Version 1/26/24

Syllabus
Soc. 365
Data Management for Social Science Research
Spring 2024

Instructor: Christine R. Schwartz
(she/her/hers)
4462 Sewell Social Science
cschwart@ssc.wisc.edu

Office hours: Thursdays & Fridays
via Zoom,
https://cschwart.youcanbook.me
or by appointment

Instructor: Gracie Venechuk (née Finnigan-Fox)
(she/her/hers)
Sewell Social Science
finniganfox@wisc.edu

Office hours: Mondays & Wednesdays
via Zoom,
finniganfox.youcanbook.me
or by appointment

Lecture time and place: Mondays 9:55-11:50am, 3218 Sewell Social Science Building

Course description: This is a course in how to manage social scientific data. Statistics courses in Sociology (and other social sciences) provide students with a solid theoretical understanding of data analysis but typically do not provide sufficient training in how to actually prepare and work with real data to apply those analytical tools. Because the large majority of research time (in any research job) is spent on data management, this is an important shortcoming in your training. By the end of this course you will understand the structure of different types of social scientific data, how to clean messy data, how to effectively document data, how to merge data from multiple sources, and how to restructure data for analysis.

You will also learn techniques for visual display of data (i.e., graphing) to identify patterns and problems and to effectively convey information to consumers of your research. To this end, we will work with a widely used software package for the management and analysis of social scientific data (Stata). Hands-on, nuts and bolts work will be supplemented throughout the semester with discussion of the bigger picture – why is careful and effective data management and preparation so essential? You will also have a chance to work with a publicly available data set of your choosing to examine a question that interests you.

This is not a statistics course but a prior or concurrently taken statistics course—e.g., Soc. 360, or the equivalent—is required so that everyone has a baseline understanding of the statistics we will be working with. Knowing how and when to use basic descriptive statistics such as two-way tables, means, and correlations will be reviewed and is a required part of the course. No previous experience with Stata is necessary, but again, any familiarity you have with the program or similar programs (SPSS, SAS, etc.) will certainly be a plus. There are many on-line resources for learning Stata and for troubleshooting – you may want to explore some of the sites at http://www.stata.com/links/resources-for-learning-stata/.

Canvas course URL: https://canvas.wisc.edu/courses/390259

Instructional Mode: Face-to-face
Prerequisites: This course is open to graduate students and upper-level undergraduates. Sociology 360 or the equivalent, or Sociology 357 or the equivalent are required. Concurrent enrollment is permitted. Exceptions may be made with the instructor consent.
Credits: 3-credits. The credit standard for this course is met by an expectation of a total of 135 hours of student engagement with the course learning activities (at least 45 hours per credit), which include regularly scheduled instructor-student meeting times during class of 115 hours per week. The additional time is spent on required readings, problem sets, a final project, significant independent engagement with a data set of the students' choice for the final project, and other student work as described in the syllabus.
Course Designations: Honors optional.

Required textbooks: *Michael N. Mitchell, Data Management Using Stata: A Practical Handbook.* Either the first or second edition.

The data sets used in this book can be downloaded from: http://www.stata-press.com/data/dmus.html.

These files will allow you to replicate all of the examples in the book. We will download these files on the first day of class.

Computing: All assignments will require manipulation of sample survey data, using the statistical package Stata. There are many other similar statistical packages but Stata offers an excellent combination of power, flexibility, and ease of use. For those of you thinking about graduate school in Sociology or another social science, this course will be a wonderful opportunity to master the leading software package.

Stata is available via Winstat, meaning that you can do your work from any computer as long as you have an internet connection. An introduction to Winstat will be provided on the first day of class. You may also download a copy of Stata for free onto your own computer from the campus software library (https://www.doit.wisc.edu/services/software/). There are a number of types of Stata you can download. I recommend Stata/SE. While it is sometimes convenient to have Stata on your own machine, all instruction will be done in Winstat and there are complications with moving back and forth between file locations. Therefore, I recommend always using Stata on Winstat until you are quite comfortable with the programming environment.

Even experienced Stata users frequently use the help manuals. These are available electronically via Winstat. To access Stata manuals, you can just click on Start and Stata in Winstat – the manuals are in pdf format. Another quick way to find help (and my preferred option) is to Google your Stata question.

In addition to Stata (and other statistical resources available via Winstat), SSCC provides statistical consulting to students in this course. You may stop by 4226 Social Science to talk with a consultant if you need help. You can always email either or us or stop by our office hours as well. If you have a STATA question, don't beat your head against a brick wall for too long, but also please attempt to figure things out on your own before seeking help.

Course Requirements:

- Class participation & attendance: At the end of most lectures, you will have an opportunity to write a very brief lecture reflection via Canvas. In class, we will present a Canvas access code, which will enable you to write a reflection for a few minutes at the end of class. You can do this by downloading the "Canvas Student" app on your phone or using Canvas on your laptop. You may have *three* "free" lecture absences, that is, if there is a lecture reflection, you do not need to do five of these to receive 100% attendance. There are no "excused" absences. Please do not email us about making up lost lecture reflections. If you are unable to attend for some reason, you should arrange to catch up on what you missed from another student (not the instructors). If you have a situation that requires you not to be in lecture for a prolonged period of time, please email Prof. Schwartz.

- Weekly Exercises: Exercises are due before class the week after they are assigned. Model answers will be posted and we will spend some class time the following week to discuss questions about the assignments. Late exercises will not be accepted. If for some reason you do not complete your assignment on time, I encourage you to complete it on your own, but I will not accept it for credit. Students will not get credit for assignment that are turned in that have substantial problems or are incomplete. Because I understand that sometimes unanticipated events occur, *your lowest homework score will be dropped.*

  These assignments will count for 30% of your final grade so it is in your best interest to complete and submit these on time. They will be graded as either receiving "full credit"=1 or "no credit"=0. Assignments that are turned in on time, are complete, and demonstrate thought and effort will be given full credit. Assignments that are not turned in on time, are incomplete, or indicate substantial lack of effort will be given 0s. Putting effort into the homework and learning from the feedback you receive is the best way to do well on the exam and final project.

  I encourage you to work together to complete the assignments, but you must turn in individual assignments. Everyone should put effort into answering all the questions, but discussion and collaboration with others is acceptable and useful. In this context, it is academically dishonest to simply copy other people's work, but working together in the spirit of learning is encouraged.

  Please submit all assignments via the Canvas "Assignments" page by 9:50am (right before class) on the date that they are due. This deadline is firm.

- Tests: There will be two open-book exams on which you will be asked to conduct a series of data manipulations similar to those covered in the text and the weekly assignments. The second test will be non-cumulative, although much of the knowledge gained in the first half of the class will be necessary to carry out the tasks in the second half. The tests will count for 50% of your final grade (25% x 2).

    - Make-up tests. If you cannot take a test because of an unavoidable scheduling conflict (e.g., religious holiday, athletic event), you must **contact the instructor** via email at least *2 weeks prior to the exam date.* If you have an emergency that prevents you from taking an exam, **contact the instructor as soon as possible**. Permission of the instructor is required to take a make-up exam. A make-up exam will be scheduled either before or after the original exam date. Be aware that the

make-up exam may be different from the original exam.

- Final project: The course will culminate in a research project in which you will use the technical and analytic skills developed in class to address a research question of your choosing. This project will involve choosing a publicly available data set (in consultation with the instructor), carefully describing those data, addressing missing data, conducting consistency checks, recoding variables, and preparing basic descriptive results to answer your question. I will provide detailed instructions for the project early in the semester. **No late projects will be accepted.**

Grading: Final grades will be allocated as follows,

- Class participation & attendance: 5%
- Weekly assignments: 27%
- Final project: 18%
- Tests I & II: 50% (25% each)

Final grades will be computed as follows: A=92-100, AB=88-91, B=80-87 BC=76-79, C=68-75, D=50-67, F=< 50. I may curve the grades up if necessary, but I never curve down.

Course Learning Outcomes: We have designed this course to achieve the following learning outcomes designated as priorities by the Department of Sociology:

- *Conduct Research and Analyze Data.* Although professional-quality research requires graduate-level training, we expect that all undergraduate majors will be able to conduct small-scale research in which they formulate a research question, collect data, analyze results, and draw conclusions.
- *Communicate Skillfully.* Sociology majors write papers and make oral presentations that build arguments and assess evidence in a clear and effective manner.
- *Prepare for Graduate School and the Job Market.* Students use their social research skills to identify opportunities for employment or further study, assess their qualifications for these opportunities, and identify strategies for gaining the necessary knowledge and experience to improve their qualifications. Students are encouraged to develop and maintain portfolios of their written work and educational experiences to aid them in preparing applications.
- *Improve Project Management Skills*. Students will improve their skills in time management, ordering and executing a series of complex and inter-related tasks, and integrating distinct components of a project into a final product.

Academic Policies

Academic calendar & religious observances. Establishment of the academic calendar for the University of Wisconsin-Madison falls within the authority of the faculty as set forth in Faculty Policies and Procedures. Construction of the academic calendar is subject to various rules and laws prescribed by the Board of Regents, the Faculty Senate, State of Wisconsin and the federal government. For additional dates and deadlines for students, see the Office of the Registrar's pages. **Students are responsible for notifying instructors within the first two weeks of classes** about any need for flexibility due to religious observances.

Academic Integrity. By virtue of enrollment, each student agrees to uphold the high academic standards of the University of Wisconsin-Madison; academic misconduct is behavior that negatively impacts the integrity of the institution. Cheating, fabrication, plagiarism, unauthorized collaboration, and helping others commit these previously listed acts are examples of misconduct which may result in disciplinary action. Examples of disciplinary sanctions include, but are not limited to, failure on the assignment/course, written reprimand, disciplinary probation, suspension, or expulsion.

According to UWS 14, academic misconduct is defined as:

- seeks to claim credit for the work or efforts of another without authorization or citation;
- uses unauthorized materials or fabricated data in any academic exercise;
- forges or falsifies academic documents or records;
- intentionally impedes or damages the academic work of others;
- engages in conduct aimed at making false representation of a student's academic performance;
- assists other students in any of these acts.

For a complete description of behaviors that violate the University's standards as well the disciplinary penalties and procedures, please see the Office of Student Conduct and Community Standards website. If you have questions about the rules for any of the assignments or exams, please ask your instructor.

Accommodations. The University of Wisconsin-Madison supports the right of all enrolled students to a full and equal educational opportunity. The Americans with Disabilities Act (ADA), Wisconsin State Statute (36.12), and UW-Madison policy (UW-855) require the university to provide reasonable accommodations to students with disabilities to access and participate in its academic programs and educational services. Faculty and students share responsibility in the accommodation process. **Students are expected to inform faculty of their need for instructional accommodations** during the beginning of the semester, or as soon as possible after being approved for accommodations. Faculty will work either directly with the student or in coordination with the McBurney Center to provide reasonable instructional and course-related accommodations. Disability information, including instructional accommodations as part of a student's educational record, is confidential and protected under FERPA.
(See: McBurney Disability Resource Center)

Course evaluations, grievance and appeal rights. The Department of Sociology regularly conducts student evaluations of all professors and teaching assistants near the end of the semester. Student participation in course evaluation is an integral component of course development, and confidential feedback is important. UW-Madison strongly encourages student participation in course evaluations at the end of the semester. Students who have more immediate concerns should report them to the instructor or to the chair, 8128 Social Science (socchair@ssc.wisc.edu).

Institutional statement on diversity & inclusion: Diversity is a source of strength, creativity, and innovation for UW-Madison. We value the contributions of each person and respect the

profound ways their identity, culture, background, experience, status, abilities, and opinion enrich the university community. We commit ourselves to the pursuit of excellence in teaching, research, outreach, and diversity as inextricably linked goals. The University of Wisconsin-Madison fulfills its public mission by creating a welcoming and inclusive community for people from every background – people who as students, faculty, and staff serve Wisconsin and the world.

Mental health and well-being statement: Students often experience stressors that can impact both their academic experience and personal well-being. These may include mental health concerns, substance misuse, sexual or relationship violence, family circumstances, campus climate, financial matters, among others. Students are encouraged to learn about and utilize UW-Madison's mental health services and/or other resources as needed. Visit uhs.wisc.edu or call University Health Services at (608) 265-5600 to learn more.

Privacy of student records & the use of audio recorded lectures statement: Lecture materials and recordings for this course are protected intellectual property at UW-Madison. Students in courses may use the materials and recordings for their personal use related to participation in class. Students may also take notes solely for their personal use. If a lecture is not already recorded, students are not authorized to record lectures without permission unless they are considered by the university to be a qualified student with a disability who has an approved accommodation that includes recording. [Regent Policy Document 4-1] Students may not copy or have lecture materials and recordings outside of class, including posting on internet sites or selling to commercial entities, with the exception of sharing copies of personal notes as a notetaker through the McBurney Disability Resource Center. Students are otherwise prohibited from providing or selling their personal notes to anyone else or being paid for taking notes by any person or commercial firm without the instructor's express written permission. Unauthorized use of these copyrighted lecture materials and recordings constitutes copyright infringement and may be addressed under the university's policies, UWS Chapters 14 and 17, governing student academic and non-academic misconduct. View more information about FERPA.

Students' rules, rights, & responsibilities: Rights & Responsibilities

Sexual harassment and misconduct: The mission of the University of Wisconsin-Madison (university) is to provide a teaching, learning and working environment in which faculty, staff, students, and guests can discover, examine critically, preserve, and transmit knowledge, wisdom, and values that will improve the quality of life for all. To promote the institutional mission, the university is committed to creating and maintaining a campus community that is free from sexual harassment and sexual violence. This policy prohibits acts of sexual harassment and sexual violence (including sexual assault, dating violence, domestic violence, stalking, and sexual exploitation) in all programs and activities of the University. Individuals who engage in such acts, hereafter referred to collectively as sexual harassment and sexual violence, are in violation of this policy and are subject to disciplinary action. Individuals who are subjected to acts of sexual harassment or sexual violence in violation of this policy are encouraged to report these incidents. If you believe that you have been harassed, contact your instructor, the chair of the Department of Sociology (socchair@ssc.wisc.edu), and/or UW-Madison Office of Compliance. For more information, see the UW-Madison Office of Compliance website.

Course content:

Chapters from Mitchell are given for version 2 of the textbook.

Week 1 Jan 29 [Schwartz]: Introduction. Getting onto Winstat. What do data look like? Why is data management so important? Basics of data structure and do files. Inputting data into Stata.
Reading: Mitchell, Chs. 1 and 2 ("Introduction" and "Reading & importing data")
Assignment: Exercise 1

Week 2 Feb 5 [Schwartz]: Data cleaning
Reading: Mitchell, Ch. 4 ("Data cleaning")
Assignment: Exercise 2
Due: Exercise 1

Week 3 Feb 12 [Venechuk]: Labeling, codebook, documentation, and fact-checking
Reading: Mitchell, Ch. 5 ("Labeling datasets"); Long Ch. 2
Assignment: Exercise 3
Due: Exercise 2

Week 4 Feb 19 [Schwartz]: Recoding, creating new variables
Reading: Mitchell, Ch. 6 ("Creating variables")
Assignment: Exercise 4
Due: Exercise 3

Week 5 Feb 26 [Venechuk]: Basic descriptive statistics (tabulation and summarization)
Reading: Kohler and Kreuter Ch. 7
Assignment: Exercise 5
Due: Exercise 4
Due 2/28 11:59pm: Final project proposal

Week 6 Mar 4 [Venechuk]: Missing data
Reading: Allison Ch.1
Assignment: Exercise 6
Due: Exercise 5

Week 7 Mar 11 [Schwartz]: Combining datasets
Reading: Mitchell, Ch. 7 ("Combining datasets")
Assignment: Exercise 7
Due: Exercise 6
Due 3/13 11:59pm: Preliminary descriptive statistics for final project

Week 8 Mar 18 [Venechuk]: Test I

Week 9 Mar 25: **Spring Break – No Class**

Week 10 Apr 1 [Schwartz]: Processing observations across subgroups

Reading: Mitchell, Ch. 8 ("Processing observations across subgroups")
Assignment: Exercise 8
Due: Exercise 7

Week 11 Apr 8 [Venechuk]: Reshaping data (longitudinal data)
Reading: Mitchell, Ch. 9 ("Changing the shape of your data")
Assignment: Exercise 9
Due: Exercise 8
Due 4/10 11:59pm: Next step for final project

Week 12 Apr 15 [Venechuk]: Programming for data management
Reading: Mitchell, Chs. 10 & 11 ("Programming for data management" Parts I & II)
Assignment: Exercise 10
Due: Exercise 9

Week 13 April 22 [Schwartz]: Paper preparation, extensions, review
Reading: None
Due: Exercise 10

Week 14 (April 29) [Venechuk]: Test II

Friday, May 7th, 11:59pm: Final term papers due to Canvas assignment folder.