The Power of Numbers:
Introduction to Quantitative Social Science

Instructor: Byron Villacis
Days and Time: T-TH 8-10am
Room: 402 Barrows
Office Hours: T 10-12 (sign-up), Barrows (TBD)

Overview[1]:

Social statistics have become more widespread and more visible in recent years. From Nate Silver's quantitatively-driven prediction of the 2008 and 2012 elections, to the book *Moneyball*, to the sometimes euphoric claims made for "big data," numbers are now everywhere in the public sphere. But despite their ubiquity, these numbers are not always well understood. Some statistics seem so transparent that we do not think much about them. Others seem so opaque that we give up. Many of the numbers that circulate as common knowledge are not even right. Yet, they have the appearance of precision, and a certain social power, and so they stay on. Wrong numbers can have important social consequences. As citizens, professionals, social scientists and activists, and civic leaders, we need to recognize bad statistics and produce better ones.

There are many facets of society that can be effectively understood quantitatively. There are even most facets that can *only* be understood quantitatively. ***This course will provide students with a set of skills to understand, evaluate, use, and produce quantitative data about the social world.*** It is intended specifically for social science majors and focuses on social science questions. ***You do NOT need a strong mathematical, statistical, or computing background to succeed in this course***. What you do need is a basic curiosity about how society is organized and a willingness to try something new. Our aim is to show you that quantitative social science is useful, can be fun, and is something that you can do.

By the end of this course, students will be able to:

1. Understand, through practical examples, basic concepts behind statistical measures in social sciences, with emphasis on: central tendency, variations, correlation and basic causality.
2. Understand, calculate, and explain basic statistical measures of central tendency, variation, and correlation.
3. Understand, evaluate, and produce basic graphs.
4. Find and manipulate relevant high-quality data.
5. Understand and apply basic concepts of sampling.

[1] I am grateful to Mao-Mei Liu, Sara Lopus and Daniel Schneider for sharing their Soc 5 and Soc 7 syllabus, original sources for this version. It borrows from them many of the course policies, general guidelines, and general formatting. I am also grateful to Linus Huang for his suggestions and comments.

6. Thinking critically about reported statistics and quantitative social science more broadly.
7. Develop basic skills to evaluate quantitative arguments regularly present in social sciences.

Books and resources:
You must have a laptop with Internet access to take this course. There is no textbook. All readings will be posted on bCourses. We will focus on working through real applications and will rely heavily on web-based resources and tools.

Expectations and evaluation:
Classroom time will be a combination of lecture, discussion, and practical work. You will be required to read in advance of class. Most of pieces will be empirical examples; we will use them to understand step-by-step how measures are conceptualized, created, and applied. Then we will discuss particular statistical concepts, their process of construction and their scope of analysis. Finally, we will apply examples with real data.

Individual homework: There will be 7 individual homework assignments, due every Tuesday by 8:00am on bCourses. Your best 6 will be counted. Late homework will not be accepted. Homework assignments are designed to introduce new topics or reinforce material covered in recent lectures. Grading of these assignments will focus more upon effort than correct answers, since the purpose is typically to prepare you for activities performed in the next class.

Group work: Over the course of the semester, you will work in groups of four or five on a comprehensive, collaborative project using a real dataset. As you learn new topics in class, you will use your dataset to apply the course material. Your group project will build as the sum of in-class activities and some home assignment. At the end of the semester, each group will do a class presentation in which you identify some interesting relationships you have discovered, presented in graphical forms. This project will be an opportunity to familiarize yourself with data, move it around, and work with it in useful ways.

Exams: There will be two exams: a mid-term and a final exam in the normally assigned slot during the final week of class. These exams will be much more conceptual than computational and will focus on your understanding of the core concepts of the course.

Grading (100 points possible):

Assignment	%	Due Date
Homework (6 of 7)	20	Tuesdays 8AM, (submit online)
Midterm	20	July 12, 8AM
Final	20	August 9, 8AM
Project	20	Final week of class
Participation	10	Semester-basis
Presentation	10	Final week of class

Course Policies

Readings and Lecture

Students are responsible for materials covered in class and in readings. I will post lecture slides on the bCourses site. The slides are aids for learning and are not a complete account of class activity. Please do all of the readings in a timely fashion. Each day usually has one or two readings associated with it. You will get the most out of lecture if you do these readings before the class for which they are assigned.

Accommodation

I will provide accommodation to any student who provides me with a written letter from a DSP Specialist. If you require accommodation, the first step is to have DSP send me an official written accommodation letter. Once I receive this letter and if I have any questions, I will contact you by email. Please arrange for me to receive the letter as early in the semester as possible. I will also provide accommodation for observation of religious rituals. University policy is that such requests should be made by the second week of the semester. Please submit them by email.

Late Work

There are several written assignments for this course. The precise due dates and where the assignments should be handed in are noted above. Assignments turned in late will be penalized one letter grade for every day late (e.g. one day late makes a B a C). If an emergency arises, email me about it at least 24 hours before the deadline. Exams will be given on the posted dates. If you know now that you will have a conflict with the scheduled exam times, speak with me as soon as possible so that we can work out an accommodation. If you have a true unforeseen emergency that prevents you from attending the exam, contact me as soon as possible to discuss an accommodation.

Grading Policy

If you wish to contest a grade, please outline in writing (1) what assignment you are contesting, (2) the grade you received on the assignment, and (3) the reason(s) why you believe the grade you received is unfair. I will consider your appeal and may decide to re-grade your assignment. Please note: a re-grade likely involves closer scrutiny of the work and so may result in an increase or a decrease in your grade. Whatever the outcome, the score from the re-grade will be final. The grade appeals process should be initiated within seven days of receiving the grade in question.

Email
Email should not be used for substantive questions about the reading or course materials. Please ask such questions in class or sign-up for office hours. I will endeavor to respond to other email messages within 24 hours.

Office Hours
I highly encourage you to sign-up for office hours. You may also just stop-by my office during office hours and if I don't have another student scheduled, I will be happy to talk with you. Office hours are a good time to introduce yourself, talk about ideas you are thinking about, and discuss problems you may have with class. I may decide to hold obligatory office hours for groups.

Group Presentations
Using PowerPoint or a short movie, present a 5-8 minute interesting, informative and effective audio- visual overview of your group project (Introduce data and variables; Descriptive Statistics; Hypotheses & Inference). Include a reflection of the problems and difficulties encountered and how the group solved them.

Honor Code:
The student community at UC Berkeley has adopted the following Honor Code:
"As a member of the UC Berkeley community, I act with honesty, integrity, and respect for others." We hope and expect that you will adhere to this code. This is Berkeley. We are too good to cheat.

There will be a good bit of collaborative work in this course. While this kind of collaboration is an important tool in learning and realistic preparation for the contemporary world of work, it also can make it hard to know what acceptable collaboration is and what is taking unfair advantage of others. Throughout the course, we will indicate whether a given assignment is to be completed alone or in cooperation with others. You will avow on each assignment that you complied with those instructions. If at any point you have any questions about how the honor code applies, or how best to fulfill your obligations as a member of the UC Berkeley community, please feel free to ask the instructor. Almost certainly, someone else has the same question.

#	W	Date		Main activity / Topic	Main source	Examples/Additional
1	1	T	19-Jun	Overview of the course, introduction to basic concepts.	Gapminder	Intro to graphs
2		W	20-Jun	Basic Concepts: arithmetic operations, objects and vectors.	Selection from Cook 2013 (reinforce "reading graphs" with info graphics)	R basic tutor
3		Th	21-Jun	Units of analysis, scales and basic operationalizations.	Eurostat, Gapminder	Unit of Analysis in Poverty Measurement
4	2	T	26-Jun	Basic graphs I and types of data.	https://www.stat.berkeley.edu/~stark/SticiGui/Text/histograms.htm#variables	Basic Graphs R Gallery
5		W	27-Jun	Central tendency measures.	http://www.stat.berkeley.edu/~stark/SticiGui/Text/location.htm#location	Central Tendency measures in the Australian Census
6		Th	28-Jun	Basic descriptive statistics I.	Seen Theory: visual introduction to probability and statistics	Quick Basic descriptive Child care early education example
7	3	T	3-Jul	Range, variance and standard deviation.	http://www.stat.berkeley.edu/~stark/SticiGui/Text/location.htm#spread	Measuring wage dispersion: pay ranges reflect industry traits
8		W	4-Jul	*Holiday*		
9		T	5-Jul	Basic descriptive statistics II.	http://www.stat.berkeley.edu/%7Estark/SticiGui/Text/scatterplots.htm	Unemployment Rate Hits 3.9%, a Rare Low
10	4	T	10-Jul	Basic Graphs II and common mistakes.	Huff "How to lie with statistics" Wheelan "Deceptive Description: "He's got a great personality!" and other true but grossly misleading statements"	How to Lie and Cheat with Statistics How to lie with statistics
11		W	11-Jul	Finding and evaluating data I and Midterm Review.	Tufte "The Visual Display of Quantitative Information"	Data Visualization in Sociology
12		Th	12-Jul	Basic manipulation of data and **Midterm.**		
13	5	T	17-Jul	Finding, evaluating and manipulating data II.	TBD	
14		W	18-Jul	Sampling I and basic probability.	http://www.stat.berkeley.edu/~stark/SticiGui/Text/sampling.htm#simple	Sampling Techniques in Social Research
15		Th	19-Jul	Sampling II and basic probability.	http://www.stat.berkeley.edu/~stark/SticiGui/Text/probabilityPhilosophy.htm#theories_of_probability	Current Population Survey (CPS)
16	6	T	24-Jul	Survey structures and basic design.	Silver Chapter 5 ("Desperately seeking signal")	General Social Survey (GSS) Overview of the General Social Survey
17		W	25-Jul	Data analysis plan: a design of analysis step by step.	https://www.stat.berkeley.edu/~stark/SticiGui/Text/samplin	Data Analysis in Excel

					g.htm#survey_bias	Example of statistical data analysis using R
18		Th	26-Jul	Normal distribution and p-value.	http://www.stat.berkeley.edu/~stark/SticiGui/Text/clt.htm#normal_curve	Normal distribution calculator
19	7	T	31-Jul	Hypothesis tests and types of error.	http://www.stat.berkeley.edu/~stark/SticiGui/Text/testing.htm#hypothesis_testing	Hypothesis testing, type I and type II errors Sociological Images
20		W	1-Aug	Correlation vs. Causality: basics.	Wheelan Chapter 4 "Correlation: How does Netflix know what movies I like?"	Intro Spurious correlations Tasty correlation
21		Th	2-Aug	Correlation: estimations and analysis.	Seen Theory: visual introduction to probability and statistics	Correlation in Excel. Correlation in R
22	8	T	7-Aug	Linear regression: basics.	https://www.stat.berkeley.edu/~stark/SticiGui/Text/regression.htm	Linear regression is not dead, and please don't call it OLS How to Read a (Quantitative) Journal Article
23		W	8-Aug	Group Presentations (50%), review of the course		
24		Th	9-Aug	Final exam		