.Revised: September 2022

Introduction to Quantitative Data Analysis

Sociology 400

Fall 2022

Professor: Quincy Thomas Stewart

Teaching Assistant
Kris Rosentel
krisrosentel@u.northwestern.edu

Stats Tutor
Alice Kang
alice.kang@u.northwestern.edu

COURSE DESCRIPTION: Social scientists use quantitative methods to explore and test hypotheses, describe patterns in survey and census data, analyze experimental findings, and dynamically model social relations among individuals and groups. The aim of this course is to introduce students to the basic concepts of quantitative methods as they relate to social scientific research, and lay the foundation for more advanced graduate-level courses in multiple regression, regression models for categorical dependent variables, computational simulations, social networks, event-history analysis, etc. Students in the course will learn to: use graphs, tables, and measures of central tendency and spread; explain random sampling using probability concepts; explain what a sampling distribution is and give a rudimentary explanation of its role in inferential statistics; calculate and explain confidence intervals; test hypotheses about means, proportions, and pairs of means and proportions; test the hypothesis of independence in a contingency table; compute and interpret correlations and regressions for pairs of variables; and use R statistical software to perform basic statistical analyses.

COURSE REQUIREMENTS: The requirements include, **class attendance and participation** (10%), **homework** (30%), two **quizzes** (20%), and a **final program** (40%).

Attendance and participation: Class attendance, a very easy variable to measure, is based on your regular presence in class. Attendance is mandatory. Students will be penalized 4% of their final grade for each unexcused absence. I will excuse absences for sickness, religious holidays, in-patient hospital admissions, and military service. Please do not bring me documentation for any other reason.

Class participation will be measured through participation in the in-class discussions/problems during the course. Participation may come in the form of asking questions and making comments, as well as work on problems in the context of the class meeting—we'll often have break-out sessions to solve a problem independently or in small groups.

Homework: I will assign homework each week to be completed by the following week. The homework is designed to have you apply and interpret the material covered in the course. There

will be both questions to be solved by hand (using formulas) and computer work. I include both types of questions to: 1) increase your comfort with manipulating equations and 2) introduce you to statistical computing. In addition to individual homework, I *may* also assign group-based homework. In the case of group-based work, you will need to download a specific data set, perform an analysis covered in class, and interpret the results. The goal of group-based work is to familiarize each of you with downloading, manipulating and analyzing publicly available data. We will assign groups when there are group-based homework questions.

Quizzes: There will be two quizzes over the course of the quarter (10% each). Each quiz will consist of a few additional questions on a homework assignment designed to assess your understanding of the material covered in class. The quiz will consist of questions that you will solve by hand (i.e., without a computer). The questions generally entail: estimating and interpreting statistics; explaining theorems; solving for unknown quantities; performing tests; and showing mastery of other relevant aspects of introductory statistics. The quizzes will focus on material covered up to the respective point in the course.

Final Program/Paper: The final program/paper for the course takes the form of an R batch program, results and interpretation. The program will consist of a batch file where you code, notate (i.e., insert comments), and analyze a data set that we provide. This program will be the culmination of your work in the class and lab over the quarter, and must include clear documentation, be able to run without errors, and produce accurate results (i.e., statistics). You will also have to interpret the results of your program in an appended document. The program/paper is an assessment of your ability to apply the methods learned in this class to a real-world problem.

CLASS AND OFFICE HOURS: The class meets on Tuesdays and Thursdays between 2:00 p.m. and 3:20 p.m. in Kresge Centennial Hall room 2-410. The lab for the course will meet on Thursdays between 3:30 and 4:50am in the Kresge Centennial Hall room 2-325.

My office is located on the third floor of 1810 Chicago Ave in Room 322. I hold Office Hours each week; I am available on Mondays between 4:30 p.m. and 5:30 p.m. CST via Zoom or by appointment. (If you should come to my office for an in-person meeting, please DO NOT wear perfume or cologne. *I will have to ask you to reschedule our meeting if you wear perfume or cologne to an office meeting*.) My office phone number is 847-491-7044. My email address is q-stewart@northwestern.edu. *Please feel free to use Canvas to communicate with me as well; the messages do end up in my email, but I check the Canvas site more often when teaching.* I will be available for talking via email during my office hours.

Kris Rosenfeld, our TA, is also hosting office hours each week of the quarter. Kris is available to meet in-person on Tuesdays from 3:30 to 4:30pm, and via Zoom on Fridays from 10:00 to 11:00am.

Alice Kang, our Stats tutor, will be hosting officer hours each week of the quarter as well. Alice is available to meet in-person on Fridays from 2:30 to 3:30pm, via Zoom on Mondays from 10:00 to 11:00am, or by appointment.

Note regarding Syllabus Changes - I reserve the right to make changes to the schedule of readings and/or lectures during the course of the quarter. I will announce any such changes in class. You are responsible for noting the changes and preparing for class appropriately.

OTHER IMPORTANT INFORMATION

Note on Academic Integrity - Students in this course are required to comply with the policies found in the booklet, "Academic Integrity at Northwestern University: A Basic Guide". All papers submitted for credit in this course must be submitted electronically unless otherwise instructed by the professor. Your written work may be tested for plagiarized content. For details regarding academic integrity at Northwestern or to download the guide, visit: https://www.northwestern.edu/provost/policies-procedures/academic-integrity/index.html.

Note on Accommodations for Students with Disabilities - Northwestern University is committed to providing the most accessible learning environment as possible for students with disabilities. Should you anticipate or experience disability-related barriers in the academic setting, please contact AccessibleNU to move forward with the university's established accommodation process (e: accessiblenu@northwestern.edu; p: 847-467-5530). If you already have established accommodations with AccessibleNU, please let me know as soon as possible, preferably within the first two weeks of the term, so we can work together to implement your disability accommodations. Disability information, including academic accommodations, is confidential under the Family Educational Rights and Privacy Act.

Class sessions for this course will occur in person. Individual students will not be granted permision to attend remotely except as the result of an Americans with Disabilities Act (ADA) accommodation as determined by AccessibleNU.

Note on COVID-19 Classroom Expectations - Students, faculty and staff must comply with University expectations regarding appropriate classroom behavior, including those outlined below and in the COVID-19 Expectations for Students. With respect to classroom procedures, this includes:

- Policies regarding masking, social distancing and other public health measures evolve as the situation changes. Students are responsible for understanding and complying with current University, state and city requirements.
- In some classes, masking and/or social distancing may be required as a result of an Americans with Disabilities Act (ADA) accommodation for the instructor or a student in the class even when not generally required on campus. In such cases, the instructor will notify the class.

If a student fails to comply with the COVID-19 Expectations for Students or other University expectations related to COVID-19, the instructor may ask the student to leave the class. The instructor is asked to report the incident to the Office of Community Standards for additional follow-up.

Note on Recording - Unauthorized student recording of classroom or other academic activities (including advising sessions or office hours) is prohibited. Unauthorized recording is unethical and may also be a violation of University policy and state law. Students requesting the use of assistive technology as an accommodation should contact AccessibleNU. Unauthorized use of classroom recordings – including distributing or posting them – is also prohibited. Under the University's Copyright Policy, faculty own the copyright to instructional materials – including those resources created specifically for the purposes of instruction, such as syllabi, lectures and lecture notes, and presentations. Students cannot copy, reproduce, display, or distribute these materials. Students who engage in unauthorized recording, unauthorized use of a recording, or unauthorized distribution of instructional materials will be referred to the appropriate University office for follow-up.

Note on Supports for Wellness/Mental Health - Northwestern University is committed to supporting the wellness of our students. Student Affairs has multiple resources to support student wellness and mental health. If you are feeling distressed or overwhelmed, please reach out for help. Students can access confidential resources through the Counseling and Psychological Services (CAPS), Religious and Spiritual Life (RSL) and the Center for Awareness, Response and Education (CARE). Additional information on all of the resources mentioned above can be found here:
https://www.northwestern.edu/counseling/
https://www.northwestern.edu/religious-life/
https://www.northwestern.edu/care/

TEXTBOOKS:
(Available either online via the NU Library or through a direct link.)

Levine, David M. and David F. Stephan. 2004. *Even You Can Learn Statistics: A Guide for Everyone Who Has Ever Been Afraid of Statistics*. Upper Saddle River, NJ: Pearson Prentice Hall.(https://search.library.northwestern.edu/permalink/01NWU_INST/h04e76/alma9981878999302441).

Wickham, Hadley and Garrett Grolemund. 2017. *R for Data Science: Visualize, Model, Transform, Tidy and Import Data*. Sebastopol, CA: O'Reilly. (https://r4ds.had.co.nz/) 0

Yau, Chi. 2022. *R Tutorial: An R Introduction to Statistics*. www.r-tutor.com. (Provides an introduction to R and serves as a simple online manual of measures and basic statistical commands.)

Supplemental Text:
Agresti, Alan and Barbara Finlay. 2009. *Statistical Methods for the Social Sciences*. 4th Edition. New York: Prentice Hall.

Note: *Earlier editions and/or similar intermediate statistics texts will suffice as a supplemental text for those students in a more detailed discussion of the respective measure/method.*

SCHEDULE:

Week 1: Introduction

9/20	*Topics*:	Variables; Graphing Distributions
	Readings:	Levine & Stephen. *Chapters 1 & 2.1*
9/22	*Topics*:	Measures of Central Tendency; Standardization
	Readings:	Levine & Stephen. *Chapter 3.1*
	Lab:	Wickham & Grolemund. *Preface.* www.r-tutor.com/r-introduction (*Introduction section*)

Week 2: Descriptive Statistics

9/27	*Topics*: Standard Deviation; Variance; Measures of Distributional Position	
	Readings:	Levine & Stephen. *Chapter 3.2 & 3.3*
9/29	*Topics*: Bivariate Plots; Slope and Intercept (Review of Algebraic Equation for a Line); Cross-Tabulations.	
	Readings:	Levine & Stephen. *Chapter 2.2 & 2.3*

Week 3: Probability I

10/4	*Topics*: Long-Run Argument; Rules of Probability; Independence; Discrete and Continuous Probability Distributions	
	Readings:	Levine & Stephen. *Chapter 4*
10/6	Extended Q&A with Kris/Alice and subsequent Lab	

Week 4: Probability II

10/11	*Topics*: Binomial Probability Distribution; Normal Probability Distribution	
	Readings:	Levine & Stephen. *Chapter 5.1, 5.2 & 5.3*
10/13	*Topics*: Samples and Populations; Expected Value and Standard Error	
	Readings:	Levine & Stephen. *Chapter 6.1*
	Quiz I: Big Homework on material from weeks 1-4	

Week 5: Statistical Inference

10/18	*Topics*: z-scores; Law of Averages; Central Limit Theorem; Normal Probability Plot	
	Readings:	Levine & Stephen. *Chapter 5.4*

10/20 *Topics*: Confidence Intervals; CIs for Means
Readings: Levine & Stephen. *Chapter 6.2 & 6.3*

Week 6: From Inference to Significance Tests

10/25 *Topics*: CIs for Proportions; Sample Size.
Readings: Levine & Stephen. *Chapter 6.4*

10/27 *Topics*: Tests for Means and Proportions
Readings: Levine & Stephen. *Chapter 7.1, 7.2, 7.4, 8.1 & 8.2*

Week 7: Significance Tests

11/1 *Topics*: Type-I and Type-II Errors; Limitations of Tests
Readings: Levine & Stephen. *Chapter 7.3*

11/3 *Topics*: Comparing Proportions; Standard Error for a Difference
Readings: Levine & Stephen. *Chapter 8.3*
Lab Conducted Online: Materials available on Canvas

Week 8: Significance Tests

11/8 *Topics*: Comparing Means; Experiments (i.e., Dependent Samples)
Readings: Levine & Stephen. *Chapter 8.3* (same as prior class)

11/10 *Topics*: Contingency Tables; The χ^2-Test;
Readings: Levine & Stephen. *Chapter 9.1*

Week 9: Introduction to Regression I

11/15 *Topics*: Tests of Independence; Tests of Association.
Readings: Levine & Stephen. *Chapter 9.2*

11/17 *Topics*: Scatterplots; The SD Line; Correlation Coefficient (Pearson)
Readings: Levine & Stephen. *Chapter 10.1, 10.2 & 10.3*

Quiz II: Big Homework on material largely from weeks 5-8

Week 10: Introduction to Regression II

11/22 *Topics*: Changing SDs; Spearman Correlation; Kendall Correlation
Readings: Levine & Stephen. *Chapter 10.3* (same as prior class)

11/24 No Class (Thanksgiving)

Week 11: Introduction to Regression III

11/29 *Topics*: Bivariate Linear Regression; Graph of Averages; Least Squares Method; Inferences for Slope Coefficients
Readings: Levine & Stephen. *Chapter 10.6*

12/1 *Topics*: Plotting Residuals; Model Assumptions
Readings: Levine & Stephen. *Chapter 10.4, 10.5 & 10.7*

FINAL PROGRAM: Evaluation will cover both statistics and coding material from the entire class.

Week 12: Final Exams Week

12/7 Final Program/Paper Due at 5:00pm, 12/7/16