

Sociology 7 (Summer 2017)
The Power of Numbers: Introduction to Quantitative Social Science

Instructor: René Patnode

Email: rpatnode@berkeley.edu

Class: TWTh 8:00-10:00 am in Barrows 56

Office Hours: TWTh 10:00-10:30 am in Barrows 477

Course Description:

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 correct. 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 some 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 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 can be useful, fun, and accessible.

Course Goals:

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

- Think critically about reported statistics
- Manipulate data in a spreadsheet
- Understand and calculate basic statistical measures of central tendency, dispersion, and correlation
- Understand and apply basic concepts of sampling and selection
- Test hypotheses using basic statistical methods
- Find good-quality and relevant data on the internet
- Understand, evaluate, and produce basic graphs
- Begin thinking quantitatively about social science questions

Required Texts and Materials:

Silver, Nate. 2012. *The Signal and the Noise: Why So Many Predictions Fail—But Some Don't*. Penguin.

Wheelan, Charles. 2013. *Naked Statistics: Stripping the Dread from the Data*. Norton.

Other assigned readings and materials will be made available on bCourses. Readings should be completed by the corresponding date on the course schedule.

Students must have access to spreadsheet software. (Cloud-based is acceptable.) Bringing a laptop or tablet to class is highly recommended in order to facilitate class participation.

Course Requirements and Grading:

Final grades will be determined according to the following scheme with a fixed curve added, if necessary, to bring the mean score into the B grade-range.

- 30% Exams. (Midterm and final; each will be 15%.) Given in-class and comprised of short and multiple-choice questions. Exams will be based on material covered in class. Questions will focus more on concepts than computation although simple arithmetic might be required. The final exam might use basic terminology learned early in the course but will otherwise not be cumulative. No makeup exams will be provided unless there are extreme extenuating circumstances; arrangements should be made in advance when possible.
- 25% Homework. (Five; each will be 5%.) Each assignment will require answering a few questions pertaining to course material covered during the preceding week. Questions will be similar to those appearing on the exams. Assignments must be submitted to bCourses on time unless there are extenuating circumstances and approval is granted.
- 20% Written group assignments. (Two; each will be 10%.) Students will form groups and select a topic that they want to explore based on an appropriate data set. Groups will then answer questions in which they apply course concepts to their topic and data. Assignments must be submitted to bCourses on time unless there are extenuating circumstances and approval is granted.
- 15% Group presentation. Groups will present on their topic and what they have discovered in their data during the final week of the course.
- 10% In-Class Participation. (Evaluated twice for 5%.) Students are expected to participate fully in class. Asking questions, answering instructor questions, engaging in class activities, and attending office hours count towards the grade.

Attendance:

Attendance is not required although participation is graded. Failure to attend could cause you to miss material that is important for the exams and assignments as well as miss out on participation credit. If you come to class, please come on time.

Technology Usage:

As described above, use of laptops and other electronic devices is encouraged when it pertains to class activities. Other uses should be avoided.

Academic Integrity:

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." You should 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, must comply with those instructions. If at any point you have questions about how the honor code applies, please ask the instructor.

SCHEDULE (subject to change):

	Tuesday	Wednesday	Thursday
<i>Week 1:</i>	6/20: Course Introduction	6/21: Spreadsheet Basics; Units of Analysis <ul style="list-style-type: none"> • Wheelan, Ch. 1 • Best, Ch. 1 ("The Public as Innumerate Audience") 	6/22: Data Types <ul style="list-style-type: none"> • Stark, Ch. 3 ("Variables")
<i>Week 2:</i>	6/27: Central Tendency <ul style="list-style-type: none"> • Wheelan, Ch. 2, pp. 15-23 • Stark, Ch. 4 ("Measures of Location") <p>Homework one due</p>	6/28: Dispersion <ul style="list-style-type: none"> • Wheelan, Ch. 2, pp. 23-35 • Stark, Ch. 4 ("Spread or Variability"; "The Range, IQR, and SD") 	6/29: Association <ul style="list-style-type: none"> • Wheelan, Ch. 4 • Stark, Ch. 5 ("Multivariate Data"; "Scatterplots")
<i>Week 3:</i>	7/4: (no class)	7/5: Probability <ul style="list-style-type: none"> • Wheelan, Ch. 5 • Wheelan, Ch. 5 ½ • Silver, Ch. 4, pp. 134-138 • Stark, Ch. 13 ("Random Events"; "Equally Likely Outcomes"; "Frequency Theory") <p>Homework two due</p>	7/6: Distributions and Sampling <ul style="list-style-type: none"> • Wheelan, Ch. 8 • Stark, Ch. 23 ("The Normal Curve") • Stark, Ch. 24 ("Sampling"; "Parameters and Statistics"; "Why Sample?")
<i>Week 4:</i>	7/11: Sampling Methods <ul style="list-style-type: none"> • Wheelan, Ch. 6 • Wheelan, Ch. 7, pp. 110-113 • Stark, Ch. 24 ("Sampling Designs"; "Ways to Draw Samples") • RadioLab, "A Very Lucky Wind" <p>Homework three due</p>	7/12: Review	7/13: Midterm Exam

<i>Week 5:</i>	<p>7/18: Hypothesis Testing</p> <ul style="list-style-type: none"> • Silver, Ch. 8 • Data Skeptic Podcast, “Type I/Type II Errors” (start at 2:30) • Stark, Ch. 27 (“Hypothesis Testing: Does Chance Explain the Results?”; “Examples of Hypothesis Testing Problems”; “Significance Level and Power”) <p>Group assignment one due</p>	<p>7/19: Hypothesis Testing: Continuous</p> <ul style="list-style-type: none"> • Wheelan, Ch. 9 • Data Skeptic Podcast, “The <i>t</i>-test” • Data Skeptic Podcast, “<i>p</i>-values” • Stark, Ch. 27 (“Test Statistics and <i>p</i>-values”) 	<p>7/20: Hypothesis Testing: Categorical</p> <ul style="list-style-type: none"> • Stark, Ch. 31 (“The Chi-Square Statistic”; “The Chi-Square Test for Goodness of Fit”) • Data Skeptic Podcast, “The Chi-Squared Test”
<i>Week 6:</i>	<p>7/25: Regression</p> <ul style="list-style-type: none"> • Wheelan, Ch. 11-12 <p>Homework four due</p>	<p>7/26: Data Sources</p> <ul style="list-style-type: none"> • Huff, Ch. 10 • Silver, Ch. 12 • Best, Ch. 1 (“Thinking about Statistics as Social Products”) • Data Skeptic Podcast, “Data Provenance” 	<p>7/27: Data Issues</p> <ul style="list-style-type: none"> • Wheelan, Ch. 10 • Berk • Stark, Ch. 24 (“Bias in Surveys”) • Data Skeptic Podcast, “Selection Bias”)
<i>Week 7:</i>	<p>8/1: Visualization</p> <ul style="list-style-type: none"> • Wheelan, Ch. 3 • Huff, Ch. 5-6 • Tufte <p>Group assignment two due</p>	<p>8/2: Pivot Tables</p> <ul style="list-style-type: none"> • Excel Pivot Table Tutorial • Steele 	<p>8/3: Big Data</p> <ul style="list-style-type: none"> • Lewis, Ch. 4 • Freedman • Silver, Ch. 5 • Wheelan, Ch. 7, pp. 113-126 <p>Homework five due</p>
<i>Week 8:</i>	8/8: Group Presentations	8/9: Review	8/10: Final Exam