UNIVERSITY OF PENNSYLVANIA
SCHOOL OF SOCIAL POLICY & PRACTICE
Master of Science in Social Policy

MSSP 630-001 Quantitative Reasoning

Fall 2021
This version: August 24 2021

Twitter hashtag: #mssp630
Class-related tweets: https://twitter.com/search?q=%23mssp630&src=typd

Ioana Marinescu
Email: ioma@upenn.edu
Office hours: use Calendly (https://calendly.com/ioma/mssp-630-meeting), or email me if you cannot find a time that works on Calendly. You will be prompted to schedule at least one meeting with me to discuss how the course applies to one of your areas of interest. More broadly, I am here to answer questions that I am in a unique position to address, including career advice and more in depth questions about course material and related topics. Come prepared: ask your question on Piazza first if it's something the TAs or other students could answer. If it's about grading and related issues, reach out to your TA first.

Synchronous time: Tuesdays 10:15am – 11:45am

Labs:
TBD

Teaching Assistants

101	102	103
Office Hours: TBD	Office Hours: TBD	Office Hour: TBD

Course Description
What do these numbers mean? Do they confirm my theory? And what is a theory anyway? In this class, we will be exploring simple theories and how to test them using data. We will also look into how data can give us clues to formulate our theories. We will discuss how to plot data to understand its contents and potential problems. Once we understand what is in the data, we can start testing some simple theories. For example, can we say that more educated people earn more than less educated people? And how confident can we be about this statement? Even if more educated people do earn more than less educated people, does this mean that increasing education will be causing people to earn more? Or is it simply that more educated people are smarter to begin with? We will see how data can allow us to solve this kind of question and advise policy makers on the benefits of increased education.

Educational Objectives
Upon completion of this course, students will be able to demonstrate:
1.   a strong foundation and competence in introductory statistics and probability
2.   their knowledge of descriptive statistics such as measures of central tendency and variability

1

3. an understanding of inferential statistics and when and how to use them including significance testing
4. knowledge of the distinction between correlational and causal inference
5. their competence with R
6. their ability to employ descriptive and inferential statistics in a policy report

Course Requirements and Expectations

Improve your writing

Writing is an essential professional skill for policy analysts and others working in areas related to social policy. Even the best writers always have room for improvement. In some cases, instructors may suggest that students seek out help with their writing; in other cases, students might decide on their own, that professional help is in order.

The Marks Family Writing Center
(http://writing.upenn.edu/critical/graduate_students/help_with_your_writing.php)
has a variety of services for those who would like help with their writing. The School also offers writing workshops.

Grading policy

Assignments have to be handed in on time. Each day an assignment is late without a substantial excuse, a full letter grade is deducted from the paper. Starting from 30 minutes after the official deadline, your paper is considered to be a day late, and two days late 24 hours after the official deadline, etc.

Class participation

You are expected to come to every class on time and fully prepared to discuss the material assigned for the week. That is, you should have completed all assigned items in the Canvas module for the week.

Academic Integrity

Students are expected to adhere to the University's Code of Academic Integrity, available at https://catalog.upenn.edu/pennbook/code-of-academic-integrity/. Care should be taken to avoid academic integrity violations, including plagiarism, fabrication of information, and multiple submissions (see descriptions below).** Students who engage in any of these actions will be referred to the Office of Student Conduct, which investigates and decides on sanctions in cases of academic dishonesty.

1. Plagiarism: using the ideas, data, or language of another person or source without specific or proper acknowledgment. Example: copying, in part or in its entirety, another person's paper, article, or web-based material and submitting it for an assignment; using someone else's ideas without attribution; not using quotation marks where appropriate; etc.

2. Fabrication: submitting contrived or altered information in any academic exercise. Example: making up data or statistics, citing nonexistent articles, contriving sources, etc.

3. Multiple submissions: submitting, without prior permission, any work submitted to fulfill another academic requirement.

**It is students' responsibility to consult the instructor if they are unsure about whether something constitutes a violation of the Code of Academic Integrity.

Required Texts

This class will use the following book as a textbook (available at the Penn bookstore):

Moore, David, George McCabe & Bruce Craig. Introduction to the Practice of Statistics, ninth edition. Macmillan, 2017.

Go to http://www.macmillanlearning.com/catalog/studentresources/ips9e# and download the datasets for R.

All other texts will be found either on the Internet (links below) or on Canvas https://canvas.upenn.edu. To access Canvas, you will need your PennKey and password.

This class will use R as its main software. Recommended R resources:
- Complete how to for beginners in R, beautifully illustrated: https://moderndive.com/preface.html (Links to an external site.)

- Once you start coding in R, you will necessarily make coding mistakes. This is perfectly normal. This page helps you remember basic coding rules and spot common mistakes: https://learn.r-journalism.com/en/how_to_use_r/syntax/r-syntax/ (Links to an external site.)

- An online textbook that largely overlaps with what you will do in labs, so it can be very useful as a complement to the class: https://cran.r-project.org/doc/contrib/Verzani-SimpleR.pdf (Links to an external site.).

Additional R resources that could be useful:

- An interactive tutorial focused on making nice charts in R with ggplot (and on importing data): https://hhsievertsen.shinyapps.io/r_introduction/#section-welcome (Links to an external site.)

- Basic R for data journalism: https://learn.r-journalism.com/en/ (Links to an external site.)

- Basic R with examples from criminology: http://crimebythenumbers.com/index.html#what-you-will-learn

- R for data science (requires a bit more programming background): https://r4ds.had.co.nz/

Assignments (see Canvas for full details)

Assignment 1 (10% of final grade): weekly lab report (pass / fail)
Using a dataset to be specified in your lab session, produce a one-page lab report in Word and submit through Canvas with the following structure:

- One table or graph, based on code you learned in that week's lab
- Code that produces this table or graph in R
- Explain what this graph or table shows and why this graph or table seemed more important than others you could have created through this week's lab

The assignment is graded pass / fail, and each pass is worth one point towards the 10 points allotted to this exercise. The lab report is due BEFORE the next lab.

Assignment 2: (15% of final grade). Describe education, race and weekly earnings from the Current Population Survey (CPS) in Pennsylvania & New Jersey. Due September 19, in Word-compatible format on Canvas by 11pm.
Using what you have learned in chapter one of the textbook, create figures to describe education, race and weekly earnings in Pennsylvania & New Jersey. Pay attention to the fact that some of the variables are categorical, and others are quantitative. For each variable, compare the results for Pennsylvania & New Jersey: to what extent are they similar?

Assignment 3 (15% of final grade). The Current Population Survey and sample selection, due October 10, in Word-compatible format on Canvas by 11pm.

Assignment 4 (20% of final grade). Group assignment. Analyze the relationship between education, earnings and race in the CPS in Pennsylvania due November 21, in Word-compatible format on Canvas by 11pm.

Assignment 5 (15% of final grade): exam on Canvas on December 2.

Assignment 6 (20% of final grade). Final at home policy report, difference-in-differences analysis for the minimum wage increase in New Jersey in 2014, using Pennsylvania as a control group, due December 12, in Word-compatible format on Canvas by 11pm.

Participation grade (5% of final grade): up to 5 points for answering well a question asked on Piazza, posting something relevant on Twitter (a post similar to posts students and instructors posted in prior years), sharing your screen during lab to ask a question about your code, and other similar actions. If you don't want to use your personal Twitter account, you can create a new account for professional uses. During lab time, TAs will ask if anybody has posted a Tweet, and if you did you can say so and briefly explain what it was about.

For further details about the assignments, see the guidelines on Canvas under assignments.

Class Schedule and Required Readings

Lecture slides will be posted on Canvas AFTER the respective lecture has taken place.

Week 1: August 31. Introduction & describing distributions with graphs
Required reading:
Moore, McCabe & Craig, chapter 1, section 1.1 and 1.2

Week 2: September 7. Describing distributions with numbers
Required reading:
Moore, McCabe & Craig, chapter 1, section 1.3 and 1.4

Week 3: September 14. Relationships between two variables
Required reading:
Moore, McCabe & Craig, chapter 2, sections 2.1-2.4

Week 4: September 21. Relationships between two variables and causality
Required reading:
Moore, McCabe & Craig, chapter 2, sections 2.5-2.7
Helpful video: https://www.youtube.com/watch?v=iPBV3BlV7jk

Week 5: September 28. Producing data
Required reading:
Moore, McCabe & Craig, chapter 3

Week 6: October 5. Using existing data
Required reading: TBD

Week 7: October 12. Probability and sampling distributions
Required reading:
Moore, McCabe & Craig, chapter 4 & 5

Week 8: October 19. Confidence intervals and significance testing
Required reading:
Moore, McCabe & Craig, chapter 6, sections 6.1 and 6.2

Week 9: October 26. Issues with significance testing
Required reading:
Moore, McCabe & Craig, chapter 6, sections 6.3 and 6.4

Week 10: November 2. Inference for means
Required reading:
Moore, McCabe & Craig, chapter 7

Week 11: November 9. Inference for proportions & for categorical data
Required reading:
Moore, McCabe & Craig, chapter 8
Moore, McCabe & Craig, chapter 9

Week 12: November 16. Inference for regression
Required reading:
Moore, McCabe & Craig, chapter 10

Week 13 November 23: NO CLASS. A Thursday class schedule is followed in anticipation of Thanksgiving

Week 14 November 30: Estimating a causal effect without an experiment: difference-in-differences
Optional reading:
Angrist, Joshua D. and Jörn-Steffen Pischke, Mostly Harmless Econometrics: An Empiricist's Companion, Princeton University Press, 2009. Chapter 5, section 5.2.

Week 15: December 7. Final exam