Simple linear regression in EViews

Econometrics for Financial Markets

Learning goals:

- Understand how to run a simple regression in EViews.
- Use EViews to interpret regression output and test basic hypotheses
- Know how to calculate hedging ratios
- Understand how to calculate asset alphas and betas and how you can test whether a portfolio has a significant abnormal return.

Literature

- Brooks, Chapter 3. A brief overview of classical linear regression

Background Reading

- Brealey, Myers and Allen, *Principles of Corporate Finance*
  Chapter 8. Portfolio Theory and Capital Asset Pricing Model (Ch8.2-8.4)
- Eviews Tutorials, http://www.eviews.com/Learning/index.html

In this eviews class we are going to use simple regressions to (i) calculate hedging ratios and (ii) calculate the beta of a stock using the Capital Asset Pricing Model (CAPM).

Data Set 1: Calculating an optimal hedge ratio.

Task 1 Start Eviews and Import Data

a) **Import** SandPhedge.xls; This file can be found on the Moodle
b) **Generate** a new series that represents the log returns for the spot and futures

$$r = \ln(P_t) - \ln(P_{t-1}) = \Delta \ln(P_t)$$

Task 2 Descriptive Analysis of Data

a) Create a **group** containing the Spot and Futures return data
b) check the **descriptive statistics** of the Spot and Futures return data
c) create line and scatter **graphs** containing the Spot and Futures return data

Task 3 Calculating the Hedge Ratio

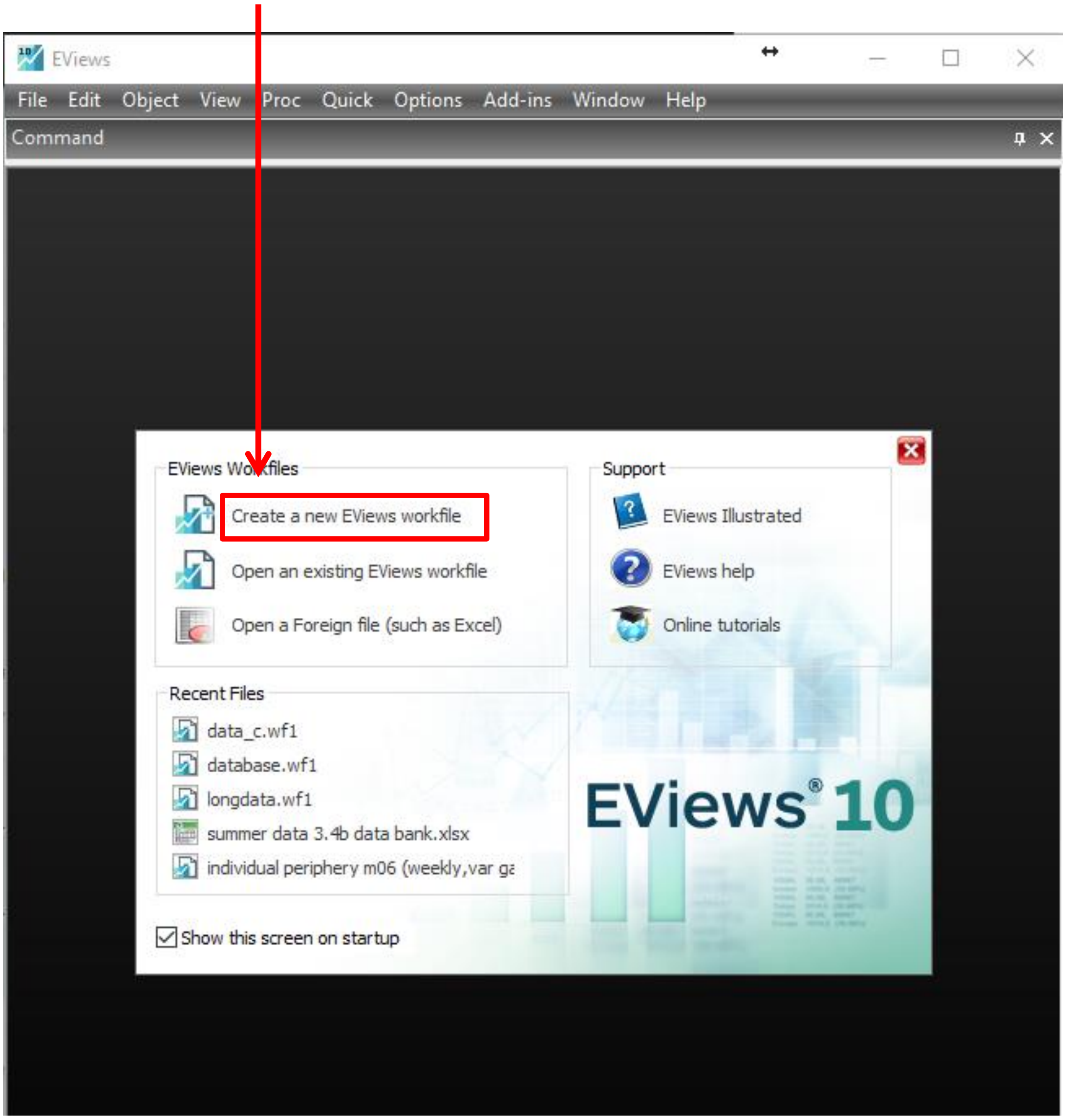
a) **Estimate** the following regression: $r_s = \alpha + \beta r_f + \epsilon$ and save the regression output, where r_s is the spot return and r_f is the futures return.
b) **Test** the null hypothesis that $\alpha = 0$ and $\beta = 0$ for both the returns regression
c) **Test** the null hypothesis that $\beta = 1$ for the returns regression
d) **Interpret** your results

Solutions 1: Calculating an optimal hedge ratio.

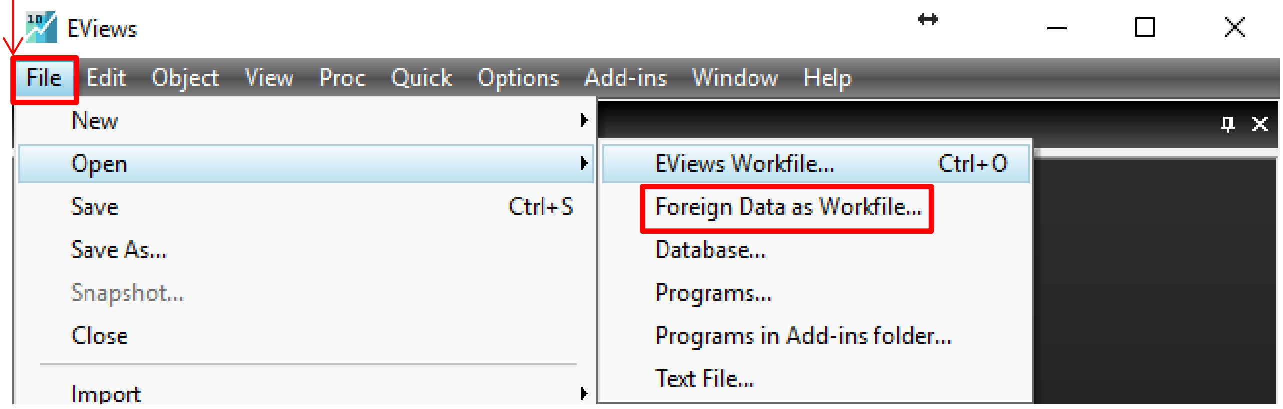
Task 1 Start Eviews and Import Data

a) Import SandPhedge.xls

To import an excel file or other format of data into eviews, you can either click `create a new Eviews workfile` on the start up screen of eviews



Or go to
File\Open\Foreign Data as Wokfile...

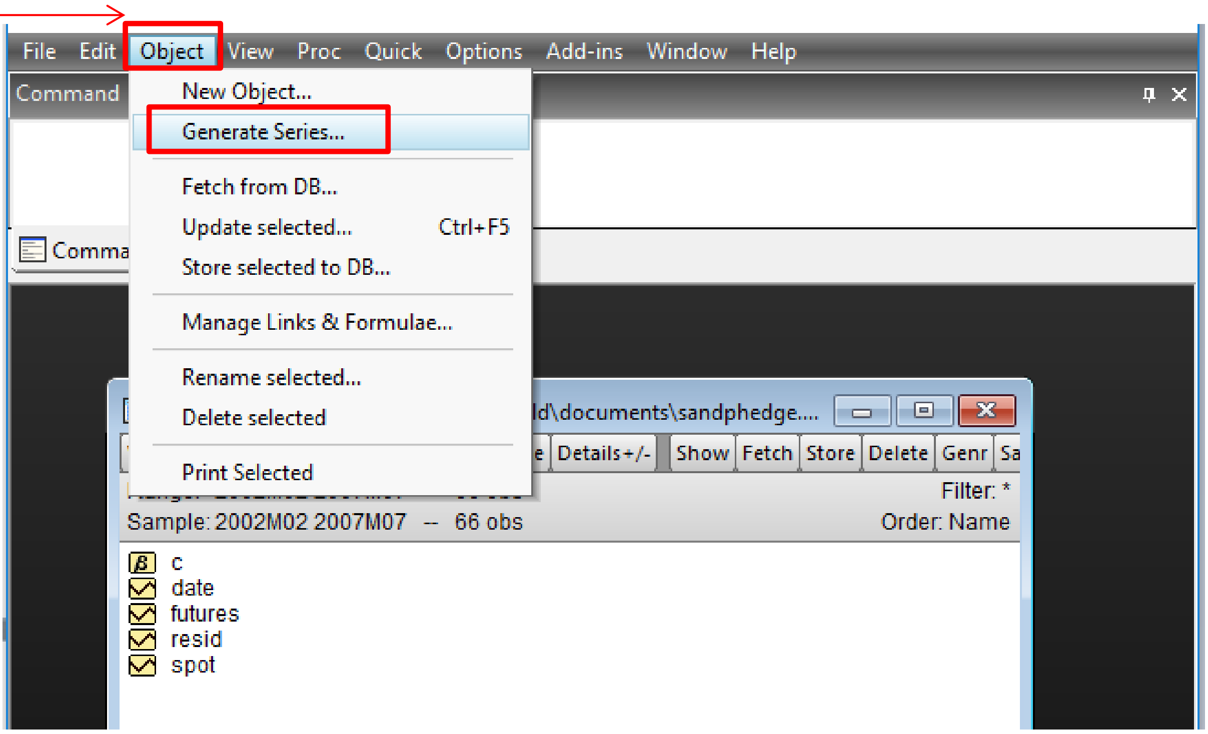


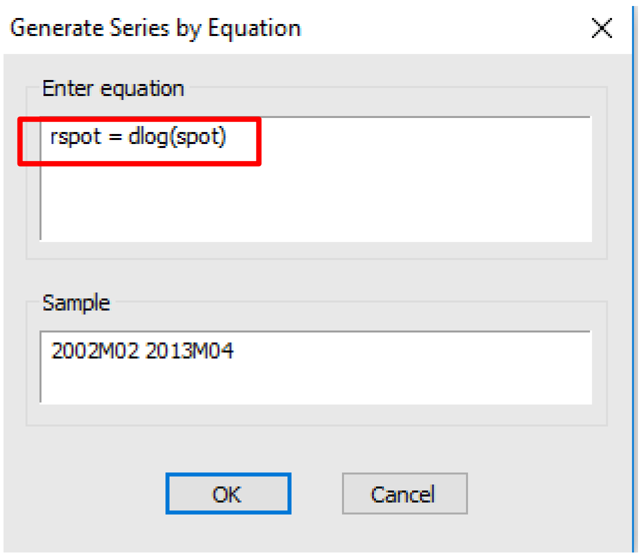
b) Generate new series that represent log returns for the spot and futures

To manipulate your data and create a transformed series from your original data go to

`Object\Generate Series\...` and type the desired equation in the windows that pops up.

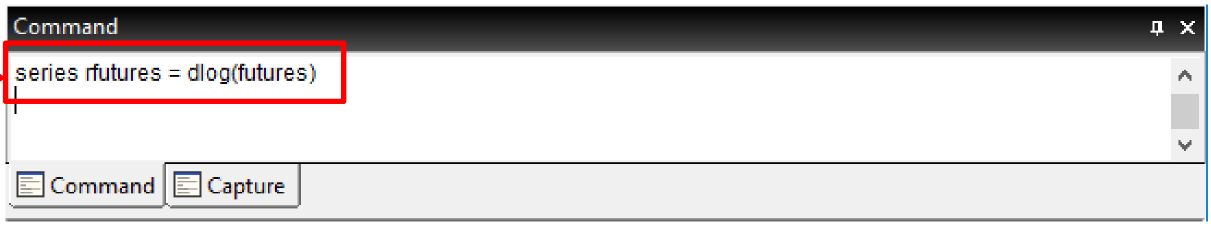
On the LHS of the equation you type the name of the variable you want to create and on the RHS the formula. Two common functions that are used are lnP = log(P) to create the natural logs of a variable and `ret = dlog(P)` which calculates $\Delta \ln P = \ln P_t - \ln P_{t-1}$.





Alternatively, you can write commands straight in the command window. The command series will create a new series which has a name the first input you give and as definition the part to the right of the equal sign. So entering

`Series rfutures = dlog(futures)` will create a new series called `rfutures` which is defined as the log return (log difference) of the series `futures`

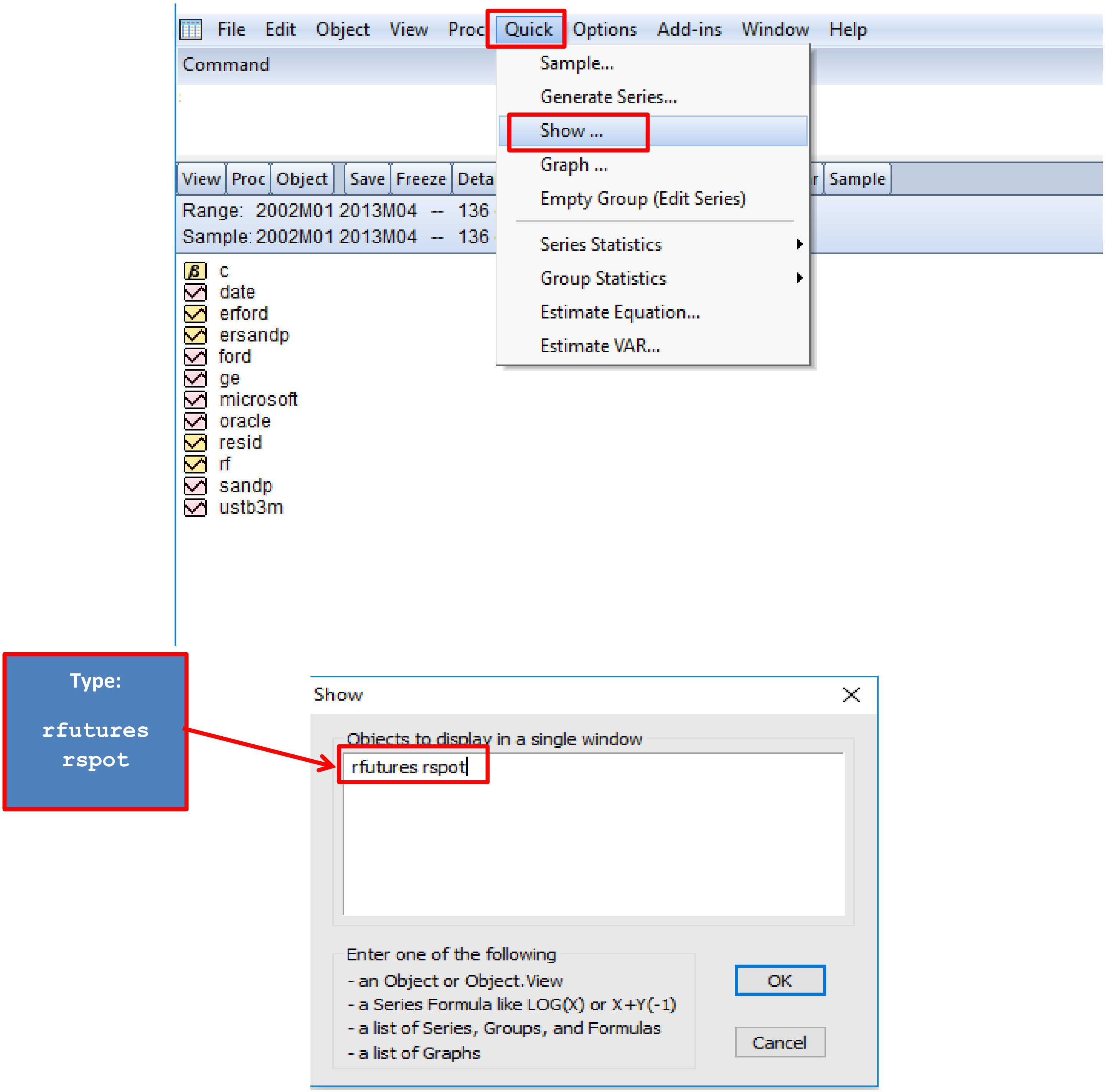


Task 2 Descriptive Analysis of Data

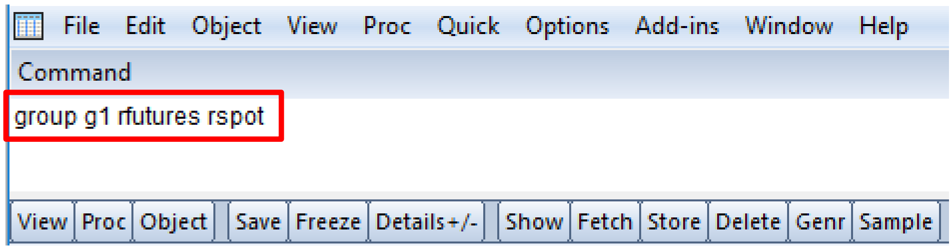
a) Create a **group** containing the Spot and Futures return data

The object `group` is very useful to perform operations on multiple series simultaneously. We will use it here to get summary statistics and graphs for all variables at the same time. Like most things in eviews, there are several ways to create a new group: via the menu or via a command.

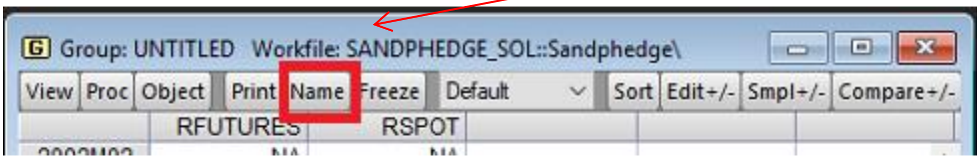
A new group can be created via the menu option **`Quick\Show`**... and type the the series you want to include in the group in the window that pops up.



Alternative you can type **`group g1 rfutures rspot`** in the command line, which will create a group object called `g1` containing the series `rftutures` and `rspot`.

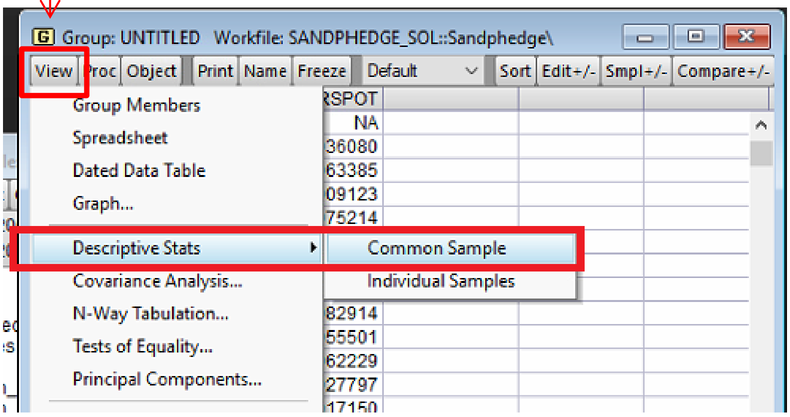


To name your series and make it a permanent object click on the **`name`** button in the group window should be open now.



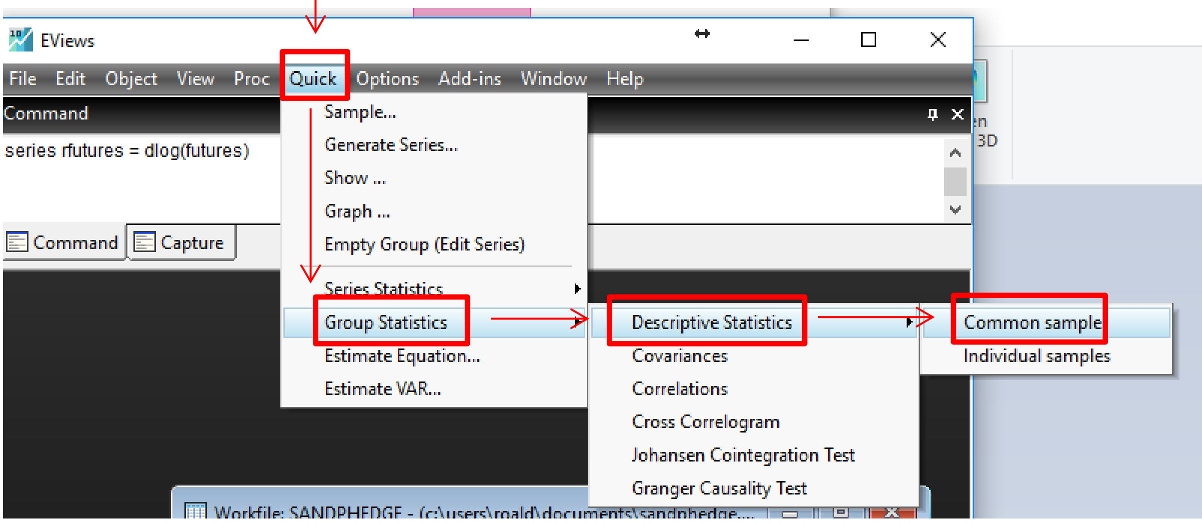
b) check the **descriptive statistics** of the Spot and Futures return data

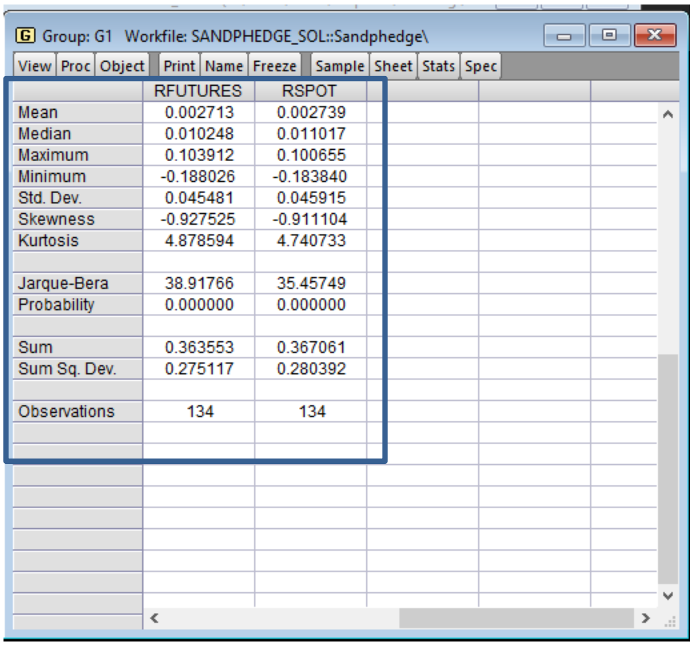
One of the advantages of a group is that you can get summary statistics of all series simulteneous. You can access the summary statistics by opening the group window and clicking on `View\Descriptive Stats\Common Sample`



or through the main ribbon via

`Quick\Group Statistics\Descriptive\Common sample`

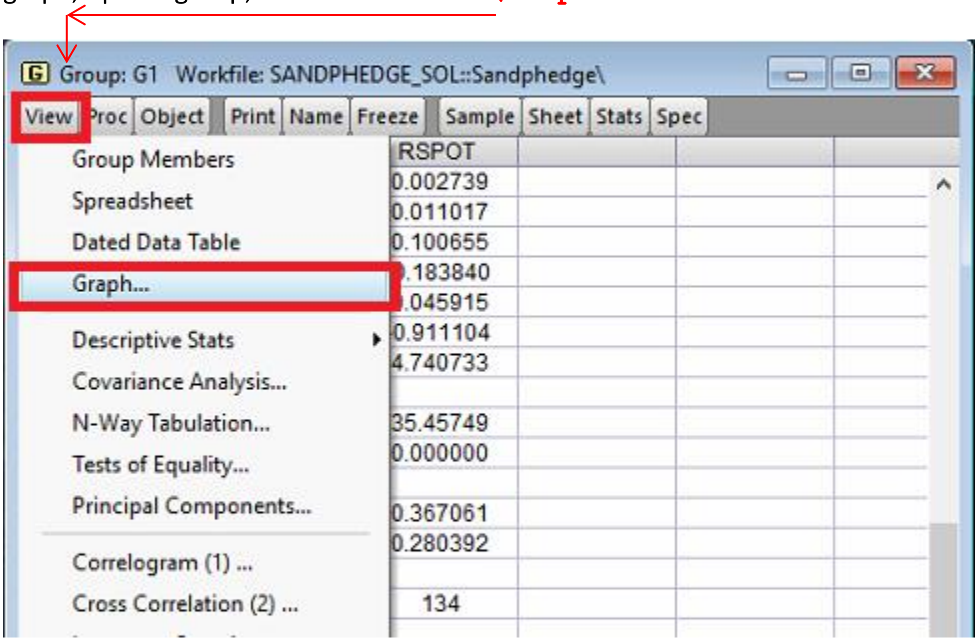




	RFUTURES	RSPOT
Mean	0.002713	0.002739
Median	0.010248	0.011017
Maximum	0.103912	0.100655
Minimum	-0.188026	-0.183840
Std. Dev.	0.045481	0.045915
Skewness	-0.927525	-0.911104
Kurtosis	4.878594	4.740733
Jarque-Bera	38.91766	35.45749
Probability	0.000000	0.000000
Sum	0.363553	0.367061
Sum Sq. Dev.	0.275117	0.280392
Observations	134	134

c) Create line and scatter **Graphs** containing the Spot and Futures return data

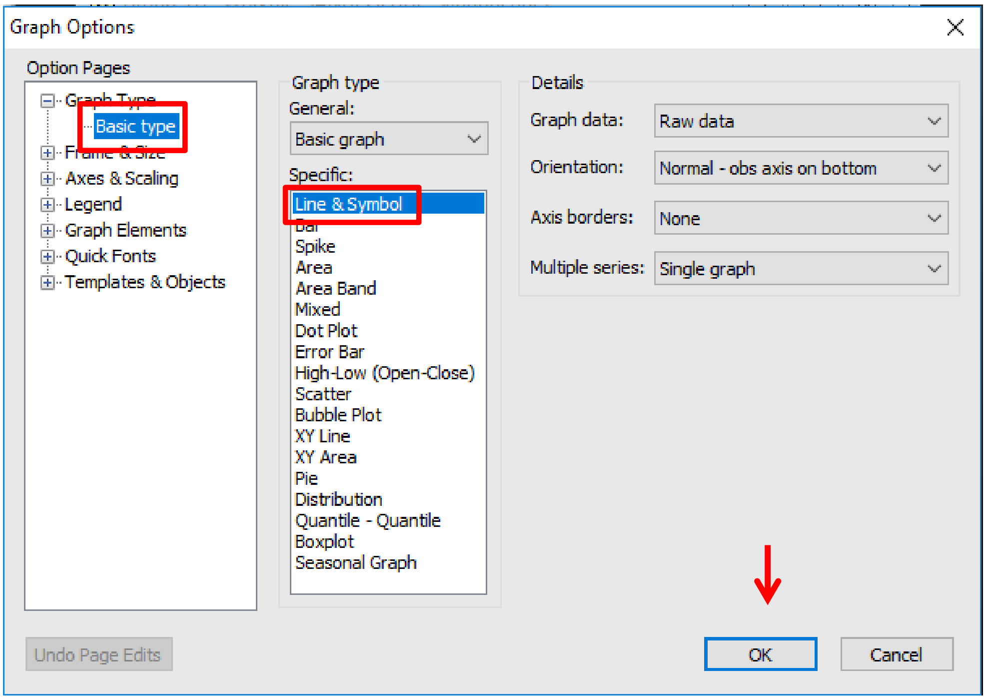
Graphs can be created through objects as well. For a graph of a single series, to create a multiple graph, open a group; then click on **View\Graph**...

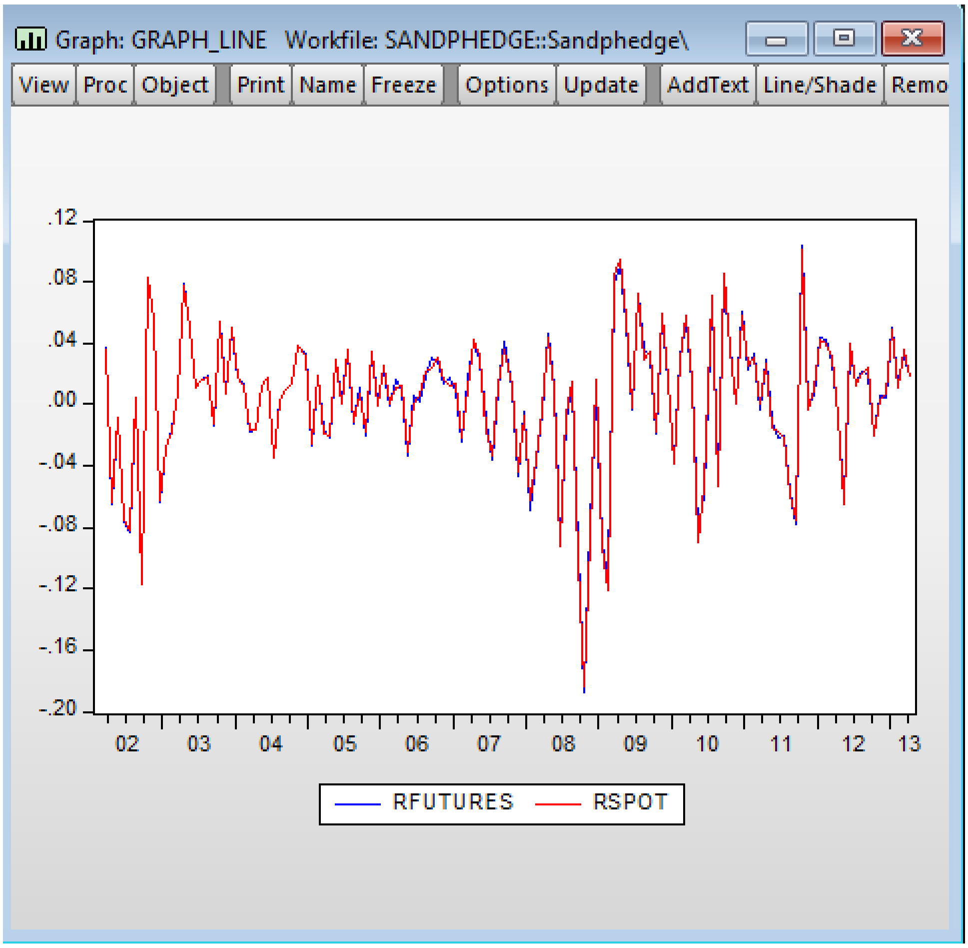


After you click on this the graph a window pops up which a large host of options to create graphs. Most important is the column in the middle which allows you to set the type of graph you want to create. This will be shown below.

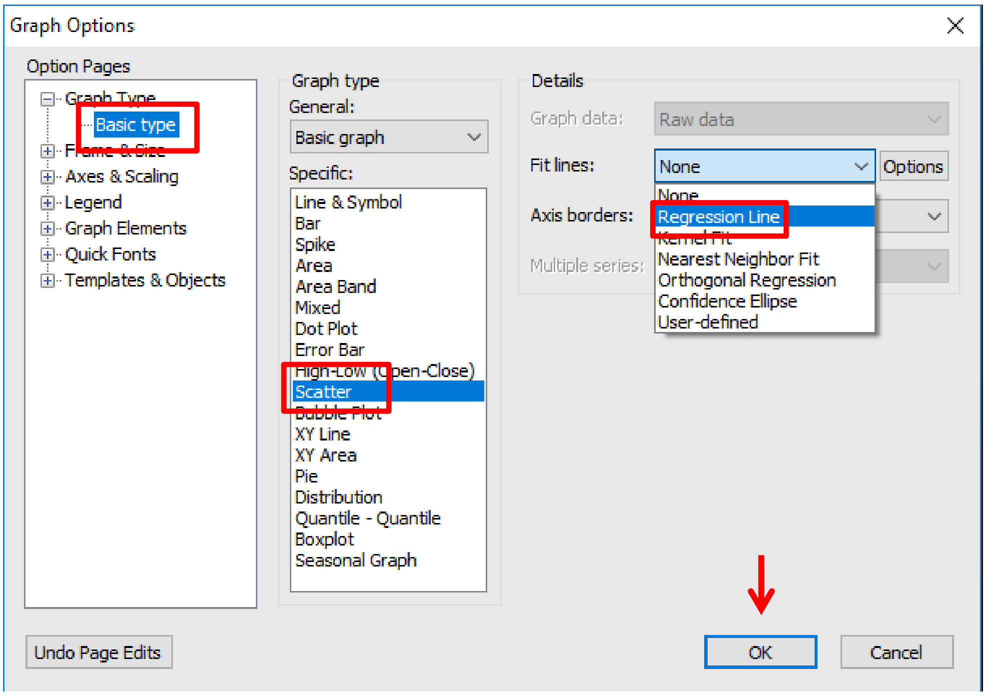
Alternatively you can use the command window to request graphs. Examples of command line commands to create graphs are

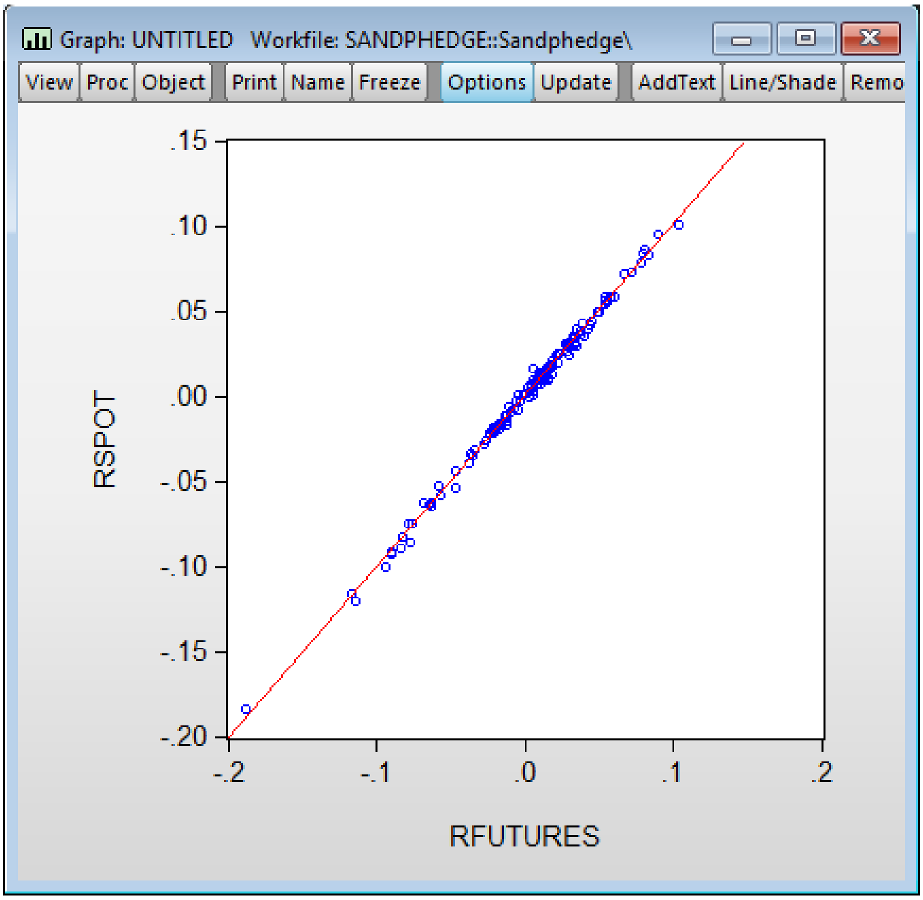
First lets create a line graph, which is selected as default.





To create a scatter select the option scatter. You can add a fit line by changing the options on the RHS (see picture below)





Task 3 Calculating the Hedge Ratio

a) **Estimate** the following regression: $r_s = \alpha + \beta r_f + \epsilon$ and save the regression output, where r_s is the spot return and r_f is the futures return.

To estimate an equation, click on **Quick\Esimate Equation\...**

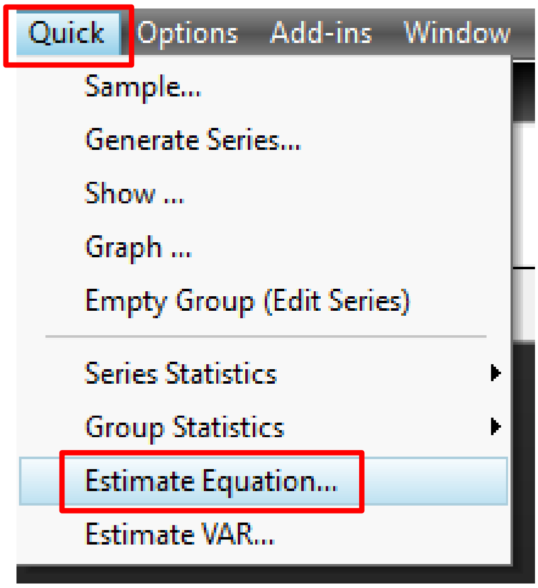
A window will open with all the estimation options. As innocent as it looks, it is probably the most important window in Eviews. The top half of the window contains the equation you are to estimate, and the bottom half allows you to change the estimation method and the sample. For now let's leave the estimation method unchanged. The equation in the top half takes some getting use to. The syntax is as follows: **[dependent variable] [constant] [independent variables]**

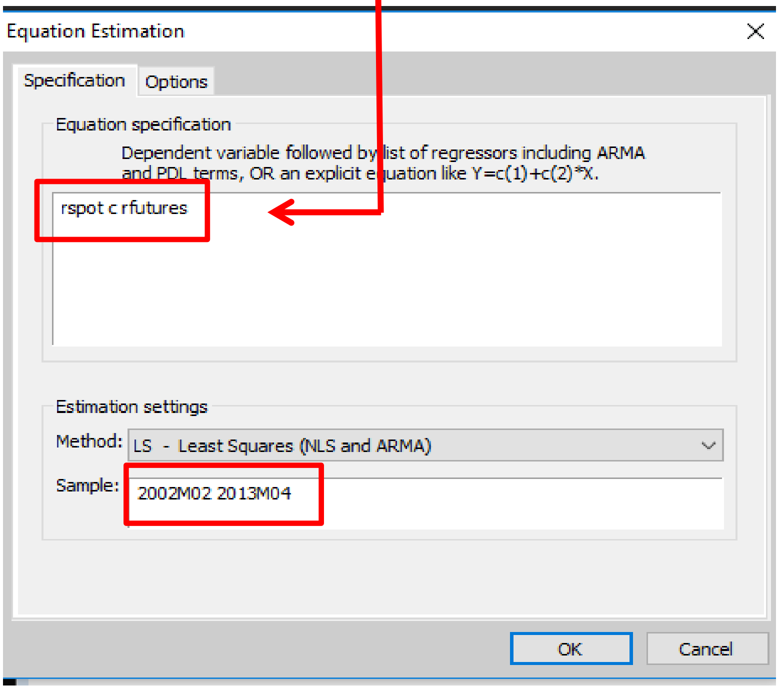
So to estimate:

$$rspot_t = \alpha + b \times rfutures_t + \epsilon_t$$

you type inside the box:

```
rspot c rfutures
```





b) Test the null hypothesis that $\alpha = 0$ and $\beta = 0$ for both the returns regression

The blue box contains all the information you need. The t-Statistic and Prob. Are the tstat and the **p-val** associated with the hypothesis that the associated coefficient is equal to zero.

As you can see, the p-value for the constant is very high (well above the threshold of 0.05) and the p-value for **RFUTURES** is very low (well below the threshold of 0.05). So we conclude that we cannot reject the null that the constant is zero, but we *can* reject the null that beta is zero.

Equation: EQ_HEDGE Workfile: SANDPHEDGE::Sandphedge\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: RSPOT
Method: Least Squares
Date: 11/01/17 Time: 15:33
Sample (adjusted): 2002M03 2013M04
Included observations: 134 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.40E-06	0.000266	0.024032	0.9809
RFUTURES	1.007291	0.005865	171.7341	0.0000

R-squared	0.995544	Mean dependent var	0.002739
Adjusted R-squared	0.995510	S.D. dependent var	0.045915
S.E. of regression	0.003076	Akaike info criterion	-8.715236
Sum squared resid	0.001249	Schwarz criterion	-8.671984
Log likelihood	585.9208	Hannan-Quinn criter.	-8.697660
F-statistic	29492.60	Durbin-Watson stat	2.947218
Prob(F-statistic)	0.000000		

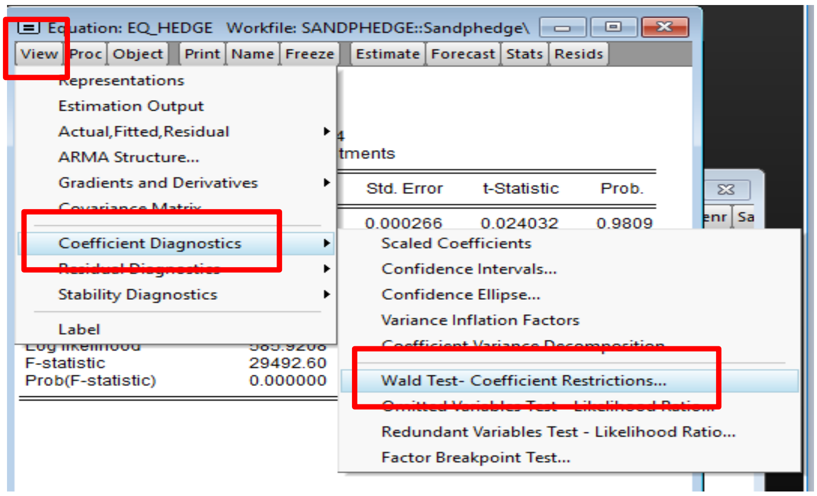
Prob = P-Value

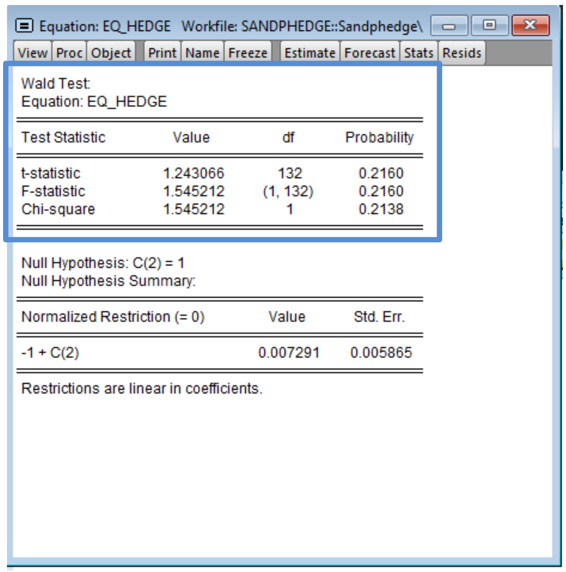
c) **Test** the null hypothesis that $\beta = 1$ for the returns regression

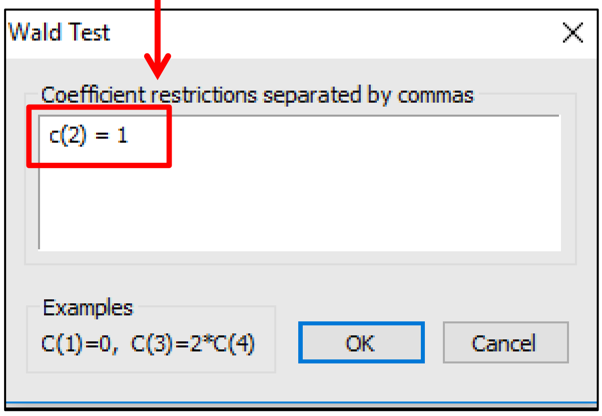
You can test other hypotheses, like for instance $\beta = 1$, via

Views\Coefficient Diagnostics\Wald Test – Coefficient Restrictions…

The way Eviews deals with coefficients takes some getting used to. Eviews numbers all the coefficients, starting at the left and going to the right, so the equation
$r_s = \alpha + \beta r_f$ (error term omitted) will be interpreted by Eviews as
`rspot = c(1) + c(2) rfutures`
So to test whether $\beta = 1$, we will have to write **c(2) = 1** in the corresponding window.







d) **Interpret** your results

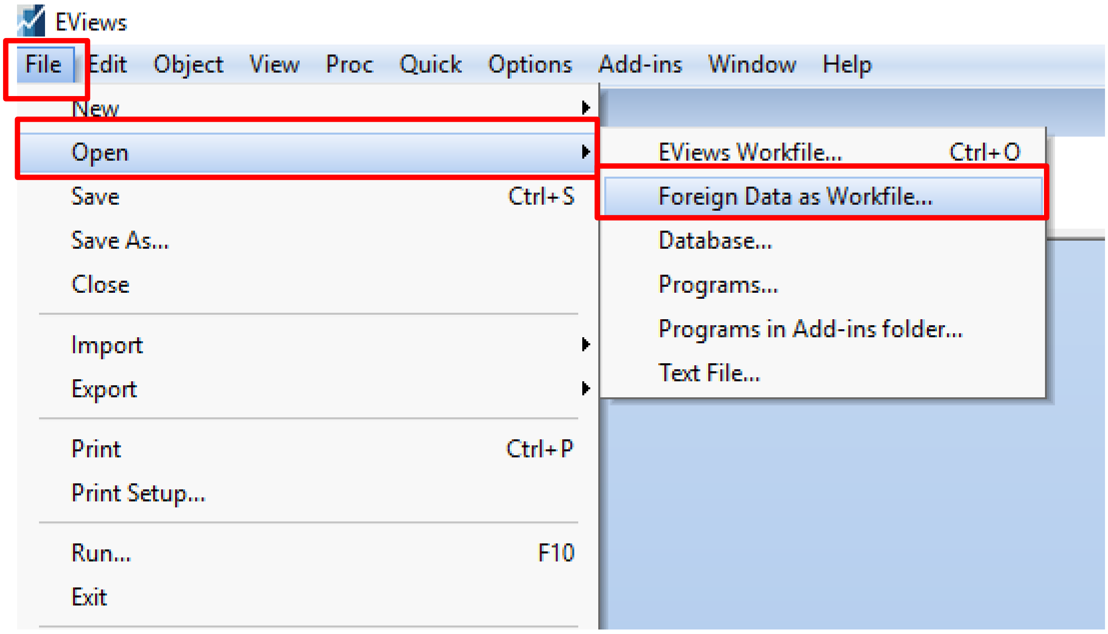
The hedge ratio is significantly different from zero, but not significantly different from one.

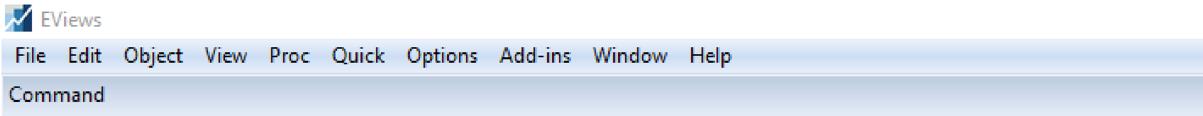
This is suggestive that: a) SandP futures are a good way to hedge spot returns and the optimal hedge ratio is close to 1.

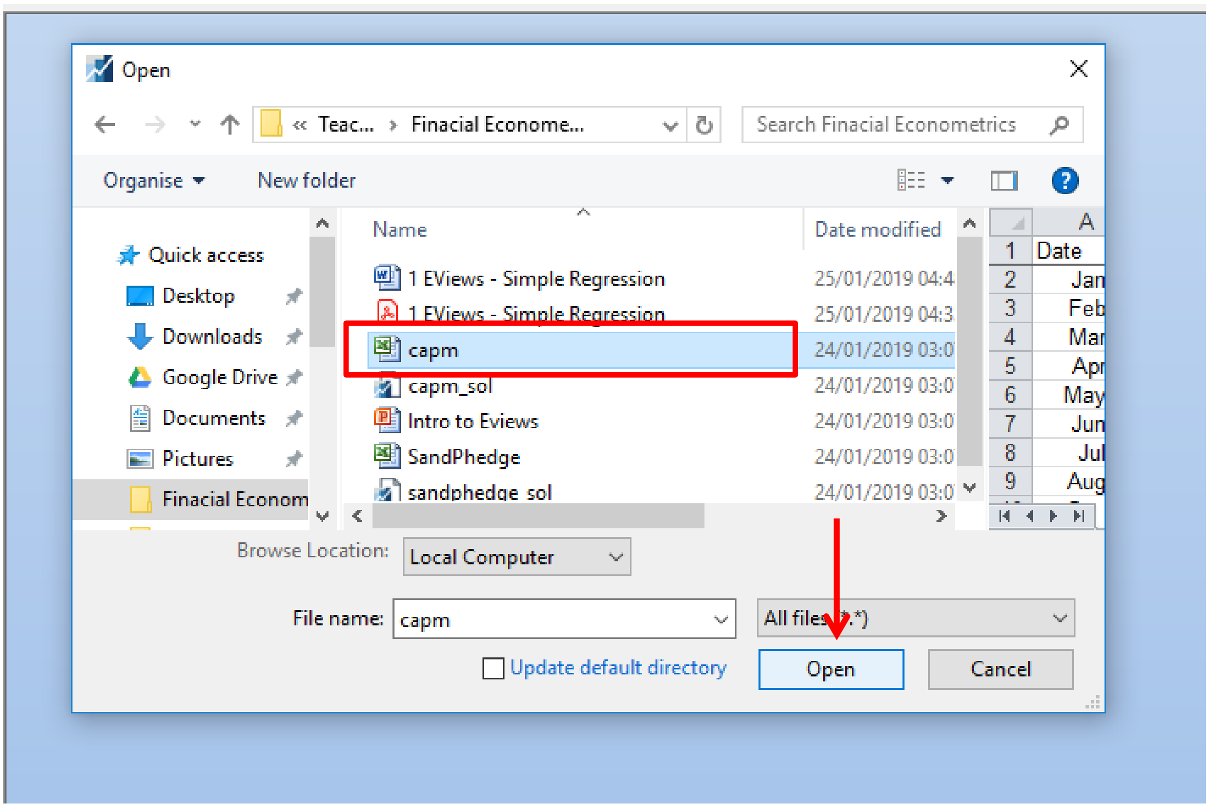
Data Set 2: CAPM

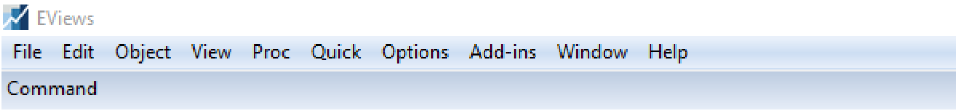
Task 1 Start Eviews and Import Data

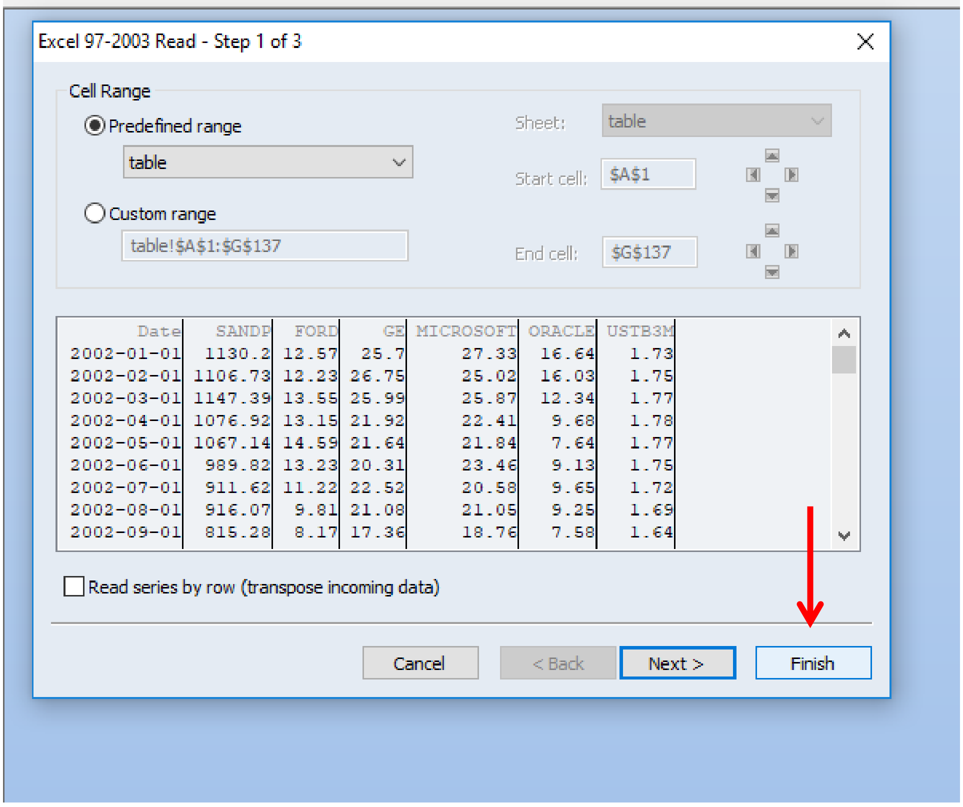
a) **Import** capm.xls



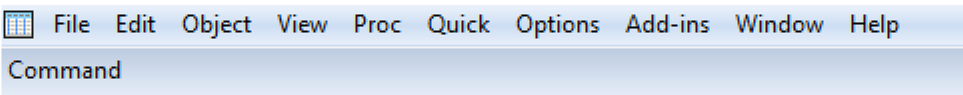


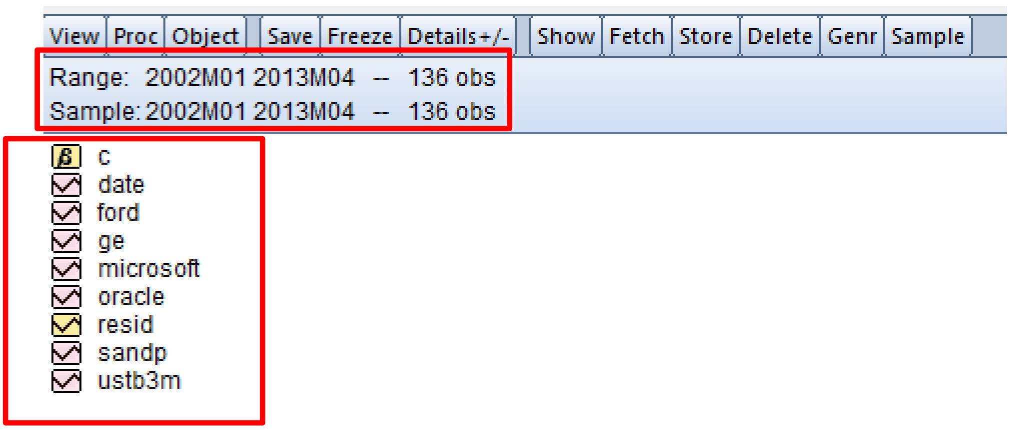






The Workfile will look like the image below:





b) **Generate** new series to calculate the montly risk-free rate and the log excess returns for the S&P 500 and Ford.

Note that you need to be careful in how you transform the interest rates. 1) the tbill data is currently expressed annually, not monthly, and 2) the data is expressed in a way such that a value of 1 is interpreted as 1%. We would like to scale it so that 0.01 is intrepreted as 1%. Thus you will have to divide the data by 12 to get monthly returns and by 100 to go from 1 to 0.01.

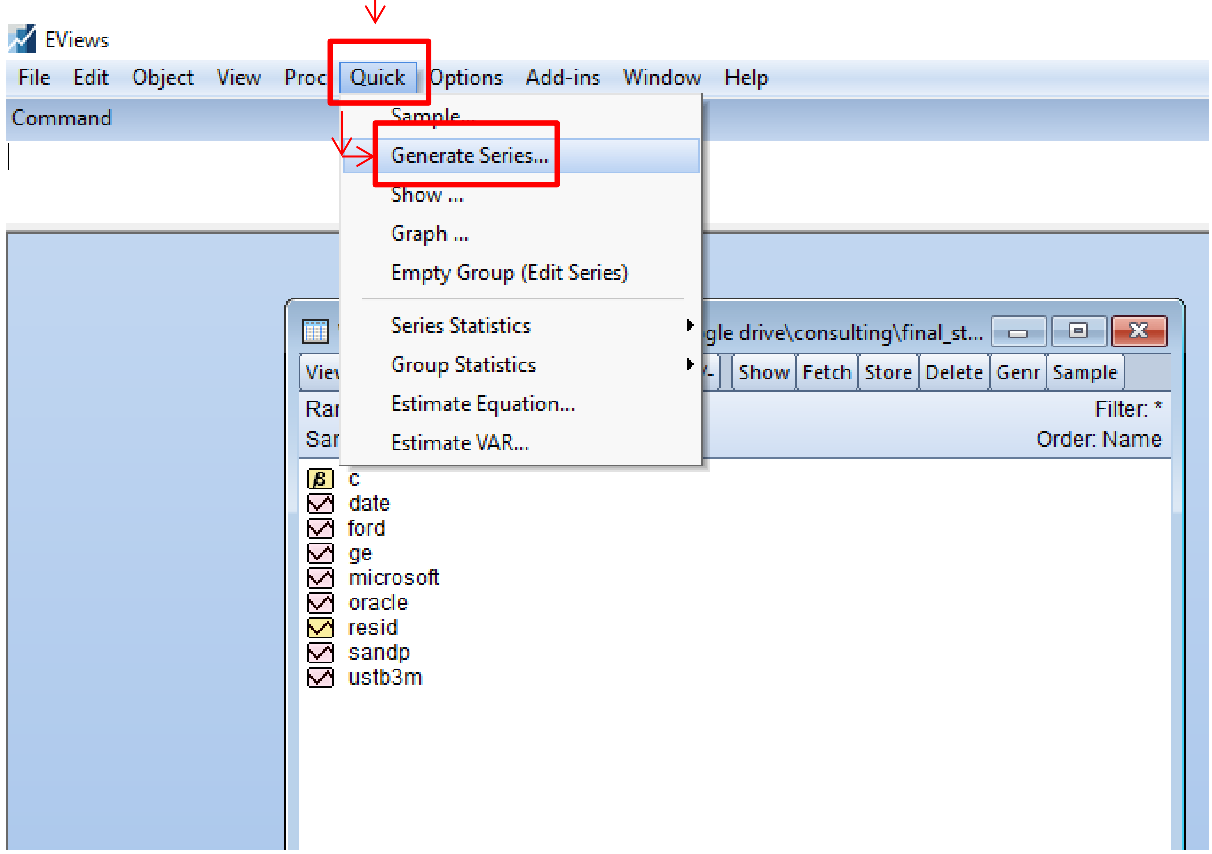
Recall that the formula for log returns of an asset is

$$r = \ln(P_t) - \ln(P_{t-1}) = \Delta \ln(P_t)$$

An excess return is defined as a return less the risk free rate.

$$r_e = r_i - r_f$$

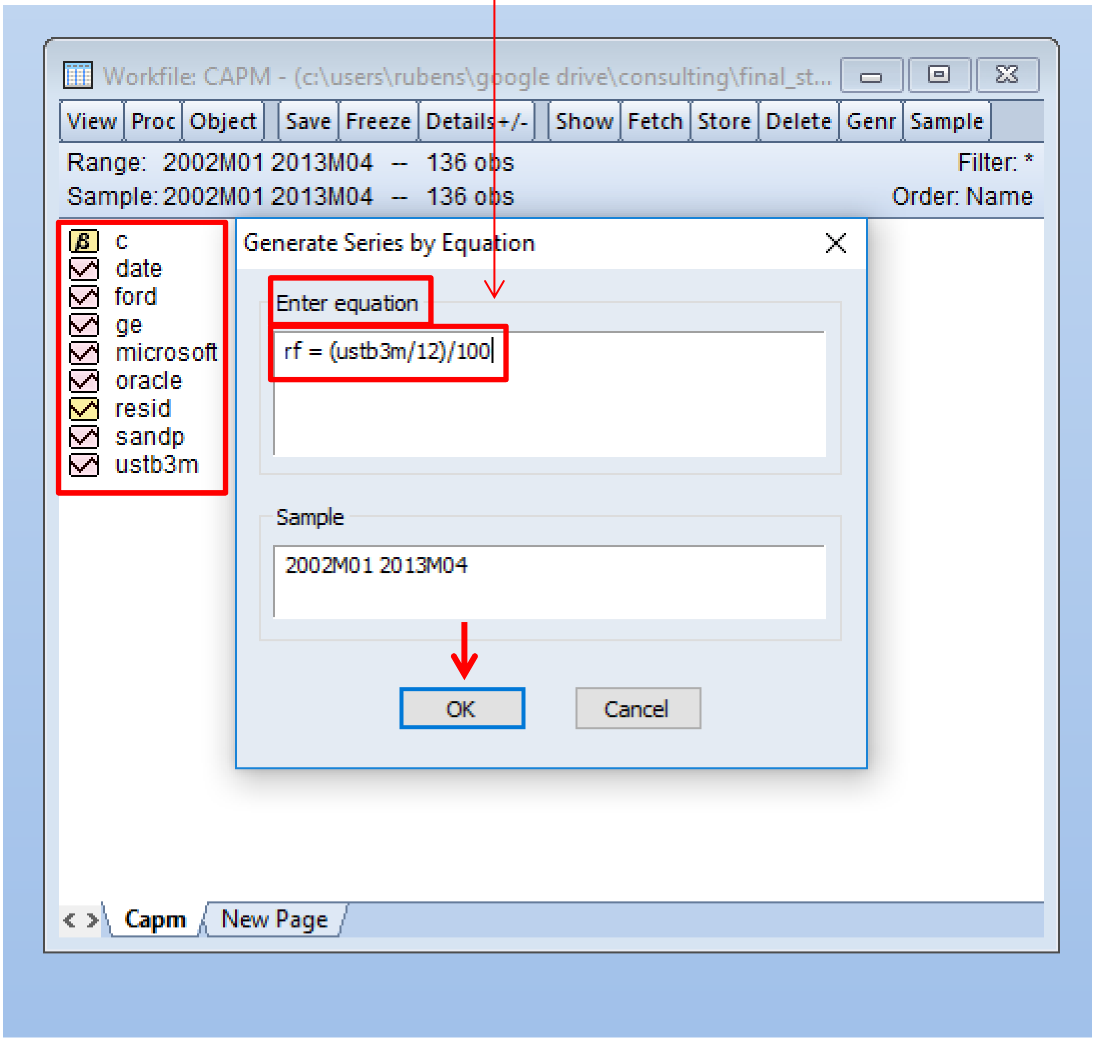
Quick/Generate Series…



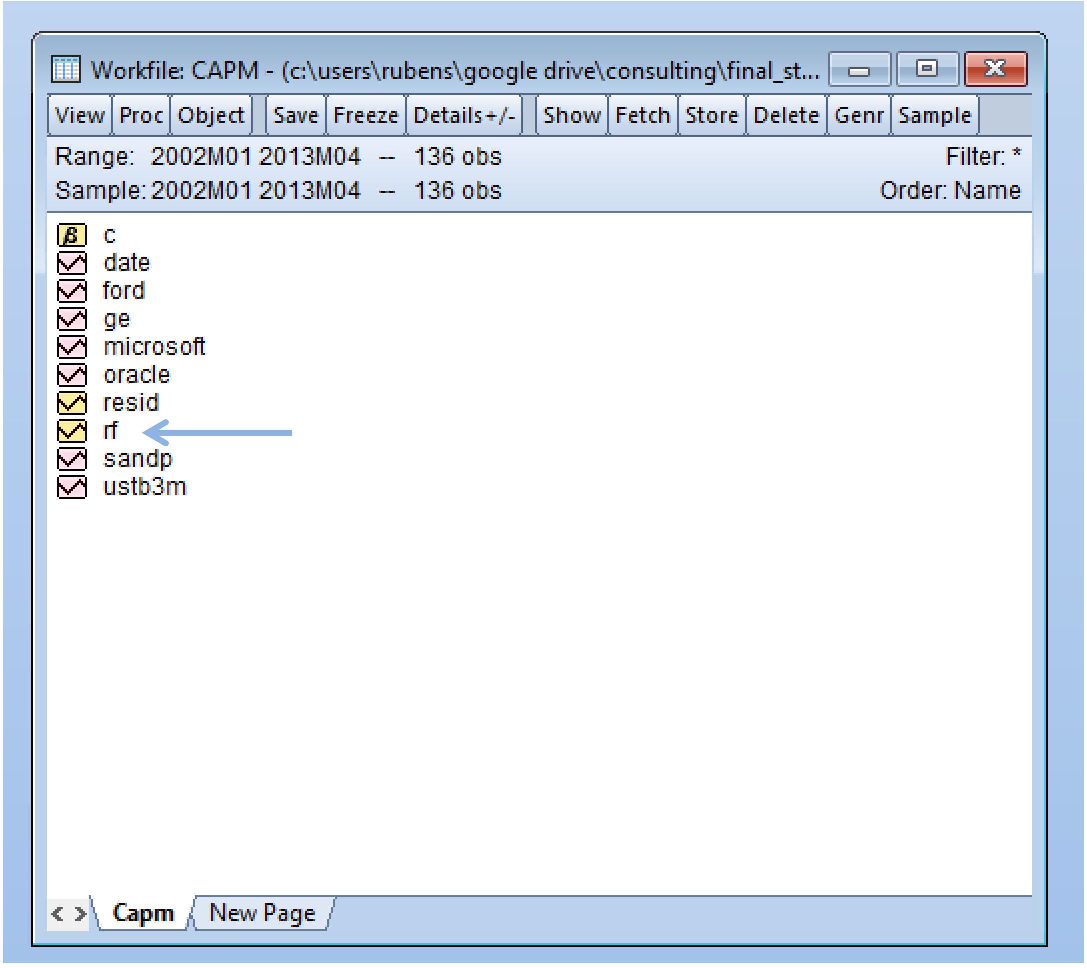
The formulas for each of the variables is given below.

Type the formulas inside the box **one by one** : **Enter equation**

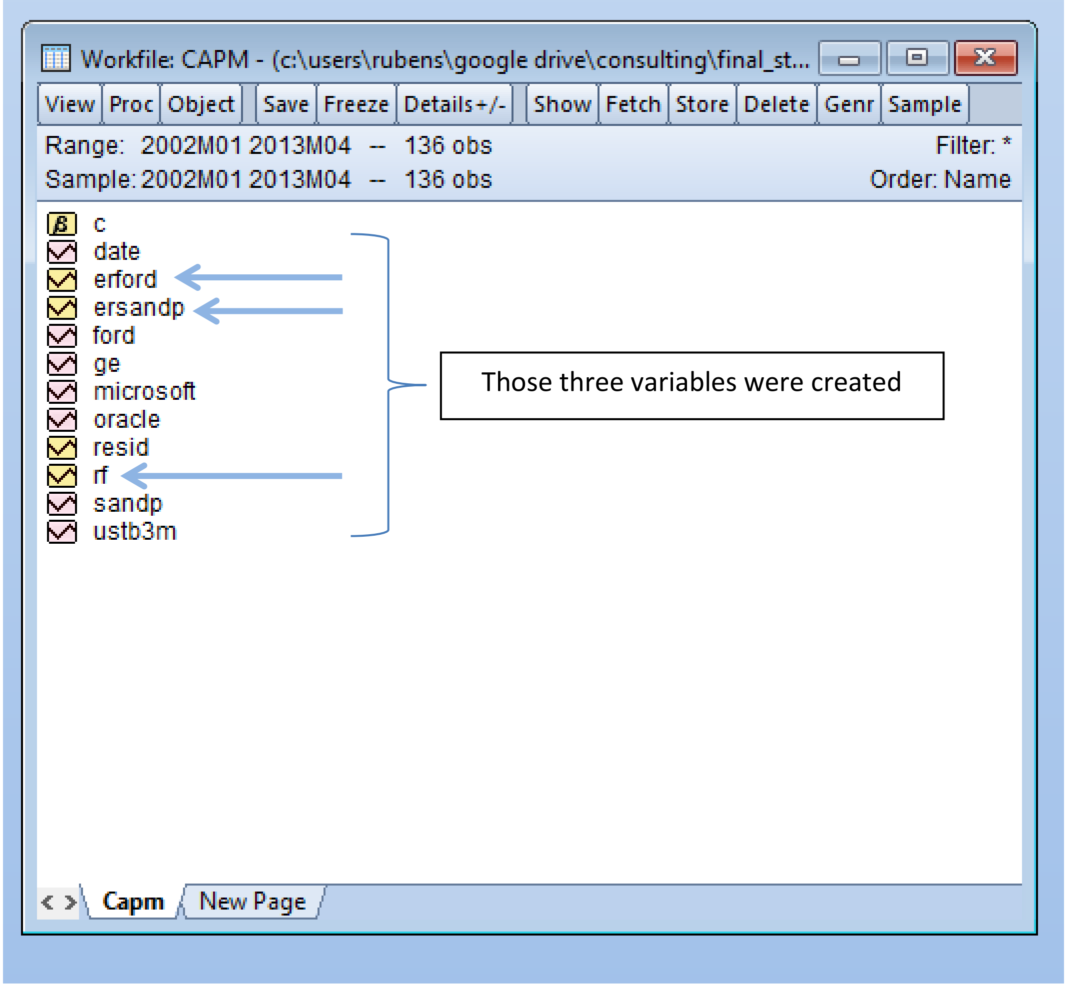
```
rf = (ustb3m/12)/100
erford = dlog(ford) - rf
ersandp = dlog(sandp) - rf
```



The new variable `rt` will be created and saved in the workspace.



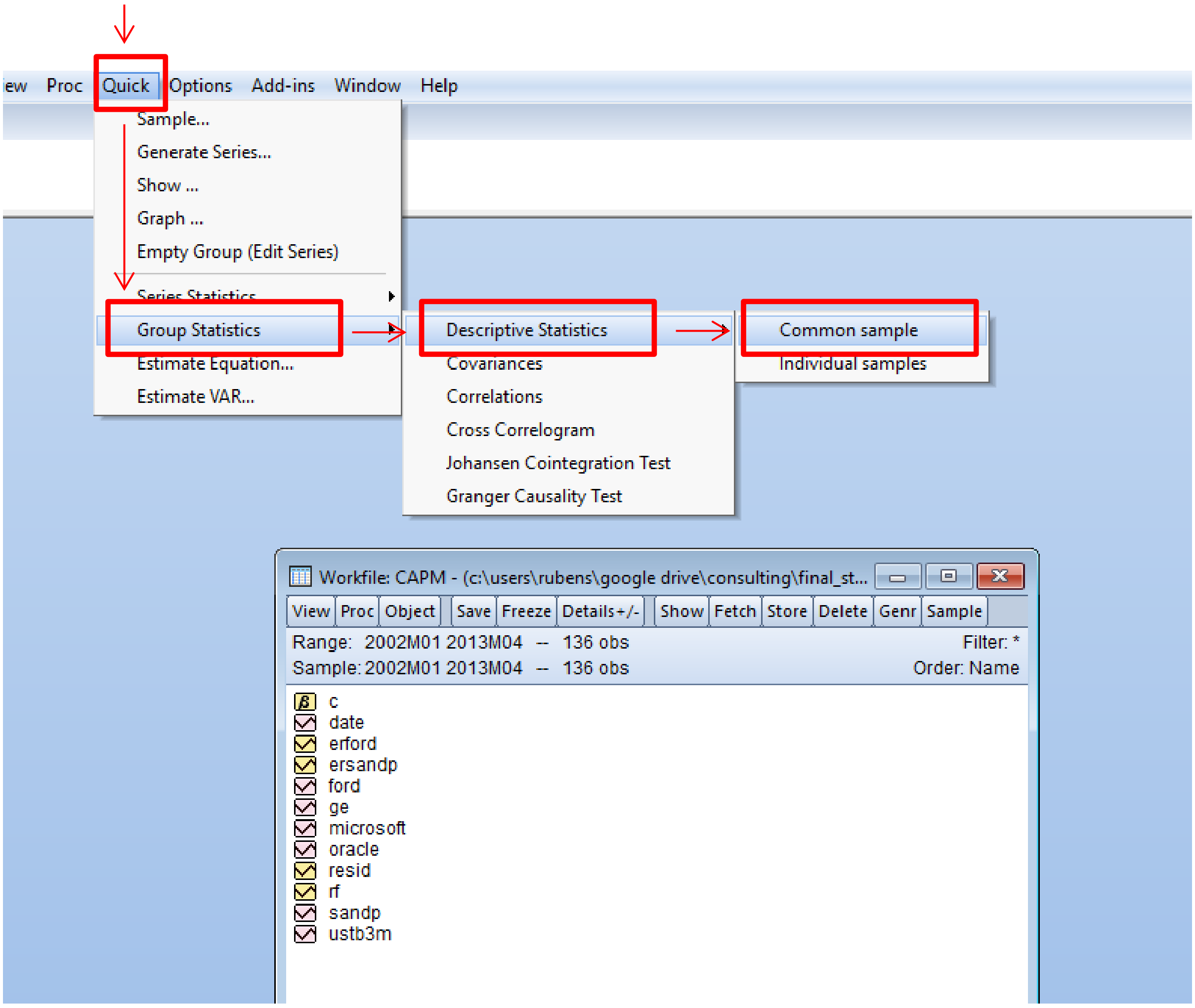
After typing all the formulas, your workspace will look like:



Task 2 Descriptive Analysis of Data

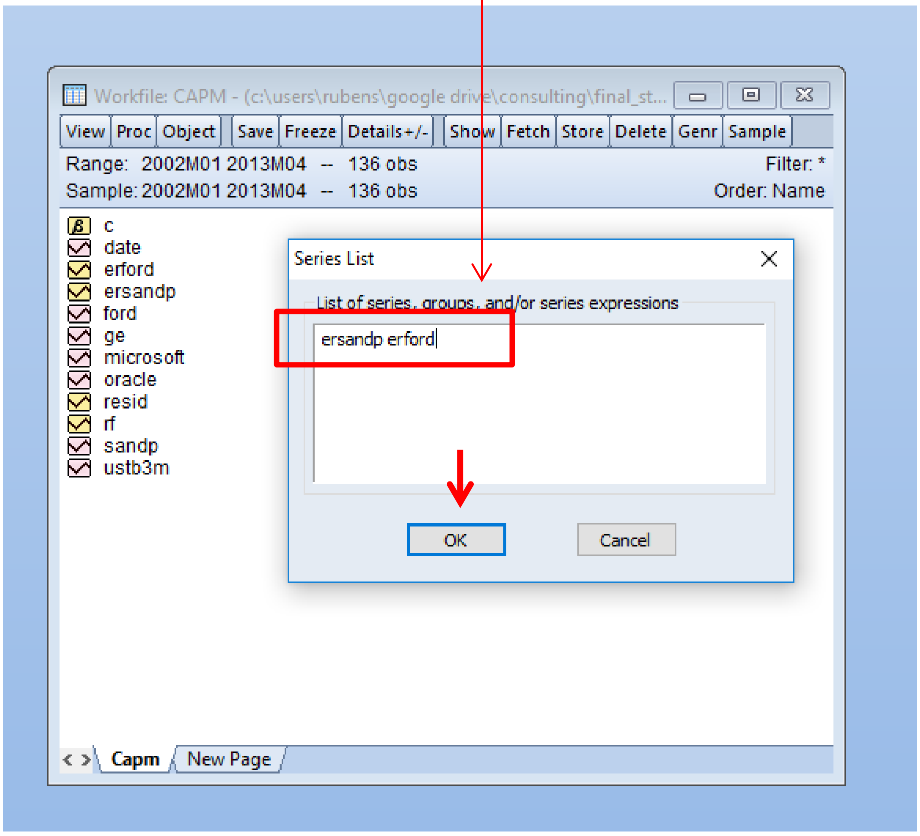
a) Create a group

`Quick\Group Statistics\Descriptive\Common sample`

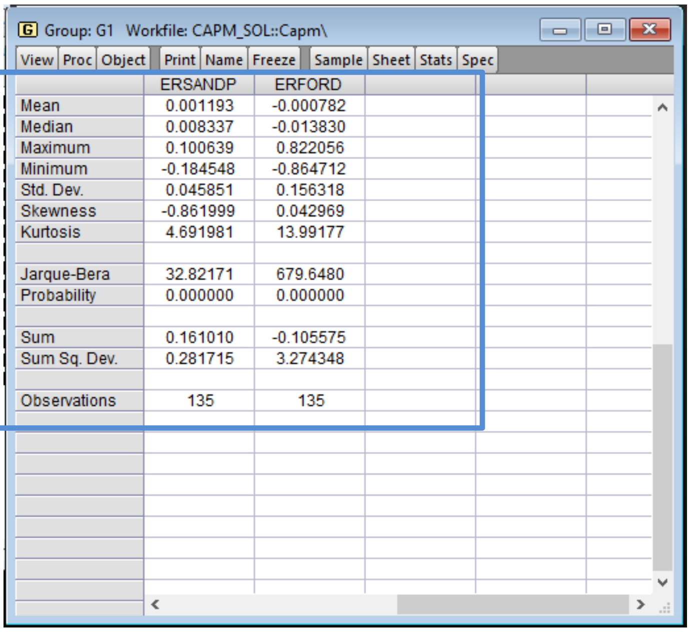


The following scream must appear to you. Then, inside the box called **Series List**, type:

ersandp erford



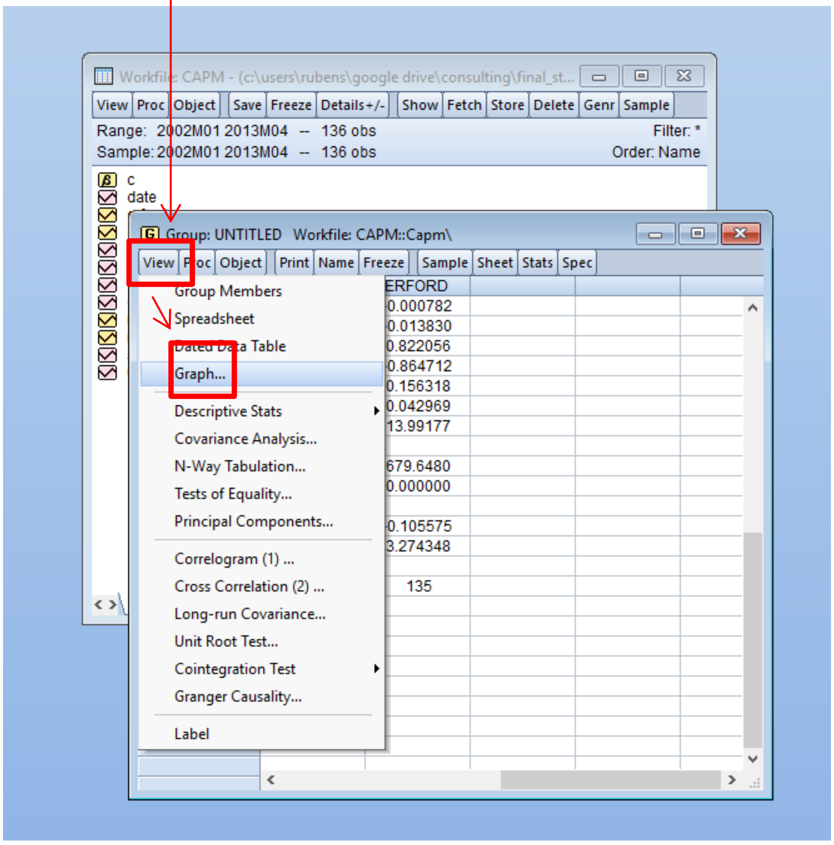
b) check the **descriptive statistics**



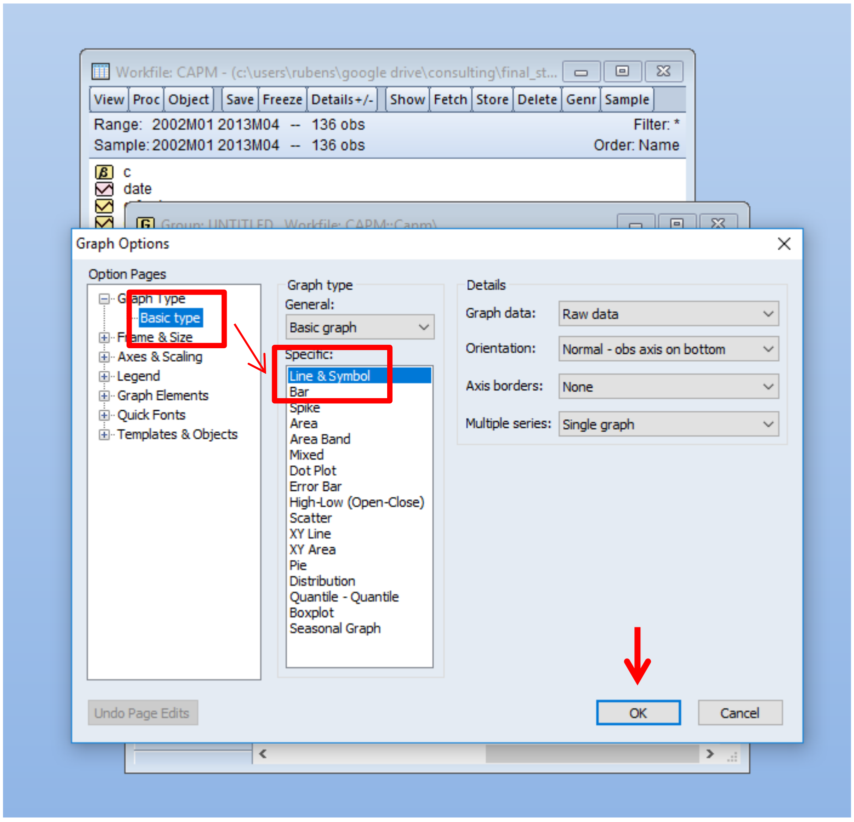
	ERSANDP	ERFORD
Mean	0.001193	-0.000782
Median	0.008337	-0.013830
Maximum	0.100639	0.822056
Minimum	-0.184548	-0.864712
Std. Dev.	0.045851	0.156318
Skewness	-0.861999	0.042969
Kurtosis	4.691981	13.99177
Jarque-Bera	32.82171	679.6480
Probability	0.000000	0.000000
Sum	0.161010	-0.105575
Sum Sq. Dev.	0.281715	3.274348
Observations	135	135

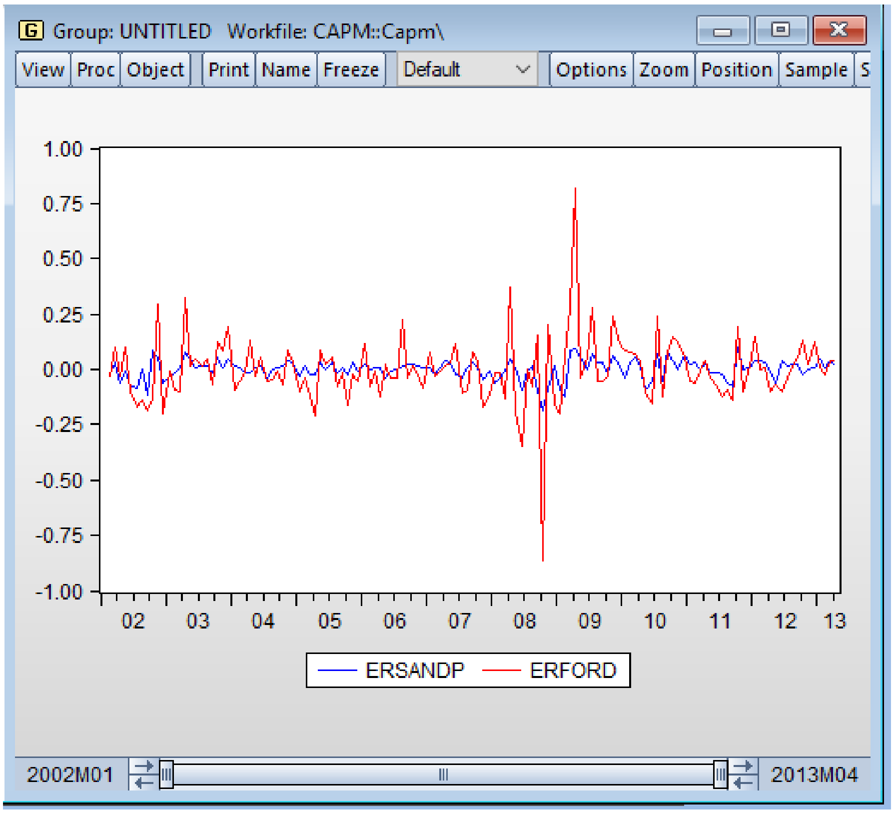
c) create line and scatter **graphs**

In **View/Graph**



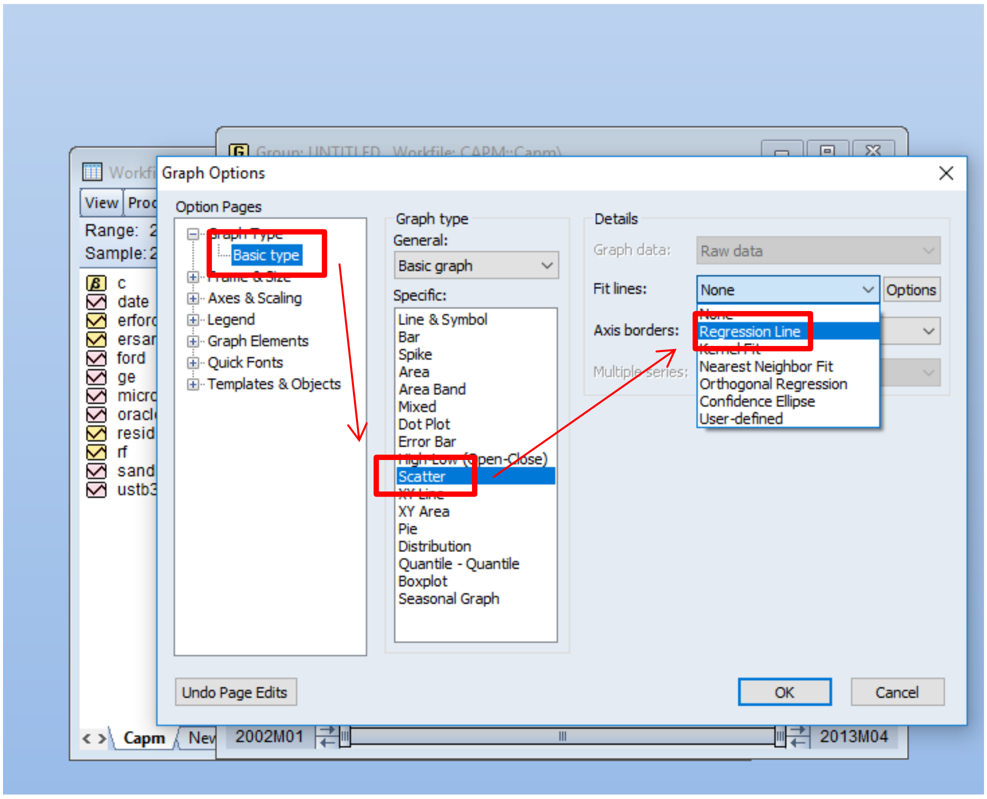
Then, the following scream must appear to you:

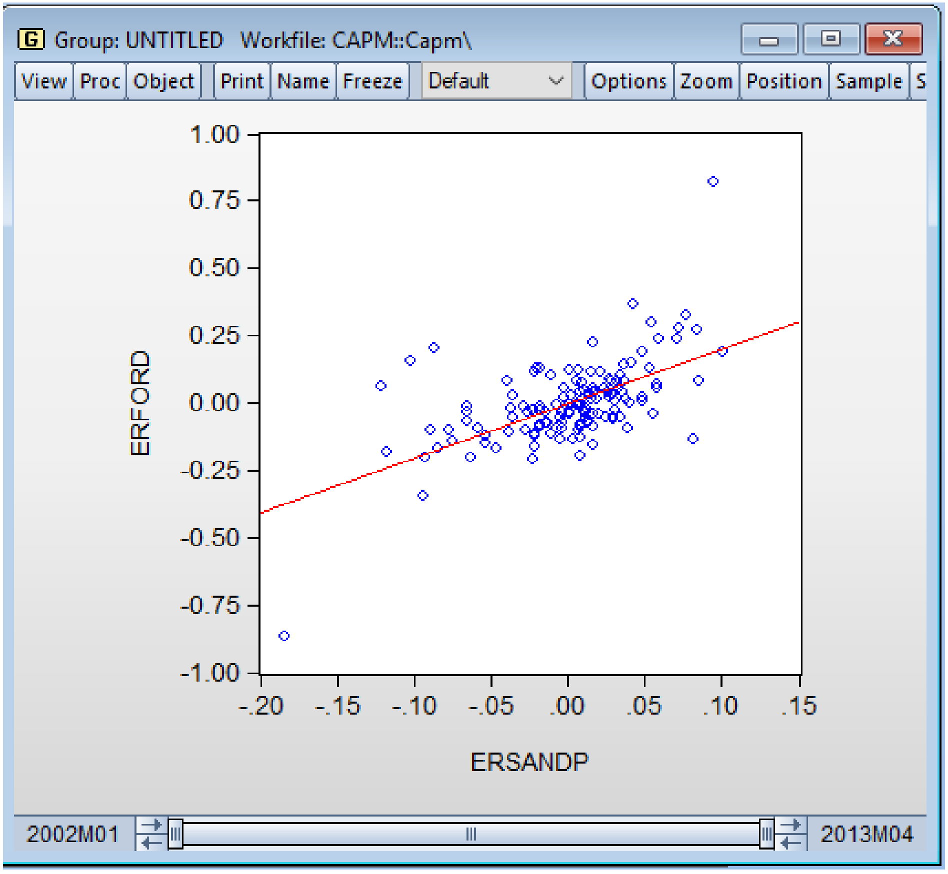




For the scatter we have:

Just select a different type of plot

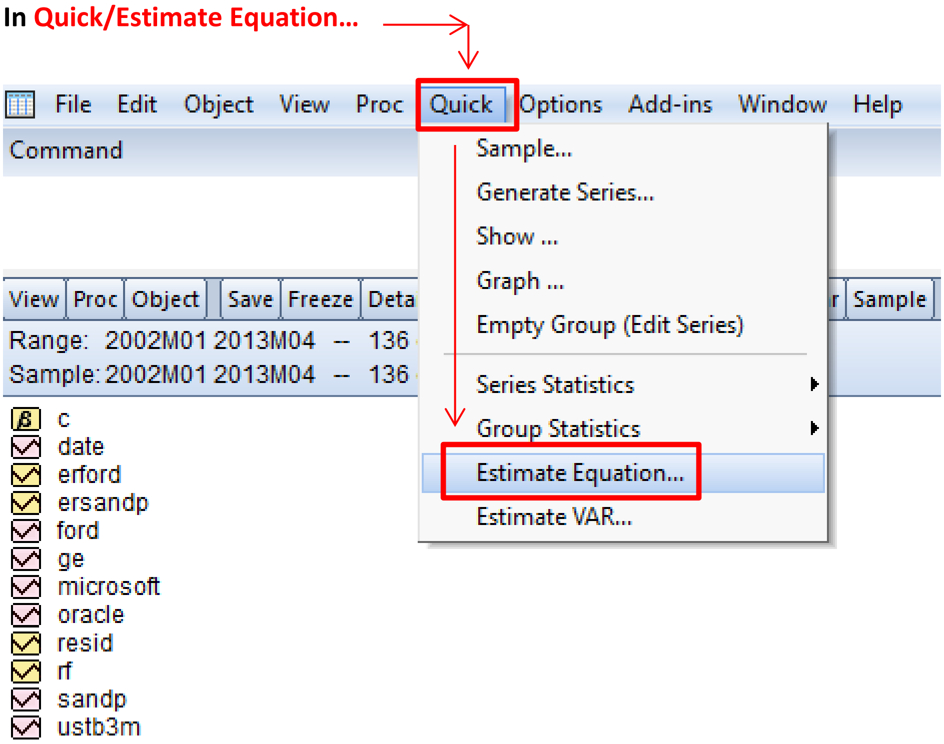




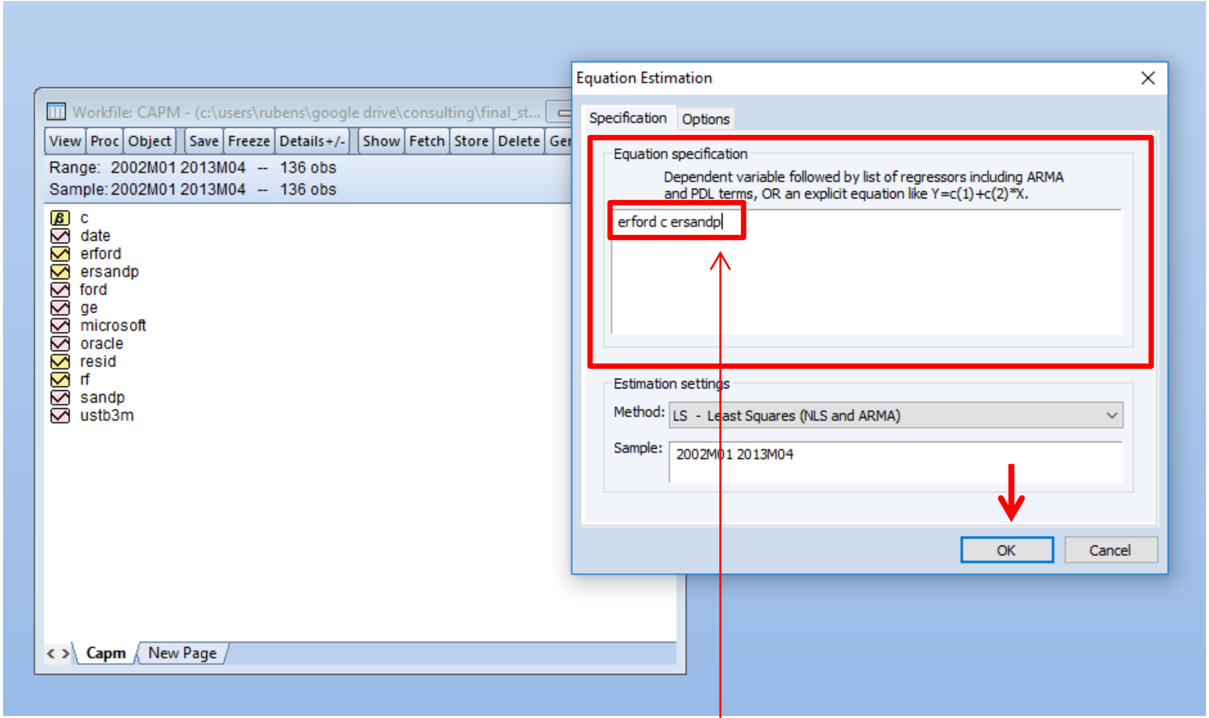
Task 3 Calculating alpha and beta for Ford

a) **Estimate** the CAPM regression for Ford, ie. $(r_i - r_f) = \alpha + \beta(r_m - r_f) + \epsilon$

b) **In Quick/Estimate Equation…**

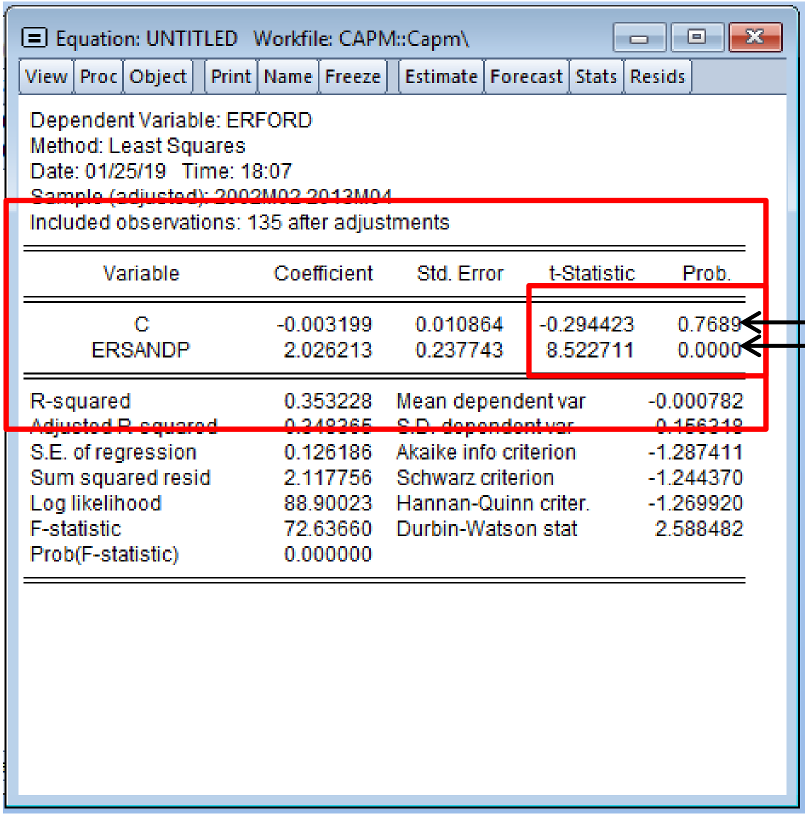


A following scream must appear:



Inside the box, **Equation specification**, type: erford c ersandp

c) **Test** the null hypothesis that $\alpha = 0$ and $\beta = 0$ for both the returns regression


Equation: UNTITLED Workfile: CAPM::Capm\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: ERFORD
Method: Least Squares
Date: 01/25/19 Time: 18:07
Sample (adjusted): 2002M02 2013M04
Included observations: 135 after adjustments

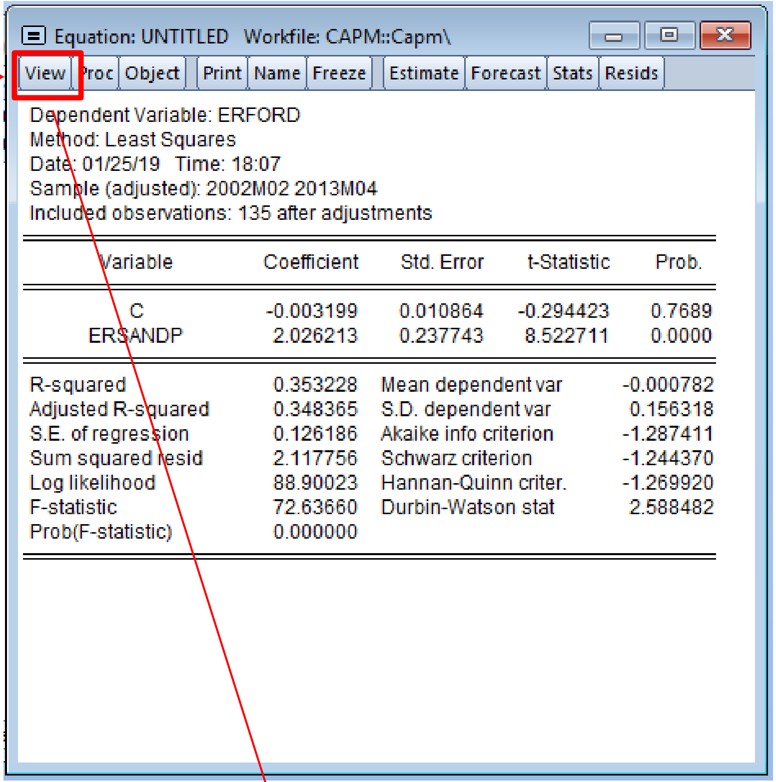
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.003199	0.010864	-0.294423	0.7689
ERSANDP	2.026213	0.237743	8.522711	0.0000

R-squared	0.353228	Mean dependent var	-0.000782
Adjusted R-squared	0.348365	S.D. dependent var	0.156318
S.E. of regression	0.126186	Akaike info criterion	-1.287411
Sum squared resid	2.117756	Schwarz criterion	-1.244370
Log likelihood	88.90023	Hannan-Quinn criter.	-1.269920
F-statistic	72.63660	Durbin-Watson stat	2.588482
Prob(F-statistic)	0.000000		

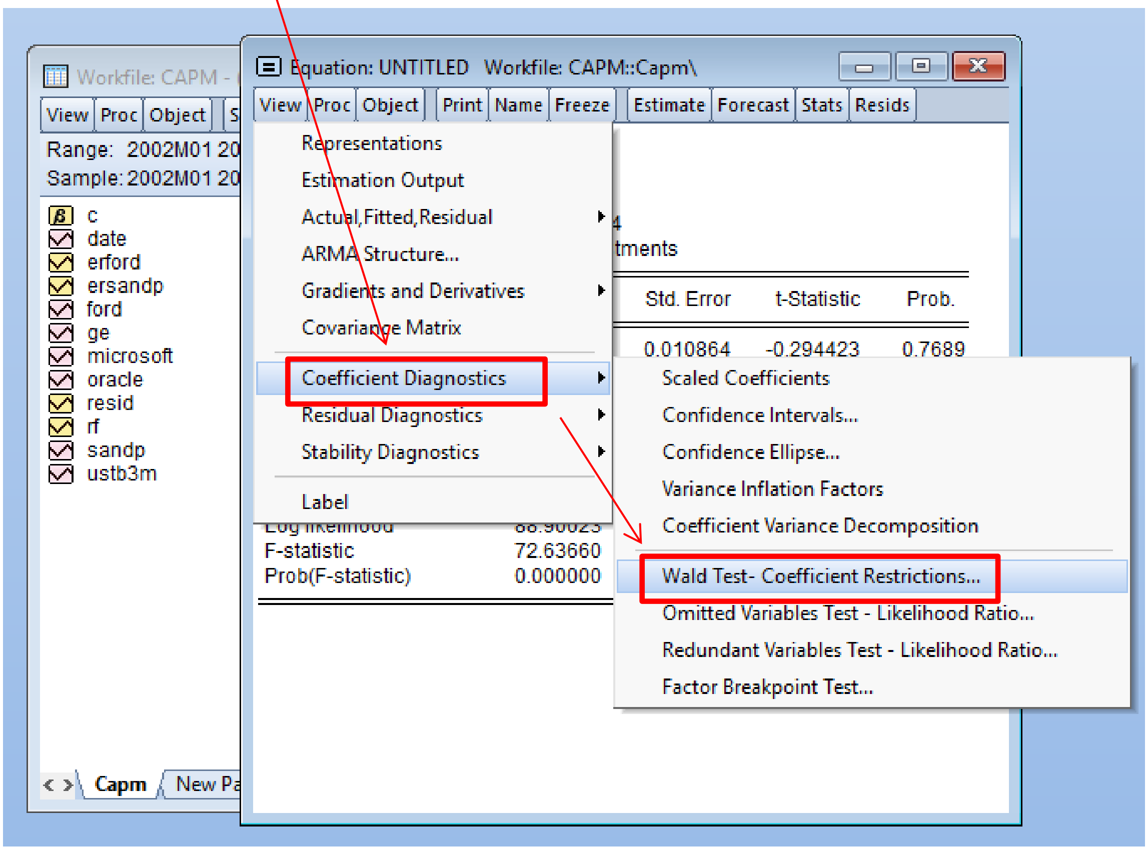
As can be seen in the output above, the intercept (alpha) is not significant from zero at the 5% level (pval 0.7689 > 0.05), but the slope coefficient (beta) is (pval 0.000 < 0.05).

d) **Test** the null hypothesis that $\beta = 1$ for the returns regression

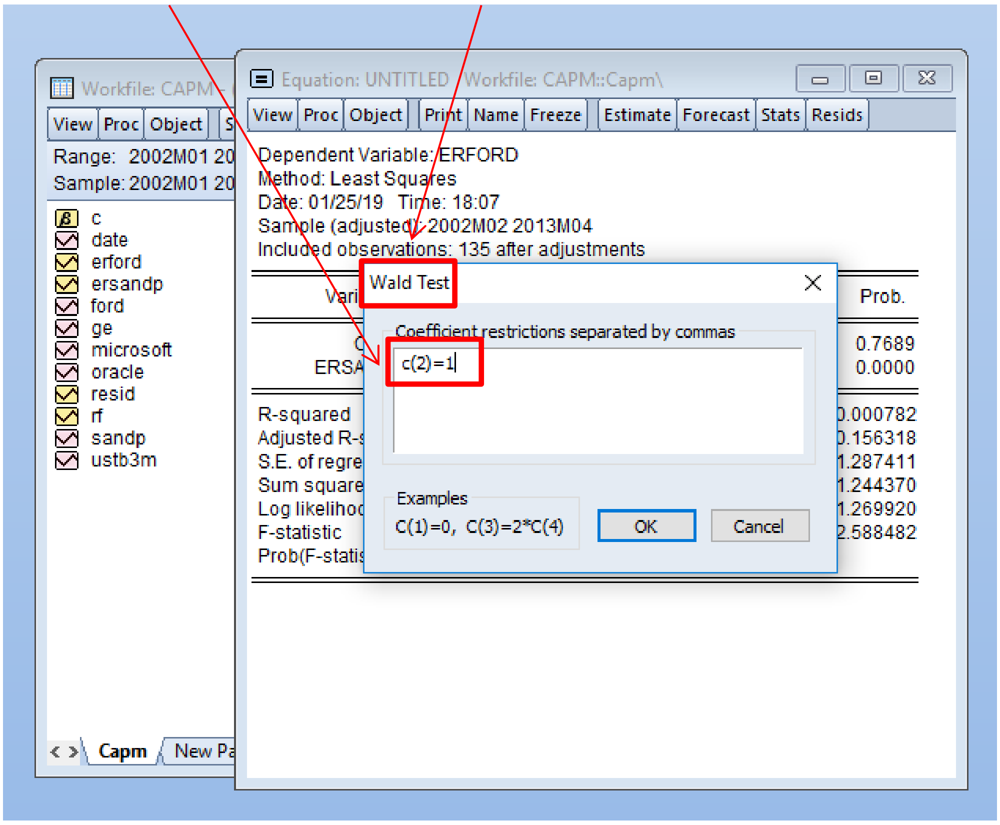
In the output window go to: **View**



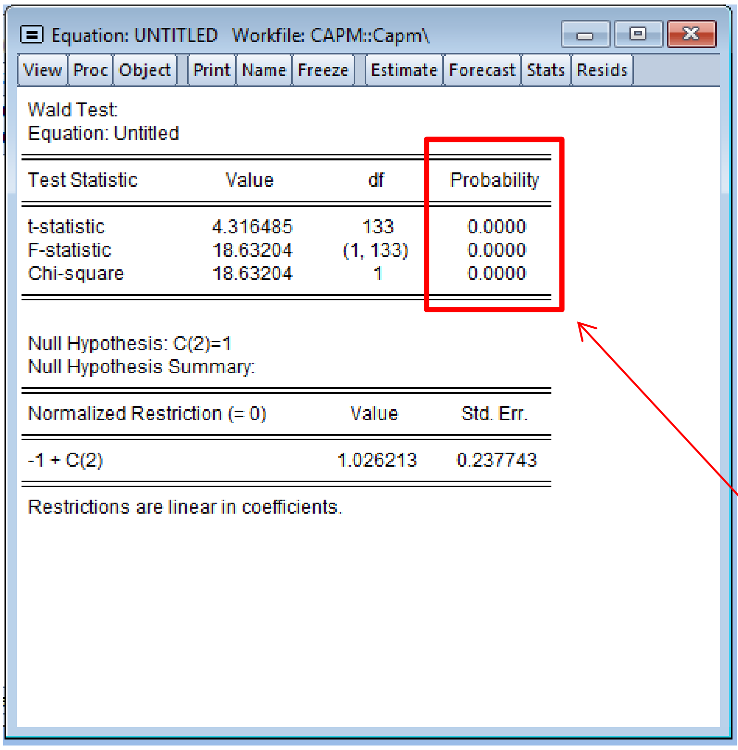
Then, **Coefficient Diagnostics/ Wald Test- Coefficient Restrictions...**



Then, type: **c(2)=1** in the box called **Wald Test**



The output is:



Test Statistic	Value	df	Probability
t-statistic	4.316485	133	0.0000
F-statistic	18.63204	(1, 133)	0.0000
Chi-square	18.63204	1	0.0000

Normalized Restriction (= 0)	Value	Std. Err.
-1 + C(2)	1.026213	0.237743

The output below shows that the t-stat of the test of significance is lower than 0.05, so reject the (two-sided) test that $\beta = 1$.

e) **Interpret** your results

i. Are the coefficients significant?

ii. Can the CAPM price Ford? (check the intercept)

iii. Is there evidence that Ford is more or less risky than the market? (check the slope)

iv. (Trickier) How much of the risk of Ford is idiosyncratic, how much is systematic? (check R2)

Alpha is not significant from zero, which could be interpreted that there is no evidence that Ford is either underperforming or overperforming.
Beta is both significantly different from zero and significantly different from 1, this implies that Ford is exposed to systematic risk (ie. $\beta \neq 0$) and has systematic risk higher than the market (ie. $\beta > 1$)

The R2 measures which fraction of the variability of Ford can be explained by the market. It is 35% here, meaning that 35% of the movements of Ford can be explained by systematic factors. The other 65% is determined by firm-specific idiosyncratic factors (that is, assuming that CAPM is the right model)