
Subject: Re: Number of respondents mismatch in dataset and published report
Posted by [Bridgette-DHS](#) on Mon, 15 May 2017 15:22:20 GMT

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

Following is a response from Senior DHS Stata Specialist, Tom Pullum and DHS Senior Research Associate, Shireen Assaf:

The difference between the 6167 and the 5047 is completely due to the use of weights. Below I will paste the results in Stata, first without weights and second with weights, using BDIR70FL.dta. I do this with iweight but you can also do it with pweight, svyset, and svy.

To produce the weighted table in R, you would need to use the survey package. The following lines should help you get started.

```
library(survey)
# install and load the survey package
options(survey.lonely.psu="adjust")
# to fix the issue with strata with single PSU
# adjust will center the stratum at the population mean

data$wt = data$v005/1000000
# create the weight variable

mydesign<-svydesign(id=data$v021, data=data, strata=data$v023, weight=data$wt, nest=T)
# set the survey design using the uploaded data (named data), the cluster (v021), the strata
(v023) , and the weight (wt).
# then use svymean or svytable to get the weighted proportions and frequencies of your variables.
See the survey package documentation for further details.
```

File Attachments

1) [tab-v025.jpg](#), downloaded 740 times
