
Subject: Are cluster values the same across different years (versions)?

Posted by [SantiagoCC](#) on Tue, 23 May 2017 14:59:34 GMT

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

Hi, I wanted to ask if the cluster values relating specific regions in a country are the same in different versions of the DHS program.

I mean, if in version 2000 a cluster with value 1 made reference to region A, then if in version 2005 a value of 1 for the same variable also refers to region A or to another one.

Thanks!

Subject: Re: Are cluster values the same across different years (versions)?

Posted by [Bridgette-DHS](#) on Wed, 24 May 2017 02:23:52 GMT

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

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

Any repeats of the same cluster could only happen by chance, and it is unlikely that there have ever been any such repeats. And if a repeat did occur, there would be no way to identify it, even by DHS staff.

Subject: Re: Are cluster values the same across different years (versions)?

Posted by [SantiagoCC](#) on Wed, 24 May 2017 03:34:01 GMT

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

In that case, is there any way in which I can match clusters across different DHS versions? I'm particularly interested in obtaining the village codes for each observation across DHS samples in years 1995, 2000 and 2005 for Colombia. I tried making a request for this data to the archive of DHS but they haven't answered yet.

Subject: Re: Are cluster values the same across different years (versions)?

Posted by [Bridgette-DHS](#) on Fri, 26 May 2017 15:19:27 GMT

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

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

Thank you for posting this question on the forum, where other users can see it. Unless a question is very specific, we prefer that it be put on the forum rather than sent to the archive email address.

It also happens that we did not receive the email you sent to the archive. Perhaps there was a typo in the address....

In general, clusters are enumeration areas (EAs) from the most recent census. After the survey

has been done, DHS does not have any way to identify which EAs appeared in the sample, let alone to link cluster id codes to specific EAs in the census sampling frame. The lists used during sampling and data collection are always destroyed. This is part of the agreement between DHS and the countries, in order to protect the privacy of the respondents. This is also the reason why the GPS coordinates of the clusters (in the more recent surveys that include them) are always displaced. DHS tries to make it impossible for anyone to identify the selected EAs.
