**SYA 7933/CCJ 5934: Computational Social Science: Programming and Principles**
University of Florida
Spring 2024

**Meeting time/place:** F, 1:55 PM - 4:55 AM / Turlington 2306

**Instructor:** AJ Alvero, aalvero@ufl.edu

**Office/Hours:** M, 12:15 PM – 1:15 PM / Turlington Hall 3344. If those times don't work, I can meet virtually at other times/days.

**Course website:** https://elearning.ufl.edu (Canvas)

**Overview:**

The world is awash in data. The social sciences have responded by developing "computational" subfields that leverage big data, faster processing, and complex algorithms to answer social questions. Our class is an introductory class into this world that focuses on both social scientific thought and theory of and in response to the big data revolution along with hands-on training of key tools. Students who take this class will therefore become conversant in ongoing conversations on this topic while learning fundamental programming skills in Python and R in order to delve deeper into the methodology behind the movement. This part seminar-part methods class does not assume past coding experience nor engagement with computational sociology literature.

**Course Objectives**

After taking this class, students will be able to:

-- Analyze, critique, and interpret computational sociological scientific literature

-- Effectively communicate the complexities of big data and society

-- Learn fundamental programming skills in Python and R

**Readings:**

All of the reading materials for the class will be made available as PDFs on Canvas. They comprise key papers in computational sociology and sociologically oriented data science. Students are expected to complete the readings and be prepared to discuss them before each class. However, since this class is not a traditional seminar and you

will have programming assignments, I will be mindful of the amount of reading required each week.

For the programming part of the class, there are a litany of free materials to help with coding. I encourage you to look around and find materials that you like. The creator of Python also wrote a free introductory guide to use. The same goes for base R and the Tidyverse. These are just a few resources though and don't begin to scratch the surface of the training materials that are available to users.

**Software:**
We will use Jupyter notebooks in class for coding in Python and RStudio for coding in R. These are not required but they are standard tools in this domain, user friendly, and great for beginners and experts alike. Other coding studio type software (eg. PyCharm, Visual Studio Code) are totally fine to use if you feel confident importing libraries/packages or are willing to learn. Same for command line coding (eg. Vim, Emacs). We will use Python versions higher than 3.0 and the latest version of R.

**General Expectations**

Come on time, attend every class, complete all readings, share in the presentation of specific readings, submit every assignment on time, and participate actively throughout the class sessions. Class discussions and activities will assume that students completed the assigned readings. We will hold coding sessions during class, so you will need to bring whatever device you use to code.

**Assignments**

**Weekly coding assignments:**

Each week, students will have coding assignments to complete before the following class. It is imperative that students complete each assignment on time as the methods will build off of each other. I will write code that will evaluate your code, so please ensure that the final version you submit can be easily downloaded and run. The assignments will build off the lessons/practice from each class.

**Writing assignments**

There will be four writing assignments due. Each writing assignment will correspond to a week of reading; you will get to pick which four weeks to focus on. The goal for the writing assignments is not to merely summarize the readings but to engage deeply with the core social tensions, theories, and arguments laid out in the readings. You will also

be expected to lay out your perspectives and theories in response to the readings. Each submission should briefly conclude with potential research questions and directions in response to the readings and your thinking/theorizing.

The written assignments should be 2-3 pages excluding additional/outside references (no need to include references for the week's reading).

**Student led in-class discussion**

In lieu of a final paper/project, we will instead have student led discussion sections throughout the semester (starting in week 4). Students will partner up, create a powerpoint presentation, and present them in class.

- Division of labor: each partner is responsible for one of the readings and explaining the strengths and weaknesses relative to the other reading. The powerpoint will then be easy to make: pick a template and each partner adds their own text. This will make the powerpoint much less stressful because you won't be as dependent on each other.

The powerpoints will contain four chunks:

1. Overview of each reading
    - Key arguments/theories/ideas/evidence/etc.
2. How they connect and do not connect/disagree
    - Strengths and weaknesses of each reading relative to the other reading.
3. Exemplify the arguments
    - Your research, other research, possible research. Just make it real!
4. Synthesis of the papers: what can be learned from pairing these two specific papers together? How are they greater than the sum of their parts?

After the powerpoint presentation, the partners will lead a group discussion in two parts.

First, each group will lead the class in some kind of discussion activity. I am extremely flexible about what this looks like. It could include things like:

- Small group activities and discussions, surveys, debates. Just keep it respectful!

Then, we will conclude with:

- Open group discussion (co-facilitated by myself and the group)
    - Goal of each group discussion: theory development and construction. Not about which paper is "better"/"worse" but rather about building theoretical explanations from the synthesis. This will shape our understanding as well as generate ideas for computationally driven social science.

**Late policy:** Aside from any approved extensions by me, assignments must be submitted by each date at 11:59PM. Assignments submitted after will be docked 10% for each late day (excluding weekends) up to 50%.

In other words, since each assignment is due on Friday, you will have until the following Friday to submit the assignment for partial credit.

If you ever foresee any issue or obstacle towards submitting your assignment on time, your default move should ***always*** be to contact me.

### Grading

A 93+, A- 90-92, B+ 88-89, B 83-87, B- 80-82, C+ 78-79, C 73-77, C- 70-72, D+ 68-69, D 63-67, D- 60-62, F 59 and lower

### Evaluation

Weekly coding assignments: 50%

Writing assignments: 20%

Student led discussion with partner: 20%

Attendance and participation: 10%

### Attendance

As noted above, attendance and participation is worth 10% of your final grade. Missing class due to illness, emergencies, etc. is permitted **as long as you contact me ahead of time.** Student-athletes competing as part of official university activities are also excused but likewise must coordinate with me ahead of time.

For more information see UF's Attendance Policies.

### Academic Integrity

UF students are bound by The Honor Pledge:

“We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honesty and integrity by abiding by the Honor Code. On all work submitted for credit by students at the University of Florida, the following pledge is either required or implied: "On my honor, I have neither given nor received unauthorized aid in doing this assignment."

The Honor Code specifies specific behaviors which violate this code along with respective sanctions.

### Academic Accommodations

"Students with disabilities requesting accommodations should first register with the Disability Resource Center (352-392-8565) by providing appropriate documentation. Once registered, students will receive an accommodation letter which must be presented to the instructor when requesting accommodation. Students with disabilities should follow this procedure as early as possible in the semester.

#### Student Evaluation and Feedback

Students are expected to provide professional and respectful feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback in a professional and respectful manner is available at https://gatorevals.aa.ufl.edu/students/. Students will be notified when the evaluation period opens, and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via https://ufl.bluera.com/ufl/. Summaries of course evaluation results are available to students at https://gatorevals.aa.ufl.edu/public-results/.

### Course Schedule (readings will be discussed in order presented on syllabus)

| Meeting Date | Overview | Readings |
|---|---|---|
| Jan 12 | Introduction, setting up software, "Hello, world!" | Lazer et al. (2009)<br>Lazer et al. (2020) |
| Jan 19 | Python: Functions, variables, types, containers/collections | Edelmann et al. (2020)<br>Lewis (2015) |
| Jan 26 | Python: Built-in functions, libraries, control flow (if/else, loops), reading in files/data | McFarland et al. (2016)<br>Miconi (2023) |
| Feb 2 | Python: Classes, iterators, generators | Brayne (2017)<br>Housley et al. (2014) |
| Feb 9 | Python: web scraping, APIs | Kitchin (2014)<br>Heiberger & Riebling (2016) |
| Feb 16 | Python: visualization and basic statistical analysis | Savage & Burrows (2007)<br>Bakir et al. (2023) |
| Feb 23 | Python: Regular expressions | McMillan Cottom (2020)<br>Marres & Weltevrede (2013) |

| | | |
|---|---|---|
| March 1 | R: Comparing basics of Python and R | boyd & Crawford (2012)<br>Kozlwoski et al. (2019) |
| March 8 | R: RStudio, markdown | Hofman et al. (2021)<br>Törnberg & Uitermark (2021) |
| March 22 | R: Tidyverse | Halford & Savage (2017)<br>Moats & Seaver (2019) |
| March 29 | R: CRAN, libraries | Joyce et al. (2021)<br>Jensen et al. (2021) |
| April 5 | R: Text analysis | Tinati et al. (2014)<br>Playford et al. (2016) |
| April 12 | R: Statistical analysis | Bail (2014)<br>Törnberg & Törnberg (2018) |
| April 19 | R: Tidyverse + statistical analysis | Vicari & Kirby (2023)<br>Bermann & Hirschman (2018) |

*I reserve the right to make small, reasonable alterations to the syllabus as we progress through the semester. Making changes to a course in this way is a common practice and will be communicated to students in a timely manner.*