

Data in Politics I
POLI 281
MW 3:35-4:50
Woollen Gym Rm 302

Instructor	Colin Case
Email	crcase@live.unc.edu
Office	312 Hamilton Hall
Office Hours	Monday 1:30 – 3:00 & Wednesday 12:00 – 1:30

COURSE DESCRIPTION

The Information Revolution has dramatically expanded the volume of information we have about the world around us. Social scientific analytical skills are transforming many sectors—business, journalism, law, public policy, health care, and finance, to name but a few—and are more valued now than ever. However, data can also be used to mislead, and without strong data skills it’s hard to be an informed data consumer. The broad learning objective for this course is to help students develop the tools they need to be better citizens, as well as informed participants and active leaders in data-driven sectors. More specifically, the learning objectives are:

1. To increase students’ comfort and facility managing data in the R statistical language, with an emphasis on versatile tools such as loops, sampling functions, merging datasets, and the GGPlot data visualization software.
2. To teach basic principles of data description, including standard descriptive plots and statistics.
3. To develop students’ ability to use data to answer important social scientific questions.

Students will leave this class with the competencies they need to conduct basic analysis on many different forms of data, as well as the foundation they need to acquire more advanced skills (such as characterizing uncertainty in data and testing formal hypotheses).

The target audience for this course is undergraduate students with interest in the social sciences (not only Political Science), who want to use quantitative approaches to solve important problems and develop marketable analytical skills. This course is a prerequisite for Poli381: Data in Politics II: Frontiers and Applications.

This course fulfills the Quantitative Intensive (QI) requirement in UNC’s Making Connections curriculum. It also fulfills counts as a course in research methods (required for completing the Political Science Bachelor’s degree). It also counts towards the Data Science minor.

REQUIRED MATERIALS

Much of the hands-on work we will do in this class requires us to use computers, so I ask that you bring your laptops to class each day. Please let me know if you do not have regular access to a laptop computer. Throughout the semester we will make use of the R statistical computing environment to analyze data. RStudio is a popular editor that allows you to open, edit, and save R text files, making it much easier to work with R. I will use RStudio to demonstrate in class. To access these programs:

- R: Download precompiled binary distributions at <http://cran.us.r-project.org>
- RStudio: Download RStudio Desktop at <http://www.rstudio.com/products/rstudio>

There is no required textbook for this course.

ASSIGNMENTS AND GRADING SCALE

A 94.00-100
A- 90.00-93.99
B+ 87.00-89.99
B 84.00-86.99
B- 80.00-83.99
C+ 77.00-79.99
C 74.00-76.99
C- 70.00-73.99
D+ 65.00-69.99
D 64.99-60.00
F 59.99-0

Grades will be calculated using the below components and weights:

Participation (20%)

Your participation grade will be based on the following factors:

1. **Attending and participating in class.** We will have many activities throughout the semester, some individual and some in groups. Full participation in class means being active during lecture, discussion, and group work. This comprises the majority of your participation grade.
2. **Participation in our class's online discussion forum.** On Piazza (see instructions to sign up on Canvas), you can benefit your participation grade either by posting your own questions or by providing thoughtful answers to other students' questions. You can also post anonymously.
3. **Completion of online quizzes throughout the semester.** These quizzes will be ungraded and will be given occasionally at the end of class. They are designed to help me gauge overall progress and comprehension in the class in case there are areas we need to spend more time on.

While I will not take attendance each class, your participation grade will suffer if you do not attend class regularly. If you anticipate missing more than 3 courses, it is your responsibility to talk to me about any extenuating circumstances that may be impacting your ability to attend class. Learning coding and statistical concepts has a steep learning curve. In many ways, it is like learning a new language and this class will cover a lot of material quickly. The best way to be successful and get the most out of this course is to be present for class and to ask questions as they come up (and you will have them).

Assignments (50%)

There will be five assignments due as noted in the schedule below. They are weighted equally. Assignments are due at 11:59pm on the days indicated. Students are allowed to work with one another on assignments, but must submit individual assignments on Canvas and must submit their own code.

Final Project (30%)

The class has a capstone final project for which students will perform and present original data analysis on an existing dataset. Your grade consists of a final paper (25% of final grade), and a final presentation on the last day of class (5% of final grade). More information on this assignment will be available later in the semester.

COMMUNICATION

I am very happy to meet with students outside of class time. Whether it be to discuss concerns about the course, questions about the material, or simply to engage further with the topic, please feel free to come to office hours. I will be holding office hours in Hamilton 312. If you are unable to meet during my office hours, which are listed at the top of this syllabus, please email me to set up an alternative time. Office hours are an important resource that should be utilized to improve understanding of material or ask more personalized questions. Office hours before an exam will be held at regular times and, unless noted by me, no additional office hours will be held – so plan accordingly. Prior to writing assignment due dates, I will be available during office hours to read over paper outlines or talk through ideas and provide feedback. Please do not bring completed papers and, be advised, office hours before paper deadlines may be busy.

Outside of office hours, e-mail is the easiest way to contact me. Please include “Poli 281” in the subject line of the e-mail. I will typically respond to email within 24 hours. If I do not respond to your email in this time frame, please send a follow up. If you have extensive questions that require a longer conversation, please come to my office hours or stay after class. I will frequently send emails relating to the course material, upcoming assignments or activities, and general reminders. I expect that you are checking your UNC email regularly to stay on top of these updates.

ACCESSIBILITY RESOURCES AND SERVICES

UNC facilitates the implementation of reasonable accommodations, including resources and services, for students with disabilities, chronic medical conditions, a temporary disability, or pregnancy complications resulting in barriers to fully accessing University courses, programs, and activities. Accommodations are determined through the Office of Accessibility Resources and Service (ARS) for individuals with documented qualifying disabilities in accordance with applicable state and federal laws. If you have an accommodation related to testing procedures, please let me know ASAP so I can provide you with the accommodation at the outset of class. For more information visit <https://ars.unc.edu> or email ars@unc.edu.

COUNSELING AND PSYCHOLOGICAL SERVICES

CAPS is strongly committed to addressing the mental health needs of a diverse student body through timely access to consultation and connection to clinically appropriate services, whether for short or long-term needs. Visit their website <https://caps.unc.edu/> or their facilities on the third floor of the Campus Health Services building for a walk-in evaluation to learn more.

TITLE IX RESOURCES

Any student who is impacted by discrimination, harassment, interpersonal (relationship) violence, sexual violence, sexual exploitation, or stalking is encouraged to seek resources on campus or in the community. Reports can be made online to the EOC at <https://eoc.unc.edu/report-an-incident/>. Additional resources are available at safe.unc.edu.

ACADEMIC HONESTY

In a class setting, cooperative work has both benefits and pitfalls. You can learn a lot by explaining things to each other. But it can also be easy to stumble into a passive mindset where you're not really understanding the concepts. To strike a balance, I will designate some activities and assignments (or parts thereof) as being cooperative, and others as individual. It is critical that you attend to this distinction, as completing individual work cooperatively would be a breach of academic integrity.

By its nature, this class has an extra matter we need to address. While discussion with other people is permitted and encouraged for work designated as cooperative, there is a distinction between discussing a problem and copying someone else's work. Writing computer code is an especially tempting activity for which to copy work. Students can discuss problem-solving strategies, clarify concepts, and point out mistakes—but ultimately each person must generate his or her own path to the solution. In our class, copying another person's computer code or written analysis is plagiarism. This is true if you copy is from any source, including a classmate or former student, an online source, or an AI product similar to ChatGPT or CoPilot. Even for work designated as cooperative, you must write your code and analysis individually.

COURSE SCHEDULE

Jan 10	Introduction <ul style="list-style-type: none">• Syllabus
Jan 15	No Class (MLK Day)
Jan 17	The R Statistical Software: Getting Set Up <ul style="list-style-type: none">• Register for Piazza• Download and install both R and RStudio• After class: Review LearnR Tutorial 1: Intro to R
Jan 22	Data Structures Part 1: Vectors <ul style="list-style-type: none">• LearnR Tutorial 2, Topics 1-3: Vectors
Jan 24	Data Structures Part 2: Data Types & Operations <ul style="list-style-type: none">• LearnR Tutorial 2, Topics 4-5: Vectors
Jan 29	Data Structures Part 3: Dataframes and Matrices <ul style="list-style-type: none">• LearnR Tutorial 3, Topics 1-3: Introduction to Dataframes
Jan 31	Summarizing Data, Part 1 <ul style="list-style-type: none">• LearnR Tutorial 3, Topic 4: Importing Data
Feb 5	Summarizing Data Part 2 <ul style="list-style-type: none">• LearnR Tutorial 4, Topics 1-3: Basics of Data Cleaning
Feb 7	Summarizing Data Part 3 <ul style="list-style-type: none">• LearnR Tutorial 4, Topics 4-6: Creating New Variables• Assignment 1 Due Friday, February 9th at 11:59 PM
Feb 12	No Class (Wellness Day)
Feb 14	Putting Data in Context: Sampling <ul style="list-style-type: none">• No reading
Feb 19	Advanced Data Tools: The TidyVerse <ul style="list-style-type: none">• LearnR Tutorial 5: Introduction to the Tidyverse
Feb 21	Data Visualization Part 1: Types of Visualizations <ul style="list-style-type: none">• LearnR Tutorial 6, Topics 1-3: Types of Visualizations and Variables• Assignment 2 Due Friday, February 23rd at 11:59 PM
Feb 26	Data Visualization Part 2: Making Graphs and Figures <ul style="list-style-type: none">• LearnR Tutorial 6, Topics 4-5: Introduction to ggplot

Feb 28	Data Visualization Part 3: Advanced Visualization <ul style="list-style-type: none"> • LearnR Tutorial 6, Topic 6: Advanced ggplot Commands
Mar 4	Data Visualization Part 4: Practice <ul style="list-style-type: none"> • LearnR Tutorial 7: Thinking through Visualization
Mar 6	Functions and Loops <ul style="list-style-type: none"> • Watch the following video: https://www.youtube.com/watch?v=p8tAQx7ijXE • Assignment 3 Due Friday, March 8th at 11:59 PM
Mar 11	No Class (Spring Break)
Mar 13	No Class (Spring Break)
Mar 18	Causality Part 1: Standard Deviations and Z scores <ul style="list-style-type: none"> • TBD
Mar 20	Causality Part 2: Correlation vs. Causation <ul style="list-style-type: none"> • TBD
Mar 25	Causality Part 3: Randomization <ul style="list-style-type: none"> • TBD
Mar 27	Causality Part 4: Observational Studies <ul style="list-style-type: none"> • TBD
Apr 1	Causality Part 5: Observational Studies 2 <ul style="list-style-type: none"> • TBD
Apr 3	Prediction Part 1 <ul style="list-style-type: none"> • TBD • Assignment 4 Due Friday, April 5th at 11:59 PM
Apr 8	Prediction Part 2 <ul style="list-style-type: none"> • TBD
Apr 10	Prediction Part 3 <ul style="list-style-type: none"> • TBD
Apr 15	Prediction Part 4 <ul style="list-style-type: none"> • TBD
Apr 17	In-class Workshop <ul style="list-style-type: none"> • No reading • Assignment 5 Due Friday, April 19th at 11:59 PM

Apr 22	In-class Workshop <ul style="list-style-type: none">• No reading
Apr 24	In-class Workshop <ul style="list-style-type: none">• No reading
Apr 29	In-class Presentations <ul style="list-style-type: none">• No reading
May 3	Final Paper Due at 11:59 PM