Computational Social Science
Rutgers University

Syllabus

Dr. Thomas Davidson

Fall 2025

CONTACT AND LOGISTICS

E-mail: thomas.davidson@rutgers.edu.

Website: https://github.com/t-davidson/SOC360-CSS and Canvas.

Class meetings: Monday & Thursday 12:10 - 1:30pm, Beck Hall 003, Livingston Campus.

Office hours: Thursday 5:00-6:30 p.m, 109 Davison Hall, Douglass Campus, or by appointment.

COURSE DESCRIPTION

This course introduces students to the growing field of computational social science. Students will learn to collect and critically analyze social scientific data using a range of techniques including natural language processing, machine learning, and agent-based modeling. We will discuss how these techniques are used by social scientists and consider the ethical implications of big data and artificial intelligence. Students will complete homework assignments involving coding in the R programming language to analyze several different datasets and a final project to create an application for data analysis and visualization.

LEARNING GOALS

- Further develop proficiency in using R and RStudio
- Competence in data merging, cleaning, analysis, and visualization
- Understand and implement various methods for online data collection, natural language processing, and machine learning
- Use RShiny to develop a web application for data analysis and visualization
- Identify important ethical issues related to the use of social data and computational methods

PREREQUISITES

Data 101 or equivalent. Enrolled students must have experience writing basic programs in a general purpose programming language, e.g. R, Python, Java, C. The course will use the R programming language. We will review the fundamentals for programming and data science in R in weeks 1-3.

ASSESSMENT

- 10% Class participation
  - Students are expected to attend all class meetings and to actively participate in class discussions
- 60% Homework assignments (4 x 15%)
  - Students will complete a series of homework assignments to gain experience using R for data science
- 30% Group project
  - Students will complete a final project involving the use of RShiny to build an interactive web application for data analysis and visualization. The project will include a written report and a presentation.

Rubric

Final grades will be determined according to the following rubric:

- A: 90-100%
- B+: 85-89%
- B: 80-84%
- C+: 75-79%
- C: 70-74%
- D: 60-69%
- F: <60%

READINGS

Most of the readings will consist of chapters from the textbooks listed below. These readings are intended to build familiarity with key concepts and programming skills. Some weeks there will be an additional reading to highlight how data science techniques are used in empirical social scientific research.

Textbooks

All textbooks are available for free online (hover over titles for links).

- Matthew Salganik. 2017. *Bit by Bit*. Princeton University Press. ISBN: 0691196109
- Wickham, Hadley, and Garrett Grolemund. 2023. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. 2nd edition (*R4DS*). O'Reilly Media, Inc. ISBN: 1491910399
- Silge, Julia, and David Robinson. 2017. *Text Mining with R: A Tidy Approach.* O'Reilly Media. ISBN: 1491981652

COURSE RESOURCES

The course will be organized using two different tools, Github and Canvas. Canvas will be used for class communication and to host readings. Github Classroom will be used for the submission of assignments.

CHAT-GPT AND AI TOOLS

Over the course of the semester we will consider how generative artificial intelligence tools can be used to advance computational research and will situate these tools in the context of earlier techniques. Not only can these technologies contribute to methodological advances, but they are a valuable pedagogical resource. I encourage you to draw on these tools to help you to learn to code and learn various computational methods.

At the same time, however, these tools come with risks. For example, they can produce misleading or incorrect results, and hence must be used with caution. I am also concerned that overreliance on AI can undermine your learning. I strongly recommend trying to solve problems yourself or by consulting the course materials before resorting to AI for assistance.

AI tools can also be used in ways that violate academic integrity policies. As such, you are prohibited for using ChatGPT or other AI tools to complete homework assignments. Of course, the definition of AI tools becomes some what blurred when we think about tools like Grammarly, which uses similar technologies to help you improve your writing. Google Docs now has similar functionality and Microsoft is rolling this out in Office. I will permit you to use these tools to help you edit your work, but you may not use them to generate sentences or paragraphs of text. Similarly, I will permit you to use AI to help edit and debug code in your final project, but you may not use them to generate entire chunks of code. *You will be asked to provide a statement in assignments and your final project detailing any ways that AI tools were used.* Please contact me if you are unsure about whether a particular use case is acceptable.

TECHNOLOGY REQUIREMENTS

Students will be required to have access to a computer to complete assignments. Ideally, students should bring a laptop computer to class, although laptops and cell phones should only be used when required.

Please visit the Rutgers Student Tech Guide page for resources available to all students. If you do not have the appropriate technology for financial reasons, please email Dean of Students deanofstudents@echo.rutgers.edu for assistance. If you are facing other financial hardships, please visit the Office of Financial Aid.

COURSE POLICIES

The Rutgers Sociology Department strives to create an environment that supports and affirms diversity in all manifestations, including race, ethnicity, gender, sexual orientation, religion, age, social class, disability status, region/country of origin, and political orientation. We also celebrate diversity of theoretical and methodological perspectives among our faculty and students and seek to create an atmosphere of respect and mutual dialogue. We have zero tolerance for violations of these principles and have instituted clear and respectful procedures for responding to such grievances.

Students must abide by the Code of Student Conduct at all times, including during lectures and in participation online.

Students must abide by the university's Academic Integrity Policy. Violations of academic integrity will result in disciplinary action. Please review this policy or contact Professor Davidson if there is something you are unsure about.

If you have a documented disability and require accommodations to obtain equal access in this course, please contact me during the *first week of classes*. Students with disabilities must be registered with the Office of Student Disability Services and must provide verification of their eligibility for such accommodations. See end of syllabus for further details.

Please see the bottom of the syllabus for information on additional information and resources.

COURSE OUTLINE

Week 1

Introduction to Computational Social Science

Readings

- Thursday 9/4:
    - *Bit by Bit*, C1
    - *R4DS*: Introduction & C1, C27 in 1st edition [Note: Chapter numbers correspond to the online book; physical book numbers are different]

Recommended

- Golder, Scott A., and Michael W. Macy. 2014. "Digital Footprints: Opportunities and Challenges for Online Social Research." *Annual Review of Sociology* 40 (1): 129–52.
- Edelmann, Achim, Tom Wolff, Danielle Montagne, and Christopher A. Bail. 2020. "Computational Social Science and Sociology." *Annual Review of Sociology* 46 (1)

Week 2

Data Structures in R & Social Networks

Readings

- Monday 9/8:
    - *R4DS*: C2, 27
    - David Easley and Jon Kleinberg. 2010. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, Chapter 1.
- Thursday 9/11:
    - *R4DS*: C3-5

Week 3

Programming Fundamentals & Agent-based Models

Readings

- Monday 9/15:
    - Macy, Michael W., and Robert Willer. 2002. "From Factors to Factors: Computational Sociology and Agent-Based Modeling." *Annual Review of Sociology* 28(1):143–66
        * *Read pages 143-150*
    - *R4DS*: C25-26
- Thursday 9/18:
    - *R4DS*: C6-8

Assignment 1 released: Programming in R.

Week 4

Data Collection I: Collecting Data Using Application Programming Interfaces

Readings

- Monday 9/22:
    - *Bit by Bit*, C2

– *R4DS*: C17, 19
- Thursday 9/25:
    – *R4DS*: C9

Week 5

Data Collection II: Scraping Data From the Web

Readings

- Monday 9/29:
    – *Bit by Bit*, C6
    – *R4DS*: C24
- Thursday 10/2:
    – *R4DS*: C14, 16

Recommended

- Fiesler, Casey, Nate Beard, and Brian C Keegan. 2020. "No Robots, Spiders, or Scrapers: Legal and Ethical Regulation of Data Collection Methods in Social Media Terms of Service." In *Proceedings of the Fourteenth International AAAI Conference on Web and Social Media*, 187–96.

Assignment 2: Collecting and storing data released.

Week 6

Data Collection III: Online Experiments and Surveys

Readings

- Monday 10/6:
    – *Bit by Bit*, C3-5
- Thursday 10/9:
    – R Shiny tutorial: https://shiny.rstudio.com/tutorial/

Week 7

Natural Language Processing I: The Vector-Space Model

Readings

- Monday 10/13:
    – Evans, James, and Pedro Aceves. 2016. "Machine Translation: Mining Text for Social Theory." *Annual Review of Sociology* 42 (1): 21–50.
- Thursday 10/16:
    – *Text Mining with R*, C1 & 3

Week 8

Natural Language Processing II: Word Embeddings

Readings

- Monday 10/20:
    – *Text Mining with R*: C5.
- Thursday 10/23:
    – Hvitfeldt, Emil and Julia Silge. 2020 *Supervised Machine Learning for Text Analysis in R.* Chapter 5.

– Kozlowski, Austin, Matt Taddy, and James Evans. 2019. "The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings." *American Sociological Review*, September, 000312241987713.

Project proposals due 10/23

Week 9

Natural Language Processing III: Topic Models

Assignment 3: Natural language processing released.

Readings

- Monday 10/27:
    – *Text Mining with R*: C6
- Thursday 10/30:
    – Roberts, Margaret, Brandon M. Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson, and David Rand. 2014. "Structural Topic Models for Open-Ended Survey Responses: Structural Topic Models for Survey Responses." *American Journal of Political Science* 58 (4): 1064–82.

Week 10

Machine Learning I: Prediction and Explanation

Readings

- Monday 11/3:
    – Molina, Mario, and Filiz Garip. 2019. "Machine Learning for Sociology." *Annual Review of Sociology* 45: 27–45.
- Thursday 11/6:
    – Salganik, Matthew J, Ian Lundberg, Alexander T Kindel, Caitlin E Ahearn, Khaled Al-Ghoneim, Abdullah Almaatouq, Drew M Altschul, et al. 2020. "Measuring the Predictability of Life Outcomes with a Scientific Mass Collaboration." *Proceedings of the National Academy of Sciences*.

Week 11

Machine learning II: Text Classification

Readings

- Monday 11/10:
    – *Text Mining with R*, Chapter 6
- Thursday 11/13:
    – Dixon, Lucas, John Li, Jeffrey Sorensen, Nithum Thain, and Lucy Vasserman. 2018. "Measuring and Mitigating Unintended Bias in Text Classification." In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society* - AIES '18, 67–73. New Orleans, LA, USA: ACM Press.

Week 12

Machine learning III: Large language models

Assignment 4: Machine learning released.

Readings

- Monday 11/17
- Thursday 11/20:
    - Manning, Christopher D. 2022. "Human Language Understanding & Reasoning." *Daedalus* 151(2):127–38.
    - Ziems, Caleb, William Held, Omar Shaikh, Zhehao Zhang, Diyi Yang, and Jiaao Chen. 2023. "Can Large Language Models Transform Computational Social Science?"

Project prototype due 11/20

Week 13

Machine learning IV: Image Classification

Readings

- Monday 11/25:
    - Gebru, Timnit, Jonathan Krause, Yilun Wang, Duyun Chen, Jia Deng, Erez Lieberman Aiden, and Li Fei-Fei. 2017. "Using Deep Learning and Google Street View to Estimate the Demographic Makeup of Neighborhoods across the United States." *Proceedings of the National Academy of Sciences* 114 (50): 13108–13.
    - Buolamwini, Joy, and Timnit Gebru. 2018. "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification." In *Proceedings of Machine Learning Research*, 81:1–15.
- Tuesday 11/26:
    - Thursday class switched to Tuesday due to break

THANKSGIVING RECESS

Week 14

Generative Artificial Intelligence

Readings

- Monday 12/1:
    - Davidson, Thomas. 2023. "Start Generating: Harnessing Generative Artificial Intelligence for Sociological Research."
- Thursday 12/4:
    - Project workshop

Week 15

Project presentations

- Monday 12/8
    - Presentations

Final projects due 12/12

Additional information

The Rutgers University Student Assembly urges that the following information be included at the end of every syllabus.

Report a Bias Incident

Bias is defined by the University as an act, verbal, written, physical, psychological, that threatens, or harms a person or group on the basis of race, religion, color, sex, age, sexual orientation, gender identity or expression, national origin, ancestry, disability, marital status, civil union status, domestic partnership status, atypical heredity or cellular blood trait, military service or veteran status.

If you experience or witness an act of bias or hate, report it to someone in authority. You may file a report online and you will be contacted within 24 hours. The bias reporting page is here.

Counseling, ADAP & Psychiatric Services (CAPS)

(848) 932-7884 / 17 Senior Street, New Brunswick, NJ 08901 / Link to website.

CAPS is a University mental health support service that includes counseling, alcohol and other drug assistance, and psychiatric services staffed by a team of professionals within Rutgers Health services to support students' efforts to succeed at Rutgers University. CAPS offers a variety of services that include: individual therapy, group therapy and workshops, crisis intervention, referral to specialists in the community, and consultation and collaboration with campus partners.

Crisis Intervention

Link to website.

Report a Concern:

Link to website.

Violence Prevention & Victim Assistance (VPVA)

(848) 932-1181 / 3 Bartlett Street, New Brunswick, NJ 08901 / Link to website. The Office for Violence Prevention and Victim Assistance provides confidential crisis intervention, counseling and advocacy for victims of sexual and relationship violence and stalking to students, staff and faculty. To reach staff during office hours when the university is open or to reach an advocate after hours, call 848-932-1181.

Disability Services

(848) 445-6800 / Lucy Stone Hall, Suite A145, Livingston Campus, 54 Joyce Kilmer Avenue, Piscataway, NJ 08854 / Link to website

Rutgers University welcomes students with disabilities into all of the University's educational programs. In order to receive consideration for reasonable accommodations, a student with a disability must contact the appropriate disability services office at the campus where you are officially enrolled, participate in an intake interview, and provide documentation: see guidelines. If the documentation supports your request for reasonable accommodations, your campus's disability services office will provide you with a Letter of Accommodations. Please share this letter with your instructors and discuss the accommodations with them as early in your courses as possible. To begin this process, please complete the Registration form on the ODS web site.