

Introduction to Gretl

Social Science Research Lab

American University, Washington, D.C.

Web. www.american.edu/provost/ctrl/pclabs.cfm

Tel. x3862 Email. SSRL@American.edu



Course Objective

This course is designed to provide students with the basic tools to work with data using the open source package gretl.

Learning Outcomes

In this tutorial you will learn how to

- Write a script file.
- Import native-type data sets and several other types of data sets.
- Explore your data.
- Run basic statistical tests.
- Run OLS regressions.
- Create graphs and plots.

Installing gretl

Go to <http://gretl.sourceforge.net/> with your browser and click the link “**gretl for Windows**” in the left column. You can download an executable file named **gretl-x.x.x.exe**, “x.x.x.”, its numbers representing the version of gretl. For the rest of this tutorial, we will use the version 1.8.4. Once the file is downloaded, double click the file and follow the instructions.

Once gretl is installed, you can find the gretl folder in All Programs. (In the CTRL computer labs gretl is located under the “math and stat applications” folder). Click “gretl”, which has the icon of a girl, and the software will launch.

Getting started with gretl

Gretl is an open-source statistical package for econometrics. It can be used with other analytical packages such as R. To get a feeling of gretl, let’s try some simple tasks using graphical user interface. (GUI).

Go to:

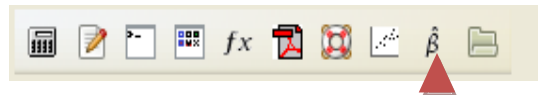
File > Open Data

Then click "**Sample Files**" with a folder icon. You will find many sample data sets. Click "**Greene**" tab, and choose the first file "**greene5_1**", which contains the



US macro data. In the main window, the variable names are listed with their descriptions. Double click "**realgdp**", and you can browse the variable. Click an icon of a small graph, and you can produce a time plot of this variable.

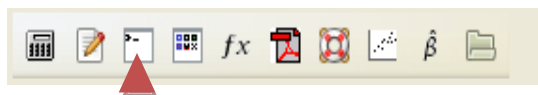
Go back to the main window. Click the " **β** " icon at the bottom to estimate a simple OLS model. Highlight a variable by clicking and use the arrow bottom to specify the independent/dependent variables. (See the figure)



Hit "OK", and you will see the output window corresponding to the regression model.

Using the Console and Command Lines

Instead of clicking the menus and buttons, you can use gretl by entering commands in the console. The console is accessible from the button at the bottom of the main window. (See the figure.)



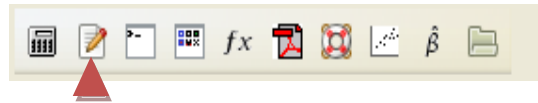
For example, in the console, type "help adf" and hit the enter key. The help menu for the command "adf" will show up.

help adf

Tip: The console is a convenient way to use gretl, but it is hard to reproduce the same result as it does not record your commands.

Why and how to start writing a script?

Script is a text file that contains commands. It is common to use a script file, instead of mouse clicks as in the previous section. There are various advantages of using script file, but the most important is that it enables you to reproduce the result from the original data set.



To create a new script file you can click the second icon at the bottom of the main window. (See the figure above) or go to

File > Script files > New Script and click “gretl script”

Once you write the commands, you can execute the script by hitting the button **Run**.

Edit the script, run it, see the result, and edit it. This is a common way of working with gretl. The script has an extension of “.inp”

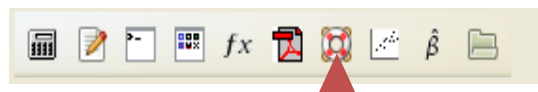
Tip: Read Ch.3 of the manual for more about the available modes.

Getting Help in Gretl: accessing the manual and references

The manual is accessible from the “pdf” icon at the bottom of the main window.



The Command Reference is accesible from the icon right to the manual icon.



The command “**help**” displays the list of commands.

Getting the data into gretl

To open a gretl native data set (**.gdt**), use “**open**” command. For example:
open greene7_8

The file needs to be placed in the “**working directory**” or “**gretl/data**” folder.

Tip: For Windows users, this is the “gretl” folder under your My Documents folder.

Alternatively, you can change the working directory.

Go to

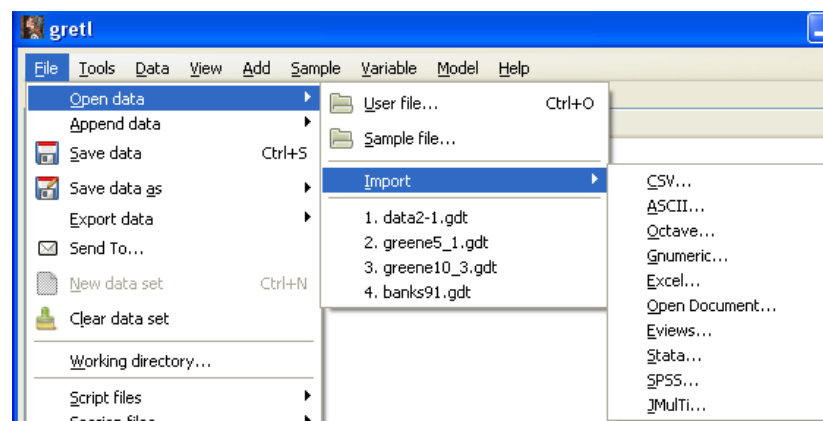
File > Working directory

and select your choice of folder. You can check the current working folder with “**workdir**” command.

To import any foreign data type, you need to do so manually through GUI. Go to

File > Open data > Import

Then choose the file type.



This illustration shows the various data types Gretl supports. For practice, we will import an Excel file.

Tip: This tutorial uses the file *Intro_gretl.xls*, which is available on our website.

Now go back to the gretl main window.

Go to

File > Open data > Import

choose Excel..., then open the excel file. Gretl will ask you the start point of the data, choose “**column 1 and row 1**” as a starting point.

gretl tries to configure the format. Confirm that the way gretl interpreted the data is correct.

If not configured correctly, you need to do so manually. In the main window, go to

Data > dataset structure

Next specify the frequency and starting date (year, quarter, month and so on).

To save the data as a gretl native data set, use “**store**” command.

```
store greene5_1.gdt
```

The data will be stored in the current working folder.

Tip: Read Ch.4 of the manual for more about the data types.

Getting descriptive statistics

Using a console (or a script) we are going to obtain descriptive statistics.

Open “**data2-1**” (SAT scores) from the sample files:

```
open data2-1
```

Use “**varlist**” to list up the variable contained in this dataset:

```
varlist
```

Use “**summary**” command to obtain the descriptive statistics. You can also specify the variable.

```
summary  
summary vsat
```

Use “**corr**” command to obtain the correlation of two variables.

```
corr vsat msat
```

Sample code1:

```
open data2-1  
varlist  
summary  
summary vsat
```

Selecting sub-groups

Open “**greene5_1**” (U.S. Macro data for 50 years) in the sample files. This is a time-series dataset.

```
open greene5_1
```

The data is quarterly recorded. We can select subsample using “**smp1**” command.

```
smp1 1950:1 1973:1
summary cpi_u
```

To select the entire sample, use the key word “**full**”.

```
smp1 full
summary cpi_u
```

Tip: For more information refer to “Dataset” in the Command Reference

Creating variables

The command “**genr**” defines a new variable with the indicated mathematical expressions, below an example:

```
genr d_realgdp = realgdp - realgdp(-1)
```

*Tip: You can generate a n-log variable by using the following parenthesis **varname(-n)***

Alternatively, built-in functions are convenient for common variable transformation such as taking the differences, lags or logarithms of a variable.

```
diff realgdp
lags realgdp
logs realgdp
```

Tip: For more information refer to “Dataset” and “Transformation” in the Command Reference.

Testing for unit root

Use “**adf**” command to test if the variable has a unit root. It computes Dickey-Fuller tests: the null hypothesis is the variable has a unit root.

```
adf 0 realgdp
```

Using **-difference** option, you can test the first difference of the variable:

```
adf 0 realgdp -difference
adf 0 d_realgdp
```

Tip: For more information refer to “Tests” in the Command Reference.

```

Sample Code: jgm-1996.inp
open jgm-data.gdt
smp1 1955 ;
# unit root tests for short-term interest rate
adf 0 r_s -c                # With constant
adf 0 r_s -ctt              # With constant and
quadratic trend
adf 1 r_s -ctt              # With one lag

```

Drawing graphs and plots

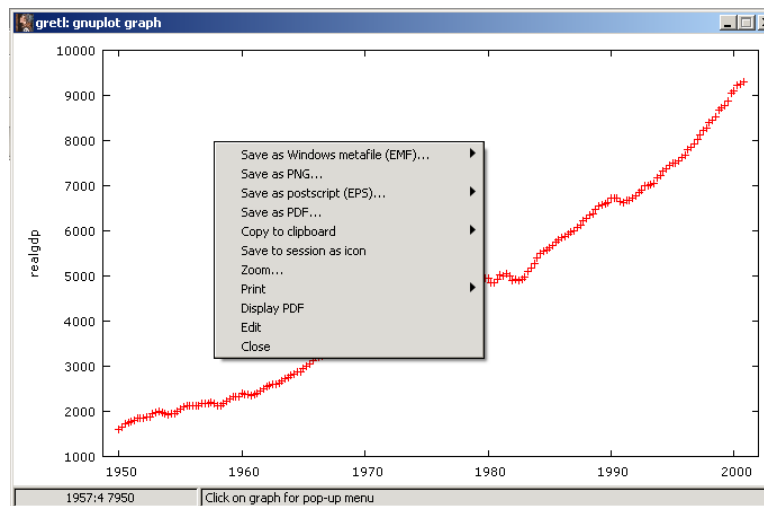
The command “**gnuplot**” provides various forms of plots such as scatter plot between two variables and plot against time (time plot).

```
gnuplot realgdp time
```

```
gnuplot realgdp --time-series
```

You can export the plotted graph by two ways: i) right click on the graph and choose the option “**Save as PNG**” (See the figure below), ii) use “**--output**” option of the “gnuplot”.

```
gnuplot realgdp --output=plot1.png
```

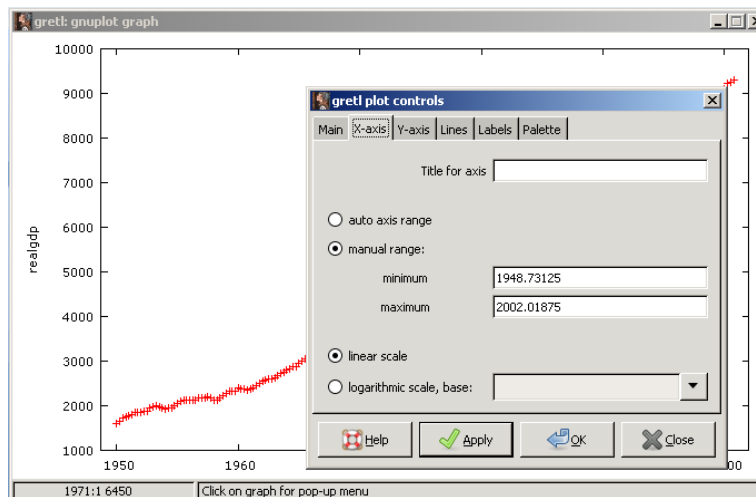


By only specifying the file name, the PNG file will be saved in the currently working folder. Alternatively, you can specify the folder to save the file by providing the full path:

```
gnuplot realgdp --output=G:/gretl/plot1.png
```

Furthermore, you can have more control of the plot either i) by choosing “**Edit**” in right click menu (See the figure, below), or ii) by using gnuplot commands. The command needs to be enclosed within braces, and each command needs to be followed by semi-colons. For example, you can set your title and axis names, with “**set title**” and “**set xlabel**” commands, as follows;

```
gnuplot realgdp time --output=plot1.png {set title 'My Title';  
set xlabel "Year"; set grid;}
```



Tip: For more information about gnuplot commands, refer to <http://www.gnuplot.info/> and its “Documentation” link.

To plot the ACF of the variable, use “**corrgram**” command.

```
corrgram realgdp
```

“**freq**” command displays the histogram and the frequency table.

```
Freqpop
```



```

# Sample code 2:
open data2-1
gnuplot vsat msat

open greene5_1
plot1 <- gnuplot realgdp time      # Record a time plot with a
                                   name (stored in the
                                   session)
plot2 <- gnuplot realgdp time --with-lines
                                   # Use line instead of marks
gnuplot realgdp time --output=plot1.png
                                   # Save a time plot as a
                                   .png file (under the
                                   working directory)
gnuplot realgdp time --output=plot1.png \
    {set title 'My Title'; set xlabel "Year"; \
    set ylabel "Real GDP";} # Write multiple lines with
                             backslashes ("\")
corrgm realgdp                  # Print the autocorrelation
                                function
freq pop                        # Print the histogram
freq pop --normal --quiet      # Test for normal
distribution, Suppressing graph

```

Tip: For more information refer to “Graphics” and “Statistics” in the Command Reference.

Estimate models

Here, vanilla ordinary least square (OLS) model estimation and autoregressive moving average (ARMA) model are explained, with “**ols**” and “**arima**” commands respectively:

For illustration, you can estimate with OLS with “**ols**” command when using the data2-1, SAT data set.

```
ols vsat msat
```

Alternatively, you can specify the variables by the id numbers:

```
ols 0 1
```

Some keywords following \$ are available for the post-estimation use:

```
genr ui = $uhat
```

```
genr coeff = $beta
```

```
# Sample Code: mrw.inp
# Solow model," in Mankiw, Romer and Weil, QJE 1992
open mrw.gdt
genr lny = log(gdp85)
genr ngd = 0.05 + (popgrow/100.0)
genr lngd = log(ngd)
genr linv = log(inv/100.0)
genr x3 = linv - lngd           # generate variable for
                                # testing Solow restriction
smp1 nonoil -dummy             # set sample to non-oil
                                # producing countries
modell1 <- ols lny const linv lngd # Recording estimation
                                # result by <-
genr essu = $ess               # Estimation result stored
                                # by specific keywords
genr dfu1 = $df
```

Autoregressive model (**AR**(1)) on the real GDP is estimated with “**arima**” command when using the dataset `greene5_1`

```
arima 1 1 0; realgdp
arima 1 0 0; d_realgdp
```

As you can see, the numbers represent the autoregressive order, the difference order, and the moving average order respectively. Independent variables may follow the independent variable. For example, the following is the **AR**(1) , **MA**(1) model on the first difference of real GDP.

```
arima 1 1 1; realgdp
```

(Note: For more information refer to “Estimation” in the Command Reference.)

Tip: *gretl* provides various “practice files”. **Go to Files > Script files > Practice files**