

STATS 216V
Introduction to Statistical Learning
Summer 2019 Syllabus

Instructor: Ismael Lemhadri

TAs: Mona Azadkia, Souvik Ray, Chenyang Zhong

Course website: <https://suclass.stanford.edu/courses/course-v1:HumanitiesSciences+STATS216V+Summer2019/course/>

Piazza: <https://piazza.com/stanford/summer2019/stats216v>

1. Course info

Overview of supervised learning, with a focus on regression and classification methods. Linear and polynomial regression, logistic regression and linear discriminant analysis; cross-validation and the bootstrap, model selection and regularization methods (ridge and lasso); nonlinear models, splines and generalized additive models; tree-based methods, random forests and boosting; support-vector machines. Some unsupervised learning: principal components and clustering (k-means and hierarchical).

Course objectives:

- to introduce some of the fundamental tools for building predictive models, including some "state-of-the-art" methods used by data scientists.
- to understand the role of model selection and assessment using cross-validation and randomization.
- to learn how to use the vast collection of tools in R to implement the methods learned.

The class is entirely online with no physical meetings.

2. Office Hours

Join Zoom at the URL <https://stanford.zoom.us/j/194614437> .

Office hours:

Mona Azadkia – Tuesday 8-9am PDT

Souvik Ray – Friday 3-4pm PDT

Chenyang Zhong – Thursday 6-7pm PDT

Zoom software instructions: <https://uit.stanford.edu/service/zoom>

3. Course requirements and grading

Quizzes (10%), homework (35%), midterm exam (15%), and final exam (40%). These weights are approximate; we reserve the right to change them later.

We'll use [Gradescope](#) for all homework and exams submissions. All students should be automatically enrolled by the end of the first week. You can also sign up manually using entry code MWJY4V.

- Quizzes

All quizzes have a submission deadline for Sunday, August 18 at 1pm (Pacific). You are allowed to wait until that day to answer all the quizzes, although that's not recommended. You are allowed only one submission attempt per question.

- Homework

There are 4 homework assignments over the quarter, spaced about bi-weekly. Homework will be graded on a scale of 0–100. **Late homework will not be accepted.** You will receive **two (2) bonus points** for submitting a typed written assignment (e.g. LaTeX, Microsoft Word). We will accept scanned handwritten assignments but they will not receive the bonus points. Once again, the homeworks are due at **5pm**.

- Midterm

The midterm will be posted under the Exams tab on the course website. You may work on the exam for *one hour*, so long as you submit it by the end of the 24-hour window. The midterm covers all material from Week 1 through Week 4. Practice problems will be posted.

- Final

The final will be posted under the Exams tab on the course website. You may work on the exam for *three hours*, so long as you submit it by the end of the 24-hour window. The final covers all the material in the course. Practice problems will be posted.

4. Textbook

An Introduction to Statistical Learning, by G. James, D. Witten, T. Hastie and R. Tibshirani. [Available online](http://www-bcf.usc.edu/~gareth/ISL/) at www-bcf.usc.edu/~gareth/ISL/, or in hardcopy at the Stanford Bookstore, Amazon, and Springer.

5. Collaboration Policy and Honor Code

Quizzes must be completed individually, but you are free to form study groups for homework. However, you must write up homework and code from scratch independently, and you must acknowledge in your submission all the students you discussed with.

The following are considered honor code violations:

- Looking at solutions from previous years' homework or exams - either official or written up by another student.
- Looking at the writeup or code of another student.
- Showing your writeup or code to another student.
- Uploading your writeup or code to a public repository so that it can be accessed by other students.
- Discussing homework problems in such detail that your solution (writeup or code) is almost identical to another student's answer.

Unless explicitly mentioned otherwise, we will assume that any submitted work is

- your own;
- created without assistance from anyone else (except possibly course staff); and
- created without consulting any resources other than the course materials.

If any work you submit in part or in whole does not adhere to these criteria, you are required to include a citation in your work explaining what additional assistance you received. If you feel like you made a mistake (which can happen, especially under time pressure!), please reach out to Ismael or to one of the TAs; the consequences will be much less severe than if we approach you.

6. Contact

Please use [Piazza](#) for all questions related to lectures, homework and exams.

Students may earn up to 3% extra credit by answering other students' questions in a substantial and helpful way.

Email is used for requests of a personal nature, such as about grades or Office of Accessible Education accommodations. Please use the email address stats216v-

2019-staff@lists.stanford.edu. In particular, please do not use the Instructor's or the TAs' direct email addresses for matters related to the course.

7. SCPD Students

SCPD student submit homework and exams on Gradescope just like non-SCPD students. In particular, there is no need to contact a proctor or the distribution team.

8. Dates

All times are Pacific.

Video release: videos are released by 10am on the release day.

- June 24: all of 1 and 2 released
- July 1: all of 3 until 4.5 released
- July 8: rest of 4, all of 5 released
- July 15: all of 6 released
- July 22: all of 7 released
- July 29: all of 8 released
- Aug 5: all of 9 and 10 released

Homework dates: homework is due by 5pm on the due day.

Homework 1: released June 24, due Friday July 5

Homework 2: released July 8, due Friday July 19

Homework 3: released July 22, due Friday August 2

Homework 4: released August 5, due Wednesday August 14

Practice exam dates: Practice midterm: released July 17. Practice final: released August 9.

Exam dates: exams are due by 10am on the due day.

Midterm: released July 24 10am, due July 25. Final: released August 16 10am, due August 17.

Other dates:

- Final study list deadline: July 8, 5pm

- Change of grading basis: August 2, 5pm