Subject: Zimbabwe 2015 Wealth Index quintiles
Posted by amyfinnegan on Sat, 18 Nov 2017 21:17:12 GMT
View Forum Message <> Reply to Message

Hello -

I am working with the 2015 Zimbabwe DHS to replicate the wealth quintiles. I want to apply the cut-off values to a wealth index we collected in rural Zimbabwe (we asked the same questions used in ZDHS 2015) to show how our sample compares to the population of Masvingo county (where our survey was fielded).

I have downloaded the principal component output for Zimbabwe 2015 from here: https://dhsprogram.com/topics/wealth-index/Wealth-Index-Cons truction.cfm.

I am trying to replicate the quintiles in Stata. But they are not exactly as in the spreadsheet provided.

In the spreadsheet, the values are:

Mean -.1347808
Std. Error of Mean .00903679
Median -.3942115
Mode .03433a
Std. Deviation .92749292
Minimum -1.63862
Maximum 3.24577
Percentiles
20 -1.0280612
40 -.6095050
60 -.0076580
80 .8723670

Here is how I calculated the HHMEMWT:

```
gen dejure = hv012
replace dejure = hv013 if hv012 == 0
gen HHMEMWT = (dejure*hv005)/1000000
```

After weighting, I have the same number of observations and min/max as the spreadsheet but a different mean and slightly different standard deviation: :

```
. sum hv271 [aw=HHMEMWT]

    Variable |     Obs      Weight       Mean   Std. Dev.       Min        Max
-------------+-------------------------------------------- -------------------
       hv271 |  10,534  42909.8382  -183250.4    927336   -1638620    3245770
```

The output from using xtile in Stata gives the below, which does not match the spreadsheet:

. _pctile hv271 [pw=HHMEMWT], nq(5)

. return list

scalars:
r(r1) = -1050290
r(r2) = -654000
r(r3) = -175310
r(r4) = 860240

This matches with the summary by quintile in the ZDHS data:

. bysort hv270: sum hv271 [aw=HHMEMWT]

-> hv270 = poorest

Variable	Obs	Weight	Mean	Std. Dev.	Min	Max
hv271	1,758	8581.44865	-1267315	133072.1	-1638620	-1050780

-> hv270 = poorer

Variable	Obs	Weight	Mean	Std. Dev.	Min	Max
hv271	1,707	8581.49289	-851790.8	111743.7	-1050290	-654100

-> hv270 = middle

Variable	Obs	Weight	Mean	Std. Dev.	Min	Max
hv271	1,774	8582.97362	-441030.8	133150	-654000	-175310

-> hv270 = richer

Variable	Obs	Weight	Mean	Std. Dev.	Min	Max
hv271	2,690	8582.03223	409289.8	322580.7	-175210	860240

-> hv270 = richest

Variable	Obs	Weight	Mean	Std. Dev.	Min	Max
hv271	2,605	8581.89077	1234530	312694.8	860620	3245770

Why does creating quintiles in Stata not match the principal component output?

Many thanks!
Amy