
Subject: Turkey mother tongue variable

Posted by [gocuk7](#) on Fri, 22 Jan 2016 13:00:15 GMT

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

Hi,

I hope it won't be a very easy question. In the 2003 Turkish DHS, I use the following data set: TRHR4AFL.DTA.

The variable I am interested is sh57_*. In the file TRHR4AFL.FRW, there is only one variable for sh57, while in the data set there are 8, sh57_1...sh57_8. When I individually run counts, I am able to get the weighted data for Kurdish count 599.

The code is:

```
generate psu=hv021
```

```
generate wgt=hv005/1000000
```

```
egen strata = group(hv024 hv025), label  
svyset psu [iweight=wgt], strata(strata)
```

```
svy: tab sh57_1, count format(%11.1fc) /* mother tongue */  
svy: tab sh57_2, count format(%11.1fc)/* mother tongue */  
svy: tab sh57_3, count format(%11.1fc)/* mother tongue */  
svy: tab sh57_4, count format(%11.1fc)/* mother tongue */  
svy: tab sh57_5, count format(%11.1fc)/* mother tongue */  
svy: tab sh57_6, count format(%11.1fc)/* mother tongue */  
svy: tab sh57_7, count format(%11.1fc) /* mother tongue */  
svy: tab sh57_8, count format(%11.1fc)/* mother tongue */
```

However, if I wanted to compile the data with rowmisscount function or other methods, I cannot reached the 559 number.

```
egen mother=rowmisscount(sh57_1 sh57_2 sh57_3 sh57_4 sh57_5 sh57_6 sh57_7 sh57_8)  
svy: tab mother, count format(%11.1fc)/* mother tongue */
```

Any idea how to do this?I hope I was clear.

Subject: Re: Turkey mother tongue variable

Posted by [Bridgette-DHS](#) on Thu, 04 Feb 2016 15:03:05 GMT

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

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

For tables you do not need the full svyset command. You only need to specify the weights. The results are more readable with something like "tab sh57_1 {iweight=hv005/1000000}".

I looked at the household questionnaire in the report on the 2003 survey. sh57 and other sh5* variables refer to never-married women (NMW) age 15-49 who are living in the household. The subscripts _1, _2, etc., are used to identify specific people (never-married women age 15-49) in the household. Values will be missing (".") if there are fewer than 8 eligible persons in the household.

Normally such a variable would also appear in the PR file, just as sh57, with no subscript, for the women 15-49 who were not eligible for the survey of women, which was limited to ever-married women. It would be easier for you to use the PR file rather than the HR file. Then you just have sh57 for the people in the household to whom it applied. The units in the HR file are households and the units in the PR file are individuals.

I looked for sh57 in the PR file. It is there, but all cases are coded ".", i.e. are missing. This is strange. There must have been an error in the construction of the PR file.

In the HR file, sh51a_*, "line number in household schedule", is included. I used it to merge sh57 with the PR file with the following steps (change path names, of course). Note that within this merge, the existing incorrect version of sh57 (with all values ".") must first be removed from the PR file, or it will override the new value.

```
* Copy sh57 from the HR file to the PR file
set more off
use e:\DHS\DHS_data\HR_files\TRHR4AFL.dta, clear
keep hv001 hv002 sh51a_* sh57_*
reshape long sh51a_ sh57_, i(hv001 hv002) j(nmw)
drop if sh51a_==.
rename sh51a hvidx
rename sh57_ sh57
drop nmw

sort hv001 hv002 hvidx
save e:\DHS\DHS_data\scratch\temp.dta, replace

use e:\DHS\DHS_data\PR_files\TRPR4AFL.dta, clear

* You must drop the existing imperfect copy of sh57 !!
drop sh57
sort hv001 hv002 hvidx
merge hv001 hv002 hvidx using e:\DHS\DHS_data\scratch\temp.dta

tab sh57 _merge,m
keep if _merge==3

tab sh57
tab sh57 [iweight=hv005/1000000]
```

The last two lines give the unweighted and weighted tabulations of sh57. The number of

never-married women who are Kurdish speakers is 887 (unweighted) or 599 (weighted). I can confirm these numbers by adding up the numbers from sh57_1 through sh57_8 in the HR file. I do not get your number, 559.

I may not have directly answered your question, but I have shown you how to do much more, because once you have a valid version of sh57 in the PR file, you can actually do some analysis to compare the Kurdish NMWs with the Turkish NMWs, etc.

There are other questions about language as part of the household interview (the language of the interview) and in the survey of ever-married women (go to the HR or IR file and enter "lookfor language").

Subject: Re: Turkey mother tongue variable
Posted by [gocuk7](#) on Fri, 05 Feb 2016 07:29:56 GMT
[View Forum Message](#) <> [Reply to Message](#)

Dear Tom,

Thank you very much for the reply. It is much more enlightening than I ever expected. You answered all the questions about the variable in my mind clearly. My aim finally is to compare other fundamentals with speaking Kurdish, you figured it out. It is like you read my mind and I think this is what being an expert means. Thanks again for your time and effort.

Brigitte, thanks to you too for this great forum.

Best Wishes