Subject: merge IR MR AND PR
Posted by SYLLA on Tue, 27 Feb 2024 04:28:31 GMT
View Forum Message <> Reply to Message

Hi,
I am working on the spatial distribution of hypertension in Côte d'Ivoire in both females and males aged 15-49. The variables i am interested are reported hypertension, physical activity, age , region and residence, smoking status, alcohol use, Bmi , wealth index, employement status and matrimonial status.
I have some of those variables in IR(female) MR(male) ans PR(household member)data set, so i realized that i need to merge them, but i'm having trouble doing that in stata.

I have tried according to the Guide_to_DHS_Statistics_DHS
but stata keep sending this error message:

variables v001 V002 v003 do not
uniquely identify observations in the using
data

Can you help me with the proper codes?

thanks

Subject: Re: merge IR MR AND PR
Posted by Bridgette-DHS on Tue, 27 Feb 2024 16:30:12 GMT
View Forum Message <> Reply to Message

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

The following Stata lines work for this merge (at least for me). I only include a few variables--you will want to revise to include more. You could put a line before "save MRtemp* such as "rename mv* v*" so the mv variables for men will be renamed as v variables. I have assumed that you are using the most recent CI survey. Hope this does what you want.

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

use "...CIIR81FL.DTA", clear
keep v0*
gen in_IR=1
tab1 in*
gen cluster=v001
gen hh=v002
gen line=v003
save IRtemp.dta, replace
```

```
use "...CIMR81FL.DTA", clear
keep mv0*
gen in_MR=1
tab1 in*
gen cluster=mv001
gen hh=mv002
gen line=mv003
save MRtemp.dta, replace

use "...CIPR81FL.DTA", clear
keep hv0* hvidx
gen in_PR=1
tab1 in*
gen cluster=hv001
gen hh=hv002
gen line=hvidx

* Merge PR with IR
merge 1:1 cluster hh line using IRtemp.dta
rename _merge merge_PR_IR

* Merge with MR
merge 1:1 cluster hh line using MRtemp.dta
rename _merge merge_PR_MR
save IRMRPR.dta, replace

keep if merge_PR_IR==3 | merge_PR_MR==3
tab in_IR in_MR,m
```

Subject: Re: merge IR MR AND PR
Posted by SYLLA on Wed, 28 Feb 2024 06:24:15 GMT
View Forum Message <> Reply to Message

Thank you so much for your help

Subject: Re: merge IR MR AND PR
Posted by tanvirpmc04 on Mon, 08 Apr 2024 19:56:54 GMT
View Forum Message <> Reply to Message

Hi
I am interested in merging IR and MR with PR, append them, then creating a common variable BMI from variables ha2, ha3, hb2 and hb3. I have used the following codes:

```
clear
set maxvar 100000

use " C:\Users\Hp\Desktop\datasets\nepal_dhs\NPIR82DT\NPIR82FL.DTA "
gen sex = 2
gen in_IR=1
tab1 in*
gen cluster=v001
gen hh=v002
gen line=v003
gen id=CASEID
sort cluster hh line id
save IRtemp.dta, replace

use " C:\Users\Hp\Desktop\datasets\nepal_dhs\NPMR82DT\NPMR82FL.DTA ", clear
gen sex = 1
gen in_MR=1
tab1 in*
gen cluster=mv001
gen hh=mv002
gen line=mv003
gen id=MCASEID
sort cluster hh line id
save MRtemp.dta, replace

use " C:\Users\Hp\Desktop\datasets\nepal_dhs\NPPR82DT\NPPR82FL.DTA "
gen in_PR=1
tab1 in*
gen cluster=hv001
gen hh=hv002
gen line=hvidx
gen id=HHID
sort cluster hh line id

* Merge PR with IR
merge 1:1 cluster hh line using IRtemp.dta
rename _merge merge_PR_IR

* Merge with MR
merge 1:1 cluster hh line using MRtemp.dta
rename _merge merge_PR_MR
save IRMRPR.dta, replace

keep if merge_PR_IR==3 | merge_PR_MR==3
tab in_IR in_MR,m
```

However, if i then use: tab hb2 or hb3, it returns no observation. Can you please help me to

identify the error?

Subject: Re: merge IR MR AND PR
Posted by Bridgette-DHS on Tue, 09 Apr 2024 11:35:10 GMT
View Forum Message <> Reply to Message

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

When you merge using cluster, hh, and line, you can ignore the string id codes. The id codes should be omitted from the "sort" lines. (Also, hhid does not include the line number, but caseid does,)

When you include "1:1" in the merge statement, you do not have to sort the files. You could omit the "sort" lines entirely. You only need to sort if you omit "1:1" from the merge statement, for an older version of the sort command. (I used to use the older version but no longer do so.)

The PR file already includes BMI. You do not have to construct it from height and weight (ha2, ha3, hb2, hb3). BMI is ha40 for women and hb40 for men. You could combine them into a single variable, call it hab40, for example, with these two lines:

```
gen     hab40=ha40 if hv104==2
replace hab40=hb40 if hv104==1
```

The IR file includes BMI as v445. The MR file should include BMI as mv445, but for some reason BMI does not appear in the MR file for this survey, with that name or any other. If you merge the MR and PR files, you can construct mv445 with "gen mv445=hb40".

Note that men were subsampled in this survey. Therefore, when you combine IR and MR data into a single file, you will need to weight up the men for any estimates of means, etc., for the combined population.

I cannot tell whether your primary goal is to have a single file of women and men that includes BMI, or something else. Please clarify, if you need more suggestions.

Subject: Re: merge IR MR AND PR
Posted by tanvirpmc04 on Fri, 12 Apr 2024 05:37:41 GMT
View Forum Message <> Reply to Message

Dear Tom,
Thank you for your quick response.

I intend to study hypertension which is in PR files, while some associated factors are in IR and MR files. So i plan to merge IR with PR, then merge MR with PR and then append resulting datasets.

Following your advice, i have used these codes:

```
clear
set maxvar 100000

use "  C:\Users\Hp\Desktop\datasets\nepal_dhs\NPIR82DT\NPIR82FL.DTA "
gen sex = "woman"
gen in_IR=1
tab1 in*
gen cluster=v001
gen hh=v002
gen line=v003
save IRtemp.dta, replace

use "  C:\Users\Hp\Desktop\datasets\nepal_dhs\NPMR82DT\NPMR82FL.DTA ", clear
gen sex = "man"
gen in_MR=1
tab1 in*
gen cluster=mv001
gen hh=mv002
gen line=mv003
save MRtemp.dta, replace

use "  C:\Users\Hp\Desktop\datasets\nepal_dhs\NPPR82DT\NPPR82FL.DTA "
gen in_PR=1
tab1 in*
gen cluster=hv001
gen hh=hv002
gen line=hvidx
gen hab40=ha40 if hv104==2
replace hab40=hb40 if hv104==1
gen bmi=hab40/100
gen bmic=1 if bmi<18.5
replace bmic=2 if bmi>=18.5 & bmi<25
replace bmic=3 if bmi>=25 & bmi<30
replace bmic=4 if bmi>=30 & bmi<50
label define bmic 1"Underweight" 2"Normal" 3"Overweight" 4"Obese"
label values bmic bmic
save PRtemp.dta, replace

* Merge PR with IR
use PRtemp.dta
merge 1:1 cluster hh line using IRtemp.dta
rename _merge merge_PR_IR
keep if merge_PR_IR==3
save IRPR.dta, replace
```

```
* Merge with MR
use PRtemp.dta
merge 1:1 cluster hh line using MRtemp.dta
rename _merge merge_PR_MR
keep if merge_PR_MR==3
save MRPR.dta, replace

* Append
use MRPR.dta
append using IRPR.dta
save IRMRPR.dta, replace
use IRMRPR.dta
tab sex
tab bmic
tab bmic sex
```

However, the outputs are as follows,

tab sex

```
       sex |      Freq.     Percent        Cum.
-----------+-----------------------------------
       man |      4,913       24.87       24.87
     woman |     14,845       75.13      100.00
-----------+-----------------------------------
     Total |     19,758      100.00
```

. tab bmic

```
      bmic |      Freq.     Percent        Cum.
-----------+-----------------------------------
Underweight |        993       13.51       13.51
    Normal |      4,471       60.81       74.32
 Overweight |      1,460       19.86       94.18
     Obese |        428        5.82      100.00
-----------+-----------------------------------
     Total |      7,352      100.00
```

. tab bmic sex

```
           |    sex
      bmic |     woman |     Total
-----------+-----------+----------
Underweight |       993 |       993
    Normal |     4,471 |     4,471
 Overweight |     1,460 |     1,460
     Obese |       428 |       428
-----------+-----------+----------
```

Total | 7,352 | 7,352

Can You tell me where I am doing it wrong? Thanks in advance.

Subject: Re: merge IR MR AND PR
Posted by Bridgette-DHS on Fri, 12 Apr 2024 12:38:55 GMT
View Forum Message <> Reply to Message

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

Your merge program was perfect. However, I see that at the end you did not have BMI scores, or hb40, for the merged men. This puzzled me too, until I went to the PR file, found hv027 ("household selected for man's interview") and entered the following two lines:

tab hv027, summarize(ha40)
tab hv027, summarize(hb40)

They show that height/weight were only obtained for women and men in households that were NOT selected for the men's interview. That is, if a household was selected for the men's interview, the heights and weights of adult women and men would not be measured. This meant that there were NO men who were (a) measured AND (b) interviewed.

DHS surveys often have similar subsampling. It is intended to reduce costs and also to help equalize the amount of time spent in each household. Inevitably, it has some impact on the analysis of covariation. Unfortunately, you will have to change your analysis plan to take this into account. For men, you cannot analyze BMI in relation to variables in the MR file.

Subject: Re: merge IR MR AND PR
Posted by tanvirpmc04 on Fri, 12 Apr 2024 17:45:59 GMT
View Forum Message <> Reply to Message

Thank you very much.
