

1. Course Description and Objectives

The digital, connected, sensor rich world is generating extraordinary amounts and variety of data (“Big Data”). Computational Social Science (CSS) is a research perspective that incorporates new methods and models for harnessing data to study human behavior from the level of neurons to collective behavior. This change in approach has already made breakthroughs possible in understanding human creativity, scientific performance, the sharing economy, human conflict, and online behavior, as well as the impacts of policy on the complex socioeconomic web that composes our world.

The goal of the course is to teach you these new methods and their applications so that you can apply them to your own research. The primary focus of this course is on unstructured data—text and networks. We will cover how to acquire and process data from the web and documents, as well as foundational and advanced methods in text processing and network analytics to mine and analyze data. After taking this course students will be able to use these methods to conduct qualitative and quantitative research with greater breadth and depth.

2. Course materials

There is no required textbook for the course. Academic papers are available at the links or, if behind a paywall, other sites **cough**.

3. Requirements and Evaluation

Attendance

Students need to attend lecture. If you are sick, please schedule office hours with Professor Pah or catch up with notes from a classmate.

Grading

The primary evaluation for this course is based on an independent research project that leverages course methods and concepts. The project is evaluated based on the quality of the writing, the clarity of presentation, and the appropriateness of the research methods and analysis.

In the first half of the semester there will be a handful of homework assignments to help you practice the basic coding mechanics. These assignments are all equally weighted. Grading of these assignments is the same as in “Coding in R” – homeworks are graded pass/fail based upon an assessment of whether you sincerely attempted the homework and answered over half of the questions correctly. This is designed to hold you accountable for the material but not create anxiety about perfection.

You will present an academic paper of your choice during the semester. You will be graded based upon the quality of your presentation and your command of the material.

<i>Assignment Category</i>	<i>Percentage</i>
<i>Homeworks</i>	40

<i>Presentation</i>	10
<i>Project Write-up</i> (Split between proposal and final paper)	50

Academic Integrity

High standards of intellectual and academic integrity will be enforced. GSU guidelines are posted at <https://deanofstudents.gsu.edu/files/2017/09/Academic-Honesty-Policy.pdf> and <https://codeofconduct.gsu.edu/>.

Students **are** allowed to use ChatGPT for their project, but **are not** allowed to submit unedited ChatGPT responses as a part of their final manuscript.

Students **are not** allowed to use ChatGPT in their homework assignments. Students **are** allowed to use StackOverflow and similar resources for homework assignments, but **are not** allowed to copy code from these sources. Students **are** allowed and encouraged to help each other understand programming concepts, errors, and how to approach problems.

Roadmap

Week	Section	Topics
1 – Jan 8	Fundamentals	Introduction, Data Types
2 – Jan 15	Fundamentals	Handling and processing data with dictionaries
3 – Jan 22	Fundamentals	Tabular data management
4 – Jan 29	Data Acquisition	APIs and Web Scraping
5 – Feb 5	Data Acquisition	Web Processing
6 – Feb 12	Data Acquisition	PDF processing
7 – Feb 19	Unstructured Data	Text Processing Fundamentals
8 – Feb 26	Unstructured Data	Networks
9 – Mar 4	Unstructured Data	Sentiment Analysis
10 – Mar 18	Unstructured Data	Social Networks and Dynamic Processes
11 – Mar 25	Structuring Data	Topic Maps and Clustering
12 – Apr 1	Structuring Data	Word2Vec and Neural Networks
13 – Apr 8	Structuring Data	NER and Disambiguation
14 – Apr 15	Structuring Data	Community Detection

Adam Pah

apah@gsu.edu

55 Park Pl NE Rm. 969

**CRJU 9610--Computational Social Science:
Methods and Applications**

In-Person

Book Office Hours at:

www.calendly.com/adampah