Subject: Re: Merging Individual and HH Member Recode Files
Posted by Bridgette-DHS on Wed, 27 Jul 2016 21:12:06 GMT
View Forum Message <> Reply to Message

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

You are probably working with successive rounds of the Peru CS, and the cluster numbers, in particular, are being re-used for different rounds. That is, cluster 1, say, does not refer to the same cluster for all rounds. That would be a major problem. The households are numbered within clusters and the line numbers are within household, so the problem is mainly with the cluster id codes, which are being re-cycled. Each round has different cases in it. The Peru CS is not longitudinal; it is successive cross-sections.

The safest thing to do would be to do these merges within each round. Add a variable which is "round" or something like that. After you have done all the merges within rounds, append them to make one long file. Then "round" will be part of the id information. If you use svyset you will have to construct unique identifiers for the strata, with something like "egen strata=group(round v023)". You may want to re-normalize the weights, as described in several other posts. But I recommend that you only use the pooled file for calculating differences and changes across rounds. If you try to calculate a mean, say, for the full file, it will not have a clear reference date and will be confusing.