

Homework 1

The problem set requires you to use one data set of your interest. We require you to download the data set, clean it up and put it in a format that you can comfortably work with. Depending on your research idea you may want to restrict the sample (say to young, married couples of a certain age). We also require you to provide summary statistics of the variables of interests.

In addition, we want you to focus on **one fact** you think it is important and can be represented in a **Figure**. The most impactful the fact and the figure the better. Then you have to write 1000 words to describe what is in the figure and why is important.

You can obtain extra points by providing more work than required (presenting new facts, sensitivity and regression analysis, etc). Prepare a 5-10 minute presentation of the results (4-5 slides) discussing sample selection, obstacles encountered, how you solved the problem, and findings.

Common US data sets used in modern macroeconomics are the CPS, PSID, SIPP, CEX, SCF, ATUS, ACS, Nielsen Retail Data, OI Economic Tracker, and NLSY. All these data sets are public. There are also other data sets that you have to apply for permission or buy a subscription.

This is not an exhaustive list and you can work with a data set that is somewhat unexplored if you think it is better suited for your research idea. For example, it is common nowadays to use cross-country micro data for example from OECDStat data repositories, or from each country's sources. Good examples of the latter are the German SOEP, the CNEF household data repository, the German LIAB or Danish matched employer-employee data sets, firm level data from the OECD, firm level and household data from developing countries in the World Bank data bank.

To obtain access to some of these data sets you need to follow an application procedure, some of them can be complicated so you have to be careful with that. As for the Nielsen data set, you are fortunate to be subscribers and you need to talk to Bent to get the access to it, it is a great opportunity for us.

Upload your results by **midnight February 5th**. To upload your homework use the following Dropbox's link <https://www.dropbox.com/request/0IktHWDdbWMXsS6sYXzY> Please upload the code you wrote in a zip folder with your name. The zip folder must contain a ReadMe file with the instructions to run the code. Try to automatize as much as you can and comment every line of the code to inform yourself and the reader what the code is doing.

There, you will find subfolders for each homework. Please upload the code you wrote in a zip folder with your name. The zip folder must contain a ReadMe file specifying the data you used, the procedure to run the code including the link to the data, etc. Try to automatize as

much as you can and comment every line of the code to inform yourself and the reader what the code is doing.