Subject: Re: Colombia Gender Scale
Posted by Janet-DHS on Fri, 29 Sep 2023 19:47:07 GMT

Following is a response from DHS staff member, Tom Pullum:

d109 is a timing variable, based on when the first event of d105a -d105h occurred (if any of them occurred). The NA cases are the women for whom d105a to d105h are NA. If you do "tab d105*,m" you will see that for each of the d105* variables there are 13,828 NA cases. In your work file you can just drop these cases (for example, with "drop if d105a==."). I would also drop the 152 women who are "don't know" for d109, but that's a judgment call. Then do "tab d109,m" and you see that there are 8,223 women who are censored, in the sense described before. They have not experienced violence up to the time of the survey but they are at risk of experiencing it in the future. You would construct a new variable that is 0 if d109 is "." (for those 8,223 censored cases) and 1 otherwise. (I have not used survival models myself recently and hope that this is not misleading.)

I found a youtube video that may help: https://www.youtube.com/watch?v=HnsJG42LxMo. I'm sure there are others, probably some using DHS data.