Subject: Median Age at Any Event
Posted by crabyou on Wed, 07 Aug 2024 16:40:26 GMT
View Forum Message <> Reply to Message

Since I had to spend a day to finally figure out how the median age variables on NFHS reports are calculated, here is the code for anyone trying the same. I do not claim originality, I have taken snippets of code from multiple sources to customise this code. I illustrate the code below with NFHS-5 data and reports.

```
program define calc_median_age

** Source:
https://userforum.dhsprogram.com/index.php?t=msg&goto=12912&&srch=svyset+individual+data
#msg_12912

replace age=99 if age==. | age==0 // age is any age var of interest. ideally these should be coded missing but the values in NFHS reports will differ slightly without this recoding

summarize age [fweight=v005], detail

scalar sp50=r(p50)

gen dummy=0
replace dummy=1 if age<sp50


summarize dummy [fweight=v005]
scalar sL=r(mean)

replace dummy=0
replace dummy=1 if age<=sp50

summarize dummy [fweight=v005]
scalar sU=r(mean)

drop dummy

scalar smedian=round(sp50+(.5-sL)/(sU-sL),.01)
scalar list sp50 sL sU smedian

* warning if sL and sU are miscalculated
if sL>.5 | sU<.5 {
display in red "ERROR IN CALCULATION OF L AND/OR U"
}

drop age
end
```

```
**************************************************
**************************************************
**************************************************
**************************************************
**************************************************
**************************************************
* EXECUTION BEGINS HERE

* sp50 is the integer-valued median produced by summarize, detail;
*   what we need is an interpolated or fractional value of the median.

* In the program, "age" is reset as age at first any event of interest.

* sL and sU are the cumulative values of the distribution that straddle the integer-valued median

set maxvar 10000
use "NFHS/IAIR7EDT/IAIR7EFL.DTA", clear // individual level women data

************************ Create Variables **********************************

**** Age at first marriage
replace s308c=. if s308c==9998 | s308c==9997 // inconsistent ( coded as 9997 ) or don't know (
coded as 9998 )
replace s309=. if s309==98 // dont know is coded as 98

**Source: suggested method for calculating age at marriage
https://userforum.dhsprogram.com/index.php?t=msg&goto=14267&S=Google#:~:text=Can%20an
yone%20illustrate%20this%20variable%3F&text=*Using%20this%20syntax%2C%20it%20has,bef
ore%2018%20years%20of%20age.

gen afm= int((s308c - v011) / 12)
replace afm=s309 if afm==.

la var afm "Age at First Marriage"

**** Age at first birth
gen afb= int((v211-v011)/12)
la var afb "Age at First Birth"


************************** Replicate Numbers on Report ************************
****** Source: https://dhsprogram.com/pubs/pdf/FR375/FR375.pdf
***** Median age at first marriage is 19.2 years among women age 20-49 - Page 208
preserve
gen age=afm

keep if v013!=1 // 20-49 olds
```

```
calc_median_age
scalar median_ans=smedian
scalar list median_ans
restore

***** Median age at first birth among women age 25-49 in India is 21.2 years - Page 114
preserve
gen age=afb

keep if v013!=1 & v013!=2 // 25-49 olds

calc_median_age
scalar median_ans=smedian
scalar list median_ans
restore
```

For calculating median age at any event for men, replace the weights (v005) with men's sample weights. I have also added the relevant links I had referred to.
