Subject: Re: Mergining PR onto a Combined IR, MR Dataset in SPSS
Posted by Trevor-DHS on Tue, 03 Feb 2015 16:03:01 GMT
View Forum Message <> Reply to Message

There are a number of issues to watch out for in combining data in the manner that you are talking about:
1) Not all women and men in the PR file will be found in the IR and MR files, for a number of reasons:
a) Age: for women this is 15-49 in almost all cases, but for men this can vary - 15-49/54/59.
b) Individual interviews not successfully completed are excluded from the IR and MR files.
c) Men's data are usually collected on a subsample of the data, typically half or one third of households. Using women's and men's data together would require recalculating weights for the combined data.

2) While generally the coding of the recode variables is done the same way for men and women, there are often differences and you would need to check the coding of the women and men's variables before combining them.

Below I am including the outline of some code for combining the IR and MR data, and then merging the PR data:
* Working directory.
cd 'C:\Data\DHS_SPSS'.
* Open the men's data.
get file='BDMR61FL.SAV'.
* rename men's variables to the same name as women's variables.
rename variables (MCASEID MV000 to MV016 MV021 to MV032=CASEID V000 to V016 V021 to V032).
rename variables (MV801 MV802 MV803 = V801 V802 V803).
rename variables (MV101 to MV107 = V101 to V107).
rename variables (MV130 MV131 MV133 to MV136 MV138 = V130 V131 V133 to V136 V138).
rename variables (MV149 to MV152 MV155 to MV159 MV167 MV168 MV190 MV191 = V149 to V152 V155 to V159 V167 V168 V190 V191).
rename variables (MV201 to MV207 MV212 MV213 MV217 MV218 MV225 MV245 to MV252 = V201 to V207 V212 V213 V217 V218 V225 V245 to V252).
* ...
rename variables (MV714 MV716 MV717 MV719 MV721 MV731 MV732 = V714 V716 V717 V719 V721 V731 V732).
* ...

* Sort men's data according to ID fields.
sort cases by V001 V002 V003.
* Save the required subset of data - adapt to your needs.
save outfile='MRsmall.sav' /keep=caseid V001 to V032 V101 to V107 V714 V716 V717 V719 V721 V731 V732.

* Open the women's data.
get file='BDIR61FL.SAV'.
* Sort women's data according to ID fields.

sort cases by V001 V002 V003.

* Add men's data to women's data.
ADD FILES /FILE=*
  /FILE='C:\data\dhs_spss\MRsmall.sav'
  /KEEP=caseid V001 to V032 V101 to V107 V714 V716 V717 V719 V721 V731 V732.

* Match PR data to combined women and men's data.
MATCH FILES /FILE=*
  /TABLE='C:\data\dhs_spss\BDPR61FL.SAV'
  /RENAME HV001=V001 HV002=V002 HVIDX=V003
  /BY V001 V002 V003.

In this code, I recommend that you select the variables that you are interested in using from each of the IR, MR and PR files. I've given an example, but you will need to adapt this.
Also note that renaming lots of variables is a pain in SPSS. You can use the TO parameter, and I've given a couple of examples, but for this to work the variable have to be in sequence with no gaps in the numbering.

I hope this helps.