Subject: Mismatch of file names within the same survey
Posted by Olena Stavrunova on Tue, 29 Mar 2022 03:53:25 GMT
View Forum Message <> Reply to Message

Dear all,
I am trying to match the BR and PR data for several surveys. I noticed that for some surveys the name of BR files does not correspond to the name of the PR file. For examples, Indonesia 2007 has IDBR51DT.ZIP and IDPR52DT.ZIP - notice different survey release code: 51 vs 52. There are many similar examples from other countries. Could you please explain what this means and whether these two files can still be merged? I attach a screenshot FYI.
Thank you
Olena

File Attachments
1) Screenshot 2022-03-29 143341.png, downloaded 393 times

Subject: Re: Mismatch of file names within the same survey
Posted by admin on Fri, 15 Apr 2022 20:01:53 GMT
View Forum Message <> Reply to Message

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

You have identified a nuisance that we often have to deal with. Column 6 of the filename is a "version" code. Sometimes (for a range of reasons) the DHS data processing staff will update a file and will increment the version code (for example from 1 to 2) for that file. If the other files for that survey are not affected, their version codes will not be changed. If you previously would have merged BRxx51 with PRxx51, you would now merge BRxx51 with PRxx52, for example.

When looping through lots of surveys, you may hit a problem like this. A related problem comes up if both files were 51 and now both are 52. You have to have enough flexibility to handle such changes/updates in filenames.

Note that you need the version number because the phase identifier in column 5 is not sufficient to identify a survey. For example, Egypt had both 51 and 5A surveys--two surveys within DHS-5.