

Day 2 - EU-SILC

Part II Practical Training Session: Cross-sectional data structure & analyses

Exercise 3 – Appending multiple years

Append cross-sectional P-files, harmonize the educational attainment variables PE020 and PE021 to single digit ISCED, and check for breaks and jumps.

Data sets needed

- all cross-sectional p files from the years 2013 to 2021
- UDB_c21P_ver_2022_release2.dta
- UDB_c20P_ver_2022_release2.dta
- ...
- UDB_c14P_ver_2022_release2.dta
- UDB_c13P_ver_2022_release2.dta

Variables needed

- PB010 Year of the survey
- PB020 Country alphanumeric
- PB030 Personal ID
- PB040 Personal cross-sectional weight
- PE020 ISCED level currently attended (for 2020 to 2013)
- PE021 Level of current/most recent formal education or training activity (for 2021)

Solution steps

- Open and append P files from the years 2013 to 2021
 - Best to start with the latest and adding the older ones
- Create harmonized education variable of first digit of ISCED level, with ISCED 5-8 combined
- Create data set of percentages of ISCED levels across years and countries
- Plot percentages of ISCED levels across years and countries