
Subject: Use of weight on subsetting data
Posted by [FLW](#) on Thu, 15 Aug 2019 18:14:18 GMT
[View Forum Message](#) <> [Reply to Message](#)

Based on my understanding, the sampling weight in DHS is provided based on the entire survey sample.

If I would like to study only a subset of population i.e. men with children.

Can the original sampling weight be applied?

Will the applied weight ensure national representativeness of subsetting data?

Or I will need to calculate the new sampling weights based on the subsetting data?

Hope someone can guide me for this.

Thanks :)

Subject: Re: Use of weight on subsetting data
Posted by [schoumaker](#) on Thu, 15 Aug 2019 19:38:12 GMT
[View Forum Message](#) <> [Reply to Message](#)

Yes, you can use the weights, but their sum should equal to the sample size for your subset. So you may have to rescale them. This is automatic with some commands and/or some software packages, but not all.

Best,
Bruno

Subject: Re: Use of weight on subsetting data
Posted by [FLW](#) on Fri, 16 Aug 2019 06:52:48 GMT
[View Forum Message](#) <> [Reply to Message](#)

I have seen this in the internet regarding subsetting data.

<https://stylizeddata.com/how-to-use-survey-weights-in-r/>

is it appropriate to do so?

Use svydesign in R to declare surveydata and weights, then only i subset the data.

Will it fix the rescaling issue?

Subject: Re: Use of weight on subsetting data
Posted by [Bridgette-DHS](#) on Tue, 03 Sep 2019 18:40:45 GMT
[View Forum Message](#) <> [Reply to Message](#)

Following is a response from DHS Research & Data Analysis Director, Tom Pullum:

Are you asking whether it is appropriate to use weights with DHS data (it is!) or are you asking how to use weights with R? The link should provide all the information you need about the latter.

Subject: Re: Use of weight on subsetting data
Posted by [LizavG](#) on Sun, 10 Nov 2019 20:52:41 GMT
[View Forum Message](#) <> [Reply to Message](#)

This is a really relevant question and I would like to extend it. I am also sub-setting data based on the availability of child measures and ages. I would like to rescale the data because I am working with STATA commands that do not allow for svy.

I know that v005 are the weights usually used for weighing in DHS but I am unsure how it is constructed, and thus how I could rescale the data. What factors are used to create v005?

Subject: Re: Use of weight on subsetting data
Posted by [Bridgette-DHS](#) on Mon, 18 Nov 2019 12:33:47 GMT
[View Forum Message](#) <> [Reply to Message](#)

Following is a response from DHS Research & Data Analysis Director, Tom Pullum:

There are some Stata commands that do not allow svyset. For some of those commands, however, you CAN use weights, or weights and clusters, so long as you don't include strata, which seem to be the most complicated part of the svyset adjustments. You can use the weights with, for example, "[pweight=v005]" or "[iweight=v005/1000000]". Those inserts always go before the comma in the estimation command. You can try to use the cluster adjustment by putting "cluster(v021)", for example, with the options AFTER the comma. That may work even if svyset is not accepted.

If you are using pweight, which is the type of weight that is required in svyset, you do not need to re-scale v005 in any way. In fact, with pweights you CANNOT rescale v005. Whatever variable you enter as a pweight will always be rescaled so the values add to 1. For that reason, for example, "[pweight=v005]" and "[pweight=v005/1000000]" will always give you the same results. Many users go through the step of defining the weight to be v005/1000000 but that's not necessary (for the pweight option). Hope this answers your question.

Subject: Re: Use of weight on subsetting data
Posted by [Wahyu dh](#) on Mon, 22 Jun 2020 23:24:00 GMT
[View Forum Message](#) <> [Reply to Message](#)

I also have a question about this. I analyze the institutional delivery in rural area and i only use the weight v005/1000000 is it right? Or i need to do other adjustment? If yes, what should i do? And i

also want to ask. The percentages of institutional delivery which i got is different from the publication of country dhs data (percentage of institutional delivery in rural), any explanation toward this? Do i use the wrong weighting or other -.-
I use spss software.

Subject: Re: Use of weight on subsetting data

Posted by [Bridgette-DHS](#) on Tue, 30 Jun 2020 20:50:15 GMT

[View Forum Message](#) <> [Reply to Message](#)

Following is a response from DHS Research & Data Analysis Director, Tom Pullum:

I am not familiar with using weights in SPSS. I have reviewed the earlier posts in this thread and don't have anything to add. If you are having trouble matching something, please tell us which survey you are using, the number of the table in the main report, and the number you cannot match (a percentage or a frequency).
