

Policies

Course Description

Foundations of Data Science combines three perspectives: inferential thinking, computational thinking, and real-world relevance. Given data arising from some real-world phenomenon, how does one analyze that data so as to understand that phenomenon? The course teaches critical concepts and skills in computer programming and statistical inference, in conjunction with hands-on analysis of real-world datasets, including economic data, document collections, geographical data, and social networks. It delves into social issues surrounding data analysis such as privacy and design.

Prerequisites

This course does not have any prerequisites beyond high-school algebra. The curriculum and format is designed specifically for students who have not previously taken statistics or computer science courses. Students with some prior experience in either statistics or computing are welcome to enroll and will find much of interest due to the innovative nature of the course. Students who have taken several statistics or computer science courses should instead take a more advanced course.

Materials & Resources

Our primary text is an online book called Computational and Inferential Thinking: The Foundations of Data Science (<http://inferentialthinking.com/>). This text was written for the course by the course instructors.

The computing platform for the course is hosted at datahub.berkeley.edu (<http://datahub.berkeley.edu/>). Students find it convenient to use their own computer for the course. If you do not have adequate access to a personal computer, we have machines available for you; please contact your GSI.

Support

You are not alone in this course; the staff and instructors are here to support you as you learn the material. It's expected that some aspects of the course will take time to master, and the best way to master challenging material is to ask questions. For online questions, use Piazza (<http://piazza.com/>). We will also hold office hours (<http://data8.org/sp18/office-hours.html>) for in-person discussions. Small-group tutoring sessions will be available for students in need. In past semesters, students who attended have found these sessions to be a great use of their time. Details about sign-ups will be available later in the term.

Labs

Weekly labs are a required part of the course. After the first week, labs will be released on Sunday night. You can get credit for each lab in one of two ways described below:

- Attend your own assigned lab section, make progress substantial enough for your work to be checked off by course staff, and submit your lab (even if it is incomplete) by the end of the lab period.
- Complete the lab independently and submit the completed lab by the Wednesday of the same week at 9:59 a.m. "Completing a lab" means passing all tests included in the lab.

Homework and Projects

Weekly homework assignments are a required part of the course. Each student must submit each homework independently, but are allowed to discuss problems with other students and course staff. See the "Learning Cooperatively" section below.

If you submit a homework or project 24 hours before the deadline or earlier, you will receive 1 bonus point on that assignment.

Data science is about analyzing real-world data sets, and so a series of projects involving real data are a required part of the course. On each project, you may work with a single partner; your partner must be from your assigned lab section.

Exams

The midterm exam will be held during the class period on Friday, March 9th. Rooms will be announced closer to the date.

The final exam will be held from 3 p.m. to 6 p.m. on Tuesday, May 8th. Rooms will be announced closer to the date.

Unless you have accommodations as determined by the university and approved by the instructor, you must take the midterm and the final at the dates and times provided here. Please check your course schedule and make sure that you have no conflicts with these exams.

Participation

Participation points must be earned by attending lecture. As a class majority attended Data 8 lecture during the first 2 weeks of attendance collection, everyone will get full credit for the 1% participation portion of the grade.

Grades

Grades will be assigned using the following weighted components:

Activity	Grade
Participation	1%
Lab	11%
Homework	21%
Projects	26%

Midterm	10%
Final	31%

Overall, in past semesters of Data 8, more than 40% of the students have received grades in the A+/A/A- range, and more than 35% have received grades in the B+/B/B- range.

Late Submission

Late submissions of labs will not be accepted under any circumstances. The same goes for homework, unless you have relevant DSP accommodations that provide a two-day extension on homework assignments.

Your lowest two homework scores and your lowest lab score will be dropped in the calculation of your overall grade.

Projects will be accepted up to 2 days (48 hours) late; a project submitted less than 24 hours after the deadline will receive 2/3 credit, a project submitted between 24 and 48 hours after the deadline will receive 1/3 credit, and a project submitted 48 hours or more after the deadline will receive no credit.

Waitlisted Students

If you are on the waitlist, please do all coursework and attend labs and lectures in accordance to the deadlines.

If you enroll or join the waitlist after Week 1, you must contact your assigned Lab TA (see Staff Page) by email as soon as you add the course. The email must include your full name, Student ID number, date you added the course, and your @berkeley.edu email address. Once we receive your email, we will explain the prorating policy.

Learning Cooperatively

With the obvious exception of exams, we encourage you to discuss all of the course activities with your friends and classmates as you are working on them. You will definitely learn more in this class if you work with others than if you do not. Ask questions, answer questions, and share ideas liberally.

If some emergency takes you away from the course for an extended period, or if you decide to drop the course for any reason, please don't just disappear silently! You should inform your GSI and your project partner (if you have one), so that nobody is expecting you to do something you can't finish.

Academic Honesty

Collaboration has a limit, however. You should not share your code or answers directly with other students. Doing so doesn't help them; it just sets them up for trouble on exams.

Make a serious attempt at the assignment yourself, and then discuss your doubts with others. In this way you, and they, will get more out of the discussion.

Please write up your answers in your own words and don't share your completed work. The exception to this rule is that you can share everything related to a project with your project partner and turn in one project between you.

Penalties for cheating are severe - they range from a zero grade for the assignment or exam up to an F in the course, or even dismissal from the University.

Rather than copying someone else's work, ask for help. You are not alone in this course! The course staff is here to help you succeed. We expect that you will work with integrity and with respect for other members of the class, just as the course staff will work with integrity and with respect for you.

A Parting Thought

The main goal of Data 8 is that you should learn, and have a fantastic experience doing so. Please keep that goal in mind throughout the semester. Welcome to Data 8.