Development Data Bootcamp for Young Economists

Summer 2025

Building Inclusive Growth (BIG) Lab, Department of Economics
University of Notre Dame

Course Details

Instructor: Coleson Weir
Email: cweir2@nd.edu
Dates: May 12-16 and May 19-23, 2025
Class Meeting Times:

- One 45-minute morning Zoom lesson from 9:00 - 9:45 a.m. EST
- One 60-minute afternoon Zoom meeting from 2:00 - 3:00 p.m. EST
- Class on May 23 will be from 9:00 - 10:45 a.m. for final project presentations

Class Location: Zoom
Class Site: Google Drive

Course Overview

This course covers the core concepts necessary for quantitative research assistants. Completion of an Introduction to Statistics course or its equivalent is a prerequisite for the bootcamp. This course stresses the importance of attention to detail and replicability in the practical implementation of quantitative research tasks in STATA. Topics include basic programming, data cleaning, plotting graphs, and making regression tables. Best practices for working collaboratively will also be covered.

There is no course credit for taking this training course, and thus there is no grading. The main payoff from this course is in the skills you will develop to work as a research assistant in development economics, and the preparation it will prove for jobs that require quantitative skills.

Zoom Meeting Link:

https://notredame.zoom.us/j/92601346660?pwd=bn0lFUtmvhbemCV3FST4nzaI7TaTNY.1

- Meeting ID: 926 0134 6660
- Meeting Passcode: 365684

Google Drive Link:

All of the class materials, including data, codes, slides and readings, will be shared with the class via Google Drive. You should review slides before class. The link to the Google drive is:
https://drive.google.com/drive/folders/13VFsUzuwYL2ySkJKNRi1EP6uyyXmQ-bS?usp=sharing

Course Material

1. The main resource for this class covers important principles that good research assistants should follow in their work. We hope this will be a resource you can come back to again and again.
    a. Gentzkow, M., & Shapiro, J. M. (2014). Code and data for the social sciences: A practitioner's guide. Chicago, IL: University of Chicago.

2. Readings (include notes and blog posts) for each section will be the class Google Drive ("Readings" folder). These will be helpful for understanding the class materials and we recommend you read them after each component of the class.
3. Accessing STATA: You will need to download a short-term licensed copy of STATA, the main statistically packaged used by applied economics researchers. The instructions on the STATA installment (including the license) will be provided right before the class.
4. Exercises that are assigned in class should be submitted to the shared class Google Drive under the folder titled "Exercises". You are encouraged to discuss exercises with your group members, but everyone should have an individual submission. The primary purpose of these practice problems is to let you get familiar with the commands we talk about in class. So just try your best and play with the codes!
5. The final exercise will involve generating a short research paper in Overleaf, using output produced in STATA. More details on this project will follow.

Course Schedule

1. Introduction and Preparation (Day 1)
    a. Programming Languages
        i. Why we use Stata
        ii. Getting access to Stata
    b. General Workflow Management
        i. What makes a good research assistant
        ii. Examples for folder structure
    c. Introduction to Randomized Controlled Trials
2. Introduction to Stata (Days 2, 3, 4, and 5)
    a. Stata quick start
        i. Getting to know the Stata interface and windows
        ii. Importing data into Stata and data types
        iii. Understanding data structure in Stata
    b. Basic commands in Stata for data management
    c. Stata files
        i. Introduction to log and do files
        ii. Writing your own do file
    d. More advanced commands
    e. Basic programming
        i. Conditional expressions
        ii. Scalars and matrices
        iii. Macros: global versus local
        iv. Algorithmic structures: loops
3. Making Effective Graphs (Day 6)
    a. Examples for useful and effective graphs
    b. Two-way graphs
        i. Basic syntax

        ii. Fine tuning graphs: shapes, colors, titles, labels, tick marks, secondary axes
    c. Other graph types
        i. Bar, line, histograms, scatter plots, pie charts
        ii. Stata help documentation
4. Linear Regression (Day 7)
    a. Understanding Ordinary Least Squares (OLS)
        i. Why use regression analysis?
        ii. Interpreting regression results
    b. Exporting regression tables
        i. How to make a regression table in Stata
        ii. Automating the process of table-making
5. Preparations for your own research (Day 8)
    a. Generating summary statistics tables
        i. What is a summary statistics table and why do we need one?
        ii. How to extract summary statistics from Stata and moving the output into other environments such as LaTex
    b. LaTex templates and Overleaf
    c. Organization and Replicability
        i. Automating every step of quantitative research production
        ii. What does replication mean
        iii. Importance of replication in economic research
            1. Review AEA replicability requirements
            2. Look at a good example of a replication package
6. Working with relational data (Day 9)
    a. Types of data
        i. Time series, panel data, cross sectional data
    b. Commands to work with relational data
        i. Reshape and collapse
        ii. Preserve and restore
        iii. Merge and append
7. Wrap Up (Day 10)
    a. String commands (if time permits)
    b. Present final projects in class

Academic Honesty

You are expected not to engage in academic dishonesty. One of the goals of this course is to prepare you to do well in advanced economics courses, and build the basic foundations of regression analysis for research projects and future careers, so it is essential that you develop a solid understanding of the material. To that end, while collaboration on problem sets is encouraged in this course, copying is not. Students are expected to understand and abide by the University's honor code.

Support for Mental Health:

The University Counseling Center (UCC) provides cost-free and confidential mental health services to help you manage personal challenges that threaten your emotional or academic well-being. Remember, getting help is a smart and courageous thing to do. For more resources, please see ucc.nd.edu. For course-related concerns, you can also speak with me at any time.

Disability Services:

It is the policy and practice of the University of Notre Dame to provide reasonable accommodations for students with properly documented disabilities. Students who have questions about the Office of Disability Services or who have, or think they may have, a disability are invited to contact the Office of Disability Services for a confidential discussion in the Sara Bea Center for Students with Disabilities or by phone at (574) 631-7157. Because the University's Academic Accommodations Processes generally require students to request accommodations well in advance of the dates when they are needed, students who believe they may need an accommodation for this course are encouraged to contact the Office of Disability Services at their earliest opportunity. Additional information about Disability Services and the process for requesting accommodations may be found at supportandcare.nd.edu.

Inclusive Classroom:

The University of Notre Dame is committed to social justice. I share that commitment and strive to maintain a positive learning environment based on open communication, mutual respect, and non-discrimination. In this class, we will not discriminate on the basis of race, sex, age, economic class, disability, veteran status, religion, sexual orientation, gender identity, color, or national origin. Any suggestions as to how to further such a positive and open environment will be appreciated and given serious consideration.

Title IX:

Lauren's Promise: I will listen and believe you if someone is threatening or harassing you.

Lauren McCluskey, a 21-year old student at the University of Utah and the daughter of economist Jill McCluskey was murdered on Oct. 22, 2018 by a man she briefly dated. I am committed to doing what I can to make sure this does not happen again.

If you are in immediate danger, call 911. If you are experiencing sexual assault, domestic violence, stalking, or harassment you can report it to me. You should be aware that I am a mandatory reporter, which means that I am required to report these instances to the university's Title IX Coordinator or Deputy Title IX Coordinator to investigate. You can also learn about your options (on-campus and off-campus, confidential and not) at http://titleix.nd.edu.

The University of Notre Dame provides services for those who have been affected by sexual assault, sexual misconduct, dating or domestic violence, stalking and any conduct that creates a hostile environment. For help and further information including contact information for on and off-campus resources, please consult https://titleix.nd.edu/support-resources.