

ENV 872L Environmental Data Analytics

Spring 2021

Synchronous Sessions: Section 1: Wednesday 10:15-11:30 am
 Section 2: Thursday 8:30-9:45 am
 Section 3: Monday 1:45-3:30 pm

Instructors:	Luana Lima	John Fay
	Email: luana.lima@duke.edu	Email: john.fay@duke.edu
	Office hours: TBA	Office hours: TBA

TAs:	Sarah Roberts	Abhishek Jain
	Email: sarah.m.roberts@duke.edu	Email: abhisheksanjay.jain@duke.edu
	Office hours: TBA	Office hours: TBA

Course Texts and Readings

R for Data Science

Garrett Grolemund and Hadley Wickham

Open source online textbook: <https://r4ds.had.co.nz/>

Hard copy available on Amazon

Additional texts and articles available on Sakai

Course Description

Given the growing focus of environmental disciplines on quantitative approaches, students entering the environmental workplace have a need to face new challenges related to data. Data analytics encompasses not only statistics and data visualization but also puts those practices in context of the acquisition, exploration, processing, and reporting of data. In this course, we work through contemporary data analyses while developing skills to integrate software applications, manage data, and effectively report results. Students will develop reproducible workflows to analyze real environmental datasets from start to finish. The setting of this course is a hands-on lab, where students will work through a series of lessons, assignments, and a final course project.

Course Objectives

1. Gain proficiency in the language and application of R software
2. Synthesize information from datasets, working from start to finish in the “data pipeline”
3. Develop skills to identify and apply appropriate statistical and graphical approaches for environmental datasets, incorporating the guidelines of experimental design and interpretation of output
4. Integrate multiple technological applications involved in contemporary data analysis, workflow, and management

Expectations

This course has a fairly steep learning curve, meaning that most of us will encounter technical and conceptual difficulties on our way to developing proficiency. This is not meant to discourage you but to help you be prepared to face challenges in the best way possible. With that in mind, the expectations for this course include:

- If you are using your own computer, software installations and upgrades must be completed before class. Most applications require updated software systems to operate correctly, and we do not have time to dedicate to this in class.
- If you run into an issue, please attempt to diagnose the issue and do some troubleshooting on your own before asking an instructor. We are not expecting you to know how to fix all problems on your own, but taking time to understand the issue will (1) help us to fix the issue with you faster and (2) help you develop independence.
- If you miss class, contact a classmate for an update on the lesson and work through the lesson on your own. We will gladly answer any questions or help to fix any troubleshooting issues with you afterward.
- Data analytics is an extremely collaborative field. We will replicate this environment in our classroom, where you will actively participate and cooperate with one another. We will do our best to cultivate this collaborative environment.

Grading

Assignments 75%

Each major unit of the course will be accompanied by an assignment, which should be completed individually (with help from classmates if needed). Each assignment will be graded for successful completion (on time, correct format), application of concepts and code, and discussion of topics.

Late work: Assignments will be accepted up to two weeks late for a maximum of 75 % of possible points. It is important that you develop your skills through assignments, regardless of whether you can complete the work by the deadline. This also affords you the flexibility to complete an assignment even if you run into a coding barrier or troubleshooting issue. If you do turn in an assignment after the due date, please email me and Njeri so that we are aware of your submission.

Grade appeals: Coding and data analysis are creative fields, often with multiple paths to a desired outcome. With that in mind, you are welcome to appeal a score on an assignment if you can provide justification for receiving a higher score. Appeals must be submitted in writing (via email) to me and Njeri within one week of receiving your assignment score.

Project 25%

Students will choose one of the example datasets from class and pose a question that can be addressed with the data analysis tools learned in class. Students will then create a data workflow protocol to answer this question and create a report that details methodology and findings. These components will be detailed in a report, a template for which will be supplied. More details and a grading rubric will be supplied in a separate instructions document.

Course Schedule

Date	Recorded content	Synchronous Activity	Assignment
Jan 20		Course introduction & setup	
Jan 27	Software Set-up;	Q&A on class set up	Assignment 1
Feb 10	Reproducibility & coding basics		Assignment 2
Feb 17	Data exploration		
Feb 24	Data wrangling		
Mar 3	Data visualization		
Mar 10	-no class-		
Mar 17	Generalized linear modeling & visualization		
Mar 24	Time series		
Mar 31	Spatial visualization & analysis		
Apr 7	-no class- (MP presentations & Wellness day)		
Apr 14	Crafting reports		

Details of this syllabus, including but not limited to grading and schedule, are subject to change at the instructor's discretion. Students will be notified of any changes in advance and in writing.

The Duke Community Standard

Duke University is a community dedicated to scholarship, leadership, and service and to the principles of honesty, fairness, respect, and accountability. Citizens of this community commit to reflect upon and uphold these principles in all academic and nonacademic endeavors, and to protect and promote a culture of integrity. To uphold the Duke Community Standard:

- I will not lie, cheat, or steal in my academic endeavors;
- I will conduct myself responsibly in all my endeavors; and
- I will act if the Standard is compromised.

If a student engages in academic dishonesty (e.g., cheating, plagiarism), they will receive an automatic zero on that assignment. Additional incidences of academic dishonesty will result in automatic failure of the course. We will also report incidences of academic dishonesty to the Office of Student Conduct, who may choose to impose additional consequences.

Diversity Statement

We all bring unique backgrounds, experiences, and perspectives to the classes we take. Collectively, these enrich the learning pathways, but only if we show each other the proper respect and allow each other to share new and different viewpoints. As your instructor, I will do my best to foster inclusive and equitable conversations on all aspects of the material covered and to overcome any implicit biases I may have. I encourage students to point out shortcomings that I or others may have on this front so we can work through them.

Native Lands Acknowledgment

It's important to contemplate that Duke University occupies land that once supported other communities, and that these communities were likely displaced by unfriendly means. Have a look at his [web mapping app](#) (or the map below) to discover what native communities once existed throughout the US. Credit to <https://native-land.ca> for the source of these data.