Sociology 106:
Quantitative Sociological Methods
Spring 2021, UC Berkeley

Instructor: Robert Pickett
rpickett@berkeley.edu
www.robertempickett.com

Online/Asynchronous

Course Website: https://bcourses.berkeley.edu/

Please note*: This syllabus is subject to change. Please access bCourses for the most up-to-date version. If anything is unclear, please send me an email.*

Course Description[1]

Sociology 106 is an intermediate undergraduate sociology research methods course. It will emphasize the motivation, computation, and interpretation of statistical tests for one or two continuous or categorical variables. The course will also introduce students to the R statistical programming language for data management and analysis. Sociology 106 is most appropriate for social science undergraduates who have some familiarity with sociological research methods and wish to learn how to carry out a quantitative research project.

Learning Objectives

By the end of the semester, you should be able to:

1) understand the basic logic of statistical inference,

2) identify the appropriate statistical test given a specific type of data,

3) visualize data and produce descriptive statistics and simple statistical tests using R, and

4) interpret the results of statistical tests and discuss their relevance in the context of a particular research question.

Prerequisites

Previous training in statistics in neither required nor expected. Successful completion of Sociology 5 is a requirement for this course, but other courses that introduce social science

[1] I am grateful to Jacob Habinek, Felipe Dias, Mao-Mei Lieu and Joe LaBriola for their work teaching this class previously, and whose materials I have adapted. I am also thankful to David Harding and Trond Peterson for their work designing the Sociology 271B class.

research design and methods may suffice. If you have not taken Sociology 5, contact the instructor to obtain permission to enroll.

Course Materials

There are no required readings for this course, though there are several good (and freely available online) textbooks that you might find useful.

For general statistics:

- David Lane's *Online Statistics*.

For statistics and R:

- Danielle Navarro's *Learning Statistics with R: A Tutorial for Psychology Students and other Beginners* and
- Beckman, Guerrier, Lee, Molinari, Orso and Rudnytskyi's *An Introduction to Statistical Programming Methods with R*. (more advanced coverage of programming in R)

Because this is an online class, you are required to have a computer with internet access to take this course. A major component of this course is learning the statistical programming language R and getting practice using R to analyze data, so a reasonably powerful computer is preferable. In the first week, we will walk through downloading R onto your computer (http://cran.rstudio.com), as well as RStudio, a free program that makes working in R much, much, easier (https://www.rstudio.com/products/rstudio/download/). Downloading and installing R and RStudio on Windows and Mac computers is simple. Installing it on a Chromebook may be more difficult; please let me know immediately if you are using a Chromebook for this course, and I will try to set you up with someone who can help you with the installation.

Evaluation

Weekly Exercises (10%)

> Each week there will be a selection of exercises to make sure you're following along with the material. These will be low stakes opportunities to check your understanding and will be graded for completion rather than for correctness (you will have as many chances to get the right answer as you'd like).

Weekly Assignments (30%)

> Weekly assignments will consist of a short problem set designed to test your comprehension of the material, and/or your ability to put the material into practice using R to analyze data of your choice. *It is highly recommended that you spend some time during the first week of the semester to choose a single dataset that interests you that you will use throughout the semester for your assignments and possibly for your final paper (see below for where to find data).*

I encourage you to work on these assignments with other students in the class. If you do so, you must report who you worked on the assignment with and still must write up your own work in your own words and submit your own R code and output. All students need practice writing up their analysis of data. Copying is plagiarism and will be treated as such (see UC Berkeley's Academic Code of Conduct).

Research Paper (40%)

Each student will develop and present a research question of their choice, address it by using the descriptive and inferential techniques presented in the course to analyze data in R, and write a paper summarizing your findings. You may use your weekly data analysis assignments to work on your research question. If you do not have an idea about a possible research question to pursue, please sign up for my office hours in the first few weeks and I can try to help you find a question you are interested in.

There will be several milestones throughout the semester to help you smoothly progress towards a final paper. I will give you feedback at each step of the way. Your overall grade on the research paper will be the sum of your grades on the following assignments:

- Paper proposal (5%): A (no more than) two-page single spaced proposal for the final paper will be due on bCourses by **Friday, February 26 at 11:59 PM**. The proposal should succinctly state the research question, why we might care about the answer to this question, hypotheses about what you think the answer is to the research question (and why), the data source you will use to answer this question, and the key independent and dependent variables in the data. At this point, you do not need to discuss what statistical techniques you will use to answer your research question.
- Annotated bibliography (5%): Next, you will identify ten scholarly sources related to your research question. Articles in academic journals, books, and book chapters from edited volumes are all considered to be scholarly sources. Blogs, internet and newspaper articles, Wikipedia pages, etc., are not considered to be scholarly sources. After identifying these ten scholarly sources, for each source, you will briefly (in one to two paragraphs) describe how this source relates to your research question. This includes, but is not limited to, the overall argument in the article or book, the types of evidence the author(s) use to support their article, and any possible weaknesses or strengths in the paper. This annotated bibliography will be due on bCourses by **Friday, March 19 at 11:59 PM.**
- Revised paper proposal with literature review (5%): Then, you will submit a revised paper proposal that incorporates feedback I gave you on your first proposal, includes a more thorough discussion of the motivation for the research question based on your annotated bibliography, and discusses what statistical technique you will use to answer your research question. This revised paper proposal should be significantly

longer than your original proposal, but no more than three pages single spaced, and will be due by **Friday, April 9 at 11:59 PM.**

- Final paper (25%): Your final research paper will be due on bCourses by **Thursday, May 13 at 11:59 PM**. This paper should be 10-12 pages double spaced, not counting a cover page, tables, or references. The paper will roughly follow the same format as your in-class presentation.

Final Exam (20%)

The Final Exam for our class will be on **Wednesday, May 12 at 3:00-6:00 PM.** The final will be open book and open notes, but due to the time constraints I strongly recommend you study the material ahead of time so that you do not need to rely on your notes for each question.

Weekly Discussion

Each week (Tuesday 11:00-Noon Pacific) we will have an optional hour of synchronous discussion time. I highly recommend attending and using this as a place to: get help with the homework, work through the week's material with others, ask questions about the final or paper, or for general office hours. If you would like to attend but the time does not work for you, please send me an email and we'll figure out an alternative arrangement.

Data for Assignments and Research Paper

Over the course of the semester, you will work with one (or more) datasets of your choosing in your assignments and your final paper. You may use any data you like and are encouraged to consult with me as you choose and begin working with your data. Below are some online repositories where data are available:

- The Survey Documentation and Analysis page at https://sda.berkeley.edu/archive.htm. This site has opinion data from the General Social Surveys (also available at https://gssdataexplorer.norc.org/), American National Election Studies (also available at https://electionstudies.org/data-center/), and various studies on racial attitudes and prejudice, among other surveys on household finance and health.
- The Inter-university Consortium for Political and Social Research (ICPSR) has thousands of surveys. You may search for surveys at https://www.icpsr.umich.edu/web/pages/ICPSR/index.html. ·
- The Association of Religion Data Archives (ARDA) also has a voluminous collection of surveys on religion, accessible at https://thearda.com/archive/browse.asp.
- The Integrated Public Use Microdata Series (IPUMS) is the most accessible way to access lots of U.S. and international census and survey data: https://www.ipums.org.
- Finally, the Roper Center at Cornell University has a massive database of public opinion polls: http://ropercenter.cornell.edu.

Keys to Success in Sociology 106

In some ways, learning statistics is like learning a language, and it is important not to be intimidated by new terms to represent quantities or concepts. It is often helpful to write in plain language the meaning of the quantities or concepts represented by a letter or symbol.

Because most of the material is cumulative, it is **essential** that you keep up with the course material. If you find yourself falling behind, seek help *immediately* from me during office hours.

The most effective way to study for the exams is to do practice problems. As such, the assignments are a critical part of the course. You are strongly encouraged to do homework assignments and study for exams in groups.

Grading Policy

If you wish to contest a grade, please e-mail me within one week of receiving the grade and outline in writing 1) the assignment you are contesting, 2) the grade you received on the assignment, and 3) why you believe your grade is unfair. I will consider your appeal and may decide to re-grade your assignment. Please note that a re-grade involves a closer scrutiny of your work and may result in you receiving a lower grade.

Academic Integrity

You can find the student honor code here. If you are unsure about the expectations of academic integrity, please feel free to talk to me about it.

Student Accommodations

It is my intention to make this an inclusive course. If you are experiencing disability related barriers in the learning environment, please contact DSP and they will assist you.

Getting Help

College is difficult, especially in these new circumstances, but I will do my best to ensure that you are supported. If you find you are struggling, please feel free to come talk to me. You can find additional resources here: counseling, crisis/suicide prevention hotline, social services counseling, health services, and food security and basic needs resources. You can find academic support services here. Please do not wait until you are suffering to reach out, although you can certainly do so even at that point.

Students who do not have access to the technology resources they need should check out the Student Tech Equity Program (STEP). Resources available include: a new laptop, Wi-Fi hotspot, or other resources to support their studies in time for Spring semester. Devices will be mailed to students' homes directly. If you have questions, please email ucbstep@berkeley.edu.

For coronavirus specific support please see https://coronavirus.berkeley.edu and/or https://technology.berkeley.edu/COVID-19.

For a longer list (borrowed from another grad student), please see the 'Resources to Support Student Well-being' pdf uploaded to bCourses.

I note that I (alongside other academic advisers and professors at UC Berkeley) am designated as a 'responsible employee', which means that I must report incidents of sexual violence and/or harassment that I am told about to the Office for the Prevention of Harassment and Discrimination. Students who are concerned about maintaining confidentiality are advised to use confidential resources such as the PATH to Care Center, UHS Social Services, Be Well at Work Employee Assistance, and the Ombuds Office for Students and Post-Doctoral Appointees.

Schedule

Week		Topic	Assignment Due
1	1/19-1/22	Finding Data	
2	1/25-1/29	Statistical Computing	Homework 1
3	2/1-2/5	Visualizing Distributions	Homework 2
4	2/8-2/12	Summarizing Distributions	Homework 3
5	2/15-2/19	Basics of Probability Theory	Homework 4
6	2/22-2/26	Probability Models of Distributions	**Paper Proposal**
7	3/1-3/5	Sampling Distributions	Homework 5
8	3/8-3/12	Estimates of Confidence Intervals	Homework 6
9	3/15-3/19	Hypothesis Testing	**Annotated Bibliography**
Spring Break	3/22-3/26		
10	3/29-4/2	Chi-Square Test	Homework 7
11	4/5-4/9	Linear Regression, i.	**Revised Paper Proposal**
12	4/12-4/16	Linear Regression, ii.	Homework 8
13	4/19-4/23	Logistic Regression	Homework 9
14	4/26-4/30	Extensions to Regression	Homework 10
RRR	5/3-5/7		
Finals Week	5/10-5/14	**Final Exam** 5/12 3:00-6:00 PM Pacific	**Final Paper** 5/13 11:59 PM Pacific