

Final Project Proposal

PSC290 - Data Management and Cleaning, Fall 2023

Emorie D Beck

Final Project Proposals

- Due at 11:59 PM PST on November 19, 2022 on Canvas
- 1-2 page (single spaced) proposal
 - Short background (What are you hacking?)
 - Approach (How you do plan to hack this? What “product” will you producing?)
 - Motivation (Why are you hacking this particular problem / issue / question / procedure?)
 - Challenges and barriers (what do you struggle with with bringing your project to life; are there specific barriers?)

Goal

The goal of this course is to build a set of tools for cleaning and managing your data. As discussed in class, a big part of that is to create a workflow for your research that is efficient, clear, and minimizes errors. The final project in the course aims to provide an opportunity for you to flesh out what that means for you and research. There’s no right or wrong way to build a sustainable workflow for yourself.

Scope

The project is open-ended, but some examples are:

- A preregistration template for your kind of data / research
- A quarto / Rmarkdown / R script / R Project template for the typical structure of research projects
- A data cleaning pipeline / template for your kind of data / research
- A set of functions meant to help you improve your research workflow
- Writing clear documentation for a new or previous data collection

- A data cleaning or analysis script that draws on different course concepts (e.g., directory structures, codebooks, functions, iteration / `purrr`, etc.)
- A “checks” / procedure checklist that documents robustness tests and aims to improve the accuracy and efficiency of your workflow
- Anything else that draws on things you learned in class and focuses on data cleaning, R, data management, procedures, GitHub, etc.

Recommendation

For the final project, you will be creating something to help improve your workflow. You can think of it sort of like a chance to “hack” your own work. I recommend choosing something you’ve been wanting to do but haven’t had time to prioritize and using this as an excuse to do so.

As mentioned above, this could be anything from finishing up cleaning some gnarly data you’ve had for awhile, redoing your documentation for a study that’s already been collected (or you plan to collect), building templates or procedures, etc. What I want from you for this project is something that is: * Useful to you, both short- and long-term * Challenging but doable (you’re better at this than you think!) * Something you feel proud of