Subject: Linking dataset in SPA_Ethiopia
Posted by gizachew on Mon, 07 Mar 2016 02:23:25 GMT
View Forum Message <> Reply to Message

Dear Liz,

I am currently working on Ethiopian SPA 2014 data after getting an official access for the FP related data from EPHI (there has been a mutual agreement between EPHI and me). I am planning to link the three datasets (facility inventory, Exit interview/Observation (already linked, and provider interview). Unfortunately, it appears that they have interviewed a random provider in each facility that makes linking the provider and the client difficult. However, it is possible to link the facility with the client and the provider. Would you please assist me in linking through using Stata 13

Best
Gizachew

Subject: Re: Linking dataset in SPA_Ethiopia
Posted by Liz-DHS on Sun, 20 Mar 2016 02:54:53 GMT
View Forum Message <> Reply to Message

Dear User,
I have submitted your query to one of our experts. However, our programs are all written in CSPro. We will post as soon as we are able to provide some guidance.
Thank you!

Subject: Re: Linking dataset in SPA_Ethiopia
Posted by Liz-DHS on Mon, 21 Mar 2016 15:05:35 GMT
View Forum Message <> Reply to Message

Dear User,
Although we do not have access or distribute the Ethiopia SPA data, here is some guidance from our technical expert Dr. Wenjuan Wang regarding SPA data in general:
Quote:
It is always possible to link between facility inventory data, provider data and client (exit and observation) data using facility id and provider id. Facility id variable should be included in facility, provider and client data files and provider id should be included in provider and client data files.

Let's say if you are interested in providers and want to merge some facility variables into the provide file. This is going to be a many-to-one merging because multiple providers are associated with one facility
See below the Stata program for this type of merging.

use providerdata.dta, clear /* you need to use the path where your data files are stored*/

merge m:1 facilityid using facilitydata.dta /* variable "facilityid" is the facility identifier(it may be a different name in ESPA files); it has to be the same variable name in facility and provider data. Otherwise you need to rename it before merging*/
drop if _merge!=3 /*drop unmatched cases*/
save mergedfile.dta, replace /*save merged file */

The program for merging facility and client data would be similar as above, just replace the provider dataset with the client dataset.

If you are interested to merge some provider variables into the client file, the identifiers would be two variables: facilityid and providerid

use clientfile.dta, clear
merge m:1 facilityid providerid using providerdata.dta /*as above, both variables should have the same name in provider and client data, otherwise, rename them before merging */
drop if _merge!=3
save mergedfile.dta, replace

Let us know if you have other questions.

Subject: Re: Linking dataset in SPA_Ethiopia
Posted by gizachew on Fri, 25 Mar 2016 04:57:12 GMT
View Forum Message <> Reply to Message

Thanks so much for Liz and Dr Wenjuan

It seems that i don't have providerid for provider data and clientdata. However, i do have facilityid in all three (facility data, provider data, client data).
Hence I am planning do merge provider data into facility data. Hence does that mean that i am merging 1 to many (1:m)

Moreover, I got different findings for some findings. For instance, the nth value for facility type is different while doing on the original document and on the merged file. can you assist please

best,
Gizachew

Subject: Re: Linking dataset in SPA_Ethiopia
Posted by Liz-DHS on Fri, 25 Mar 2016 14:50:41 GMT
View Forum Message <> Reply to Message

Dr. Wang responded:

Quote:
1) It is strange that provider id variable is not included in provider data and client data. It may have a different name and usually is labeled as "provider serial number".

2) yes, if you are merging provider data into facility data, you are dealing with one to many merging. In this cases, facility dataset is the master file and provider dataset is the using file.

3) Without access to the ESPA data, it is difficult to figure out the problem.
Please note when you merge provider data into facility data, the number of cases(facilities) in your master file (facility dataset) will proliferate because multiple provider records are associated with one facility. Maybe this causes the nth value no longer the same?

Subject: Re: Linking dataset in SPA_Ethiopia
Posted by gizachew on Sat, 26 Mar 2016 00:51:40 GMT
View Forum Message <> Reply to Message

Thanks again for the notes,

I will double check the presence of provider serial number data in the client data file (I got it in the provider data file)

Moreover, I have now taken the impression that the descriptive statistics to characterize the facilities and (when necessary the providers/clients) need to be done before merging Does my understanding sounds correct?

Good day

Subject: Re: Linking dataset in SPA_Ethiopia
Posted by gizachew on Sat, 26 Mar 2016 01:57:55 GMT
View Forum Message <> Reply to Message

Thanks again for the notes,

I will double check the presence of provider serial number data in the client data file (I got it in the provider data file)

Moreover, I have now taken the impression that the descriptive statistics to characterize the facilities and (when necessary the providers/clients) need to be done before merging Does my understanding sounds correct?

While doing merging all the three, is it okay to do a step by step merging. This means first merge facility and provider (i.e facprov) and then to the client (facprov Vs client) as the final intention is to conduct analysis at the client level

I have also seen that the weighting variable is available in all the three datasets. However, I am unsure which of the weighing variable is important after merging (my expectation was that all the three weighting variables maybe similar). In this regard, is it okay to use gen wt = wtvar/1000000

I highly appreciate your assistance

Subject: Re: Linking dataset in SPA_Ethiopia
Posted by Liz-DHS on Wed, 30 Mar 2016 13:24:58 GMT
View Forum Message <> Reply to Message

Response from Dr. Wenjuan Wang:
Quote:
Yes, your understanding is correct!
Regarding the merging: You can only merge two files at a time.
If the final unit of the analysis is clients, I would merge the client file with the provider file first (with client file as the master file), then use this merged file as the new master file to merge with the facility file.
Regarding the weights: Which weight to use depends on the unit of analysis. If you analyze facilities, for example, to look at facilities characteristics, you would use facility weight. When analyzing clients, you would use the client weight. They are not the same.

Subject: Re: Linking dataset in SPA_Ethiopia
Posted by gizachew on Mon, 04 Apr 2016 04:39:09 GMT
View Forum Message <> Reply to Message

Thanks so much! I found your assistance for crucial.

Cheers,
Gizachew