

---

Subject: Re: DHS-MIS

Posted by [Bridgette-DHS](#) on Fri, 01 Mar 2024 16:12:35 GMT

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

---

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

First, you did not identify the survey you are working with, but I believe it is the Tanzania 2022 DHS (which is not an MIS). 15,705 households were selected. These are the cases in the HR file. The number of people listed in the households, as either de jure or de facto residents, is 40,394. These are the cases in the PR file. They include men, women, children. Eligible women for the interviews are women age 15-49 who are de facto residents. Some households have no such women, some have 1, some have 2, etc. The total number of women who were eligible AND were interviewed is 15,254. These are the cases in the IR file. There is an average of about one woman per household, but the ratio of women to households does not have to be this close to 1 and it certainly does not need to be exactly 1. The cluster ID is v001. Households are numbered within clusters, with v002 as the household ID. Individuals within the household are given a line number, and that's v003. "caseid" is a string variable that includes columns for v001, v002, and v003. It takes 15,254 unique values.

If you open the IR file in Stata and enter "codebook v001 v002 v003 caseid" and "list v001 v002 v003 caseid if \_n<=100, table clean" you will get a list of the ID variables for the first 100 women in the file. It may help you to see how the data are structured. As I said, there is no problem.