Subject: Merging data files
Posted by priscambah on Wed, 25 Oct 2023 18:07:12 GMT
View Forum Message <> Reply to Message

Please i need help merging the women's file with their HIV status.
My research will assess contraceptive use in women living with HIV/AIDS in Cameroon using the 2018 data.

Please what is the unique identifier used to merge the data sets using SPSS?
This is my first time using the DHS data so any help is really appreciated.
Thank you

Subject: Re: Merging data files
Posted by Janet-DHS on Thu, 26 Oct 2023 21:01:48 GMT
View Forum Message <> Reply to Message

Following is a response from DHS staff member, Tom Pullum:

The lines below show how to do this merge in Stata. The DHS analysis team does not use SPSS, but the logic should be similar in SPSS.

* Stata lines to merge the IR and AR files for the Cameroon 2018 DHS survey

* Specify a workspace
cd e:\DHS\DHS_data\scratch

* Prepare the AR file
use "...CMAR71FL.DTA", clear
gen cluster=hivclust
gen hh=hivnumb
gen line=hivline
sort cluster hh line
save CMARtemp.dta, replace

* Prepare the IR file
use "...CMIR71FL.DTA", clear
gen cluster=v001
gen hh=v002
gen line=v003
sort cluster hh line

* Do the merge
merge 1:1 cluster hh line using CMARtemp.dta

* Cases that are in both files have _merge=3
tab _merge
keep if _merge=3

drop _merge

* Save the merged cases with a new file name

Subject: Re: Merging data files
Posted by priscambah on Wed, 08 Nov 2023 21:30:29 GMT
View Forum Message <> Reply to Message

Thank you for your reply.
we are trying to merge CMCR71FL, Household data with CMAR71FL which is the HIV data using SPSS.
Please what is the unique identifier to use to link these data sets?

Subject: Re: Merging data files
Posted by Janet-DHS on Thu, 16 Nov 2023 19:31:40 GMT
View Forum Message <> Reply to Message

Following is a response from DHS staff member, Tom Pullum:

As with the earlier merge, I will give the code in Stata. I hope you can follow the logic and convert to SPSS. Note that the AR file is used twice, once to merge with the men in the CR file and once to merge with the women. The CR file is sorted differently for each merge

* Stata lines to merge the CR and AR files for the Cameroon 2018 DHS survey
* Note that the cases in the CR file are couples
* The cases in the AR file are individuals, with sex not specified
* In the CR file, v001=mv001 and v002=mv002

* Specify a workspace
cd e:\DHS\DHS_data\scratch

* Prepare the AR file
use "...CMAR71FL.DTA" , clear

gen cluster=hivclust
gen hh=hivnumb

* w for woman, m for man
gen wline=hivline
gen mline=hivline
sort cluster hh wline
save CMARtemp.dta, replace

* Prepare the CR file

```
use "...CMCR71FL.DTA"
gen cluster=v001
gen hh=v002
gen wline= v003
gen mline=mv003
gen in_CR=1
sort cluster hh wline
save CMCRtemp.dta, replace

* Merge the women
merge 1:1 cluster hh wline using CMARtemp.dta
keep if in_CR==1
rename hiv* w_hiv*

* women who are in both files have _wmerge=3
rename _merge _wmerge

* Must re-sort both files
save CMCRtemp.dta, replace

use CMARtemp.dta, clear
sort cluster hh mline
save CMARtemp.dta, replace

* Go back to the CR file that has been merged for women
use CMCRtemp.dta
sort cluster hh mline

* Merge the men
* This is not a 1:1 merge because husbands can appear more than once
merge m:1 cluster hh mline using CMARtemp.dta
keep if in_CR==1
rename hiv* m_hiv*

rename _merge _mmerge
tab _wmerge _mmerge,m
tab _wmerge _mmerge if in_CR==1,m
tab w_hiv03 m_hiv03,m
drop in_CR *merge

label list HIV03
replace w_hiv03=10 if w_hiv03==.
replace m_hiv03=10 if m_hiv03==.
label define HIV03 10 "Not tested", modify
tab w_hiv03 m_hiv03,m

* Save the merged cases with a new file name
```