Subject: Re: Pooled datasets - weighting data problem
Posted by Bridgette-DHS on Thu, 17 Apr 2014 18:14:47 GMT
View Forum Message <> Reply to Message

Following are comments by our Senior Sampling Specialist, Ruilin Ren:
Quote:Now I read the one-pager by Ruilin Ren but I still don't know where to find "total number of females age 15-49 in the whole country at the time of the survey".
Comment: You need to find this piece of information from other sources, such as the country's statistical office or census office, or from the UN's World Population Prospects: http://esa.un.org/wpp/unpp/panel_indicators.htm

Quote:I guess the only way is to go and check all the censuses that have been used for my DHS surveys at hand and look for that specific number there? Is there a database at DHS maybe where all that information is stored already?

Comment: DHS does not store the population information for any country.

Quote:...now that I have them pooled and want to specify my svyset, do I just use the weight variable as it is, e.g. "svyset psu [pweight=v005], strata(strat_id)" ?
Or do I have to make any other adjustments?
Comment: The way you pool the data together is fine. You do not need to rename the weight variable, but you do need some treatment for the PSU/cluster (HV001 or V001 or MV001) and the stratification variables (HV022, V022, MV022, or HV024, V024, MV024 if you use region cross urban/rural for stratification). The idea is that each PSU/cluster and each stratum from each survey must stand alone. If you do not treat them properly, for example, the software cannot distinguish cluster 001 (HV001=001) from survey one and survey two, the system will simply merge them as one cluster. This is not true. The easiest way to do is to add 10000 to HV001 from survey one, add 20000 to HV001 from survey two. So cluster 10001 and 20001 will be treated as two different PSUs/ clusters. Similarly you need to treat HV022, or HV024 if you use HV024 cross HV025 for stratification.