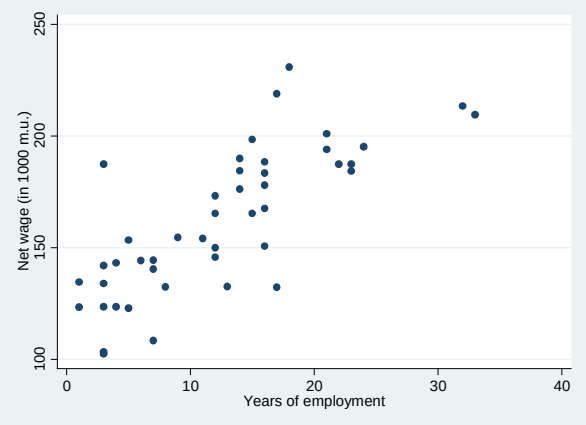
Example 1: For 45 employees we gathered data in the file `wage1.dta` on their net wage (in 1,000 m.u.), period of employment (in years) and gender (a variable that equals 1 for males and 2 for females). For gender we introduce a dummy variable *D*, which has value of 1 for males and value of 0 for females.

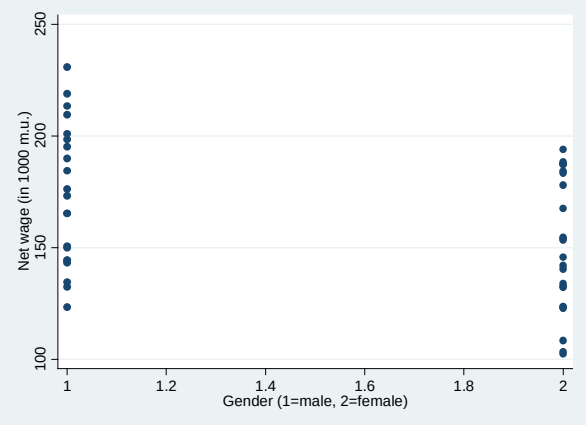
Estimate the regression models in which the dummy variable will appear in various ways and explain the obtained regression coefficients.

Computer printout of the results in Stata:

```
. scatter wage employment
```




```
. scatter wage gender
```




```
. gen d=1
. replace d=0 if gender==2
(24 real changes made)

. label variable d "Gender of the employee (1=male, 0=female)"

. gen dalt=1
. replace dalt=0 if gender==1
(21 real changes made)

. label variable dalt "Gender of the employee (1=female, 0=male)"

. regress wage d

      Source |       SS       df       MS              Number of obs =      45
-------------+------------------------------           F(  1,    43) =    5.90
       Model |  5549.44883     1  5549.44883           Prob > F      =  0.0194
    Residual |  40416.5491    43  939.919746           R-squared     =  0.1207
-------------+------------------------------           Adj R-squared =  0.1003
       Total |  45965.9979    44  1044.68177           Root MSE      =  30.658

------------------------------------------------------------------------------
        wage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           d |   22.25953   9.160863     2.43   0.019     3.784886    40.73417
       _cons |   151.3167   6.258061    24.18   0.000     138.6961    163.9372
------------------------------------------------------------------------------
```

```
. tab d, sum(wage)

  Gender of |
        the |
   employee |
   (1=male, | Summary of Net wage (in 1000 m.u.)
  0=female) |        Mean   Std. Dev.       Freq.
------------+------------------------------------
          0 |   151.31667   29.742015          24
          1 |   173.57619   31.678887          21
------------+------------------------------------
      Total |   161.70444   32.321537          45

. regress wage employment d

      Source |       SS       df       MS              Number of obs =      45
-------------+------------------------------           F(  2,    42) =   36.07
       Model |     29052.7     2    14526.35           Prob > F      =  0.0000
    Residual |  16913.2979    42  402.697568           R-squared     =  0.6320
-------------+------------------------------           Adj R-squared =  0.6145
       Total |  45965.9979    44  1044.68177           Root MSE      =  20.067

------------------------------------------------------------------------------
        wage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  employment |   2.975623   .3894964     7.64   0.000     2.189588    3.761659
           d |   12.67731   6.126038     2.07   0.045      .314466    25.04015
       _cons |   118.9568    5.89244    20.19   0.000     107.0653    130.8482
------------------------------------------------------------------------------

. gen demployment=d*employment

. regress wage employment d demployment

      Source |       SS       df       MS              Number of obs =      45
-------------+------------------------------           F(  3,    41) =   23.48
       Model |  29052.7495     3  9684.24985           Prob > F      =  0.0000
    Residual |  16913.2484    41  412.518253           R-squared     =  0.6320
-------------+------------------------------           Adj R-squared =  0.6051
       Total |  45965.9979    44  1044.68177           Root MSE      =  20.311

------------------------------------------------------------------------------
        wage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  employment |   2.970835   .5886187     5.05   0.000     1.782095    4.159574
           d |   12.57033   11.56706     1.09   0.284    -10.78982    35.93047
 demployment |    .008684   .7926433     0.01   0.991    -1.592092     1.60946
       _cons |   119.0088   7.626532    15.60   0.000     103.6067    134.4109
------------------------------------------------------------------------------

. list wage employment d demployment

     +---------------------------------+
     |  wage   employ~t   d   demplo~t |
     |---------------------------------|
  1. | 150.6         16   1         16 |
  2. | 213.4         32   1         32 |
  3. | 108.4          7   0          0 |
  4. | 123.6          4   0          0 |
  5. |   194         21   0          0 |
     |---------------------------------|
  6. | 154.1         11   0          0 |
  7. | 184.5         14   1         14 |
  8. | 173.2         12   1         12 |
  9. | 167.5         16   0          0 |
```

```
 10. | 144.3          6   1          6 |
     |----------------------------------|
 11. | 165.3         15   1         15 |
 12. | 103.4          3   0          0 |
 13. | 154.7          9   0          0 |
 14. |   219         17   1         17 |
 15. | 188.4         16   0          0 |
     |----------------------------------|
 16. |   201         21   1         21 |
 17. | 153.4          5   0          0 |
 18. | 132.7         13   0          0 |
 19. | 183.4         16   0          0 |
 20. | 165.3         12   1         12 |
     |----------------------------------|
 21. | 187.4         22   0          0 |
 22. | 132.4         17   0          0 |
 23. | 123.6          3   0          0 |
 24. | 176.3         14   1         14 |
 25. | 187.4         23   0          0 |
     |----------------------------------|
 26. | 134.5          1   1          1 |
 27. | 102.3          3   0          0 |
 28. | 198.4         15   1         15 |
 29. |   150         12   1         12 |
 30. | 140.4          7   0          0 |
     |----------------------------------|
 31. | 184.3         23   0          0 |
 32. | 143.1          4   1          4 |
 33. | 187.5          3   0          0 |
 34. | 132.5          8   1          8 |
 35. |   190         14   1         14 |
     |----------------------------------|
 36. | 145.7         12   0          0 |
 37. | 123.4          1   1          1 |
 38. |   142          3   0          0 |
 39. | 195.3         24   1         24 |
 40. |   123          5   0          0 |
     |----------------------------------|
 41. | 144.4          7   1          7 |
 42. |   134          3   0          0 |
 43. |   231         18   1         18 |
 44. |   178         16   0          0 |
 45. | 209.6         33   1         33 |
     +----------------------------------+
```

. regress wage d dalt

```
note: dalt omitted because of collinearity

      Source |       SS       df       MS              Number of obs =      45
-------------+------------------------------           F(  1,    43) =    5.90
       Model |  5549.44883     1  5549.44883           Prob > F      =  0.0194
    Residual |  40416.5491    43  939.919746           R-squared     =  0.1207
-------------+------------------------------           Adj R-squared =  0.1003
       Total |  45965.9979    44  1044.68177           Root MSE      =  30.658

------------------------------------------------------------------------------
        wage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           d |   22.25953   9.160863     2.43   0.019     3.784886    40.73417
        dalt |  (omitted)
       _cons |   151.3167   6.258061    24.18   0.000     138.6961    163.9372
------------------------------------------------------------------------------
```

```
. regress wage employment d dalt
note: dalt omitted because of collinearity

      Source |       SS       df       MS              Number of obs =      45
-------------+------------------------------           F(  2,    42) =   36.07
       Model |     29052.7     2    14526.35           Prob > F      =  0.0000
    Residual |  16913.2979    42  402.697568           R-squared     =  0.6320
-------------+------------------------------           Adj R-squared =  0.6145
       Total |  45965.9979    44  1044.68177           Root MSE      =  20.067

------------------------------------------------------------------------------
        wage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  employment |   2.975623   .3894964     7.64   0.000     2.189588    3.761659
           d |   12.67731   6.126038     2.07   0.045      .314466    25.04015
        dalt |  (omitted)
       _cons |   118.9568    5.89244    20.19   0.000     107.0653    130.8482
------------------------------------------------------------------------------
```

■

Example 2: We gathered a sample of data for 32 European countries for the year 2003. We have the following variables available (the data are provided in Stata Data file `health.dta`, while the programming code is given in Stata Do file `health-commands-l12.do`):

- life expectancy at birth (*LIFE*; in years);
- health expenditure per capita (*EXP*; in U.S. dollars);
- percentage of smokers among adults (*TOBACCO*);
- consumption of alcohol per capita (*ALCO*; in liters of distilled spirits).

We divided the countries into two groups based on whether they are EU15 member states (in this case the dummy variable *DEU* equals 1) or not (in this case the dummy variable *DEU* equals 0).

We estimated separately for each group the following regression model:

$$LIFE_i = \beta_1 + \beta_2 EXP_i + \beta_3 ALCO_i + \beta_4 TOBACCO_i + u_i,$$

and found based on the Chow test that the analyzed regression function differs between the aforementioned groups of countries. Fill in the findings with the use of dummy variables.

Computer printout of the results in Stata:

```
. regress life exp alco tobacco

      Source |       SS       df       MS              Number of obs =      32
-------------+------------------------------           F(  3,    28) =   26.30
       Model |  413.850212     3  137.950071           Prob > F      =  0.0000
    Residual |  146.874565    28  5.24552017           R-squared     =  0.7381
-------------+------------------------------           Adj R-squared =  0.7100
       Total |  560.724777    31   18.087896           Root MSE      =  2.2903
```

```
------------------------------------------------------------------------------
        life |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         exp |   .0018569   .0004023     4.62   0.000     .0010329    .0026809
        alco |  -.6493606   .2805689    -2.31   0.028     -1.22408   -.0746412
     tobacco |  -.2238391   .0837702    -2.67   0.012    -.3954346   -.0522436
       _cons |   81.42053   2.720683    29.93   0.000     75.84746    86.99359
------------------------------------------------------------------------------
```

```
. gen dexp=deu*exp
. gen dalco=deu*alco
. gen dtobacco=deu*tobacco
```

```
. regress life exp alco tobacco deu
```

```
      Source |       SS       df       MS              Number of obs =      32
-------------+------------------------------           F(  4,    27) =   24.56
       Model |  439.854433     4  109.963608           Prob > F      =  0.0000
    Residual |  120.870344    27  4.47667939           R-squared     =  0.7844
-------------+------------------------------           Adj R-squared =  0.7525
       Total |  560.724777    31   18.087896           Root MSE      =  2.1158

------------------------------------------------------------------------------
        life |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         exp |   .0015803   .0003889     4.06   0.000     .0007823    .0023783
        alco |  -.4965685   .2668332    -1.86   0.074    -1.044065     .050928
     tobacco |  -.2039382   .0778272    -2.62   0.014    -.3636265   -.0442499
         deu |   2.144349   .8897162     2.41   0.023     .3188019    3.969896
       _cons |   79.81336   2.600354    30.69   0.000     74.47788    85.14885
------------------------------------------------------------------------------
```

```
. regress life exp alco tobacco deu dexp dalco dtobacco
```

```
      Source |       SS       df       MS              Number of obs =      32
-------------+------------------------------           F(  7,    24) =   23.59
       Model |  489.576762     7  69.9395375           Prob > F      =  0.0000
    Residual |  71.1480148    24  2.96450062           R-squared     =  0.8731
-------------+------------------------------           Adj R-squared =  0.8361
       Total |  560.724777    31   18.087896           Root MSE      =  1.7218

------------------------------------------------------------------------------
        life |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         exp |   .0015388   .0003798     4.05   0.000     .0007548    .0023227
        alco |  -.5335537   .2317374    -2.30   0.030    -1.011836   -.0552712
     tobacco |  -.4019988   .0917503    -4.38   0.000    -.5913622   -.2126354
         deu |  -7.934472   4.162727    -1.91   0.069    -16.52592    .6569745
        dexp |  -.0014602   .0007901    -1.85   0.077    -.0030908    .0001705
       dalco |   .6710069   .8493576     0.79   0.437    -1.081981    2.423995
    dtobacco |   .4129869   .1349227     3.06   0.005     .1345201    .6914537
       _cons |   85.90512   3.096202    27.75   0.000     79.51487    92.29537
------------------------------------------------------------------------------
```

```
. test deu=dexp=dalco=0
```

```
 ( 1)  deu - dexp = 0
 ( 2)  deu - dalco = 0
 ( 3)  deu = 0

       F(  3,    24) =    4.20
            Prob > F =    0.0159
```

```
. test dexp=dalco=0

 ( 1)  dexp - dalco = 0
 ( 2)  dexp = 0

       F(  2,    24) =    2.20
            Prob > F =    0.1332

. regress life exp alco tobacco deu dtobacco

      Source |       SS       df       MS              Number of obs =      32
-------------+------------------------------           F(  5,    26) =   29.44
       Model |  476.562498     5  95.3124995           Prob > F      =  0.0000
    Residual |  84.1622793    26  3.23701074           R-squared     =  0.8499
-------------+------------------------------           Adj R-squared =  0.8210
       Total |  560.724777    31   18.087896           Root MSE      =  1.7992

------------------------------------------------------------------------------
        life |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         exp |   .0012305   .0003467     3.55   0.001      .000518    .0019431
        alco |  -.5723521    .228013    -2.51   0.019     -1.04104   -.1036647
     tobacco |  -.4311841   .0945176    -4.56   0.000    -.6254678   -.2369004
         deu |  -10.55756   3.847028    -2.74   0.011    -18.46524    -2.64988
    dtobacco |   .4550991   .1351442     3.37   0.002     .1773063    .7328919
       _cons |   87.19745   3.114081    28.00   0.000     80.79636    93.59853
------------------------------------------------------------------------------
```

■