Subject: Household crowding
Posted by j_lewis98 on Mon, 17 Aug 2020 12:28:57 GMT
View Forum Message <> Reply to Message

I am using the SADHS 2016 dataset. I want to create a household crowding variable from the PR file. What existing variables could I use for this; and what would the STATA code be?

Many thanks

Subject: Re: Household crowding
Posted by Bridgette-DHS on Mon, 17 Aug 2020 14:40:00 GMT
View Forum Message <> Reply to Message

Following is a response from DHS Senior Analysis & Research Manager, Shireen Assaf:

Dear user,
The following Stata code can be used to compute the crowding index from the HR file. You can then recode this index to a binary variable if you like as I have shown with the variable "crowded".
Thank you.
Best,
Shireen

Stata code:

```
* calculating crowding index
gen hhusual =hv012
replace hhusual=hv013 if hhusual==0
gen crowd=.
replace crowd=trunc(hhusual/hv216) if hv216>0
replace crowd=hhusual if hv216==0
replace crowd=. if hv216>=99
label var crowd "crowding index"

recode crowd (0/2=0 "No") (3/max=1 "Yes"), gen(crowded)
label var crowded "Three or more people living in the same room"
```

Subject: Re: Household crowding
Posted by j_lewis98 on Mon, 17 Aug 2020 18:03:50 GMT
View Forum Message <> Reply to Message

Many thanks for your reply- this was very helpful for me!

I have another question, if you may be able to assist me. Using the SADHS 2016 dataset, I want to run a multilevel logistic regression model with self-reported presence TB as the outcome of

interest (generated from the variables s1410 and sm1105 from the female and male adult health recode files, respectively). I dropped the 'I don't know' observations from this variable as they were very few and therefore the resulting TB variable is binary and appropriate to use in a logistic regression model.
However, for the three levels I intend to use individuals (adult health respondents) as level 1, households as level 2 and MUNICIPALITIES as level 3. Given that cluster shall not be the intended level 3, how do I format this in the svyset command? Could someone help me out with what STATA code I would need to use?

NB: I have matched clusters to municipalities, using GADM and merged this with the PR file so I know what households are in what municipalities, and I have h_munic as a municipality variable.

Best,

Jadene

Subject: Re: Household crowding
Posted by Bridgette-DHS on Thu, 20 Aug 2020 17:53:36 GMT
View Forum Message <> Reply to Message

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

I don't believe the distinction between level 1 and level 2 will be useful. The only leverage you will get on estimating an intra-class correlation between individuals in the same household will come from households in which more than one person has TB. If the number of persons with TB in the same household is always 0 or 1, you have no information. If there are very few exceptions to 0 and 1, the estimate will be very uncertain. The number of persons tested in the household is also relevant

To use municipality, rather than cluster, at level 3 (or 2) requires constant weights and I don't think you have them. Clusters, not municipalities, were the PSUs in the sample design.

We just issued a methodological report related to multilevel models, and I recommend that you look at it: https://www.dhsprogram.com/pubs/pdf/MR27/MR27.pdf. I hope other users have suggestions.
