

```

-----
-----
-----
name: <unnamed>
log: E:\nlsy97_2022\mar_edu_age.log
log type: text
opened on: 14 Mar 2022, 10:29:15

```

```

.
.
. *****
. * Read in the data
. *****
.
.
. clear
.
. set more 1
.
. infile using "E:\nlsy97_2022\mar_edu_age.dct"

```

```

infile dictionary {
  E7013912 "2019 MAR: MAR STATUS MO L12"
  R0000100 "PUBID - YTH ID CODE 1997"
  R0536300 "KEY!SEX (SYMBOL) 1997"
  R0536401 "KEY!BDATE M/Y (SYMBOL) 1997"
  R0536402 "KEY!BDATE M/Y (SYMBOL) 1997"
  R1193900 "CV_AGE(MONTHS)_INT_DATE 1997"
  R1194000 "CV_AGE_12/31/96 1997"
  R1194100 "CV_AGE_INT_DATE 1997"
  R1235800 "CV_SAMPLE_TYPE 1997"
  R1236101 "R1 SAMPLE WEIGHT CC 1997"
  R1236201 "R1 SAMPLE WEIGHT PANEL 1997"
  R1482600 "KEY!RACE_ETHNICITY (SYMBOL) 1997"
  R1489700 "VARIANCE STRATUM 1997"
  R1489800 "VARIANCE PSU 1997"
  U1845400 "CV_HGC_EVER 2017"
  U1846000 "CV_HIGHEST_DEGREE_EVER 2017"
  U3455500 "R19 SAMPLE WEIGHT CC 2019"
  U3455600 "R19 SAMPLE WEIGHT PANEL 2019"
}

```

(8,984 observations read)

```

. do "E:\nlsy97_2022\mar_edu_age-value-labels.do"
.
. label define v1E7013912 0 "Never Married, Not Cohabiting" 1 "Never
Married, Cohabiting" 2 "Married" 3 "Legally Separated" 4 "Divorced" 5
"Widowed"
. label values E7013912 v1E7013912

```

```

.
. label define vlR0000100 0 "0" 1 "1 TO 999" 1000 "1000 TO 1999" 2000
"2000 TO 2999" 3000 "3000 TO 3999" 4000 "4000 TO 4999" 5000 "5000 TO
5999" 6000 "6000 TO 6
> 999" 7000 "7000 TO 7999" 8000 "8000 TO 8999" 9000 "9000 TO 9999"

. label values R0000100 vlR0000100

.
. label define vlR0536300 1 "Male" 2 "Female" 0 "No Information"

. label values R0536300 vlR0536300

.
. label define vlR0536401 1 "1: January" 2 "2: February" 3 "3: March"
4 "4: April" 5 "5: May" 6 "6: June" 7 "7: July" 8 "8: August" 9 "9:
September" 10 "10: O
> ctober" 11 "11: November" 12 "12: December"

. label values R0536401 vlR0536401

.
. label define vlR1194000 0 "0 TO 11: LESS THAN 12" 12 "12" 13 "13" 14
"14" 15 "15" 16 "16" 17 "17" 18 "18" 19 "19 TO 999: GREATER THAN 18"

. label values R1194000 vlR1194000

.
. label define vlR1194100 0 "0 TO 11: LESS THAN 12" 12 "12" 13 "13" 14
"14" 15 "15" 16 "16" 17 "17" 18 "18" 19 "19 TO 999: GREATER THAN 18"

. label values R1194100 vlR1194100

.
. label define vlR1235800 1 "Cross-sectional" 0 "Oversample"

. label values R1235800 vlR1235800

.
. label define vlR1236101 0 "0" 30000 "30000 TO 59999: 300.00-599.99"
60000 "60000 TO 99999: 600.00-999.99" 100000 "100000 TO 149999: 1000.00-
1499.99" 150000 "1500
> 00 TO 199999: 1500.00-1999.99" 200000 "200000 TO 249999: 2000.00-
2499.99" 250000 "250000 TO 299999: 2500.00-2999.99" 300000 "300000 TO
349999: 3000.00-3499.99" 3
> 50000 "350000 TO 399999: 3500.00-3999.99" 400000 "400000 TO 449999:
4000.00-4499.99" 450000 "450000 TO 499999: 4500.00-4999.99" 500000
"500000 TO 549999: 5000.00-
> 5499.99" 550000 "550000 TO 599999: 5500.00-5999.99" 600000 "600000 TO
649999: 6000.00-6499.99" 650000 "650000 TO 699999: 6500.00-6999.99"
700000 "700000 TO 74999
> 9: 7000.00-7499.99" 750000 "750000 TO 799999: 7500.00-7999.99" 800000
"800000 TO 849999: 8000.00-8499.99" 850000 "850000 TO 9999999: 8500.00+"

```

```
. label values R1236101 v1R1236101
```

```
.
. label define v1R1236201 0 "0" 30000 "30000 TO 59999: 300.00-599.99"
60000 "60000 TO 99999: 600.00-999.99" 100000 "100000 TO 149999: 1000.00-
1499.99" 150000 "1500
> 00 TO 199999: 1500.00-1999.99" 200000 "200000 TO 249999: 2000.00-
2499.99" 250000 "250000 TO 299999: 2500.00-2999.99" 300000 "300000 TO
349999: 3000.00-3499.99" 3
> 50000 "350000 TO 399999: 3500.00-3999.99" 400000 "400000 TO 449999:
4000.00-4499.99" 450000 "450000 TO 499999: 4500.00-4999.99" 500000
"500000 TO 549999: 5000.00-
> 5499.99" 550000 "550000 TO 599999: 5500.00-5999.99" 600000 "600000 TO
649999: 6000.00-6499.99" 650000 "650000 TO 699999: 6500.00-6999.99"
700000 "700000 TO 74999
> 9: 7000.00-7499.99" 750000 "750000 TO 799999: 7500.00-7999.99" 800000
"800000 TO 849999: 8000.00-8499.99" 850000 "850000 TO 99999999: 8500.00+"
```

```
. label values R1236201 v1R1236201
```

```
.
. label define v1R1482600 1 "Black" 2 "Hispanic" 3 "Mixed Race (Non-
Hispanic)" 4 "Non-Black / Non-Hispanic"
```

```
. label values R1482600 v1R1482600
```

```
.
. label define v1R1489700 0 "0" 1 "1 TO 9" 10 "10 TO 19" 20 "20 TO 29"
30 "30 TO 39" 40 "40 TO 49" 50 "50 TO 59" 60 "60 TO 69" 70 "70 TO 79"
80 "80 TO 89" 90
> "90 TO 99" 100 "100 TO 109" 110 "110 TO 119" 120 "120 TO 999999999:
120+"
```

```
. label values R1489700 v1R1489700
```

```
.
. label define v1R1489800 1 "1" 2 "2"
```

```
. label values R1489800 v1R1489800
```

```
.
. label define v1U1845400 0 "NONE" 1 "1ST GRADE" 2 "2ND GRADE" 3 "3RD
GRADE" 4 "4TH GRADE" 5 "5TH GRADE" 6 "6TH GRADE" 7 "7TH GRADE" 8
"8TH GRADE" 9 "9TH GRAD
> E" 10 "10TH GRADE" 11 "11TH GRADE" 12 "12TH GRADE" 13 "1ST YEAR
COLLEGE" 14 "2ND YEAR COLLEGE" 15 "3RD YEAR COLLEGE" 16 "4TH YEAR
COLLEGE" 17 "5TH YEAR COLLE
> GE" 18 "6TH YEAR COLLEGE" 19 "7TH YEAR COLLEGE" 20 "8TH YEAR COLLEGE
OR MORE" 95 "UNGRADED"
```

```
. label values U1845400 v1U1845400
```

```
.
```

```

. label define v1U1846000 0 "None" 1 "GED" 2 "High school diploma
(Regular 12 year program)" 3 "Associate/Junior college (AA)" 4
"Bachelor's degree (BA, BS)" 5 "M
> aster's degree (MA, MS)" 6 "PhD" 7 "Professional degree (DDS, JD,
MD)"

. label values U1846000 v1U1846000

.
. label define v1U3455500 0 "0" 30000 "30000 TO 59999: 300.00-599.99"
60000 "60000 TO 99999: 600.00-999.99" 100000 "100000 TO 149999: 1000.00-
1499.99" 150000 "1500
> 00 TO 199999: 1500.00-1999.99" 200000 "200000 TO 249999: 2000.00-
2499.99" 250000 "250000 TO 299999: 2500.00-2999.99" 300000 "300000 TO
349999: 3000.00-3499.99" 3
> 50000 "350000 TO 399999: 3500.00-3999.99" 400000 "400000 TO 449999:
4000.00-4499.99" 450000 "450000 TO 499999: 4500.00-4999.99" 500000
"500000 TO 549999: 5000.00-
> 5499.99" 550000 "550000 TO 599999: 5500.00-5999.99" 600000 "600000 TO
649999: 6000.00-6499.99" 650000 "650000 TO 699999: 6500.00-6999.99"
700000 "700000 TO 74999
> 9: 7000.00-7499.99" 750000 "750000 TO 799999: 7500.00-7999.99" 800000
"800000 TO 849999: 8000.00-8499.99" 850000 "850000 TO 9999999: 8500.00+"

. label values U3455500 v1U3455500

.
. label define v1U3455600 0 "0" 30000 "30000 TO 59999: 300.00-599.99"
60000 "60000 TO 99999: 600.00-999.99" 100000 "100000 TO 149999: 1000.00-
1499.99" 150000 "1500
> 00 TO 199999: 1500.00-1999.99" 200000 "200000 TO 249999: 2000.00-
2499.99" 250000 "250000 TO 299999: 2500.00-2999.99" 300000 "300000 TO
349999: 3000.00-3499.99" 3
> 50000 "350000 TO 399999: 3500.00-3999.99" 400000 "400000 TO 449999:
4000.00-4499.99" 450000 "450000 TO 499999: 4500.00-4999.99" 500000
"500000 TO 549999: 5000.00-
> 5499.99" 550000 "550000 TO 599999: 5500.00-5999.99" 600000 "600000 TO
649999: 6000.00-6499.99" 650000 "650000 TO 699999: 6500.00-6999.99"
700000 "700000 TO 74999
> 9: 7000.00-7499.99" 750000 "750000 TO 799999: 7500.00-7999.99" 800000
"800000 TO 849999: 8000.00-8499.99" 850000 "850000 TO 9999999: 8500.00+"

. label values U3455600 v1U3455600

. /* Crosswalk for Reference number & Question name
> * Uncomment and edit this RENAME statement to rename variables for
ease of use.
> * This command does not guarantee uniqueness
> */
. /* *start* */
. /*
> rename E7013912 MAR_STATUS_2019_12_XRND // MAR_STATUS_2019.12
> rename R0000100 PUBID_1997
> rename R0536300 KEY!SEX_1997

```

```

> rename R0536401 KEY!BDATE_M_1997
> rename R0536402 KEY!BDATE_Y_1997
> rename R1193900 CV_AGE(MONTHS)_INT_DATE_1997
> rename R1194000 CV_AGE_12_31_96_1997 // CV_AGE_12/31/96
> rename R1194100 CV_AGE_INT_DATE_1997
> rename R1235800 CV_SAMPLE_TYPE_1997
> rename R1236101 SAMPLING_WEIGHT_CC_1997
> rename R1236201 SAMPLING_PANEL_WEIGHT_1997
> rename R1482600 KEY!RACE_ETHNICITY_1997
> rename R1489700 VSTRAT_1997
> rename R1489800 VPSU_1997
> rename U1845400 CV_HGC_EVER_EDT_2017
> rename U1846000 CV_HIGHEST_DEGREE_EVER_EDT_2017
> rename U3455500 SAMPLING_WEIGHT_CC_2019
> rename U3455600 SAMPLING_PANEL_WEIGHT_2019
>
> */
. /* *end* */
. /* To convert variable names to lower case use the TOLOWER command
> * (type findit tolower and follow the links to install).
> * TOLOWER VARLIST will change listed variables to lower case;
> * TOLOWER without a specified variable list will convert all
variables in the dataset to lower case
> */
. /* tolower */
.
end of do-file

.
.
.
. rename *, lower

. save "E:\nlsy97_2022\mar_edu_age.dta", replace
file E:\nlsy97_2022\mar_edu_age.dta saved

.
.
. *****
. * Analyze the data
. *****

.
. *****
. * Code the variable
. *****

.
. *****
. * Marital Status
. *****
. tab1 e7013912

-> tabulation of e7013912

```

2019 MAR: MAR STATUS MO L12	Freq.	Percent	Cum.
-4	6,033	67.15	67.15
-3	21	0.23	67.39
Never Married, Not Cohabiting	780	8.68	76.07
Never Married, Cohabiting	412	4.59	80.65
Married	1,309	14.57	95.22
Legally Separated	54	0.60	95.83
Divorced	370	4.12	99.94
Widowed	5	0.06	100.00
Total	8,984	100.00	

```
. tab1 e7013912, nol
```

```
-> tabulation of e7013912
```

2019 MAR: MAR STATUS MO L12	Freq.	Percent	Cum.
-4	6,033	67.15	67.15
-3	21	0.23	67.39
0	780	8.68	76.07
1	412	4.59	80.65
2	1,309	14.57	95.22
3	54	0.60	95.83
4	370	4.12	99.94
5	5	0.06	100.00
Total	8,984	100.00	

```
.
```

```
. gen mar = .
```

```
(8,984 missing values generated)
```

```
. replace mar = 1 if inlist(e7013912,2,3)
```

```
(1,363 real changes made)
```

```
. replace mar = 0 if inlist(e7013912,0,1,4,5)
```

```
(1,567 real changes made)
```

```
. label variable mar "e7013912: marital status as of December, 2019"
```

```
. label define mar 1 "married" 0 "not married"
```

```
. label value mar mar
```

```
.
```

```
. tab2 e7013912 mar, mis
```

```
-> tabulation of e7013912 by mar
```

		e7013912: marital status as of			
2019 MAR: MAR STATUS		December, 2019			
MO L12		not marri	married	.	Total
-4		0	0	6,033	6,033
-3		0	0	21	21
Never Married, Not Co		780	0	0	780
Never Married, Cohabi		412	0	0	412
Married		0	1,309	0	1,309
Legally Separated		0	54	0	54
Divorced		370	0	0	370
Widowed		5	0	0	5
Total		1,567	1,363	6,054	8,984

```
.
.
. *****
. * Age at the baseline
. *****
. tab1 r1194000, mis
```

-> tabulation of r1194000

CV_AGE_12/31/96 1997		Freq.	Percent	Cum.
12		1,771	19.71	19.71
13		1,807	20.11	39.83
14		1,841	20.49	60.32
15		1,874	20.86	81.18
16		1,691	18.82	100.00
Total		8,984	100.00	

```
.
. clonevar age_base = r1194000
. label variable age_base "r1194000:CV_AGE_12/31/96 1997"
.
.
. tab2 r1194000 age_base, mis
```

-> tabulation of r1194000 by age_base

CV_AGE_12/31/96 1997	r1194000:CV_AGE_12/31/96 1997					Total
	12	13	14	15	16	
12	1,771	0	0	0	0	1,771
13	0	1,807	0	0	0	1,807
14	0	0	1,841	0	0	1,841
15	0	0	0	1,874	0	1,874
16	0	0	0	0	1,691	1,691
Total	1,771	1,807	1,841	1,874	1,691	8,984

```

.
. *****
. * Education
. *****
.
. clonevar college = u1846000
. label variable college "recoded u1846000 for whether R's received a
college degree"
. recode college (-5/-3 =.) (0/2=0) (3/7=1)
(college: 8469 changes made)
. label define college 0 "without college degree" 1 "with college degree"
. label value college college
.
. tab2 u1846000 college, mis

```

-> tabulation of u1846000 by college

CV_HIGHEST_DEGREE_EVE	recoded u1846000 for whether R's received a college degree			Total
	R 2017	without c	with coll	
-5	0	0	2,250	2,250
-3	0	0	26	26
None	515	0	0	515
GED	862	0	0	862
High school diploma (2,692	0	0	2,692
Associate/Junior coll	0	598	0	598
Bachelor's degree (BA	0	1,352	0	1,352
Master's degree (MA,	0	540	0	540
PhD	0	51	0	51
Professional degree (0	98	0	98
Total	4,069	2,639	2,276	8,984

```

.
. *****
. * Weight variables
. *****
.
. gen vstrat = r1489700
. label variable vstrat "r1489700: strata weight"
.
. gen vpsu = r1489800
. label variable vpsu "r1489800: psu weight"
.

```

```

. gen weight_2019 = u3455500
. label variable weight_2019 "u3455500: cross-sectional sample weight for individual at 2019"
.
. gen panel_weight_2019 = u3455600
. label variable panel_weight_2019 "u3455600: panel sample weight for individual at 2019"
.
. *****
. * Identify the size of analytic sample
. *****
.
. egen missing = rowmiss(mar age_base college vstrat vpsu
panel_weight_2019)
. label variable missing "the number of variables with missing values"
. tab1 missing, mis
-> tabulation of missing

the number |
  of       |
variables  |
with       |
missing    |
values     |      Freq.    Percent    Cum.
-----+-----
          0 |      2,328    25.91    25.91
          1 |      4,982    55.45    81.37
          2 |      1,674    18.63   100.00
-----+-----
        Total |      8,984   100.00

.
.
. gen valid = 0
. replace valid = 1 if missing ==0
(2,328 real changes made)
. label variable valid "if the Rerspondent is in the analytic sample"
. label define valid 0 "excluded" 1 "included"
. label value valid valid

```

```

. tab1 valid, mis
-> tabulation of valid

      if the |
Rerspondent |
is in the   |
analytic    |
sample      |      Freq.      Percent      Cum.
-----+-----
excluded    |      6,656      74.09      74.09
included    |      2,328      25.91     100.00
-----+-----
Total       |      8,984     100.00

.
.
.
. *****
. * svyset command, using the cross-sectional weights at 2019
. *****
.
. svyset [pweight=weight_2019] , strata(vstrat) psu(vpsu)
singleunit(scaled)

      pweight: weight_2019
          VCE: linearized
Single unit: scaled
  Strata 1: vstrat
      SU 1: vpsu
      FPC 1: <zero>

.
. *****
. * analysis
. *****
.
.
. svy, subpop(valid):tab mar
(running tabulate on estimation sample)

Number of strata =      117      Number of obs =      8,984
Number of PSUs  =      234      Population size = 1,937,845,282
Subpop. no. obs =      2,328
Subpop. size    =    659,264,984
Design df       =      117

-----
e7013912: |
marital   |
status as |
of        |
December, |
2019     | proportion
-----+-----
not marr |      .491

```

```

married |      .509
      |
Total |      1
-----

```

Key: proportion = cell proportion

```

. svy, subpop(valid):tab age_base
(running tabulate on estimation sample)

```

Number of strata	=	117	Number of obs	=	8,984
Number of PSUs	=	234	Population size	=	1,937,845,282
			Subpop. no. obs	=	2,328
			Subpop. size	=	659,264,984
			Design df	=	117

```

-----
r1194000: |
CV_AGE_12 |
/31/96 |
1997 | proportion
-----
      12 |      .2005
      13 |      .1959
      14 |      .2037
      15 |      .2046
      16 |      .1953
      |
Total |      1
-----

```

Key: proportion = cell proportion

```

. svy, subpop(valid):tab college
(running tabulate on estimation sample)

```

Number of strata	=	117	Number of obs	=	8,984
Number of PSUs	=	234	Population size	=	1,937,845,282
			Subpop. no. obs	=	2,328
			Subpop. size	=	659,264,984
			Design df	=	117

```

-----
recoded |
u1846000 |
for |
whether |
R's |
received |
a college |
degree | proportion
-----
without |      .5884
with col |      .4116
      |
Total |      1
-----

```

Key: proportion = cell proportion

```

. svy, subpop(valid):logit mar age_base college
(running logit on estimation sample)

```

Survey: Logistic regression

Number of strata	=	117	Number of obs	=	8,984
Number of PSUs	=	234	Population size	=	1,937,845,282
			Subpop. no. obs	=	2,328
			Subpop. size	=	659,264,984
			Design df	=	117
			F(2, 116)	=	44.17
			Prob > F	=	0.0000

```
-----
```

	mar	Coef.	Linearized Std. Err.	t	P> t	[95% Conf. Interval]	
	age_base	.1165866	.0342751	3.40	0.001	.0487066	.1844665
	college	.8675147	.103883	8.35	0.000	.6617797	1.07325
	_cons	-1.948684	.4804396	-4.06	0.000	-2.90017	-.9971987

```
-----
```

```
.
.
. *****
. * svyset command, using the panel weights at 2019
. *****
```

```
. svyset [pweight=panel_weight_2019] , strata(vstrat) psu(vpsu)
singleunit(scaled)
```

```
      pweight: panel_weight_2019
           VCE: linearized
Single unit: scaled
  Strata 1: vstrat
      SU 1: vpsu
     FPC 1: <zero>
```

```
. svy, subpop(valid):tab mar
(running tabulate on estimation sample)
```

```
Number of strata   =      117           Number of obs     =      8,984
Number of PSUs    =      234           Population size   = 1,937,845,320
                                           Subpop. no. obs  =      1,424
                                           Subpop. size     = 636,121,055
                                           Design df        =      117
```

```
-----
e7013912: |
marital   |
status as |
of        |
December, |
2019     | proportion
-----+-----
not marr  |      .479
married  |      .521
         |
Total    |      1
-----
```

```
Key: proportion = cell proportion
```

```
. svy, subpop(valid):tab age_base
(running tabulate on estimation sample)
```

```
Number of strata   =      117           Number of obs     =      8,984
Number of PSUs    =      234           Population size   = 1,937,845,320
                                           Subpop. no. obs  =      1,424
                                           Subpop. size     = 636,121,055
                                           Design df        =      117
```

```
-----
r1194000: |
CV_AGE_12  |
/31/96    |
1997     | proportion
-----+-----
12       |      .2093
```

```

      13 |      .1974
      14 |      .1994
      15 |      .2016
      16 |      .1923
      |
Total |      1
-----
Key:  proportion = cell proportion

. svy, subpop(valid):tab college
(running tabulate on estimation sample)

Number of strata =      117      Number of obs =      8,984
Number of PSUs  =      234      Population size = 1,937,845,320
                                   Subpop. no. obs =      1,424
                                   Subpop. size  =    636,121,055
                                   Design df     =      117

-----
recoded |
u1846000 |
for |
whether |
R's |
received |
a college |
degree | proportion
-----
  without |      .5586
  with col |      .4414
  Total |      1
-----
Key:  proportion = cell proportion

.
. svy, subpop(valid):logit mar age_base college
(running logit on estimation sample)

Survey: Logistic regression

Number of strata =      117      Number of obs =      8,984
Number of PSUs  =      234      Population size = 1,937,845,320
                                   Subpop. no. obs =      1,424
                                   Subpop. size  =    636,121,055
                                   Design df     =      117
                                   F( 2, 116) =      31.47
                                   Prob > F     =      0.0000

-----
      |      Linearized
      mar |      Coef.   Std. Err.   t   P>|t|   [95% Conf. Interval]
-----+-----
  age_base |   .1322662   .0360473   3.67  0.000   .0608764   .203656
  college |   .8208087   .1238954   6.63  0.000   .5754404   1.066177
  _cons |  -2.120497   .5029905  -4.22  0.000  -3.116643  -1.124351
-----

.
.
.
.
. log close
  name: <unnamed>
  log: E:\nlsy97_2022\mar_edu_age.log
  log type: text
  closed on: 14 Mar 2022, 10:29:16

```