

Syllabus
Autumn 2019: PPHA 31002 Statistics for Data Analysis I

Dr. Eyal Frank
eyalfrank@uchicago.edu

Dr. Daniel Moskowitz
dmoskowitz@uchicago.edu

Dr. James Saxon
jsaxon@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) Probability Theory
- Topic (3) Random Variables
- 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*. We also include several additional materials, either PDFs, or links to websites that we think can be useful as additional review materials. They are by no means necessary to review, but they can be used as an additional source of reference for topics you think you need more practice with.

*** 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/>

Before each class, 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. To further incentivize you to exert effort on these mini-assignments, we have designed them to be helpful for your problem sets, and we will award you with a bonus to your final exam grade. If you complete more than 50% of the mini-assignments, you will receive a 2-points bonus. If you complete more than 90%, the bonus will be 5 points.

Software:

This course will require you to follow lectures and complete assignments using the statistical software R (and its companion software RStudio) and Stata. The course will start with using R, so make sure it is fully working on your computer before the quarter begins.

Stata is available on the computers in the Harris School Computer Lab and on the student servers. While it is not necessary to purchase, students wishing to buy 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 RStudio with the `tidyverse` libraries installed on their machines. There is no charge for either RStudio 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, of which there will be a total of eight. In the sixth week, we will reshuffle groups. This means that you will work on four homework assignments with each group. The first three of those assignments will be submitted as a group and each group member will receive the same grade. The last of those four assignments will be submitted individually. Although you are allowed to work on the individual assignments with your group, **each group member must still submit their own write-up.** 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, 8 homework assignments, and a final exam.

The Final Exam is on Monday, December 9th from 8:00am to 10:50am. It is your responsibility to verify that you have no scheduling conflicts that will prevent you from taking the exam. Ensure you are not booking flights or signing up for programs that will create such conflicts. Should a conflict occur, it needs to be resolved with Academic Student Affairs. Do not approach one of the instructors asking for guidance on how to resolve such a conflict.

Do not schedule travel, interviews, medical appointments, hosting of relatives, 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.**

In general, we will try to make the homework assignments available to you 10 days before they are due (see a full list of due dates below). This should allow you to have two weekends to work on each assignment. Notice that this means that there will be some overlap between the time periods in which assignments are available to work on. Use your time effectively. Also, make sure to use the office hours (listed below) to ask questions about the material. **Late submissions within 24 hours of the due date/time will be penalized by a 10 point reduction, and late assignments submitted 24-48 hours after the due date/time will be penalized by a 20 point reduction. Assignments submitted more than 48 hours after the due date/time will receive a score of 0.** The only exceptions to this policy will be due to medical reasons (with documentation). **Homework submission is due by 11:59 pm on the date indicated below:**

Timing of the Assignments/Quizzes/Exams

- Homework 1: Due on October 9 (group submission)
- Homework 2: Due on October 14 (group submission)
- Homework 3: Due on October 21 (group submission)
- Quiz 1: Week of October 23
- Homework 4: Due on October 28 (individual submission)
- Homework Groups Reshuffle --
- Homework 5: Due on November 4 (group submission)
- Homework 6: Due on November 13 (group submission)
- Quiz 2: Week of November 18
- Homework 7: Due on November 25 (group submission)
- Homework 8: Due on December 2 (individual submission)
- Final Exam: December 9 8:00am-10:50am

Re-Grades

A re-grade request must be submitted within 7 days after the assignment or exam is returned to you. To request a re-grade, you and your group members (if applicable) must print, complete, sign, and email a scanned re-grade request form (available on Canvas) to your instructor (the instructor of the person submitting the request) and cc the head TA. All members of a group must agree to have the assignment re-graded (in the case of assignments that are submitted as a group). In all cases, **the entire assignment will be re-graded**, not just the question or specific part that pertains to your grievance. As a result, the re-grade can (and often does) result in a lower overall grade on the assignment or exam.

Communication:

Communication from instructors to students will happen through posting of materials on Canvas, including postings to Announcements and Discussions.

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 the head TA: Maria Martinez, mariaadelaidamc@uchicago.edu (and cc Mythili Vinnakota mvinnakota@uchicago.edu).

Questions regarding course material may be posted on Canvas Discussions, 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. TA office hours as follows:

Keller 2050

Tuesday 9:00am to 10:00am,
Wednesday 3:30pm to 4:30pm
Thursday 9:00am to 10:00am

Keller 2058

Friday 11:00am to 12:00pm

Your instructors are available for office hours at the following times and locations:

Dr. Frank: Wednesdays, 11:00 am – 12:30 pm, Keller 2057.

Dr. Moskowitz: Fridays, 10:00 am – 11:30 am, Keller 3103.

Dr. Saxon: Tuesdays, 1:00 pm – 2:30 pm, [Searle](#) 230

(take the elevator to the second floor to bypass keycard)

Coding Workshop

Harris has dedicated two instructors – Ari Anisfeld (anisfeld@uchicago.edu) and Angela Li (ali6@uchicago.edu) – to provide instruction in coding with R, literally every weekday, in the first five weeks of the quarter. The sessions are M-Th 5-6:20, and Friday at both 9:30-10:50am and 11:00am-12:20pm. Please use this resource! We will expect that you know the material that they cover – both in class and for homework assignments – throughout the quarter.

<https://harris-coding-lab.github.io>

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 instructor for R is Rei Bertoldi (rbertoldi@uchicago.edu). Your instructor for Stata is Fernando Regalado (fernandoregalado@uchicago.edu). These resources are in addition to, and independent of, the coding workshop sessions. We highly advise that you attend the coding sessions as they will prove useful for many other classes at Harris.

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

All University of Chicago students are expected to uphold the highest standards of academic integrity and honesty. Among other things, this means that students shall not represent another's work as their own, use impermissible materials during exams, or otherwise gain an unfair academic advantage. All students suspected of academic dishonesty will be reported to the Harris Dean of Students for investigation and adjudication. The disciplinary process can result in sanctions up to and including suspension or expulsion from the University. In addition to disciplinary sanctions, the student will receive a grade of 0 on the exam or assignment in question and cannot earn higher than a B- in the course, regardless of their performance on other assignments and exams. At the instructors' discretion, the student may receive a failing grade for the course regardless of their performance on other elements of the course.

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. Please note that grading decisions are fully at the discretion of the instructor, who may decide to impose harsher grade penalties.
- 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.
- Students found in violation of the academic honesty policy are not permitted to withdraw from the course to avoid grade penalties from the instructor.
- In cases where plagiarism or academic dishonesty is egregious, the case may be referred to the Area Disciplinary Committee even on a first offense. The Dean makes all decisions about which cases will go before the Area Disciplinary Committee.

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](#). 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 transcript notes, disciplinary probation, suspension or expulsion from the University.

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](#).

Full Harris Academic Integrity Procedures from online student Handbook accessible at <https://harris.uchicago.edu/gateways/current-students/policies>

University of Chicago Policy on Academic Honesty & Plagiarism accessible at <https://studentmanual.uchicago.edu/academic-policies/academic-honesty-plagiarism/>

To clarify ethical academic conduct within the boundaries of your homework assignments:

1. Homework assignments are team assignments. You are allowed to work on the assignments within your assigned group. Copying the homework of another group and passing code from group to group is cheating. Providing another group with your group's assignment to copy is cheating. Here are some guidelines to clarify what is considered collaboration but not cheating:
2. All work must be your own. Do **NOT**:
 - a) show other students your code (outside of your group)
 - b) ask for another student's code

- c) use online solutions to textbook questions
- d) copy large portions of code from online repositories (e.g. replication code)
- 3. Every submission needs to include information on who you collaborated with (their names).
- 4. So how can I collaborate?
 - a) In-person collaboration
 - i. clarify ambiguities in problem set questions
 - ii. discuss conceptual aspects of problem sets (e.g. at the whiteboard)
 - iii. show output on screen (e.g. a graph or table)
 - iv. show helpful documentation files
 - b) Electronic collaboration
 - i. Canvas Discussions message board
 - 1. ask questions
 - 2. share error messages (but not code)
 - ii. Code from an online forum or resource (other than documentation files)
 - 1. cite all code you use, even a one-line snippet
- 5. How do these rules change for problem sets working in groups?
 - a) You and your group members will submit a single problem set.
 - b) If you work collaboratively with other students, but turn in your own problem set
 - c) You can talk to your group members as needed and look at other members work to facilitate that discussion
 - d) Your problem set should be solely your authorship (written up by yourself, in your own language, including your own code.)
 - e) Your code should have a comment at the top listing the members of your group.
 - f) Any part of your code that was substantially altered because of your group discussion should cite others' contributions with names and descriptions in a comment at the place where it is applicable.
- 6. Unsure about some aspect of this policy? Please ask your instructor.

Please [see example](#).

(source: This policy draws heavily on the CS 12100 academic honesty [policy](#) and CMSCC 23300 [policy](#))

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:

| | |
|----------------------|--|
| Angel Aliseda Alonso | aaliseda@uchicago.edu |
| Julian Athaide | jda1@uchicago.edu |
| Margaret Yu | margaretjyu@uchicago.edu |
| Marissa Slowey | mslowey@uchicago.edu |
| Mengyuan (Jenny) Guo | mengyuan@uchicago.edu |
| Mythili Vinnakota | mvinnakota@uchicago.edu |
| Nur Aisyah Safitri | safitri@uchicago.edu |
| Pietong Jin | pjing@uchicago.edu |
| Saurabh Kalra | kalras@uchicago.edu |
| Shambhavi Sawney | ssawhney@uchicago.edu |
| Shuai Yuan | shuaiyuan@uchicago.edu |

Head TA:

| | |
|----------------|--|
| Maria Martinez | mariaadelaidamc@uchicago.edu |
|----------------|--|

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) Probability Theory
Naked Statistics Chp 5,6

Topic (3) Random Variables
The Lady Tasting Tea Chp 1-2

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*

David Diez, Christopher Barr, and Mine Cetinkaya-Rundel (2012) Open Intro Statistics.

<http://www.openintro.org/>

Michael Lavine, *Introduction to Statistical Thought* (2013)

Edward Tufte *The Visual Display of Quantitative Information*