Syllabus

Soc. 365

Computing in Sociological Research

Fall 2015

General information:

Time: Friday 9:00-10:55am

Room: 3218 Sewell Social Science Building

Instructor: Christine Schwartz

Office: 4458 Social Science Building

Tel: 262-5791

Email: cschwart@ssc.wisc.edu

Office Hours: 9:00-10:30am Thursday

Class Page: Class materials will be posted on Learn@UW

Required textbooks:

Michael N. Mitchell, *Data Management Using Stata: A Practical Handbook.* (Available at UBS)

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.

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. 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/.

Course website: I have set up a page for the course on Learn@UW, which will contain this syllabus, slides from class, weekly assignments, and additional readings. Each Friday afternoon I will put up an "Illustrative Answer" to the exercise you turn in that day, which you may also download or print.

Computing: All assignments will require manipulation of sample survey data, using the statistical package Stata. There are many other similar statistical packages but I think that Stata offers the best 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 an excellent 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 Russell Dimond, Doug Hemken, or Mark Banghart if you need help. You can always email me or stop by my 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: You are expected to complete the assigned reading before each class. Lectures and in-class work will assume that you have read the assigned materials. Valuable class time is reserved for hands-on work, examples, discussion, and clarification. Class participation will not factor into your grade directly but you should assume that it will have an indirect impact (via your ability to complete the homework assignments and the quality of

your final project and exam). If you are unable to attend for some reason, you should arrange to get class notes and catch up on what you missed from another student.

- Weekly Exercises: Exercises are due before class the week after they are assigned. We do not have a TA or grader. Thus, I will post model answers and allocate 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.

  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.

  These assignments will count for 33% 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.

  Please submit all assignments via the Learn@UW Dropbox by 8:55am (right before class) on the date that they are due. The 8:55am deadline is firm.

- Test: On the last day of class, we will have an open-book exam on which you will be asked to work with a prepared data set to conduct a series of data manipulations similar to those covered in the text and the weekly assignments. The test will count for 33% of your final grade.

- 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. This project will count for 34% of your final grade in the course. 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,
- Weekly assignments: 33%
- Final project: 33%

- Final exam: 34%

I assign grades based on the following cut points,

A 100-92
AB 91-88
B 87-80
BC 79-76
C 75-66
D 65-50
F < 50

If necessary, I will curve grades up. I never curve grades down.

Accommodations. Please send the instructor an email <u>by the end of the second week of the course</u> if you are eligible for special arrangements or accommodations for testing, assignments, or other aspects of the course. This may be the case if English is your second language or you experience a physical or psychological condition that makes it difficult for you to complete assignments and/or exams without some modification of those tasks. Accommodations are provided for students who qualify for disability services through the McBurney Center. Their website has detailed instructions about how to qualify: http://www.mcburney.wisc.edu/. Provide a copy of your accommodations request (VISA) to the instructor by the end of the second week of class. We try to reserve rooms and proctors by the third week in class, so we must know of all accommodations by then.

If you wish to request a scheduling accommodation for religious observances, send an email by the end of the second week of the course stating the specific date(s) for which you request accommodation; campus policy requires that religious observances be accommodated if you make a timely request early in the term. See the university's web page for details: https://kb.wisc.edu/page.php?id=21698

Academic honesty. As with all courses at the University of Wisconsin, you are expected to follow the University's rules and regulations pertaining to academic honesty and integrity. The standards are outlined by the Office of the Dean of Students at http://www.students.wisc.edu/doso/academic-integrity/

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 Dean of Students website. If you have questions about the rules for any of the assignments or exams, please ask your instructor.

Departmental notice of grievance and appeal rights. The Department of Sociology regularly conducts student evaluations of all professors and teaching assistants near the end of the semester. Students who have more immediate concerns about this course should report them to the instructor or to the chair, 8128 Social Science (Pamela.oliver@wisc.edu). I will also ask for anonymous feedback from the class about 1/3rd of the way through the semester as a way to check in to see how things are going.

Department learning objectives. Beyond the specific substantive and methodological content I will cover in this course, I have designed this course to achieve the following instructional objectives 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.

Course content:

Week 1 (Sept 4): Introduction. Getting onto Winstat. What do data look like? Why is data management so important? Real world applications.

Week 2 (Sept 11): Inputting data into Stata
    Reading: Mitchell, Chapters 1 and 2; NYT articles ("How Companies Learn Your Secrets," "Sizing up Big Data," "How to Get a Job at Google" (parts 1 & 2))
    Assignment: Exercise 1

Week 3 (Sept 18): Data cleaning
    Reading: Mitchell, Chapter 3
    Assignment: Exercise 2
    Due: Exercise 1

Week 4 (Sept 25): Labeling, codebook, documentation, and fact-checking
Reading: Mitchell, Chapter 4; Long Chapter 2
Assignment: Exercise 3
Due: Exercise 2

Week 5 (Oct 2): Graphics in Stata (Guest lecture: Doug Hemken)
Reading: Mitchell Visual Guide Chapter 1
Assignment: Exercise 4
Due: Exercise 3

Week 6 (Oct 9): Recoding, creating new variables
Visit from Aaron Anderson, Advanced Analytics, Procter & Gamble (10:00-10:15am)
Reading: Mitchell, Chapter 5
Assignment: Exercise 5
Due: Exercise 4

Week 7 (Oct 16): Combining data sets
Reading: Mitchell, Chapter 6
Assignment: Exercise 6
Due: Exercise 5

Week 8 (Oct 23): Basic descriptive statistics (tabulation and summarization)
Reading: Kohler and Kreuter Chapter 7
Assignment: Exercise 7
Due: Exercise 6

Week 9 (Oct 30): Missing data
Reading: Allison Chapter 1
Assignment: Exercise 8
Due: Exercise 7
Due: Final project proposal

Week 10 (Nov 6): Processing observations across subgroups
Reading: Mitchell, Chapter 7
Assignment: Exercise 9
Due: Exercise 8

Week 11 (Nov 13): Reshaping data (longitudinal data)
Reading: Mitchell, Chapter 8
Assignment: Exercise 10
Due: Exercise 9
Due: Preliminary descriptive statistics for final project

Week 12 (Nov 20): Programming for data management 1
Reading: Mitchell, Chapter 9
Assignment: Exercise 11
Due: Exercise 10

Week 13 (Nov 27): **No class – Thanksgiving holiday**

Week 14 (Dec 4): Programming for data management 2
        Reading: TBA
        Due: Exercise 11
        Due: <u>Preliminary table & figure for final project</u>

Week 15 (Dec 11): Test

Friday, Dec 18th, 12:30pm: Final draft of term papers due to learn@uw Dropbox.