

Syllabus

GEO511 Spatial Data Science

Course Description

This course introduces students to the fundamentals of spatial data science. The first part of the course introduces students to a high-level programming language (currently R). The second part covers methods to incorporate spatial data into data science workflows. The third part addresses the generation of dynamic, reproducible research output including figures, maps, manuscripts, and websites. The course includes a project for students to conduct spatial analysis related to their research. Familiarity with basic GIS concepts (raster, vector, geographic projection, etc.) will be assumed, but no prior coding experience is required.

Logistics

Professor Adam M. Wilson (wilsonlab.io)

Office Hours: Thursdays 9:00-11:00am

Schedule

Tuesdays/Thursdays 2-3:20pm virtually via Zoom.

Course Structure

The course will focus on programming in the R language. Typical class sessions will consist of a short (<30 minute) lecture followed by interactive exercises and activities. All class activities will use RStudio.

Course Communication (UBLearns & Slack)

Course announcements and other materials will be distributed through UBLearns and/or our Slack Channel. Please check the sites regularly (or enable notifications).

Computer requirements

During the course we will complete class exercises on your personal laptop (under any Mac, Linux, or Windows). If you do not have access to a laptop, please let the professor know as soon as possible.

Email Policy

During the week, I will attempt to respond to emails within 48 hours of receiving them (not including weekends). Do not expect an immediate response (please plan accordingly). For example, do not send an email with a question about an assignment the same day that the assignment is due. If you send an email over the weekend, do not expect any response until Monday or Tuesday.

Student Learning Outcomes

Successful completion of this course will enable the student to:

1. convert data from varied formats/structures to desired format for analysis and visualization
2. clean, transform, and merge data attributes/variables appropriately
3. effectively display and communicate meaning from spatial, temporal, and textual data
4. use current analysis, presentation, and collaboration tools in the spatial data science field.

These learning outcomes are related to those expected of students completing the Geography program.

Course Requirements

The major course components are as follows:

Data Camp courses

Several mini-courses will be assigned via DataCamp throughout the semester. These assignments will be graded as pass/fail (pass if you finish the course, fail if you don't). These typically take about 3-5 hours per course, but some people report taking much longer (up to 8 hours). You can use the 'hints' provided to complete the exercises (but try not to!). See full DataCamp Description for more details.

Tasks

The course includes many tasks that are performed both in and out of class (see the tasklist). You will 'commit' evidence of completing these tasks to your course repository on GitHub.

Case Studies

Most weeks we will work on a 'case study' project alone or in small groups. Typically these are open-ended mini-projects in which you use your new skills to perform a task related to spatial data science.

Team Leadership

Most weeks we will spend 15-30 minutes in a team meeting where you will discuss the previous case study in small groups. Each group will have a 'leader' who facilitates the meeting and shares his/her/their solution to the case study. To successfully perform as a leader, you must:

- 1) complete the class tasks / case studies before class starts
- 2) be prepared to lead the discussion of the case study

See here for more information about case study leadership.

Resource presentation

Each student will have the opportunity to introduce a R-related resource in a 5 minute presentation during class. Most students choose to describe a R package that does something they are interested in, but you could also introduce us to other kinds of resources (useful online forums, web resources, online textbooks, etc). See [here](#) for more information about the resource presentation.

Final Project

The final project will consist of a poster-length reproducible analysis published in html format. This project can be related to the student's own research or a separate topic.

Weekly Rhythm

1. **Preparation**
 - a. Read Assigned Material
 - b. Work on class tasks and DataCamp assignments
 - c. Submit questions for class discussion on Slack
2. **Class Time Tuesday**
 - a. Meet with your Case Study Teams
3. **Continue Preparation**
 - a. Finish case studies and prepare to present
 - b. Work on class tasks and DataCamp assignments
 - c. Submit questions for class discussion on Slack
4. **Class Time Thursday**
 - a. Updates & Questions from reading and daily class tasks [~10 minutes]
 - b. Student Resource Presentation(s) [20 minutes]
 - c. Case Study Presentation [30 minutes]
 - One group selected to share solution
 - Other groups share other approaches / solutions
 - General discussion about methods
 - d. Case Study Introduction (for following week) [20 minutes]
5. **Rinse and Repeat**

Grading

Individual tasks in the class will not be traditionally graded. If your work meets the specified criteria you will get full credit and only then (there is no partial credit on tasks).

In a specifications-grading system all tasks are evaluated on a high-standards pass/fail basis using checklists of task requirements and expectations. Letter grades are earned by passing marks on a set of tasks. This system provides for a variety of choice and is closer to how learning, and work, is done in the real world. It will be easy to tell if work is complete, done in good faith, and consistent with the requirements. The definitive word is “complete”. Starting them or getting them almost done is not completing.

Grade

Class Tasks

Case Studies
 Team Leader
 Data Camp
 Resource Presentation
 Semester Project
 A
 12
 11
 3
 10
 yes
 yes
 A-
 11
 10
 3
 10
 yes
 yes
 B+
 10
 9
 2
 9
 yes
 yes
 B
 9
 8
 2
 9
 no
 no
 B-
 8
 7

1
8
no
no
C
7
6
0
7
no
no
C-
6
5
0
6
no
no
D
5
4
0
5
no
no

Coding Challenge

On the last day of class you will be asked to complete a coding challenge. You will not be able to work with others. The coding challenge will use the techniques and programming developed from the course and is meant to demonstrate a foundational understanding of the course material.

1. **You must pass the challenge to keep grade earned from your completed tasks.**
2. Failure to pass the challenge will lower your grade 1-2 steps.

There will be no final exam.

Semester Deliverables

- GitHub Profile
 - Course Repository
 - * Tasks

- * Case Studies
- Grade Request
 - Before the last day of classes, you will submit a cover letter (<1 page) stating:
 - * the key concepts and techniques that you learned during the course (~500-1,000 words)
 - * a semester task form that records your completed tasks during the semester
 - * a grade request based your completion of course tasks (using the table above)
- Final project website

Textbook

We will read parts of *R for Data Science* and *Geocomputation with R* which are both available online. All additional materials will be available through the course website.

On-time Completion

There is not a strict definition of *on-time* in this course. In general, on-time means that you have come to class with the reading and tasks complete so that you can actively participate in the conversation. You have to define **prepared** for class. You should note that the workload in this course does not allow you to fall behind. If you blow off a week, it will be challenging to catch back up.

Attendance Policy

This class will include ample opportunities for in-class discussion and you are expected to attend every class session unless you have a valid excuse (as defined by the University at Buffalo's class attendance policy:

Students may be justifiably absent from classes due to religious observances, illness documented by a physician or other appropriate health care professional, conflicts with university-sanctioned activities documented by an appropriate university administrator, public emergencies, and documented personal or family emergencies. The student is responsible for notifying the instructor in writing with as much advance notice as possible.

If you miss a class session, you are still responsible for completing the class content/assignments. Please consult with a classmate to see if there was any important information not included in the online materials.

See the University website for cancellations/delays due to weather or other unforeseen events (<http://emergency.buffalo.edu/>).

Academic Integrity

Academic integrity is critical to the learning process. It is your responsibility as a student to complete your work in an honest fashion, upholding the expectations your individual instructors have for you in this regard. The ultimate goal is to ensure that you learn the content in your courses in accordance with UB's academic integrity principles, regardless of whether instruction is in-person or remote. Thank you for upholding your own personal integrity and ensuring UB's tradition of academic excellence. Examples of academic dishonesty include: submitting work from another course, plagiarism, cheating, falsification, misrepresentation, and usage of confidential documents.

Writing computer code often involves use of existing code chunks (e.g. copying an example from the documentation) which complicates identification/definition of academic dishonesty. The primary goal of the course is to learn how to **program and think** as a data scientist concerning data wrangling and visualization. I want you to use your time as efficiently as possible to meet this goal. With this goal in mind here are some guiding principles.

- You can see others code in the classroom repositories on Github. Think of your classmate's code as a resource but not a crutch.
 - If you look at other's code to get help solving a problem, you need help with then that is ok. Just put **# Got help for the next three lines of code from Jason's Task 12 script** where you copied code.
 - If you copy and paste other's code to complete a task and can't recreate the script on your own or understand what it is doing **you are cheating**.

If there is reason to believe that submitted code was simply copied from elsewhere, the student will be asked to verbally (and specifically) explain the code used in the analysis to ensure comprehension.

If a student is suspected of academic dishonesty, then a three-step consultative resolution will be employed. First, the instructor will notify the student of the incident and arrange a meeting. Second, the instructor will orally inform the student of the sanction, which could include: warning, revision, reduction in grade, or failure of course. Third, the instructor will provide the student with a written copy of the decision. See the university policy for more information (<https://catalog.buffalo.edu/policies/integrity.html>). Please review it and ask if you have any questions.

Accessibility Resources

If you have any disability which requires reasonable accommodations to enable you to participate in this course, please contact the Office of Accessibility Resources in 60 Capen Hall, 716-645-2608 and also the instructor of this course during the first week of class. The office will provide you with information and review appropriate arrangements for reasonable accommodations, which can be found on the web at: <http://www.buffalo.edu/studentlife/who-we-are/departments/accessibility.html>.

Course Content

Course content is designed to be flexible to accommodate student interest and abilities. The order and timing of course topics may change as the semester progresses. See the course schedule on the website for detailed course content.