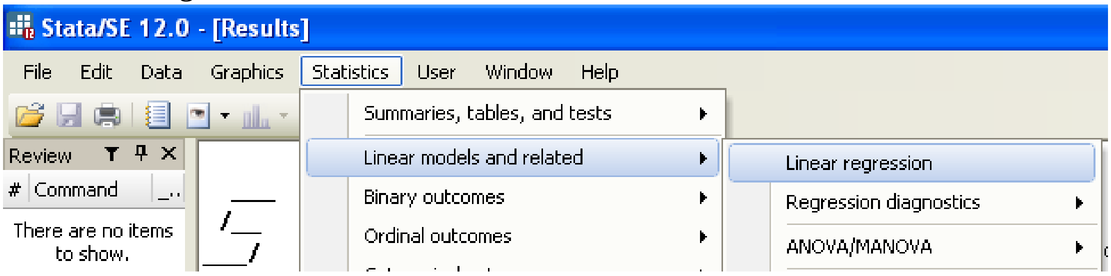
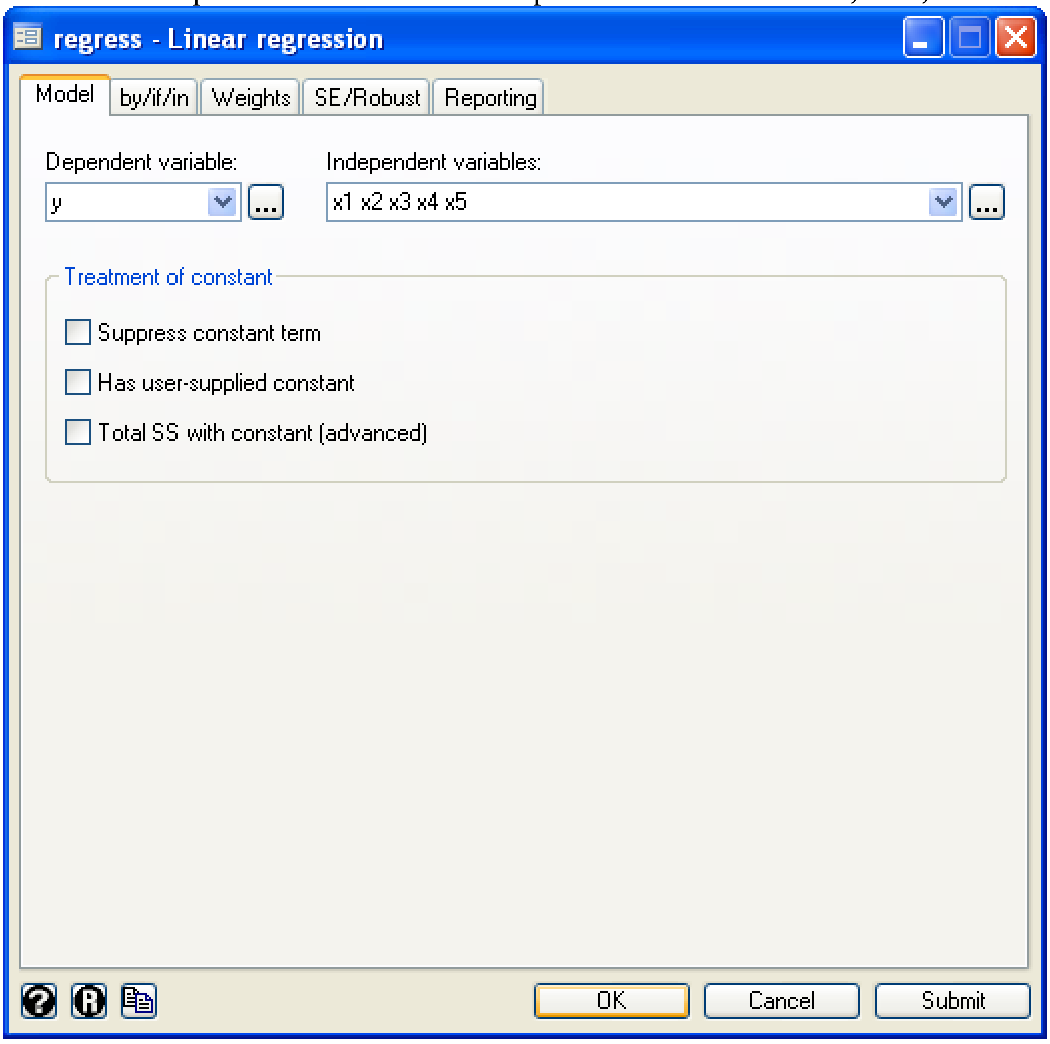
Multiple Regression Model

The model: $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + u_i$

To run a regression, from menu bar go to Statistics, choose Linear models and related, select Liner regression.



From regress – Linear regression window, specify dependent and independent variables in Dependent variable: and Independent variables: boxes, then, click OK.



The results will be as follows:

```
. reg y x1 x2 x3 x4 x5

      Source |       SS       df       MS              Number of obs =      23
-------------+------------------------------           F(  5,    17) =   57.63
       Model |  1129.30649     5  225.861299           Prob > F      =  0.0000
    Residual |   66.622243    17  3.91895547           R-squared     =  0.9443
-------------+------------------------------           Adj R-squared =  0.9279
       Total |  1195.92874    22  54.3603971           Root MSE      =  1.9796

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .0048894   .0049619     0.99   0.338    -.0055794    .0153581
          x2 |  -.6518875   .1743999    -3.74   0.002    -1.019839   -.2839359
          x3 |   .2432418   .0895442     2.72   0.015       .05432    .4321636
          x4 |   .1043175   .0706436     1.48   0.158    -.0447275    .2533625
          x5 |  -.0711103   .0983811    -0.72   0.480    -.2786762    .1364556
       _cons |   38.59691   4.214487     9.16   0.000     29.70512     47.4887
------------------------------------------------------------------------------
```

Hypotheses Testing

To test whether we should eliminate x_4 and x_5 out of the model, or H_0: $\beta_4 = \beta_5 = 0$.

```
. test (x4=0) (x5=0)

 ( 1)  x4 = 0
 ( 2)  x5 = 0

       F(  2,    17) =    1.17
            Prob > F =    0.3354
```

To test whether $\beta_4 = \beta_5$.

```
. test x4 = x5

 ( 1)  x4 - x5 = 0

       F(  1,    17) =    1.27
            Prob > F =    0.2746


. regress y x1 x2 x3

      Source |       SS       df       MS              Number of obs =      23
-------------+------------------------------           F(  3,    19) =   93.65
       Model |  1120.17021     3  373.390069           Prob > F      =  0.0000
    Residual |  75.7585287    19  3.98729099           R-squared     =  0.9367
-------------+------------------------------           Adj R-squared =  0.9267
       Total |  1195.92874    22  54.3603971           Root MSE      =  1.9968

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .0108762   .0023812     4.57   0.000     .0058922    .0158602
          x2 |  -.5410846   .1579697    -3.43   0.003    -.8717189   -.2104503
          x3 |   .1740546   .0625307     2.78   0.012     .0431764    .3049328
       _cons |    38.6472     3.6496    10.59   0.000      31.0085     46.2859
------------------------------------------------------------------------------
```

To test whether $\beta_2 + \beta_3 = 1$

```
. test x2+x3=1

 ( 1)  x2 + x3 = 1

       F(  1,    19) =  132.06
            Prob > F =    0.0000
```

Chow Test

```
. g dummy=(obs>1981)

. g d_inc=inc*dummy

. reg sav inc

      Source |       SS       df       MS              Number of obs =      26
-------------+------------------------------           F(  1,    24) =   79.10
       Model |  76621.7867     1  76621.7867           Prob > F      =  0.0000
    Residual |     23248.3    24  968.679166           R-squared     =  0.7672
-------------+------------------------------           Adj R-squared =  0.7575
       Total |  99870.0867    25  3994.80347           Root MSE      =  31.124

------------------------------------------------------------------------------
         sav |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         inc |   .0376791   .0042366     8.89   0.000     .0289353     .046423
       _cons |   62.42267   12.76075     4.89   0.000     36.08578    88.75957
------------------------------------------------------------------------------

. reg sav inc if obs<=1981

      Source |       SS       df       MS              Number of obs =      12
-------------+------------------------------           F(  1,    10) =   92.19
       Model |  16456.2587     1  16456.2587           Prob > F      =  0.0000
    Residual |  1785.03254    10  178.503254           R-squared     =  0.9021
-------------+------------------------------           Adj R-squared =  0.8924
       Total |  18241.2912    11   1658.2992           Root MSE      =  13.361

------------------------------------------------------------------------------
         sav |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         inc |   .0803319   .0083665     9.60   0.000     .0616901    .0989737
       _cons |   1.016115   11.63771     0.09   0.932    -24.91432    26.94655
------------------------------------------------------------------------------

. reg sav inc if obs>1981

      Source |       SS       df       MS              Number of obs =      14
-------------+------------------------------           F(  1,    12) =    3.14
       Model |  2614.39647     1  2614.39647           Prob > F      =  0.1020
    Residual |  10005.2214    12  833.768451           R-squared     =  0.2072
-------------+------------------------------           Adj R-squared =  0.1411
       Total |  12619.6179    13  970.739837           Root MSE      =  28.875

------------------------------------------------------------------------------
         sav |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         inc |   .0148624   .0083932     1.77   0.102    -.0034248    .0331496
       _cons |   153.4947   32.71227     4.69   0.001     82.22075    224.7686
------------------------------------------------------------------------------

. reg sav inc dum d_inc

      Source |       SS       df       MS              Number of obs =      26
-------------+------------------------------           F(  3,    22) =   54.78
       Model |  88079.8327     3  29359.9442           Prob > F      =  0.0000
    Residual |  11790.2539    22  535.920634           R-squared     =  0.8819
-------------+------------------------------           Adj R-squared =  0.8658
       Total |  99870.0867    25  3994.80347           Root MSE      =   23.15

------------------------------------------------------------------------------
         sav |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         inc |   .0803319   .0144968     5.54   0.000     .0502673    .1103964
       dummy |   152.4786   33.08237     4.61   0.000     83.86992    221.0872
       d_inc |  -.0654694   .0159824    -4.10   0.000     -.098615   -.0323239
       _cons |   1.016115   20.16483     0.05   0.960    -40.80319    42.83542
------------------------------------------------------------------------------

. test (dummy=0) (d_inc=0)

 ( 1)  dummy = 0
 ( 2)  d_inc = 0

       F(  2,    22) =   10.69
            Prob > F =    0.0006
```

Dummy Variable Regression Models

ANOVA Model

In the study of seasonal effect (Monday and Friday effects), ANOVA model is estimated using both models with and without intercept term:

Model with Intercept term: $Y_t = \beta_1 + \beta_2 D_{2t} + \beta_3 D_{3t} + \beta_4 D_{4t} + \beta_5 D_{5t} + u_t$

Model without Intercept term: $Y_t = \beta_1 D_{1t} + \beta_2 D_{2t} + \beta_3 D_{3t} + \beta_4 D_{4t} + \beta_5 D_{5t} + u_t$

where: Y_t = Return on asset at time t.
D_{1t} = 1 on Monday and = 0 otherwise.
D_{2t} = 1 on Tuesday and = 0 otherwise.
D_{3t} = 1 on Wednesday and = 0 otherwise.
D_{4t} = 1 on Thursday and = 0 otherwise.
D_{5t} = 1 on Friday and = 0 otherwise.

```
. reg y d2 d3 d4 d5

      Source |       SS       df       MS              Number of obs =     995
-------------+------------------------------           F(  4,   990) =    7.60
       Model |  .00681682     4  .001704205            Prob > F      =  0.0000
    Residual | .221894717   990  .000224136            R-squared     =  0.0298
-------------+------------------------------           Adj R-squared =  0.0259
       Total | .228711537   994  .000230092            Root MSE      =  .01497

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          d2 |   .0044637   .0015009     2.97   0.003     .0015184    .0074089
          d3 |   .0044351   .0015009     2.96   0.003     .0014898    .0073803
          d4 |   .0040519   .0015009     2.70   0.007     .0011067    .0069972
          d5 |   .0082592   .0015009     5.50   0.000      .005314    .0112045
       _cons |  -.0040211   .0010613    -3.79   0.000    -.0061037   -.0019385
------------------------------------------------------------------------------

. reg y d1 d2 d3 d4 d5, noconstant

      Source |       SS       df       MS              Number of obs =     995
-------------+------------------------------           F(  5,   990) =    6.13
       Model | .006865358     5  .001373072            Prob > F      =  0.0000
    Residual | .221894717   990  .000224136            R-squared     =  0.0300
-------------+------------------------------           Adj R-squared =  0.0251
       Total | .228760075   995   .00022991            Root MSE      =  .01497

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          d1 |  -.0040211   .0010613    -3.79   0.000    -.0061037   -.0019385
          d2 |   .0004426   .0010613     0.42   0.677    -.0016401    .0025252
          d3 |    .000414   .0010613     0.39   0.697    -.0016687    .0024966
          d4 |   .0000308   .0010613     0.03   0.977    -.0020518    .0021134
          d5 |   .0042381   .0010613     3.99   0.000     .0021555    .0063207
------------------------------------------------------------------------------
```

Interpretations of the estimated results of the two models are as follows:

Model	With intercept term	Without intercept
Monday	$\hat{Y}_t = \hat{\beta}_1$ = -0.004021	$\hat{Y}_t = \hat{\beta}_1$ = -0.004021
Tuesday	$\hat{Y}_t = \hat{\beta}_1 + \hat{\beta}_2$ = -0.004021+0.004464 = 0.000443	$\hat{Y}_t = \hat{\beta}_2$ = 0.000443
Wednesday	$\hat{Y}_t = \hat{\beta}_1 + \hat{\beta}_3$ = -0.004021+0.004435 = 0.000414	$\hat{Y}_t = \hat{\beta}_3$ = 0.000414
Thursday	$\hat{Y}_t = \hat{\beta}_1 + \hat{\beta}_4$ = -0.004021+0.004052 = 0.000031	$\hat{Y}_t = \hat{\beta}_4$ = 0.000031
Friday	$\hat{Y}_t = \hat{\beta}_1 + \hat{\beta}_5$ = -0.004021+0.008259 = 0.004238	$\hat{Y}_t = \hat{\beta}_5$ = 0.004238

For the purpose of testing only Monday and Friday effect, the model can be stated as:

Model to test Monday and Friday effect: $Y_t = \beta_0 + \beta_1 D_{1t} + \beta_5 D_{5t} + u_t$

```
. reg y d1 d5

      Source |       SS       df       MS              Number of obs =     995
-------------+------------------------------           F(  2,   992) =   15.19
       Model |  .006795783     2  .003397892           Prob > F      =  0.0000
    Residual |  .221915754   992  .000223705           R-squared     =  0.0297
-------------+------------------------------           Adj R-squared =  0.0278
       Total |  .228711537   994  .000230092           Root MSE      =  .01496

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          d1 |  -.0043169   .0012243    -3.53   0.000    -.0067194   -.0019144
          d5 |   .0039423   .0012243     3.22   0.001     .0015399    .0063448
       _cons |   .0002958   .0006121     0.48   0.629    -.0009055     .001497
------------------------------------------------------------------------------
```

Interpretation of the estimated result of the model is as follow:

Model	
Tuesday **Wednesday** **Thursday**	$\hat{Y}_t = \hat{\beta}_0$ = 0.000296
Monday	$\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1$ = 0.000296 - 0.004317 = -0.004021
Friday	$\hat{Y}_t = \hat{\beta}_1 + \hat{\beta}_5$ = 0.000296 - 0.003942 = 0.004238

ANCOVA Model

If the model also includes quantitative independent variable *(X_t)*, the model can be stated as:

Model with intercept dummy: $Y_t = \beta_0 + \beta_1 X_{1t} + \gamma_0 D_{1t} + \lambda_0 D_{5t} + u_t$

This model can be interpreted as:

Model for Tuesday, Wednesday and Friday: $Y_t = \beta_0 + \beta_1 X_{1t} + u_t$
Model for Monday: $Y_t = (\beta_0 + \gamma_0) + \beta_1 X_{1t} + u_t$
Model for Friday: $Y_t = (\beta_0 + \lambda_0) + \beta_1 X_{1t} + u_t$

```
. reg y x1 d1 d5

      Source |       SS       df       MS              Number of obs =     995
-------------+------------------------------           F(  3,   991) =   11.68
       Model |  .007808227     3  .002602742           Prob > F      =  0.0000
    Residual |  .220903311   991  .000222909           R-squared     =  0.0341
-------------+------------------------------           Adj R-squared =  0.0312
       Total |  .228711537   994  .000230092           Root MSE      =  .01493

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   1.90e-06   8.91e-07     2.13   0.033     1.50e-07    3.65e-06
          d1 |  -.0043131   .0012221    -3.53   0.000    -.0067113   -.0019149
          d5 |   .0039385   .0012221     3.22   0.001     .0015403    .0063368
       _cons |  -.0708331   .0333809    -2.12   0.034    -.1363383   -.0053278
------------------------------------------------------------------------------
```

Model with intercept and slope dummy variables:

$$Y_t = \beta_0 + \gamma_0 D_{1t} + \lambda_0 D_{5t} + \beta_1 X_{1t} + \gamma_1 D_{1t} X_{1t} + \lambda_1 D_{5t} X_{1t} + u_t$$

This model can be interpreted as:

Model for Tuesday, Wednesday and Friday: $Y_t = \beta_0 + \beta_1 X_{1t} + u_t$
Model for Monday: $Y_t = (\beta_0 + \gamma_0) + (\beta_1 + \gamma_1) X_{1t} + u_t$
Model for Friday: $Y_t = (\beta_0 + \lambda_0) + (\beta_1 + \lambda_1) X_{1t} + u_t$

```
. g x1d1 = x1*d1

. g x1d5 = x1*d5

. reg y d1 d5 x1 x1d1 x1d5

      Source |       SS       df       MS              Number of obs =     995
-------------+------------------------------           F(  5,   989) =    7.19
       Model |  .008022894     5  .001604579           Prob > F      =  0.0000
    Residual |  .220688644   989  .000223143           R-squared     =  0.0351
-------------+------------------------------           Adj R-squared =  0.0302
       Total |  .228711537   994  .000230092           Root MSE      =  .01494

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          d1 |  -.0541205    .086225    -0.63   0.530    -.2233254    .1150844
          d5 |   .0576593   .0862319     0.67   0.504    -.1115592    .2268777
          x1 |   1.92e-06   1.15e-06     1.67   0.096    -3.39e-07    4.18e-06
        x1d1 |   1.33e-06   2.30e-06     0.58   0.564    -3.19e-06    5.85e-06
        x1d5 |  -1.43e-06   2.30e-06    -0.62   0.533    -5.95e-06    3.08e-06
       _cons |  -.0716146   .0431142    -1.66   0.097    -.1562204    .0129911
------------------------------------------------------------------------------
```

Chow Test vs Dummy Variable Technique

If we would like to check whether the model should be separated into two sub-periods (2000-2001 and 2002-2004), the models can be stated as:

Model 2000-2001: $Y_t = \beta_0 + \beta_1 X_{1t} + u_{1t}$

Model 2002-2004: $Y_t = \gamma_0 + \gamma_1 X_{1t} + u_{2t}$

Chow Test: H_0: $\beta_0 = \gamma_0$ and $\beta_1 = \gamma_1$

Dummy Variable Regression Model

Model with intercept and slope dummy:

$$Y_t = \beta_0 + \delta_0 D_{YEARt} + \beta_1 X_{1t} + \delta_1 D_{YEARt} X_{1t} + u_t$$

where: $D_{YEARt} = 1$ for 2002-2004 and $= 0$ for 2000-2001.

The model can be interpreted as:

Model 2000-2001: $Y_t = \beta_0 + \beta_1 X_{1t} + u_{1t}$

Model 2002-2004: $Y_t = (\beta_0 + \delta_0) + (\beta_1 + \delta_1) X_{1t} + u_{2t}$

H_{01}: $\delta_0 = 0$ and $\delta_1 = 0$

```
. g x1dyear = x1*dyear

. reg y dyear x1 x1dyear

      Source |       SS       df       MS              Number of obs =     995
-------------+------------------------------           F(  3,   991) =    2.69
       Model |  .001849335     3  .000616445           Prob > F      =  0.0450
    Residual |  .226862202   991  .000228923           R-squared     =  0.0081
-------------+------------------------------           Adj R-squared =  0.0051
       Total |  .228711537   994  .000230092           Root MSE      =  .01513

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       dyear |   .2727136   .1521381     1.79   0.073    -.0258363    .5712635
          x1 |   8.49e-06   3.60e-06     2.36   0.019     1.42e-06    .0000156
     x1dyear |  -7.38e-06   4.10e-06    -1.80   0.072    -.0000154    6.66e-07
       _cons |  -.3139704   .1329766    -2.36   0.018    -.5749185   -.0530223
------------------------------------------------------------------------------

. test dyear x1dyear

 ( 1)  dyear = 0
 ( 2)  x1dyear = 0

       F(  2,   991) =    1.80
            Prob > F =    0.1657
```

Distinction between the two methods is that while Chow test is to test whether all estimators are the same or not, dummy variables technique is to separately test equality of each estimator for different sub-period -- some estimators may be different while some may not.

Interaction Effect

Model with interaction effect dummy:

$$Y_t = \beta_0 + \gamma_0 D_{1t} + \lambda_0 D_{JANt} + \delta_0 D_{1t} D_{JANt} + \beta_1 X_{1t} + u_t$$

where: D_{1t} = 1 for Monday and = 0 otherwise.
D_{JANt} = 1 for January and = 0 otherwise.

The meaning of $D_{1t}D_{JANt}$

$D_{1t}D_{JANt}$ = 1 for Monday in January and = 0 otherwise.

```
. g d1djan = d1*djan

. reg y d1 djan d1djan x1

      Source |       SS       df       MS              Number of obs =     995
-------------+------------------------------           F(  4,   990) =    7.55
       Model |  .006773371     4  .001693343           Prob > F      =  0.0000
    Residual |  .221938166   990   .00022418           R-squared     =  0.0296
-------------+------------------------------           Adj R-squared =  0.0257
       Total |  .228711537   994  .000230092           Root MSE      =  .01497

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          d1 |  -.0059289   .0012667    -4.68   0.000    -.0084147   -.0034432
        djan |   .0019612   .0017065     1.15   0.251    -.0013876    .0053099
      d1djan |    .004773   .0036262     1.32   0.188    -.0023429    .0118889
          x1 |   2.08e-06   8.98e-07     2.32   0.021     3.21e-07    3.85e-06
       _cons |  -.0769402    .033645    -2.29   0.022    -.1429639   -.0109166
------------------------------------------------------------------------------
```