POLI 281 Data in Politics I: An Introduction
Spring 2024

Instructor: Silviya Nitsova
Email: snitsova@live.unc.edu

Class meeting days, times, and location:
Section 004: Mondays & Wednesdays at 4.40-5.55pm in Gardner Hall–Room 308

Office hours:
Wednesdays at 12.00-3.00pm in Hamilton Hall–Office 300 or via Zoom (see Sakai for details)

Section Sakai site: https://sakai.unc.edu/portal/site/poli281.004.sp24

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. The broad learning objective for this course is to help students develop the tools they need to be 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 computing environment, with an emphasis on versatile tools such as conditionals, loops, and data visualization with ggplot;

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; and

4. To learn how to acquire new quantitative skills autonomously, such as by reading software documentation.

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 *POLI 381 Data in Politics II: Frontiers and Applications.*

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

Course Requirements and Evaluation

Course assessment is broken down as follows.

PROBLEM SETS (50%) There will be five problem sets due as noted in the schedule below. They are weighted equally. Problem sets are due at 11.55 pm Eastern Time (NC time) on the days indicated.

DATACAMP ASSIGNMENTS (10%) During the course of the semester, you will be required to complete several DataCamp chapters with accompanying exercises. These will be assessed for completion.

FINAL PROJECT (25%) The class has a capstone final project for which students, working in groups, conduct and present an original data analysis on an existing dataset. This project consists of a paper (20%) and a presentation (5%). A separate document specifies final project requirements more completely.

As a default, all final project group members receive the same grade. However, different grades can be assigned when the work was not shared equitably. Additionally, if one group member is significantly disrupting progress (such as by not providing timely responses to communication or by missing scheduled meetings), this person can be removed from the group, and required to complete an alternative final project alone.

PARTICIPATION (15%) Your participation grade has three components. The first component is attendance and participation 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.

The second component is participation in our class's online Piazza discussion forum (see Sakai for a link). On Piazza, you can benefit your participation grade **either** by posting your own questions **or** by providing thoughtful answers to other students' questions. The website keeps track of your activity. These statistics are advisory to the summary participation grade.

The final component is completion of ungraded online checks/quizzes throughout the semester. These checks/quizzes may not happen, but I reserve the right to assign them using Poll Everywhere. The checks/quizzes are ungraded, and are designed to help me gauge overall progress and comprehension in the class.

Missing one or two classes will generally not negatively affect your participation score. **If you find yourself missing more than 3 classes, it is your responsibility to come and talk to me about what's going on.**

Note also that, the weighting scheme above notwithstanding, all assignments must be submitted. If you have not submitted an assignment, you are at risk of receiving a course grade of I (Incomplete) until it is complete.

Final grades for the course will be based on the following grading scale. I reserve the right to make adjustments to individual grades based on overall performance in the course and/or extenuating circumstances.

Percentage Grade	Letter Grade
100-93	A
92-90	A-
89-87	B+
86-83	B
82-80	B-
79-77	C+
76-73	C
72-70	C-
69-67	D+
66-60	D
<60	F

Course Texts and Software

This course has one required textbook:

1. Imai, Kosuke. 2017. Quantitative Social Science: An Introduction. Princeton: Princeton University Press.

In addition, there are software requirements. Students must download and install R , a free statistical computing environment available at http://cran.r-project.org/, as well as RStudio (also free), which is available at www.rstudio.com. (Note that RStudio is rebranding to Posit in an effort to expand beyond the R community.)

Students must also register for DataCamp at www.datacamp.com, a resource that provides tutorial videos and interactive training modules to help learn programming skills. DataCamp is free for students enrolled in this class (I will provide you with login information). Please do not pay for DataCamp access!

Helpful Resources

Below are a number of additional resources you might find useful during the course of the semester or at a later stage in your academic or professional career. I want to emphasize that you do not need to consult any of these additional resources to succeed in POLI 281.

1. UNC's R Open Labs: http://ropenlabs.web.unc.edu.

2. Quantitative Social Science (textbook) online resources: http://qss.princeton.press/student-resources-for-quantitative-social-science.

3. Wickham, Hadley and Garrett Grolemund. 2023. R for Data Science (2e). Sebastapol, CA: O'Reilly. You can access the book for free at https://r4ds.hadley.nz/.

4. Wickham, Hadley. 2019. Advanced R. Second Edition. Boca Raton, FL: Taylor & Francis. You can access the book for free at https://adv-r.hadley.nz/.

Class Policies

Technology in the classroom. Please use technology only for purposes directly related to class. If I see you using your computer for extraneous purposes–e.g. chat, social media, ESPN, or checking the news–this will negatively impact your participation grade. This policy is motivated by evidence that extraneous technology use decreases learning not only for the user, but also for the people around them.

Cooperation and academic integrity. In a class setting, cooperative work has both benefits and pitfalls. Peers learn a lot by explaining things to each other. But it can also be easy to stumble into a passive mindset where you are not really *assimilating* 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 (mind that 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 their own path to the solution. **In our class, copying and pasting another person's computer code is potentially tantamount to plagiarism. This is true if you copy from any source, including a classmate or former student, an online source, or ChatGPT or a similar AI product.** Even for work designated as cooperative, you must write your code individually. Unless I have given you explicit permission for some special reason, do not do it. Copied code is surprisingly easy to detect. Be assured that if I identify a case of cheating or plagiarism, I will handle it 100% "by the book."

Class environment. I aim to make this class a safe and open environment. Part of that is respecting all questions, comments, and experiences that others choose to share. Not everyone will share your view, but if we all agree to listen to and respect each other, then this class will be a positive environment where all opinions are welcome. Offensive language (derogatory, racist, sexist, homophobic, etc.) will not be tolerated.

Students with disabilities. If you think you need an accommodation for a disability, please let me know. Some aspects of the course and its assignments may be modified to facilitate your success. I will work with the Office of Accessibility Resources and Services to determine appropriate accommodations. I will treat any information you provide as confidential. Barring unusual circumstances, I require notice of a need for accommodation within the first two weeks of the semester.

Deadlines. Late submission or completion of assigned work will result in a 10-point grade penalty for each additional day late. That being said, I want to emphasize that I will strive to be as flexible as possible. **If you think you may need an extension on any assignment, please reach out to me. As long as I hear from you before the deadline, I would be happy to grant an extension.**

Grade grievances. Requests for regrades have a time window. They cannot be submitted until at least 48 hours have passed since the assignment was returned (a cool-down period), and then they will only be accepted within three weeks of an assignment being returned (a statute of limitations). To request a regrade, you must submit a written memo (two pages max) explaining what aspect of your original grade you think was in error.

Absences. Requests for an excused absence should come via email (for clear documentation). The request must come as far in advance of the absence as possible if the absence is foreseen, or as soon after as possible if it is not foreseen. Per university policy, only your academic advisor can provide an official final exam excused absence.

This course includes several class periods focused on workshopping final projects. Absences on these days are especially disruptive, since they interfere with your group's ability to work on the final project together. For this reason, absences on these dates (noted below) will be excused only for reasons of religious observance, illness, or family emergencies. Unexcused absences on these days cause your individual final project grade to receive an automatic 5-point penalty.

Schedule changes. I occasionally modify the schedule to accommodate lesson plans that took more or less time than expected and to address relevant group needs. Assignment due dates also occasionally change, though they are only ever postponed–not advanced.

Email. I usually respond to emails within 48 hours (except on weekends). However, please limit your use of email to issues that are private, or at least specific to you. For matters that are not private and where other students might want to see the response, please use Piazza. Note that Piazza permits anonymous posting (the post will be anonymous to other students, but not to me).

I typically will not use email to repeat information that was missed because of an absence; I will direct you to correspond with a classmate.

Communicating concerns. If you have any questions or concerns regarding the course as a whole or the progress you are making in it, please talk to me. I am always happy to go over comments and strategies to improve. The earlier you make your concerns known, the better I will be able to (help you) address them.

Course Schedule

Week	Date	Topic	Readings	Assignments
1 W	1/10	Course Introduction		• Register for Poll Everywhere
Part I: Using Data to Describe the World				
2 M	1/15	No Class (MLK Day)		
				Continued on next page

Week	Date	Topic	Readings	Assignments
2 W	1/17	R: Getting Set Up I		• Download and install both R and RStudio • Register for DataCamp and Piazza
3 M	1/22	R: Getting Set Up II	• <u>Skim</u> *QSS*, Ch. 1	
3 W	1/24	R: Basic Tools I		• DataCamp: Introduction to R, Ch. 1-3
4 M	1/29	R: Basic Tools II		• DataCamp: Introduction to R, Ch. 4-6
4 W	1/31	R: Intermediate Tools I		• DataCamp: Intermediate R, Ch. 1 (Conditionals)
5 M	2/5	R: Intermediate Tools II		• DataCamp: Intermediate R, Ch. 2 (Loops) • **Problem Set 1 due**
5 W	2/7	Variable Types and Descriptive Statistics	• <u>Read</u> *QSS*. Ch. 1	
6 M	2/12	No class (Well-being day)		
6 W	2/14	Review		
7 M	2/19	Data Visualization I		• DataCamp: Introduction to the Tidyverse, Ch. 1-4 • **Problem Set 2 due**
7 W	2/21	Data Visualization II		
8 M	2/26	Data Visualization III		
8 W	2/28	Data Visualization IV		
Part II: Data Analysis as Problem Solving				
9 M	3/4	Causality I	• Bertrand & Mullainathan, pp. 991-997 (Sakai) • *QSS*, pp. 32-48	

Schedule – continued from previous page

Part II: Data Analysis as Problem Solving

Week	Date	Topic	Readings	Assignments
9 W	3/6	Causality II	• Gerber et al., pp. 33-38 (Sakai) • *QSS*, pp. 48-54	• **Problem Set 3 due**
10 M	3/11	No Class (Spring Break)		
10 W	3/13	No Class (Spring Break)		
11 M	3/18	Causality III	• Card & Krueger, pp. 772-778 (Sakai) • *QSS*, pp. 54-69	
11 W	3/20	Causality IV	• Mosteller (Sakai)	
12 M	3/25	Review/Catch up day		
12 W	3/27	Measurement	• *QSS*, pp. 75-111	
13 M	4/1	Prediction I	• *QSS*, pp. 123-139	• **Problem Set 4 due**
13 W	4/3	Prediction II	• *QSS*. pp. 139-148	
14 M	4/8	Prediction III	• *QSS*, pp. 148-161	
14 W	4/10	Prediction IV	• *QSS*, pp. 161-170	
Part III: Putting Data to Work				
15 M	4/15	Storytelling with Data	• TBD	• **Problem Set 5 due**
15 W	4/17	In-class Final Project Workshopping		
16 M	4/22	In-class Final Project Workshopping		
16 W	4/24	In-class Final Project Workshopping		
17 M	4/29	Review and Conclusions		
Final	**5/9 12 pm**	Final Project Presentations (a "non-traditional format" final exam)		• **Final projects due**

Schedule – continued from previous page