Subject: Re: Mismatch in educational attainment and education in single years
Posted by Bridgette-DHS on Wed, 24 Jan 2024 18:31:37 GMT
View Forum Message <> Reply to Message

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

You are asking about the Kenya 2022 survey. An erratum is being issued on the coding of education in this survey.

I suggest that you look at the relationship among v106, v107, v133, and v149 in the previous survey, the Kenya 2014 survey, because the relationship was ok in that survey. In the 2014 survey, you will see how v133 is supposed to be constructed as a recode of v106 and v107. In the IR file, do this:

. tab v107 v106, summarize(v133) means

                  Means of education in single years

highest | year of | education	highest educational level			
	primary	secondary	higher |	Total
no years |	0	8	12 |	8.7214076
1 |	1	9	13 |	9.2506473
2 |	2	10	14 |	10.554584
3 |	3	11	15 |	9.9238281
4 |	4	12	16 |	11.115502
5 |	5	.	17 |	5.3082022
6 |	6	14	18 |	6.2527945
7 |	7	.	19 |	7.0332318
8 |	8	.	20 |	8.0017804
Total |	6.6739256	10.794648	14.333705 |	8.7562835"

Ignore the "total" row and column. You will see that each combination of v106 and v107 gives a value of v133. (There are two combinations that give v133=12 and two that give v133=14.) That is, v133 is a recode of v106 and v107.

If you enter "tab v106 v149" you will see that v149 is a recode of v106, which just distinguishes between whether the person completed the primary or secondary levels. That's based on the value of v107 within each level of v106.

That's all there is to it. v106 and v107 are the key variables; v133 and v149 are constructed variables and they were constructed incorrectly in the Kenya 2022 survey. You can either reconstruct them yourself, for your use, or wait for the next version of the Kenya 2022 data files.