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

Fall 2019 | UC Berkeley

Instructor: Professor Linus Huang
Office Hours (drop-in): Wednesdays, 2:15-4:00 PM, 487 Barrows
E-mail: lbhuang@berkeley.edu
Final Exam: Monday, December 16th, 8:00-11:00 AM

Course Overview[1]

Numbers abound at all levels of 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 ubiquity, however, numbers are not always 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 course will introduce you to the basic concepts central to quantitative sociology and equip you to become more savvy consumers of numbers in society today. It seeks to give you an intuitive overview of quantitative tools employed by social scientists and hands-on opportunities to use these tools to examine the world. It is intended specifically for social science majors, and focuses on social science questions. 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 curiousity about how numbers "work" (what they do and do not say, why some numbers are misleading, etc.) and a desire to try something new. My aim is to show you that quantitative science is useful, and is something that you can do.

To put our lessons into practice, we will use data from real-world datasets including Gapminder, a country-level dataset that sheds light on global inequalities; CalEnviroScreen, a dataset maintained by the State of California that gives detailed information on environmental health hazards at the level of census tracts (roughly speaking, the neighborhood level); and the General Social Survey (GSS), probably the premier dataset for assessing Americans' attitudes on social issues.

[1] This course overview, and much of the general plan of the course overall, draws heavily from previous iterations of the course offered by Sara Lopus, Michael Schultz, and Mao-Mei Liu.

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

Required Readings and Resources

There is one required text for this course, Charles Wheelan's *Naked Statistics: Stripping the Dread from the Data* (2013). It is available at the ASUC Store, or wherever savvy consumers today can go to buy books. I will also place the book on 1-day reserve at Moffitt Library.

All other readings for this course are in PDF format online in the `FILES` section of bCourses.

Important: **it is highly advisable for you to have a laptop with internet access and Microsoft Excel or equivalent to take this course**, as you will need it to participate in class.

Note that all UC Berkeley students are entitled to free use of Microsoft Office (of which Excel is one component of the "productivity suite") on their personal machines. If you do not already have Office/Excel, go to https://software.berkeley.edu/productivity-software#Microsoft.

Assignments / Grading

Your course grade will consist of four different components: 1) individual homework assignments; 2) in-class small group exercises; 3) an in-class midterm exam; and 4) an in-class final exam.

1) Individual homework

There will be eight individual homework assignments designed to accompany current course topics. They will be due roughly (but not exactly) weekly. Check on bCourses for the exact due dates for each of the assignments.

Each of the eight homework assignments will be worth 4% of the course grade, or 32% for all eight of them.

2) In-class small group exercises

Throughout the semester, we will form small groups of 3 to 4 people to collaborate on an in-class exercise related to the topic. On days where we do this, we will form the groups at the beginning

of class, and the collaboration will last just that meeting. It won't be necessary to maintain the same groups from exercise to exercise.

It is during these group exercises where it will be particularly important to have a laptop at your disposal. The plan is to have eight of them. Each will be worth 4% of the course grade, or 32% for all eight of them.

3) In-class midterm exam

We will have an in-class midterm exam scheduled for Wednesday, October 9th. It will be worth 18% of your course grade.

4) In-class final exam

The final exam will be held during the University's official timeslot for our course, which is on Monday, December 16th, from 8-11 AM. It will be cumulative.

The final exam will be worth 18% of your course grade.

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.

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

Readings for this course will be either from the required Wheelan text, or in PDF format in the **Readings** folder of the **Files** section on bCourses.

Introduction			
Introduction	Aug 28	W	
	Aug 30	F	*Read:* Wheelan ch. 1
Labor Day: no class	Sep 2	M	Labor Day: no class
Basic spreadsheet operations	Sep 4	W	*No readings.*
	Sep 6	F	
Our datasets: Gapminder, CalEnviroScreen 3.0, and the General Social Survey	Sep 9	M	*No readings.*

Descriptive Statistics			
Types of data	Sep 11	W	*Read:* Agresti 2.1 (PDF)
Measures of centrality	Sep 13	F	*Read:* • Wheelan ch. 2, pp. 15-23 • Agresti 3.2 (PDF)
	Sep 16	M	Homework #1 due on bCourses
	Sep 18	W	In-class small group exercise #1
Measures of dispersion	Sep 20	F	*Read:* • Wheelan ch. 2, pp. 23-35 • Agresti 3.3-3.4 (PDF)
	Sep 23	M	Homework #2 due on bCourses
	Sep 25	W	In-class small group exercise #2
Association	Sep 27	F	*Read:* • Wheelan ch. 4 • Agresti 3.5
	Sep 30	M	Homework #3 due on bCourses
	Oct 2	W	In-class small group exercise #3
Some do's and don't's of data visualization	Oct 4	F	*No readings.*
	Oct 7	M	
Midterm Exam, in-class	Oct 9	W	

Introduction to probability			
Introduction to probability, and why we should care about it	Oct 11	F	*Read:* • Wheelan ch. 5, 5 ½ • Silver, *The Signal and the Noise* ch. 2 (PDF)
	Oct 14	M	Homework #4 due on bCourses

	Oct 16	W	In-class small group exercise #4
Distributions and the Central Limit Theorem	Oct 18	F	*Read:* • Wheelan ch. 8 • Agresti 4.3-4.5 (PDF)
	Oct 21	M	Homework #5 due on bCourses
	Oct 23	W	In-class small group exercise #5

Inference			
Samples and populations	Oct 25	F	*Read:* Wheelan ch. 6, 7
	Oct 28	M	
Confidence intervals	Oct 30	W	*Read:* • Wheelan ch. 10 • Agresti 5.1-5.4 (PDF)
	Nov 1	F	Homework #6 due on bCourses
	Nov 4	M	In-class small group exercise #6
Hypothesis testing	Nov 6	W	*Read:* • Wheelan ch. 9 • Agresti 6.1-6.6 (PDF)
	Nov 8	F	Homework #7 due on bCourses
	Nov 11	M	Veteran's Day: no class
	Nov 13	W	In-class small group exercise #7
Comparing two groups	Nov 15	F	*Read:* Agresti 7.1-7.3 (PDF)
	Nov 18	M	
Association between categorical variables	Nov 20	W	*Read:* Agresti 8.1-8.4 (PDF)
	Nov 22	F	Homework #8 due on bCourses
	Nov 25	M	In-class small group exercise #8
Turkey Day: no class	Nov 27	W	Turkey Day: no class
	Nov 29	F	
Regression	Dec 2	M	*Read:* • Wheelan ch. 11 • Agresti 9.1-9.5 (PDF)
	Dec 4	W	
	Dec 6	F	

Reading, Recitation, and Review	Dec 9	M	
	Dec 11	W	
	Dec 13	F	
Final Exam, in-class	Dec 16	M	We're in the 8 AM-11 AM timeslot.