STAT 133: Concepts in Computing with Data

Instructor: Deborah Nolan, 395 Evans
deborah_nolan@berkeley.edu
OH: Wed 8:30 - 10:30

GSI: TBA
OH: TBA

Lectures: Mon/Wed/Fri 11:10 am – noon; 10 EVANS
Lab: Thursday 9-11, 11-1, 1-3, 3-5, and 4-6; most in 342 Evans

Lab:
Attendance in lab is expected. Lab time is spent working on practice problems, which must be uploaded to bcourses by midnight Thursday. Some of the lab assignments and HWs require an account on the Statistical Computing Facility machines. Students can sign up for an account from a personal device in section beginning the second week of classes at https://scf.berkeley.edu/cgi-bin/class-account.cgi

Course Work:
Homework

Homework is posted on bcourses. All homework is due on bcourses. Typically, you are to upload an Rmd file. At times, you may be asked to upload additional files. There are 8 HW assignments, and the lowest score is dropped in calculating the total HW score. HW is due at the posted time. You may have a one-hour grace period for handing in each lab that does not count against the 24-hour late-bank (see below).

Lab

There are 9 lab assignments, and the lowest score is dropped in calculating the total Lab score. Lab is due at midnight of the day of section (Thursday). There is a one-hour grace period for handing in lab reports. This grace period does not count against the 24-hour late-bank.

Projects

There are 2 projects. Projects are to be uploaded to bcourses **AND** printed and turned in during lecture. The second project will most likely be done in teams. Each project has a 24-hour grace period for submission (which does not count toward the late bank).

Exams

There is one midterm, scheduled for 8-9pm, Wednesday, Oct 5 in 155 Dwinelle.
The Final Exam is scheduled for Monday, Dec 12 11:30-2:30.

Participation

We are using clickers in the classroom, and participation is based on being present for 75%

of the classes (beginning 9/7).

Overall score

- Lab Assignments: 12% (1.5% each lab)
- Participation: 6%
- Homework: 20% (3.33% each HW)
- Projects: 26% (12% for project 1 and 14% for project 2)
- Exams: The higher of 10% for Midterm + 26% for Final OR 36% for Final

Academic Honesty Policy

- Homework must be done independently. You may discuss the HW with other students, but you must independently write your code and solutions.
- Projects must be done independently. That is, groups must work independently, and within a group you may discuss and share code with students in your group.
- Exams are in-class written tests. You are provided with a set of notes that you can use during the exam. You may not bring or use any written or electronic materials into an exam other than those provided by the instructor.

Any evidence of cheating will result in a minimum penalty of a score of zero (0) on the assignment or examination. Depending on the severity of the infraction, cheating may result in an F for the course grade.

Disability

If you need accommodations for any learning disability, please speak to me after class or during office hours. Please make arrangements in a timely manner (through DSP) so that I can make the appropriate accommodations.

Late Policy

You will be given a late bank of 24 hours. There is no penalty for the first 24 hours of accumulated late times. After 24 hours in the "late bank" have been expended, then points will be lost as follows: 1 point for the first 8 hours or any fraction thereof, 1 point for the next 8 hours; and all 3 points for the HW after 16 hours. Note the grace periods for HWs, labs, and projects are not counted as part of the late bank.

Topics

Week	Topics	Lab	HW Due	Reading
Aug 22	Calling functions and working with vectors in R			Chap 1
Aug 29	Simple data structures and subsetting & EDA	Lab #1	HW #1	Chap 2.1-3, 2.5, 2.7
Sep 5	Graphics	Lab #2		Chap 3
Sep 12	Graphics & Data in nonrectangular formats	Lab #3	HW #2	4.1 & 4.2
Sep 19	Writing Functions	Lab #4	HW #3	Chap 5
Sep 26	Writing Functions	Lab #5		
Oct 3	Review + midterm	No lab	HW #4	Chap 7.1-7.6
Oct 10	Simulation & Representation of Information	Proj Help		
Oct 17	Regular Expressions	Proj Help		Chap 8
Oct 24	Text Mining	Lab #6	Project #1	
Oct 31	Web Scraping and XML	Lab #7	HW #5	Chap 11
Nov 7	Structured Query Language (MW only)	Lab #8	HW #6	Chap 9
Nov 14	Special Topic*	Lab #9	HW #7	
Nov 21	Special Topic* (Monday only)	No lab		
Nov 28	Special Topic*	Proj help		
Dec 5	RRR week	Review	Project #2	

This syllabus is only a guide, and there is likely to be slight variation as the semester progresses. It will be updated as we progress through the semester.

*The Special Topics have not yet been decided on. They may include JavaScript, interactive visualization, parallel computing, and statistical topics, such as the bootstrap and machine-learning topics, e.g., supervised and unsupervised learning methods.