
Subject: Re: survey set preparation for Pakistan DHS 2012-13
Posted by [Reduced-For\(u\)m](#) on Mon, 25 Aug 2014 23:01:05 GMT
[View Forum Message](#) <> [Reply to Message](#)

This is a little tricky since I haven't used the Pakistan data, but the usual method, which seems to be applicable here, is to define strata as the combination of region and urban/rural (that is how the sample is stratified according to the Pakistan report (<http://dhsprogram.com/pubs/pdf/FR290/FR290.pdf>)).

First, your PSU is defined wrong - use v021. v001 is something else

As for the strata, the usual method is to choose region, separated by urban/rural, which is done thusly:

```
egen strata = group(v024 v025), label
```

This tells Stata that the survey is stratified by region and urban/rural. That sounds like what the Pakistan report is saying in the Appendix: <http://dhsprogram.com/pubs/pdf/FR290/FR290.pdf>

Check out the DHS FAQ (<http://dhsprogram.com/faq.cfm>) under "using data files"

There is a small chance that the Pakistan survey works differently. You may want to investigate how v022 relates to the strata as I defined them above. But v023 is almost certainly the wrong variable.
