Subject: Re: Stratification in pooled datasets - Tanzania Posted by Bridgette-DHS on Tue, 19 Sep 2017 18:01:34 GMT

View Forum Message <> Reply to Message

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

For all the surveys, you will be safe if you use the combinations of region and urban/rural as the strata. You need a line such as this: egen stratumid=group(v024 v025), and then put stratumid as the stratum variable in svyset.

In svyset the type of weight is always a pweight. You can easily verify that multiplying the weights by a constant, any constant, will have no effect whatever on the results. You can define the weight as v005 or as v005/1000000, for example, and the results will be exactly the same. That's because Stata always normalizes pweights to have an average value of 1.

What you may be thinking of is scaling v005 up to be fweight expansion weights, which is only useful if you are trying to estimate, for example, the number of children that are stunted, using svy and tabulate. Another consideration if you do that is that fweights must be integers. DHS does not recommend this because there is an implication of more accuracy than is justified.