

# **500 Syllabus for Spring 2024**

Thomas E. Love

2024-01-11

# Table of contents

<b>Key Information</b>	<b>4</b>
Course Website . . . . .	4
Working With This Document . . . . .	4
Course Text . . . . .	4
Getting Help . . . . .	5
Emergency? . . . . .	5
<b>1 Description</b>	<b>6</b>
1.1 Prerequisites . . . . .	6
1.2 Everything is on the Web . . . . .	6
<b>2 Professor Love</b>	<b>8</b>
2.1 A More Complete Biography . . . . .	8
2.2 Email . . . . .	9
2.3 Name and Pronouns . . . . .	9
2.4 Web . . . . .	10
2.5 Offices . . . . .	10
<b>3 Teaching Assistant</b>	<b>11</b>
3.1 About Leila . . . . .	11
3.2 TA Office Hours . . . . .	11
<b>4 Assignments</b>	<b>12</b>
4.1 Deadlines . . . . .	12
4.2 The Course Project . . . . .	12
4.3 An “Observational Studies in Action” presentation . . . . .	12
4.4 Lab Assignments . . . . .	13
4.5 A Note on Feedback . . . . .	13
4.6 Grading . . . . .	13
<b>5 Advice for Graduate Students</b>	<b>14</b>
5.1 On Graduate School . . . . .	14
5.2 On Seeking a Job . . . . .	14
<b>6 On Writing, Presenting &amp; Communicating</b>	<b>16</b>
6.1 Campus Resources . . . . .	16

6.2	Advice from Other People . . . . .	16
6.3	A Few Tips from Professor Love . . . . .	17
<b>7</b>	<b>Other Policies</b>	<b>19</b>
7.1	Statement on Disability . . . . .	19
7.2	Incomplete Grades . . . . .	19
7.3	Laptop and cell phone use . . . . .	20
7.4	Diversity and Inclusion . . . . .	20
7.5	Academic Integrity . . . . .	20
7.6	Plagiarism . . . . .	21

# Key Information

This is the Spring 2024 syllabus for PQHS / CRSP 500: Observational Studies, taught by Professor Thomas Love. The course is given on Thursday mornings, in room 1217 of the Wolstein Research Building at Case Western Reserve University. The first class begins at 8:30 AM on Thursday 2024-01-18.

## Course Website

The course website, with links to everything else you'll need, is at <https://thomaseLove.github.io/500-2024>.

## Working With This Document

1. This document includes multiple sections. Use the table of contents to navigate, and the search box as needed.
2. A PDF download of this document is available by clicking the Acrobat icon next to the title at the top left of the screen.
3. The document will be updated occasionally through the semester.

## Course Text

During the course, we will read Paul Rosenbaum's 2023 book *Causal Inference*, published by MIT Press, which is available [as an e-book or in paperback](#) for under \$15. The ISBN-13 number is 978-0262545198. The schedule for reading the book is part of the Course Calendar.

Additional reading materials are scheduled on [the Calendar](#), and are gathered on the [Sources page](#) on our website.

## Getting Help

To get help with anything related to the course, email Dr. Love and the TA, Leila Hojat, at **500-help at case dot edu**.

## Emergency?

What if my situation changes (for any reason) during the semester?

If you need to miss a session, just review the materials (including the class recording, if available) and move on, and please feel encouraged to ask for help.

If an emergency arises that will keep you from timely completion of an assignment, send an email to us saying the following:

I have an emergency that will keep me from completing (*list things you're worried about completing*) in a timely fashion. I will contact you again when I am able to make plans.

- We will respond with “OK” and temporarily excuse you from things until **you** send a “re-connection” email to make plans for making up the work you missed, at which time he’ll evaluate the situation and together you can make that plan. At no time do you owe us any information about your health or the nature of the emergency.

# 1 Description

PQHS 500 (cross-listed as CRSP 500) is a one-semester advanced course in modern design and analysis of observational studies, usually in the context of comparative effectiveness research, with a special emphasis on issues that arise in clinical and health services research.

An observational study is an empirical investigation of treatments, policies or exposures and the effects that they cause, but it differs from an experiment in that the investigator cannot control the assignment of treatments to subjects. This course is designed to introduce design, data collection and analysis methods appropriate for scientists engaged in observational studies, and will prepare students to design and interpret their own studies, as well as those of others in their field. Technical formalities will be minimized, and the presentations will focus on the practical application of methods and strategies.

## 1.1 Prerequisites

People take this course with a wide range of backgrounds and a common interest in using data effectively in research related to biology or medicine. All CWRU students who feel up to it are welcome, regardless of their field of study or prior experience. Students with a working knowledge of R, multiple regression, and some familiarity with logistic regression, should be well prepared.

I strongly encourage people to first complete the PQHS 431-432 sequence, or at least to be substantially familiar with the use of R and RStudio (as well as, ideally, R Markdown and Quarto) on the coding side, and with multivariate regression modeling and in particular logistic regression on the statistical side. Those who are unsure if they are well prepared for the course should contact the instructor for some advice.

## 1.2 Everything is on the Web

<https://thomaselove.github.io/500-2024> is the place to go for everything related to this course. Please visit any time you need something. I update the web site frequently. You'll find links on the top menu for the individual elements of the course, including:

- The Course Syllabus (this document)

- The [Course Calendar](#)
  - The Calendar includes links to the slides and other materials for each class.
  - The Calendar is also the final word on deadlines for all assignments.
- Material about R and Data, including:
  - details on [software](#), specifically installing R and RStudio and the R packages we'll use.
  - the [500-Data](#) page for downloading data, code and templates for assignments.
- The [Sources page](#) for additional readings assigned through the semester.
- Material and instructions for your Assignments, including:
  - The [Course Project](#)
  - The [Labs](#)
  - The [Observational Studies in Action](#) presentations
- A link to the [Canvas system](#), where you'll turn in some of your work and get connected to the Zoom sessions for the class.
- Information on how to [Contact Us](#) if you need to.

## 2 Professor Love



Figure 2.1: Thomas E. Love, Ph.D.

- Professor of Medicine, Population and Quantitative Health Sciences, [CWRU](#).
- Director of [Biostatistics and Data Science](#), [Population Health Research Institute](#) and [Center for Health Care Research and Policy](#), [The MetroHealth System](#).
- Chief Data Scientist, [Better Health Partnership](#)
- Track Lead for Health Care Analytics, MS in Biostatistics, [Department of Population and Quantitative Health Sciences](#), CWRU
- Fellow, [American Statistical Association](#)

### 2.1 A More Complete Biography

Hi. I am Thomas E. Love, Ph.D. and I have at least three different jobs.

- I am a Professor in the Departments of Medicine and Population & Quantitative Health Sciences at Case Western Reserve University. I teach three courses per year there (PQHS 431, 432 and 500) and also lead the Health Care Analytics track of the MS program in Biostatistics.
- I direct [Biostatistics and Data Science](#) at the [Population Health Research Institute](#) and at the [Center for Health Care Research and Policy](#) at [The MetroHealth System](#).



- For ten years, I was the (founding) Data Director for [Better Health Partnership](#), an alliance of people who provide, pay for and receive care in Northeast Ohio. I now serve as Chief Data Scientist there.
- I am a Fellow of the American Statistical Association, and have won numerous awards for my teaching and my research, including the 2018 [John S. Diekhoff Award for Graduate Teaching](#) from CWRU.
- I have been teaching at CWRU since 1994, and have taught every manner of CWRU student over the years, especially students in biostatistics, medicine, and management.

In research, I use statistical methods to look at questions in health policy and in particular the provision of health services. I mostly work with observational data, rather than data that emerge from randomized clinical trials, and I have a special interest in working with data from electronic health records.

- You may be interested in a [recent study in Health Affairs](#) showing the impact of a [Medicaid-like expansion plan on care and outcomes of poor patients in Cleveland](#).
- Or you might be interested in our [New England Journal of Medicine study](#) of the effect of electronic health records on the care and outcomes of people with diabetes.
- In 2011, [James O'Malley](#) and I chaired the [Ninth International Conference on Health Policy Statistics](#), here in Cleveland. Here's a [recap](#).
- I've also worked on many projects involving the use of propensity scores to make causal inferences from observational studies, particularly in heart failure.

If you want to see a [pretty complete list of my publications](#), knock yourself out.

I hold degrees from Columbia University in the City of New York and from the University of Pennsylvania. My dissertation adviser was Paul Rosenbaum. I am married to a brilliant woman who is an attorney at Savant Systems, and we are raising two terrific sons. The elder lives and works in Pittsburgh and the younger is in his third year at Columbia University, my alma mater. We live in Shaker Heights. I sing and act occasionally in [community theater](#).

## 2.2 Email

- Thomas dot Love at case dot edu (for anything)
- Professor Love is available after class from 11 AM to Noon for office hours. Just let him know during class that you wish to speak with him afterwards.

## 2.3 Name and Pronouns

- Professor Love uses he/him/his pronouns.
- Most students refer to him either as Professor Love or Dr. Love.

- He prefers his given name to be written “Thomas” as opposed to “Tom”.
- Most of his friends and colleagues call him “Tom”. You are welcome to do so, as well, if that makes you more comfortable.

## 2.4 Web

- Dr. Love’s GitHub name is [THOMASELOVE](#).

## 2.5 Offices

My offices are:

- Wood WG-82J on the ground floor of the Wood building on the CWRU Campus.
- Rammelkamp R-229A at MetroHealth Medical Center.

## 3 Teaching Assistant

Leila Hojat will serve as a teaching assistant for this course, for which I am extremely grateful. Leila has taken the course in the past, and will be happy to help you with questions you have about R, about propensity scores, and about other course materials when I cannot.

### 3.1 About Leila

Leila Hojat is a former PQHS 500 student who completed the course in Spring 2022 as part of the Masters in Biostatistics Program. She is an Assistant Professor in the Department of Medicine/ Division of Infectious Diseases at CWRU and a physician at University Hospitals. Her work focuses on antimicrobial stewardship in the hospital setting, and some of her research interests include antimicrobial resistance, rapid diagnostics, and health equity. In addition to R, Leila has experience working with the VA and University Hospitals' electronic data warehouses and SQL for data extraction, and she has modest familiarity with natural language processing tools such as EMERSE and other medical databases including Premier, TriNetX, and CMS. She is looking forward to working with this class while refreshing her propensity score analysis skills. Feel free to reach out to her to discuss class topics as well as any of her personal interests such as cats, wellness, and creative sustainability approaches.

### 3.2 TA Office Hours

Leila's Office Hours for 500 in Spring 2024 will be held **via Zoom** from **5 to 6 PM on most Mondays and Wednesdays**, beginning Monday 2024-01-22 and continuing through Wednesday 2024-05-01.

TA office hours are canceled on 2024-03-11 and 2024-03-13 (Spring Break) as well as 2024-04-01 and 2024-04-03, 2024-04-08 (due to the eclipse) and 2024-04-17.

Zoom meeting details for TA office hours are posted as an announcement on [Canvas](#) and also to our Shared Google Drive.

## 4 Assignments

The three types of assignments required of you in this course are:

1. A semester-long project
2. An “observational studies in action” presentation
3. Four “labs” using R to perform and present analyses.

### 4.1 Deadlines

All deadlines will be found on [the Course Calendar](#).

### 4.2 The Course Project

You will do a small observational study as a capstone project for the course. Your deliverables include:

1. a **proposal** which you will submit for my approval before proceeding further.
2. an **update** verifying that you have the data and are proceeding appropriately.
3. a final **presentation** to the class about your results accompanied by an **abstract**, and by Quarto and HTML files that describe the work you did.

Instructions for the Course Project [are available here](#).

### 4.3 An “Observational Studies in Action” presentation

During the semester, you will be responsible for presenting the methods and results of an observational study from the literature that uses propensity scores. Your choice of manuscript must be accepted in advance by Dr. Love. After you’ve claimed a study, you’ll give a presentation of it, and you’ll also act as “second reviewer” for one of the studies selected by your colleagues. Some presentations will be in class, and others will be recorded.

Instructions for the OSIA work [are available here](#).

## 4.4 Lab Assignments

Four lab assignments require you to do some analyses (using R and Quarto) on data we will provide to you.

Details on the lab assignments [are available in the Labs section of our website](#).

## 4.5 A Note on Feedback

500 is a very different course from 431-432 in terms of the “hands on” assistance that I make available to you as you’re working on a lab. In particular, I’m not going to review your code to be sure you’re going in the right direction, even though I understand some of you have come to expect that from 431-432.

Instead, we discuss the homework in class and we provide a detailed answer sketch after the fact. So if you have questions, please feel free to send them to us at **500-help at case dot edu**, or ask about them in office hours, but what we are likely to say to anything we cannot resolve quickly for you (or think is worthy of deeper discussion as a group) is:

1. That sounds like an excellent question to bring up on Thursday in class, and
2. We’re going to let you (and everyone else) flounder a bit between now and our discussion of this homework in class.

## 4.6 Grading

- If you complete all deliverables **on time**, and your project and OSIA presentations are solid, you will receive an A in the course.
  - If you mostly meet that standard but don’t quite (either because more than one thing is late or because you have substantial project problems that linger), either an A or B is possible.
- This is an advanced graduate school course. I don’t anticipate that anyone who makes a concerted effort will fall below the B standard, as no one has in the past.
- If you have any concerns, raise them with Dr. Love (email to **500-help** is fine) directly as soon as possible.

## 5 Advice for Graduate Students

My most important piece of general advice to people is to be kind. That's not always the thing I do as well as I'd like.

### 5.1 On Graduate School

For graduate students, I urge you to take as much advantage of this learning experience as you can. While I'll refer to some of the pieces below during the course, I've gathered a few favorites here.

1. From [matt.might.net](http://matt.might.net)...
  - [12 Resolutions for Grad Students](#)
  - [How to get a great letter of recommendation](#)
  - [How to send and reply to email](#). You might also want to look at [Email Etiquette: Guidelines for Writing to Your Professors](#)
2. [Four Things You Should Do When You're Bored, on YouTube](#) (the four things are Exercise / Read / Meditate / Find and Engage a Hobby with Passion). Also, [Get Up and Move. It May Make You Happier](#) by Gretchen Reynolds, in the *New York Times*, 2017-01-25.
3. [Why academics need to focus on structuring their time](#) from *University Affairs*.
4. Most people are really bad at meetings, including Dr. Love. [Here are some extremely useful suggestions from Greg Wilson](#). In a tweet, [Wilson argues that](#) "the single most useful training you can give an adult is how to run a meeting and how to participate in someone else's."
5. Some people need help taking notes. You might be interested in [Cornell Note Taking](#) or the [5 other methods described here](#).
6. You may be interested in the American Statistical Association, and its [This is Statistics](#) program.

### 5.2 On Seeking a Job

If I have a job or internship to offer, I'll be noisy about it in class. In the meantime, I'd consider joining the [American Statistical Association as a student member](#) and perhaps joining

the [Greater Cleveland R Meetup Group](#).

Here are some gathered thoughts from other folks that you might enjoy:

1. [General Advice on an Academic Career Path](#) (which is filled with useful advice, especially for those studying biostatistics.)
2. [Advice for Applying to Data Science Jobs](#) from [Emily Robinson](#)
3. [Academic job search advice](#) from [matt.might.net](#)
4. [Career Advice for Data Scientists](#) was a panel at [rstudio::conf 2020](#) focused on building a career around R.

## 6 On Writing, Presenting & Communicating

I write all the time. It's my job. It's yours, too. You'll do more of it here than you may be expecting in this class, and in life. So you'll need to take advantage of every opportunity you have to write more effectively tomorrow than you do today.

Here's what I find to be a compelling argument from George Cobb ...

If you want your work with data to make a difference, devote time and effort to choosing the words and pictures you use to present your evidence and conclusions. If you teach or supervise, seek to reward those - they who learn from you, and they who report to you - when they spend time crafting their message.

Think back to the last "report" you were expected to read. Is it easy to recall the main points? The answer, yes or no, depends not so much on the quality of the data, the effort, and the thinking that went into the report, nor on your own dutiful diligence in reading the report, but rather, and mainly, on whether the people who wrote the report had learned and practiced the skills of how to use words and pictures, first, to claim attention, and second, to claim retention: to deliver a message that sticks in the mind.

### 6.1 Campus Resources

The [University Resources for Student Support][University Resources for Student Support] section in this Syllabus lists several important resources for improving your English and your communication skills, most of which will be available remotely after the first week of the semester. We strongly encourage you to take advantage of these resources.

### 6.2 Advice from Other People

1. [Preparing Manuscripts for Submission to Medical Journals: The Paper Trail](#) by H. Gilbert Welch, from *Effective Clinical Practice* in 1999.
  - Start early, focus on high-visibility components, develop a systematic approach to the body of the paper, finish strong.



- Use [thinkchecksubmit](#) and [see this tweet](#) for some related suggestions when selecting a journal or publisher for your work.
- 2. [Writing a Scientific Paper in Four Easy Steps](#) from Claus Wilke at [The Serial Mentor blog](#).
- 3. [Rules to write a good research paper](#) from Daniel Lemire.
- 4. [Hey-here are some tips on communicating data and statistics!](#) from Andrew Gelman 2017-06-02.
- 5. [Writing Pet Peeves: Correctness, References, and Style](#) from Tamara Munzner.
- 6. Frank Harrell's [Checklist for Authors of Statistical Problems to Document and to Avoid](#)

## 6.3 A Few Tips from Professor Love

1. Statistics is a “getting the details right” business - we care deeply about details, and this applies to writing code or complete English sentences. RStudio has a spell-checker. To use it, click F7.
2. Nothing impresses us as much as a clear and concise argument, presented using well-written English sentences, effective and well-labeled figures and tables.
3. Don't parrot back material that Professor Love wrote or said. State ideas in your own words. Stating them in my words is, technically, plagiarism.
4. Edit your more adventurous output; don't present everything you know how to do in R, and don't forget that someone is trying to read both your code and your results.
5. Make your work easy to evaluate. In responding to an assignment, be sure to answer the question that was asked, restating it as necessary.
6. Clearly label everything: graphs, tables, your answer to a specific question. Everything. Again, make your work easy to evaluate.
7. Simplify. Emphasize ideas in plain language. Avoid jargon. Use English well.
8. Data are plural. Use “the data **are** ...” rather than “the data *is* ... ”
9. A paragraph must contain more than one sentence.
10. Don't switch tenses. If you want to write in the present tense, stick to it throughout.
11. Don't write or say random sample unless you used a random number generator. If you used haphazard sampling or convenience sampling, call it what it is, and indicate whether any problems could have cropped up as a result.
12. Similarly, don't defend a method of data collection because it is random. Most of the time we want to represent some population, and a random sample is just one way to ensure that certain types of biases have a low probability of creeping in.
13. If you want to write that you used  $\alpha = 0.05$  as your significance level, then state that your results were obtained using a 95% confidence level, not a 95% confidence interval, unless you are actually interpreting a confidence interval.

14. If you find yourself in the appalling situation of writing about a  $p$ -value, then you should state something like:
  - [1] We're using a 95% confidence level.
  - [2] We're using a 5% significance level.
  - [3] We're using  $\alpha = 0.05$ .
  - Don't use more than one of these expressions.
15. Again, don't use  $p$ -values in most settings, but if you must, refer to all  $p$ -values that are less than 0.001 or perhaps less than 0.0001 as  $p < 0.001$ , rather than, for instance,  $p = 0.00000001$  or, worse yet,  $p = 0$ . In a similar vein, write all  $p$ -values that exceed 0.99 as  $p > 0.99$  instead of, for instance,  $p = 1$ .
16. To the extent possible, don't use **computer-ese** to label variables, plots or tables. R and Markdown allow you to change the labels on graphs and tables to meaningful things – do so. Use meaningful abbreviations, as necessary, explaining what they mean on the first usage.
17. When in doubt, err on the side of clarity. Clear thinking, clear writing.

## 7 Other Policies

These are things I must include in the syllabus in this format. If you have questions, email me.

### 7.1 Statement on Disability

In accordance with federal law, if you have a documented disability, you may be eligible to request accommodations from Disability Resources. Students with a disability, please make an appointment with the instructor to discuss your needs at the earliest convenience. The necessary adjustments will be provided to facilitate the learning experience. Additionally, please be in contact with the Coordinator of Disability Resources, Educational Services for Students (ESS). ESS is located in 470 Sears Building. The office phone number is 216-368-5230, and the website is: <http://studentaffairs.case.edu/education/disability/policy.html>

Please keep in mind that accommodations are not retroactive.

### 7.2 Incomplete Grades

The grade of Incomplete is assigned at the discretion of an instructor provided that:

1. There are extenuating circumstances, explained to the instructor before the assignment of the grade, which clearly justify an extension of time beyond the requirements established for other students in the class. It is the student's responsibility to notify the instructor of the circumstances preventing completion.
2. The student has been passing the course and only a small segment of the course remains to be completed, such as a term paper, for which the extenuating circumstances justify a special exception.

An Incomplete grade may not be assigned if a student is absent from a final examination, unless the dean has authorized the absence. Unauthorized absence from a final examination will result in a failing grade. When the student completes the work, the Incomplete is changed to an A, B, C, D, P, F, or NP.

All work for the incomplete grade must be made up and the change of grade recorded in the Office of the University Registrar by the date specified by the instructor, but no later than the

last day of class in the semester following the one in which the Incomplete was received. A student who has a permanent Incomplete for a required course must retake the course in a later term. If the student cannot complete the work by the end of the following semester, he or she must petition for an extension which must be endorsed by the instructor, explain the reasons why the work has not been completed, and include a new date for completion. Students will be allowed only one extension of no more than one additional semester to complete the work for an Incomplete grade.

### **7.3 Laptop and cell phone use**

Although having a laptop in class opens up new learning possibilities for students, it can be used in ways that are inappropriate. It is easy for your laptop to become a distraction to you and to those around you. Laptops are to be used only when essential to the task at hand. Please turn off or silence all cell/smart phones, tablets, and other electronic devices for the duration of the course. Inappropriate uses will be noted and may affect the final grade.

### **7.4 Diversity and Inclusion**

It is the intent that all students regardless of their background and perspective be well-served by this class. Further, we intend to present material whose content is respectful of diversity (gender identity, sexuality, disability, age, socioeconomic status, ethnicity, race, nationality, religion, and culture) and deliver it in a way that respects these differences as well. We expect that all students, instructors and guests will help foster an atmosphere of respect, trust and safety in the classroom.

If you have suggestions for how to make the class content or environment more inclusive, or have specific incidents to report, please reach out to the instructor. If you are not comfortable reaching out to the instructor, feel free to reach out to someone else, such as the School of Medicine Graduate Education Office (som-geo@case.edu) or the Office of Inclusion, Diversity and Equal Opportunity (OIDEO) (oideo@case.edu).

More information on University policy and resources are available on OIDEO's website at <https://case.edu/diversity/>.

### **7.5 Academic Integrity**

Any violation of the University's Code of Ethics will not be tolerated. All forms of academic dishonesty including cheating, plagiarism, misrepresentation, and obstruction are violations of academic integrity standards and will result in a minimum penalty of receiving a zero for

the assignment, the potential for failing the entire course. Cheating includes copying from another's work, falsifying problem solutions or laboratory reports, or using unauthorized sources, notes or computer programs. Plagiarism includes the presentation, without proper attribution, of another's words or ideas from printed or electronic sources. It is also plagiarism to submit, without the instructor's consent, an assignment in one class previously submitted in another. Misrepresentation includes forgery of official academic documents, the presentation of altered or falsified documents or testimony to a university office or official, taking an exam for another student, or lying about personal circumstances to postpone tests or assignments. Obstruction occurs when a student engages in unreasonable conduct that interferes with another's ability to conduct scholarly activity. Destroying a student's computer file, stealing a student's notebook, and stealing a book on reserve in the library are examples of obstruction.

In addition, the incident will be reported to the Dean of Undergraduate Studies and Academic Review Board for undergraduates or Senior Associate Dean of Graduate Studies, for Graduate Students.

The CWRU Statement of Ethics for graduate students can be found here: <http://case.edu/gradstudies/about-the-school/policies-procedures/>

## 7.6 Plagiarism

Unless specifically stated otherwise, the faculty of the PQHS expect and require original writing for all assignments given. Submitting plagiarized work for an academic requirement is a violation of the academic integrity standards set forth by the University. Plagiarism is the representation of another's work or ideas as one's own; it includes the unacknowledged, word-for-word use and/or paraphrasing of another person's work, and/or the inappropriate unacknowledged use of another person's ideas. Submitting substantially the same work to satisfy requirements for one course that has been submitted in satisfaction of requirements for another course, without permission of the instructor of the course for which the work is being submitted, is also prohibited.