Chapter 3: Learning To Use Regression Analysis

In this chapter:

1. Displaying the spreadsheet view of a group of variables (*UE*, Table 3.1)
2. Displaying the descriptive statistics for a group of variables (*UE*, Table 3.1)
3. Displaying the simple correlation coefficients between all pairs of variables in a group (*UE*, Table 3.1)
4. Running a simple regression for Woody's Restaurants example (*UE*, Table 3.2)
5. Documenting the results
6. Displaying the actual, fitted, residual, and a plot of the residuals (*UE*, Table 3.2)

The Woody's Restaurants data will be used to show how EViews can be used to do the items listed above. The steps required to display the 'spreadsheet view', descriptive statistics, and simple correlation coefficients between all pairs of variables in a group are as follows:

Displaying the spreadsheet view of a group of variables (*UE*, Table 3.1):

Step 1. Open the EViews workfile named *Woody3.wf1* by clicking **File/Open/Workfile** on the main menu and double click on the file named *Woody3.wf1*.

Step 2. To Create an EViews group for the Woody's Restaurants example, hold down the **Ctrl** button, click on *Y, N, P & I*, select **Show** from the workfile toolbar, and click **OK**.

Step 3. When you click **OK**, EViews displays the spreadsheet view of the data for the variables included in the group (i.e., *Y, N, P & I* in this case). EViews allows you to change the view with the click of a button. If you are in another view, you can display the spreadsheet view by clicking **View/Spreadsheet** on the group window menu bar.

Step 4. Click **Name** on the group menu bar and enter *GROUP01* in the Name to identify object: window.[1]

Displaying the descriptive statistics for a group of variables (*UE*, Table 3.1):

Step 1. Open the group object created in Display the spreadsheet view of a group of variables by double clicking the group01 icon in the workfile window.

Step 2. Click **View/Descriptive Stats/Individual Samples** on the group window menu bar to view descriptive statistics for *GROUP 01*.[2]

---

[1] You must name a group object if you want to keep its results. Unnamed objects are labeled “UNTITLED” and the results are lost when the object window or the workfile is closed. To name a group, click **Name** on the group menu bar and enter the name in the Name to identify object: window. Once named, a 'group object' is saved with the workfile and can be viewed by double clicking its icon (i.e., group01) in the workfile window.

[2] The group view drop-down menu is divided into four blocks:

- The views in the first block provide various ways of looking at the actual data in the group.
- The views in the second block display various basics statistics.
- The views in the third block are for specialized statistics for time series data.
- The fourth block contains the label view, which provides information regarding the group object.

The descriptive statistics printed in the table below are defined in footnote 3 at the bottom of this page.[3] See Chapter 16 to gain a better understanding of these statistics.

Step 3. Select **View/Spreadsheet** on the group window menu bar to restore the spreadsheet view for *GROUP 01*.

	Y	N	P	I
Mean	125634.6	4.393939	103887.5	20552.58
Median	122015.0	4.000000	95120.00	19200.00
Maximum	166755.0	9.000000	233844.0	33242.00
Minimum	91259.00	2.000000	37852.00	13240.00
Std. Dev.	22404.09	1.919300	55884.51	5141.865
Skewness	0.355246	0.555101	0.672915	0.933694
Kurtosis	1.920334	2.359612	2.280488	3.161758
Jarque-Bera	2.296908	2.258639	3.202315	4.830791
Probability	0.317127	0.323253	0.201663	0.089332
Observations	33	33	33	33

Displaying the simple correlation coefficients between all pairs of variables in a group (*UE*, Table 3.1):

Step 1. Open the group object created in Display the spreadsheet view of a group of variables by double clicking the group01 icon in the workfile window.

Step 2. Click **View/Correlations** on the group window menu bar to display the simple correlation coefficients between all pairs of variables included in the group object (see the table below).[4]

	Y	N	P	I
Y	1.000000	-0.144225	0.392568	0.537022
N	-0.144225	1.000000	0.726251	-0.031534
P	0.392568	0.726251	1.000000	0.245198
I	0.537022	-0.031534	0.245198	1.000000

[3] All of the statistics are calculated using observations in the current sample.
Mean is the average value of the series, obtained by adding up the series and dividing by the number of observations. Median is the middle value (or average of the two middle values) of the series when the values are ordered from the smallest to the largest. The median is a robust measure of the center of the distribution that is less sensitive to outliers than the mean. Max and Min are the maximum and minimum values of the series in the current sample. Std. Dev. (standard deviation) is a measure of dispersion or spread in the series. Skewness is a measure of asymmetry of the distribution of the series around its mean. Kurtosis measures the peakedness or flatness of the distribution of the series. Jarque-Bera is a test statistic for testing whether the series is normally distributed. The test statistic measures the difference of the skewness and kurtosis of the series with those from the normal distribution. Under the null hypothesis of a normal distribution, the Jarque-Bera statistic is distributed as with 2 degrees of freedom. The reported Probability is the probability that a Jarque-Bera statistic exceeds (in absolute value) the observed value under the null—a small probability value leads to the rejection of the null hypothesis of a normal distribution.

[4] **View/Correlations** displays the correlation matrix of the series in the selected group. Observations for which any one of the series has missing data are excluded from the calculation.

Running a simple regression for Woody's Restaurants example (*UE* 3.2):

To regress the number of customers served (*Y*) on the Constant (*C*), the number of direct market competitors within a two-mile radius of the Woody's location (*N*), the number of people living within a three-mile radius of the Woody's location (*P*), and the average household income of people living within a three-mile radius of the Woody's location (*I*):

Step 1. Open the EViews workfile named *Woody3.wf1*.
Step 2. To estimate *UE*, Equation 3.6, p. 74, select **Objects/New Object/Equation** on the workfile menu bar.[5] You can name the equation object now by deleting *Untitled* in the Name for Object: window and typing a name for the equation you are about to estimate, or you can skip this step and name the equation later (if you find that it is worth saving). Click **OK** to reveal the Equation Specification: window.
Step 3. Enter *Y C N P I* in the Equation Specification: window.
Step 4. Do not change the default settings for Method: and Sample:.
Step 5. Click **OK** to get the computer output printed in the top box of *UE*, Table 3.2, p. 76.
Step 6. To name the equation for later use, select **Name** on the equation window menu bar, enter *EQ01* in the Name to identify object: window, and click **OK**.[6]

Documenting the results:

All of the data necessary to write the results printed in (*UE*, Equation 3.7, p. 77) are printed in the EViews regression output. EViews provides a variety of views for regression results. For example, if you are documenting the results in a word processing file, you can facilitate the process by clicking **View/Representations** on the equation window menu bar to get the following:

```
Estimation Command:
====================
LS Y C N P I

Estimation Equation:
====================
Y = C(1) + C(2)*N + C(3)*P + C(4)*I

Substituted Coefficients:
====================
Y = 102192.4277 - 9074.674399*N + 0.3546683674*P + 1.287923391*I
```

You can copy the last line and paste it into your word processing file to get the first line of *UE*, Equation 3.7.

---

[5] Alternately, select **Quick/Estimate Equation** from the main menu. If this method is used, you must name the equation to save it. Click **Name** on the equation menu bar and enter the desired name in the Name for object: window and click **OK**.

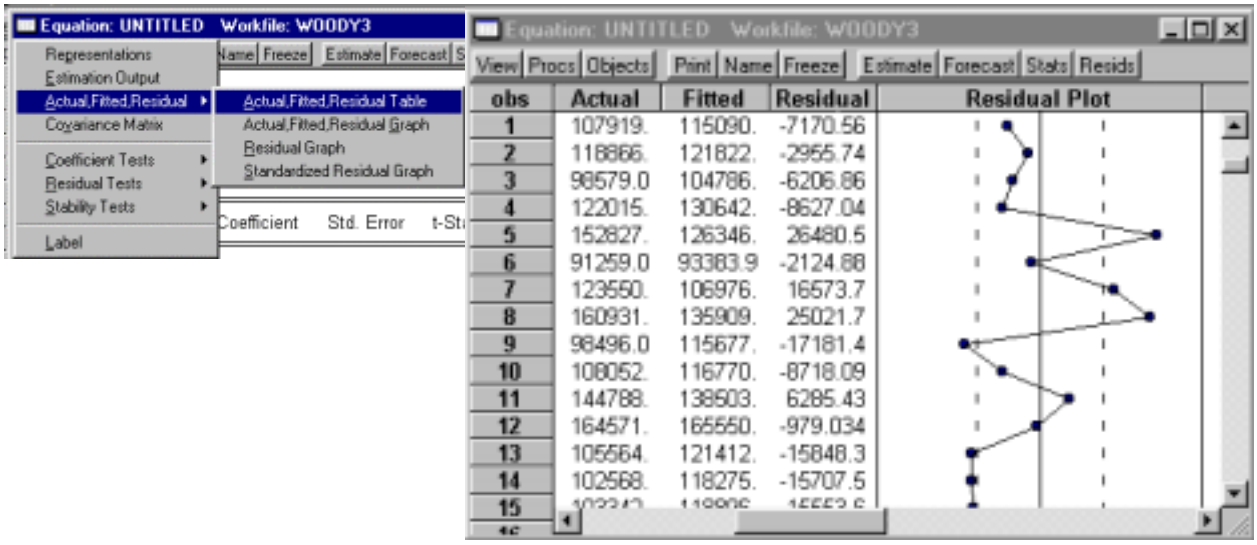
[6] Restrictions on naming an object include, the name cannot exceed 16 characters and it cannot be a reserved name. The following names are reserved and should not be used for series: ABS, ACOS, AR, ASIN, C, CON, CNORM, COEF, COS, D, DLOG, DNORM, ELSE, ENDIF, EXP, LOG, LOGIT, LPT1, LPT2, MA, NA, NRND, PDL, RESID, RND, SAR, SIN, SMA, SQR, and THEN.

Select **View/Estimation Output** on the group window menu bar to restore the estimation output view for *EQ01*.

Displaying the actual, fitted, residual, and a plot of the residuals for a regression (see *UE*, Table 3.2, p. 76):

Step 1. Open the EViews workfile named *Woody3.wf1* and open the equation named *EQ01* by double clicking the equation icon in the workfile window.

Step 2a. Click **View/Actual,Fitted,Residual/Actual,Fitted,Residual Table** on the equation window menu bar (see figure below left). Click **OK** to reveal the figure below right.



Step 2b. Alternately, to display a graph of the actual, fitted, and residuals for a regression, click **View/Actual,Fitted,Residual/Actual,Fitted,Residual Graph** (see figure below left). Click **OK** to reveal the figure below right. Other views available in the equation window will be explained in future chapters of this guide.