Subject: Re: What is the underlying criteria for merging datasets from various waves?
Posted by Bridgette-DHS on Tue, 14 Jun 2016 14:28:18 GMT
View Forum Message <> Reply to Message

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

The fifth character of the file name is the phase of DHS (0 through 7, but 0 refers to phase 1; we are currently in phase 7). The sixth character is the version. It is usually 0 for the first version of the data file, when the survey is first released, and then if corrections or additions (almost always of negligible importance) are made, the version is updated to 1, 2, etc. CMAR61FL and CMMR60FL are completely compatible. The AR file had one update but the MR file did not.

Sometimes this happens--one of the files from a survey has had more updates than another one of the files from the same survey. This should not affect a merge.

If there was a second survey in the same phase of DHS, then the first version of the second survey will be A, which is updated to B, etc. If there was a third survey, then the first version of the third survey will be G, which is updated to H, etc.

Obviously the file naming system is sometimes awkward. It goes back to a time when the file name (before the dot) could be at most 8 characters long. That's no longer a restriction, and I expect the whole system will be revised before long.....