

---

Subject: Re: Data name and data do not match

Posted by [Bridgette-DHS](#) on Mon, 21 Jun 2021 13:23:01 GMT

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

---

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

The 5th character in the file name is the phase of the DHS coding manual and the 6th character is the "version" if there are updates. For example, when the files are first issued, the 5th-6th characters for a Phase 6 survey could have been "60". If there was some correction or update, the files were re-issued with "61". Next time, "62". Usually, ALL the files from a survey (HR, PR, IR, MR, CR, BR, KR) will have their file names updated, even if the update does not affect the specific file. Sometimes (more in the past than in the present), if a file is not affected, it will not be renamed. In the examples you give, the MR file must not have been affected, so that file was not re-issued or re-named.

This is annoying for all of us who use the data, but when it happens, you have to adapt. For example, if you have a combination of IR62 and MR61, you just have to treat the MR61 file as if it were labelled MR62. The two files come from the same survey.

When there are two surveys in the same phase of DHS, the 6th character will begin as a letter, and then any updates will be the next letter. For example (these may not be the actual letters that would be used) the first survey may be in a sequence 60, 61, 62, as described above, but a second survey would be 6H, 6I, 6J... Then if there is a third survey within the same phase of DHS, the sequence will be 6R, 6S, 6T...

This numbering system leaves much to be desired. It goes back to decades ago, when file names were restricted to 8 characters. Eventually, I am sure, a more flexible and intuitive naming system will take over.