Subject: Trend analysis with different number of 'PSU' and 'Strata'
Posted by Mason92 on Mon, 07 Mar 2022 11:59:03 GMT
View Forum Message <> Reply to Message

I plan to do trend analysis for some indices using two Pakistan DHSs conducted in 2012-13 and 2017-18. The sampling frame of PDHS 2017-18 was based on the 2017 census which included the population of AJK and FATA areas. While the PDHS 2012-13 was based on 1998 census; the PDHS 2012-13 did not include the AJK and FATA areas.

Secondly, the PDHS 2017-18 has 580 PSUs and 16 strata, while the PDHS 2012-13 has 500 PSUs and 10 strata. Would it be statistically okay to append both data files and conduct trend analysis?

Kindly guide!

Thank you.

Subject: Re: Trend analysis with different number of 'PSU' and 'Strata'
Posted by Bridgette-DHS on Mon, 07 Mar 2022 15:45:06 GMT
View Forum Message <> Reply to Message

Following is a response from DHS Lead Statistician, Ruilin Ren:

The Pakistan surveys are not easily combined because of their differences. First, you cannot simply append the two surveys data because of the normalization of the sampling weight, especially with the different coverage and different normalization polices. The 2012-13 survey did not include AJK and FATA regions, but the two regions are small, they together represent only about 5% of Pakistan. Second, the 2017-18 survey had a separate normalization policy, requested by the Government of Pakistan. GB and AJK were normalized separately and independently. If you want to use the 2017-18 survey data for all Pakistan, the denormalization of must be done in 3 parts, separately for GB, AJK, and the rest of Pakistan. We cannot help with that.

The recommendation is that you analyze trends using just the areas that were included in both the surveys. You can append the data files but note that the PSUs (clusters) in the two surveys were different.

Subject: Re: Trend analysis with different number of 'PSU' and 'Strata'
Posted by Mason92 on Mon, 07 Mar 2022 20:38:56 GMT
View Forum Message <> Reply to Message

Thank you so much for the prompt and very clear guidance, I am very grateful.

So, using the appended datafile, would it be correct to look at trends by looking at each province/region individually (or two provinces/regions combined) that were collected in both surveys by using the Stata commands:

. svy: logit i.sample d105a if v024 == 1
. svy: logit i.sample d105a if v024 == 1 2

* The newly created variable in the appended file "sample" will be coded as 0 if data from PDHS 2012-13, and 1 if data from PDHS 2017-18
* Variable v024 to be recoded as binary (0/1)

Finally, would it be statistically correct to analyze trends even if, for the same province/region, the number of clusters is different in the two surveys? Is there any rule-of-thumb when the number of clusters is different in two surveys?

Thank you.

Subject: Re: Trend analysis with different number of 'PSU' and 'Strata'
Posted by Bridgette-DHS on Tue, 08 Mar 2022 15:41:54 GMT
View Forum Message <> Reply to Message

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

I agree with the approach you describe. It's what I do when looking at trends. Regarding the question about clusters---these are census enumeration areas drawn at random from a sampling frame. The number of clusters within a region, as well as the specific clusters themselves, varies from one DHS survey to the next. The specification of svyset will build in adjustments for sample weights, clusters, and strata. After including svyset and svy, you don't need any further adjustments related to the clusters.

Subject: Re: Trend analysis with different number of 'PSU' and 'Strata'
Posted by Mason92 on Wed, 09 Mar 2022 13:36:18 GMT
View Forum Message <> Reply to Message

Thanks a lot, I am very grateful.

Subject: Re: Trend analysis with different number of 'PSU' and 'Strata'
Posted by David34 on Sat, 04 Jun 2022 11:01:07 GMT

View Forum Message <> Reply to Message

Hi,

Using the approach described below for trend analysis (.svy: logit i.sample d105a if v024 == 1); is it possible that 95% confidence intervals of two proportions overlap but the p-value of the trend is statistically significant? If not, then what could be the possible explanation please?

Thanks

Subject: Re: Trend analysis with different number of 'PSU' and 'Strata'
Posted by Janet-DHS on Mon, 06 Jun 2022 20:16:39 GMT
View Forum Message <> Reply to Message

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

You are correct. Two means (or proportions) can be significantly different even if the confidence intervals overlap. The reason for this is that the standard error of a difference is less than the sum of the two standard errors (the widths of the confidence intervals are proportional to the standard errors). In most situations, the standard error of a difference is equal to the square root of the sum of the squares of the two standard errors. It's like a right triangle, in which the hypotenuse is always shorter than the sum of the other two sides. Check a statistics text on tests and confidence intervals for a difference.

Subject: Re: Trend analysis with different number of 'PSU' and 'Strata'
Posted by David34 on Tue, 07 Jun 2022 16:24:01 GMT
View Forum Message <> Reply to Message

Thank you so much!

Subject: Re: Trend analysis with different number of 'PSU' and 'Strata'
Posted by jayevid on Thu, 14 Jul 2022 08:18:52 GMT
View Forum Message <> Reply to Message

The purpose of this analysis is to make statistically valid statements such as, "there was a significant linear decrease in the prevalence of high school aged americans who have ever smoked a cigarette across the period 1999-2011" with complex sample survey data.
vidmate.app saveinsta