

## Individual Final Project:

### Data

- Explore one of the three datasets we have discussed: the replication folder, South African labor force data in 2008, or the 2022 General Social Survey (GSS). Propose a research question that the variables in the chosen dataset can answer. Here are a few examples to inspire you:
  - From the replication folder: In this sample, whether students have a tv at home and can spend time watching tv have a higher gender attitude index?
  - From the South African Labor force data: Will a low educational level increase a person's possibility of working in an informal sector?
  - From the GSS: Would older generations tend to have more children?
- It is totally okay if you would like to use other dataset, but please provide a brief introduction to the dataset.

### What you need to do?

1. Briefly explain your research question and propose your hypothesis.
2. Find the relevant variables that can answer your research question in the dataset, and provide some summary statistics for these variables. (Remember that One graph is better than ten tables!)
3. Test your hypothesis by a simple regression model.
4. Explain your regression results, mainly about the coefficients and the significant level. If it is consistent with your hypothesis, explain the channels. If it is not consistent with your previous hypothesis, provide possible reasons.

### A possible roadmap:

1. [After learning Topic 4: regression] Explore one of the datasets, propose a research question that you are interested and write it down. The research question is usually in the form of "What is the effect of X on Y?", so please list the X and Y, and also other variables besides X in that database that might affect Y.
  - Research question:
  - X: Y:
  - Other control variables:
2. [After learning Topic 6: Prepare for your own research] Generate the summary statistics table for your key variables.
3. Run regression and export the regression tables.
4. [After learning Topic 7: Latex] Write your paper in Overleaf

**Submission:**

1. A 2-4 page final paper written in Latex. A possible outline could be:
  1. research question and hypothesis
  2. brief data and variables description (better with graphs)
  3. discussion of results (the main part)
  4. discussion of challenges
2. Do file and log file.