The SPSS environment

3

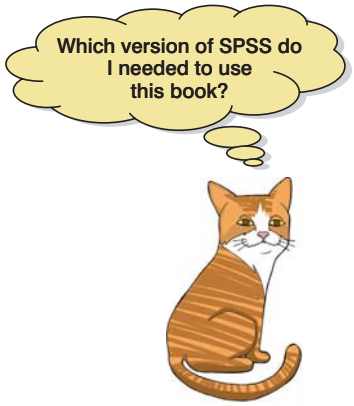
FIGURE 3.1
All I want for Christmas is … some tasteful wallpaper

3.1. What will this chapter tell me? ①

At about 5 years old I moved from nursery (note that I moved, I was not 'kicked out' for showing my …) to primary school. Even though my older brother was already there, I remember being really scared about going. None of my nursery school friends were going to the same school and I was terrified about meeting all of these new children. I arrived in my classroom and, as I'd feared, it was full of scary children. In a fairly transparent ploy to make me think that I'd be spending the next 6 years building sand castles, the teacher told me to play in the sand pit. While I was nervously trying to discover whether I could build a pile of sand high enough to bury my head in it, a boy came and joined me. He was Jonathan Land,

and he was really nice. Within an hour he was my new best friend (5 year olds are fickle …) and I loved school. Sometimes new environments seem more scary than they really are. This chapter introduces you to a scary new environment: SPSS. The SPSS environment is a generally more unpleasant environment in which to spend time than your normal environment; nevertheless, we have to spend time there if we are to analyse our data. The purpose of this chapter is, therefore, to put you in a sand pit with a 5 year old called Jonathan. I will orient you in your new home and everything will be fine. We will explore the key windows in SPSS (the **data editor**, the **viewer** and the **syntax editor**) and also look at how to create variables, enter data and adjust the properties of your variables. We finish off by looking at how to load files and save them.

3.2. Versions of SPSS ①



This book is based primarily on version 17 of SPSS (at least in terms of the diagrams); however, don't be fooled too much by version numbers because SPSS has a habit of releasing 'new' versions fairly regularly. Although this makes SPSS a lot of money and creates a nice market for people writing books on SPSS, there are few differences in these new releases that most of us would actually notice. Occasionally they have a major overhaul (version 16 actually looks quite different to version 15, and the output viewer changed quite a bit too), but most of the time you can get by with a book that doesn't explicitly cover the version you're using (the last edition of this book was based on version 13, but could be used easily with all subsequent versions up to this revision!). So, this third edition, although dealing with version 17, will happily cater for earlier versions (certainly back to version 10). I also suspect it'll be useful with versions 18, 19 and 20 when they appear (although it's always a possibility that SPSS may decide to change everything just to annoy me). In case you're very old school there is a file called **Field2000(Chapter1).pdf** on the book's website that covers data entry in versions of SPSS before version 10, but most of you can ignore this file.

3.3. Getting started ①

SPSS mainly uses two windows: the data editor (this is where you input your data and carry out statistical functions) and the viewer (this is where the results of any analysis appear).[1] There are several additional windows that can be activated such as the SPSS Syntax Editor (see section 3.7), which allows you to enter SPSS commands manually (rather than using the window-based menus). For many people, the syntax window is redundant because you can carry out most analyses by clicking merrily with your mouse. However, there are various additional functions that can be accessed using syntax and in many situations it can save time. This is why sick individuals who enjoy statistics find numerous uses for syntax and start dribbling excitedly when discussing it. Because I wish to drown in a pool of my own excited dribble, there are some sections of the book where I'll force you to use it.

Once SPSS has been activated, a start-up window will appear (see Figure 3.2), which allows you to select various options. If you already have a data file on disk that you would like to open then select *Open an existing data source* by clicking on the ○ so that it looks like ◉: this is the default option. In the space underneath this option there will be a list of recently used data files that you can select with the mouse. To open a selected file click on OK. If you want to open a data file that isn't in the list then simply select *More Files* …

[1] There is also the SmartViewer window, which actually isn't very smart, but more on that later.

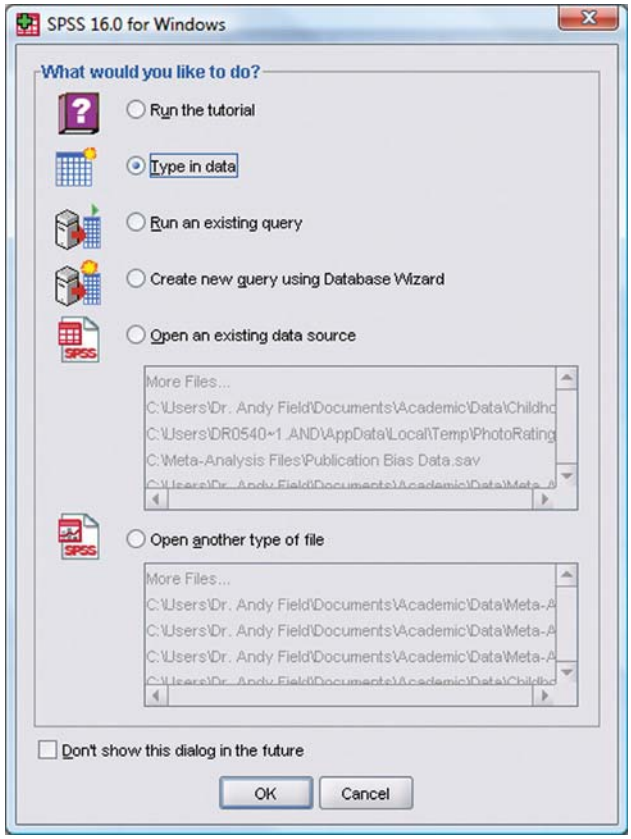


FIGURE 3.2
The start-up window of SPSS

with the mouse and click on OK. This will open a standard explorer window that allows you to browse your computer and find the file you want (see section 3.9). Now it might be the case that you want to open something other than a data file, for example a *viewer* document containing the results of your last analysis. You can do this by selecting *Open another type of file* by clicking on the ○ (so that it looks like ◉) and either selecting a file from the list or selecting *More Files* … and browsing your computer. If you're starting a new analysis (as we are here) then we want to type our data into a new data editor. Therefore, we need to select *Type in data* (by again clicking on the appropriate ○) and then clicking on OK. This will load a blank data editor window.

3.4. The data editor ①

The main SPSS window includes a data editor for entering data. This window is where most of the action happens. At the top of this screen is a menu bar similar to the ones you might have seen in other programs. Figure 3.3 shows this menu bar and the data editor. There are several menus at the top of the screen (e.g. File Edit View) that can be activated by using the computer mouse to move the on-screen arrow onto the desired menu and then pressing the left mouse button once (I'll refer to pressing this button as *clicking*). When you have clicked on a menu, a menu box will appear that displays a list of options that can be activated by moving the on-screen arrow so that it is pointing at the desired option and then clicking with the mouse. Often, selecting an option from a menu makes a window appear; these windows are referred to as *dialog boxes*. When referring to selecting options in a menu I will use images to notate the menu paths; for example, if I were to say that you should select the *Save As* … option in the *File* menu, you will see File Save As....

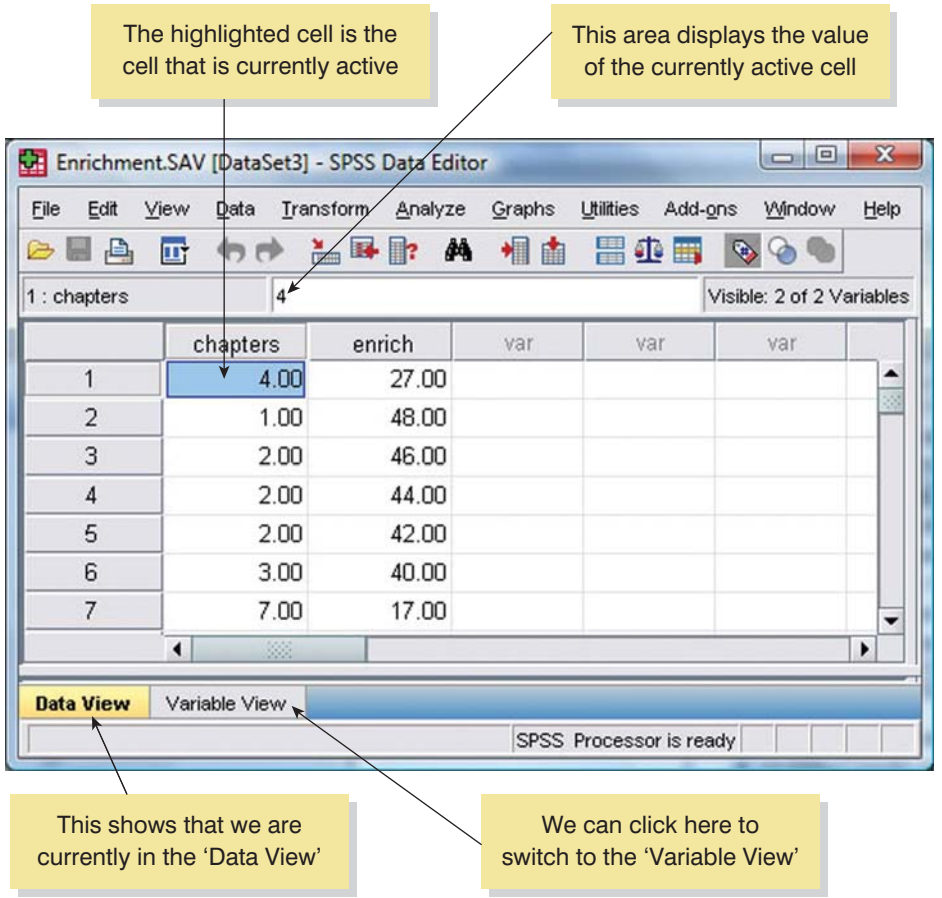


FIGURE 3.3
The SPSS Data Editor

The data editor has two views: the **data view** and the **variable view**. The data view is for entering data into the data editor, and the variable view allows us to define various characteristics of the variables within the data editor. At the bottom of the data editor, you should notice that there are two tabs labelled 'Data View' and 'Variable View' (Data View Variable View) and all we do to switch between these two views is click on these tabs (the highlighted tab tells you which view you're in, although it will be obvious). Let's look at some general features of the data editor, features that don't change when we switch between the data view and the variable view. First off, let's look at the menus.

In many computer packages you'll find that within the menus some letters are underlined: these underlined letters represent the *keyboard shortcut* for accessing that function. It is possible to select many functions without using the mouse, and the experienced keyboard user may find these shortcuts faster than manoeuvring the mouse arrow to the appropriate place on the screen. The letters underlined in the menus indicate that the option can be obtained by simultaneously pressing *Alt* on the keyboard and the underlined letter. So, to access the *Save As…* option, using only the keyboard, you should press *Alt* and F on the keyboard simultaneously (which activates the *File* menu), then, keeping your finger on the *Alt* key, press A (which is the underlined letter).[2]

Below is a brief reference guide to each of the menus and some of the options that they contain. This is merely a summary and we will discover the wonders of each menu as we progress through the book:

[2] If these underlined letters are not visible (in Windows XP they seemed to disappear, but in Vista they appear to have come back) then try pressing *Alt* and then the underlined letters should become visible.

- File This menu allows you to do general things such as saving data, graphs or output. Likewise, you can open previously saved files and print graphs data or output. In essence, it contains all of the options that are customarily found in *File* menus.
- Edit This menu contains edit functions for the data editor. In SPSS it is possible to *cut* and *paste* blocks of numbers from one part of the data editor to another (which can be very handy when you realize that you've entered lots of numbers in the wrong place). You can also use the Options... to select various preferences such as the font that is used for the output. The default preferences are fine for most purposes.
- View This menu deals with system specifications such as whether you have grid lines on the data editor, or whether you display value labels (exactly what value labels are will become clear later).
- Data This menu allows you to make changes to the data editor. The important features are Insert Variable, which is used to insert a new variable into the data editor (i.e. add a column); Insert Cases, which is used to add a new row of data between two existing rows of data; Split File..., which is used to split the file by a grouping variable (see section 5.4.3); and Select Cases..., which is used to run analyses on only a selected sample of cases.
- Transform You should use this menu if you want to manipulate one of your variables in some way. For example, you can use *recode* to change the values of certain variables (e.g. if you wanted to adopt a slightly different coding scheme for some reason) – see SPSS Tip 7.1. The *compute* function is also useful for transforming data (e.g. you can create a new variable that is the average of two existing variables). This function allows you to carry out any number of calculations on your variables (see section 5.7.3).
- Analyze The fun begins here, because the statistical procedures lurk in this menu. Below is a brief guide to the options in the statistics menu that will be used during the course of this book (this is only a small portion of what is available):
  - Descriptive Statistics This menu is for conducting descriptive statistics (mean, mode, median, etc.), frequencies and general data exploration. There is also a command called *crosstabs* that is useful for exploring frequency data and performing tests such as chi-square, Fisher's exact test and Cohen's kappa.
  - Compare Means This is where you can find *t*-tests (related and unrelated – Chapter 9) and one-way independent ANOVA (Chapter 10).
  - General Linear Model This menu is for complex ANOVA such as two-way (unrelated, related or mixed), one-way ANOVA with repeated measures and multivariate analysis of variance (MANOVA) – see Chapters 11, 12, 13, 14, and 16.
  - Mixed Models This menu can be used for running multilevel linear models (MLMs). At the time of writing I know absolutely nothing about these, but seeing as I've promised to write a chapter on them I'd better go and do some reading. With luck you'll find a chapter on it later in the book, or 30 blank sheets of paper. It could go either way.
  - Correlate It doesn't take a genius to work out that this is where the correlation techniques are kept! You can do bivariate correlations such as Pearson's *R*, Spearman's rho (ρ) and Kendall's tau (τ) as well as partial correlations (see Chapter 6).
  - Regression There are a variety of regression techniques available in SPSS. You can do simple linear regression, multiple linear regression (Chapter 7) and more advanced techniques such as logistic regression (Chapter 8).
  - Loglinear Loglinear analysis is hiding in this menu, waiting for you, and ready to pounce like a tarantula from its burrow (Chapter 13).

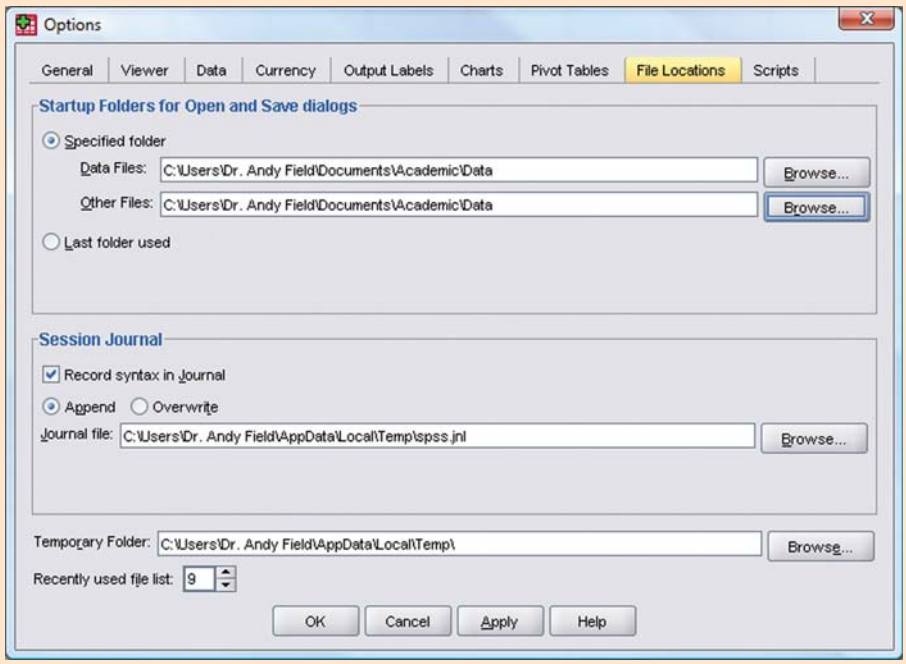
- Data Reduction You'll find factor analysis here (Chapter 17).
- Scale Here you'll find reliability analysis (Chapter 17).
- Nonparametric Tests There are a variety of non-parametric statistics available such as the chi-square goodness-of-fit statistic, the binomial test, the Mann–Whitney test, the Kruskal–Wallis test, Wilcoxon's test and Friedman's ANOVA (Chapter 15).

- **Graphs** SPSS has some graphing facilities and this menu is used to access the Chart Builder (see Chapter 4). The types of graphs you can do include: bar charts, histograms, scatterplots, box–whisker plots, pie charts and error bar graphs to name but a few.
- **Window** This menu allows you to switch from window to window. So, if you're looking at the output and you wish to switch back to your data sheet, you can do so using this menu. There are icons to shortcut most of the options in this menu so it isn't particularly useful.
- **Utilities** In this menu there is an option, Data File Comments..., that allows you to comment on your data set. This can be quite useful because you can write yourself notes about from where the data come, or the date they were collected and so on.
- **Add-ons** SPSS sells several add-ons that can be accessed through this menu. For example, SPSS has a program called *Sample Power* that computes the sample size required for studies, and power statistics (see section 2.6.5). However, because most people won't have these add-ons (including me) I'm not going to discuss them in the book.
- **Help** This is an invaluable menu because it offers you online help on both the system itself and the statistical tests. The statistics help files are fairly incomprehensible at times (the program is not designed to teach you statistics) and are certainly no substitute for acquiring a good book like this – erm, no, I mean acquiring a good knowledge of your own. However, they can get you out of a sticky situation.

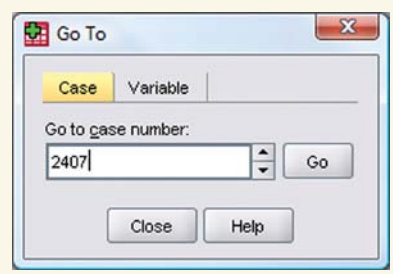
SPSS TIP 3.1 Save time and avoid RSI ①

By default, when you try to open a file from SPSS it will go to the directory in which the program is stored on your computer. This is fine if you happen to store all of your data and output in that folder, but if not then you will find yourself spending time navigating around your computer trying to find your data. If you use SPSS as much as I do then this has two consequences: (1) all those seconds have added up and I have probably spent weeks navigating my computer when I could have been doing something useful like playing my drum kit; (2) I have increased my chances of getting RSI in my wrists, and if I'm going to get RSI in my wrists I can think of more enjoyable ways to achieve it than navigating my computer (drumming again, obviously). Well, we can get around this by telling SPSS where we'd like it to start looking for data files. Select Edit Options... to open the *Options* dialog box below and select the *File Locations* tab.

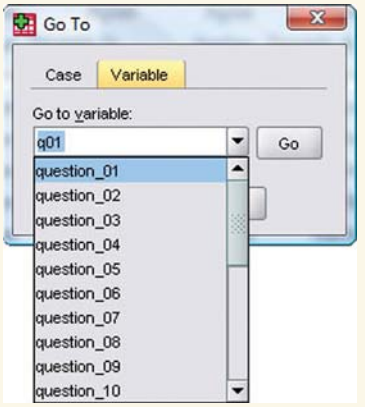
This dialog box allows you to select a folder in which SPSS will initially look for data files and other files. For example, I keep all of my data files in a single folder called, rather unimaginatively, 'Data'. In the dialog box here I have clicked on Browse... and then navigated to my data folder. SPSS will use this as the default location when I try to open files and my wrists are spared the indignity of RSI. You can also select the option for SPSS to use the *Last folder used,* in which case SPSS remembers where you were last time it was loaded and uses that folder when you try to open up data files.



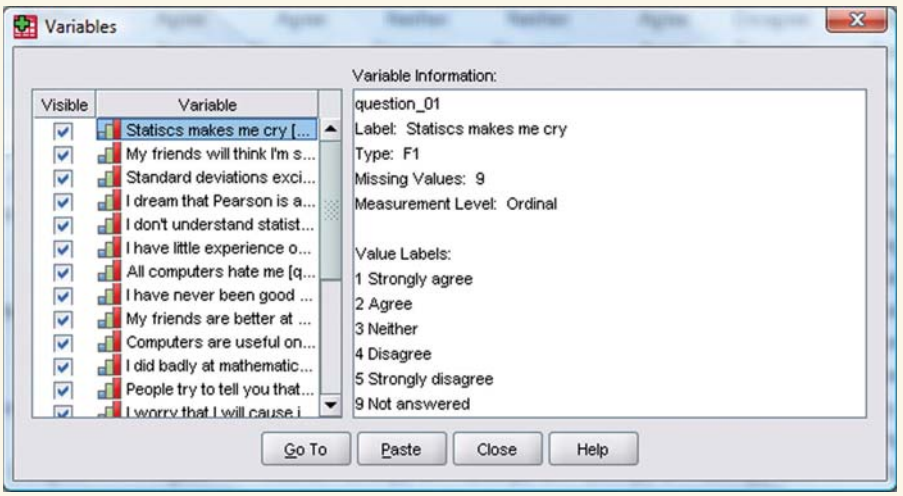
As well as the menus there is also a set of *icons* at the top of the data editor window (see Figure 3.3) that are shortcuts to specific, frequently used, facilities. All of these facilities can be accessed via the menu system but using the icons will save you time. Below is a brief list of these icons and their functions:

	This icon gives you the option to open a previously saved file (if you are in the data editor SPSS assumes you want to open a data file; if you are in the output viewer, it will offer to open a viewer file).
	This icon allows you to save files. It will save the file you are currently working on (be it data or output). If the file hasn't already been saved it will produce the *Save Data As* dialog box.
	This icon activates a dialog box for printing whatever you are currently working on (either the data editor or the output). The exact print options will depend on the printer you use. By default SPSS will print everything in the output window so a useful way to save trees is to print only a selection of the output (see SPSS Tip 3.4).
	Clicking on this icon will activate a list of the last 12 dialog boxes that were used. From this list you can select any box from the list and it will appear on the screen. This icon makes it easy for you to repeat parts of an analysis.
	This icon looks a bit like your data have had a few too many beers and have collapsed in the gutter by the side of the road. The truth is considerably less exciting: it enables you to go directly to a case (a case is a row in the data editor and represents something like a participant, an organism or a company). This button is useful if you are working on large data files: if you were analysing a survey with 3000 respondents it would get pretty tedious scrolling down the data sheet to find participant 2407's responses. This icon can be used to skip directly to a case. Clicking on this icon activates a dialog box in which you type the case number required (in our example 2407): 

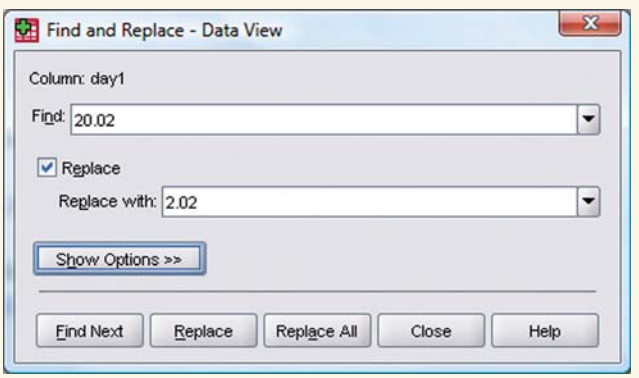
This icon activates a function that is similar to the previous one except that you can skip directly to a variable (i.e. a column in the data editor). As before, this is useful when working with big data files in which you have many columns of data. In the example below, we have a data file with 23 variables and each variable represents a question on a questionnaire and is named accordingly (we'll use this data file, **SAQ.sav**, in Chapter 17). We can use this icon to activate the *Go To* dialog box, but this time to find a variable. Notice that a drop-down box lists the first 10 variables in the data editor but you can scroll down to go to others.



This icon implies to me (what with the big question mark and everything) that you click on it when you're looking at your data scratching your head and thinking 'how in Satan's name do I analyse these data?' The world would be a much nicer place if clicking on this icon answered this question, but instead the shining diamond of hope is snatched cruelly from you by the cloaken thief that is SPSS. Instead, clicking on this icon opens a dialog box that shows you the variables in the data editor and summary information about each one. The dialog box below shows the information for the file that we used for the previous icon. We have selected the first variable in this file, and we can see the variable name (question_01), the label (Statistics makes me cry), the measurement level (ordinal), and the value labels (e.g. the number 1 represents the response of 'strongly agree').



This icon doesn't allow you to spy on your neighbours (unfortunately), but it does enable you to search for words or numbers in your data file and output window. In the data editor it will search within the variable (column) that is currently active. This option is useful if, for example, you realize from a graph of your data that you have typed 20.02 instead of 2.02 (see section 4.4): you can simply search for 20.02 within that variable and replace that value with 2.02:



	Clicking on this icon inserts a new case in the data editor (so it creates a blank row at the point that is currently highlighted in the data editor). This function is very useful if you need to add new data at a particular point in the data editor.
	Clicking on this icon creates a new variable to the left of the variable that is currently active (to activate a variable simply click once on the name at the top of the column).
	Clicking on this icon is a shortcut to the Data > Split File... function (see section 5.4.3). There are often situations in which you might want to analyse groups of cases separately. In SPSS we differentiate groups of cases by using a coding variable (see section 3.4.2.3), and this function lets us divide our output by such a variable. For example, we might test males and females on their statistical ability. We can code each participant with a number that represents their gender (e.g. 1 = female, 0 = male). If we then want to know the mean statistical ability of each gender we simply ask the computer to split the file by the variable **Gender**. Any subsequent analyses will be performed on the men and women separately. There are situations across many disciplines where this might be useful: sociologists and economists might want to look at data from different geographic locations separately, biologists might wish to analyse different groups of mutated mice, and so on.
	This icon shortcuts to the Data > Weight Cases... function. This function is necessary when we come to input frequency data (see section 18.5.2) and is useful for some advanced issues in survey sampling.
	This icon is a shortcut to the Data > Select Cases... function. If you want to analyse only a portion of your data, this is the option for you! This function allows you to specify what cases you want to include in the analysis. There is a Flash movie on the companion website that shows you how to select cases in your data file.
	Clicking on this icon will either display or hide the value labels of any coding variables. We often group people together and use a coding variable to let the computer know that a certain participant belongs to a certain group. For example, if we coded gender as 1 = female, 0 = male then the computer knows that every time it comes across the value 1 in the **Gender** column, that person is a female. If you press this icon, the coding will appear on the data editor rather than the numerical values; so, you will see the words *male* and *female* in the **Gender** column rather than a series of numbers. This idea will become clear in section 3.4.2.3.

3.4.1. Entering data into the data editor ①

When you first load SPSS it will provide a blank data editor with the title *Untitled1* (this of course is daft because once it has been given the title 'untitled' it ceases to be untitled!). When inputting a new set of data, you must input your data in a logical way. The SPSS Data Editor is arranged such that *each row represents data from one entity while each column represents a variable*. There is no discrimination between independent and dependent variables: both types should be placed in a separate column. The key point is that each row represents one entity's data (be that entity a human, mouse, tulip, business, or water sample). Therefore, any information about that case should be entered across the data editor. For example, imagine you were interested in sex differences in perceptions of pain created by hot and cold stimuli. You could place some people's hands in a bucket of very cold water for a minute and ask them to rate how painful they thought the experience was on a scale of 1 to 10. You could then ask them to hold a hot potato and again measure their perception of pain. Imagine I was a participant. You would have a single row representing my data, so there would be a different column for my name, my gender, my pain perception for cold water and my pain perception for a hot potato: Andy, male, 7, 10.

SPSS TIP 3.2 Entering data ①

There is a simple rule for how variables should be placed in the SPSS Data Editor: data from different things go in different rows of the data editor, whereas data from the same things go in different columns of the data editor. As such, each person (or mollusc, goat, organization, or whatever you have measured) is represented in a different row. Data within each person (or mollusc etc.) go in different columns. So, if you've prodded your mollusc, or human, several times with a pencil and measured how much it twitches as an outcome, then each prod will be represented by a column.

In experimental research this means that any variable measured with the same participants (a repeated measure) should be represented by several columns (each column representing one level of the repeated-measures variable). However, any variable that defines different groups of things (such as when a between-group design is used and different participants are assigned to different levels of the independent variable) is defined using a single column. This idea will become clearer as you learn about how to carry out specific procedures. This golden rule is broken in mixed models but until Chapter 19 we can overlook this annoying anomaly.

The column with the information about my gender is a grouping variable: I can belong to either the group of males or the group of females, but not both. As such, this variable is a between-group variable (different people belong to different groups). Rather than representing groups with words, in SPSS we have to use numbers. This involves assigning each group a number, and then telling SPSS which number represents which group. Therefore, between-group variables are represented by a single column in which the group to which the person belonged is defined using a number (see section 3.4.2.3). For example, we might decide that if a person is male then we give them the number 0, and if they're female we give them the number 1. We then have to tell SPSS that every time it sees a 1 in a particular column the person is a female, and every time it sees a 0 the person is a male. Variables that specify to which of several groups a person belongs can be used to split up data files (so in the pain example you could run an analysis on the male and female participants separately – see section 5.4.3).

Finally, the two measures of pain are a repeated measure (all participants were subjected to hot and cold stimuli). Therefore, levels of this variable (see SPSS Tip 3.2) can be entered in separate columns (one for pain to a hot stimulus and one for pain to a cold stimulus).

The data editor is made up of lots of *cells*, which are just boxes in which data values can be placed. When a cell is active it becomes highlighted in blue (as in Figure 3.3). You can move around the data editor, from cell to cell, using the arrow keys ← ↑ ↓ → (found on the right of the keyboard) or by clicking the mouse on the cell that you wish to activate. To enter a number into the data editor simply move to the cell in which you want to place the data value, type the value, then press the appropriate arrow button for the direction in which you wish to move. So, to enter a row of data, move to the far left of the row, type the value and then press → (this process inputs the value and then moves you into the next cell on the right).

The first step in entering your data is to create some variables using the 'Variable View' of the data editor, and then to input your data using the 'Data View' of the data editor. We'll go through these two steps by working through an example.

3.4.2. The 'Variable View' ①

Before we input any data into the data editor, we need to create the variables. To create variables we use the 'Variable View' of the data editor. To access this view click on the 'Variable View' tab at the bottom of the data editor (Data View | Variable View); the contents of the window will change (see Figure 3.4).

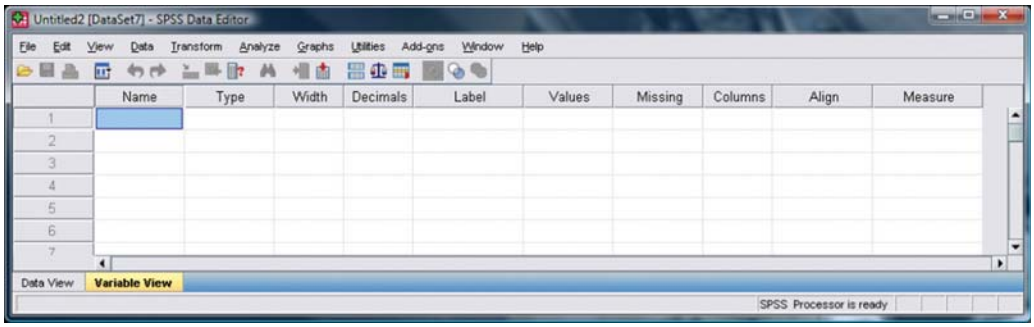


FIGURE 3.4 The 'Variable View' of the SPSS Data Editor

Every row of the variable view represents a variable, and you set characteristics of a particular variable by entering information into the labelled columns. You can change various characteristics of the variable by entering information into the following columns (play around and you'll get the hang of it):

Name	You can enter a name in this column for each variable. This name will appear at the top of the corresponding column in the data view, and helps you to identify variables in the data view. In current versions of SPSS you can more or less write what you like, but there are certain symbols you can't use (mainly symbols that have other uses in SPSS such as +, –, $, &) , and you can't use spaces. (Many people use a 'hard' space in variable names, which replaces the space with an underscore, for example, i.e. Andy_Field instead of Andy Field.) If you use a character that SPSS doesn't like you'll get an error message saying that the variable name is invalid when you click on a different cell, or try to move off the cell using the arrow keys.
Type	You can have different types of data. Mostly you will use **numeric variables** (which just means that the variable contains numbers – SPSS assumes this data type). You will come across **string variables**, which consist of strings of letters. If you wanted to type in people's names, for example, you would need to change the variable type to be string rather than numeric. You can also have **currency variables** (i.e. £s, $s, euro) and **date variables** (e.g. 21-06-1973)
Width	By default, when a new variable is created, SPSS sets it up to be *numeric* and to store 8 digits, but you can change this value by typing a new number in this column in the dialog box. Normally 8 digits is fine, but if you are doing calculations that need to be particularly precise you could make this value bigger.
Decimals	Another default setting is to have 2 decimal places displayed. (You'll notice that if you don't change this option then when you type in whole numbers to the data editor SPSS adds a decimal place with two zeros after it – this can be disconcerting initially!) If you want to change the number of decimal places for a given variable then replace the 2 with a new value or increase or decrease the values using ⇕, when in the cell that you want to change.
Label	The name of the variable (see above) has some restrictions on characters, and you also wouldn't want to use huge long names at the top of your columns (they become hard to read). Therefore, you can write a longer variable description in this column. This may seem pointless, but is actually one of the best habits you can get into (see SPSS Tip 3.3).
Values	This column is for assigning numbers to represent groups of people (see Section 3.4.2.3 below).
Missing	This column is for assigning numbers to missing data (see Section 3.4.3 below).
Columns	Enter a number into this column to determine the width of the column that is how many characters are displayed in the column. (This differs from Width , which determines the width of the variable itself – you could have a variable of 10 characters but by setting the column width to 8 you would only see 8 of the 10 characters of the variable in the data editor.) It can be useful to increase the column width if you have a string variable (section 3.4.2.1) that exceeds 8 characters, or a coding variable (section 3.4.2.3) with value labels that exceed 8 characters.
Align	You can use this column to select the alignment of the data in the corresponding column of the data editor. You can choose to align the data to the Left , Right or Center .
Measure	This is where you define the level at which a variable was measured (*Nominal*, *Ordinal* or *Scale* – section 1.5.1.2).

SPSS TIP 3.3 Naming variables ①

Why is it a good idea to take the time and effort to type in long variable names for your variables? Surely it's a waste of my time? In a way, I can understand why it would seem to be so, but as you go through your course accumulating data files, you will be grateful that you did. Imagine you had a variable called 'number of times I wanted to shoot myself during Andy Field's statistics lecture'; then you might have called the column in SPSS 'shoot'. If you don't add a label, SPSS will use this variable name in all of the output from an analysis. That's all well and good, but what happens in three weeks' time when you look at your data and output again? The chances are that you'll probably think 'What did shoot stand for? Number of shots at goal? Number of shots I drank?' Imagine the chaos you could get into if you had used an acronym for the variable 'workers attending new kiosk'. I have many data sets with variables called things like 'sftg45c', and if I didn't give them proper labels I would be in all sorts of trouble. Get into a good habit and label all of your variables!

Let's use the variable view to create some variables. Imagine we were interested in looking at the differences between lecturers and students. We took a random sample of five psychology lecturers from the University of Sussex and five psychology students and then measured how many friends they had, their weekly alcohol consumption (in units), their yearly income and how neurotic they were (higher score is more neurotic). These data are in Table 3.1.

TABLE 3.1 Some data with which to play

Name	Birth Date	Job	No. of Friends	Alcohol (Units)	Income	Neuroticism
Leo	17-Feb-1977	Lecturer	5	10	20,000	10
Martin	24-May-1969	Lecturer	2	15	40,000	17
Andy	21-Jun-1973	Lecturer	0	20	35,000	14
Paul	16-Jul-1970	Lecturer	4	5	22,000	13
Graham	10-Oct-1949	Lecturer	1	30	50,000	21
Carina	05-Nov-1983	Student	10	25	5,000	7
Karina	08-Oct-1987	Student	12	20	100	13
Doug	16-Sep-1989	Student	15	16	3,000	9
Mark	20-May-1973	Student	12	17	10,000	14
Mark	28-Mar-1990	Student	17	18	10	13

3.4.2.1. Creating a string variable ①

The first variable in our data set is the name of the lecturer/student. This variable consists of names; therefore, it is a *string variable*. To create this variable follow these steps:

1. Move the on-screen arrow (using the mouse) to the first white cell in the column labelled *Name*.
2. Type the word *Name*.
3. Move off this cell using the arrow keys on the keyboard (you can also just click on a different cell, but this is a very slow way of doing it).

You've just created your first variable! Notice that once you've typed a name, SPSS creates default settings for the variable (such as assuming it's numeric and assigning 2 decimal places). The problem is that although SPSS has assumed that we want a numeric variable (i.e. numbers), we don't; we want to enter people's names, namely a *string* variable. Therefore, we have to change the variable type. Move into the column labelled Type using the arrow keys on the keyboard. The cell will now look like this Numeric ... Click on ... to activate the dialog box in Figure 3.5. By default, SPSS selects the numeric variable type (Numeric) – see the left panel of Figure 3.5. To change the variable to a string variable, click on String and the dialog box will change to look like the right panel of Figure 3.5. You can choose how many characters you want in your string variable (i.e. the maximum number of characters you will type for a given case of data). The default is 8, which is fine for us because our longest name is only six letters; however, if we were entering surnames as well, we would need to increase this value. When you have finished, click on OK to return to the variable view.

Now because I want you to get into good habits, move to the cell in the Label column and type a description of the variable, such as 'Participant's First Name'. Finally, we can specify the level at which a variable was measured (see section 1.5.1.2) by going to the column labelled *Measure* and selecting either *Nominal*, *Ordinal* or *Scale* from the drop-down list. In this case, we have a string variable, so they represent only names of cases and provide no information about the order of cases, or the magnitude of one case compared to another. Therefore, we need to select Nominal.

Once the variable has been created, you can return to the data view by clicking on the 'Data View' tab at the bottom of the data editor (Data View Variable View). The contents of the window will change, and you'll notice that the first column now has the label *Name*. To enter the data, click on the white cell at the top of the column labelled *Name* and type the first name, 'Leo'. To register this value in this cell, we have to move to a different cell and because we are entering data down a column, the most sensible way to do this is to press the ↓ key on the keyboard. This action moves you down to the next cell, and the word 'Leo' should appear in the cell above. Enter the next name, 'Martin', and then press ↓ to move down to the next cell, and so on.

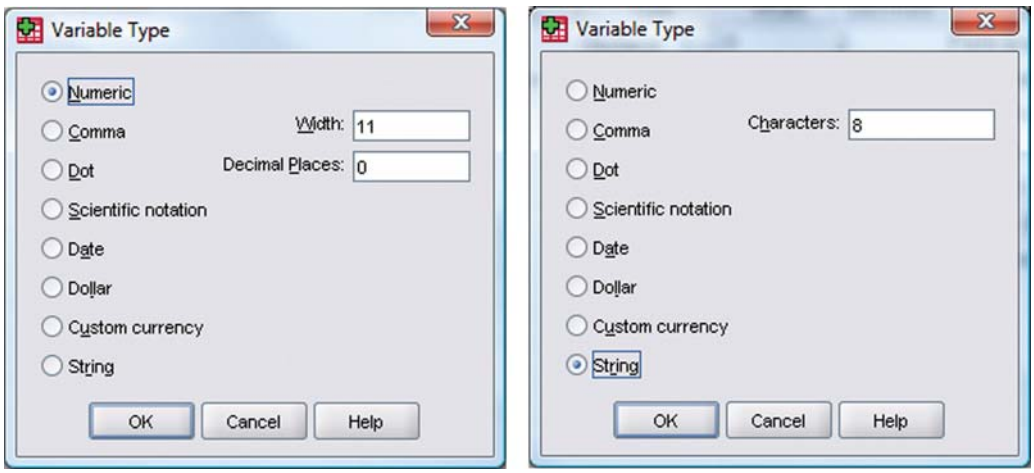


FIGURE 3.5 Defining a string variable in SPSS

3.4.2.2. Creating a date variable ①

Notice that the second column in our table contains dates (birth dates to be exact). To enter date variables into SPSS we use the same procedure as with the previous variable, except that we need to change the variable type. First, move back to the 'Variable View' using the tab at the bottom of the data editor (Data View | Variable View). As with the previous variable, move to the cell in row 2 of the column labelled *Name* (under the previous variable you created). Type the word 'Birth_Date' (note that we have used a hard space to separate the words). Move into the column labelled Type using the → key on the keyboard (SPSS will create default settings in the other columns). The cell will now look like this Numeric ... . Click on ... to activate the dialog box in Figure 3.6. By default, SPSS selects the numeric variable type (Numeric) – see the left panel of Figure 3.6. To change the variable to a date, click on Date and the dialog box will change to look like the right panel of Figure 3.6. You can then choose your preferred date format; being British, I am used to the days coming before the month and I have stuck with the default option of dd-mmm-yyyy (i.e. 21-Jun-1973), but Americans, for example, will be used to the month and date being the other way around and could select mm/dd/yyyy (06/21/1973). When you have selected a format for your dates, click on OK to return to the variable view. Finally, move to the cell in the column labelled *Label* and type 'Date of Birth'.

Now that the variable has been created, you can return to the data view by clicking on the 'Data View' tab (Data View | Variable View) and input the dates of birth. The second column now has the label *Birth_Date*; click on the white cell at the top of this column and type the first value, 17-Feb-1977. To register this value in this cell, move down to the next cell by pressing the ↓ key on the keyboard. Now enter the next date, and so on.

3.4.2.3. Creating coding variables ①

A coding variable (also known as a grouping variable) is a variable that uses numbers to represent different groups of data. As such, it is a *numeric variable*, but these numbers represent names (i.e. it is a nominal variable). These groups of data could be levels of a treatment variable in an experiment, different groups of people (men or women, an experimental group, or a control group, ethnic groups, etc.), different geographic locations, different organizations, etc.

In experiments, coding variables represent independent variables that have been measured between groups (i.e. different participants were assigned to different groups). If you were to run an experiment with one group of participants in an experimental condition and a different group of participants in a control group, you might assign the experimental group a code

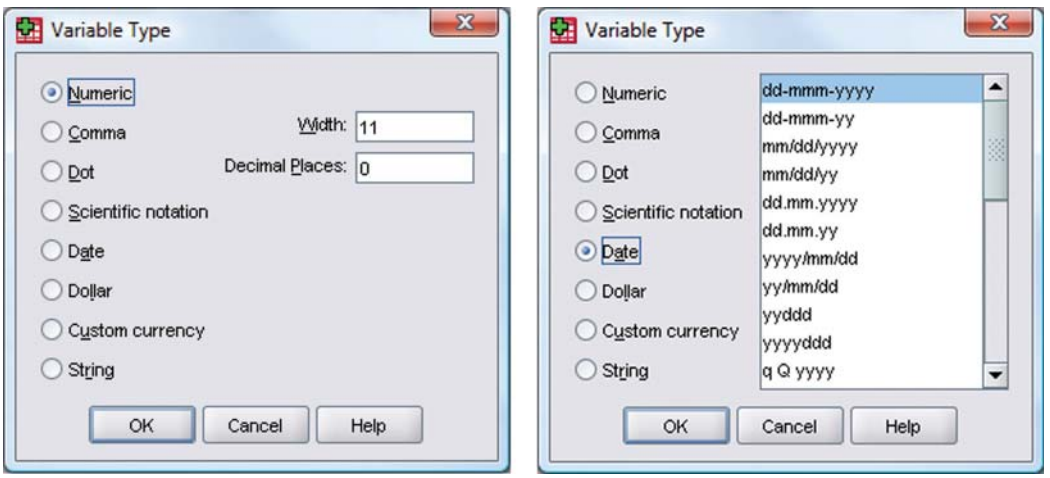


FIGURE 3.6 Defining variable types in SPSS

of 1 and the control group a code of 0. When you come to put the data into the data editor, you would create a variable (which you might call **group**) and type in the value 1 for any participants in the experimental group, and 0 for any participant in the control group. These codes tell SPSS that all of the cases that have been assigned the value 1 should be treated as belonging to the same group, and likewise for the cases assigned the value 0. In situations other than experiments, you might simply use codes to distinguish naturally occurring groups of people (e.g. you might give students a code of 1 and lecturers a code of 0).

We have a coding variable in our data: the one describing whether a person was a lecturer or student. To create this coding variable, we follow the steps for creating a normal variable, but we also have to tell SPSS which numeric codes have been assigned to which groups. So, first of all, return to the variable view (Data View **Variable View**) if you're not already in it and then move to the cell in the third row of the data editor and in the column labelled *Name* type a name (let's call it **Group**). I'm still trying to instil good habits, so move along the third row to the column called *Label* and give the variable a full description such as 'Is the person a lecturer or a student?' Then to define the group codes, move along the row to the column labelled Values and into this cell: None ... . Click on ... to access the *Value Labels* dialog box (see Figure 3.7).

The *Value Labels* dialog box is used to specify group codes. This can be done in three easy steps. First, click with the mouse in the white space next to where it says *Value* (or press *Alt* and *u* at the same time) and type in a code (e.g. 1). These codes are completely arbitrary; for the sake of convention people typically use 0, 1, 2, 3, etc., but in practice you could have a code of 495 if you were feeling particularly arbitrary. The second step is to click the mouse in the white space below, next to where it says *Value Label* (or press *Tab*, or *Alt* and *e* at the same time) and type in an appropriate label for that group. In Figure 3.7 I have already defined a code of 1 for the lecturer group, and then I have typed in 2 as my code and given this a label of *Student*. The third step is to add this coding to the list by clicking on Add . When you have defined all of your coding values you can click on Spelling... and SPSS will check your variable labels for spelling errors (which can be very handy if you are as bad at spelling as I am). To finish, click on OK ; if you click on OK and have forgotten to add your final coding to the list, SPSS will display a message warning you that any pending changes will be lost. In plain English

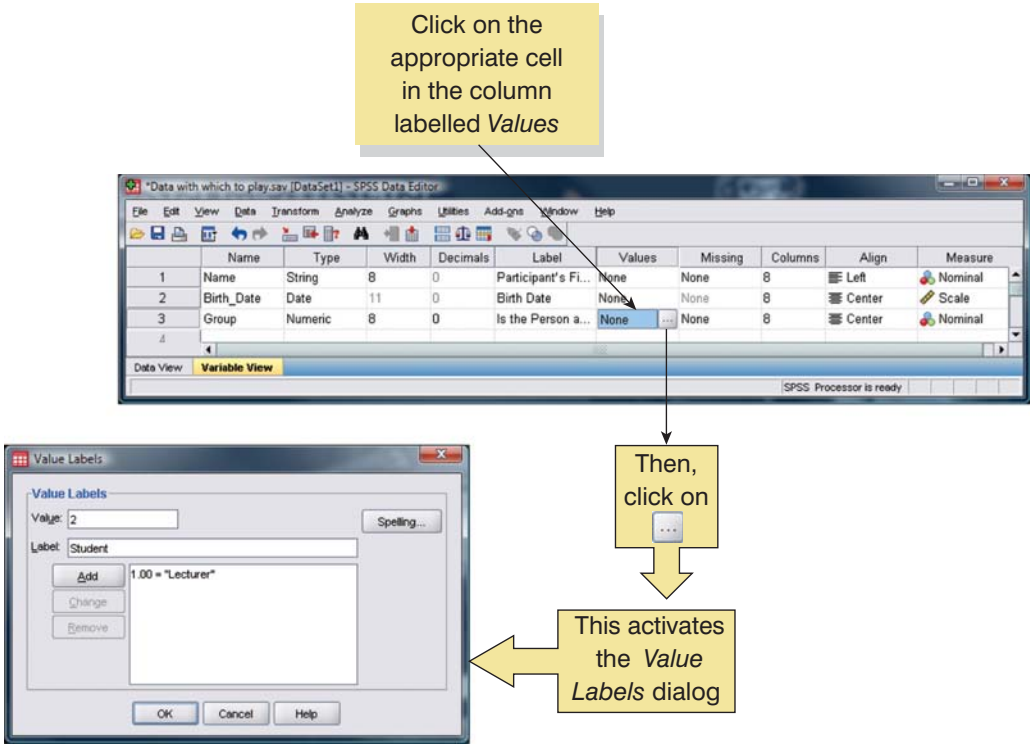


FIGURE 3.7 Defining coding variables and their values in SPSS

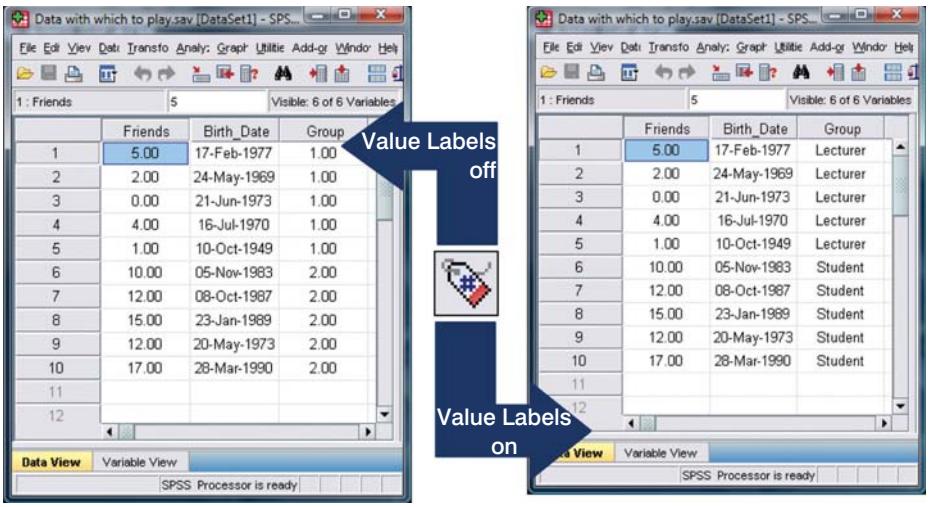


FIGURE 3.8
Coding values in the data editor with the value labels switched off and on

this simply tells you to go back and click on Add before continuing. Finally, coding variables always represent categories and so the level at which they are measured is nominal (or ordinal if the categories have a meaningful order). Therefore, you should specify the level at which the variable was measured by going to the column labelled *Measure* and selecting Nominal (or Ordinal if the groups have a meaningful order) from the drop-down list (see earlier).

Having defined your codes, switch to the data view and type these numerical values into the appropriate column (so if a person was a lecturer, type 1, but if they were a student then type 2). You can get SPSS to display the numeric codes, or the value labels that you assigned to them by clicking on (see Figure 3.8), which is pretty groovy. Figure 3.8 shows how the data should be arranged for a coding variable. Now remember that each row of the data editor represents data from one entity and in this example our entities were people (well, arguably in the case of the lecturers). The first five participants were lecturers whereas participants 6–10 were students.

This example should clarify why in experimental research grouping variables are used for variables that have been measured between participants: because by using a coding variable it is impossible for a participant to belong to more than one group. This situation should occur in a between-group design (i.e. a participant should not be tested in both the experimental and the control group). However, in repeated-measures designs (within subjects) each participant is tested in every condition and so we would not use this sort of coding variable (because each participant does take part in every experimental condition).

3.4.2.4. Creating a numeric variable ①

Numeric variables are the easiest ones to create because SPSS assumes this format for data. Our next variable is **No. of friends**; to create this variable we move back to the variable view using the tab at the bottom of the data editor (Data View Variable View). As with the previous variables, move to the cell in row 4 of the column labelled *Name* (under the previous variable you created). Type the word 'Friends'. Move into the column labelled Type using the → key on the keyboard. As with the previous variables we have created, SPSS has assumed that this is a numeric variable, so the cell will look like this Numeric ... . We can leave this as it is, because we do have a numeric variable.

Notice that our data for the number of friends has no decimal places (unless you are a very strange person indeed, you can't have 0.23 of a friend). Move to the Decimals column and type '0' (or decrease the value from 2 to 0 using) to tell SPSS that you don't want any decimal places.

Next, let's continue our good habit of naming variables and move to the cell in the column labelled *Label* and type 'Number of Friends'. Finally, we can specify the level at which a variable was measured (see section 1.5.1.2) by going to the column labelled *Measure* and selecting Scale from the drop-down list (this will have been done automatically actually, but it's worth checking).

SELF-TEST Why is the 'Number of Friends' variable a 'scale' variable?

Once the variable has been created, you can return to the data view by clicking on the 'Data View' tab at the bottom of the data editor (Data View | Variable View). The contents of the window will change, and you'll notice that the first column now has the label *Friends*. To enter the data, click on the white cell at the top of the column labelled *Friends* and type the first value, 5. To register this value in this cell, we have to move to a different cell and because we are entering data down a column, the most sensible way to do this is to press the ↓ key on the keyboard. This action moves you down to the next cell, and the number 5 should appear in the cell above. Enter the next number, 2, and then press ↓ to move down to the next cell, and so on.

SELF-TEST Having created the first four variables with a bit of guidance, try to enter the rest of the variables in Table 3.1 yourself.

3.4.3. Missing values ①

Although as researchers we strive to collect complete sets of data, it is often the case that we have missing data. Missing data can occur for a variety of reasons: in long questionnaires participants accidentally (or, depending on how paranoid you're feeling, deliberately just to piss you off) miss out questions; in experimental procedures mechanical faults can lead to a datum not being recorded; and in research on delicate topics (e.g. sexual behaviour) participants may exert their right not to answer a question. However, just because we have missed out on some data for a participant doesn't mean that we have to ignore the data we do have (although it sometimes creates statistical difficulties). Nevertheless, we do need to tell SPSS that a value is missing for a particular case. The principle behind missing values is quite similar to that of coding variables in that we choose a numeric value to represent the missing data point. This value tells SPSS that there is no recorded value for a participant for a certain variable. The computer then ignores that cell of the data editor (it does not use the value you select in the analysis). You need to be careful that the chosen code doesn't correspond to any naturally occurring data value. For example, if we tell the computer to regard the value 9 as a missing value and several participants genuinely scored 9, then the computer will treat their data as missing when, in reality, they are not.

To specify missing values you simply click in the column labelled Missing in the variable view and then click on ... to activate the *Missing Values* dialog box in Figure 3.9. By default SPSS assumes that no missing values exist, but if you do have data with missing

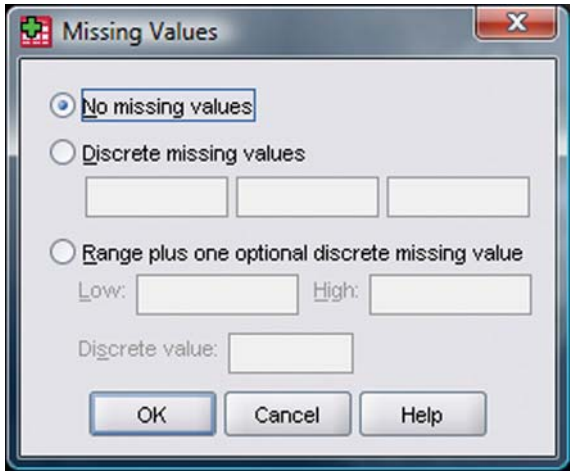


FIGURE 3.9
Defining missing values

values you can choose to define them in one of three ways. The first is to select discrete values (by clicking on the circle next to where it says *Discrete missing values*) which are single values that represent missing data. SPSS allows you to specify up to three discrete values to represent missing data. The reason why you might choose to have several numbers to represent missing values is that you can assign a different meaning to each discrete value. For example, you could have the number 8 representing a response of 'not applicable', a code of 9 representing a 'don't know' response, and a code of 99 meaning that the participant failed to give any response. As far as the computer is concerned it will ignore any data cell containing these values; however, using different codes may be a useful way to remind you of why a particular score is missing. Usually, one discrete value is enough and in an experiment in which attitudes are measured on a 100-point scale (so scores vary from 1 to 100) you might choose 666 to represent missing values because (1) this value cannot occur in the data that have been collected and (2) missing data create statistical problems, and you will regard the people who haven't given you responses as children of Satan! The second option is to select a range of values to represent missing data and this is useful in situations in which it is necessary to exclude data falling between two points. So, we could exclude all scores between 5 and 10. The final option is to have a range of values and one discrete value.

3.5. The SPSS Viewer ①

Alongside the SPSS Data Editor window, there is a second window known as the *SPSS Viewer*. The viewer window has come a long way; not that you care, I'm sure, but in days of old, this window displayed statistical results in a rather bland and drab font, graphs appeared in a window called the carousel (not nearly as exciting as it sounds), and all of the computations were done by little nerdy dinosaurs that lived inside your computer. These days, the SPSS Viewer displays everything you could want (well, OK, it doesn't display photos of your cat) and it's generally prettier to look at than it used to be. Sadly, however, my prediction in previous editions of this book that future versions of SPSS will include a tea-making facility in the viewer have not come to fruition (SPSS Inc. take note!).

Figure 3.10 shows the basic layout of the output viewer. On the right-hand side there is a large space in which the output is displayed. SPSS displays both graphs and the results of statistical analyses in this part of the viewer. It is also possible to edit graphs and to do this

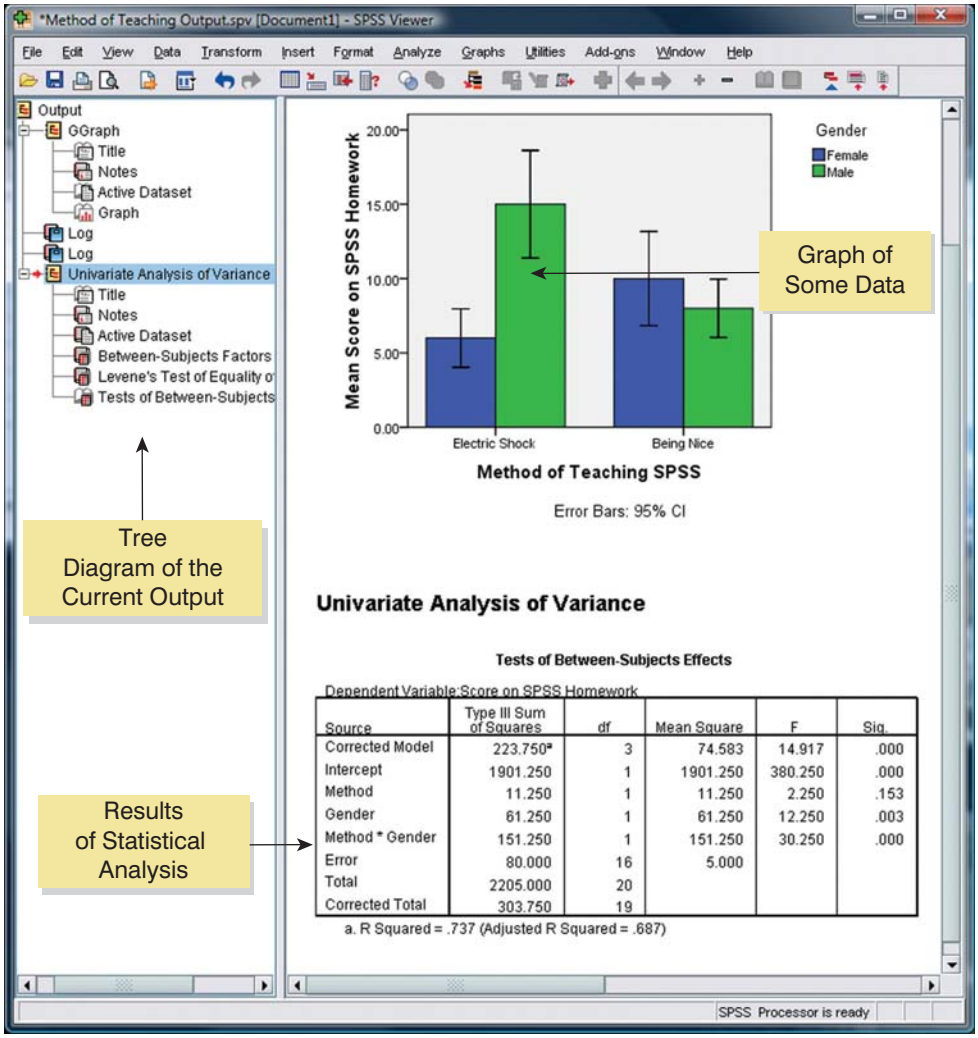


FIGURE 3.10
The SPSS Viewer

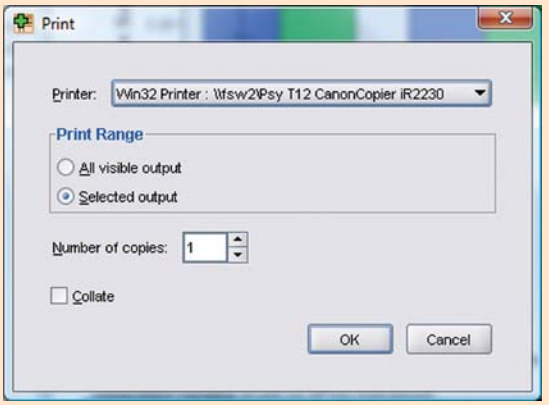
you simply double-click on the graph you wish to edit (this creates a new window in which the graph can be edited – see section 4.9). On the left-hand side of the output viewer there is a tree diagram illustrating the structure of the output. This tree diagram is useful when you have conducted several analyses because it provides an easy way of accessing specific parts of the output. The tree structure is fairly self-explanatory in that every time you do something in SPSS (such as drawing a graph or running a statistical procedure), it lists this procedure as a main heading.

In Figure 3.10 I conducted a graphing procedure followed by a univariate analysis of variance (ANOVA) and so these names appear as main headings. For each procedure there are a series of sub-headings that represent different parts of the analysis. For example, in the ANOVA procedure, which you'll learn more about later in the book, there are several sections to the output such as Levene's test (see section 5.6.1) and a table of the between-group effects (i.e. the *F*-test of whether the means are significantly different). You can skip to any one of these sub-components of the ANOVA output by clicking on the appropriate branch of the tree diagram. So, if you wanted to skip straight to the between-group effects you should move the on-screen arrow to the left-hand portion of the window and click where it says *Tests of Between-Subjects Effects*. This action will highlight this part of the output in the main part of the viewer (see SPSS Tip 3.4).

SPSS TIP 3.4 Printing and saving the planet ①

Rather than printing all of your SPSS output on reams of paper, you can help the planet by printing only a selection of the output. You can do this by using the tree diagram in the SPSS Viewer to select parts of the output for printing. For example, if you decided that you wanted to print out a graph but you didn't want to print the whole output, you can click on the word *GGraph* in the tree structure and that graph will become highlighted in the output. It is then possible through the print menu to select to print only the selected part of the output.

It is worth noting that if you click on a main heading (such as *Univariate Analysis of Variance*) then SPSS will highlight not only that main heading but all of the sub-components as well. This is useful for printing the results of a single statistical procedure.



There are several icons in the output viewer window that help you to do things quickly without using the drop-down menus. Some of these icons are the same as those described for the data editor window so I will concentrate mainly on the icons that are unique to the viewer window:

	As with the data editor window, this icon activates the print menu. However, when this icon is pressed in the viewer window it activates a menu for printing the output (see SPSS Tip 3.4).
	This icon returns you to the data editor in a flash!
	This icon takes you to the last output in the viewer (so it returns you to the last procedure you conducted).
	This icon *promotes* the currently active part of the tree structure to a higher branch of the tree. For example, in Figure 3.10 the *Tests of Between-Subjects Effects* are a sub-component under the heading of *Univariate Analysis of Variance*. If we wanted to promote this part of the output to a higher level (i.e. to make it a main heading) then this is done using this icon.
	This icon is the opposite of the above in that it *demotes* parts of the tree structure. For example, in Figure 3.10 if we didn't want the *Univariate Analysis of Variance* to be a unique section we could select this heading and demote it so that it becomes part of the previous heading (the *Graph* heading). This button is useful for combining parts of the output relating to a specific research question.
	This icon collapses parts of the tree structure, which simply means that it hides the sub-components under a particular heading. For example, in Figure 3.10 if we selected the heading *Univariate Analysis of Variance* and pressed this icon, all of the sub-headings would disappear. The sections that disappear from the tree structure don't disappear from the output itself; the tree structure is merely condensed. This can be useful when you have been conducting lots of analyses and the tree diagram is becoming very complex.
	This icon expands any collapsed sections. By default all of the main headings are displayed in the tree diagram in their expanded form. If, however, you have opted to collapse part of the tree diagram (using the icon above) then you can use this icon to undo your dirty work.
	This icon and the following one allow you to show and hide parts of the output itself. So you can select part of the output in the tree diagram and click on this icon and that part of the output will disappear. It isn't erased, but it is hidden from view. This icon is similar to the collapse icon listed above except that it affects the output rather than the tree structure. This is useful for hiding less relevant parts of the output.

	This icon undoes the previous one, so if you have hidden a selected part of the output from view and you click on this icon, that part of the output will reappear. By default, all parts of the output are shown, so this icon is not active; it will become active only once you have hidden part of the output.
	Although this icon looks rather like a paint roller, unfortunately it does not paint the house for you. What it does do is to insert a new heading into the tree diagram. For example, if you had several statistical tests that related to one of many research questions you could insert a main heading and then demote the headings of the relevant analyses so that they all fall under this new heading.
	Assuming you had done the above, you can use this icon to provide your new heading with a title. The title you type in will actually appear in your output. So, you might have a heading like 'Research question number 1' which tells you that the analyses under this heading relate to your first research question.
	This final icon is used to place a text box in the output window. You can type anything into this box. In the context of the previous two icons, you might use a text box to explain what your first research question is (e.g. 'My first research question is whether or not boredom has set in by the end of the first chapter of my book. The following analyses test the hypothesis that boredom levels will be significantly higher at the end of the first chapter than at the beginning').

SPSS TIP 3.5 Funny numbers ①

You might notice that SPSS sometimes reports numbers with the letter 'E' placed in the mix just to confuse you. For example, you might see a value such as 9.612 E–02 and many students find this notation confusing. Well, this notation means 9.61×10^{-2} (which might be a more familiar notation, or could be even more confusing). OK, some of you are still confused. Well think of E–02 as meaning 'move the decimal place 2 places to the left', so 9.612 E–02 becomes 0.09612. If the notation read 9.612 E–01, then that would be 0.9612, and if it read 9.612 E–03, that would be 0.009612. Likewise, think of E+02 (notice the minus sign has changed) as meaning 'move the decimal place 2 places to the right'. So 9.612 E+02 becomes 961.2.

3.6. The SPSS SmartViewer ①

Progress is all well and good, but it usually comes at a price and with version 16 of SPSS the output viewer changed quite dramatically (not really in terms of what you see, but behind the scenes). When you save an output file in version 17, SPSS uses the file extension **.spv**, and it calls it an SPSS Viewer file. In versions of SPSS before version 16, the viewer documents were saved as SPSS output files (**.spo**). Why does this matter? Well, it matters mainly because if you try to open an SPSS Viewer file in a version earlier than 16, SPSS won't know what the hell it is and will scream at you (well, it won't open the file). Similarly in versions after 16, if you try to open an output file (**.spo**) from an earlier version of SPSS you won't be able to do it. I know what you're thinking: 'this is bloody SPSS madness – statistics is hard enough as it is without versions of SPSS not being compatible with each other; is it trying to drive me insane?' Well, yes, it probably is.

This is why when you install SPSS on your computer you can install the SPSS **SmartViewer**. The SPSS SmartViewer looks suspiciously like the normal SPSS Viewer but without the icons. In fact, it looks so similar that I'm not even going to put a screenshot of it in this book – we can live on the edge, just this once. So what's so smart about the SmartViewer? Actually bugger all, that's what; it's just a way for you to open and read your old pre-version 17 SPSS output files. So, if you're not completely new to SPSS, this is a useful program to have.

The Syntax Editor ③

SMART ALEX ONLY

I've mentioned earlier that sometimes it's useful to use SPSS syntax. This is a language of commands for carrying out statistical analyses and data manipulations. Most of the time you'll do the things you need to using SPSS dialog boxes, but SPSS syntax can be useful. For one thing there are certain things you can do with syntax that you can't do through dialog boxes (admittedly most of these things are fairly advanced, but there will be a few places in this book where I show you some nice tricks using syntax). The second reason for using syntax is if you often carry out very similar analyses on data sets. In these situations it is often quicker to do the analysis and save the syntax as you go along. Fortunately this is easily done because many dialog boxes in SPSS have a Paste button. When you've specified your analysis using the dialog box, if you click on this button it will paste the syntax into a syntax editor window for you. To open a syntax editor window simply use the menus File > New > Syntax and a blank syntax editor will appear as in Figure 3.11. In this window you can type your syntax commands into the command area. The rules of SPSS syntax can be quite frustrating; for example, each line has to end with a full stop and if you forget this full stop you'll get an error message. In fact, if you have made a syntax error, SPSS will produce an error message in the Viewer window that identifies the line in the syntax window in which the error occurred; notice that in the syntax window SPSS helpfully numbers each line so that you can find the line in which the error occurred easily.

As we go through the book I'll show you a few things that will give you a flavour of how syntax can be used. Most of you won't have to use it, but for those that do this flavour will hopefully be enough to start you on your way. The window also has a navigation area (rather like the Viewer window). When you have a large file of syntax commands this navigation area can be helpful for negotiating your way to the bit of syntax that you actually need. Once you've typed in your syntax you have to run it using the Run menu. Run > All will run all of the syntax in the window (clicking on ▸ will also do this), or you can highlight a selection of your syntax using the mouse and use Run > Selection to process the selected syntax. You can also run the current command by using Run > Current (or press *Ctrl* and *R* on the keyboard), or run all the syntax from the cursor to the end of the syntax window using Run > To End. Another thing to note is that in SPSS you can have several data files open at once. Rather than have a syntax window for each data file, which could get confusing, you can use the same syntax window, but select the data set that you want to run the syntax commands on before you run them using the drop-down list DataSet3.

EVERYBODY

OLIVER TWISTED

Please, Sir, can I have some more … syntax?

'I want to drown in a puddle of my own saliva' froths Oliver, 'tell me more about the syntax Window'. Really? Well, OK, there is a Flash movie on the companion website that guides you through how to use the syntax window.

3.8. Saving files ①

Although most of you should be familiar with how to save files, it is a vital thing to know and so I will briefly describe what to do. To save files simply use the icon (or use the

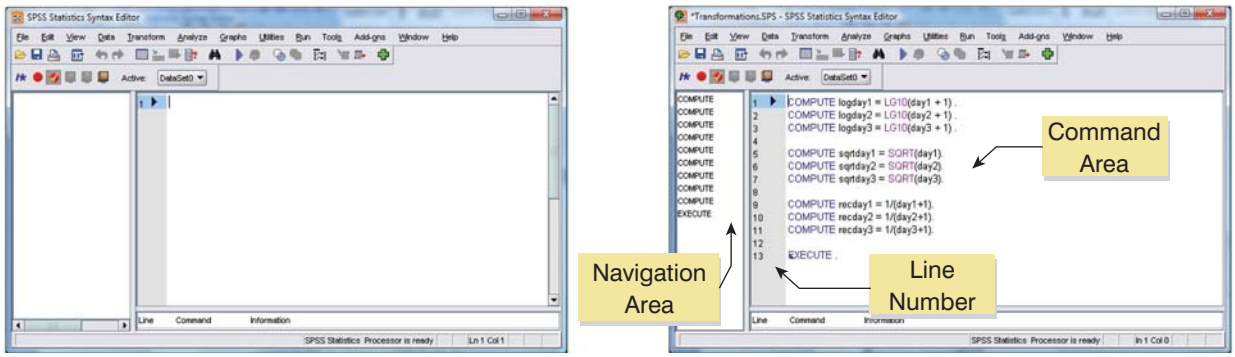


FIGURE 3.11 A new syntax window (top) and a syntax window with some syntax in it (bottom)

menus File Save or File Save As...). If the file is a new file, then clicking on this icon will activate the *Save As* ... dialog box (see Figure 3.12). If you are in the data editor when you select *Save As* … then SPSS will save the data file you are currently working on, but if you are in the viewer window then it will save the current output.

There are several features of the dialog box in Figure 3.12. First, you need to select a location at which to store the file. There are many types of locations where you can save data: the hard drive (or drives), a USB drive, a CD or DVD, etc. (you could have many other choices of location on your particular computer). The first thing to do is select a main location by double-clicking on it: your hard drive ( ), a CD or DVD ( ), or a USB stick or other external drive ( ). Once you have chosen a main location the dialog box will display all of the available folders on that particular device. Once you have selected a folder in which to save your file, you need to give your file a name. If you click in the space next to where it says *File name*, a cursor will appear and you can type a name. By default, the file will be saved in an SPSS format, so if it is a data file it will have the file extension **.sav**, if it is a viewer document it will have the file extension **.spv**, and if it is a syntax file it will have the file extension

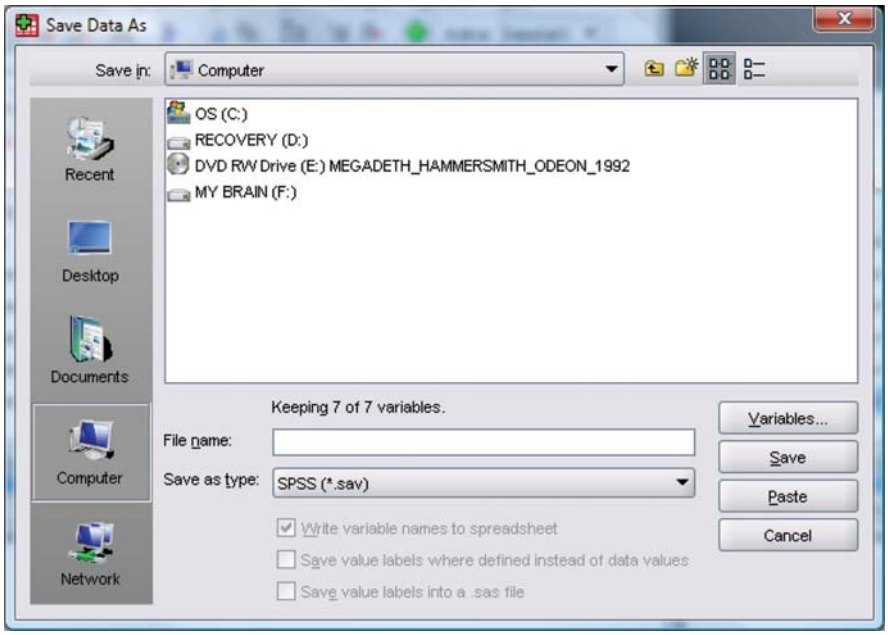


FIGURE 3.12 The *Save Data As* dialog box

.sps. However, you can save data in different formats such as Microsoft Excel files and tab-delimited text. To do this just click on SPSS (*.sav) and a list of possible file formats will be displayed. Click on the file type you require. Once a file has previously been saved, it can be saved again (updated) by clicking on . This icon appears in both the data editor and the viewer, and the file saved depends on the window that is currently active. The file will be saved in the location at which it is currently stored.

3.9. Retrieving a file ①

Throughout this book you will work with data files that you need to download from the companion website. It is, therefore, important that you know how to load these data files into SPSS. The procedure is very simple. To open a file, simply use the icon (or use the menus File Open Data...) to activate the dialog box in Figure 3.13. First, you need to find the location at which the file is stored. Navigate to wherever you downloaded the files from the website (a USB stick, or a folder on your hard drive). You should see a list of files and folders that can be opened. As with saving a file, if you are currently in the data editor then SPSS will display only SPSS data files to be opened (if you are in the viewer window then only output files will be displayed). If you use the menus and used the path File Open Data... then data files will be displayed, but if you used the path File Open Output... then viewer files will be displayed, and if you used File Open Syntax... then syntax files will be displayed (you get the general idea). You can open a folder by double-clicking on the folder icon. Once you have tracked down the required file you can open it either by selecting it with the mouse and then clicking on Open, or by double-clicking on the icon next to the file you want (e.g. double-clicking on). The data/output will then appear in the appropriate window. If you are in the data editor and you want to open a viewer file, then click on SPSS (*.sav) and a list of alternative file formats will be displayed. Click on the appropriate file type (viewer document (***.spv**), syntax file (***.sps**), Excel file (***.xls**), text file (***.dat**, ***.txt**)) and any files of that type will be displayed for you to open.

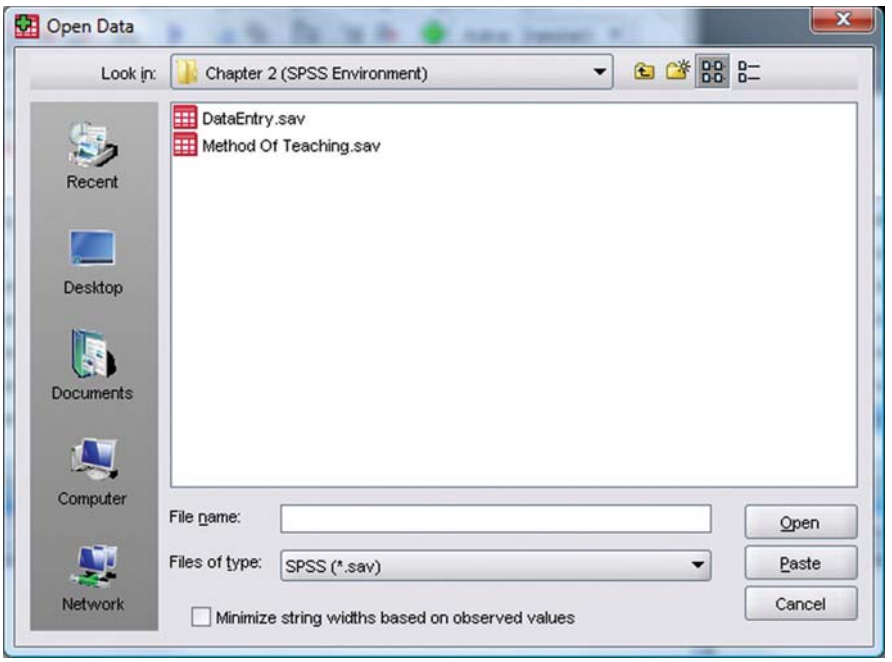


FIGURE 3.13
Dialog box to open a file

What have I discovered about statistics? ①

This chapter has provided a basic introduction to the SPSS environment. We've seen that SPSS uses two main windows: the data editor and the viewer. The data editor has both a data view (where you input the raw scores) and a variable view (where you define variables and their properties). The viewer is a window in which any output appears, such as tables, statistics and graphs. You also created your first data set by creating some variables and inputting some data. In doing so you discovered that we can code groups of people using numbers (coding variables) and discovered that rows in the data editor represent people (or cases of data) and columns represent different variables. Finally, we had a look at the syntax window and were told how to open and save files.

We also discovered that I was scared of my new school. However, with the help of Jonathan Land my confidence grew. With this new confidence I began to feel comfortable not just at school but in the world at large. It was time to explore.

Key terms that I've discovered

Currency variable
Data editor
Data view
Date variable
Numeric variable
SmartViewer
String variable
Syntax editor
Variable view
Viewer

Smart Alex's tasks

- **Task 1**: Smart Alex's first task for this chapter is to save the data that you've entered in this chapter. Save it somewhere on the hard drive of your computer (or a USB stick if you're not working on your own computer). Give it a sensible title and save it somewhere easy to find (perhaps create a folder called 'My Data Files' where you can save all of your files when working through this book).
- **Task 2**: Your second task is to enter the data that I used to create Figure 3.10. These data show the score (out of 20) for 20 different students, some of whom are male and some female, and some of whom were taught using positive reinforcement (being nice) and others who were taught using punishment (electric shock). Just to make it hard, the data should not be entered in the same way that they are laid out below:

Male		*Female*	
Electric Shock	*Being Nice*	*Electric Shock*	*Being Nice*
15	12	6	10
14	10	7	9
20	7	5	8
13	8	4	8
13	13	8	7

- **Task 3**: Research has looked at emotional reactions to infidelity and found that men get homicidal and suicidal and women feel undesirable and insecure (Shackelford, LeBlanc, & Drass, 2000). Let's imagine we did some similar research: we took some men and women and got their partners to tell them they had slept with someone else. We then took each person to two shooting galleries and each time gave them a gun and 100 bullets. In one gallery was a human-shaped target with a picture of their own face on it, and in the other was a target with their partner's face on it. They were left alone with each target for 5 minutes and the number of bullets used was measured. The data are below; enter them into SPSS and save them as **Infidelity.sav** (clue: they are not entered in the format in the table!).

Male		Female	
Partner's Face	Own Face	Partner's Face	Own Face
69	33	70	97
76	26	74	80
70	10	64	88
76	51	43	100
72	34	51	100
65	28	93	58
82	27	48	95
71	9	51	83
71	33	74	97
75	11	73	89
52	14	41	69
34	46	84	82

Answers can be found on the companion website.

Further reading

There are many good introductory SPSS books on the market that go through similar material to this chapter. Pallant's *SPSS survival manual* and Kinnear and Gray's *SPSS XX made simple* (insert a version number where I've typed XX because they update it regularly) are both excellent guides for people new to SPSS. There are many others on the market as well, so have a hunt around.

Online tutorials

The companion website contains the following Flash movie tutorials to accompany this chapter:

- Entering data
- Exporting SPSS output to Word
- Importing text data to SPSS
- Selecting cases
- The Syntax window
- The Viewer window
- Using Excel with SPSS