
Subject: Variable for weighting
Posted by [ICEH](#) on Mon, 02 Apr 2018 12:41:13 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi all,

I would appreciate some clarification about the strata used for weighting in the surveys. As far as I understand, we are supposed to use in most of the cases the domains stratified by urban/rural. This seems to be represented by the variable v023. Although, the variable v022 seems to be the most appropriate in this method as it is named "Sample strata for sampling errors". In several cases, v022 and v023 are similar. But in many cases the variable v022 seems to have too many stratas, going over 100, even 200.

Given such scenario, I have the following questions:

- What is the appropriate variable for weighting (svyset v021 [pw=v005], strata(?))
- If it is v022, why does some surveys have so many strata levels, such as Rwanda 2010, Albania 2008, Phillipines 2008?

I hope I was able to make myself clear and I apologize if I have a misunderstanding of the situation.

Thanks in advance,

ICEH

Subject: Re: Variable for weighting
Posted by [Bridgette-DHS](#) on Mon, 02 Apr 2018 13:39:21 GMT
[View Forum Message](#) <> [Reply to Message](#)

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

The correct variable is usually v022, but you are right that it is often the same as v023 and also right to be suspicious if there are, say, 200 different values for v022. Usually the strata are the combinations of v024 and v025, i.e. what you would get from "egen strata=group(v024 v025)". If in doubt, check the report on the survey. Chapter 1 of the report should have a section called "Sample Design" that will define the strata.

Subject: Re: Variable for weighting
Posted by [ICEH](#) on Mon, 02 Apr 2018 13:45:10 GMT
[View Forum Message](#) <> [Reply to Message](#)

Thank you so much for that information.
