

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 also delves into social issues surrounding data analysis such as privacy and study 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 discussion 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 questions, use Piazza (<https://piazza.com/class/kjukuuch9zp6xi>). We will also hold virtual office hours (<office-hours.html>) for real-time discussions.

Your discussion TA will be your main point of contact for all course related questions/grade clarifications. The TAs are here to support you so please lean on your discussion TA if you need more support in the class or have any questions/concerns.

Small-group tutoring sessions will be available for students in need of additional support to develop confidence with core concepts. 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.

Disclaimer

The rest of this page details the policies that will be enforced in the Spring 2021 offering of this course. These policies are subject to change until the beginning of the semester and throughout the remainder of the course, at the judgement of the course staff.

All times listed below are in Pacific Standard Time (PST).

Waitlisted Students

If you are on the waitlist, you must still do all coursework and complete labs and homework by the deadlines. We will not be offering extensions if you are admitted into the course later. So it is your responsibility to stay up to date on the assignments.

Unfortunately, doing all the coursework is not a guarantee of enrollment. You will only be enrolled if there is space in lecture. Enrollment for lecture will proceed by CalCentral.

Live Lecture Sessions

Live lecture sessions will be held on Mondays, Wednesdays, and Fridays from 10am to 11am Pacific time over Zoom. These lectures will be used to highlight and review vital concepts of the course. Accompanying notebooks with examples will typically be provided to students. Recordings of these sessions will be provided, though students are highly encouraged to attend in real time.

Vitamins

To help you stay on top of lecture material, you will complete a short Gradescope "vitamin" - a mini-quiz - each week. The vitamin will be released on the website after lecture on Friday and due by 11:59pm PST the following Wednesday. Vitamins will cover the previous week's lectures and include one question that can only be answered if you attend discussion section.

Your lowest two vitamin scores will be dropped.

Labs

This course has a 1-hour weekly lab. You are required to attend lab in the first week of classes.

The weekly lab session consists of a lab assignment that develops skills with computational and inferential concepts. These lab assignments are a required part of the course and will be released on Monday nights.

Lab sessions are not webcast.

You can get credit for each lab assignment in one of two ways described below:

- Attend your own assigned lab section, make progress substantial enough for your work to be checked off, and submit your lab (even if it is incomplete) by the end of the lab period. Note that your submitted work need not be complete in order to receive full credit if you were checked off. However, your TA will only check you off if you have either finished the lab notebook or towards the end of the lab. You must both be checked off and submit your lab notebook to receive credit.
- Complete the lab on your own and submit the completed lab by Wednesday morning at 8:59 a.m PST. If you choose this route, you must finish the entire lab and pass all autograder tests (100% of tests passed) to receive credit.

Discussions

You will also attend a weekly discussion section. Discussions will be 1 hour between Monday and Wednesday and will give you a chance to practice with material covered in lecture and practice for exams. You will enroll in a discussion section, and you are only eligible to attend the discussion section you signed up for. **We will be recording attendance for discussions; attendance is mandatory.**

We will drop the two lowest discussion scores in the calculation of your final grade.

If you have questions related to discussion materials, post on the relevant Piazza thread or attend office hours.

Homework and Projects

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

Homeworks will be released on Friday after lecture and due the following Thursday night. If you submit a homework or project 24 hours before the deadline (Wednesday night) or earlier, you will receive 1 bonus point on that assignment.

Data science is about analyzing real-world data sets, and so you will also complete three projects involving real data. On each project, you may work with a single partner; your partner must be from the lab you enrolled in. Both partners will receive the same score.

Exams

The midterm exam will be held on **March 12th, 7-9pm PST**. Please note the date and time carefully.

The final exam is required for a passing grade, and will be held on Tuesday, May 11, 3-6pm. Please double check your course schedule to make sure that you have no conflicting finals.

There will be alternate exams for both the midterm and final. More information will be posted on Piazza closer to the exam dates.

Grades

Grades will be assigned using the following weighted components:

Activity	Grade
Vitamins	5%
Discussion Attendance	5%
Lab	5%
Homework	20%
Projects	20%
Midterm	15%
Final	30%

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

Instructors and TAs will not release grade bins during the semester, and they will be created after all grades come in by the instructors at the end of the semester. No staff members know this information, including the instructors, so please do not ask us.

Grades for Homeworks, Projects, and Labs will be posted about 1 week after the assignment's due date. There are 2 scores, OK and Gradescope, that add up together to make your total assignment score. Solutions to the assignment and common mistakes will also be posted on Piazza. It is up to you to check the solutions and request

a regrade request before the regrade deadline. Regrade requests for coding OK questions should be emailed to your discussion TA and regrade requests for written Gradescope questions can be made on Gradescope. **Any regrade request past the deadline will not be looked at;** this is to enforce the same deadline across all students, so please do not delay in reviewing your work.

For the midterm exam, there will be a regrade request submission window. Please review the solutions and common mistakes before submitting a regrade request. Requests where a rubric item was incorrectly selected or not selected will be reviewed, but any regrade requests that ask to change the rubric or for partial credit will be ignored.

Submitting Assignments

After running the submit cell for Homeworks, Projects, and Labs, please click on the OkPy link and make sure that your submission is not empty (i.e. it shows an .ipynb file that is your completed notebook). We cannot guarantee that your submission will be valid, so please double-check after you submit.

Late Submission

Late submissions of labs will not be accepted under any circumstances. The same goes for homeworks, unless you have relevant university accommodations filed with the DSP office.

Your two lowest homework scores, your two lowest lab scores, your two lowest vitamin scores, and your two lowest discussion attendance scores will be dropped in the calculation of your overall grade. There will be no alternate due dates for assignments missed due to illness, other commitments, and so on. The drops are intended to cover those situations.

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

Learning Cooperatively

We encourage you to discuss course content with your friends and classmates as you are working on your assignments. No matter your academic background, you will learn more if you work alongside others than if you work alone. 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 discussion TA and your project partner (if you have one) immediately, so that nobody is expecting you to do something you can't finish.

Academic Honesty

You must write your answers in your own words, and you must not share your completed work. The exception to this rule is that you can share everything related to a project with your project partner (if you have one) and turn in one project between the two of you, and if you are attending a lab session and have a lab partner you can share everything related to that lab with your lab partner.

Make a serious attempt at every assignment yourself. If you get stuck, read the textbook and go over the lectures and lab discussion. After that, go ahead and discuss any remaining doubts with others, especially the course staff. That way you will get the most out of the discussion.

It is important to keep in mind the limits to collaboration. As noted above, you and your friends are encouraged to discuss course content and approaches to problem solving. But you are not allowed to share your code or answers with other students. Doing so is academically dishonest, and it doesn't help them either. It sets them up for trouble on upcoming assignments and on the midterm exam.

You are also not permitted to turn in answers or code that you have obtained from others. Not only is such copying dishonest, it circumvents the pedagogical goals of an assignment. You must solve problems with the resources made available in the course. You should never look at or have in your possession solutions from another student or another semester.

Please read Berkeley's Code of Conduct (<https://sa.berkeley.edu/code-of-conduct>) carefully. Penalties for cheating in Data 8 are severe and include reporting to the Center for Student Conduct (<https://sa.berkeley.edu/conduct>). They might also include a F in the course or even dismissal from the university. It's just not worth it!

When you need help, reach out to the course staff using Piazza, in office hours, and/or during live labs and discussions. You are not alone in Data 8! Instructors and staff are 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.

Finally, know that it's normal to struggle. Berkeley has high standards, which is one of the reasons its degrees are valued. Everyone struggles even though many try not to show it. Even if you don't learn everything that's being covered, you'll be able to build on what you do learn, whereas if you cheat you'll have nothing to build on. You aren't expected to be perfect; it's ok not to get an A.

A Parting Thought

The main goal of the course 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!