Subject: Re: CR-KR Merger
Posted by Janet-DHS on Mon, 09 Sep 2024 14:15:43 GMT
View Forum Message <> Reply to Message

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

I don't know what survey you are using, but in the Kenya 2022 survey, for example, the d105 variables take 5 values, as follows:

. label list D105A

D105A:

0 never

1 often

2 sometimes

3 yes, but not in the last 12 months

4 yes, but frequency in last 12 months missing

These are not on an interval-level scale, and are not even ordinal. A z-score does not make sense--that is, is not interpretable--for such a variable. So my first recommendation would be that you re-code these variables before constructing a z score. However, in my opinion, a usable recode is not possible. You could reduce to binary (0/1) form, but in general a 0/1 variable is much easier to interpret in that form than in a standardized form.

I think the specific problem you have is due to a variable taking only one value in the data, such as "never". In such a situation the standard deviation is 0, and you cannot divide by 0, so the standard score is not defined.

The files include a variable d106, which is a summary of all the d105 variables. I don't think you can do much better than that for a summary. I really don't think your sum of z scores is interpretable.