

Chapter 1: An Overview of Regression Analysis

In this chapter:

1. A simple example of regression analysis (*UE* 1.4):
 - a) [Creating an EViews workfile](#)
 - b) [Entering data into an EViews workfile](#)
 - c) [Creating a group in EViews](#)
 - d) [Graphing with EViews](#)
 - e) [Generating new variables in EViews](#)
2. [Exercises](#)

Section 1.4 describes how a weight guesser can use regression analysis to make better guesses about a person's weight (Y) based only on a person's height (X). The data for this example are printed in *UE*¹, Table 1.1, p. 20, and available in the EViews workfile named *htwt1.wfl*. However, since some data sets may not be in EViews file format, the process of creating an EViews workfile and entering data into the new workfile will be reviewed here. The process of creating an EViews workfile and importing data into the new workfile from an Excel file will be reviewed in Chapter 2.

Creating an EViews workfile:

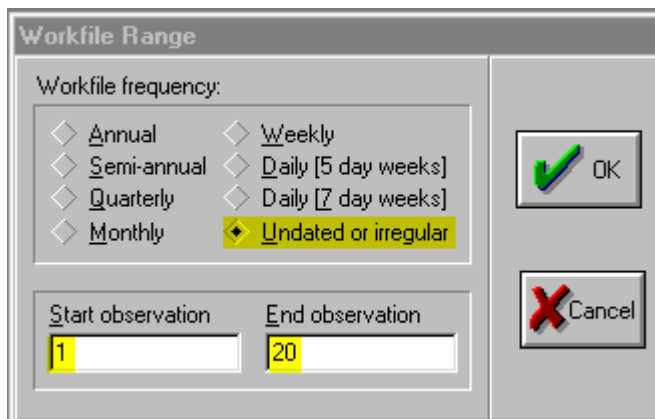
To create a new workfile for the weight (Y) versus height (X) data set printed in columns (2) and (3) of (*UE*, Table 1.1, p. 20):

Step 1. Select **File/New/Workfile** on the EViews main menu bar.

Step 2. Set the **Workfile frequency:** to **Undated or irregular**.

Step 3. Since there are 20 observations in the data set, set the **Start observation** to 1 and **End observation** to 20 (refer to areas highlighted in yellow in the figure on the right).²

Step 4. Once you have selected the appropriate frequency and entered the information for the workfile range, click **OK**. EViews will create an untitled workfile, and will display the workfile window in the main work area of the EViews screen. The workfile window displays two pairs of numbers highlighted in yellow: one for the **Range**: of data contained in the workfile, and the second for the current workfile **Sample**:. Both the workfile range and sample can be

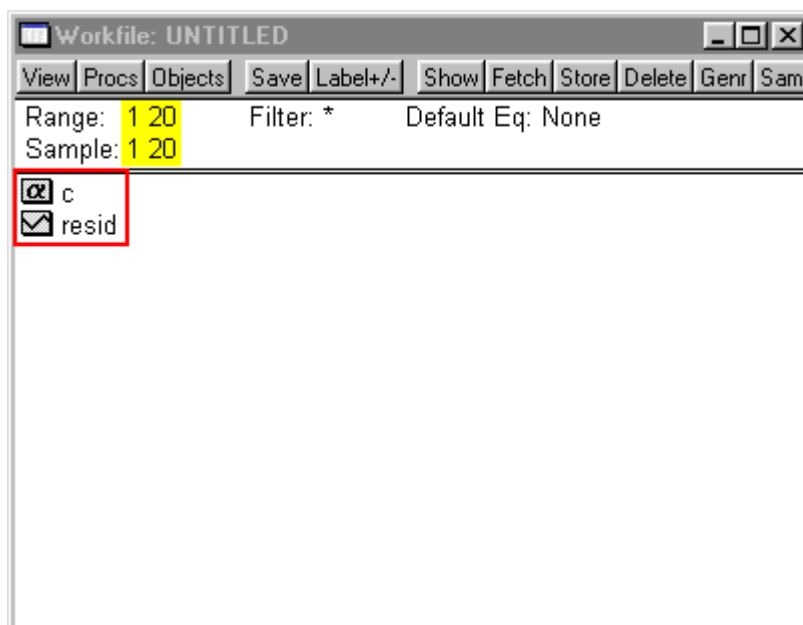


¹ Using *Econometrics, A Practical Guide* (fourth edition), by A. H. Studenmund will be referred to as (*UE*) when referenced in this handbook.

² For monthly, quarterly, or annual time series data, with observations ranging from the beginning of 1980 to the end of 1996, select from the following options: for monthly data, set the **Workfile frequency:** to *monthly*, and specify the **Start date**, 1980:01, and the **End date**, 1996:12. For quarterly data, set the **Workfile frequency:** to *quarterly*, and specify the **Start date**, 1980:1, and the **End date**, 1996:4. For annual data, set the **Workfile frequency:** to *annual*, and specify the **Start date**, 1980, and the **End date**, 1996. See **Help/Contents/EViews Basics/Workfile Basics** for more information about creating and using workfiles in EViews.

changed after the workfile has been created. Note that all new workfiles will contain two objects boxed in red: the coefficient vector named c and the residual series named $resid$. (see the figure below).

Step 5. To save your workfile, select **Save** on the workfile menu bar or **File/Save** or **File/Save As** on the main menu bar and enter *htwt1.wfl* in the **File name:** window, select the destination drive in the **Save in:** window and click **OK**. It is advisable to save the workfile after it is created and frequently thereafter.



Entering data into an EViews workfile:

Complete the section entitled [Creating an EViews workfile](#) before attempting this procedure. To enter the data for height (inches above 5') and weight from columns (2) and (3) of Table 1.1, (*UE*, p. 20) into the newly created workfile, follow the steps below.

Step 1. To create a new series for the weight (Y) variable, select **Objects/New Object/Series** from the main menu or the workfile menu, enter Y in the **Name for Object:** window and click **OK**. All of the observations in the series will be assigned the missing value code 'NA'.³

Step 2. To enter data into the newly created series, double click the series in the workfile window and click **edit+/-** on the series window menu bar. The numbers from the table can be entered to replace the NA's in the spreadsheet, pressing **Enter** after each entry. After the numbers have been entered, click **edit+/-** on the series window menu bar to save the changes and exit the edit function. The series window can be closed by clicking the **X** button in the upper right corner of the series window.

Step 3. Repeat the process for the height (X) variable.

Step 4. To save changes to your workfile, click **Save** on the workfile menu bar.

Most data are available in spreadsheet file or ASCII text file format, which can be imported directly into the workfile. The procedure for importing the weight and height data from an ASCII text file will be explained in [Chapter 2](#). The edit procedure described above is mostly used to adjust data series after they are imported. In cases where expressions can be used to assign values for the series, click on **Quick/Generate Series** or click on **Genr** on the workfile window menu bar and enter the expression defining the series. For more information on how to generate a series using expressions, see **Help/Contents/EViews Basics/Working with Series**.

³ A series can also be created by clicking on **Quick/Generate Series** on the main menu bar or by clicking on **Genr** on the workfile window menu bar and enter the expression $Y = NA$ or any other mathematical expression.

Creating a group in EViews:

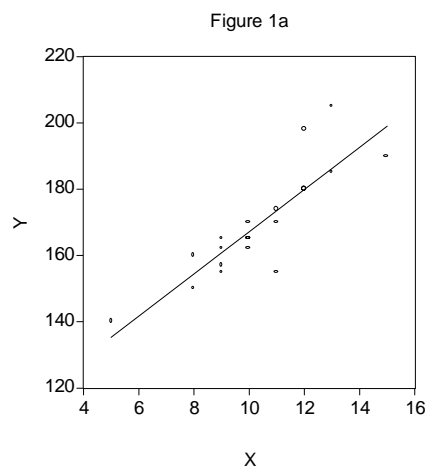
EViews provides specialized tools for working with groups of series. Follow these steps to create a group object containing the Y & X series:


- Step 1.** Open the EViews workfile named *htwt1.wfl* by selecting **File/Open/Workfile** on the main menu and click on the file name.
- Step 2.** To create a group object for the X and Y variables, hold down the **Ctrl** button and click on X and Y and then select **Show** from the workfile toolbar.
- Step 3.** To name a group, click **Name** or **Object/Name** on the group menu bar and enter *GROUP01* in the **Name to identify object:** window.⁴ Some features of groups are described in the footnote at the bottom of this page.⁵
- Step 4.** To save changes to your workfile, click **Save** on the workfile menu bar.

Graphing with EViews:

The weight guesser in section 1.4 (*UE*) hypothesized a positive relationship between weight and height. As a first step, plot Y against X in order to visually confirm the hypothesis.

- Step 1.** Open the EViews workfile named *htwt1.wfl*.
- Step 2.** To plot Y against X , [open the two series in a group window](#) (enter X before Y because EViews specifies the first variable selected in a group as the "X" axis variable and the second as the "Y" axis variable).
- Step 3.** Select **View/Graph/Scatter/Scatter with Regression** and click **OK** to reveal Figure 1a. EViews shows a graph with Y on the vertical axis and X on the horizontal axis as well as a line representing the regression equation between the two variables (i.e., plots *UE*, Equation 1.21, p. 21). As you can see in Figure 1a, a positive relationship between Y and X is visually confirmed but the relationship is distorted because EViews uses **Optimal-Linear Scaling** as the graph default setting.
- Step 4.** To change the graph scaling, click the right mouse button anywhere in the graph and select **Options**.
- Step 5.** Change the **Graph Scaling** option to **Linear - Force Through Zero**.



⁴ 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.

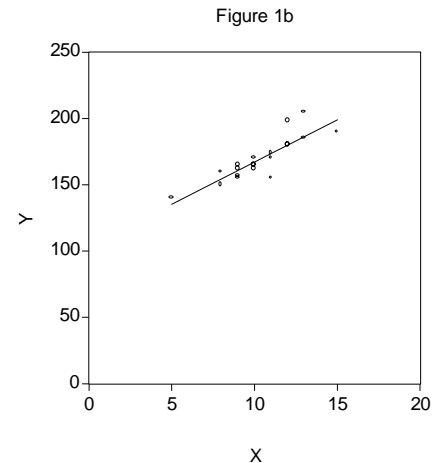
⁵ **Features of Groups:** A group is simply a list of series identifiers, not a copy of the data in the series. If you change the data for one of the series in the group, you will see the changes reflected in the group. If you delete a series from the workfile, it will disappear from any group that included the series. If the deleted series is the only series in a group, the group will also be deleted. Renaming a series changes the reference in every group containing the series. Groups, like other EViews objects, contain their own views and procedures. A detailed description can be found in **Help/Contents/Statistical Views and Procedures/Group Views and Procs**.

Step 6. Click **OK** to get the graph in Figure 1b.

Note that the graph no longer gives a distorted view of the relationship between Y and X . The equation represented in Figures 1a & 1b has the following formula (see *UE* Equation 1.21, p. 21): *Estimated weight* (Y) = $103.40 + 6.38 * \text{Height in inches above five feet}$ (X)

Generating new variables in EViews:

Follow the steps below to generate the data in columns (4) through (7) of *UE*, Table 1.1, p. 20:



Step 1. Open the EViews workfile named *htwt1.wf1*.

Step 2. To generate a new variable named *predicted*, with the data printed in column (4) of *UE*, Table 1.1, click **Genr** on the workfile menu and enter the formula ($\text{predicted} = 103.4 + 6.38 * X$) in the **Enter Equation:** window. Click **OK** and a new variable named *predicted* appears in the workfile window with the values in column (4) of Table 1.1.

Step 3. To generate a new variable named *residual*, with the data printed in column (5) of Table 1.1, click **Genr** on the workfile menu and enter the formula ($\text{residual} = Y - \text{predicted}$) in the **Enter Equation:** window. Click **OK** and a new variable named *residual* will appear in the workfile window with the values in column (5) of Table 1.1.

Step 4. To generate a new variable named *gainorloss*, with the data printed in column (6) of Table 1.1, click **Genr** on the workfile menu and enter the formula $\text{gainorloss} = .5 * (\text{residual} >= -10 \text{ and } \text{residual} <= 10) - .6 * (\text{residual} < -10 \text{ or } \text{residual} > 10)$ in the **Enter Equation:** window.⁶ Click **OK** and a new variable named *gainorloss* will appear in the workfile window with the values in column (6) of Table 1.1.

Step 5. [Create a group](#) with *x*, *y*, *predicted*, *residual*, and *gainorloss* in order. Click **View/Spreadsheet** to view the data in Table 1.1.

Step 6. Type $=@sum(\text{gainorloss})$ in the command window, press **Enter**, view the net amount gained on the status line in the lower left of the screen (i.e., **Scalar = 6.7**).


⁶ The conditional value typed within a set of parenthesis is equal to one when the condition is true and zero if the condition is false. Thus, the condition ($\text{residual} >= -10 \text{ and } \text{residual} <= 10$) is equal to one if the residual falls within the range -10 to +10 and zero if it is outside of this range. The condition ($\text{residual} < -10 \text{ or } \text{residual} > 10$) is equal to one if the residual is less than -10 or greater than +10 and zero if it falls within the range -10 to +10.

Exercises:

13.

- a.
- b.
- c.
- d. The data for this exercise are available in an EViews file named *help1.wfl*. However, it is a good idea to take this opportunity to make sure that you understand how to create a workfile and enter data. If you want to skip these steps, open the EViews workfile named *help1.wfl* and skip steps 1 through 6 below.

Follow these steps to create a new workfile for the help-wanted advertising index (*hwi*) and unemployment rate (*ur*) data set printed in the table accompanying exercise 13.

- Step 1.** Select **File/New/Workfile** on the EViews main menu bar. Set the workfile frequency to *Quarterly*. Enter *1962:1* for the **Start date:** and *1967:4* for the **End date:** (see footnote 2). Click **OK** to open a new untitled workfile.
- Step 2.** To save your workfile, select **Save** on the workfile menu bar or **File/Save** or **File/Save As** on the main menu bar and enter *help1.wfl* in the **File name:** window, select the destination drive in the **Save in:** window, and click **OK**.
- Step 3.** To create a new series for the help-wanted advertising index (*hwi*) variable, select **Objects/New Object/Series** from the main menu or the workfile menu, enter *hwi* in the **Name for Object:** window and click **OK**. All of the observations in the series will be assigned the missing value code 'NA'.
- Step 4.** To enter data into the newly created series, double click the series in the workfile window and click **edit+/-** on the series window menu bar. The numbers from the table can be entered to replace the NA's in the spreadsheet, pressing **Enter** after each entry. After the numbers have been entered, click **edit+/-** on the series window menu bar to save the changes and exit the edit function. The series window can be closed by clicking the  button in the upper right corner of the series window.
- Step 5.** Repeat the process for the unemployment rate (*ur*) variable.
- Step 6.** To save changes to your workfile, click **Save** on the workfile menu bar.

The process of running a regression is explained, in detail, in [Chapter 2](#), but a short description of the process is presented here. Follow these steps to run the regression between *hwi* and *ur*:

- Step 1.** Open the EViews workfile named *Help1.wfl*.
- Step 2.** Select **Objects/New Object/Equation** and click **OK** (or **Quick/Estimate Equation** from the main menu).
- Step 3.** Enter the dependent variable (*hwi*), the constant (*c*) and the independent variable (*ur*) in the **Equation Specification:** window that appears, using spaces between each term. It is important to enter the dependent variable first (*hwi* in this case).
- Step 4.** Select the estimation method {LS - Least Squares (NLS and ARMA)}. EViews uses this as the default setting because it is selected most of the time.
- Step 5.** The workfile sample range is automatically entered but it can be changed if another sample range is desired.
- Step 6.** Click **OK** when finished to reveal the regression output generated by EViews.