Introduction to Statistics and Data Science (STAT 202)

Fall 2024

Northwestern
Department of **Statistics and Data Science**

> Please note that the specifics of this course syllabus are subject to change in the case of unforeseen circumstances. Instructors will notify students of any changes as soon as possible. Students will be responsible for abiding by the changes.

Course Description

THIS COURSE INTRODUCES students to the discipline of statistics as a science of understanding and analyzing data. Students will learn the importance of data collection and sampling, methods to analyze data, and how to use data to make inferences and conclusions about real world phenomena. Students will be introduced to the free statistical programming software, Posit Cloud, to apply both descriptive and inferential statistics to real data sets.

When:
Mon, Wed, Fri, 9:00 - 9:50 am,

Location:
Lutkin Hall

Prerequisite: High School Algebra

Instructional Team

Arvind Krishna (Krish)
Instructor
Office Hours: After class or by appointment;
Location: In-class or zoom
✉ krish@northwestern.edu

Molly Chou (grader)
Graduate TA
Office Hours (for grading queries only): By appointment
✉ hway-minchou2026u.northwestern.edu

Zizheng Wang (grader)
Graduate TA
Office Hours (for grading queries only): By appointment
✉ zizhengwang2025u.northwestern.edu

Sarah Sajjad
UG TA
Office Hours: TBA
✉ sarahsajjad2025u.northwestern.edu

Krupal Patel
UG TA
Office Hours: TBA
✉ krupalpatel2025u.northwestern.edu

Maggie Johnson
UG TA
Office Hours: TBA
✉ maggiejohnson2025u.northwestern.edu

Jordan Popoola
UG TA
Office Hours: TBA
✉ sjordanpopoola2026u.northwestern.edu

UG TA Office Hours - held virtually: 📹 998 6392 3813

MAKE INDIVIDUAL APPOINTMENT/MEETING requests through Ed Discussion (discussed below) by selecting *Private (Visible to you and staff only).*

Course Goals

- USE STATISTICAL SOFTWARE to manage and process data.

- USE STATISTICAL SOFTWARE to perform exploratory data analyses. That is, explore data numerically and visually to gain understanding through data and generate hypotheses and inferences to later test.

- RECOGNIZE THE IMPORTANCE of data collection, identify limitations in data collection methods, and determine how they affect the scope of inference.

- BUILD A CONCEPTUAL UNDERSTANDING of the unified nature of statistical inference.

- APPLY ESTIMATION AND TESTING METHODS to analyze single variables or the relationship between two variables in order to understand natural phenomena and make data-based decisions.

- MODEL NUMERICAL RESPONSE VARIABLES using a single or multiple explanatory variables.

- INTERPRET RESULTS in context without relying on statistical jargon.

- CRITIQUE AND EVALUATE data-based claims and decisions.

"Numbers don't lie. That's where we come in."

Course Structure

A part of the class time will be dedicated to discussion/lecture, while another part will be dedicated to working on activities. Students will either work by themselves or in groups. Throughout the class we will discuss and review the work on the activities. In many cases we may come together to work on parts of an activity as a class.

Students are expected to prepare for class time by working through the indicated learning material prior to each class meeting.

Textbook (Free Online)

WE WILL BE USING Introduction to Statistics and Data Science which is a free online book that we have been developing for this course.

Software

WE WILL BE USING/INTRODUCING the free statistical software Posit Cloud.

Posit Cloud (formerly called RStudio Cloud), https://posit.cloud/

Hardware

STUDENTS WILL NEED a laptop or Chromebook to be able to follow lectures and to work with Posit Cloud to complete activities. If access to a laptop is an issue, then please contact the course instructor and we will work to find an accommodation. [1]

[1] This requirement will not prevent students from taking this course.

Asking Questions

We will use Ed Discussion as our preferred platform for questions about activities, reading checks or general course questions. The system is highly catered to getting you help quickly and efficiently from classmates and the instructional team. Rather than emailing questions to the instructional team, you should post your questions on Ed Discussion.

Questions concerning individual grades or appointments should be addressed through email.

The instructional team will check Ed Discussion periodically and answer questions[2], but we strongly encourage students to answer each other's questions.

[2] Please do not expect answers during weekends and evenings.

Tips for Success

- DEDICATE YOURSELF to being an active and engaged learner.

- PREPARE FOR CLASS by *reading and working* through code *before* class.[3]

  [3] Complete the reading checks!

- WORK IN GROUPS to learn and complete activities. [4].

  [4] Don't just copy, help each other.

- ASK QUESTIONS! Ask them during class, office hours, or on Ed. [5].

  [5] Be active on Ed!

- SEARCH THE INTERNET for help, especially for software issues.

- CONTRIBUTE TO a welcoming and inclusive learning environment.

- DON'T BE AFRAID to make mistakes, you learn from mistakes.

.
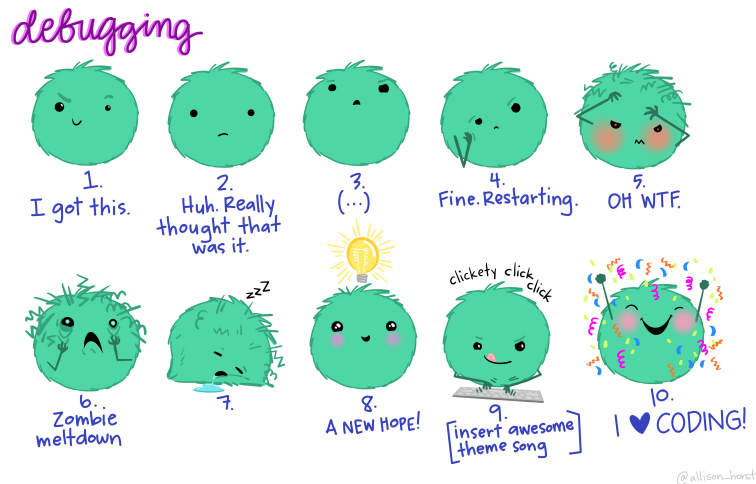
Debugging code is something we all struggle with and go through. It is a process and you will get better over time, but no matter your skill level there will always be struggles. Illistration by Allison Horst

Evaluation

- STUDENTS WILL BE EVALUATED through (1) reading checks; (2) activities/small assignments; (3) 3 exams; and a (4) final project.

- Reading checks will be completed using "Tutorials" on Posit Cloud and uploaded to the course Canvas page. All daily activities will be graded for completion. This doesn't mean that your solutions are correct, so make sure that you check them against the solutions.

Each reading check will be scaled to be worth 10 points.

- There will be 3 in-class exams; they will be structured very similarly to your reading checks. Students will be allowed one 8.5 x 11 inch cheat sheet (front & back) on each in-class exams. The exams are not cumulative.

- There will be 1 final group project. You will pick your own group of 3-6 people. More information will be provided later in the quarter.

Exam Improvement Policy

WE HAVE WORKED TO DEVELOP a policy geared towards a growth mindset. That is, we want a policy where students clearly demonstrate that they have used the exam as a diagnostic tool to learn from and improve their understanding of statistics. There is NO final cumulative exam, instead you may choose to retake 1 exam during the designated final exam time. This exam will replace your old score — only in cases where it is an improvement.

Missed Exam Policy

There are no make-up exams. If you miss an exam due to sleeping in, travel, etc., you will need to take the exam during the final exam period as your re-take exam.[6]

[6] Extenuating circumstances apply for proof of University excused absence.

Late Policy

ANY WORK SUBMITTED AFTER the due date will incur a 10% late penalty. Absolutely no assignments will be accepted more than three days after the due date without prior approval. To account for these strict due dates, we will drop your 3 lowest reading check scores **and** your 3 lowest activities.

Grading

GRADING SCALE	
93.0 - 100%	A
90.0 - 92.9%	A-
87.0 - 89.9%	B+
83.0 - 86.9%	B
80.0 - 82.9%	B-
77.0 - 79.9%	C+
73.0 - 76.9%	C
70.0 - 72.9%	C-
60.0 - 69.9%	D
Below 59.9%	F

CATEGORY	WEIGHT
Reading Checks	15%
Activities/Small Assignments	10%
Exam 1	20%
Exam 2	20%
Exam 3	20%
Final Project	15%

Final grades will be rounded to nearest tenth of a percent. There is NO curve to this course. Take advantage of the exam improvement policy and any extra credit opportunity.

Students opting to take this course Pass/No Pass must (1) take and pass the final exam with at least a D and (2) earn a course grade of at least a D.

Drop-In Support (No Appointment Needed)

Drop-In Peer Tutoring is set up such that students can drop in to study alone or with others and ask questions of a peer leader who has done well in the class. Tutoring is provided for many of the introductory courses in Biology, Chemistry, Economics, Engineering, Mathematics, Physics, and **Statistics**. Check their website for a complete list of supported courses.

Contact Valerie Wolf (valerie.wolf@northwestern.edu) with any questions.

Attendance Expectation/Policy

WHILE WE DO NOT collect formal attendance, implicit in the course design it is expected that you attend class to benefit from working with others — either by helping others or by helping others learn.

Syllabus statements, generative AI, and academic integrity

This course follows the Northwestern University Syllabus Standards. Students are responsible for familiarizing themselves with this information.

In addition to the aforementioned standards, **if you copy code, idea, or text from another team, or from some other resource without citing it, you will fail the course. Plagiarism will result in the F grade**.

Academic work and classes missed for medical reasons

In partnership with Student Health (Northwestern Medicine), the undergraduate schools have devised a system for students to request an excuse note. Any student who becomes ill must make use of Northwestern Medicine's process for missing academic work for medical reasons.

Tentative Course Schedule

Posted on the Course Website!

THERE IS ALMOST ALWAYS something due on Tuesday, Thursday, and Sunday.