

SOCS-SHU 190

Introduction to the Computational Social Science Methods

Subject to Change Statement

The syllabus and course schedule may be subject to change. Changes will be communicated via email and in NYU LMS (Brightspace). It is your responsibility to check email messages and course announcements to stay current in the course.

Instructor Information*

- Instructor: Zixi Chen, Ph.D.
- Office: N717
- Office hours: TBA
 - For additional contact hours, please schedule an appointment by email
- Email: zc3154@nyu.edu

Course Information*

- **Course Description:** This course invites students from diverse backgrounds and provides an accessible introduction to the burgeoning field of Computational Social Science (CSS). We aim to explore research design and data skills at the intersection of social sciences and digital innovations. This course emphasizes using big and rich social digital data to understand and explain societies and human behaviors. We introduce fundamental topics in CSS, including digital trace data collection, social network analysis, and text-as-data, using the R programming language. Besides discussing the new opportunities of CSS, this course also highlights critical topics, such as data ethics and data-driven bias.
- **Prerequisite:**
 - Prior coding experience is not required. We will learn R to reach an introductory level in this course.
 - Students are encouraged to come with a project interest or are willing to develop a project interest along with the class.
- **Class meeting days and times:** 2:15 PM - 3:30 PM M W
- **Class room number:** S219

Course Overview and Learning Outcomes

In this entry-level CSS course, we introduce four modules: 1) CSS: Opportunities and Challenges, 2) Data in Digital Age, 3) Text-as-Data, and 4) Social Network Analysis.

Each module has three main parts:

- An introduction of the weekly topics led by the instructor.
- In-class discussions and activities are designed to introduce two to three classic or recent applied social science research that uses that method. These materials are labeled as background readings.
- R lab sessions: the instructor will demonstrate how to use the method with R and provide R instructions and exercises in Quarto (.qmd).

The instructor will preview the next module at the end of each module.

Learning Outcomes

	Course Learning Outcomes	STS Learning Outcomes	AT Learning Outcomes
1	Understand the basic principles of computational methods and describe their applications in social science.	STS 1, STS 2	AT 1
2	Apply basic CSS methods to answer social science questions.	STS 1, STS 2	AT 2, AT 3
3	Evaluate the data sources and models from the perspectives of social science and data science.	STS 1, STS 2	AT 4, AT 5
4	Earn hands-on experience using R to collect, process, and analyze big and rich social digital data.	STS 1	AT 5

In addition, the first two course learning outcomes correspond with the STS's additional learning outcomes related concerning communication and critical thinking. Meanwhile, the last two course learning outcomes align with the STS's additional learning outcomes in terms of information literacy and critical thinking.

Course Requirements*

Class Participation

Active engagement and participation are essential components of this course. You will have one point deducted from “In-class discussion and activities” grade for each class you missed without a legitimate reason. Also see details in Absences and Grades.

Getting-to-know-you survey

The information will help the instructor to know the students better and to accommodate the learning materials to meet the students’ needs. It is due in Week 2. Click [the survey link here](#).

- a) Describe your CSS interests in five sentences. They can be your general idea of applying CSS methods in your work or any learning objectives you have for this class.
- b) Answer survey items about your academic background, such as research interest and research methods you have learned in other courses.

Assignments x 2

The assignments are served to tight CSS applications with real-life experiences. Students can work in groups (no more than three people per group) or individually.

- a) Asgmt. 1.
 Topic option1: Discuss one type of digital data that emerges during a specific event of interest (e.g., the Spring Break, announcement of a new LLM model, a significant political or social news event).
 Topic option 2: Analyze the data privacy policy of one or two generative AI platforms (e.g., ChatGPT, Gemini, Deepseek, etc.)

Example questions to discuss will be provided. Write down your observations (3-5 pages, double-spaced, 12-point font of Times New Roman, Chicago or APA 7 style) and share your observations in the next class. Examples of observational questions could be: What are the emerging data characteristics? What social phenomena does it inform, might explore and/or explain?

- b) Asgmt. 2.
 Find a list of CSS-relevant job posts in your field ($4 < \text{posts} < 10$ for individual work and $14 < \text{posts} < 26$ for group work) and compose a job description corpus. Identify three to five textual data or measures you think are most informative to you. Submit the corpus data (and code scripts, if any) and write the introduction and explanation of the textual data/measures in the *Word* or *R* (.qmd or .rmd) files. Introduce your work to the class, including the tools you used and why you think those data/measures are informative to you.

R assignments x4

Each R assignment will have 3-5 exercises (see the due dates in the Course Schedule below). The exercises will be assigned one week ahead of the due date. Students can work on them in groups (no more than three people per group) or individually.

Class project

The class project aims to allow students to apply CSS methods to solve their interested research or real-life questions. The project idea can be developed from their previous assignments and work.

- a) Project proposal
 - i. Structure:
 - Topic-specific questions (no more than three questions) you are interested in exploring;
 - The background of the questions (i.e., why they are important to answer); and
 - CSS approaches that you plan to employ.
 - ii. Format: no more than *three* pages (excluding figures, tables, and references), double-spaced, 12-point font of Times New Roman, Chicago or APA 7 style
- b) Final paper
 - i. Structure:
 - Develop the three components in the proposal with the instructor's feedback.
 - Describe the analytical process and results. Notice that a formal and complete project is not required. Using a sample or toy data set for preliminary analysis and results is sufficient. Students can also provide hypotheses or expected results if no data is available, or analyses are incomplete.
 - ii. Format: no more than *eight* pages (excluding figures, tables, and references), double-spaced, 12-point font of Times New Roman, Chicago or APA 7 style
- c) In-class presentation (In Week 14)

Assigned Readings

As outlined in the Course Schedule and Course Materials, students should complete the *required* readings before each week's instruction.

The background readings are *optional*, and students are not required to complete them. Students are encouraged to skim the background readings to know the basic ideas, as these readings will be used to enrich in-class discussions and activities. The instructor uses these background readings to show the applications of CSS methods in various disciplines.

Grading of Assignments

The grade for this course will be determined according to the following formula:

Assignments/Activities	% of Final Grade
------------------------	------------------

Getting-to-know-you survey	5%
In-class discussion and activities	15%
Assignments	16% (8% for each assignment)
R assignments	24% (6% for each exercise)
Class project proposal	10%
Class project presentation	10%
Class project write-up	20%

Letter Grades

Letter grades for the entire course will be assigned as follows:

Letter Grade	Points	Percent
A	4.00	92.5% and higher
A-	3.67	90.0 – 92.49%
B+	3.33	87.5 – 89.99%
B	3.00	82.5 – 87.49%
B-	2.67	80.0 – 82.49%
C+	2.33	77.5 – 79.99%
C	2.00	72.5 – 77.49%
C-	1.67	70.0 – 72.49%
D+	1.33	67.5 – 69.99%
D	1.00	60 – 67.49%
F	.00	59.99 and lower

Course Schedule

Topics and Assignments

Week/ Date	Module	Topic	Reading	Assignment Due
Week 1, Aug 31 & Sep 2	CSS: Opportunities & Challenges	Introduction to CSS & course overview.	Required readings (14 pages): <ul style="list-style-type: none"> • BiB, Ch. 1.1- 1.3 (pp. 1-6) • CTS, Ch. 10.1-10.5 (pp.251-258) 	
Week 2, Sep 7, 9	CSS: Opportunities & Challenges	Intro. to CSS (cont'd) R session 1: R basics (link)	Background readings (24 pages): <ul style="list-style-type: none"> • Golder, S. A., & Macy, M. W. (2014). Digital footprints: Opportunities and challenges for online social research. <i>Annual Review of Sociology</i>, 40, 129-152. 	Getting-to-know-you survey due on Sep. 9
Week 3, Sep 14, 16	Data in Digital Age	Research Ethics in CSS Intro. to GenAI for coding learning R session 1 (cont'd) (link)	Required readings (19 pages): <ul style="list-style-type: none"> • CTS, Ch. 11 (pp. 268-286) Background readings (34 pages): <ul style="list-style-type: none"> • BiB, Ch. 6.1, 6.4-6.7 (pp.281-283; pp. 294-324) • 	
Week 4, Sep 21, 23	Data in Digital Age	Data in Digital Space R session 2: data collection and wrangling (link)	Required readings (8 pages): <ul style="list-style-type: none"> • BiB, Ch. 2.1-2.2 (pp. 13-16) • BiB, Ch. 3.1-3.2 (pp. 85-88) Background readings (22 pages): <ul style="list-style-type: none"> • Tufekci, Z. (2014, May). Big questions for social media big data: Representativeness, validity and other methodological pitfalls. In <i>Proceedings of the international AAAI conference on web and social media</i> (Vol. 8, No. 1, pp. 505-514). • CTS, Ch. 9.2-9.4 (pp. 224-235) 	R Ex. 1 due on Sep. 23
Week 5, Sep 28, 30	Data in Digital Age	Data collection and linkage	Required readings (10 pages):	Asgmt 1. due on Sep. 30

		R session 2 (cont'd) (link)	<ul style="list-style-type: none"> BiB, Ch.3.5 (pp.107-116) Background readings (22 pages): <ul style="list-style-type: none"> CTS, Ch. 9.5-9.6 (pp. 235-245) 	
National Day recess (no class the week of Oct 5)				
Week 6, Oct 12, 14	Data in Digital Age	Online experiments LLM for silicon sampling R session 2 (Viz) (link)	Required readings (18 pages): <ul style="list-style-type: none"> BiB, Ch. 4.1-4.3 (pp. 147-157) CTS, Ch. 12.1-12.3 (pp. 292-298) 	
Week 7, Oct 19, 21	Text-as-Data	Overview of textual data and text analysis methods R session 3.a: Textual data basics (link)	Required readings (10 pages): <ul style="list-style-type: none"> TAD, Ch. 1 (pp. 1-10) Background readings (7 pages): <ul style="list-style-type: none"> TAD, Ch. 2.1-2.5 (pp. 13-19) 	R Ex. 2 due on Oct. 21
Week 8, Oct 26, 28	Text-as-Data	Bag-of-words and similarity Metrics. R session 3.a (cont'd) (link)	Background readings (read the Intro. and Discussion): <ul style="list-style-type: none"> Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. <i>Political analysis</i>, 21(3), 267-297. 	Class project proposal due on Oct. 28
Week 9, Nov 2, 4	Text-as-Data	Word Embeddings.	Background readings (skim): <ul style="list-style-type: none"> Kozlowski, A. C., Taddy, M., & Evans, J. A. (2019). The geometry of culture: Analyzing the meanings of class through word embeddings. <i>American Sociological Review</i>, 84(5), 905-949. 	
Week 10, Nov 9, 11	Text-as-Data	Topic Models. R session 3.b (Topic modeling) (link)	Required readings (8 pages): <ul style="list-style-type: none"> TAD, Ch. 13.1-13.2 (pp. 147-152) TAD, Ch. 13.5-13.6 (pp. 160-161) Background readings (19 pages):	R Ex. 3 due on Nov. 11

			<ul style="list-style-type: none"> Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., ... & Rand, D. G. (2014). Structural topic models for open-ended survey responses. <i>American journal of political science</i>, 58(4), 1064-1082. 	
Week 11, Nov 16, 18	Text-as-Data	Exploring LLM for text analysis	Background readings (skim) <ul style="list-style-type: none"> Caleb Ziems, William Held, Omar Shaikh, Jiaao Chen, Zhehao Zhang, Diyi Yang; Can Large Language Models Transform Computational Social Science?. <i>Computational Linguistics</i> 2024; 50 (1): 237–291. 	Asgmt 2. due on Nov.18
Week 12, Nov 23, 25	Social Network Analysis	An overview of basic concepts and topics in social network analysis	Required readings (11 pages): <ul style="list-style-type: none"> CTS, Ch.4.1-4.4 (pp. 74-84) Background readings (32 pages): <ul style="list-style-type: none"> Stephen P. Borgatti et al., "Network Analysis in the Social Sciences," <i>Science</i> 323, no. 5916 (2009): 892–895. 	
Week 13, Nov 30, Dec 2	Social Network Analysis	SNA overview (cont'd) R session 4 (SNA) (link)	<ul style="list-style-type: none"> McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. <i>Annual review of sociology</i>, 27(1), 415-444. 	R Ex. 4 due on Dec. 2
Week 14, Dec 7, 9	Pitch Days	Students present and work on their class projects.	<ul style="list-style-type: none"> No readings 	
Final Week				Class project paper due on Dec. 16 (No late penalty)

Course Materials

Required Textbooks & Materials

- Salganik, M. J. (2019). *Bit by bit: Social research in the digital age*. Princeton University Press. [\[Free online version\]](#)

- Students are required to read the assigned chapters. It is referred to as “*BiB*” in the Course Schedule.
- R sessions are posted on our course website:
<https://zixichen.quarto.pub/r-session-of-socs-sh-190-intro-to-css/>

Supplemental Materials

- Nelimarkka, M. (2022). *Computational Thinking and Social Science: Combining Programming, Methodologies and Fundamental Concepts*. SAGE Publications
 - This textbook is relatively new. We will select a few sections to read alongside *BiB* and adapt some exercises from this book into in-class activities. It is referred to as “*CTS*” in the Course Schedule.
- Grimmer, J., Roberts, M. E., & Stewart, B. M. (2022). *Text as data: A new framework for machine learning and the social sciences*. Princeton University Press.
 - In this class, we only read the first two chapters for a general overview of text-as-data applications in social science. It is referred to as “*TAD*” in the Course Schedule.
- Supplemental R resources.
We encourage students to use the following materials for R preview.
 - [R bootcamp](#) from [Data Services](#) of NYU Shanghai Library
 - [Rstudio](#) official educational resources.
 - Learn the [R workflow and tidyverse](#) (Read *Introduction*, *Explore*, and *Wrangle*). The [tidyverse](#) collection will be the backbone of our R labs.
 - Basic text analysis in [Quanteda](#) and [tidy](#) approach.
 - Basic network analysis in [igraph](#).
 - A curated resource pool with many popular R books and packages if you like to dig more: [Awesome R](#).
- Supplemental CSS textbooks and resources.
 - Scott, J., & Carrington, P. J. (2011). *The SAGE handbook of social network analysis*. SAGE publications.
 - A curated CSS resource pool: [Awesome Computational Social Science](#)

Course Policies

Attendance and Tardiness

You are expected to attend *all* scheduled classes. If you are unable to attend a class, notify me *before* that class.

Absences and Grades

- 5 absences will lead to an F for your participation grade.
- 8 absences will lead to a 40% reduction in your final grade.
- 10 absences will lead to failure of the course.

Absence Exceptions

Observance of Religious Holidays: You may miss class for the observance of religious holidays. If you anticipate being absent because of religious observance, notify me in advance so we can create a plan for making up missed work. For more on this policy:

<https://www.nyu.edu/about/policies-guidelines-compliance/policies-and-guidelines/university-calendar-policy-on-religious-holidays.html>

Competitions, Conferences, Presentations: You are permitted to be absent from classes to participate in competitions, conferences, and presentations, either at home or out of town, as approved by the Associate Provost for Academic Affairs. Review the Undergraduate Bulletin for the conditions you must meet to obtain approval for this kind of absence.

Illness: When students are ill, they are expected to notify professors in advance of class, if at all possible. If the instructor determines that it is an excused absence then the student should negotiate with the professor the time and place for make-up of assignments, tests and/or examinations missed. Students who have been seriously ill, hospitalized and/or miss more than a week of classes due to medical reasons, should contact their academic advisor to discuss appropriate options for missed classes and/or coursework.

A student with an injury or medical condition that requires ongoing accommodations (temporary or permanent) should contact the NYU Moses Center for Student Accessibility (CSA). If an accommodation is recommended by the Moses Center, then Academic Affairs may communicate on behalf of students to advocate for excused absences/extensions. Reasonable accommodations, considering the course objectives, student learning, and fair standards, are ultimately decided by the professor.

Students who, in the judgment of the instructor, have not substantially met the requirements of the course or who have been excessively absent are not considered to have withdrawn from the course if they remain on the roster and may be given the final grade of F.

Tardiness

Punctual arrival is mandatory for this class. Be on time. Please do not leave in the middle of class unless it is an emergency.

Late Assignments

Late Penalty: Should you need to submit an assignment late, the penalty is a deduction of one-third of a letter grade for each 24-hour period it is late (e.g., an A paper submitted one day late would earn an A-, two days late a B+, and so on).

Final Project Paper: The deadline for the final project paper is firm. To ensure I can grade all projects and submit final course grades to the university on time, late final project papers cannot be accepted and will receive a failing grade for the assignment. Please plan accordingly and feel free to discuss your progress with me well in advance of this deadline.

Electronic Devices

Mobile Devices: You may not use mobile devices in class unless otherwise indicated.

Recording Class: To ensure the free and open discussion of ideas, you may not record classroom lectures, discussion and/or activities without my advance written permission; any such recording can be used solely for your own private use. If you have approved accommodations from the Office of Disability Resources permitting the recording of class meetings, you must present the accommodation letter to me in advance of any recording. Distribution or sale of class recordings is prohibited without the written permission of the instructor and other students who are recorded.

Instructional Technology

Email Communication: I will contact you regularly via the NYU LMS (Brightspace). You should check for emails from me (reminders, logistics, updates, etc.) at the email address connected to NYU LMS (Brightspace). Please note that I will try to respond to your emails within 24 hours. Do not expect immediate responses to emails sent late at night or on the weekends.

Assignment Notification: All assignments will be posted to our class NYU LMS (Brightspace) Site. You are responsible for looking at NYU LMS (Brightspace) after each class period to learn about the next homework assignment.

Instructional Technology Tools and Assistance: If you need background on specific instructional technology tools, such as Zoom, NYU LMS (Brightspace) and Voicethread, check the [RITS Student Toolkit](#). You may also email shanghai.rits@nyu.edu for assistance.

Academic Honesty/Plagiarism*

Carefully read NYU Shanghai's *Statement on Academic Integrity* (in the [Undergraduate Bulletin](#)). Breaches of academic integrity could result in failure of an assignment, failure of the course, or other sanctions, as determined by the Academic Affairs office.

Disability Disclosure Statement*

NYU Shanghai is committed to providing equal educational opportunity and participation for students with disabilities. It is NYU Shanghai's policy that no student with a qualified disability be excluded from participating in any NYU Shanghai program or activity, denied the benefits of any NYU Shanghai program or activity, or otherwise subjected to discrimination with regard to any NYU Shanghai program or activity. Any student who needs a reasonable accommodation

based on a qualified disability should register with the [Moses Center for Student Accessibility](#) for assistance. Students can [register online](#) through the Moses Center and can contact the Academic Accommodations Team at shanghai.academicaccommodations@nyu.edu with questions or for assistance.

Title IX Statement

Title IX of the Education Amendments of 1972 (Title IX) prohibits discrimination on the basis of sex in educational programs. It protects victims of sexual or gender-based bullying and harassment and survivors of gender-based violence. Protection from the discrimination on the basis of sex includes protection from being retaliated against for filing a complaint of discrimination or harassment. NYU Shanghai is committed to complying with Title IX and enforcing University policies prohibiting discrimination on the basis of sex. Mary Signor, Executive Director of the Office of Equal Opportunity, serves as the University's Title IX Coordinator. The Title IX Coordinator is a resource for any questions or concerns about sex discrimination, sexual harassment, sexual violence, or sexual misconduct and is available to discuss your rights and judicial options. University policies define prohibited conduct, provide informal and formal procedures for filing a complaint and a prompt and equitable resolution of complaints.

Links to the Title IX Policy and related documents:

- [Sexual Misconduct, Relationship Violence, and Stalking Policy](#)
- [Procedures for Complaints Against Students](#)
- [Procedures for Complaints Against Employees](#)
- [Resource Guide for Students](#)
- [Resource Guide for Employees](#)

Policy on the Use of AI/GenAI Tools

Use of AI or GenAI tools, such as ChatGPT, is permitted in this course. To adhere to NYUSH's guidelines for academic integrity and our scholarly values, you must disclose any AI-generated material they use and properly attribute it, including in-text citations, quotations, and references.

Required Disclosure

You should include the following statement in assignments to indicate use of a GenAI tool: "I/We acknowledge the use of [Generative AI Tool Name] (version [if applicable]), in the preparation of this assignment. Specifically, [Generative AI Tool Name] was utilized for [list specific purposes, e.g., brainstorming ideas, suggesting coding syntax, proofreading, formatting citations]." Failure to acknowledge the use of a GenAI tool will result in penalties.

If you used AI or GenAI tools for the R assignments, you must submit the log file documenting your prompts and chat with the tools. This log will be used solely by the instructor to understand students' learning processes and improve instruction. Your privacy will be respected, and the content will not be shared or used for any other purpose.

Be aware that all large language models still have a tendency to make up incorrect facts and fake citations and code generation models have a tendency to produce inaccurate outputs. Critical evaluation of AI-generated content is essential.

Responsible Uses

When engaging with AI or GenAI tools, do not provide information such as name, University ID#, phone number, address etc. which is personal information. Doing so is a violation of privacy regulations.

Permitted and Prohibited Uses

Students are welcome to use AI tools to learn coding, including annotating scripts and asking for coding syntax and debugging suggestions. Use of GenAI/AI Tools is strictly prohibited in the following areas: writing research questions or paper statements; writing the project proposal or outline; writing any part of the paper or paper elements, except for proofreading or critiquing.

Suspected violations of academic integrity will be referred to Academic Affairs. If uncertain, students should consult their professor for guidance.

Academic Resources

ARC Services

The Academic Resource Center (ARC) offers a wide range of academic support services for students enrolled in NYU Shanghai courses. Whether you're tackling challenging course content, working on a paper, preparing for a presentation, or developing study and interpersonal skills, the ARC is here to support your academic success. You are welcome to walk into the ARC to meet with Discipline Learning Assistants (LAs) for support in STEM, Business, Economics, and Chinese Language courses. You can also meet with Writing & Speaking Learning Assistants, who can provide assistance with specific questions related to writing, speaking, reading, or listening in English. Additionally, the ARC offers a variety of academic skills workshops to help you develop effective study strategies and enhance your interpersonal skills. All tutoring sessions are drop-in only. Check the ARC Learning Assistants' [tutoring schedule](#) or [register](#) for upcoming academic skills workshops. Visit the [ARC website](#) to learn more about academic support resources.

Library Support

Library Services

The Library is available to support your research needs. They have access to over 27,000 print resources, 2,000 DVDs, and 1,000 databases (including over a million e-books, as well as streaming audio and video and image databases).

Librarians with expertise in your research topic are available to meet either *in person* or *online* [by appointment](#) or [by email](#) to help you navigate the research process. Our library team features

experts in Business, Arts & Humanities, STEM, Social Sciences & Economics, and data tools & resources. Ask us how we can assist you in developing a research question and formulating a research strategy, to selecting databases, requesting materials, and citing your sources. Visit shanghai.nyu.edu/library for more information on:

- 24/7 access to e-books, e-journals, streaming media, and databases
- Booking one-on-one consultations for research help
- [Asking the Library](#) questions via chat or email

Electronic Reserves

Students can access course readings using their NYU credentials for courses they currently enrolled in at <https://ares.library.nyu.edu/>.

Interlibrary Loan Service

For materials not available to you immediately, you can request scanned copies of a book chapter or journal article through our [Interlibrary Loan \(ILL\) service](#). If you don't know which chapter you need, you can request a Table of Content through ILL.