Syllabus
Autumn 2018: PPHA 31002 Statistics for Data Analysis I
Drs. Eyal Frank, Amir Jina, and Carolyn Sloane

email: eyalfrank@uchicago.edu amirjina@uchicago.edu csloane@uchicago.edu

Course Materials:  
(1) Course slides posted weekly on Canvas  
(2) Texts (recommended): David Salsburg The Lady Tasting Tea, Charles Wheelan Naked Statistics: Stripping the Dread from Data  
(3) Supplemental readings and exercises provided on Canvas

Course Objective:
This course aims to provide students with a basic understanding of statistical analysis for policy research, basic computation skills for using data, and some experience working with data. We will place a strong emphasis on “statistical thinking”.

Specifically, the course is divided into 10 topics:

Topic (1) Statistics, Models and Causation (& a preamble about programming)  
Topic (2) Random Variables  
Topic (3) Probability Theory  
Topic (4) Summary Statistics  
Topic (5) Data Problems  
Topic (6) Hypothesis Testing  
Topic (7) Simulations  
Topic (8) Resampling: the Magic of the Bootstrap  
Topic (9) Experiments  
Topic (10) Data Visualization

A Word on Course Materials:

There is no required textbook for this class. The course materials are posted on Canvas. We recommend the following texts as additional resources: David Salsburg The Lady Tasting Tea, Charles Wheelan Naked Statistics: Stripping the Dread from Data.

* At the close of the syllabus, we have included a suggested reading schedule by topic. *

Class Preparation:

You should come to each class with a fully charged laptop computer that has software (see below) loaded. If you need to borrow a laptop, you can do so from the
TechBar. There are TechBar locations at the Regenstein Library (1st floor) and in the Polsky Center:

https://its.uchicago.edu/techbar/

For weeks 1-5 of the quarter, we will ask you to complete mini-assignments in R prior to lecture. These assignments are not graded or collected but are required as preparation for lecture. You will find these mini-assignments posted in Canvas.

Software:

This course will require you to follow lectures and complete assignments using the statistical software STATA, R (and its companion software R Studio), and a spreadsheet program we will generically refer to as Excel.

Stata is available on the computers in the Harris School Computer Lab and on the student servers. Students wishing to purchase Stata may do so through the University at a substantial discount. Stata SE is $235 for a one-year license; $395 for a perpetual license.

http://www.stata.com/order/new/edu/gradplans/student-pricing/

Students should have R-studio with the “tidyverse” libraries installed on their machines. There is no charge for either R-Studio or tidyverse.

Homework Groups

You will be randomly assigned to a group. You will meet with this group to review course material and to complete the homework assignments. All assignments will be submitted as a group and each group member will receive the same grade. Your homework groups are internal labor markets to which your instructors and teaching assistants are not a party. You are responsible for extracting maximum work effort and quality from your team. Collaborating effectively with your group may be one of the most important things you learn from this course.

Course Grading

Here is a breakdown of the grading for the course

- 10% Quizzes
- 60% Homework assignments
- 30% Final Exam

There will be 2 in-class quizzes, 7 homework assignments, and a final exam.
Assignments that are not turned in on the due date will not be accepted and will be counted as a 0. There are no exceptions.

The Final Exam is on December 10 from 11am-1:50pm.

**Timing of the Assignments/Quizzes/Exams**

- Homework 1: October 9
- Homework 2: October 16
- Homework 3: October 23
- Quiz 1: Week of October 22
- Homework 4: November 6
- Quiz 2: Week of November 12
- Homework 5: November 13
- Homework 6: November 27
- Homework 7: December 4
- Final Exam: December 10 11am-1:50pm

Do not schedule travel, interviews, medical appointments, hosting of distant relatives, television appearances, etc. that conflict with the final exam date. In the event of a personal, life-threatening health incident (hospitalization with documentation), or a death in your immediate family, you may have an opportunity to make up the exam. Otherwise, **there are no make-up final exams.**

**Re-Grades**

Any item for which there is a re-grade request must be done within 7 days after we return the assignment or exam to you. The request for re-grade MUST be done in writing and attached to the assignment. All of your group members must agree to have the assignment re-graded. In such cases, the teaching assistants will re-grade the whole assignment— not just the question you identified. As a result, your grade may be lower.

**Communication:**

Communication from instructors to students will happen through posting of materials on Canvas, and occasionally through instructor participation on Piazza.

As there are many students in this sequence, emailing your instructor directly is the least effective way to have either a logistical or a pedagogical issue resolved. Therefore, we suggest and request that communication from students take the following forms:

**Questions regarding scheduling** should be directed to TBA.

**Questions regarding course material** may be posted on Piazza, a forum that is monitored by the teaching assistants and instructors.
**Recitations:**

Each week, your teaching assistants dedicate formal instruction time through the weekly recitation. You must attend the recitation section for which you are assigned.

**Office Hours:**

Each week, your teaching assistants dedicate time for additional instruction through their office hours. Please consult with your assigned TA for their office hours schedule.

Your instructors are available for office hours:
Dr. Frank: Tuesdays 4:00-5:30pm Office 187
Dr. Jina: TBA
Dr. Sloane: TBA

**Stata and R Support Bar:**

The Harris School has dedicated additional resources for teaching programming in R and Stata through the Stata and R support bars. Your instructors for Stata and R Bar are Cory Rand (crand@uchicago.edu) and Marissa Block (blockml@uchicago.edu).

**Tutoring**

If you would like to employ a tutor for additional instruction, please contact your academic advisor or the Assistant Director for Student Affairs, Jen Lombardo (lombardo@uchicago.edu).

**Ethical Academic Conduct**

The University’s Academic Policies and Procedures and guidance regarding Civil Conduct apply to all activity in our course. If you need to review the University’s policies, please see:

https://studentmanual.uchicago.edu/Policies#Honesty
https://studentmanual.uchicago.edu/university

By taking this course, you explicitly pledge your honor that you will not cheat (or help others to cheat) in any way on the assignments/exams.

**We adhere to the official Harris School protocol for ethical violations:**

“First Violation

If a student is accused by an instructor or teaching assistant of plagiarism, cheating, or any other form of academic dishonesty, the student will be summoned to meet with the Dean of Students and the instructor. In the meeting, the student and instructor both present information about the situation. If it is determined by the instructor and the Dean
of Students that the student has, in fact, plagiarized or cheated, the following sanctions will be imposed for the first violation:

The student will generally receive a grade of 0 on the assignment or exam in question (subject to the discretion of the instructor)
The student may be asked to re-do the assignment or retake the exam (without credit) to ensure that the student has learned how to properly cite sources or demonstrate that he or she has command of material covered.
A formal letter of finding is sent to the student stating that the student has been found in violation of the code of academic honesty and what the sanctions were. The letter, along with any evidence presented, is archived in Harris Student Affairs records until the student graduates if the student has no other violations.

Second Violation

If a student who has already been found in violation academic dishonesty is again accused of academic dishonesty, the case will be sent to the Harris Area Disciplinary Committee. Details about the Area Disciplinary Committee procedures can be found in the University Student Manual (https://studentmanual.uchicago.edu/area).

Information about the first violation, including the formal letter of finding any evidence, will be presented to the Area Disciplinary Committee, along with evidence of the current allegation. If the student is found in violation of academic honesty a second time, the Area Disciplinary Committee can assign sanctions including suspension or expulsion from the University.

Just to summarize what the linked content says: In the case where we need to convene an Area Disciplinary Committee to hear a case, Dean Baicker would select 3 faculty and at least one student of her choosing to sit on the committee. The faculty selected must be full-time faculty, not lecturers. A representative from central University campus and student life would also be on the committee. The Harris Dean of Students will also sit on the committee, but as a non-voting representative. The committee would hear the case and decide what sanctions, if any, to apply to the student. These sanctions can include suspension or expulsion.

Academic Dishonesty Appeals

If a student has been found in violation of academic honesty and does not believe that either the finding or the sanction is fair or correct, the student has the right to appeal the finding by requesting a hearing from the Area Disciplinary Committee. More information about the Area Disciplinary Committee is available here.”

To clarify ethical academic conduct within the boundaries of your homework assignments:
Homework assignments are team assignments. You are allowed to work on the homeworks within your assigned group. Copying the homework of another group/passing code from group to group is cheating. Providing another group with your group’s assignment to copy is cheating.

*Copyrights and Course Content (Use of Course Hero and similar websites):*

This course is a work of original authorship. All course materials (including, but not limited to, class lectures and discussions, handouts, examinations, study guides and web materials) and the intellectual content of the course itself are protected by United States Federal Copyright Law. Students are permitted to make notes solely for their own private educational use. Students and all other persons are expressly forbidden from recording lectures or discussions and from distributing or selling lectures notes and all other course materials without the prior written permission of the instructors. Because the instructors own the copyright to the classroom presentations and all course materials, any notes taken during those presentations and subsequently sold or distributed to others would constitute an unauthorized derivative work and expose the person or persons involved to individual copyright infringement actions by the instructors.

*Teaching Assistants:*

TBA

*A Suggested Reading Schedule by Topic:*

Topic (1) Statistics, Models and Causation (& a preamble about programming)
Krugman, The Accidental Theorist ([here](#))
The Lady Tasting Tea Chp 3-6
Naked Statistics Chp 1

Topic (2) Random Variables
The Lady Tasting Tea Chp 1-2

Topic (3) Probability Theory
Naked Statistics Chp 5,6

Topic (4) Summary Statistics
Naked Statistics Chp 2, 3
The Lady Tasting Tea Chp 9 & 10

Topic (5) Data Problems
Naked Statistics Chp 7

Topic (6) Hypothesis Testing
Naked Statistics Chp 4, 9
Topic (7) Simulations
The Lady Tasting Tea Chp 11-12, 17-18

Topic (8) Resampling: the Magic of the Bootstrap
The Lady Tasting Tea Chp 28
Naked Statistics Chp 13

Topic (9) Experiments
The Lady Tasting Tea Chp 27

Topic (10) Data Visualization
The Visual Display of Quantitative Information

Additional resources:

James Stock and Mark Watson Introduction to Econometrics

http://www.openintro.org/

Michael Lavine, Introduction to Statistical Thought (2013)

Edward Tufte The Visual Display of Quantitative Information