STA304H1/STA1003H1: Surveys, Sampling and Observational Data
Fall 2021 (Sept-Dec)

Sections L0101 & L0201
Instructor: Prof. Samantha-Jo Caetano
Office hours: Wednesday 11am - 12pm ET and Wednesdays 3pm - 4pm ET (on Zoom)
Preferred pronouns: she/her

Course administrative email

Use the email address sta304@utoronto.ca for all administrative inquiries, including missed assessments and re-mark requests. Please note that this email address will not be monitored after December 31, 2021.

Course web page

All materials will be posted on Quercus https://q.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 2 hours per week (per lecture section) to be delivered online synchronously. We will be using a flipped class for most lessons in this class. Most weeks, there will be a set of lecture videos available early in the week (posted Monday morning) to watch prior to the Wednesday synchronous class. The 2 hour time slot (on Wednesdays) will be held synchronously through Zoom (links are available in our Quercus page). Typically the 2-hour class meeting on Wednesday will be one hour of synchronous lecture, followed by a second hour scheduled for synchronous, online extra help/office hour.

Suggested Weekly Routine

Monday & Tuesday	Wednesday	**Thursday**
Watch weekly asynchronous videos and work on assigned homework.	Attend synchronous lecture and office hours and extra help.	Complete weekly quiz. Work on homework and/or upcoming assessment.

Note: All synchronous lectures will be recorded and made available to all students in the course.

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

Reference Materials

We will be relying on material from the following textbooks. Please note, that access to the textbooks are not mandatory in this course, but having them for reference is recommended.

1. Wu, Changbao and Mary E. Thompson, 2020, Sampling Theory and Practice, Springer.
   This is the primary reference for the course.
   Available for free as in pdf through the UofT library here.

2. 2. Gelman, Andrew, Jennifer Hill and Aki Vehtari, 2020, Regression and Other Stories, Cambridge University Press.

3. Kohavi, Ron, Diane Tang, and Ya Xu, 2020, Trustworthy Online Controlled Experiments: A Practical Guide to A/B Testing, Cambridge University Press.

4. McElreath, Richard, 2020, Statistical Rethinking, 2nd Edition, CRC Press.

Practice Problems

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

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.

Communication

Tentatively, we will be using Zoom for most synchronous online meetings 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/fall2021/sta3041010110201.
Students can post anonymously to classmates on Piazza, but the identity of the author of all posts is view able 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@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@utoronto.ca. Please note that email will not be monitored on evenings or weekends (Toronto time), as well as Piazza. Depending on the amount of emails please allow a reasonable amount of time for email (and Piazza) responses. Please note that this email address will not be monitored after December 31, 2021.

Announcements and other course information will be posted on Quercus.

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.

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.

Assessment

Assessment	Weight	Date
Weekly Quiz	Best 6 out of 10 3% (0.5% each)	every Thursday (except week of the Test) timed quiz due at 11:59pm ET Thursday
Assignment 1 (Individual)	15%	Friday October 1 at 11:59pm ET
Assignment 2 (Individual)	15%	Friday October 22 at 11:59pm ET
Assignment 3 (Group)	15%	Friday November 5 at 11:59pm ET
Test	20%	Wednesday November 24 (overlaps with lecture time)
Final Project (Proposal)	1%	Friday December 3, 2021 at 11:59pm ET
Final Project (Peer Review)	1%	Wednesday December 8, 2021 at 11:59pm ET
Final Project (Final Report)	30%	Friday December 17, 2021 at 11:59pm ET

Notes:

- No accommodation for missed weekly quizzes beyond the flexibility already built into the grading scheme (i.e., best 6 of 10).
- Assignments will be posted on Quercus at least one week in advance of the due date.
- The test will be held to overlap with the Wednesday November 24 lecture period window, once the test is started you will have a time limit within that window.
- Extensions for assignments may be granted, to a maximum of 3 days. Requests must be made in advance of the assignment due date via the course email (sta304@utoronto.ca). Please note that just because a request is made does not mean the request will be granted. Additionally, if a request is made on the due date it may not be granted. Thus it is

recommended to make a submission before the due date regardless of the extension being granted or not. Requests for extensions made after the due date will not be considered.

- Extensions granted on assignments will not exceed 3 days, to avoid overlapping with upcoming assessments and to ensure grading is completed in a reasonable time frame.
- Extensions for up to 3 days for the final report of the final project may be granted. Requests must be made in advance of the final report due date via the course email. Again, just because a request is made, does not mean that the extension will be granted. Thus it is recommended to make a submission before the due date regardless of the extension being granted or not. Requests for extensions made after the due date will not be considered.
- We will NOT accept email submissions for assessments.
- There will be no extensions granted for the Proposal or Peer Review portion of the Final Project as it will contain a peer-review and other peer-related components which will require on-time submissions.
- Late submissions (without a granted extension) will receive a mark of 0.
- If the test is missed you must contact me via the course email within 72 hours of the missed test. For consideration your email must:
  - be received within 72 hours of the test date,
  - must include 'STA304/1003 Reporting Test Absence' in the subject line,
  - must include your full name and student number,
  - must include a screenshot/photo of your self-declared absence on Acorn, and
  - must include the following two sentences:
    1. "I affirm that I am experiencing an illness or personal emergency and I understand that to falsely claim so is an offence under the Code of Behaviour on Academic Matters."
    2. "I understand that an alternative assessment will be arranged at the instructor's discretion (including an oral exam and/or a make-up assessment in December, after the lecture period)."
- If you miss the test and complete the accommodation procedure correctly (described above), an alternative assessment will be arranged at the instructor's discretion. Note that this alternative assessment may have a different format (e.g., oral assessment) and may be scheduled in December after classes end.
- Mistakes occasionally happen when marking. If you feel there is an issue with the marking of a test/assignment, 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 incorrectly calculated). Requests to correct such mistakes must by form which will be available via a Quercus announcement when the grades are released. For consideration, any re-mark request:
  - 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 accepted.

Please note that your entire test/assignment may be re-marked when submitting a re-marking 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 project, the re-mark process is handled by the Department of Statistical Sciences.

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 test, and the final project. 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.

Writing Resources

Again, communication and writing will play a 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-

COVID-19 & Mental Health Resources

This iteration for STA304/1003 will be running during the COVID-19 pandemic, and will be completely online. There may be times where extensions for students are needed, and/or instructors and TAs may take longer than usual to respond to emails and/or marking needs. It is recommended to please stay active in the course as much as possible (attend lectures, visit office hours, post on Piazza, etc.) and please notify us of needs for extensions or other course related content as early as possible.

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, learning online can be more challenging than learning in-person. If you need help regarding mental health, please do not hesitate to find support. 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