

CHL8010H F2: Statistical Programming and Computation for Health Data

Dalla Lana School of Public Health

Fall 2024

1 Details

Contact Information

Instructor

NAME: Aya Mitani

EMAIL: aya.mitani@utoronto.ca

OFFICE: Health Sciences Building 628

OFFICE HOURS: Thursdays 10:30-11:30am via [Zoom](#)

Course Meeting Information

DATE AND TIME: Monday 10:00am – 1:00pm

LOCATION: TBD

WEBSITE: <https://q.utoronto.ca/courses/369245>

Acknowledgement of Territory

We wish to acknowledge this land on which the University of Toronto operates. For thousands of years, it has been the traditional land of the Huron-Wendat, the Seneca, and the Mississaugas of the Credit. Today, this meeting place is still home to many Indigenous people from across Turtle Island and we are grateful to have the opportunity to work on this land.

Pre-requisites

Open to MSc and PhD Biostatistics students. Others with permission from the instructor. Students are expected to have some experience with a statistical programming language, preferably R, and a basic understanding of linear regression and logistic regression.

Course Description

This is a pilot course that covers essential R programming and computational tools in health sciences research. Topics include data manipulation, data visualization, loops and functions, optimization, and package development with an emphasis on reproducibility and replicability.

Learning Outcomes

Students who complete this course will be able to:

- Develop a reproducible workflow while integrating version control for conducting research in health data sciences.
- Use R to prepare an analytical data set and perform descriptive analysis.
- Produce a visual (graphical or tabular) display of the data that effectively communicates the trend or pattern based on the research question of interest.

- Write efficient reproducible code throughout the data science project.
- Write an efficient simulation program.
- Develop an R package and deploy it on GitHub.
- Create reproducible professional grade documents (reports, articles, blog posts, presentation slides) using R.
- Use computational tools (bootstrap, multiple imputation, etc.) to aid in statistical analyses.

Computing

We will use R for all statistical programming and computation. R is a free open-source statistical software available for Macs, Windows, and Linux machines. R can be downloaded free from <https://www.r-project.org/>. Other software you need for the course are:

- Quarto, which is an open-source scientific and technical publishing system. Download for free from <https://quarto.org/>.
- RStudio, which is an integrated development environment for R. It can be downloaded for free from <https://rstudio.com/products/rstudio/>. See the course website for more information, including links to tutorials.
- Git, which is a free and open source distributed version control system. It can be downloaded from <https://git-scm.com/>.
- Download the Rtools package from <https://cran.r-project.org/bin/windows/Rtools/>

Finally, create a GitHub account (<https://github.com/>) If this is your first time creating a GitHub account, here are a few tips on choosing your username, <https://happygitwithr.com/github-acct.html>

Notes and Texts

Electronic copies of lecture notes and assignments, as well as code and data sets, will be posted on the course website. Although there are no required texts, the following are very good references for the material that will be covered in class:

1. Peng, R. D. *R Programming for Data Science* <https://bookdown.org/rdpeng/rprogdatascience>.
2. Wickham, H. & Grolemund, G. *R for Data Science* <https://r4ds.had.co.nz/>.
3. Alexander, R. *Telling Stories with Data* <https://tellingstorieswithdata.com/>.
4. Baumer, B. S., Kaplan, D. T. & Horton, N. J. *Modern Data Science with R* <https://mdsr-book.github.io/mdsr2e/>.

Absences

If you have to miss class

If you have to miss a class, then try to complete the in-class activities in your own time. If you need to talk to me, please schedule a meeting with me. I would be happy to fill in the details of the missed lecture(s), within reason (i.e., I will not be your personal tutor every week). Due to the structure of this course that employs active learning, attendance is part of the evaluation. However, you are allowed to skip two classes without penalty.

If I have to miss class

If I need to stay home due to my own or a family member's illness, then I will move the class to Zoom. I have two young children, and unfortunately they get sick often. I will do my best to run this course smoothly but please understand if we experience any disruptions.

2 Evaluation

- **Attendance and participation** (10% of final grade): This course uses a hands-on active learning style, which makes attendance and participation crucial. Class time will be dedicated to individual and group work. You are allowed to skip two classes without penalty. Your in-class contributions will be assessed weekly on a scale of 0 to 2, given in Appendix A.
- **Pre-class assignments and quizzes** (40% of final grade): Each week (except for Week 1), there will be a pre-class assignment, either in the form of reading or video watching. There will be a total of ten short online quizzes, each due before class time. Each quiz will be based on the pre-class assignment. The pre-class assignments and quizzes will be posted a week before they are due. You will have one attempt to complete the quiz in one hour at any time during the week. You can choose to skip two quizzes without penalty. If you choose to complete all ten quizzes, then the two lowest marked quizzes will be excluded from calculating the final grade.
- **Final project** (30% of final grade): You will develop an R package for the final project. In late October, you will give me a short proposal on the R package you plan to develop. I will provide some feedback. The proposal is not part of the final grade, but is required. The evaluation criteria for the project are given in Appendix B.
- **Final presentation** (20% of final grade): You will give a short tutorial on how to use your R package that you developed for your final project. The tutorial can be in any format; you can give a live demonstration, create slides, go through a vignette document, etc. The evaluation criteria for the presentation are given in Appendix C.

3 Course Schedule

Week	Date	Topic	Assignment due
1	Sep 9	Introduction – Quarto for R	
2	Sep 16	Reproducible workflow – R project, Git, and Github	Quiz 1
3	Sep 23	Preparing data for analysis – functions and loops	Quiz 2
4	Sep 30	Exploratory analysis and data visualization	Quiz 3
5	Oct 7	Table of baseline characteristics	Quiz 4
6	Oct 14	Fall break – no class	
7	Oct 21	Developing an R package	Quiz 5
8	Oct 28	Presenting results from multiple regression models	Quiz 6
9	Nov 4	Multiple imputation for handling missing data	Quiz 7
10	Nov 11	Conducting a simulation study	Quiz 8
11	Nov 18	Creating a website	Quiz 9
12	Nov 25	Bootstrap standard errors and confidence intervals	Quiz 10
13	Dec 2	Final presentations I	
14	Dec 9	Final presentations II	Final project due on Dec 13

*The schedule is subject to change.

4 Course Policies

Collaboration and external resources

You may use external resources such as textbooks, journal articles, or the internet to complete the quizzes, but you may not collaborate or consult with your classmates. You may collaborate and consult with your classmates to complete the final project and prepare for the final presentation. It may be a good idea to have your classmates test out the package before the due date.

Assignment submission

Please submit each assignment (quizzes and final project) on time. You are responsible for ensuring that I receive each assignment before it is due. If you need a deadline extension for the quizzes, let me know as soon as possible, but before the deadline. In general, late quizzes will not be accepted without prior approval from me. For the final project, 5% will be deducted for each day it is late, and it will not be accepted after four days from the due date.

Children in the classroom

Dependents, including children of any age (bottle-feeding, nursing, etc) are welcome in class as often as necessary. Please feel free to take breaks to feed or express milk for your child as needed, either in the classroom or elsewhere: <https://familycare.utoronto.ca/childcare/breastfeeding-at-u-of-t/>. Please let me know your needs so that I can make proper accommodations. There may be times when I need to bring *my* children to class too.

Talk to me!

I want for you to learn the material in a fun, collaborative, and stimulating environment. Please provide feedback – positive or negative – at any time during the semester. If you are having trouble learning, come talk to me before it is too late. If you enjoyed something we did in class, tell me so that we can do it again. This is the first time I’m teaching this course in-person. Your feedback will also benefit future students taking this course. The feedback you provide will *not* affect your grade in any way.

5 UofT Policies

Satisfactory performance and good academic standing

Students are reminded they must achieve satisfactory performance to remain in good academic standing and continue with their program. According to the School of Graduate Studies (SGS), “Satisfactory performance in a degree program requires the completion of every course taken for graduate credit with a grade of at least a B– “. For a full description of these requirements please see the SGS Calendar section “*7 Good Academic Standing and Satisfactory Academic Progress, Time Limits, Supervision, and Candidacy*” and section “*7.4 Satisfactory Completion of Graduate Courses*”. (<https://sgs.calendar.utoronto.ca/general-regulations#7>)

Respect for classmates

The University of Toronto is committed to equity, human rights, and respect for diversity. All members of the learning environment in this course should strive to create an atmosphere of mutual respect where all members of our community can express themselves, engage with each other, and respect one another’s differences. U of T does not condone discrimination or harassment against any persons or communities.

Academic integrity

Students must adhere to the [Code of Behaviour on Academic Matters](#). **It is your responsibility to know what constitutes appropriate academic behaviour.** You are responsible for ensuring that you do not act in such a way that would constitute cheating, misrepresentation, or unfairness, including but not limited to, using unauthorized aids and assistance, personating another person, and committing plagiarism. For more information see U of T [Academic Integrity website](#).

Academic integrity includes understanding appropriate research and citation methods. If you are uncertain about this, please seek out additional information from the instructors or from other institutional resources including the following:

- This [tip sheet](#) provides clear and helpful information about appropriate academic citation.
- [This site](#) offers a series of scenarios to help students understand how to prevent themselves from being subject to academic offence allegations.

- Before handing in assignments students can also review this [academic integrity checklist](#) provided by the UofT Centre of Teaching Support & Innovation:
 - I have acknowledged the use of another’s ideas with accurate citations.
 - If I used the words of another (e.g., author, instructor, information source), I have acknowledged this with quotation marks (or appropriate indentation) and proper citation.
 - When paraphrasing the work of others, I put the idea into my own words and did not just change a few words or rearrange the sentence structure.
 - I have checked my work against my notes to be sure I have correctly referenced all direct quotes or borrowed ideas.
 - My references include only the sources used to complete this assignment.
 - This is the first time I have submitted this assignment (in whole or in part) for credit.
 - Any proofreading by another was limited to indicating areas of concern which I then corrected myself.
 - This is the final version of my assignment and not a draft.
 - I have kept my work to myself and did not share answers/content with others, unless otherwise directed by my instructor.
 - I understand the consequences of violating the University’s Academic Integrity policies as outlined in the [Code of Behaviour on Academic Matters](#).

Accessibility

Students with diverse learning styles and needs are welcome in this course. If you have a disability or health consideration that may require accommodations, please feel free to approach me/us and/or the Accessibility Services Office as soon as possible. The Accessibility Services staff are available by appointment to assess specific needs, provide referrals and arrange appropriate accommodations. The sooner you let them and me know your needs, the quicker we can assist you in achieving your learning goals in this course. For more information, or to register with Accessibility Services, please visit: <http://studentlife.utoronto.ca/as>.

6 Student Support Services

Student Mental Health Resources

- U of T [Central Hub for Student Mental Health Resources](#)
- [SGS Graduate Wellness Services](#)
- [Student Life Health and Wellness](#)
- MySSP: 1-844-451-9700 (or use the app)
- Good2Talk: Call: 1-866-925-5454 or Text: GOODTOTALKON to 686868
- 8th Floor Lounge on Quercus

Tri-Campus Sexual Violence Prevention and Support Centre (“The Centre”)

- The Centre has a location on all three campuses to help students who disclose to access supports and, in cases where the student chooses to formally make a report, the Centre will explain the process, and facilitate the making of the report to the University and/or to the Police. The Centre can be contacted through its confidential phone at 416-978-2266 or thesvpcentre@utoronto.ca.
- Students are encouraged to call or email for an appointment, which will be booked as soon as possible. **Students in crisis or in need of immediate support related to sexual violence can contact Women’s College Hospital Sexual Assault Care Centre at 416-323-6040.**

Resource	Summary Description
The Office of Graduate Affairs	Provides a variety of services, including academic, program and personal advising. DLSPH students that require any assistance or information regarding coursework extension, program requirements, etc..
Graduate Department of Public Health Sciences Student (GDPHS) Handbook	This resource provides comprehensive information on getting started, enrolment, policies and procedures, financial matters, awards and funding opportunities, student services and more.
DLSPH Student Resources	This resource site includes information for incoming students, the GDPHS Student Handbook, program requirements, policies and forms, online learning resources, timetables, course database, information for international students, professional opportunities, mentorship program, health & well-being, public health students' association, and PhD Final Oral Exams.
U of T Graduate Student Union	The UTGSU is a voice for over 18 500 students as well as a platform for community building and services. UTGSU supports and advocates on behalf of graduate students.
Health Sciences Writing Centre (for DLSPH PhD students)	<p>The Health Sciences Writing Centre provides free individualized, confidential writing instruction to:</p> <ul style="list-style-type: none"> • Develop your writing skills • Improve your capacity to plan, organize, write, and revise academic papers (in any subject!) • Manage ESL/EFL language challenges <p>The Centre works with all students, for all assignments, at all stages of the writing process. Visit the website to book an appointment or for more information.</p>
UofT Academic Success Centre	Offers group workshops and individual counselling to develop strategies for a range of learning challenges such as time management, stress and anxiety, memory, exams, note taking, textbook reading, concentration.
UofT Career Services	Provides opportunities to meet employers, industry experts and alumni; strategies to Identify goals and navigate career decisions; and resources: Improve your resume, interviews, and online presence.

Office of the Ombudsperson

- As part of the University's commitment to ensuring that the rights of its individual members are protected, the University Ombudsperson is devoted to ensuring procedural fairness and just and reasonable outcomes.
- The Ombudsperson offers confidential advice and assistance for complaints and concerns and can recommend changes in academic or administrative procedures where this seems justified. For information, see <https://governingcouncil.utoronto.ca/ombudsperson>.

Other Resources and Supports for DSLPH Graduate Students

- [U of T Graduate Student Union](#)
- [Biostatistics Union of Graduate Students \(BUGS\)](#)
- [DLSPH Covid Information](#)
- [DLSPH Student Resources \(Policies, Financial Aid, Health and Wellness, etc.\)](#)
- [DLSPH Student Handbook](#)

Appendix A: Weekly assessment for attendance and participation

- 0 = Absent
- 1 = Present but not respectfully participating in class discussion or individual/group activities
- 2 = Present and respectfully participating in class discussion and individual/group activities

Appendix B: Evaluation criteria for the final project

Item	Details	Score per item (1-10)
Accuracy	Package does what it intends to do, correct methods, user-friendly, clear outputs and results with correct labels	
Code	Free of bugs, organized, separate scripts when necessary, good documentation within code, efficient use of functions and loops	
Originality and creativity	Name of package, name of function, uniqueness of idea	
Documentation	Describes each component of package, each function has a help file, clear and concise, provides examples, free of typos and grammar errors	
Repository	User can download package without error, describes package, has simple instructions, organized	
Total		/50

Appendix C: Evaluation criteria for the final presentation

Item	Details	Score per item (1-10)
Delivery	Speed, eye contact, audibility, tone, composure, posture, facial expressions, completes within allotted time frame	
Content and structure	Objectives and goals clearly stated, organization, knowledge of material	
Materials	Easy to read, tables and figures are clearly described and labeled, free of typos and grammar errors	
Response to questions	Understands question that was asked, answers questions appropriately	
Total		/40