DALHOUSIE UNIVERSITY

Faculty of Science Course Syllabus (Section A)
Department of Biology
BIOL4062
Analysis of Biological Data
Fall 2022

Dalhousie University is located in Mi'kma'ki, the ancestral and unceded territory of the Mi'kmaq. We are all Treaty people.

We acknowledge the histories, contributions, and legacies of the African Nova Scotian people and communities who have been here for over 400 years.

Instructors: *Hal Whitehead* *hwhitehe@Dal.ca* *LSC3076*

Derek Tittensor *derek.tittensor@dal.ca* *LSC7060*

Office hours: *Tues, Thurs 14:35-15:00 (HW first half of course; DT second half); if set time does not work email instructors to set up time.*

Lectures: *Tues, Thurs 13:05-14:25* *LSC C234*

Laboratories: -

Tutorials: *TBA*

Course delivery: *In-person*

Course Description

Biologists are increasingly using quantitative techniques to analyze larger and larger data sets. A command of the available analytical techniques is an important part of the set of skills which are expected of a trained biologist, especially those working in the broad area of ecology. The class will introduce techniques available for the analysis of biological data, including correlation, regression, multivariate, Bayesian and hierarchical methods. Emphasis will be on the practical use and abuse of these techniques rather than derivations or mathematical formulae; the idea being that students will learn a suite of approaches that will enable them to select suitable techniques for multiple data types. Students will explore real and realistic data sets, as well as simulated data, and learn good practices for hands-on application of approaches.

Course Prerequisites

Prerequisites: STAT 2080.03 or ECON 2280.03 or MATH 2080.03, AND BIOL 3872.03. Students should have familiarity with R (preferred), or some other statistical, command-line, programming language (e.g. Python, MATLAB; but support is only in place for R).

Course Exclusion

None

Learning Objectives
Knowledge of methods of analyzing biological data

Course Materials
Course Brightspace page

Course Assessment

Assessment	Weight (% of final grade)	Date
Assignments		
Report of analysis of data set 1a	***10%***	***7 Oct***
Report of analysis of data set 1b	***10%***	***11 Oct***
Report of analysis of data set 1c	***10%***	***18 Oct***
Report of analysis of data set 1d	***10%***	***25 Oct***
Report of analysis of data set 1e	***10%***	***15 Nov***
Description of data set 2 and proposed analysis	***5%***	***13 Oct***
Report of analysis of data set 2	***30%***	***8 Dec***
Tests/quizzes		
In class test (80 min)	***15%***	***6 Dec***

Other course requirements
None

Conversion of numerical grades to Final Letter Grades follows the Dalhousie Common Grade Scale

A+ (90-100)	**B+** (77-79)	**C+** (65-69)	**D** (50-54)
A (85-89)	**B** (73-76)	**C** (60-64)	**F** (<50)
A- (80-84)	**B-** (70-72)	**C-** (55-59)	

Course Policies on Missed or Late Academic Requirements

Papers must be sent in through Brightspace by 16:30 on due date

10% off for each weekday late, without medical or other legitimate excuse (use the *Student Declaration of Absence* form; up to 3 times in course)

Just put your B00.... number on paper

Single-spaced, 12pt or more font

Page limits do not include reference lists, figures, tables, etc.

If an assignment deadline is unavoidably missed, students may submit it later

Course Policies related to Academic Integrity

No collaboration on assignments

Plagiarism detection software may be used

Course Content

Inference in Biology
R--a refresher
Correlation, and linear regression
Introduction to multivariate analysis and multivariate distances, association measures
Principal component analysis
Multivariate analysis with grouped units or grouped variables
Multivariate analysis of association matrices; Cluster analysis
Categorical Data: Contingency Tables and Log-Linear Models
Introduction to likelihood
Multiple linear regression and path analysis
Generalized linear models
Logistic regression
Bayesian data analysis using STAN
Hierarchical models
Simulating data to check your models and cross-validation
Diversity analysis, sampling effort, and bootstrapping
Analysing temporal data
Spatial data analysis and good modelling practice