

Syllabus
Sociology 541: Analysis of Sociological Data I

Spring 2018

Instructor: Julie Phillips
Office: B039 Davison Hall, Douglass Campus
Email: jphillips@sociology.rutgers.edu (Best way to reach me)
Phone: (848) 932-7873 (Sociology office)
Office Hours: Thursdays, 1:00-2:00 pm or by appointment
Course URL: <https://sakai.rutgers.edu>

Lab Assistant: Laura Callejas
Office: Computer Lab, Douglass Campus
Email: callejas@sociology.rutgers.edu
Office Hours: Tuesdays. 12:30-1:30 pm; Wednesdays, 9:30-11:30 am.

Course Description

This course is the first part of a two-semester sequence (541 and 542) designed to introduce you to methods of research and elementary statistics. Through this course, you will be introduced to a range of standard statistical techniques used in sociological analysis. Indeed, a considerable amount of sociological inquiry relies on such techniques; research using large surveys, public opinion polls, and census data document, describe and explain a wide range of sociologically motivated research questions. Thus, this course provides a critical foundation for students in the social sciences. The course will be taught under the assumption that registered students have little or no statistical background.

Learning Goals

1. To introduce you to the basic concepts, terminology and procedures of data analysis, as well as to the logic underlying those procedures.
2. To understand how to calculate basic descriptive and inferential statistics and interpret them.
3. To acquire statistical literacy and be able to determine when, why, and how various statistical tests are used.
4. To learn a statistical software package to perform analyses of quantitative data.
5. To foster the ability to think critically about scientific and media reports of research findings based on quantitative data.

By the end of the semester, you should have a solid understanding of a variety of statistical concepts and techniques and be prepared to tackle multivariate regression, the starting point of the second course in the Sociology graduate statistics sequence.

Required Texts

Frankfort-Nachmias, Chava and Anna Leon-Guerrero. 2014. *Social Statistics for a Diverse Society*. 7th edition. Pine Forge Press.

Miller, Jane E. 2013. *The Chicago Guide to Writing about Multivariate Analysis*. Second edition. University of Chicago Press.

Recommended Readings

Healey, Joseph F. 2015. *Statistics: A Tool for Social Research*. 10th Edition. Thomson Wadsworth.

Acock, Alan. 2012. *A Gentle Introduction to Stata*, Fourth Edition. Stata Press.

On the course website, you can find articles that have employed some of the statistical techniques you will be learning in this course. I strongly encourage you to read these articles to get a sense of how these techniques are applied in sociological research. All additional assigned articles referenced below can be found on Sakai.

Other References

Agresti, A. and B. Finlay. *Statistical Methods for the Social Sciences*. Prentice Hall.
Devore, J. and R. Peck. *Statistics: The Exploration and Analysis of Data*.
Freedman, D., R. Pisani, and R. Purves. *Statistics*. (Less mathematical presentation)
Tanur, J. *Statistics: A Guide to the Unknown*. (Full of examples and a useful source for projects)
Wonnacott, T. and R. Wonnacott. *Introductory Statistics*. (More mathematical presentation)

Computing

There are a variety of powerful packages available to conduct statistical analyses; in my own work, I tend to use SAS and STATA (each one offers certain advantages). In this class, we will use STATA for Windows. I will provide handouts with information on the relevant STATA commands for this class and have included a couple of introductory manuals on the Sakai website. You may also find the following links helpful as you master the program.

<http://data.princeton.edu/stata/>
http://dss.princeton.edu/online_help/stats_packages/stata/
http://www.cpc.unc.edu/research/tools/data_analysis/statatutorial
<http://www.ats.ucla.edu/stat/stata/modules/>

Course Website

Please make sure that you check the course website before you come to class each week. I will often make available handouts that we will discuss in class on the website. I will also post important class announcements here.

Course Requirements

Readings: Readings from the text will be assigned each week and it is expected that you will have completed the reading before the class. I recommend that you read the assigned material once before class and then again afterwards. It is also strongly encouraged that you complete the review exercises at the end of each chapter to ensure that you've understood the concepts. Learning statistics requires lots of practice.

Problem Sets/Computer Assignments: Weekly assignments will be handed out at the end of each class and are due at the beginning of the next class. These assignments account for 25% of your final grade.

Class presentations: Students are expected to participate fully in this class. We will typically begin each meeting by reviewing homework assignments and I will ask students to present solutions to assigned problems at that time. In addition, we will often break up into small groups to review certain statistical concepts, and I will ask a representative from the groups to present solutions and conclusions to the class. Your participation in these presentations constitutes 15% of your final grade.

Mid-Term Examination: There will be a midterm exam addressing concepts covered in the first half of the course. The exam constitutes 25% of your final grade.

Final Oral/Written Report: You will be required to acquire your own secondary data set and apply the techniques learned in this course to analyze the data. You have several options with regard to the data set you choose to use. I will make available two data sets that you can analyze. One contains information on the fifty states of the United States and the other is the 2012 General Social Survey. These data sets contain a series of variables and you can construct a smaller data set from this information, focusing on a topic that is of greatest interest to you. If you prefer, you may analyze another data set with which you are familiar, formulating a question of interest. I ask that you each turn in a one-page proposal of your topic (see last page of syllabus for details), due on March 9. The last class will be devoted to oral reports of these findings. Each class member will be allotted approximately ten minutes toward this end. A short paper (10 pages) detailing the analysis will also be required. This report and paper account for 35% of your grade.

Statement on Diversity

The Rutgers Sociology Department strives to create an environment that supports and affirms diversity in all manifestations, including race, ethnicity, gender, sexual orientation, religion, age, social class, disability status,

region/country of origin, and political orientation. We also celebrate diversity of theoretical and methodological perspectives among our faculty and students and seek to create an atmosphere of respect and mutual dialogue. We have zero tolerance for violations of these principles and have instituted clear and respectful procedures for responding to such grievances.

Tentative Class Schedule

Week 1: January 16

Topic: Introduction to Statistics
Reading: FN and LG, Chapter 1 and Appendix F
Healey, Chapter 1
Recommended: Zerubavel, Eviatar 2007. "Generally Speaking: The Logic and Mechanics of Social Pattern Analysis". Sociological Forum 22(2): 131-145. (on Sakai)

Week 2: January 23

Topic: Basic Concepts
Displaying and Describing Data
Reading: FN and LG, Chapters 2 and 3
Miller, Chapter 2 (pp. 13-27)
Miller, Chapter 4 (pp. 49-60); Chapter 5 (pp. 77-89); Chapter 6 (pp. 113-120; Table 6.1)
"In calculation of Military Rates, the Numbers are not all Straightforward", *New York Times*, May 16, 2013 (on Sakai)
Healey, Chapter 2

Week 3: January 30

Topic: Introduction to Stata (Meet in computer lab)
Reading: Review materials in the Stata Resources Folder on Sakai (Stata Guides and Cheat Sheets)
Check out Stata learning modules (especially Fundamentals of Using Stata) both before and after the lab session at: <http://www.ats.ucla.edu/stat/stata/modules/>

Week 4: February 6

Topic: Measures of Central Tendency
Measures of Dispersion
Reading: FN and LG, Chapters 4 and 5
Miller, Chapter 4 (pp. 61-64)
Miller, Chapter 14 (pp. 297-302)
Healey, Chapters 3 and 4
"When Numbers Mislead", *New York Times*, May 25, 2013 (on Sakai)

Week 5: February 13

Topic: Sampling, Probability and the Normal Distribution
Reading: FN and LG, Chapters 6 and 7
Healey, Chapters 5 and 6

Week 6: February 20

Topic: Basis of Statistical Inference (Point Estimates and Confidence Intervals)
Reading: FN and LG, Chapter 8
Healey, Chapter 7

Week 7: February 27

Topic: Hypothesis Testing and Significance Tests
Reading: FN and LG, Chapter 9 (pp. 267- 281)
Baker, M. 2016. "Statisticians issue warning on p-values." *Nature* 531: 151. (on Sakai)
Healey, Chapter 8

Homework assigned this week will be due by noon on March 5.

Week 8: March 6

Topic: Review and midterm exam handed out (open-book). **Midterm due by noon on March 9.**
One-page summary of proposed paper topic is due on March 6

HAPPY SPRING BREAK! March 10-18

Week 9: March 20

Topic: Comparing Two Groups (t tests)
Reading: FN and LG, Chapter 9 (pp. 281-294)
Miller, Chapter 11 (pp. 235-242)
Miller, Chapter 13 (pp. 302-306)
Healey, Chapter 9

Week 10: March 27

Topic: ANOVA
Reading: FN and LG, Chapter 12
Healey, Chapter 10

Week 11: April 3

Topic: Bivariate Regression and Correlation/Linear Association
Reading: FN and LG, Chapter 13
“Eight (No, Nine!) Problems with Big Data” *New York Times*, April 6, 2014 (on Sakai)
Miller, Chapter 3
Healey, Chapters 12 and 15

Week 12: April 10

Topic: Cross-Tabulations and Elaboration
The Chi-Square Test
Reading: FN and LG, Chapter 10 and Chapter 11 (pp. 347-363)
Miller, Chapter 5 (pp. 89-92), Chapter 6 (pp. 120-156)
Healey, Chapter 11 and 16

Week 13: April 17

In lieu of class, I will be available for individual appointments with those students wishing to discuss their final report.

Reading: Miller, Chapters 7, 12-14
Presenting and talking about numbers (PowerPoint on “Writing about distributions and associations”)
Review Miller, Chapters 4 and 5 and bring the book to class

At your leisure (but before you take Soc. 542!), read the following:

Measures of Association for Nominal and Ordinal Variables (FN and LG, Chapter 11, pp. 363-374)
FN and LG, Chapter 12; Healey, Chapter 13 and 14

Week 14: April 24

Topic: Final Reports

Oral reports will be given in class. Written versions are due at the conclusion of class. No late papers will be accepted.

Sociology 541

Short Paper (due April 24, 2018)

One course requirement is to write a short paper (approximately 10 written pages, plus tables/graphs) on a topic of your choice. I recommend that you use one of the following data sets:

States07.dta

GSS2012.dta <http://www3.norc.umd.edu/GSS+Website/>

These data sets are available on the Sakai website and on the Sociology network (W:\jp-soc541). These data sources and the variables they contain are described in codebooks located on Sakai and on the Sociology network.

If you have access to other data sets that are more relevant to your research interests, you should feel free to use them if you prefer. The primary goal of the paper is for you to learn the statistical methods, but there's no reason why the work can't be related to your primary areas of research. Feel free to come to me if you are interested in doing work of this kind but are having trouble locating an appropriate data set.

Due on March 6, 2018: A one-page summary of your paper topic.

Please be sure to consider the following points in the summary:

1. What is your research question? What is the central relationship you are interested in? Identify your outcome variable and the explanatory or independent variables. You want to identify 4-5 variables with which you'll work. You can select one outcome variable and 4-5 independent variables or you can select multiple outcome variables and 2-3 independent variables.
2. What hypotheses can you draw about the relationship? Be sure to link these hypotheses to sociological theory. What guides your thinking and leads to these hypotheses?
3. How will you operationalize your hypotheses? What data set will you use? Which variables will you use from the data set to test your hypotheses?

I encourage you to come to me with any questions you may have about the paper early on in the semester so that we can make sure you're on the right track.