Please note that the following is subject to change. Changes will be announced via Canvas or via the course page in Ed.

Course Ed Page:
- [https://edstem.org/us/courses/22852/discussion/](https://edstem.org/us/courses/22852/discussion/)

**TA/Office Hours:**
- **Ying Jin**  
  Email: ying531@stanford.edu  
  OH: 5:00 pm - 7:00 pm, Zoom
- **Will Hartog**  
  Email: whartog@stanford.edu  
  OH: 9:45 am - 11:45 am, Zoom

**Course Info:** Overview of supervised learning, with a focus on regression and classification methods. Syllabus includes: 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). Some course objectives are outlined as follows:

- 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.

Computing is done in R, through tutorial sessions and homework assignments. **This math-light course is offered remotely only via video segments (MOOC style).**

**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 will use Gradescope for all homework and exams submissions. All students should be automatically enrolled by the end of the first week.

- **Quizzes:** Quizzes along with videos will be released on a week by week basis (tentative schedule provided below). The due dates will be spread throughout the week to keep your progress roughly in line with expectations, but you can feel free to complete all the quizzes at the beginning of each week. These quizzes are short 1-3 question quizzes related to the videos assigned for the week. You are only allowed 1 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.**

- **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.
Textbook: The primary text for this course is ISLR2: An Introduction to Statistical Learning, with applications in R (2d Edition) by Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani. We will use both the second edition, available as pdf online at: https://www.statlearning.com/.

Schedule/Important Dates: All times are Pacific.

Video release: Videos will be released at midnight the Monday beginning each week
- Week 1: 1 and 2 released
- Week 2: 3 and up to 4.5 released
- Week 3: Rest of 4 and 5 released
- Week 4: 6 released
- Week 5: 7 and 8
- Week 6: 9 and 12 released
- Week 7: 10 released

Homework: Homeworks will be due at 5pm on the due day
- HW1: Released June 20, due July 1
- HW2: Released July 5, due July 15
- HW3: Released July 18, due July 29
- HW4: Released August 1, due August 10

Exam dates: Exams will be released at 10am on release date and due at 10am on due date. Practice exams will be released approximately 1 week in advance.
- Midterm: Practice released July 13, Midterm released July 20, due July 21
- Final: Practice released August 4, Exam released August 11, due August 12
Course Policy:

- Please use Ed 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.

- I recommend personal emails to the instructor and TAs only for requests of a personal nature, such as questions about grades or OAE accommodations. Please do not use the Instructor’s or the TAs’ direct email addresses for matters related to the course.

Collaboration Policy/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 the instructor or to one of the TAs; the consequences will be much less severe than if we approach you.