UNIVERSITY OF WATERLOO

Published January 5, 2025

STAT 341 Winter 2025

Computational Statistics and Data Analysis

Section 001, 002, 101

Class Schedule

Course	Meet Days	Meet Time	Location	Instructor(s)
STAT 341 001 [LEC]	**Mon, Wed** *Jan 6 - Apr 4*	04:00PM - 05:20PM	EIT 1015	Reza Ramezan rramezan@uwaterloo.ca
STAT 341 002 [LEC]	**Mon, Wed** *Jan 6 - Apr 4*	10:00AM - 11:20AM	EIT 1015	Reza Ramezan rramezan@uwaterloo.ca
STAT 341 101 [TUT]	**Thursdays** *Jan 6 - Apr 4*	04:30PM - 05:20PM	EXP 1689	Reza Ramezan rramezan@uwaterloo.ca

schedule data automatically refreshed daily

Instructor & TA (Teaching Assistant) Information

Instructor:

- Reza Ramezan
- Office: M3 3107
- Email: rramezan@uwaterloo.ca
- Office Hours:
  - Tuesday 9:30-11:00 am
  - Thursday 2:30-4:00 pm
  - Or by appointment

TAs:

- Bryn Crandles
- Megan French
- Rohit Gajendragadkar
- Jessica Lauren Leclair
- Yicheng Qin
- Yan Yu
- Daniel Zhang

(TAs' office hours will be posted on LEARN):

Lectures and Tutorials

Course content will be delivered through two in-person lectures and one in-person tutorial each week, as outlined in the schedule above. Lectures will feature the instructor presenting and annotating typeset notes, along with demonstrations of analyses in R. Tutorials will provide a deeper focus on working through examples to help you with your assignments. Tutorials may be delivered by the instructor and/or teaching assistants. Some tutorials may be used as lectures. All relevant notes, datasets, and R files for each week will be posted on LEARN by Monday of that week at the latest.

Course Homepage

The course homepage is on LEARN (https://learn.uwaterloo.ca [1]). You are expected to visit this webpage regularly to download course material. Important announcements may be posted on LEARN.

Course Communication

The online platform Piazza will be used for class discussions and announcements. This platform is designed to provide quick and efficient support from classmates, TAs, and the instructor. While personal matters can be communicated via email, Piazza should be the primary platform for Q&A outside of class. The teaching team will respond to questions Monday to Friday, 9 AM to 5 PM (excluding holidays), typically within 24 hours. Click the link below to access Piazza.

Click here to access Piazza for STAT 341 [2]

Course Description

Calendar Description for STAT 341

A computationally focused approach to statistical reasoning in the context of real data. Functional programming in R and algorithms will be used to define interesting attributes of finite populations and their sampling characteristics. Computational approaches to inductive inference and the assessment of predictive accuracy.

View requirements for STAT 341 [3]

The background knowledge from STAT 230/231 and MATH 237 is necessary for this course.

Learning Outcomes

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

Understand the fundamentals of population-based statistical analyses, as well as implicitly/explicitly defined population attributes
Understand the basics of sampling and the underlying ideas/mechanisms of probability sampling and related issues
Understand the basics of inductive inference including target and study populations, measurements, comparing subpopulations, random intervals and coverage probabilities, and resampling methods such as bootstrap
Understand the fundamentals of prediction and how its accuracy is measured
Generate high-quality visualization and inferential statistics R output for the topics above
Prepare reports that communicate your statistical analyses using LaTeX and R Markdown.

Tentative Course Schedule

Week	Topics	Deadlines
1. Jan. 6	Introduction to the course Understanding populations Explicitly defined attributes (properties) location and scale invariance/equivariance	
2. Jan. 13	Influence and sensitivity Graphical attributes Power transformations and linear relationships	
3. Jan. 20	Order, rank, and quantiles Implicitly defined Attributes Outliers and robust regression	Assignment #1 will be released Jan. 20
4. Jan. 27	Gradient Descent and its variants Gradient descent and robust regression Solving system of equations	Assignment #1 Due Jan. 31
5. Feb. 3	Newton-Raphson method Iteratively re-weighted least-squares Samples: all possible samples, consistency Comparison across attributes	
6. Feb. 10	Selecting samples Quantifying sample error Sampling mechanisms and inclusion probabilities Estimating totals Horvitz-Thompson estimator	Tutorial Test #1 Feb. 13
7. Feb. 17	**Reading Week** (No classes, No Office Hours)	
8. Feb. 24	HT estimator (cont'd) Sampling design Inductive inference Comparing sub-populations	Assignment #2 will be released Feb. 24
9. Mar. 3	Anatomy of tests of significance A t-like discrepancy measure	Assignment #2 Due March 7
10. Mar. 10	Multiple testing Sampling distributions Random intervals Student-t-based intervals	
11. Mar. 17	Introduction to resampling methods Non-parametric bootstrap Bootstrap confidence/percentile intervals	Tutorial Test #2 March. 20
12. Mar. 24	Variations of bootstrap Parametric bootstrap Bootstrap in regression Accuracy of prediction	Assignment #3 will be released Mar. 24

13. Mar. 31	Measuring inaccuracy Multiple samples and prediction Decomposition of APSE, bias-variance trade-off Training and Test sets K-fold cross validation	Assignment #3 Due Apr. 4

Texts / Materials

Note: Any prices provided in course outlines are best estimates based on recent online prices and do not include shipping or taxes. Prices may vary between retailers.

No materials required.

There is no required text for this course. An older version of the notes is posted here [(4)], which do not include the most recent updates of the material. The extensive lecture notes posted on LEARN have been broken into weeks, sections, and subsections. To facilitate everyone's different approach to learning. I have tried to create various ways to interact with the notes. Each folder will contain five files, e.g. for section **2.1- Populations** we have:

2.1 – Populations.html	A complete set of lecture notes, ideally for somebody who likes to read or listen during lectures.
2.1 – Populations.pdf	A partially filled set of lecture notes, ideally for somebody who likes to write during lectures.
2.1 – Populations_annotated.pdf	The annotated slides from lectures.

Note that the material discussed on the board will NOT be included in the posted annotated slides.

Exercises: There are problems with solutions posted on LEARN grouped by the following sections; Inference, Populations Explicit, Populations Implicit, Samples, and Prediction. These problems are partially compiled from previous semesters' assignments and tests.

All material (lectures, slides, exercises, assignments, R codes, solutions, extra materials, handouts, data sets, and grades) will be posted on LEARN. If LEARN is down, an announcement will posted on Piazza and instructions will be posted on the instructor's webpage at www.neuroinformatics.ca [(5)].

Student Assessment

Component	Value
Assignments	18%
Tutorial Tests	32%
Final Exam	50%

To earn a grade of 50% or higher in the course, students must achieve at least 50% in both the assignment category and the proctored tests category. The assignment category grade is the average of the three assignments, while the test category grade is a weighted average of 1/3 x (average score from tutorial tests) + 2/3 x (final exam score).

Assignments (18%):

- There will be three (3) assignments, each worth 6% of your final grade.
- You will have about two weeks to complete each assignment – the release dates and due dates are shown in the tentative schedule above.
- Assignments will be submitted electronically via Crowdmark by 10:00 pm ET on the due dates.
- While you can discuss the questions of the assignment with other students, you are expected to complete these assignments **independently** and submit your own work. Cheating is a serious offence and will be treated as such. Please refer to the Academic Integrity section below.
- Assignments must be completed using **RMarkdown**.
- If you fail to submit an assignment and have a **valid** reason with supporting documentation, the weight from that assignment will be shifted to the final exam. In the case of a 2-day self-declared absence (see below), a 2-day extension will be provided.
- If your reason for incompletion is not deemed valid, you do not have supporting documentation, or you haven't self-declared an absence, you will receive a zero grade. Note that for your circumstance to be deemed valid, you must have been incapacitated for the majority of the timeframe between the assignment's release date and the due date.
- Assignments submitted up to 24 hours late will receive a 50% late penalty (i.e., the grade will be halved). Assignments submitted more than 24 hours late will receive a grade of zero.
- If you submit your assignment to Crowdmark on time but make any errors in the submission (e.g., uploading your Question 3 solution in the Question 1 box), a 5% deduction will be applied to your total assignment score.

Tutorial Tests (32%):

- There will be two (2) closed-book tutorial tests, each 50 minutes long and each worth 16% of your final grade. These tests are held during the Thursday tutorial class on the dates listed in the tentative schedule above.
- The tests will evaluate your comprehension of the course material and may consist of a series of short answer calculations, short answer written responses, multiple choice questions, proofs and R output interpretation.
- If a test is missed due to a self-declared absence, or for a **valid** reason with supporting documentation, your test score will be imputed using the method described in the next bullet point. If your reason for missingness is invalid or you do not have supporting documentation, you will receive a zero.
- If a missed test is being accommodated, your test score of 0 will be replaced by the quantile of that test's grade distribution matching *your* quantile in the final exam grade distribution. For instance, if your final exam score is the 75th percentile and you missed Test 1, your Test 1 score will be the 75th percentile of the Test 1 scores. Relative to simply shifting a missed test's weight to the final exam, this ensures that students who miss a test are not unfairly advantaged or disadvantaged if the final exam is much easier or harder than the missed test.
- **Your instructor will not offer "make-up" tests for any reason.**

Final Exam (50%)

- There will be a 2.5-hour, closed book, cumulative final exam during the Final Examination Period: **April 9 – April 25**. Please refrain from booking end-of-term travel before the actual exam date is scheduled.
- The format of the final exam will mimic that of the two tests. Specifically, short answer calculations, written responses, multiple choice questions, proofs, and R output interpretation can all be expected.
- If the final exam is missed for a **valid** reason, and you provide supporting documentation, you may be provided an INC grade or scheduled to write a make-up exam during the Registrar's Office Winter 2025 Make-up Exam Session. An INC grade is rarely granted and will be considered on a case-by-case basis, with decisions based largely on your performance throughout the semester.
- If your reason for missing the final exam is not valid and/or you do not provide supporting documentation within 24 hours of the final exam, you will receive a "did not write" (DNW) grade, which counts as a final grade of 32 in transcript averages.

Other Comments about Assessments:

- **Remark Policy:** If you have a dispute with your grade on an assignment or test, it may be submitted to be remarked within 1 week of the assignment/ test being returned to you. Note that the entire assignment/ test is then subject to be remarked. Remark requests must be directed to our Instructional Support Coordinator, Pam Hitihamillage (pbandara@uwaterloo.ca) and the instructor (rramezan@uwaterloo.ca) must be copied on your email. The subject line of remark request emails must be **STAT 341: Remark Request for Assessment (your userID)**, for example, STAT 341: Remark Request for A1 (rramezan) or STAT 341: Remark Request for Test2 (rramezan). Please note that an automatic filter for such emails is used, so failure to follow the subject line instructions may result in your request being ignored. In your email, you must explain clearly which questions are marked unfairly and why you believe that you deserve more points for those questions. A sample solution will be posted on LEARN. Please review these solutions **before** emailing a remark request.
- **Self-Declared COVID-19 Related Absences:** You may self-declare one absence per term during the formal lecture period for influenza-like illnesses. You must both self-declare this in Quest **and** email your instructor within 48 hours of the self-declaration to confirm that you seek accommodation. Accommodations for missed assessments are outlined above.
- **Self-Declared Short-Term Absences:** You may self-declare one short-term (2-day) absence per term for any reason. You must both self-declare this in Quest **and** email your instructor within 48 hours of the self-declaration to confirm that you seek accommodation. Accommodations for missed assessments are outlined above.
- **Other Illnesses:** For accommodations due to all other kinds of illness, a University of Waterloo Verification of Illness Form [(6)] (VIF) is the required documentation. See the Math Faculty's Accommodation for Illness or Extenuating Circumstances [(7)] webpage for more information.

Assignment Screening

No assignment screening will be used in this course.

Administrative Policy

Course Software: R is a language and a free statistical computing and graphics software that we will use extensively throughout the course. All computational lecture examples, tutorial exercises and assignment problems will be done in R (via the RStudio development environment). R and RStudio are available on all Math Faculty Computing Facility environments. If you have a personal computer or laptop it is in your best interest to install both R and RStudio. You can download both of them for free here [(8)].

- You are welcome to use R libraries and/or packages such as tidyverse, lubridate, purrr, e.g. dplyr in addition to the base functions. Any packages like those, that help with managing data, etc., are fine. The only time you may run into trouble is if you call a function from a package to handle a task you have been asked to code out explicitly. For example, if you are asked to code a function that calculates the mean of a vector in R, and your solution is to just call the function mean(), then you will not receive any marks.

Communication: When you email your instructor, make sure you include the following information in your email:

- Course: STAT 341
- The name with which you are registered in the course (your official name)
- Your student ID number AND your UW/Quest user ID

I know this sounds like a lot, but it helps me find information on Crowdmark or LEARN for you faster.

Including non-English characters in your email (sometimes imposed by installing a non-English operating system) prompts UW servers to spam your emails for the instructor. These characters most often appear in (not exclusively) the

subject line, From, To, and Date/Time portions of previous communications if you reply to an email. [N.B. If you have read this far in the course syllabus, I'd like it if you could send me an email with the subject line 'STAT 341 MEMES' and one of your favourite pictures from the Internet. This will make me happy to know some people read the course syllabus.] As a result, I may not receive your email if you fall in this category. You must use the web-based email platform if you suspect that you have non-English characters in your emails imposed by your computer, phone, tablet, etc.

Course Material IP and Generative AI: At no point should any student post any course material including, but not limited to, recorded lectures, slides, practice problems, assignments, tests, solutions, etc. to any website except LEARN and Piazza. All course material is the intellectual property of the instructor and the University of Waterloo.

This course includes the independent development and practice of specific skills related to computational statistics and data analysis. Therefore, the use of Generative artificial intelligence (GenAI) trained using large language models (LLMs) or other methods to produce text, images, or code, like Chat GPT, DALL-E, or GitHub CoPilot, is not permitted in this course. Unauthorized use in this course, such as running course materials through GenAI or using GenAI to complete a course assessment is considered a violation of Policy 71 [(9)] (plagiarism or unauthorized aids or assistance). Work produced with the assistance of AI tools does not represent the author's original work and is therefore in violation of the fundamental values of academic integrity including honesty, trust, respect, fairness, responsibility and courage (ICAI [(10)], n.d.).

You should be prepared to show your work. To demonstrate your learning, you should keep your rough notes, including research notes, brainstorming, and drafting notes. You may be asked to submit these notes along with earlier drafts of their work, either through saved drafts or saved versions of a document. If the use of GenAI is suspected where not permitted, you may be asked to meet with your instructor or TA to provide explanations to support the submitted material as being your original work. Through this process, if you have not sufficiently supported your work, academic misconduct allegations may be brought to the Associate Dean. In addition, you should be aware that the legal/copyright status of generative AI inputs and outputs is unclear. More information is available from the Copyright Advisory Committee: https://uwaterloo.ca/copyright-at-waterloo/teaching/generative-artificialintelligence [(11)]. Students are encouraged to reach out to campus supports if they need help with their coursework including:

- Student Success Office [(12)] for help with skills like notetaking and time management
- Writing and Communication Centre [(13)] for assignments with writing or presentations
- AccessAbility Services [(14)] for documented accommodations

Continuity Plans: In the event of a University-wide mandate to transition courses from in-person to online instruction, lectures and tutorials will be live-streamed virtually on Teams. All office hours will similarly become virtual (at their originally scheduled times), and assignments will carry on as planned. Should the tests and/or final exam be impacted, the following adjustments will be made. Online tests will replace any tests that are missed and will collectively account for the missed test's (tests') weight(s) in the final grade calculation. A take-home final exam worth 50% will replace the in-person final exam and will be due during the final exam period. In the event of a short-term absence of the instructor due to health or other emergencies, lectures may be delivered via Teams. You should already have access to the STAT 341 Teams channel. Any lectures held on Teams will be recorded and available on the same platform.

Mental Health Support: The Faculty of Math encourages students to seek out mental health support if needed.

On-campus Resources:

- Campus Wellness: https://uwaterloo.ca/campus-wellness/ [(15)]
- Counselling Services: counselling.services@uwaterloo.ca 519-888-4567 ext 32655

- MATES: one-to-one peer support program offered by Federation of Students (FEDS) and Counselling Services: mates@uwaterloo.ca
- Health Services: located across the creek from the Student Life Centre, 519-888- 4096.

Off-campus Resources:

- Good2Talk (24/7): Free confidential help line for post-secondary students. Phone: 1-866-925-5454
- Here 24/7: Mental Health and Crisis Service Team. Phone: 1-844-437-3247
- OK2BME: set of support services for lesbian, gay, bisexual, transgender or questioning teens in Waterloo. Phone: 519-884-0000 extension 213

University Policy

Academic integrity: In order to maintain a culture of academic integrity, members of the University of Waterloo community are expected to promote honesty, trust, fairness, respect and responsibility. [Check the Office of Academic Integrity [16] for more information.]

Grievance: A student who believes that a decision affecting some aspect of their university life has been unfair or unreasonable may have grounds for initiating a grievance. Read Policy 70, Student Petitions and Grievances, Section 4 [17]. When in doubt, please be certain to contact the department's administrative assistant who will provide further assistance.

Discipline: A student is expected to know what constitutes academic integrity to avoid committing an academic offence, and to take responsibility for their actions. [Check the Office of Academic Integrity [18] for more information.] A student who is unsure whether an action constitutes an offence, or who needs help in learning how to avoid offences (e.g., plagiarism, cheating) or about "rules" for group work/collaboration should seek guidance from the course instructor, academic advisor, or the undergraduate associate dean. For information on categories of offences and types of penalties, students should refer to Policy 71, Student Discipline [19]. For typical penalties, check Guidelines for the Assessment of Penalties [20].

Appeals: A decision made or penalty imposed under Policy 70, Student Petitions and Grievances [21] (other than a petition) or Policy 71, Student Discipline [22] may be appealed if there is a ground. A student who believes they have a ground for an appeal should refer to Policy 72, Student Appeals [23].

Note for students with disabilities: AccessAbility Services [24], located in Needles Hall, Room 1401, collaborates with all academic departments to arrange appropriate accommodations for students with disabilities without compromising the academic integrity of the curriculum. If you require academic accommodations to lessen the impact of your disability, please register with AccessAbility Services at the beginning of each academic term.

Turnitin.com: Text matching software (Turnitin®) may be used to screen assignments in this course. Turnitin® is used to verify that all materials and sources in assignments are documented. Students' submissions are stored on a U.S. server, therefore students must be given an alternative (e.g., scaffolded assignment or annotated bibliography), if they are concerned about their privacy and/or security. Students will be given due notice, in the first week of the term and/or at the time assignment details are provided, about arrangements and alternatives for the use of Turnitin in this course.

It is the responsibility of the student to notify the instructor if they, in the first week of term or at the time assignment details are provided, wish to submit alternate assignment.

Reference: Links from Document

1. https://learn.uwaterloo.ca/
2. https://piazza.com/uwaterloo.ca/winter2025/stat341
3. https://acal.fast.uwaterloo.ca/course/1251/STAT/341
4. https://www.math.uwaterloo.ca/~rwoldfor/courses/ComputationalStatDraft/Outline.html
5. http://www.neuroinformatics.ca/
6. https://uwaterloo.ca/math/vif-submission
7. https://uwaterloo.ca/math/accommodations
8. https://rstudio.com/products/rstudio/download/#download
9. https://uwaterloo.ca/secretariat/policies-procedures-guidelines/policy-71
10. https://academicintegrity.org/images/pdfs/20019_ICAI-Fundamental-Values_R12.pdf
11. https://uwaterloo.ca/copyright-at-waterloo/teaching/generative-artificial-intelligence
12. https://uwaterloo.ca/student-success/resources
13. https://uwaterloo.ca/writing-and-communication-centre/services-0/services-undergraduate-students
14. https://uwaterloo.ca/accessability-services/students
15. https://uwaterloo.ca/campus-wellness/
16. https://uwaterloo.ca/academic-integrity/
17. https://uwaterloo.ca/secretariat/policies-procedures-guidelines/policy-70
18. https://uwaterloo.ca/academic-integrity/
19. https://uwaterloo.ca/secretariat/policies-procedures-guidelines/policy-71
20. https://uwaterloo.ca/secretariat/guidelines/guidelines-assessment-penalties
21. https://uwaterloo.ca/secretariat/policies-procedures-guidelines/policy-70
22. https://uwaterloo.ca/secretariat/policies-procedures-guidelines/policy-71
23. https://uwaterloo.ca/secretariat/policies-procedures-guidelines/policy-72
24. https://uwaterloo.ca/accessability-services/