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

Posted by [Isabelle](#) on Sat, 24 Aug 2019 13:56:26 GMT

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

Thank you Bridgette for your response and help.

However, I have a follow up question after merging the datasets including forth identifier v024/hv024.

When merging the datasets IR and PR using

household member:

```
gen int v001 = hv001
```

```
gen int v002 = hv002
```

```
gen byte v003 = hv003
```

```
gen int v024 = hv024
```

```
sort v001 v002 v003 v024
```

individual:

```
sort v001 v002 v003 v024
```

```
merge 1:m v001 v002 v003 v024 using "C:\Users\Isabelle\Desktop\DHS Data India\Single  
Datasets Recode\1998-99Recode\V21998-99HHMSort.dta"
```

I get the following result:

```
not matched: 337,481
```

```
from master 49,203
```

```
from using: 288,278
```

```
matched: 229,101
```

The household dataset has 517,379 observations

The individual dataset has 90,582 observation

Does this mean that from my master (individual) file, only (90,582-49,203) 41,379 individuals are not matched to a household; meaning I cannot use them in my further analysis?

Thank you in advance for your help!

Best

Isabelle