

---

Subject: Re: How to Weight Data in R

Posted by [Bridgette-DHS](#) on Mon, 25 Mar 2013 14:52:33 GMT

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

---

Here is a response to your question, from one of our DHS experts, Tom Pullum.

We cannot offer much support for R.

Yes, V005 is always the weight variable.

The psu or cluster variable is V001 or V021. These are generally exactly the same--that is, they are duplicates. If in any doubt, use V001. There will typically be several hundred clusters. Your code used v021 as the stratification variable, and that would be a mistake.

The stratification variable is not always clearly identified, but in virtually all surveys the strata are the combinations of region (the first subnational unit) and urban/rural (always v025). Region and strata are usually given by v022, v023, or v024. (v101 is a duplicate of region.) Take a quick look at those three variables. There will typically be about twice as many strata as regions--often one less than twice as many, because the capital region may be completely urban. The number of strata will typically be in the range of 20 to 40.

If you have difficulty identifying the stratification variable for a specific survey, please contact DHS.

I hope this helps.

Bridgette-DHS

---