

Department of Statistics, University of Manitoba
**STAT 7350 – Statistical Analysis and Visualization
of Biological Data in R**

Tentative Course Outline, Spring Term 2019

Course Details

Course Number & Title: STAT 7350, Statistical Analysis and Visualization of Biological Data in R
Section & CRN: TBA
Course Schedule: Tuesday/Thursday, 13:00 to 14:15
Location for Lectures: TBA
Prerequisites: There are no formal prerequisites for this course, but experience with the R programming language will be an important asset.

Instructor Contact Information

Instructor: Aleeza Gerstein
Office Location: 364 Machray Hall
Phone: (204) 272-1681
Email: Aleeza.Gerstein@umanitoba.ca
Office Hours: Wednesday, from 11:30 to 12:30, or by appointment.

General Goals for this Course

This course will be focusing on learning

- some theory on the principles and methods of effective data visualization
- the "tidyverse" suite of packages for working with data in R
- some biological subject knowledge
- appropriate visualization and analysis techniques for biological data

Each of these topics in isolation could be studied over one or many semesters of coursework. The overarching aim is to provide students with a broad overview the type of data that is commonly collected in biological research and analysis techniques powered by the R programming language and the tidyverse packages. Data visualization is increasingly being appreciated as an important but under-discussed aspect of data analysis, and we will spend a considerable amount of time thinking about how to utilize data visualization to understand and convey a message about the what the data shows us. Specific biological topics will likely include ecological research data, empirical lab data, and next-generation sequencing (genomics) data. Instruction in specific statistical techniques is from an applied rather than theoretical perspective and is driven by the data.

Outline of Covered Topics

1. Introduction to biological data
2. R and R Markdown review
3. Data visualization with ggplot2
Assignment 1
4. Importing data, tidying data, wrangling data
5. Ecological research data
Assignment 2
6. The central dogma (DNA → RNA → Protein)
7. Empirical lab data
Assignment 3
8. Next generation sequencing (genomics) data

Textbook and Other Materials

Textbook: There is no formal textbook for this course. My course notes will draw in part from the following textbooks available online.

- Claus Wilke, Fundamentals of Data Visualization, <https://serialmentor.com/dataviz/>
- Chester Ismay and Albert Y. Kim, An Introduction to Statistical and Data Sciences via R <https://moderndive.com/>
- Susan Holmes and Wolfgang Huber, Modern Statistics for Modern Biology <http://web.stanford.edu/class/bios221/book/>

Online resources: The following are very useful online resources for R:

- <http://www.rstudio.com> (for downloading R Studio)
- <http://rmarkdown.rstudio.com> (to learn about R Markdown)
- <https://www.tidyverse.org/> (to learn about the tidyverse packages)
- <http://r-pkgs.had.co.nz/style.html>
(a simple R style guide that is generally good to follow)

Assignments, Test and Exams

Assignments: There will be three assignments and one final project. All should be completed and submitted as R Markdown files (RMD & html or pdf).

You will have one week to complete each assignment and at least two and a half weeks to complete the final project.

Exams: There will be no exams.

Readings/Homework: You will be required to do some readings outside of class. They are not intended to be particularly strenuous. This will free up additional class time for discussion and hands-on work.

Grading timeline: Under normal circumstances, assignments should be graded a week after submission.

Course Evaluation and Grading Scheme

Final Mark: The final mark for the course will be obtained from the following:

Assignments (3)	60%	(20% each)
Final project	40%	

Letter Grade: I normally use the following cutoffs when assigning letter grades:

Letter Grade	Mark out of 100
A+	90-100
A	80-90
B+	75-80
B	70-75
C+	65-70
C	60-65
D	50-60
F	below 50

Technology in the Classroom

It is the general University of Manitoba policy that all technology resources are to be used in a responsible, efficient, ethical and legal manner. Students should restrict their use of technology to those approved by the instructor *for educational purposes only*. Electronic messaging, email, social networking, gaming, etc. should be avoided during class time. Cell phones should be on silent. If a student is on call for emergencies, their cell phone should be on vibrate mode and the student should leave the classroom before using it.

Important Dates

The following dates are important to how the course will progress throughout the term.

Date	Information
February 26	First lecture
March 20	Last day for VW
April 9	Last lecture

Class Communications

The University requires all students to activate an official U of M email account, which should be used for all communications between yourself and the university (including all your instructors). All these email communications should comply with the University's policy on electronic communication with students, which can be found at: http://umanitoba.ca/admin/governance/governing_documents/community/electronic_communication_with_students_policy.html

Copyrights

Copyrighted Materials: We will use copyrighted content in this course. I have ensured that the content I use is appropriately acknowledged and is copied in accordance with copyright laws and University guidelines. Copyrighted works, including those created by me, are made available for private study and research and must not be distributed in any format without permission.

Lectures: No audio or video recording of lectures or presentations is allowed in any format, openly or surreptitiously, in whole or in part without my permission.

More details are available online at <http://umanitoba.ca/copyright/>.

Academic Integrity

Students are encouraged to discuss course material, including assignments and the final project. However, each student must hand in his or her own copy of each assignment/project with personalized solutions, including comments, discussions, explanations and interpretations, and R code. Copying from anywhere, including other students, books, or the internet constitutes a case of academic dishonesty and could have serious consequences.

The goal in this class (as in all academic pursuits) is to learn. If you are unclear on what is acceptable, please ask me or visit the Faculty of Science page on cheating and plagiarism at: <http://umanitoba.ca/faculties/science/undergrad/resources/webdisciplinedocuments.html>

ROASS Schedule A

Schedule “A” of the *Responsibilities of Academic Staff with regards to Students* (ROASS) policies of the University of Manitoba lists resources and policies for students. It is important that you familiarize yourself with these resources and policies. This document is available from the Department of Statistics web page at: <http://umanitoba.ca/science/statistics/>.

University of Manitoba Acknowledgement of Traditional Territories

The University of Manitoba campuses are located on original lands of Anishinaabeg, Cree, Oji-Cree, Dakota, and Dene peoples, and on the homeland of the Métis Nation.

We respect the Treaties that were made on these territories, we acknowledge the harms and mistakes of the past, and we dedicate ourselves to move forward in partnership with Indigenous communities in a spirit of reconciliation and collaboration.