

---

Subject: Households within clusters

Posted by [zafferano](#) on Mon, 21 Aug 2017 20:50:35 GMT

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

---

Hello there,

As far as I understood, most of the DHS data covers around 1/4 of the households within a cluster (e.g., 25 out of 100 households in a cluster). However, I require a dataset that covers a higher proportion of households in a cluster (i.e., more than 1/4). I was wondering whether you know any dataset that fits to this description.

Furthermore, I realized that the continuous Peru datasets also use the same clusters in the latter waves. So, I was also wondering whether the same households are chosen in the latter waves or new households are interviewed in the latter waves. In other words, if I merge 4 continuous datasets from Peru, would I be able to include more households within a cluster in my final sample? Thanks in advance for your help!

---

---

Subject: Re: Households within clusters

Posted by [Bridgette-DHS](#) on Fri, 17 Nov 2017 17:19:22 GMT

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

---

An email was sent to the user.

---

---

Subject: Re: Households within clusters

Posted by [Bridgette-DHS](#) on Mon, 08 Jan 2018 20:30:11 GMT

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

---

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

We apologize for the extreme delay in this response to your question. However, we are not able to be very helpful. DHS typically takes about 30 households from each cluster or enumeration area (EA). The target number is generally the same for every EA, and any deviation from that target number is mainly related to vacant households, refusals, etc., and has no relationship at all to the number of households in the EA. We do not retain any information about the number of households in the EA, either from the sampling frame or from the listing that precedes the selection of specific households. Therefore we cannot calculate the fraction of households in the EA that are included in the sample. It is probably much less than 1/4, on the average. In the Peru CS, if you combine rounds as you describe, you will definitely get a larger fraction than in a single round, but it is impossible to tell what that fraction is.

---