
Subject: Re: STATA codes to merge women and household datasets

Posted by [Bridgette-DHS](#) on Tue, 03 Sep 2019 19:37:55 GMT

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

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

Hi Isabelle--Here is how I would do the merge. I use the old version of the merge command but you would get the same thing if you used 1:1. I also introduce a variable called "in_IR", which is coded 1 for every case in the IR file. It just clarifies the "_merge" code, which is described in your results with different terms.

ALL of the 90,303 cases in the IR file are also in the PR file. 427,076 people in the PR file are NOT also in the IR file. Hope this makes sense. Tom

```
cd e:\DHS\DHS_data\scratch
```

```
use "C:\Users\26216\ICF\Analysis - Shared Resources\Data\DHSdata\IAIR42FL.DTA" , clear
gen hv024=v024
gen hv001=v001
gen hv002=v002
gen hvidx=v003
gen in_IR=1
sort hv024 hv001 hv002 hvidx
save IAIR42_temp.dta, replace
```

```
use "C:\Users\26216\ICF\Analysis - Shared Resources\Data\DHSdata\IAPR42FL.DTA" , clear
sort hv024 hv001 hv002 hvidx
merge hv024 hv001 hv002 hvidx using IAIR42_temp.dta
replace in_IR=0 if in_IR==.
tab _merge
tab _merge in_IR
```

The following table is produced:

File Attachments

1) [table.PNG](#), downloaded 4186 times
