
Subject: Re: Calculating median age at first sex and percentage of respondents with sex before the age of 15

Posted by [Bridgette-DHS](#) on Mon, 07 Aug 2017 13:35:52 GMT

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

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

You need to use weights, but the rest of svyset is not relevant for calculating a point estimate. The main issue is that in Stata, medians (and other percentiles) are calculated as integers. You have to interpolate to get anything to the right of the decimal place. Another issue is that you need to recode 0, 98, and 99. The following program will match DHS calculations. It gives 21.46, which rounds to 21.5. To run it, you must change the path to the data file.

* Calculation of median age at first sex

```
set more off
```

```
*****
```

```
program define calc_median_age
```

```
summarize age [fweight=v005] if v012>=25 & v012<=49, detail
```

```
scalar sp50=r(p50)
```

```
gen dummy=.
```

```
replace dummy=0 if v012>=25 & v012<=49
```

```
replace dummy=1 if v012>=25 & v012<=49 & age<sp50
```

```
summarize dummy [fweight=v005]
```

```
scalar sL=r(mean)
```

```
replace dummy=.
```

```
replace dummy=0 if v012>=25 & v012<=49
```

```
replace dummy=1 if v012>=25 & v012<=49 & 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 {
```

```
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 cohabitation or age at first birth;
* with modifications, other possibilities would require modifications.

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

* v011 date of woman's birth (cmc)
* v211 date of first child's birth (cmc)
* v511 age at first cohabitation

```
set maxvar 10000
use e:\DHS\DHS_data\IR_files\PHIR61FL.dta, clear
```

```
* age at first sex calculated from v531
gen afs=v531
replace afs=99 if v531==. | v531==0
replace afs=. if v531==98 | v531==99
gen age=afs
calc_median_age
scalar safs_median=smedian
scalar list safs_median
```