# Course Information

**Lecture Schedule:** M-W 9-10AM, 502 Davis

**Lab Schedule:** W 1-3 (3-5, 5-7) PM, 345 Davis Hall

**Catalog Description**:
Application of the concepts and methods of probability theory and statistical inference to Civil & Environmental Engineering (CEE) problems and data; graphical data analysis and sampling; elements of set theory; elements of probability theory; random variables and expectation; simulation; statistical inference. Applications to a wide range of CEE problems involving real data will be developed, using both pre-existing and student-prepared Python codes.

**Prerequisites**: Data 8 (Python). No credit will be given after taking Stat28.

**Units**: 3

**Course Changes in 2019:**
While CE 93 has been taught for many years, beginning this semester there are some important changes. CE 93 is now part of a two-semester sequence intended for freshman, building on Data 8. Labs will use Python instead of Matlab. Unlike Data 8, CE 93 focusses on data from the engineering domain, introduces the concepts of probability and random variables, and uses these concepts to rigorously quantify the uncertainty associated with results obtained from data analysis.

**Course Objective:**
Introduce the student to the concepts and methods of probability theory and statistical inference by way of their application to CEE problems involving real-world data. Graphical and computational methods, using Python, will be emphasized. The course also serves to introduce the student to a variety of CEE problems and data through their statistical/probabilistic analysis.

**Required Textbook**:
William Navidi, ***Statistics for Engineers and Scientists***, Fourth Edition, McGraw Hill (the 3rd edition is cheaper and could also work, but you will need to correlate the reading and problem set assignments with the 4th edition).

**Course Websites:**
On bCourses (https://bcourses.berkeley.edu/). Contains assignments, labs, solution sets, lecture notes, supplementary readings.

On Piazza (https://piazza.com/berkeley/spring2019/ce93/home), hosts current discussions on material. This link will guide you how to enroll our CEE93 class: piazza.com/berkeley/spring2020/ce93. When posting questions on Piazza, please locate your questions under the correct folders. For example, all questions about HW#1, please post under the folder of hw1.

**Class Sessions:**
The classroom, Wheeler 212, is designed to be an Active Learning Classroom. In general student will sit in groups of four and, on many class days, part of the class time will be devoted to active learning exercises involving these groups. Other portions of the class will be for conventional lecturing, using the unique projection capabilities that enable all students to see what is being projected.

We encourage students to vary the seats they choose so that they can have the experience of working in different groups. Toward that end, individual students should not choose a seat in the same seat group in two consecutive classes.

**Homework:**
Assignments will be given weekly. See bCourses for the assignments and due dates. Assignments are due Monday at the beginning of class. 20% is subtracted from the grade of any assignment turned in late, up to the following Wednesday at the beginning of class. We will not accept assignments turned in after that time.

**Labs:**
Weekly two-hour sessions where students are trained on statistical and probabilistic manipulation of data using computer software (Python). Topics covered include histogram analysis, distribution fitting and plotting of all needed graphs. Lab assignments will be posted on bCourses.

- We recommend that bring your own laptop for the labs;
- All necessary files available at 12:00pm the day before the lab session;
- Lab assignments are given as .docx on bCourses, fill out and submit it in the format of PDF (filename format: Name _ Lab #. pdf) as you go;
- The assignments should be submitted electronically by the end of the lab section, but no later than 12 midnight of same day;
- Reader will grade for completion, but partial credits generously given for your efforts during the labs, so try your best.

**Exams:**
There will be two midterm exams and a final exam for this course. See the course schedule.

**Quizzes:**
Starting 1/30, there will be a 1/3 probability of a quiz at the beginning of each class, based on the throw on a dice throw. Students in the same group may collaborate on the quiz, but each student should hand in their own answer.

**Grading**:
30pts for final exam, 30 pts for midterms, 20 pts for HW assignments and lab reports, 20 pts for quizzes. From N quizzes, N-2 will be used for grading.

**Academic Integrity:**
Berkeley Campus Code of Student Conduct (http://sa.berkeley.edu/student-code-of-conduct):

"The Chancellor may impose discipline for the commission or attempted commission (including aiding or abetting in the commission or attempted commission) of the following types of violations by students, as well as such other violations as may be specified in campus regulations:

102.01 Academic Dishonesty: All forms of academic misconduct including but not limited to cheating, fabrication, plagiarism, or facilitating academic dishonesty."

For CE93, instances of academic dishonesty include, and are not limited to, the following:

- Homework and labs: You may discuss problems together, but all written work must be original and each student must do their own Python programming. Copying of solutions from any source IS NOT acceptable (and… not in your interest). Be very careful about plagiarism, particularly; all text must be in your own words and properly cited. See http://www.plagiarism.org for more information.
- Exams: Pop-quizzes, midterm and final exam are closed book other than the cheat sheet. No discussion, collaboration, or copying allowed.

**Instructors:**

| Name | Contact | Office | Office Hours |
|---|---|---|---|
| Prof. Mark Hansen, Instructor | mhansen@ce.berkeley.edu | 114 McLaughlin | Mon 2-3:30, Th 2-3:30 + Piazza |
| Ke Liu, GSI | liuke126@berkeley.edu When sending emails, please format as [CEE93: HW#] | 305 Davis | Fri 9-10 + Piazza |
| Meghna Thomas, GSI | meghnathomas@berkeley.edu When sending emails, please format as [CEE93: HW#] | 305 Davis | Tue 2-4 + Piazza |

## Course Schedule

| Week (Date) | Topic | Reading Assignment | Homework Due Date | Lab Schedule |
|---|---|---|---|---|
| 1 (1/20-1/24) | Introduction, course organization and objectives. Populations and samples. Types of data. Types of Experiments, Summary statistics: central tendency, dispersion, percentiles. | N* 1.1- 1.3 | | |
| 2 (1/27-1/31) | Graphical summaries: histograms, cumulative frequency diagrams, box plots, scatter plots, correlations Probability: Experiments, sample space, events, algebra of events. Axioms of probability. Combinatorics. | N 2.1-2.2 | | Lab 1 Graphical Data Analysis |
| 3 (2/3-2/7) | Conditional probability, total probability theorem, Bayes' formula. Independent events and the multiplication rule. | N 2.3 | Set 1 | Lab 2 Numerical Summaries of Data |
| 4 (2/10-2/14) | Random variables. Probability distributions for discrete and continuous RVs: PMF, PDF, CDF. Mean and variance of an RV. Linear functions of RVs. | N 2.4-2.5 | Set 2 | Lab 3 Elements of Probability Theory |
| 5 (2/17-2/21) | Jointly distributed RVs. Marginal and conditional distributions. Correlation, covariance, and independence**.** | | Set 3 | Lab 4 Random Variables |
| 6 (2/24-2/28) | **Midterm this week on Wednesday 2/26** | N 2.6 | Set 4 | |
| 7 (3/2-3/6) | Special random variables: Bernoulli, binomial, Poisson, hypergeometric. Uniform, normal, exponential, gamma, Central limit theorem. | N 4.1-4.8 | Set 5 | Lab 5 Seismic Hazard Analysis I |
| 8 (3/9-3/13) | Point estimation. Maximum likelihood. | N 4.9-4.12 | Set 6 | Lab 6 Distributions |
| 9 (3/16-3/20) | Confidence intervals for means and proportions. Large sample and small sample cases. Confidence Interval for population differences and paired data | N 5.1-5.7 | Set 7 | Lab 7 Seismic Hazard Analysis II |
| 10 (3/30-4/3) | Hypothesis testing for means and proportions. | N 6.1-6.4 | Set 8 | Lab 8 Parameter Estimation |
| 11 (4/6-4/10) | Hypothesis testing, simulations | N 6.5-6.11, 6.15 | Set 9 | Lab 9 Central Limit Theorem |
| 12 (4/13-4/17) | **Midterm this week on Wednesday 4/15** | | Set 10 | |
| 13 (4/20-4/24) | Regression. | N 7.1-7.4 | Set 11 | Lab 10 Hypothesis Testing |
| 14 (4/27-5/1) | Multiple Regression | N 8.1-8.3 | Set 12 | Lab 11 Regression |
| 15 (5/4-5/8) | Recitation Week | | | |
| 5/13 7-10pm | **Final Exam** | | | |

*N-William Navidi, Statistics for Engineers and Scientists, Fourth Edition