```
      name:  <unnamed>
       log:  D:\Jason\workshop\Regression 2019\final\regression1.log
  log type:  text
 opened on:  20 Feb 2019, 14:16:10

. 
. set more 1

. 
. use "D:\Jason\workshop\Regression 2019\final\regression.dta", clear

. 
. ************************************
. * Continuous dependent variable
. ************************************
. 
. ******************
. * Questions 1:
. ******************
. 
.         ***************************************
.         * 1.1 Regression without complex survey data
.         ***************************************
. 
.                 ******************************************************
.                 * 1.1.0 Regression without using the weight variable
.                 ******************************************************
. 
.                 reg illness sex age

      Source |       SS           df       MS      Number of obs   =    10,351
-------------+----------------------------------   F(2, 10348)     =   1424.18
       Model |  3718.57668         2  1859.28834   Prob > F        =    0.0000
    Residual |  13509.5022    10,348  1.30551818   R-squared       =    0.2158
-------------+----------------------------------   Adj R-squared   =    0.2157
       Total |  17228.0788    10,350  1.66454868   Root MSE        =    1.1426

------------------------------------------------------------------------------
     illness |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         sex |   1.171145   .0224904    52.07   0.000     1.127059     1.21523
         age |   .0073366   .0006524    11.25   0.000     .0060578    .0086155
       _cons |  -.2641429   .0474133    -5.57   0.000    -.3570822   -.1712036
------------------------------------------------------------------------------

. 
.                 ***************************************************************************
.                 * 1.1.1. Regression using a personal weight variable for the whole sample
.                 ***************************************************************************
. 
.                 reg illness sex age [pw=finalwgt]
(sum of wgt is   1.1716e+08)

Linear regression                               Number of obs     =     10,351
                                                F(2, 10348)       =    1258.32
                                                Prob > F          =     0.0000
                                                R-squared         =     0.2429
                                                Root MSE          =     1.1101

------------------------------------------------------------------------------
             |               Robust
     illness |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         sex |   1.237503   .0255848    48.37   0.000     1.187352    1.287654
         age |    .006566   .0007423     8.84   0.000     .0051108    .0080211
       _cons |  -.3691744   .0500926    -7.37   0.000    -.4673656   -.2709832
------------------------------------------------------------------------------

. 
. 
```

```
.                 ********************************************************************* *********
.                 * 1,1,2. Regression using a personal weight variable and adjusted for
.                 *dependent observations for the whole sample
.                 ************************************************************************
.
.
.                 reg illness sex age [pw=finalwgt], cluster(company_id)
(sum of wgt is   1.1716e+08)

Linear regression                               Number of obs     =     10,351
                                                F(2, 61)          =    1028.84
                                                Prob > F          =     0.0000
                                                R-squared         =     0.2429
                                                Root MSE          =     1.1101

                            (Std. Err. adjusted for 62 clusters in company_id)
------------------------------------------------------------------------------
             |               Robust
     illness |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         sex |   1.237503   .0291051    42.52   0.000     1.179304    1.295702
         age |    .006566   .0009264     7.09   0.000     .0047136    .0084184
       _cons |  -.3691744   .0720101    -5.13   0.000    -.5131675   -.2251813
------------------------------------------------------------------------------


.
.
.                 ************************************************************************
.                 * 1.1.3. Regression using a personal weight variable and adjusted for
.                 *dependent observations for white respondents
.                 ************************************************************************
.
.                 reg illness sex age if race ==1 [pw=finalwgt], cluster(company_id)
(sum of wgt is   1.0300e+08)

Linear regression                               Number of obs     =      9,065
                                                F(2, 61)          =    1084.03
                                                Prob > F          =     0.0000
                                                R-squared         =     0.2460
                                                Root MSE          =     1.0866

                            (Std. Err. adjusted for 62 clusters in company_id)
------------------------------------------------------------------------------
             |               Robust
     illness |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         sex |   1.219168   .0270341    45.10   0.000      1.16511    1.273226
         age |   .0066663   .0009175     7.27   0.000     .0048316     .008501
       _cons |  -.3914528   .0711509    -5.50   0.000    -.5337279   -.2491778
------------------------------------------------------------------------------


.
.
.          ***************************************
.          * 1.2. Regression with complex survey data
.          ***************************************
.          svydes

Survey: Describing stage 1 sampling units

      pweight: finalwgt
          VCE: linearized
  Single unit: missing
     Strata 1: stratid
         SU 1: psuid
        FPC 1: <zero>

                                    #Obs per Unit
                              ----------------------------
Stratum    #Units     #Obs      min       mean      max
--------  --------  --------  --------  --------  --------
       1         2       380       165     190.0       215
```

```
       2         2       185        67      92.5       118
       3         2       348       149     174.0       199
       4         2       460       229     230.0       231
       5         2       252       105     126.0       147
       6         2       298       131     149.0       167
       7         2       476       206     238.0       270
       8         2       338       158     169.0       180
       9         2       244       100     122.0       144
      10         2       262       119     131.0       143
      11         2       275       120     137.5       155
      12         2       314       144     157.0       170
      13         2       342       154     171.0       188
      14         2       405       200     202.5       205
      15         2       380       189     190.0       191
      16         2       336       159     168.0       177
      17         2       393       180     196.5       213
      18         2       359       144     179.5       215
      20         2       285       125     142.5       160
      21         2       214       102     107.0       112
      22         2       301       128     150.5       173
      23         2       341       159     170.5       182
      24         2       438       205     219.0       233
      25         2       256       116     128.0       140
      26         2       261       129     130.5       132
      27         2       283       139     141.5       144
      28         2       299       136     149.5       163
      29         2       503       215     251.5       288
      30         2       365       166     182.5       199
      31         2       308       143     154.0       165
      32         2       450       211     225.0       239
--------  --------  --------  --------  --------  --------
      31        62    10,351        67     167.0       288

.         svyset, clear

.         svyset psuid [pweight=finalwgt], strata(stratid)

      pweight: finalwgt
          VCE: linearized
  Single unit: missing
     Strata 1: stratid
         SU 1: psuid
        FPC 1: <zero>

.
.
.                 ******************************************
.                 * 1.2.1. Regression using complex survey data  for the whole sample
.                 ******************************************
.
.                 svy: reg illness sex age
(running regress on estimation sample)

Survey: Linear regression

Number of strata   =        31                Number of obs     =       10,351
Number of PSUs     =        62                Population size   =  117,157,513
                                              Design df         =           31
                                              F(   2,     30)   =       798.46
                                              Prob > F          =       0.0000
                                              R-squared         =       0.2429

------------------------------------------------------------------------------
             |             Linearized
     illness |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         sex |   1.237503   .0312795    39.56   0.000     1.173708    1.301298
         age |    .006566   .0010271     6.39   0.000     .0044713    .0086607
       _cons |  -.3691744   .0757518    -4.87   0.000    -.5236712   -.2146775
------------------------------------------------------------------------------
```

```
.
.
.                 ************************************************************************
.                 * 1,2,2. Regression using complex survey data and adjusted for
.                 *dependent observations for the whole sample
.                 ************************************************************************
.
.                 *Cluster option is not alllowed with the svy command
.
.
.                 ************************************************************************
.                 * 1.2.3. Regression using a personal weight variable and adjusted for
.                 *dependent observations for white respondents only
.                 ************************************************************************
.
.                 *correct specification*
.
.                 gen white = 0

.                 replace white =1 if race ==1
(9,065 real changes made)

.                 svy, subpop(white): reg illness sex age
(running regress on estimation sample)

Survey: Linear regression

Number of strata   =        31                Number of obs     =       10,351
Number of PSUs     =        62                Population size   =  117,157,513
                                              Subpop. no. obs   =        9,065
                                              Subpop. size      =  102,999,549
                                              Design df         =           31
                                              F(   2,     30)   =       878.21
                                              Prob > F          =       0.0000
                                              R-squared         =       0.2460

------------------------------------------------------------------------------
             |             Linearized
     illness |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         sex |   1.219168   .0292974    41.61   0.000     1.159416     1.27892
         age |   .0066663   .0010312     6.46   0.000     .0045632    .0087694
       _cons |  -.3914528   .0745836    -5.25   0.000    -.5435671   -.2393386
------------------------------------------------------------------------------


.
.                 *incorrect specification*
.                 svy: reg illness sex age if race ==1 /* incorrect specification */
(running regress on estimation sample)

Survey: Linear regression

Number of strata   =        31                Number of obs     =        9,065
Number of PSUs     =        62                Population size   =  102,999,549
                                              Design df         =           31
                                              F(   2,     30)   =       878.21
                                              Prob > F          =       0.0000
                                              R-squared         =       0.2460

------------------------------------------------------------------------------
             |             Linearized
     illness |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         sex |   1.219168   .0292974    41.61   0.000     1.159416     1.27892
         age |   .0066663   .0010312     6.46   0.000     .0045632    .0087694
       _cons |  -.3914528   .0745836    -5.25   0.000    -.5435671   -.2393386
------------------------------------------------------------------------------


.
.
.
```

```
. ********************************************************************************
. * Questions 2 and 3: the equality and linear combination of regression coefficients
. ********************************************************************************
.
.
.         ***********************************************************************
.         * 2.1 Modifying Mdoel 1.1.3
.         ***********************************************************************
.         reg illness sex age if race ==1 [pw=finalwgt], cluster(company_id)
(sum of wgt is   1.0300e+08)

Linear regression                               Number of obs     =      9,065
                                                F(2, 61)          =    1084.03
                                                Prob > F          =     0.0000
                                                R-squared         =     0.2460
                                                Root MSE          =     1.0866

                           (Std. Err. adjusted for 62 clusters in company_id)
------------------------------------------------------------------------------
             |               Robust
     illness |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         sex |   1.219168   .0270341    45.10   0.000      1.16511    1.273226
         age |   .0066663   .0009175     7.27   0.000     .0048316     .008501
       _cons |  -.3914528   .0711509    -5.50   0.000    -.5337279   -.2491778
------------------------------------------------------------------------------

.
.         *testing the equlaity of two regression coefficients
.         test sex = age

 ( 1)  sex - age = 0

       F(  1,    61) = 1996.46
            Prob > F =    0.0000

.
.         *testing the hypothesis that variables cancel out each other
.         test sex = -age

 ( 1)  sex + age = 0

       F(  1,    61) = 2066.99
            Prob > F =    0.0000

.
.
.         ***********************************************************************
.         * 2.2 Regression Analysis Using Results from Model 1.2.3.
.         ***********************************************************************
.
.         svy, subpop(white): reg illness sex age hlthstat
(running regress on estimation sample)

Survey: Linear regression

Number of strata   =        31                Number of obs     =       10,349
Number of PSUs     =        62                Population size   =  117,131,111
                                              Subpop. no. obs   =        9,063
                                              Subpop. size      =  102,973,147
                                              Design df         =           31
                                              F(   3,     29)   =       721.51
                                              Prob > F          =       0.0000
                                              R-squared         =       0.2663

------------------------------------------------------------------------------
             |             Linearized
     illness |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         sex |   1.205585   .0290865    41.45   0.000     1.146263    1.264908
         age |   .0024924   .0011262     2.21   0.034     .0001955    .0047893
```

```
    hlthstat |   .1656665   .0128492    12.89   0.000     .1394603    .1918726
       _cons |   -.581727    .070513    -8.25   0.000    -.7255391   -.4379148
------------------------------------------------------------------------------

.
.         *testing the equlaity of two or three regression coefficients
.         test sex =  age

Adjusted Wald test

 ( 1)  sex - age = 0

       F(  1,    31) = 1708.35
            Prob > F =    0.0000

.
.         *testing the equlaity of two or three regression coefficients
.         test sex = -age

Adjusted Wald test

 ( 1)  sex + age = 0

       F(  1,    31) = 1722.43
            Prob > F =    0.0000

.
.
.
. ****************************************************************************
. * Questions 4: testing the interaction effects
. ****************************************************************************
.
.
.         ****************************************************************************
.         * 4.1.1 A two-way interaction: Modify the model 1.1.2
.         ****************************************************************************
.         reg illness i.sex##c.age, cluster(company_id)

Linear regression                               Number of obs     =     10,351
                                                F(3, 61)          =    1010.15
                                                Prob > F          =     0.0000
                                                R-squared         =     0.2219
                                                Root MSE          =     1.1382

                            (Std. Err. adjusted for 62 clusters in company_id)
------------------------------------------------------------------------------
             |               Robust
     illness |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         sex |
     Female  |   1.725664   .0720129    23.96   0.000     1.581665    1.869663
         age |   .0134893   .0009387    14.37   0.000     .0116122    .0153664
             |
   sex#c.age |
     Female  |  -.0116584   .0014024    -8.31   0.000    -.0144626   -.0088541
             |
       _cons |   .6152195   .0653779     9.41   0.000     .4844882    .7459508
------------------------------------------------------------------------------

.         reg,  coeflegend

Linear regression                               Number of obs     =     10,351
                                                F(3, 61)          =    1010.15
                                                Prob > F          =     0.0000
                                                R-squared         =     0.2219
                                                Root MSE          =     1.1382

                            (Std. Err. adjusted for 62 clusters in company_id)
------------------------------------------------------------------------------
     illness |      Coef.  Legend
```

```
------------+----------------------------------------------------------------
        sex |
     Female |   1.725664  _b[2.sex]
        age |   .0134893  _b[age]
            |
   sex#c.age |
     Female |  -.0116584  _b[2.sex#c.age]
            |
      _cons |   .6152195  _b[_cons]
------------------------------------------------------------------------------

.         test _b[2.sex#c.age] =0

 ( 1)  2.sex#c.age = 0

       F(  1,    61) =   69.11
            Prob > F =    0.0000


.
.         margins i.sex,  at(age=(20(10)80))

Adjusted predictions                            Number of obs     =     10,351
Model VCE    : Robust

Expression   : Linear prediction, predict()

1._at        : age             =          20

2._at        : age             =          30

3._at        : age             =          40

4._at        : age             =          50

5._at        : age             =          60

6._at        : age             =          70

7._at        : age             =          80

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     _at#sex |
     1#Male  |   .8850057   .0538992    16.42   0.000     .7772275    .9927839
   1#Female  |   2.377502   .0603358    39.40   0.000     2.256853    2.498151
     2#Male  |   1.019899   .0498489    20.46   0.000     .9202198    1.119578
   2#Female  |   2.395811   .0540707    44.31   0.000      2.28769    2.503932
     3#Male  |   1.154792   .0473384    24.39   0.000     1.060133    1.249451
   3#Female  |    2.41412   .0495605    48.71   0.000     2.315018    2.513223
     4#Male  |   1.289685   .0466173    27.67   0.000     1.196468    1.382902
   4#Female  |    2.43243   .0473099    51.41   0.000     2.337828    2.527032
     5#Male  |   1.424578   .0477667    29.82   0.000     1.329063    1.520093
   5#Female  |   2.450739   .0476401    51.44   0.000     2.355476    2.546001
     6#Male  |   1.559471   .0506594    30.78   0.000     1.458171    1.660771
   6#Female  |   2.469048   .0505006    48.89   0.000     2.368066     2.57003
     7#Male  |   1.694364   .0550212    30.79   0.000     1.584343    1.804386
   7#Female  |   2.487357   .0555014    44.82   0.000     2.376375    2.598339
------------------------------------------------------------------------------

.         marginsplot, yline(0)

  Variables that uniquely identify margins: age sex

.
.         ***********************************************************************
.         * 4.1.2 A three-way interactdion: Modify the model 1.2.3.
.         ***********************************************************************
.         reg illness i.sex##c.age##i.race [pw=finalwgt], cluster(company_id)
(sum of wgt is   1.1716e+08)
```

```
Linear regression                               Number of obs     =     10,351
                                                F(11, 61)         =     281.36
                                                Prob > F          =     0.0000
                                                R-squared         =     0.2603
                                                Root MSE          =     1.0977

                               (Std. Err. adjusted for 62 clusters in company_id)
------------------------------------------------------------------------------
             |               Robust
     illness |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         sex |
      Female |   1.659972   .0707731    23.45   0.000     1.518452    1.801491
         age |   .0121751   .0011102    10.97   0.000      .009955    .0143952
             |
     sex#c.age |
      Female |  -.0103753   .0014578    -7.12   0.000    -.0132904   -.0074603
             |
        race |
       Black |   .1250714   .1535165     0.81   0.418    -.1819039    .4320466
       Other |   .3733645    .611498     0.61   0.544    -.8494014     1.59613
             |
      sex#race |
 Female#Black |   .4175091    .254642     1.64   0.106     -.091679    .9266972
 Female#Other |  -.8134471   .8232953    -0.99   0.327    -2.459728    .8328336
             |
    race#c.age |
       Black |   .0070698   .0032279     2.19   0.032     .0006153    .0135243
       Other |  -.0090257   .0113165    -0.80   0.428    -.0316544     .013603
             |
sex#race#c.age |
 Female#Black |  -.0071912   .0054496    -1.32   0.192    -.0180882    .0037059
 Female#Other |   .0191441   .0171505     1.12   0.269    -.0151504    .0534387
             |
        _cons |   .5957215   .0687846     8.66   0.000     .4581781    .7332649
------------------------------------------------------------------------------

.         reg,  coeflegend

Linear regression                               Number of obs     =     10,351
                                                F(11, 61)         =     281.36
                                                Prob > F          =     0.0000
                                                R-squared         =     0.2603
                                                Root MSE          =     1.0977

                               (Std. Err. adjusted for 62 clusters in company_id)
------------------------------------------------------------------------------
     illness |      Coef.  Legend
-------------+----------------------------------------------------------------
         sex |
      Female |   1.659972  _b[2.sex]
         age |   .0121751  _b[age]
             |
     sex#c.age |
      Female |  -.0103753  _b[2.sex#c.age]
             |
        race |
       Black |   .1250714  _b[2.race]
       Other |   .3733645  _b[3.race]
             |
      sex#race |
 Female#Black |   .4175091  _b[2.sex#2.race]
 Female#Other |  -.8134471  _b[2.sex#3.race]
             |
    race#c.age |
       Black |   .0070698  _b[2.race#c.age]
       Other |  -.0090257  _b[3.race#c.age]
             |
sex#race#c.age |
 Female#Black |  -.0071912  _b[2.sex#2.race#c.age]
```

```
 Female#Other  |   .0191441  _b[2.sex#3.race#c.age]
               |
         _cons |   .5957215  _b[_cons]
------------------------------------------------------------------------------

.         test _b[2.sex#c.age] =0

 ( 1)  2.sex#c.age = 0

       F(  1,    61) =   50.65
            Prob > F =    0.0000

.
.
.         margins i.sex , over(i.race)  at(age=(20(10)80))

Predictive margins                              Number of obs     =     10,351
Model VCE    : Robust

Expression   : Linear prediction, predict()
over         : race

1._at        : 1.race
                   age             =          20
               2.race
                   age             =          20
               3.race
                   age             =          20

2._at        : 1.race
                   age             =          30
               2.race
                   age             =          30
               3.race
                   age             =          30

3._at        : 1.race
                   age             =          40
               2.race
                   age             =          40
               3.race
                   age             =          40

4._at        : 1.race
                   age             =          50
               2.race
                   age             =          50
               3.race
                   age             =          50

5._at        : 1.race
                   age             =          60
               2.race
                   age             =          60
               3.race
                   age             =          60

6._at        : 1.race
                   age             =          70
               2.race
                   age             =          70
               3.race
                   age             =          70

7._at        : 1.race
                   age             =          80
               2.race
                   age             =          80
               3.race
                   age             =          80
```

	Delta-method					
	Margin	Std. Err.	t	P>|t|	[95% Conf.	Interval]
_at#race#sex						
1#White#Male	.8392234	.0533896	15.72	0.000	.7324643	.9459825
1#White#Female	2.291688	.0538363	42.57	0.000	2.184036	2.399341
1#Black#Male	1.10569	.1165164	9.49	0.000	.8727009	1.338679
1#Black#Female	2.831841	.1875228	15.10	0.000	2.456866	3.206816
1#Other#Male	1.032073	.3870692	2.67	0.010	.2580806	1.806066
1#Other#Female	2.053974	.1628213	12.61	0.000	1.728393	2.379555
2#White#Male	.9609743	.0477473	20.13	0.000	.8654978	1.056451
2#White#Female	2.309686	.0474384	48.69	0.000	2.214827	2.404545
2#Black#Male	1.298138	.1049704	12.37	0.000	1.088237	1.50804
2#Black#Female	2.848624	.1518188	18.76	0.000	2.545044	3.152205
2#Other#Male	1.063567	.2762399	3.85	0.000	.5111912	1.615943
2#Other#Female	2.173155	.122288	17.77	0.000	1.928625	2.417685
3#White#Male	1.082725	.0442231	24.48	0.000	.9942958	1.171155
3#White#Female	2.327683	.0435479	53.45	0.000	2.240604	2.414763
3#Black#Male	1.490587	.1000821	14.89	0.000	1.29046	1.690714
3#Black#Female	2.865408	.1252804	22.87	0.000	2.614894	3.115921
3#Other#Male	1.095061	.1673311	6.54	0.000	.7604614	1.42966
3#Other#Female	2.292337	.113883	20.13	0.000	2.064614	2.52006
4#White#Male	1.204476	.0433369	27.79	0.000	1.117819	1.291134
4#White#Female	2.345681	.0428533	54.74	0.000	2.25999	2.431371
4#Black#Male	1.683036	.1028059	16.37	0.000	1.477462	1.888609
4#Black#Female	2.882191	.1144703	25.18	0.000	2.653294	3.111089
4#Other#Male	1.126554	.0699734	16.10	0.000	.986634	1.266475
4#Other#Female	2.411519	.1433727	16.82	0.000	2.124827	2.69821
5#White#Male	1.326227	.045244	29.31	0.000	1.235756	1.416698
5#White#Female	2.363678	.0455011	51.95	0.000	2.272693	2.454663
5#Black#Male	1.875484	.1125906	16.66	0.000	1.650345	2.100623
5#Black#Female	2.898975	.1235873	23.46	0.000	2.651847	3.146103
5#Other#Male	1.158048	.0836551	13.84	0.000	.9907694	1.325327
5#Other#Female	2.5307	.1941998	13.03	0.000	2.142374	2.919027
6#White#Male	1.447978	.0496234	29.18	0.000	1.34875	1.547206
6#White#Female	2.381676	.0509732	46.72	0.000	2.279749	2.483603
6#Black#Male	2.067933	.1278249	16.18	0.000	1.812331	2.323534
6#Black#Female	2.915758	.1490176	19.57	0.000	2.617779	3.213737
6#Other#Male	1.189542	.1852172	6.42	0.000	.8191771	1.559906
6#Other#Female	2.649882	.2538559	10.44	0.000	2.142266	3.157498
7#White#Male	1.569729	.0558969	28.08	0.000	1.457956	1.681502
7#White#Female	2.399673	.058482	41.03	0.000	2.282731	2.516615
7#Black#Male	2.260381	.146822	15.40	0.000	1.966792	2.55397
7#Black#Female	2.932542	.1841215	15.93	0.000	2.564368	3.300715
7#Other#Male	1.221035	.2946489	4.14	0.000	.6318486	1.810222
7#Other#Female	2.769063	.3174011	8.72	0.000	2.134381	3.403746

```
. marginsplot, yline(0)

  Variables that uniquely identify margins: age sex race

.
.
.       ****************************************************************************
.       * 4.2.1 A two-way interaction Modify the model 1.2.1.
.       ****************************************************************************
.       svy: reg illness i.sex##c.age
(running regress on estimation sample)

Survey: Linear regression

Number of strata   =        31                  Number of obs     =       10,351
Number of PSUs     =        62                  Population size   =  117,157,513
                                                Design df         =           31
                                                F(   3,     29)   =       600.17
                                                Prob > F          =       0.0000
                                                R-squared         =       0.2472
```

```
             |             Linearized
     illness |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+------------------------------------------------------------------
         sex |
      Female |   1.697105   .0714118    23.77   0.000      1.55146     1.84275
         age |     .01235   .0007846    15.74   0.000     .0107498    .0139501
             |
   sex#c.age |
      Female |  -.0108862   .0014354    -7.58   0.000    -.0138137   -.0079588
             |
       _cons |   .6259047   .0553246    11.31   0.000     .5130694    .7387399
------------------------------------------------------------------------------

.         reg,  coeflegend

Survey: Linear regression

Number of strata   =        31                Number of obs     =       10,351
Number of PSUs     =        62                Population size   =  117,157,513
                                              Design df         =           31
                                              F(   3,     29)   =       600.17
                                              Prob > F          =       0.0000
                                              R-squared         =       0.2472

------------------------------------------------------------------------------
     illness |      Coef.  Legend
-------------+----------------------------------------------------------------
         sex |
      Female |   1.697105  _b[2.sex]
         age |     .01235  _b[age]
             |
   sex#c.age |
      Female |  -.0108862  _b[2.sex#c.age]
             |
       _cons |   .6259047  _b[_cons]
------------------------------------------------------------------------------

.         test _b[2.sex#c.age] =0

Adjusted Wald test

 ( 1)  2.sex#c.age = 0

       F(  1,    31) =   57.52
            Prob > F =    0.0000

.
.         margins i.sex,  at(age=(20(10)80)) vce(unconditional)

Adjusted predictions                          Number of obs     =     10,351

Expression   : Linear prediction, predict()

1._at        : age             =          20

2._at        : age             =          30

3._at        : age             =          40

4._at        : age             =          50

5._at        : age             =          60

6._at        : age             =          70

7._at        : age             =          80

------------------------------------------------------------------------------
             |             Linearized
             |     Margin   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
```

```
     _at#sex |
     1#Male  |   .8729037   .0471375    18.52   0.000     .7767662    .9690413
   1#Female  |   2.352285    .066818    35.20   0.000     2.216008    2.488561
     2#Male  |   .9964032   .0445776    22.35   0.000     .9054865     1.08732
   2#Female  |   2.366922   .0567169    41.73   0.000     2.251247    2.482597
     3#Male  |   1.119903    .043307    25.86   0.000     1.031578    1.208228
   3#Female  |   2.381559   .0493509    48.26   0.000     2.280907    2.482211
     4#Male  |   1.243402   .0434388    28.62   0.000     1.154808    1.331996
   4#Female  |   2.396197   .0460517    52.03   0.000     2.302273     2.49012
     5#Male  |   1.366902   .0449608    30.40   0.000     1.275204      1.4586
   5#Female  |   2.410834   .0476713    50.57   0.000     2.313608     2.50806
     6#Male  |   1.490401   .0477402    31.22   0.000     1.393035    1.587768
   6#Female  |   2.425471   .0537669    45.11   0.000     2.315813     2.53513
     7#Male  |   1.613901    .051574    31.29   0.000     1.508715    1.719087
   7#Female  |   2.440109   .0630535    38.70   0.000      2.31151    2.568707
-------------------------------------------------------------------------------

.         marginsplot, yline(0)

  Variables that uniquely identify margins: age sex

.
.         ************************************************************************
.         * 4.2.1 A three-way interactdion: Modify the model 1.2.1
.         ************************************************************************
.
.         svy: reg illness i.sex##c.age##i.race
(running regress on estimation sample)

Survey: Linear regression

Number of strata   =        31                  Number of obs     =       10,351
Number of PSUs     =        62                  Population size   =  117,157,513
                                                Design df         =           31
                                                F(  11,     21)   =       230.20
                                                Prob > F          =       0.0000
                                                R-squared         =       0.2603

------------------------------------------------------------------------------
             |             Linearized
     illness |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         sex |
      Female |   1.659972   .0621101    26.73   0.000     1.533297    1.786646
         age |   .0121751   .0009509    12.80   0.000     .0102358    .0141144
             |
     sex#c.age |
      Female |  -.0103753    .001262    -8.22   0.000    -.0129493   -.0078014
             |
        race |
       Black |   .1250714   .1690141     0.74   0.465    -.2196352    .4697779
       Other |   .3733645   .5887048     0.63   0.531    -.8273068    1.574036
             |
      sex#race |
 Female#Black |   .4175091   .2748922     1.52   0.139    -.1431373    .9781555
 Female#Other |  -.8134471   .8108853    -1.00   0.324    -2.467259    .8403644
             |
    race#c.age |
       Black |   .0070698   .0036788     1.92   0.064    -.0004332    .0145728
       Other |  -.0090257   .0109807    -0.82   0.417    -.0314211    .0133696
             |
sex#race#c.age |
 Female#Black |  -.0071912    .005996    -1.20   0.239    -.0194202    .0050378
 Female#Other |   .0191441    .016423     1.17   0.253    -.0143507     .052639
             |
        _cons |   .5957215    .056524    10.54   0.000       .48044     .711003
------------------------------------------------------------------------------

.         reg,  coeflegend

Survey: Linear regression
```

```
Number of strata   =        31                Number of obs     =       10,351
Number of PSUs     =        62                Population size   =  117,157,513
                                              Design df         =           31
                                              F(  11,     21)   =       230.20
                                              Prob > F          =       0.0000
                                              R-squared         =       0.2603

------------------------------------------------------------------------------
      illness |      Coef.  Legend
--------------+---------------------------------------------------------------
          sex |
       Female |   1.659972  _b[2.sex]
          age |   .0121751  _b[age]
              |
      sex#c.age |
       Female |  -.0103753  _b[2.sex#c.age]
              |
         race |
        Black |   .1250714  _b[2.race]
        Other |   .3733645  _b[3.race]
              |
      sex#race |
 Female#Black |   .4175091  _b[2.sex#2.race]
 Female#Other |  -.8134471  _b[2.sex#3.race]
              |
    race#c.age |
        Black |   .0070698  _b[2.race#c.age]
        Other |  -.0090257  _b[3.race#c.age]
              |
sex#race#c.age |
 Female#Black |  -.0071912  _b[2.sex#2.race#c.age]
 Female#Other |   .0191441  _b[2.sex#3.race#c.age]
              |
        _cons |   .5957215  _b[_cons]
------------------------------------------------------------------------------

.
.         margins i.sex , over(i.race)  at(age=(20(10)80)) vce(unconditional)

Predictive margins                              Number of obs     =     10,351

Expression   : Linear prediction, predict()
over         : race

1._at        : 1.race
                   age             =          20
               2.race
                   age             =          20
               3.race
                   age             =          20

2._at        : 1.race
                   age             =          30
               2.race
                   age             =          30
               3.race
                   age             =          30

3._at        : 1.race
                   age             =          40
               2.race
                   age             =          40
               3.race
                   age             =          40

4._at        : 1.race
                   age             =          50
               2.race
                   age             =          50
               3.race
```

```
                    age             =          50

5._at         : 1.race
                    age             =          60
                2.race
                    age             =          60
                3.race
                    age             =          60

6._at         : 1.race
                    age             =          70
                2.race
                    age             =          70
                3.race
                    age             =          70

7._at         : 1.race
                    age             =          80
                2.race
                    age             =          80
                3.race
                    age             =          80
```

	Margin	Linearized Std. Err.	t	P>|t|	[95% Conf.	Interval]
_at#race#sex						
1#White#Male	.8392234	.045837	18.31	0.000	.7457383	.9327085
1#White#Female	2.291688	.0575763	39.80	0.000	2.174261	2.409116
1#Black#Male	1.10569	.1130932	9.78	0.000	.8750349	1.336345
1#Black#Female	2.831841	.2060305	13.74	0.000	2.411639	3.252043
1#Other#Male	1.032073	.3766672	2.74	0.010	.2638555	1.800291
1#Other#Female	2.053974	.1691387	12.14	0.000	1.709013	2.398935
2#White#Male	.9609743	.0427235	22.49	0.000	.8738391	1.04811
2#White#Female	2.309686	.0486548	47.47	0.000	2.210454	2.408918
2#Black#Male	1.298138	.0955762	13.58	0.000	1.103209	1.493067
2#Black#Female	2.848624	.161878	17.60	0.000	2.518472	3.178777
2#Other#Male	1.063567	.2672403	3.98	0.000	.5185268	1.608607
2#Other#Female	2.173155	.1127393	19.28	0.000	1.943222	2.403089
3#White#Male	1.082725	.0415981	26.03	0.000	.9978854	1.167565
3#White#Female	2.327683	.0424001	54.90	0.000	2.241208	2.414159
3#Black#Male	1.490587	.0867754	17.18	0.000	1.313607	1.667567
3#Black#Female	2.865408	.1263719	22.67	0.000	2.607671	3.123145
3#Other#Male	1.095061	.1595531	6.86	0.000	.76965	1.420471
3#Other#Female	2.292337	.0818491	28.01	0.000	2.125405	2.459269
4#White#Male	1.204476	.0426184	28.26	0.000	1.117555	1.291397
4#White#Female	2.345681	.0400808	58.52	0.000	2.263935	2.427426
4#Black#Male	1.683036	.0893059	18.85	0.000	1.500895	1.865176
4#Black#Female	2.882191	.1083729	26.60	0.000	2.661163	3.103219
4#Other#Male	1.126554	.0632262	17.82	0.000	.9976037	1.255505
4#Other#Female	2.411519	.1028363	23.45	0.000	2.201783	2.621254
5#White#Male	1.326227	.0456409	29.06	0.000	1.233142	1.419312
5#White#Female	2.363678	.0423484	55.82	0.000	2.277308	2.450048
5#Black#Male	1.875484	.1023305	18.33	0.000	1.66678	2.084189
5#Black#Female	2.898975	.1163143	24.92	0.000	2.66175	3.136199
5#Other#Male	1.158048	.0838957	13.80	0.000	.9869416	1.329155
5#Other#Female	2.5307	.1560071	16.22	0.000	2.212522	2.848879
6#White#Male	1.447978	.0503058	28.78	0.000	1.345379	1.550578
6#White#Female	2.381676	.0485646	49.04	0.000	2.282628	2.480724
6#Black#Male	2.067933	.1225477	16.87	0.000	1.817995	2.31787
6#Black#Female	2.915758	.1460233	19.97	0.000	2.617942	3.213575
6#Other#Male	1.189542	.185957	6.40	0.000	.8102799	1.568804
6#Other#Female	2.649882	.2190617	12.10	0.000	2.203102	3.096661
7#White#Male	1.569729	.0562058	27.93	0.000	1.455097	1.684362
7#White#Female	2.399673	.0574619	41.76	0.000	2.282479	2.516867
7#Black#Male	2.260381	.1470196	15.37	0.000	1.960533	2.56023
7#Black#Female	2.932542	.1874194	15.65	0.000	2.550297	3.314786
7#Other#Male	1.221035	.2943163	4.15	0.000	.6207733	1.821298
7#Other#Female	2.769063	.2855253	9.70	0.000	2.186731	3.351396

```
.
.         marginsplot, yline(0)

  Variables that uniquely identify margins: age sex race

.
.
. ******************************************************************************
. * Questions 5: testing the total effects of X
. ******************************************************************************
.
.         ************************************************
.         * 5.1 Regression using a simple survey data and a three-way interaction term
.         ************************************************
.
.         reg illness i.sex##c.age##hlthstat [pw=finalwgt], cluster(company_id)
(sum of wgt is   1.1713e+08)

Linear regression                               Number of obs     =     10,349
                                                F(23, 61)         =     227.88
                                                Prob > F          =     0.0000
                                                R-squared         =     0.2774
                                                Root MSE          =     1.0855

                                  (Std. Err. adjusted for 62 clusters in company_id)
------------------------------------------------------------------------------------
                   |               Robust
           illness |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------------+----------------------------------------------------------------
               sex |
            Female |   1.481147   .1563845     9.47   0.000     1.168437    1.793858
               age |   .0046442   .0020626     2.25   0.028     .0005198    .0087686
                   |
         sex#c.age |
            Female |  -.0062903   .0034037    -1.85   0.069    -.0130964    .0005158
                   |
          hlthstat |
                 2 |  -.0039889   .1382512    -0.03   0.977    -.2804393    .2724615
                 3 |   .1427757   .1664162     0.86   0.394    -.1899941    .4755455
                 4 |   .0946943   .2019677     0.47   0.641     -.309165    .4985536
                 5 |   .4744487   .3804952     1.25   0.217    -.2863985    1.235296
                 8 |  -2.912394   1.314039    -2.22   0.030    -5.539978   -.2848096
                   |
      sex#hlthstat |
         Female#2  |   .1868007   .2166193     0.86   0.392    -.2463563    .6199577
         Female#3  |   .1873244      .2165     0.87   0.390     -.245594    .6202429
         Female#4  |   .4056525   .2879938     1.41   0.164    -.1702266    .9815317
         Female#5  |   .0864362   .5316279     0.16   0.871    -.9766196    1.149492
         Female#8  |  -.5682743   1.507553    -0.38   0.708    -3.582812    2.446264
                   |
    hlthstat#c.age |
                 2 |   .0000467   .0030854     0.02   0.988    -.0061229    .0062163
                 3 |   .0025419   .0035405     0.72   0.476    -.0045376    .0096215
                 4 |   .0076151   .0039138     1.95   0.056    -.0002111    .0154413
                 5 |   .0103682   .0062929     1.65   0.105    -.0022153    .0229517
                 8 |   .0582859   .0189807     3.07   0.003     .0203315    .0962402
                   |
sex#hlthstat#c.age |
         Female#2  |  -.0025453   .0049583    -0.51   0.610      -.01246    .0073694
         Female#3  |  -.0041849   .0045901    -0.91   0.366    -.0133634    .0049936
         Female#4  |   -.008245   .0055853    -1.48   0.145    -.0194135    .0029234
         Female#5  |  -.0054305   .0091741    -0.59   0.556    -.0237752    .0129141
         Female#8  |   .0011581   .0237083     0.05   0.961    -.0462496    .0485657
                   |
             _cons |   .7755973    .107274     7.23   0.000     .5610897    .9901048
------------------------------------------------------------------------------------

.         margins i.sex, atmeans

Adjusted predictions                            Number of obs     =     10,349
```

```
Model VCE    : Robust

Expression   : Linear prediction, predict()
at           : 1.sex           =    .4794583 (mean)
               2.sex           =    .5205417 (mean)
               age             =    42.25401 (mean)
               1.hlthstat      =    .2747975 (mean)
               2.hlthstat      =    .2747034 (mean)
               3.hlthstat      =    .2793058 (mean)
               4.hlthstat      =    .1227706 (mean)
               5.hlthstat      =      .04728 (mean)
               8.hlthstat      =    .0011428 (mean)

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         sex |
        Male |   1.134917   .0422296    26.87   0.000     1.050473     1.21936
      Female |   2.374649   .0484046    49.06   0.000     2.277858     2.47144
------------------------------------------------------------------------------

.         margins, dydx(i.sex) atmeans

Conditional marginal effects                    Number of obs     =     10,349
Model VCE    : Robust

Expression   : Linear prediction, predict()
dy/dx w.r.t. : 2.sex
at           : 1.sex           =    .4794583 (mean)
               2.sex           =    .5205417 (mean)
               age             =    42.25401 (mean)
               1.hlthstat      =    .2747975 (mean)
               2.hlthstat      =    .2747034 (mean)
               3.hlthstat      =    .2793058 (mean)
               4.hlthstat      =    .1227706 (mean)
               5.hlthstat      =      .04728 (mean)
               8.hlthstat      =    .0011428 (mean)

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         sex |
      Female |   1.239733   .0299038    41.46   0.000     1.179936    1.299529
------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.

.
.
.         ********************************************************************************
.         * 5.1 Regression using complex survey data and a three-way interaction term
.         ********************************************************************************
.
.         svy: reg illness i.sex##c.age##hlthstat
(running regress on estimation sample)

Survey: Linear regression

Number of strata   =        31                  Number of obs     =       10,349
Number of PSUs     =        62                  Population size   =  117,131,111
                                                Design df         =           31
                                                F(  23,      9)   =       153.93
                                                Prob > F          =       0.0000
                                                R-squared         =       0.2774

------------------------------------------------------------------------------
             |             Linearized
     illness |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         sex |
```

```
          Female |   1.481147   .1678344     8.83   0.000     1.138847    1.823448
             age |   .0046442   .0021159     2.19   0.036     .0003288    .0089595
                 |
         sex#c.age |
          Female |  -.0062903   .0036226    -1.74   0.092    -.0136786     .001098
                 |
          hlthstat |
               2 |  -.0039889   .1583813    -0.03   0.980    -.3270096    .3190318
               3 |   .1427757   .1875964     0.76   0.452    -.2398297    .5253812
               4 |   .0946943   .2138854     0.44   0.661    -.3415279    .5309165
               5 |   .4744487   .3596022     1.32   0.197    -.2589649    1.207862
               8 |  -2.912394   1.297575    -2.24   0.032    -5.558815    -.265973
                 |
      sex#hlthstat |
         Female#2 |   .1868007   .2438318     0.77   0.449    -.3104976     .684099
         Female#3 |   .1873244   .2346484     0.80   0.431    -.2912441    .6658929
         Female#4 |   .4056525   .3090278     1.31   0.199    -.2246138    1.035919
         Female#5 |   .0864362   .5165054     0.17   0.868    -.9669836    1.139856
         Female#8 |  -.5682743   1.403271    -0.40   0.688    -3.430264    2.293715
                 |
    hlthstat#c.age |
               2 |   .0000467    .003441     0.01   0.989    -.0069713    .0070647
               3 |   .0025419   .0038458     0.66   0.514    -.0053016    .0103855
               4 |   .0076151   .0042862     1.78   0.085    -.0011266    .0163568
               5 |   .0103682   .0060514     1.71   0.097    -.0019737    .0227101
               8 |   .0582859   .0187797     3.10   0.004     .0199844    .0965873
                 |
sex#hlthstat#c.age |
         Female#2 |  -.0025453   .0056584    -0.45   0.656    -.0140857     .008995
         Female#3 |  -.0041849   .0048587    -0.86   0.396    -.0140942    .0057244
         Female#4 |   -.008245   .0060516    -1.36   0.183    -.0205873    .0040972
         Female#5 |  -.0054305   .0088294    -0.62   0.543    -.0234381    .0125771
         Female#8 |   .0011581   .0223065     0.05   0.959    -.0443364    .0466525
                 |
             _cons |   .7755973   .1143987     6.78   0.000     .5422795    1.008915
------------------------------------------------------------------------------

. margins i.sex, atmeans vce(unconditional)

Adjusted predictions                            Number of obs     =     10,349

Expression   : Linear prediction, predict()
at           : 1.sex           =    .4794583 (mean)
               2.sex           =    .5205417 (mean)
               age             =    42.25401 (mean)
               1.hlthstat      =    .2747975 (mean)
               2.hlthstat      =    .2747034 (mean)
               3.hlthstat      =    .2793058 (mean)
               4.hlthstat      =    .1227706 (mean)
               5.hlthstat      =      .04728 (mean)
               8.hlthstat      =    .0011428 (mean)

------------------------------------------------------------------------------
             |            Linearized
             |     Margin   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         sex |
        Male |   1.134917   .0397368    28.56   0.000     1.053873     1.21596
      Female |   2.374649   .0475868    49.90   0.000     2.277595    2.471703
------------------------------------------------------------------------------

. margins, dydx(i.sex) atmeans vce(unconditional)

Conditional marginal effects                    Number of obs     =     10,349

Expression   : Linear prediction, predict()
dy/dx w.r.t. : 2.sex
at           : 1.sex           =    .4794583 (mean)
               2.sex           =    .5205417 (mean)
               age             =    42.25401 (mean)
               1.hlthstat      =    .2747975 (mean)
```

```
              2.hlthstat      =    .2747034 (mean)
              3.hlthstat      =    .2793058 (mean)
              4.hlthstat      =    .1227706 (mean)
              5.hlthstat      =      .04728 (mean)
              8.hlthstat      =    .0011428 (mean)

------------------------------------------------------------------------------
             |            Linearized
             |      dy/dx   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         sex |
      Female |   1.239733   .0300214    41.29   0.000     1.178504    1.300962
------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.

.
.
. log close
      name:  <unnamed>
       log:  D:\Jason\workshop\Regression 2019\final\regression1.log
  log type:  text
 closed on:  20 Feb 2019, 14:16:17
-------------------------------------------------------------------------------------------------------
```