Dear members,
I am replicating the "table 2.26 School attendance by state/union territory" as presented in India Country Report (Page No. 64). However, I am getting slightly different statistics and the results are not exactly matching. The code is stated below. As well, the STATA do file is attached with this topic. I request this community to look into codes and suggest me way forward. Any help will be highly appreciated. Thank you very much in advance.

*Replication of the "Table 2.26 School attendance by state/union territory" in the India National Report (NFHS-5)

* Setting maximum variable
set maxvar 12000

* Prepare IR File
* Import data
use "E:\NIUA\3_Data\NFHS-5_Unit Level Data\IAIR7ADT\IAIR7AFL.DTA", clear
tab v133
tab v140
tab v140 [iweight = v005]
keep v024 v001 v002 v003 v011 v024
gen in_IR=1
rename v024 hv024
rename v001 hv001
rename v002 hv002
rename v003 hvidx
sort hv024 hv001 hv002 hvidx
save "E:\NIUA\3_Data\NFHS-5_Unit Level Data\Calculation for school attendance and new age\IAIR7AFL.DTA", replace

* Prepare MR file for merge
use "E:\NIUA\3_Data\NFHS-5_Unit Level Data\IAMR7ADT\IAMR7AFL.DTA", clear
keep mv024 mv001 mv002 mv003 mv011 mv024
gen in_MR=1
rename mv024 hv024
rename mv001 hv001
rename mv002 hv002
rename mv003 hvidx
sort hv024 hv001 hv002 hvidx
save "E:\NIUA\3_Data\NFHS-5_Unit Level Data\Calculation for school attendance and new age\IAMR7AFL.DTA", replace

```
 * Prepare BR file for merge
use "E:\NIUA\3_Data\NFHS-5_Unit Level Data\IABR7ADT\IABR7AFL.DTA", clear
keep v024 v001 v002 b16 b1 b2 b3
gen in_BR=1
rename v024 hv024
rename v001 hv001
rename v002 hv002
rename b16 hvidx
sort hv024 hv001 hv002 hvidx
save "E:\NIUA\3_Data\NFHS-5_Unit Level Data\Calculation for school attendance and new
age\IABR7AFL.DTA", replace

* Prepare PR file for merge
use "E:\NIUA\3_Data\NFHS-5_Unit Level Data\IAPR7BDT\IAPR7BFL.DTA", clear
gen in_PR=1
sort hv024 hv001 hv002 hvidx
save "E:\NIUA\3_Data\NFHS-5_Unit Level Data\Calculation for school attendance and new
age\IAPR7BFL.DTA", replace

 *------------------------------------------------------------ -------------
* program define merge_files

* Using PR File
use "E:\NIUA\3_Data\NFHS-5_Unit Level Data\Calculation for school attendance and new
age\IAPR7BFL.DTA", clear
* Merge with BR
sort hv024 hv001 hv002 hvidx

merge hv024 hv001 hv002 hvidx using "E:\NIUA\3_Data\NFHS-5_Unit Level Data\Calculation for
school attendance and new age\IABR7AFL.DTA"
rename _merge _merge_PR_BR
tab1 in_*,m
drop if in_PR==.
sort hv024 hv001 hv002 hvidx

* Merge with IR
merge hv024 hv001 hv002 hvidx using "E:\NIUA\3_Data\NFHS-5_Unit Level Data\Calculation for
school attendance and new age\IAIR7AFL.DTA"
rename _merge _merge_PR_IR
tab1 in_*,m
drop if in_PR==.
sort hv024 hv001 hv002 hvidx

* Merge with MR
merge hv024 hv001 hv002 hvidx using "E:\NIUA\3_Data\NFHS-5_Unit Level Data\Calculation for
school attendance and new age\IAMR7AFL.DTA"
rename _merge _merge_PR_MR
tab1 in_*,m
```

```
drop if in_PR==.
sort hv024 hv001 hv002 hvidx
gen cmc_of_birth=.
replace cmc_of_birth=b3    if in_BR==1
replace cmc_of_birth=v011  if in_IR==1
replace cmc_of_birth=mv011 if in_MR==1
save "E:\NIUA\3_Data\NFHS-5_Unit Level Data\Calculation for school attendance and new
age\merged_with_PR.dta", replace


  *------------------------------------------------------------ -------------------
* Month of the start of the academic year (April: 40
gen m_acadyr= 4

* Determining the CMC of the academic year
gen cmc_acadyr = ((hv007 -1900)*12+m_acadyr) if m_acadyr >= hv006
replace cmc_acadyr = ((hv007-1-1900)*12+m_acadyr) if m_acadyr < hv006

* Generating age
gen age_schl= .
replace age_schl = int((cmc_acadyr - cmc_of_birth) / 12)


* Age groups
gen age_schl_group = .
replace age_schl_group = 1 if age_schl >=6 & age_schl <=14
replace age_schl_group = 2 if age_schl >=15 & age_schl <=17

* value level define of age groups
label define age_schl_group 1 "6-14 years" 2 "15-17 years", modify
label values age_schl_group age_schl_group


* Share of children attending school withing age group of 6 to 17 years
tab age_schl_group hv121 [aweight = hv005/1000000] if hv103==1, row nofreq

* Share of children attending school withing age group of 6 to 17 years across states
tab hv024 hv121 [aweight = hv005/1000000]if hv103==1 & age_schl_group ==1 | age_schl_group
==2, row nofreq
```

File Attachments

Subject: Re: Replication of table 2.26 School attendance by state/union territory in the India Country Report
Posted by Bridgette-DHS on Thu, 11 Aug 2022 12:48:56 GMT
View Forum Message <> Reply to Message

Following is a response from DHS Senior Analysis & Research Manager, Shireen Assaf:

Please check our standardized code on GitHub at
https://github.com/DHSProgram/DHS-Indicators-Stata .

Specifically check the code for Chapter 2 (Chap02_PH) and the PH_SCHOL.do file. Please read the notes for this do file since you need to make some country-specific adjustments to get the correct year and month of the school calendar and age range for school attendance from the UNESCO website.