topics/miscellany/evidence-policy-and-politics.html.

October 25 & 30: Statistical Inference (Comparison of Two Groups)

▷ Problem Set 6 due Wednesday, October 25.

- Agresti, chapter 7.
- Melissa Healy (2018). "Coroner sent letters to doctors whose patients died of opioid overdoses. Doctors' habits quickly changed." *Los Angeles Times*, August 9, 2018. See Canvas.
- Additional resource: *OpenIntro Statistics*, sections 6.2 and 7.2–7.4, or *Introductory Statistics*, chapter 10.

November 1 & 6: Association Between Categorical Variables

▷ Problem Set 7 due Monday, November 6.

- Agresti, chapter 8.
- Additional resource: *OpenIntro Statistics*, sections 6.3–6.4, or *Introductory Statistics*, chapter 11.

November 8: ANOVA

- Agresti, sections 12.3–12.4.
- Additional resource: *OpenIntro Statistics*, section 7.5, or *Introductory Statistics*, chapter 13.

November 13: Correlation Analysis

⇒ Quiz on Monday, November 13. ⇐

- Agresti, section 9.4.
- Additional resource: *OpenIntro Statistics*, section 8.1.4, or *Introductory Statistics*, sections 12.6–12.7.

November 15 & November 20: Linear Regression

▷ Problem Set 8 due Monday, November 20.

- Agresti, sections 9.1–9.3 and 9.5.
- James H. Stock and Mark W. Watson. 2018. *Introduction to Econometrics*, 4th ed, chapter 4.
- Additional resource: *Intro Statistics*, sections 8.1–8.2 and 8.4, or *Introductory Statistics*, sections 12.1–12.5 and 12.8–12.9.

November 22: Thanksgiving Break

November 27 & November 29: Linear Regression continued.

▷ Problem Set 9 due Wednesday, November 29.

- Agresti, chapter 10, sections 11.1–11.3.
- James H. Stock and Mark W. Watson. 2018. *Introduction to Econometrics*, 4th ed, chapter 5.
- *OpenIntro Statistics*, sections 8.3.

December 4: Power Analysis.

- *OpenIntro Statistics*, section 7.4.

December 6: Review

▷ Problem Set 10 due Wednesday, December 6.

Wednesday, December 13, 4:00–6:00 pm: Final Exam