

Final Project Documentation

Author: Kristóf Reizinger

Date: 4th April, 2022

Files

My files

- create_folder_structure.txt
- data_manipulation.py
- data_manipulation_stata.do
- 1_data_preparation.do
- 2_merge_files.do
- 3_for_cycle.do
- 4_regression_plotting.do

Data files

- health2019.csv
- health2019_part1.csv
- health2019_part2.csv

Generated files

- health2019.dta
- health2019_new.dta
- vaccinated.rtf
- vaccinated.txt

Coding design

Folder structure Firstly, I have created the Folder structure to my project using command in `create_folder_structure.txt` file. I moved the data files into the relevant folders.

Python exercise Considering the Python related exercise, I recommend to navigate to *codes* and *python_codes* folder. You can run the python file in this folder. (All necessary data are available here.) You will need Internet connection, because I use finance.yahoo.com to get some financial data, which has dictionary format to demonstrate relevant dictionary related skills. (This is the reason, why the python code `data_manipulation.py` takes longer time to run.) All functions are defined step by step, you can run the whole file and it will print out the results.

Stata exercise To reproduce the Stata exercise, you should go back to the main folder (*CEU_coding_project*). You can run the do files from this folder. When I read or write data, I have defined relative path. (Versus to the Python part, all paths are absolute. So, you should take it into consideration, when running the code.)

I splitted my code into four parts:

- 1_data_preparation.do
- 2_merge_files.do
- 3_for_cycle.do
- 4_regression_plotting.do

The numbers indicates the recommended order to run the files.