Subject: Re: Comprehensive HIV Knowledge
Posted by Janet-DHS on Fri, 27 May 2022 17:44:58 GMT
View Forum Message <> Reply to Message

Following is response from DHS Research & Data Analysis Director, Tom Pullum:

I will describe the construction of this variable in terms of the women's (IR) file for this survey and Stata. First you have to find the relevant variables. Here they are, in terms of the "describe" command and what it produces:

```
. describe v754cp v754dp v756 v754jp v754wp

              storage   display    value
variable name   type    format     label      variable label
------------------------------------------------------------- ----------------------------------------------------------- ---
v754cp          byte    %8.0g      V754CP     reduce risk of getting hiv: always use condoms during sex
v754dp          byte    %8.0g      V754DP     reduce risk of getting hiv: have 1 sex partner only, who has no other partners
v756            byte    %8.0g      V756       a healthy looking person can have hiv
v754jp          byte    %8.0g      V754JP     can get hiv from mosquito bites
v754wp          byte    %8.0g      V754WP     can get hiv by sharing food with person who has aids
```

You could find the exact wording of the questions in the questionnaire, which is included in an appendix. Then you have to find the codes for these five variables. They are all coded the same way, like this:

```
. label list V754CP
V754CP:
           0 no
           1 yes
           8 don't know
```

Then, to match table 12.3.1 in the report, the variable is constructed with these two lines:

```
gen     comp_know=0
replace comp_know=1 if v754cp==1 & v754dp==1 & v756==1 & v754jp==0 & v754wp==0
```

To confirm that this matches the table, do this:

```
. tab comp_know [iweight=v005/1000000]

  comp_know |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 | 4,069.3388       61.46       61.46
          1 | 2,551.6612       38.54      100.00
------------+-----------------------------------
      Total | 6,620.9999      100.00
```

This gives the number of cases (6621) and the relevant percentage in the bottom row of the table (38.5%).

I believe you were thinking that you would add up 1's and 0's and get a total, and then find a cutoff for that table. This variable has a more complicated definition. You have to specify that the first three components are exactly 1 and the last two are exactly 0. You can't construct it just with addition.
