Subject: Matching most recent sexual activity (women) in India_2019/21
Posted by mahfuz.ru.pops@gmail.com on Wed, 03 Aug 2022 11:46:57 GMT
View Forum Message <> Reply to Message

Hi there,
I calculated the percentage of women whose last sexual intercourse was within the last 4 weeks in India using the NFHS-5 (2019-21). While calculating the percentage, I applied weight (V005/1000000) and tried different variables (such as V528 and V536). Nevertheless, results from none of the variables matched the result published in the report of NFHS-5 (please see the Table 6.9.1 in the report). Can anyone please provide me the code (preferably SPSS code) for calculating the percentage of women whose last sexual intercourse was within the last 4 weeks for India from NFHS-5 data? Thanks in advance.

Subject: Re: Matching most recent sexual activity (women) in India_2019/21
Posted by Bridgette-DHS on Fri, 19 Aug 2022 11:13:15 GMT
View Forum Message <> Reply to Message

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

I'll answer in Stata. I don't use SPSS but I think you will be able to see the logic. This variable is based on v527, which is one of the DHS variables that has the units in the first column and the number in the next two columns. Here are the possible units of time in the first column: 1 for days, 2 for weeks, 3 for months, 4 for years. The code 998 is used for "don't know / refused to answer". "tab v527" and "label list V527" will help you to see how v527 is coded. You also have to look at v536 to get "never".

Here are the Stata lines that I would use:

```
gen time=6 if v536==0
replace time=1 if v527<107
replace time=2 if v527>=107 & v527<128
replace time=3 if v527>=128 & v527<400
replace time=4 if v527>=400 & v527<998
replace time=5 if v527==998

label define time 1 "<1 week" 2 ">1 week, <1 month" 3 ">1 month, <1 year" 4 "1+ years" 5
"Missing" 6 "Never"
label values time time

gen time12=0 if time<=6
replace time12=1 if time<=2
label variable time12 "< 1 month"

gen time123=0 if time<=6
replace time123=1 if time<=3
label variable time123 "< 1 year"
```

tab1 time time12 time123 [iweight=v005/1000000]

This includes the construction of time12, which combines categories time=1 and time=2 to get "within last 4 weeks". Also, time123 combines categories 1, 2, and 3 of time to get "within 1 year". The table is a little unusual because of how it combines categories within the distribution. The values of v527 and time are mutually exclusive, so the combinations of categories must be constructed. It would take some special formatting to get the columns as they are in the report.

When I do this, the n is slightly off. I get n=107,956, whereas the report has n=108,014. The difference of 18 cases is very small and I can't take the time to reduce it. The distribution I get differs from the report, but this has to do with how the values of v527 are interpreted. The number of days (leading digit 1) goes from 0 to 90. The number of weeks (leading digit 2) goes from 1 to 50. The number of months (leading digit 3) goes from 1 to 11 (which is ok!); the number of years (leading digit 4) goes from 1 to 90. There are inconsistencies and ambiguities. I don't have the original program for this table and don't know how those cases were handled, but however they were handled, there is underlying ambiguity in the data. It is likely that in some combinations of unit and number it is the number that is wrong, and in other cases it is the unit that has been entered incorrectly. You could tinker with how v527 is recoded into the categorical variable I call "time".

Subject: Re: Matching most recent sexual activity (women) in India_2019/21
Posted by mahfuz.ru.pops@gmail.com on Mon, 22 Aug 2022 12:32:38 GMT
View Forum Message <> Reply to Message

Thanks a lot Bridgette, the logics make sense and I can use that. Thank you very much.

Subject: Re: Matching most recent sexual activity (women) in India_2019/21
Posted by mahfuz.ru.pops@gmail.com on Wed, 31 Aug 2022 13:27:32 GMT
View Forum Message <> Reply to Message

Hi Bridgette,
I ran the code exactly you gave me, but my output did not match with that you explained after the code (i don't know why). Although you suggested to use tab1 for calculating the frequencies and percentages, the tabe1 was not running in presence of "iweight". For this reason I used Table for calculating the frequencies and percentages and using "iweight". My entire output and codes are presented in the attachement.

Wednesday August 31 14:31:18 2022 Page 1
___ ____ ____ ____ ____(R)
/__ / ____/ / ____/
___/ / /___/ / /___/

Statistics/Data Analysis

1 . gen time=6 if v536==0
(723892 missing values generated)
2 .
3 . replace time=1 if v527<107
(33603 real changes made)
4 .
5 . replace time=2 if v527>=107 & v527<128
(5277 real changes made)
6 .
7 . replace time=3 if v527>=128 & v527<400
(30185 real changes made)
8 .
9 . replace time=4 if v527>=400 & v527<998
(7422 real changes made)
10 .
11 . replace time=5 if v527==998
(6172 real changes made)
12 .
13 .
14 .
15 . label define time 1 "<1 week" 2 ">1 week, <1 month" 3 ">1 month, <1 year" 4 "1+ years" 5 "Missing" 6 "Ne
label time already defined
r(110);
16 .
17 . label values time time
18 .
19 .
20 .
21 . gen time12=0 if time<=6
(641233 missing values generated)
22 .
23 . replace time12=1 if time<=2
(38880 real changes made)
24 .
25 . label variable time12 "< 1 month"
26 .
27 .
28 .
29 . gen time123=0 if time<=6
(641233 missing values generated)
30 .
31 . replace time123=1 if time<=3
(69065 real changes made)
Wednesday August 31 14:31:32 2022 Page 2
32 .
33 . label variable time123 "< 1 year"

34 . table time [iweight=v005/1000000]
time Freq.
<1 week 32,975.4
>1 week, <1 month 5,419.79
>1 month, <1 year 31,956.7
1+ years 8,070.33
Missing 4,909.21
Never 237.8367
35 . table time12 [iweight=v005/1000000]
< 1 month Freq.
0 45,174.1
1 38,395.2
36 .
37 . table time123 [iweight=v005/1000000]
< 1 year Freq.
0 13,217.4
1 70,351.9
38 .

File Attachments

Subject: Re: Matching most recent sexual activity (women) in India_2019/21
Posted by Bridgette-DHS on Wed, 31 Aug 2022 14:57:44 GMT
View Forum Message <> Reply to Message

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

This is mysterious! The "table" command in Stata is new to me, but when I run it I get exactly the same numbers as with "tab1". I do NOT get the numbers you get. I will paste my results below. Perhaps there is a difference in our versions of Stata? I am using 16 (I also have 17 but have not installed it). With tab1 you can also use "[fweight=v005]" to get the correct percentages, although the frequencies will have a factor of 1000000.

Your original posting asked for SPSS code. If you adapt this to SPSS, I hope it will work ok.

File Attachments

Subject: Re: Matching most recent sexual activity (women) in India_2019/21
Posted by mahfuz.ru.pops@gmail.com on Wed, 31 Aug 2022 17:26:14 GMT
View Forum Message <> Reply to Message

Actually I tried the same logic in SPSS, My Stata (13) and SPSS gave the same result. Now I'll try Stata 16 or 17. You have spent a lot of your valuable time and energy.....thanks a lot for your effort.

Subject: Re: Matching most recent sexual activity (women) in India_2019/21
Posted by Bridgette-DHS on Wed, 31 Aug 2022 18:59:25 GMT
View Forum Message <> Reply to Message

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

You are welcome.

If you are able to figure this out, please provide a brief post saying how you did it. Good luck!

Subject: Re: Matching most recent sexual activity (women) in India_2019/21
Posted by mahfuz.ru.pops@gmail.com on Tue, 27 Sep 2022 05:42:26 GMT
View Forum Message <> Reply to Message

Hi Brigette,
Currently I have used the updated NFHS-5 data and my calculations have matched with your calculations. Thanks for your assistance, and sorry for making confusion.

Best wishes
Mahfuz

Subject: Re: Matching most recent sexual activity (women) in India_2019/21
Posted by Mlue on Tue, 27 Sep 2022 23:53:21 GMT
View Forum Message <> Reply to Message

Hello,

Try this one. It will hopefully point you in the right direction. The denominator matches (not sure if my approach is correct), but not all the percentages (for the categories) match.

cls
clear all
set matsize 800
set maxvar 32000
set mem 1g

```
cd "C:\Users\..."
use caseid v000 v001 v002 v003 v004 v005 v006 v007 v008 v008a v009 v010 v011 v012 v013
v021 v022 v023 v024 v025 sdist v106 v107 v130 v131 v149 v150 v151 v152 v155 v157 v158
v159 v190 v312 v313 v384a v384b v384c v501 v502 v511 v513 v531 v527 v528 v529 v536 v714
v716 v717 v171a using"IAIR7DFL.dta", clear
set more off


****************************************************************************

sort v001 v002 v003

***
========================================================================
***

gen hhid =substr(caseid,1,12)
gen wt=v005/1000000
gen weight=wt
svyset v021 [pweight = wt], strata(v023) singleunit(centered)

***
========================================================================
***

** Timing of last sexual intercourse
cap label drop recent_sex
cap drop recent_sex
gen recent_sex=9
replace recent_sex = 0 if v527<107
replace recent_sex = 1 if v528 <= 28 & !inlist(recent_sex,0)
replace recent_sex = 2 if v527>=128 & v527<400 & !inlist(recent_sex,0,1)
replace recent_sex = 3 if v527>=400 & v527<994 & !inrange(recent_sex,1,2)
replace recent_sex = 4 if v531 == 0
replace recent_sex = 9 if v527==998
label var recent_sex "Timing of last sexual intercourse"
label define recent_sex 0"Within the last week" 1"Within the past 4 weeks" 2"Within 1 year" 3"One
or more years" 4"Never had sexual intercourse" 9"Missing"
label val recent_sex recent_sex

tab recent_sex if inlist(v171a,0,3) [iw=wt], m
svy: tab v190 recent_sex if inlist(v171a,0,3), percent format(%9.1f) miss row

svy: tab v190 recent_sex if inlist(v171a,0,3), count format(%9.0f) miss //Table 3.6 is the first time
we get "108,014": https://dhsprogram.com/pubs/pdf/FR375/FR375.pdf

***
========================================================================
***
```

```
cap label drop time
cap drop time*
gen time=6 if v536==0
replace time=1 if v527<107
replace time=2 if v527>=107 & v527<128
replace time=3 if v527>=128 & v527<400
replace time=4 if v527>=400 & v527<998
replace time=5 if v527==998
replace time=5 if time==.

label define time 1 "<1 week" 2 ">1 week, <1 month" 3 ">1 month, <1 year" 4 "1+ years" 5
"Missing" 6 "Never"
label values time time

gen time12=0 if time<=6
replace time12=1 if time<=2
label variable time12 "< 1 month"

gen time123=0 if time<=6
replace time123=1 if time<=3
label variable time123 "< 1 year"

tab1 time time12 time123 [iweight=v005/1000000], m

**************************

svy: tab v190 time if inlist(v171a,0,3), percent format(%9.1f) miss row

*******************************************************************************

keep if inlist(v171a,0,3)

***************************

svy: tab v190 time, percent format(%9.1f) miss row
svy: tab v190 recent_sex, percent format(%9.1f) miss row

*******************************************************************************
*******************************************************************************

exit
```

File Attachments