
STA303 Methods of Data Analysis 2

Department of Statistical Sciences, University of Toronto
Winter 2025

SECTION DETAILS

Prof. Patrick Brown	Prof. Katherine Dagnault	
<u>L0101</u> Lecture: Mondays 9:00AM - 11:00AM in AH100 Office hour: Wednesdays 9:00AM - 10:00AM	<u>L0201</u> Lecture: Mondays 11:00AM - 1:00PM in AH100 Office hour: Wednesday 12:00PM - 1:00PM	<u>L5101</u> Lecture: Mondays 5:00PM - 7:00PM in AH100 Office hour: Wednesday 2:00PM - 3:00PM

**** All communication with the course instructors should be directed to the course email at sta303@course.utoronto.ca. This email is monitored during regular business hours only (i.e. Monday-Friday from 9am - 5pm). Please allow up to 24-48 hours for a response.**

OVERVIEW

This course introduces a variety of statistical models for situations in which the assumptions of Normal linear regression models do not hold, with a focus on the application of these models to data using the R statistical software. The statistical models covered include: one- and two-way analysis of variance, generalized linear models, linear and general linear mixed models, and generalized additive models.

Prerequisites: STA302/STAC67 and all prerequisites of STA302/STAC67

LEARNING OUTCOMES

By the end of this course, students should be able to:

1. Select an appropriate statistical model for specific datasets and research goals.
2. Write and execute R code for the statistical models covered.
3. Correctly interpret the results of the statistical models covered for a variety of audiences.
4. Recognize limitations and benefits of the statistical models covered as applicable in real-world contexts.
5. Identify, describe and implement estimation methods and algorithms for different statistical models covered.

COURSE FORMAT AND MATERIALS

Course Components:

Overall Format: All three sections of this course are offered in a flipped classroom format with materials created collaboratively between the two instructors. Students are responsible for watching up to 1 hour of pre-class videos prior to in-person class time on Mondays. The asynchronous materials will be available on Tuesdays. Class time on Mondays will involve a combination of coding demonstrations, additional lecturing, and student activities/exercises. It will be the student's responsibility to come to class prepared and to keep up with the content from week to week.

Weekly Office Hours: Instructors and teaching assistants will hold weekly office hours in a combination of in-person and online formats. Confirmed schedules and format (online or in person) will be posted on Quercus early in the semester. You are encouraged to visit office hours whenever you have questions about content or assessments.

Discussion Forum: We will be using the Quercus Discussion Board as an online discussion forum, which can be accessed through the Quercus course page. **All questions about course material should be posted here or asked during TA/instructor office hours.** The instructor and TAs will monitor the board and will help answer questions but students are encouraged to answer posts and help their fellow classmates.

Sample Weekly Schedule: Find below a suggested workflow for each week in this course. We recommend creating your own schedule for the course to ensure you don't fall behind on the asynchronous material.

Mondays	Tuesdays	Wednesdays	Thursdays	Fridays	Saturdays	Sundays
In-person lectures Complete Post-Class Reflection by 11:59 PM ET	<ul style="list-style-type: none">- Watch Pre-Class Videos and Work on Pre-Class exercises- Attend instructor and TA office hours- Post questions on discussion forum- Work on any assessments or practice problems				Complete Pre-Class Quiz Questions by 11:59 PM ET	

Course Materials and Resources

Course page: There is a [common Quercus page](#) for all three sections of this course. All materials will be posted here weekly. Students should visit the course page regularly to remain updated on the course material and all important announcements.

Statistical software: The R statistical software will be used to illustrate and apply the course content and will be required for a variety of assessments. Students are encouraged to access R and RStudio through the [UofT](#)

[JupyterHub](#). You will need your UofT credentials to log in. All code needed for activities and assessments will be available in lecture slides/videos, in-class demonstrations, activities and notes uploaded to Quercus.

Suggested (but not required) Textbooks: There is no formal textbook for this course. However, should a student wish to refer to a textbook for additional explanation or preparation, we recommend [Beyond Multiple Linear Regression by Roback and Legler](#). Any additional suggested reading will be provided each week on Quercus.

GRADING SCHEME

Each student's final grade will be computed according to the below grading scheme. No special rounding rules or individual grade adjustments (e.g. to meet GPA cut-offs, minimal requirements for programs, etc.) will be used to calculate course grades. No special reweighting of assessments or extra work will be accepted to account for perceived poor performance, nor to account for any assessment(s) that have been missed without accommodation. There are no exceptions to these policies.

Assessment	Date Due	Weight
<u>Engagement Activities:</u>		<u>Total: 5%</u>
Pre-Class Quizzes (best 7 out of 9)	Saturdays by 11:59 PM	2%
Post-Class Reflections (best 7 out of 9)	Mondays by 11:59 PM	2%
Ethics Pre/Post-Survey	Pre-Survey available from January 17-20 Post-Survey available from March 27 to April 4	1%
<u>Assignments:</u>		<u>Total: 20%</u>
Assignment 1	Friday February 7th, 2025 at 8:00 PM	10%
Assignment 2	Friday March 21st, 2025 at 8:00 PM	10%
Term Test	TBD (as per departmental request)	25%
Final Exam	TBD (scheduled by FAS in April 2025)	50%

EVALUATION BREAKDOWN

Participation

Pre-Class Quizzes: After completing the pre-class video content (see [Course Components](#)), students will complete a short set of questions and exercises with the intention of checking one's understanding of the material prior to attending in-person class time. Students will need to submit the answers to the questions/exercises through a Quercus quiz. The deadline for submitting the quiz each week is Saturday by 11:59 PM. Each quiz is graded on completion only, meaning that correctness of the answers is not assessed. However, an answer demonstrating a reasonable attempt to complete the work must be submitted for the quiz

to be considered complete. Only 7 out of 9 pre-class quizzes will be counted towards this portion of the grading scheme.

Post-Class Reflections: Following each in-person class on Mondays, students will be asked to complete a short reflection survey through Quercus which will be due by 11:59 PM each Monday. The survey will ask questions prompting students to reflect on the week's material and in-class activities, and to think about their studying and comfort with the materials each week. This practice has been shown to be beneficial to student learning. Only 7 out of 9 reflection surveys will be counted towards this portion of the grading scheme.

Ethics Pre-/Post-Survey Completion: Two surveys will be given out (one in the second week and one in the last week of the semester) to students in this course and in another STA course. These surveys are intended to collect information about student's understanding of ethics in statistics and are being collected as part of a research study. Students must complete both surveys in their entirety to receive the 1% completion point. Students may opt out of allowing their survey responses to be used as part of the research study and for any published work that may result from this project. Opting out of participating in the research study will have no impact on receiving the completion marks.

Assignments

There will be two assignments due during the term. Each will take the form of a short data analysis in which students will demonstrate their ability to select an appropriate model from the course content, apply it to a real dataset, and answer a research question using the results. A short report summarizing the analysis and results will be submitted to Quercus for evaluation. Specific details, instructions and rubrics will be provided at a later date. Students will have the option to complete the assignment in pairs (i.e., groups of two).

Term Test

A common term test will take place on **February 14th, 2025 from 5:00 - 7:00 PM**. The test will cover all course content up until the last class meeting on February 10th. The test will be approximately 1 hour and 40 minutes in duration. More details will be provided regarding the format and coverage closer to the test date.

Any student with an academic conflict (e.g. regularly scheduled tutorial or lecture) must contact the teaching team no later than **January 24th, 2025 to advise us of this conflict. No accommodations can be guaranteed for a conflict reported after this date.**

Final Exam

A common final exam will take place during the April 2025 Final Exam Period. The exam will be scheduled by the Faculty of Arts and Science and becomes available in March 2025. The exam will be cumulative (i.e., covering all content throughout the term). More details will be available at the end of the semester.

ASSESSMENT POLICIES

Late Assessments/Extension Requests

Engagement Activities: Due to the flexible scheme (counting only 7 out of 9), long availability period, and nature of these activities (completion only, not correctness), there are no extensions or late acceptance of these activities.

Assignments: Students have two (2) extensions available to them to use on the assignments. Each extension allows students to take an additional 3 days (72 hours) to submit an assignment. Extensions cannot be combined (i.e. cannot create one 6-day extension) and any unused time from an extension will be forfeited (i.e. cannot use 2 days on assignment 1 and then use 4 days on assignment 2). Students should NOT notify the instructors when they intend to use an extension. No additional extensions will be granted and no late submissions beyond the extension will be accepted.

Missed Assessments

If you experience a prolonged absence due to illness or emergency that prevents you from completing any number of assessments, please contact your College Registrar as soon as possible so that any necessary arrangements or academic advising can be made as quickly as possible.

Engagement Activities: Due to the flexible scheme (counting only 7 out of 9), long availability period, and nature of these activities (completion only, not correctness), there are no additional accommodations available for a missed engagement activity.

Assignments: Only one (1) missed assignment will be accommodated by shifting the weight of the assignment equally to the remaining assignment (15%) and to the final exam (55%). In the event that a student does not turn in the other assignment, they will receive a zero on that assignment. There will be no make-up opportunities for missing both assignments.

Term Test: If a student is experiencing a serious personal illness or emergency on the date of the test, the student must submit [an Absence Declaration form on ACORN](#) or [other accepted documentation](#) (such as Verification of Illness form, letter from College Registrar, or Letter of Academic Accommodation from Accessibility Services), and notify the teaching team at sta303@course.utoronto.ca no later than one week after the date of the test. Only one make-up test will be available for students who provide legitimate documentation for missing the original test. The date of the make-up midterm is tentatively **February 28th, 2025 from 5:00 PM - 7:00 PM.**

IMPORTANT: students may only use the Absence Declaration once per academic term (e.g., the Winter term) for a maximum period of 7 consecutive calendar days. See [A&S Student Absences](#) for additional information on eligibility. The once-per-term limit is set by the ACORN Absence Declaration Tool. Once a declaration is

submitted, students will be restricted from using the tool to declare any further absences in that term. Students may then submit a Verification of Illness form completed by a physician or a letter from a College Registrar.

Final Exam: Should a student miss the final exam, they should follow the instructions outlined on the [Faculty of Arts and Science Final Exam](#) webpage. All arrangements for a missed final exam are handled by the Faculty, not the instructor.

Regrade Requests

Regrade requests will be accepted for assignments and the term test. Regrade requests must provide a justification for where there exists a grading error and/or how the work meets the grading rubric. These justifications must further be backed up with concrete references to the course material. All regrade requests will be accepted through a form available on the Quercus course page and will be accepted no later than one week after the grade for that assessment is released. No regrade requests will be accepted by email or after the 1 week deadline. The instructor further reserves the right to re-evaluate the assessment in its entirety (i.e. grades can go up, down, or remain unchanged). Please allow a few weeks for regrade requests to be processed by the instructor.

ACCEPTABLE USES OF GENERATIVE AI

ChatGPT and other generative AI are freely available tools that can perform a variety of functions for us. However, it's important to understand how such tools are allowed to be used in this course. Acceptable uses of generative AI in this course include:

- Editing or rephrasing written work that has already been written by the student to improve the syntax, grammar and overall readability of the work.
- Synthesizing or explaining course concepts while learning and studying to contribute to their understanding of the course material
- Looking up appropriate syntax of individual R functions for use in a data analysis or for understanding errors that may arise when running R code.

However, the work turned in by students must ultimately be their own and students will therefore be accountable for the work they turn in. Unacceptable uses of generative AI in this course include:

- Copying from any generative artificial intelligence applications, including ChatGPT and other AI writing and coding assistants, for the purpose of completing assignments in this course.
- Producing an entire data analysis, written report, or any other piece of work meant for grades.

In summary, generative AI like ChatGPT can be really helpful in your learning process and to improve skills valued in the workplace. However, it cannot be used as a substitute for learning and material produced from these tools should not be passed off as your own. This would be considered academic misconduct (see below).

The instructor therefore reserves the right to ask students to explain their work and their process for creating their assignment.

ACADEMIC INTEGRITY

The University treats cases of plagiarism and cheating very seriously. It is the students' responsibility for knowing the content of the University of Toronto's [Code of Behaviour on Academic Matters](#). All suspected cases of academic dishonesty will be investigated following procedures outlined in the above document. If you have questions or concerns about what constitutes appropriate academic behaviour or appropriate research and citation methods, you are expected to seek out additional information on academic integrity from your instructor or from other institutional resources (see <http://academicintegrity.utoronto.ca/>). Here are a few guidelines regarding academic integrity:

- Using ChatGPT and other generative AI for any purpose not outlined above.
- Being dishonest when reporting an illness or personal emergency to get an extension or accommodation is an academic offence.
- You may consult class notes/lecture slides during take-home assessments, however sharing or discussing questions or answers with other students is an academic offence.
- Students must complete all assessments individually. Working together is not allowed unless otherwise specified.
- Paying anyone else to complete your assessments for you is academic misconduct.
- Completing assessments for another student is academic misconduct.
- Sharing your answers/work/code with others is academic misconduct.
- Using sources external to the course (anything not on Quercus) on an assessment is an academic offence.

All work that you submit must be your own! You must not copy mathematical derivations, computer output and input, or written answers, etc. from anyone or anywhere else. Unacknowledged copying or unauthorised collaboration will lead to severe disciplinary action, beginning with an automatic grade of zero for all involved and escalating from there. Please read the UofT Policy on Cheating and Plagiarism, and don't plagiarise.

ACCESSIBILITY NEEDS

The University of Toronto offers academic accommodations for students with disabilities. If you require accommodations, 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://accessibility.utoronto.ca>.

TENTATIVE SCHEDULE

Find below a tentative schedule of topics for the course. Note that this schedule is subject to change at the instructor's discretion.

****In this course, a week runs from Tuesday to the following Monday.**

Week #	Pre-Work Available	In-Class Debrief	Content
0	NA	Jan. 6	Course Introduction and Motivation
1	Jan. 7 - Jan. 12	Jan. 13	Analysis of Variance (one-way and two-way)
2	Jan. 14 - Jan 19	Jan. 20	Generalized Linear Models 1
3	Jan. 21 - Jan. 26	Jan. 27	Generalized Linear Models 2
4	Jan. 28 - Feb. 2	Feb. 3	Using Contrasts and Presenting Results
	NA	Feb. 10	Term Test Review
5	Feb. 11 - Feb. 23	Feb. 24	Mixed Models 1
6	Feb. 25 - Mar. 2	Mar. 3	Mixed Models 2
7	Mar. 4 - Mar. 9	Mar. 10	Generalized Additive Models 1
8	Mar. 11 - Mar. 16	Mar. 17	Generalized Additive Models 2
9	Mar. 18 - Mar. 23	Mar. 24	Case-Control and Biostatistics Models
	NA	Mar. 31	Final Exam Review