
Subject: Re: Merging Women-Men-HIV Data: Different Countries, Different Years
Posted by [Bridgette-DHS](#) on Fri, 15 Mar 2019 12:26:10 GMT

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

Following is another response from Senior DHS Stata Specialist, Tom Pullum:

You are welcome.

Sorry for the double "=" sign. I should have said "gen sex=2". The reason for putting that variable ("sex") into the appended IR+MR file is that you will want to change the "mv" variables in the MR file to "v" variables, but once you do that, you need a way to distinguish whether the case comes from the IR file or the MR file. The IR and MR files do not contain a variable that is explicitly the sex of the respondent. The coding I suggest is the same as for hv104, with 1 for males and 2 for females. At some point I would also put

```
label define sex 1 "male" 2 "female"  
label values sex sex
```

As for numbering the surveys, you can do that however you want, but the most important thing is also to have a string label such as "survey_label" because the DHS files do not explicitly include such a label. You cannot rely on v000, for example, to identify the survey. It often happens that two successive surveys have the same value of v000. You can construct the label afterwards using the first two characters of v000, which are the country code, and the year(s) of fieldwork, v007, but I always insert the survey_label during the construction of the large file, rather than in a later step.