Quantitative Analysis of Social Data
SOCY 10—Winter 2017
Course meets in Room 105, Thornton Hall

Professor: Kristin Smith
Kristin.E.Smith@Dartmouth.edu
Kristin.smith@unh.edu
Office: 105 Silsby Hall

10A, T/TH 10-11:50PM
X-period: W 3:00-3:50
Office hours: T/TH 12-1:00PM
Wednesdays, by appointment only

STATA help: Jianjun Hua Jianjun.Hua@Dartmouth.edu
Email Jianjun to schedule an appointment

Course description:
Every day we hear statements in the media citing statistics like:

Americans are more worried than ever about the economy.
Teens in families that eat dinners together have a lower risk of drug use.
Schools with smaller class sizes have better outcomes.
Youth who grow up in good neighborhoods are more likely to go to college.

Where do these statements come from? If someone else has conducted a study, how can you tell whether to trust the results? How would you test these hypotheses using quantitative data? What are the appropriate techniques to use for the analysis?

This course provides 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?”
“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. The emphasis of the class is on understanding the concepts and appropriate uses of quantitative methods rather than deriving and memorizing the formulas. A rudimentary familiarity with basic arithmetic and a few algebraic functions will suffice for material presented in the course. 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. I reserve the right to use X hours as an impromptu review session if we feel that some issues need further clarification.

Course Readings

Textbooks:
There is one required textbook which can be purchased from the Dartmouth Bookstore or Wheelock Books. It is also on reserve in Baker/Berry library.

Required Text

*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.

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:
**If you are having trouble with STATA, this is an excellent resource.


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/](http://familyinequality.wordpress.com/)
- Five Thirty Eight (Nate Silver et al.): [http://www.fivethirtyeight.com/](http://www.fivethirtyeight.com/)
- Vox (Ezra Klein et al.) [www.vox.com](http://www.vox.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:


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

In most classes, we will be using STATA in order to demonstrate statistical techniques. This will generally occur for the last 35 minutes of class every day. For this reason, you are encouraged to bring your laptop to class to use during the second half of class so that you may follow along.

A brief note on STATA: Learning a new statistical analysis program and programming language is **hard.** To be successful at learning something new, you must practice outside of class to hone your skills. There are many ways for you to get help with STATA: You may
ask another student for help with STATA, or ask Juanjun Hua, or come see me during office hours with 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 calculator 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.

Study group
The Academic Skills Center is arranging a study group for this class. This is a great resource for students who are less confident in the class. If you are interested in joining the study group contact the Skills Center.

Course Requirements

Problem Sets (20%). There will be 7 problem sets assignments this term. Assignments will be due at the beginning of class on Tuesday. Assignments must be handed in during the class period—do not email them to me. Late assignments (even those assignments handed in on Tuesday after class) will be graded according to the late assignment policy below. For all problem set assignments (with the exception of HW#1) you will need to turn in your STATA output (LOG files and DO files). 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 in their own words. Assignments will be posted on Canvas by Tuesday 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. Please do not email me with questions about the individual assignments.
Exams (20% each). There will be three exams during the course of the term. The exams are not 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.

Group Project (Paper and Presentation 20%). Students will be required to complete a group project, which includes a paper and a presentation. 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:

<table>
<thead>
<tr>
<th>Component</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group Project Assignment 1</td>
<td>5%</td>
</tr>
<tr>
<td>Group Project Assignment 2</td>
<td>5%</td>
</tr>
<tr>
<td>Final Presentation*</td>
<td>15%</td>
</tr>
<tr>
<td>Final Paper</td>
<td>75%</td>
</tr>
</tbody>
</table>

*attendance is required for all final presentation dates

I will use the following scale in assigning grades:

<table>
<thead>
<tr>
<th>Grade</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>95%-100%</td>
<td>A</td>
</tr>
<tr>
<td>92%-94.9%</td>
<td>A-</td>
</tr>
<tr>
<td>89%-91.9%</td>
<td>B+</td>
</tr>
<tr>
<td>83%-88.9%</td>
<td>B</td>
</tr>
<tr>
<td>80%-82.9%</td>
<td>B-</td>
</tr>
<tr>
<td>77%-79.9%</td>
<td>C+</td>
</tr>
<tr>
<td>73%-76.9%</td>
<td>C</td>
</tr>
<tr>
<td>70%-72.9%</td>
<td>C-</td>
</tr>
<tr>
<td>60%-69.9%</td>
<td>D</td>
</tr>
<tr>
<td>Below 60%</td>
<td>F</td>
</tr>
</tbody>
</table>

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

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

The final grade will be determined based on the following weighting scheme:

- Problem sets: 20%
- Exams: 60% (20% each)
- Group project: 20% (paper and presentation)
Course Guidelines and Policies

Course attendance:
Although class attendance and participation are not part of your final grade, students who attend class regularly tend to do better in this course. I assume that you are serious about your commitment to this class and thus presume that missed classes are due to a good reason (you are very sick, you are caring for someone who is very sick). In the event that you miss a class for any reason, you are responsible for getting class notes from another student. As a general policy, I do not make my notes available. After you review your classmate’s notes, I would be happy to meet with you to talk about the material or answer questions about anything that is unclear.

Exceptions: If your absences become chronic or you anticipate missing several classes, then it is time to consult with me. All students are required to attend class for group presentation days.

Honor principle:
Examinations and assignments will be conducted in accord with the principles of academic honor detailed in the Dartmouth Organization, Regulations and Courses. Academic misconduct will result in a failing grade and a report to your academic dean. Students are encouraged to read Dartmouth’s statement on Sources and Citations: http://www.dartmouth.edu/~writing/sources

Religious holidays:
Some students may wish to take part in religious observances that occur during the academic year. If you have a religious observance that conflicts with your participation in the course, please meet with me as soon as possible to discuss appropriate accommodations.

Students with disabilities:
Students with learning, physical, or psychiatric disabilities who will be taking this course and may need disability-related classroom accommodations are encouraged to make an appointment to see me as soon as possible, by the end of the second week of classes. All discussions will remain confidential, although the Student Disability Services office may be consulted to discuss appropriate implementation of any accommodation requested. Also, stop by the Academic Skills Center in 301 Collis Center to register for support services.

For individual peer tutorial assistance with papers, research, and new media projects, students may use RWIT, the Student Center for Research, Writing, and Information Technology. Make appointments online at www.dartmouth.edu/~rwit, or drop by the Center at Baker-Berry Library, Level One, from 4 to 6PM and 7 to 10PM, S-Th.

Late submissions and make-up exams:
The submission of assignments, exams and papers is expected on the indicated “due date” at the beginning of that day's class meeting (unless otherwise stated). On any assignment,
your grade will be reduced by a 8 percentage points for each day late (e.g., 1 day late, a grade of 100 will be reduced to a 92; 2 days late, 84; and so on). Assignments that are more than 5 days late will 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.

There will be no make-up exams except in extremely rare cases in which some unforeseen crisis or 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. If the make-up exam is missed for any reason the student will receive a zero for the exam.

Also note that computer failure is never considered a justifiable excuse for a paper extension. Computers sometimes freeze or breakdown; power outages happen. The best way to deal with these possibilities is to save your work, and save often.

Classroom etiquette:
Disruptions to class negatively affect everyone’s ability to learn. Everyone should arrive to class on time and not leave early. If you must arrive late or leave early, please notify the instructor in advance and sit by the door to minimize disruptions. It is inappropriate to use cell phones or use laptops for activities outside of taking notes (such as email, facebook, shopping, etc.). Sidebar discussions with fellow students are disruptive to the professor. Students will be offered a short break (5 minutes) during the course. Students should respect the time frame and return to class within the time allotted.
## Course Schedule

This syllabus is a working document; it may change over the course of the term. You are advised to check the Canvas Website regularly for changes.

<table>
<thead>
<tr>
<th>Class</th>
<th>Topic</th>
<th>Readings</th>
<th>Assignment Due Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Th 1/5</td>
<td>Course Introduction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tu 1/10</td>
<td>The Research Process &amp; Intro to Stata; STATA LAB</td>
<td>Healey Ch. 1; Wheelan Ch. 1</td>
<td></td>
</tr>
<tr>
<td>Th 1/12</td>
<td>Working with and displaying data; Measures of Central Tendency</td>
<td>Healey Ch. 2; Healey Ch. 3</td>
<td></td>
</tr>
<tr>
<td>Tu 1/17</td>
<td>Measures of Central Tendency (Cont.); Measures of Dispersion; STATA LAB</td>
<td>Wheelan Ch. 2 (CV); Healey Ch. 4</td>
<td>HW #1 Due</td>
</tr>
<tr>
<td>W 1/18</td>
<td>X Hour: Optional STATA BASICS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Th 1/19</td>
<td>Normal Curve and Z scores; STATA LAB</td>
<td>Healey Ch. 5</td>
<td>Group Project Assignment 1 Due</td>
</tr>
<tr>
<td>Tu 1/24</td>
<td>Intro to Sampling; Confidence Intervals; STATA LAB</td>
<td>Healey Ch. 6: Healey Ch. 7</td>
<td>HW #2 Due</td>
</tr>
<tr>
<td>W 1/25</td>
<td>X Hour: Optional Exam Review Session</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Th 1/26</td>
<td>EXAM 1 STATA LAB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tu 1/31</td>
<td>Intro to Hypothesis Testing; STATA LAB</td>
<td>Healey Ch. 8</td>
<td>HW #3 Due</td>
</tr>
<tr>
<td>Th 2/2</td>
<td>Two Sample Hypothesis Testing; STATA LAB</td>
<td>Healey Ch. 9</td>
<td>Group Project Assignment 2 due</td>
</tr>
<tr>
<td>Tu 2/7</td>
<td>Analysis of Variance; STATA LAB</td>
<td>Healey Ch. 10</td>
<td>HW #4 Due</td>
</tr>
<tr>
<td>Th 2/9</td>
<td>Chi-Square; STATA LAB</td>
<td>Healey Ch. 11</td>
<td></td>
</tr>
<tr>
<td>Tu 2/14</td>
<td>Correlation; STATA LAB</td>
<td>Healey Ch. 13; Wheelan Ch. 11 (CV)</td>
<td>HW #5 Due</td>
</tr>
</tbody>
</table>

(CV) indicates course materials.
<table>
<thead>
<tr>
<th>Date</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>W 2/15</td>
<td>X Hour: Optional Exam Review Session</td>
</tr>
<tr>
<td>Th 2/16</td>
<td>EXAM 2 STATA LAB</td>
</tr>
<tr>
<td>Tu 2/21</td>
<td>Bivariate Regression; STATA LAB</td>
</tr>
<tr>
<td>W 2/22</td>
<td>X Hour: Group Project Workshop / Review</td>
</tr>
<tr>
<td>Th 2/23</td>
<td>Multivariate Analysis; STATA LAB</td>
</tr>
<tr>
<td>Tu 2/28</td>
<td>Promises and Perils of Inferential Statistics; Wrap up &amp; Advanced Topics</td>
</tr>
<tr>
<td>W 3/1</td>
<td>X Hour: Group Project Workshop / Review</td>
</tr>
<tr>
<td>Th 3/2</td>
<td>Group Presentations</td>
</tr>
<tr>
<td>Tu 3/7</td>
<td>Group Presentations</td>
</tr>
</tbody>
</table>

**Exam #3 During Final Exam Week: Sunday 3/12, 3pm (Room TBA)**
**Final Group Project Papers Due Tuesday 3/7 by 10am**

**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