Statistics/Data Analysis

```
      name:  <unnamed>
       log:  E:\nhats\nhats_code\NHATS data setup\logs\/1-combine_waves_1_Mar_20
> 19.smcl
  log type:  text
 opened on:   1 Mar 2019, 12:31:48

. 
. 
. /*Combines rounds 1 and 2 sample person SP interview files,
> sensitive demo files and the round 2 cumulative tracker file case
> status 1 into a single file
> 
> Data format is multiple observations per subject, one for each round
> */
. ******************************************
. 
. forvalues w=1/$w {
  2. //round w
. use "${r`w'raw}NHATS_Round_`w'_SP_File.dta"
  3. //check to make sure sample ids are unique
. sort spid
  4. quietly by spid: gen dup = cond(_N==1,0,_n)
  5. tab dup
  6. capture gen wave=`w'
  7. la var wave "Survey wave"
  8. save round_`w'_1.dta, replace
  9. clear
 10. }

        dup |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |      8,245      100.00      100.00
------------+-----------------------------------
      Total |      8,245      100.00
file round_1_1.dta saved

        dup |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |      7,075      100.00      100.00
------------+-----------------------------------
      Total |      7,075      100.00
file round_2_1.dta saved

        dup |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |      5,799      100.00      100.00
------------+-----------------------------------
      Total |      5,799      100.00
file round_3_1.dta saved

        dup |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |      4,737      100.00      100.00
------------+-----------------------------------
      Total |      4,737      100.00
file round_4_1.dta saved

        dup |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |      8,334      100.00      100.00
------------+-----------------------------------
      Total |      8,334      100.00
file round_5_1.dta saved

        dup |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |      7,276      100.00      100.00
------------+-----------------------------------
      Total |      7,276      100.00
file round_6_1.dta saved
```

```
        dup |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |      6,312      100.00      100.00
------------+-----------------------------------
      Total |      6,312      100.00
file round_7_1.dta saved

. 
. //round 1
. forvalues w=1/$w{
  2. use round_`w'_1.dta
  3. 
. if `w'!=1 local pd pd*
  4. 
. //keep selected variables only
. local keepallwaves spid wave r`w'dresid w`w'varunit w`w'anfinwgt0 w`w'varstrat
>   ///
>         mo* r`w'd2intvrage hh`w'martlstat ///
>         ip`w'cmedicaid ip`w'mgapmedsp ip`w'nginsnurs ip`w'covmedcad ip`w'covtr
> icar ///
>         hh* hc* ss* pc* cp* cg* ha* sc* mc* sd* pa* hw* ///
>         is`w'* ht`w'placedesc fl`w'* ir* cm* ew* hp* sn* dt* `pd' gr* wa* r`w'
> dorigwksc ///
>         r`w'dnhatswksc r`w'dnhatsgrav r`w'dnhatsgrb wb* ho* cs*
  5. 
. if `w'==1 {
  6. keep `keepallwaves' r`w'dgender rl`w'dracehisp rl`w'spkothlan rl`w'condspan
> h el`w'higstschl ///
>         ia`w'*  re`w'resistrct re1dcensdiv va1serarmfor
  7. 
. rename (r`w'* w`w'* mo`w'* hh`w'* ip`w'* hc`w'* ss`w'* pc`w'* cp`w'* cg`w'* ha
> `w'* ///
> sc`w'* mc`w'* sd`w'* pa`w'* hw`w'* is`w'* ht`w'* fl`w'* ir`w'* cm`w'* ew`w'* h
> p`w'* ///
> sn`w'* dt`w'* gr`w'* wa`w'* wb`w'* ho`w'* cs`w'* rl`w'* el`w'* ia`w'* re`w'* v
> a`w'*) ///
> (*)
  8. destring dfavact, replace
  9. }
 10. 
. if `w'==2 {
 11. keep `keepallwaves' re2intplace re2newstrct re2spadrsnew re2dresistrct ///
>         re2dadrscorr re2dcensdiv ip2nginslast ep2eoltalk ep2poweratty ep2livng
> will
 12. 
. rename (r`w'* w`w'* mo`w'* hh`w'* ip`w'* hc`w'* ss`w'* pc`w'* cp`w'* cg`w'* ha
> `w'* ///
> sc`w'* mc`w'* sd`w'* pa`w'* hw`w'* is`w'* ht`w'* fl`w'* ir`w'* cm`w'* ew`w'* h
> p`w'* ///
> sn`w'* dt`w'* gr`w'* wa`w'* wb`w'* ho`w'* cs`w'* re`w'* pd`w'* ep`w'*) ///
> (*)
 13. }
 14. 
. if `w'==3 {
 15. keep `keepallwaves' re3intplace re3newstrct re3spadrsnew re3dresistrct ///
>         re3dcensdiv ip3nginslast ia*
 16. 
. rename (r`w'* w`w'* mo`w'* hh`w'* ip`w'* hc`w'* ss`w'* pc`w'* cp`w'* cg`w'* ha
> `w'* ///
> sc`w'* mc`w'* sd`w'* pa`w'* hw`w'* is`w'* ht`w'* fl`w'* ir`w'* cm`w'* ew`w'* h
> p`w'* ///
> sn`w'* dt`w'* gr`w'* wa`w'* wb`w'* ho`w'* cs`w'* re`w'* ia`w'* pd`w'*) ///
> (*)
 17. }
 18. if `w'==4 {
 19. keep `keepallwaves' re4intplace re4newstrct re4spadrsnew re4dresistrct ///
>         re4dcensdiv ip4nginslast
 20. 
```

```
21. rename (r`w'* w`w'* mo`w'* hh`w'* ip`w'* hc`w'* ss`w'* pc`w'* cp`w'* cg`w'* ha
  > `w'* ///
  > sc`w'* mc`w'* sd`w'* pa`w'* hw`w'* is`w'* ht`w'* fl`w'* ir`w'* cm`w'* ew`w'* h
  > p`w'* ///
  > sn`w'* dt`w'* gr`w'* wa`w'* wb`w'* ho`w'* cs`w'* re`w'* pd`w'*) ///
  > (*)
   21. }
   22. if `w'==5 {
   23. keep `keepallwaves' r`w'dgender rl`w'dracehisp rl`w'spkothlan rl`w'condspan
  > h el`w'higstschl re5intplace re5newstrct re5spadrsnew re5dresistrct ///
  >         re5dcensdiv ip5nginslast ia`w'* va5serarmfor w5an2011wgt0
   24.
22. rename (r`w'* w`w'* mo`w'* hh`w'* ip`w'* hc`w'* ss`w'* pc`w'* cp`w'* cg`w'* ha
  > `w'* ///
  > sc`w'* mc`w'* sd`w'* pa`w'* hw`w'* is`w'* ht`w'* fl`w'* ir`w'* cm`w'* ew`w'* h
  > p`w'* ///
  > sn`w'* dt`w'* gr`w'* wa`w'* wb`w'* ho`w'* cs`w'* rl`w'* el`w'* ia`w'* re`w'* v
  > a`w'* pd`w'*) ///
  > (*)
   25. }
   26. if `w'==6 {
   27. keep `keepallwaves' re6intplace re6newstrct re6spadrsnew re6dresistrct ///
  >         re6dcensdiv ip6nginslast w6an2011wgt0
   28.
23. rename (r`w'* w`w'* mo`w'* hh`w'* ip`w'* hc`w'* ss`w'* pc`w'* cp`w'* cg`w'* ha
  > `w'* ///
  > sc`w'* mc`w'* sd`w'* pa`w'* hw`w'* is`w'* ht`w'* fl`w'* ir`w'* cm`w'* ew`w'* h
  > p`w'* ///
  > sn`w'* dt`w'* gr`w'* wa`w'* wb`w'* ho`w'* cs`w'* re`w'* pd`w'*) ///
  > (*)
   29. }
   30. if `w'==7 {
   31. keep `keepallwaves' re7intplace re7newstrct re7spadrsnew re7dresistrct ///
  >         re7dcensdiv ip7nginslast ia`w'* w7an2011wgt0
   32.
24. rename (r`w'* w`w'* mo`w'* hh`w'* ip`w'* hc`w'* ss`w'* pc`w'* cp`w'* cg`w'* ha
  > `w'* ///
  > sc`w'* mc`w'* sd`w'* pa`w'* hw`w'* is`w'* ht`w'* fl`w'* ir`w'* cm`w'* ew`w'* h
  > p`w'* ///
  > sn`w'* dt`w'* gr`w'* wa`w'* wb`w'* ho`w'* cs`w'* re`w'* ia`w'* pd`w'*) ///
  > (*)
   33. }
   34.
25. save round_`w'_ltd.dta, replace
   35. }
  dfavact: all characters numeric; replaced as int
  file round_1_ltd.dta saved
  file round_2_ltd.dta saved
  file round_3_ltd.dta saved
  file round_4_ltd.dta saved
  file round_5_ltd.dta saved
  file round_6_ltd.dta saved
  file round_7_ltd.dta saved

26.
27. //check sensitive data files, keep only some variables, merge with ltd dataset
  > s
28. forvalues w=1/$w{
    2.         use "${r`w's}NHATS_Round_`w'_SP_Sen_Dem_File.dta"
    3.         sort spid
    4.         quietly by spid: gen dup = cond(_N==1,0,_n)
    5.         tab dup
    6.         clear
    7. }

        dup |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |      8,245      100.00      100.00
------------+-----------------------------------
      Total |      8,245      100.00
```

```
        dup |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |      7,075      100.00      100.00
------------+-----------------------------------
      Total |      7,075      100.00

        dup |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |      5,799      100.00      100.00
------------+-----------------------------------
      Total |      5,799      100.00

        dup |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |      4,737      100.00      100.00
------------+-----------------------------------
      Total |      4,737      100.00

        dup |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |      8,334      100.00      100.00
------------+-----------------------------------
      Total |      8,334      100.00

        dup |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |      7,276      100.00      100.00
------------+-----------------------------------
      Total |      7,276      100.00

        dup |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |      6,312      100.00      100.00
------------+-----------------------------------
      Total |      6,312      100.00

29.
30. //combine waves into single dataset
31. use round_1_ltd.dta

32. forvalues w=2/$w{
    2. append using round_`w'_ltd.dta
    3. }
  (note: variable dproxyid was str4, now str12 to accommodate using data's
         values)
  (note: variable dspouseid was str4, now str12 to accommodate using data's
         values)
  (label pa1dv2favact already defined)
  (label pa1dv2favact already defined)
  (note: variable dfavact was int, now double to accommodate using data's
         values)
  (label pa1dv2favact already defined)
  (label pa5dfavact already defined)
  (label pa5dfavact already defined)

33.
34. preserve

35.
36. forvalues w=1/$w{
    2. use "${r`w's}NHATS_Round_`w'_SP_Sen_Dem_File.dta", clear
    3. gen wave=`w'
    4. rename (r`w'* hh`w'*) (*)
    5. if `w'==1 {
    6. rename (rl`w'* hc`w'*) (*)
    7. }
    8. if `w'!=1 {
    9. rename (pd`w'*) (*)
   10. }
   11. save "${r`w's}NHATS_Round_`w'_SP_Sen_Dem_File_new.dta", replace
   12. }
  file E:\nhats\data\NHATS Sensitive\r1_sensitive\NHATS_Round_1_SP_Sen_Dem_File_ne
```

```
  > w.dta saved
  file E:\nhats\data\NHATS Sensitive\r2_sensitive\NHATS_Round_2_SP_Sen_Dem_File_ne
  > w.dta saved
  file E:\nhats\data\NHATS Sensitive\r3_sensitive\NHATS_Round_3_SP_Sen_Dem_File_ne
  > w.dta saved
  file E:\nhats\data\NHATS Sensitive\r4_sensitive\NHATS_Round_4_SP_Sen_Dem_File_ne
  > w.dta saved
  file E:\nhats\data\NHATS Sensitive\r5_sensitive\NHATS_Round_5_SP_Sen_Dem_File_ne
  > w.dta saved
  file E:\nhats\data\NHATS Sensitive\r6_sensitive\NHATS_Round_6_SP_Sen_Dem_File_ne
  > w.dta saved
  file E:\nhats\data\NHATS Sensitive\r7_sensitive\NHATS_Round_7_SP_Sen_Dem_File_ne
  > w.dta saved

37.
38. use "${r1s}NHATS_Round_1_SP_Sen_Dem_File_new.dta", clear

39.
40. restore

41.
42. //merge in sensitive data, use r1 as basis, added in cancer vars
43. merge m:1 spid wave using "${r1s}NHATS_Round_1_SP_Sen_Dem_File_new.dta", ///
  >         keepusing(dbirthmth dbirthyr dintvwrage modob yrdob dspousage ///
  >         primarace hisplatno cancerty*) nogen

      Result                           # of obs.
      -----------------------------------------
      not matched                        39,533
          from master                    39,533
          from using                          0

      matched                             8,245
      -----------------------------------------

44.
45. //merge in additional r2 sensitive data
46. merge m:1 spid using "${r2s}NHATS_Round_2_SP_Sen_Dem_File_new.dta", ///
  >         keepusing(dintvwrage dspousage ddeathage mthdied yrdied) nogen
  (label r1dbirthmth already defined)

      Result                           # of obs.
      -----------------------------------------
      not matched                        12,036
          from master                    12,036
          from using                          0

      matched                            35,742
      -----------------------------------------

47.
48. //merge in additional r3 sensitive data
49. merge m:1 spid using "${r3s}NHATS_Round_3_SP_Sen_Dem_File_new.dta", ///
  >         keepusing(dintvwrage dspousage ddeathage mthdied yrdied) nogen
  (label r1dbirthmth already defined)

      Result                           # of obs.
      -----------------------------------------
      not matched                        14,588
          from master                    14,588
          from using                          0

      matched                            33,190
      -----------------------------------------
```

```
50.
51. //merge in additional r4 sensitive data
52. merge m:1 spid using "${r4s}NHATS_Round_4_SP_Sen_Dem_File_new.dta", ///
  >         keepusing(dintvwrage dspousage ddeathage mthdied yrdied) nogen
  (label r1dbirthmth already defined)
  (label r1dbirthyr already defined)

      Result                           # of obs.
      -----------------------------------------
      not matched                        17,774
          from master                    17,774
          from using                          0

      matched                            30,004
      -----------------------------------------

53.
54. //merge in additional r5 sensitive data
55. merge m:1 spid using "${r5s}NHATS_Round_5_SP_Sen_Dem_File_new.dta", ///
  >         keepusing(dintvwrage spageall ddeathage mthdied yrdied) nogen

      Result                           # of obs.
      -----------------------------------------
      not matched                         9,248
          from master                     9,248
          from using                          0

      matched                            38,530
      -----------------------------------------

56.
57. //merge in additional r6 sensitive data
58. merge m:1 spid using "${r6s}NHATS_Round_6_SP_Sen_Dem_File_new.dta", ///
  >         keepusing(dintvwrage spageall ddeathage mthdied yrdied) nogen
  (label r5dbirthmth already defined)
  (label r5dbirthyr already defined)

      Result                           # of obs.
      -----------------------------------------
      not matched                        12,214
          from master                    12,214
          from using                          0

      matched                            35,564
      -----------------------------------------

59.
60. //merge in additional r7 sensitive data
61. merge m:1 spid using "${r7s}NHATS_Round_7_SP_Sen_Dem_File_new.dta", ///
  >         keepusing(dintvwrage spageall ddeathage mthdied yrdied) nogen
  (label r5dbirthmth already defined)
  (label r5dbirthyr already defined)

      Result                           # of obs.
      -----------------------------------------
      not matched                        15,926
          from master                    15,926
          from using                          0

      matched                            31,852
      -----------------------------------------
```

```
62.
63. /*
  > tempfile nhats
  > save "`nhats'"
  >
  >
  > //Going to tracker to rename variables.
  >
  > use "${r7raw}NHATS_Round_7_Tracker_File", clear
  >
  >
  > keep spid yearsample ///
  > w7trfinwgt0 w7tr2011wgt0 r7status r7spstat r7spstatdtmt r7spstatdtyr r7fqstatd
  > tmt ///
  > w6trfinwgt0 w6tr2011wgt0 r6status r6spstat r6spstatdtmt r6spstatdtyr r6fqstatd
  > tmt ///
  > w5trfinwgt0 w5tr2011wgt0 r5status r5spstat r5spstatdtmt r5spstatdtyr r5fqstatd
  > tmt ///
  > w4trfinwgt0                      r4status r4spstat r4spstatdtmt r4spstatdtyr r
  > 4fqstatdtmt ///
  > w3trfinwgt0                      r3status r3spstat r3spstatdtmt r3spstatdtyr r
  > 3fqstatdtmt ///
  > w2trfinwgt0                      r2status r2spstat r2spstatdtmt r2spstatdtyr r
  > 2fqstatdtmt ///
  > w1trfinwgt0                      r1status r1spstat r1spstatdtmt r1spstatdtyr r
  > 1fqstatdt*
  >
  > forvalues w=1/$w{
  >
  > rename (r`w'*  w`w'*) (*w`w')
  > }
  >
  > reshape long trfinwgt0w tr2011wgt0w statusw spstatw spstatdtmtw spstatdtyrw fq
  > statdtmtw fqstatdtyrw, i(spid) j(wave)
  >
  > rename (*w) (*)
  >
  > tempfile tracker
  > save "`tracker'"
  >
  > use "`nhats'", clear
  >
  > //3B?
  > //merge in tracker status information
  > merge 1:1 spid wave using "`tracker'"
  > */
64.
65. merge m:1 spid using "${r7raw}NHATS_Round_7_Tracker_File", keepusing( yearsamp
  > le ///
  > w7trfinwgt0 w7tr2011wgt0 r7status r7spstat r7spstatdtmt r7spstatdtyr r7fqstatd
  > tmt ///
  > w6trfinwgt0 w6tr2011wgt0 r6status r6spstat r6spstatdtmt r6spstatdtyr r6fqstatd
  > tmt ///
  > w5trfinwgt0 w5tr2011wgt0 r5status r5spstat r5spstatdtmt r5spstatdtyr r5fqstatd
  > tmt ///
  > w4trfinwgt0                      r4status r4spstat r4spstatdtmt r4spstatdtyr r
  > 4fqstatdtmt ///
  > w3trfinwgt0                      r3status r3spstat r3spstatdtmt r3spstatdtyr r
  > 3fqstatdtmt ///
  > w2trfinwgt0                      r2status r2spstat r2spstatdtmt r2spstatdtyr r
  > 2fqstatdtmt ///
  > w1trfinwgt0                      r1status r1spstat r1spstatdtmt r1spstatdtyr r
  > 1fqstatdt*)

    Result                           # of obs.
    -----------------------------------------
    not matched                         7,103
        from master                         0  (_merge==1)
        from using                      7,103  (_merge==2)
```

```
      matched                            47,778  (_merge==3)
      -----------------------------------------

66.
67. gen trfinwgt0=.
  (54,881 missing values generated)

68. gen tr2011wgt0=.
  (54,881 missing values generated)

69. forvalues w=1/$w{
    2. replace trfinwgt0= w`w'trfinwgt0 if wave==`w'
    3. capture replace tr2011wgt0 = w`w'tr2011wgt0 if wave==`w'
    4. capture drop w`w'trfinwgt0 w`w'tr2011wgt0
    5. }
  (8,245 real changes made)
  (7,075 real changes made)
  (5,799 real changes made)
  (4,737 real changes made)
  (8,334 real changes made)
  (7,276 real changes made)
  (6,312 real changes made)

70.
71.
72. //merge in nsoc tracker information
73. merge m:1 spid using "${r1s}NSOC_Round_1_SP_tracker_file", nogen

      Result                           # of obs.
      -----------------------------------------
      not matched                        17,969
          from master                    17,969
          from using                          0

      matched                            36,912
      -----------------------------------------

74. //drop obs that are in tracker file but not sp file
75. drop if _merge==2
  (7,103 observations deleted)

76. drop _merge

77. save round_1_to_$w.dta, replace
  file round_1_to_7.dta saved

78.
79. //old version of stata
80. forvalues w=1/$w{
    2. clear all
    3. use "E:\nhats\data\NHATS Public\round_`w'\NHATS_Round_`w'_SP_File.dta"
    4. saveold "E:\nhats\data\NHATS Public\round_`w'\NHATS_Round_`w'_SP_File_stata
  > 12.dta", replace version(12)
    5. }
  (saving in Stata 12 format, which can be read by Stata 11 or 12)
  file E:\nhats\data\NHATS Public\round_1\NHATS_Round_1_SP_File_stata12.dta saved
  (saving in Stata 12 format, which can be read by Stata 11 or 12)
    note: variable label "R2 F SS DEAF PRIOR OR CURRENT ROUND " contains unicode
          and thus may not display well in Stata 13.
  file E:\nhats\data\NHATS Public\round_2\NHATS_Round_2_SP_File_stata12.dta saved
  (saving in Stata 12 format, which can be read by Stata 11 or 12)
  file E:\nhats\data\NHATS Public\round_3\NHATS_Round_3_SP_File_stata12.dta saved
  (saving in Stata 12 format, which can be read by Stata 11 or 12)
  file E:\nhats\data\NHATS Public\round_4\NHATS_Round_4_SP_File_stata12.dta saved
  (saving in Stata 12 format, which can be read by Stata 11 or 12)
  file E:\nhats\data\NHATS Public\round_5\NHATS_Round_5_SP_File_stata12.dta saved
  (saving in Stata 12 format, which can be read by Stata 11 or 12)
  file E:\nhats\data\NHATS Public\round_6\NHATS_Round_6_SP_File_stata12.dta saved
  (saving in Stata 12 format, which can be read by Stata 11 or 12)
  file E:\nhats\data\NHATS Public\round_7\NHATS_Round_7_SP_File_stata12.dta saved
```

```
81.
82. ******************************************
83. log close
       name:  <unnamed>
        log:  E:\nhats\nhats_code\NHATS data setup\logs\/1-combine_waves_1_Mar_20
  > 19.smcl
    log type:  text
   closed on:   1 Mar 2019, 12:32:11
  -------------------------------------------------------------------------------
```