

---

Subject: Ethnicity values across surveys

Posted by [Daniel Pérez](#) on Fri, 28 Oct 2022 07:29:58 GMT

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

---

Dear DHS team,

I'm currently using DHS 2000, 20005, and 2011 for my research. When I merged them I realized that, apparently, the ethnicity value (number) changes between DHS 2005 and DHS 2011, so the amhara ethnicity value is 5 (DHS 2011) instead of 4 (DHS 2000, DHS 2005).

I'm wondering if the ethnicity values also change between DHS 2000 and DHS 2005 (this is difficult to say since only the major ethnicities are labeled in DHS 2000, their values coincide but I don't know if this is the same case for the rest). All in all, using Stata, I would like to have the three surveys merged with each ethnicity defined by a single value. If one ethnicity is numbered differently across surveys I would have 2 or more values for the same ethnicity and that wouldn't be accurate for my analysis.

Perhaps you can help me with this issue,

Thank you in advance,

Daniel.

---

---

Subject: Re: Ethnicity values across surveys

Posted by [Bridgette-DHS](#) on Fri, 28 Oct 2022 15:04:29 GMT

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

---

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

I just looked at the IR files for these surveys (ET41 and ET51) and did "label list V131" for both of them. I also did "numlabel, add" and "tab v131". As you say, the label in the 2000 survey is incomplete but it appears that the classification was the same in both surveys. Some categories do not appear in both surveys, but when they do appear, the frequencies are low, so it's plausible that in one of the surveys no one with such ethnicities would appear. I don't see evidence that the codes changed, but there's no way that we could confirm this. Perhaps you could check whether there was a census around the same time and find what codes it used. The surveys would probably have been consistent with the census.

---