

POLITICAL SCIENCE 281:
DATA IN POLITICS I: AN INTRODUCTION
Syllabus¹

Spring 2021

Instructor: Simon Hoellerbauer
Email: hoellers@unc.edu

Class Meetings: T/Th 11:00 AM – 12:15 PM (ET)
Remote-only Course - Synchronous Lectures

Office Hours (Virtual): T 1:30-3:00 PM (ET); Th 3:00 - 4:30 PM (ET); or by appointment.

Zoom links can be found on Sakai page.

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 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; and
3. To develop students' ability to use data to answer important social scientific questions.
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).

¹Text, ideas, and topics for this syllabus adopted and adapted from Stephanie Shady, Lucy Britt, Jacob Smith, Lucy Martin, and Tim Ryan.

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 fulfills counts as a course in research methods (required for completing the Political Science Bachelor's degree).

Class Structure

This class is comprised of lectures, in-class activities, (some) readings, and assignments. It is highly interactive. I will rarely lecture for a full class-period — we learn best by doing, after all. We will use Poll Everywhere for for some in- and out-of-class activities, mainly comprehension checks. We will use Piazza as a Wiki-style question and answer/discussion forum. This makes it easier for the instructor and ULAs to answer common questions and also allows students to crowd-source answers. There are some readings — most often from the textbook (see [below](#)), but also newspaper articles and some academic journal articles.

Unless indicated, you are expected to have completed the readings and assignments by the date they are listed in the course schedule (see [below](#)).

Assignments and Grading

Course assessment is broken down as follows:

HOMEWORK (50%) There will be five homework assignments due as noted in the schedule [below](#). They are weighted equally. These homework assignments are due by 11:59pm on the days indicated, unless we decide something different in class. Each is weighted equally.

DATA CAMP (10%) During the course of the semester, you will be required to complete several Data Camp (see Class Texts and Software section [below](#)) lessons with accompanying exercises. These will be assessed for completion.

FINAL PROJECT (25%, broken up as follows) 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 preparation for (and actual attendance in) class. The best way to do this is to come to our synchronous class. While I will be recording lectures, watching recorded lectures asynchronously is at best a partial substitute for this, as a large part of in-class participation is in group work.

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, too, are advisory to the summary participation grade.

The final component is completion of ungraded online quizzes throughout the semester. These quizzes may not happen, but I reserve the right to assign them. These 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.

Grading Scale

I will use the following grading scale:

- A: 100-92.5; A-: less than 92.5-89.5
- B+: less than 89.5-86.5; B: less than 86.5-82.5; B-: less than 82.5-79.5
- C+ less than 79.5-76.5; C: less than 76.5-72.5; C-: less than 72.5-69.5
- D+: less than 69.5-66.5; D: less than 66.5-59.5
- F: less than 59.5-0

Some professors make subjective decisions about rounding up or down in certain ranges (92-93, for example). This has always struck me as unfair and, well, subjective. This grading scale makes it clear exactly what percentage you need to get for a particular letter grade. I will not do any further rounding beyond this.

Class Texts and Software

This course has one required textbook:

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

QSS in the [schedule](#) refers to this textbook.

This book has additional online resources, including more R-learning support, which you can access here:

<http://qss.princeton.press/student-resources-for-quantitative-social-science>

I usually do not assign textbooks, but they are very useful for this kind of course. This text will be a good resource for you going forward, and not just in this class. This book is thankfully relatively inexpensive. Professor Imai is one of the best, well-known political methodologists out there (he's also a pretty nice guy!). You will not need this book until the second part of the course (indicated in the schedule at the end of this syllabus).

In addition, there are software requirements. Students must download and install R, a free statistical program available at <http://cran.r-project.org/>, as well as RStudio (also free), which is available at www.rstudio.com. They also must register for Data Camp (www.datacamp.com), a resource that provides tutorial videos and interactive training modules to help learn programming skills. Data Camp is free for students enrolled in this class. (I will provide you with login information.) **Please do not pay for Data Camp access!**

Helpful Resources

You are in no way expected or required to read and/or use these resources.

For Learning R:

- Learn R at UNC through R Open Labs/beginR: <http://ropenlabs.web.unc.edu/>
- Wickham, Hadley² and Garrett Golemund. 2017. *R for Data Science*. Sebastopol, CA: O'Reilly. — **DO NOT BUY THIS BOOK**; you can access its contents for free online here: <https://r4ds.had.co.nz/>.

For Learning More About R As A Programming Language:

- Wickham, Hadley. 2019. *Advanced R*. Second Edition. Boca Raton, FL: Taylor & Francis. — **DO NOT BUY THIS BOOK**; you can access its contents for free online here: <https://adv-r.hadley.nz/>.

For Statistics:³

- Diez, David, Mine Cetinkaya-Rudel, Christopher Barr, and OpenIntro. 2019. *OpenIntro Statistics*. Fourth Edition. — **DO NOT BUY THIS BOOK**; you can get a PDF version for free here: <https://www.openintro.org/book/os/>.

²Hadley Wickham is a statistician and the creator of the `ggplot2` package and the tidyverse. He's from New Zealand, hence the `.nz` in the links here and below.

³We won't be covering statistics or probability in much depth in this class, so this is only if you are interested.

General Tips:

If you ever get an error message and do not know why or what it is trying to tell you, it is always a good strategy to enter the error message into Google. It is very likely that someone has had a similar problem. Doing so will very often bring you to [Stack Overflow](#), so you can also just start there if you want.

Class Policies

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're 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. (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. **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 (there is software designed to detect it).

Academic dishonesty—violating the standards of integrity all students pledge to adhere to as a condition of enrollment at UNC—results in being referred to the Honor Court. Students with questions about what qualifies as academic dishonesty should ask the instructor and can also consult the Honor Court Website at <http://honor.unc.edu>. Students may use any citation style they choose when writing papers so long as they remain consistent within any one paper.

Late Work Policy/Deadlines

Deadlines are given in Eastern Time (so UNC time), but I realize that some of you may be in different time zones. Please try to get assignments in by the proper time, but I will be flexible.

That said, **if you think you may need an extension on any assignment, please do not hesitate to reach out.** This has always been my policy, but it is especially critical during the current crisis. This is a stressful time; I do not want this class to add to your stress unnecessarily. The reason for an extension does not have to be a COVID-related; I

will not ask you to explain or to present any documents or anything. Please do try to let me know as far ahead as possible, purely for administrative purposes, but I recognize that this will not always be possible.

If you have not talked to me beforehand⁴, I will deduct one letter grade (10 percentage points) per day that an assignment is late from the maximum grade you can receive. I will then grade your paper as normal and weight it so that it could not exceed this new maximum grade. As such, if you turn in an assignment one day late, the highest grade you can receive is a 90. If you then receive an 85 on the assignment, your actual grade will be $.85 * 90 = 76.5$.⁵ The reason I do this is because it makes it seem less arbitrary and helps me separate out where you lose points, in ways that are not related to the lateness of your paper.

Contact Policy

You are encouraged to come to my office hours, listed above, and to contact me with any questions you may have, even if you just want to chat. I will make use of the scheduling system Calendly. Please use this link: https://calendly.com/poli281_sp21/office-hours to schedule office hours (link is also on Sakai page).

I am also available by appointment if you are unable to meet during the official office hours, but please do try to give me as much advance notice as possible. I really want to be a resource for you this semester, so if you have anything you want to talk about, please do not hesitate to come to my office hours. If I have to change my office hours for any reason, I will let you know.

My email is hoellers@unc.edu. I will try to respond as soon as possible, although I cannot guarantee same day response. Therefore, I encourage you to ask me questions about assignments and papers as far in advance as possible, which will hopefully help you get in the habit of working on assignments well before they are due.

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).

If you miss one or two classes, it could set you back significantly, or if you do not start working on the assignments soon enough, you may find yourself running out of time when the due date comes around. Please reach out to me if you have any issues, problems, or concerns. In addition, if anything or anyone makes you feel uncomfortable, and you feel comfortable doing so, please come talk to me so that I can do my best to assist you. Finally, **in line with university policy, I cannot discuss grades over email**. If you have questions about grades, you must come to office hours or set up an appointment.

⁴AKA, please reach out if you need an extension, I cannot emphasize this enough.

⁵If I were not doing the weighting, you would get 75 (100 - 10 - 15). This format will always benefit you, even if only by a little.

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.

Teaching Philosophy

I view my role as a teacher as a support person for you, my students. Because of my background and education, I have knowledge that I will strive to communicate with my you, which is why lectures do form an important part of this course. My primary goal as a teacher, however, is to make you feel engaged and active and to help you learn skills that you will be able to use outside of the contexts of this course and even of this field of study. As such, I believe that active engagement with the course material is essential to helping you learn, and I structure the course in such a way that there are plenty of ways in which to participate and be active, as I recognize that not all students learn in the same way. At the same time, I do not believe that surface-level skimming of a topic is all that useful; therefore, this class is more detail-oriented than other introductory courses may be, without being overwhelming. Finally, I am always open to feedback—I want to make sure that you are getting both what you want and need from this course.

Student Accessibility

Any student with a documented condition who needs to arrange reasonable accommodations should contact the instructor and Accessibility Resources (<https://accessibility.unc.edu/>). Accessibility Resources can be reached by phone at 919-962-8300 or by email at accessibility@unc.edu.

Disclaimer

I reserve the right to make any and all necessary changes to this syllabus. If I do make any changes, I will of course inform you about them. Please always refer to the most updated syllabus, which I will upload to the course Sakai page under “Syllabus,” when seeking information.

Course Schedule

Week	Date	Lecture Topic	Readings	Assignments
1 Tu	1/19	Course Introduction	<ul style="list-style-type: none"> • Lohr 2009 • Vance 2009 • Rogers Cook 2019 	<ul style="list-style-type: none"> • Register for Poll Everywhere
1 Th	1/21	The Possibilities and Perils of Data Science	<ul style="list-style-type: none"> • Christensen et al. 2019, Introduction 	<ul style="list-style-type: none"> • Register for Piazza • Complete "Getting to Know You Questionnaire"
2 Tu	1/26	R: Getting Set Up I		<ul style="list-style-type: none"> • Register for DataCamp • Download and install both R and RStudio
2 Th	1/28	R: Getting Set Up II	<ul style="list-style-type: none"> • Skim QSS, Ch. 1 	
3 Tu	2/2	R: Basic Tools I		<ul style="list-style-type: none"> • DataCamp: Introduction to R, Ch. 1-3
3 Th	2/4	R: Basic Tools II		<ul style="list-style-type: none"> • DataCamp: Introduction to R, Ch. 4-6
4 Tu	2/9	R: Intermediate Tools I		<ul style="list-style-type: none"> • DataCamp: Intermediate R, Ch. 1 (Conditionals)
4 Th	2/11	R: Intermediate Tools II		<ul style="list-style-type: none"> • DataCamp: Intermediate R, Ch. 2 (Loops) • Homework 1 due
5 Tu	2/16	No class, No office hours (Wellness Day)		
5 Th	2/18	No class, Yes office hours (Ice Storm)		
6 Tu	2/23	R: Intermediate Tools III		<ul style="list-style-type: none"> • DataCamp: Intermediate R, Ch. 3 (Functions)
6 Th	2/25	R: Wrap-Up	<ul style="list-style-type: none"> • Read QSS, Ch. 1 	

Continued on next page

Calendar – continued from previous page				
Week	Date	Lecture Topic	Readings	Assignments
7 Tu	3/2	Causality I	<ul style="list-style-type: none"> • Bertrand & Mul-lainathan, pp. 991-997 (Sakai) • <i>QSS</i>, pp. 32-48 	• Homework 2 due
7 Th	3/4	Causality II	<ul style="list-style-type: none"> • Gerber et al., pp. 33-38 (Sakai) • <i>QSS</i>, pp. 48-54 	
8 Tu	3/9	Causality III	<ul style="list-style-type: none"> • Card & Krueger, pp. 772-778 (Sakai) • <i>QSS</i>, pp. 54-69 	
8 Th	3/11	No class, No office hours (Wellness Day)		
9 Tu	3/16	Causality IV	<ul style="list-style-type: none"> • Mosteller, entire (Sakai) 	
9 Th	3/18	Causality Workshop	TBD	
10 Tu	3/23	Data Visualization I		<ul style="list-style-type: none"> • DataCamp: Introduction to the Tidyverse, Ch. 1-4
10 Th	3/25	Data Visualization II		• Homework 3 due
11 Tu	3/30	Data Visualization III		
11 Th	4/1	Measurement	<ul style="list-style-type: none"> • <i>QSS</i>, pp. 75-111 	
12 Tu	4/6	Prediction I	<ul style="list-style-type: none"> • <i>QSS</i>, pp. 123-139 	
12 Th	4/8	Prediction II	<ul style="list-style-type: none"> • <i>QSS</i>, pp. 139-148 	• Homework 4 due
13 Tu	4/13	Prediction III	<ul style="list-style-type: none"> • <i>QSS</i>, pp. 148-161 	
13 Th	4/15	Prediction IV	<ul style="list-style-type: none"> • <i>QSS</i>, pp. 161-170 	• Homework 5 due
14 Tu	4/20	Flex Day — Topic TBD		
14 Th	4/22	In-class Final Project Workshopping		
15 Tu	4/27	In-class Final Project Workshopping		

Continued on next page

Calendar – continued from previous page					
Week	Date	Lecture	Topic	Readings	Assignments
15	Th	4/29	In-class Workshopping	Final Project	
16	Tu	5/4	Review and Conclusions		
Final	5/7, 12pm	Final Project Presenta- tions			• Final projects due