Data Science for a Living Planet
BIOL 3383

Instructor: Keegan Lutek (he/him)
Hear my name
Office: MacAskill Hall 108
Office hours: Wed 10:30-11:30 & 12:30-14:00
Thurs 12:30-13:00
E-mail: keegan.lutek@acadiau.ca

Lecture: Thurs 13:00 – 15:50
CAR 207
In-person

Evaluation	Percentage	Date
Participation	10%	Assessed by instructor & random checks
Learning Log	5%	Sept. 17, Sept. 24, Oct. 1, Oct. 29, Nov. 12 and Nov. 26
Exercise 1: Data Wrangling	20%	Oct. 8
Exercise 2: Working with Maps	20%	Nov. 5
Exercise 3: Images & Statistical Analysis	15%	Nov. 19
Workgroup Presentation & Code	20%	Dec. 3
Final Exam (take home)	10%	Dec. 20

Part 1: Course Information

Course Description

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

Prerequisite(s):

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

Course Materials

There is no text to buy; there is an online textbook. There are also many useful FREE texts and resources though that will be outlined in class. All of these materials will be made accessible on Moodle.

Course Requirements

1. Access to Moodle. Moodle is the central portal for all course material.
2. Computer with R and RStudio installed.
3. Attendance at all lectures.
4. Enthusiasm.

Course Structure

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

Lecture Recordings

Students may not create audio or video recordings without written permission from the instructor. Permission for such recordings normally 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, 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 and via presentation.

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: Course Plan

The instructor reserves the right to amend the course plan with reasonable notice, and in consultation with the class.

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

Students must indicate what was accomplished most weeks through a short survey/questionnaire. Be honest with your responses as this is the best way for us to gauge course speed!

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. 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!

Lecture Schedule:

Date	Topic(s)	Synopsis
Sept. 10	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
Sept. 17	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
Sept. 17	Learning Log 1 due	Log 1 of 6; collectively worth 5%
Sept. 24	10: Tidyverse – lubridate 11: Summarizing data 12: Working with tables	Working with dates and times Summarizing data using descriptive statistics Making tables in kable, gt, dt Saving tables as files
Sept. 24	Learning Log 2 due	Log 2 of 6; collectively worth 5%
Oct. 1	13: Tidyverse – ggplot2 14: Designing plots 15: Multiple elements	Introduction, aesthetics, basic plots (histograms, dot plots, scatter plot, boxplots), labels, saving plots Practicing plots, layering plots, adding lines and dots Colors, groups, fills, factors, themes, facets
Oct. 1	Learning Log 3 due	Log 3 of 6; collectively worth 5%
Oct. 8	16: Scales 17: Legends, annotations 18: Equations, text manipulation	Scales, dates and times Legends, annotations, adding images Equations, advanced text
Oct. 8	Exercise 1: Data Wrangling due	worth 20%
Oct. 15	*NO CLASS; THANKSGIVING BREAK.*	
Oct. 22	19: Reproducible research Reproducible workflows 20: Quarto 21: Advanced RMarkdown/Quarto	RMarkdown syntax, formatting, code chunks, inline text Expanding RMarkdown RMarkdown equations (Latex), figures, tables, images and links
Oct. 29	22: Mapping 23: Creating basemaps 24: Annotating maps 25: Exporting maps	
Oct. 29	Learning Log 4 due	Log 4 of 6; collectively worth 5%
Nov. 5	26: Image Processing 27: Working with image types 28: Publication quality images	It's Magick!
Nov. 5	Exercise 2: Working with Maps due	worth 20%
Nov. 12	29: Statistical Modeling 30: ggeffects or easystats 31: Fitting a linear regression 32: GLMs	What is it? New packages! Review – it's a model! Generalizing and building models
Nov. 12	Learning Log 5 due	Log 5 of 6; collectively worth 5%
Nov. 19	33: Model validation 34: Model selection	Residual plots Selecting the "best" model

	35: Model equations 36: Model testing	Writing model equations Testing effects
Nov. 19	Exercise 3: Images and Statistical Analysis due	worth 15%
Nov. 26	37: Workgroup Projects	Time to finalize projects
Nov. 26	Learning Log 6 due	Log 6 of 6; collectively worth 5%
Dec. 3	38: Workgroup Project Presentations 39: Workgroup Project Reports	worth 10% worth 10%
Dec. 20	Take Home Exam (Final Report) due	worth 10%

Part 3: Assessment and Grading

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.

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

There are several ways to enhance participation marks. There will be random opportunities to show what you know. Additionally, participation will be assessed in-class by 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 required at regular intervals (see course schedule) and marked for completion alone.

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 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 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

Equity, Diversity and Inclusion

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 policy against harassment and discrimination, and resources for students who believe they may have experienced, or witnessed, discrimination or harassment, are available here: https://www2.acadiau.ca/student-life/equity-judicial/equity.html

Last Drop Day

Last day to drop a course and receive a "W" is November 27. Please check the Acadia University calendar dates, which are available here: https://registrar.acadiau.ca/AcademicCalendars.html

Inform Your Instructor of Accommodations

Acadia University is dedicated to improving access to campus life for all students with disabilities. While we attempt to ensure that all courses are accessible, we recognize that there are barriers that need to be addressed on an individual basis. Students who require accommodations to complete coursework or otherwise fully participate in class should contact Accessible Learning Services directly as soon as possible. https://www2.acadiau.ca/student-life/accessiblelearning.html

Acadia's Policy for Students with Disabilities can be found here: https://senate.acadiau.ca/tl_files/sites/senate/Senate%20Policies/Jan%202015%20Revised%20Policy%20Regarding%20Support%20for%20Students%20with%20Disabilities.pdf

The Use of Animals in Teaching and Research

The use of animals in teaching and research at Acadia University is done in accordance with guidelines on the care and use of animals published by the Canadian Council on Animal Care (CCAC). For more information on the CCAC, please visit their website at http://www.ccac.ca

Commitment to Integrity

It is standard practice in Biology to check exams and assignments for cheating and plagiarism. Cheating in the class 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

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

The spoken and written course content (including the syllabus, handouts, lectures, presentations, labs, assignments, 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. The instructor's consent shall not be unreasonably withheld.

Use of artificial intelligence tools by students to assist with their learning is allowed only if: prior permission has been given by the course professor/instructor, and the student properly credits/cites how the AI technology was used in their work. AI must be used in an ethical and responsible manner as a positive learning tool. AI may not be used in ways that violate Acadia's Academic Integrity Policy. Students are reminded that plagiarism, or other forms of cheating, are both types of academic dishonesty.

Share the Air Policy

In consideration of the difficulties that exposure to scented products causes individuals with sensitivities and allergies, all students, faculty, staff, employees of any companies working on university property, visitors, and guests of Acadia University, or of members of the University community are asked to refrain from wearing scented personal care products such as perfumes / aftershave, lotions, hair spray and deodorant. In addition, users of tobacco and cannabis are asked to be aware that odours associated with product use may impact individuals with sensitivities and allergies. Acadia University in consultation with its contracted cleaning staff, have agreed to use products that do not leave residual odors that may cause difficulties for individuals with sensitivities and allergies.
https://hr.acadiau.ca/employment/university-policies.html

Part 6: Program Learning Outcomes

Foundations of knowledge	**Explanation**	**Course specific examples**	**Proficiency** 1-Introduction 2-Reinforcement 3-Proficient
Scientific method, inquiry and hypothesis testing	Find, understand and apply information from the literature; understand how to use the scientific method to examine problems from different perspectives	Exercises will be based around the scientific method	2
Historical concepts and contributions by important figures	Explain foundational concepts in biology, Two-eyed Seeing, and ethical implications of scientific discoveries	N/A	N/A
Biodiversity and ecology	Understand the genetic, taxonomic and ecosystem levels of biodiversity; focus on SW Nova including the Acadian	N/A	N/A

	Forest and Bay of Fundy ecosystems		
Genetics and evolution	Understand the chemical basis of heredity, genetics and genomics; integrate concepts across disciplines to understand evolution	N/A	N/A
Human and environmental health	Understand form and function in health and disease within a One Health framework, integrating human and environmental health	N/A	N/A
Lab and field skills	**Explanation**	**Course specific examples**	**Proficiency** 1-Introduction 2-Reinforcement 3-Proficient
Experimental design	Gain experience in applying the scientific method	N/A	N/A
Safety	Work safely and productively in lab and field settings	N/A	N/A
Lab skills	Gain experience with basic and advanced lab techniques and understand their application in research, health science and industry	N/A	N/A
Field skills	Gain experience in basic and advanced field skills and understand their application in ecology, conservation biology and environmental change	N/A	N/A
Data acquisition, analysis and interpretations	Collect data, present results both qualitatively and quantitatively, and interpret outcomes in light of the literature	This course works through the cleaning, analysis and presentation of datasets and outlines best practices in data collection	3
Statistical analysis	Use R and or other programs to analyze biological data	Use of R is essential throughout the course for data manipulation, presentation and analysis	3

Professional skills	**Explanation**	**Course specific examples**	**Proficiency** 1-Introduction 2-Reinforcement 3-Proficient
Ethical practices	Demonstrate ethical conduct, apply principles of academic integrity, and understand the principles of EDI in science	All assessments require application of academic integrity	3
Collaboration and group work	Work effectively in groups within and across disciplines	Group assessments and exercises used throughout the course	3
Critical thinking	Analyze and evaluate information to make science-based decisions	All assessments will require critical thinking	2
Computer proficiency	Use common and discipline-specific software	R is essential throughout this course	3
Scientific communication	Communicate science effectively to both scientific and general audiences	Data presentation in various formats is practiced and assessed	2