
Subject: PSU and Strata identifiers - DV module Nepal

Posted by [Rashu](#) on Sun, 04 Aug 2019 00:50:10 GMT

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

I am doing research using the data from DHS 2016. I am only using a sample of women who participated in the domestic violence module.

I am trying to use the svyset option to weight my data. I found that v021 is labeled PSU and v022 is labeled as sample strata. So that I used the following command, but I am not sure if I used the correct one or not?

```
gen wgt=d005/1000000
svyset [pw=wgt], psu(v021) strata(v022)
```

```
    pweight: wgt
      VCE: linearized
Single unit: missing
  Strata 1: v022
    SU 1: v021
    FPC 1: <zero>
```

Thank You.

Subject: Re: PSU and Strata identifiers - DV module Nepal

Posted by [Bridgette-DHS](#) on Fri, 16 Aug 2019 18:21:24 GMT

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

Following is a response from DHS Research & Data Analysis Director, Tom Pullum:

In this survey, v022 and v023 are the same. They have different variable labels, but if you do "tab1 v022 v023" you will get identical distributions. Your specification of svyset is correct. However, you would get the same results with this:

```
svyset [pw=d005], psu(v021) strata(v022)
```

It isn't necessary to divide d005 by 1000000; Stata will automatically normalize the weights to have a mean of 1.
