Subject: Re: Weighting variables in DHS India data (1992 and 1998)
Posted by user_rm on Fri, 14 Aug 2015 18:07:20 GMT

Many thanks for your reply.

I am using the individual level data (ie for kids) for 1992 and 1998 for rural households. In stata, I have the 1998 data appended below the 1992 data for the same set of variables (ie it's a pooled cross-section analysis).

I wanted to specify the survey design in stata to make the data nationally representative. The NFHS manual states that the survey is a two stage stratified sample design for rural areas. The first level of stratification consists of identifying districts within each state. Districts were further sub-divided into regions and households for villages were identified for each region. In terms of specifying the sampling design in stata, do I need to specify the cluster number (v001) and household number (v002)? The weights used would be sWeight = v005/100000. Is the code below correct for this purpose? Or am I missing something here?

svyset v001 [aweight=sWeight]|| v002 [aweight=sWeight]

Please clarify.

Many thanks
R