Subject: Pakistan DHS 2012 district codes
Posted by Magda on Tue, 23 May 2023 14:33:27 GMT
View Forum Message <> Reply to Message

Hello,
I am interested in using Pakistan DHS 2012 and DHS 2017. In the DHS2017 individual recode dataset PKIR71FL we have a variable "sdist" with district codes and district names reported in the do file. The DHS 2012 dataset PKIR61FL has the same variable "sdist" with what look to be largely the same codes but we are not given the names corresponding to each code in the do file. Is it safe to assume the codes in 2012 are the same as the codes in 2017? Alternatively, is there other way to tell location in the 2012 data that is more specific than just regions?
Thank you for any tips you may have.
Best,
Magda

Subject: Re: Pakistan DHS 2012 district codes
Posted by Janet-DHS on Thu, 25 May 2023 12:53:56 GMT
View Forum Message <> Reply to Message

Following is a response from DHS staff member, Tom Pullum:

We are checking with Pakistan's National Institute of Population Studies (NIPS) and will post a response after hearing them. Please re-post the question if we have not provided an answer by June 15.

Subject: Anemia in Adolescents
Posted by Damilola on Thu, 25 May 2023 15:28:18 GMT
View Forum Message <> Reply to Message

Hello Team,

I would like to inquire about the appropriateness of using the Household member recode file for my analysis on anemia in adolescent boys and girls. Upon investigation, I discovered that only hemoglobin (Hb) levels of females are recorded in this file. Consequently, I am uncertain if this is the correct file for my intended unit of analysis.

If the Household member recode file is not suitable, I kindly request your guidance on which files I should consider instead. However, if it is indeed the correct file, I would appreciate your assistance in incorporating the Hb details for males into my analysis.

Thank you for your attention to this matter.

Subject: Re: Pakistan DHS 2012 district codes

Posted by Janet-DHS on Tue, 30 May 2023 21:01:23 GMT
View Forum Message <> Reply to Message

Following is a response from DHS staff member, Tom Pullum:

When users have a question that is completely different from one they posted before, they are asked to post it as a separate question with a new title.

Are you asking again about the Pakistan 2012 survey? None of the DHS surveys in Pakistan have included hemoglobin / anemia. The variables for children and women are included in the PR files but they are empty. The variable labels include "na" and the cases are all coded with a dot, which in Stata means "not applicable".

If a survey includes hemoglobin, the priorities are children age 6-59 months, then women 15-49, then men 15-49 (the prevalence of anemia is usually much lower for men than for women). The variables for children are hc52-hc57. For women they are ha52-ha57, and for men they are hb52-hb57. PR files may include those variables even if they are empty.

When data are not collected during the fieldwork we cannot put them into the data files.

Subject: Re: Pakistan DHS 2012 district codes
Posted by Janet-DHS on Wed, 31 May 2023 12:25:47 GMT
View Forum Message <> Reply to Message

Following is a response from DHS staff member, Tom Pullum:

I have heard back from the implementing agency regarding the names of districts in the 2012 survey. They say that the names were intentionally not released, in order to discourage users from making district-level estimates. The labels in the 2017-18 survey cannot be used for the 2012 survey.

The 2017-18 survey was only slightly larger than the 2012 survey, but district names were released for the 2017-18 survey. The justification for not releasing district names for the 2012 survey seems inconsistent. However, I agree that district-level estimates would be statistically very unstable.

I believe your plan was to use district as a statistical control, in an analysis of national-level or province-level change between the two surveys, rather than to produce district-level estimates. That was a good idea.

The districts are nested within provinces. The total numbers of districts within provinces is usually close but not the same in the two surveys. With some detective work it might be possible to establish equivalences but perhaps some other geographic structuring of the data would be safer.

Subject: Re: Pakistan DHS 2012 district codes
Posted by Magda on Wed, 31 May 2023 16:49:43 GMT
View Forum Message <> Reply to Message

Thank you very much for your response. Just a quick follow-up: based on your message (and the response from the Pakistan team), am I correct in thinking that the survey sampling for the 2012 data is not representative at the district level, but the 2017 data is representative at the district level? I am not planning on calculating district-level estimates but I was hoping to use variation across districts as a the source of identifying variation.
Thanks again.
Magda

Subject: Re: Pakistan DHS 2012 district codes
Posted by Janet-DHS on Fri, 02 Jun 2023 20:43:17 GMT
View Forum Message <> Reply to Message

Following is a response from DHS staff member, Tom Pullum:

District-level estimates are no more representative in the second survey than in the first survey. There was no relevant difference in the survey designs. The 2012 survey had about 94,000 household members and the 2017-18 had about 101,000, including some geographic areas that were not in the 2012 survey, so the number of cases per district was about the same.

The estimates are unbiased at the district level but DHS would say they are "not representative" because the standard errors would be so high.

What you want to do, looking at variation across districts, sounds good, and you can do that even if you do not have names for the districts. But you have to be cautious about over-interpreting the extreme values. Statistical instability because of small numbers of cases in many of the districts can lead to exaggeration of the estimated range across districts. You could google something like "Bayesian adjustment for hot spots", for example, to find ways adjust for this effect.

Subject: Re: Pakistan DHS 2012 district codes
Posted by owraza on Mon, 25 Sep 2023 02:49:13 GMT
View Forum Message <> Reply to Message

Janet-DHS wrote on Sat, 03 June 2023 01:13Following is a response from DHS staff member, Tom Pullum:

The estimates are unbiased at the district level but DHS would say they are "not representative" because the standard errors would be so high.

What you want to do, looking at variation across districts, sounds good, and you can do that even if you do not have names for the districts. But you have to be cautious about over-interpreting the extreme values. Statistical instability because of small numbers of cases in many of the districts

can lead to exaggeration of the estimated range across districts. You could google something like "Bayesian adjustment for hot spots", for example, to find ways adjust for this effect.

So it would be safe to say that using geographically weighted regression, one can create a spatial surface to depict variation across districts where aim is not to estimate at district level.

Subject: Re: Pakistan DHS 2012 district codes
Posted by Janet-DHS on Wed, 27 Sep 2023 13:28:08 GMT
View Forum Message <> Reply to Message

Following is a response from DHS staff member, Tom Pullum:

Yes, "geographically weighted regression, [with] a spatial surface to depict variation across districts" should be ok. Smoothing is needed to avoid over-estimating the highs and lows.

Subject: Re: Pakistan DHS 2012 district codes
Posted by owraza on Thu, 28 Sep 2023 02:37:57 GMT
View Forum Message <> Reply to Message

Thanks.
