

# STA304H1 F

## Surveys, Sampling and Observational Data (formerly STA322H1)

### Fall 2024 Syllabus

#### Course Meetings

##### STA304H1 F

Section	Day & Time	Delivery Mode & Location
LEC0101	Thursday, 3:00 PM - 6:00 PM	In Person: AH 100

Refer to ACORN for the most up-to-date information about the location of the course meetings.

#### Course Contacts

**Instructor:** Rohan Alexander

**Email:** [rohan.alexander@utoronto.ca](mailto:rohan.alexander@utoronto.ca)

#### Course Overview

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.

The work of applied statisticians, regardless of their specific job title and area of application, is the most important and exciting work in the world right now. The ability to gather data, analyse it, and communicate your understanding of the underlying process is incredibly valuable. In this course you will learn and apply the essentials of this.

We focus on surveys, sampling and observational data. The very stuff of statistical science! We will approach these topics from a practical perspective. You will actually run surveys and learn how messy it is to put one together. You will learn how to think about sampling, how to implement it, and why the details matter. You will forecast an election. And you will conduct original research. More generally, you will learn how to obtain and analyse data and use it to make sensible claims about the world.

## Marking Scheme

Assessment	Percent	Details	Due Date
<b>In-class group reflection exercise</b>	12%	These largely draw on end-of-chapter “Tutorial” questions. Done in groups randomly created by the instructor. Only best three of six count.	2024-09-05,2024-09-12,2024-09-19,2024-10-03,2024-10-10,2024-11-14
<b>Term papers</b>	30%	Only best one of two term papers counts, but both should be submitted. Term paper 1 is individual. Term paper 2 is groups of 1-3 and will involve forecasting the US election.	2024-11-04,2024-09-27
<b>Conduct peer review</b>	6%	You are expected to provide peer review, via GitHub Issues, for three peers. Papers will be distributed by a spreadsheet—add a link to the Issue/PR to a paper and submit your review to Quercus. You will only have 24 hours to do this. Students are not assigning grades to other students, but are instead getting the mark based on the quality of the feedback they provide to other students.	2024-09-25,2024-10-23,2024-11-27
<b>LLMs prompting competition</b>	1%	Experiment to develop a sampling-based approach to pick an optimal LLM prompt.	2024-10-04
<b>Mid-term</b>	10%	An in-class mid-term. Questions largely draw on end-of-chapter “Quiz” questions.	2024-10-17
<b>Complete survey</b>	1%	Complete a survey to better understand how you used LLMs in this course.	2024-11-29
<b>Final paper</b>	40%	Due at 5pm.	2024-11-29

Please be aware that there are two industry-sponsored awards for final papers written in this class:

- "Investigative Journalism Foundation Award for Best Paper" - The Investigative Journalism Foundation is sponsoring a prize for the best final paper written using their procurement dataset. An IJF reporter will come to class to discuss the details of this later in the term.
- "Open Data Toronto Award for Best Paper" - The City is sponsoring a prize for the best final paper written using data from their Open Data Portal.

## **Late Assessment Submissions Policy**

If no extension has been granted and no accommodation applies, then late submissions will not be accepted.

## **Policies & Statements**

### **Late/Missed Assignments**

#### **Late submissions**

If no extension has been granted and no accommodation applies, then late submissions will not be accepted.

#### **Missed submissions**

##### *In-class group reflection exercise*

Only the best three of six count, and so if some are missed then it is not an issue. If you have a accommodation or situation that makes even this impossible, then please consult your faculty/department/college advisor and then email me.

##### *Term Paper I*

Please submit something for Term Paper I, even if it gets zero. That one is done individually, while the other term paper is (optionally) done in groups. If you have a situation or accommodation that makes it impossible for you to submit something for Term Paper I, then in that situation Term Paper II must be done individually to ensure fairness with the rest of the class. Again, please consult with your faculty/department/college advisor and then email me.

##### *Term Paper II*

If you have an accommodation or situation which means you cannot submit Term Paper II, then the mark for Term Paper I will be used for this 30%.

If you do not submit Term Paper I and Term Paper II, then the 30% will be distributed: 15% to the mid-term, and 15% to the final paper.

##### *Peer review*

There are many opportunities to get the peer review marks through Term Papers I and II and the Final Paper, so if you cannot do any for a particular paper, then just do the others. If you

have a situation or accommodation that makes this impossible, then please consult your faculty/department/college advisor and then email me.

### *Mid-term*

If you miss the mid-term, due to a medical situation or accommodation, then you will have the opportunity to do it in Week 8.

### *Final paper*

The Final Paper is a critical piece of assessment. If you have a situation or accommodation that makes it impossible for you to submit the final paper, then please email me before Week 11, and cc your faculty/department/college advisor so that we can work out an alternative plan. If you have an emergency situation, then please consult your faculty/department/college advisor and then email me.

## **Plagiarism Detection Tool**

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>).

## **Students with Disabilities or Accommodation Requirements**

Students with diverse learning styles and needs are welcome in this course. If you have an acute or ongoing disability issue or accommodation need, you should register with Accessibility Services (AS) at the beginning of the academic year by visiting <https://studentlife.utoronto.ca/department/accessibility-services/>. Without registration, you will not be able to verify your situation with your instructors, and instructors will not be advised about your accommodation needs. AS will assess your situation, develop an accommodation plan with you, and support you in requesting accommodation for your course work. Remember that the process of accommodation is private: AS will not share details of your needs or condition with any instructor, and your instructors will not reveal that you are registered with AS.

## **Re-marking Policy - Timeline and Protocol**

Marking mistakes happen and I want to correct those. Requests to have your work re-graded will **not** be accepted within 24 hours of the release of grades. This is to give you a chance to reflect. Similarly, requests to have your work re-graded more than two weeks after the release of the grades will not be accepted. This is to ensure the course runs smoothly.

Inside that period if you would like to request a re-grade, please email [rohan.alexander@utoronto.ca](mailto:rohan.alexander@utoronto.ca) and use the subject line "STA304: re-grade request". Please specify where the marking mistake was made. The entire assessment will be re-marked and it is possible that your grade could reduce.

Plenty of students get 0 on the first paper, but go on to get an A+ overall in the course. The nature of the work in this course requires students to adjust from what is expected in other courses, and the forgiving assessment weighting is designed to allow this.

### **Accommodation for Personal Reasons**

There may be times when you are unable to complete course work on time due to non-medical reasons. If you have concerns, speak to me or to an advisor in your College Registrar's office; they can help you to decide if you want to request an extension or other forms of academic consideration. They may be able to email your instructors directly to provide a College Registrar's letter of support and connect you with other helpful resources on campus.

### **Attendance and Absence Declaration**

A reminder that FAS has a policy on declaring absence which may be relevant for assessment purposes: <https://www.artsci.utoronto.ca/current/academics/student-absences>