
Subject: Re: Aggregating household level data
Posted by [Bridgette-DHS](#) on Mon, 15 Aug 2016 12:33:17 GMT
[View Forum Message](#) <> [Reply to Message](#)

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

The following Stata lines will calculate the proportions of men in each cluster who are in each category of mv717. The crucial lines are the one that begins with "xi" and the one that begins with "collapse".

```
use e:\DHS\DHS_data\MR_files\KEMR70FL.dta, clear
keep mv001 mv005 mv717
tab mv717,m
```

```
* Give the NA cases a code of 99
replace mv717=99 if mv717==.
```

```
xi i.mv717, noomit
gen n=1
collapse (mean) _l* (sum) n (first) mv005, by(mv001)
rename _l* *
save e:\DHS\DHS_data\scratch\KEMRtemp.dta, replace
```

```
* This cluster-level file has the number of cases, weight, and the proportions
* of men in each category of mv717.
* You will need to attach variable names to the mv717* variables
```

```
* Proceed to the following steps if you want to merge back with the MR file
drop mv005
sort mv001
```

```
use e:\DHS\DHS_data\MR_files\KEMR70FL.dta, clear
```

```
* I would reduce to just the variables you actually need in the MR file
* example:
keep mv0*
```

```
sort mv001 mv002 mv003
merge mv001 using e:\DHS\DHS_data\scratch\KEMRtemp.dta
drop _merge
```

```
* save this file with another name
```
