EDS 214: Analytical Workflows and Scientific Reproducibility

"Making reproducibility of your work by peers a realistic possibility sends a strong signal of quality, trustworthiness, and transparency." - Sandve et al. (2013) PLOS Computational Biology

👩‍🏫 Instructor: Max Czapanskiy (they/them)
📨 Email: maxczap@ucsb.edu
🏢 Office hours: M-F 4:30-5:00PM

👩‍🏫 TA: Ale Vidal Meza (she/her/ella)
📨 Email: avidalmeza@ucsb.edu

⏱️Class Time: Monday - Friday, 10:00 AM - 4:30 PM
🏫Class Location: NCEAS Classroom

I. 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, data scientists use tools that document and help to organize workflows to ensure reproducibility, shareability, and transparency of the results. This course will introduce students to the conceptual organization of workflows (including code, documents, and data) as a way to conduct reproducible analyses.

These concepts will be combined with the practice of various software tools and collaborative coding techniques to develop and manage multi-step analytical workflows as a team.

Course Website

https://eds-214-reproducibility-m26.github.io/

All course content / information will be posted to this course website. Be sure to bookmark it somewhere you can easily find it!

II. Learning Goals

The goal of this course is to enable MEDS students to collaboratively create reproducible analyses. Essential skills and concepts are:

- 🤖***Automate the steps in an analytical workflow using scripts***
- 🗂️***Organize workflow components modularly***
- 📝***Document individual components and their relationships***
- 🐘***Scale workflows for computational performance and large datasets***
- 🤝***Collaborate in a team to develop a workflow***

III. Course Format

As a MEDS summer course, EDS 214 meets Monday through Friday 10am - 4:30pm. Each day will be divided into a morning session (10-12), an afternoon session (1-3), and an end-of-day session (3:30-4:30). We'll have a longer break for lunch from 12-1 and a shorter break after the afternoon session 3-3:30.

> *Breaks are critical to your success in this class*! Brains need downtime to integrate new knowledge. During the breaks, I encourage you to close your laptops and take a walk around town. A walk down State Street will do more for your learning than looking at a screen.

The **morning and afternoon sessions** cover new content with a variety of activities. You will be actively participating in live-coding demos, paper-and-pencil problem sets, and mini-lectures throughout these sessions. The **end-of-day session** will be office hours with your TA.

There is no homework or reading assigned outside of class. I will provide you with links to additional resources on the course website that you can use to review or deepen your understanding. These resources include online articles, tutorials, and textbooks. Many students find them helpful later in their degrees when they review material covered in earlier classes.

IV. Course Content

Below is a table of the topics and sessions covered over the ten days of EDS 221. Slides and activities will be posted to the course website.

☀️Day	🔎 Topic	📑Sessions
1	Organize	Reproducible workflows Workflow organization
2	Automate	Command line and remote server Documentation
3	Collaborate	Version control with git Collaboration with GitHub
4	Implement	Draft Quarto analysis Peer feedback and contributions
5	Review	Meet your next instructor Instructor feedback

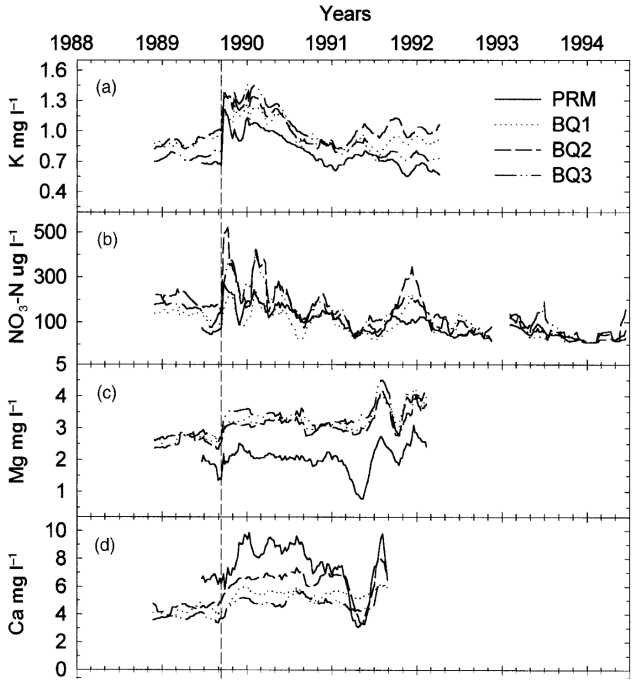
V. Grading

Grades for EDS 214 are assigned based on ***attendance, reflections, and your final project***. You must earn a C or better satisfactory grade to fulfill the requirements of your MEDS degree. Attendance each day is required to earn a passing grade. Absences for illness are excused; email your instructor to let them know you will be missing class. Please see the MEDS summer absence policy, below, 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 "pass") for the course. If you are feeling ill, please do not come to class and contact the instructor to arrange accommodations.

- ☑**Attendance**. Attend all five classes to earn an A, four classes for a B, and three for a C. If you have to miss a class for an excused reason (e.g., illness, family obligation, or medical appointment) let your instructor know in advance.

- 🤔**Reflections**. Each session (10 total) ends with a brief reflection to connect the session's content to the broader objectives of the class. Complete 9 or more reflections for an A, 7 or more reflections for a B, and 5 or more for a C.
- 🎯**Final project**. The final project consists of reproducing a figure from a publication. Specifically, figure three of Schaefer et al. (2000), part of which is shown to the right. Using the original data from the study and the tools/skills from this course, you will create a fully reproducible workflow. This final project has eight *specifications*, describing criteria for a successful workflow. Meet at least 7 specifications for an A, 6 for a B, and 4 for a C.



💯Grade	☑Attendance	🤔Reflections	🎯Final project
A	5 days	9+ reflections	7+ specifications
B	4 days	7+ reflections	6+ specifications
C	3 days	5+ reflections	4+ specifications

VI. AI Policy

In this core summer course, ***the use of AI tools to complete coursework is not permitted***. The course is designed to help you develop a strong foundation in the essential mathematical skills needed for your environmental data science classes this Fall. Generative AI shortcuts that formation by offloading the mental struggle required to build understanding. While it may be helpful in the short term, studies highlight that using AI to generate answers during early skill-building can hinder long-term skill development and independent problem-solving (Bastani et al., PNAS, 2025). Practicing involves making mistakes, and exploring solutions to gain understanding. We will move step-by-step, making sure you gain confidence, independence, and mastery through your own problem-solving and collaboration.

Submitting work you don't understand or can't explain or justify will be considered plagiarism. If there are concerns about AI use in your work, your instructor will ask you to meet and discuss it. If it appears that understanding and ownership of your work are incomplete, you will have the opportunity to resubmit the assignment with a reflection on your process. Repeated issues, however, will be referred to the UCSB Office of Student Conduct.

Your instructional team is here to support you every step of the way. By tackling challenges directly, practicing, and asking questions, you'll gain the confidence and skills to succeed both in this course and in your future work!

VII. Campus Resources

There are many on-campus resources that support graduate students in navigating challenges, building community, and staying well during the quarter. Your instructors are also available to discuss individual needs and help connect you with a campus resource that may be best suited to your situation.

- 🍎**Food Insecurity:** UCSB has several food resources available. Explore options on their website, including the CalFresh Program (aka SNAP benefits) and the Associated Students food bank.
- 🏡**Housing Needs:** As a student of UCSB and a member of the community, you should have access to stable housing. Browse UCSB's resources, including emergency transitional housing support.
- 🧠**Mental Health:** UCSB's Counseling and Psychological Services (CAPS) provides confidential mental health services and crisis support. Their 24/7 phone number is (805) 893-4411.
- 💵**Financial Crisis:** If you are experiencing a financial crisis, email financialcrisis@sa.ucsb.edu. Within 1-2 days, a financial advisor will reply. UCSB also has short-term emergency loan programs.
- 🦋**Undocumented Students:** Undocumented Student Services (USS) offers workshops, helps students find scholarships and financial support, and works to provide a campus community.
- ❤️**Interpersonal Violence:** The Campus Advocacy, Resources, & Education (CARE) office has 24/7 confidential support and advocacy for sexual assault, dating and domestic violence, and stalking.
- 🏳️‍🌈**Gender and Sexuality:** The Resource Center for Sexual and Gender Diversity (RCSGD) office and the Queer & Trans Graduate Student Union (QTGSU) hosts events & resources for LGBTQ+ students.
- 🏥**Medical Health:** All UCSB students, including master's students, have full access to the campus health center, including a free after-hours nurse line at (877) 351-3457 and social work services.
- ❌**Interpersonal Conflict:** The Office of the Ombuds is a safe and confidential place to discuss workplace issues, interpersonal conflict, academic concerns, and other campus-related problems.
- 🤲**Addictive Behaviors and Recovery:** The Addictive Behaviors & Recovery (ABR) promotes safe, healthy choices around substance use and offers free, confidential counseling to all students to reduce high-risk behavior.
- 🌎**Cultural Community:** The UCSB Multicultural Center (MCC) is a community space that provides critical dialogue and upholds a safer, welcoming space that validates marginalized identities.
- 🤎**Black Community:** The Black Graduate Student Association (BGSA) is a UCSB organization that provides academic support and social resources support that meet the needs of Black students.
- 🧡**Indigenous Community:** The American Indian and Indigenous Cultural Resource Center (AIICRC) provides professional development opportunities, scholarships, and community.
- 🩵**Asian & Pacific Islander Community:** The Asian Pacific Islander Graduate Student Alliance (APIGSA) strives to support graduate students of Asian and Pacific Islander panethnicity.
- 💚**Hispanic & Latinx Community:** The Society for Advancement of Chicanx and Native Americans in Science (SACNAS) includes a graduate chapter for the UCSB community that hosts events.
- 🚺**Women in STEM:** UCSB has a graduate chapter of Women and Science in Engineering (WISE). They host events, career panels, and regular K-12 outreach events with the community.
- 🎓**Graduate Advocacy:** The Graduate Student Association (GSA) is the elected representative government for students and includes funding for childcare support, emergency grants, & travel.

VIII. Instructional Team

Instructor: Max Czapanskiy is an Assistant Teaching Professor in Marine Data Science, jointly appointed to UCSB's Bren School of Environmental Science & Management and the College of Creative Studies. A former software engineer turned marine ecologist, they apply the tools and problem-solving strategies from

computer science to questions about the ecology and conservation of large marine vertebrates. They earned degrees in Computer Science, Geographic Information Systems, and Marine Biology while working for corporations (Microsoft), federal agencies (USGS and NOAA), and non-governmental organizations (Point Blue Conservation Science). When not at a keyboard or a whiteboard, they can most often be found in a kitchen baking bagels or trying a new recipe (preferably involving dried chiles). At UCSB, they teach MEDS courses and advise Capstone Projects in the Bren School, and teach undergraduate Marine Science majors in the College of Creative Studies.

Teaching Assistant: Ale Vidal Meza (she/her/ella) is a MEDS alumnus (Class of 2023) and 3rd-year PhD student advised by Dr. Ruth Oliver. After MEDS, she was a Dangermond Fellow at the National Audubon Society in partnership with Esri, using GIS to protect birds and the places they need. She notably mapped out the demographies near Audubon centers and sanctuaries to get clear, actionable data on which communities do and do not have access to bird conservation programs. She then worked as a data scientist at Oak Ridge National Laboratory before returning to Bren. Now, as a PhD student, her research interests sit at the intersection of macroecology and social-ecological systems in urban environments. She integrates animal movement analysis with econometrics to better understand how social inequity underlies animal behavior in cities. At UCSB, she is also an organizer for the graduate chapter of the Society for Advancement of Chicanos and Native Americans in Science (SACNAS).