STAT 133: Concepts in Computing with Data

Instructor: Deborah Nolan, 395 Evans,
deborah_nolan@berkeley.edu
OH: Wed 1:30-3:30

GSI: Bradly, Christine, Inna
OH: Mon 10-11, 5-6, Tue 2-4, Wed 4-5, and Thu 5-7

Lectures: Tue/Thu 12:30-2pm; 2050 VLSB
Lab: Monday 9-10, 10-11, 11-12, 12-1, and 1-2; 342 Evan

Lab:
Attendance in lab is required. Lab time will be spent working on practice problems, which must be uploaded to bcourses by midnight Monday. You ay have a one-hour grace period for handing in each lab that does not count against the 24-hour bank (see below).

Course Work:
Homework

Homework will be posted to bcourses. All homework is due on bcourses. Upload the Rmd file. There will be 8 HW assignments, and the lowest score will be dropped in calculating the total HW score. HW is due at the posted time.

Lab

There will be 10 lab assignments, and the two lowest scores will be dropped in calculating the total Lab score. Lab is due at the posted time. There is a one-hour grace period for handing in lab reports.

Projects

There will be 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 groups. Each project will have a 24-hour grace period for submission.

Exams

There will be one midterm, scheduled for Thursday, Mar 19. Review the clicker questions when studying for the midterm.

There will be a Final Exam, scheduled for Thursday, May 14.

Participation

We will use clickers in the classroom, and participation will be based on being present for75% of the classes.

Overall score

- Lab Assignments: 10%
- Participation: 5%
- Homework: 20%
- Projects: 25%
- Exams: The higher of 10% for Midterm + 30% for Final OR 40% 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.
- For exams, cheating includes, but is not limited to, bringing written or electronic materials into an exam or quiz, using written or electronic materials during an exam or quiz, copying off another person's exam or quiz, allowing someone to copy off of your exam or quiz, and having someone take an exam or quiz for you.

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. We will keep a running tally of the number of hours late your HW assignments and projects are. There will be 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 that there is no penalty for lateness on HW1 or Labs 1 and 2.)

Topics

Week	Topics	Lab	HW Due
Jan 19	Calling functions and working with vectors in R		
Jan 26	Data structures and subsetting & EDA	Lab #1	HW #1
Feb 2	Graphics	Lab #2	HW #2
Feb 9	Apply functions & data in nonrectangular formats	Lab #3	
Feb 16	Writing Functions		HW #3
Feb 23	Simulation and Probability Distributions	Lab #4	HW #4
Mar 2	Representation of Information & UNIX	Lab #5	
Mar 9	HTML & Regular Expressions	Lab #6	Project #1
Mar 16	Text Mining + Midterm		
Mar 23	SPRING BREAK		
Mar 30	Web Scraping and XML	Lab #7	HW #5
Apr 6	Structured Query Language	Lab #8	HW #6
Apr 13	Special Topic*	Lab #9	HW #7
Apr 20	Special Topic*	Lab #10	HW #8
Apr 27	Special Topic*		
May 4	RRR week		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.