QUANTITATIVE ANALYSIS OF SOCIAL DATA
SOCY 10
Fall Quarter 2015
(MWF, 10am-11:05am; x-period, Th 12-12:50pm)
Location: 105 Thornton Hall

Professor: Jason Houle   *Office*:104 Silsby Hall
Email: Jason.Houle@Dartmouth.edu   *Phone:* 646-3524
Office Hours: Wednesday, 3:15pm-5:15pm, or by appointment
TA (STATA help): Jianjun Hua Jianjun.Hua@Dartmouth.edu (Office Hours: Monday 2pm-4pm, 010 Silsby)

Course Description:

This course is intended to provide an introduction to basic statistical techniques used by researchers to investigate social, economic, and political phenomena. The statistical techniques you will learn are really just tools to effectively organize data and present information about the world. For example, researchers have recently employed these statistical techniques to answer the following kinds of questions:

> "How have societal opinions about same-sex marriage changed over time?"
> "Do people who grow up in impoverished communities have better or worse health than those who did not?"
> "Are birth defects related to local levels of water pollution?"
> "Is a history of depression associated with premature death?"
> "Does getting a college degree have the same labor market benefits for blacks and whites?"
> "Is student loan debt leading the current generation of young adults to forgo home-buying?"

By the end of this quarter, you too will be skilled enough in the art of statistics to address these types of questions. Just as important, however, this course is designed to help you become informed consumers of social statistics. This will involve attaining an awareness of the many uses of statistics in everyday life, gaining the ability to think critically about the use of statistics, and developing the skills necessary to interpret and critique social scientific research employing these techniques.

This course is designed to introduce you to the logic of statistical analysis, not to test your mathematical abilities. A rudimentary familiarity with basic arithmetic and a few algebraic functions will suffice for material presented in the course. This means that, even if you are insecure about your math prowess, this class should not intimidate you. More than anything, success in this course requires diligence. The material covered in the course is very incremental, with the presentation of new material building on what you learn in earlier segments of the course. You will have the greatest chance of success if you attend class regularly, take meticulous notes, keep up on your reading, and make sure to visit office hours to get clarification on issues about which you are unsure. In the past, students who have not taken these steps to be proactive in this course have not done well. I reserve the right to use X hours as an impromptu review session if we feel that some issues need further clarification. In addition, please do not hesitate to ask for x-hour review sessions!

Course Readings

Required Texts
Healey, Joseph. 2014. *Statistics: A Tool for Social Research, 10th edition.* Stamford, CT: Cengage.

*I recognize that the latest edition of this book is relatively expensive. Thus, you are welcome to purchase a previous edition of the book at a reduced cost. That said, you are responsible for the material covered in Healey 2014. If you cannot afford the book, please see me as soon as possible. A copy will also be available on reserve at Baker-Berry.

Below is a list of optional texts that you may find useful to have on your bookshelf. I will be drawing from these readings (and will occasionally assign chapters on Canvas) throughout the course.

Optional Texts:

Acock, Alan. 2012. *A Gentle Introduction to Stata* (Revised Third Edition). College Station, TX:
Stata Press. ****If you are having trouble with STATA, this is an excellent resource.**

Best, Joel. 2001. *Damned Lies and Statistics: Untangling Numbers from the Media, Politicians, and Activists.*
Berkeley: University of California Press.

Firebaugh, Glenn. 2008. *Seven Rules for Social Research.* Princeton: Princeton University Press

Silver, Nathan. 2012. *The Signal and the Noise: Why So Many Predictions Fail—But Some Don't.* New York:
Penguin.

Wheelan, Charles (Dartmouth '88). 2013. *Naked Statistics: Stripping the Dread from the Data.* New
York: WW Norton.

Ziliak, Stephen T. and Deirdre N. McCloskey. 2008. *The Cult of Statistical Significance: How the Standard
Error Costs Us Jobs, Justice, and Lives.* Ann Arbor: University of Michigan Press

For a good application of statistics to everyday life, I highly recommend that you add these blogs to your daily reading:

> Family Inequality (Phil Cohen): http://familyinequality.wordpress.com/
> Five Thirty Eight (Nate Silver et al.): http://www.fivethirtyeight.com/
> Vox (Ezra Klein et al.) www.vox.com
> The Upshot (NYTimes) http://www.nytimes.com/upshot/
> Wonkblog (WaPo): http://www.washingtonpost.com/blogs/wonkblog/
> Statistical Modeling, Causal Inference, Social Science (Andy Gelman) http://andrewgelman.com/

Statistical Software and Electronics

We will be using STATA 13 for in-class demonstrations, problem sets, and the final project. STATA is available for free on the Dartmouth Network (Mac/PC) and is also available on most public computers on campus. For more information on STATA resources and installing STATA on your computer, see the following websites:

http://tech.dartmouth.edu/its/services-support/help-yourself/knowledge-base/software

MAC: http://tech.dartmouth.edu/its/services-support/help-yourself/knowledge-base/stata-macintosh-os-x
PC: http://tech.dartmouth.edu/its/services-support/help-yourself/knowledge-base/stata-windows

Please use these resources and the helpdesk for any installation questions.

In some classes, we will be using STATA in order to demonstrate statistical techniques. This will generally occur every Friday. For this reason, you are encouraged to bring your laptop to class so that you may follow along on Fridays.

A brief note on STATA: Learning a new statistical analysis program and programming language is **hard**. I highly recommend that you practice outside of class to hone your skills. I recommend you come see me during office hours with all STATA related questions. In addition, here are some resources that will help you as you learn to code in STATA:

Jianjun Hua's research guides http://researchguides.dartmouth.edu/content.php?pid=316205&sid=2587562

UCLA Stata Resources
http://www.ats.ucla.edu/stat/stata/

Princeton Stata Resources:
http://data.princeton.edu/stata/
http://dss.princeton.edu/training/StataTutorial.pdf
http://dss.princeton.edu/online_help/stats_packages/stata/

You will also need to have a <u>calculator</u> to complete your assignments and exams, and to participate in in-class exercises. A basic and inexpensive calculator that has square-root and squaring functions is best for use in this class. *Make sure to bring your calculator to class every day and to all exams.* You are not permitted to use cell phone calculators for exams, but you are welcome to use them during regular class periods.

Evaluation and Grading

Exam 1,2,3 (20% each) There will be three exams during the course of the term. The exams <u>are not</u> open book. For exams, students are permitted to bring in one *hand-written* 8.5x11 sheet of paper (front and back) with notes and equations. Other materials that will be needed during the exams (such as a copy of the standard normal table) will be provided by the instructor when applicable. For Exams 1 and 2, you will be allotted 65 minutes to complete the exam. For Exam 3 you will be allotted 120 minutes to complete the exam.

Homework/Problem Sets (20%) There will be 8 homework assignments this term. Assignments will be due at the beginning of class <u>on Monday</u>. **Assignments must be handed in during the class period—do not email them to me.** Late assignments (even those assignments handed in on Monday after class) will be graded according to the late assignment policy below. **For all homework assignments (with the exception of HW#1) you <u>will need to turn in your STATA output (LOG files and DO files)</u>.** If you do not do so, you will lose a great deal of points.

For most of these assignments you will use STATA. Students are welcome to work together and help each other on assignments, **but each person MUST turn in their own assignment and provide answers <u>in their own words</u>**. Assignments will be posted on Canvas by Monday of the week before they are due (at the latest). Please post any questions about the assignments to the "Homework Discussions" forum on Canvas. I will be checking this forum occasionally, but its primary purpose is to provide you a place to help one another understand the material. **<u>Please do not email me with questions about the individual assignments.</u>**

Group Project (Report and Presentation) (20%) The group project is one of the most important, and perhaps the most rigorous aspect of this course. See the description in Canvas. These projects will require a significant time commitment, so it's important to choose a topic that you're excited about, and get started early! At the end of the term, you will be asked to evaluate the contribution of your fellow group members to the project. These evaluations will be factored into your project grade. In addition, the group project has several components, with the final report worth the lion's share of the points. The breakdown for the group project grade is as follows:

Group Project Assignment 1:	5%
Group Project Assignment 2:	5%
Final Presentation:	15%
Final Report:	75%

<u>*I will use the following scale in assigning grades:*</u>

93%-100%:	A	83%-86.9%:	B	73%-76.9%:	C	Below 60%:	F
90%-92.9%	A-	80%-82.9%	B-	70%-72.9%	C-		
87%-89.9%:	B+	77%-79.9%:	C+	60%-69.9%:	D		

Please note the following about grades: 1) I do not round grades; 2) I **<u>do not negotiate final grades</u>** unless an error was made.

<u>One final note:</u> To receive full credit on homework assignments and exams, you **<u>must</u>** show your work. If you write down the answer without showing how you got to that answer, <u>even if that answer is correct</u>, you will not receive full credit.

Attendance Policy
As a general rule, I do not believe in attendance policies for college-aged students. If you must miss class, <u>you do not need to provide me with an explanation</u>. More importantly, do not ask whether you missed anything

important (if the material you missed wasn't important, it wouldn't have been covered in class). Do not ask me for lecture notes—ask your classmates. I will assume that you are serious about your commitment to this class. As such, I presume that if you miss class you have a good reason (e.g., you are very sick, you are caring for someone who is very sick, or you are trapped under something very heavy). One exception: If your absences become chronic, or if you anticipate chronic absences, then it's time to consult with me.

All that said, if you blow off class and do so on a regular basis, you will likely fall behind and find it very difficult to catch up. Fair warning: if this happens, I am unlikely to sympathize with your plight.

Late Policy

On any assignment, your grade will be reduced by 8 percentage points for each day late (e.g., 1 day late a grade of 100 would be reduced to a 92; 2 days, 84; and so on). Assignments that are more than 5 days late will not be not be graded and will receive a 0. Given that assignments are posted far in advance, no exceptions to the late policy will be made, including for planned and unplanned absences.

Make-up exams: There will be no make-up examinations except in extremely rare cases in which some unforeseen crisis/emergency arises. If you know ahead of time that you have a conflict with the exam schedule, discuss this with me as soon as possible to make arrangements for the exam. Do not expect to arrange different exam schedules simply because it is more convenient.

Academic Integrity

Academic integrity is the pursuit of scholarly activity in an open, honest and responsible manner and all members of the Dartmouth community are expected to act in accordance with this principle. Academic integrity includes a commitment not to engage in or tolerate acts of falsification, misrepresentation or deception. Such acts of dishonesty violate the fundamental ethical principles of the Dartmouth community and compromise the worth of work completed by others. As such, dishonesty of any kind will not be tolerated and students found in violation of the Dartmouth Honor Code will be notified and reported to the appropriate authorities. Cheating and other forms of dishonesty (such as plagiarizing) often result when students feel too much pressure to perform and that they do not have the tools to achieve their goals. If you are falling behind or feeling overwhelmed, please come sit down and chat with me BEFORE you decide to cheat.

Disability Needs

Any student with a disability-related need for modifications or reasonable accommodations in this course must let me know as soon as early in the term as possible and contact the Student Disabilities Coordinator. I will keep this information strictly confidential, but please note that the Student Disabilities Coordinator may be consulted to verify the disability. See http://www.dartmouth.edu/~accessibility/ for more information.

TENTATIVE COURSE SCHEDULE

Class	Topic	Readings	Assignment Due Dates
W 9/16	Course Introduction		
Th 9/17	X hour: The Research Process	Healey Ch 1; Wheelan Ch. 1 (CV)	**make-up class (originally scheduled for Sat 10/24)
F 9/18	Working with and displaying data; Intro to STATA	Healey Ch 2	
M 9/21	Measures of Central Tendency	Healey Ch 3; Wheelan Ch. 2 (CV)	HW #1 Due
W 9/23	Measures of Dispersion	Healey Ch 4	
Th 9/24	X hour: STATA Basics		

F 9/25	STATA Lab		
M 9/28	Normal Curve and Z scores	Healey Ch 5	HW #2 Due; Group Project Assignment 1 Due
W 9/30	STATA Lab and Catch-up		
Th 10/1	X hour: Optional exam review		
Fri 10/2	**EXAM I**		
M 10/5	Intro to Sampling	Healey Ch 6	HW #3 Due
W 10/7	Confidence Intervals	Healey Ch 7	
F 10/9	STATA Lab		
M 10/12	Intro to Hypothesis Testing	Healey Ch 8	HW #4 Due
W 10/14	Two Sample Hypothesis Testing	Healey Ch 9	
Th 10/15	X Hour: STATA LAB		
F 10/16	NO CLASS (class held @10/15 X hour)		Group Project Assignment 2: Select Data, Schedule Meeting (via email) by 10/16
M 10/19	Analysis of Variance	Healey Ch 10	HW #5 Due
W 10/21	Chi-Square	Healey Ch 11	
Th 10/22	Optional Exam Review Session		
F 10/23	STATA Lab		
M 10/26	**EXAM 2**		HW #6 Due
W 10/28	Correlation	Healey Ch 13	
F 10/30	STATA Lab	Wheelan Ch 11(CV)	
M 11/2	Bivariate Regression	Healey Ch. 13 (continued)	HW #7 Due
W 11/4	Multivariate Analysis	Healey Ch 15 Wheelan Ch 12 (CV)	
Th 11/5	X Hour: Group Project Workshop / Review		
F 11/6	STATA Lab	Barkan (CV)	
M 11/9	Wrap up/Advanced Topics		HW # 8 Due
W 11/11	Promises and Perils of Inferential Statistics	Ziliak & McCloskey; Siegfried & Membrillas; Silver; Smith & Pell; Flam (CV)	
Th 11/12	X Hour: Group Project Workshop / Review		
F 11/13	Group Presentations		
M 11/16	Group Presentations		
Exam #3 During Final Exam Week: Tues 11/24, 8am (Room TBA) **Final Group Report Due Fri 11/20 by 3pm**			

**NOTE: We do not have the time to cover Healey Ch 12 or 14, though I recommend you read these chapters on your own.
CV=Reading on Canvas