STAT 2430 Data Visualization

This is the syllabus for the Data Visualization (STAT 2430) course offered in Fall 2022 at Dalhousie University.

- Instructor: Andrew Irwin, a.irwin@dal.ca
- Meeting times: Tuesday and Thursday 1:00-2:20 pm Atlantic in LSC-C244 and on Collaborate
- Office hours: Wednesday 1:00 pm Atlantic on Collaborate, or by appointment.

Course Description

Data visualization is the art and science of turning data into readable graphics. We will explore how to design and create data visualizations. This exploration will include both the principles and techniques of data visualization. Students will learn the value of visualization, specific techniques, and how to design and improve a data visualization.

Course Prerequisites

At least one MATH or STAT course at or above the 1000 level. No experience with R or computer programming is required, but a desire to learn about both is essential.

Course Materials

Course notes:

- Andrew J Irwin (2022) Data visualization.

Textbooks:

- Kieran Healy (2018) Data Visualization: A Practical Introduction. Princeton University Press.
- Claus O. Wilke (2019) Fundamentals of data visualization. O'Reilly.
- Garret Grolemund & Hadley Wickham (2017) R for data science. O'Reilly.

The textbooks are available in printed form from bookstores and online. The online versions are free to use. You are not required to buy the books on paper. All three are excellent books with very distinct goals. The books by Healy and Wilke focus on the design and interpretation of data visualizations. The book by Grolemund and Wickham describes the use of R, particularly the set of packages known as the "tidyverse" for data visualization and the broader tasks of data science.

Course Structure

Each week will be structured around the following components:

- Reading course notes and suggested material from the textbooks.
- In class meetings which will be a mixture of lectures, demonstrations, and skills practice.

  - Some class time will be be dedicated to computer exercises, so you should bring your computer to these classes. I will alert you to these activities in the detailed course outline and by making an announcement in the previous class meeting.

- Opportunities to develop and demonstrate your knowledge in course work to be evaluated (tasks, assignments, and a term project).
- Scheduled office hours. Come to office hours if you have questions about the classes, readings, or work to be submitted for grading. If you can't participate in office hours, but you have a question, please get in touch with me and arrange another time to meet.

I recorded mini-lectures for most lessons the last time this course was offered. You don't need to watch these, but if you wish additional material you can find them here.

Follow the detailed outline for each week's plan to keep you on track.

Evaluation

Your goal in this course is to learn the principles and skills of data visualization. Most people benefit from a mixture of collaborative and independent work. The general guideline is that work you submit should be your own—your ideas, your thoughts, your choices, your code, your typing. You are encouraged to discuss your work with the instructor and students. When you help another student, be careful you don't help them so much that you inhibit their learning.

Tasks are opportunities to test your understanding of the key concepts from the lessons and demonstrate you have developed basic proficiency with essential computing skills. Most lessons will have a task for you to complete. You are expected to submit one task to be graded each week. *Tasks* are primarily designed to help you learn and you are encouraged to seek assistance from classmates, but work you submit must be your own.

Assignments combine skills from multiple lessons into a meaningful data visualization activity. These are opportunities to apply the content of lessons in thoughtful and sometimes creative ways. *Assignments* are assessments of your skills and should be done independently.

The *Term Project* is an opportunity to combine many of the skills learned in the course. You will explore, analyze, and present to a reader an analysis of a data set of your choice. Your proposal will select a data set and describe some of your planned analyses. The oral presentation will be a 5 minute overview of your data and key visualizations. The report will be a polished reproducible document demonstrating many of the ideas of the course using your

selected data set. The *term project* is designed to be done collaboratively in a group. The work on the project should be the work of members of the group and not outside collaborators or internet sources.

- Tasks (30%, 1 per week, due weekly, equally weighted)
- Assignments (40%, roughly 1 every two weeks, equally weighted)
- Term project, divided into three components

    - Proposal, due Thursday 27 October, 5%
    - Oral presentation, due Monday 5 December, 10%
    - Report, due Wednesday 7 December, 15%

Tasks are due on Friday. Assignments are due on Wednesday. (Office hours are also on Wednesday for last-minute question.)

Copying work from any other sources is not allowed and will be considered an academic offense for this course. You are encouraged to **learn** from many different sources. If you make use of material outside of course materials on assignments or the project report, include references to the material and a description of what you used in a "Sources used" section at the end of your work.

Late policy. Tasks will not be accepted late, but you can submit two "catch-up" tasks to replace any tasks you may have missed, or if you want to improve your grade on a submitted task. Assignments will be accepted without penalty until the work is graded or solutions are posted. After that, no late assignments will be accepted. If an additional accommodation is requested and granted due to illness or other unforeseeable circumstance, the value of that work will be redistributed to other tasks or assignments. *If you anticipate not being able to submit any component of the term project (proposal, oral presentation, or report) on time, please contact me by email.*

Grades will be reported on Brightspace

Software

We will use the statistical software R and RStudio and version control software git. No prior experience with R, RStudio or git is assumed. We will take class time to learn the software.

- To download and install R go to r-project.org and click on the link to download R
- To download and install Rstudio, go to Rstudio.com and click on the link to download Rstudio
- To download and install git:

    - on Windows, go to git-scm.org and click on the link to download a version for Windows

– on Macintosh, use the Terminal app and type `xcode-select --install` to download and install git.

These tools can be installed on Linux computers as well. Contact me if you have trouble. If you have a Chromebook you can use all these tools through the cloud service rstudio.cloud. Everyone should learn to use the cloud service as a backup in case of problems with setting up the software on their own computer.

R and Rstudio are available on Dalhousie computer labs, but the git version control software must be installed following the instructions for Windows computers above. Since all your user files are erased from lab computers when you log out, this process must be repeated on each login. For this reason, I don't recommend using Dalhousie computer labs for this course, unless you are using rstudio.cloud.

You should install this software on your computer during the first week of the course. If you are unable to install the software, get in touch with me for assistance.

Course Policies

COVID-19

Please do not come on campus if you have symptoms similar to COVID-19 such as a fever or cough, if you have had a close contact with someone who has COVID-19, or if you have COVID-19.

The course is intended to be delivered in person, on campus at Dalhousie. I hope that most of you will be physically present in the classroom for most classes. Since you can be infected and transmit the virus without experiencing any symptoms, I encourage everyone to wear a high-quality mask while on campus to reduce the probability of virus transmission. I anticipate that some students will need to miss some classes following a close contact with an infected person, developing symptoms similar to COVID-19, or contracting COVID-19. All course meetings will be available live on Brightspace and recorded for watching later. This means that you can participate in the whole course even if you can't attend a particular class. You are not required to be physically present on campus for classes, tests, or a final exam. This design is to keep you safe and to help you keep your classmates safe.

If I am unable to be on campus, the class will be delivered remotely.

Data retention policy

During the course you will create work on a third-party website (github). These files will be shared only with the instructor, TAs, and student team members. All files will be kept for 6 months after the end of the course. Repositories will be deleted after that time. If you wish to

keep files after that date, either clone the repositories to your computer or fork them to your own GitHub account separate from the course.

Important dates

- Tuesday 6 September: Term begins
- Monday 10 October: Thanksgiving day, University closed
- November 7-11: Study break
- Fri 11 November: Remembrance Day
- Tue 6 December: Last Tuesday/Thursday class

If you know you will be unable to meet certain deadlines in advance, for *bona fide* reasons such as scheduled medical appointments or religious observances, please contact me as soon as practical.

I will never require you to provide a medical note or to use the Student Declaration of Absence form.

For due dates for evaluations, see the detailed course outline and evaluation schedule.

Letter grades

Your numerical grade will be converted to a letter grade for the course using the Dalhousie Common Grade Scale.

First your numerical grade will be rounded up to the nearest integer, then it will be converted to a letter using this table.

Letter grade	Grade range
A+	90-100
A	85-89
A-	80-84
B+	77-79
B	73-76
B-	70-72
C+	65-69
C	60-64
C-	55-59
D	50-54
F	0-49

University Policies and Statements

This course is governed by the academic rules and regulations set forth in the University Calendar and by Senate

Missed or Late Academic Requirements due to Student Absence. Information on the university policy, including the use of the Student Declaration of Absence can be found here.

Academic Integrity. At Dalhousie University, we are guided in all of our work by the values of academic integrity: honesty, trust, fairness, responsibility and respect (The Center for Academic Integrity, Duke University, 1999). As a student, you are required to demonstrate these values in all of the work you do. The University provides policies and procedures that every member of the university community is required to follow to ensure academic integrity. More information.

Accessibility. The Advising and Access Services Centre is Dalhousie's centre of expertise for student accessibility and accommodation. The advising team works with students who request accommodation as a result of a disability, religious obligation, or any barrier related to any other characteristic protected under Human Rights legislation (Canada and Nova Scotia). More information.

Student Code of Conduct. Everyone at Dalhousie is expected to treat others with dignity and respect. The Code of Student Conduct allows Dalhousie to take disciplinary action if students don't follow this community expectation. When appropriate, violations of the code can be resolved in a reasonable and informal manner—perhaps through a restorative justice process. If an informal resolution can't be reached, or would be inappropriate, procedures exist for formal dispute resolution. Code.

Diversity and Inclusion - Culture of Respect. Every person at Dalhousie has a right to be respected and safe. We believe inclusiveness is fundamental to education. We stand for equality. Dalhousie is strengthened in our diversity. We are a respectful and inclusive community. We are committed to being a place where everyone feels welcome and supported, which is why our Strategic Direction prioritizes fostering a culture of diversity and inclusiveness. Statement.

Recognition of Mi'kmaq Territory. Dalhousie University would like to acknowledge that the University is on Traditional Mi'kmaq Territory. The Elders in Residence program provides students with access to First Nations elders for guidance, counsel and support. Visit or e-mail the Indigenous Student Centre (1321 Edward St) (elders@dal.ca). Information.

- Important Dates in the Academic Year (including add/drop dates)
- University Grading Practices

Student Resources and Support

Please follow the following links for additional resources and support.

Advising: General Advising, Science Program Advisors, Indigenous Student Centre, Black Advising Centre, International Centre

Academic supports: Library, Writing Centre, Studying for Success, Copyright Office, Fair Dealing Guidelines

Other supports and services: Student Health & Wellness Centre, Student Advocacy, Ombudsperson

Safety: Biosafety, Chemical Safety, Radiation Safety, Scent-Free Program