

---

Subject: Re: Weighting data after merging survey rounds with different levels of representation

Posted by [Bridgette-DHS](#) on Fri, 13 Jan 2017 16:45:52 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

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

Sorry about that mistake--"table clean" is an option with "list" and not with "tab". Don't know why I added that--it definitely would not run.

You definitely need something like "egen clusters=group(survey v021)" to get unique identifiers for the clusters across the three surveys and the strata across the three surveys.

The change you observe in the number of strata from one survey to the next is not implausible (although usually the definitions are the same from one survey to the next.) In general, the number of strata is in the range of 20 to 60. I suggest you look at the report. I HOPE it will confirm what I passed on to you.

DHS estimates at the stratum level are always representative, in terms of being unbiased. It is true that some older documentation mentioned a lack of representativeness, but that actually refers just to higher standard errors when there are fewer cases. That can be an issue, but bias is NOT an issue. The sampling is designed so that small strata tend to be over-sampled (large strata correspondingly tend to be under-sampled) in order to get more stable estimates.

The generic term we use for the first national sub-division is "region". The generic term for the second level is "district". In general the strata are the combinations of region and urban/rural.

---