Dear Brigette from DHS

I am posting a query regarding the South Africa DHS 2016 Domestic Violence module.
More specifically, my query pertains to domestic violence module and I am unable to replicate results from SA DHS report page 356 paragraph 2:

"... module was completed for 6,620 women, including 5,865 (weighted) ever-partnered women."

The SA DHS report is attached. I try to replicate SA DHS report to make sure that my methods is correct, but can not replicate this 5,865.
This 5,868 is the key since it is the denominator for much of the DV results in the DHS report.

According to the page 18 paragraph 2 of the SA DHS 2016 INDIVIDUAL RECODE DOCUMENTATION (attached):

" 4. Selection for domestic violence included all women 18+ even in households where the adult health questionnaire was not applied. Women 49+ selected for domestic violence in households not selected for "Adult health", were asked question related to the adult health study, but that information wasn't used in the report. Questions on violence from partner were asked not only to ever married women, but to never married women that currently have or had a boyfriend (Question Q1502A in questionnaire). Tables in the final report were produced for all women selected for the module regardless of age. Caution should be observed when analyzing adult health variables or domestic violence because the variables may be present in the file (i.e., in the woman's file) but they don't include all individuals that were selected for those studies. In general, if the intended analysis includes any of these variables it is advised to use the adult data files or the special file produced for domestic violence."

Constructing DV and deriving 5,865 from various SA DHS dataset is indeed challenging after several months of trying.
Consistent with the last sentence of the above paragraph, would it be possible to obtain a STATA do file to produced a file for domestic violence?

I obtained permission to download SA DHS 2016 dataset.
 https://dhsprogram.com/data/dataset/South-Africa_Standard-DH S_2016.cfm?flag=1
Based on going to above url, there is no DV dataset to the best of my knowledge but I could be mistaken.
Any assistance with constructing DV dataset would be much appreciated.

File Attachments
1) DHS Questionnaire.pdf, downloaded 465 times
2) DHS Individual Recode.doc, downloaded 252 times

Hello,

For the domestic violence indicators you will need to merge a DV and IR file. This is unique to the South Africa 2016 survey. Usually in other DHS surveys the domestic violence variables are only in the IR file. I provide Stata code below that merges the two files then creates the first domestic violence indicator in the final report (ever experienced physical violence). The denominator I obtain is 9 cases over what is reported in the final report but I have not been able to find a way to exclude these cases. However, the percentage matches since it's only 9 cases off.

Hope you find this useful.

Best,
Shireen Assaf
The DHS Program

Stata code:

```
*first open the DV file
use "ZADV71FL.DTA" , clear

*removing missing values from the first domestic violence indicator in the file
keep if  d101a<.

gen in_DV=1
sort v001 v002 v003
save ZAtemp.dta, replace

*now open IR file
use "ZAIR71FL.DTA" , clear

*remove the same missing values
keep if  d101a<.

gen in_IR=1
sort v001 v002 v003

*merge the files using the temporary DV file you saved.
merge 1:1 v001 v002 v003 using ZAtemp.dta
tab _merge
* this is to see which women are coming from each file
tab in_DV in_IR,m

*generate an age variable
gen age10=int((v012-5)/10)
replace age10=6 if age10>6
```

```
label define age10 1 "18-24" 2 "25-34" 3 "35-44" 4 "45-54" 5 "55-64" 6 "65+"
label values age10 age10

*tabulate age variable to see denominator
tab age10  [iweight=d005/1000000]

*generate variable for every experienced physical violence
gen PVever = (inrange(d105a,1,4) | inrange(d105d,1,4) | inrange(d105e,1,4) | inrange(d105f,1,4) |
inrange(d130a,1,4) )

*the percentage matches the final report
tab PVever [iw=d005/1000000]
```