I am working with the NFHS-3 survey KR dataset in Stata, studying the determinants of child health at the all-India level, and have a doubt about the appropriate stratification variable for the survey declaration. Looking at various answers on the Forum, I see recommendations by experts for both

1. v022, "sample stratum number" (73 distinct values)

e.g, by Trevor-DHS here

   http://userforum.dhsprogram.com/index.php?t=msg&goto=326 8&S=c572d650be37c4375bff4d100ab9ced0&srch=NFHS+strat a#msg_3268

2. a grouping of v024, "state" (29 values) with v025, "type of place of residence" (2 values), i.e.,

 egen float stratum = group(v024 v025), resulting in 58 distinct values,

e.g., from Ruilin Ren here

   http://userforum.dhsprogram.com/index.php?t=msg&goto=213 4&S=c572d650be37c4375bff4d100ab9ced0&srch=NFHS+strat a#msg_2134

I was using option 2 so far, but revisiting this issue, I checked the NFHS publication

'National Family Health Survey (NFHS-3), 2005-06: India: Volume I', section on Sample Design, p61, (copy uploaded)

and it seems the stratification scheme is more complex than just the simple grouping of State & place of residence.

I would really appreciate a definitive answer to this from one of the experts familiar with the NFHS data- I think this would serve many researchers now & in future as well.

File Attachments
1) 2SamplingDesign_from_NFHS_FRIND3-Vol1AndVol2 copy.pdf, downloaded 556 times