
PSCI 205
Data Analysis II

Spring 2022
Tues/Thurs 11:05-12:20, Dewey 2110E

Prof Curt Signorino
303 Harkness Hall
Office Hours: Wed 10am-12
curt.signorino@rochester.edu

TAs:
Fred Liu
Office hours: Tues 2-3pm
sliu85@u.rochester.edu

Ha Nguyen
Office hours: TBD
hnguy28@u.rochester.edu

COURSE DESCRIPTION: This course builds on PSCI 200, Data Analysis I, taking the linear regression model as its starting point. We will explore various statistical techniques for analyzing a world of data that is relevant to political science in particular, and to the social sciences more broadly. In addition to the linear regression model, we will examine models for binary data, durations, counts, censoring and truncation, self-selection, discrete choice, and strategic choice, among others. These models will be applied to topics such as international conflict, civil war onset, parliamentary cabinet survival, international sanctions, campaign contributions, and voting. Students will be taught how to (1) frame research hypotheses, (2) analyze data using the appropriate statistical model, and (3) interpret and present their results. Statistical analysis will be conducted using R.

COURSE MEETING & CREDITS. This course follows the College credit hour policy for four-credit courses. We will meet twice a week (Tues & Thurs) for 1.5 hour sessions. These will be held in person in Dewey 2110E or via Zoom, depending on University Covid policy. There is no separately designated day for labs. Rather, the normal Tues/Thurs sessions will be a mix of lecture and labs. During the labs, students will receive computer instruction, analyze data, and start on new homework problems. The remaining credit hour is fulfilled through independent reading and completion of the homeworks.

All classes (including Zoom sessions) will be recorded and uploaded to Blackboard.

PREREQUISITES: Students should have taken a course (such as PSC 200, ECO 230, STT 212, STT 213, or STT 214) that introduces them to probability, hypothesis tests, confidence intervals, and bivariate regression. Students should have a basic familiarity with R. Calculus and matrix algebra are not required. If you need to refresh yourself concerning R and/or introductory statistics, I have uploaded review material to Blackboard.

GRADING: Course grades will be based on a series of homeworks (70%), a final exam (25%), and participation/attendance (5%). Taking the final exam is required to pass the course.

Homeworks. Homeworks will typically be handed out (via Blackboard) in class, at which point a short “lab” will be held. The lab allows students to read over the HW, download any data, begin working on the HW, and ask me questions about the HW before taking it home to complete it. Homeworks will normally be due by the start of lecture one week after they are handed out. Students should submit their homework answers, properly formatted, via Blackboard. Homework grades will also be posted on Blackboard. All assignments are to be completed individually. Be sure to read the PSCI 205 course academic honesty policy concerning HW completion.

Late Homework Submissions. It is important that students submit their HWs on time. We do our best to provide grades and answer keys in a timely manner. Late HW submissions can hold up that process, in which case the class will not have as much time to review previous HW answers before starting new HWs.

That said, life happens. If you need to attend a major event – e.g., a conference, a job interview, an athletic tournament for UR, etc – contact me ahead of time. As long as this is a one-time occurrence, you will likely be given permission to turn in the current HW after the deadline. Similarly, if you fall ill, email me immediately and we’ll try to work something out. In either case, you should expect that the deadline extension will be no more than 7 days after the original due date, usually less.

In all other cases, late assignments will be penalized (as a percentage of total points possible) as follows:

Lateness	Penalty
Up to 10 hrs late	5%
10hrs to 24hrs (≤ 1 day)	10%
24hrs to 48hrs (≤ 2 days)	20%
48hrs to 72hrs (≤ 3 days)	30%
>3 days but ≤ 4 days	40%
>4 days but ≤ 5 days	50%
>5 days but ≤ 6 days	60%
>6 days but ≤ 7 days	70%
>7 days	100%

Participation/Attendance. Students who regularly attend class generally perform better on homeworks and exams. Class attendance will be noted and factored into the course grade. During Zoom classes, students are expected to have their video on.

READINGS: There is no perfect text for this course. Instead, I will assign readings from various texts and articles and supplement those with lecture notes. Except for the Wooldridge text, all texts and articles will be available as pdf’s on Blackboard. Texts used for this course will include

- John Verzani. [*SimpleR: Using R for Introductory Statistics*](#). This is an open source pdf that introduces students to using R for statistics. There is little to no math. It focuses on the mechanics of data analysis, hypothesis testing, and linear regression using R.

If you feel rusty concerning R, please work through pages 1-24 before the end of the first week of class.

- G. Jay Kerns. *Introduction to Probability and Statistics using R*. 3rd ed. The topics overlap quite a bit with Verzani. However, Kerns is much more mathematical, including the use of calculus. The open source pdf is available as part of R's IPSUR package.
- David M. Diez, Christopher D. Barr, and Mine Cetinkaya-Rundel. *OpenIntro Statistics*. 3rd ed. For some, this will be a more user-friendly version of Kerns, without any calculus or more advanced math.
- Jeffrey M. Wooldridge. *Introductory Econometrics*. 7th ed. More advanced than *OpenIntro Statistics*. Does not demonstrate with R. Does not use calculus or linear algebra, but does provide some proofs.
- Marco R. Steenbergen. 2008. *Discrete Choice Models for Political Analysis*. Advanced Political Methodology Lecture Notes. (pdf on Blackboard)

STATISTICAL SOFTWARE: We will use R and RStudio for our statistical analysis. R is open source and free. There are versions for Mac OSX, Windows, and Linux. You can download it from <https://cran.r-project.org/>. Additionally, we will use [RStudio](#) as a graphical interface for R. RStudio is free for students to download. Instructions for downloading R and RStudio are provided on Blackboard under the Prerequisite R Review folder on the Course Home Page.

IMPORTANT DATES:

First lecture: Thursday, Jan 13

Spring break: March 5-13

Last lecture: Tuesday, Apr 26

Final exam: TBD, finals week

COURSE OUTLINE:

1. Course Introduction

2. R, RStudio, & RNotebook

R vs RStudio, calculations, variables, classes, vectors, matrices, logical operations, data frames, loading data sets, descriptive statistics, tables, plots, help, knitting an R script, RNotebook

Required (R): Verzani, pp. 1-24. IPSUR, Ch 2.

3. Bivariate Linear Regression: Estimation & Inference

It's a line!, estimating the coefficients, t tests, CIs, r^2

Required (math): *OpenIntro*, Ch 7. Wooldridge, Ch 2.1-2.4. *IPSUR*, Ch 11.1-11.3.3.
Required (R): Verzani, pp 24-31, 77-83.

4. Multiple Regression I: Estimation & Inference

It's a plane!, research hypotheses, estimating and interpreting coefficients and standard errors

Required (math): *OpenIntro*, Ch 8.1-8.3. Wooldridge, Ch 3.1-3.2, Ch 4.2.

Required (R): Verzani, pp 84-89.

Optional: *IPSUR*, Ch 12.1-12.3

5. Multiple Regression II: Model Fit & Joint Hypotheses

R^2 , F test of the regression, nested models, multicollinearity, statistical vs substantive significance

Required (math): *OpenIntro*, Ch 8.1-8.3. Wooldridge, Ch 3.2h, Ch 4.5-4.6.

Required (R): Verzani, pp 84-89.

Optional: *IPSUR*, Ch 12.1-12.3

6. More on R^2

Is R^2 evil?

Required: Wooldridge, Ch 6.3.

7. Dummy variables

Factor variables with >2 levels, omitted/baseline category

Required: Wooldridge, Ch 7.1-7.3

8. Polynomial & log transformations

Polynomial regression, log-linear & linear-log regression models

Required: Wooldridge, Ch 2.4, Ch 6.2.

9. Predictions & Interactions

Plotting fitted values for more complicated regressions

Required: Wooldridge, Ch 6.2c, Ch 7.4a.

10. Diagnostics

Checking whether our modeling assumptions are incorrect, linearity, heteroskedasticity, outliers

Required: *OpenIntro* Ch 8.3. *IP SUR*, 11.4-11.5. Wooldridge, Ch 9.5.

11. Maximum Likelihood

Intuition, one parameter model, multiple parameters, Normal, Binomial

Required: Gary King. 1998. *Unifying Political Methodology*. Ch 2 & 4.

12. Binary Data

Logit, probit, research hypotheses part II, derivations, nonlinear $E(Y)$, interpretation, implicit interaction effect, CIs

Required: *OpenIntro*, Ch 8.4. Steenbergen, Ch 2. Wooldridge, Ch 17.1.

13. Count Data

Poisson, negative binomial regression, hurdle models

Required: Wooldridge, Ch 17.3.

Beaujean & Morgan. 2016. "Tutorial on Using Regression Models with Count Outcomes..."

Zeileis et al. "Regression Models for Count Data in R."

14. Survival Models

Exponential, Weibull, Kaplan Meier, Cox proportional hazard model

Required: TBD

15. Censoring and Truncation

Tobit model, truncated normal

Required: Wooldridge, Ch 17.2, Ch 17.4.

Arne Henningsen. "Estimating Censored Regression Models in R using the censReg..."

Possibly, but unlikely...

16. Ordered Logit/Probit

Required: Steenbergen, Ch 3.

17. Discrete Choice

Random utility, multinomial logit, conditional logit

Required: Steenbergen, Ch 4-5.

18. If Time Permits...

- Writing your own function in R
 - Maximum Likelihood Estimation using `maxLik()`
-

OTHER IMPORTANT ITEMS

Course Organization. The course organization may be adjusted/optimized during the semester according to the pace of learning and the priority of topics. Students are responsible for attending lectures and maintaining an awareness of any changes to the course materials, homework requirements, or exam dates.

Student Disability Accommodation. I am happy to work with any student who requires an accommodation due to a disability. However, I am not authorized to grant any accommodations on my own. It is important that students first contact the Office of Disability Resources. They will discuss any barriers a student is experiencing, explain the process for establishing academic accommodation, and then authorize me to provide a specific level of accommodation. You can reach the Office of Disability Resources at disability@rochester.edu or (585) 276-5075.

Academic Honesty. Students are expected to be familiar with the University’s policies on [academic honesty](#). I have provided additional course-specific academic honesty policies on Blackboard under the Course Academic Honesty tab. If I suspect a student has violated any of these policies, I am required to report the violation. Punchline: don't cheat. If in doubt about what is acceptable behavior concerning completing an exam or homework, just ask me.

During the first week of class, please review both the University policies and the course policies. You must confirm that you have read and accept these policies by completing the Acceptance of

Academic Honesty Policy activity at the bottom of the Course Academic Honesty page on Blackboard.

University COVID-19 Statement: In-Person Classes

Most of this is taken verbatim from the University syllabi requirements (<https://www.rochester.edu/college/faculty/resources/syllabi-requirements.html>) and pertains to in-person classes.

The University is committed to protecting the health and safety of the entire community — students, faculty and staff. For this reason, it is mandatory that everyone wear a mask in University buildings and observe appropriate social distancing, including classrooms. Masks have been provided to students, faculty and staff and classrooms have been specifically assigned to allow for social distancing to support these requirements. You must wear a mask appropriately — e.g. over the nose and mouth — if you are attending class in person, and you must do this for every class session and for the entire duration of each class session. If you fail to do this, you will be politely reminded of the requirement and then asked to leave if you do not comply.

If you do not want to wear a mask, you may consider taking the course remotely (online). This may require you to complete a set of online requirements different from the in-person requirements, although these will be equivalent in their learning objectives.

Students who refuse to adhere to requirements for mask wearing or social distancing the course will be in violation of the COVID-19 Community Commitment and will be referred to the Student Conduct system through a COVID-19 Concern Report. Such referrals will lead to student conduct hearings and may result in disciplinary action.

Students who feel unable to wear a mask may contact the Office of Disability Resources to explore options for accommodations. Students requiring accommodations may be asked to participate in the course through synchronous or asynchronous learning as part of this accommodation.

Finally, until further notice, eating is prohibited in the classroom.

Attending Classes via Zoom

The online version of class is always a work in progress. However, there are a few things you can do to help it run more smoothly for everyone.

- Find a quiet place to use Zoom. Try to find a spot in your room or home that is quiet, doesn't get interrupted by others, and doesn't have distractions.
- Make sure your audio and video are working. Before you join the Zoom session, take a second to verify that your mic, audio output (e.g., computer speakers or headphones), and video are all working.

- Calibrate your volume. This will take some work on everyone's part – including mine. I want to hear you, but I also don't want my ears blown out any more than they already are!
- Join the class a few minutes early to get yourself situated.
- MUTE your audio. In general, unless you are asking or answering a question, mute your mic. Small background noises can cause the audio and spotlight to jump around.
- Attend with your video turned on. It's very helpful for me to have visual feedback. Students who attend with the video on also tend to be more engaged during class.
- Use the Chat feature while I'm lecturing. If you have a question or want to speak while I'm lecturing, type it in the chat area. I will do my best to get to it promptly. If, for whatever reason, I appear not to have seen it – or forgot about it – copy and paste it again. If you really, really need to get my attention quickly, unmute your audio and just interrupt me.
- Remember that we can see you. When attending class, please minimize distracting movements. Want to enjoy a cup of coffee or tea? No problem. But please don't start your Zumba workout or hold a conversation with your roommate.
- If you need to use the restroom or refill your coffee, no worries – and no need to notify me.

Updated: 1/7/22