# Environment 175 - Syllabus

Programming with Big Environmental Datasets

## Contents



**Last Updated: April 4, 2025**

# Instructors

**Instructor:** Gleb Satyukov
**E-mail:** gleb@ucla.edu

**Teaching Assistant:** Kaitlynn Sandstrom-Mistry
**E-mail:** ksandstrom@ucla.edu

# Office Hours

As this is a very practical and hands-on type of course it is important to stay up-to-date on all of the class materials. However, **it is up to you to reach out** if you think that you are falling behind and/or need additional help. We are going to provide you all with continuous support throughout the class during our office hours, again - please do not hesitate to reach out if you have any questions or concerns.

## Instructor

Gleb will host two separate office hours throughout the week:

**Wednesdays** between **5pm** and **6pm** on Zoom
**Fridays** between **3pm** and **4pm** on Zoom

Gleb's personal Zoom link is: https://ucla.zoom.us/j/6935808910

Please do not hesitate to reach out for help if you are stuck with anything in this class, even if it is outside the office hours. We do not want you to fall behind! You can message me on Slack in the #tech-support channel or send me an email (gleb@ucla.edu) in order to set up a quick call if you need urgent help.

## Teaching Assistant

Kaitlynn will host two separate office hours throughout the week:

**Mondays** between **12pm (noon)** and **1pm** on Zoom
**Wednesdays** between **11am** and **12pm (noon)** on Zoom

Kaitlynn's personal Zoom link is: https://ucla.zoom.us/j/8321830416

# General Information

**Time:** TR 3:30pm - 4:45pm
**Room:** Fowler Museum at UCLA Room A139

**First day of classes is Tuesday April 1st, 2025**

| Type | Time | Day | Location |
|---|---|---|---|
| Lecture | 3:30pm - 4:45pm | Tuesday | Fowler Museum A139 |
| Lecture | 3:30pm - 4:45pm | Thursday | Fowler Museum A139 |

---

Link to the class schedule: Environ 175 Schedule of Classes

## Course Pages

**Course Slides**: https://environ-175.com/slides/<week number 1 through 10>
**Canvas:** https://bruinlearn.ucla.edu/courses/204902
**Slack:** https://environ-175.slack.com

Please use this link to join our Slack environment:
https://join.slack.com/t/environ-175/shared_invite/zt-3316phuc3-xJWLftbvgk6LB0lj1rq15A

# Class Requirements

## Software

Students should have access to a computer outside of class time with the following software installed:

- R and R Studio
- Slack
- Zoom

*Download and install R first and R Studio second, in that order.*

Zoom and Slack will be used predominantly for any communication outside of class, alongside our Canvas module for class-wide announcements.

## Hardware

We **strongly recommend** that you bring your laptop with you to class and follow along as we go through the examples together. Please make sure that your laptop is **fully charged** before class as access to power outlets in the classroom can be limited.

If you don't have a laptop that you can use on hand, you can check one out at the library for short periods of time. Please refer to the CLICC equipment lending program.

## Literature

We will almost exclusively be using our own examples and tutorials on Canvas and as such do not require any additional literature. However, if having a textbook gives you comfort, check out this free online version of R for Data Science, 2nd Edition by Hadley Wickham and Garrett Grolemund.

List of other suggested textbooks that can be used alongside this class:

- R for Data Science, 2nd Edition by Hadley Wickham and Garrett Grolemund

- R Cookbook, 2nd Edition by James Long and Paul Teetor

### Other Courses

Recommended prerequisites for this course are Statistics 12 or 13, or Life Science 40.

## Course Description

The overarching goal is for you to acquire the skills necessary to conduct environmental research with big data. There is no one definition of what "big data" is, one way you can think about "big data" is as data that is too large to be analyzed with software like Excel or Google Sheets or weaker laptops. Big data is exciting to people since larger sample sizes can provide more credibility to the findings in terms of both generalized application and statistical precision.

We will learn how to program using the statistical software R. Many researchers use R because it is user friendly, powerful, and, importantly, open-source and free. R is especially popular among ecologists, public health experts, and political scientists. These days R is starting to become more popular in commercial sectors, too. Learning how to program in R will have overlapping benefits with other programming languages because the basic structure of R code is similar. Throughout the class, we will compare and contrast R with Google Sheets so you can better appreciate R's value.

This course is broken up into several modules built around meaningful research questions. For example, we'll examine how the Acid Rain Program impacted sulfur dioxide emissions at coal-fired power plants in the United States. Across the modules, we'll use real-world data such as well-being surveys, income and poverty data, pollution data, fertility data, and satellite images. We will focus on producing visual results, like graphs and maps. We will not be spending too much time on statistical tests, e.g. T stats, or regressions. Visual products deserve our full attention.

## Course Objectives

By the end of this course, you will be able to:

1. Write your own R scripts
2. Construct convincing visualizations (e.g. scatter plots, maps, etc)
3. Import various types of data (e.g. csv, txt, tif, shp)
4. Calculate group-level averages
5. Quantify changes in environmental factors over time
6. Match populations to environmental factors
7. Compile separate data files into one usable data object
8. Reshape wide data to long format
9. Automate commands using loops
10. Aggregate pixel values by polygon boundaries
11. Navigate your file system and manage files in R
12. Run any R scripts from the command line/terminal app

# Course Schedule

**Table 1:** Weekly outline *(this is subject to change)*

| Week | Date | Tuesday Class | Date | Thursday Class |
|---|---|---|---|---|
| **Week 1** | Apr-01 | Introduction | Apr-03 | R Basics 1: Import |
| **Week 2** | Apr-08 | R Basics 2: Collapse | Apr-10 | R Basics 2: Collapse |
| **Week 3** | Apr-15 | R Basics 3: Generate vars | Apr-17 | R Basics 4: Multi-var plots |
| **Week 4** | Apr-22 | Escape room 1 (Zoom) | Apr-24 | Project 1 (no class) |
| **Week 5** | Apr-29 | R Advanced 1: Join | May-01 | R Advanced 2: Append |
| **Week 6** | May-06 | R Advanced 3: Reshape | May-08 | R Expert 1: Loops |
| **Week 7** | May-13 | Escape room 2 (Zoom) | May-15 | Project 2 (no class) |
| **Week 8** | May-20 | R Spatial 1: Import | May-22 | R Spatial 2: Choropleth |
| **Week 9** | Mar-27 | R Spatial 3: CSV points | Mar-29 | R Spatial 4: Raster math |
| **Week 10** | Jun-03 | R Spatial 5: Zonal avg | Jun-05 | R Spatial 6: Mosaic |
| **Finals** | Jun-09 | Final project released | Jun-11 | Final project due |

## Class Assignments

All of the assignments will need to be turned in as valid R code in Canvas.

## Attendance

Attendance will be registered by taking in-class surveys and answering questions using a smartphone or laptop. For this, we will be using Slido app, or a similar app.

## Final Project

The Final is a take-home project and must be submitted by **Wednesday June 11 at 5pm PT**. The Final is intended to take no more than 2 hours for students who have mastered all of the materials in our class.

We will release the Final project on **Monday June 9 by 5pm PT** so you have more than 48 hours to work on it. Note that our originally scheduled Final date and time is Monday June 9th at 6:30pm - 9:30pm, but we won't be needing to use that date for our final.

A link to the official class page on the UCLA website is here: Link

# Course Evaluation

There are a handful of graded components to this course, with the assignments revolving around different modules:

- Ten days of attendance x 1 points each = 10 points
- Ten highest scores on the short data assignments x 4 points each = 40 points
- Two remote "escape rooms" x 5 points each = 10 points
- Two midterm projects x 10 points each = 20 points
- One final project x 20 points = 20 points

## Grading Scheme

Grades will range from A through F. We do not award A+ grades.
There are a total of 100 possible points. Here are the exact cutoffs for each grade.

| **Points** | **Grade** |
|---|---|
| 93 - 100 | = A |
| 90 - 92 | = A- |
| 85 - 89 | = B+ |
| 72 - 84 | = B |
| 65 - 71 | = B- |
| 60 - 64 | = C+ |
| 55 - 59 | = C |
| 50 - 54 | = C- |
| 45 - 49 | = D+ |
| 40 - 44 | = D |
| 35 - 39 | = D- |
| < 35 | = F |

## More on Grading

As you can see, this is not your typical point system. This point system is set up so you could earn a B before you even take the final. I'm hopeful this will remove some of the stress of finals week, while still motivating you to work and learn. I do not grade on a curve. I don't want to create a competitive learning environment.

My goal is to encourage you to work at a continuous pace, which I find is the best for learning how to code. With this in mind, I expect you to come to at least 10 days of in-person classes. There will be in-class survey questions to measure attendance, with students needing to take at least 75% of the survey questions to count as attending. There will only be 14 days of in-person classes where I record attendance, so you can skip four days and still get full credit.

There will be a short data assignment after most every class for a total of 14 short data assignments. But, we will only take the 10 highest scores on these data assignments. So, this means you can skip four data assignments and still score perfectly.

The "escape room" is a multi-task in-class exercise where you need to show your work to the professor and TA to move from one task to the next. Students completing the exercise by end of day will receive full credit. Students needing to make up the escape room will have one opportunity to do so without penalty.
On the data assignments, we will give you full credit if you produce the correct output, follow best practices, and turn in the assignment on time. Minor errors will be penalized one point. If you make any substantive error (including failing to follow instructions, best practices, or include extraneous code), then your assignment will be penalized multiple points. There is a 20% late penalty (0.8 points) per hour. These assignments should be relatively straightforward since the template for the code will be provided to you.

The two midterm projects and final project will be graded stringently. Your grade is a function of the appropriateness of the code regardless of whether you have the correct output or not. I penalize code that is superfluous because it shows you do not understand the material. I penalize code that is messy and poorly organized because part of being a good coder is producing code that others can easily use. You can earn partial credit even if you can't complete the projects or final, so you should turn in whatever work you have done before the submission window ends.

There is a 20% late penalty (2 points) for every hour the midterm projects are late, with five hours late being no credit since there are only 10 points possible to begin with. Projects are due at 5pm on Fridays, so that means anything later than 10pm will receive a zero. You will have a little over three days to do the projects and the dates are known well in advance to help with scheduling conflicts.

I will automatically use your final score to replace one midterm project score, should you score unusually low on that midterm project for whatever reason. For example, if you score 5/10 (50%) on the second midterm project, but 18/20 (90%) on the Final, I will give you a 9/10 on that midterm project. This can only be used for one midterm project score.

The final project will be due Wednesday of Finals Week at 5pm. Each hour the final project is turned in after the deadline will result in a 20% penalty (4 points). Given the final is worth 20 points, final projects turned in more than 5 hours late (10pm) will receive automatic zeros. For the final, you will have approximately 48 hours to complete them from the time you have access to the material to when it is due. There is no makeup Final Exam.

Grade modifications for missed assignments are given at the professor's discretion. I understand that life is often complicated by physical illness, mental health challenges, technological barriers, and economic problems. I don't want to penalize unpredictable misfortunes. Reasonable excuses related to missed assignments will result in your final score percentage being used in place of the missing item.

You may not collaborate on the final project. But, you may collaborate on the assignments and other projects. You will still need to turn in your own work with personalized code.

Students taking the class Pass/No Pass must earn a C in total points.

There is no extra credit work in this course.

Regrades: Any request for regrade on a project or the Final will result in all your work being regraded. This is to ensure that I did not make any errors in your favor that would cancel out any requested point changes. Also, this helps encourage only serious regrade requests.

You are free to use ChatGPT or other AI Tools for coding help. But, I don't recommend copying and pasting the code without careful thought. ChatGPT might give you code that doesn't follow best practices or, worse yet, gives you the incorrect code.

# How to Succeed in This Course

## Growth Mindset

Having a "growth mindset" means viewing mistakes and challenges as opportunities for growth, and not an indication of your innate ability.

## Consistency

You should expect to work at a constant pace throughout the semester. There are routine assignments for each class period so I can check in on your progress in real time.

## Communication

The material builds on itself, so communicate any difficulties you are having to the professor so that we can avoid you falling too far behind.

## Collaboration

Individual work is important since you will learn by making mistakes. But, working with teammates is fun and can help you overcome hard-to-spot mistakes. This is crucial when it comes to coding, since even small typos can cause programs to crash. We encourage you to find someone you can collaborate with through out the course.

## Additional Resources

YouTube, Codecademy, DataCamp, and Data Carpentry all have free R Courses that may have additional material.

Though we will touch and maybe even use some of these technologies in our class they are consdered to be out of scope. For those interested in continuing to develop their data-computing skills, we recommend learning about the following programs and platforms at some point in the future:

- GitHub - For code revision history and more
- Overleaf - For pretty PDF files and such
- SQL - For relational databases
- MongoDB - For non-relational databases
- Python - For everything else

# Course Policies

## Academic Integrity

Academic integrity and good student behavior are a fundamental expectation of all students in this course. It enhances the quality of education, and is critical to the learning process. Be cognizant of the fact that your behavior (good or bad) impacts others in the class. Respect and kindness for others enhances the learning environment for everyone. Please familiarize yourself with the discussion of plagiarism and student copyright here: https://registrar.ucla.edu/registration-classes/enrollment-policies/class-policies/plagiarism-and-student-copyright

Plagiarism – presenting someone else's ideas as your own, either verbatim or recast in your own words (i.e. without giving appropriate attribution or citation) – is a serious academic offense with serious consequences. Instances of plagiarism or cheating on assignments will result in an automatic grade of 0 for the assignment in addition to a letter grade deduction from your final course grade, and that student may be referred to the Office of the Dean of Students. However, sharing separate sections or parts of code with each other is allowed and can be used with permission from the original author.

## UCLA Policies

It is your responsibility to be familiar with UCLA's Code of Conduct here:
https://deanofstudents.ucla.edu/student-conduct-code

UCLA's Office for Equity, Diversity, and Inclusion provides resources, events, and information about current initiatives at UCLA to support equality for all members of the UCLA community. Please talk with the professor or TA if you experience anything in this course that does not support an inclusive environment, and you can also report any incidents you may witness or experience on campus to the Office of Equity, Diversity, and Inclusion on their website here: https://equity.ucla.edu/

## Covid-19 Instructions

Ensuring a safer campus depends on each of us following the latest UCLA health and safety guidelines. While the campus policies may need to be modified to address changing local, state, and national orders and guidance, the most current information is available at https://covid-19.ucla.edu/information-for-students/.

We will support any students who are staying away from campus if symptoms present themselves. Please reach out to me (gleb@ucla.edu) if you need to stay home, and we will make alternative arrangements for you to make up for any missed classwork. You will be given automatic extensions on all of the assignments as necessary.

Be advised that refusing to comply with the current campus directions and guidelines pertaining to Covid-19 mitigation may result in dismissal from the classroom and a direct referral to the Office of Student Conduct. Please refer to the official UCLA guidelines regarding Covid-19 if you have any questions or concerns.

Thank you for protecting your fellow Bruins!