

Data Science in Ecology

BIOL4253

Instructor: Trevor Avery
Office: BIO410
Office Hours: Online via Teams or in person by appointment
E-mail: trevor.avery@acadiau.ca

Slot: 31
DateTime: Mon 1:00–4:00 PM
Room: HSH 019

TA: Paige Levangie
Office: BIO404
Office Hours: By appointment
E-mail: paigelevangie@acadiau.ca



	Mark	Due Date
Participation	10 %	Assessed by Instructor/TA/random checks
Learning Log	5 %	Random checks
Exercise 1	20 %	27 Jan
Exercise 2	20 %	24 Feb
Exercise 3	15 %	10 Mar
Workgroup Presentation + Code	20 %	31 Mar
Final Exam (Take home)	10 %	End of exam period

Part 1: Course Information

Course Description

Core topics include data collection, visualization and wrangling, reproducible workflows, basic applied statistical modeling and reporting.

Prerequisite: Undergraduates MATH2233/1523/2243/2253 or a similar introductory statistics course
 Permission of the instructor required

Textbook & Course Materials

A text is not required. There are some useful FREE texts and resources though that will be outlined in class.

Course Requirements

1. Access to Moodle and Microsoft Teams. Moodle is the central portal for all course material. Teams may be used to share code and analyses, and can be used to interact with the instructor and TA for appointments/help.
2. Access to Teams is provided through Office356 available to all students.
3. Computer with R and RStudio installed.
4. Enthusiasm.

Course Structure

The course will be presented in-person. Exercises and code/scripts will be posted to Moodle and GitHub.

Lecture Recordings

Students may not create audio or video recordings without written permission from the instructor. Permission for such recordings will not be withheld for students with accommodation needs. The reality is this is a hands-on, participation-driven course that is not served well by non-interactive modes of teaching.

Student Learning Outcomes

1. Understand how to use your computer (for more than watching streaming content, running Instagram and videos) including files, folders, syncing to online content (Teams in this case), downloading, and online data connections.
2. Be able to work through a complex dataset, clean it, wrangle it, and produce summary information from it in a table.
3. Be able to visualize data by creating figures and diagrams to summarize data in a visual format.

4. Be able to organize data and analysis repeatable workflow to accomplish tasks.
5. Analyze data using general/generalized linear models, understand the model framework, validate and interpret models, and report the outcomes.
6. Be able to report your workflow through dynamic documents such as RMarkdown and/or Quarto.

How to Meet the Learning Outcomes

1. Attend lectures/work sessions, take notes, and **participate**.
2. Read required readings and complete tasks on Moodle before lectures/work sessions
3. Complete exercises. Some of these will be in a blended learning format where you create the works and then these are reflected on during class.
4. Work with R/RStudio/Excel/Quarto/RMarkdown tools regularly and often.
5. Participate in group activities such as those organized by your working group.

Part 2: Topic Outline/Schedule

Tasks

Tasks are meant to enhance the classroom activities and provide training in Excel and R/RStudio beyond what is taught in the classroom and/or to practice what you learned. This approach is a variant of blended learning; think of it as homework. Tasks are required, choosing not to do tasks is generally very noticeable in performance and will result in poor performance on exercises. Tasks become part of exercises in some form.

Tasks will be posted to Moodle and reflect what was done in class each week. Tasks will typically be a series of activities to accomplish.

Learning Log

Undergraduate students must provide a detailed learning log on what tasks were completed, task outcomes, etc. Learning logs will be setup on Moodle to be completed weekly. A learning log is a record of what you accomplish and must be updated before each class. It can have tables, visualizations, text, code, etc.. As you get better at using R/RStudio/Quarto, you may wish (strongly consider) migrating your Learning Log to a this format.

Lecture Schedule

Basically, we'll try to stick to a schedule (see below). Obviously, some events may alter the schedule, but it can also be altered by shared interest. As well, timelines are subject to change due to weather, availability of course materials, and random occurrences beyond our control like pandemics.

Date	Topic #: Topic	Synopsis	Notes
6 Jan	1: Introduction	Course intro, expectations, etc.	
	2: R/RStudio	Confirmation of R/RStudio installation. Intro, review, and application. Projects and writing scripts	
	3: Files and Folders	Where are my files? Relative file structure. Sharing/syncing files via online resources (Teams, etc.)	
	4: File Formats + Excel	There are many file formats. Dos and Don'ts	
	5: Variable names	Variable naming conventions	
	6: Importing/Exporting files	Importing/exporting/connecting to files to/from/in R. File validation/verification. R data formats	
13 Jan	7: Workgroup Projects	Outline workgroups and projects	
	8: Tidyverse – dplyr	Understanding the grammar of the tidyverse	
	9: Cleaning and Wrangling data	Zeros and NAs, cases, mixed data types, saving cleaned data	
20 Jan	10: Tidyverse – lubridate	Working with dates and times	
	11: Summarizing data	Summarizing data	
	12: Working with tables	Making tables in kable, gt, dt Saving tables as files	
27 Jan	Exercise 1 Due	Data Wrangling	20 %

27 Jan	13: Tidyverse – ggplot2	Introduction, aesthetics, basic plots (histograms, dot plots, scatter plot, boxplots), axis and title labels, saving plots	
	14: Designing plots	Practicing plots, layering plots, adding lines and dots	
	15: Multiple elements	Colors, groups, fills, factors, themes, facets	
3 Feb	16: Scales	Scales, dates and times	
	17: Legends, annotations	Legends, annotations, adding images	
	18: Equations, text manipulation	Equations, advanced text	
10 Feb	19: Reproducible research Reproducible workflows	RMarkdown syntax, formatting, code chunks, inline text	
	20: Quarto	Expanding RMarkdown	
	21: Advanced RMarkdown/Quarto	RMarkdown equations (Latex), figures, tables, images and links	
17 - 21 Feb	No Classes	Study Break	
24 Feb	Exercise 2 Due	Making a Report	20 %
24 Feb	22: Mapping		
	23: Creating basemaps		
	24: Annotating maps		
	25: Exporting maps		
3 Mar	26: Images	It's Magick!	
	27: Working with image types		
	28: Publication quality images		
10 Mar	Exercise 3 Due	Adding a Map and Images	15%
10 Mar	29: Statistical Modeling	What is it?	
	30: Fitting a linear regression	Review – it's a model!	
	31: GLMs	Generalizing stuff	
	32: Model equations	Writing model equations	
	33: Model validation	Residual plots	
17 Mar	34: General Linear Model Demo	Continuous and Categorical Gaussian (normal) Fitting a model	
	35: Model validation	Residual plots	
	36: Model selection	Selecting the “best” model	
24 Mar	37: ggeffects	New package!	
	38: Model testing	Testing effects	
	39: GLM visualization	Using ggeffects	
	40: Other model types	Normal is not normal...	
31 Mar	41: Workgroup Project	Presentations	10 %
31 Mar	42: Workgroup Project	Reports	10 %
Take Home Exam	Final Exam	Final Report	10 %

Exercises

Data Science is an intensive, immersive, hands-on experience. Thus, the course will consist of several escalating exercises. Exercises are aimed to put into practice the lessons learned, coding, syntax, and examples. Exercises will build on each other so **save your scripts** because you can recycle much of your code!

All exercises will be posted on Moodle and/or GitHub. There will be a way to share code with your instructors, but that's currently being sorted.

Datasets

Datasets will be provided for course work. Workgroup datasets will need to be found!

Part 3: Grading Policy

Due Dates

Due dates are absolute, therefore, stay ahead of the work.

Exercises

Exercises are designed to investigate what you have learned and apply it to real data. What are the main concepts? Are you able to apply concepts with examples? Do you understand the material, and can you integrate terminology, theoretical concepts, and apply these in some sensible way? Can you integrate the required readings/tasks to support your work? Is your code clean, well commented, and does it work from a 'fresh start'? Are all your supporting files, folders, data and scripts organized and in the correct formats? Have you done everything asked in each exercise?

Assessment of exercises will be done in class (e.g. participation, review, feedback) and through submitting code at the end of class. It is straightforward to 'see' what each person is doing during class time. Feedback will be provided so be prepared to take notes on how you corrected mistakes, improved coding/syntax/organization etc., and clearly document what was revised and how you did it (this can become part of your Learning Log).

Feedback will be applied both to the current exercise and carried forward for future exercises. Failure to implement feedback from one exercise to future exercises will be viewed much more critically.

Workgroup Projects

Each student will belong to a working group. Each workgroup will be responsible to find at least one dataset and apply what is learned in class to that dataset in addition to the datasets provided in class. A presentation and report are required to be produced as a group and clearly detailed as to who undertook what tasks. All students will present during their workgroup presentation. Details on working groups will be provided on Moodle.

Participation/Learning Logs

Participation will be assessed by in-class participation (e.g. answering questions, interacting with classmates, completing work during class), participation in organized workgroup (or other) study/work sessions, and through random checks of weekly code uploads.

Learning logs will be assessed randomly through weekly random checks.

Random checks will consist of randomly checking a proportion of students each week for each of participation and learning log. Each week a random selection of students will be made and will include previously unchecked students (higher proportion) and previously checked students (lower proportion).

Make-up Grades

Make-up exercises for either absence, or poor performance, will not be provided as an option. If you have a valid excuse from the registrar for missing any exercises or other items, and extension will be provided, but no longer than when marks for an item are returned to the class. Missed items without a valid excuse from the registrar will result in a mark of zero.

Part 4: Course Policies

1. **Complete tasks and exercises as scheduled.** If you are unable to do so for an approved reason, please see me as soon as possible.
2. **Come to all classes.** There is considerable material presented in class that is not provided on lecture slides or in other material.
3. **Take notes.** Normally notes/slides will not be provided. This is an immersive course, not one built on repeating/regurgitating materials provided.
4. Login to Moodle and Teams **at least once per day** to check for news, participate in discussion forums, and/or check for updated material. From time-to-time I will update course slides if errors are found, or if I find ancillary material that will help you to study.
5. Act accordingly in class.
6. Visit/contact the instructor or TA when you do not know the material. It is impossible to help if you do not visit/contact.

Part 5: University Policies

University policies are available in the Acadia University Academic Calendar or through the Registrar's website:

<https://registrar.acadiau.ca/welcometotheregistrarsoffice.html>

Covid-19

All students must read and adhere to covid-19 policies found here:

<https://www2.acadiau.ca/covid-19.html>

Declaration of Cause

University policies on missing classes, etc. can be found here:

<https://registrar.acadiau.ca/RecordsandOtherRequests.html>

Last Drop Day

Last day to drop a course and receive a "W". Please check the Acadia University calendar dates, which are available here:

<https://registrar.acadiau.ca/AcademicCalendars.html>

Accommodations

If you are a student with documentation for accommodations or if you anticipate needing supports or accommodations, please contact Marissa McIsaac, Accessibility Resource Facilitator at 902-585-1520, or Emily Duffett, Accessibility Officer, disability.access@acadiau.ca. Accessible Learning Services is in Rhodes Hall, rooms 111-115.

Commitment to Integrity

Cheating in the lecture and/or lab, including plagiarism, will not be tolerated. Please read the appropriate sections of the current Acadia University Academic Calendar.

<https://registrar.acadiau.ca/AcademicCalendars.html>

If caught cheating you will automatically receive a grade of zero on the exercise/exam, and your name will be submitted to the registrar. If this is not the first occurrence you will either receive a mark of zero for the course (2nd occurrence) or be expelled from the University (3rd occurrence).

Equity, Diversity and Inclusion

EDI Acadia University is committed to becoming a culturally safe and anti-oppressive community. This can only be achieved where there are simultaneous efforts to eliminate all forms of discrimination and harassment from our campus community, including the elimination of all discrimination, harassment and violence based on one's identity, including but not limited to, gender, race, class, ethnicity, sexual orientation, disability, gender identity, gender expression, and Indigeneity.

The Equity, Diversity and Inclusion Officer is available to students, staff, and faculty. The fundamental objective of the Equity Office is to prevent discrimination, sexual harassment, and personal harassment from occurring, in part by managing Acadia's Policy Against Harassment and Discrimination. For more information, as well as for resources for students who believe they may have experienced or witnessed discrimination, sexual harassment, or personal harassment please contact Acadia's Equity, Diversity and Inclusion Officer, Polly Leonard, MSW, RSW (she/her/hers) at equity@acadiau.ca, and check out the [website](#).

Copywrite

Information on copy-write and course content from Acadia University is available through the Vaughan Memorial Library:

<http://libguides.acadiau.ca/c.php?g=433650&p=5027078>

Course Content Intellectual Property

The spoken and written and recorded course content (including the syllabus, handouts, lectures, presentations, labs, assignments, lecture videos, quizzes, tests, and exams) are the intellectual property of the instructor and may only be copied for personal use. Sharing these materials or uploading them where they may be accessed by others is a violation of copyright. If you wish to make audio, video, or photographic recordings in class, you must first obtain the consent of the instructor and of any other persons (e.g. guest speakers, other students) who may be captured in such recordings. In the case of personal use by students with disabilities, the instructor's consent shall not be unreasonably withheld. Students with disabilities who wish to request accommodation should contact Accessible Learning.