Small Projects

This repository contains information and prompts for people working on research projects for me. In general these are "summer projects". This repository is public and multiple people working on projects for me will be interacting with it at any given time.

If you are interested in working on a project, then you need only sign up (see next sub-section), assign yourself one of the "Issues", and get to it.

The terms of this are very hands off from my perspective. I do not have the time to do careful coaching of anyone through these summer projects. These are entirely self-driven. See below for how to communicate questions or issues about the project to me.

If you complete a substantial set of work on one of the projects, I will be happy to discuss what kind of help I can give you going forward. Often that includes a reference letter for undergrad, grad, or post-doc positions.

Projects as issues

Each project I'm interested in is tracked as an "Issue" in Github. If you are going to work on one, you'll need to:

1. Have a Github account. You can sign up for one for free.
2. Assign yourself to the issue associated with the project you are going to work on. This ensures I know who is working on what, and allows others to see that a project is "claimed" already.

Communication and questions

All questions or comments on your project should go through your "Issue" in Github. You can leave comments there and I'll get notified. You can/should leave links, working examples, or copies of files that clarify your question.

I get notified of the comments, and I'll reply in the Issue itself. I don't monitor this continuously, so it is likely to be several days before I reply.

In general, trust yourself. There are no clear "right answers" for these projects. The most important thing is that you keep clear documentation of what you did (both in the code, and possibly in a write-up) so that we can replicate it, see whether it works, and modify if necessary.

Files

There are three folders that all your work should go into

1. Code. For, you know, any code that you write
2. Data. This is for raw data only. Any file that goes in here should be untouched by you in any way. Your code reads or manipulates this into a form that we can use.
3. Output. This is where those manipulated files get stored.

The basic idea is that Data -> Code -> Output.

Sub-folders

Within Code, Data, and Output, you should create a sub-folder for your particular work. That sub-folder should have the same name. For example, there should be Code/Geography, Data/Geography, and Output/Geography, if you are working on an economic geography project.

Excel and CSV

Excel is for raw data only. You can store an Excel file you download or find into the Data folder for your project.

But all output files of data have to be in CSV format. Excel is not viable.

If you have an Excel file as a raw input, then there must be something in Code that reads that Excel file and writes the CSV file to the Output folder. There is no cutting and pasting, or hand calculation involved or allowed here.

If you don't know how to write code to read Excel and write CSV, you need to figure that out for yourself

Code

You can use Python, R, or Stata. I'm not conversant in other languages and I don't want to learn them.

CoPilot, ChatGPT, and Stack Overflow are where you can find help. *I am not a help desk*