INTRODUCTION TO SOCIAL RESEARCH II

SOCI 206-001
Spring (Jan 7, 2019–May 3, 2019)
Tuesday, Thursday 9:30 AM – 10:45 AM
Avery Hall 111
University of Nebraska-Lincoln

Professor: Dr. Jeffrey A. Smith
Office: 706 Oldfather Hall
Email: jsmith77@unl.edu
Office Hours: Tuesday, Thursday 11:00 AM – 12:00 or by appointment

TA: Joseph Jochman
Office: 735 Oldfather Hall
Email: joseph.jochman@gmail.com
Office Hours: Wednesday 1–2PM and Friday 2–3 PM or by appointment

Prerequisite: SOCI 205: Introduction to Social Research I

Course Description
This course covers the basics of data analysis. By the end of the course you should have gained a set of practical skills: you should be able to manipulate data sets, produce graphics, and, in general, be able to answer substantive questions using real data. Students will gain experience using a statistical software package (Stata), while learning how to interpret statistical tests and measures. Students will thus learn new computer skills (specifically data management skills) as well as the basics of statistical inference; both are necessary to analyze data in a sophisticated manner. Note that the mathematical parts of the course are designed with sociologists in mind and little mathematical background is needed to do well.

The hope is that this class will prove useful in your future endeavors. At UNL, this course will prepare you for SOCI 495 (Senior Seminar), where you will produce a full research project on your own. Research skills will also be essential if you work with faculty on a research project (which I encourage you to do) through the UCARE or USTARS programs. Basic research/data skills are also very attractive on the job market, particularly given the increasing proliferation of data. Finally, this class will serve you well if you go on to future study in graduate school.

Achievement-Centered Education (ACE) Student Learning Outcomes
The University of Nebraska-Lincoln seeks to provide quality education to all of its students. To that end, it has designated certain classes as ACE certified. These classes provide and assess specific learning outcomes. As an ACE class, Sociology 206 - Introduction to Social Research II will facilitate Learning Outcome #3: Use mathematical, computational, statistical, or formal reasoning (including reasoning based on principles of logic) to solve problems, draw inferences, and determine reasonableness. This course introduces students to the basic statistical analysis of social data, including descriptive statistics, inferential statistics, bivariate measures of association, and multivariate techniques.
Course objectives

• Provide practical skills to analyze and draw conclusions from quantitative social science data. Emphasis will be placed on understanding, computing, and interpreting basic statistics; interpreting and evaluating survey research findings; and analyzing quantitative data with a statistical software program.

• Provide opportunities to acquire knowledge of how to answer social science questions and problems with quantitative and statistical data and make inference to a population of interest using descriptive statistics, bivariate associations, and multivariate techniques through lectures, homework assignments, hands-on activities, and course projects.

• Offer opportunities to acquire knowledge on drawing inference to a population and determining whether those inferences are reasonable through lectures, homework assignments, hands-on activities, and course projects.

• Instill an appreciation for the uses of quantitative data analytic techniques and their relevance to social issues and social life.

• Provide students with practical skills to analyze and draw conclusions from quantitative data with a particular emphasis on social science theories and questions.

• Provide students the necessary skills to become informed consumers of statistics in everyday life.

• Facilitate critical reading, interpretation, and understanding of published quantitative social science research literature while deepening the student’s appreciation for the empirical basis of sociology.

• Facilitate students’ understanding how social science researchers use the scientific method and hypothesis testing for testing sociological theories and questions that are appropriate for quantitative analysis.

• Facilitate students in the developing their statistical literacy, including descriptive statistics, inferential statistics, bivariate measures of association, and multivariate techniques, and developing their practical skills using a statistical software program (Stata).

ACE learning outcomes in this class will be assessed by:

• Homework assignments
• Quizzes
• Exams

Format of Class

The class will consist of lectures and labs. It is highly recommended that students attend lectures and look over the book prior to class. The material is cumulative (i.e. builds on earlier ideas) and it is in the best interest of the class if everyone attends regularly and no one falls too far behind. The class will include in-class problems and activities as well as quizzes. Students are expected to come to class prepared for the day. This includes bringing a simple hand calculator to class. You must bring the calculator to every class. This will make it easier for you to participate, as we will be working practice problems together.

The labs will introduce students to Stata. They will take place in Selleck 9046, a computer lab in the basement of Selleck Quadrangle. It is highly recommended that students attend the labs as basic Stata skills are necessary to complete the homework and exams.
Text and Required Supplies


- Software: the course will make use of the statistical software package Stata. Stata is available on the sociology server, which you may access from any computer on UNL’s campus. We will go over how to access the server and Stata in our first lab. You do not need to buy any software for this course, although you are welcome to buy a copy for your own computer if you like. If you want a copy on your computer, you should purchase the Student version of Stata, or Small Stata, which is 38 dollars: [http://www.stata.com/order/new/edu/gradplans/student-pricing/](http://www.stata.com/order/new/edu/gradplans/student-pricing/)

- The following website also provides many helpful examples concerning Stata: [https://stats.idre.ucla.edu/stata/](https://stats.idre.ucla.edu/stata/)

- The course webpage on Canvas will be used regularly for posting material, answer keys, and for discussion.

FAQ About Stata

Q) Why do we have to learn a statistical program?
A) The goal of this course is for you to gain experience doing data analysis. In order to actually do data analysis (i.e. answering a substantive question using real data) you must be able to use the tools of the trade. If you wanted to be a dentist, you would have to learn how to use a drill. It is no different for sociologists. Learning a statistical program is important even if you are a qualitative scholar. Qualitative scholars must be able to understand and engage with quantitative work, including the complex decisions made when variables are coded.

Q) Why do we have to learn Stata?
A) Good question, as there are many possible programs one could use to analyze data. I find Stata to be an ideal choice for this class. Stata is based on a set of simple commands. You will gain practice writing syntax (which is an important skill!) but the syntax is simple enough to pick up in a semester, unlike other programs. Stata is also widely used by sociologists, as it can handle almost any analysis you would wish to do. We will cover but a fraction of Stata’s potential.

Q) How hard is Stata to learn?
A) Like any new skill, picking up Stata may be difficult at first. My experience is that students tend to get the hang of it after the second or third homework. I promise that you will be successful if you come to lab, practice, do the homework, and ask questions when you are stuck.

Q) Can I do my homework in Excel instead?
A) No, it is an important component of the class to learn Stata. The skills learned through Stata are different than Excel (also useful to learn), and you cannot pick them up using Excel.
Q) Will I ever use Stata again?
A) You are likely to use Stata in future classes or in research conducted here at UNL. You may also use Stata in a job or internship. Given that, I believe that learning Stata is a useful endeavor even if you never use Stata again. The key here is not Stata itself, but rather the base programming skills picked up while learning Stata. Once you learn Stata, it will be much easier to learn *any* other program (R, SAS, python, SPSS, etc.); and that is really the goal of learning the software in the first place—to learn the logic of statistical programming so you can apply those skills to new settings and contexts. The goal, then, is to expand your ability to learn new things in the future (on the job, in grad school, etc.).

Q) Do I need to bring my computer to lab days?
A) Yes, I would recommend bringing your computer to lab days if it is laptop (otherwise that would be fairly impractical). The advantage of bringing your own laptop is that you learn how to use Stata on your own machine. For those who do not have a laptop, you can use the available computers in the lab.

Q) Can I pass this course without learning Stata?
A) Very unlikely.

Q) Are you making us learn Stata as some kind of punishment?
A) No.

**Grading**

**Option 1 Grading System**

This is the default grading system for all students. See alternative grading system below for those who have taken a very similar course in analogous departments.

1. Attendance/class participation: 10%

   Attendance is highly encouraged and will contribute to your grade. One must attend class *and* actively participate to earn full marks. You are expected to come to class on time and to be actively engaged. This includes paying attention, taking notes, working on problems in groups (when prompted) and asking questions. Leaving class early will be counted the same as not attending (unless there is a valid excuse and you let me know before class that you have to leave early). Coming late will also be reflected in lower participation grades. Arriving late to class is very disruptive and is disrespectful to your fellow students.

2. Exams: Midterm Exam 20%; Final Exam 30%

   There will be two exams making up half of the course grade. The second exam will be taken at the end of the course but will only cover the second half of the material (but note that this material is dependent on understanding the first half of the course…). Both exams will be take-home exams. The midterm exam will be passed out on Tuesday, Mar 5th and will be due the next
class, on Thursday, Mar 7th. The final exam will be passed out on the last day of class, Thursday Apr 25th, and will be due on our final exam day, Tuesday, Apr 30th at noon.

3. Homework Assignments: 25%

There will be 5 homework assignments throughout the course. Each assignment will be posted on the Canvas website and you will have one week to hand in the assignment (i.e. 7 days). The assignment will be docked 25% for each day it is late. Printing difficulties are not an excuse for turning in homework late. The assignments will require the use of Stata and provide a hands-on opportunity to use the statistical tools we talk about in class. Your homework must be printed out and handed to me at the beginning of class. If you cannot attend class, please hand in your homework in my box in Oldfather Hall. I will not accept emailed homework assignments. You must complete the assignments on your own, although you may ask Joe or I for help. If you simply copy another student’s assignment then I will consider that plagiarism and you will receive a 0 for the assignment. In general you will be graded on the accuracy and clarity of your answers. Note that the homework will be graded based on the accuracy of your answers (as well as completeness, clarity etc.). You do not get full marks for turning it in. In general, it is a good idea to start the homework early. This way if you run into trouble you can come into class and ask me questions. Students who start the homework earlier tend to do better in the class. Note that some of the questions will require Stata, while others will not.

4. Stata Lab Quizzes: 15%

We will have three in-class Stata quizzes. You will be given a hands-on assignment where you will be asked to analyze a data set, run various statistical tests and interpret the results. The idea is to get some practice doing Stata and running actual analyses, while being able to ask questions and get some help as you do it. You will have the entire class period to work on the quiz but you must turn in whatever you have at the end of the period. You must hand in your own assignment but you can ask Joe or myself for help. You may ask your fellow student for Stata help if you get stuck, but you must try it on your own.

<table>
<thead>
<tr>
<th>Final Grade</th>
<th>97-100</th>
<th>93-96</th>
<th>90-92</th>
<th>87-89</th>
<th>83-86</th>
<th>77-79</th>
<th>73-76</th>
<th>67-69</th>
<th>63-66</th>
<th>60-62</th>
<th>0-59</th>
</tr>
</thead>
<tbody>
<tr>
<td>A+</td>
<td>87-89</td>
<td>B+</td>
<td>77-79</td>
<td>C+</td>
<td>67-69</td>
<td>D+</td>
<td>60-62</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>83-86</td>
<td>B</td>
<td>73-76</td>
<td>C</td>
<td>63-66</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-</td>
<td>80-82</td>
<td>B-</td>
<td>70-72</td>
<td>C-</td>
<td>60-62</td>
<td>D-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There will be no opportunities for extra credit.
Option 2 Grading System

This is an alternative grading system for those who have taken a very similar course in analogous departments and feel like they understood the material very well and I agree that you can do this version of the class. The idea is that you should have already picked up the basics in similar courses and it doesn’t really make sense for you sit through it all again. The grading system lets you do a more independent project and take the exams but skip some of the other items, like homework and Stata quizzes. Of course you should still come to class and are welcome to do the homework and quizzes, but you do not have to complete these assignments. Let me emphasize that you are still responsible for the material presented in class as you are required to take both the exams. This means that you must learn Stata, as you will need to utilize Stata for the project and the exams. See details below.

Note that you must get my permission to do this grading system. If you do not speak to me, I will assume that you are doing grading system 1 and will grade you accordingly. I reserve the right tell you no, that you must do the option 1 grading system. If you are interested in this grading system, you must discuss this with me by January 15th. By Jan 15th, you must present me with the syllabus of the already taken course so I can decide if there is a high degree of overlap. You must also provide clear evidence that you have picked up the material (e.g., evidenced by getting a high mark in the previous class or an example exam).

Grading:

1. Project: 50% (40% for final paper and 10% for initial submission of proposal and pieces of project)

The main assignment is a research paper, in which you perform an empirical analysis on real data and write up the results. You must employ (at least) one of the methods from this course in your paper.

The research paper must include: a short introduction with a substantive justification for the project; a clear research question; a set of testable hypotheses; a description of the data and models (including variables used in the analysis); a description of the methods employed to answer the research question; tables/figures describing the results; a summary of results; and a conclusion. You must also include your Stata syntax used to complete the project. Note that the paper does not require a lit review or theory section. I want you to focus on the data, methods, and results sections. The paper should be between 10-15 pages, double-spaced, normal margins, etc.

I leave it you to decide what question you want to answer and what data you would like to use but you must get my approval for the project before you begin. I can also help you find data and a research question if you are struggling on your own. One strong constraint you will face during this class is time. It is not a good idea to leave the project to the very last minute, hoping to get it all done in a mad rush. You should be working on the project throughout the entire semester. This means that you cannot pick a dataset that will only be available late in the semester (or even worse, after the semester is over). You need to plan ahead and make sure the data is available
early enough to make the project feasible. With this in mind, you should have a usable dataset by mid February. Note that you need to have descriptive statistics by February 19th. You need to pick a different project if that is not possible.

You will be required to turn in initial results and pieces of the project as the semester moves along. This is designed to ensure that you are making sufficient progress. I am also requiring that you meet with me three times during the semester. This will ensure that you are making proper progress and that I am up to date on where the paper stands. It also gives me an opportunity to give you feedback face-to-face.

**Key Dates for Project**

Feb 7: Deadline for getting my approval for your proposed research project. This includes the core research question and the data you will use. After this date, I will take 1% off your final grade every day until you discuss your proposed project with me.

Feb 19: Descriptive statistics on key variables due

Mar 7: Draft of Introduction and Data Section

March 28: Deadline for meeting with me (a second time) to give an update of the project and to get feedback on initial drafts. After this date, I will take 1% off your final grade every day until you discuss your project with me.

April 2: Tables and figures showing results

April 16: Deadline for meeting with me (a third time) to give an update of the project and to get feedback on results. After this date, I will take 1% off your final grade every day until you discuss your project with me.

Apr 23: Final Paper Due.

2. Exams: Midterm Exam 20%; Final Exam 30%
You will also be required to take both exams. This is the same as above:

There will be two exams making up half of the course grade. The second exam will be taken at the end of the course but will only cover the second half of the material (but note that this material is dependent on understanding the first half of the course…). Both exams will be take-home exams. The midterm exam will be passed out on Tuesday, Mar 5th and will be due the next class, on Thursday, Mar 7th. The final exam will be passed out on the last day of class, Thursday Apr 25th, and will be due on our final exam day, Tuesday, Apr 30th at noon.
## Course Schedule (Subject to Change)

<table>
<thead>
<tr>
<th>Week</th>
<th>Date</th>
<th>Topic</th>
<th>Reading</th>
<th>DUE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Part I of Course: Describing Data and Introducing Stata</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Tue – Jan 8</td>
<td>Introduction to the Course</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thu – Jan 10</td>
<td>Distributions, Graphs and Measures of Central Tendency and Spread</td>
<td>Chapter 1, 2</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Tue – Jan 15</td>
<td>Stata Lab: Introduction to Stata (class in Selleck 9046)</td>
<td>Stata Help</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thu – Jan 17</td>
<td>Distributions, Graphs and Measures of Central Tendency and Spread</td>
<td>Chapter 2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Tue – Jan 22</td>
<td>Stata Lab: Introduction to Stata (class in Selleck 9046)</td>
<td>UCLA tutorial</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thu – Jan 24</td>
<td>Stata Lab: Summarizing Variables (class in Selleck 9046)</td>
<td>UCLA tutorial</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Tue – Jan 29</td>
<td>Measuring the Relationship between Categorical Variables</td>
<td>Chapter 3.1; 5.1-5.3</td>
<td>HW 1 due</td>
</tr>
<tr>
<td></td>
<td>Thu – Jan 31</td>
<td>Measuring the Relationship between Categorical Variables</td>
<td>Chapter 5.1-5.3</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Tue – Feb 5</td>
<td>Stata Lab: Measuring the Relationship between Categorical Variables (class in Selleck 9046)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thu – Feb 7</td>
<td>Stata Quiz/Practice Day 1 (class in Selleck 9046)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Tue – Feb 12</td>
<td>Stata Lab: Measuring the Relationship between One Quantitative and One Categorical Variable (class in Selleck 9046)</td>
<td>HW 2 due</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thu – Feb 14</td>
<td>Measuring the Relationship between Two Quantitative Variables</td>
<td>Chapter 3.2</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Tue – Feb 19</td>
<td>Stata Lab: Measuring the Relationship between Two Quantitative Variables (class in Selleck 9046)</td>
<td>UCLA tutorial</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thu – Feb 21</td>
<td>Bivariate Regression</td>
<td>Chapter 12.1-12.2</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Tue – Feb 26</td>
<td>Bivariate Regression</td>
<td>Chapter 12.1-12.2</td>
<td>HW 3 due</td>
</tr>
<tr>
<td></td>
<td>Thu – Feb 28</td>
<td>Stata Lab: Bivariate Regression (class in Selleck 9046)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Tue – Mar 5</td>
<td>Exam Review and Pass out Midterm</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Part II of Course: Statistical Inference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Tue – Mar 7</td>
<td>Samples and Populations</td>
<td>Chapter 1.2, 6.2, 7.2</td>
<td>Exam Due</td>
</tr>
<tr>
<td></td>
<td>Thu – Mar 12</td>
<td>Confidence Intervals</td>
<td>Chapter 7.1-7.2</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Tue – Mar 14</td>
<td>Confidence Intervals Continued</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thu – Mar 19</td>
<td>Spring Break</td>
<td>Chapter 8.1-8.4</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Tue – Mar 21</td>
<td>Spring Break</td>
<td>Chapter 8.1-8.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thu – Mar 26</td>
<td>Stata Lab: Confidence Intervals (class in Selleck 9046)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Tue – Apr 2</td>
<td>Hypothesis Testing</td>
<td>Chapter 9.1-9.5</td>
<td>HW 4 due</td>
</tr>
<tr>
<td></td>
<td>Thu – Apr 4</td>
<td>Hypothesis Testing</td>
<td>Chapter 9.1-9.5</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Tue – Apr 9</td>
<td>Hypothesis Testing: Comparing Groups</td>
<td>Chapter 10.1-10.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thu – Apr 11</td>
<td>Stata Lab: Hypothesis Testing (class in Selleck 9046)</td>
<td>UCLA tutorial</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Tue – Apr 16</td>
<td>Stata Quiz/Practice Day 3 (class in Selleck 9046)</td>
<td></td>
<td>HW 5 due</td>
</tr>
<tr>
<td></td>
<td>Thu – Apr 18</td>
<td>Hypothesis Testing: Relating Two Categorical Variables</td>
<td>Chapter 11.1-11.4</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Tue – Apr 23</td>
<td>Stata Lab: Hypothesis Testing: Relating Two Categorical Variables (class in Selleck 9046)</td>
<td>UCLA tutorial</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thu – Apr 25</td>
<td>Review and Pass Out Exam</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tue Apr 30</td>
<td>Final Exam due at noon</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Course Policies

Classroom Conduct
Students are expected to be engaged in the course. Out of respect for you fellow students, I ask that cell phones be turned off and that students refrain from distracting behavior (i.e. inappropriate talking or web browsing). In general the course will be much more enjoyable if everyone is considerate of the other individuals in the room. Excessive talking or distracting behavior will not be tolerated. Students who participate in distracting behavior will receive a 0 for the day for participation, as will any student who is found texting or talking on a phone.

Academic Misconduct (or Don’t Cheat)
“The maintenance of academic honesty and integrity is a vital concern of the University community. Any student found guilty of academic dishonesty shall be subject to both academic and disciplinary sanctions. Academic dishonesty includes, but is not limited to, the following: Cheating; Fabrication or Falsification; Plagiarism; Abuse of Academic Materials; Complicity in Academic Dishonesty; Falsifying Grade Reports; Misrepresentation to Avoid Academic Work.”
Quoted from the UNL Student Code of Conduct

Disabilities
“It is the policy of the University of Nebraska-Lincoln to provide flexible and individualized accommodation to students with documented disabilities that may affect their ability to fully participate in course activities or to meet course requirements. To receive accommodation services, students must be registered with the Services for Students with Disabilities (SSD) office, 132 Canfield Administration, 472-3787 voice or TTY.”

If you need accommodations it is your responsibility to discuss this with me early on in the semester.

Exam and Paper Policy
The midterm is a take home and must be turned in on Thursday, Mar 7th at the beginning of class. The final exam must be turned in by the end of the official exam period for the class, stipulated by the university (Tuesday Apr 30th at noon). There will be no exams accepted after the fact unless there are extraordinary circumstances and the student has received permission from me to turn in the exam late.