STA304H1/STA1003H1: Surveys, Sampling and Observational Data
Winter 2025 (Jan-Apr)

Sections L0101 & L0201
Instructors: Dr. Samantha-Jo Caetano
Office hours: Thursdays 11am - 12pm ET (on Zoom) and Thursdays 4pm - 5pm ET (on Zoom)
Preferred pronouns: she/her

Course Information

Course administrative email

Use the email address sta304@course.utoronto.ca for all administrative inquiries or personal inquiries (e.g., notifying the teaching team of a personal circumstance). To ensure your email inquiries do not go to spam/junk, please use your University of Toronto email address (i.e., email addresses from "@mail.utoronto.ca") Please note that the sta304@course.utoronto.ca inbox will not be monitored after April 30, 2025.

Course web page

All materials will be posted on Quercus sta304@course.utoronto.ca. Course materials provided on Quercus are for the use of students currently enrolled in this course only. Distributing course materials to anyone outside of the course is considered unauthorized use.

Teaching assistants

See the course Quercus page for information about TAs, office hours and contact.

Calendar description

Design of surveys, sources of bias, randomized response surveys. Techniques of sampling; stratification, clustering, unequal probability selection. Sampling inference, estimates of population mean and variances, ratio estimation. Observational data; correlation vs. causation, missing data, sources of bias.

Required prerequisites

ECO227Y1 /STA255H1 /STA261H1 /STA248H1 /STA238H1 /STAB57H3 /STA258H5 / STA260H5 /ECO227Y5

Please note that all prerequisites for all STA courses are strictly enforced and your instructor cannot waive them. Any questions about prerequisites should be directed to ug.statistics@utoronto.ca.

Class format

The course is scheduled for 3 hours of lecture per week (per lecture section). The sessions will be delivered as such: Tuesday lectures will be synchronous and in-person, Thursdays will be a synchronous online office hour (on Zoom). We will be using a partial flipped class for most modules in this class. Most weeks, there will be a set of lecture videos or readings available mid-week (posted by Thursday morning 11am ET) to watch prior to the Tuesday synchronous in-person class. The 2 hour time slot (on Tuesdays) will be held synchronously in-person.

Suggested Weekly Routine

Tuesday	Thursday	Friday-Monday
Attend synchronous lecture and work. on upcoming assessment	Watch weekly asynchronous videos and attend office hours.	Complete weekly quiz. Work on homework and/or upcoming assessment.

Note: The instructor will do her best to make all in-person synchronous lectures available virtually too. When possible the synchronous Tuesday lectures will be recorded and made available to all students in the course.

Reference Materials

We will *not* be relying on material from external textbooks. All materials for the course will be provided by the instructor on Quercus.

Below are some supplementary textbooks the may be helpful for referencing throughout the course. Please note, that access to the textbooks are not mandatory in this course. All the textbooks listed below are *freely available* to University of Toronto students, via the UofT library website.

1. Wu, Changbao and Mary E. Thompson, 2020, Sampling Theory and Practice, Springer. ($186 hard copy or FREE digitally on the library website)
2. Lohr, Sharon, 2019, Sampling: Design and Analysis, Taylor & Francis. ($132 hard copy or FREE digitally on the library website)

Course Content

The course will consider the following overarching themes in statistical theory and data analysis:

- Designing a survey or sample that is appropriately gathering information of interest.
- Carrying out a variety of statistical analyses in R to make inference on the data collected from a survey/sample.
- Identifying and implementing different sampling techniques and different study designs and the trade-offs involved in each.
- Identifying sources of bias within a study and comment on a study's design, including its weaknesses, strengths, and appropriate analyses.
- Clearly communicating results of statistical analyses to technical and non-technical audiences.
- Understanding best practices and ethical considerations in survey design, sampling, and observational data collection and analysis.

We will consider various perspectives on these themes, including Bayesian, frequentist, and likelihood approaches. We will consider methods that rely on mathematical thinking and methods that rely on computational thinking, with particular emphasis on computational approaches to analyzing data and understanding statistical methods.

Assessments

The assessments for STA304 Winter 2025 are described below. Your final grade will be calculated as the maximum of Weight Scheme A and Weight Scheme B.

Assessment	Weight Scheme A	Weight Scheme B	Date
Weekly Quiz/Survey	Best 6 out of 10 12% (2% each)	Best 6 out of 10 12% (2% each)	due Mondays (except Jan 6 & Feb 10) at 8pm ET
Weekly In-class Attendance	Best 6 out of 10 12% (2% each)	Best 6 out of 10 12% (2% each)	due Tuesdays (except Jan 7 & Feb 11) during class
Assignment 1	8%	8%	Thurs Jan 30 8pm ET
Midterm Test	20%	0%	Tues Feb 11 (in-person) during lecture
Assignment 2	8%	8%	Thurs Mar 13 8pm ET
Final Exam	40%	60%	TBA

Policy for Quizzes:

- There will be no additional accommodation for missed weekly quizzes beyond the flexibility already built into the grading scheme (i.e., best 6 of 10). Quizzes will be due Mondays and are intended to be a quick way to check your knowledge and understanding of the material posted in the recent module (i.e., released the previous Thursday) before going to the Tuesday lecture.
- Each week there will be either a quiz or a survey assigned. Quizzes will be short (no longer than 15 minutes) and will be each worth a graded total of 2 points. Surveys will not have a time limit (other than being due on the Monday) and may consist of more than 3 questions, but will graded for completion points. Weekly quizzes will take place on Quercus. The quizzes can consist of multiple choice, tick all that apply, matching, numerical answer, fill in the blanks, and/or short answer questions. Please note that most of these questions will be auto-graded by Quercus in its default manner. For example, please note that for tick all that apply questions it will deduct partial points if you select incorrect options and add partial points for selecting correct answers.
- For support on writing Quercus quizzes please visit the link below and read under the "Best Practices and Troubleshooting while taking a Quiz" section.
  https://q.utoronto.ca/courses/46670/pages/student-guide

Policy for In-class attendance:

- There will be no additional accommodation for missed in-class attendance beyond the flexibility already built into the grading scheme (i.e., best 6 of 10). In-class attendance will be due Tuesdays during the lecture times and are intended to be a quick way to check your knowledge and understanding of the material covered in the lecture.
- Each week the attendance will be taken by completing an onlince component on Quercus (i.e., quiz or survey), MarkUs (i.e., submitting code), or using pen and paper.
- please bring a writing utensil (pen or pencil) to each lecture.

Policy for Assignments:

Both Assignments will be open-ended reports. There will be strong focus on communication and writing throughout these assessments. Thus, it is recommended to seek writing support, if needed. Please see the "Writing Resources" section of the syllabus for more information on this.

- Assignment instructions will be posted on Quercus at least one week in advance of the due date.
- We will NOT accept email submissions for Assignments; Assignments must be submitted through Quercus.
- Sometimes life throws unexpected curve balls. We like to be as flexible as possible, within the limits that are inherent to teaching a course this large and with a team of TAs who need to balance their workloads. Do try your best to meet the Assignment deadlines. These deadlines are planned to spread your work out and allow TAs to schedule their grading time. For Assignments, there is a 7-day grace period to submit without penalty. You do not have to request this 7-day extension from our teaching team, you just need to submit your work when you are finished (within the extra 7 days). Please note, if you submit during the grace period, you will receive your grade later than those who submit on time, because your TA will mark it when their schedule allows. If you need an extension beyond 7 days, then you will need to contact your registrar's office and notify them of your personal situation and have them contact our teaching team. Assignments not submitted within the 7 day grace period will receive a grade of 0, unless your registrar's office or the accessibility office made alternative arrangements with the STA304 teaching team.
- If you miss an assignment deadline and need an extension beyond the 7 day grace period already provided, then you must notify the teaching team via email (sta304@course.utoronto.ca) no later than one week after the date of the assessment. Valid reasons for missing an assignment include:
  - Absence declaration via ACORN (emailed to sta304@course.utoronto.ca within 48 hours of completing it). For information on Absence Declaration Tool for A&S students, please see https://www.artsci.utoronto.ca/absence
  - U of T Verification of Illness or Injury Form (VOI) emailed to sta304@course.utoronto.ca within 48 hours of completing it). Please see here for more information.
  - College Registrar's letter
  - Letter of Academic Accommodation from Accessibility Service
- Since Assignments are open ended and writing-based, we will be using a Plagiarism Detection Tool, Turnitin, on the written components of the Assignments. Please see the "Academic Integrity" section of the syllabus for more information on the Conditions of Use of this tool.

Policy for the Midterm Test:

The midterm test will take place in-person in lecture on Tuesday February 11. The midterm test will be 75 minutes and you must write in your registered lecture section. If you do not write in the lecture section you are currently enrolled in, you will receive

a grade of 0. If you miss the midterm, the weight will be shifted to the final exam (i.e., Weight Scheme B). There is no need to contact the instructors/teaching team in cases of a missed test. Please note that if you use the Absence Declaration tool for the for the Midterm Test, then the weight of the midterm will automatically be shifted onto the final exam.

Policy for the Final Exam:

The final exam will be scheduled by the Faculty of Arts and Sciences sometime in the final assessment period in December. The final exam will be cumulative and will be a proctored in-person timed examination. Any scheduling conflicts or petitions to write a deferred exam must go through the Faculty of Arts and Sciences standard procedures. https://www.artsci.utoronto.ca/current/faculty-registrar/exams-assessments Be sure not to make travel plans until the final exam schedule has been officially released by the Faculty of Arts and Sciences

Regrade Requests:

- Mistakes occasionally happen when marking. If you feel there is an issue with the marking of an assessment, you may request that it be re-marked. The course re-mark policy exists to correct mistakes, and any request should clearly identify the error (for example, a question that was not marked, or a total that was incorrectly calculated). Regrade requests must be made via a form (the link will be shared in a Quercus announcement when the grades are posted). Regrade requests will NOT be accepted by email and must satisfy the following requirements:
  - must not be sent within the first 24 hours of the release of the assessment grade,
  - must be received within one week of the date that the marks for the assessment became available,
  - must include your full name and student number, (and group number if applicable), and
  - must give a specific, clear, and concise reason for each request, referring to a possible error or omission by the marker. Re-mark requests without a specific reason will not be considered.

  Please note that your entire test/Assignment may be remarked when submitting a re-mark request. Keep in mind that it is possible for your assessment grade to go down if the regraded mark is lower than your original assessment grade.
- For the final exam, the remark process is handled by the Faculty of Arts and Sciences.
- All regrading will be done by the instructor.

Generative Artificial Intelligence Tools

Generative Artificial Intelligence (AI), and specifically foundational models that can create writing, computer code, and /or images using minimal human prompting, are proliferating and becoming ubiquitous.

Cases in which students *can use Generative AI* in this course:

- Students are encouraged to make use of technology, including generative artificial intelligence tools, to contribute to their understanding of course materials.
- Students may use artificial intelligence tools, including generative AI, in this course as

learning aids or to help produce assignments. However, students are ultimately accountable for the work they submit.

- Students must submit, as an appendix (and in an additional Generative AI statement) with their Assignments, any content produced by an artificial intelligence tool, and the prompt used to generate the content.
- Any content produced by an artificial intelligence tool must be cited appropriately. Many organizations that publish standard citation formats are now providing information on citing generative AI (e.g., MLA: https://style.mla.org/citing-generative-ai/.).
- Students may choose to use generative artificial intelligence tools as they work through the Assignments in this course; this use must be documented in an appendix for each Assignment (more information on how to do this will be included in the Assignment instructions).The documentation should include what tool(s) were used, how they were used, and how the results from the AI were incorporated into the submitted work.
- Note that if you choose to use applications like ChatGPT, the terms of use with any external program may change without notice during the semester.

Cases in which students *can NOT use Generative AI* in this course:

- The knowing use of generative artificial intelligence tools, including ChatGPT and other AI writing and coding assistants, for the completion of, or to support the completion of invigilated, closed book assessments, including the final examination, midterm test, and (in some cases) weekly quizzes, may be considered an academic offense in this course.
- This course policy is designed to promote your learning and intellectual development and to help you reach course learning outcomes.

Academic integrity

Academic integrity is fundamental to learning and scholarship at the University of Toronto. Participating honestly, respectfully, responsibly, and fairly in this academic community ensures that the University of Toronto degree that you earn will be valued as a true indication of your individual academic achievement, and will continue to receive the respect and recognition it deserves.

Familiarize yourself with the University of Toronto's Code of Behaviour on Academic Matters available at http://academicintegrity.utoronto.ca.

Discussion about lecture materials, textbook concepts and course concepts with your classmates and the teaching team is encouraged, but it is expected that you work independently on all individual assessments. Please note, you may not submit for credit any work that was completed by another student (or person). This includes, but is not limited to, partially or fully completed code, communication of solutions, and plagiarism. In particular, you are expected to complete and submit independent work for assignments (that are not group work), the midterm, quizzes and the final exam. Specifically, you are expected to work on individual work, individually. You may discuss lecture materials and general course concepts, but it is expected that you work individually through assessments. You may use code provided by your STA304/1003 instructors without providing a citation. If you use code from any other source, you must provide the source. To protect yourself from potential academic integrity offences, do not share your code and written submissions.

Since assignments are open ended and writing based, we will be using a Plagiarism Detection Tool, Turnitin, on the written components of the assignments. The Conditions of Use is stated as follows: "Normally, students will be required to submit their course essays to the University's plagiarism detection tool for a review of textual similarity and detection of possible plagiarism. In doing so, students will allow their essays to be included as source documents in the tool's reference database, where they will be used solely for the purpose of detecting plagiarism. The terms that apply to the University's use of this tool are described on the Centre for Teaching Support & Innovation web site (https://uoft.me/pdt-faq)."

General Course Policy

Accessibility needs
The University of Toronto is committed to accessibility. If you require accommodations for a disability, or have any accessibility concerns about the course, the classroom, or course materials, please contact Accessibility Services as soon as possible:
accessibility.services@utoronto.ca or http://www.accessibility.utoronto.ca.

Computing
Computational work is a central part of developing statistical thinking and developing facility in the use of computational tools for carrying out simulations and data analysis is a core objective of this course.

We will use R, the R Studio IDE, and R Markdown. All of these are freely available. You need to first install R, and then R Studio. R can be downloaded for free from http://cran.r-project.org. R Studio can be downloaded for free from http://www.rstudio.com/products/rstudio/download/. Additionally, you can also use R Studio through the U of T Jupyterhub, by selecting the RStudio option and logging in with your utorID and password, available here: https://jupyter.utoronto.ca

Some resources for using R and R Markdown:

- The course supplementary notes give guidelines on installing and getting started with R and R Studio.
- A short intro to R workshop is available here:
  https://awstringer1.github.io/ssu-r-workshop/ssu-r-workshop.html
- Hands-On Programming with R by Garrett Grolemund, available here:
  https://rstudio-education.github.io/hopr
- R for Data Science by Hadley Wickham and Garrett Grolemund, available here:
  https://r4ds.had.co.nz
- An R Markdown Cheat Sheet is available at https://rstudio.com/resources/cheatsheets

Practice Problems
There will be weekly quiz problems for practice. Some recommended practice problems may be assigned from the textbooks, but these will not be graded.

Communication
Tentatively, we will be using Zoom for online synchronous meetings (e.g., office hours) in the course. Please ensure you are able to access your UofT Zoom account as you will need to be signed in this way to enter any STA304/1003 course related online Zoom meetings. https://utoronto.zoom.us/.

We will be using Piazza as the platform for discussions related to the course material and assessments. You can find our course page at:
piazza.com/utoronto.ca/fall2023/sta3041010110201.
Students can post anonymously to classmates on Piazza, but the identity of the author of all posts is viewable by instructors.

Be sure to read Piazza's Privacy Policy and Terms of Service carefully. Take time to understand and be comfortable with what they say. They provide for substantial sharing and disclosure of your personal information held by Piazza, which affects your privacy. When you use Piazza, only provide content that you are comfortable sharing under the terms of the Privacy Policy and Terms of Use. With that being said Piazza will still be considered a part of our class and thus all posts and conduct on Piazza must remain professional. Posts regarding personal matters such as inquiries about grades, reporting absences, regrade requests, etc. should be communicated via email (at sta304@course.utoronto.ca) and *NOT* be posted on Piazza. Piazza is intended for students to receive support regarding course information and content and thus should be an overall positive and professional environment.

Again, email is appropriate only for personal matters that can not be shared with the rest of the class. To be fair to all students, we are not able to answer questions about the course material by email. These questions should be asked on the discussion forum (publicly) or during office hours.

Inquires about administrative matters, such as missed tests and re-mark requests, should be sent to sta304@course.utoronto.ca. Please note that email will not be monitored on evenings, holidays or weekends (Toronto time). Additionally, Piazza will not be monitored by the teaching team on on evenings, holidays or weekends (Toronto time). Thus, the teaching team would like to encourage students to respond to each other's questions on Piazza to allow for quicker response times. Depending on the amount of emails please allow a reasonable amount of time for email (and Piazza) responses.

Announcements and other course information will be posted on Quercus.

Course Materials
All course materials are copyrighted. If they are from the textbook, the copyright belongs to the textbook publisher. If they are provided by an instructor (for example, lecture notes, computer code, assignments, tests, solutions) the copyright belongs to the instructor. Distributing materials online or sharing them in any way is a copyright violation and, in some situations, an academic offence.

Additional Resources

Writing Resources
Communication and writing will play an important role in this course, thus I wanted to emphasize some of the writing resources that the university has made available to its students:

- https://writing.utoronto.ca/writing-centres/arts-and-science/
- https://www.artsci.utoronto.ca/current/academic-advising-and-support/english-language-learning

COVID-19 & Mental Health Resources

The Faculty of Arts and Science have put together the following list of Frequently Asked Questions (FAQs) regarding COVID-19:
https://www.artsci.utoronto.ca/covid19-artsci-student-faqs.

Additionally, transitioning from fully online to in-person learning can challenging. Life sometimes throws curve balls and it can be hard to move forward. If you need help regarding mental health, it is good to ask for help. Here are some UofT mental health resources:

- https://prod.virtualagent.utoronto.ca/.
- https://studentlife.utoronto.ca/department/health-wellness/.
- Call Good2Talk. Free, confidential helpline with professional counselling, information and referrals for mental health, addictions and well-being, 24/7/365 1-866-925-5454