HDSP/LRN_SCI 410: Quantitative Methods I
Probability and Statistics
Fall 2021

[Note: This is a public version of the syllabus for HDSP/LS 410. As such, many links (underlined text) have been deactivated and some information removed]

Instructor: Cora Wigger
Office Hours: Calendly Link (all appointments will be virtual by default)
E-mail: corawigger2021@u.northwestern.edu
E-mail Policy: I will respond to your email within 24 hours M-F. I do not respond to email on weekends.

Teaching Assistant: Matias Martinez
Section: Wednesdays, 2-3 p.m.
Office Hours: Calendly Link
E-Mail: [Removed from Public Posting]

Course objectives: Welcome to Probability and Statistics! This course is an introduction to the quantitative methods sequence. Whether you plan to do quantitative, qualitative or mixed-methods research, it is important for you to understand how data can inform important policy-relevant questions. The course content will help you to be a well-informed producer and consumer of quantitative research, and will be important preparation for the other courses in the sequence. Some of the topics we will cover are descriptive statistics, probability, sampling and sample size estimation, hypothesis testing, techniques for analyzing categorical data and ANOVA.

Learning Statistics, again and again: If you have never taken a statistics course, do not be frustrated if you don't "get" everything right away. You are not supposed to, nor do you need to to succeed in this course. Statistics is a topic where you will learn a bit more about every time you review and use it. With that in mind, if you have extensive experience with statistics, you should still approach this course with a beginner's mind. Every time you go through this material or do applied statistical work, you will learn something new, see something in a new way, or develop greater intuition for the material.

Required readings: There is **no required text** for you to purchase for this course. I will occasionally assign readings and videos to review prior to class. These will be available on Canvas. That said, you may find it useful to have reference materials. One option is the text you will use next quarter, *Introduction to Econometrics* by Stock and Watson (ISBN-10: 0133595420). The first few chapters of the book are a *very* terse review of probability and statistics – not easy reading, but a good resource.

Other resources: There are many resources available for free online that explain the concepts we will focus on in this course in different ways than I do. I will share some of these additional resources in this document, and I encourage all of you to add to the document as you find materials that may be useful to your classmates.

Computer software: You will use Stata (and perhaps occasionally Excel) to analyze data for homework assignments. We will discuss how to access and navigate Stata on the first day of class. If you would like to use a different software (e.g. R or Python), please discuss with me first.

Assignments & Grading: Your grade will be based on two take home exams (50%), completion and reflection on problem sets (20%), a descriptive statistics presentation (20%), and class participation (10%).

I strongly encourage you to form study groups: you are not competing for a limited number of high grades! Rather, you are fostering an intellectual community with your classmates that will continue through your time at Northwestern and beyond.

Problem Sets – Completion & Self-Evaluation (20%) [Due ~Weekly]
I will assign (roughly) weekly problem sets with practice problems similar to those we cover in class. The purpose of these problem sets is twofold: 1) For you to gain more practice implementing the material on your own, and 2) To provide feedback to myself and the TA about what is going well and what is not. **Problem sets will <u>not</u> be graded based on how many questions you get right.** Instead, you will be graded on completing the problem sets *and* providing feedback about how they are going.

Problem sets will be released Thursdays after class, along with the answer key and a self-evaluation form. You are responsible for completing the following steps prior to the following Tuesday at 11:59 pm:

1. Complete as much of the problem set as you can (preferably with a study group) *without* looking at the answer key, using whatever resources you would like.
2. Compare your completed (or as completed as possible) problem set to the answer key and note which questions you got right and wrong.
3. Re-do any questions you got wrong or could not complete, using the answer key as a guide.
4. Complete the self-evaluation form and submit it along with evidence of your completed problem set (word document, annotated do-file, or picture of written work are all acceptable). You will receive **full credit** so long as you submit both the self-evaluation and a completed problem set.

Descriptive Data Presentation (20%) [Submit Topic & Dataset Name by 10/14; Presentations Tentatively scheduled for 11/4]

You will be responsible for downloading a publicly available dataset, loading it into analysis software and cleaning it, creating 2-4 descriptive tables and figures, and presenting them to the class. <u>This document</u> contains a SESP-crowdsourced list of publicly available datasets and data repositories. The document is editable so please continue adding to the list if you would like.

Your presentation will cover the following:

1. **Motivation:** What are your high-level, guiding questions? What are you interested in that encouraged you to look at this dataset?
2. **Research Question(s):** What is your specific descriptive research question(s)? OR, What is a bigger research question you are interested in, and what descriptive statistics will help set the stage for answering it?
3. **Data Description:** Complete a data biography (template sourced from We All Count) and then briefly describe the data you're using.
4. **Results:** Present and describe 2-4 summary tables and/or figures.
5. **Reflections and Challenges:** Briefly summarize some reflections you had from doing this assignment. You should touch on at least one of the following:
    1. Challenges you ran into and lessons learned from gathering and cleaning the data.
    2. How this data and your approach to analyzing it intersect with principles of QuantCrit, such as: certainty of racism, non-neutrality of numbers, categories are never natural, voice and insight are vital, and using numbers for social justice (Garcìa, Lòpez, and Vèlez (2018); Gillborn, Warmington, & Demack (2018); Suzuki, Morris, & Johnson (2021)).
    3. Reflections on your positionality to this data and question, and how it shapes the way you approach, analyze, and interpret the data (See Secules et al. (2021)).
    4. New research questions this assignment and process raised.

After you present, submit your data biography and presentation slides (which will include the tables/figures) via Canvas. Optionally, you may also submit your code for cleaning the data to be shared with your classmates.

Midterm Exam (20%) [Assigned 10/14, Due 10/21]
Complete at home, individually, but use whatever resources you like. Expect to spend about 2-4 hours on the exam.

Final Exam (30%) [Assigned 12/2, Due 12/10]
Complete at home, individually, but use whatever resources you like. Expect to spend about 4-6 hours on the exam.

Class Participation (10%)

Extra Credit options may be made available at the instructor's discretion. Any extra credit opportunities will be provided to the entire class.

Grading Scale

A 94%-100%	B+ 87%-89%	C+ 77%-79%	D 66%-69%
A- 90%-93%	B 83%-86%	C 73%-76%	F Under 65%
	B- 80%-82%	C- 70%-72%	

Policy on Late Work: Course material accumulates quickly and each week builds on the last. Therefore, it is important to stay on track. Making up a lecture or material late will take up more of your time than it will to complete it on time. That said, we are in a time filled with extenuating circumstance, and we will do our best to help students get caught up when such circumstances occur.

Please reach out as soon as possible if you need to request additional time for an assignment due to personal circumstances. You do not need to provide documentation to receive an extension on assignments but do need to reach out in advance to make appropriate arrangements. If you request an extension, you must specify your plan for completing the assignment(s), including your preferred updated deadline. If you fail to communicate ahead of the assignment due date, late assignments will be docked the equivalent of 1/3 letter grade for each day late (e.g. An "A" assignment turned in 1 day late will receive an "A-" grade), or 3 percentage points. Exceptions will be made for serious emergencies, but I ask that you communicate early, often, and honestly whenever possible.

Policies Around Covid-19: Circumstances around the Covid-19 pandemic are changing rapidly. Before our first class meeting, please complete this anonymous survey about your comfort level and preferences. We will discuss this the first week of class and decide together on classroom policies to aid learning during the pandemic.

Policy on Children in the Classroom

When Remote: All exclusively breastfeeding babies are welcome in the class as often as is necessary. For older children and babies, I understand that unforeseen disruptions in childcare often place caregivers in the position of having to miss class. You are welcome to leave your camera on or intermittently turn it off when your child is in the room or when you have to attend to them. Just let us know in the chat that you're still there and listening.

When In Person: Due to the ongoing challenges of the pandemic and the lack of availability of vaccines for children under age 12, I ask that students not bring children to the classroom when we are meeting in person. I will, however, provide a remote option for cases when students must stay home with children, though this is not meant to be a permanent substitution for in-person classes.

I ask that all students work with us to create a welcoming environment that is respectful of all forms of diversity, including diversity in parenting status.

I maintain the same standards and expectations for all students. However, please contact us if you are having difficulty with school-parenting balance.

Health and Wellness: Your physical and mental health are more important than anything we will do in this class. This course (and your Ph.D.) should challenge you – and I will work hard to push you to grow as a scholar – but it should not cost you your wellbeing.

If you need extra help or more time with this course or your studies more broadly, please ask. You are welcome to talk with myself and the TA about ongoing challenges, if you are comfortable, but you never owe us any personal information.

"Working until exhaustion is NOT a badge of honor; it shows that you are out of balance." – Nicole Gonzalez Van Cleve

Accommodations for students with disabilities: Any student requesting accommodations related to a disability or other condition is required to register with AccessibleNU (accessiblenu@northwestern.edu; 847-467-5530) and provide professors with an accommodation notification from AccessibleNU, preferably within the first two weeks of class. All information will remain confidential.

If you have a reason for requesting additional accommodations that falls outside of AccessibleNU's purview, please email me directly with a description of the requested accommodation (or of the barrier, if you are not sure of the appropriate accommodation) within the first two weeks of class, along with whatever information about your situation you would like to provide. I will do my best to quickly approve any reasonable accommodation requests, or I will work with you to find an appropriate alternative. If an issue arises after the 2nd week of class that requires additional attention, please reach out to me as soon as possible.

Academic Integrity: Students in this course are required to comply with the policies found in the booklet, "Academic Integrity at Northwestern University: A Basic guide". All papers submitted for credit in this course must be submitted electronically unless otherwise instructed by the professor. Your written work may be tested for plagiarized content. For details regarding academic integrity at Northwestern or to download the guide, visit: http://www.northwestern.edu/provost/policies/academic-integrity/index.html

Course Calendar: The exact topics and pacing for this course will depend on the math and statistics background of the class as a whole. This schedule is my best guess for what we will do, but it is tentative and subject to change.

Class	Topic
1 (9/23)	Describing data (center, spread, distribution) and the normal distribution
2 (9/30)	Probability, conditional probability, correlations, scatterplots
3 (10/7)	Distributions and sampling distributions; Bayes' Rule
4 (10/14)	Sampling, bias, Central Limit Theorem
5 (10/21)	Inference, hypothesis testing
6 (10/28)	Hypothesis testing
7 (11/4)	**Descriptive Data Presentations**
8 (11/11)	ANOVA
9 (11/18)	ANOVA; chi squared tests; simple regression
10 (12/2)	Multiple regression