

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

The tricky thing about these merges is that the HW files have the household id as a string. With "describe" you can find that hwhhid is a 12-character string. You then need to extract v001 and v002 from hwhhid. Below I show a trick to help find out which columns are v001 and which are v002. You use destring and substring. I show how to do it with the 2000 survey but the alignment is the same in the 2005 survey.

```
use "C:\Users\26216\ICF\Analysis - Shared Resources\Data\DHSdata\EGHW41FL.DTA"  
rename hwwline b16
```

```
local li=1  
while `li'<=12 {  
  gen col`li'=substr(hwhhid,`li',1)  
  local li=`li'+1  
}
```

```
list hwhhid col* if _n<=20, table clean
```

```
drop col*  
gen v001=substr(hwhhid,3,7)  
gen v002=substr(hwhhid,10,3)
```

```
destring v001 v002, replace  
sort v001 v002 b16  
save e:\DHS\DHS_data\scratch\EGtemp.dta, replace
```

```
use "C:\Users\26216\ICF\Analysis - Shared Resources\Data\DHSdata\EGKR42FL.DTA"  
drop if b16==0 | b16==.  
sort v001 v002 b16
```

```
merge v001 v002 b16 using e:\DHS\DHS_data\scratch\EGtemp.dta  
tab _merge
```
