Computational social science for organizational research

Russell J. Funk

Carlson School of Management
University of Minnesota

March 13, 2024

Roadmap

- Introductions
- Syllabus
- Assignments
- Information technology
- Course overview
- Discussion of readings
- Introduction to Python

Introductions

About me

Russell J. Funk

Position

- Associate Professor
- Strategic Management and Entrepreneurship
- Carlson School of Management
- University of Minnesota

Background

- From Plainfield, Illinois
- AB from University of Chicago
- PhD from University of Michigan
- At Minnesota since 2014

Research

- Networks
- Innovation
- SciSci

Personal

- 2 sons
- Love the outdoors

Contact

- E-mail: rfunk@umn.edu
- Phone: (612) 626-1598
- Office: CSOM #3-354
- Meetings: By appointment

About you

- What's your name?
- What's your degree program?
- What are your research interests?
- What do you hope to get out of the class?

Syllabus

MGMT 8404

Computational Social Science for Organizational Research

Spring 2024

Instructor

Russell J. Funk
Associate Professor
Strategic Management and Entrepreneurship
Carlson School of Management
University of Minnesota

Office: CSOM 3-354
Office Hours: By apointment
Phone: (612) 626-1598
E-mail: rfunk@umn.edu

Meeting Times

Section 001 Wednesday 2:30 p.m.–5:30 p.m.

Meeting Location

CSOM 3-377

Assignments

Overview

Assignment	**Points**
Readings, discussion, and memos	40
Research paper	50
Research presentations	10
Total	**100**

Readings, discussion, and memos

Readings and discussion

- You should complete the readings before each class session and come prepared to discuss.

Critical memos

- For **4 class sessions**, you should write a 1 page critical response memo.
- The format should be typed, singled spaced, 1 inch margins, 11 or 12 point font, PDF.
- Submission should be via Canvas.
- Your memo should be based on the readings for the day.

Code memos

- In place of writing a critical response memo, you may also write a "code memo."
- For a code memo, you'll apply a method from the day's session or a prior session.
- Walk the reader through what you did, any interesting findings, questions or roadblocks.
- Over the course of the term, you should **aim to do at least 2 code memos.**

So to clarify, you should write 4 memos total (at least 2 and up to 4 code memos).

Research paper

(approximately 15 pages)

Apply a method from the course to a problem in organizational theory, strategic management, or entrepreneurship. Your paper should include the following components.

1. An introduction that positions the paper with respect to prior work, articulates a research question, and suggests a solution that uses computational social science.
2. A data and methods section that gives an overview of the methods you'll use, your data and so forth.
3. A results section, where you present the findings from your analysis.
4. A discussion section, where you review the implications of your findings and discuss the strengths and weaknesses of your approach.
5. A "next steps" section, where you discuss how you would build out and strengthen the paper and analysis.
6. A section with relevant tables and figures as necessary.

In short, your term paper should more or less include the components of a regular academic article, sans the theory and/or hypothesis development section.

You are encouraged to build off your existing papers, ideas, and/or code memos.

Research paper presentations

- We will carve out time for presentations relating to the research paper.
- The idea will be to use this as an opportunity for early feedback.
- More details will be given in class.

Information technology

Information technology

We'll be using two main information technologies for our class. . .

Canvas

- Assignment submission
- Grades
- Communications
- http://canvas.umn.edu

Github

- Jupyter notebooks
- Slide decks
- http://github.com/russellfunk

Course overview

7 sessions (Subject to small changes. . .)

Introduction

- Preliminaries
- What is CSS?
- Crash course on Python + SQL

Machine learning

- Supervised/Unsupervised learning
- Deep learning
- Prediction
- Causal inference
- Data work

Networks

- Introduction to networks
- Foundational network models
- Generative models
- Graph neural networks

Documents and text

- Corpora
- Vector space model
- Document similarity
- Topic models
- Text networks

Language and words

- Linguistics
- Word embeddings
- Sentiment analysis

Simulation

- Understanding real world data
- Creating your own worlds
- Best practices

Geography and space

- Spatial autocorrelation
- Sociometric badges
- Mobile phone metadata
- Social media and images

A few more details...

Caveat on session order

- I may (and likely will) change the order of sessions around.
- Will do this based on our pace and so forth.
- Please bear with me.

Session design

- Part 1—a short introductory overview from me.
- Part 2—a discussion of the readings.
- Part 3—a hands on "lab" session.

Lab sessions

- Will be based on custom Jupyter notebooks made for this class.
- Not all methods studied in the readings will be covered.
- Instead, focus will be on getting you experience with the basics.
- Focus on what you need to know to learn more on your own.

Software

- You will get the most out of this class by working through the notebooks.
- If you have trouble getting Python up and running, let me know.
- You may also want to check in with your IT office.

Discussion of readings

What is computational social science?

Is computational social science a new **kind** of science or a new **approach** to science?

If the former...

- What benefits might it have over the traditional scientific method?
- What are the limitations?

If the latter...

- How do we define CSS vis-a-vis established approaches?
- Do we include within CSS any application of computation?
- If so, then do we consider things like regression (or even word processing) as CSS?
- Or perhaps we want to only include as CSS things we can't do without computers?
- But defined that way, some things we think of as CSS won't meet our definition.
  - We can run simple simulations with coin flips and random number tables.
- Perhaps CSS just means scale (e.g., of data) or computational intensity?
- But then does what counts at CSS change over time?
  - Was regression CSS in the 1960s, 1970s, 1980s?

Required readings

- Matthes, Eric. (2023) "Getting started," "Variables and simple data types," and "Introducing lists." In *Python crash course, 3rd edition: A hands-on, project-based introduction to programming*: pp. 3-14, pp. 15-32, pp. 33-49. San Francisco: No Starch Press.
- Foster, I., Ghani, R., Jarmin, R. S., Kreuter, F., & Lane, J. (2016). "Databases" pp. 75-108. In *Big data and social science: A practical guide to methods and tools*. CRC Press.

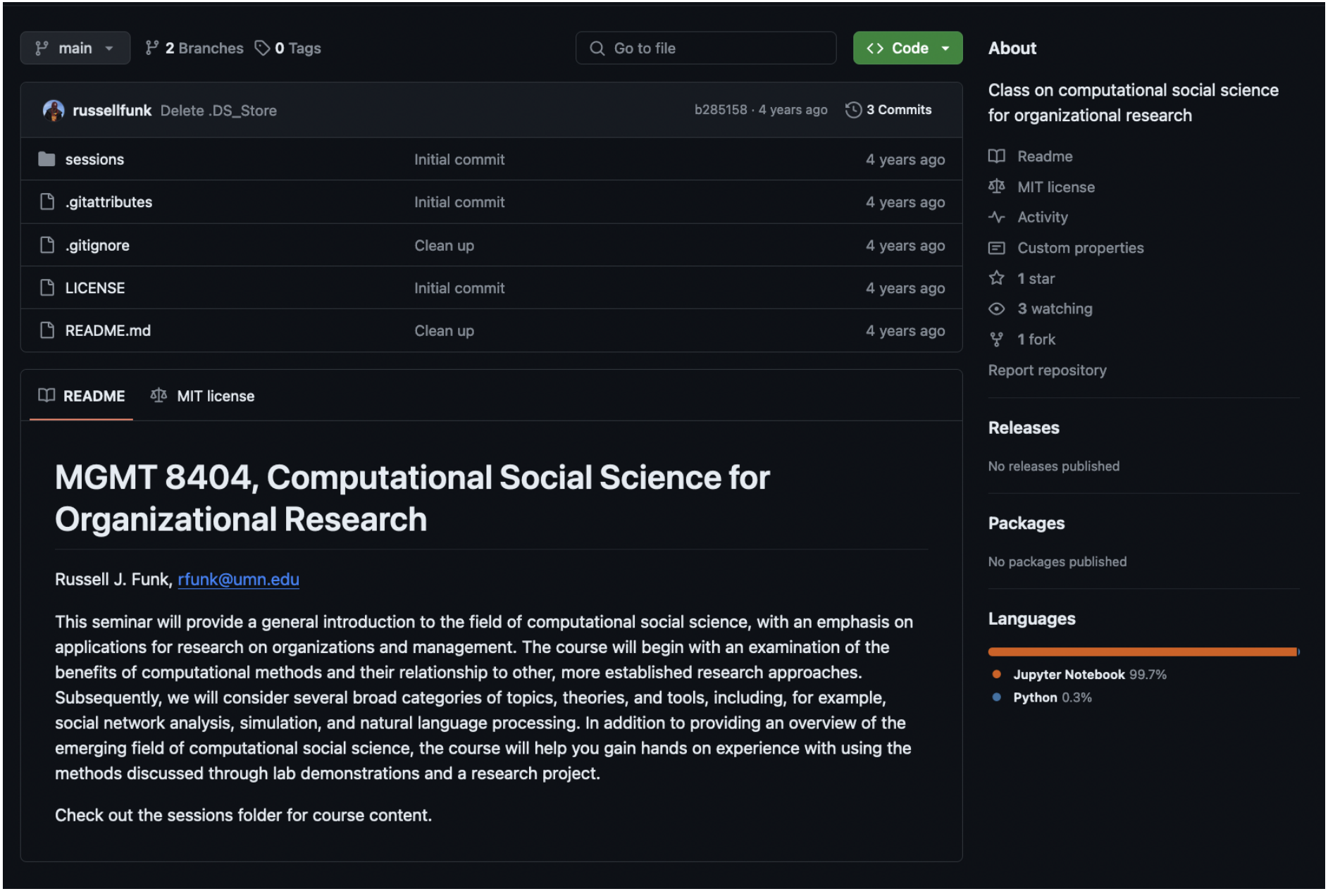
  `https://textbook.coleridgeinitiative.org/chap-db.html`.
  - See also: Luna, Javier Canales. "MySQL Tutorial: A Comprehensive Guide for Beginners." *DataCamp*, Accessed March 4, 2024. `https://www.datacamp.com/tutorial/my-sql-tutorial`.
- Anderson, Chris. (2008) "The end of theory: The data deluge makes the scientific method obsolete." *Wired*, June 23.
- Edelmann, Achim, Tom Wolff, Danielle Montagne, and Christopher A. Bail. (2020) "Computational Social Science and Sociology." *Annual Review of Sociology* 46(1)1: 61-81.
- Halevy, Alon, Peter Norvig, and Fernando Pereira. (2009) "The unreasonable effectiveness of data." *IEEE Intelligent Systems* 24(2): 8-12.
- Ziems, Caleb, William Held, Omar Shaikh, Jiaao Chen, Zhehao Zhang, and Diyi Yang. (2023) "Can Large Language Models Transform Computational Social Science?" *arXiv preprint arXiv:2305.03514*: `https://arxiv.org/abs/2305.03514`.

Optional readings

- Bradshaw, Gary F., Patrick W. Langley, and Herbert A. Simon. (1983) "Studying scientific discovery by computer simulation." *Science* 222(4627): 971-975.
- Evans, James, and Andrey Rzhetsky. (2010) "Machine science." *Science* 329(5990): 399-400.
- George, Gerard, Ernst C. Osinga, Dovev Lavie, and Brent A. Scott. (2016) "Big data and data science methods for management research." 59(5): 1493-1507.
- Hey, Tony, Stewart Tansley, and Kristin Tolle (2009) "Jim Grey on eScience: A transformed scientific method." In *The Fourth Paradigm: Data-Intensive Scientific Discovery*: pp. 281-354. Redmond, WA: Microsoft Research.
- Lazer, David, Ryan Kennedy, Gary King, and Alessandro Vespignani. (2014) "The parable of Google Flu: Traps in big data analysis." *Science* 343(6176): 1203-1205.
- Lazer, David, Alex Pentland, Lada Adamic, Sinan Aral, Albert-László Barabási, Devon Brewer, Nicholas Christakis et al. (2009) "Computational social science." *Science* 323(5915): 721-723.
- Lazer, David M. J., Alex Pentland, Duncan J. Watts, Sinan Aral, Susan Athey, Noshir Contractor, Deen Freelon, Sandra Gonzalez-Bailon, Gary King, Helen Margetts, Alondra Nelson, Matthew J. Salganik, Markus Strohmaier, Alessandro Vespignani, and Claudia Wagner. (2020) "Computational Social Science: Obstacles and Opportunities. Data Sharing, Research Ethics, and Incentives Must Improve." *Science* 369(6507): 1060-1062.
- Salganik, Matthew (2018) "Ethics." In *Bit-by-bit: Social research in the digital age*: pp. . Redmond, WA: Microsoft Research.
- Truhn, Daniel, Jorge S. Reis-Filho, and Jakob Nikolas Kather. (2023) "Large Language Models Should Be Used as Scientific Reasoning Engines, Not Knowledge Databases." *Nature Medicine* 29: 2983–2984.

Introduction to Python

Let's dig in...


main
2 Branches
0 Tags
Go to file
Code
russellfunk Delete .DS_Store
b285158 · 4 years ago
3 Commits
sessions Initial commit 4 years ago
.gitattributes Initial commit 4 years ago
.gitignore Clean up 4 years ago
LICENSE Initial commit 4 years ago
README.md Clean up 4 years ago
README
MIT license
MGMT 8404, Computational Social Science for Organizational Research
Russell J. Funk, rfunk@umn.edu
This seminar will provide a general introduction to the field of computational social science, with an emphasis on applications for research on organizations and management. The course will begin with an examination of the benefits of computational methods and their relationship to other, more established research approaches. Subsequently, we will consider several broad categories of topics, theories, and tools, including, for example, social network analysis, simulation, and natural language processing. In addition to providing an overview of the emerging field of computational social science, the course will help you gain hands on experience with using the methods discussed through lab demonstrations and a research project.
Check out the sessions folder for course content.
About
Class on computational social science for organizational research
Readme
MIT license
Activity
Custom properties
1 star
3 watching
1 fork
Report repository
Releases
No releases published
Packages
No packages published
Languages
Jupyter Notebook 99.7%
Python 0.3%

Appendix