Subject: Merging SPA datasets- Nepal SPA 2015
Posted by rkchettri on Mon, 05 Feb 2018 21:59:18 GMT
View Forum Message <> Reply to Message

Dear DHS experts,

I am trying to merge health facility dataset and provider dataset of Nepal SPA 2015. My unit of analysis is health facilities,so I have to consider health facility audit datasest as master data.
I used the command _merge m:1 facil using provider data, but I got the message -variable facil does not uniquely identify observations in the using data. Here 1:m command worked but it considered the provider as master dataset.
While I considered provider dataset as master dataset and used same command, it worked but I found many changes in the variables responses in health facility datasets.

So I would like to request how to solve this problem?

Best regards,
RK

Subject: Re: Merging SPA datasets- Nepal SPA 2015
Posted by Bridgette-DHS on Wed, 07 Feb 2018 12:17:07 GMT
View Forum Message <> Reply to Message

Following is a response from DHS Senior Research Associate, Wenjuan Wang:

I guess you'd like to borrow some provider-level information into the facility file. If this is the case, we usually collapse provider variables to the facility level and then merge with the facility file and it is going to be a one-to-one merging. For example, you'd like a variable on % of providers at the facility who were trained in breastfeeding. Since the training information is stored in the provider file, you can first collapse this variable to the facility level and then merge with the facility file. The Stata commands are:

```
use NPPV71FLSP.DTA, clear
g train_bf = 0
replace train_bf = 1 if p513_2==1

collapse (mean) per_trbf=train_bf [pw=provwt], by(facil)
sort facil
save train_bf.DTA, replace

use NPFC71FLSP.DTA, clear
merge 1:1 facil using train_bf.DTA
```

Please try and let me know if this works.

Subject: Re: Merging SPA datasets- Nepal SPA 2015
Posted by rkchettri on Thu, 08 Feb 2018 01:33:37 GMT
View Forum Message <> Reply to Message

Thank you, Wang, for ideas and command. I want to do same what you mentioned.
I tried with your code and it worked for a single variable at a time. When I tried to recode and run but it did not work. Actually, I want to recode and run all ANC/ delivery training variables and want to merge with health facility dataset.

Here is the code that I worked to identify collapse variables of two ANC training, these codes work when I run one at a time but not working when I run together. Is any ideas to collapse all training related variables at a single entry?. I have 20 training variables that I want to recode per facility.

```
 gen tp502rec_1=0
replace tp502rec_1 = 1 if p502_1==1
collapse (mean) per_tp502rec_1=tp502rec_1 [pw=provwt], by(facil)
recode per_tp502rec_1 0=0 .=0  0.00001/max=1, gen( p502rec_1)
codebook p502rec_1

gen tp502rec_2=0
replace tp502rec_2 = 1 if p502_2==1
collapse (mean) per_tp502rec_2=tp502rec_2 [pw=provwt], by(facil)
recode per_tp502rec_2 0=0 .=0  0.00001/max=1, gen( p502rec_2)
codebook p502rec_2
```

Subject: Re: Merging SPA datasets- Nepal SPA 2015
Posted by Bridgette-DHS on Thu, 08 Feb 2018 14:43:47 GMT
View Forum Message <> Reply to Message

Here is another response from Wenjuan:

Yes, you can collapse multiple variables at a time, just add them all in the command. For example, if you need to collapse x1, x2, x3 at a time, do this following:

```
collapse (mean) per_x1=x1 per_x2=x2 per_x3=x3 [pw=provwt], by(facil)
```

Hope this works.

Subject: Re: Merging SPA datasets- Nepal SPA 2015
Posted by rkchettri on Thu, 08 Feb 2018 20:00:06 GMT
View Forum Message <> Reply to Message

Thank you,Wenjuan.Now, it worked, problem solved.