Data science for studying language & the mind

Fall 2025

Syllabus

Course description: Data Sci for Lang & Mind is an entry-level course designed to teach basic principles of statistics and data science to students with little or no background in statistics or computer science. Students will learn to identify patterns in data using visualizations and descriptive statistics; make predictions from data using machine learning and optimization; and quantify the certainty of their predictions using statistical models. This course aims to help students build a foundation of critical thinking and computational skills that will allow them to work with data in all fields related to the study of the mind (e.g. linguistics, psychology, philosophy, cognitive science, neuroscience).

Prerequisites: There are *no prerequisites beyond high school algebra*. No prior programming or statistics experience is necessary, though you will still enjoy this course if you already have a little. Students who have taken several computer science or statistics classes should look for a more advanced course.

Professor: Dr. Katie Schuler (likes to be called Katie). I really enjoy getting to know students outside of class. Here are a few ways to connect with me:

- **Office Hours**: Day/Time TBD, 3401-C Walnut, 314C. Drop in for questions/ideas or just to chat.
- **Snack & Share**: Join me weekdays at 1:30pm in the Linguistics Library (15 min free snack break)

TAs: Brittany Zykoski and Mingyang Bian, both graduate students in the Department of Linguistics.

Lectures: Tuesdays and Thursdays from 12 - 1:29pm in COHN 402.

Labs: Hands-on practice and exam prep guided by TAs.

- 402: Fri at 3:30p in WILL 316 with Brittany
- 403: Thu at 1:45p in WILL 204 with Brittany
- 405: Fri at 12:00p in WILL 321 with Mingyang

- 406: Thu at 5:15p in BENN 322 with Mingyang

Office Hours: You are welcome to attend any office hours that fit your schedule. The linguistics department is located on the 3rd floor of 3401-C Walnut street, between Franklin's Table and Modern Eye.

- Katie Schuler: TBD in 314C
- Brittany Zykoski: TBD
- Mingyang Bian: TBD

Grading:

- 20% Homework (equally weighted, lowest dropped)
- 20% Lab Attendance/Participation (up to 2 absences excused without penalty)
- 60% Exams (equally weighted, final is optional to replace lowest exam)

Collaboration: Collaboration on problem sets and labs is highly encouraged! If you collaborate on psets, you need to write your own code/solutions, name your collaborators, and cite any outside sources you consulted (you don't need to cite the course material).

Accomodations: We will fully support any accommodations arranged through Disability Services via the Weingarten Center. If class conflicts with a religious holiday you observe, we are happy to make alternate arrangements. Please let us know early in either case so we can plan ahead.

Extra credit: There is no extra credit in the course. However, students can submit any missed problem set or exam by the end of the semester for half credit (50%) to boost their grade.

Regrade requests Regrade requests should be submitted through Gradescope within one week of receiving your graded assignment. Please explain why you believe there was a grading mistake, given the posted solutions and rubric

Assignments

Assignments for the Fall 2025 semester will be posted below. Last years materials are avilable here

Problem sets

There are 9 problem sets, due to Gradescope by noon on the following Mondays. You may request an extension of up to 3 days for any reason. After solutions are posted, late problem sets can still be submitted for half credit (50%). Your lowest score will be dropped.

- Pset 1 (Monday Sep 8)
- Pset 2 (Monday Sep 15)
- Pset 3 (Monday Sep 22)
- Pset 4 (Monday Sep 29)
- Pset 5 (Monday Oct 20)
- Pset 6 (Monday Oct 27)
- Pset 7 (Monday Nov 3)
- Pset 8 (Monday Nov 17)

Exams

There are 2 midterm exams, taken in class on the following dates. Exams cannot be rescheduled, except in cases of genuine conflict or emergency (documentation and a Course Action Notice are required). However, you can submit any missed exam by the end of the semester for half credit (50%). You may also replace your lowest midterm exam score with the optional final exam.

- Midterm 1 (Thursday, October 2)
- Midterm 2 (Thursday November 20)
- Optional Final Exam (TBD)

Lab exercises

Lab exercises are designed for practice and skill-building. They are not collected or graded. However, lab attendance is required and will count toward your course grade. You may miss up to 2 labs (for reasons such as illness or travel) without penalty.Notify your section TA if you must miss a lab.

- Lab 1 R Basics & Data Visualization
- Lab 2 Data wrangling
- Lab 3 Sampling
- Lab 4 Hypothesis testing
- Lab 5 Model specification
- Lab 6 Model fitting
- Lab 7 Model evaluation
- Lab 8 Feature engineering (topic may change)
- Lab 9 (optional)

Schedule

All coding demos can be found here.

Week	Begins	Topic	Materials	Lab (Thu/Fri)	Homework
1	Aug 25	R Basics	slides, demo	No lab!	
2	Sep 3	Data visualization		Lab 1	
3	Sep 8	Data wrangling		Lab 2	Pset 1
4	Sep 15	Sampling distribution		Lab 3	Pset 2
5	Sep 22	Hypothesis testing		Lab 4	Pset 3
6	Sep 29	Midterm 1 reviewMidterm 1			Pset 4
7	Oct 6	Make-up date Midterm 1**Fall break, no class/lab**			
8	Oct 13	Model specification		Lab 5	
9	Oct 20	Model fitting		Lab 6	Pset 5
10	Oct 27	Model evaluation		Lab 7	Pset 6
11	Nov 3	**Katie conference, no class/lab**			Pset 7
12	Nov 10	Feature engineering (??)		Lab 8	
13	Nov 17	Miterm 2 reviewMidterm 2			Pset 8
14	Nov 24	**Thanksgiving break, no class/lab**			
15	Dec 1	Careers in datasci lunch!		Lab 9	

Week	Begins	Topic	Materials	Lab (Thu/Fri)	Homework
16	Dec 8	**Last day of classes, no class/lab**			
17	TBD	Final exam (optional)			