

Example 3.2

Modifying Gretl data files

Pilar González and Susan Orbe

Dpt. Applied Economics III (Econometrics and Statistics)

- 1 3.2.1. Adding new data.
 - More observations.
 - Adding variables.
- 2 3.2.2. Importing data from other files.
- 3 3.2.3. Generating new variables.
- 4 3.2.4. Simulating variables.

- 1 3.2.1. Adding new data.
 - More observations.
 - Adding variables.
- 2 3.2.2. Importing data from other files.
- 3 3.2.3. Generating new variables.
- 4 3.2.4. Simulating variables.

Example 3.2.1. Adding new data.

More observations.

New data are available on the variables included in the file `chicken.gdt`.

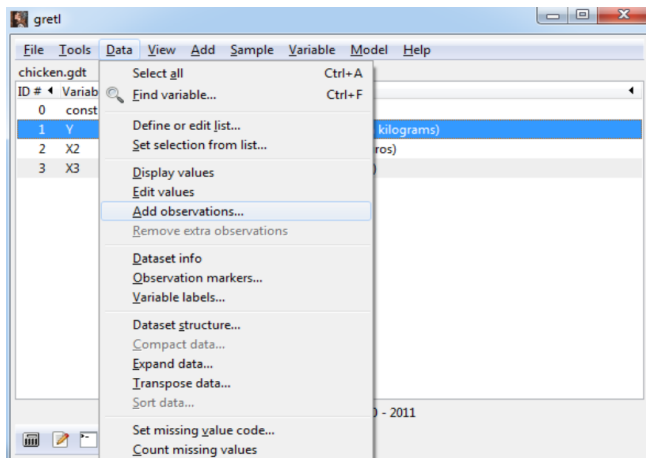
year	2012
Y	52.9
X2	2478.7
X3	2.82

Enter this new observation in the file `chicken.gdt` manually and save all the changes.

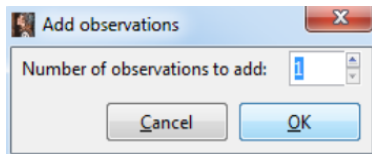
Example 3.2.1. Adding new data. Observations.

To **add** new observations, select the option *Add observations* from the **Data** pulldown menu:

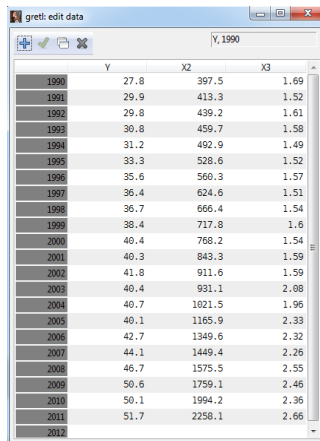
Data -> Add observations...



Example 3.2.1. Adding new data. Observations.



Number of new observations: 1



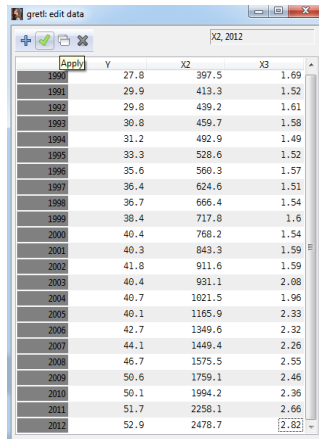
The 'gretl: edit data' window showing a table of data. The title bar includes a plus icon, a checkmark, a save icon, and a close button. The window title is 'gretl: edit data'. The table has columns labeled 'Y', 'X2', and 'X3'. The rows are indexed from 1990 to 2012. A new empty row has been added at the bottom, corresponding to the year 2012.

	Y	X2	X3
1990	27.8	397.5	1.69
1991	29.9	413.3	1.52
1992	29.8	439.2	1.61
1993	30.8	459.7	1.58
1994	31.2	492.9	1.49
1995	33.3	528.6	1.52
1996	35.6	560.3	1.57
1997	36.4	624.6	1.51
1998	36.7	666.4	1.54
1999	38.4	717.8	1.6
2000	40.4	768.2	1.54
2001	40.3	843.3	1.59
2002	41.8	911.6	1.59
2003	40.4	931.1	2.08
2004	40.7	1021.5	1.96
2005	40.1	1165.9	2.33
2006	42.7	1349.6	2.32
2007	44.1	1449.4	2.26
2008	46.7	1575.5	2.55
2009	50.6	1759.1	2.46
2010	50.1	1994.2	2.36
2011	51.7	2258.1	2.66
2012			

A new empty row is added

Example 3.2.1. Adding new data. Observations.

Enter data manually and click on the icon *Apply*.



The screenshot shows the 'gretl: edit data' window with a table of data. The 'Apply' button is highlighted. The table has columns for years (1990 to 2012) and variables Y, X2, and X3. The data is as follows:

	Y	X2	X3
1990	27.8	397.5	1.69
1991	29.9	413.3	1.52
1992	29.8	439.2	1.61
1993	30.8	459.7	1.58
1994	31.2	492.9	1.49
1995	33.3	528.6	1.52
1996	35.6	560.3	1.57
1997	36.4	624.6	1.51
1998	36.7	666.4	1.54
1999	38.4	717.8	1.6
2000	40.4	768.2	1.54
2001	40.3	843.3	1.59
2002	41.8	911.6	1.59
2003	40.4	931.1	2.08
2004	40.7	1021.5	1.96
2005	40.1	1165.9	2.33
2006	42.7	1349.6	2.32
2007	44.1	1449.4	2.26
2008	46.7	1575.5	2.55
2009	50.6	1759.1	2.46
2010	50.1	1994.2	2.36
2011	51.7	2258.1	2.66
2012	52.9	2478.7	2.82

Example 3.2.1. Adding new data.

More variables.

The prices of substitute goods, such as pork and beef, could be relevant to explain chicken consumption. Let's denote:

X_4 : price of pork (in euros/kilogram).

X_5 : price of beef (in euros/kilogram).

Furthermore, the avian flu epidemic that lasted from 1999 to 2003 (both years included) may have affected as well the temporal evolution of chicken consumption. A binary variable has been defined to quantify this qualitative information:

X_6 : = 1 if no avian flu epidemic, = 2 if avian flu epidemic

Example 3.2.1. Adding new data. Variables.

The table below shows data on variables $X4$, $X5$ and $X6$.

year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
X4	2.028	2.08	2.16	2.212	2.188	2.548	2.792	2.636	2.58	2.8	2.928
X5	3.132	3.168	3.168	3.168	3.096	3.208	3.216	3.356	3.42	3.748	4.244
X6	1	1	1	1	1	1	1	1	1	2	2

year	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
X4	2.712	3.164	3.816	3.768	4.94	5.196	4.704	5.236	5.192	5.12	5.64	6.728
X5	4.192	4.56	4.964	5.104	5.716	5.744	5.568	6.62	8.132	8.784	8.864	9.304
X6	2	2	2	1	1	1	1	1	1	1	1	1

Example 3.2.1. Adding new data. Variables.

Questions.

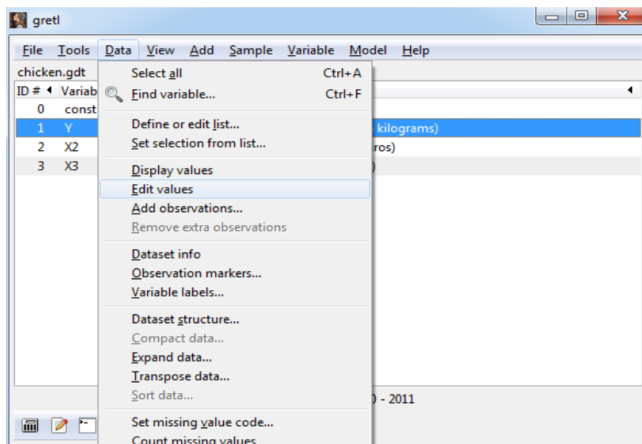
- Enter data on variables $X4$, $X5$ and $X6$ in the file `chicken.gdt` manually.
- Label the new variables.
- Define $X6$ as a discrete variable.
- Write this information about the data file in the Gretl session:

Data come from file Table 7.9. (Gujarati sample folder).

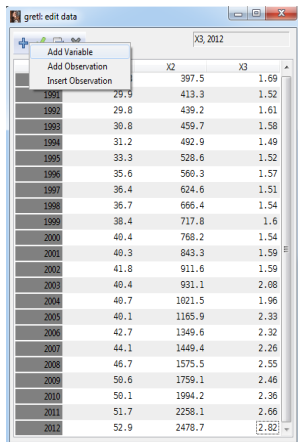
Example 3.2.1. Adding new data. Variables.

To **add** more variables, select the option *Edit values* from the **Data** pulldown menu.

Data -> Edit values



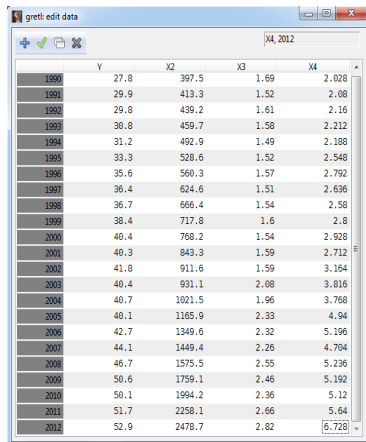
Example 3.2.1. Add new data. Variables.



The screenshot shows the 'gretl: edit data' window with a menu open over the '+' icon. The menu options are 'Add Variable', 'Add Observation', and 'Insert Observation'. The data table has columns X2, X3, and X4. The year 2012 is highlighted in the first column.

	X2	X3	X4
1990	29.9	413.3	1.52
1991	29.8	439.2	1.61
1992	30.8	459.7	1.58
1993	31.2	492.9	1.49
1994	33.3	528.6	1.52
1995	35.6	560.3	1.57
1996	36.4	624.6	1.51
1997	36.7	666.4	1.54
1998	38.4	717.8	1.6
1999	40.4	768.2	1.54
2000	40.3	843.3	1.59
2001	41.8	911.6	1.59
2002	40.4	931.1	2.08
2003	40.7	1021.5	1.96
2004	40.1	1165.9	2.33
2005	42.7	1349.6	2.32
2006	44.1	1449.4	2.26
2007	46.7	1575.5	2.55
2008	50.6	1759.1	2.46
2009	50.1	1994.2	2.36
2010	51.7	2258.1	2.66
2011	52.9	2478.7	2.82
2012			

Click on the icon 
and select **Add Variable**



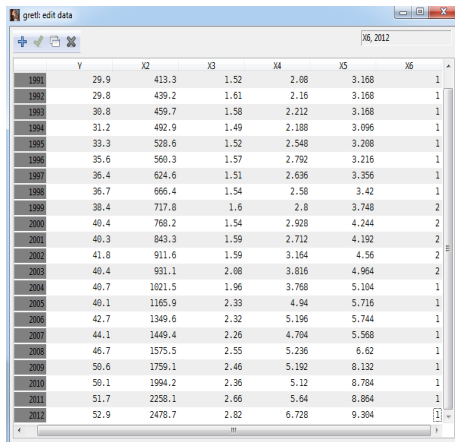
The screenshot shows the 'gretl: edit data' window with the 'Add Variable' menu option selected. The data table now has columns Y, X2, X3, and X4. The year 2012 is highlighted in the first column.

	Y	X2	X3	X4
1990	27.8	397.5	1.69	2.028
1991	29.9	413.3	1.52	2.08
1992	29.8	439.2	1.61	2.16
1993	30.8	459.7	1.58	2.212
1994	31.2	492.9	1.49	2.188
1995	33.3	528.6	1.52	2.548
1996	35.6	560.3	1.57	2.792
1997	36.4	624.6	1.51	2.636
1998	36.7	666.4	1.54	2.58
1999	38.4	717.8	1.6	2.8
2000	40.4	768.2	1.54	2.928
2001	40.3	843.3	1.59	2.712
2002	41.8	911.6	1.59	3.164
2003	40.4	931.1	2.08	3.816
2004	40.7	1021.5	1.96	3.768
2005	40.1	1165.9	2.33	4.94
2006	42.7	1349.6	2.32	5.196
2007	44.1	1449.4	2.26	4.704
2008	46.7	1575.5	2.55	5.236
2009	50.6	1759.1	2.46	5.192
2010	50.1	1994.2	2.36	5.12
2011	51.7	2258.1	2.66	5.64
2012	52.9	2478.7	2.82	6.728

Once the name of the variable is written,
enter the data

Repeat this procedure for variables X5 and X6.

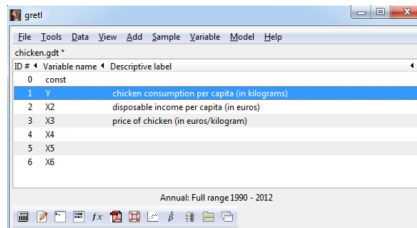
Example 3.2.1. Adding new data. Variables.



The screenshot shows the 'gretl: edit data' window. The title bar says 'gretl: edit data'. The window contains a table with columns labeled Y, X2, X3, X4, X5, and X6. The rows represent years from 1991 to 2012. The data is as follows:

	Y	X2	X3	X4	X5	X6
1991	29.9	413.3	1.52	2.08	3.168	1
1992	29.8	439.2	1.61	2.16	3.168	1
1993	30.8	459.7	1.58	2.212	3.168	1
1994	31.2	492.9	1.49	2.188	3.096	1
1995	33.3	528.6	1.52	2.548	3.208	1
1996	35.6	560.3	1.57	2.792	3.216	1
1997	36.4	624.6	1.51	2.636	3.356	1
1998	36.7	666.4	1.54	2.58	3.42	1
1999	38.4	717.8	1.6	2.8	3.748	2
2000	40.4	768.2	1.54	2.928	4.244	2
2001	40.3	843.3	1.59	2.712	4.192	2
2002	41.8	911.6	1.59	3.164	4.56	2
2003	40.4	931.1	2.08	3.816	4.964	2
2004	40.7	1021.5	1.96	3.768	5.104	1
2005	40.1	1165.9	2.33	4.94	5.716	1
2006	42.7	1349.6	2.32	5.196	5.744	1
2007	44.1	1449.4	2.26	4.704	5.568	1
2008	46.7	1575.5	2.55	5.236	6.62	1
2009	50.6	1759.1	2.46	5.192	8.132	1
2010	50.1	1994.2	2.36	5.12	8.784	1
2011	51.7	2258.1	2.66	5.64	8.864	1
2012	52.9	2478.7	2.82	6.728	9.304	1

Once the data are entered,
click on the icon *Apply*



The screenshot shows the main Gretl window. The title bar says 'gretl'. The menu bar includes File, Tools, Data, View, Add, Sample, Variable, Model, and Help. The main area shows a list of variables:

ID #	Variable name	Descriptive label
0	const	
1	Y	chicken consumption per capita (in kilograms)
2	X2	disposable income per capita (in euros)
3	X3	price of chicken (in euros/kilogram)
4	X4	
5	X5	
6	X6	

At the bottom, it says 'Annual: Full range 1990 - 2012'.

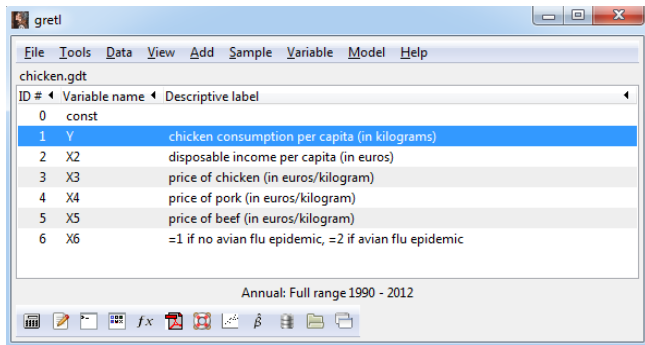
The main window shows the three new
variables without any descriptive label

Example 3.2.1. Adding new data. Variables.

To [edit](#) the variable characteristics, highlight the variable, right-click and select the option *Edit attributes* from the pulldown menu.

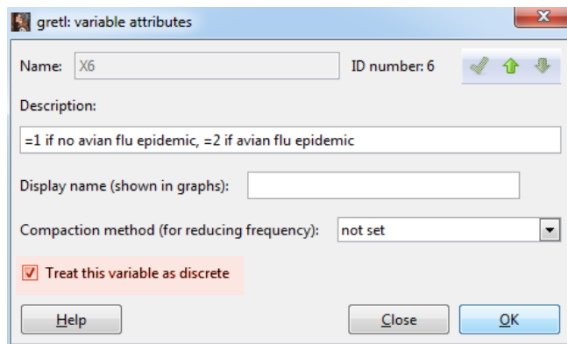
Then follow the procedure explained in [Example 3.1.3](#).

The result is shown in the figure below.



Example 3.2.1. Adding new data. Variables.

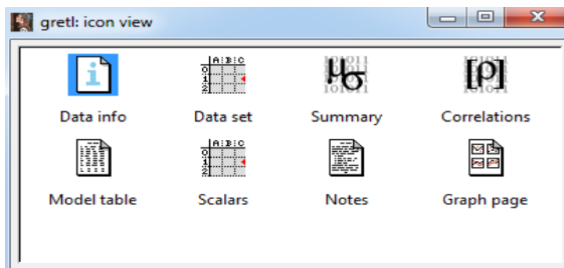
To define the variable X_6 as a **discrete variable**, highlight the variable X_6 , right-click and select the option *Edit Attributes* from the pulldown menu. Then, mark Treat this variable as discrete.



Don't forget to save all the changes in
the file `chicken.gdt` before quitting Gretl !

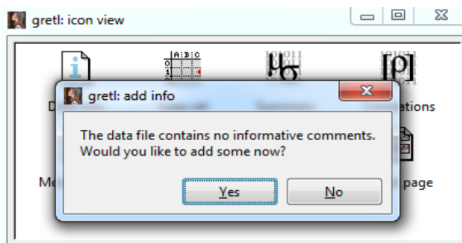
Example 3.2.1. Adding new data. Variables.

To **enter and save information** about the data file, click on the icon *session icon* view in the main window Toolbar. Then, double-click on the icon *Data Info*.



Example 3.2.1. Adding new data. Variables.

Given that there is no previous information about this data file, you are asked whether you want to add some.



Clicking **Yes** brings up an empty window where you can write some notes. To save this information click on the left icon (*Save*) in the menu bar of the **edit data info** window.



- 1 3.2.1. Adding new data.
 - More observations.
 - Adding variables.
- 2 3.2.2. Importing data from other files.
- 3 3.2.3. Generating new variables.
- 4 3.2.4. Simulating variables.

Example 3.2.2. Importing data from other files.

Visitors to Bilbao.

Consider that both the Hospitality Price Index and the economic situation may be relevant factors to explain the temporal evolution of the number of visitors to Bilbao.

The Excel data file `tourism.xls` (sheet called `Other variables`) contains information on:

HPIR: monthly variation rate of the Hospitality Price Index in Biscay.

IPI: Spanish Industrial Production Index (raw data).

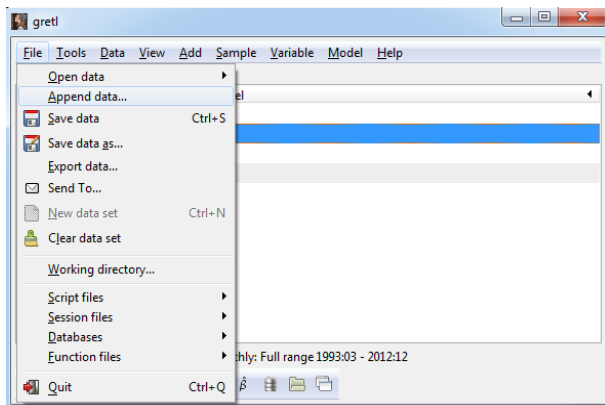
IPIR: annual variation rate of Spanish IPI.

Add these data to the Gretl file `tourism.gdt`. Label the new variables and save all the changes in the data file `tourism.gdt`.

Example 3.2.2. Importing data from other files.

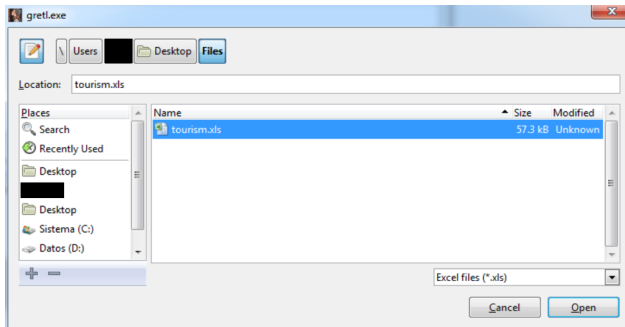
You may **import** data from an Excel file to an existing Gretl file, selecting the option *Append data* from the **File** pulldown menu.

File -> Append data...



Example 3.2.2. Importing data from other files.

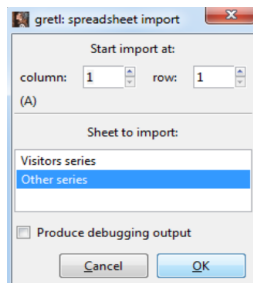
Select the folder where your Excel file is (in this case, the Desktop), choose the file `turismo.xls` where the data are and open it using the **Open** button at the bottom of the window.



Example 3.2.2. Importing data from other files.

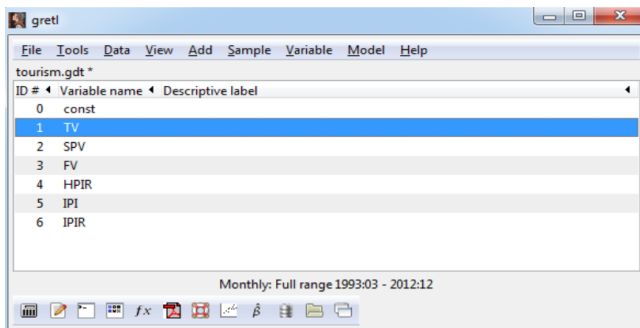
From this dialog box, select:

- the sheet of the Excel file where the data are (Other variables).
- the column and the row where the data start (including the row with the names of the variables).



Example 3.2.2. Importing data from other files.

The main window shows the three new variables imported.

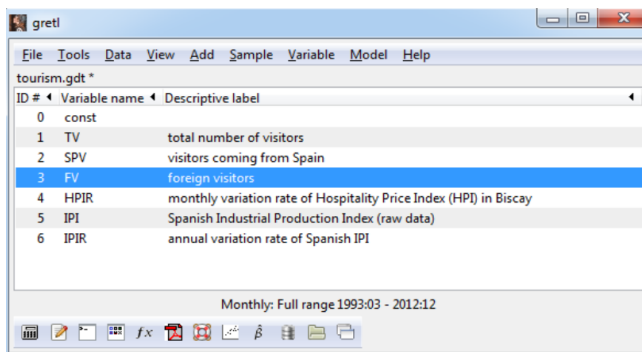


Example 3.2.2. Importing data from other files.

To **edit** the characteristics of the variables, select a variable, go up to the Menu Bar and click **Variable -> Edit attributes**.

Replicate the procedure explained in **Example 3.1.3**.

The result is shown in the figure below.



Don't forget to save all the changes in
the file **tourism.gdt** before quitting Gretl!

- 1 3.2.1. Adding new data.
 - More observations.
 - Adding variables.
- 2 3.2.2. Importing data from other files.
- 3 3.2.3. Generating new variables.
- 4 3.2.4. Simulating variables.

Example 3.2.3. Generating new variables.

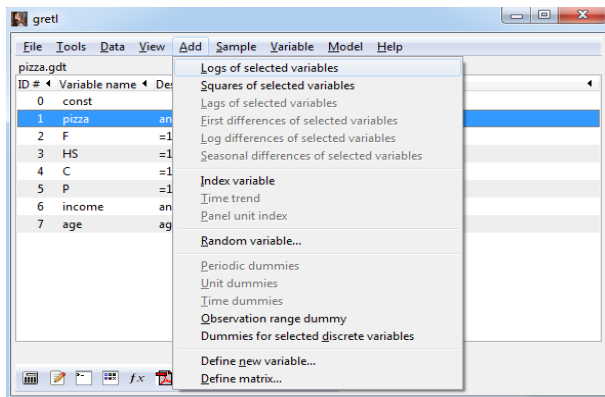
Pizza consumption.

1. Load your file `pizza.gdt`.
2. Add the logarithm of pizza consumption.
3. Add the squares of income and age.
4. Add the variables $FI = (F \times \text{income})$ and $IA = (\text{income} \times \text{age})$.
5. Add the variable pizza consumption measured in thousands of dollars instead of in dollars (`pizzaT`).
6. Save all the changes in your Gretl file `pizza.gdt`.

Example 3.2.3. Generating new variables.

To **add** logarithms, highlight the desired variables, and select the option *Logs of selected variables* from the **Add** pulldown menu.

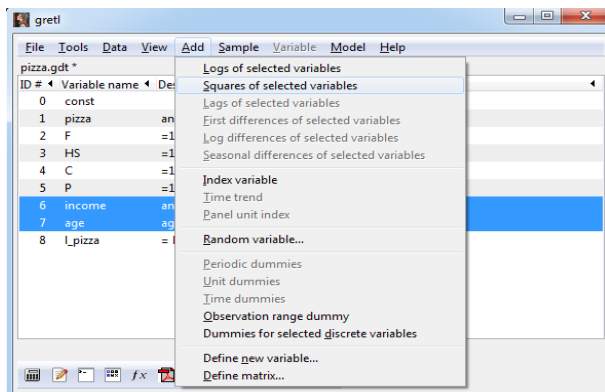
Add -> Logs of selected variables



Example 3.2.3. Generating new variables.

To **add** squares, highlight the desired variables and select

Add -> Squares of selected variables

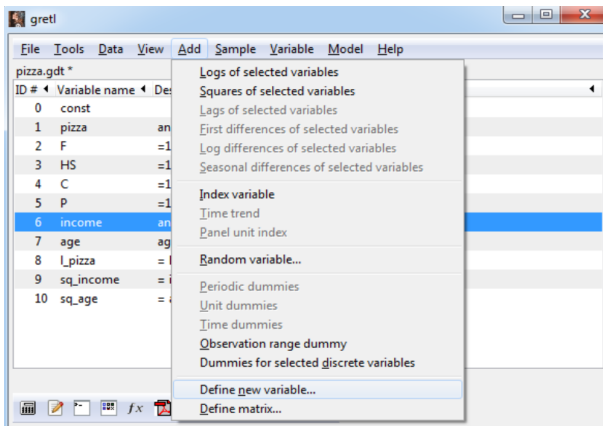


Note that the logarithm of pizza consumption, generated previously, is shown in the main window with the name assigned by default, l_pizza.

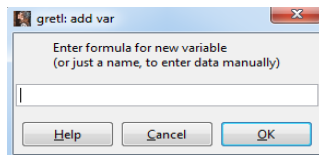
Example 3.2.3. Generating new variables.

To **generate** a new variable from others included in the data file, go up to the Menu Bar and click

Add -> Define new variable...



Example 3.2.3. Generating new variables.

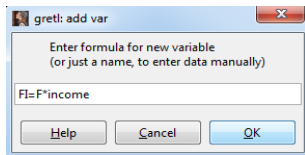


Write down the formula for the new variable in the dialog box using the following operators:

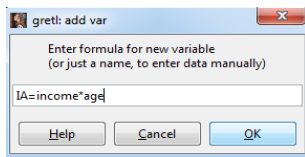
Operators

Addition	+	Subtraction	-
Multiplication	*	Division	/
Square root	sqrt(x)	Exponent	**
Logarithms	log(x)	Exponential (e^x)	exp(x)

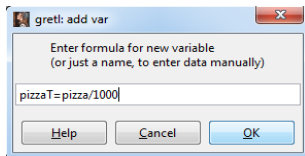
Example 3.2.3. Generating new variables.



Variable FI



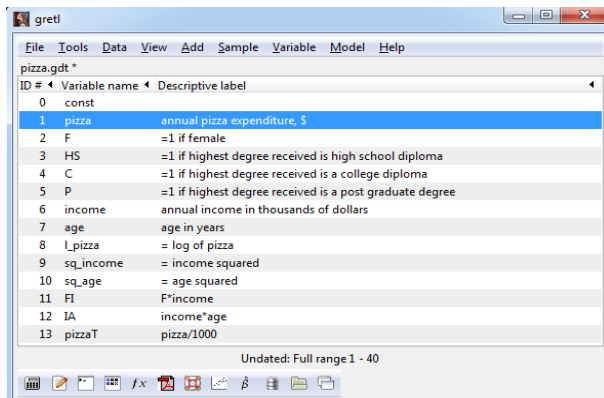
Variable IA



Variable pizzaT (thousands of dollars)

Example 3.2.3. Generating new variables.

The main window shows the three new variables with the names assigned in the formulas. The label of each variable is given by the formula used to generate it and it may be changed editing the attributes.



jj Do no forget to save all the changes in
the data file pizza.gdt before quitting Gretl!

- 1 3.2.1. Adding new data.
 - More observations.
 - Adding variables.
- 2 3.2.2. Importing data from other files.
- 3 3.2.3. Generating new variables.
- 4 3.2.4. Simulating variables.

Example 3.2.4. Simulating variables.

New data file.

1. Create a new data set of 500 observations (cross-section).
2. Simulate u from a normal distribution of mean 0 and variance 16.
3. Simulate X_2 from a uniform distribution between 20 and 50, and X_3 from a Poisson distribution of mean 3.
4. Simulate variable Y as follows:

$$Y = -4 + X_2 - 5X_3 + u$$

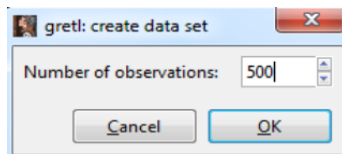
5. Write down some notes about the simulation of the data set and save them in the session.
6. Save the data in a Gretl file called `simulate.gdt`.

Example 3.2.4. Simulating variables.

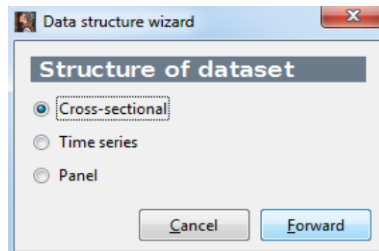
To **create** a new data set, go up to the Menu Bar and click

File -> New data set

From the dialog windows shown below, you tell Gret the number of observations and the structure of the data set.



Write the number of observations

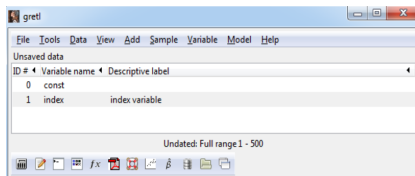


Select the type of data

Example 3.2.4. Simulating variables.

Confirm the characteristics of the file (number of observations and type of data), clicking **Apply**. DO NOT mark start entering data values because you are not going to enter any data. The data will be artificially generated.

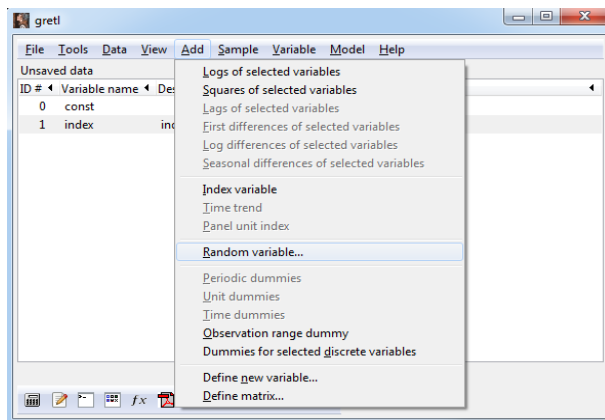
The main window shows only an index that goes from 1 to 500.



Example 3.2.4. Simulating variables.

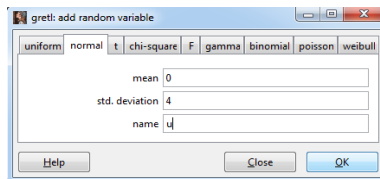
To **simulate** a random sample from a specific distribution, select the option *Random variable* from the **Add** pulldown menu.

Add -> Random variable...



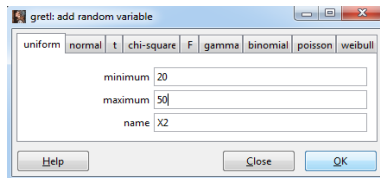
Example 3.2.4. Simulating variables.

Select the normal distribution tab, write down its parameters (mean = 0, std. deviation = 4) and give a name to the new variable (u).



The screenshot shows the 'gretl: add random variable' dialog box. The 'normal' tab is selected. The 'mean' field is set to 0, the 'std. deviation' field is set to 4, and the 'name' field is set to u . The 'OK' button is highlighted.

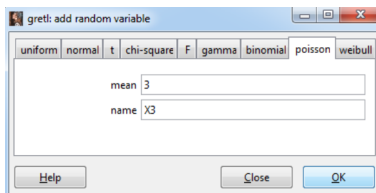
Select the uniform distribution tab, write down its parameters (minimum = 20, maximum = 50) and give a name to the new variable ($X2$).



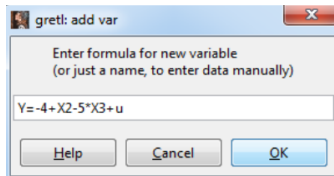
The screenshot shows the 'gretl: add random variable' dialog box. The 'uniform' tab is selected. The 'minimum' field is set to 20, the 'maximum' field is set to 50, and the 'name' field is set to $X2$. The 'OK' button is highlighted.

Example 3.2.4. Simulating variables.

Select the Poisson distribution tab, write down its parameter (mean = 3) and give a name to the new variable (X_3).

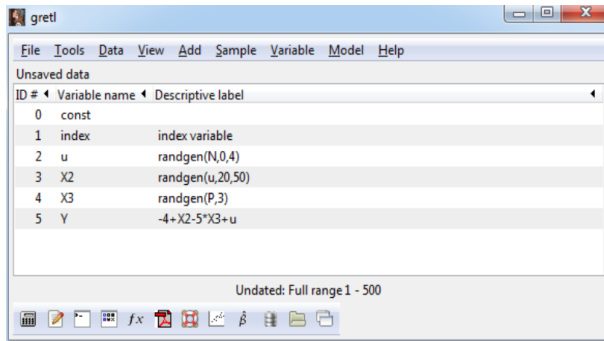


To generate variable Y , go up to the Menu Bar, select **Add -> Define new variable** and write down the formula.



Example 3.2.4. Simulating variables.

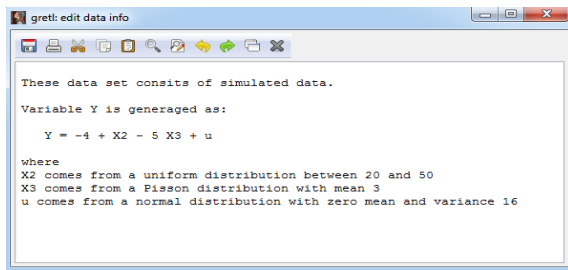
The main window shows all the simulated variables with their assigned names. The label of the variables is given by the distribution they come from or the formula used to generate them.



Example 3.2.4. Simulating variables.

Click on the icon *session icon view* in the main window Toolbar. Then click on the icon *Data info*.

Write down the information in the empty window and then click on the left icon (*Save*).



Example 3.2.4. Simulating variables.

To save these data in a Gretl file, click

File -> Save data as ...

Select the folder to save the file (in this case, the Desktop), write the name of the file, `simulate.gdt`, and click **Save**.

