

---

Subject: Re: Duplicates in IR file when merging

Posted by [Bridgette-DHS](#) on Wed, 13 Jan 2016 13:18:12 GMT

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

---

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

If you list out caseid and v003 for the first few cases in the IR file, and list out hhid for the first few cases in the HR file, you will see that caseid is simply a combination of hhid and v003.

Both caseid and hhid are character strings (str15 and str12, respectively); v003 is numeric.

Before you do the merge, when you open the IR file, use this line: `gen hhid=substr(caseid,1,12)`. Then merge with the IR file using hhid. I think this will work for what you want to do, but let me know if it does not.

---