Subject: Re: Sample weights and stratification - Nigeria 2008 and 2018
Posted by Bridgette-DHS on Wed, 20 May 2020 21:44:49 GMT
View Forum Message <> Reply to Message

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

Mostly I agree, but with these differences. First, the actual clusters are not the same in successive surveys. Whatever the original codes are for enumeration areas in the sampling frame, they are replaced with new numbers. If another researcher did not construct new distinct id codes, they made a mistake (although the impact may be small).

If you are using Stata and svyset, then yes, you should make the adjustments for weights, clusters, and strata. All three. It's painless. But include the "singleunit" option.

Yes, v023 does not always give the strata. Usually, since 2013 or so, but not necessarily earlier. The DHS website gives the correct stratum variable for every survey. I believe that for both of these surveys, the strata are the combinations of state and urban/rural, v025. I believe state is given by shstate. I recommend the following lines:

* In the 2008 survey
egen stratum_ID_2008=group(shstate v025)
gen tempvar=stratum_ID_2008

* In the 2018 survey
gen stratum_ID_2018=v023
gen tempvar=stratum_ID_2018

* Append and construct "survey" using v007

* In the combined file
egen stratum_ID=group(tempvar survey)
drop tempvar

egen cluster_ID=group(v001 survey)

svyset cluster_ID [pweight=v005], strata(stratum_ID) singleunit(centered)

Let us know if you have other questions or this is not clear.