

# Getting Started with Stata

## **WELCOME TO STATA**

So you've decided you want to learn how to use Stata. Great! Whether you're a completely new user to statistical software, or you've been using other packages, we're sure this survival manual will be able to help you with many of the questions you're bound to have.

You've probably come across Stata because your university or employer uses it. Stata has been less popular than its market competitors, such as SPSS and SAS, but is gaining in popularity every year. Stata is now used by medical researchers, biostatisticians, epidemiologists, economists, sociologists, political scientists, geographers, psychologists, social scientists, and other research professionals needing to analyse statistical data. One reason for this is that this software is particularly user-friendly when it comes to analysing complicated data sets, such as those where several data files need to be linked together.

The first version of Stata was released over twenty years ago in 1985. Since then, Stata has changed and developed according to user requests. Like its competitors, Stata is used for analysing quantitative data; but, unlike its competitors, Stata has several features which make it stand out as considerably more desirable. What are these reasons? They will be covered throughout this manual, but we will first turn to some of the resources that are available to Stata users.

## **RESOURCES**

There are considerable resources out there, in addition to this survival manual, to answer questions you might have about Stata. It should be noted that only a minority of these resources are

officially tied to StataCorp (the corporation that creates, sells, and distributes the Stata software, along with other products).

### **Stata website**

If you go to [www.stata.com](http://www.stata.com) you will find yourself at the official StataCorp website. On this site, there is information about StataCorp products, Stata technical support, versions of Stata, and what's new in Stata 'news'. Stata 'news' covers a wide range of topics, including (but not limited to) Stata user group meetings, training courses (such as the ones we teach), publications, and technical updates.

### **Timberlake website**

Another useful website is [www.timberlake.co.uk](http://www.timberlake.co.uk). Timberlake Consultants is a statistical consultancy company that also distributes and sells Stata in the UK. They have a lot of information on Stata on their website, and this is another place to go if you are looking for conferences and specialized training courses.

### **UCLA Stata 'portal'**

The University of California at Los Angeles has a remarkable web 'portal' at <http://statcomp.ats.ucla.edu/stata/> which anyone can access. This site is a virtual 'help desk' for statistical and Stata questions that is provided free of charge by the UCLA Academic Technology Service Stata Consulting Group. It is a remarkably rich resource archiving course notes, tutorials, and detailed annotated examples which include Stata commands, the output, and discussion about what the output means. We explicitly draw on some of these resources in Chapter 8.

### **Statalist**

Another resource is Statalist, which is a mass subscriber-based email list. You can sign up for it by going to [www.hsph.harvard.edu/statalist](http://www.hsph.harvard.edu/statalist) and following the instructions. You should note that there is a lot of traffic on this email list and you will get many messages. To give you an idea of just how much communication there is on this list, according to the archives, from 28 February 2007 to 8 March 2007 there were 303 individual messages! And this is just over nine days – you do the maths! There is a digest version available in which emails are batched together, so that you get fewer individual messages in your inbox. There are also Statalist archives (see the website) which you should browse before posting to the list. It is likely that your question has been asked before!

## Stata Journal

*Stata Journal* is a peer-reviewed journal about Stata that is published quarterly. It is available in hard copy and as an electronic version. You can find out more about *Stata Journal* at [www.stata-journal.com](http://www.stata-journal.com). The journal contains articles written about Stata as well as user-written software additions. Many of the changes that have occurred throughout the various versions of Stata have been as a result of user input. Some users give feedback to Stata by writing software additions which are programs that do specific tasks that aren't (yet) incorporated into the Stata software. Such user-written programs are often released with subsequent versions of Stata. These additions are profiled in the *Stata Journal* for those who are interested. It may sound like an enormous bore, but we have brought copies of the journal to the classes we teach to show students that the journal is appealing to advanced users and beginners alike. It really is worthwhile taking a look.

## Stata help files

Without trying to be ironic, we really want to impress upon you that the Stata help files are really very helpful. If you want to find out about a specific command, going to the **Help** menu and querying a command or searching for a keyword (which will lead to the command you need) provides a lot of information. There are explanations of what specific commands do, as well as the options that go along with each command. Often, there are examples which can help you set up your analysis. Much of the content of help files comes from the Stata user manuals, to which we now turn.

## Stata *User's Guide* and reference manuals

If you purchase Stata from a licensed vendor, you have the choice of purchasing the reference manuals at the same time. The *User's Guide*, while a small book, gives you basic introductory information on Stata. A detailed table of contents is found on the Stata website (address above). The reference manuals are a full set of books that, alphabetically, give detailed information about all the commands included in the version you are running (not including user-written additions). The reference manuals are excellent sources of statistical information, as detailed examples often are included, which include annotated discussion of results. There are also subject-specific reference manuals, although these vary by the version of Stata. In Stata 10, the reference manuals include a separate manual on data management, graphics, programming,

longitudinal panel data, multivariate statistics, survey data, survival analysis, and time-series analysis. The utility of these books cannot be overstated. If you want to become a regular user, you should have access to a set of these manuals!

## HOW IS STATA DIFFERENT?

New and potential users often ask us how Stata is different. Usually, this actually means ‘How is Stata different from SPSS?’. As we were both SPSS users for many years before using Stata, we think we can give some straight answers about the differences. But before we do so, it should be noted that, apart from writing this book, the authors have no vested interest in Stata as a product. We are university teachers who have had to learn as well as teach both programs. We both have a strong interest in research methods, and particularly, we both have an interest in answering research questions in the most effective manner possible.

We can tell you that, compared to SPSS, the commands in Stata are much more intuitive and less fussy regarding punctuation. If you fear ‘syntax’ because of your experiences in SPSS, we are fairly confident that you will find writing command language much less onerous in Stata.

Another strength of Stata is that it is user-driven. When there is a flaw or something that could be improved in the software, Stata listens to its users. In fact, many of the new applications incorporated into newer versions were written originally by users. Related to this point is that Stata, if you are connected to the web and there aren’t restrictions on what you can download, is ‘web active’. This means that you can download new applications that were written by users to perform specific tasks, and use them as commands. These additional user-written programs cover vast numbers of applications and simply searching for them within Stata and easily downloading and installing these bolt-ons is something that helps Stata stand out.

We both originally moved to using Stata because we were working with large complex data sets and found working with them in SPSS to be very cumbersome. If you have experience of working with longitudinal data sets with various different types of file structures, you will see that dealing with these in Stata is much quicker and easier.

But it isn’t the case that SPSS has no strengths over Stata. The editing of output and the labelling of variables, for example, is

much easier in SPSS. SPSS users who are used to having a data window open at the same time as running commands will also find it frustrating that this is simply not possible in Stata. For those who prefer to work entirely with pull-down menus, the newer versions of Stata have comprehensive pull-down menus (although we believe that it is a good thing to move on to saveable command files).

Although by no means a complete list of the advantages and disadvantages of Stata and SPSS, it will give you an idea of the reasons why people may or may not change their particular software preference. One of the biggest reasons for not changing is the market dominance of SPSS, although the number of Stata users continues to grow. We personally made the switch when we started working with large and complex data sets – but we haven't looked back!

## GETTING STARTED

In this book we assume that you are working with either version 9 or version 10 of Stata, but the graphics we use to accompany the text are done in version 10. We highlight the major differences between versions 9 and 10 later in this chapter. If you have an earlier version, much of what we discuss in this book will be relevant. As this is an introductory book, we will be focusing on the beginnings of data analysis, and these basic functions apply to the earlier versions of Stata as well. We also assume you will be using Windows as your operating system, although Stata can be run on Unix, Macintosh, and Linux.

The first thing you should do when you are getting started with Stata is get acquainted with the various windows that you are presented with when you launch it. Obviously, you need to begin by launching Stata on your computer. Depending on how your computer is set up, this is done by either double-clicking on the Stata icon or selecting **Stata** from the **Programs** menu, which you get to by clicking on **Start** in the left-hand corner of your screen and, then selecting **All Programs**, and then looking for **Stata** in the list of programs.

When you first launch Stata, you may find it off-putting. If you are familiar with other statistical software programs, particularly SPSS, you could think that Stata looks a lot less user-friendly. What are all these boxes? Where are my data? And what's with that black window that looks suspiciously like a DOS screen from



way back when? This may instil apprehension if you immediately associate this appearance with DOS programming language. But rest assured, you don't have to be a master of computing languages to use Stata. What we are going to show you is how these windows are customizable, and have features that were introduced as a response to user requests. As soon as you get used to the appearance and set up the windows the way you like them, you will find that Stata is extremely user-friendly indeed. So don't let first impressions put you off.

## FLAVOURS OF STATA

Stata comes in four basic 'flavours' – this is their terminology, not ours! The flavours differ according to the size of data sets and numbers of variables they can handle. The first flavour of Stata is Stata/SE, which is especially for large data sets. Stata/SE can handle over 32,000 variables and can fit models including over 10,000 independent variables (we can't imagine a model like this, however). The number of cases or observations it can handle is only limited by the capabilities of the computer on which it is installed. The second flavour of Stata is Stata/MP, which is simply a parallel-processor equivalent of Stata/SE. So if you have a computer with multiple processors and run estimations that take a lot of time on Stata/SE, you have the option of switching to Stata/MP, which is essentially the same but much faster.

The third flavour of Stata is Intercooled Stata, which is the 'standard'. If you are using Stata at a university lab or have purchased a single license, you probably have Intercooled Stata. This flavour can handle over 2000 variables and include just under 800 independent variables in a single model. Like Stata/SE, the number of observations in the data is only limited by the capabilities of your machine.

The fourth flavour is Small Stata, which is very limited compared to the other flavours. Small Stata can only accommodate 99 variables, is limited to 1000 observations, and can only use a maximum of 38 independent variables in an estimation.

What flavour do you have? It will say in the top left-hand corner of Stata when it is launched. For example it could be  Intercooled Stata 9.2 which shows that you have main version 9 (updated to 9.2) in Intercooled flavour. Or you may have  Stata/IC 10.1 which shows you have main version 10 (10.1) also in Intercooled flavour.

## ORGANIZING WINDOWS IN STATA

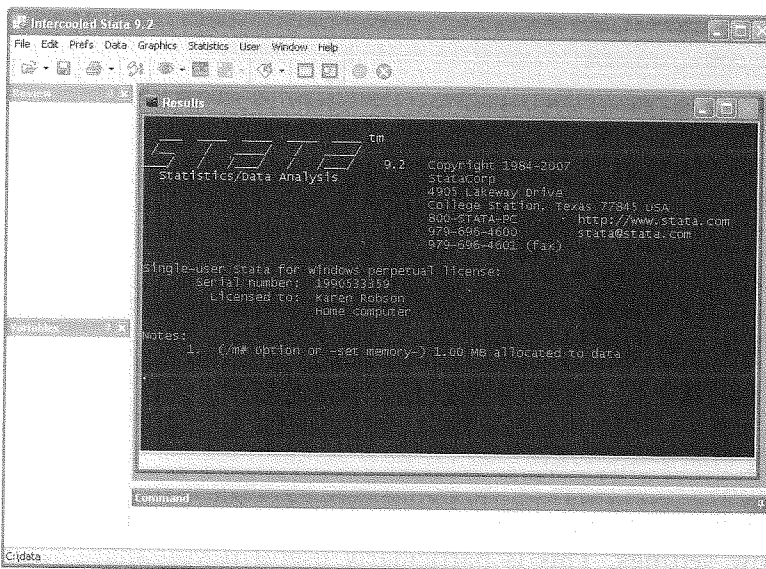
Stata launches with four windows: Review, Variables, Results and Command. There are considerable differences between versions 9 and 10. Prior to version 9, the window functions were fairly simple and mostly required you to manually change the size and position. In version 9 the windows got substantially more complicated in the default setting with pinning, docking and floating. In version 10 the default window settings returned to simple stretch and drag but with pinning, docking and floating as options. As there are quite a few differences between versions 9 and 10 at start-up, we tackle each in turn.

### Version 9

If the four windows (Review, Variables, Results and Command) are not all present, or just for the sake of starting on the same page as us, select from the menus at the top of the screen:

**Prefs → Manage preferences → Load preferences  
→ Factory settings**

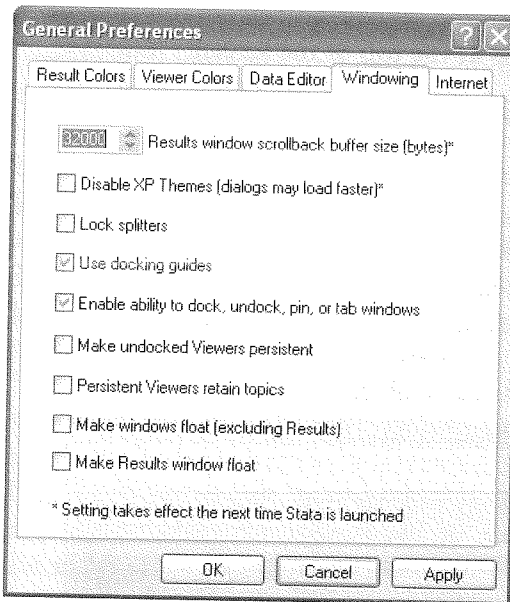
If you follow the above steps, you will get windows that look much like those below. Of course, your ‘flavour’ might be different (as indicated in the top left-corner), but even if it is, it will look the same upon launching. And obviously, your licensing and serial numbers will be different.



When starting out with version 9 we strongly recommend you disable the pinning, docking and floating functions so that you can avoid problems with windows disappearing or turning into tabs when you didn't intend them to. You can always come back to this later when you want to learn more about changing the display options available to you (see Box 1.1 for more information on this). Select

### Prefs → General Preferences

Then click on the **Windowing** tab and uncheck **Use docking guides** and **Enable ability to dock, undock, pin, or tab windows** and then click **Apply**. You will have to restart Stata for the effects to take place.



It is possible to resize the windows to suit your own preferences by dragging and stretching them. If you resize your windows, you can select

### Prefs → Manage preferences → Save preferences → new preferences set

Then a dialogue box will appear which allows you to 'name' your window settings. This is particularly useful if you share your



computer with another Stata user or if you have different preferences for your windowing for different projects you are working on. If you relaunch Stata and the window preferences have changed, provided you have saved your preferences, you just go to

**Prefs → Manage preferences → Load preferences →**  
**<name of your preferences>**

Our preferred window organization is shown below but you can size the windows, change the background colour, and change the size, colour and font of the text to whatever you find suits you best.








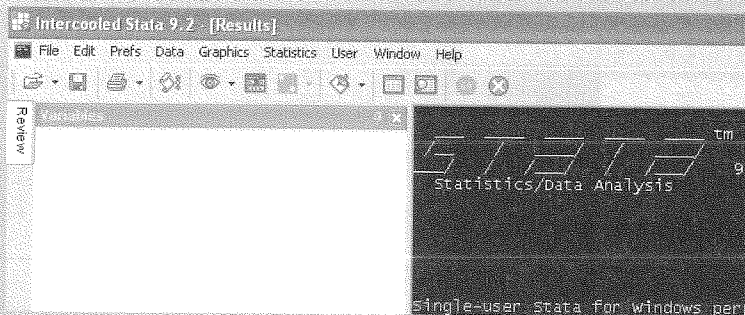
You may choose different settings depending on what you are using Stata for. For example, you may choose a different set-up when you are exploring data by mainly using the Command window compared to the set-up that you prefer when running analyses from do files. Similarly, if you go on to use Stata for presentations the windows set-up may be different again. However, we recommend you start with a windows set-up that you like and stick with it until you become familiar with the software. There's no point getting tied up in window preferences at the expense of getting to know the analytical capabilities of Stata. After all, this is the main purpose of using Stata.



### Box 1.1: Pinning, docking and floating windows

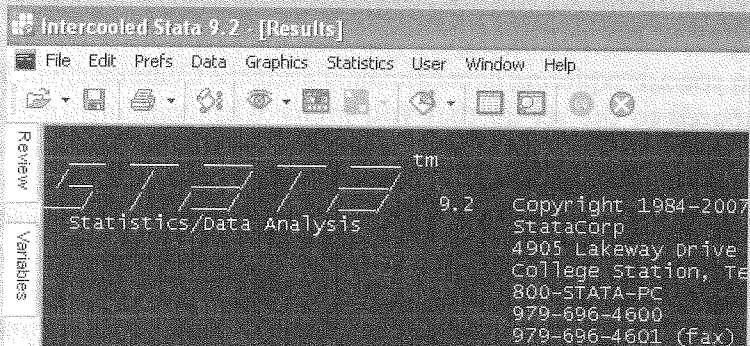
Pinning, docking, and floating windows are a default feature in version 9.

#### Pinning

If you have version 9 and look in the top right-hand corner of the Review, Variables, and Command windows, you will see  . The X symbol closes the window (as in many programs), but the  symbol is used for 'pinning' (hence, it is shaped like a pin). If you click on the  in the Review window, you will see that the Review window vanishes and becomes a tab on the left-hand side. If you click  on the Variables window, it too will become a tab.



If you click on the tabs for Review and Variables, you will see that they reappear, but quickly vanish once you click in any other Stata window. If you look carefully on the Review and Variables windows when they 'reappear', you will notice that the  symbol has changed to . This sideways position means that the window is not 'pinned'; as soon as you click elsewhere, it will become a tab again.



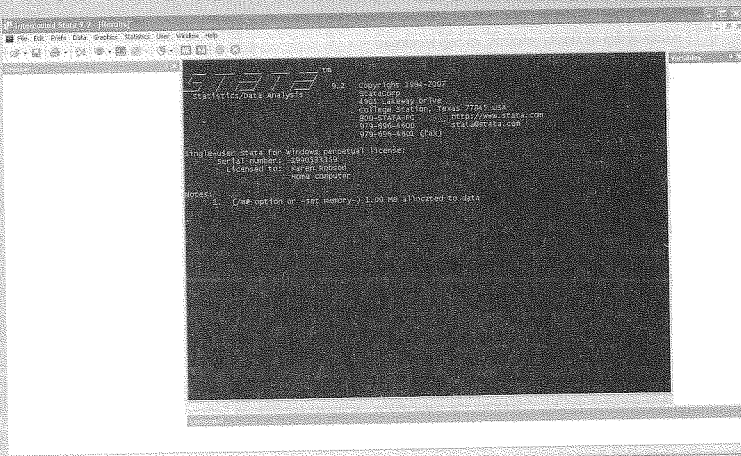
## Docking

If you drag the windows around in Stata, some arrows will appear:



These are docking tools that help you position your windows. If you drag the variables window across the screen – and by dragging we just mean clicking the blue area (in the Variables window in the above screen capture), holding down the left-click button on your mouse, and pulling the box somewhere else on the screen – a number of arrows appear. A compass-like set of arrows has appeared in the middle of the screen and additional arrows appear in at the top, bottom, and right and left sides of the Stata screen. What do these do?

If we select either of the right arrows (either the centre one or the one on the far right), the window repositions as below on the far right-hand side:



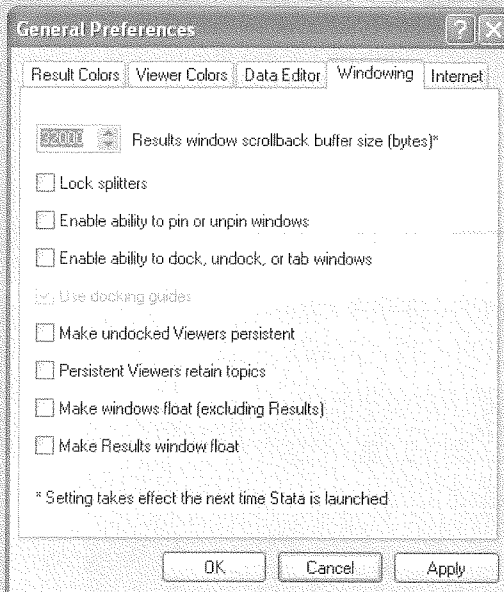
- ▶ You could drag it to different positions, docking it on top, below, or beside the Review window. The best way to familiarize yourself with these functions is to play around with them.

### Floating

Windows can 'float' – in other words, only appear on screen when you click on them, and at other times appear as tabs in your Windows task bar. If you double-click on the blue parts of the individual Variables, Review, and Command windows, they will float. They will appear in your task bar and only become visible when you click on them. We have had many students do this by mistake and get very frustrated when they can't get their windows 'back!' All you need to do to restore them to non-floating is click on them so that they reappear in the Stata window, and then *double-click* anywhere on the blue bar of the window to lock them into 'non-floating'.

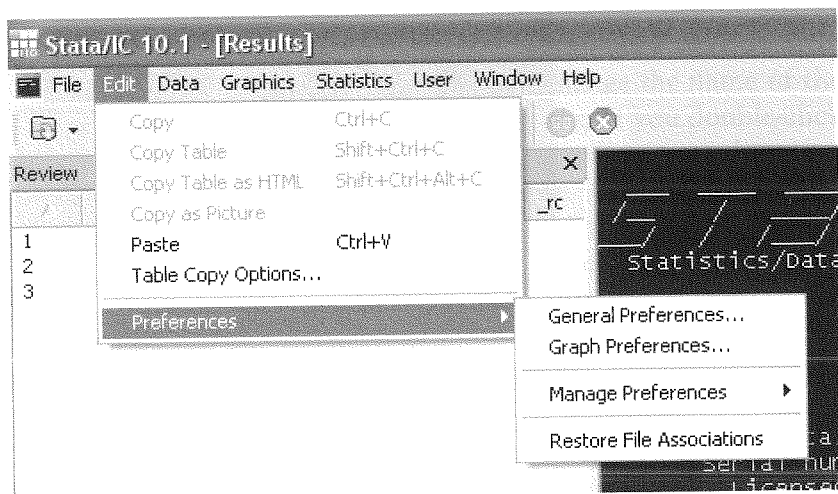
In version 10 you need to enable these features by ticking the boxes in the *Windowing* tab in the pull-down menu:

Edit → Preferences → General Preferences



## Version 10

Version 10 launches with a windows layout that is the same as when it was last closed. This is fine if you are the only user but if you share a computer with another Stata user, or want to save different layouts for different tasks, then you can save these preferences. To get to the layout you like simply stretch and drag the windows then save your preference using this pull-down menu:



To save your own preferences select:

**Edit** → **Preferences** → **Manage preferences** →  
**Save preferences** → **New preferences set**

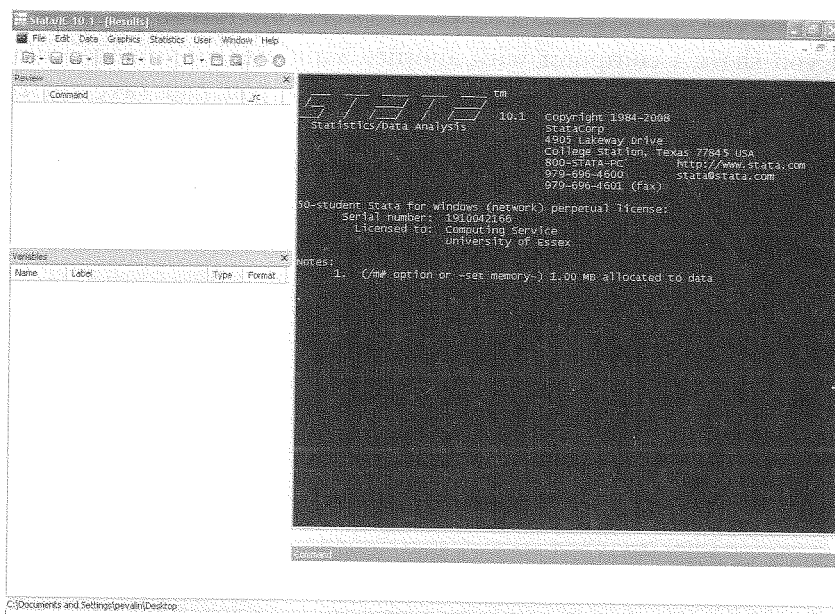
Then a dialogue box will pop up which allows you to 'name' your window settings. Then if you launch Stata and the window preferences have changed you just go to

**Edit** → **Preferences** → **Manage preferences** →  
**Load preferences** → <name of your preferences>

If you would like to start again with the factory settings follow:

**Edit** → **Preferences** → **Manage preferences** →  
**Load preferences** → **Factory settings**

Our preferred layout is shown below:



## WHERE DID MY WINDOW GO?

Sometimes, by accident, you may close one of the windows on your Stata screen. If, for example, your Variables window ‘vanishes’, it is quite easy to get it back – just go to the menu at the top of the screen and click

Window → Variables

## WINDOWS IN STATA AND WHAT THEY DO

The four windows of Stata that you see upon start-up all have very different purposes.

### The Command window

The Command window opens when Stata is launched and is a quick way of running commands, but only if you are familiar with the basic commands. By default, this window is located at the bottom of the Stata screen. We will talk more about typing commands into this window in the next section on ‘working interactively with Stata’.

## The Review window

The Review window is, by default, positioned in the upper left-hand corner of Stata. All of your commands are recorded here. If you type commands into the Command window, they will appear in the Review window. The Review window is particularly handy when you are exploring your data because if you, for example, do a frequency distribution of a variable, the command will appear in the Review window. All you need to do to repeat the command is either hit the ‘page-up’ button with the cursor in the Command window or click on the command in the Review window. The command will reappear in the Command window and you can even modify the command (for example, change the name of the variable you are examining) to save on typing. If you double-click on commands in the Review window, Stata will execute them.

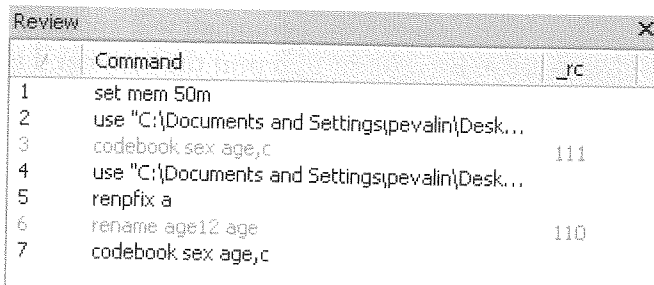
In version 10 you can select multiple lines in the Review window by holding down the Shift key and then double-clicking to get them all to run. Also in version 10 you can list the commands in either ascending or descending order by clicking on the **Command** heading. The order changes and this is indicated by the arrow next to the **Command** heading and the numbers next to the commands. A simple illustration of three commands is shown below, first with descending order and then with ascending order.

Review	
	Command
1	use "C:\Documents and Settings\...
2	ta d_ghq
3	su ghqscale

Review	
	Command
3	su ghqscale
2	ta d_ghq
1	use "C:\Documents and Settings\...

Also in version 10, the Review window indicates which commands returned an error by displaying them in red with the error code on the right. It is possible to sort according to the error codes by clicking on the **\_rc** heading. In this way you can easily delete the commands that caused an error if you want to convert your

Review window to a do file. To save the commands in the Review window as a do file simply right-click when one of the commands is selected and you will see an option to **Save Review Contents** which will automatically save them in a do file format.



	Command	
1	set mem 50m	
2	use "C:\Documents and Settings\pevalin\Desk...	
3	codebook sex age, c	111
4	use "C:\Documents and Settings\pevalin\Desk...	
5	renprefix a	
6	rename age12 age	110
7	codebook sex age, c	

### The Variables window

When you have a data set loaded, the variable names and labels appear in this window. Sometimes you need to stretch out the window widthwise to see the variable labels – if the variables are labelled in your particular data set. In version 10 you will see that the type and format of each variable is also listed in the window.

### The Results window

In this window, the results of the commands are presented. You should note that Stata does not ‘save’ the results here indefinitely and that by default, only 32,000 bytes of memory are allocated to the ‘buffer’. To permanently save your results you will need to use a log file that we introduce later in this chapter and tackle in more detail in Chapter 2. You can change the buffer to up to 500,000 bytes either by typing in the Command window:

```
set scrollbufsize 100000
```

or by using the pull-down menu:

**Prefs** → **General Preferences** (version 9)  
**Edit** → **Preferences** → **General Preferences** (version 10)

Then click on the **Windowing** tab and at the top you will see the up and down arrows to change the size of the buffer. Whatever size you allocate (between 10,000 and 500,000 bytes) will be



remembered by Stata and then become the default every time Stata is launched.

MASARYKOVA UNIVERZITA  
Fakulta sociálních studií  
Joštova 10  
602 00 BRNO



## INTERACTING WITH STATA

There are three ways to ‘interact’ with Stata to tell the program what you would like to do with the data. Here we have called these ways using pull-down menus, ‘working interactively’, and writing do files.

Before we get on to these different ways of interacting with Stata, it is important that we clarify our terminology. There is a good deal of confusing terminology surrounding the ways to interact – some comes from the language used by Stata, but other users carry over language from other software packages. This can be a little confusing for the beginner. For example, in SPSS the written text files for commands are called ‘syntax files’ but often referred to as ‘code’. The Stata name for a text file of commands is a *do* file, which has a *.do* extension (see Box 1.2 for other file extensions). That distinguishes it from the SPSS syntax files, but how do you refer to the actual text in the file? Some users carry over the word syntax but we prefer to use commands. So, in a *do* file we have a series of commands and these commands have options.

### Through pull-down menus

If you have used the Windows versions of SPSS then using pull-down menus to manipulate and analyse data will be familiar to you. Version 8 of Stata was the first version to incorporate pull-down menus and, while we strongly recommend the use of *do* files, the pull-down menus are useful when getting to know Stata, especially for graphing. You will see in the pull-down menus under **Statistics**, for example, a series of options for various types of statistical tests. It should be pointed out early on to the new user that there are more statistical possibilities in Stata than there are options in the pull-down menu which open dialogue boxes. It is not possible to access some types of statistical analyses in the pull-down menu – they must be entered as commands in either a *do* file or the Command window.

### Through the Command window

A second way of using Stata is through writing commands in the Command window and hitting the Enter key. The command then



appears in the Review window, and as discussed earlier, can be saved as a do file.

The Command window can be used to directly call up the dialogue boxes in the same way as the pull-down menus use them. For example, typing **db summarize** in the Command window then hitting the Enter key will bring up the dialogue box used by the pull-down menu to produce summary statistics (see Chapter 5 and Box 5.4).

In the Command window the typed commands can run over one line without any problem. However, in a do file if a command runs over one line (or you want to break up the command for ease of editing) then there needs to be a `///` at the end of each line except the last. This `///` tells Stata to ignore the line break and continue reading the command as if on a single line. We use this `///` notation in this manual when our commands run over a single line. If you are using the Command window to follow or adapt these commands, then you can ignore the continuation notation `///`.

### Through the construction of do files

Instead of just typing commands into the Command window or using the pull-down menus, it is likely that you will want to keep a record of your commands so that you can refer to them (and run them again) later. For those familiar with the syntax window in SPSS, the do file editor in Stata is much the same.


If you click on the do file editor icon at the top of the Stata window –  in version 9 and  in version 10 – a do file editor box appears. The do file editor is a simple text editor (similar to Notepad) in which you can type, copy, cut and paste text. In version 10 you may have more than one do file editor open; we suggest you start by only having one open, but note this function for when you are more familiar with the software.

In here you can type your commands and save them, keeping a record of how you have configured your files and manipulated your data. We will focus more on the use of do files later. We recommend using do files as soon as possible but to start out using either the pull-down menus or typing commands into the Command window.

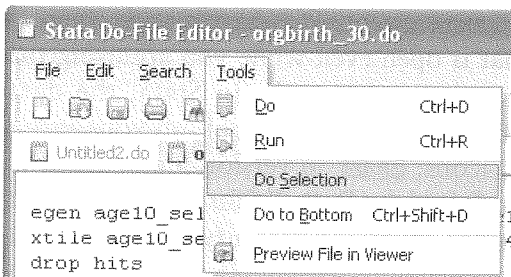
The do file editor in version 10 can have numerous do files open. The open do files are listed as tabs below the toolbar. In this example there are three do files open in the editor: *analysis\_10.do*, *org10\_1.do*, and *orgbirth\_30.do*. The do file currently shown (or at the front) is ***analysis\_10.do*** as the name is in bold, while the

other two are dimmed. To bring one of the others to the front simply click on the name.





Stata distinguishes between being asked to **do** a command and being asked to **run** a command. Both execute the commands but **do** tells you what command has been used and produces any results as it does them while **run** is silent or not showing the results. This can be a little frustrating when getting to know Stata. To start with we suggest that you use **do** all the time. The **do** icon is the farthest right on the do file editor tool bar . The one to its left is the **run** icon. If you look closely you can see that the **do** icon has lines on it which indicate output whereas the **run** icon is blank showing that it runs silently.

Under the **Tools** pull-down menu are some other options for **do** and **run**. It is important to understand that if you just click the **do** icon on the toolbar (or choose **Do** from the pull-down menu) then Stata will **do** the whole do file starting from the top. If you only want to **do** a line or two then you can select those lines in the do file then either click the **do** icon on the toolbar (or choose **Do** from the pull-down menu) and Stata will only execute those commands. The last option is to position the cursor in the do file and then use **Do to Bottom** from the pull-down menu. In this case Stata will do all commands from where the cursor is to the end of the do file.



It should be noted here that there are other text editors that people use to save their commands, but for the purposes of this book, we will use the editor provided in Stata.

## LOG FILES





Because you can't save the contents of the results screen in Stata (like the output window in SPSS, if you have used that program before), you should quickly get into the habit of using log files. Like the name suggests, log files make a record of your Stata session and include all your commands and results in one document. You will find an icon –  in version 9 and  in version 10 – on the toolbar which is the button for beginning a log file. We will return to log files in the next chapter, but for now, just remember that these are the types of files we use in Stata to make a record of our session.

### Box 1.2: File extensions

There are four file extensions in Stata that you should get to know:

- data files have the extension `.dta` – this corresponds to `.sav` files if you might be familiar with in SPSS;
- do files have the extension `.do` – this is similar to the `.spss` extension in SPSS for syntax;
- log files have the extension `.log` or `.smcl` (Stata Markup and Control Language) – this is most like the output file in SPSS, which has the extension `.spo`;
- graphs have the extension `.gph`

## WHERE ARE MY DATA?

If you are accustomed to other statistical software, you might be puzzled as to how to view your data. In SPSS, for example, there is a Data Viewer window that is open even while you are using the pull-down menus to select your commands. In Stata, the data is 'behind the scenes', like the log file. If you look on your menu bar, you will see these two icons:   in version 9 and   in version 10. The one on the left is the **Data Editor** and the one on the right is the **Data Browser** (with the magnifying glass). You can physically change the data in your data file in the **Data Editor**, but *not* in the **Data Browser**. Only the **Data Editor** or the **Data Browser** may be open at one time (not both). Also, you must close the **Data Editor** and **Data Browser** before Stata will run any commands.

Now that we have an idea of what all the windows are for, it is time to turn to how to really 'get started'. We begin by preparing Stata to work with data and setting up our directories.

## ALLOCATING MEMORY TO STATA

When Stata launches, there is a note in the Results window that tells you how much space has been allocated to data. This varies by the flavour of Stata, network and the default settings on your computer/network. On the particular computer that we use, Stata opens with 1 Mb allocated to memory. This may seem like plenty, but a lot of data sets are much bigger than this and therefore it is necessary to increase the allocation of memory. To use your available memory efficiently we suggest that you set the memory to only slightly larger than the data set to be used. If, for example, we want to change the allocated memory to 50 Mb, we just use the command:

```
set mem 50m
```

You can type whatever amount of memory you require here depending on the size of the data set you are using and the specifications of your computer. If you move on to analysing large survey-type data then you may need to specify 250–500 Mb of memory. If you decide, after some practice, that you typically use 50 Mb (or some other amount) of memory every time you use Stata, you can type

```
set mem 50m, permanently
```

This memory allocation will take effect by default every time you start Stata.

## SETTING DIRECTORIES

You can, and should, make sure that all of your files for a particular project are saved to the same directory on your computer. This will make it easier to retrieve any file related to a particular project and save you retyping long file locations every time you want to open or save data.

You can tell Stata to create a directory using the command **mkdir** (make directory) – a new directory is created where you

can store all of our data, syntax, and log files. So if you type in the command window:

```
mkdir c:/projectname
```

Stata will make a folder in your C drive called `projectname`. Of course, you would want to change this to where you want your data and do files stored. It might not be on the C drive and you probably would want a more meaningful name for your folder! Alternatively, you can create the folder you want to use first through your software and its file and folder management functions.

You can then use the command `cd` to tell Stata to change directory so that your files are stored there. Type:

```
cd c:/projectname
```

In version 10 you can change the working directory using the pull-down menu:

**File → Change working directory**

Then browse to the location of the directory/folder you wish to use.

Now if you save any files without explicitly defining another directory, your files will be saved in this directory. So if you have a data set open and you want to save it, you can simply just type **save nameofdata** and this file would now appear in your new directory.