Subject: How to combine data from different countries for a period of years?
Posted by kinsukmanisinha@gmail.com on Mon, 02 Mar 2015 15:50:09 GMT

Dear all,

I am using the DHS database for the first time and also I have never worked with survey data before. My aim is a bit broad, it is to understand which variables have an impact on child mortality/under five mortality.

In order to achieve my aim I need to combine datasets from different countries for different year and build a pooled database.

And, this is the first problem I face -- I don't know how to combine databases from different countries for different year. Do I need to add a weight? If yes, how do I calculate this weight and how do I take it into account? Or take another factor into account.

I do understand that my question is a bit naive but if anyone could point me in the right direction, that would be simply great..!!!!

Many thanks..!!!

Regards
Kinsuk

Subject: Re: How to combine data from different countries for a period of years?
Posted by Trevor-DHS on Thu, 05 Mar 2015 21:00:02 GMT

1) If you are using DHS data for the first time and have never used survey data before, then you are probabbly taking on too much. I suggest learning to use a single survey dataset first and being able to reproduce published results before you try to do anything as complicated as you are suggesting.
2) It is not clear that you need to combine datasets based on the aim you have stated. You can analyze each dataset individually.
3) You should know that each individual dataset includes its own weight so you should be weighting data even if you are only analyzing one survey.
4) If you do need to pool data, you will need to denormalize the weights and then use the denormalized weights. There are several posts on the forum related to this.
5) Once you have the denormalized weights for your pooled data then you would apply weights just as you do for any other analysis, depending on the software you are using.

Subject: Re: How to combine data from different countries for a period of years?
Posted by kinsukmanisinha@gmail.com on Tue, 10 Mar 2015 12:52:00 GMT

Many Thanks for the detailed reply.

I agree with you, I should start with one data set first.

I read your points and I have a question:-

Point 3 says:-
3) You should know that each individual data set includes its own weight so you should be weighting data even if you are only analyzing one survey.

So, for example if I use only Egypt for year xxxx, I don't think I need to use weights. But if I use Egypt and Sudan for the same year, same survey, then I need to worry about weight. Is this what you meant?

Thanks once again..!!!

Subject: Re: How to combine data from different countries for a period of years?
Posted by Trevor-DHS on Mon, 16 Mar 2015 23:46:44 GMT
View Forum Message <> Reply to Message

If you are using only Egpyt xxxx survey you should be using weights. The reason for this is that the probability of selection of households as used in the sample design varies for different parts of the country. To correct for this it is necessary to use weights. You will find the relative weights in one of the following variables:
v005 - for women and children's datasets
hv005 - for household and household member datasets
mv005 - for men's datasets

If you are combining surveys for, say, Sudan and Egypt, then you have to not only use the weights, but also denormalize the weights (described in other posts on the forum) before pooling the data.

Subject: Re: How to combine data from different countries for a period of years?
Posted by kinsukmanisinha@gmail.com on Tue, 17 Mar 2015 10:50:19 GMT
View Forum Message <> Reply to Message

Thanks once again..!!!