BIO380
Biological Data Analysis
Draft Course Syllabus
Term Year

Class Location & Time	Tue, 11:10 AM – 12:25 PM (Lecture) Wed, 1:50 PM – 4:50 PM (Lab)
Instructor	David Murray-Stoker
Office Location	DV2014
Office Hours	Mon/Wed, 11:00 AM – 1:00 PM Tue/Thu, 1:00 PM – 3:00 PM
E-mail Address	d.stoker@mail.utoronto.ca
Course Web Site	TBD
Teaching Assistant	TBD
E-mail Address	TBD

Course Overview
Students will learn the fundamentals of biological data analysis through a combination of lectures, computer labs using R, and group projects. The course will begin with a foundation in hypothesis testing, experimental design, and interpreting results. It will then take students from correlation, linear regression, and multiple regression to analysis of variance, mixed-effect models, and structural equation modeling. Students will apply their knowledge of these concepts and analyze biological data. Additionally, students will build a proficiency in using R for data management and analysis. No prerequisite programming experience is required.

Prerequisites: STAT215H5 or BIO259H5

Learning Objectives
As you participate and engage in the course, you will be able to:
- Build a strong foundation in statistical concepts and best practices.
- Develop biological and statistical reasoning.
- Analyze data and interpret results from case studies.
- Complete a research project that addresses a research question or hypothesis using statistical analyses.
- Communicate the relevance of the statistical methods and results from the project.

Course Instructors
Primary Instructor: David Murray-Stoker, Ph.D. Candidate (he/they)
Email: d.stoker@mail.utoronto.ca
Office Location: DV2020

Teaching Assistant: TBD
Email: TBD
Office Location: TBD

Student Hours
Join your instructors and your peers to discuss material being covered in class, raise any questions or concerns you might have, and any other topics that will help you and your learning in the course. You are welcome to join these student hours even if you do not have a question, as listening to the conversation can still be helpful (and might raise a question for you to ask).

David Murray-Stoker	Mon/Wed, 11:00 AM – 1:00 PM Tue/Thu, 1:00 PM – 3:00 PM If these times do not work with your schedule, please email me so we can arrange a time.
Teaching Assistant	TBD

Course Resources

Required Readings
Chapter readings will be uploaded to the course website at the start of the term. **Reading the posted chapters before attending and engaging in class is essential.**

In addition to the chapter readings, these books are great resources:
Hector, Andy. 2021. The New Statistics with R: An Introduction for Biologists. Oxford University Press.
Touchoun, J. C. 2021. Applied Statistics with R: A Practical Guide for the Life Sciences. Oxford University Press.

Supplemental Readings
If you would like to supplement the chapter readings with a textbook, I recommend the following:
Quinn, G. P., and M. J. Keough. 2002. Experimental Design and Data Analysis for Biologists. Cambridge University Press.

Reading Assignments
Papers for the assigned readings will be posted to the course website at the beginning of the term. Each assigned reading is associated with a specific tutorial and class discussion (see the course schedule below).

Technology
You will need access to a device (e.g., laptop, tablet) with internet access for lecture activities, email correspondence, using the course website, and completing assignments. You are also encouraged to use a laptop or equivalent device with Microsoft Office installed (software subscription included with your university email) or use Google Docs to complete the coursework (described below). Students may also borrow laptops from the UTM library. **If you do not have reliable access to the internet and/or a suitable device, please contact me so we can find a positive solution.**

R Statistical Software
R is a flexible, powerful, open-source program for statistical analysis that runs on all operating systems. R should be downloaded and installed before the first lecture. You can download R by following this link: https://cran.r-project.org/. I will provide an instructional video on Quercus to demonstrate the installation process.

RStudio
RStudio is a graphical user interface that helps to write code and analyze data. RStudio also allows for easy writing of scripts R Markdown files, which will be used to illustrate some concepts in lecture. You can download RStudio by following this link: https://posit.co/download/rstudio-desktop/ . I will provide an instructional video on Quercus to demonstrate the installation process and user interface.

Teaching Methods
BIO380 is an active learning class where you are part of the learning process. You can really only learn statistics and how to use R by doing the work and doing it consistently. You are therefore expected to come to lecture and lab ready to engage in the material, ask questions, participate in activities, and collaborate with your peers.

Coursework
You should expect to complete 10-12 hours of study and work each week for this course, including time spent in lecture. In other words, there will be ~6-8 hours of work outside of lectures and computer labs for you to complete the readings and course assignments. The course schedule is at the end of the syllabus, but the table below highlights the assignments and their weight towards your final grade.

Type	Description	Due Date	Weight
Reflections	Reflection on the lecture and computer lab (4 total)	Ongoing	10%
Computer Labs	Practical application of lecture material to biological data (11 total)	Weekly	40%
Project Proposal	Proposed data and research question and hypotheses; proposal to be submitted in week 9	TBD	10%
Group Presentation	10-minute presentation summarizing the research project	TBD	20%
Group Project Report	Report detailing the statistical methods from the group project	TBD	20%
		Total	100%

Notes on Coursework

Reflections: You will complete a weekly reflection on the lecture topic and computer lab (~250-500 words, more if desired). These reflections will have directed prompts to guide your reflection, but there will also be an open field for you to expand on any component. You only need to submit 4 reflections, so you can choose from the list of topics and content.

Computer Labs: Computer labs are most important component of this course. In the computer labs, you will learn essential skills for data management and presenting results. Additionally, you will develop proficiency and confidence when conducting statistical analyses. Detailed outlines for each computer lab will be posted on the course website.

Group Project Proposal: In week 9, your group will submit a project proposal describing which of the provided datasets you would like to analyze. You will also submit the main research question(s) or hypotheses you would like to test with the data. I will then return your proposal with feedback and advice on potential statistical analyses you could use.

Group Presentation: Each group will give a 10-minute presentation that briefly describes the research question/hypotheses, statistical analyses, and results from the analyses.

Group Project Report: You will write a report detailing the statistical methods and results from your group project. Detailed assignment instructions will be provided on the course website, but the report will consist of: (1) a description of the statistical methods, (2) why the test(s) were appropriate for your question(s) or hypotheses and your data, (3) report of how the data management and analysis was conducted, and (4) results of the study using a combination of text, tables, and figures. Reports will be written using R Markdown, and all code and summary reports will be submitted for the final report.

Evaluation

We will be using the 'ungrading' approach to all evaluations rather than traditional grading systems. Ungrading is a fairly complex approach, but the main point is to make evaluation and assessment more of a conversation between you and me We are able to do this through a combination of feedback, self-evaluation, and reflection. Below I will expand on the evaluation for each type of coursework and how ungrading will be applied.

Reflections (10%): Lecture reflections are designed for you to articulate what you learned from the lecture and computer lab. I will provide comments and feedback on each reflection, offering advice, clarification, and encouragement as appropriate. I will also be using these reflections to help identify common challenges, misconceptions, or misunderstandings, so it is important that reflections also discuss challenging topics. Reflections will be evaluated for completion and addressing the directed reflection prompts.

Computer Labs (40%): Each computer lab is designed to develop your proficiency in R and your ability and confidence to conduct statistical analyses. Computer labs will be related to lecture content presented that week, with analyses applied to case studies across ecology, evolution, and biomedical fields. Computer labs must be submitted in both the R Markdown script file and the rendered summary report documenting the analyses and results. Each computer lab will be due by the end of the weekend (i.e., 11:59pm on Sunday night). I will return 'graded' computer lab assignments with a summarized feedback form. I will not provide any written scores, but I will maintain a spreadsheet of scores that each student earned on the assignment. You will evaluate your work and determine the number of points you think you earned. I will then compare the points I think you earned, your self-assessment, and the average of your score and my score. If my score is higher than your point total, we will typically use my point total. We can discuss any discrepancies if you would like to do so.

Group Presentation (20%): The group presentation will be a 15-minute presentation (plus 5 minutes for any questions) on your group project. Not every group member must speak during the presentation, but all group members must contribute to the presentation. The core checklist for the presentation will be posted on the course website. I will evaluate you presentation based on the core checklist and provide feedback. You will use this feedback and evaluate how many points you think you earned. I will then compare the points I think you earned, the points you think you earned, and the average of the two scores. If my score is higher than your score, we will typically use my point score.

Group Project Report (20%): You will work in groups of 2-3 people on a data analysis project. You will select a dataset, develop your own research questions and/or hypotheses, and then analyze those data. The project report will (1) provide an overview of the dataset, (2) list the research questions, hypotheses, and/or predictions, (3) describe the statistical methods, (4) present the results of the statistical analyses, and (5) interpret the statistical and biological relevance of the results; detailed instructions and a general core checklist will be provided on the course website. Groups will be formed by the end of week 9, and groups must submit their code of conduct and expected responsibilities by the end of week 10. The project report will be due at the beginning of week 15. I will evaluate your presentation based on the core checklist and provide feedback. You will use this feedback and evaluate how many points you think you earned. I will then compare the points I think you earned, the points you think you earned, and the average of the two scores. If my score is higher than your score, we will typically use my point score.

To control overinflation from students inflating their grade, if your score is within 1 standard deviation of my overall score, you will receive 5 bonus points. If your score is outside of 3 standard deviations of my overall score, I will deduct 5-10 points. I want to encourage accurate and honest self-assessment, which means fair evaluation of yourself and respecting my evaluation (which includes the performance by the entire class).

Teaching Methods and Academic Supports

BIO380 is an active learning class where you are part of the learning process. You are expected to come to class ready to engage in the material by participating in lecture activities, collaborating with your peers, and applying the concepts learned to case studies. **Learning can also bring about discomfort, and I will be challenging you in this course. I will challenge you because I know we all have the potential to grow and learn**.

Ungrading is central to this course. While ungrading does require work from both you and me, that work has lasting benefits beyond any single lecture or discussion. I want to help you learn about and have fun with building data literacy, but I am also here to help you grow as a learner. **Through the process of ungrading, we will stress less on any grade and focus more on learning.**

BIO380 is designed to build a strong conceptual understanding and the ability to conduct the appropriate statistical analyses using R. We will not be focusing on the mathematics underlying the statistics, but the broader meanings and applications.

Lectures

Lectures will expand on aspects of the assigned readings by going into great depth and applying knowledge to case studies and examples. You are responsible for reading the assigned readings before class to get the most out of the lectures. All lectures will be recorded and posted to Quercus within 24 hours.

Computer Labs

You will build your proficiency and confidence when using R by working through the computer labs.

Group Project

Group projects are a great way to learn how to work collaboratively and effectively with others. If you are facing any challenges in the group work, please contact me so we can work through the problem and find a solution.

Time Management and Learning Practices

If you find you are struggling with time management or keeping up with the material, please come to student (office) hours or we can schedule a private, one-on-one meeting. You may also talk to your academic advisor or the Robert Gillespie Academic Skills Center for guidance and advice on time management and effective learning practices. **I know that every student can succeed in this course, but sometimes the learning environment and support systems just need to be restructured to make that happen**.

Procedures & Policies

E-Mail Policy

The University's official method of correspondence with students is through their University of Toronto e-mail accounts. It is the student's responsibility to keep his/her @mail.utoronto.ca account active and check it on a regular basis.

All e-mails from students must include your full name and student number as well as have the course code in the subject line.

Re-Mark Policy

Requests for re-evaluation of course work must be made in writing to the instructor no later than one month following the return of the work. Re-evaluation may result in a grade increase, decrease, or no change.

Further Notes on the Re-Mark Policy

By using the ungrading approach, we will actively discuss evaluations and should therefore reduce the likelihood of any remark. That being said, we will still follow the official University of Toronto remark policy discussed above.

Further Notes on the E-Mail Policy

To help me and your TA better respond to emails, please include BIO205 in the subject line and then your student number either in the text or signature of your email. I also ask for patience when responding to emails. I will try to respond as quickly as possible but give me at least 24 hours to respond to any message. I likely will not respond to emails over the weekend, but I will aim to respond to by 5 PM the following Monday.

Attendance and Participation

Attendance is essential for your learning, as is your participation in active learning during lectures and computer labs. I will not take attendance during lecture, but attendance will be taken during computer labs.

Absences
Absences from computer labs must be communicated to me by email before that class period is over. Please send the email with a brief explanation for the absence. For an absence to be excused, it must meet University-approved and beyond-your-control criteria. Absences beyond University guidelines may be excused on a case-by-case basis.

Religious Observance
You are encouraged to observe and express your religious identity. I will make reasonable accommodations to allow any student to observe their religious practices without penalty. Please look at the course schedule below and let me know if there are any potential conflicts. Accommodations do not absolve students of responsibility for the coursework, but they can result in extensions.

Information about the University's Policy on Scheduling of Classes and Examinations and Other Accommodations for Religious Observances is at http://www.viceprovoststudents.utoronto.ca/publicationsandpolicies/guidelines/religiousobservances.htm

Classroom Management
You and I are expected to come to lectures and computer labs prepared, on time, and with our cell phones and other devices on silent and only to be used for lecture or computer labs activities. All lectures will be recorded over Zoom and posted on the course website by the next day following the lecture.

I expect you to treat yourself and others with respect in our learning community so we can engage, learn, and grow throughout the course. We each bring our own identities and experiences from our everyday lives, and that diversity will be celebrated.

Late Policy and Extensions
You are expected to complete and submit all assignments on time, although extensions and accommodations can be provided.

Late Policy
Group project reports will have a penalty of 15% for each day the assignment is late up to a maximum of 3 days, after which late submissions will not be accepted. Only group project reports will be accepted with a late penalty; no other assignments be accepted after the due date except for extreme circumstances. I have this policy to encourage you to stay on top of the material, which is to your benefit and that of your peers.

Extensions
If you require an extension to complete an assignment due to injury, illness, or accessibility, please let me know as soon as possible and preferably at least 24 hours advance of the due date. Extensions beyond accessibility and illness will be granted on a case-by-case basis.

Academic Integrity
The code of Behaviour on Academic Matters states that:

The University and its members have a responsibility to ensure that a climate that might encourage, or conditions that might enable, cheating, misrepresentation or unfairness not be tolerated. To this end all must acknowledge that seeking credit or other advantages by fraud or misrepresentation, or seeking to disadvantage others by disruptive behaviour is unacceptable, as is any dishonesty or unfairness in dealing with the work or record of a student.

- University of Toronto Mississauga Academic Calendar

It is your responsibility as a student at the University of Toronto, to familiarize yourself with, and adhere to, both the Code of Student Conduct and the Code of Behaviour on Academic Matters.

Notes on Academic Integrity

We will not be using Ouriginal for plagiarism detection in this course; however, generative AI (e.g., ChatGPT) will not be permitted in any form during this class and its use will be considered a violation of academic integrity. I am looking to see how you grow and learn throughout the course by looking at your work, not that of an algorithm or text-mining program.

Additional Notes

Personal Health Resources

There are many resources available through UTM that can benefit students and I would encourage you to use these resources discussed below to help you make the most of your time at UTM. These include:

The Equity, Diversity, and Inclusion Office: https://www.utm.utoronto.ca/equity-diversity/

The Indigenous Centre: https://www.utm.utoronto.ca/indigenous-centre/welcome-indigenous-centre The Health and Counselling Centre: https://www.utm.utoronto.ca/health/health-counselling-centre
This also includes the My Student Support Program or MySSP which provides University of Toronto students with immediate and/or ongoing confidential, 24-hour support for any school, health, or general life concern at no cost to students. You can call or chat with a counsellor directly from your phone whenever, wherever you are for a range of concerns. Students who use MySSP still have access to existing campus and community mental health services; MySSP is an additional support service. You can also access the service 24/7 by calling 1-844-451-9700. Outside of North America, call 001-416-380-6578. There is also an App you can use to access this service (https://myssp.app/ca/home).

Another number to have saved is Good2Talk (https://good2talk.ca/), which is a free, confidential support service for post-secondary students in Ontario. To talk, call 1-866-925-5454. To text, text GOOD2TALKON to 686868.

Other Resources AccessAbility
The University accommodates students with disabilities who have registered with the AccessAbility Resource Centre. Please let me know in advance, preferable in the first week of class, if you will require any accommodation on these grounds. To schedule a registration appointment with a disability advisor, please call the centre at 905-569-4699 or e-mail at: access.utm@utoronto.ca. http://www.utm.utoronto.ca/access/

Robert Gillespie Academic Skills Centre
Students can visit the Academic Skills Centre to consult with one of its strategists about understanding learning style, developing study plans for upcoming tests/exams, or discussing papers. Special Diagnostic Assessments are also offered and are designed to help you learn exactly where you stand with respect to critical academic skills.
http://www.utm.utoronto.ca/asc

UTM Library (Hazel McCallion Academic Learning Centre)
The University of Toronto boasts the biggest academic library in Canada and the second biggest in North America. Various services are available to students at the UTM Library and across the UofT library system. Services including borrowing, interlibrary loans, online references, laptop loans and the RBC Learning Commons. For more information, visit http://library.utm.utoronto.ca.

Course Schedule

Week	Lecture	Readings	Computer Lab
1	L1: Syllabus & Why Data Literacy Matters	Syllabus	Lab 1: Introduction to R
2	L2: Research Questions & Hypothesis Testing	Chapter 1	Lab 2: Description & Estimation
3	L3: Experimental Design	Chapter 2 -Kennedy-Shaffer 2019	Lab 3: Data Management in the tidyverse
4	L4: Interpretation of Results	Chapter 3 -Nakagawa & Cuthill 2007 -Muff et al. 2022 -Berner & Amrhein 2022	Lab 4: Data Visualization
5	L5: Correlation & Regression	Chapter 4	Lab 5: Linear Regression
6	L6: Model Selection	Chapter 5 -Johnson & Omland 2004 -Grueber et al. 2011	Lab 6: Multiple Regression & Model Selection
7	L8: ANOVA	Chapter 6	Lab 7: Simple & Complex ANOVAs
8	L9: ANCOVA	Chapter 7	Lab 8: ANCOVA
9	L10: Linear Mixed-Effects Models	Chapter 8 -Bolker et al. 2009 -Harrison et al. 2018 -Silk et al. 2020	Lab 9: Linear Mixed-Effects Models
10	L11: Generalized Linear Mixed-Effects Models	Chapter 9	Lab 10: Generalized Linear Mixed-Effects Models
11	L12: Structural Equation Modeling I	Chapter 10 -Grace et al. 2010 -Grace et al. 2012 -Grace 2020	Lab 11: Global SEM
12	L12: Structural Equation Modeling II	Chapter 11 -Lefcheck 2016 -Laubach et al. 2021 -Rohrer 2018	Lab 12: Local SEM
13	L13: Moving Forward with Best Practices	Wasserstein & Lazar 2016 Wasserstein et al. 2019	Lab 13: Project Work
14	L14: Project Work		Lab 14: Group Presentations
15	No Class (Thanksgiving Break)		
16	L15: Group Presentations		Lab 15: Group Presentations
17	No Class (Final Exams)		

Reading List

Berner, D., and V. Amrhein. 2022. Why and how we should join the shift from significance testing to estimation. Journal of Evolutionary Biology 35:777–787.

Bolker, B. M., M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. H. Stevens, and J.-S. S. White. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends in Ecology and Evolution 24:127–135.

Grace, J. 2020. A "Weight of Evidence" approach to evaluating structural equation models. One Ecosystem 5:e50452.

Grace, J. B., T. M. Anderson, H. Olff, and S. M. Scheiner. 2010. On the specification of structural equation models for ecological systems. Ecological Monographs 80:67–87.

Grace, J. B., D. R. Schoolmaster Jr., G. R. Guntenspergen, A. M. Little, B. R. Mitchell, K. M. Miller, and E. W. Schweiger. 2012. Guidelines for a graph-theoretic implementation of structural equation modeling. Ecosphere 3:art73.

Harrison, X. A., L. Donaldson, M. E. Correa-Cano, J. Evans, D. N. Fisher, C. E. D. Goodwin, B. S. Robinson, D. J. Hodgson, and R. Inger. 2018. A brief introduction to mixed effects modelling and multi-model inference in ecology. PeerJ 6:e4794.

Johnson, J. B., and K. S. Omland. 2004. Model selection in ecology and evolution. Trends in Ecology and Evolution 19:101–108.

Kennedy-Shaffer, L. 2019. Before $p < 0.05$ to beyond $p < 0.05$: using history to contextualize p-values and significance testing. The American Statistician 73:82–90.

Laubach, Z. M., E. J. Murray, K. L. Hoke, R. J. Safran, and W. Perng. 2021. A biologist's guide to model selection and causal inference. Proceedings of the Royal Society B: Biological Sciences 288:20202815.

Lefcheck, J. S. 2016. piecewiseSEM: Piecewise structural equation modelling in r for ecology, evolution, and systematics. Methods in Ecology and Evolution 7:573–579.

Muff, S., E. B. Nilsen, R. B. O'Hara, and C. R. Nater. 2022. Rewriting results sections in the language of evidence. Trends in Ecology and Evolution 37:203–210.

Nakagawa, S., and I. C. Cuthill. 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. Biological Reviews 82:591–605.

Rohrer, J. M. 2018. Thinking clearly about correlations and causation: graphical causal models for observational data. Advances in Methods and Practices in Psychological Science 1:27–42.

Silk, M. J., X. A. Harrison, and D. J. Hodgson. 2020. Perils and pitfalls of mixed-effects regression models in biology. PeerJ 8:e9522.

Wasserstein, R. L., and N. A. Lazar. 2016. The ASA statement on p-values: context, process, and purpose. The American Statistician 70:129–133.

Wasserstein, R. L., A. L. Schirm, and N. A. Lazar. 2019. Moving to a world beyond "$p < 0.05$." The American Statistician 73:1–19.