

BIOL 343

Advanced Data Analysis for Biologists

Instructor: Dr. Robert Colautti, 4325 BioSciences, robert.colautti@queensu.ca

Lectures: MON 12:30-1:30, WED 11:30-12:30, THR 13:30-14:30; See OnQ for more detail

Tutorial: See SOLUS & OnQ for your scheduled Tutorial Time

Prerequisites: BIOL 243

Textbook: **Required:** *Essential Biostatistics. A Nonmathematical Approach* by Motulsky (Oxford University Press)

Suggested:
R for Data Science by Hadley Wickham

Assessments: 70% Weekly Assignments & Pop Quizzes
30% In-class midterm exams (2x 15%)

Course Learning Outcomes:

Students completing this course shall be able to:

1. Produce graphs and figures to visualize and explore biological data
2. Choose and apply appropriate statistical models to test scientific hypotheses
3. Interpret output of statistical tests and translate to biological inference
4. Translate real-world observations into data appropriate for analysis
5. Compare and contrast common data types in statistical models and computer science
6. Navigate R Studio and write analysis code in R that is **OPEN** and **REPRODUCIBLE**

Course Outline

This course builds on the foundations of BIOL 243 and introduces students to the management, visualization, and analysis of biological data using the R coding environment. Emphasis is on hands-on/active coding relevant to students who are interested in pursuing a career in data science, or analyzing original data sets in field courses, honours theses, and graduate research.

Readings combined with lectures/tutorials over all fundamental statistics, including frequency distributions, skewness, measures of central tendency, nominal, ordinal, interval measurements, summary measures of variation, probability theory, normal distribution, sampling theory, confidence intervals, tests of significance (null hypothesis, t-test, one-tail vs. two-tail tests), analyses of proportions and categories, analyses for group means, correlation, regression, and linear models.

Philosophy

This course introduces basic computational skills and applications for the analysis of biological data. In contrast to many biology courses, coding and quantitative skills learned in this course must be developed through extensive practice, trial and failure. The philosophy of this course is that you won't learn to code and analyze data by reading and memorization – only through extensive application and practice. Some fundamental concepts are explored in the course textbook, which is tested in two in-class midterm exams. Regular assignments and quizzes are designed to reinforce skill development and problem-solving. In addition, you are **STRONGLY ENCOURAGED** to find opportunities to practice coding wherever possible. For example, write code instead of messing around with spreadsheet program like Microsoft Excel, use [R Markdown](#) instead of Microsoft Word to render final reports for your other classes. It will take longer at first, but it will save a lot of time in the long run. Finally, be prepared to get frustrated – you will make

many errors and most of your coding time will be debugging and searching for answers on the internet. It is important to know that **this is COMPLETELY NORMAL** and self-directed research to solve coding problems is perhaps the most crucial skill you will learn in this course.

Assistance

This course has no lectures in the traditional sense. Weekly 'lectures' include short introductions to the material, with most of your time devoted to independent study and hands-on tutorial sessions where you interact with your group and your TA. Outside of synchronous meetings (i.e. lecture and help sessions), a very common and useful approach to solving errors or other problems is to search Google or Stack Overflow. Often, simply copying and pasting an error into an online search will produce a helpful link. Very often you can just type 'How do I X in R' (or the R package name like ggplot2, dplyr) into Google and look for links to similar questions answered on the Stack Overflow website.

Grade Assessment

Weekly Assignments:

Individual assignments account for a relatively small portion of your final mark, so don't worry if you don't do well on a few of them. They are designed to motivate you to practice coding and reinforce learning of the lecture and tutorial material. For this reason, it is important that you begin the assignment as soon as possible and work on it independently. Assignment due dates are posted on OnQ, and a **late assignment receives 0%**. However, students who require accommodation for any reason will receive a 5-day grace period without penalty, **as long as official requests are submitted to the appropriate website** (see "Accommodations" and "Extenuating Circumstances", below).

Tests:

There are two in-class midterms in this course, and periodic quizzes to complete via OnQ. The midterms ensure that you are following (and understanding) the readings in the textbook, while the quizzes assess both the textbook readings and tutorials. **There are no make-up quizzes or midterms**, however students requiring accommodation for any reason may miss up to two quizzes, and missed midterms will be pro-rated based on other grading items, **as long as official requests are submitted to the appropriate website** (see "Accommodations" and "Extenuating Circumstances", below).

Other Considerations

Grievances:

It is the student's responsibility to raise any concerns about grading errors within 5 business days of receiving a grade on any quiz or assignment. Requests must first be made in writing via email to the TA, and must include a very clear explanation of the reason for concern. If the situation is not resolved, a follow-up request can be made to the course coordinator to re-grade the assignment, which may result in a lower or higher grade. Other grade adjustments will NOT BE CONSIDERED without prior approval of accommodations or extenuating circumstances, as follows:

Accommodations:

Queen's University is committed to achieving full accessibility for people with disabilities. Part of this commitment includes arranging academic accommodations for students with disabilities to ensure they have an equitable opportunity to participate in all of their academic activities. If you are a student with a disability and think you may need accommodations, you are strongly encouraged to contact the Queen's Student Accessibility Services (QSAS) and register as early as possible. For more information, including deadlines and application forms, please visit the QSAS website:

<http://www.queensu.ca/studentwellness/accessibilityservices/>

1. Submit a request
2. Email the course co-ordinator and TAs when a request has been submitted

Extenuating circumstances:

For excused absences due to short-term extenuating circumstances (e.g. flu, injury, short personal leave), please submit a formal Academic Consideration Request Portal (ACRP) on the Arts & Sciences accommodations website: <https://www.queensu.ca/artsci/accommodations>

1. Submit a request
2. Email the course co-ordinator and TAs when a request has been submitted

Academic Integrity (<http://www.academicintegrity.org>):

Queen's students, faculty, administrators and staff all have responsibilities for supporting and upholding the fundamental values of academic integrity. Academic integrity is constituted by the five core fundamental values of honesty, trust, fairness, respect and responsibility and by the quality of courage. These values and qualities are central to the building, nurturing and sustaining of an academic community in which all members of the community will thrive. Adherence to the values expressed through academic integrity forms a foundation for the "freedom of inquiry and exchange of ideas" essential to the intellectual life of the University. Students are responsible for familiarizing themselves with and adhering to the regulations concerning academic integrity. General information on academic integrity is available at **Academic Integrity @ Queen's University**, along with Faculty or School specific information. Departures from academic integrity include, but are not limited to, plagiarism, use of unauthorized materials, facilitation, forgery and falsification. Actions which contravene the regulation on academic integrity carry sanctions that can range from a warning, to loss of grades on an assignment, to failure of a course, to requirement to withdraw from the university

Turnitin:

Queen's University has partnered with the third-party application Turnitin to help maintain our standards of excellence in academic integrity. Turnitin is a suite of tools that provide instructors with information about the authenticity of submitted work and facilitates the process of grading. Submitted files are compared against an extensive database of content, and Turnitin produces a similarity report and a similarity score for each assignment. A similarity score is the percentage of a document that is similar to content held within the database. Turnitin does not determine if an instance of plagiarism has occurred. Instead, it gives instructors the information they need to determine the authenticity of work as a part of a larger process.

Terms: http://turnitin.com/en_us/about-us/privacy.

Privacy Policy: http://turnitin.com/en_us/about-us/privacy#policy

Copyright:

The material in this course is the property of the instructor unless otherwise stated. Third party copyrighted materials (such as book chapters and articles) have either been licensed for use in this course or fall under an exception or limitation in Canadian Copyright law. Copying this material for distribution (e.g. uploading material to a commercial third-party website) can lead to a violation of Copyright law. More information: <http://library.queensu.ca/copyright/students>