
Subject: Re: Merging Men, Women, Child dataset into Household level unit analysis
Posted by [Bridgette-DHS](#) on Wed, 07 Sep 2016 23:17:24 GMT

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

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

For most merges and appends It helps to think of a "case" as a unique individual and a "record" as a line of data. When you merge two files, you are consolidating or combining two records that refer to the same case. That's why you have to identify the cases in both files with id codes such as hv001, hv002, hvidx in the PR file and v001, v002, v003 in the IR file. This is done when you want to attach information about a woman to the information about her household by merging the IR and PR files. Or maybe the case in the merged file will be a couple, a man and a woman, so you merge the IR and MR files to make a CR file, using the stated line numbers of the partners.

You append one file to another if the cases are similar but different (that phrase could be made more precise!). For example, you may have a 2010 survey and a 2015 survey from the same country. The cases are completely different but you can simplify some of the data processing if you append or combine into a single file (keeping an identifier for which survey is which). You would never append an IR file to a PR file, for example.

When I think about manipulating two files, I usually have a physical image in my mind, of two stacks of paper. Do I want to put one stack of paper on top of the other one (append), or do I want to transfer the information in one stack to the other stack, sheet by sheet (merge). Most computer procedures are just a faster way of doing what could be done manually (if we had a LOT of time!).

Sometimes it can be efficient to combine appending and merging, in succession. Let me know if you want to be more specific.
