POLSCI 514: The Use of Social Science Computer Programs

Mike Thompson-Brusstar

Winter, 2022

E-mail: `mrthomp@umich.edu`
Office Hours: Tuesday 12-2PM link   Class Hours: Friday 9-11:00am

Course Description

This course aims to complete the taught component of your computational education, with an eye towards one goal, especially: ensuring you have the *literacy* necessary for applying extant and future computational tools for your research. This will mean learning (including by practice) more about how your computer is structured, tools used by developers of statistical and research software, and how to manage research projects and files. It will also entail high-level surveys of some methods that might be appropriate for your developing research agendas.

Required Materials

- There are no required textbooks for this course.

- The course materials, including optional readings, will be made available on Canvas and via Slack.

Course Objectives

By the end of this course, successful students will:

1. be comfortable with some basic command line tools (`cd`,`ls`,`mv`,`mkdir`,`rm`,`touch`,`rsync`,etc.).

2. demonstrate facility with basic data manipulation (tidyverse and/or base R).

3. be able to organize and author analysis files that will replicate and run when used by others.

4. be able to *read, locate, and understand* documentation for statistical software and tools (official docs and online advice).

5. be familiar with a broad suite of computational tools for applied researchers, and their connections to research design (to be determined).

Course Structure

Class Structure

The plan is to divide time between:

- computational tasks required for POLSCI 699

- a set of 'necessary software skills'

- a survey of other computational methods and approaches

Some review will also be conducted as necessary, based on the self-evaluation conducted at the beginning of the semester.

Assessments

We will focus on a *project-based* course approach, in which students will largely complete "project milestones" instead of problem-sets. Completing applied tasks will be better for cementing what we learn in lecture and will also potentially benefit your other applied or methodological coursework, while more closely matching the research-project based workflows you will likely employ elsewhere.

Lecture

The two-hours of course time will be divided between *lecture* and *lab*, roughly one hour each. In 'lecture,' the core tools and concepts of the week will be presented. The script (likely in R) that I work with will be provided before class so that students can comment and code along during lecture.

Lab

The second hour of instructional time will be spent in a series of guided exercises that require the application of the concepts and tools covered in class. I will roam through the class to answer questions while you complete these exercises, which are designed to be completed before class time ends or with bounded time outside of class. When possible, these exercises will be connected to the projects of the students (e.g. "Identify a static website whose content is relevant to your project. Annotate your strategy for scraping that content from the website responsibly. Do so, and attach the data in rectangular format as a .csv").

Evaluation

Assignments will be graded on a 'check-plus, check, check-minus, zero' system, corresponding to 100%, 85%, 50%, 0% respectively.

Regular assignments will be accepted: a 'check' level will be lost for each 3 day period (72 hours) that have *passed in their entirety* since the due date. Note that this means that all assignments are accepted without penalty for 3 days after their due date. Note also that 'regular assignments' excludes *only* the final project milestone, so that grades can be turned in on time.

Attendance of lab and lecture is *required* and forms a component of your participation grade. If you must miss class, please let me know as far in advance as possible and I will excuse you. Students who miss class should attend the subsequent office hours or make an appointment so we can discuss the content that you missed.

Grading Breakdown

- **60%** of your grade will be determined by four (4) project milestones. (15% each).

- **20%** of your grade will be determined by two (2) problem-sets centered on content unsuited for project milestones (10% each).

- **20%** of your grade will be determined by in-class participation.

Problem set Guidelines

The problem sets will be placed on Canvas. Problem set 1 is due after class on week 2. Problem set 2 is due after class on week 4.

Project Milestone Assignment Guidelines

The guidelines for the project will be several-fold: most importantly, they must demonstrate one or several of the tools or approaches covered since the last assignment (after *consultation with the instructor,* the *depth of engagement* with covered tools can substitute for the *number of tools* used). As projects develop, other computational approaches not covered in the course may (after consultation with Mike) be substituted for those covered in class if those covered are unsuitable for some reason. The purpose of the later milestones is not the mastery of the 'topics' sections but to practice locating, understanding, and applying computational techniques.

Projects also should be chosen in consultation with Mike to ensure they present sufficient opportunities to demonstrate understanding of course material.[1] The best projects will demonstrate the potential for use in multiple project milestones, but 'one-off' projects are also likely fine.

The three pillars for evaluation of project milestones are:

1. input

2. output

3. documentation

That means that project milestones should have: an input, output, and an explanation for why you did what you did and how you did it. That's probably three or four files.

Example: Say you are extracting text from a pdf of individual biographies, processing that information (e.g. extracting birthdays), and storing the results in a data-frame. That milestone would have: the original pdf file, a file that contains the code that parses, extracts, and organizes the information and writes the resulting dataframe to file, and the final output file, at least. The code could be well commented to explain the decision processes and how things work (that's

---

[1]Note that in general, this is a terrible criterion for selecting a project. As elsewhere, research questions and research designs should guide your efforts in computational research.

three files) or you could just attach a quick txt file that has that content in it (that's four files). I don't care as long as the three pillars are well represented, the input produces the output, and I can follow your logic as you code.

The due dates of the project milestones are as follows:

1. First discussion: by Week 7 (Feb 18)

2. Milestone 1: Week 9 (Mar 4)

3. Milestone 2: Week 11 (Mar 18)

4. Milestone 3: Week 13 (April 1)

5. Milestone 4 (presentation): Week 15 (April 15)

Topics to be covered:

Computing essentials

- Command line tools

- Data wrangling: reshaping, recoding, grouping, summarising

- Data visualization: `plot` and `ggplot`, diagnostic summaries

- Filepaths and project organization

699 Related topics

- Necessary review

- Optimization

- Monte Carlo simulation

- Bootstrap

- linear regression (lm(), glm() etc)

- hypothesis testing

Topics in computational analysis (tentative!)

- Web scraping

- Parallel computing (cluster and locally)

- Text analysis

- Network analysis

week	date	514 topic	comment
2	1/14	Command line	`ls,cd,mkdir,touch,head,rm,`
3	1/21	Data manipulation	`tidyverse`, recoding
4	1/27	Data visualization	`ggplot()`
5	2/4	Advanced loops, distributions	`if(),for(),r/q/d/norm`
6	2/11	Regression and hypothesis testing	`lm(), glm()`
7	2/18	Web scraping	`wget` and `rvest`
8	2/25	Project organization	`git,here()` etc.
9	3/4	*Spring Break*	*No class*
10	3/11	Text analysis	`stm, quanteda`
11	3/18	Parallel computing	greatlakes, CLI (reprise)
12	3/25	Text analysis 2	`quanteda`
13	4/1	Web scraping 2	Scraping dynamic sites
14	4/8	Network analysis	???
15	4/15	Presentations	milestones

Course Policies

- **Student Sexual Misconduct Policy:** Title IX prohibits sex discrimination to include sexual misconduct: harassment, domestic and dating violence, sexual assault, and stalking. If you or someone you know has been harassed or assaulted, you can receive confidential support and academic advocacy at the Sexual Assault Prevention and Awareness Center (SAPAC). SAPAC can be contacted on their 24-hour crisis line, 734–936–3333 and online at sapac.umich.edu. Alleged violations can be reported non-confidentially to the Office for Institutional Equity (OIE) at institutional.equity@umich.edu. Reports to law enforcement can be made to University of Michigan Police Department at 734–763–3434.[2]

- **Accommodations for Students with Disabilities:** If you think you need an accommodation for a disability, please let me know at your earliest convenience. Some aspects of this course, the assignments, the in-class activities, and the way the course is usually taught may be modified to facilitate your participation and progress. As soon as you make me aware of your needs, we can work with the Services for Students with Disabilities (SSD) office to help us determine appropriate academic accommodations. SSD (734–763–3000; http://ssd.umich.edu) typically recommends accommodations through a Verified Individualized Services and Accommodations (VISA) form. Any information you provide is private and confidential and will be treated as such.[3]

- **Religious-Academic Conflicts:** While the university does not observe religious holidays, it is the policy of the University of Michigan to make every reasonable effort to allow members of the university community to observe their religious holidays without academic penalty. Absence from classes or examinations for religious reasons does not relieve students from responsibility for any part of the course work required during the period ob absence. Students who expect to miss classes as a consequence of their religious observance shall be provided with a reasonable alternative opportunity to make-up missed academic work. It

---

[2]This statement is taken from: https://sapac.umich.edu/article/faculty-resources-sample-syllabus-language.
[3]This statement is taken from: https://ssd.umich.edu/article/syllabus-statement.

514 Winter 2022

is the obligation of students to provide faculty with reasonable notice of the dates on which they will be absent. When the absence coincides with an exam or other assignment due date, the options to make up that missed work may be limited and will be determined by the instructor within the boundaries of the respective class.[4]

- **Academic Misconduct:** The University of Michigan community functions best when its members treat one another with honesty, fairness, respect, and trust. The college promotes the assumption of personal responsibility and integrity, and prohibits all forms of academic dishonesty and misconduct. All cases of academic misconduct will be referred to the Office of the Assistant Dean for Undergraduate Education. Being found responsible for academic misconduct will usually result in a grade sanction, in addition to any sanction from the college. For more information, including examples of behaviors that are considered academic misconduct and potential sanctions, please see `https://lsa.umich.edu/lsa/academics/academic-integrity.html`.[5]

- **Student Mental Health and Well-being:** The University of Michigan is committed to advancing the mental health and well-being of its students. If you or someone you know is feeling overwhelmed, depressed, and/or in need of support, services are available. For help, contact *Counseling and Psychological Services (CAPS)* at (734) 764-8312 and `https://caps.umich.edu/` during and after hours, on weekends and holidays, or through its counselors physically located in schools on both North and Central Campus. You may also consult *University Health Service (UHS)* at (734) 764-8320 and `https://www.uhs.umich.edu/mentalhealthsvcs`, or for alcohol or drug concerns, see `https://www.uhs.umich.edu/aodresources`. For a listing of other mental health resources available on and off campus, visit: `http://umich.edu/~health`.[6]

---

[4]This statement is taken from: *Handbook for Faculty and Instructional Staff 2018*, p. 17.
[5]This statement is taken from: *Handbook for Faculty and Instructional Staff 2018*, p. 16.
[6]This statement is taken from: *Handbook for Faculty and Instructional Staff 2018*, p. 16.

6/6