Subject: Re: Are the merging commands and results correct
Posted by Bridgette-DHS on Tue, 24 Mar 2020 18:18:08 GMT
View Forum Message <> Reply to Message

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

I recommend that when you change the names of the matching variables, you consistently use EITHER the PR variable names OR the IR variable names. For example, rename hv001 v001; rename hv002 v002; rename hvidx v003. Then sort both files on v001 v002 v003. The way you did it will work but it's not as systematic.

I think the merge has worked ok. My rule of thumb is that about a quarter of the people in the PR file will be women age 15-49. That is, the cases with _merge=1 will be about 3 times as large as the cases with _merge=3. If that's what you got, you are ok.

You could consider reducing the PR file to women age 15-49, right after the "use ...PR...". This will give you both the de fact and de jure residents.