

# SPA 613: CONDUCT OF INQUIRY II

## AMERICAN UNIVERSITY — SPRING 2024

Version 1.0

### Instructor Information

*Name:* ANDREW R. FLORES

*Office:* Kerwin Hall 217

*Email:* aflores@american.edu

*Office Hours:* Wednesdays 3 PM – 5 PM, or by appointment

### Graduate Assistant Information

*Name:* DAKOTA STRODE

*Email:* ds4090a@american.edu

*Office Hours:* Thursdays 3 PM – 4 PM, or by appointment

### Class Information

*Dates:* Wednesday, Jan. 17 – May 8

*Time:* 11:20 AM – 2:10 PM

*Classroom:* Kerwin Hall 6

### Course Description

This course is intended as an introduction to research design and regression analysis. We will discuss the underlying principles of research design and how to frame a research question. We will learn the basic tools of statistical inference and apply this knowledge to the practice of regression analysis. The course will emphasize the relationship of these statistical tools for drawing causal inferences, however prediction and description will also be covered. To achieve this we will focus on:

- Basic probabilistic thinking
- Describing distributions
- The difference between design-based inference and model-based inference and their roles in research design
- The principles of statistical inference
- The building blocks of causal inference
- Characterization of the regression model
- Interpretation of the regression model
- Diagnostics and extensions of the regression model

- Threat to validity and sensitivity analysis for regression models
- Flexible regression modeling

Student should be aware that this course will be a lot of work. Lectures will be aimed at introducing and reinforcing materials, but much of the learning of the material must be done by you.

You will learn by doing. To that end, you will apply the lessons learned in lecture in numerous ways—through occasional problems in class, analytical and applied problems in problem sets, and many coding applications. Learning statistics is like learning another language—you will most likely not be an expert at the end of these fifteen weeks, but you will have the basic building blocks to continue practicing and learning beyond the scope of the course. At times it will be confusing and possibly overwhelming. Sustained effort is key to success in this class.

### Course Website

The course website will be hosted on Canvas. This will be the main source for all course materials, including problem sets, lecture note, and worksheets. A Discussion Board will be enabled on Canvas for students to ask questions about lectures, problem sets, and other course materials. This allows all students to benefit from the discussion and help each other understand materials. If you have a question, chances are one of your colleagues does too. *Students are encouraged and expected to participate in answering questions and helping each other on the forum.*

### Grading

Grading student work can often promote an antagonistic or stressful dynamic in a course that distracts from the business of learning. The goal of this course is to help you begin to build the toolkit you need to do quantitative methods in your own, individual research. This requires that you try new skills, even if you do not get it right on the first try. Because of these potential tensions, this course will instead rely on a form of evaluation called “contract grading” that encourages responsible student behavior and rewards persistence and effort. My hope is that by clearly outlining the effort necessary to earn an ‘A’ or ‘B’ in this class, you will feel empowered to take risks when attempting new material and concepts and seek out help without fearing the impact on your grade. It also mitigates the wide range in quantitative backgrounds that students may have and does not merely reward previous experience with the material.

To that end, the following table outlines the contract for receiving an ‘A’ or ‘B’ in this course. Deviations of  $\pm$  are at the discretion of the instructor.

	<u>B</u>	<u>A</u>
Participation	≥ 3 participation points in at least 11 weeks. Completed pre-test, post-test, midterm, and final.	≥ 5 participation points in at least 11 weeks, with a cumulative of at least 110 pts. On-time completion of the pre-test, post-test, midterm, and final.
Coding	≥ 75% of coding assignments completed on time (18).	≥ 87.5% of coding assignments completed on time (21).
Problem Sets	≥ 6 problem sets that receive a complete.	≥ 6 problem sets that receive a complete, and 5 problem set reflections that receive a complete.
Midterm	Passing grade on the midterm.	Passing grade on the midterm and completed midterm resubmission opportunity.
Final	Passing grade on the final.	Passing grade on the final.
Term Paper	Passing grade on the term paper.	Passing grade on the term paper.

Everyone in the course will be participating in the contract unless you explicitly notify me by email, and receive a written confirmation. *You may opt-out of the contract at any time before the final. If you have reverted to traditional grading, you will do so for the remainder of the term.*

The traditional course grade is determined by the following components:

Participation	5%
Coding Exercises	15%
Problem Sets	35%
Midterm	15%
Final	15%
Term Paper	15%

## Course Assignments

Grading in this class will be based on the components described below. **Late work will not be accepted without prior permission.** Makeup exams will not be given, and students who miss exams will receive a score of 0 absent extraordinary circumstances.

### *Participation*

There are numerous ways to participate in the course. Within a given week, you can accumulate participation points by:

- Attending lecture (2 pts)
- Attending office hours (2 pts)
- Posting a thoughtful question in the discussion forum (1 pt each, max 2 pts)
- Respectfully and constructively answering a question in the discussion forum (1 pt each, max 2 pts)

In addition to weekly participation, you must complete the following assignments for full participation credit:

- Pre-test
- Midterm Evaluation
- Post-test
- Final

We will use Canvas discussion boards for all online discussion for the course. Do not email questions about course material to the professor. You should post and answer questions on the forum. If you email questions to the instructor or GA, we will request you post them (optionally anonymously) to the forum.

Anonymous posts will be anonymous to your peers but known by your instructors.

#### *Coding Exercises*

Learning to code in R is a key component of this course. To do this, we will rely on [datacamp.com](https://datacamp.com) lessons. To ensure that you stay on top of coding, there will be a coding lesson that must be completed *by midnight before each class period*. Due dates are available in the [datacamp.com](https://datacamp.com) module, and are listed in the course schedule of this syllabus.

#### *Problem Sets*

Problem sets will be assigned approximately bi-weekly, with each contributing equally toward this portion of the final grade. Each problem set will contain a mix of analytical and coding problems.

Problem sets will be evaluated as being complete or incomplete. Each problem within a problem set will be evaluated on the following (note that problems can have sub-parts):

- ‘O - Outstanding’: Attempted all of the problems. Truly exceptional work indicating mastery of the material.
- ‘E - Exceeds Expectations’: Attempted all of the problems, mastered most concepts, but room for growth remains on the material.
- ‘A - Acceptable’: Attempted all of the problems. Demonstrated clear understanding of core concepts expected for the course, but lacking mastery or nuance on some concepts.
- ‘N - Needs Improvement’: Attempted all of the problems. Clear holes in understanding of core concepts. Not meeting expectations for the course.
- ‘U - Unsatisfactory’: Little discernible effort or no attempt.

The determination of complete/incomplete will be a holistic assessment of the performance on the problem set problems. For example, receiving several U’s would likely result in a problem set being evaluated as incomplete.

**Reflections:** For each problem set, there will be detailed solutions. You may choose to do the following set reflection assignments:

- Pick at least one of your lowest scoring problems, for which you received lower than an 'O'. For each problem, explain what you did incorrectly, and what you have subsequently learned. Outline any remaining questions you have. These should be no more than 2 pages for each problem set.
- If you received an 'O' on everything, pick at least one problem. Write a related homework problem with solution. Note that this cannot be merely a marginal change to the problem, it needs to be a related but sufficiently different problem.

**Collaboration Policy:** Collaboration is allowed and encouraged with students currently enrolled in the course. You may not collaborate with students outside of the course. *All persons and resources consulted in completion of the problem set must be clearly listed.* This means that if you discuss a problem with a colleague, you should list their name at the beginning of that problem. If you copy code from stackexchange.com or other sources, you should cite them.

Note that while collaboration is an integral part of this course, you should not copy, or rely too heavily on others, when completing problem sets. See the section on Academic Honesty and ensure that you're following university guidelines.

All questions should be posted to the Canvas discussion forum.

**Submission Policy:** All submissions should be uploaded to the Canvas. Unless otherwise noted, solutions should be typed in RMarkdown. The .Rmd and final PDF should be uploaded to Canvas before the class period the problem set is due.

Guidelines for submission:

- The problem being addressed should be replicated at the beginning of the solution.
- Code should be clearly commented, or described in the markdown file.
- Figures should be neatly labeled and annotated. Labels should be sufficiently large to be readable.
- Any proofs should be easy to follow, with any properties or theorems cited as needed.
- Discussions should be concise and correct. Points will be deducted for verbose discussions that include incorrect answers, even if they also discuss the correct answer.
- There is no exception to the RMarkdown policy.

If you're struggling with RMarkdown, you can check out the datacamp.com module here:

<https://www.datacamp.com/courses/reporting-with-r-markdown>.

To submit your assignments in PDF form, you will need to have installed  $\text{\LaTeX}$ . Please follow the steps and guidance here to embed  $\text{\LaTeX}$  into RMarkdown (if you do not have a compiler on your computer already):

<https://everyday.codes/tutorials/how-to-use-latex-in-rmarkdown/>

A helpful reference for  $\text{\LaTeX}$ math symbols can be found here: <https://en.wikibooks.org/wiki/LaTeX/Mathematics>.

Users of Stata will still need to submit these assignments with RMarkdown. You should install the Statamarkdown package from R to enable functionality. You can see the following reference for further information:

<https://users.ssc.wisc.edu/~hemken/Stataworkshops/Statamarkdown/stata-and-r-markdown.html>

### *Midterm Exam*

The midterm exam will be take-home over the course of Weeks 7–8 and will cover the material discussed in class up to that point. While the exam is not timed, we will only meet for half the class time on Wednesday, February 28.

**Resubmission:** You will have a resubmission opportunity for the midterm exam.

### *Final Exam*

There will be a comprehensive take home final exam due May 8, 2024 (administered online, open note, and will begin on May 2, 2024).

### *Final Paper*

I want to promote your engagement with empirical political science and the development of your ideas. For the major paper in this class, I would like you to draft a paper around an analysis of quantitative data: 10–20 pages of text with additional tables as needed and references (thus, the paper may be up to 30 double-spaced pages). The paper will have a literature review, develop an argument (theory), present expectations / hypotheses and show data and analyses to test them. The discussion should envision how to move the manuscript forward, toward publication.

You may use a paper that you are drafting for a related course (e.g., Conduct I). However, **all papers should contain detailed appendices diagnosing the analyses performed**. For example, if linear regression is used, there should be appendices assessing heteroscedasticity, multicollinearity, and other types of violations. There is no page limit on these appendices, but please be judicious with the use of figures and tables. At times, these diagnostics may alter your primary analyses in your final draft (e.g., if robust standard errors are used because heteroscedasticity was found to be a problem).

Ideally, this project is something that you can consider taking to an academic conference and eventually publishing. You should use developmentally and field-appropriate analytical tools, including cross-tabulation, linear and generalized linear regression. While the assignment is individual for this course, it may lead to some collaborative research (with me, your peers, or others). I will provide more details about this project and my expectations during the next few weeks. The paper should follow APSA citation style, and students are encouraged to use  $\text{\LaTeX}$ .

### **Required Books**

**Aronow PM, Miller B. 2019. *Foundations of Agnostic Statistics*. Cambridge University Press.**

This book is technical, and it will provide clear and explicit ways of understanding the mathematics

behind statistics. You may not fully understand all the content in the text with a single read, and you may want to keep a copy of the glossaries in the back of the book handy for mathematical notation.

**Wooldridge, Jeffrey M. 2020. *Introductory Econometrics: A Modern Approach*. 7th Ed. Boston, MA: Cengage.**

This book is less technical, and it will provide clear ways to think about statistical methods and models. Earlier editions can suffice. An online Version 5 should be [here](#).

**Heiss, F. 2020. *Using R for Introductory Econometrics*. 2nd Ed. Heiss.**

This book is applied, and it is freely available [here](#).

### **Optional Book**

For users of Stata - you will find this book as helpful (earlier versions should be OK):

**Cameron, A. Colin and Pravin K. Trivdei. 2022 *Microeconometrics using Stata*, 2nd Ed. Vol. I. New York: Routledge.**

## AU Student Support Services

### *Academic Support Services*

All students may take advantage of the [Academic Support and Access Center \(ASAC\)](#) for individual academic skills counseling, workshops, Tutoring and Writing Lab appointments, peer tutor referrals, and Supplemental Instruction. The ASAC is located in Mary Graydon Center 243.

Additional academic support resources available at AU include the Bender Library, the Department of Literature's Writing Center (located in the Library), the Math Lab in the Department of Mathematics & Statistics, and the Center for Language Exploration, Acquisition, & Research (CLEAR) in Asbury Hall. A more complete list of campus-wide resources is available in the ASAC.

### *Accommodations for Students with Disabilities*

AU is committed to making reasonable accommodations for qualified students with disabilities. The ASAC assists students with disabilities and promotes full participation in academic programs and other campus activities.

Students are not required to notify the university or any of its offices or personnel of a disability either prior to or subsequent to admission; however, if a student plans to request accommodations, documentation of the disability must be provided. As accommodations are not retroactive, timely notification at the beginning of the semester, if possible, is strongly recommended.

To register with a disability or for questions about disability accommodations, contact the Academic Support and Access Center at 202-885-3360 or [asac@american.edu](mailto:asac@american.edu), or drop by MGC 243.

For more information, visit AU's [Services for Students with Disabilities web page](#).

### *Center for Diversity & Inclusion*

[CDI](#) is dedicated to enhancing LGBTQ, multicultural, first-generation, and women's experiences on campus and to advancing AU's commitment to respecting and valuing diversity by serving as a resource and liaison to students, staff, and faculty on issues of equity through education, outreach, and advocacy. It is located on the 2nd floor of Mary Graydon Center. (202-885-3651, MGC 201 & 202)

### *Counseling Center*

The [Counseling Center](#) offers counseling and consultations regarding personal concerns, self-help information, and connections to off-campus mental health resources. (202-885-3500, MGC 214)

### *Dean of Students Office*

The [Dean of Students Office](#) offers one-on-one meetings to discuss academic, adjustment, and personal issues that may be interfering with a student's ability to succeed academically. The office also verifies documentation for students who have medical or mental health issues that cause them to be absent from class. (202-885-3300, Butler Pavilion 408)

### *International Student & Scholar Services*



[International Student & Scholar Services](#) has resources to support academic success and participation in campus life including academic counseling, [support for second language learners](#), response to questions about visas, immigration status and employment and intercultural programs, clubs and other campus resources. (202-885-3350, Butler Pavilion 410)

#### *Office of Advocacy Services for Interpersonal & Sexual Violence*

[OASIS](#) provides free and confidential advocacy services for anyone in the campus community who experiences sexual assault, dating or domestic violence, or stalking. Advocacy is survivor-driven and intended to empower survivors to make informed decisions about their health, emotional well-being, and the adjudication process. (202-885-7070, Wellness Center – McCabe Hall 123)

#### *Writing Center*

[Writing Center](#) offers free, individual coaching sessions to all AU students. In your 45-minute session, a student writing consultant can help you address your assignments, understand the conventions of academic writing, and learn how to revise and edit your own work. (202-885-2991, Bender Library – 1st Floor Commons).

#### *University Library*

The reference librarian for political science, Olivia Ivey ([oivey@american.edu](mailto:oivey@american.edu)) may be an exceptional resource for you as you develop research papers.

#### *Quantitative Support*

American University students can receive additional [Quantitative Support](#) with a Math & Stats Lab and Statistical Software Support.

### **University Policies**

#### *Academic Integrity*

All students are required to follow the University's Academic Integrity Code. If you have not already done so, please familiarize yourself with the standards & requirements of the University's Academic Code of Conduct. Violations of the Code of Conduct will not be tolerated & will be reported appropriately. You can find more information about the University's Academic Integrity Code here: <http://www.american.edu/academics/integrity/code.cfm>.

#### *Artificial Intelligence*

The use of Artificial Intelligence (AI) may not be used to originally create text for a written submission. AI tools may be used to revise originally written text, and students should cite or note such use in their submission. Students should exercise caution when using AI for assistance with code. If AI is used, then students should cite the source for the code. Similar to finding answers of Stack-Exchange, students should cite their sources for ideas. Violations of the AI policy will be considered as violations of the Academic Integrity Code of Conduct and will follow the same procedures for such violations.

#### *Discrimination and Harassment (Title IX)*

American University expressly prohibits any form of discriminatory harassment including sexual harassment, dating and domestic violence, sexual assault, and stalking. The university is an equal opportunity, affirmative action institution that operates in compliance with applicable laws and regulations. AU does not discriminate on the basis of race, color, national origin, religion, sex (including pregnancy), age, sexual orientation, disability, marital status, personal appearance, gender identity and expression, family responsibilities, political affiliation, source of income, veteran status, an individual's genetic information, or any other bases under federal or local laws in its programs and activities.

If you experience any of the above, you have the option of filing a report with the [AU Department of Public Safety](#) (202-885-2527) or the [Office of the Dean of Students](#) (dos@american.edu or 202-885-3300). Please keep in mind that all faculty and staff – with the exception of counselors in the Counseling Center, staff in the Office of Advocacy Services for Interpersonal and Sexual Violence (OASIS), medical providers in the Student Health Center, and ordained clergy in the Kay Spiritual Life Center – who are aware of or witness this conduct are required to report this information to the university, regardless of the location of the incident. For more information, including a list of supportive resources on and off-campus, contact [OASIS: The Office of Advocacy Services](#) for Interpersonal and Sexual Violence (oasis@american.edu or 202-885-7070) or the Office of the Dean of Students.

For information about your rights, see the [Title IX Information](#) page on the AU website.

#### *Emergency Preparedness*

In the event of an emergency, American University will implement a plan for meeting the needs of all members of the university community. Should the university be required to close for a period of time, we are committed to ensuring that all aspects of our educational programs will be delivered to our students. These may include altering and extending the duration of the traditional term schedule to complete essential instruction in the traditional format and/or the use of distance instructional methods. Specific strategies will vary from class to class, depending on the format of the course and the timing of the emergency. Faculty will communicate class-specific information to students via AU email and Canvas, while students must inform their faculty immediately of any emergency-related absence. Students are responsible for checking their AU email regularly and keeping themselves informed of emergencies. In the event of an emergency, students should refer to the AU Student Portal, the [AU website](#), and the AU information line at (202) 885-1100 for general university-wide information, as well as contact their faculty and/or respective dean's office for course and school/college specific information.

#### *Religious Observances*

Students will be provided the opportunity to make up any examination, study, or work requirements that may be missed due to a religious observance, provided they notify their instructors before the end of the second week of classes. Please send this notification through email to the professor. For additional information, see American University's [religious observances policy](#).

#### *Sharing of Course Content*

Students are not permitted to make visual or audio recordings, including live streaming, of classroom lectures or any class-related content, using any type of recording devices (e.g., smart phone,

computer, digital recorder, etc.) unless prior permission from the instructor is obtained, and there are no objections from any of the students in the class. If permission is granted, personal use and sharing of recordings and any electronic copies of course materials (e.g., PowerPoints, formulas, lecture notes, and any classroom discussions—online or otherwise) is limited to the personal use of students registered in the course and for educational purposes only, even after the end of the course.

Exceptions will be made for students who present a signed Letter of Accommodation from the Academic Support & Access Center. Further details are available from the [ASAC website](#).

To supplement the classroom experience, lectures may be audio or video recorded by faculty and made available to students registered for this class. Faculty may record classroom lectures or discussions for pedagogical use, future student reference, or to meet the accommodation needs of students with a documented disability. These recordings are limited to personal use and may not be distributed (fileshare), sold, or posted on social media outlets without the written permission of faculty.

Unauthorized downloading, file sharing, distribution of any part of a recorded lecture or course materials, or using information for purposes other than the student's own learning may be deemed a violation of American University's Student Conduct Code and subject to disciplinary action (see Student Conduct Code VI. Prohibited Conduct).

#### *Use of Student Work*

The professor will use academic work that you complete for educational purposes in this course during this semester. Your registration and continued enrollment constitute your consent.

#### *Changes to the Syllabus*

Any substantive changes to the syllabus in terms of assignments and policies will be announced on Canvas and a revised syllabus will be posted to Canvas. If there are minor changes to due dates, then those changes will be announced to Canvas but not revised in the syllabus. Major changes in due dates will be announced on Canvas, revised in the syllabus, and posted to Canvas.

## Schedule

Outside of required books, all other readings will be posted to Canvas.

### Before Classes Begin – Math Review

- > Algebra
- > Functions

### Readings:

- > Moore WH, Siegel DA. 2013. *A Mathematics Course for Political & Social Research*. Princeton: Princeton University Press. Ch. 1–3.
- > Wooldridge. Appendix A.

### DataCamp:

- > Introduction to R

## UNIT ONE: Probability and Distributions

### January 17 – Introductions, Probability, & Summarizing Distributions

- > random events
- > basic properties of probability
- > joint and conditional probability
- > Bayes' Rule
- > Law of total probability
- > Independence
- > Random variables
- > Probability Density Functions (PDFs) and Cumulative Density Functions (CDFs)
- > expected values and their properties
- > linearity in expectations
- > variance and its properties
- > normal distributions
- > Mean Squared Error (MSE)
- > summaries of joint distributions: covariance and correlation
- > the Conditional Expectation Function (CEF) and the law of iterated expectations
- > law of total variance

#### Readings:

- > Aronow and Miller. Ch. 1–2.
- > Wooldridge. Appendix B.
- > Heiss. Ch. 1 §§1–5.6, 6-7.

#### *Stata Users:*

- > Cameron and Trivedi. Ch. 1.

#### *Suggested:*

- > Moore WH, Siegel DA. 2013. *A Mathematics Course for Political & Social Research*. Princeton: Princeton University Press. Ch. 9.

#### DataCamp:

- > Introduction to the Tidyverse
  - Data wrangling
  - Data visualization
  - Grouping and Summarizing
- > Reporting with R Markdown
  - Getting Started with R Markdown

Problem Set:

- > Problem Set 1 posted to Canvas

## UNIT TWO: Research Design and Learning from Data

### January 24 – Learning and Inferring from Data

- > random samples, estimation
- > sampling distribution
- > Weak Law of Large Numbers (WLLN) and Central Limit Theorem (CLT)
- > asymptotics
- > plug-in principle
- > Mean Squared Error (MSE)
- > intervals and hypothesis testing
- > p-values
- > bootstrap
- > permutation inference

#### Readings:

- > Aronow and Miller. Ch. 3.
- > Wooldridge. Ch. 19; Appendix C.
- > Heiss. Ch. 1  
SS 5.4, 6, 8; Ch. 19 §§1–3.
- > Leek J., et al. 2017. Five Ways to Fix Statistics. *Nature* 551: 557–559. <https://media.nature.com/original/magazine-assets/d41586-017-07522-z/d41586-017-07522-z.pdf>
- > Benjamin DJ., et al. Redefine Statistical Significance. *Nature Human Behavior* 2: 6–10. <https://www.nature.com/articles/s41562-017-0189-z.pdf>
- > Amrhein V., Sander Greenland. 2018. Remove, Rather than Redefine, Statistical Significance. *Nature Human Behavior* 2: 4. <https://www.nature.com/articles/s41562-017-0224-0.pdf>
- > Imbens GW, Rubin DB. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Ch. 5–6.

#### Stata Users:

- > Cameron and Trivedi. Ch. 2.

#### DataCamp:

- > Reporting with R Markdown
  - Adding Analyses and Visualizations

- Improving the Report
- > Data Manipulation in R with dplyr:
  - Transforming Data with dplyr
  - Aggregating Data
  - Selecting and Transforming Data



## UNIT THREE: Simple Linear Regression

### January 31 – The Conditional Expectation Function (CEF) and Best Linear Predictor (BLP) and Bivariate Regression

- > Conditional Expectation Functions (CEFs)
- > kernel estimation
- > Best Linear Predictors (BLPs)
- > derivation of Ordinary Least Squares (OLS)
- > mechanics of the Simple Linear Regression (SLR)
- > properties of the SLR
- > inference with regression

#### Readings:

- > Wooldridge. Ch. 2 §§1–3.
- > Heiss. Ch. 2 §§1–4.

#### *Stata Users:*

- > Cameron and Trivedi. Ch. 3 §§1–4.

#### DataCamp:

- > Foundations of Inference
  - Introduction to the Ideas of Inference
  - Completing a Randomization Test: Gender Discrimination
  - Hypothesis Testing Errors: Opportunity Cost
  - Confidence Intervals
- > Hypothesis Testing in R
  - Introduction to Hypothesis Testing
  - Two-Sample and ANOVA Tests

#### Problem Set:

- > Problem Set 1 **due Friday, February 2**
- > Problem Set 2 posted to Canvas

## February 7 – The Conditional Expectation Function (CEF) and Best Linear Predictor (BLP) and Bivariate Regression (Day 2)

- > Conditional Expectation Functions (CEFs)
- > kernel estimation
- > Best Linear Predictors (BLPs)
- > derivation of Ordinary Least Squares (OLS)
- > mechanics of the Simple Linear Regression (SLR)
- > properties of the SLR
- > inference with regression

### Readings:

- > Wooldridge Ch. 2 §§4–6.
- > Heiss. Ch. 2 §§6–7.

### *Stata Users:*

- > Cameron and Trivedi. Ch. 3. §5

### DataCamp:

- > Introduction to Regression in R
  - Simple Linear Regression
  - Assessing Model Fit

## February 14 – Causal Inference and Regression

- > design- vs. model-based inference
- > introduction to potential outcomes
- > OLS as difference-in-means

### Readings:

- > Aronow and Miller. Ch. 7 §2.2.
- > Wooldridge. Ch. 2 §7.
- > Imbens GW, Rubin DB. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Ch. 1–2, 5.

### Stata Users:

- > Cameron and Trivedi. Ch. 24

### Problem Set:

- > Problem Set 2 **due Friday, February 16**
- > Problem Set 3 posted to Canvas

## UNIT THREE: Multiple Regression

### Feb 21 and 28 – Multiple Regression

- > Simpson's Paradox
- > interpretation with two regressors
- > matrix notations
- > derivation / mechanics of OLS
- > geometry of OLS
- > multiple regression and causality

#### Readings:

- > Aronow and Miller. Ch. 4.
- > Wooldridge. Ch. 3.
- > Heiss. Ch. 3.

#### *Stata Users:*

- > Cameron and Trivedi. Ch. 3–4.

#### DataCamp:

- > Modeling with Data in the tidyverse
  - Introduction to Modeling
  - Modeling with Basic Regression

#### Problem Set:

- > Problem Set 3 **due Friday, March 1**
- > Problem Set 4 posted to Canvas

**STUDY FOR MIDTERM**

## March 6 and 13 – Inference in Regression

- > finite and asymptotic inference
- > hypothesis testing
- > joint hypothesis testing

### Readings:

- > Aronow and Miller. Ch. 4.
- > Wooldridge. Ch. 4–5.
- > Heiss. Ch. 4–5.

### *Stata Users:*

- > Cameron and Trivedi. Ch. 11.

### DataCamp:

- > Intermediate Regression in R
  - Parallel Slopes

### Problem Set:

- > Problem Set 4 **due Friday, March 15**
- > Problem Set 5 posted to Canvas

## March 20 and 27 – Flexible Regression Modeling

- > categorical and dummy variables
- > interactions
- > polynomials
- > overfitting
- > Quantities of Interest (QIs)
- > flexible modeling: Generalized Additive Models (GAMs) and Kernel Regularized Least Squares (KRLS)

### Readings:

- > Aronow and Miller. Ch. 4.
- > Wooldridge. Ch. 6; Ch. 7 §1–4.
- > Heiss. Ch. 6–7.

### *Stata Users:*

- > Cameron and Trivedi. Ch. 3–4; 14.

### DataCamp:

- > Intermediate Regression in R
  - Interactions
  - Multiple Linear Regression

### Problem Set:

- > Problem Set 5 **due Friday, March 29**
- > Problem Set 6 posted to Canvas

## UNIT FOUR: Diagnostics and Violations in Multiple Regression

### April 3 – Violations - Non-Linearity and Multicollinearity

- > omitted variables
- > data transformation
- > multicollinearity and Variance Inflation Factors (VIFs)

#### Readings:

- > Aronow and Miller. Ch. 3–4.
- > Wooldridge. Ch. 8–9.
- > Heiss. Ch. 1 §5; Ch. 8–9.

#### *Stata Users:*

- > Cameron and Trivedi. Ch. 5.

#### DataCamp:

- > Modeling with Data in the tidyverse
  - Modeling with Multiple Regression
  - Model Assessment and Selection

## April 10 – Violations - Exogeneity and Spherical Error Variance

- > outliers, influence, leverage, Cook's Distance
- > Standardized residuals, QQ-plots, and residual plots
- > Robust Standard Errors and Sandwich estimators
- > clustering: clustered standard errors
- > serial correlation: autocorrelated errors

### Readings:

- > Aronow and Miller. Ch. 3–4.
- > Wooldridge. Ch. 8–9; Ch. 12 §1–4.
- > Heiss. Ch. 1 §1.5; Ch. 8–9.

### *Stata Users:*

- > Cameron and Trivedi. Ch. 6.

### DataCamp:

- > Introduction to Data Visualization with ggplot2
  - Introduction
  - Aesthetics

### Problem Set:

- > Problem Set 6 **due Friday, April 12**
- > Problem Set 7 posted to Canvas



## April 17 – Sensitivity Analysis

- > sensitivity of causal effects
- > bounds
- > unobserved confounding
- > robustness

### Readings:

- > Cinelli C, Hazlett C. 2020. Making Sense of Sensitivity: Extending Omitted Variable Bias. *Journal of the Royal Statistical Society: Series B* 82(1): 39–67.

### DataCamp:

- > Introduction to Data Visualization with ggplot2
  - Geometrics
  - Themes
  - Statistics
  - Facets

## April 24 – Generalized Linear Models

- > binary choice models
- > Maximum Likelihood Estimation (ML or MLE)
- > mixture models

### Readings:

- > Aronow and Miller. Ch. 5.
- > Wooldridge. Ch. 7 §§5–7; Ch. 17.
- > Heiss. Ch. 17.

### *Stata Users:*

- > Cameron and Trivedi. Ch. 10; 17.

### *Suggested:*

#### DataCamp:

- > Cleaning Data in R
  - Common Data Problems
  - Categorical and Text Data
- > Introduction to Writing Functions in R
  - How to Write a Function
  - All about Arguments
  - Return Values and Scope

### Problem Set:

- > Problem Set 7 **due Friday, April 26**

**Acknowledgments:**

This syllabus was structured in part by the syllabi of Erin Hartman and Silvia Kim, among others.