<u>**SOCY-6111: Data II**</u>
Spring Semester 2015

Instructor: Ryan K. Masters
Office: Institute of Behavioral Science 1B33
Office Hours: Wednesdays, 12:00-2:30 pm
Email: ryan.masters@colorado.edu

Course Overview

This course is intended to develop your understanding and use of statistical techniques for answering sociological questions. The course has three general aims:

(1) Introduce key matters pertaining to effective research design in the social sciences.
(2) Expand your understanding of statistical techniques for sociological inquiry.
(3) Develop efficient and sound use of Stata for quantitative analyses of sociological questions.

Overall, my hope is for you to develop understandings of quantitative thinking, learn skills necessary to effectively describe and analyze quantitative data, and adopt a critical approach when reading and assessing others' empirical claims.

Research Design and Causal Thinking

Some course time will be devoted to reading and discussing theories of quantitative understandings of sociological phenomena. H.L. Mencken once derisively said, "Sociology is the outhouse in the grove of academe." I believe there is some truth to this depiction of the field, but it's one we ought to embrace. To help us muddle through all the $#*% we will read various perspectives on methods and claims-making. Convincingly establishing a "relationship" in the sociological world is an incredibly difficult thing to do. I hope to provide a few readings that will show this to be the case, but also present a way forward for you as young researchers.

Statistics

The core of this course will be devoted to expanding your understanding and use of statistical techniques. I will review ordinary least squares (OLS) and maximum likelihood estimation (MLE), as well as introduce you to a few extensions of generalized linear models (GLM). We will also cover issues related to measurement, model building, interpretation of results, collinearity, omitted variable bias, violations of B.L.U.E., two-way effects, and other important considerations in quantitative research.

Computer Programming

Statistical analyses in the social sciences are now entirely performed by computer programs. Gone are the days of hand computations, bean counting, and clunky card-reading machines. Thus, this course will feature a heavy dose of statistical analyses using Stata programming, as well as interpreting Stata-generated output. All Stata scripts will be provided so that you can load, edit, and analyze raw data on your own time/terms.

A required Stata lab will be a central component of Data II.

Course Meetings

Class: Mondays, 3:30-5:00 pm   Hellems 196
Stata Lab: Wednesdays, 6:00-7:00 pm   Environmental Sustainability Wing, ECES 107

Course Material

Statistical Methods for the Social Sciences, 4th Edition. By Alan Agresti and Barbara Finlay. Referred to as "A&F" in the schedule.

Counterfactuals and Causal Inference: Methods and Principles for Social Research. By Stephen L. Morgan and Christopher Winship. Referred to as "M&W" in the schedule. Either 1st or 2nd Edition is fine.
http://www.cambridge.org/us/academic/subjects/sociology/sociology-general-interest/counterfactuals-and-causal-inference-methods-and-principles-social-research

Understanding Regression Assumptions. Sage Series: Quantitative Applications in the Social Sciences. By William D. Berry. Purchase online:
http://www.sagepub.com/books/Book3056?seriesId=Series486&sortBy=defaultPubDate+desc&rows=50&pager.offset=50&fs=1
Referred to as "Berry" in the schedule.

Logistic Regression: A Primer. Sage Series: Quantitative Applications in the Social Sciences. By Fred Pampel. Purchase online:
http://www.sagepub.com/books/Book10146?seriesId=Series486&rows=50&sortBy=defaultPubDate%20desc&fs=1#tabview=toc
Referred to as "Pampel" in the schedule.

Online Readings: I will occasionally post readings and Stata material to D2L.

Recommended Texts:

Regression: A Second Course in Statistics. By Thomas H. Wonnacott and Ronald J. Wonnacott.
http://www.amazon.com/Regression-A-Second-Course-Statistics/dp/0898749700

A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences. By Gary Goertz and James Mahoney.
http://press.princeton.edu/titles/9898.html

Course Requirements and Assessment

There will be 300 possible points in this course, broken down as follows:

Assessment	Points
Homework Assignments (5)	100
Midterm (in-class)	50
Final (take-home)	100
Response Papers (2)	30
Participation (Discussion, Stata Lab)	20

POLICIES & ACCOMMODATIONS

Accommodations

1. Appropriate academic accommodations will be provided to students with disabilities. Please contact the Disability Services office located in Center for Community as soon as possible to obtain documentation: N200 (303-492-8671) http://disabilityservices.colorado.edu/ . Guidelines for addressing temporary medical conditions and/or injuries can be found at http://disabilityservices.colorado.edu/general-information/temporary-injuries

2. The University of Colorado at Boulder has both legal and moral obligations to accommodate students who choose to abstain from classes and/or miss scheduled examinations in order to observe holidays. If you plan to be absent from class to observe a holiday, please notify me of any scheduling conflicts, in writing, by February 3.

Course Expectations and Honor Code

1. I expect academic integrity (and the university requires it!). While I encourage you to collaborate with one another on exercises, support each other in studying, and edit each other's work, you are expected to turn in original work and complete all exams on your

own. Students caught cheating will be reported to the Honor Code Council, and will also have their course grade justly penalized. Information about the Honor Code can be found at http://www.colorado.edu/policies/student-honor-code-policy

2. Every homework assignment shall be turned in during the first five minutes of class on the scheduled due date. All assignments turned in after this time on the same day or emailed on the due date will be punished two points. Please type your homework assignments, exams, and papers.

3. Please know that the syllabus and course schedule are not set in stone. I reserve the right to change the basic course requirements, due dates, and overall content and schedule with adequate notice to students via D2L, class announcements, and/or email.

Classroom Etiquette

Please refrain from conversing with your neighbors during class. This can be quite disruptive to fellow students around you.

Laptops are not needed in class.

Turn off all cell phones before class begins.
You and I both have the responsibility for maintaining a professional learning environment. Those who fail to adhere to basic modicum of adult behavior may be subject to discipline. Please be courteous and sensitive to alternative perspectives, especially when dealing with topics pertaining to race, culture, religion, sexuality, political ideology, nationality, gender identity & expression, age, and disability.

Please know that the University provides me a class roster containing your picture and legal name. I will happily honor your request to remove your picture and/or address you by an alternative name if you like. Please notify me early in the semester.

Email Policy

Include "6111" in the subject line of all course-related emails.

Respect, Discrimination, and/or Harassment

Please respect your classmates. Topics discussed in class may be interpreted as contentious by some, and I would like everyone to feel comfortable enough to freely and openly participate. I will do my best to present the material and discuss the topics in an open and inclusive manner.

Spring 2015 Data II Schedule

January 12

Syllabus, Course Overview, and Expectations: What you Know and What We'll Cover

January 14

Stata Basics 1: Managing Do-files

January 19 – MLK Day: No Class

January 21

Stata Basics 2: Loading and Cleaning Data, Labeling and Editing Measures

---

January 28

Approaches to Asking and Answering Quantitative Questions in Social Sciences

Abbott (1997)   "Of Time and Place"
Gelman and Imbens (2013)   "Why Ask Why?"
Lieberson & Lynn (2002)   "Barking Up the Wrong Branch"
Platt (1964)   "Strong Inference"
Xie (2007)   "Otis Dudley Duncan's Legacy"
M&W   Introduction, Chapter 1, and Chapter 2

Response Paper

January 30

OLS and Observational Data: Berkeley Admissions Example

February 2

Regression Models and Measurement: Means, Covariates, and "Error"

A&F   Chapter 2 and Chapter 9

February 4

OLS and Coefficients: Interpreting Xs and Ys

February 9

Review of OLS and B.L.U.E.

A&F   Chapter 9 (yes, again!)
Berry   Pages 1-18

February 11

OLS and B.L.U.E: Interpreting Model Output

February 16

Violations of B.L.U.E. and Estimators: Unbiasedness, Efficiency, and Consistency.

Berry   Pages 18-41

Homework #1 Due

February 18

Relaxing OLS Assumptions I

February 23

Violations of B.L.U.E. and Model Diagnostics

Berry   Pages 41-83

February 25

Relaxing OLS Assumptions II

March 2

Regression Estimators of Causal Effects: Multivariate OLS and Identification

F&A   Chapter 10
M&W   Chapters 3-5

Homework #2 Due

February 25

Controls: Omitted Variable Bias

March 9

Model Building Using Regression

F&A   Chapter 14

Homework #3 Due

March 11

Mechanisms: Mediators and Moderators I

March 16

***************************** MIDTERM ********************************

March 18

Mechanisms: Mediators and Moderators II (Optional Lab)

March 23

SPRING BREAK

March 25

SPRING BREAK

March 30

Natural Experiments

Jackson & Cox (2013)	"Experiments in Sociology"
Kirk (2009)	"Recidivism in Post-Katrina NO"
Lauderdale (2006)	"Arabic Birth Outcomes in Post-9/11 America"
Sharkey (2010)	"Violence and Cognitive Development"
Zahran et al. (2010)	"Lead-Soil Levels in Post-Katrina NO"

Response Paper

April 1

Advanced Model Fitting: Nonlinear Functional Forms, Fixed Effects, and Multiple Interactions

---

April 6

Categorical Data Analysis I: Contingency Tables, Odds, and Probabilities

Pampel   Chapter 1 and Chapter 2

April 8

Contingency Tables

April 13

Categorical Data Analysis II: Maximum Likelihood Estimation and Logistic Regression

Pampel   Chapter 3 (skip Probit) through Conclusion

April 15

Logit Models: Fitting, Interpreting, and Presenting

April 20

RCTs and the "Gold Standard" Myth: Selection, Selection, and *Selection*

Sampson (2010)	"Gold Standard Myths"
Sampson and Sharkey (2008)	"Selection as Process"
Hernandez-Diaz et al. (2008)	"Smoker Paradox?"
Dowd and Hamoudi (2014)	"Is US Life Expectancy Falling?"
Brand and Xie (2010)	"Who Benefits From College?"

Homework #4 Due
Response Paper

April 22

Selection: "Seeing" Data and Propensity Scores

Final Assigned

April 27

Categorical Data Analysis III

| Powers (2006) | "Handout on Categorical Data and Logit Models" |

April 29

Model Building with Categorical Data

Homework #5 Due

May 6

Final Due