Statistics/Data Analysis

```
1 .
2 .
3 . * import 2012 data
4 . import delimited "2012/NFCS 2012 State Data 130503.csv", ///
  >         case(preserve) clear
  (123 vars, 25509 obs)

5 .
6 . * keep only id and education since education is missing in the tracking data
7 . keep ID A5_2012

8 . save merge_2012.dta, replace
  file merge_2012.dta saved

9 .
10. * import 2018 tracking data
11. import delimited "2018/NFCS 2018 State Tracking Data 190623.csv", ///
  >         case(preserve) clear
  (113 vars, 108310 obs)

12.
13. * year
14. gen year = TRACK

15. drop if year == 2009 // drop 2009 data because education cannot be merged
  (28,146 observations deleted)

16.
17. * national weight
18. gen weights = wgt_n2

19.
20. * state
21. gen state_cate = STATEQ

22. tab state_cate, gen(state_dummy_)

   state_cate |      Freq.     Percent        Cum.
  ------------+-----------------------------------
            1 |      1,504        1.88        1.88
            2 |      1,501        1.87        3.75
            3 |      1,503        1.87        5.62
            4 |      1,501        1.87        7.50
            5 |      2,000        2.49        9.99
            6 |      1,501        1.87       11.86
            7 |      1,501        1.87       13.74
            8 |      1,504        1.88       15.61
            9 |      1,502        1.87       17.49
           10 |      1,500        1.87       19.36
           11 |      1,506        1.88       21.24
           12 |      1,501        1.87       23.11
           13 |      1,501        1.87       24.98
           14 |      2,015        2.51       27.49
           15 |      1,502        1.87       29.37
           16 |      1,505        1.88       31.24
           17 |      1,500        1.87       33.12
           18 |      1,504        1.88       34.99
           19 |      1,502        1.87       36.87
           20 |      1,505        1.88       38.74
           21 |      1,504        1.88       40.62
           22 |      1,501        1.87       42.49
           23 |      1,506        1.88       44.37
           24 |      1,502        1.87       46.24
           25 |      1,503        1.87       48.12
           26 |      1,504        1.88       50.00
           27 |      1,501        1.87       51.87
           28 |      1,505        1.88       53.74
           29 |      1,504        1.88       55.62
           30 |      1,502        1.87       57.49
           31 |      1,510        1.88       59.38
           32 |      1,501        1.87       61.25
```

33	2,000	2.49	63.75
34	1,505	1.88	65.62
35	1,500	1.87	67.49
36	1,501	1.87	69.37
37	1,500	1.87	71.24
38	2,251	2.81	74.05
39	1,508	1.88	75.93
40	1,501	1.87	77.80
41	1,504	1.88	79.68
42	1,502	1.87	81.55
43	1,507	1.88	83.43
44	2,000	2.49	85.92
45	1,502	1.87	87.80
46	1,501	1.87	89.67
47	1,514	1.89	91.56
48	2,254	2.81	94.37
49	1,500	1.87	96.24
50	1,512	1.89	98.13
51	1,501	1.87	100.00
Total	80,164	100.00	

```
23.
24. * census division and region (for re-weighting)
25. gen cen_div_cate = CENSUSDIV

26. gen cen_reg_cate = CENSUSREG

27.
28. * age (group mean)
29. gen age_cate = A3Ar_w

30. gen age = 20 if age_cate == 1
  (71,738 missing values generated)

31. replace age = 30 if age_cate == 2
  (13,983 real changes made)

32. replace age = 40 if age_cate == 3
  (13,397 real changes made)

33. replace age = 50 if age_cate == 4
  (14,912 real changes made)

34. replace age = 60 if age_cate == 5
  (14,557 real changes made)

35. replace age = 70 if age_cate == 6
  (14,889 real changes made)

36.
37. * gender
38. gen female_dummy = A3 == 2

39.
40. * age/gender
41. gen age_gender_cate = A3B

42.
43. * race
```

```
44. gen nonwhite_dummy = A4A_new_w == 2

45.
46. * marital status
47. gen marital_dummy = A6 == 1

48.
49. * education
50. *** merge with 2012 data
51. merge 1:1 ID using "merge_2012.dta", keep(1 3) nogen

    Result                           # of obs.
    -----------------------------------------
    not matched                         54,655
        from master                     54,655
        from using                           0

    matched                             25,509
    -----------------------------------------

52. *** generate a new variable to unite different codings
53. gen educ_cate = 1 if A5_2015 == 1 | A5_2012 == 1
  (76,973 missing values generated)

54. replace educ_cate = 2 if A5_2015 == 2 | A5_2012 == 2
  (14,313 real changes made)

55. replace educ_cate = 3 if A5_2015 == 3 | A5_2012 == 3
  (5,212 real changes made)

56. replace educ_cate = 4 if A5_2015 == 4 | A5_2012 == 4
  (23,421 real changes made)

57. replace educ_cate = 5 if A5_2015 == 5 | A5_2015 == 6 | A5_2012 == 5
  (23,454 real changes made)

58. replace educ_cate = 6 if A5_2015 == 7 | A5_2012 == 6
  (10,573 real changes made)

59. *** generate high school, college, and graduate dummy
60. gen high_school_dummy = 0 if educ_cate == 1
  (76,973 missing values generated)

61. replace high_school_dummy = 1 if educ_cate > 1 & educ_cate < 7
  (76,973 real changes made)

62. gen college_dummy = 0 if educ_cate < 5
  (34,027 missing values generated)

63. replace college_dummy = 1 if educ_cate > 4 & educ_cate < 7
  (34,027 real changes made)

64. gen graduate_dummy = 0 if educ_cate < 6
  (10,573 missing values generated)

65. replace graduate_dummy = 1 if educ_cate == 6
  (10,573 real changes made)

66.
67. * income (group mean)
```

```
68. gen income_cate = A8

69. gen income = 7500 if income_cate == 1
  (70,578 missing values generated)

70. replace income = 20000 if income_cate == 2
  (8,773 real changes made)

71. replace income = 30000 if income_cate == 3
  (8,808 real changes made)

72. replace income = 42500 if income_cate == 4
  (11,716 real changes made)

73. replace income = 62500 if income_cate == 5
  (15,776 real changes made)

74. replace income = 87500 if income_cate == 6
  (10,690 real changes made)

75. replace income = 125000 if income_cate == 7
  (9,672 real changes made)

76. replace income = 200000 if income_cate == 8
  (5,143 real changes made)

77.
78. * precautionary saving (treat DK and Refused as do not have precautionary saving)
79. gen precaution_dummy = J5 == 1

80.
81. * retirement plan
82. destring(J8), replace
  J8 has all characters numeric; replaced as byte
  (17195 missing values generated)

83. gen retire_young_dummy = J8 == 1

84. destring(J9), replace
  J9 has all characters numeric; replaced as byte
  (62969 missing values generated)

85. gen retire_old_dummy = J9 == 1

86. gen retire_dummy = retire_young_dummy

87. replace retire_dummy = retire_old_dummy if retire_young_dummy == .
  (0 real changes made)

88.
89. * financial market participation (treat missing as do not participate)
90. destring(B14), replace
  B14 has all characters numeric; replaced as byte
  (4194 missing values generated)

91. gen fin_par_dummy = B14 == 1

92. replace fin_par_dummy = 0 if B14 == .
  (0 real changes made)
```

```
93.
94. * perceived financial literacy (treat DK and Refused as neutrual)
95. gen math_perceived_cate = M1_2 if M1_2 != 98 & M1_2 != 99
  (761 missing values generated)

96. replace math_perceived_cate = 4 if M1_2 == 98 | M1_2 == 99
  (761 real changes made)

97. gen fin_perceived_cate = M4 if M4 != 98 & M4 != 99
  (2,080 missing values generated)

98. replace fin_perceived_cate = 4 if M4 == 98 | M4 == 99
  (2,080 real changes made)

99.
100 * true financial literacy (1 - correct; 2 - DK/Refused; 3 - incorrect)
101 *** interest rate question
102 gen interest_q = 1 if M6 == 1
  (19,176 missing values generated)

103 replace interest_q = 2 if M6 == 98 | M6 == 99
  (9,393 real changes made)

104 replace interest_q = 3 if M6 == 2 | M6 == 3
  (9,783 real changes made)

105 *** inflation question
106 gen inflation_q = 1 if M7 == 3
  (31,101 missing values generated)

107 replace inflation_q = 2 if M7 == 98 | M7 == 99
  (16,443 real changes made)

108 replace inflation_q = 3 if M7 == 1 | M7 == 2
  (14,658 real changes made)

109 *** bond price question
110 gen bond_q = 1 if M8 == 2
  (57,172 missing values generated)

111 replace bond_q = 2 if M8 == 98 | M8 == 99
  (30,430 real changes made)

112 replace bond_q = 3 if M8 == 1 | M8 == 3 | M8 == 4
  (26,742 real changes made)

113 /*
  > *** compounded interest rate question
  > gen compound_q = 1 if M31 == 2
  > replace compound_q = 2 if M31 == 98 | M31 == 99
  > replace compound_q = 3 if M31 == 1 | M31 == 3 | M31 == 4
  > */
114 *** mortgage question
115 gen mortgage_q = 1 if M9 == 1
  (18,435 missing values generated)

116 replace mortgage_q = 2 if M9 == 98 | M9 == 99
  (12,284 real changes made)

117 replace mortgage_q = 3 if M9 == 2
  (6,151 real changes made)
```

```
118 *** mutual funds question
119 gen mutual_q = 1 if M10 == 2
  (41,300 missing values generated)

120 replace mutual_q = 2 if M10 == 98 | M10 == 99
  (34,258 real changes made)

121 replace mutual_q = 3 if M10 == 1
  (7,042 real changes made)

122
123 * summary statistics
124 summ age female_dummy nonwhite_dummy marital_dummy high_school_dummy ///
  >         college_dummy graduate_dummy income precaution_dummy retire_dummy ///
  >         fin_par_dummy math_perceived_cate fin_perceived_cate interest_q inflation_q
  > ///
  >         bond_q mortgage_q mutual_q [aw = weights]

    Variable |     Obs      Weight        Mean   Std. Dev.       Min        Max
-------------+-----------------------------------------------------------------
         age |  80,164       80164    46.34164   16.52453         20         70
female_dummy |  80,164       80164    .5136688   .4998162          0          1
nonwhite_d~y |  80,164       80164    .3500791   .4769974          0          1
marital_du~y |  80,164       80164    .5226527   .4994897          0          1
high_schoo~y |  80,164       80164    .9537342   .2100614          0          1
-------------+-----------------------------------------------------------------
college_du~y |  80,164       80164    .3553699   .4786282          0          1
graduate_d~y |  80,164       80164    .1049297   .3064647          0          1
      income |  80,164       80164    62054.31   49231.67       7500     200000
precaution~y |  80,164       80164    .4490168    .497397          0          1
retire_dummy |  80,164       80164    .3089863   .4620784          0          1
-------------+-----------------------------------------------------------------
fin_par_du~y |  80,164       80164     .313935   .4640932          0          1
math_perce~e |  80,164       80164    5.570517    1.63434          1          7
fin_percei~e |  80,164       80164    5.132684   1.300811          1          7
  interest_q |  80,164       80164    1.390386   .7069398          1          3
 inflation_q |  80,164       80164    1.613064   .7957813          1          3
-------------+-----------------------------------------------------------------
      bond_q |  80,164       80164    2.070511   .7839531          1          3
  mortgage_q |  80,164       80164    1.342482     .63016          1          3
    mutual_q |  80,164       80164    1.640837   .6545402          1          3

125
126 * export data
127 keep year ID weights state_cate state_dummy_* cen_div_cate cen_reg_cate ///
  >         age_cate age female_dummy age_gender_cate nonwhite_dummy marital_dummy ///
  >         educ_cate high_school_dummy college_dummy graduate_dummy ///
  >         income_cate income precaution_dummy retire_dummy fin_par_dummy ///
  >         math_perceived_cate fin_perceived_cate interest_q ///
  >         inflation_q bond_q mortgage_q mutual_q

128
129 save processed_NFCS.dta, replace
  file processed_NFCS.dta saved

130 export delimited using "processed_NFCS.csv", nolabel replace
  file processed_NFCS.csv saved

131
132 * stop capturing log and translate into pdf
133 log close process_NFCS
       name:  process_NFCS
        log:  D:\□□□□\□□□□□\Perspective\NFCS\process_log.smcl
   log type:  text
  closed on:   2 Jun 2020, 14:11:16
-------------------------------------------------------------------------------
  > ------------
```