Subject: separating string variables for pregnancy complication questions
Posted by chr8850 on Fri, 27 Jan 2017 03:09:31 GMT

Hi there,

I am working with the Afghanistan 2010 Mortality Survey. In the women's file there are a few questions (q406 and q427) from the maternity section that ask women what problems they had when they first had antenatal care and during delivery. As multiple responses are allowed the variable in the dataset is a long string of letters and there are several combinations of these (attached here a snapshot of part of the table generated when i run tab). Could you please advise on how these can be separated out in STATA so that I can count how many women had each of the symptoms mentioned for each pregnancy?

thanks

Ali

File Attachments
1) Screen Shot 2017-01-27 at 2.07.32 pm.png, downloaded 504 times

Subject: Re: separating string variables for pregnancy complication questions
Posted by schoumaker on Fri, 27 Jan 2017 08:41:14 GMT

Hello,
The few lines below will create new dichotomous variable (Q406_A, q406_B,...). There may be other ways to do this, but this should work.
Best,
Bruno

***

```
local resp "A B C D E F G H I J K L M N O P Q R S T U V W X Y Z"
foreach var of local resp {
gene q406_`var'=1 if regexm(q406,"`var'")==1
replace q406_`var'=0 if regexm(q406,"`var'")!=1 & q406!=""
}
```

Subject: Re: separating string variables for pregnancy complication questions
Posted by chr8850 on Thu, 16 Feb 2017 01:25:57 GMT

Many thanks Bruno, that worked perfectly