Social Diagnosis 2011 - Objective and Subjective Quality of Life in Poland

Details of the variables are in the synthpop help file for SD2011

For details do `help(SD2011)`

Note: Some numeric variables may have negative values to stand for missing or not applicable codes (e.g. -8 or -9). All factors have codes. The variable **bmi** is body mass index (weight in Kg divided by height in metres squared).

Practical 1

Select a subset of up to around 8 variables to synthesise and form a new data frame with just these variables.

The practical consists of using different methods to synthesise this smaller data set = you could call it `mydata`. You can find some sample code in **synthesis_SD2011, but you would learn more if you tried to write your own code, using the help files, perhaps with some help from the sample code.**

Suggested tasks

- Take a look at the data and try to understand the variables a little.   Use codebook.syn to check the data types and number of variables
- Are there any variables with many levels that you might want to put at the end of the visit.sequence?
- Are there an y special values for numeric variables that might need a `cont.na` parameter?
- Are there any rules that the variables should obey, or any that could be derived from others,
- Choose a full conditional synthesis method this method. Perhaps change parameters to speed up synthesis and/or improve utility. You can try:
    - Method (either a single value or a vector of same length as number of variables)
    - visit.sequence
    - predictor.matrix
    - other parameters for each method e.g. cart.minbucket ot cart.cp see help files
- Use the functions *compare.synds* and *utility.tables* to evaluate the synthesis. Are any rules preserved with the synthetic data?
- When you have a synthetic data set that seems OK, use the `sdc()` function to make it ready to send to a user.