# SOC 7

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

## Overview

Social statistics have become more widespread and visible in recent years. From Nate Silver's 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 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 opportunities for students to build skills to understand, evaluate, use, and produce quantitative data about the social world.*** It is intended 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***. Our aim is to learn how quantitative social science can be useful, fun, and accessible.

## Materials

You must have access to a spreadsheet (software or cloud-based) and the internet to take this course. Access to a laptop will also aid participation.

There are two textbooks for this course:

- **Silver, Nate. (2012) The Signal and the Noise: Why so many predictions fail—but some don't. Penguin Books.**

- **Wheelan, Charles. (2013) Naked Statistics: Stripping the dread from the data. Norton.**

## Goals

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

- Understand, evaluate, and produce basic graphs

- Manipulate data in a spreadsheet, including using pivot tables

- Understand and calculate basic statistical measures of central tendency, variation, and correlation

- Understand and apply basic concepts of sampling and selection

- Begin thinking quantitatively about social science questions

## Expectations and Evaluation

The instructor evaluates each student based on their progress on each of the following areas and milestones:

1. *Homework (25% of final grade):* There will be 5 homework assignments, each worth 5% of your grade. Sometimes you will be required to read, watch a video lecture, or try out a web-based application in advance of class.

2. *Participation (10% of final grade):* Classroom time will typically be a combination of lecture, discussion, and practical work. A lack of preparation for class will be reflected in your participation grade.

3. *Data Dives (20% of final grade)*: Data Dives provide the opportunity to "crunch" numbers using the skills you are developing in class to interpret data. You are expected to produce work that is uniquely yours although you may review course concepts with peers and have them review your assignment prior to submission.

4. *Team Talk (15% of final grade)*: You will work in teams on a comprehensive, collaborative project using a dataset to give you experience manipulating and interpreting data. Your project will culminate with an in-class group presentation in which you identify what you have discovered, presented in graphical form.

5. *Exams (30% of final grade)*: There will be two exams: a mid-term and a final exam. These exams will be much more conceptual than computational, and they will focus on your understanding of the core concepts of the course.

## Milestones

**DATA DIVE: Sept 30, 2016**
Paper due by 9AM on Friday 9/30/16

**MIDTERM: Oct 12, 2016**
Midterm Examination scheduled for Wednesday 10/12/16 8:10-9AM

**DATA DIVE: Nov 11, 2016**
Paper due by 9AM on Friday 11/18/16

**TEAM TALK: Nov 28, 2016**
Group presentations during week of November 28- December 2
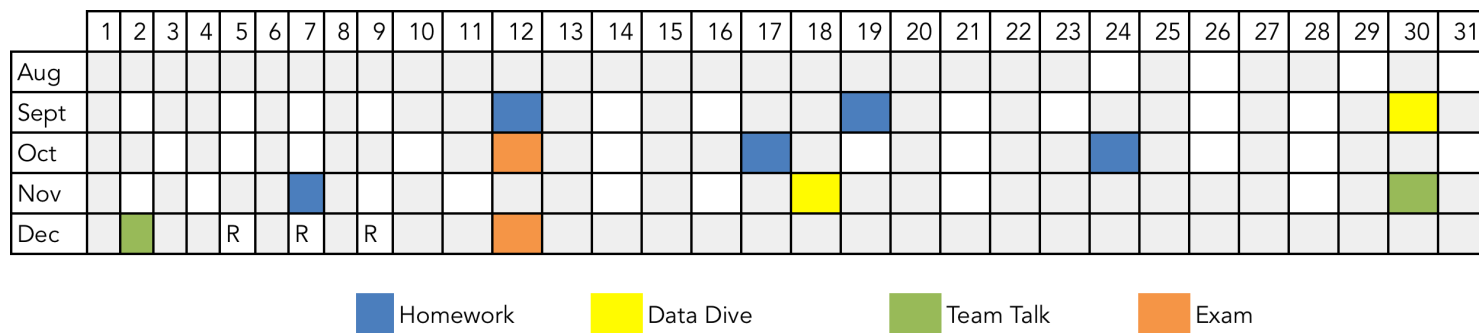
**FINAL: Dec 12, 2016**
Final Examination scheduled for Monday 12/12/16 7-10PM

## 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.

There will be collaborative work in this course. While collaboration is an important part of learning and preparation for the world of work, it can make it hard to know what is acceptable. Throughout the course, the instructor will indicate whether a given assignment is to be completed alone or in cooperation with others. You, the student, will avow on each assignment that you complied with those instructions. If at any point you have questions about how the honor code applies, please ask the instructor.

# SCHEDULE OVERVIEW

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 |
|------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Aug  |   |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| Sept |   |   |   |   |   |   |   |   |   |    |    | ■Homework |    |    |    |    |    |    | ■Homework |    |    |    |    |    |    |    |    |    |    | ■Data Dive |    |
| Oct  |   |   |   |   |   |   |   |   |   |    |    | ■Exam |    |    |    |    | ■Homework |    |    |    |    |    |    | ■Homework |    |    |    |    |    |    |    |
| Nov  |   |   |   |   |   |   | ■Homework |   |   |    |    |    |    |    |    |    |    | ■Data Dive |    |    |    |    |    |    |    |    |    |    |    | ■Team Talk |    |
| Dec  |   | ■Team Talk |   |   | R |   | R |   | R |    |    | ■Exam |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |

Legend: ■ Homework   ■ Data Dive   ■ Team Talk   ■ Exam

| Week of | Topic | Course Material for Week |
|---------|-------|--------------------------|
| 8/22/2016 | Introduction | Syllabus |
| 8/29/2016 | The Basics of Spreadsheets, Data Types, & Units of Analysis | • Best Chapter 1, section "The public as innumerate audience" <br> • Wheelan Chapter 1 (What's the Point?) <br> • Stark Chapter 3 (Variables; Exercise 3-1) |
| 9/5/2016 | Summarizing Data: Standard Measures of Centrality & Dispersion | • Wheelan Chapter 2 - "Descriptive Statistics" <br> • Stark Chapter 4 (Measures of Location; Videos of Exercises; Exercises 4-1 through 4- 5) (online) <br> • Stark Chapter 4 (Spread or Variability; The Range, IQR, and SD; Videos of Exercises; Exercises 4-6 to 4-8) (online) |
| 9/12/2016 | Association; Histograms & Probability | • Wheelan Chapter 4: "Correlation" <br> • Stark Chapter 5 (Multivariate Data, Scatterplots, Exercises 5-1 to 5-4) <br> • Wheelan Chapter 5: Basic Probability <br> • Wheelan Chapter 5 ½: The Monty Hall Problem <br> • Stark Chapter 13 (Theories of Probability, Random Events, Equally Likely Outcomes, Frequency Theory, Exercises 13-1 to 13-5) |

| | | |
|---|---|---|
| 9/19/2016 | Distribution & Central Limit Theorem; Intro to Sampling | • Stark Chapter 23 (The Normal Curve)<br>• Stark Chapter 24 (Sampling, Parameters, Exercises 24-1 and 24-2)<br>• Wheelan Chapter 8 ("central limit theorem")<br>• Wheelan Chapter 6: Problems with Probability<br>• Radio Lab- "A very lucky wind" from Stochasticity (online) |
| 9/26/2016 | Sampling, Counterfactuals, & Inference | • Stark Chapter 24 (Simple Random Samples, Systematic Random Samples, Exercise 2 4-4)<br><br>• Wheelan Chapter 7 "The Importance of Data: 'Garbage in, garbage out'" pp110-113<br><br>• Wheelan Chapter 9 ("Inference")<br>The Data skeptic podcast #4 [p-values] (online) |
| 10/3/2016 | Review; Midterm; | |
| 10/10/2016 | Hypothesis Testing: Continuous & Categorical | • Silver Chapter 8 ("Less and less and less wrong")<br>• The Data Skeptic Podcast: #2 [Type I/Type II errors] (start with minute 2:30) (online)<br>• The Data Skeptic Podcast: #24 [The T-Test]<br><br>• Stark Chapter 27 (Hypothesis Testing: Does Chance Explain the Results?, Examples of Hypothesis Testing Problems, Significance Level and Power, Test Statistics and P Values, Exercises 27-1 to 27-3) (online)<br><br>• Stark Chapter 31 (Chi-square Statistic, Chi-square test Exercise 31-2 and 31-4)<br>• The Data Skeptic Podcast # 40 |
| 10/17/2016 | Regression & Data Sources | • Wheelan Chapter 11 ("Regression Analysis")<br>• Wheelan Chapter 12 ("regression mistakes")<br>• Huff Chapter 10 ("Talk Back to a Statistic")<br>• Silver Chapter 12 ("Healthy Skepticism")<br>• The Data Skeptic Podcast: #36 [Data Provenance] (online) |

| | | |
|---|---|---|
| 10/24/2016 | Good Graphs, Credible Sources, & Pivot Tables | • Wheelan Chapter 10 "Polling: How we know that 64 percent of Americans support the death penalty (with a sampling error +/- 3 percent)" pp180-183<br>• Best Chapter 1, "statistics as social products"<br>• Steele Chapter 2 ("Identity and Performance")<br>• Excel Pivot Table Tutorial |
| 10/31/2016 | Selection Biases & Sampling; Big Data/ Data Science | • Berk " introduction to sample selection bias"<br>• Stark Chapter 24 (Bias in Surveys, Exercise 24-3)<br>• Wheelan Chapter 10 "Polling: How we know that 64 percent of Americans support he death penalty (with a sampling error +/- 3 percent)" pp169-180<br>• The Data Skeptic Podcast: #21 [Selection Bias] |
| 11/7/2016 | Lying with graphs & Visualization | • Huff "How to lie with statistics" pp. 60-73<br>• Wheelan Chapter 3: "Deceptive Description"<br>• Tufte "The Visual Display of Information" |
| 11/14/2016 | Big Data / Data Science | • Lewis, *Moneyball*, Chapter 4<br>• Freedman "Statistical Models and Shoe Leather"<br>• Silver Chapter 5 ("Desperately seeking signal")<br>• Wheelan Chapter 7 "The Importance of Data: 'Garbage in, garbage out'" pp113-126 |
| 11/21/2016 | Mostly Thanksgiving Break; Review and Team Preparation | |
| 11/28/2016 | Presentations & Review | |
| 12/5/2016 | RRR Week | |
| 12/12/2016 | Exam | |