

## Soc 7 The Power of Numbers: Quantitative Data in the Social Sciences

Fall 2020 | UC Berkeley

**Instructor:** Professor Linus Huang

**Office hours:** Tuesdays on Zoom, 1:00-3:00 PM; sign up on bCal

**E-mail:** [lbhuang@berkeley.edu](mailto:lbhuang@berkeley.edu)

**Final Exam:** See below.

### What will this course be about?

Numbers are all around us in our everyday lives. Some numbers tell us how the economy is doing overall, who is likely to win an upcoming election, or how many people attended a large gathering like an inauguration ceremony. Other numbers tell us who qualifies for a low-interest loan to go to college or to buy a home. Numbers tell us whether it's likely to rain tomorrow, or how long we might expect to sit in traffic if we go out. Certain numbers tell college applicants which schools are better than others. Other numbers tell schools which applicants are better than others. And once a college-bound person enrolls in a school, yet other numbers will represent the quality of that student's academic performance.

Despite their presence everywhere, however, numbers are not always well understood. Some seem so transparent that we don't question them. Others are so hopelessly complex that we don't even try to understand them. Many of the numbers that are widely accepted as common knowledge are not even right. Yet, since numbers have the appearance of precision, they continue to influence the way we understand the world. This is part of their double-edged power. As citizens, professionals, social activists, and civic leaders, we need to develop the numerical literacy to recognize bad numbers and either demand, or produce ourselves, better numbers.

This is an introductory course in dealing with quantitative data—finding it; handling it; analyzing and interpreting it; and reporting it, including in visual form. The core of the course is learning introductory statistical methods, and the content here will not be unlike what you would encounter in introductory statistics courses elsewhere in the University. However, this course is geared toward social science applications and will comparatively speaking move at a more deliberate pace. There will be math! However, you do not necessarily need a strong mathematical, statistical, or computing background to succeed in this course. What you *do* need is curiosity about how numbers “work” (what they do and do not say, why some numbers are misleading, etc.) and a desire to try something new.

By the end of the semester, you will be able to understand, evaluate, use, and produce quantitative data about the social world by:

- critiquing and producing basic graphs
- manipulating and analyzing data in spreadsheets
- calculating and explaining basic statistical measures of central tendency, variation, and association
- applying and explaining basic concepts of sampling and selection
- thinking critically about reported statistics and quantitative social science more broadly

## The Coronavirus

Soc 7 will meet entirely online this Fall 2020 semester. It will, *technically*, be possible to take the course **asynchronously**. Lectures will be recorded and made available off-line; and all graded assignments either are designed to be completed on your time, or accommodations can be made where they are not. (See below.)

However, I intend this course to be given **synchronously**, as if we were meeting in-person. I will lecture live, over Zoom, on the regular MWF 10:00AM-11:00AM schedule, as if we were meeting in person.\* Furthermore, we will periodically use class meetings as “lab” sessions, where you will work on problems given during class interactively, with both myself and others in the course, employing Zoom’s breakout rooms.

\*I will continue to observe the “Berkeley time” convention even though we are online, so we will actually begin at 10:10 AM instead of 10:00 AM.

Even when we are not using class sessions for lab work, much of the course will have an interactive flavor, especially in the second half when we will be primarily focused on inferential statistics. Throughout the semester, I intend to take full advantage of Zoom’s interactive features, including both normal audio/visual interaction, as well as chat.

I will hold weekly office hours during the Fall 2020 term, on Tuesdays from 1:00-3:00 PM. These will be held over Zoom, at the same URL as for the lecture sessions. Students can sign up for 20-minute appointments over bCal.

## Readings and Resources

All course readings are available in PDF format on the bCourses site. There are no books or paper readers to purchase.

All readings should be completed *prior* to the first class we begin discussing them.

*Important:* You will need Microsoft Excel on your PC or Mac, if you do not already have it. Note that all UC Berkeley students are entitled to free use of Microsoft Office (which includes Excel) on their personal machines. If you do not already have Office/Excel, go to the following URL to get it: <https://software.berkeley.edu/productivity-software#Microsoft>

## Grading

Your course grade will be determined by the following components:

- **8 homework assignments** (5% each; 40% of the course grade total), designed to accompany individual topics in the course. The due dates for these homeworks have been targeted to balance giving you an incentive to keep up with the material as we consider it in class, but also allowing us at least one class session to talk about new material before the homework is due. As the scheduling works out at the moment, they are due weekly every Wednesday in the first half of the course (before the midterm exam), and every Monday in the second half of the course (following the midterm exam). There is no homework due the week of the take-home midterm, although we will otherwise proceed with new material in class.

- **6 “laboratory” sessions** (4% each; 24% of the course grade total), where we will work on problems related to current course topics interactively using Zoom’s breakout rooms feature. There will be deliverables necessary for each of these sessions, either individual or group submissions (this has yet to be precisely determined). Accuracy will not be unimportant in these submissions, but the emphasis is more on participation compared to the homeworks.

If circumstances prevent you from participating live in these lab sessions, contact me. As stated above, synchronous course participation is recommended, but accommodations can be made where necessary.

- **a take-home midterm exam** (18% of the course grade) at roughly the halfway point in the course. Under normal circumstances this would have been an in-class exam, but for logistical reasons it is being made a take-home to be done outside of class. More details on this will be given as the time approaches.
- **a take-home final exam** (18% of the course grade). This will be due on bCourses on Wednesday, December 16<sup>th</sup>, 2020, at 11:59 PM PST. Note that the official timeslot for an in-person exam would have been Monday, December 14<sup>th</sup>, from 8-11 AM. I will give basically the same exam I would have if we were holding the exam in-person, but I will release it on Monday, Dec. 14<sup>th</sup> at 8 AM, and it will be due 64 hours later.

The course grading scale is as follows:

|    |       |   |       |    |       |
|----|-------|---|-------|----|-------|
| A+ | 97+   | A | 93-96 | A- | 90-92 |
| B+ | 87-89 | B | 83-86 | B- | 80-82 |
| C+ | 77-79 | C | 73-76 | C- | 70-72 |
| D+ | 67-69 | D | 63-66 | D- | 60-62 |
|    |       | F | 0-59  |    |       |

When it comes time to compute overall course grades, I will round to the nearest whole number using standard rounding conventions. It doesn’t really matter what the letter grade on the individual assignments are.

There are no surprises in how I calculate course grades. The **GRADES** section on bCourses will incorporate the weightings above and will accurately keep you apprised of your course progress.

## Recommendation Letters

Writing recommendation letters is an aspect of an instructor’s job which I embrace readily. However, it is difficult for me, and unhelpful to all parties involved, to write letters for students I do not know very well. I do have, and have exercised, the discretion to decline to write letters when I don’t know students beyond what grades they’ve earned in the course. If you anticipate asking me to write a letter, it is in your interest to participate in class discussion, visit me during my office hours, etc., and let me know what your plans and interests are above and beyond course material.

## Academic Honesty

The UC Berkeley Honor Code states that “As a member of the UC Berkeley community, I act with honesty, integrity, and respect for others” (<https://teaching.berkeley.edu/berkeley-honor-code>). I expect you will follow these principles. You may not copy specific text or ideas from others, whether from fellow students, from authors of our readings or other material you find, without specific attribution. To do otherwise is to plagiarize. You may not cheat on any of the homework assignments or exams by bringing in illicit outside material, copying from fellow students, or engaging in other dishonest practices. Violation of these rules will result in an immediate **-0-** on the *entire* assignment in question, plus a report to the Office of Academic Affairs at my discretion.

## Reading, Assignment, and Exam Schedule

All readings for this course are available on bCourses in the READINGS folder of the FILES section. Readings associated with a date are to be completed prior to the class meeting on that day.

Due dates for graded assignments are **highlighted** below.

| Introduction                 |        |   |                     |
|------------------------------|--------|---|---------------------|
| Introduction                 | Aug 26 | W | <i>No readings.</i> |
|                              | Aug 28 | F |                     |
|                              | Aug 31 | M |                     |
| Basic spreadsheet operations | Sep 2  | M | <i>No readings.</i> |
|                              | Sep 4  | F |                     |
|                              | Sep 7  | M | Labor Day: no class |
|                              | Sep 9  | W | <i>No readings.</i> |

| Descriptive Statistics |        |    |   |
|------------------------|--------|----|---|
| Types of data          | Sep 11 | F  | <i>Read:</i> Agresti §2.1   |
| Measures of centrality | Sep 14 | M  | <i>Read:</i><br>• Wheelan ch. 2 pp. 15-23<br>• Agresti §3.2                       |
|                        | Sep 16 | W  | HW #1 due on bCourses<br>Lab session #1   |
| Measures of dispersion | Sep 18 | F  | <i>Read:</i><br>• Wheelan ch. 2 pp. 23-35<br>• Agresti §3.3-§3.4                  |
|                        | Sep 21 | M  |   |
|                        | Sep 23 | W  | HW #2 due on bCourses<br>Lab session #2   |
| Association            | Sep 25 | F  | <i>Read:</i><br>• Wheelan ch. 4<br>• Agresti §3.5                                 |
|                        | Sep 28 | M  |   |
|                        | Sep 30 | W  | HW #3 due on bCourses   |
|                        | Oct 2  | F  | Lab session #3  |
| “Calling bullshit”     | Oct 5  | M  | Midterm exam distributed on bCourses  |
|                        | Oct 7  | W  |   |
|                        | Oct 8  | Th | Midterm exam due on bCourses,<br>Thursday, October 8 <sup>th</sup> , 11:59 PM PST |

| <b>Introduction to Probability</b>                           |        |   |  |
|--|--------|---|--|
| Introduction to probability, and why we should care about it | Oct 9  | F | <i>Read:</i> Wheelan ch. 5   |
|  | Oct 12 | M |  |
|  | Oct 14 | W |  |
| Samples and populations                                      | Oct 16 | F | <i>Read:</i> Wheelan ch. 6   |
|  | Oct 19 | M | HW #4 due on bCourses  |
| Distributions and the Central Limit Theorem                  | Oct 21 | W | <i>Read:</i>   |
|  | Oct 23 | F | <ul style="list-style-type: none"> <li>• Wheelan ch. 8</li> <li>• Agresti §4.3-§4.5</li> </ul> |
|  | Oct 26 | M | HW #5 due on bCourses  |

| <b>Inferential Statistics</b>             |        |   |   |
|---|--------|---|---|
| Hypothesis testing                        | Oct 28 | W | <i>Read:</i>  |
|   | Oct 30 | F | <ul style="list-style-type: none"> <li>• Wheelan ch. 9</li> <li>• Agresti §6.1-§6.6</li> </ul>  |
|   | Nov 2  | M | HW #6 due on bCourses<br>Lab session #4   |
| Confidence intervals                      | Nov 4  | W | <i>Read:</i>  |
|   | Nov 6  | F | <ul style="list-style-type: none"> <li>• Wheelan ch. 10</li> <li>• Agresti §5.1-§5.4</li> </ul> |
|   | Nov 9  | M | HW #7 due on bCourses<br>Lab session #5   |
|   | Nov 11 | W | Veteran's Day: no class   |
| Comparing two groups                      | Nov 13 | F | <i>Read:</i> Agresti §7.1-§7.3  |
|   | Nov 15 | M |   |
| Association between categorical variables | Nov 18 | W | <i>Read:</i> Agresti §8.1-§8.4  |
|   | Nov 20 | F |   |
|   | Nov 23 | M | HW #8 due on bCourses<br>Lab session #6   |
|   | Nov 25 | W | Turkey Day: no class  |
|   | Nov 27 | F |   |
| Regression                                | Nov 30 | M | <i>Read:</i>  |
|   | Dec 2  | W | <ul style="list-style-type: none"> <li>• Wheelan ch. 11</li> </ul>                              |
|   | Dec 4  | F | <ul style="list-style-type: none"> <li>• Agresti §9.1-§9.5</li> </ul>                           |

|                                 |        |   |   |
|---------------------------------|--------|---|---|
| Reading, Recitation, and Review | Dec 7  | M |   |
|                                 | Dec 9  | W |   |
|                                 | Dec 11 | F |   |
|                                 | Dec 14 | M | Final exam distributed on bCourses, Monday, Dec 14, 8:00 AM PST |
|                                 | Dec 16 | W | Final exam due on bCourses, Wednesday, Dec 16, 11:59 PM PST     |