BIOL 343 – Syllabus

Advanced Data Analysis for Biologists

Course Information

Fall 2024 Semester
3.0 Credits
In-Person
Pre-requisites: BIOL 343
Lectures: Tuesdays, 8:30 – 10:30, MACKINTOSH-CORRY RM D201
Tutorials: Fridays, 10:30 – 11:30, MACKINTOSH-CORRY RM D201

Instructor

Name: Dr. Robert I. Colautti (he/him)
Office Address: 4325 BioSciences, 116 Barrie St.
Office Hours: Tuesdays, 10:30 – 11:30 (Mac-Corry D201) + 11:30 – 12:20 (BioSci 4325)
Phone: 613-533-2353; Please use only during office hours.
Email: robert.colautti@queensu.ca
Website: https://ecoevogeno.org/
Twitter: @ColauttiLab

About Me: My teaching takes a student-centered approach that is supportive of diverse learners. I set high expectations and provide extensive resources to encourage students to continue to learn beyond the constraining demands of a university course. I try to emphasize a growth mindset that is focused on effort, personal development, and quality of work, rather than mastery or excellence. I try to teach students how to recover and learn from failure, which I believe is essential for a successful career in any discipline. My teaching philosophy draws on two decades of research and mentorship experience, resulting in primary research in top journals (e.g. Science, PNAS, PRSB) and dozens of former students employed in the public and private sectors or continuing to pursue advanced degrees. I have a broader perspective than many biologists, drawing from an MSc in aquatic ecology from the University of Windsor and a PhD in quantitative genetics at the University of Toronto, followed by bioinformatics and computational biology research at Duke University (North Carolina), the University of British Columbia, and the University of Toronto. Fascinated by genetic engineering and genome sequencing to disease epidemiology and the global biodiversity crisis, the skilled analysis of biological data is foundational to discovery and innovation in biological systems. The result of this effort is the course material for BIOL 343, BIOL 432, BIOL 860 and BIOL 812.

Equity, Diversity, and Inclusivity Statement

Equity and diversity are central to our educational mission and standards of excellence in this course and at Queen's University. It is critical that we work together to dismantle direct, indirect, and systemic discrimination that still exists within our institutional structures, policies and practices -- and in our community. These take many forms and work to differentially advantage and disadvantage persons across social identities defined by race, ethnicity, disability, gender identity, sexual orientation, faith and socioeconomic status, among other factors. As students and educators, we all have important roles to play to identify and address systemic discrimination for the benefit of science and society.

Land Acknowledgement

As a descendant of uninvited colonists, I feel tremendously privileged to live, learn, work, and play on these lands. As Queen's University is situated on traditional lands of the Anishinaabe and Haudenosaunee, I invite you all to be mindful with me about the many lessons we learn while on these lands, and how we might apply our newfound skills and knowledge for the benefit of all.

Even as a well-established scientist, I admit I have much to learn from the teachings and traditions of the Anishinaabe and Haudenosaunee, who have lived on these lands since time immemorial. While researching some of this history, I was moved to learn about the Seven Grandfathers in the Anishinaabe tradition, which, as I understand, demonstrate principles for living a "good" life. These include Dabaadendiziwin (humility/compassion/patience), Gwayakwaadiziwin (bravery to be honest), Minaadendamowin (respect for all creation), Nibwaakaawin (wisdom/knowledge to help people) and Zaagi'idiwin (unconditional love given and received). It is worth reflecting on how well these principles resonate with other cultures and traditions around the world, suggesting a deep truth

It can be difficult and even overwhelming as a student to struggle through the stresses and demands of a university degree and life more generally. When you feel this way, I encourage you to learn or return to these Indigenous teachings, and/or teachings from your own cultural traditions, to recall what really matters in life, and to let these insights guide you through difficult decisions

Important University Dates

Key dates (first day of class, tuition due date, last day to add/drop courses) are important to your academic success. Please find them at the Important Dates website.

“In some cases the lack of background doing this kind of assignments required much more effort and time from my side which was sometimes quite frustrating even though the answer was not really complicated.”

“The assignments sometimes were very quick and sometimes took a long time.”

“The weekly assignments were also extremely time consuming, often taking up to 7-10 hours for each.”

Lesson Learned: This course requires a lot of dedicated focus time, problem solving, and practice. This is a very different kind of effort than most biology courses and there are not many shortcuts.

To accommodate variability in learning, we will make the main content available in complementary forms including two original textbooks and prerecorded videos with annotated scripts. The textbooks lean heavily on a tutorial style, with step-by-step instructions that are reiterated in the online videos. The videos and textbook are designed to be complementary, with overlap emphasizing important skills and techniques.

You will make mistakes, both in coding and in learning about coding. Everyone makes mistakes, and coding is particularly prone to error, especially when there are distractions. We will use these opportunities to demonstrate how to troubleshoot errors by carefully reading the warning messages and running smaller subsets of code to identify where the problem is arising. **Learning how to troubleshoot code is perhaps the most important skills you can learn in this course.** But it requires a different kind of mindset (see 'For Students' Section).

The entire teaching team (instructor + teaching assistants) is committed to establishing and maintaining a healthy and inclusive learning environment. We recognize that mistakes and errors are an important part of your learning process. We respect and value students who are not afraid to take risks or try things that might be 'wrong'. Above all, we value students who are not afraid to fail. We will use frequent assignments and testing to limit the impact of mistakes on your final grade. We will provide timely feedback – usually within two to three weeks. This represents a very large time and energy investment from the teaching team. We do it to help you learn from your mistakes, focus on learning, and succeed on future assignments.

We will communicate twice per week during lecture and tutorial. Lectures will cover only part of the assigned readings, so that there will be ample time available for questions or assistance.

For Students

It is expected that you will attend weekly lectures and tutorials, though we understand this may not be possible for everyone, all the time, particularly in the post-COVID era. Therefore, everything you need to succeed in this course will also be available online.

You are expected to bring a laptop capable of running Windows, MacOS, or Linux programs, and you must be able to access Queen's wireless network during lecture and tutorial sessions (see also "Technology Requirements"). Be sure to charge your laptop battery as there may not be enough plugs for everyone who needs one. When working in class or following recorded lectures, **you will code along in real time**. The only way to effectively learn to program is to practice, and you are expected to practice as much as possible!

You are expected to complete the assigned readings each week, write down any questions that you have, and complete the online quizzes before the posted deadline. Then, review your answers to check for sources of confusion. You are expected to organize your thoughts into questions to ask during class. Please do not email questions that can be addressed during lecture or tutorial/office hours. If you aren't comfortable asking questions verbally, you may hand in written questions to the instructor or TA during class or tutorial. These steps help to ensure that you are organized and prepared before attending lectures and tutorials.

Weekly assignments are also submitted online and generally due within 72-48 hours of being posted online. Working through the assigned chapters and quizzes will prepare you for these short deadlines, which are essential to reinforce and build on what you have learned each week. Except where explicitly stated, you must complete quizzes and assignments alone, without communicating with other students. **Any attempts to communicate about quiz or assignment answers will be treated as a breach of academic integrity.** Plan to devote ~5 hours to learning the lecture material and up to 10 hours to complete the assignments.

You are expected to check the course website regularly (or use alerts) to keep track of deadlines. **Late assignments are scored as zero** (but see below regarding accommodations).

Any questions or concerns about the course should be raised in lecture or tutorial, or privately during weekly office hours (no appointment needed). Email is generally not an effective tool for course material, and questions that can be addressed in person will not receive an email response. However, email is encouraged for urgent issues, (e.g. medical or other personal emergencies, broken/incorrect website links, and other time-sensitive issues).

Course Learning Outcomes

Students completing this course shall be able to:

1. Identify different data types to enable coding for visualization and analysis.
2. Translate real-world observations into appropriate data types to produce visualization and analysis.
3. Reflect on how positionality may bias one's experimental design and data interpretation by exploring historical and contemporary biases on scientific progress.
4. Contrast the use of fixed vs random effects and linear vs generalized linear models to ensure appropriate interpretation of statistical output for real-world questions.
5. Simulate data relevant to sustainable development goals to explore assumptions of statistical models.
6. Develop a robust strategy for quality assurance and quality control to assess the reliability of statistical models.
7. Write clean and coherent code in R markdown to create reports with professional formatting and an analysis that is a

Weekly Quizzes (20%)

- Weekly quizzes are completed before each lecture and are graded on a pass/fail basis. You will receive a full grade if you complete the quizzes on time.
- These quizzes are self-assessments of the weekly assigned readings to support learning of the background knowledge needed to complete the weekly assignments.

Weekly Assignments (30%)

- Weekly assignments are assigned in each lecture and due by the end of class.
- These assignments reinforce coding knowledge learned in assigned readings and support development of coding skills that are tested in the final exam.
- A mix of group and individual projects will be assigned. Students who are absent or unable to complete group work may submit individual assignments to avoid grade penalties.
- All students receive a grace period (an extension of up to 24 hours) without the need to make a request through the Academic Considerations Portal. You do not need to send an email or explain in person; simply take the time if you need it.
- Longer extensions are discouraged because of the cumulative nature of the course. Delays will prevent learning of new content and the frequency of assignments can quickly become overwhelming when deadlines overlap. For students who have the need for longer extensions, please submit a request through the Academic Considerations Portal (see below).

Participation & Peer Review (10%)

- The participation and peer review grade has two main components.
- The first part of the grade is assigned by the instructor and TAs, and it is designed to motivate attendance and active participation in lectures and tutorials.
- The second part of the grade is based on peer evaluation forms, following criteria: in

converting your numerical course average to a letter grade according to Queen's Official Grade Conversion Scale:

Queen's Official Grade Conversion Scale

The second part of the course applies to the concepts and tools from the R Crash Course to analyze biological data

- Mistakes are encouraged in readings and quizzes because learning to identify and deal

intellectual property, unauthorized collaboration, failure to abide by academic rules, departure from the core values of academic integrity, and falsification, and are antithetical to the development of an academic community at Queen's. Given the seriousness of these matters, actions which contravene the regulation on academic integrity carry sanctions appropriate to the severity of the departure that can range from a warning or the loss of grades on an assignment to the failure of a course to a requirement to withdraw from the university.

Plagiarism is a form of cheating and includes copying code written by students. There is no 'right answer' for the assignments in this class – there are often many potential coding solutions. You will also develop your own coding style, which will make it obvious when code has been copied. To avoid potential for plagiarism, ALWAYS COMPLETE ASSIGNMENTS ON YOUR OWN. As a bonus, you will learn to code better. On the other hand, it is completely fine to ask others to help you troubleshoot an error message or help you figure out why your code isn't working properly. If you become aware of anyone trying to share or solicit code for the assignments, please point them to this passage and inform the teaching team immediately.

Queen's Student Academic Success Services (SASS) offers a self-directed, online academic integrity module which we encourage all students to take which will help with:

-

•