University of Alberta

EAS 405 (505)
Geoscience Data Analysis
LEC X01 - Winter 2025

Instructor: Christopher Mallon
Office: CCIS 3-279
E-mail: chris.mallon@ualberta.ca
Web Page: https://apps.ualberta.ca/directory/person/mallon1
Office Hours: immediately after the class, i.e., after 9pm Wednesdays, or by appointment.

Instructor: Eric Timmer
Office: CCIS 3-279
E-mail: etimmer@ualberta.ca
Web Page: https://apps.ualberta.ca/directory/person/etimmer
Office Hours: immediately after the class, i.e., after 9pm Wednesdays, or by appointment.

Lecture Room & Time: CAB 345 - Wednesdays, 6:00pm - 9:00pm

LAND ACKNOWLEDGEMENT

The land powers our computers, gives us food and air, and provides us with homes and places to gather. Our ability to gather and learn together is a result of agreements between the Indigenous peoples of this land, and all others who have since settled here. Treaty Four, Treaty Six, Treaty Seven, Treaty Eight, and Treaty Ten cover the entire province, making all Albertans Treaty people. With gratitude and respect, we acknowledge the privilege and responsibility of being Treaty people. We acknowledge the long and ongoing history of the Indigenous peoples of Turtle Island, their unique and profound connection to the land, and their invaluable role in present-day society.

COURSE CONTENT

Course Description:
Modern scientists are often confronted with a historically novel problem: enormous, complex datasets that require advanced tools to produce insights. This course aims to enable students to explore, manipulate, and efficiently analyze datasets of virtually any size, using the R programming language. The course, designed for students without any experience in programming, will teach a combination of computational theory, basic statistics, spatial data analysis, data visualization, data science techniques, machine learning methods, and more complex tasks. In this highly practical course, the focus will be practicing coding and data analysis. Students are optionally invited to bring their own data as a part of their final project. Any advanced EAS student can benefit from this course, regardless of their field of study or their research subject.

Course Prerequisites:
Any 300/400-level EAS course and permission of the department. At least one statistics course (any level), and a basic understanding of algebra. No prior experience in programming is needed. The course is intended to improve the programming and data analysis capabilities of students without existing programming experience.

Course Objectives and Expected Learning Outcomes:
By the end of this course, participants should be able to:

- Understand and describe the basics of R programming:
    - History, strengths, and limitations.
    - Comparison with other programming languages.
    - Use of integrated development environments (e.g., RStudio) and explain their utility and limitations.
    - Use and application of unit tests for debugging code.
- Use R to:
    - Read, clean, sort, group, and analyze datasets of almost any size.
    - Perform simple statistical analysis on data such as summary statistics.
    - Visualize data in different forms such as points, lines, box plots, bar charts, etc.
    - Create interactive dashboards for dynamic visualizations and communication of results.
- Use Git to:
    - Build local and remote repositories.
    - Track changes and keep a history of work.
- Understand advanced analytical methods:
    - Describe, import, and analyze spatial data; perform basic geoprocessing and raster processing operations.
    - Formulate questions as machine learning problems.
    - Describe and apply machine learning techniques to solve research problems.
    - Be familiar with parallel processing for large dataset operations.

COURSE SCHEDULE

Week	Date (2024)	Topic	Assignments
1	January 8	**Introduction** • Course content, policy, and expectations • Explanation of assignment structure and project proposals • Exploring the R development environment - IDE, console, etc. • Demo using built-in example datasets	
2	January 15	**Logic, coding, data structures, and functions** • Logic • Loading data from CSV files • Coding basics – using an IDE, writing your first scripts • Basic objects in R – defining variables and assigning different values • Writing functions and basic scripts	Assignment 1 due*
3	January 22	**Version control and R Markdown** • Creating an R Project • Introduction to Git • Code annotation and report creation with RMarkdown	Assignment 2 due*
4	January 29	**Introduction to Data Cleaning** • Introduction to package management • Data cleaning • Summarizing, grouping, and filtering data • Using the pipe operator	Assignment 3 due*
5	February 5	**Introduction to Data Wrangling** • Definition of “tidy” data model • Demonstration of data wrangling techniques • Application of functions	Assignment 4 due*
6	February 12	**Data Visualization** • Data visualization using R basic plot and ggplot2 • Chart types and different graphs for different types of data	Assignment 5 due* Project proposal submission deadline: Feb 14.

Week	Date (2024)	Topic	Assignments
7	February 19	• No lecture - Reading Week	
8	February 26	**Shiny (Part 1)** • Inputs, Outputs • Debugging	Assignment 6 due*
9	March 5	**Shiny (Part 2)** • Introduction to reactive programming in Shiny • Layout and themes • Introduction shinyapps.io	Assignment 7 due*
10	March 12	**Use and Analysis of Spatial Data** • Introduction to vector- and raster-based spatial data • Basic analytical tools (sf and terra packages) • Integrating spatial data into workflows	Assignment 8 due*
11	March 19	**Machine Learning** • How to think like a machine • Big Data • Parallel processing	Assignment 9 due*
12	March 26	**Unit Testing and Use of AI** • Introduction to unit testing • Introduction to AI tools for coding • Using AI in test-driven development	No Assignment
13	April 2	**Working session/review of concepts for final report**	No Assignment
14	April 9	**Presentation of dashboards**	Final project presentations, report, and code due April 9th

* due by the start of the next lecture (i.e., 6pm)

LEARNING RESOURCES

Required Textbook and/or Other Major Course Materials:
For lecture notes and papers to be distributed to the class, we will upload the materials on e-class or email the documents to the students at least one week before the notes are discussed.

Recommended Learning Resources:

- R for Data Science (https://r4ds.hadley.nz/)
- ggplot2: Elegant Graphics for Data Analysis (3e) (https://ggplot2-book.org/)
- Mastering Shiny: https://mastering-shiny.org/
- Advanced R (https://adv-r.hadley.nz/)
- Learn Git Branching (https://learngitbranching.js.org/)
- R-Spatial: https://www.rspatial.org, https://r-spatial.org

On-Line Homework Disclaimer

Due to the nature of the course, which involves homework learning technologies that require third party services, online homework is a component of this course and is provided by at least two third-party companies. However, these companies will not have access to any assessment information. They will have access to your homework, but not to your grades. If you have any concerns about this, please contact the instructor of this course.

1. Registration in the system and any monetary transactions are of your own accord and not the responsibility of the University.
2. Students should be mindful of protecting their personal information and should be aware of how their personal information might be used and/or shared.
3. Students MUST NOT use their @ualberta email address or CCID to register into the system and instead should use a non-identifying email address or account.

Academic Success Centre:
The Academic Success Centre provides professional academic support to help students strengthen their academic skills and achieve their academic goals. Individual advising, appointments, and group workshops are available year round in the areas of Accessibility, Communication, Learning, and Writing Resources. Modest fees apply for some services.

Faculty of Science Student Services:

The Faculty of Science Student Services office is located on the main floor of the Centennial Centre for Interdisciplinary Sciences (CCIS). This office can assist with the planning of Your Academics, and provide information related to Student Life & Engagement, Internship & Careers, and Study Abroad opportunities. Please visit Advising for more information about what Faculty Academic Advisors in the Student Services Office can assist you with.

GRADE EVALUATION

Assessment	Weight	Date
Attendance and participation	10%	-
Final Project Proposal	5%	February 14th, 2024
Weekly Assignments	40%	Weekly
Final Project	45%	April 9th, 2024

Grades are unofficial until approved by the Department and/or Faculty offering the course.

Weekly Assignments
There will be weekly assignments based on the week's lecture contents. Assignments will include using Git, RMarkdown, and various data analysis tasks (all these tools will be covered in class before the student is expected to use them). Assignments are due by the beginning of the lecture (i.e., 6pm) following the lecture in which they've been assigned.

Assignments will generally be graded as follows:

- Error-free execution of code (30%)
- Accurate results and responses (45%)
- Commit history (25%)

Complexity of assignments will increase to match the students' abilities. Students will generally have to submit their assignments, at the latest, by the beginning of the lecture following their assignment. Time will be given in class to work on the assignments, with support from instructors and colleagues.

Final Project
The final project will consist of the construction of an interactive dashboard, and a report (both due April 9th, 2024). The student will be expected to select a topic relevant to them, and it must be approved by the instructors (by reading week).

The project will be expected to use concepts from the course, and involve novel techniques that each student will apply uniquely to their research problem. In the project, a greater emphasis will be put on reproducibility, clarity of analysis, and communication of the results. The project proposal (due February 14th, 2024) is intended to ensure that the student selects a research focus that is suitable for use in the course, and that their plan will reflect sufficient depth of investigation to warrant it being a final project. The students are encouraged to make use of any publicly-available data for use in the project.

Statement of Expectations for AI Use:
Students **must not** submit work by LLMs as their own, this is a form of **plagiarism.** The use of large language models (ChatGPT, Bard/Gemini, Microsoft Bing/Copilot, Llama, etc.) will be discussed in the course. While students are not prohibited from using these tools, it is

strongly encouraged that students refrain from their use in most of the assignments. The course will discuss prudent usage of these tools, and how to integrate them into a data analysis workflow; it is critical that the student reaches a level of necessary proficiency through the course before using these tools.

Any material submitted from a large language model must be cited. Include the author (corporation) of the model, the name of the model, the subject queried, and the date.

These models are wrong a lot. Stack Overflow is a much more reliable resource for questions on how to do things in R.

POLICIES FOR LATE AND MISSED WORK

Late Policies:
Assignments that are submitted late will be reduced by 10% for each day the assignment is not submitted, to a maximum of five days (including weekends and holidays). After five days, the assignment will be given a grade of zero. Note that the student *must complete* the final project proposal in order to pass the course.

A student who cannot complete any assignment due to incapacitating illness, severe domestic affliction or other compelling reasons must contact the instructor within two working days of missing the assessment, or as soon as possible, to request an excused absence. If an excused absence is granted, then the student and instructors will assess the situation and determine a resolution on a case-by-case basis. An excused absence is a privilege and not a right. There is no guarantee that an absence will be excused. Misrepresentation of facts to gain an excused absence is a serious breach of the Code of Student Behaviour. In all cases, instructors may request adequate documentation to substantiate the reason for the absence at their discretion.

Missed Term Work or Final Exam Due to Non-medical Protected Grounds (e.g., religious beliefs):
When a term assessment or final exam presents a conflict based on non-medical protected grounds, students must apply to the Academic Success Centre for accommodations via their Register for Accommodations website. Students can review their eligibility and choose the application process specific for *Accommodations Based on Non-medical Protected Grounds*.

It is imperative that students review the dates of all course assessments upon receipt of the course syllabus, and apply ***AS SOON AS POSSIBLE*** to ensure the timely application of the accommodation. Students who apply later in the term may experience unavoidable delays in the processing of the application, which can affect the accommodation.

STUDENT RESPONSIBILITIES

Privacy:
This course asks students to push code to Github, and shinyapps.io. Students may also collaborate with their instructors and peers on this platform. Please produce work knowing that the material submitted to these third-party companies will be stored on servers owned by a private third-party, housed outside of Canada. It is the responsibility of the student to observe good online security and privacy practices, such as using strong passwords, unique usernames, and never storing any personally identifiable information.

Academic Integrity and Student Conduct:
The University of Alberta is committed to the highest standards of academic integrity and honesty, as well as maintaining a learning environment that fosters the safety, security, and the inherent dignity of each member of the community, ensuring students conduct themselves accordingly. Students are expected to be familiar with the standards of academic honesty and appropriate student conduct, and to uphold the policies of the University in this respect.

Students are particularly urged to familiarize themselves with the provisions of the Student Academic Integrity Policy and the Student Conduct Policy, and avoid any behaviour that could potentially result in suspicions of academic misconduct (e.g., cheating, plagiarism, misrepresentation of facts, participation in an offence) and non-academic misconduct (e.g., discrimination, harassment, physical assault). Academic and non-academic misconduct are taken very seriously and can result in suspension or expulsion from the University.

All students are expected to consult the Academic Integrity website for clarification on the various academic offences. All forms of academic dishonesty are unacceptable at the University. Unfamiliarity of the rules, procrastination or personal pressures are not acceptable excuses for committing an offence. Listen to your instructor, be a good person, ask for help when you need it, and do your own work -- this will lead you toward a path to success. Any academic integrity concern in this course will be reported to the College of Natural and Applied Sciences. Suspected cases of non-academic misconduct will be reported to the Dean of Students. The College, the Faculty, and the Dean of Students are committed to student rights and responsibilities, and adhere to due process and administrative fairness, as outlined in the Student Academic Integrity Policy and the Student Conduct Policy. Please refer to the policy websites for details on inappropriate behaviours and possible sanctions.

The College of Natural and Applied Sciences (CNAS) has created an Academic Integrity for CNAS Students eClass site. Students can self enroll and review the various resources provided, including the importance of academic integrity, examples of academic misconduct & possible sanctions, and the academic misconduct & appeal process. They can also complete assessments to test their knowledge and earn a completion certificate.

"Integrity is doing the right thing, even when no one is watching." -- C.S. Lewis

Contract Cheating and Misuse of University Academic Materials or Other Assets:

Contract cheating describes the form of academic dishonesty where students get academic work completed on their behalf, which they then submit for academic credit as if they had created it themselves. Contract cheating may or may not involve the payment of a fee to a third party, who then creates the work for the student.

Contract cheating companies thrive on making students believe that they cannot succeed without their help; they attempt to convince students that cheating is the only way to succeed. Do not believe them. You already have everything you need to succeed.

Uploading the instructor's teaching materials (e.g., course outlines, lecture slides, assignment or exam questions, etc.) to tutorial, study or note-sharing websites or public servers is a copyright infringement and constitutes the misuse of University academic materials or other assets. Receiving assignment solutions or answers to exam questions from an unauthorised source puts you at risk of receiving inaccurate information.

Appropriate Collaboration:
Students need to be able to recognize when they've crossed the line between appropriate collaboration and inappropriate collaboration. If students are unsure, they need to ask the instructors to clarify what's allowed and what's not allowed.

Here are some tips to avoid copying on assessments:

1) Do not write any code, or any concepts, that you cannot explain to your instructor.

2) When you are helping other students, avoid showing them your work directly. Instead, explain your solution verbally. Allowing your work to be copied is also considered inappropriate collaboration.

3) It is possible that verbally discussing the solution in too much detail may result in written responses that are too similar. Try to keep discussions at a general or high level.

4) If you find yourself reading another student's solution, do not write anything down. Once you understand how to solve the problem, remove the other person's work from your sight and then write up the solution to the question yourself. Looking back and forth between someone else's paper and your own paper is almost certainly copying and considered inappropriate collaboration.

5) If the instructor or TA writes down part of a solution in order to help explain it to you or the class, you cannot copy it and hand it in for credit. Treat it the same way you would treat another student's work with respect to copying, that is, remove the explanation from your sight and then write up the solution yourself.

6) There is often more than one way to solve a problem. Choose the method that makes the most sense to you rather than the method that other students happen to use. If none of the ideas in your solution are your own, there is a good chance it will be flagged as copying.

For more information, see "How to Avoid Inappropriate Collaboration."

Cell Phones:
Cell phones are to be turned off during lectures, labs and seminars.

Students Eligible for Accessibility-Related Accommodations:
In accordance with the University of Alberta's Discrimination, Harassment, and Duty to Accommodate policy, accommodation support is available to eligible students who encounter limitations or restrictions to their ability to perform the daily activities necessary to pursue studies at a post-secondary level due to medical conditions and/or non-medical protected grounds. Accommodations are coordinated through the Academic Success Centre, and students can learn more about eligibility on the Register for Accommodations website.

It is recommended that students apply as early as possible in order to ensure sufficient time to complete accommodation registration and coordination. Students are advised to review and adhere to published deadlines for accommodation approval and for specific accommodation requests (e.g., exam registration submission deadlines). Students who request accommodations less than a month in advance of the academic term for which they require accommodations may experience unavoidable delays or consequences in their academic programs, and may need to consider alternative academic schedules.

Recording and/or Distribution of Course Materials:
Audio or video recording, digital or otherwise, of lectures, labs, seminars or any other teaching environment by students is allowed only with the prior written consent of the instructor or as a part of an approved accommodation plan. Student or instructor content, digital or otherwise, created and/or used within the context of the course is to be used solely for personal study, and is not to be used or distributed for any other purpose without prior written consent from the content author(s).

Learning and Working Environment:
The Faculty of Science is committed to ensuring that all students, faculty and staff are able to work and study in an environment that is safe and free from discrimination, harassment, and violence of any kind. It does not tolerate behavior that undermines that environment. This includes virtual environments and platforms.

If you are experiencing harassment, discrimination, fraud, theft or any other issue and would like to get confidential advice, please contact any of these campus services:

- Office of Safe Disclosure & Human Rights*: A safe, neutral and confidential space to disclose concerns about how the University of Alberta policies, procedures or ethical standards are being applied. They provide strategic advice and referral on matters such as discrimination, harassment, duty to accommodate and wrong-doings. Disclosures can be made in person or online using the Online Reporting Tool*.
- University of Alberta Protective Services: *Peace officers dedicated to ensuring the safety and security of U of A campuses and community. Staff or students can contact UAPS to make a report if they feel unsafe, threatened, or targeted on campus or by another member of the university community.*
- Office of the Student Ombuds*: A confidential and free service that strives to ensure that university processes related to students operate as fairly as possible. They offer information, advice, and support to students, faculty, and staff as they deal with academic, discipline, interpersonal, and financial issues related to student programs.*

- Office of the Dean of Students*: They can assist students in navigating services to ensure they receive appropriate and timely resources. For students who are unsure of the support they may need, are concerned about how to access services on campus, or feel like they may need interim support while you wait to access a service, the Dean of Students office is here to help.*

Feeling Stressed, Anxious, or Upset?
It's normal for us to have different mental health experiences throughout the year. Know that there are people who want to help. You can reach out to your friends and access a variety of supports available on and off campus at the Need Help Now webpage or by calling the 24-hour Distress Line: 780-482-4357 (HELP).

Student Self-Care Guide:
This Self-Care Guide, originally designed by the Faculty of Native Studies, has broader application for use during students' learning. It provides some ideas and strategies to consider that can help navigate emotionally challenging or triggering material.

Policy about course outlines can be found in Course Requirements, Evaluation Procedures and Grading of the University Calendar.

Disclaimer:
Any typographical errors in this Course Outline are subject to change and will be announced in class. The date of the final examination is set by the Registrar and takes precedence over the final examination date reported in this syllabus.