LAST UPDATED: 2025-08-12

SOC 501
Computational Social Science

Description of the Course
This course will introduce you to the burgeoning methodological subfield called computational social science: An evolving bundle of computationally sophisticated techniques that are increasingly used to answer research questions in the social sciences. The goal of the course is for you to learn how to produce your own computational social science research proposal aimed at a scholarly outlet. The course will survey the latest published research and train you in research design, methods, and techniques in computational social science. This might include topics such as big data, data visualization, social network analysis, agent-based modeling, natural language processing, and artificial intelligence (AI), as well as identifying and developing variables, mechanisms, and research questions motivated by social science research. The course will also delve into the growing ethical challenges and considerations associated with computational social science research. The course is set up to welcome people from many different backgrounds, in particular those with or without prior exposure to programming, statistics, or the social sciences.

Location and Time
In-person Location: Social Science Building, Room 415
In-person Meeting Time: Thursday, 9:00 am to 11:30 am

Course Prerequisites or Co-requisites
None

Instructor and Contact Information
Charles J. Gomez (He/Him)
Email: cjgomez@arizona.edu
Please use the "**SOC 501 | Your Subject Here**" header to ensure a quick response.
Office: Social Sciences Building, Room 426
Office Hours: By appointment on Thursdays in person or on Zoom.

Python Help
Data & Viz Drop-in
https://lib.arizona.edu/about/events/data-viz-drop-2
Tuesdays, 9:00 a.m. to 11:00 a.m.

Expected Learning Outcomes
By the end of this course, you will be able to:
- Design academically-oriented computational social science research aimed at top disciplinary or interdisciplinary outlets.
- Develop familiarity with canonical methods and frameworks used in computational social science research.
- Critically evaluate assumptions and methods used in sound computational social science research in its different stages.
- Apply the ethical considerations involved in computational social science research.

Course Objectives: *Choose Your Own Adventure!*

The objective of this course is to train you to do academic research using computational social science methods and approaches. We review labs written in Python that cover computational social science methods. We discuss and analyze the latest empirical research in the social sciences that uses these methods. And we practice developing research designs that emphasize how to think like a computational social scientist. At some basic level, everyone engages with all three aspects of the class. However, this class also attracts students from very different backgrounds and with diverse skillsets. That's what makes it so unique and enriching. Everyone might not have sufficient background in the social sciences, research experience, programming, or statistics to go in-depth on all fronts.

So, to be as accommodating to everyone and their interests, I adopt a "choose your own adventure" format for the class: If you want to focus on applying the methods, you can. If you want to focus on research design, you can do that too. If you want to focus more on how social scientists who use these techniques think and conduct research, that's availed to you as well.

On coding: My goal is to introduce you to a smattering of techniques that you can use to engage in your own computational social science research. We use Python 3.9 in this class for the labs, but we use it as a means to an end. In other words, this class will **NOT** teach you how to code as its primary objective. So, the boundaries of support and training in this class do not extend further than that. (A computer science class will do a much better job in that regard.) The class does not require you to become anywhere near proficient in programming. However, if you choose to dive deep into programming, you should, but it will require a significant amount of self-study and effort on your part. So, plan accordingly. We will allocate the last part of class to review the labs.

Course Format and Teaching Methods

This class is dedicated to activities and discussion. Active engagement through activities, discussion, and review will pay high dividends as you acquire and hone the skills covered in this course. This is also beneficial since everyone approaches this class with different styles of learning and learning at different rates.

We meet once a week on Thursdays. A typical session should go as follows:

First 10 Minutes (9:00 a.m. to 9:10 a.m.) – Welcome, administrivia, and announcements.
Next 40 Minutes (9:10 a.m. to 9:50 a.m.) – Discussion of the required readings.
Next 10 Minutes (9:50 a.m. to 10:00 a.m.) – Short break and transition.
Next 40 Minutes (10:00 a.m. to 10:40 a.m.) – CSS research design activities.
Next 10 Minutes (10:40 a.m. to 10:50 a.m.) – Short break and lab check.
Last 40 Minutes (10:50 a.m. to 11:30 a.m.) – Lab review.

Absence and Class Participation Policy

UArizona's policy concerning Class Attendance, Participation, and Administrative Drops is available at: https://deanofstudents.arizona.edu/policies/attendance-policies-and-practices.

UArizona's policy regarding absences for any sincerely held religious belief, observance or practice will be accommodated where reasonable, http://policy.arizona.edu/human-resources/religious-accommodation-policy.

Absences pre-approved by the UArizona Dean of Students (or Dean Designee) will be honored. See: https://deanofstudents.arizona.edu/absences

Participating in the course and attending course sessions are vital to the learning process. As such, attendance is required for all sessions. Multiple absences may affect your final course grade (see class engagement). If you anticipate being absent, are unexpectedly absent, or are unable to participate in class activities, please contact me as soon as possible.

To request a disability-related accommodation to this attendance policy, please contact the Disability Resource Center at (520) 621-3268 or drc-info@email.arizona.edu. If you are experiencing unexpected barriers to your success in your courses, the Dean of Students Office is a central support resource for all students and may be helpful. The Dean of Students Student Assistance page can be found here: https://deanofstudents.arizona.edu/support/student-assistance. In addition, Student Success resources can be found here: https://studentsuccess.arizona.edu/student-resources.

Makeup Policy for Students Who Register Late
Students who register after the first class meeting may make up one week of missed assignments and must do so by the end of the second week.

Course Communications
Official UArizona e-mail address and D2L.

Required Texts or Readings
There is no single textbook for this class, and I've placed all the readings and course materials on D2L.

Required or Special Materials
A laptop that runs a version of MacOS, Linux, or Windows.

Required Extracurricular Activities
None

Requirements for the Course
None

Grading and Assessments
Your final course grade is comprised of four components:
(1) A research proposal (30%) and an in-class presentation (10%) are worth **40**% of your final grade.
(2) Memos are worth **30**% of your final grade.
(3) Labs are worth **20**% of your final grade.

(4) Class engagement and citizenship are worth **10**% of your final grade.

Once each of the four components of your final course grade is calculated and weighted, your final course grade is converted to your final course **letter grade** using the scheme below:

90 to 100: A | Exceeds all requirements and expectations.
80 to 89: B | Meets all requirements and expectations.
70 to 79: C | Misses some of the requirements and expectations.
60 to 69: D | Significantly falls short of requirements and expectations.
< 60: E | Repeat of the course needed for credit.

University policy regarding grades and grading systems is available at
http://catalog.arizona.edu/policy/grades-and-grading-system

Research Proposal (40%)

Everyone will submit a research proposal at the end of the semester. The aim is for you to eventually develop and submit your research proposal to a computational social science or traditional social science journal in the near future. The expectation is that the proposal should be between 15-and-20 pages, double-spaced, including references. The proposal should be the basis for a professional journal submission in the field. These include but are not limited to *American Sociological Review*, *American Journal of Sociology*, *Social Forces*, *Nature Human Behaviour*, *Science Advances*, *Journal of Computational Social Science*, *Scientific Reports*, *PLOS One, Journal of Informetrics*, *Socius: Sociological Research for a Dynamic World*, *Poetics: Journal of Empirical Research on Culture, the Media, and the Arts*, and *Quantitative Science Studies.*

Your research proposal may be completed individually or in groups of no more than **two** people, where you will both receive the same grade.

You will also present an overview of your proposal on the **last day of the class**, worth 10% of your final grade. You should cover what your research question is, the motivation behind it, the data and methods you used or intend to use, and if it's relevant to your main results and why these findings are important. Everyone should outline the next steps to progress your project. Presentations that are clear and well-organized (and do not run over time allotment) exceed expectations and earn full credit. Presentations that are muddled, confusing, or lack polish will not receive credit.

Memos (30%)

Each week, you will submit a memo. **For the first three memos in Modules 1, 2, and 3**, everyone will submit a reaction memo to a specific prompt that asks you to think about your research topic and research question, how you might model it, and what the social mechanisms or processes could be.

However, starting with the memo in **Module 4**, you may choose to:

(1) Submit a memo responding to the weekly prompts. Memos MUST be at least 1 page, double-spaced, 12-point font.

(2) Submit a data analysis memo that applies method(s) from the lab. These memos are straightforward. Write a brief description of what you did, why you did it, how you did it, and what you found. This simply could be taking one or two of the lab's checkpoints and using your own data to explore something you thought would be interesting. There is no need to submit the data or code, but you should include any relevant images or figures. Do not use screenshots of output and instead format your work in a professional fashion. Since figures add to the page length, memos should be longer than two pages, double-spaced, 12-point font.

This is where the choose-your-own-adventure begins! There are no restrictions on the type of memos you submit. If you want to write a memo reacting to the prompt one week and then switch to a data analysis memo the next week, that's perfectly fine. If you want to submit only reaction memos or only data analysis memos for the entire semester, that also works.

Keep in mind that you are **NOT** required or obligated to work with the same research question or topic for the semester. In fact, many students switch topics and questions frequently in this class. If you decide to change your question or topic, you **DO NOT** need to go back and rewrite any of your past memos if you switch topics or questions. In addition, you **DO NOT** need to tell me or get my approval, either.

The memos are meant to be an opportunity for you to explore different ideas by putting words on paper. **Take advantage of that**.

Memos are graded as **High Pass**, **Pass**, or **No Credit**.

A **Pass** is the baseline and modal grade for the memos, evaluated as the quality of work expected of a student in this class. A grade of **High Pass** is awarded to recognize an exceptional submission, but these grades are rare and should not be expected.

Memos that earn a grade of **No Credit** miss the standards expected for students in this class. However, you may **resubmit** a revised version of these memos to earn a **Pass**.

I provide collective feedback at the start of each class. If you would like individual feedback on your submissions, please schedule a time to speak with me.

There are **twelve** opportunities to submit a memo. I take the **number of pass** grade memos as your overall memo grade following this scheme, earning:

- Submitting at least **ten** "pass" (e.g., high pass or pass) memos earns you a **100 (A)**.
- Submitting at least **nine** "pass" memos earns you a **95 (A)**.
- Submitting at least **eight** "pass" memos earns you a **90 (A)**.
- Submitting at least **seven** "pass" memos earns you an **85 (B)**.
- Submitting at least **six** "pass" memos earns you an **80 (B)**.
- Submitting at least **five** "pass" memos earns you a **75 (C)**.
- Submitting at least **four** "pass" memos earns you a **70 (C)**.
- Submitting at least **three** "pass" memos earns you a **65 (D)**.
- Submitting at least **two** "pass" memos earns you a **60 (D)**.

- Submitting **one** "pass" memo or **no** "pass" memos will earn you a **55 (E)**.

To get an **A** (90) for the memos, you only need to complete **eight of the twelve** memos with a "pass" grade. Submitting **ten of the twelve** memos with a "pass" grade guarantees a score of 100.

On finding data for your data analysis memos: Today, there is an abundance of readily available datasets for you to use. For instance, Google provides us with a repository of over 25 million public datasets that can be accessed here:

https://datasetsearch.research.google.com/

Other popular repositories for publicly available data are Kaggle (https://www.kaggle.com/), Data.gov (https://www.data.gov/), Harvard Dataverse (https://dataverse.harvard.edu), and the UC Irvine Machine Learning Repository (https://archive.ics.uci.edu/ml/index.php).

Labs (20%)
There are **twelve labs**: eleven Jupyter Notebooks, one lab that covers how to use SQL and the high-performance computer cluster (HPC).

Jupyter Notebooks are interactive documents that contain executable Python code, videos, texts, equations, and visualizations. You will download them from D2L each week and complete them. This means running the code (i.e., executing and running **all** the cells in the Notebook) and attempting **all** checkpoints.

I will check the labs during our class time. Labs will be assessed as a completion grade. If you run all the cells and **attempt all the checkpoints**, you earn full credit on the lab. **A partial submission does not earn you any credit.**

Labs must be completed individually, but I highly encourage you to work collaboratively.

For Lab 12 (SQL and HPC), you will submit a certificate of completion from DataCamp and a memo with screenshots indicating the completion of the tasks.

Your composite grade for the labs is calculated as follows:

- Submitting at least **ten** fully attempted labs (i.e., all cells are executed and at least half of the checkpoints attempted for the Jupyter Notebook labs or fully attempting the special labs) earns you a **100 (A)**.
- Submitting at least **nine** fully attempted labs earns you a **95 (A)**.
- Submitting at least **eight** fully attempted labs earns you a **90 (A)**.
- Submitting at least **seven** fully attempted labs earns you an **85 (B)**.
- Submitting at least **six** fully attempted labs earns you an **80 (B)**.
- Submitting at least **five** fully attempted labs earns you a **75 (C)**.
- Submitting at least **four** fully attempted labs earns you a **70 (C)**.
- Submitting at least **three** fully attempted labs earns you a **65 (D)**.
- Submitting at least **two** fully attempted labs earns you a **60 (D)**.

- Submitting **one lab or no** labs will earn you a **55 (E)**.

To get an A (90) on the lab grade, you only need to complete **eight out of the twelve** labs successfully. Submitting **ten** labs guarantees a perfect score of 100. Partial submissions of a lab do not receive any credit.

Class Engagement and Citizenship (10%)
My expectation is that you will have done all the readings before class and completed the labs in earnest. I expect you to arrive on time and leave at the end of our session. I expect you to be engaged with the class and its activities, and that your attention is not elsewhere. I also want to be as accommodating as possible to people who are not accustomed to a U.S.-style classroom that focuses on discussion and may feel shy or uncomfortable speaking in public. Come talk to me, and let's work something out together. For me, citizenship is quality over quantity.

You will also serve as a **discussion leader** and a **lab leader** for at least two modules.

For at least two modules, you will sign up to lead a discussion covering the assigned readings in the syllabus designated "**Readings for Discussion**" but moderated by me. You will offer an analysis and overview of the topic readings. The objective is to explain and explore the concepts and themes of the readings, connecting them to wider issues of becoming a researcher in the computational social science arena.

You will present (with my guidance) for the first 10 minutes or so, and then we will jointly moderate our wider class discussion for the remaining time. You should assume everyone has read the material, so you should **very briefly** summarize the readings for the first ten minutes or so. You should then allocate the remaining time to what you want to discuss. This could involve comparing the pieces, using the works to discuss a broader issue, or highlighting a larger problem you've identified.

You will need to coordinate with your fellow leads in advance. You should run your ideas by me in advance to get feedback. **Please use Google Slides and share the link with me so that setting up is quick and efficient before our class starts.**

In addition, for at least two modules, you will also sign up to lead our review of the lab checkpoints at the end of class. We will go through the lab and review specific checkpoints where people have questions. As such, you will become our subject matter experts for the lab and will help field these questions. You will share your lab and show your work.

You can choose to be a discussion leader and a lab leader for different modules.

Incomplete (I) or Withdrawal (W)
Requests for incomplete (I) or withdrawal (W) must be made in accordance with University policies, which are available at http://catalog.arizona.edu/policy/grades-and-grading-system#incomplete and http://catalog.arizona.edu/policy/grades-and-grading-system#Withdrawal, respectively.

Dispute of Grade Policy
If you feel there was a mistake in grading your work or calculating your grade, you are entitled to ask me to review my grading. (My policy is that you wait at least a week before appealing.) A regrade might result in a lowered grade.

Extra Credit
As there are plenty of opportunities to do well, no extra credit is offered.

Important University Policies
https://catalog.arizona.edu/syllabus-policies

Important Graduate College Policies
https://grad.arizona.edu/policies/academic-policies/graduate-college-syllabus-policy

Use of LLMs
Large language models (LLMs) like ChatGPT, Gemini, etc. are a type of artificial intelligence (AI) engine that can look like it generates text, the code you need for labs, and short answer questions. You are welcome to experiment with these LLMs. However, abuse of LLMs can be traced (e.g., failing to give credit or cite LLMs when it is used), which could result in your suspension or termination from the course and even your program of study. Keep in mind, too, that while the code may appear legitimate, studies are showing LLMs are not all that accurate with sophisticated coding or writing. Exercise your scholarly discretion and maintain a sense of integrity in your computational social science training.

Subject to Change Statement
Information contained in the course syllabus may be subject to change with advance notice as deemed appropriate by the instructor.

LAST UPDATED: 2025-08-12

Scheduled Topics, Activities, and Due Dates

Part I | An Introduction to Computational Social Science Research

Module 1 | Thursday, September 4th 2025

[Required] Readings for Discussion

What Is (and Isn't) Computational Social Science

- Lazer, David et al. 2009. "Computational Social Science." Science 323(5915):721–23.
- Lazer, D. M., Pentland, A., Watts, D. J., Aral, S., Athey, S., Contractor, N., ... & Wagner, C. 2020. Computational social science: Obstacles and opportunities. *Science*, *369*(6507), 1060-1062.

[Required] Background and How to Readings

How to Define a Research Question from a Research Topic

- Mullaney, T.S. and Rea, C., 2022. Chapters 1 and 2 in *Where Research Begins: Choosing a Research Project That Matters to You (and the World)*. University of Chicago Press.

Assessments Due

- [Optional] If you have no prior coding experience or experience with Python, then completing DataCamp's free Introduction to Python Module (4 hours) is highly recommended before starting Lab 1 https://www.datacamp.com/courses/intro-to-python-for-data-science
- Download and install Positron.
- Review Lab 1: Coding for Data I.
- Module 1 memo due on D2L by 9 a.m. (Tucson/MST) Thursday. Respond to the following: What is your research topic, and what might be your research question?

Module 2 | Thursday, September 11th 2025

[Required] Readings for Discussion

A Preview of CSS Research: Stratification, Hierarchy, and Inequality

- Morgan, A.C., Economou, D.J., Way, S.F. and Clauset, A., 2018. Prestige drives epistemic inequality in the diffusion of scientific ideas. *EPJ Data Science*, *7*(1), p.40.
- Lee, E., Clauset, A. and Larremore, D.B., 2021. The dynamics of faculty hiring networks. *EPJ Data Science*, 10(1), p.48.

[Required] Background and How to Readings

Defining Our Models

- Lave, C.A. and March, J.G., 1993. An introduction to models in the social sciences. University Press of America. Chapters 1 and 2
- Rockmore, Dan. 2024. "How Much of the World Is It Possible to Model?" *The New Yorker*, January 15. https://www.newyorker.com/culture/annals-of-inquiry/how-much-of-the-world-is-it-possible-to-model

Assessments Due

- Review Lab 2: Coding for Data II.
- Lab 1: Coding for Data I due by 9 a.m. (Tucson/MST) Thursday.
- Module 2 memo due on D2L by 9 a.m. (Tucson/MST) Thursday. Respond to the following: What is your conceptual model for your research question or research topic?

Module 3 | Thursday, September 18th 2025

[Required] Readings for Discussion

A Preview of CSS Research: Innovation and Diversity

- Hofstra, B., Kulkarni, V.V., Munoz-Najar Galvez, S., He, B., Jurafsky, D. and McFarland, D.A., 2020. The diversity–innovation paradox in science. *Proceedings of the National Academy of Sciences*, 117(17), pp.9284-9291.
- Park, M., Leahey, E. and Funk, R.J., 2023. Papers and patents are becoming less disruptive over time. *Nature*, 613(7942), pp.138-144.

[Required] Background and "How to" Readings

Capturing Social Mechanisms and Processes with Variables

- https://understandingsociety.blogspot.com/2014/06/craver-on-mechanisms-methodology.html
- Ylikoski, P., 2021. Understanding the Coleman boat. In Research handbook on analytical sociology (pp. 49-63). Edward Elgar Publishing.

Due Assessments

- Review Lab 3: Exploratory Data Analysis with Visualizations.
- Lab 2: Coding for Data II due by 9 a.m. (Tucson/MST) Thursday.
- Module 3 memo due on D2L by 9 a.m. (Tucson/MST) Thursday. Respond to the following: What is your explanatory variable that describes the social mechanisms/social processes associated with your research topic or question?

Part II | New Applications and Approaches: Machine Learning

Module 4 | Thursday, September 25th 2025

[Required] Readings for Discussion

Designing Conceptual Spaces

- Davidson, Thomas. 2019. "Black-Box Models and Sociological Explanations: Predicting High School Grade Point Average Using Neural Networks." *Socius: Sociological Research for a Dynamic World* 5 (January): 237802311881770.
- Negro, G., Kovács, B. and Carroll, G.R., 2022. "What's next? Artists' music after Grammy awards." *American Sociological Review*, *87*(4), pp.644-674.

[Required] Background and "How to" Readings

Theory Development

- Swedberg, Richard. 2014. *The Art of Social Theory*. Princeton, NJ: Princeton University Press. Chapters 1, 2, and 3, Pp. 13 - 79

Due Assessments

- Review Lab 4: Exploratory Data Analysis with Machine Learning I.
- Lab 3: Exploratory Data Analysis with Visualizations due by 9 a.m. (Tucson/MST) Thursday.
- Module 4 memo due on D2L by 9 a.m. (Tucson/MST) Thursday. Choose one of the following:

- Respond to the prompt: What is the theory you might propose with your research topic or question? What are you expecting and why? Try to think of ongoing scholarly conversations or controversies that already exist.
- Data analysis memo.

Module 5 | Thursday, October 2nd 2025

[Required] Readings for Discussion

Machine Learning and Ethics in Research

- Salganik, M.J., Lundberg, I., Kindel, A.T., Ahearn, C.E., Al-Ghoneim, K., Almaatouq, A., Altschul, D.M., Brand, J.E., Carnegie, N.B., Compton, R.J. and Datta, D., 2020. Measuring the predictability of life outcomes with a scientific mass collaboration. *Proceedings of the National Academy of Sciences*, 117(15), pp.8398-8403.
- Card, Dallas, Serina Chang, Chris Becker, Julia Mendelsohn, Rob Voigt, Leah Boustan, Ran Abramitzky, and Dan Jurafsky. "Computational analysis of 140 years of US political speeches reveals more positive but increasingly polarized framing of immigration." *Proceedings of the National Academy of Sciences* 119, no. 31 (2022): e2120510119.

[Required] Background and "How to" Readings

Ethics in Research

- Salganik, Matthew J. 2017. Bit by Bit. Princeton University Press. Chapter 6: Ethics.

Due Assessments

- Review Lab 5: Exploratory Data Analysis with Machine Learning II.
- Lab 4: Exploratory Data Analysis with Machine Learning I due by 9 a.m. (Tucson/MST) Thursday.
- Module 5 memo due on D2L by 9 a.m. (Tucson/MST) Thursday. Choose one of the following:
  - Respond to the prompt: What are the ethical concerns that may result from your burgeoning research topic or question? For instance, are you using data from people? If so, what context are these data coming from? Will you need IRB approval? Or, what are the ethical issues that you hope to address with your research topic or question?
  - Data analysis memo.

Part III | New Applications and Approaches: Natural Language Processing (NLP)

Module 6 | Thursday, October 9th 2025

[Required] Readings for Discussion

Grounded Theory and Sentiment

- Nelson, L.K., 2020. Computational grounded theory: A methodological framework. *Sociological Methods & Research*, 49(1), pp.3-42.
- Voigt, R., Camp, N.P., Prabhakaran, V., Hamilton, W.L., Hetey, R.C., Griffiths, C.M., Jurgens, D., Jurafsky, D. and Eberhardt, J.L., 2017. Language from police body camera

footage shows racial disparities in officer respect. *Proceedings of the National Academy of Sciences*, 114(25), pp.6521-6526.

[Required] Background and "How to" Readings

- Evans, James A., and Pedro Aceves. "Machine translation: mining text for social theory." Annual Review of Sociology 42 (2016): 21-50.

Due Assessments

- Review Lab 6: NLP I.
- Lab 5: Exploratory Data Analysis with Machine Learning II due by 9 a.m. (Tucson/MST) Thursday.
- Module 6 memo due on D2L by 9 a.m. (Tucson/MST) Thursday. Choose one of the following:
    - Respond to the prompt: What are three scholarly areas of research that your research topic furthers with its potential findings? Give a name to each area. What are the main questions or controversies in each area? How does your project advance or question each one?
    - Data analysis memo.

Module 7 | Thursday, October 16th 2025

[Required] Readings for Discussion

Knowledge

- Gomez, Charles J., Andrew C. Herman, and Paolo Parigi. 2022. "Leading Countries in Global Science Increasingly Receive More Citations than Other Countries Despite Doing Similar Research." *Nature Human Behaviour*.
- Evans, Eliza, Charles J. Gomez, and Daniel A. McFarland. 2016. "Measuring Paradigmaticness of Disciplines Using Text." *Sociological Science*.

[Required] Background and "How to" Readings

What Are Topic Models

- Mohr, J.W. and Bogdanov, P., 2013. Introduction—Topic models: What they are and why they matter. *Poetics*, *41*(6), pp.545-569.

Due Assessments

- Review Lab 7: NLP II.
- Lab 6: NLP I due by 9 a.m. (Tucson/MST) Thursday.
- Module 7 memo due on D2L by 9 a.m. (Tucson/MST) Thursday. Choose one of the following:
    - Respond to the prompt: What data do you plan to use, and what methods might you consider using? Spell out the "pipeline" of how you would use these data with your particular methods.
    - Data analysis memo.

Module 8 | Thursday, October 23rd 2025

[Required] Readings for Discussion

Culture

- Kozlowski, A.C., Taddy, M., and Evans, J.A., 2019. The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings. *American Sociological Review*.
- Garg, N., Schiebinger, L., Jurafsky, D., & Zou, J. 2018. Word Embeddings Quantify 100 Years of Gender and Ethnic Stereotypes. *Proceedings of the National Academy of Sciences*, 115(16), E3635–E3644

[Required] Background and "How to" Readings

- Arseniev-Koehler, A. 2022. Theoretical Foundations and Limits of Word Embeddings: What Types of Meaning Can They Capture? *Sociological Methods & Research*, 0(0).

Due Assessments

- Review Lab 8: NLP III.
- Lab 7: NLP II due by 9 a.m. (Tucson/MST) Thursday.
- Module 8 memo due on D2L by 9 a.m. (Tucson/MST) Thursday. Choose one of the following:
    - Respond to the prompt: Create three hypothetical figures of results you would *hope* to expect from your research question. Write a few paragraphs describing the figures and the hypothetical takeaways from each figure.
    - Data analysis memo.

Part IV | Classical Applications and Approaches: Social Network Analysis (SNA) and Agent-based Models (ABM)

Module 9 | Thursday, October 30th 2025

[Required] Readings for Discussion

The Micro-to-Marco Link

- Gomez, C.J. and Lazer, D.M., 2019. Clustering knowledge and dispersing abilities enhances collective problem-solving in a network. *Nature Communications*, *10*(1), pp.1-11.
- Schelling, T. C. 1971. Dynamic models of segregation. Journal of Mathematical Sociology, 1(2), 143-186.

[Required] Background and "How to" Readings

- Bruch, E. and Atwell, J., 2015. Agent-based models in empirical social research. *Sociological methods & research*, *44*(2), pp.186-221.

Due Assessments

- Review Lab 9: Agent-based Models
- Lab 8: NLP III due by 9 a.m. (Tucson/MST) Thursday.
- Module 9 memo due on D2L by 9 a.m. (Tucson/MST) Thursday. Choose one of the following:
    - Respond to the prompt: You've sketched out a proto-research proposal so far, but you have likely changed course multiple times. (Or you might have taken the opportunity to do a data analysis memo and might have skipped a step in this process.) Take this memo as an opportunity to summarize where you are with your research topic and question, theoretical contributions to ongoing scholarly conversations, your hypotheses, the data and methods you would use, the mockup figures you plan to use, etc. This memo might recycle bits and pieces from

previous memos, which is perfectly fine if that's the case. N.B., this memo might be a bit longer.
    - Data analysis memo.

Module 10 | Thursday, November 6th 2025

[Required] Readings for Discussion

Social Structure

- McPherson, Miller, and Jeffrey A. Smith. 2019. "Network effects in Blau space: imputing social context from survey data." *Socius* 5 (2019): 2378023119868591.
- Rajkumar, K., Saint-Jacques, G., Bojinov, I., Brynjolfsson, E. and Aral, S., 2022. A causal test of the strength of weak ties. *Science*, 377(6612), pp.1304-1310.

[Required] Background and "How to" Readings

- Kadushin, C., 2012. Understanding social networks: Theories, concepts, and findings. Oxford University Press. Introduction and Chapter 1 (Introduction) and Chapter 11 (Ethical Dilemmas of Social Network Research)

Due Assessments

- Review Lab 10: Social Network Analysis I.
- Lab 9: Agent-based Models due by 9 a.m. (Tucson/MST) Thursday.
- Module 10 memo due on D2L by 9 a.m. (Tucson/MST) Thursday. Choose one of the following:
    - Respond to the prompt: What are some of the limitations of the research proposal you have been sketching out so far over this semester? Think of limitations as both methodological and theoretical. What are the limits to the data or methods you would plan to use? Does your approach *really* answer your research question?
    - Data analysis memo.

Module 11 | Thursday, November 13th 2025

[Required] Readings for Discussion

Diffusion and Communities

- Bail, Christopher, Taylor Brown, and Andreas Wimmer. 2019. "Prestige, Proximity, and Prejudice: How Google Search Terms Diffuse across the World." *American Journal of Sociology* 124 (5): 1496–1548. https://doi.org/10.1086/702007.
- Charles Gomez. 2025. "The Growing Concentration of National Influence in Global Science and its Impact on Future Research." [I will share before class.]

[Required] Background and "How to" Readings

- Newman, M.E., 2006. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23), pp.8577-8582.

Due Assessments

- Review Lab 11: Social Network Analysis II.
- Lab 10: Social Network Analysis I due by 9 a.m. (Tucson/MST) Thursday.
- Module 11 memo due on D2L by 9 a.m. (Tucson/MST) Thursday. Choose one of the following:

- Respond to the prompt: Now that your research proposal is nearing its final shape, consider its eventual publication. What communities, subfields, or public practitioners would be most interested in your findings? Describe the "so what" of your project and outline the broader impact you hope it will have. Which journals or outlets are you thinking of as a target?
- Data analysis memo.

Part V | The Future with A.I. and Managing Big Data

Module 12 | Thursday, November 20th 2025

[Required] Readings for Discussion

Artificial Intelligence and the Future of Computational Social Science

- Davidson, Thomas, and Daniel Karell. 2025. "Integrating Generative Artificial Intelligence into Social Science Research: Measurement, Prompting, and Simulation." *Sociological Methods & Research* 54(3):775–93. doi:10.1177/00491241251339184.
- Kozlowski, Austin C., and James Evans. 2025. "Simulating Subjects: The Promise and Peril of Artificial Intelligence Stand-Ins for Social Agents and Interactions." *Sociological Methods & Research* 54(3):1017–73. doi:10.1177/00491241251337316.

[Required] Background and "How to" Readings

- Read materials on "How to Write a Computational Social Science Paper."

Due Assessments

- Review Labs 12: SQL and HPC.
- Lab 11: Social Network Analysis II due by 9 a.m. (Tucson/MST) Thursday.
- Module 12 memo due on D2L by 9 a.m. (Tucson/MST) Thursday. Choose one of the following:
  - Respond to the prompt: Looking beyond this single proposal, what are the next steps for this research agenda? Identify one or two future projects that could logically follow your current one. How might emerging computational methods or new forms of data enable you to answer deeper questions or overcome the limitations you've identified in your current project?
  - Data analysis memo.

Thanksgiving | Thursday, November 27th 2025
NO CLASS

Part VI | Research Proposal

Last Day of Class | Thursday, December 4th 2025
Class Party and Final Presentations

[Required] Readings for Discussion

- None.

[Required] Background and "How to" Readings

- None.

Due Assessments

- Lab 12: SQL and HPC due by 9 a.m. (Tucson/MST) Thursday.
- Research proposal presentations are due today, worth 10% of your final grade.
- Research proposal due on D2L in one week.