



Syllabus

Introduction to Quantitative Social Science

Fall 2022 | TR 9:00 - 10:15 | Stone 345

Instructor: Shawn Bauldry (sbauldry@purdue.edu)

Office Hours: by appointment

Course Description

Quantitative methods are an integral component of contemporary social science. In this course we will learn about different forms of quantitative research questions, research designs, and data structures; how to manage and manipulate various forms of quantitative data; tools to describe variables and the relationships between variables; and how to quantify uncertainty in our estimates.

Facility with statistical software and programming is essential in conducting quantitative analyses. Writing readable, documented, and well-organized code to read data, prepare variables for analyses, and conduct analyses has many benefits. It is time-efficient, less prone to errors or mistakes, and provides a record that can help ensure analyses are reproducible and transparent. Much like writing papers, writing good code involves outlines, drafts, and revisions before settling on a final product that accomplishes an intended goal. In this course, learning how to program statistical software will be integrated into the textbook, the material we discuss in class, and the assignments.

Learning Outcomes

By the end of the semester, you will be able to:

- Understand and work with a variety of forms of quantitative data
- Employ statistical techniques to describe variables and their relationships to one another
- Write code to prepare data and conduct basic analyses

Course Details

Texts

The primary textbook for this course is Imai and Bougher (2021) *Quantitative Social Science: An Introduction in Stata* (Amazon link).

Statistical Software

Quantitative social science can be conducted using a variety of statistical and computer programming software packages. Most Sociology courses rely on Stata, so that will be the focus of this course. Stata is available for free to Purdue students through ITaP (link). I will also provide support for students interested in learning R (available for free download) or other statistical software packages.

Grades

Your grade for the course will be based on two components: (1) five homework assignments and (2) a replication analysis.

Assignments (30 points per assignment). During the semester there will be a series of assignments with a few exercises related to the material we cover in the textbook. The data for the exercises will be made available on Brightspace. Your responses to the exercises should be written in Word documents (or some other word processing software). The code you write to complete the exercises should be included at the end of your responses. The assignments should be submitted on Brightspace. You are welcome to work collaboratively on the assignments, but I expect that everyone will submit their own assignment. Due dates: indicated in the course schedule below.

Replication Analysis (50 points). By the end of the semester you will have the tools to carry out at least a partial replication of a published study that uses publicly available data. I will provide an assignment sheet with details about replicating an analysis of stigma surrounding mental health using data from the General Social Survey. Due date: Dec 16.

Accommodations

If you are a student with a disability or an ongoing medical condition, we can discuss appropriate accommodations. For additional information and resources, please visit the Disability Resource Center.

Course Schedule

Introduction

- **Reading** – Chp 1
- **Assignment 1** – due Sep 5 (Mon)
- **Aug 23 (Tue)** – Introduction to the course
- **Aug 25 (Thr)** – no class
- **Aug 30 (Tue)** – Introduction to data and programming
- **Sep 1 (Thr)** – Univariate descriptive statistics

Causality

- **Reading** – Chp 2
- **Assignment 2** – due Sep 19 (Mon)
- **Sep 6 (Tue)** – Causal research questions and counterfactual framework
- **Sep 8 (Thr)** – Programming: conditional operators and loops
- **Sep 13 (Tue)** – Experimental research designs
- **Sep 15 (Thr)** – Observational research designs

Measurement

- **Reading** – Chp 3
- **Assignment 3** – due Oct 12 (Wed)
- **Sep 20 (Tue)** – Concepts and measures
- **Sep 22 (Thr)** – no class
- **Sep 27 (Tue)** – Missing data and visualizing univariate distributions
- **Sep 29 (Thr)** – Bivariate descriptive statistics
- **Oct 4 (Tue)** – Illustrating bivariate relationships
- **Oct 6 (Thr)** – Putting it all together: descriptives analysis

Prediction

- **Reading** – Chp 4
- **Assignment 4** – due Oct 31 (Mon)

- **Oct 11 (Tue) – no class (October Break)**
- **Oct 13 (Thr) – Prediction, prediction error, and bivariate linear regression**
- **Oct 18 (Tue) – Bivariate linear regression estimation and indicator variables**
- **Oct 20 (Thr) – Multiple linear regression**
- **Oct 25 (Tue) – Nonlinear relationships and interactions**
- **Oct 27 (Thr) – Putting it all together: regression analysis**

Probability

- **Reading – Chp 5**
- **No assignment**
- **Nov 1 (Tue) – Definition of probability and conditional probability**
- **Nov 3 (Thr) – Expectations, variance, and large sample theorems**

Uncertainty

- **Reading – Chp 6**
- **Assignment 5 – due Nov 30 (Wed)**
- **Nov 8 (Tue) – Population vs samples and statistical inference**
- **Nov 10 (Thr) – Standard errors and confidence intervals**
- **Nov 15 (Tue) – Hypotheses and significance tests**
- **Nov 17 (Thr) – no class**
- **Nov 22 (Tue) – Statistical power and caution with NHST framework**
- **Nov 24 (Thr) – no class (Thanksgiving Break)**

Discovery

- **Reading – Chp 7**
- **No assignment**
- **Nov 29 (Tue) – Network data and analyses**
- **Dec 1 (Thr) – Spatial data and analyses**

- **Dec 6 (Tue)** – Textual data and analyses
- **Dec 8 (Thr)** – Course conclusion