Practicum in Reproducible Research Methods

Homework Two
Second Term 2020–21
Professor Miriam Golden
Assignment due: February 9 by 15:00

For this assignment, each research team write a pre-analysis plan (PAP) for your study of the dataset you have selected. You will also write your Data Management Plan and prepare your EGAP pre-registration form.

If you have not clearly written out your research question, you will have the opportunity to do so when you write your PAP.

Before proceeding, you should review a sample of PAPs on the EGAP website. As you skim the various PAPs, consider which seem particularly complete and thorough — but no longer than absolutely necessary. If you had the data, which PAPs would give you enough information to complete the entire analysis without having to talk to the P.I.s? Which PAPs answer all your questions about how to construct variables and estimate a statistical model?

Here are suggestions for the material you will want to include in your own PAP:

1. Research motivation (the question) [1 para]
2. Hypothesis(-es) and causal chain [2–3 para]
3. Dataset to be used (describe all variables)
4. Variables to be used (independent, dependent, and any controls)
5. Estimation strategy (equations)
6. Results expected, if known
7. Robustness and sensitivity checks to be performed

DMP: As part of preparing your analysis, you also have to prepare a Data Management Plan. You will do this on the DMPonline site.

Pre-registration: After you complete your PAP, you will also prepare your pre-registration

form. The required document is available at the EGAP registry website. The form requires you to provide details about your research design.

Due as three separate .md documents by 15:00 February 9 on your team's GitHub homework repo:

1. PAP
2. DMP
3. EGAP pre-registration form

As always, please feel free to post questions on Brightspace if you encounter obstacles. Other students may be able to help you out, or I may.