

CHAPTER 1

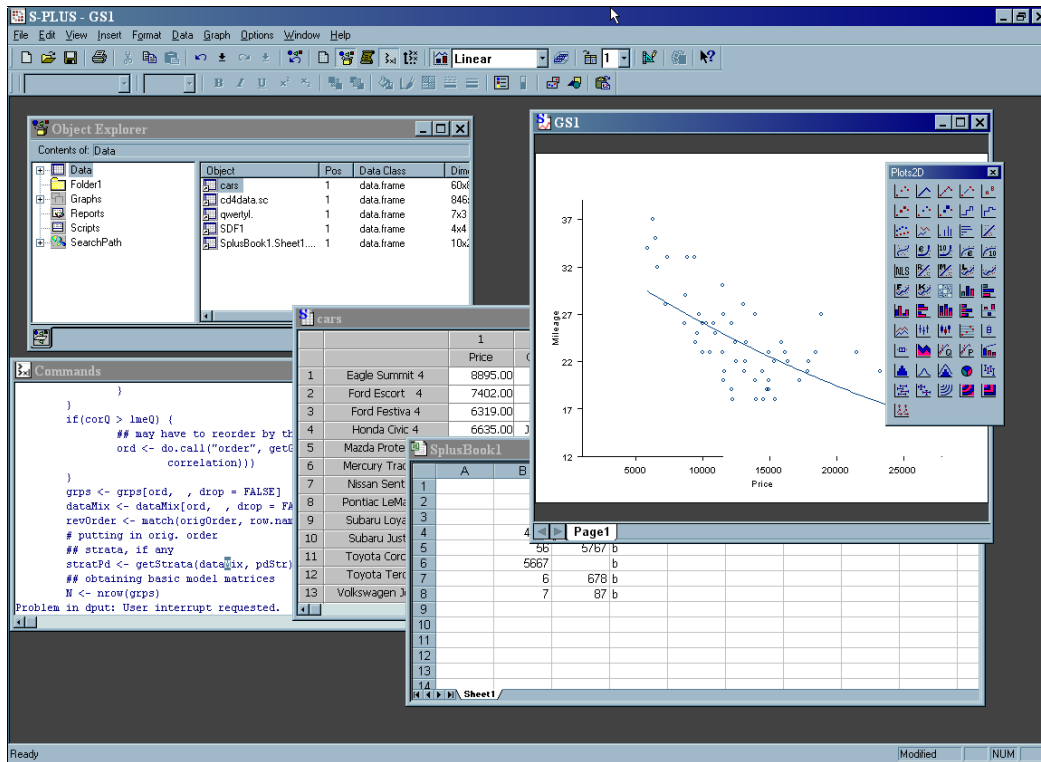
1. Introduction

The name S-PLUS consists of two parts, S and PLUS. At the end of the seventies the programming language S was developed at AT&T Bell Labs (nowadays Lucent Technologies), which is incidentally the same lab that developed the programming language C and the operating system UNIX. You can find a number of similarities between S, C and UNIX. One of the goals of S was to create a language flexible enough to interact with data and to support the research activities of the statistical department at Bell Labs. Because of its flexibility, statisticians and data analysts became more and more interested in S. The S-PLUS user's group is very active and has contributed substantially to S-PLUS.

To popularize S and to ensure a professional quality professor Doug Martin (University of Washington, Seattle) founded the company Statistical Sciences (StatSci) in 1987. With an exclusive license from Bell Labs to commercially develop S, StatSci extended S by adding numerous new functions. The system was made available for Windows platforms and most UNIX platforms. The extended version of S is called S-PLUS. In 1994 StatSci became part of MathSoft, the producers of mathematical software since 1984. In March 2001 MathSoft became Insightful.

S-PLUS is object oriented. This means that everything is considered as an object of a certain class. Creating data or graphs, performing statistical analysis and writing functions all boils down to creating or manipulating an S-PLUS object. Until 1997 S-PLUS was purely command driven, which was perfectly alright for most professional statisticians and programming minded data analysts. It gave them a way to express their "statistical feelings". To make S-PLUS more accessible for non-programming minded users S-PLUS version 4 for Windows was released in 1997. A complete graphical user interface was built around the S programming language. It enabled non-programming minded users to perform powerful analyses and create great publication quality graphs by pulling down menus and clicking buttons.

In S-PLUS you can import data from virtually any source, manipulate and manage data in data windows, create publication quality graphs by dragging and dropping, and perform modern statistical analysis. If your favorite routine is not available in S-PLUS, you can either program it yourself, or incorporate utilities that other users have programmed. All buttons are provided with explanatory tool tip texts, which become visible when you hover your mouse cursor over the button. Other toolbars are available for specific tasks, for example a 2D-plot palette to create 2D graphs.



The emphasis of this introductory book to S-PLUS will be on the graphical user interface. Chapter 2 gives an overview of the different help resources. In Chapter 3 to 7 you will find step-by-step procedures to perform common tasks such as click here or go to menu a and b. Chapter 3 describes importing data files into S-PLUS. Reading Chapter 4 is strongly recommended. It gives an overview of how to manage objects and projects in S-PLUS. Chapter 5 describes data manipulation techniques and Chapter 6 shows you how to create and modify graphs in S-PLUS. Performing statistical analysis using the S-PLUS GUI is described in Chapter 7. The last chapters provide a short introduction to the S language. Chapter 8 and 9 give an overview of the S programming language. Extending the user interface will be covered in Chapter 10.

1.1 Installation

The installation procedure of S-PLUS is straightforward. A wizard will guide you through the steps. There are a few things to consider.

1.1.1 System requirements

- A PC with Pentium processor
- 32 MB internal memory (64 MB recommended)
- 75 MB of free hard disk space
- CD-ROM drive
- Microsoft Windows 95/98/Me, Windows NT (4.0 or higher) or Windows 2000

1.1.2 The S-PLUS network version

The network version of S-PLUS helps you to keep track of the use of S-PLUS and its modules. It requires the installation of a license manager on a server and client installations of S-PLUS (and the modules) on client PC's. To install the license manager you require a license key, which can be obtained via Insightful or one of its local S-PLUS distributors.

1.1.3 Directory choices

During the installation of S-PLUS you will be asked to specify the installation directory. By default the installation procedure chooses something like “C:\Program Files\splus6”. However, this can be modified for your own convenience.

The installation procedure also asks you to specify a so-called working directory. In this working directory the subdirectories “.Data” and “.Prefs” will be created to store S-PLUS objects and personal preferences. Chapter 4 will explain this further.

1.1.4 Excel add-in and SPSS Link

If you have Excel installed S-PLUS will install the S-PLUS Add-in for Excel. If you have SPSS installed S-PLUS will install the SPSS Add-in. These add-ins make it easy to create and modify S-PLUS graphs within Excel or SPSS.

1.1.5 Acrobat Reader

You will need Acrobat Reader to read the on-line manuals of S-PLUS. A copy of Acrobat Reader is provided on the S-PLUS CD.

1.1.6 S-PLUS-Libraries

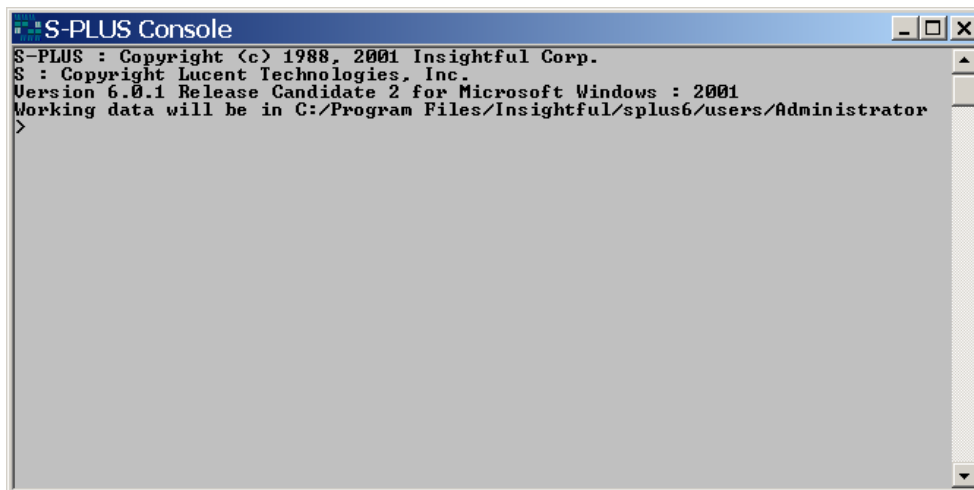
Three ‘external’ S-PLUS libraries are provided on the S-PLUS CD. An S-PLUS library is a collection of S-PLUS functions, which extend the functionality of S-PLUS. The three libraries are:

- MASS library, written by W.N. Venables and B.D. Ripley
- Design library and Hmisc library, written by F.E. Harrell

Note that Insightful does not support these libraries. To use the functions in these libraries you must first attach the libraries to S-PLUS. Start S-PLUS and go to the menu **File? Load Library**. The MASS library includes a graphical user interface to run some of the functions in the library. The Design and Hmisc libraries are purely command line driven.

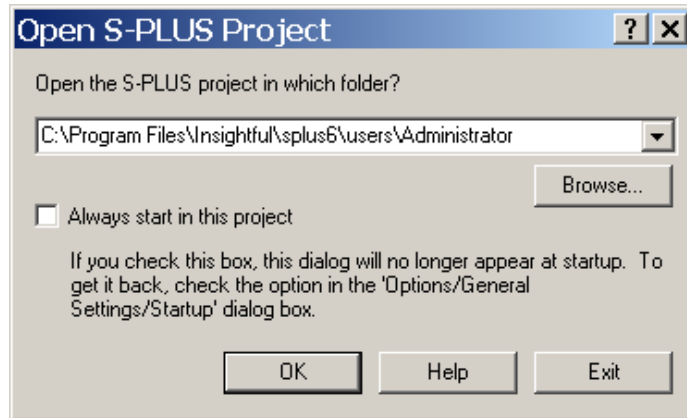
1.2 Starting S-PLUS

You can start S-PLUS in either the normal S-PLUS mode or the S-PLUS console mode. In the console mode you have to access the S language via a DOS box, as you can see in the next figure.



As old-fashioned as it may look, the console mode is a convenient way to quickly enter a few commands and it is invaluable for those users who never use the GUI anyway. In the console mode S-PLUS uses about 20 MB of internal memory less than in the GUI mode.

The normal mode gives access to the graphical user interface and the S language. When you start S-PLUS in the normal mode a small dialog will appear first, asking in which chapter or folder you want to start S-PLUS.



Project folders and chapters will be discussed in detail in Chapter 4.

1.3 Other products based on S and S-PLUS

1.3.1 S-PLUS for UNIX and Linux

The title of this book already suggest that there is an S-PLUS version for non-Windows based platforms, and indeed S-PLUS 6 is available on most UNIX (and Linux) platforms. Below you'll find a short list of similarities and differences between the UNIX and Windows version:

- Both the UNIX and Windows versions are based on the same version of the S language and functions you write in S-PLUS work on both platforms.
- The graphical user interface of the UNIX version is JAVA-based. The two versions look similar, but the UNIX version contains less menus and dialogs.
- To connect to other applications you use JAVA technology on the UNIX platform, whereas you use (OLE) Automation or C++ under Windows.
- The Windows version has two graphical systems: the object oriented graphs (see Chapter 6) and the traditional graphs (see Chapter 9). The UNIX version only has the traditional graphs.

1.3.2 The S-PLUS standard edition

The standard edition of S-PLUS differs from the S-PLUS professional version in that you cannot access the S programming language. This means that you cannot open the command window and you cannot open a script file. You can, however, run functions written by users of the professional version.

1.3.3 S-PLUS modules

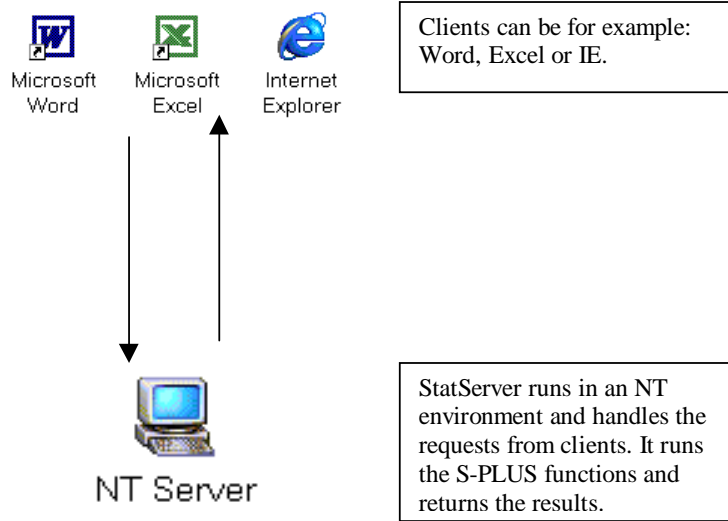
The following commercial add-on modules are available to extend the functionality of S-PLUS:

- S+Wavelets, for wavelet analysis of signals and images.
- S+GARCH, for analysis of time series with non-constant volatility.
- S+NuParam, for solving large scale (constrained) optimization problems.
- S+SpatialStats, for analyzing spatial data.
- S+EnvStats, for analyzing environmental data.
- S+ArcViewLink (Windows only), links S-PLUS and ArcView GIS.
- S+SeqTrial, for analyzing, monitoring, and evaluating clinical trials using group sequential methods.

1.3.4 StatServer and S-PLUS analytical Server

StatServer provides S-PLUS functionality via the internet or an intranet. S-PLUS and StatServer must be installed on a (central) Windows NT server, but the StatServer clients run on any platform, either a

web browser, Microsoft Excel or a custom application. The clients connect to StatServer and run pre-defined functions. The results can be graphs, numerical reports data or data files.



UNIX and Linux platforms offer the S-PLUS analytical server to prepare S-PLUS reports, graphs and analyses.

