
Subject: Re: weighting data for a subset of main data file
Posted by [Bridgette-DHS](#) on Mon, 28 Mar 2016 14:28:05 GMT
[View Forum Message](#) <> [Reply to Message](#)

Following is a response from Senior DHS Stata Specialist, Tom Pullum:

Some of the recommendations about weighting in the DHS Guide to Statistics are misleading or out of date and will be modified in the next version. You should use the sample weights no matter what subset of the data you are using. Otherwise your estimates will be biased and will over-represent the subpopulations that were over-sampled and under-represent the subpopulations that were under-sampled.

You do not have to worry about the normalization of the weights. If you are using Stata, with the "pweight" option (as in svyset) then the weights are automatically re-normalized so that the total weight equals the number of actual cases in the analysis. (The other three weight options in Stata do NOT re-normalize.) Other packages such as SPSS may also re-normalize the weights, although I cannot say for sure. If you have any doubt whether the weights are automatically re-normalized within the package and procedure you are using, you can run something with v005, which includes a factor of 1,000,000, and see whether the factor of 1,000,000 has been removed. If it has, then the weights were automatically re-normalized.
