

---

Subject: Survey Sampling weights  
Posted by [rwmungai](#) on Wed, 08 Apr 2020 17:43:36 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Dear Forum Administrator,

I understand it is important to post questions in the Forum but I am unable to figure this out. That is why I sent an email.

```
gen wt=hv005/1000000
```

```
ta shstate [iw=wt]
```

This gives me a population of 189.7K. I expect to see a population in millions. So I decided as below

```
gen wt=hv005/1000
```

```
ta shstate [iw=wt]
```

This gives me a population of 189.7 million.

I want to get the "N" in terms of the total population and not the sample "n".

Thanks,  
Rose

---

---

Subject: Re: Survey Sampling weights  
Posted by [Liz-DHS](#) on Wed, 08 Apr 2020 18:45:50 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

A response from Dr. Tom Pullum:

Quote:

The weight variable in DHS surveys is normalized, which means that mean weight is 1, or, equivalently, the weighted total is equal to the unweighted total. HOWEVER, in order to get rid of a decimal point, the calculated weight is multiplied by one million and then rounded to the nearest integer.

When you use `iweights` you need to divide by one million. When you use `pweights`, as in a statistical model, you CAN divide by one million but you don't need to, because Stata automatically normalizes with `pweights`.

If you want inflation weights, which will give estimates of the population counts, you need to define new weights, and to do that you have to provide the population total. Say that the total number of households or persons in households or women 15-49 (depending on whether you are using the HR or PR or IR file) is N. You then multiply the weight by  $N/n$ , where n is the sample size. You also divide by one million and round to the nearest integer and then use `fweight`. Note that you cannot use standard errors, confidence intervals, or tests with `fweight`, only with `pweight`.

DHS advises against inflation weights. It can be difficult to find what is the population total on the date of the survey (the median date of interview, say). The sample includes only the household population and census estimates may be larger than the household population. It's unusual to

see any use of inflation weights, but if you really want them, that's how you calculate them.