Subject: Strata, PSU, weights for Sengal Continuous Survey
Posted by LeahBevis on Sat, 22 Aug 2020 02:45:19 GMT
View Forum Message <> Reply to Message

I am working with the continuous DHS survey from Senegal, 2013-2018. We downloaded each dataset individually and then appended them all. I can see from 2 links below that 2013 and 2014 were sampled from the 2002 census, and 2015 and 2016 from the 2013 census. I'm not sure about the sampling frame for 2017 or 2018. Overall, in order to properly employ weights, I need to know more about the sampling design, but the documents are in French and I don't speak French.

So, I want to know 3 things:

(1) Do weights reflect population change, such that they can be used without adjustment?

(2a) Are PSU always given by v021?
(2b) Which PSU were unique? That is, in which years were PSU locations new? PSU values do NOT repeat across 2013 and 2014, but they DO repeat between 2015 and 2016. They also repeat between 2015/16 and 2017. At the moment, I am assuming that *every* year new PSU were drawn (even if IDs are repeated; i'm assuming that's an "accident"), but it also seems possible that the 2016 survey actually re-visited the 2015 PSUs. And I have no idea about 2017 and 2018.

(3a) Are sampling design strata always given by v023?
(3b) One of the links below explains that the definition of "rural/urban" changed in 2015-16, from the previous 2013-14 definition. So, this means the strata changed, right? Yet the 24 values of v023 do not change in 2015 onwards. I'd like to know in what years the strata truly changed, so that I can create a "sup-strata" variable for weighting (using svyset in Stata or Procsurveyfreqwith in SAS), where strata IDs change when the actual sampling strata change.

Thank you!
Leah

Links:
 https://userforum.dhsprogram.com/index.php?t=msg&goto=16 347&&srch=senegal#msg_16347
https://dhsprogram.com/pubs/pdf/DM34/DM34.pdf

Subject: Re: Strata, PSU, weights for Sengal Continuous Survey
Posted by Bridgette-DHS on Mon, 24 Aug 2020 21:04:26 GMT
View Forum Message <> Reply to Message

Following is a response from DHS Lead Statistician, Ruilin Ren:

I do not know your purpose of study, and why it is necessary to put all the data together. When you put the different surveys together, you need to adjust the survey weights, even if the surveys were from the same country, because the sampling fraction, the response rates vary from survey to survey and the sampling weights are normalized weights. See below for detailed explanations.

1. Firstly, the 2013-2014 survey was based on 2002 census. 2015-2016 surveys were based on 2013 census. The 2017 survey was based on an updated frame from 2013 census. We have combined data for 2013 and 2014, 2015 and 2016, it is better to use the combined data because the combined weights are calculated and provided in the data set. The 2017 survey was a stand alone survey. These different survey cycles differ on the urban-rural definition.
2. The weights must be adjusted if combining datasets from different surveys. You may find the weight adjustment topics on the DHS data user forum because this is a frequently asked question.
3. The PSU number is unique within survey cycle, but not unique across survey cycles, see my point #1 for the survey cycles.
4. As for sampling stratum, it can be coded in v022 or v023, but you can always use v024 cross v025 to have the sampling stratum
5. The urban-rural changes from different survey circles , they differ on the ground, some places were defined as rural in the previous survey and changed to urban in a subsequent survey, but the number of codes stay the same because it is urban-rural crosses regions. The number of stratum codes does not reflect the urban-rural definition change.
6. The number of regions (14) in Senegal have not been changed since 2013, so the number of stratum should be the same from 2013, it should have 28 codes because it is region (14) crosses urban-rural (2).