Subject: NFHS - 5
Posted by neera.vohra@competitivene on Sun, 15 May 2022 06:12:58 GMT
View Forum Message <> Reply to Message

Hello,
 I wanted to extract district-wise data on the percentage of women who have a bank account from the Household members (or Persons) Recode state file. I am not sure whether I have set the syntax right; I need clarity on the same :

gen pwt=.
replace pwt= hv005/1000000
svyset hv021 [pw=pwt], strata(hv023)

Since the NFHS-5 sample is a stratified two-stage sample. therefore shouldn't the syntax look like :
svyset hv021 [pw=pwt], strata(hv023) || hhid , strata(hv001)

My understanding is that in the second stage: about 25-30 households are looked after, identifying the villages (in the first stage)

However, I have no experience working on NFHS data sets, so I am not sure what's the right thing to do. Kindly help me and respond to it as soon as anyone can.

Thank you for your help.

Subject: Re: NFHS - 5
Posted by Janet-DHS on Thu, 19 May 2022 14:30:20 GMT
View Forum Message <> Reply to Message

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

First, for pweights you do not need to divide hv005 by 1000000. Many people do this, but it is not necessary because Stata automatically re-scales pweights to have a mean of 1.

Second, you do not need a multi-level version of svyset. Your first version of svyset is sufficient.

Third, the clusters and strata were numbered separately in each state. You need commands such as "egen cluster_id=group(v021 v024)" and "egen stratum_id=group(v023 v024)".

Then "svyset cluster_id [pw=hv005], strata(stratum_id)" should be correct for svyset.