EFOP 2422 Data Analysis Using Computer Packages Syllabus, Fall 2024

Professor Joshua Bleiberg (jbleiber@pitt.edu)

5404 Posvar Hall; Thursday; 3:00 to 5:40

Appointments and Office Hours: By appointment and via Calendly

Overview

The goal of this class is to enhance your ability to use programming for conducting research. Programming is a skill set whose practitioners are too often guilty of "gatekeeping". Ironically statistical programming languages (e.g., STATA) are laborously designed to accommodate users of all experience levels. We will focus on demystifying coding, collecting, and cleaning data. Statistics courses primarily focus on teaching a narrow set of quantitative analysis tools (e.g., OLS, t-test). This class will focus on all other aspects of programming that are applicable for all methodological approaches. The goal is to improve the efficiency and speed with which you can conduct research.

The content of this class will focus on finding data and preparing it for analysis. The course has four modules: introduction, data collection, data cleaning, and analysis. The introduction will orient students towards the course structure and describe STATA's programming capabilities. In the data collection module we will explore available datasets, access them, and transform them into files usable for analysis. We will learn how to clean raw data for analysis. Finally we will explore how to make high quality figures by incorporating the principles of design. By the conclusion of the course you will have a program that will render a dataset ready for analysis.

This course references materials created by Doctors Richard Blissett, Benjamin Skinner, Daphna Bassok, Carly Robinson, and Katherine Meyer. I am indebted to each of them for their contribution.

Course Texts and Materials

There are no required course texts. All materials will be made available via Canvass. You are required to use STATA, which is available for free to all Pitt students.

Suggested Resources:

Acock, A. C. (2023). *A gentle introduction to Stata*. Stata press.

Mitchell, M. N. (2015). *Stata for the behavioral sciences*. College Station, TX: Stata Press.
EPI. (2024). *Translating R to STATA.*

Collaboration with Peers

Working with colleagues is the norm in research. All assignments may be completed in collaboration with a peer or peers. There is tremendous value in peer learning and collaboration is permitted with students enrolled in the class. However, seeking help from other PhD students outside of this class, professors, or other skilled programmers is not permitted. For the exercises you may choose to work with a peer then you are still responsible for your own work and must submit your own assignment. You must also describe the peers that you are working with as a comment in your program. You are also permitted to ask peers for advice on the Final Program, but you must submit your own program for a distinct project.

Plagiarism

Plagiarism in the context of programming differs from other creative works. There is a strong tradition of open source programming. Open source means that a programmer has proactively given permission for other users to copy the code for the good of the community at large. For this reason copying and adapting code is a common practice in programming in a way that differs from writing or graphical presentations. Another practical reason for this norm is that there may only be a single way to write a line of code. A simple example would be if I asked you to summarize variable y in STATA. There is one line of code that will execute that command. Plagiarism requires copying a unique or creative element that is frequently not present in coding. The importance of providing a correct attribution is the same as other media. If you copy code from a classmate, website, or Large-Language Model then you ought to describe the source in your code. Failure to attribute the source of code for a program is a form of plagiarism.

Grading and Student Responsibilities

Assignment	**Weight**
Class attendance, participation, and readings	35%
Exercises	20%

Draft Program I	10%
Draf Program II	10%
Final Program	25%
Code Review	5%

Class attendance, participation, and readings

Attendance is required for each session. If extenuating circumstances do not allow you to attend class in person, then provide me as much advance notice as possible. I will provide accommodations if I am able. Each class will include lecture, discussion, and a lab. I have assigned three readings throughout the semester. They are meant to provide insight into the principles that are important to incorporate into your regular research routines. More specifically the readings will explore critical quantitative methods and research reproducibility.

Lectures will introduce new programming concepts and provide an overview of solutions. Programming classes are unique in the way that students will encounter content that is highly familiar and novel. For this reason I am particularly in need of feedback when a concept is unclear. Always assume that any confusion you are feeling about the content is shared among your peers. We will add impromptu discussion throughout the class.

Each class session will also include a lab with coding demonstrations. During the lab students will have the opportunity to work on that week's exercise or a final project. Students are encouraged to work on assignments during class and to ask me questions as they emerge. I am also able to answer questions via email and to meet outside of class. However, debugging code often works best in person as it allows me to see the data and diagnose the problem iteratively.

Exercises

Most weeks there will be an exercise due. Assignments will be submitted in the form of a STATA program or "Do file". All exercises are graded on the basis of being complete or incomplete. If an assignment is incomplete then I will provide additional feedback and time to complete the exercise. It is my expectation that all students will receive full credit for all of the exercises. The goal of the exercises is to provide an opportunity to directly apply your learning. Additionally, the exercises will provide a code repository that you can reference when writing future programs.

Final Program

The summative project for this class will be a "final" program. You have wide flexibility to choose your own project for the final program (e.g., milestone, GSRA work, personal project). The summative assignment has three different components: Draft Program I, Draft Program II, and Final Program. The structure of the assignments will allow me to provide feedback and ensure that sufficient progress is made. Each program assignment is due at the completion of a module and requires you to implement those skills in each iteration of the program. The final program will not require extensive coding of results; this will allow students to focus on finding and cleaning data. Student's whose focus is on qualitative research methods can use the topic to create sample descriptions, sentiment analysis, or graphical analysis of survey/interview data. The final program will execute code that begins by downloading/importing data, prepares the data for analysis, and analyzes the data.

During the last class meeting you will present your program in a code review. Code review is akin to an informal presentation. You do not need to prepare remarks or slides. Each student will share their screen with the class and narrate each section of the code. Class members will then have a chance to provide feedback.

Grading Scale

Grade	Point Range	Grade	Point Range
A	93-100	C	73-75
A-	90-92	C-	70-72
B+	86-89	D+	67-69
B	83-85	D	64-66
B-	80-82	D-	60-63

C+	76-79	F	Less than 60

A grade of A+ will only be granted in cases of truly exceptional work that considerably exceeds the expectations for each assignment.

University Policies

Several university policies are pertinent to this course. As an instructor I am committed to pedagogy that is anti-racist, non-sexist, non-classist, non-heterosexist, non-ableist, and non-gender normative. This includes fostering an environment that is as safe and inclusive as possible. It is my intention to name and correct as best as possible any actions on my part that fall short of these commitments. If everyone in the class is likewise committed, the work should be more easily facilitated. I am also committed to making the class accessible to anyone wishing to participate.

University COVID Precautions.

The University of Pittsburgh updates its COVID protocols in response to changing conditions. See the latest information at coronavirus.pitt.edu. In this course we will adapt as needed. Please reach out and we'll make it work.

Availability of Instructor

The instructor will respond promptly to all emails and you may also request a face-to-face meeting with the instructor. It's my hope to be able to meet with everyone in person at least once this semester. Email for appointment or via Calendly.

Departmental Grievance Procedures

The purpose of grievance procedures is to ensure the rights and responsibilities of faculty and students in their relationships with each other. When a student in the EFOP Department or an EFOP course believes that a faculty member has not met his or her obligations (as an instructor or in another capacity) as described in the *School of Education Academic Integrity Guidelines*, the student should follow the procedure described in the Guidelines by (1) first trying to resolve the matter with the faculty member directly; (2) then, if needed, attempting to resolve the matter through conversations with Dr. Max Schuster, the associate chair of the Department of Educational Foundations, Organizations, and Policy; (3) if needed, next talking to the academic integrity officer of the school, Assistant Dean Dr. Andrea Zito; and (4) if needed, filing a written statement of charges with the academic integrity officer.

Academic Integrity

Students in this course will be expected to comply with the University of Pittsburgh's Policy on Academic Integrity. Any student suspected of violating this obligation for any reason during the semester will be required to participate in the procedural process, initiated at the instructor level, as outlined in the University Guidelines on Academic Integrity. This may include, but is not limited to, the confiscation of the examination of any individual suspected of violating University Policy. Furthermore, no student may bring any unauthorized materials to an exam, including dictionaries and programmable calculators.

Disability Services

If you have a disability for which you are or may be requesting an accommodation, you are encouraged to contact both the instructor and Disability Resources and Services (DRS) as early as possible in the term: 140 William Pitt Union, (412) 648-7890, drsrecep@pitt.edu, (412) 228-5347 for P3 ASL users. DRS will verify your disability and determine reasonable accommodations for these courses.

Statement on Classroom Recording

To ensure the free and open discussion of ideas, students may not record classroom lectures, discussion and/or activities without the advance written permission of the instructor, and any such recording properly approved in advance can be used solely for the student's own private use.

Accessibility

Ensuring an accessible and pleasant experience to all users, regardless of disability, is a key focus of Canvas. The Canvas platform was built using the most modern HTML and CSS technologies, and is committed to W3C's Web Accessibility Initiative and Section 508 guidelines. Read more about Accessibility within Canvas.

Copyright Notice

These materials may be protected by copyright. United States copyright law, 17 USC section 101, et seq., in addition to University policy and procedures, prohibit unauthorized duplication or retransmission of course materials. See Library of Congress Copyright Office and the University Copyright Policy

Religious Observances

The observance of religious holidays (activities observed by a religious group of which a student is a member) and cultural practices are an important reflection of diversity. As your instructor, I am committed to providing equivalent educational opportunities to students of all belief systems. At the beginning of the semester, you should review the course requirements to identify foreseeable conflicts with assignments, exams, or other required attendance. If at all possible, please contact me (your course coordinator/s) within the first two weeks of the first class meeting to allow time for us to discuss and make fair and reasonable adjustments to the schedule and/or tasks.

Your Well Being Matters

College/Graduate school can be an exciting and challenging time for students. Taking time to care for yourself and seeking appropriate support can help you achieve your academic and professional goals. You are encouraged to maintain a healthy lifestyle by eating a balanced diet, exercising regularly, avoiding drugs and alcohol, getting enough sleep, and taking time to relax.

It can be helpful to remember that we all benefit from assistance and guidance at times, and there are many resources available to support your well-being while you are at Pitt. If you or anyone you know experiences overwhelming academic stress, persistent difficult feelings and/or challenging life events, you are strongly encouraged to seek support. In addition to reaching out to friends and loved ones, consider connecting with a faculty member you trust for assistance connecting to helpful resources. The University Counseling Center is also here for you. You can call 412-648-7930 at any time to connect with a clinician.

If you or someone you know is feeling suicidal, please call the University Counseling Center at any time at 412-648-7930. You can also contact Resolve Crisis Network at 888-796-8226. If the situation is life threatening, call Pitt Police at 412-624-2121 or dial 911.