Course Syllabus

Contents

This course is an introduction to computing and data use for research in the social sciences. We designed the course to prepare first-year students for research experiences. Using software tools including Stata and Python, students will learn how to organize, store, and document project assets on computer resources, how to link various sources of data, and how to detect and fix data problems. Further, the class will engage in case studies for summarizing and visualizing data for analyzing specific questions.

- **Term**: Spring 2024
- **Class Time**: MW 10:00 - 10:50am
- **Class Room**: 219 Chancellors Hall

Instructor

	Robert Hicks
Office	252 Chancellors Hall
Email	rob.hicks@wm.edu
Office Hours	Thu 2:30pm-3:30pm
	Fri 10:00-11am

I am available outside of class at the times listed above or by appointment.

Course Notes

Course notes for this semester's Stata and Python modules are found at

- https://econ.pages.code.wm.edu/160/stata
- https://econ.pages.code.wm.edu/160/python

Computing Resources

All computing in Python and Stata will be on `https://jupyterhub.wm.edu`.

When logging into `jupyterhub`, choose

- `Stata + SciPy: The default notebook with Stata available` for Stata work (note limited to 20 licenses)
- `Default Notebook: A Python3 SciPy Notebook` for Python work (not limited by licenses)

I will demo these computing resources early in the semester.

Important Dates

Event	Date
First day of class	Jan 24
MidTerm	March 6
Spring Break	Mar 9-17
Last day of this class	May 1
Final	May 13: 2-5pm

Assignments

The course will include two types of work:

- *A midterm and a final exam* (40%). The midterm will cover the Stata and the final will cover the Python portion of the course. This means the final is not comprehensive. The exams will be closed book and handwritten and will draw from in-class exercises and the assignments (discussed in the next item), and will ask students to discuss output from code examples, or students will be asked to provide small code snippets for doing work. These exams will test students understanding of core concepts and won't explore novel coding problems.
- *8 assignments spread through the semester* (60%): They will involve writing code in either Stata or Python along with writing explanations of your work. Assignment prompts will be in the form of Jupyter notebooks. You will edit the Jupyter notebook with your answers, save it, and upload it to Blackboard's Assignments interface. We'll explain more details about assignments–including what in the world is a Jupyter notebook–as we progress through the course.
  - *Assignment due dates* : Assignments will be due by 11:59pm on Tuesdays.
  - *Collaboration* : You will work on assignments and submit them in groups of three students. We will match you into those groups using a randomization device (using Python!). Groups will change for each assignment. Collaborate within your group. Your group may talk with members of other groups about general coding and approaches to the assignment, but for the specifics of the assignment, collaborate only with your assigned partners.
    - While group-members will collaborate on the assignments, each group member must submit their work to blackboard for grading.
    - Group members will grade their group peers on a simple three point scale. Why are the assignments completed by groups?
      - Group work is extremely common throughout the workforce and organizations you join. Developing collaboration skills–with new people– is of long-term value.
      - It's easy to get stuck in the middle of a coding process. A partner helps a lot.
      - We suspect prior coding experience might cluster in friend groups. So if you pair up with friends, you might learn less overall as a class.

- *A note on coding and assignments* : Some of you don't have a lot of coding experience. That's okay: we designed this course for you. You will find the early parts of this class frustrating as you struggle to translate your logic into workable code. The curious student who is willing to experiment (and creatively search google) will keep frustration levels to a minimum. To facilitate the learning process, you can
    - Begin assignments early. We can't help students when they get stuck at 1am.
    - Ask anyone to help solve specific coding syntax errors.
    - Come to our office hours for help tackling these types of problems.
    - Google is a great resource for syntax problems. In particular, we find **stackoverflow.com** to have the best Python advice anywhere. For Stata, **statalist** is a great resource. **ChatGPT** is a great resource for either Stata or Python.

- *Policy on Late Assignments* : Course assignments should be turned in on time. Late work tends to pile up and become unmanageable. Late submissions will get a letter grade deduction for each late day. Late in-class assignments will not be accepted.

- *Hardcopy Policy* : No hardcopies of assignments are accepted under any circumstances. All assignments are submitted via the Blackboard Assignments interface and will be submitted as Jupyter notebooks.

- *Asking Questions*: Substantive questions about course material (but not specifics about assignments) or coding in Stata or Python should be submitted to the Course issue tracker at https://code.wm.edu/econ/160/issues/-/issues.

- *Grade Discrepancies and Grade Questions* : We are happy to discuss questions you have about your grade on class assignments. Any questions you have regarding a potential grade change on an assignment must be cleared up within 1 week of receiving your work back from us. The only exception to this policy is if we made a data entry or error in adding your score up.

Other Logistics

- *Email Policy* : We will respond to emails please include the the tag `[160]` in the subject line. Emails that don't have this tag will not be flagged as important and will take extra time to process.

Class Schedule

Topic	Approx. Duration
Research protocols and computer landscapes	2 weeks
Stata module: inputting data	1 week
Stata module: inspecting data	1 week
Stata module: shaping data	2 weeks
Python module: Python Basics	1 week
Python module: Introduction to Pandas	2 weeks
Python module: Data cleaning and manipulation	2 weeks
Python modules: Analyzing and Summarizing data	2 weeks
Using Python and Stata Together	As time allows

By Rob Hicks and Peter McHenry

© Copyright 2022.