

# Introduction to Statistics for Social Science

---

## Course Overview

This class will cover the core statistical concepts used in quantitative analysis. Its main goal is to develop a strong understanding of Ordinary Least Squares (OLS) regression, the most foundational statistical tool in quantitative social science research. We will also work in the statistical software R to illustrate each of the concepts we cover with real data about income inequality in the United States. At the end of the class, you will be able to interpret the results of statistical analysis, understand how important statistical tools work and what their limitations are, and perform analyses yourself with real data.

## Expectations – For You

- **Attendance** – The spring semester has a tight schedule. The concepts from one week will lead into the next. It's important not to miss meetings.
- **Videos and Readings** – I will make pre-recorded videos and choose textbook chapters to be as short as possible. They're important to read or watch carefully and fully before meetings.
- **Engagement/Communication** – Be present in meetings. Have your camera on if possible. Ask questions when you're not sure about something. Help your classmates when you can.
- **Do Your Own Work** – Your grade will be penalized if any work you turn in isn't fully your own.

## Expectations – For Me

- **Clarity** – One of the most frustrating things about any class is when the assignments are confusing or unexpected. I'll make sure you understand what you need to know to do well on all of them. They will all have a standard format and won't come with any surprises.
- **Purpose** – Everything will intentionally build toward the goal of the class. I won't cover random topics or give busywork assignments.
- **Flexibility/Respect** – Things are weird and unsettled in the world. Things may go wrong and make this class more difficult to focus on. I'll respect you as a person and work with you so you can still do well in the class if something disruptive happens.
- **Feedback** – I'll ask for your feedback during the semester. I'll make whatever adjustments I can that will benefit the class.

## How to Do Well in the Class

- **Be Organized** – Having a clear organization system will make computer-based assignments much easier. Having a weekly habit for the class will make learning concepts and completing assignments easier.
- **Ask Questions** – I'll do my best to explain concepts as well as I can the first time around. Nobody learns everything the first time around, though. Rewatch parts of videos, come to student hours, and ask me to repeat or clarify things during meetings. Other students will benefit from the questions you have.

- **Study Together** – It helps keep a schedule if you have a normal time you study with someone else. You can help each other learn the concepts you understood more quickly. Explaining a concept to someone else is the best way to learn it well yourself.
- **Let me know if you feel like you're falling behind** – I'm committed to working with you to make sure you can keep up with the material and do well in the course. Let me know what's going wrong as soon as you can when you're worried

## Grading

### Attendance – 20%

Attendance will be based on Zoom reports of who was in the meeting for at least 75% of the length of each meeting. There are only 15 total synchronous meetings of this class during the semester. You're required to attend all of them. Contact me ahead of time if there is a reason you have to miss all or part of one.

### Problem Sets – 60%

There will be three problem sets worth 20% of your grade each. They will each have a by-hand component, an R component, and an interpretation component where you describe your answers conceptually.

### Weekly Quizzes – 20%

You will complete a quiz on Canvas at the end of every week. You'll have two chances to submit correct answers to receive 100% on each quiz. You'll be told after the first submission which are correct and incorrect. Make sure you note which answers are correct so you don't have to re-do them when you try the quiz again.

**Extra Credit** – You'll have two extra credit opportunities worth 1-2% of your overall class grade each.

**Late Assignments** – Late assignments will be penalized by 10% for every 24 hours they are late. I'll adjust this based on how late the submission is. A submission that is 12 hours late will be penalized 5%, while one that is only 6 hours late will be penalized 2.5%. You have one 24-hour grace period for both problem sets and weekly quizzes.

## Required Materials

The books and other resources for the class are available to you online for free; you'll just need to know how to access and use them.

[OpenIntro Statistics](#) – This will be the main textbook for the course. You can download this for free or choose to pay the authors however much you want.

[Regression Analysis by Example](#) is available online through the University website. I'll upload the relevant parts to the course website.

R and RStudio. I'll have detailed instructions on downloading these available to you on the Canvas site. If you want to now, you can download the base R program [here](#) (choose the option for the type of computer/operating system you use), then the free version of RStudio Desktop [here](#).

## Course Outline

### Week One: Introduction to R and Summarizing Variables

**Goal:** Get familiar with R and a sense of the course's direction

**Wednesday:** Overview and Measures of Centrality

**Readings:** OpenIntro Statistics 2.1  
(2.1.7 and 2.1.8 not required, don't focus on variance & standard deviation)

**Synchronous Lecture:** Introduction, syllabus, and centrality

**Friday:** Tutorial and Applications in R

**Reading:** R Startup Document                      **Videos:** Starting with R  
(Open this, then start the first video)                      R Basics

**Synchronous Tutorial:** Interacting with data and calculating statistics in R

---

### Week Two: Measuring Distributions and Relationships

**Goal:** Quantify characteristics of distributions and place specific values in them

**Monday:** Measuring Distributions

**Reading:** Review 2.1.4 through 2.1.6                      **Videos:** Variance & Standard Deviation  
OpenIntro Statistics 4.1    The Gaussian/Normal Distribution

**Wednesday:** Relationships between Variables

**Synchronous Lecture:** Covariance and correlation

**Friday:** Tutorial and Applications in R

**Synchronous Tutorial:** Visualizing distributions and Z-scores

---

### Week Three: Learning about Society by Sampling

**Goal:** Relate values from samples to the larger population and understand statistical confidence

**Monday:** Inferences for Proportions

**Reading:** OpenIntro Statistics 5.1-2                      **Videos:** Samples and Populations  
Confidence Intervals for Proportions

**Wednesday:** Inferences for Means

**Readings:** OpenIntro Statistics 7.1

**Synchronous Lecture:** Confidence Intervals for Means

**Friday:** Tutorial and applications in R

**Synchronous Tutorial:** Simulating samples and calculating standard errors

---

#### **Week Four: Testing Hypotheses**

**Goal:** Understand the steps of hypothesis testing and apply them in three situations

**Monday:** The basics of hypothesis testing

**Reading:** OpenIntro Statistics 5.3

**Videos:** Hypotheses about Proportions  
Hypotheses about Means

**Wednesday:** Differences between groups

**Reading:** OpenIntro Statistics 7.2-3

**Synchronous Lecture:** Hypothesis Testing for Differences in Means

#### **PROBLEM SET ONE DUE THURSDAY**

**Friday:** Tutorial and applications in R

**Synchronous Tutorial:** Examining differences between groups in R

---

#### **Week Five: Calculating Bivariable Regressions**

**Goal:** Calculate the relationship between variables using least squares regression and interpret results

**Monday:** Memorial Day!

**Wednesday:** Calculating Regressions

**Reading:** OpenIntro Statistics 8.1-2

**Synchronous Lecture:** Regression concepts and math

**Friday:** Tutorial and applications in R

**Synchronous Tutorial:** Interpreting regression output and checking assumptions

---

#### **Week Six: Bivariable Regression Hypothesis Tests and Fit**

**Goal:** Calculate hypothesis tests with bivariable OLS and identify issues with them

**Monday:** Testing Hypotheses with Regression Coefficients

**Reading:** Regression Analysis 2.6-7, 9  
WATCH VIDEOS FIRST  
Skip “A Test using Correlation Coefficient” section

**Videos:** Regression Hypothesis Tests  
Standard Errors for Regressions

**Wednesday:** Assessing Regression Assumptions

**Synchronous Lecture:** Assessing the fit of bivariable OLS

## **PROBLEM SET TWO DUE THURSDAY**

**Friday:** Tutorial and applications in R

**Synchronous Tutorial:** Interpreting OLS hypothesis tests and fit

---

## **Week Seven: Multivariable Regression**

**Goal:** Interpret and assess fit of OLS with multiple predictor variables

**Monday: Reading:** OpenIntro Statistics 9.1

**Videos:** Confounding Variables  
Interpreting Multivariable Regressions

**Wednesday:** Real applications of multivariable regression

**Reading:** Budig, Michelle J., and Paula England. 2001. “The Wage Penalty for Motherhood.”  
American Sociological Review 66(2):204–25.

**Synchronous Lecture:** Understanding multivariable regression in action

**Friday:** Tutorial and applications in R

**Reading:** OpenIntro Statistics 9.3

**Synchronous Tutorial:** Interpreting multivariable OLS and checking assumptions

---

## **Week Eight: Review,**

**Monday:** Synchronous Review

## **PROBLEM SET THREE DUE THURSDAY**

