
Software Requirements Specification

for

Gretl

Version 1.0 approved

Prepared by Ioannis Spyridonidis 2805

Aristotle University of Thessaloniki

January 2019

Table of Contents

Table of Contents	ii
Revision History	ii
1. Introduction.....	1
1.1 Purpose	1
1.2 Document Conventions.....	1
1.3 Intended Audience and Reading Suggestions.....	1
1.4 Product Scope	1
1.5 References.....	2
2. Overall Description.....	3
2.1 Product Perspective	3
2.2 Product Functions	3
2.3 User Classes and Characteristics	9
2.4 Operating Environment.....	9
2.5 Design and Implementation Constraints.....	9
2.6 User Documentation	9
2.7 Assumptions and Dependencies	9
3. External Interface Requirements	10
3.1 User Interfaces	10
3.2 Hardware Interfaces.....	15
3.3 Software Interfaces	15
3.4 Communications Interfaces	15
4. System Features.....	16
4.1 Basic Statistics	16
4.2 Estimators	18
4.3 Limited dependent variables.....	19
4.4 Time series methods.....	20
5. Other Nonfunctional Requirements.....	21
5.1 Performance Requirements.....	21
5.2 Safety Requirements	21
5.3 Security Requirements.....	21
5.4 Software Quality Attributes.....	21

Revision History

Name	Date	Version
gretl	2002-08-25	0.9
gretl	2002-11-15	1.0
gretl	2003-07-23	1.1.0
gretl	2003-12-14	1.2.0
gretl	2004-11-15	1.3.0
gretl	2005-06-30	1.4.1
gretl	2005-12-16	1.5.0

gretl	2006-09-13	1.6.0
gretl	2007-11-28	1.7.0
gretl	2009-01-24	1.8.0
gretl	2010-05-02	1.9.0
gretl	2015-04-02	1.10.0
gretl	2015-10-19	2015d
gretl	2016-03-10	2016a
gretl	2017-05-26	2017b
gretl	2018-06-20	2018a
gretl	2018-08-11	2018b
gretl	2018-09-03	2018c
gretl	2018-12-21	2018d

1. Introduction

1.1 Purpose

The purpose of this document is to present a detailed description of the open-source software Gretl. It will explain the purpose and features of the software, the interfaces of the software, what the software will do and the constraints under which it must operate. This document is intended for users of the software and also potential developers.

1.2 Document Conventions

This Document was created based on the IEEE template for System Requirement Specification Documents.

1.3 Intended Audience and Reading Suggestions

- Typical Users, such as students, who want to use Gretl as a basic econometrics tool.
- Advanced/Professional Users, such as economists or researchers, who want to use Gretl for more demanding graph analysis.
- Programmers who are interested in working on the project by further developing it or fix existing bugs

1.4 Product Scope

Gretl is a cross-platform software package for econometric analysis; it is very easy to get up and running with econometric analysis. Thanks to its association with the econometrics textbooks by Ramu Ramanathan, Jeffrey Wooldridge, and James Stock and Mark Watson, the package offers many practice data files and command scripts. These are well annotated and accessible. Two other useful resources for gretl users are the available documentation and the gretl-users mailing list.

1.5 References

Gretl's website:

<http://gretl.sourceforge.net/>

Gretl Conferences

gretlconference.org

Sourceforge link of project

<https://sourceforge.net/projects/gretl/>

Gretl twitter for updates

https://twitter.com/gretl_stats

Git access and browsing

<https://sourceforge.net/p/gretl/git/ci/master/tree/>

2. Overall Description

2.1 Product Perspective

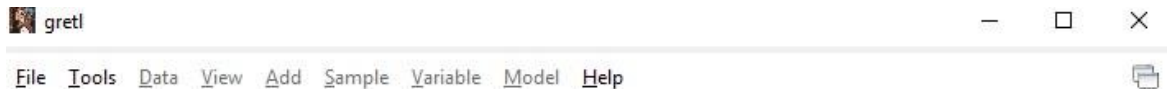
Gretl is an econometrics package, including a shared library, a command-line client program and a graphical user interface.

You can choose your preferred point on the spectrum from interactive point-and-click to complex scripting, and can easily combine these approaches.

Users can enhance gretl by writing their own functions and procedures in gretl's scripting language, which includes a wide range of matrix functions.

It is an open source project and it has a very active developer team to support it and provide feedback to users. It was developed to run on Windows, Mac OS X and Linux.

2.2 Product Functions



- File menu
 - Open data: Open a native gretl data file or import from other formats.
 - Append data: Add data to the current working data set, from a gretl data file, a comma-separated values file or a spreadsheet file.
 - Save data: Save the currently open native gretl data file.
 - Save data as: Write out the current data set in native format, with the option of using gzip data compression
 - Export data: Write out the current data set in Comma Separated Values (CSV) format, or the formats of GNU R or GNU Octave.
 - Send to: Send the current data set as an e-mail attachment.
 - New data set: Allows you to create a blank data set, ready for typing in values or for importing series from a database. See below for more on databases.
 - Clear data set: Clear the current data set out of memory. Generally you don't have to do this (since opening a new data file automatically clears the old one) but sometimes it's useful.
 - Script files: A "script" is a file containing a sequence of gretl commands. This item contains entries that let you open a script you have created previously ("User file"), open a sample script, or open an editor window in which you can create a new script.

- Session files: A “session” file contains a snapshot of a previous gretl session, including the data set used and any models or graphs that you saved. Under this item you can open a saved session or save the current session.
- Databases: Allows you to browse various large databases, either on your own computer or, if you are connected to the internet, on the gretl database server.
- Exit: Quit the program. You’ll be prompted to save any unsaved work.
- Tools menu
 - Statistical tables: Look up critical values for commonly used distributions (normal or Gaussian, t, chi-square, F and Durbin–Watson).
 - P-value finder: Look up p-values from the Gaussian, t, chi-square, F, gamma, binomial or Poisson distributions. See also the pvalue command in the Gretl Command Reference.
 - Distribution graphs: Produce graphs of various probability distributions. In the resulting graph window, the pop-up menu includes an item “Add another curve”, which enables you to superimpose a further plot (for example, you can draw the t distribution with various different degrees of freedom).
 - Test statistic calculator: Calculate test statistics and p-values for a range of common hypothesis tests (population mean, variance and proportion; difference of means, variances and proportions).
 - Nonparametric tests: Calculate test statistics for various nonparametric tests (Sign test, Wilcoxon rank sum test, Wilcoxon signed rank test, Runs test).
 - Seed for random numbers: Set the seed for the random number generator (by default this is set based on the system time when the program is started).
 - Command log: Open a window containing a record of the commands executed so far.
 - Gretl console: Open a “console” window into which you can type commands as you would using the command-line program, gretlcli (as opposed to using point-and-click).
 - Start Gnu R: Start R (if it is installed on your system), and load a copy of the data set currently open in gretl.
 - Sort variables: Rearrange the listing of variables in the main window, either by ID number or alphabetically by name.
 - Function packages: Handles “function packages” which allow you to access functions written by other users and share the ones written by you.

- NIST test suite: Check the numerical accuracy of gretl against the reference results for linear regression made available by the (US) National Institute of Standards and Technology.
- Preferences: Set the paths to various files gretl needs to access. Choose the font in which gretl displays text output. Activate or suppress gretl’s messaging about the availability of program updates, and so on.
- Data menu
 - Select all: Several menu items act upon those variables that are currently selected in the main window. This item lets you select all the variables.
 - Display values: Pops up a window with a simple (not editable) printout of the values of the selected variable or variables.
 - Edit values: Opens a spreadsheet window where you can edit the values of the selected variables.
 - Add observations: Gives a dialog box in which you can choose a number of observations to add at the end of the current dataset; for use with forecasting.
 - Remove extra observations: Active only if extra observations have been added automatically in the process of forecasting; deletes these extra observations.
 - Read info, Edit info: “Read info” just displays the summary information for the current data file; “Edit info” allows you to make changes to it (if you have permission to do so).
 - Print description: Opens a window containing a full account of the current dataset, including the summary information and any specific information on each of the variables.
 - Add case markers: Prompts for the name of a text file containing “case markers” (short strings identifying the individual observations) and adds this information to the data set.
 - Remove case markers: Active only if the dataset has case markers identifying the observations; removes these case markers.
 - Dataset structure: invokes a series of dialog boxes which allow you to change the structural interpretation of the current dataset. For example, if data were read in as a cross section you can get the program to interpret them as time series or as a panel.

- Compact data: For time-series data of higher than annual frequency, gives you the option of compacting the data to a lower frequency, using one of four compaction methods (average, sum, start of period or end of period).
- Expand data: For time-series data, gives you the option of expanding the data to a higher frequency.
- Transpose data: Turn each observation into a variable and vice versa (or in other words, each row of the data matrix becomes a column in the modified data matrix); can be useful with imported data that have been read in “sideways”.
- View menu
 - Icon view: Opens a window showing the content of the current session as a set of icons.
 - Graph specified vars: Gives a choice between a time series plot, a regular X–Y scatter plot, an X–Y plot using impulses (vertical bars), an X–Y plot “with factor separation” (i.e. with the points colored differently depending to the value of a given dummy variable), boxplots, and a 3-D graph. Serves up a dialog box where you specify the variables to graph.
 - Multiple graphs: Allows you to compose a set of up to six small graphs, either pairwise scatter-plots or time-series graphs. These are displayed together in a single window.
 - Summary statistics: Shows a full set of descriptive statistics for the variables selected in the main window.
 - Correlation matrix: Shows the pairwise correlation coefficients for the selected variables.
 - Cross Tabulation: Shows a cross-tabulation of the selected variables. This works only if at least two variables in the data set have been marked as discrete.
 - Principal components: Produces a Principal Components Analysis for the selected variables.
 - Mahalanobis distances: Computes the Mahalanobis distance of each observation from the centroid of the selected set of variables.
 - Cross-correlogram: Computes and graphs the cross-correlogram for two selected variables.

- Add menu
Offers various standard transformations of variables (logs, lags, squares, etc.) that you may wish to add to the data set. Also gives the option of adding random variables, and (for time-series data) adding seasonal dummy variables (e.g. quarterly dummy variables for quarterly data).
- Sample menu
 - Set range: Select a different starting and/or ending point for the current sample, within the range of data available.
 - Restore full range: self-explanatory.
 - Define, based on dummy: Given a dummy (indicator) variable with values 0 or 1, this drops from the current sample all observations for which the dummy variable has value 0.
 - Restrict, based on criterion: Similar to the item above, except that you don't need a predefined variable: you supply a Boolean expression (e.g. $\text{sqft} > 1400$) and the sample is restricted to observations satisfying that condition. See the entry for `genr` in the Gretl Command Reference for details on the Boolean operators that can be used.
 - Random sub-sample: Draw a random sample from the full dataset.
 - Drop all obs with missing values: Drop from the current sample all observations for which at least one variable has a missing value.
 - Count missing values: Give a report on observations where data values are missing. May be useful in examining a panel data set, where it's quite common to encounter missing values.
 - Set missing value code: Set a numerical value that will be interpreted as "missing" or "not available". This is intended for use with imported data, when `gretl` has not recognized the missing-value code used.
- Variable menu
Most items under here operate on a single variable at a time. The "active" variable is set by highlighting it (clicking on its row) in the main data window. Most options will be self-explanatory. Note that you can rename a variable and can edit its descriptive label under "Edit attributes". You can also "Define a new variable" via a formula (e.g. involving some function of one or more existing variables). For the syntax of such formulae, look at the online help for "Generate variable syntax" or see the `genr` command in the Gretl Command Reference. One simple example: $\text{foo} = \text{x1} * \text{x2}$ will create a new variable `foo` as the product of the existing variables `x1` and `x2`. In these formulae, variables must be referenced by name, not number.

- Model menu

For details on the various estimators offered under this menu please consult the Gretl Command Reference.

- Help menu

Please use this as needed! It gives details on the syntax required in various dialog entries.

The gretl toolbar



The icons have the following functions, reading from left to right:

1. Launch a calculator program. A convenience function in case you want quick access to a calculator when you're working in gretl. The default program is calc.exe under MS Windows, or xcalc under the X window system. You can change the program under the "Tools, Preferences, General" menu, "Programs" tab.
2. Start a new script. Opens an editor window in which you can type a series of commands to be sent to the program as a batch.
3. Open the gretl console. A shortcut to the "Gretl console" menu item (Section 2.3 above).
4. Open the session icon window.
5. Open a window displaying available gretl function packages.
6. Open this manual in PDF format.
7. Open the help item for script commands syntax (i.e. a listing with details of all available commands).
8. Open the dialog box for defining a graph.
9. Open the dialog box for estimating a model using ordinary least squares.
10. Open a window listing the sample datasets supplied with gretl, and any other data file collections that have been installed.

2.3 User Classes and Characteristics

- Typical Users, such as students, who want to use Gretl as a basic econometrics tool.
- Advanced/Professional Users, such as economists or researchers, who want to use Gretl for more demanding graph analysis.
- Programmers who are interested in working on the project by further developing it or fix existing bugs

2.4 Operating Environment

- “Home Platform” is Linux
- MS Windows
- Mac OS X
- Any unix-like system that has the appropriate basic libraries

2.5 Design and Implementation Constraints

Gretl is written in the C programming language. It is free, open-source software. You may redistribute it and/or modify it under the terms of the GNU General Public License (GPL) as published by the Free Software Foundation.

2.6 User Documentation

- User Manual
<http://gretl.sourceforge.net/gretl-help/gretl-guide.pdf>
- Feature Requests and Bug Reports
<https://sourceforge.net/projects/gretl/support>

2.7 Assumptions and Dependencies

- **For Linux**

On the Linux platform you have the choice of compiling the gretl code yourself or making use of a pre-built package. Building gretl from the source is

necessary if you want to access the development version or customize gretl to your needs, but this takes quite a few skills; most users will want to go for a pre-built package.

- **MS Windows**

The MS Windows version comes as a self-extracting executable. Installation is just a matter of downloading gretl_install.exe and running this program. You will be prompted for a location to install the package. Nothing else is needed.

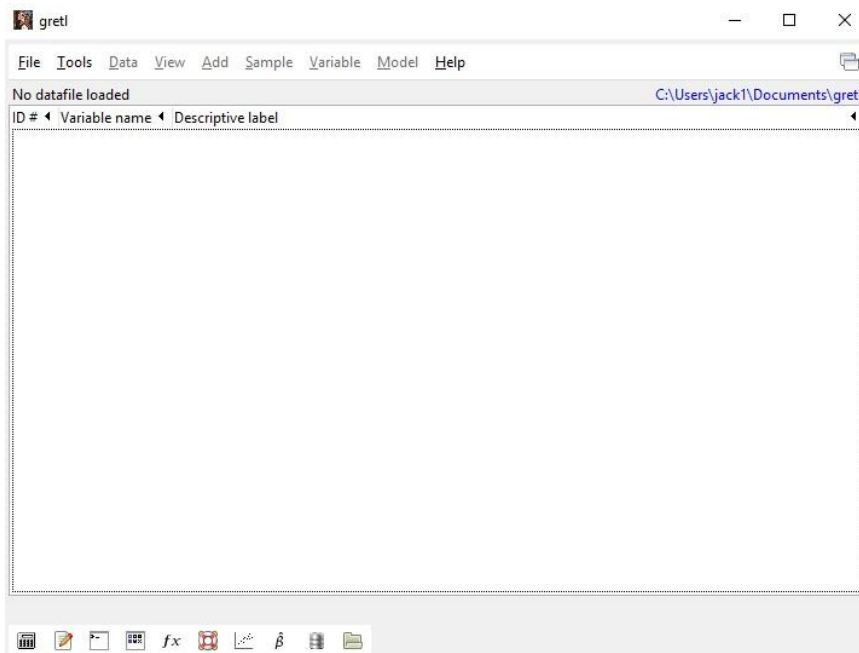
- **Mac OS X**

The Mac version comes as a gzipped disk image. Installation is a matter of downloading the image file, opening it in the Finder, and dragging Gretl.app to the Applications folder. However, when installing for the first time two prerequisite packages must be put in place first; details are given on the gretl website.

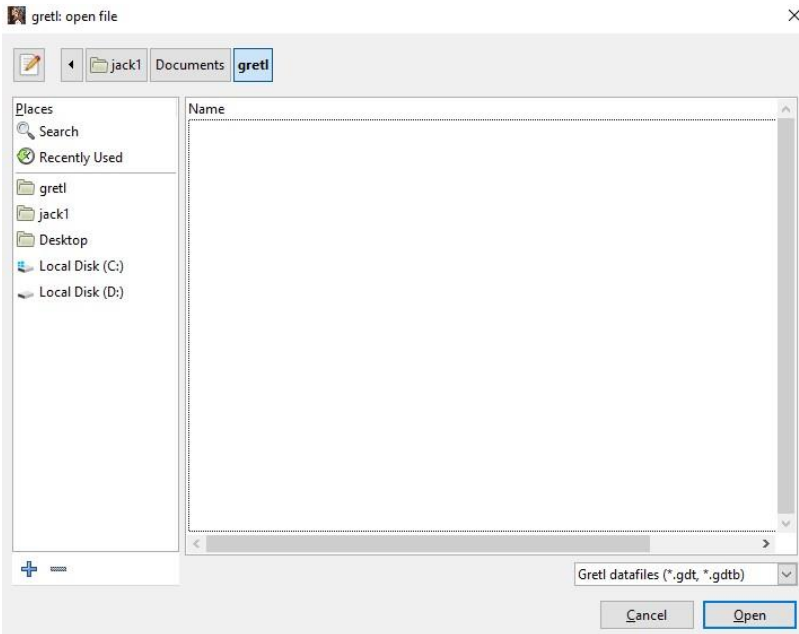
3. External Interface Requirements

3.1 User Interfaces

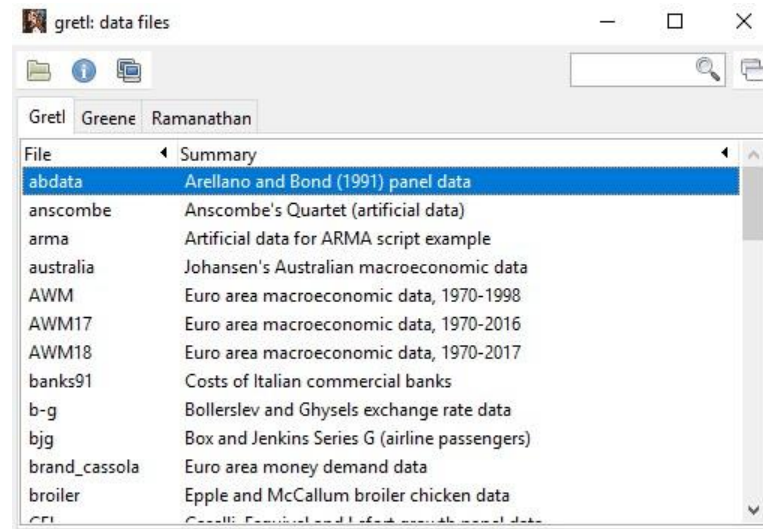
1. Starting Screen



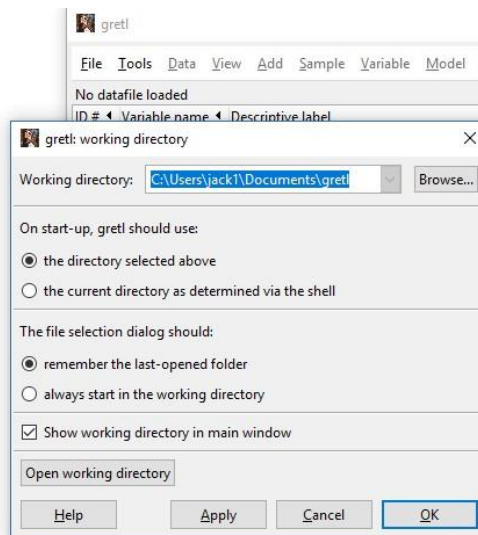
2. Import File



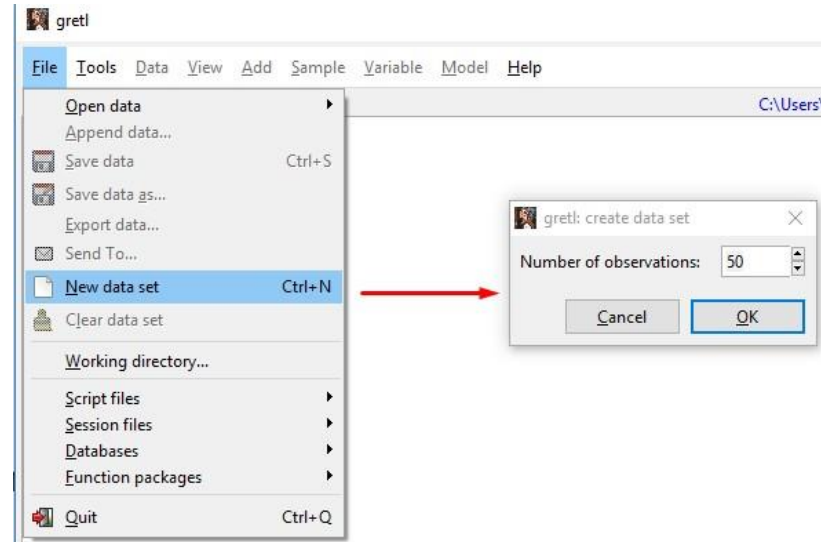
3. Import Sample



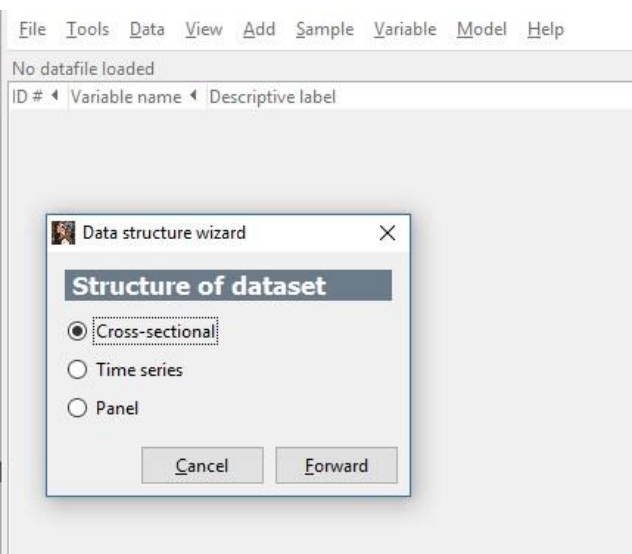
4. Working Directory



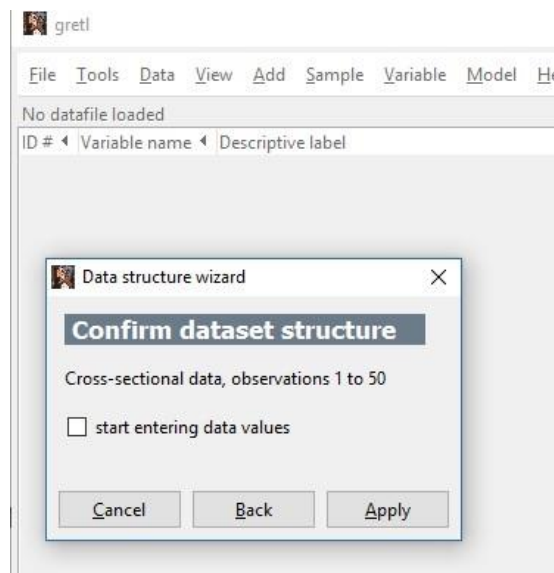
5. New Data Set(1/4)



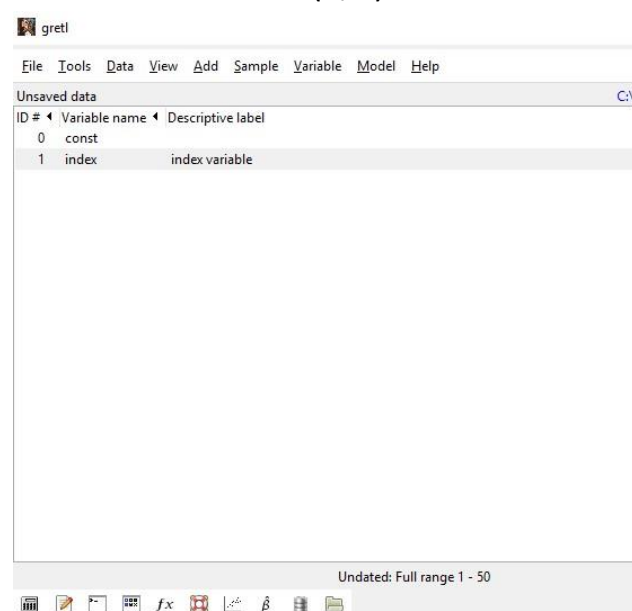
6. New Data Set(2/4)



7. New Data Set(3/4)



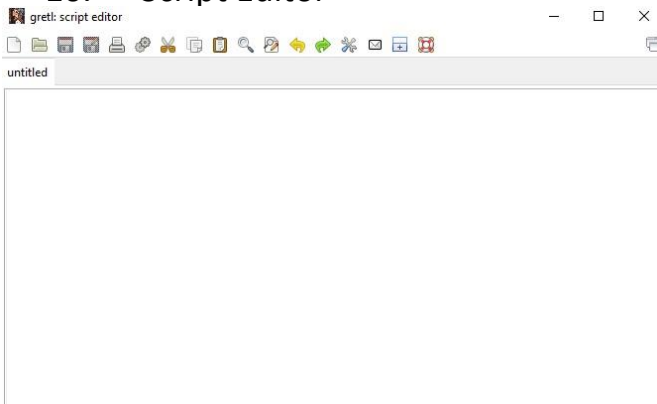
8. New Data Set(4/4)



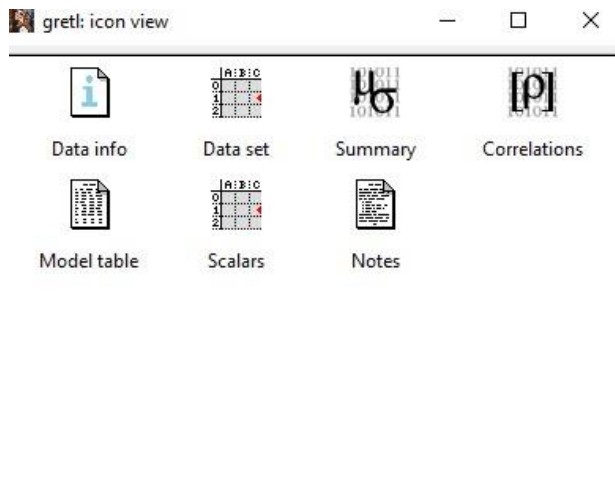
9. Console



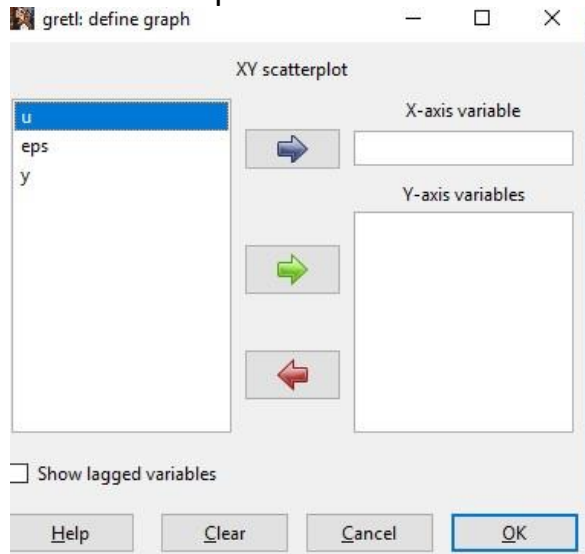
10. Script Editor



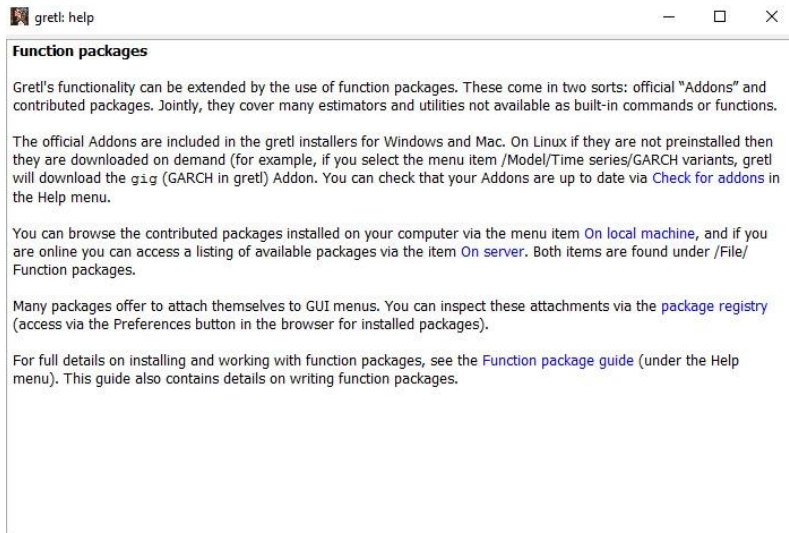
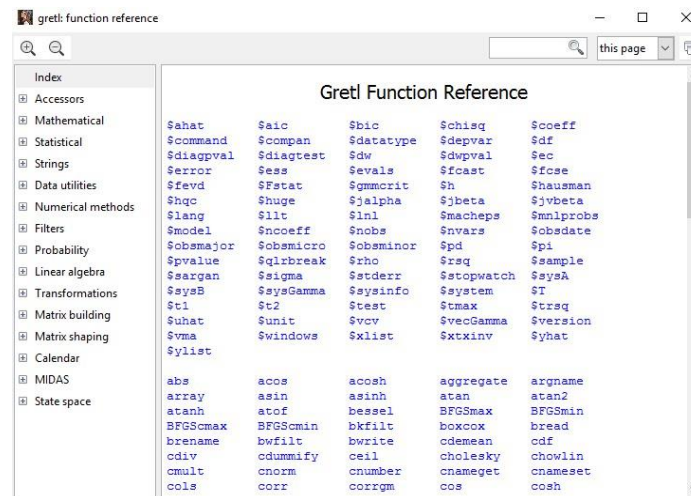
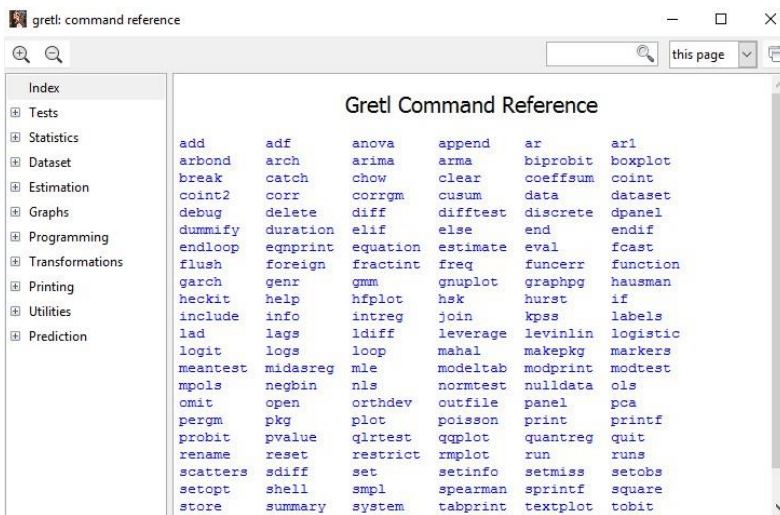
11. Icon View



12. X-Y Graph Menu



13. Help



14. Closing



3.2 Hardware Interfaces

- Mac OS X

First you need to determine if your Mac has an Intel or a PPC processor. Newer Macs (since 2006) all run Intel. If you have an older Mac and are not sure, click on the Apple icon at the top left of your screen and select "About this Mac". In the window that appears you'll find the OS X version number and also the processor type. If you have an intel Mac running OS X 10.6 (Snow Leopard) or higher, we recommend the "quartz" download [below](#). Some users may prefer to use the X11 version of gretl.

- MS Windows

As of version 1.9.4, gretl requires Windows XP or higher and a processor that supports the SSE2 instruction set. SSE2 support is found in all modern processors; it is absent in AMD CPUs prior to the Athlon 64, and in Intel CPUs prior to the Pentium 4.

3.3 Software Interfaces

If you have no administrator rights on Windows, you will need to have a program to unzip the installation folder.

3.4 Communications Interfaces

Gretl can fetch materials such databases, collections of textbook datafiles and addon packages over the internet.

4. System Features

This section demonstrates Gretl's most prominent features and explains how they can be used and the results they will give back to the user

4.1 Basic Statistics

1. Summary Statistics

With the given variables as an example (Price, Square-footage, Age, Utown, Pool, Fplace) from a house sale, using all 6 of them, we get the following Summary Statistics:

ID #	Variable name	Descriptive label
0	const	
1	Price	Price in \$
2	Sqft	Square footage of house
3	Age	Age of house (years)
4	Utown	If = 1 the house is located close to a university
5	Pool	If = 1 the house has at least one pool
6	Fplace	If = 1 the house has at least one fireplace

Undated: Full range 1 - 1000

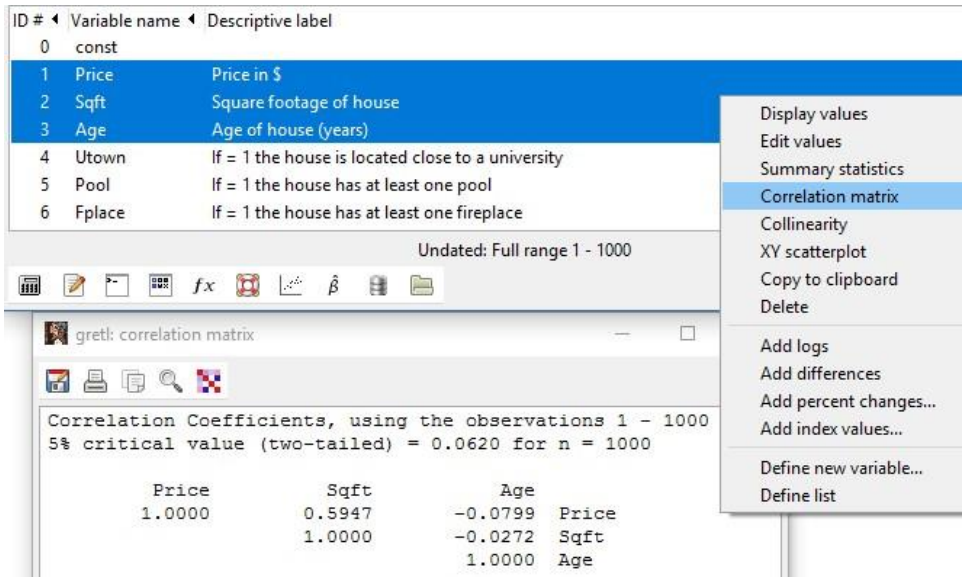
	Mean	Median	Minimum	Maximum
Price	2.4766e+005	2.4583e+005	1.3432e+005	3.4520e+005
Sqft	2521.0	2536.0	2003.0	3000.0
Age	9.3920	6.0000	0.0000	60.0000
Utown	0.51900	1.0000	0.0000	1.0000
Pool	0.20400	0.0000	0.0000	1.0000
Fplace	0.51800	1.0000	0.0000	1.0000

	Std. Dev.	C.V.	Skewness	Ex. kurtosis
Price	42193	0.17037	0.090562	-0.66743
Sqft	291.85	0.11577	-0.092835	-1.1850
Age	9.4267	1.0037	1.6475	3.0146
Utown	0.49989	0.96318	-0.076055	-1.9942
Pool	0.40317	1.9763	1.4691	0.15824
Fplace	0.49993	0.96511	-0.072047	-1.9948

	5% perc.	95% perc.	IQ range	Missing obs.
Price	1.8232e+005	3.1950e+005	62885	0
Sqft	2047.0	2956.9	492.75	0
Age	0.0000	29.950	10.000	0
Utown	0.0000	1.0000	1.0000	0
Pool	0.0000	1.0000	0.0000	0
Fplace	0.0000	1.0000	1.0000	0

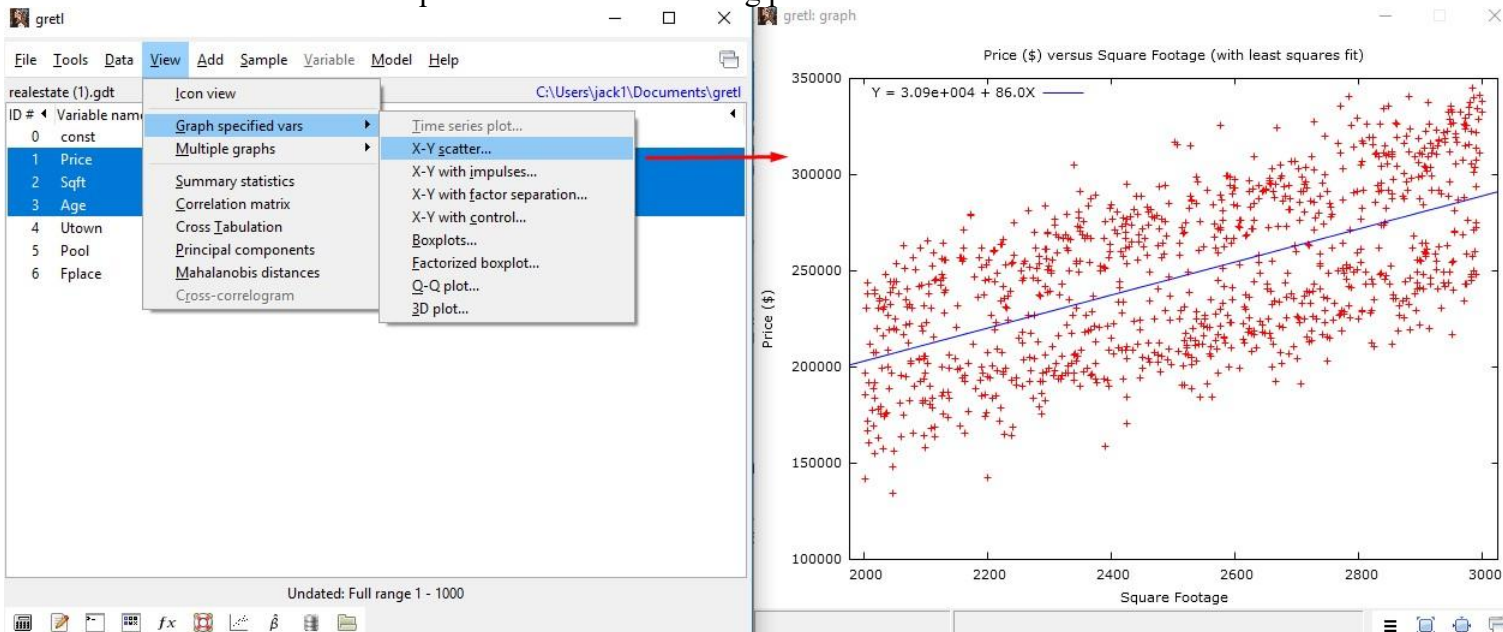
2. Correlation Matrix

With the same data, using the Price, Sqft and Age variables, we find the Correlation Coefficient between Square-footage and Price is 0.5947, the Correlation Coefficient between Age and Price is -0.0799, and the Correlation Coefficient between Age and Square-footage is -0.0272



3. X-Y Scatter Graph

Using the same data, square-footage will be the X-variable, and Price will be the Y-variable. The X-Y Scatter Graph is useful for visualizing patterns.



As shown, as Square-footage is increasing, so does the Price.

4. Statistical Tables

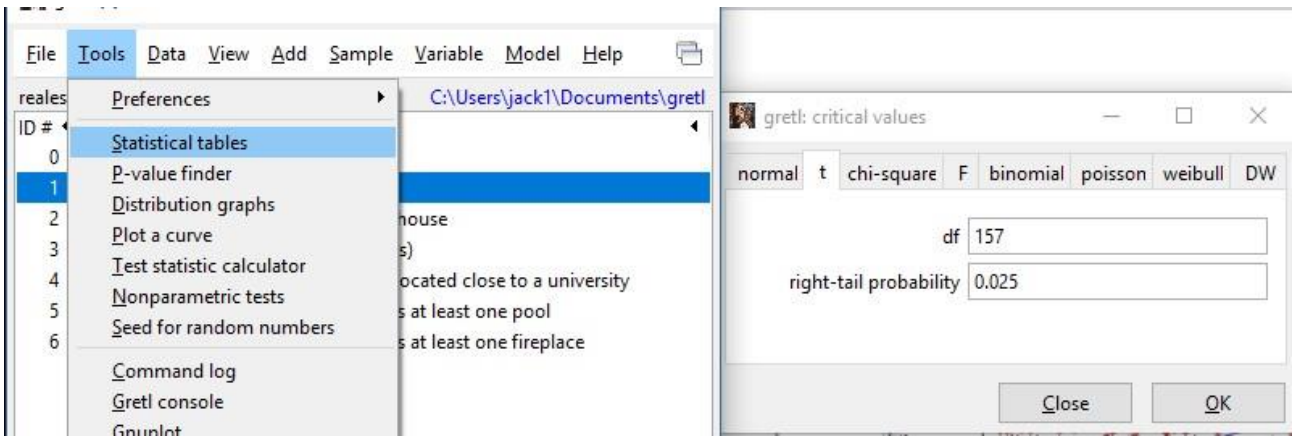
Example:

Two-tailed t test

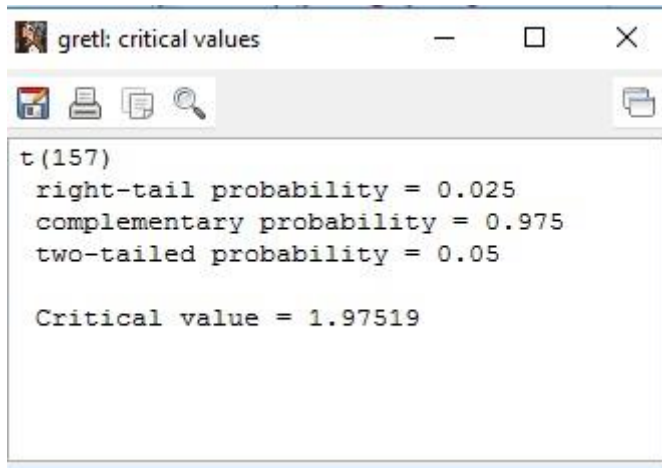
Right-tail propability = 0.025

df(degrees of freedom) = 157

Critical t value = ?



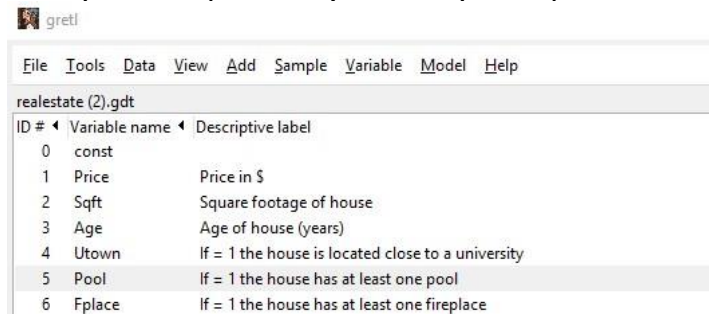
With Output



4.2 Estimators

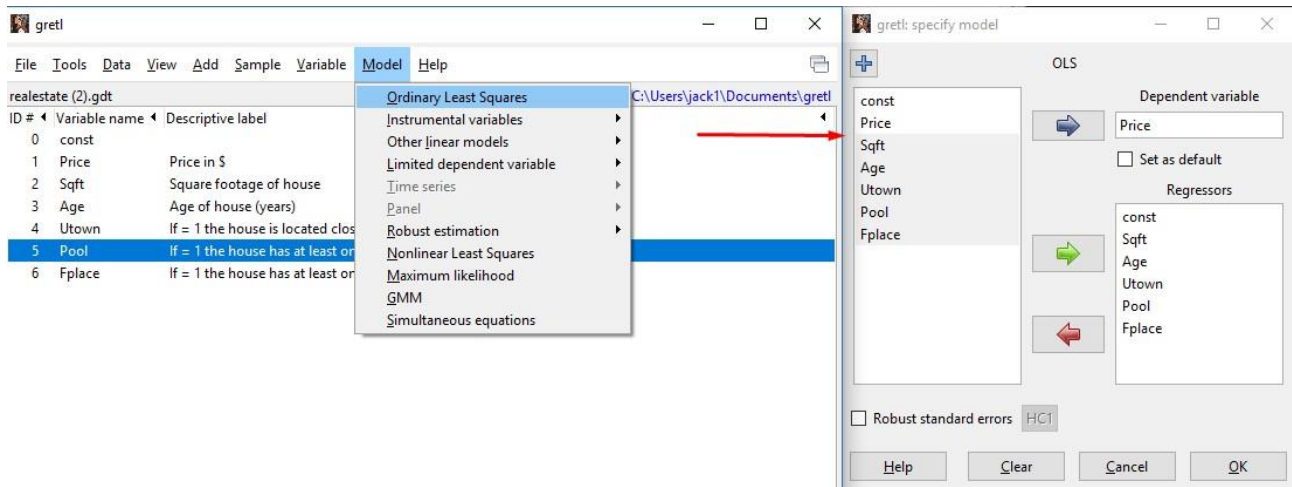
Gretl provides a wide variety of estimators: least squares, maximum likelihood, GMM; single-equation and system methods.

Example: OLS(ordinary least squares) Estimation with the given variables

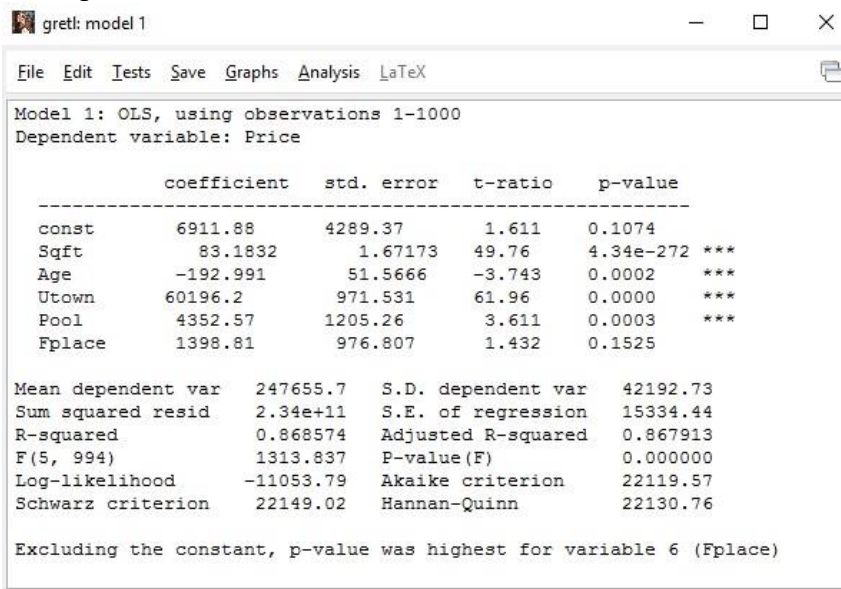


OLS Estimation

$$Price_i = \beta_1 + \beta_2 Sqft_i + \beta_3 Age_i + \beta_4 Utown_i + \beta_5 Pool_i + \beta_6 Fplace_i + \epsilon_i$$

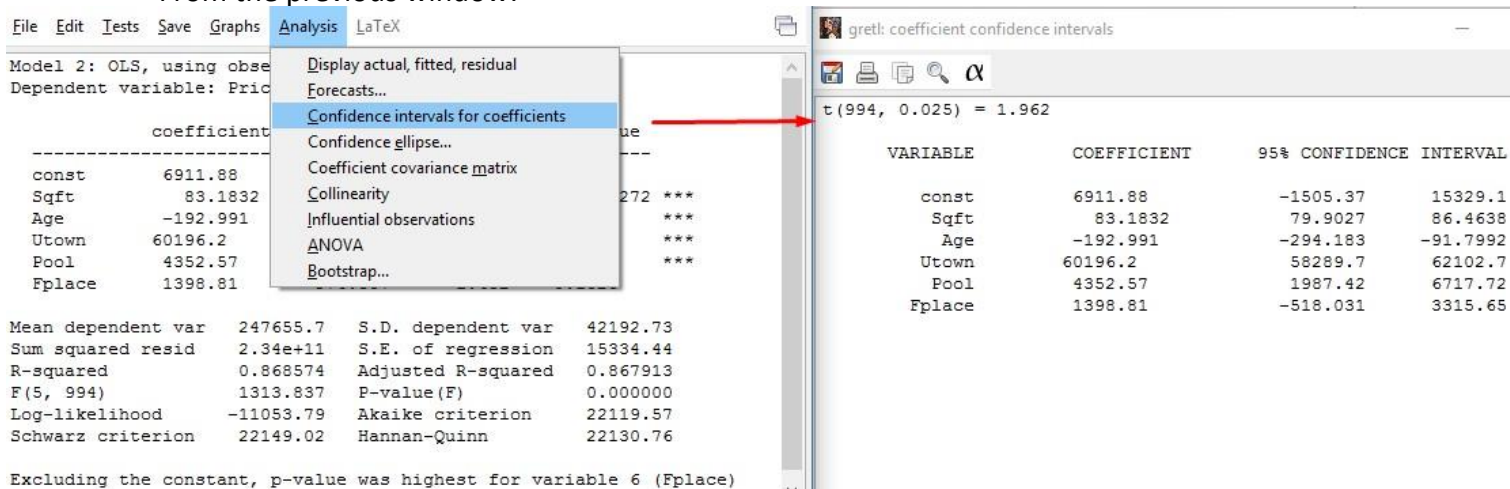


Output:



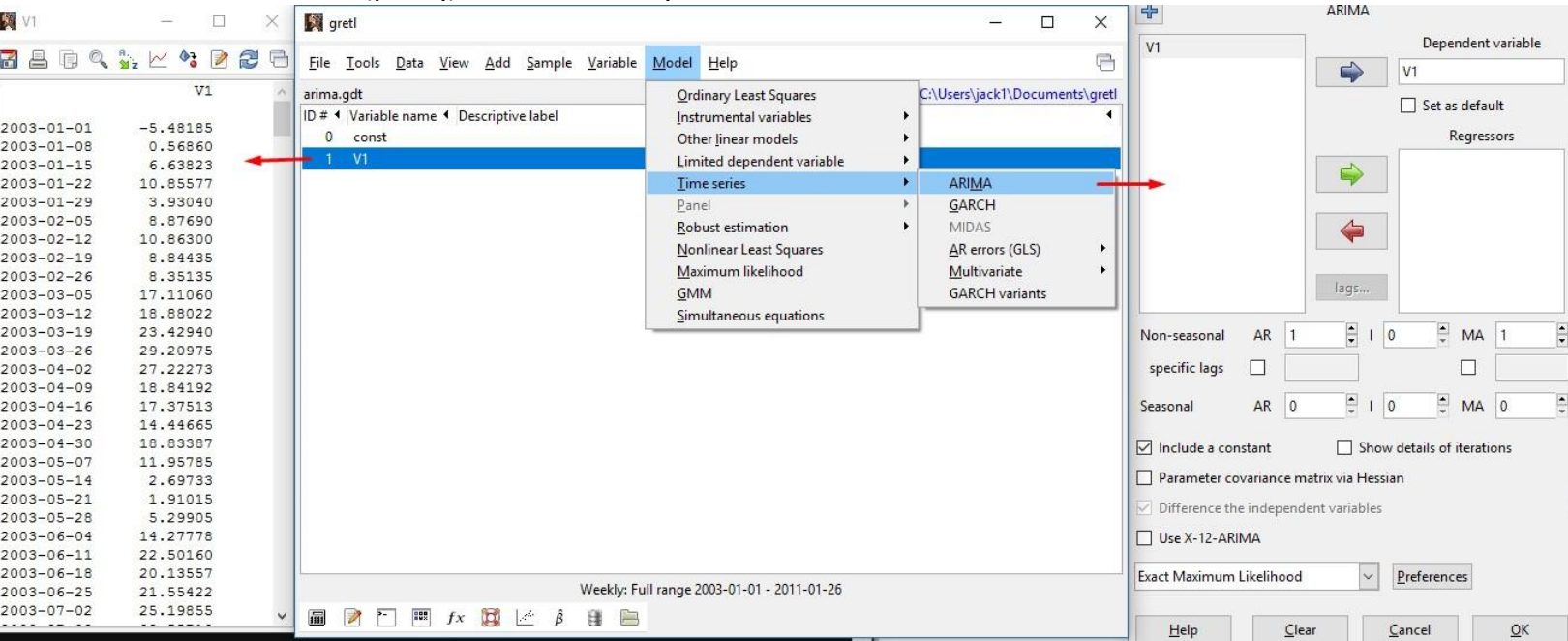
4.3 Limited dependent variables: logit, probit, tobit, sample selection, interval regression, models for count and duration data, etc.

Confidence Intervals Example:
From the previous window:

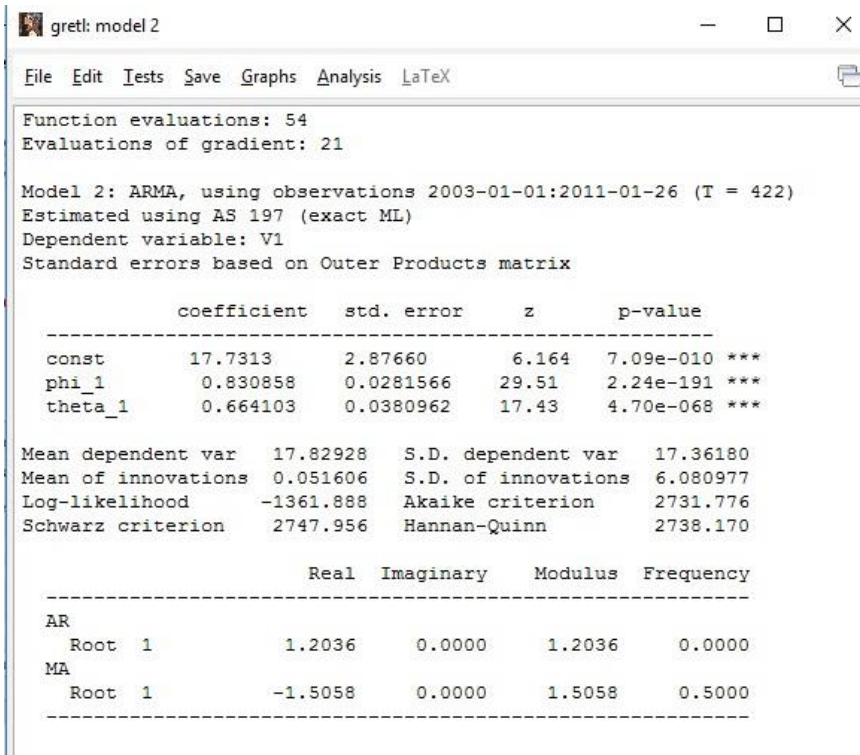


4.4 Time series methods: ARIMA, a wide variety of univariate GARCH-type models, VARs and VECMs (including structural VARs), unit-root and cointegration tests, Kalman filter, etc.

ARIMA(p,d,q) Models Example:



With Output:



5. Other Nonfunctional Requirements

5.1 Performance Requirements

Check 2.7 Assumptions and Dependencies.

5.2 Safety Requirements

To ensure that no one of Gretl's users loses any data while using Gretl (due to a crash or a bug of some kind) the developer team updates Gretl regularly. There is a bug tracker available where users can report any bugs they have encountered so that the developers can fix it in the next release.

5.3 Requirements

Gretl does not have any security requirements and thus any type of user can use it without any additional privileges.

5.4 Software Quality Attributes

Gretl provides the users with both simple and advanced features. Due to its well designed and easy to use interface it can be used by both experts and typical users. However, users must already have a basic knowledge of graphs before using it.

Glossary

References: https://en.wikipedia.org/wiki/Main_Page

- **Econometrics**: is the the branch of economics concerned with the use of mathematical methods (especially statistics) in describing economic systems.
- **Comma-separated values (CSV)**: is a delimited text file that uses a comma to separate values.
- **Summary Statistics**: are used to summarize a set of observations, in order to communicate the largest amount of information as simply as possible.
- **Correlation matrix**: is a table showing correlation coefficients between variables.
- **Scatter graph**: is a type of plot or mathematical diagram using Cartesian coordinates to display values for typically two variables for a set of data.
- **Two-tailed test**: is an alternative way of computing the statistical significance of a parameter inferred from a data set, in terms of a test statistic.
- **Heavy-tailed distributions**: are probability distributions whose tails are not exponentially bounded: that is, they have heavier tails than the exponential distribution. In many applications it is the **right tail** of the distribution that is of interest, but a distribution may have a heavy left tail, or both tails may be heavy.
- The number of **degrees of freedom** is the number of values in the final calculation of a statistic that are free to vary.
- The **critical values** of a statistical test are the boundaries of the acceptance region of the test.
- **Logit**: is the logarithm of the odds $p/(1 - p)$ where p is the probability.
- **Probit**: is a type of regression where the dependent variable can take only two values, for example married or not married.
- **Tobit**: The Tobit model is a statistical model proposed by James Tobin (1958) to describe the relationship between a non-negative dependent variable y_i and an independent variable (or vector) x_i .
- **Regression analysis** is a set of statistical processes for estimating the relationships among variables.
- **Time series** is a series of data points indexed (or listed or graphed) in time order. Most commonly, a time series is a sequence taken at successive equally spaced points in time.
- **Autoregressive integrated moving average (ARIMA) model**: is a generalization of an autoregressive moving average (ARMA) model. Both of these models are fitted to time series data either to better understand the data or to predict future points in the series (forecasting).

