

---

Subject: Re: Merging Egypt 2000 KR and HW files  
Posted by [Bridgette-DHS](#) on Tue, 10 Jan 2017 02:43:20 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

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

Quote: This merge is non-standard because caseid in the KR file for this survey is left-adjusted and has no leading blanks. You have to move the trailing blanks to the beginning, and remove the woman's line number in order to get a hhid that will match with the HW file. This was a mistake but unfortunately it comes up in a few other surveys too. To figure this out I had to do quite a bit of column-by-column deciphering of caseid. The following lines will work. You have to change the paths. I am using an older version of the merge command, which I much prefer to the current version.

\* How to merge the KR and HW files from the Egypt 2000 DHS survey

```
set more off
```

\* Prepare the KR file

```
use e:\DHS\DHS_data\KR_files\EGKR42FL.dta, clear
```

\* caseid is left-adjusted with no leading blanks, but has trailing blanks;

\* caseid is str15, but longest is str13. It should be hhid plus 3 cols for v003

\* I must convert it to be right-adjusted and without the woman's line number

```
use e:\DHS\DHS_data\KR_files\EGKR42FL.dta, clear
```

```
codebook caseid
```

```
recast str13 caseid
```

```
gen length=strlen(caseid)
```

```
tab length
```

\* remove the last three non-missing columns of caseid, which are the line number of the woman, to get hhid

```
gen str2 blank2=" "
```

```
gen str3 blank3=" "
```

```
gen str4 blank4=" "
```

```
gen str12 caseidrev=" "
```

```
replace caseidrev=blank4+substr(caseid,1,8) if length==11
```

```
replace caseidrev=blank3+substr(caseid,1,9) if length==12
```

```
replace caseidrev=blank2+substr(caseid,1,10) if length==13
```

```
rename caseidrev hhid
```

```
rename b16 hvidx
```

```
sort hhid hvidx
```

save e:\DHS\scratch\temp.dta, replace

\* Prepare the HW file and do the merge  
use e:\DHS\DHS\_data\HW\_files\EGHW41FL.dta, clear

\* hwhhid is right adjusted with no trailing blanks, but has leading blanks, as is normal

\* hwhhid is str12

rename hwhhid hhid

rename hwline hvidx

sort hhid hvidx

merge hhid hvidx using e:\DHS\scratch\temp.dta

tab \_merge

\* then keep if \_merge==3, drop \_merge, and save the file