Subject: v001 v002 v003 doesnot uniquely identify observation
Posted by kalfikirsisay@gmail.com on Mon, 13 Oct 2014 03:32:36 GMT
View Forum Message <> Reply to Message

Hi All,
I am using Ethiopian DHS2011 for my analysis. I need to merge birth recode file (using file) with HH member recode(master file). I tried to see the DHS merging guide and related topics in the FAQ. I was aware of the fact that HIDX is the household line number and used the following STATA command to merge,
gen gen v001= hv001
gen v002 = hv002
gen v003= hvidx
sort v001 v002 v003
merge m:1 v001 v002 v003 using "the birth recode file"
neverthless, I could'nt succed in merging and STAT propmted me the messgae " v001 v002 v003 doesnot uniquely identify the observation in using file". would you please quickly tip me on how to merge the two recodes?
Thanks in advance,
Kal

Subject: Re: v001 v002 v003 doesnot uniquely identify observation
Posted by Bridgette-DHS on Mon, 13 Oct 2014 12:47:33 GMT
View Forum Message <> Reply to Message

Dear User,

Please reference: post # 3011

Thanks - Bridgette-DHS

Subject: Re: v001 v002 v003 doesnot uniquely identify observation
Posted by kalfikirsisay@gmail.com on Mon, 13 Oct 2014 13:31:12 GMT
View Forum Message <> Reply to Message

Dear Bridgette,
thanks for your help, but still I tried both methods(gen long v001 = hv001,gen v002 = hv002,gen v003 = hvidx)and the rename approach, both turned out not to work and STATA gave me similar results as before"variables v001 v002 v003 do not uniquely identify observations in the using data". What is wrong with my case? your quick elaboration is highly appreciated,
Regards,
Kalkidan Sisay

Subject: Re: v001 v002 v003 doesnot uniquely identify observation

Posted by Trevor-DHS on Mon, 13 Oct 2014 14:22:44 GMT
View Forum Message <> Reply to Message

You should be using a 1:1 merge, not many to one - try that. If not, send your exact code you are using.

Subject: Re: v001 v002 v003 doesnot uniquely identify observation
Posted by kalfikirsisay@gmail.com on Mon, 13 Oct 2014 15:43:52 GMT
View Forum Message <> Reply to Message

dear Trevor,
thanks for quick response
I tried using 1:1 merge command only to get simialr message from STATA as before. I couldnot figure out why this is so, the following was the exact code I used.

```
use "C:\Users\mademsei\Desktop\CSA data and reports\Biomaker manual\merging files for bio\HH member recode\ETPR61FL.DTA", clear
gen v001 = hv001
gen v002 = hv002
gen v003= hvidx
sort v002 v002 v003
merge 1:1 v001 v002 v003 using "C:\Users\mademsei\Desktop\birth 1 to 1 merge.dta"
variables v001 v002 v003 do not uniquely identify observations in the using data
```

The master file I am using is HH member and the using file is birthrecode. Neverthless, I used the 1:m merge and get matched data albeit many observations(over 66000) were not matched. This is the table I obtained from 1:m merge command.

```
merge 1:m v001 v002 v003 using "C:\Users\mademsei\Desktop\birth 1 to 1 merge.dta"

    Result                      # of obs.
    -----------------------------------------
    not matched                    66,848
        from master                66,848  (_merge==1)
        from using                     0  (_merge==2)

    matched                        45,540  (_merge==3)
    -----------------------------------------
```

Is this the right way? I am a bit sceptic of the merging approach, I thought either m:1 or 1:1 to be proper, ofcourse I merged all observations in my using file and this what I need to appear in the maser file. your last comment on this is of prior value,Thanks in advance,
Kalkidan Sisay

Subject: Re: v001 v002 v003 doesnot uniquely identify observation
Posted by Trevor-DHS on Mon, 13 Oct 2014 16:07:31 GMT
View Forum Message <> Reply to Message

Your merge doesn't work as you have it because you are trying to match the births recode (BR) file to the PR file. For each respondent there are many births, so consequently the 1:1 merge doesn't work. If you tried merging using the IR file it would work. What are you trying to achieve? If you can explain what you want your final dataset to look like then I can tell you the best way of achieving that.

Subject: Re: v001 v002 v003 doesnot uniquely identify observation
Posted by kalfikirsisay@gmail.com on Mon, 13 Oct 2014 16:41:04 GMT
View Forum Message <> Reply to Message

Dear Trevor,
Many thanks for your extended helps,
My objective is to have a data set containing HIV test result, anemia test result ,other biomarkers(nutrition and vaccine) and other diseasses like diarrhoea prevalence and treatements (I also needed malaria but thanks to DHS experts, one expert told me the Eth. DHS has no data on it).I choose birth recode, hiv recode and HH memeber recode files as candisate files to merge them so that I can have one file contining the aforementioned inforamtion for my analysis (disease prevalence and treatements).
Thanks in advance
Kalkidan S.

Subject: Re: v001 v002 v003 doesnot uniquely identify observation
Posted by Trevor-DHS on Mon, 13 Oct 2014 17:44:49 GMT
View Forum Message <> Reply to Message

HIV test results are only available for women and men age 15-49. Diarrhoea prevalence and treatment, and nutrition and vaccine information are only available for children under 5. Anemia data are available for children under 5 and for women age 15-49 and for men age 15-49. I suggest constructing separate datasets for your analysis for children and for women and men. I also recommend thoroughly reviewing the questionnaires first to understand who has been asked which set of questions before deciding how to structure your data.

Subject: Re: v001 v002 v003 doesnot uniquely identify observation
Posted by kalfikirsisay@gmail.com on Mon, 13 Oct 2014 19:42:08 GMT
View Forum Message <> Reply to Message

Dear Trevor,
I thank you very much for your helpfull advices, I will structure the data and do the anlaysis separately.
Regards,
Kalkidan S.
