

**Statistics 4750**  
**Sample Survey Methods**  
**Syllabus**  
**Spring 2023**

**Professor:** Dr. Elaine Zanutto

**E-mail:** [zanutto@wharton.upenn.edu](mailto:zanutto@wharton.upenn.edu)  
[ezanutto@gmail.com](mailto:ezanutto@gmail.com)

**Office Hours:** Wednesdays at 8pm via Zoom

**Teaching Assistant:** Aarushi Parikh [aaruship@sas.upenn.edu](mailto:aaruship@sas.upenn.edu)

**TA Office hours:** TBA

**Class Website:** Canvas website

**Text:** *Sampling: Design and Analysis, Third Edition*\*\*\*, Sharon H. Lohr, 2021, Brooks/Cole Cengage Learning. \*\*\*note new edition

**Course Description:** This course will focus on statistical methods for the design and analysis of sample surveys. We plan to cover Chapter 1-5 and parts of Chapters 6-9, and 11 of the textbook, and other related topics as time permits. The core topics are

- simple random sampling
- stratified sampling
- cluster sampling
- ratio estimation
- sampling with unequal probabilities
- sampling weights
- analyzing complex surveys
- dealing with nonresponse
- some advice for questionnaire design
- pros and cons of online panels
- linear and logistic regression with survey data
- graphing survey data with R
- analysis of categorical data from a complex survey
- small area estimation

**Course Requirements:** There will be five homework assignments, a midterm exam, a take-home midterm assignment, and a final project.

**Homework Policy:** Homeworks must be turned in individually. You may speak to each other about the homework. You may ask me and our TA about the homework. But you must turn in your own work. The midterm will be an in-class test. The final project may be completed individually or in groups of up to 2 people (3 people with special permission).

**Late Policy for Homework:** Extensions to homework deadlines are often granted. Homeworks are due on Wednesday nights. Generally anyone who asks can have an extension until Saturday midnight. Send me an email to ask and cc our TA.

**Missed Midterm Policy:** Missing the midterm is not easily accommodated. Accommodation will be made without penalty ONLY in the event of an excused absence including serious illness, court appearance, or personal crisis. Specific accommodations are granted on a case-by-case basis and may include a make-up midterm or shifting grade weight to the final exam or other assignment.

**Grading:** Homework (20%, all 5 are weighted equally), Midterm (40% -- in-class portion 25%, take-home portion 15%), Final Project (40%).

**Attendance:** Attendance at lecture is not mandatory. I will record all lectures and post them on the canvas website. However, it is your responsibility to watch the lectures. Do not come to office hours unprepared.

**Computing software:** We will make use of the statistical computing software R via the RStudio interface in class and on homework assignments. R can be downloaded from <http://www.r-project.org/> and RStudio can be downloaded from <https://rstudio.com/>. It would be beneficial if you already have some experience with R, but not absolutely necessary. I will give you instructions for all the R commands you will need for the course, but if you have never used R before expect a learning curve to get up to speed.

**Course Prerequisites:** Stat 102, Stat 112, Stat 431 or comparable courses. Experience with R is very helpful but not required.

**Academic Integrity:** Students are expected to be familiar with and comply with Penn's Code of Academic Integrity, which is available in the Pennbook, or online at <https://catalog.upenn.edu/pennbook/code-of-academic-integrity/>. I generally have a zero-tolerance policy for cheating, and all violations will result in substantial penalties. The final project will be submitted via the Turnitin feature on Canvas. If you use ideas from other people you must provide proper citations. If you have any doubts or questions about what constitutes academic misconduct, please do not hesitate to contact me.