
Subject: Merging PR onto a Combined IR, MR Dataset in SPSS

Posted by [cudis](#) on Mon, 02 Feb 2015 03:05:44 GMT

[View Forum Message](#) <> [Reply to Message](#)

We are interested in conducting an analysis of determinants of employment and hence would like to create a combined IR and MR file of all interviewed individuals, with relevant household member data on background and household characteristics from the PR file merged onto the individual records, in SPSS.

We have found an analogous procedure outlined in a previous topic (quoted below).

Bridgette-DHS wrote on Wed, 20 February 2013 11:46 Here is a response from one of our STATA experts Tom Pullum, that should answer your question.

Assuming you want a single file that includes men and women, as individuals, and has the HIV data merged onto the individual records. In Stata, I would go through the following steps:

Open the IR file and construct a new variables, sex=2 (female), for all cases. Save as file 1.

Open the MR file and construct a new variable, sex=1 (male), for all cases. For all variables you need that have an mv prefix, drop the m, so the prefix becomes v. Save as file 2.

Open the AR file and change the variable names exactly as you did. Sort on v001 v002 v003. Save as file 3.

Open file 1. Then APPEND file 2 to the end of file 1, getting a file with all the men and women as observations. Sort on v001 v002 v003. Merge with file 3, on v001 v002 v003. Drop any cases with hiv05 missing (these are cases with no HIV result). Save as file 4. This will be your working file.

In file 4, the preferred weight will be hiv05, rather than v005. The cluster variable will be v001 (it is duplicated in the v020's but we always use v001). There may be a variable that is identified as a stratum variable, e.g. v022, but we recommend that you use what is identified as the domain variable, e.g. v023. If domains are not given, you can construct a domain variable for virtually all the surveys as all combinations of region and urban/rural.

However, in our case, is it necessary to construct a new sex variable in the individual files, as this information is available in the household member recode and could be included when the PR data is merged onto the combined IR and MR file?

Also, we wonder whether the above steps need to be modified specifically for SPSS and/or the household member recode, and if an example syntax could please be provided?

Subject: Re: Merging 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.

Subject: Re: Merging PR onto a Combined IR, MR Dataset in SPSS
Posted by [cudis](#) on Fri, 06 Feb 2015 19:13:02 GMT

[View Forum Message](#) <> [Reply to Message](#)

Thank you very much for the advice, Trevor, it is very useful indeed!

We have two follow-up questions, if you do not mind:

Given that, as you note, only a subsample (half or one-third) of men within a certain age range are interviewed, compared to typically all women of the same (or similar) ages, we would like to confirm whether the interviewed subsamples (for both genders and combined) remain nationally representative for the relevant age groups?

Secondly, how should the new weights for the combined data be calculated?

Subject: Re: Merging PR onto a Combined IR, MR Dataset in SPSS
Posted by [Trevor-DHS](#) on Sat, 21 Feb 2015 20:48:17 GMT

[View Forum Message](#) <> [Reply to Message](#)

The two samples of women and of men are individually representative, but the combined sample of women and men is not representative because only a subsample of men was selected. To adjust for this you need to de-normalize the weights. See this message for a note on de-normalization from our sampling statistician.

Subject: Re: Merging PR onto a Combined IR, MR Dataset in SPSS

Posted by [Gazimbi](#) on Wed, 01 Apr 2015 22:00:50 GMT

[View Forum Message](#) <> [Reply to Message](#)

Hi, Trevor

Thank you for your custom. I hope this time I found my saviour. There is quite a significant amount of discussion on merging datasets in this forum but I don't seem to get it. Here is my problem, I want to merge Zimbabwe DHS datasets (women, men and HIV). I know I should sort and match files, however, I don't know how to drop m prefix on the male variables. I am just wondering whether the renaming procedure you provided below in another discussion could be applicable to my situation in Zimbabwe

* 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.

Your help will be greatly appreciated

Subject: Re: Merging PR onto a Combined IR, MR Dataset in SPSS

Posted by [Gazimbi](#) on Wed, 01 Apr 2015 22:15:12 GMT

[View Forum Message](#) <> [Reply to Message](#)

Hi, Trevor

Thank you for your custom. I hope this time I found my saviour. There is quite a significant amount of discussion on merging datasets in this forum but I don't seem to get it. Here is my problem, I want to merge Zimbabwe DHS datasets (women, men and HIV). I know I should sort and match files, however, I don't know how to drop m prefix on the male variables. I am just wondering whether the renaming procedure you provided below in another discussion could be applicable to my situation in Zimbabwe

* 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.

Your help will be greatly appreciated

Subject: Re: Merging PR onto a Combined IR, MR Dataset in SPSS

Posted by [Trevor-DHS](#) on Thu, 02 Apr 2015 15:25:20 GMT

[View Forum Message](#) <> [Reply to Message](#)

Yes, it sounds like you should:

- 1) follow the procedure given to rename the men's variables to the same name as the women's variables,
 - 2) then concatenate a selected set of women;s and men's variables using the Add CASES command in SPSS,
 - 3) sort the combined data using SORT CASES and
 - 4) then merge the HIV data onto the combined women and men's data using MATCH FILES /FILE=* /TABLE=hiv dataset ...
-

Subject: Re: Merging PR onto a Combined IR, MR Dataset in SPSS

Posted by [Gazimbi](#) on Sun, 05 Apr 2015 11:32:39 GMT

[View Forum Message](#) <> [Reply to Message](#)

Thank you very much Trevor. Lastly, do I have to construct sex=2 (female) in IR and Sex=1 (male) in MR files.

Thank you again

Subject: Re: Merging PR onto a Combined IR, MR Dataset in SPSS

Posted by [Trevor-DHS](#) on Sat, 11 Apr 2015 20:47:41 GMT

[View Forum Message](#) <> [Reply to Message](#)

Yes, it is a good idea to do that.

Subject: Re: Merging PR onto a Combined IR, MR Dataset in SPSS

Posted by [Hassen](#) on Sun, 06 May 2018 10:06:42 GMT

[View Forum Message](#) <> [Reply to Message](#)

Thank you Bridgette,Tom Pullum,Trevor,Liz,I love you so much all!! I wish you have an endless happiness and success in your Life all!!
