Syllabus

Data Science for Studying Language and the Mind

Welcome to Data Science for Studying Language & the Mind! The Fall 2023 course information and syllabus are below. Course materials for previous semesters are archived here.

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).

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.

People

- **Instructor**: Dr. Katie Schuler
- **TAs**: June Choe, Avinash Goss, Ravi Arya

Lectures

Tuesdays and Thursdays at 10:15am in COHN 402.

Labs

Hands-on practice, quiz prep, and problem set work guided by TAs.

- 402: Thu at 1:45p in WILL 4 with TA
- 403: Thu at 3:30p in TOWN 305 with TA
- 404: Fri at 10:15a in WILL 24 with Ravi
- 405: Fri at 12:00p in TOWN 307 with Avinash

Office hours

The linguistics department is on the 3rd floor of the C-wing at 3401 Walnut street, between Franklin's Table and Modern Eye.

- Katie: Tue 12:30-1:30p in 314 ling dept
- June: TBD in ling dept
- Avinash: TBD
- Ravi: TBD

Quizzes

There are 4 quizzes, taken in class on Tuesdays. Missed quizzes cannot be made up except in cases of genuine conflict or emergency (documentation and a Course Action Notice are required). Instead, you will be invited to submit a missed quiz for half credit (50%).

Problem sets

There are 6 problem sets, due on Sundays to Gradescope by 11:59pm. Students may request an extension of up to 3 days. Extensions beyond this are not permitted, because delaying the release of solutions would negatively impact other students. After solutions are posted, late problem sets can be submitted for half credit (50%).

Grading

- 60% problem sets (equally weighted, lowest dropped)
- 40% quizzes (equally weighted)
- Letter grade minimums: 97% A+, 93% A, 90% A-, 87% B+, 84% B, 80% B-, 77% C+, 74% C, 70% C-, 67% D+, 64% D, 61% D-, else F
- All problems will be graded according to this rubric.

Collaborations

Collaboration on problem sets is highly encouraged! If you collaborate, 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 support any accommodations arranged through Disability Services via the Weingarten Center.

Extra credit

There is no extra credit in the course. However, students can submit any missed problem set or quiz by the end of the semester for half credit (50%). To ensure fair treatment across all students, all students will receive a 1% "bonus" to their final course grade: 92.54% will become 93.54%.

Regrade requests

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

Resources

In addition to the course website, we will use the following:

- google colab (r kernel) - for computing
- canvas - for posting grades
- gradescope - for submitting problem sets
- ed discussion - for announcements and questions

Please consider using these Penn resources this semester:

- Weingarten Center for academic support and tutoring.
- Wellness at Penn for health and wellbeing.