EDS 214 - Analytical Workflows and Scientific Reproducibility

EDS 214 is an intensive 1-week long 2-unit course. Students should plan to attend **all** scheduled sessions. All course requirements will be completed between 10am and 4:30pm PT (M - F), i.e. you are not expected to do additional work for EDS 214 outside of those hours, unless you are with the Teaching Assistant during student hours.

Meeting time & location

Dates: Monday (2024-08-26) - Friday (2024-08-30)
Schedule: 10am to 4:30pm PT (M - F); 9AM early start on Wed
Location: NCEAS 1st floor classroom

Instructor & TA information

Instructor: Julien Brun (https://brunj7.github.io/)
TA: Brian Lee (https://bren.ucsb.edu/people/brian-lee)
MEDS Slack is the best way to communicate with us.

Course description

The generation and analysis of environmental data is often a complex, multi-step process that may involve the collaboration of many people. Increasingly tools that document and help to organize workflows are being used to ensure reproducibility, shareability, and transparency of the results. This course will introduce students to the conceptual organization of analytical workflows (including code, documents, and data) as a way to conduct reproducible analyses.

These concepts will be combined with the practice of various tools and collaborative coding techniques to develop and manage multi-step analytical workflows as a team through students' first group projects. Students will also be introduced to the use of a remote server to conduct data-intensive analyses.

Course website

https://brunj7.github.io/EDS-214-analytical-workflows/

Recommended readings

- R for Data Science, 2nd edition: https://r4ds.hadley.nz/
- The Practice of Reproducible Research: http://www.practicereproducibleresearch.org/

Class structure

Daily class structure may vary a bit day-to-day, however, you can generally expect two lectures, two interactive sessions, and an end-of-day activity to practice / reinforce concepts learned. There is no take-home homework.

Grades

Grades for EDS 214 are assigned using the Letter-based scale (A-F). Attendance each day is required with exceptions to be requested ahead of time except for illness and family emergencies. Please see the MEDS summer absence policy for more details.

For summer Environmental Data Science (EDS) courses, graduate students are expected to be present in person for every day of class. The intensive nature of the summer session means that missing even a day or two can significantly hinder your progress in the program -- each day is equivalent to a week of instruction during which you will be learning crucial fundamentals. There is no online option to attend class remotely. Students should only miss a class if they are experiencing an illness or family emergency. If a student will miss part or all of a class, they must inform the instructor as soon as possible and devise a strategy to complete the coursework or risk not receiving a grade (including a "Satisfactory" grade) for the course. If you are feeling ill, please do not come to class and contact the instructor to arrange accommodations.

MEDS summer absence policy

For summer Environmental Data Science (EDS) courses, graduate students are expected to be present in person for every day of class. The intensive nature of the summer session means that missing even a day or two can significantly hinder your progress in the program -- each day is equivalent to a week of instruction during which you will be learning crucial fundamentals. There is no online option to attend class remotely. Students should only miss a class if they are experiencing an illness or family emergency. If a student misses part or all of a class, they must inform the instructor as soon as possible and devise a strategy to complete the coursework or risk not receiving a grade (including a "pass") for the course. If you are feeling ill, please do not come to class and contact the instructor to arrange accommodations.

Computing

- Minimum MEDS device requirements
- Have a ready-to-be-used GitHub Account (https://github.com/)
- MEDS server Taylor