```
.                /***********************************************/
>         /*  Program to run Analysis on Wage data set    */
>         /*  Create by Debra Whitney, 2/98               */
>         /***********************************************/
>
> /*  Compute Descriptive Statistics
>         ==============================*/
>         summarize wage lnwage exper age educ, detail;
```

Hourly Wages ($)

Percentiles		Smallest		
1%	2.884615	1.923077		
5%	4.625	2.076923		
10%	5.769231	2.115385	Obs	1003
25%	8.173077	2.403846	Sum of Wgt.	1003
50%	12.5		Mean	14.7697
		Largest	Std. Dev.	9.257249
75%	19.23077	48.07644		
90%	26.92308	48.07644	Variance	85.69666
95%	33.65385	48.07644	Skewness	1.514996
99%	48.07644	48.07644	Kurtosis	5.599439

Log Wages

Percentiles		Smallest		
1%	1.059392	.6539264		
5%	1.531476	.7308875		
10%	1.752539	.7492366	Obs	1003
25%	2.100845	.8770701	Sum of Wgt.	1003
50%	2.525729		Mean	2.514337
		Largest	Std. Dev.	.6045773
75%	2.956511	3.872792		
90%	3.292984	3.872792	Variance	.3655137
95%	3.516127	3.872792	Skewness	-.0934076
99%	3.872792	3.872792	Kurtosis	2.844012

Work Experience (years)

Percentiles		Smallest		
1%	2	0		
5%	4	0		
10%	6	0	Obs	1003
25%	11	0	Sum of Wgt.	1003
50%	19		Mean	20.14506
		Largest	Std. Dev.	11.16894
75%	28	48		
90%	36	48	Variance	124.7452
95%	40	49	Skewness	.3133469
99%	45	50	Kurtosis	2.284639

```
                            Age
-------------------------------------------------------------
      Percentiles      Smallest
 1%           20             18
 5%           24             18
10%           26             19       Obs                1003
25%           31             19       Sum of Wgt.        1003

50%           39                      Mean           39.59721
                        Largest       Std. Dev.      10.98771
75%           48             66
90%           55             66       Variance       120.7298
95%           59             67       Skewness       .2617562
99%           64             68       Kurtosis       2.231103

                     Education (years)
-------------------------------------------------------------
      Percentiles      Smallest
 1%           5.5            2.5
 5%            10            5.5
10%            12            5.5       Obs                1003
25%            12            5.5       Sum of Wgt.        1003

50%            13                      Mean           13.45214
                        Largest       Std. Dev.       2.60904
75%            16             20
90%            16             20       Variance       6.807089
95%            18             20       Skewness       .1719891
99%            20             20       Kurtosis       4.169621

. tab1 sex race edlevel industry occupat marrstat union region;

-> tabulation of sex

Sex 1=Male  |
2=Female    |      Freq.     Percent        Cum.
------------+-----------------------------------
       Male |        537       53.54       53.54
     Female |        466       46.46      100.00
------------+-----------------------------------
      Total |       1003      100.00

-> tabulation of race

Race 1=oth  |
2=blk 3=hisp|
4=wh        |      Freq.     Percent        Cum.
------------+-----------------------------------
      Other |         52        5.18        5.18
      Black |        102       10.17       15.35
   Hispanic |         61        6.08       21.44
      White |        788       78.56      100.00
------------+-----------------------------------
```

```
      Total |       1003      100.00

-> tabulation of edlevel

Education   |
level       |
achieved    |      Freq.     Percent        Cum.
------------+-----------------------------------
    0-12yrs |         86        8.57        8.57
    HSDiplm |        362       36.09       44.67
   ColAssoc |        280       27.92       72.58
   BA/AB/BS |        175       17.45       90.03
    Masters |         70        6.98       97.01
         MD |         14        1.40       98.40
        PHD |         16        1.60      100.00
------------+-----------------------------------
      Total |       1003      100.00

-> tabulation of industry

   Industry |      Freq.     Percent        Cum.
------------+-----------------------------------
     Constr |         53        5.28        5.28
        Mfg |        216       21.54       26.82
   TransCom |         95        9.47       36.29
     Retail |        137       13.66       49.95
    Finance |         75        7.48       57.43
    Medical |         97        9.67       67.10
     Educat |         80        7.98       75.07
   PublAdmn |         60        5.98       81.06
      Other |        190       18.94      100.00
------------+-----------------------------------
      Total |       1003      100.00

-> tabulation of occupat

Occupation  |      Freq.     Percent        Cum.
------------+-----------------------------------
   Managemt |        153       15.25       15.25
    Profess |        197       19.64       34.90
      Sales |        111       11.07       45.96
   Clerical |        170       16.95       62.91
    Service |         95        9.47       72.38
   Bluecoll |        277       27.62      100.00
------------+-----------------------------------
      Total |       1003      100.00

-> tabulation of marrstat

Marital     |
Status      |      Freq.     Percent        Cum.
------------+-----------------------------------
    Nevmarr |        185       18.44       18.44
   Wid/Divr |        184       18.34       36.79
```

```
    Married |        634       63.21      100.00
------------+-----------------------------------
      Total |       1003      100.00

-> tabulation of union

Union Member|
(1=Yes,0=No)|      Freq.     Percent        Cum.
------------+-----------------------------------
         No |        824       82.15       82.15
        Yes |        179       17.85      100.00
------------+-----------------------------------
      Total |       1003      100.00

-> tabulation of region

Region of   |
Residence   |      Freq.     Percent        Cum.
------------+-----------------------------------
   Northeas |        247       24.63       24.63
    Midwest |        244       24.33       48.95
      South |        317       31.61       80.56
       West |        195       19.44      100.00
------------+-----------------------------------
      Total |       1003      100.00

.         /*  Determine Transformation
>         ========================*/
>         regress wage sex;

  Source |       SS       df       MS                  Number of obs =    1003
---------+------------------------------                F(  1,  1001) =   75.22
   Model |  6001.77616     1  6001.77616                Prob > F      =  0.0000
Residual |   79866.276  1001  79.7864896                R-squared     =  0.0699
---------+------------------------------                Adj R-squared =  0.0690
   Total |  85868.0522  1002  85.6966589                Root MSE      =  8.9323

------------------------------------------------------------------------------
    wage |      Coef.   Std. Err.       t     P>|t|      [ 95% Conf. Interval]
---------+--------------------------------------------------------------------
     sex |  -4.904675   .5655032     -8.673   0.000      -6.014383   -3.794968
   _cons |   21.95312   .8749448     25.091   0.000       20.23618    23.67006
------------------------------------------------------------------------------

.         regress wage exper;

  Source |       SS       df       MS                  Number of obs =    1003
---------+------------------------------                F(  1,  1001) =   19.45
   Model |  1637.06073     1  1637.06073                Prob > F      =  0.0000
Residual |  84230.9915  1001  84.1468446                R-squared     =  0.0191
---------+------------------------------                Adj R-squared =  0.0181
   Total |  85868.0522  1002  85.6966589                Root MSE      =  9.1732

------------------------------------------------------------------------------
```

```
    wage |      Coef.   Std. Err.       t     P>|t|       [ 95% Conf. Interval]
---------+------------------------------------------------------------------
   exper |   .1144423   .0259462      4.411   0.000       .0635272    .1653574
   _cons |   12.46425    .597576     20.858   0.000       11.29161     13.6369
------------------------------------------------------------------------------

.         regress wage educ;

  Source |       SS       df       MS                  Number of obs =    1003
---------+------------------------------               F(  1,  1001) =  206.20
   Model |  14666.8401     1  14666.8401               Prob > F      =  0.0000
Residual |  71201.2121  1001   71.130082               R-squared     =  0.1708
---------+------------------------------               Adj R-squared =  0.1700
   Total |  85868.0522  1002  85.6966589               Root MSE      =  8.4339

------------------------------------------------------------------------------
    wage |      Coef.   Std. Err.       t     P>|t|       [ 95% Conf. Interval]
---------+------------------------------------------------------------------
    educ |   1.466404   .1021203     14.360   0.000        1.26601    1.666799
   _cons |  -4.956578   1.399311     -3.542   0.000      -7.702497   -2.210659
------------------------------------------------------------------------------

.         /*      Univariate Associations with Log Wages
>         =====================================*/
>         correlate lnwage age sex exper educ;
(obs=1003)

        |   lnwage      age      sex    exper     educ
--------+---------------------------------------------
  lnwage|   1.0000
     age|   0.2529   1.0000
     sex|  -0.2636   0.0384   1.0000
   exper|   0.1548   0.9724   0.0405   1.0000
    educ|   0.4025   0.0487  -0.0117  -0.1857   1.0000


.         tab sex,summarize(wage);

Sex 1=Male  |     Summary of Hourly Wages ($)
2=Female    |        Mean   Std. Dev.       Freq.
------------+------------------------------------
       Male |   17.048445   10.240742         537
     Female |    12.14377    7.132306         466
------------+------------------------------------
      Total |   14.769702    9.257249        1003

.         tab sex,summarize(lnwage);

Sex 1=Male  |       Summary of Log of wages
2=Female    |        Mean   Std. Dev.       Freq.
------------+------------------------------------
       Male |   2.6627056    .6033977         537
     Female |   2.3433633   .55966571         466
------------+------------------------------------
```

```
      Total |   2.5143372   .60457733        1003

.         regress lnwage sex;

  Source |       SS       df       MS                  Number of obs =    1003
---------+------------------------------               F(  1,  1001) =   74.73
   Model |  25.4432335     1  25.4432335               Prob > F      =  0.0000
Residual |  340.801541  1001   .34046108               R-squared     =  0.0695
---------+------------------------------               Adj R-squared =  0.0685
   Total |  366.244774  1002  .365513747               Root MSE      =  .58349

------------------------------------------------------------------------------
  lnwage |      Coef.   Std. Err.       t     P>|t|       [ 95% Conf. Interval]
---------+--------------------------------------------------------------------
     sex |  -.3193424   .0369406     -8.645   0.000      -.3918323   -.2468524
   _cons |   2.982048   .0571544     52.175   0.000       2.869892    3.094204
------------------------------------------------------------------------------

.         tab race,summarize(wage);

Race 1=oth  |     Summary of Hourly Wages ($)
2=blk 3=hisp|
4=wh        |        Mean   Std. Dev.       Freq.
------------+------------------------------------
      Other |   13.879308   8.2283586          52
      Black |   12.467623   7.5165665         102
   Hispanic |   9.8324086   6.9178947          61
      White |   15.508646   9.5208053         788
------------+------------------------------------
      Total |   14.769702    9.257249        1003

.         tab race,summarize(lnwage);

Race 1=oth  |       Summary of Log of wages
2=blk 3=hisp|
4=wh        |        Mean   Std. Dev.       Freq.
------------+------------------------------------
      Other |   2.4775528   .55915946          52
      Black |    2.340565    .6264733         102
   Hispanic |   2.1285081   .53193239          61
      White |   2.5691255   .59550002         788
------------+------------------------------------
      Total |   2.5143372   .60457733        1003

.         anova lnwage race;

                           Number of obs =    1003     R-squared     =  0.0399
                           Root MSE      = .593296     Adj R-squared =  0.0370

                  Source |  Partial SS    df       MS           F     Prob > F
              -----------+----------------------------------------------------
                   Model |  14.5965218     3  4.86550727      13.82     0.0000
                         |
                    race |  14.5965218     3  4.86550727      13.82     0.0000
```

```
                    |
            Residual |  351.648253   999  .352000253
           ----------+-----------------------------------------------------
               Total |  366.244774  1002  .365513747

.         anova,regress detail;

  Factor      Value          Value          Value          Value
  ------------------------------------------------------------------------
  race        1 1            2 2            3 3            4 4

  Source |       SS       df       MS                 Number of obs =    1003
---------+------------------------------              F(  3,   999) =   13.82
   Model |  14.5965218     3  4.86550727               Prob > F      =  0.0000
Residual |  351.648253   999  .352000253              R-squared     =  0.0399
---------+------------------------------              Adj R-squared =  0.0370
   Total |  366.244774  1002  .365513747              Root MSE      =   .5933

------------------------------------------------------------------------------
  lnwage        Coef.   Std. Err.       t     P>|t|       [ 95% Conf. Interval]
------------------------------------------------------------------------------
_cons        2.569125   .0211353    121.556   0.000       2.527651      2.6106
race
       1    -.0915726   .0849467     -1.078   0.281       -.258267    .0751217
       2    -.2285605   .0624314     -3.661   0.000      -.3510723   -.1060487
       3    -.4406173   .0788492     -5.588   0.000      -.5953464   -.2858883
       4    (dropped)
------------------------------------------------------------------------------

.         regress lnwage exper;

  Source |       SS       df       MS                 Number of obs =    1003
---------+------------------------------              F(  1,  1001) =   24.56
   Model |   8.7714897     1   8.7714897               Prob > F      =  0.0000
Residual |  357.473285  1001  .357116168              R-squared     =  0.0239
---------+------------------------------              Adj R-squared =  0.0230
   Total |  366.244774  1002  .365513747              Root MSE      =  .59759

------------------------------------------------------------------------------
  lnwage |      Coef.   Std. Err.       t     P>|t|       [ 95% Conf. Interval]
---------+--------------------------------------------------------------------
   exper |    .008377   .0016903      4.956   0.000       .0050601    .0116939
   _cons |   2.345581   .0389295     60.252   0.000       2.269188    2.421974
------------------------------------------------------------------------------

.         regress lnwage age;

  Source |       SS       df       MS                 Number of obs =    1003
---------+------------------------------              F(  1,  1001) =   68.39
   Model |  23.4215197     1  23.4215197               Prob > F      =  0.0000
Residual |  342.823255  1001  .342480774              R-squared     =  0.0640
---------+------------------------------              Adj R-squared =  0.0630
   Total |  366.244774  1002  .365513747              Root MSE      =  .58522
```

```
------------------------------------------------------------------------------
  lnwage |      Coef.   Std. Err.       t     P>|t|       [ 95% Conf. Interval]
---------+--------------------------------------------------------------------
     age |   .0139145   .0016826      8.270   0.000       .0106127    .0172163
   _cons |   1.963363   .0691407     28.397   0.000       1.827686    2.099041
------------------------------------------------------------------------------

.         regress lnwage educ;

  Source |       SS       df       MS                  Number of obs =    1003
---------+------------------------------               F(  1,  1001) =  193.52
   Model |  59.3347547     1  59.3347547               Prob > F      =  0.0000
Residual |   306.91002  1001  .306603416               R-squared     =  0.1620
---------+------------------------------               Adj R-squared =  0.1612
   Total |  366.244774  1002  .365513747               Root MSE      =  .55372

------------------------------------------------------------------------------
  lnwage |      Coef.   Std. Err.       t     P>|t|       [ 95% Conf. Interval]
---------+--------------------------------------------------------------------
    educ |   .0932696   .0067046     13.911   0.000       .0801129    .1064263
   _cons |   1.259661   .0918705     13.711   0.000       1.079381    1.439942
------------------------------------------------------------------------------

.         tab edlevel,summarize(wage);

Education   |     Summary of Hourly Wages ($)
level       |
achieved    |        Mean   Std. Dev.       Freq.
------------+------------------------------------
    0-12yrs |    9.310918   6.0386959          86
    HSDiplm |   12.522641   6.9705726         362
   ColAssoc |   13.961516   8.0862407         280
   BA/AB/BS |   17.896819   10.012366         175
    Masters |   21.820186   11.431568          70
         MD |   33.114526   13.480527          14
        PHD |   27.993269   7.6974816          16
------------+------------------------------------
      Total |   14.769702    9.257249        1003

.         anova lnwage edlevel;

                           Number of obs =    1003     R-squared     =  0.1752
                           Root MSE      = .550718     Adj R-squared =  0.1702

                  Source |  Partial SS    df       MS           F     Prob > F
              -----------+----------------------------------------------------
                   Model |  64.1680423     6  10.6946737      35.26     0.0000
                         |
                 edlevel |  64.1680423     6  10.6946737      35.26     0.0000
                         |
                Residual |  302.076732   996  .303289892
              -----------+----------------------------------------------------
                   Total |  366.244774  1002  .365513747
```

```
.         anova,regress detail;

  Factor      Value          Value          Value          Value
  ------------------------------------------------------------------------
  edlevel     1 1            2 2            3 3            4 4
              5 5            6 6            7 7

  Source |       SS       df       MS                 Number of obs =    1003
---------+------------------------------              F(  6,   996) =   35.26
   Model |  64.1680423     6  10.6946737              Prob > F      =  0.0000
Residual |  302.076732   996  .303289892             R-squared     =  0.1752
---------+------------------------------              Adj R-squared =  0.1702
   Total |  366.244774  1002  .365513747             Root MSE      =  .55072

------------------------------------------------------------------------------
  lnwage        Coef.   Std. Err.       t     P>|t|       [ 95% Conf. Interval]
------------------------------------------------------------------------------
_cons        3.298107   .1376794     23.955   0.000       3.027932    3.568282
edlevel
       1    -1.203179   .1499408     -8.024   0.000      -1.497415   -.9089431
       2    -.9119391   .1406891     -6.482   0.000       -1.18802   -.6358579
       3    -.8218106   .1415585     -5.805   0.000      -1.099598   -.5440236
       4    -.5731721   .1438357     -3.985   0.000      -.8554279   -.2909164
       5    -.3379098   .1526051     -2.214   0.027      -.6373743   -.0384453
       6     .1099699    .201542      0.546   0.585      -.2855257    .5054656
       7    (dropped)
------------------------------------------------------------------------------

.         tab industry,summarize(wage);

           |     Summary of Hourly Wages ($)
  Industry |        Mean   Std. Dev.       Freq.
-----------+------------------------------------
     Constr |   13.856577   8.2433776          53
        Mfg |   16.661392   10.541315         216
   TransCom |     17.3108   8.6511829          95
     Retail |   10.390213   6.7267431         137
    Finance |   15.479026   10.520977          75
    Medical |   14.941064   9.6665909          97
     Educat |   14.641352   6.7929294          80
   PublAdmn |   16.647628   8.7676782          60
      Other |   13.854691   9.0388441         190
-----------+------------------------------------
      Total |   14.769702    9.257249        1003

.         anova lnwage industry;

                           Number of obs =    1003     R-squared     =  0.0686
                           Root MSE      = .585829     Adj R-squared =  0.0611

                  Source |  Partial SS    df       MS           F     Prob > F
              -----------+----------------------------------------------------
                   Model |  25.1087443     8  3.13859303       9.15     0.0000
                         |
```

```
    industry |  25.1087443     8  3.13859303       9.15     0.0000
             |
    Residual |   341.13603   994  .343195201
-------------+------------------------------------------------------
       Total |  366.244774  1002  .365513747

. anova,regress detail;

  Factor      Value          Value          Value          Value
  ------------------------------------------------------------------------
  industry    1 1            2 2            3 3            4 4
              5 5            6 6            7 7            8 8
              9 9

  Source |       SS       df       MS                 Number of obs =    1003
---------+------------------------------              F(  8,   994) =    9.15
   Model |  25.1087443     8  3.13859303              Prob > F      =  0.0000
Residual |   341.13603   994  .343195201             R-squared     =  0.0686
---------+------------------------------              Adj R-squared =  0.0611
   Total |  366.244774  1002  .365513747             Root MSE      =  .58583

------------------------------------------------------------------------------
  lnwage        Coef.   Std. Err.       t     P>|t|       [ 95% Conf. Interval]
------------------------------------------------------------------------------
_cons        2.449945   .0425005     57.645   0.000       2.366544    2.533346
industry
       1       .01193   .0910037      0.131   0.896      -.1666514    .1905114
       2     .1820484    .058268      3.124   0.002        .067706    .2963907
       3     .2613796    .073613      3.551   0.000       .1169249    .4058343
       4     -.272788    .065661     -4.154   0.000       -.401638   -.1439379
       5     .0913972   .0798888      1.144   0.253      -.0653729    .2481673
       6     .0795135   .0731053      1.088   0.277       -.063945    .2229719
       7     .1083118   .0780784      1.387   0.166      -.0449056    .2615291
       8     .2323117   .0867537      2.678   0.008       .0620703    .4025532
       9    (dropped)
------------------------------------------------------------------------------

. tab occupat,summarize(wage);

            |     Summary of Hourly Wages ($)
Occupation  |        Mean   Std. Dev.       Freq.
------------+------------------------------------
   Managemt |   20.567792    11.69531         153
    Profess |   18.814216   9.8798015         197
      Sales |   14.144551   9.6114193         111
   Clerical |   10.711768   4.6514721         170
    Service |   10.409413   6.5159441          95
   Bluecoll |   12.927072   7.1680352         277
------------+------------------------------------
      Total |   14.769702    9.257249        1003

. anova lnwage occupat;

                           Number of obs =    1003     R-squared     =  0.1608
```

```
                            Root MSE      = .555236     Adj R-squared =  0.1566

                    Source |  Partial SS    df       MS           F     Prob > F
                -----------+----------------------------------------------------
                     Model |  58.8825167     5  11.7765033      38.20     0.0000
                           |
                   occupat |  58.8825167     5  11.7765033      38.20     0.0000
                           |
                  Residual |  307.362258   997  .308287119
                -----------+----------------------------------------------------
                     Total |  366.244774  1002  .365513747

.         anova,regress detail;

  Factor      Value          Value          Value          Value
  ------------------------------------------------------------------------
  occupat     1 1            2 2            3 3            4 4
              5 5            6 6

  Source |       SS       df       MS                  Number of obs =    1003
---------+------------------------------               F(  5,   997) =   38.20
   Model |  58.8825167     5  11.7765033               Prob > F      =  0.0000
Residual |  307.362258   997  .308287119               R-squared     =  0.1608
---------+------------------------------               Adj R-squared =  0.1566
   Total |  366.244774  1002  .365513747               Root MSE      =  .55524

------------------------------------------------------------------------------
  lnwage        Coef.   Std. Err.       t     P>|t|       [ 95% Conf. Interval]
------------------------------------------------------------------------------
_cons        2.414303   .0333609     72.369   0.000       2.348838    2.479769
occupat
       1     .4465032   .0559276      7.984   0.000       .3367538    .5562525
       2     .3847819    .051748      7.436   0.000       .2832343    .4863294
       3     .0336476   .0623724      0.539   0.590      -.0887486    .1560438
       4    -.1436518   .0540962     -2.655   0.008      -.2498073   -.0374962
       5     -.243126   .0660157     -3.683   0.000      -.3726718   -.1135803
       6    (dropped)
------------------------------------------------------------------------------

.         tab marrstat,summarize(wage);

Marital     |     Summary of Hourly Wages ($)
Status      |        Mean   Std. Dev.       Freq.
------------+------------------------------------
    Nevmarr |   11.651915   7.5255505         185
   Wid/Divr |   13.747714   8.1282427         184
    Married |   15.976069   9.7717832         634
------------+------------------------------------
      Total |   14.769702    9.257249        1003

.         anova lnwage marrstat;

                            Number of obs =    1003     R-squared     =  0.0425
                            Root MSE      = .592195     Adj R-squared =  0.0405
```

```
                Source |  Partial SS    df       MS           F     Prob > F
           ------------+----------------------------------------------------
                 Model |  15.5499524     2  7.77497622      22.17     0.0000
                       |
              marrstat |  15.5499524     2  7.77497622      22.17     0.0000
                       |
              Residual |  350.694822  1000  .350694822
           ------------+----------------------------------------------------
                 Total |  366.244774  1002  .365513747

.         anova,regress detail;

  Factor      Value          Value          Value          Value
  ------------------------------------------------------------------------
  marrstat    1 1            2 2            3 3

  Source |       SS       df       MS                  Number of obs =    1003
---------+------------------------------               F(  2,  1000) =   22.17
   Model |  15.5499524     2  7.77497622               Prob > F      =  0.0000
Residual |  350.694822  1000  .350694822               R-squared     =  0.0425
---------+------------------------------               Adj R-squared =  0.0405
   Total |  366.244774  1002  .365513747               Root MSE      =  .59219

------------------------------------------------------------------------------
  lnwage        Coef.   Std. Err.       t     P>|t|       [ 95% Conf. Interval]
------------------------------------------------------------------------------
_cons        2.600061   .0235191    110.551   0.000       2.553909    2.646214
marrstat
       1     -.321209   .0494853     -6.491   0.000      -.4183159    -.224102
       2    -.1443358   .0495893     -2.911   0.004      -.2416467   -.0470248
       3    (dropped)
------------------------------------------------------------------------------

.         tab union,summarize(wage);

Union Member|     Summary of Hourly Wages ($)
(1=Yes,0=No)|        Mean   Std. Dev.       Freq.
------------+------------------------------------
         No |   14.581807   9.6517007         824
        Yes |    15.63465   7.1257612         179
------------+------------------------------------
      Total |   14.769702    9.257249        1003

.         regress lnwage union;

  Source |       SS       df       MS                  Number of obs =    1003
---------+------------------------------               F(  1,  1001) =    8.75
   Model |  3.17543538     1  3.17543538               Prob > F      =  0.0032
Residual |  363.069339  1001  .362706632               R-squared     =  0.0087
---------+------------------------------               Adj R-squared =  0.0077
   Total |  366.244774  1002  .365513747               Root MSE      =  .60225

------------------------------------------------------------------------------
```

```
  lnwage |      Coef.   Std. Err.       t     P>|t|       [ 95% Conf. Interval]
---------+------------------------------------------------------------------
   union |   .1469474   .0496636      2.959   0.003       .0494907    .2444041
   _cons |   2.488112   .0209804    118.592   0.000       2.446942    2.529283
------------------------------------------------------------------------------

.         tab region,summarize(wage);

Region of   |     Summary of Hourly Wages ($)
Residence   |        Mean   Std. Dev.       Freq.
------------+------------------------------------
   Northeas |    16.47508   10.044439         247
    Midwest |   14.092022   8.1207815         244
      South |   14.066194   9.1102953         317
       West |   14.601179   9.5797632         195
------------+------------------------------------
      Total |   14.769702    9.257249        1003

.         anova lnwage region;

                           Number of obs =    1003     R-squared     =  0.0124
                           Root MSE      = .601725     Adj R-squared =  0.0094

                  Source |  Partial SS    df       MS           F     Prob > F
              -----------+----------------------------------------------------
                   Model |  4.53408514     3  1.51136171       4.17     0.0060
                         |
                  region |  4.53408514     3  1.51136171       4.17     0.0060
                         |
                Residual |  361.710689   999  .362072762
              -----------+----------------------------------------------------
                   Total |  366.244774  1002  .365513747

.         anova,regress detail;

  Factor      Value          Value          Value          Value
  ------------------------------------------------------------------------
  region      1 1            2 2            3 3            4 4

  Source |       SS       df       MS                  Number of obs =    1003
---------+------------------------------              F(  3,   999) =    4.17
   Model |  4.53408514     3  1.51136171               Prob > F      =  0.0060
Residual |  361.710689   999  .362072762               R-squared     =  0.0124
---------+------------------------------              Adj R-squared =  0.0094
   Total |  366.244774  1002  .365513747               Root MSE      =  .60172

------------------------------------------------------------------------------
  lnwage        Coef.   Std. Err.       t     P>|t|       [ 95% Conf. Interval]
------------------------------------------------------------------------------
_cons        2.485293   .0430904     57.676   0.000       2.400734    2.569851
region
       1     .1456967   .0576426      2.528   0.012       .0325824    .2588111
       2    -.0011166   .0577987     -0.019   0.985      -.1145374    .1123041
       3    -.0207661   .0547628     -0.379   0.705      -.1282295    .0866973
```

```
       4    (dropped)
------------------------------------------------------------------------------

.         /*      Relationships Among Covariates
>         ==============================*/
>         regress lnwage educ exper;

  Source |       SS       df       MS                 Number of obs =    1003
---------+------------------------------              F(  2,  1000) =  138.21
   Model |  79.3139552     2  39.6569776              Prob > F      =  0.0000
Residual |  286.930819  1000  .286930819              R-squared     =  0.2166
---------+------------------------------              Adj R-squared =  0.2150
   Total |  366.244774  1002  .365513747              Root MSE      =  .53566

------------------------------------------------------------------------------
  lnwage |      Coef.   Std. Err.       t     P>|t|       [ 95% Conf. Interval]
---------+--------------------------------------------------------------------
    educ |   .1034977   .0066008     15.680   0.000       .0905448    .1164507
   exper |   .0128666   .0015419      8.345   0.000       .0098408    .0158924
   _cons |   .8628728   .1007955      8.561   0.000       .6650779    1.060668
------------------------------------------------------------------------------

.                               regress lnwage sex educ;

  Source |       SS       df       MS                 Number of obs =    1003
---------+------------------------------              F(  2,  1000) =  148.54
   Model |   83.883979     2  41.9419895              Prob > F      =  0.0000
Residual |  282.360795  1000  .282360795              R-squared     =  0.2290
---------+------------------------------              Adj R-squared =  0.2275
   Total |  366.244774  1002  .365513747              Root MSE      =  .53138

------------------------------------------------------------------------------
  lnwage |      Coef.   Std. Err.       t     P>|t|       [ 95% Conf. Interval]
---------+--------------------------------------------------------------------
     sex |   -.313703   .0336436     -9.324   0.000      -.3797231    -.247683
    educ |   .0925705   .0064345     14.387   0.000       .0799438    .1051973
   _cons |   1.728516   .1014949     17.031   0.000       1.529349    1.927684
------------------------------------------------------------------------------

.         regress lnwage sex educ exper;

  Source |       SS       df       MS                 Number of obs =    1003
---------+------------------------------              F(  3,   999) =  135.02
   Model |  105.659586     3  35.2198619              Prob > F      =  0.0000
Residual |  260.585189   999  .260846035              R-squared     =  0.2885
---------+------------------------------              Adj R-squared =  0.2864
   Total |  366.244774  1002  .365513747              Root MSE      =  .51073

------------------------------------------------------------------------------
  lnwage |      Coef.   Std. Err.       t     P>|t|       [ 95% Conf. Interval]
---------+--------------------------------------------------------------------
     sex |  -.3252252    .032361    -10.050   0.000      -.3887285   -.2617218
    educ |   .1032311   .0062936     16.402   0.000       .0908808    .1155813
   exper |   .0134428   .0014713      9.137   0.000       .0105556      .01633
```

```
   _cons |   1.331179   .1068058     12.464   0.000        1.12159    1.540769
------------------------------------------------------------------------------

.         /*      Collinearity Among Covariates
>         ============================*/
>         regress lnwage age exper;

  Source |       SS       df       MS                  Number of obs =    1003
---------+------------------------------               F(  2,  1000) =  138.21
   Model |  79.3139552     2  39.6569776               Prob > F      =  0.0000
Residual |  286.930819  1000  .286930819               R-squared     =  0.2166
---------+------------------------------               Adj R-squared =  0.2150
   Total |  366.244774  1002  .365513747               Root MSE      =  .53566

------------------------------------------------------------------------------
  lnwage |      Coef.   Std. Err.       t     P>|t|       [ 95% Conf. Interval]
---------+--------------------------------------------------------------------
     age |   .1034977   .0066008     15.680   0.000       .0905448    .1164507
   exper |  -.0906311   .0064937    -13.957   0.000      -.1033739   -.0778884
   _cons |   .2418865   .1386307      1.745   0.081      -.0301539    .5139269
------------------------------------------------------------------------------

.         regress lnwage age educ exper;

  Source |       SS       df       MS                  Number of obs =    1003
---------+------------------------------               F(  2,  1000) =  138.21
   Model |  79.3139552     2  39.6569776               Prob > F      =  0.0000
Residual |  286.930819  1000  .286930819               R-squared     =  0.2166
---------+------------------------------               Adj R-squared =  0.2150
   Total |  366.244774  1002  .365513747               Root MSE      =  .53566

------------------------------------------------------------------------------
  lnwage |      Coef.   Std. Err.       t     P>|t|       [ 95% Conf. Interval]
---------+--------------------------------------------------------------------
     age |   .1034977   .0066008     15.680   0.000       .0905448    .1164507
    educ |  (dropped)
   exper |  -.0906311   .0064937    -13.957   0.000      -.1033739   -.0778884
   _cons |   .2418865   .1386307      1.745   0.081      -.0301539    .5139269
------------------------------------------------------------------------------

.  /*     Gender Bias Example
>         ===================*/
>                 tab sex,summarize(wage);

Sex 1=Male  |     Summary of Hourly Wages ($)
2=Female    |        Mean   Std. Dev.       Freq.
------------+------------------------------------
       Male |   17.048445   10.240742         537
     Female |    12.14377    7.132306         466
------------+------------------------------------
      Total |   14.769702    9.257249        1003

.         regress lnwage sex;
```

```
  Source |       SS       df       MS                 Number of obs =    1003
---------+------------------------------              F(  1,  1001) =   74.73
   Model |  25.4432335     1  25.4432335              Prob > F      =  0.0000
Residual |  340.801541  1001   .34046108              R-squared     =  0.0695
---------+------------------------------              Adj R-squared =  0.0685
   Total |  366.244774  1002  .365513747              Root MSE      =  .58349

------------------------------------------------------------------------------
  lnwage |      Coef.   Std. Err.       t     P>|t|      [ 95% Conf. Interval]
---------+--------------------------------------------------------------------
     sex |  -.3193424   .0369406     -8.645   0.000      -.3918323   -.2468524
   _cons |   2.982048   .0571544     52.175   0.000       2.869892    3.094204
------------------------------------------------------------------------------

.         regress lnwage sex educ exper;

  Source |       SS       df       MS                 Number of obs =    1003
---------+------------------------------              F(  3,   999) =  135.02
   Model |  105.659586     3  35.2198619              Prob > F      =  0.0000
Residual |  260.585189   999  .260846035              R-squared     =  0.2885
---------+------------------------------              Adj R-squared =  0.2864
   Total |  366.244774  1002  .365513747              Root MSE      =  .51073

------------------------------------------------------------------------------
  lnwage |      Coef.   Std. Err.       t     P>|t|      [ 95% Conf. Interval]
---------+--------------------------------------------------------------------
     sex |  -.3252252    .032361    -10.050   0.000      -.3887285   -.2617218
    educ |   .1032311   .0062936     16.402   0.000       .0908808    .1155813
   exper |   .0134428   .0014713      9.137   0.000       .0105556      .01633
   _cons |   1.331179   .1068058     12.464   0.000        1.12159    1.540769
------------------------------------------------------------------------------

.         /*      Determinants of Wages
>         ====================*/
>         anova lnwage sex educ exper marrstat industry occupat union region,co
> nt(educ exper);

                           Number of obs =    1003     R-squared     =  0.4064
                           Root MSE      = .470993     Adj R-squared =  0.3931

                  Source |  Partial SS    df       MS           F     Prob > F
              -----------+----------------------------------------------------
                   Model |  148.847129    22   6.7657786      30.50     0.0000
                         |
                     sex |  15.8373427     1  15.8373427      71.39     0.0000
                    educ |  20.4331455     1  20.4331455      92.11     0.0000
                   exper |  9.52811313     1  9.52811313      42.95     0.0000
                marrstat |  4.89259269     2  2.44629634      11.03     0.0000
                industry |  15.8810112     8  1.98512639       8.95     0.0000
                 occupat |  19.0558187     5  3.81116374      17.18     0.0000
                   union |  2.64439485     1  2.64439485      11.92     0.0006
                  region |  2.02296603     3   .67432201       3.04     0.0282
                         |
                Residual |  217.397645   980  .221834332
```

```
           -----------+------------------------------------------------------
                Total |  366.244774  1002  .365513747

. anova, regress detail;

  Factor      Value          Value          Value          Value
  ------------------------------------------------------------------------
  sex         1 1            2 2

  marrstat    1 1            2 2            3 3

  industry    1 1            2 2            3 3            4 4
              5 5            6 6            7 7            8 8
              9 9

  occupat     1 1            2 2            3 3            4 4
              5 5            6 6

  union       1 0            2 1

  region      1 1            2 2            3 3            4 4

  Source |       SS       df       MS                 Number of obs =    1003
---------+------------------------------              F( 22,   980) =   30.50
   Model |  148.847129    22   6.7657786              Prob > F      =  0.0000
Residual |  217.397645   980  .221834332              R-squared     =  0.4064
---------+------------------------------              Adj R-squared =  0.3931
   Total |  366.244774  1002  .365513747              Root MSE      =  .47099

------------------------------------------------------------------------------
  lnwage        Coef.   Std. Err.       t     P>|t|      [ 95% Conf. Interval]
------------------------------------------------------------------------------
_cons        1.115323   .1201722      9.281   0.000       .8794984    1.351147
sex
       1     .2984633   .0353235      8.449   0.000       .2291448    .3677817
       2    (dropped)
educ          .071712    .007472      9.597   0.000        .057049     .086375
exper        .0098973   .0015102      6.554   0.000       .0069337    .0128608
marrstat
       1    -.1890965   .0426461     -4.434   0.000      -.2727847   -.1054083
       2    -.0935827   .0407044     -2.299   0.022      -.1734605   -.0137049
       3    (dropped)
industry
       1     .0795472   .0771797      1.031   0.303      -.0719094    .2310037
       2      .237637   .0494519      4.805   0.000       .1405932    .3346808
       3     .2123923   .0618934      3.432   0.001       .0909335     .333851
       4    -.1256094   .0570295     -2.203   0.028      -.2375234   -.0136954
       5     .1080547   .0663295      1.629   0.104      -.0221095    .2382189
       6     .0977911   .0618852      1.580   0.114      -.0236516    .2192338
       7    -.1846039   .0689094     -2.679   0.008      -.3198308   -.0493769
       8     .0940225   .0718175      1.309   0.191      -.0469113    .2349563
       9    (dropped)
occupat
       1     .4142106   .0558632      7.415   0.000       .3045854    .5238358
```

```
       2     .3550925   .0590835      6.010   0.000       .2391479    .4710372
       3     .2395311   .0641775      3.732   0.000         .11359    .3654723
       4     .0780173   .0547025      1.426   0.154      -.0293302    .1853648
       5     -.015846   .0626048     -0.253   0.800      -.1387008    .1070089
       6    (dropped)
union
       1    -.1461691   .0423357     -3.453   0.001      -.2292482     -.06309
       2    (dropped)
region
       1      .087824   .0459206      1.913   0.056        -.00229     .177938
       2    -.0246037   .0461045     -0.534   0.594      -.1150786    .0658711
       3    -.0199084   .0436797     -0.456   0.649      -.1056248    .0658081
       4    (dropped)
------------------------------------------------------------------------------

. /*      Interactions
>         ============*/
>         anova lnwage sex educ exper marrstat industry occupat union region se
> x*educ,cont(educ exper);

                           Number of obs =    1003     R-squared     =  0.4095
                           Root MSE      = .470001     Adj R-squared =  0.3956

                  Source |  Partial SS    df       MS           F     Prob > F
              -----------+----------------------------------------------------
                   Model |  149.982477    23  6.52097728      29.52     0.0000
                         |
                     sex |  3.51225774     1  3.51225774      15.90     0.0001
                    educ |  21.5679622     1  21.5679622      97.64     0.0000
                   exper |   9.6988186     1   9.6988186      43.91     0.0000
                 marrstat |   5.0875474     2   2.5437737      11.52     0.0000
                industry |  16.6974899     8  2.08718623       9.45     0.0000
                 occupat |  18.9493176     5  3.78986351      17.16     0.0000
                   union |  2.45276273     1  2.45276273      11.10     0.0009
                  region |  2.05486362     3   .68495454       3.10     0.0260
                sex*educ |  1.13534824     1  1.13534824       5.14     0.0236
                         |
                Residual |  216.262297   979  .220901223
              -----------+----------------------------------------------------
                   Total |  366.244774  1002  .365513747

.         anova lnwage sex educ exper marrstat industry occupat union region se
> x*exper,cont(educ exper);

                           Number of obs =    1003     R-squared     =  0.4146
                           Root MSE      = .467963     Adj R-squared =  0.4009

                  Source |  Partial SS    df       MS           F     Prob > F
              -----------+----------------------------------------------------
                   Model |  151.854338    23  6.60236251      30.15     0.0000
                         |
                     sex |  .536725148     1  .536725148       2.45     0.1178
                    educ |  20.8475373     1  20.8475373      95.20     0.0000
                   exper |  9.31013309     1  9.31013309      42.51     0.0000
```

```
                marrstat |  4.55683782     2  2.27841891      10.40     0.0000
                industry |   14.816241     8  1.85203012       8.46     0.0000
                 occupat |  18.7144437     5  3.74288873      17.09     0.0000
                   union |  2.34281683     1  2.34281683      10.70     0.0011
                  region |  2.31578067     3   .77192689       3.52     0.0146
                sex*exper |  3.00720853     1  3.00720853      13.73     0.0002
                          |
                Residual |  214.390437   979   .21898921
           -----------+----------------------------------------------------
                   Total |  366.244774  1002  .365513747

.         anova lnwage sex educ exper marrstat industry occupat union region se
> x*marrstat,cont(educ exper);

                           Number of obs =    1003     R-squared     =  0.4096
                           Root MSE      = .470201     Adj R-squared =  0.3951

                  Source |  Partial SS    df       MS           F     Prob > F
           -----------+----------------------------------------------------
                   Model |  150.019983    24  6.25083264      28.27     0.0000
                         |
                     sex |  11.5437561     1  11.5437561      52.21     0.0000
                    educ |  19.7984556     1  19.7984556      89.55     0.0000
                   exper |  9.75819173     1  9.75819173      44.14     0.0000
                marrstat |  4.32516445     2  2.16258223       9.78     0.0001
                industry |  15.0405416     8   1.8800677       8.50     0.0000
                 occupat |  19.2112221     5  3.84224441      17.38     0.0000
                   union |  2.55563936     1  2.55563936      11.56     0.0007
                  region |  2.06211328     3  .687371093       3.11     0.0257
            sex*marrstat |  1.17285424     2  .586427121       2.65     0.0710
                         |
                Residual |  216.224791   978  .221088743
           -----------+----------------------------------------------------
                   Total |  366.244774  1002  .365513747

.         anova lnwage sex educ exper marrstat industry occupat union region se
> x*region,cont(educ exper);

                           Number of obs =    1003     R-squared     =  0.4094
                           Root MSE      = .470538     Adj R-squared =  0.3943

                  Source |  Partial SS    df       MS           F     Prob > F
           -----------+----------------------------------------------------
                   Model |  149.930757    25  5.99723029      27.09     0.0000
                         |
                     sex |  15.4114238     1  15.4114238      69.61     0.0000
                    educ |  20.2535835     1  20.2535835      91.48     0.0000
                   exper |  9.27145529     1  9.27145529      41.88     0.0000
                marrstat |  5.09302412     2  2.54651206      11.50     0.0000
                industry |  15.4350581     8  1.92938227       8.71     0.0000
                 occupat |  18.9433503     5  3.78867007      17.11     0.0000
                   union |  2.32426007     1  2.32426007      10.50     0.0012
                  region |  2.34729189     3  .782430631       3.53     0.0144
              sex*region |    1.083628     3  .361209332       1.63     0.1804
```

```
                    |
           Residual |  216.314017   977  .221406363
          ----------+-----------------------------------------------------
              Total |  366.244774  1002  .365513747

.         tab occupat*sex,summarize(lnwage);

        Means, Standard Deviations and Frequencies of Log of wages

           | Sex 1=Male 2=Female
Occupation |      Male     Female |     Total
-----------+----------------------+----------
  Managemt | 2.9782528  2.7178281 | 2.8608064
           |  .6133342  .52977117 | .58980804
           |        84         69 |       153
-----------+----------------------+----------
   Profess | 2.9750666  2.6510632 | 2.7990851
           | .54732101  .48693462 | .53892232
           |        90        107 |       197
-----------+----------------------+----------
     Sales | 2.6561594  2.1425782 | 2.4479508
           | .61028957  .55250281 | .63745511
           |        66         45 |       111
-----------+----------------------+----------
  Clerical | 2.5266198  2.2201788 | 2.2706514
           | .49940685   .4528144 | .47319296
           |        28        142 |       170
-----------+----------------------+----------
   Service | 2.4296997  1.9573989 | 2.1711772
           | .59639105  .47970747 | .58277632
           |        43         52 |        95
-----------+----------------------+----------
  Bluecoll |  2.484136  2.1048479 | 2.4143032
           | .53407544  .50854772 | .54869832
           |       226         51 |       277
-----------+----------------------+----------
     Total | 2.6627056  2.3433633 | 2.5143372
           |  .6033977  .55966571 | .60457733
           |       537        466 |      1003

.         anova lnwage sex occupat sex*occupat;

                           Number of obs =    1003     R-squared     =  0.2377
                           Root MSE      =  .53076     Adj R-squared =  0.2293

                  Source |  Partial SS    df       MS           F     Prob > F
              -----------+----------------------------------------------------
                   Model |  87.0735468    11  7.91577698      28.10     0.0000
                         |
                     sex |  26.3028947     1  26.3028947      93.37     0.0000
                 occupat |  54.9837252     5   10.996745      39.04     0.0000
             sex*occupat |  1.44653423     5  .289306846       1.03     0.4003
                         |
                Residual |  279.171228   991  .281706587
```

```
           -----------+---------------------------------------------------
                Total |  366.244774  1002  .365513747

.         /*      Racial Bias Example
>         ==================*/
>         tab race,summarize(wage);

Race 1=oth  |     Summary of Hourly Wages ($)
2=blk 3=hisp|
4=wh        |        Mean   Std. Dev.       Freq.
------------+------------------------------------
      Other |   13.879308   8.2283586          52
      Black |   12.467623   7.5165665         102
   Hispanic |   9.8324086   6.9178947          61
      White |   15.508646   9.5208053         788
------------+------------------------------------
      Total |   14.769702    9.257249        1003

.         anova lnwage race;

                           Number of obs =    1003     R-squared     =  0.0399
                           Root MSE      = .593296     Adj R-squared =  0.0370

                  Source |  Partial SS    df       MS           F     Prob > F
           --------------+----------------------------------------------------
                   Model |  14.5965218     3  4.86550727      13.82     0.0000
                         |
                    race |  14.5965218     3  4.86550727      13.82     0.0000
                         |
                Residual |  351.648253   999  .352000253
           --------------+----------------------------------------------------
                   Total |  366.244774  1002  .365513747

.                anova,regress detail;

  Factor      Value          Value          Value          Value
  ------------------------------------------------------------------------
  race        1 1            2 2            3 3            4 4

  Source |       SS       df       MS                  Number of obs =    1003
---------+------------------------------              F(  3,   999) =   13.82
   Model |  14.5965218     3  4.86550727              Prob > F      =  0.0000
Residual |  351.648253   999  .352000253             R-squared     =  0.0399
---------+------------------------------              Adj R-squared =  0.0370
   Total |  366.244774  1002  .365513747             Root MSE      =   .5933

------------------------------------------------------------------------------
  lnwage        Coef.   Std. Err.       t     P>|t|       [ 95% Conf. Interval]
------------------------------------------------------------------------------
_cons        2.569125   .0211353    121.556   0.000       2.527651      2.6106
race
       1    -.0915726   .0849467     -1.078   0.281       -.258267    .0751217
       2    -.2285605   .0624314     -3.661   0.000      -.3510723   -.1060487
       3    -.4406173   .0788492     -5.588   0.000      -.5953464   -.2858883
```

```
       4    (dropped)
------------------------------------------------------------------------------

.         anova lnwage race sex;

                           Number of obs =    1003     R-squared     =  0.1053
                           Root MSE      = .573021     Adj R-squared =  0.1017

                  Source |  Partial SS    df       MS           F     Prob > F
              -----------+----------------------------------------------------
                   Model |  38.5483685     4  9.63709213      29.35     0.0000
                         |
                    race |   13.105135     3  4.36837833      13.30     0.0000
                     sex |  23.9518467     1  23.9518467      72.95     0.0000
                         |
                Residual |  327.696406   998  .328353112
              -----------+----------------------------------------------------
                   Total |  366.244774  1002  .365513747

.                 anova,regress detail;

  Factor      Value          Value          Value          Value
  ------------------------------------------------------------------------
  race        1 1            2 2            3 3            4 4

  sex         1 1            2 2

  Source |       SS       df       MS                 Number of obs =    1003
---------+------------------------------              F(  4,   998) =   29.35
   Model |  38.5483685     4  9.63709213              Prob > F      =  0.0000
Residual |  327.696406   998  .328353112              R-squared     =  0.1053
---------+------------------------------              Adj R-squared =  0.1017
   Total |  366.244774  1002  .365513747              Root MSE      =  .57302

------------------------------------------------------------------------------
  lnwage        Coef.   Std. Err.       t     P>|t|       [ 95% Conf. Interval]
------------------------------------------------------------------------------
_cons        2.397275   .0286627     83.638   0.000       2.341029    2.453521
race
       1    -.0873486   .0820452     -1.065   0.287      -.2483495    .0736524
       2    -.1787909   .0605788     -2.951   0.003      -.2976675   -.0599144
       3    -.4422817   .0761549     -5.808   0.000      -.5917238   -.2928397
       4    (dropped)
sex
       1     .3113055   .0364492      8.541   0.000       .2397797    .3828313
       2    (dropped)
------------------------------------------------------------------------------

.                 anova lnwage race sex educ exper,cont(educ exper);

                           Number of obs =    1003     R-squared     =  0.2963
                           Root MSE      = .508678     Adj R-squared =  0.2921

                  Source |  Partial SS    df       MS           F     Prob > F
```

```
-----------+----------------------------------------------------
     Model |  108.526139     6  18.0876898      69.90     0.0000
           |
      race |  2.86655312     3  .955517707       3.69     0.0116
       sex |  24.8878915     1  24.8878915      96.18     0.0000
      educ |  59.2169815     1  59.2169815     228.85     0.0000
     exper |  21.9313012     1  21.9313012      84.76     0.0000
           |
  Residual |  257.718635   996   .25875365
-----------+----------------------------------------------------
     Total |  366.244774  1002  .365513747

.         anova,regress detail;

  Factor      Value          Value          Value          Value
  ------------------------------------------------------------------
  race        1 1            2 2            3 3            4 4

  sex         1 1            2 2

  Source |       SS       df       MS                 Number of obs =    1003
---------+------------------------------              F(  6,   996) =   69.90
   Model |  108.526139     6  18.0876898              Prob > F      =  0.0000
Residual |  257.718635   996   .25875365              R-squared     =  0.2963
---------+------------------------------              Adj R-squared =  0.2921
   Total |  366.244774  1002  .365513747              Root MSE      =  .50868

------------------------------------------------------------------------------
  lnwage        Coef.   Std. Err.       t     P>|t|       [ 95% Conf. Interval]
------------------------------------------------------------------------------
_cons         .7676203   .1028365      7.464   0.000       .5658192    .9694215
race
       1    -.0668473   .0728629     -0.917   0.359      -.2098298    .0761352
       2    -.1525117    .054221     -2.813   0.005      -.2589121   -.0461113
       3    -.1383799   .0704215     -1.965   0.050      -.2765715   -.0001884
       4    (dropped)
sex
       1     .3174755   .0323712      9.807   0.000       .2539518    .3809992
       2    (dropped)
educ          .0988998   .0065376     15.128   0.000       .0860708    .1117288
exper         .0135874   .0014759      9.206   0.000       .0106912    .0164836
------------------------------------------------------------------------------

.         anova lnwage race sex educ exper marrstat occupat industry region uni
> on,cont(educ exper);

                           Number of obs =    1003     R-squared     =  0.4097
                           Root MSE      = .470409     Adj R-squared =  0.3946

                  Source |  Partial SS    df       MS           F     Prob > F
              -----------+----------------------------------------------------
                   Model |  150.050045    25  6.00200181      27.12     0.0000
                         |
                    race |  1.20291601     3  .400972003       1.81     0.1433
```

```
                    sex |    15.60298     1    15.60298      70.51     0.0000
                   educ |  17.3336535     1  17.3336535      78.33     0.0000
                  exper |  9.48502631     1  9.48502631      42.86     0.0000
                marrstat |  4.35481743     2  2.17740872       9.84     0.0001
                 occupat |  18.0189269     5  3.60378538      16.29     0.0000
                industry |   15.828276     8  1.97853451       8.94     0.0000
                  region |   1.9230824     3  .641027467       2.90     0.0342
                   union |  2.78293087     1  2.78293087      12.58     0.0004
                         |
                Residual |  216.194729   977  .221284267
           -----------+----------------------------------------------------
                   Total |  366.244774  1002  .365513747

.         anova,regress detail;

  Factor      Value          Value          Value          Value
  ------------------------------------------------------------------------
  race        1 1            2 2            3 3            4 4

  sex         1 1            2 2

  marrstat    1 1            2 2            3 3

  occupat     1 1            2 2            3 3            4 4
              5 5            6 6

  industry    1 1            2 2            3 3            4 4
              5 5            6 6            7 7            8 8
              9 9

  region      1 1            2 2            3 3            4 4

  union       1 0            2 1

  Source |       SS       df       MS                 Number of obs =    1003
---------+------------------------------              F( 25,   977) =   27.12
   Model |  150.050045    25  6.00200181              Prob > F      =  0.0000
Residual |  216.194729   977  .221284267             R-squared     =  0.4097
---------+------------------------------              Adj R-squared =  0.3946
   Total |  366.244774  1002  .365513747             Root MSE      =  .47041

------------------------------------------------------------------------------
  lnwage        Coef.   Std. Err.       t     P>|t|      [ 95% Conf. Interval]
------------------------------------------------------------------------------
_cons        1.199561   .1271574      9.434   0.000       .9500279    1.449094
race
       1    -.0728806   .0695287     -1.048   0.295      -.2093235    .0635622
       2    -.0688875   .0540475     -1.275   0.203      -.1749501    .0371751
       3    -.1284594   .0682329     -1.883   0.060      -.2623593    .0054404
       4    (dropped)
sex
       1     .2964133   .0352996      8.397   0.000       .2271416     .365685
       2    (dropped)
educ         .0682501   .0077114      8.851   0.000       .0531173    .0833829
```

```
exper         .0099197   .0015151      6.547   0.000       .0069464     .012893
marrstat
       1    -.1800136   .0430294     -4.183   0.000      -.2644544   -.0955729
       2    -.0918757   .0411279     -2.234   0.026      -.1725849   -.0111664
       3    (dropped)
occupat
       1     .4045487   .0559685      7.228   0.000       .2947163     .514381
       2     .3600429   .0593901      6.062   0.000       .2434961    .4765896
       3     .2316799   .0642125      3.608   0.000       .1056697    .3576901
       4      .074774   .0546834      1.367   0.172      -.0325364    .1820843
       5    -.0067188   .0635746     -0.106   0.916      -.1314773    .1180396
       6    (dropped)
industry
       1     .0717064   .0771935      0.929   0.353      -.0797778    .2231906
       2     .2325472   .0494642      4.701   0.000       .1354788    .3296155
       3      .208118   .0619676      3.358   0.001       .0865131    .3297229
       4    -.1302036   .0570393     -2.283   0.023      -.2421372   -.0182699
       5     .1086729   .0662985      1.639   0.102       -.021431    .2387767
       6     .0957432   .0621041      1.542   0.123      -.0261295    .2176159
       7    -.1899612     .06914     -2.747   0.006      -.3256412   -.0542811
       8     .0971525   .0720495      1.348   0.178      -.0442372    .2385422
       9    (dropped)
region
       1     .0698249   .0471064      1.482   0.139      -.0226165    .1622663
       2    -.0477362   .0479407     -0.996   0.320      -.1418147    .0463424
       3    -.0280159   .0455055     -0.616   0.538      -.1173158    .0612839
       4    (dropped)
union
       1    -.1507956   .0425219     -3.546   0.000      -.2342404   -.0673508
       2    (dropped)
------------------------------------------------------------------------------

.         tab sex*race,summarize(lnwage);

        Means, Standard Deviations and Frequencies of Log of wages

Sex 1=Male | Race 1=oth 2=blk 3=hisp 4=wh
2=Female   |     Other      Black   Hispanic      White |     Total
-----------+--------------------------------------------+----------
      Male | 2.6435375  2.5329479  2.2209492  2.7103993 | 2.6627056
           | .57418304  .59177306  .54653608   .5965086 |  .6033977
           |        28         40         34        435 |       537
-----------+--------------------------------------------+----------
    Female | 2.2839041   2.216447  2.0121009  2.3950345 | 2.3433633
           | .48331628  .62129102  .49859648  .54688872 | .55966571
           |        24         62         27        353 |       466
-----------+--------------------------------------------+----------
     Total | 2.4775528   2.340565  2.1285081  2.5691255 | 2.5143372
           | .55915946   .6264733  .53193239  .59550002 | .60457733
           |        52        102         61        788 |      1003

.         anova lnwage sex race sex*race exper educ industry occupat union regi
> on, cont(exper educ);
```

```
                            Number of obs =    1003     R-squared     =  0.3981
                            Root MSE      = .475247     Adj R-squared =  0.3821

                   Source |  Partial SS    df       MS           F     Prob > F
              ------------+------------------------------------------------------
                    Model |  145.805381    26  5.60789928      24.83     0.0000
                          |
                      sex |  6.77861667     1  6.77861667      30.01     0.0000
                     race |  1.65604811     3  .552016038       2.44     0.0627
                 sex*race |  .110153539     3  .036717846       0.16     0.9216
                    exper |  15.2500611     1  15.2500611      67.52     0.0000
                     educ |  17.6863427     1  17.6863427      78.31     0.0000
                 industry |  16.1366448     8  2.01708059       8.93     0.0000
                  occupat |  17.9162433     5  3.58324867      15.86     0.0000
                    union |  2.46585957     1  2.46585957      10.92     0.0010
                   region |  1.58228225     3  .527427415       2.34     0.0724
                          |
                 Residual |  220.439393   976  .225860034
              ------------+------------------------------------------------------
                    Total |  366.244774  1002  .365513747

. 	/*	Drop if yhat less then 1.5
> 	==========================*/
> 	fit lnwage sex educ exper marrstat industry occupat union region;

  Source |       SS       df       MS                  Number of obs =    1003
---------+------------------------------               F(  8,   994) =   69.91
   Model |  131.868203     8  16.4835254               Prob > F      =  0.0000
Residual |  234.376571   994  .235791319               R-squared     =  0.3601
---------+------------------------------               Adj R-squared =  0.3549
   Total |  366.244774  1002  .365513747               Root MSE      =  .48558

------------------------------------------------------------------------------
  lnwage |      Coef.   Std. Err.       t     P>|t|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
     sex |  -.3332264   .0320681    -10.391   0.000      -.3961554   -.2702974
    educ |   .0737519   .0071861     10.263   0.000       .0596502    .0878535
   exper |   .0109027   .0014605      7.465   0.000       .0080367    .0137687
marrstat |   .0962107   .0204508      4.704   0.000        .056079    .1363424
industry |  -.0279774   .0060103     -4.655   0.000      -.0397717   -.0161831
 occupat |  -.0868315   .0104206     -8.333   0.000      -.1072804   -.0663825
   union |   .1475521   .0412211      3.580   0.000       .0666617    .2284425
  region |  -.0350369   .0145282     -2.412   0.016      -.0635464   -.0065275
   _cons |   2.075167   .1557963     13.320   0.000       1.769439    2.380894
------------------------------------------------------------------------------

. 	predict yhat;

. 	drop if yhat<1.5;
(2 observations deleted)

. 	anova lnwage sex educ exper marrstat industry occupat union region,co
> nt(educ exper);
```

```
                            Number of obs =    1001     R-squared     =  0.4027
                            Root MSE      = .471285     Adj R-squared =  0.3893

                    Source |  Partial SS    df       MS           F     Prob > F
                -----------+----------------------------------------------------
                     Model |  146.448119    22  6.65673268      29.97     0.0000
                           |
                       sex |  15.6152103     1  15.6152103      70.30     0.0000
                      educ |  19.8211764     1  19.8211764      89.24     0.0000
                     exper |   9.5369837     1   9.5369837      42.94     0.0000
                  marrstat |  4.77586664     2  2.38793332      10.75     0.0000
                  industry |  15.9300972     8  1.99126214       8.97     0.0000
                   occupat |  19.1251053     5  3.82502105      17.22     0.0000
                     union |  2.61115764     1  2.61115764      11.76     0.0006
                    region |  2.01643888     3  .672146294       3.03     0.0287
                           |
                  Residual |  217.223208   978   .22210962
                -----------+----------------------------------------------------
                     Total |  363.671327  1000  .363671327

. anova, regress detail;

  Factor      Value          Value          Value          Value
  ------------------------------------------------------------------------
  sex         1 1            2 2

  marrstat    1 1            2 2            3 3

  industry    1 1            2 2            3 3            4 4
              5 5            6 6            7 7            8 8
              9 9

  occupat     1 1            2 2            3 3            4 4
              5 5            6 6

  union       1 0            2 1

  region      1 1            2 2            3 3            4 4

  Source |       SS       df       MS                 Number of obs =    1001
---------+------------------------------              F( 22,   978) =   29.97
   Model |  146.448119    22  6.65673268              Prob > F      =  0.0000
Residual |  217.223208   978   .22210962              R-squared     =  0.4027
---------+------------------------------              Adj R-squared =  0.3893
   Total |  363.671327  1000  .363671327              Root MSE      =  .47129

------------------------------------------------------------------------------
  lnwage        Coef.   Std. Err.       t     P>|t|       [ 95% Conf. Interval]
------------------------------------------------------------------------------
_cons         1.125143   .1216981      9.245   0.000       .8863233    1.363962
sex
       1      .2969846   .0354196      8.385   0.000       .2274774    .3664918
       2     (dropped)
educ          .0711323   .0075298      9.447   0.000       .0563558    .0859087
```

exper	.0099026	.0015112	6.553	0.000	.006937	.0128683
marrstat						
1	-.1866665	.0427751	-4.364	0.000	-.2706081	-.1027249
2	-.0939289	.0407437	-2.305	0.021	-.173884	-.0139739
3	(dropped)					
industry						
1	.0783552	.0773551	1.013	0.311	-.0734458	.2301563
2	.239158	.0496305	4.819	0.000	.1417635	.3365526
3	.2126268	.0620064	3.429	0.001	.0909459	.3343078
4	-.1261923	.0571284	-2.209	0.027	-.2383006	-.0140839
5	.1083631	.0663964	1.632	0.103	-.0219327	.2386588
6	.0980023	.0619615	1.582	0.114	-.0235904	.219595
7	-.1831428	.0689824	-2.655	0.008	-.3185132	-.0477723
8	.0950347	.0718892	1.322	0.186	-.0460402	.2361096
9	(dropped)					
occupat						
1	.4138555	.0559168	7.401	0.000	.3041248	.5235863
2	.3551804	.0591269	6.007	0.000	.2391501	.4712106
3	.2388882	.0642535	3.718	0.000	.1127977	.3649787
4	.0761378	.0548508	1.388	0.165	-.031501	.1837767
5	-.0175453	.0627493	-0.280	0.780	-.1406841	.1055934
6	(dropped)					
union						
1	-.1452958	.042376	-3.429	0.001	-.2284541	-.0621374
2	(dropped)					
region						
1	.0855722	.046066	1.858	0.064	-.0048274	.1759718
2	-.0267552	.0462439	-0.579	0.563	-.1175039	.0639935
3	-.0222971	.0438503	-0.508	0.611	-.1083485	.0637544
4	(dropped)					