
Subject: Discrepancy in Estimates using NFHS5

Posted by [aarushikataria](#) on Thu, 09 Jun 2022 11:10:09 GMT

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

Hello,

I was trying to obtain an estimate of the proportion of women that consume alcohol using the NFHS5 Individual Recode. The below is the code I was running on STATA:

```
sampleweight = v005/1000000  
rename (s720 s721) (wifealc wdrinkfreq)  
egen strata = group(v024 v025), label  
* tab strata  
svyset [pweight=sampleweight], psu(v021) strata(strata)  
svy: tab wifealc
```

This gives me that 0.74% of women consume alcohol. However, the DHS and NFHS5 report say this number is 1.3%. On closer inspection, I found that the DHS report (<https://www.dhsprogram.com/pubs/pdf/FR375/FR375.pdf>) notes that the total number of observations in this dataset are 1,044,919 to arrive at 1.3%. However, the Individual recode has only 724,115 observations. Could the difference in observations be the reason for discrepancy?

Would be very grateful for any leads.
