

Louise Marston

Introductory Statistics for Health and Nursing Using SPSS

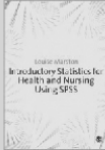
companion
website 



Introductory Statistics for Health and Nursing Using SPSS

COMPANION WEBSITE

Visit the companion website at www.uk.sagepub.com/marston to find datasets and additional questions.

Homepage	<h2>Introductory Statistics for Health and Nursing Using SPSS</h2>
About the Book	
Description	
Author Details	
Student Resources	
SPSS data sets	
Additional Questions	
Book Details	
	Louise Marston Welcome to the Companion Website for <i>Introductory Statistics for Health and Nursing Using SPSS</i> by Louise Marston This website contains the datasets in SPSS format used for the exercises in the book, with an explanation of the variables contained within them. About the Book This section contains details on the book and its author. Student Resources This section links to a variety of resources which have been developed to accompany <i>Introductory Statistics for Health and Nursing Using SPSS</i> , including: <ul style="list-style-type: none">• SPSS data sets• Additional Questions
Author Louise Marston	
Pub Date: Dec 2009 Pages: 264	
Click here for more information.	

© Sage Publications Ltd.

Introductory Statistics for Health and Nursing Using SPSS

Louise Marston



Los Angeles | London | New Delhi
Singapore | Washington DC

© Louise Marston 2010

First published 2010

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act, 1988, this publication may be reproduced, stored or transmitted in any form, or by any means, only with the prior permission in writing of the publishers, or in the case of reprographic reproduction, in accordance with the terms of licences issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to the publishers.

SAGE Publications Ltd
1 Oliver's Yard
55 City Road
London EC1Y 1SP

SAGE Publications Inc.
2455 Teller Road
Thousand Oaks, California 91320

SAGE Publications India Pvt Ltd
B 1/I 1 Mohan Cooperative Industrial Area
Mathura Road
New Delhi 110 044

SAGE Publications Asia-Pacific Pte Ltd
33 Pekin Street #02-01
Far East Square
Singapore 048763

Library of Congress Control Number: 2009927204

British Library Cataloguing in Publication data

A catalogue record for this book is available from
the British Library

ISBN 978-1-84787-482-5
ISBN 978-1-84787-483-2 (pbk)

Typeset by C&M Digital (P) Ltd, Chennai, India
Printed by MPG Books Group, Bodmin, Cornwall
Printed on paper from sustainable resources



Mixed Sources

Product group from well-managed
forests and other controlled sources
www.fsc.org Cert no. SA-COC-1565
© 1996 Forest Stewardship Council

For my father, Roger Marston
and
my mother, Joy Marston.

CONTENTS

Preface	viii
1 Getting Started with Data and SPSS	1
2 Data Management	22
3 Study Designs	42
4 Probability	54
5 Summary Statistics for Continuous Data	60
6 Summary Statistics for Categorical Data	90
7 Samples and Populations	102
8 Comparing Two Categorical Variables	112
9 Comparing Means	136
10 Non-Parametric Tests	154
11 Assessing Associations with a Continuous Outcome	173
12 Assessing Associations with a Categorical Outcome	200
Appendix 1: Datasets	213
Appendix 2: Solutions	215
Glossary	234
References	243
Index	247

PREFACE

This book came about through teaching MSc students in nursing and other health related disciplines, where I found there was no book that described the statistics I was teaching and set it with SPSS whilst giving examples related to health. This book assumes no knowledge of SPSS or statistics, and starts by describing features of data before introducing SPSS and then combines the two with data entry and management. As background to the research process, study designs and the concepts of samples and populations are explained and illustrated using published literature. Chapter 4, the final background chapter, is on probability, showing how simple concepts work before linking probability to health statistics. This book continues by describing the statistical techniques and tests commonly taught to students of nursing and other health related disciplines – both undergraduate and postgraduate. It gives examples of how each technique and test is carried out using SPSS. Each example shown uses data from real studies; some of which were collected by the students I taught. Real data are important because they show the patterns and deviations that students may see in their datasets. There are exercises at the end of each chapter so that students can test their understanding of the chapter. Some of the exercises are designed to be used with SPSS by analysing the data; giving practice with SPSS, data analysis and interpretation of the results.

I would like to express my thanks to Kiran Katikaneni, Carol Morant, Mathew Alfred, Maria Stein and Aminollah Ferdowsian for allowing me to use their data in this book. Also thank you to Alison Sherwin for allowing me to use a screenshot of her variable names. I would also like to thank Janet Peacock for allowing me to use data we have both been very close to for a number of years, and also for encouraging me to write the book. Finally, thank you to Julie Gilg and Gita Thakur for reading the book; and for spotting inconsistencies and typos. However, any that remain are my own.

Louise Marston
February 2009

1

GETTING STARTED WITH DATA AND SPSS

INTRODUCTION

When undertaking research studies, it is likely that data will have been collected. Much of this will be in a form suitable to be analysed statistically; that is, responses from a questionnaire or data collection sheet can be coded so that each response is represented by a number. When data have been coded, it can be entered to a data analysis program such as SPSS. SPSS is menu driven, making it easy to use.

This chapter will start by distinguishing between types of data. It will then move onto the situation of a pile of questionnaires that need to be coded before data are entered into SPSS. It will show the conventions for assigning codes to questionnaires before the data entry process. It will then go onto opening SPSS and giving a tour of features that will be explained during the course of this book. This will lead to how to set up a datasheet so that data can be entered into SPSS. The chapter then moves onto saving data and an introduction to SPSS syntax, which is mainly used as a means of recording the commands used for future analysis. Finally, actions associated with closing SPSS, encompassing saving, printing and exporting output are explained.


This chapter uses a questionnaire asking hotel employees about their knowledge of diabetes as well as the data collected from the resulting study.

THE AIMS OF THIS CHAPTER ARE:

- To distinguish between types of data.
- To learn how to code questionnaires/data collection sheets so the resulting data can be entered into SPSS.
- To show how to open SPSS.
- To guide the user round the windows and menus that comprise SPSS, so that data can be opened, entered and ultimately analysed.
- To explain how to save data, output and command syntax.

2 INTRODUCTORY STATISTICS FOR HEALTH AND NURSING USING SPSS

PRELIMINARY TO THIS CHAPTER

This  is a radio button. SPSS uses them when it requires one option to be selected.

TYPES OF DATA

Paramount to all data entry and analysis is the knowledge of what type of data a given variable is because different types of data are coded and ultimately analysed in different ways. SPSS gives three options regarding types of data: nominal, ordinal and scale.

Nominal data

These are categorical data that have no order. The categories within each variable are mutually exclusive: respondents can only fall into one category. For example, respondents can only be one ethnicity from a given list. Where a nominal variable has two categories, it is often referred to as dichotomous or binary.

Examples of nominal variables:

Gender – respondents can be male or female

Disease/health status – respondents can either have a disease or not

Marital status – respondents can only have one marital status at a given time: single, married, separated, divorced, widowed

Ordinal data

These are also categorical variables in which the categories are ordered.

Examples of ordinal variables:

Age group – for example, 30–39, 40–49, 50–59, 60+

Likert scales – strongly agree, agree, neither agree nor disagree, disagree, strongly disagree

Scale data

In SPSS this covers discrete and continuous data. Discrete data comprise variables that can only take integers (whole numbers).

Examples of discrete data:

Number of nights spent in hospital
 Number of courses of a given drug prescribed during the study period
 Age at last birthday
 Number of cigarettes smoked in a week

Continuous data can (in theory at least) take any value. However, this is usually restricted by the accuracy of the equipment used for measuring. For example, scales for weighing adult human weights rarely measure more accurately than whole kilograms and occasionally to one decimal place. This is also for practical reasons; there is little need to weigh adult humans to greater precision than the nearest kilogram or 100 grams (1 decimal place).

Examples of continuous data:

Blood pressure
 Body mass index (BMI)
 Lung function, for example peak expiratory flow rate (PEFR)

CODING QUESTIONNAIRES

Once data collection is complete, then the next task is to decide how to code each question so that it can easily be seen which values should be inputted into SPSS. This is necessary because numerical values are needed representing answers to questions on a questionnaire or other data collection sheet for SPSS to analyse the data. The initial task of deciding on coding is best done using an unused questionnaire so that all possible codes can be written on the questionnaire without confusion. In addition, it is a good idea to write the variable names on this questionnaire. These steps help the coding process so that individual codes do not have to be remembered and also provides a permanent record of the coding of the dataset.

Figures 1.1a, 1.1b and 1.1c show excerpts from a questionnaire aimed at hotel employees to discover their knowledge of diabetes and to find out whether employees consider hotels (as a workplace) to be appropriate places to conduct health promotion specifically aimed at type 2 diabetes. These excerpts have not been annotated with possible variable names, but have been annotated with coding. Variable names should not be long (ideally eight characters), but should be as descriptive as possible (when setting up an SPSS datasheet it is possible to give each variable a longer label, this will be explained later in this chapter). Variable names must be unique within a dataset.

Questions 2 and 3 (Figure 1.1a) produced nominal data; there is no ordering. Question 2 lists three options and participants were asked to select one, therefore a different code is needed for all options. As the data are nominal the numbers given

4 INTRODUCTORY STATISTICS FOR HEALTH AND NURSING USING SPSS

2. To the best of your knowledge can diabetes type 2 be delayed?

- | | |
|----------------------------------|---|
| <input type="radio"/> Yes | 1 |
| <input type="radio"/> No | 2 |
| <input type="radio"/> Don't know | 3 |

**3. Which of the following do you recognise as symptoms of diabetes type 2?
(tick all that apply)**

- | | | | |
|---|---|---|---|
| <input type="checkbox"/> Tiredness | 1 | <input type="checkbox"/> Excessive thirst | 1 |
| <input type="checkbox"/> Tension and depression | 1 | <input type="checkbox"/> Loss of weight | 1 |
| <input type="checkbox"/> Frequent urination | 1 | <input type="checkbox"/> Giddiness | 1 |
| <input type="checkbox"/> Slow healing of wounds | 1 | <input type="checkbox"/> Excessive hunger | 1 |

FIGURE 1.1A EXCERPT FROM A QUESTIONNAIRE – QUESTIONS GIVING NOMINAL DATA

10. Please state your weight in either kgs, lbs or stones and your height in cm or inches.

	Weight	Height
Weight in kgs, lbs, stns; Height in cm or inches	<input type="text"/>	<input type="text"/>

FIGURE 1.1B EXCERPT FROM A QUESTIONNAIRE – QUESTIONS GIVING CONTINUOUS DATA

as the codes do not necessarily have to be 1, 2 and 3, they could have equally been 0, 1, and 2 or any other three unique numbers. The convention is to code nominal data using consecutive numbers, starting with 0 or 1.

Question 3 allows the participant to select as many options as they want. Therefore, each potential response needs to be coded and entered as a separate variable. Therefore, although it is presented on the questionnaire as one question, when the data are entered into SPSS for analysis, there will be eight variables. With Question 3, if the participant has not selected a given option, it is assumed that they do not recognise that option as a symptom of type 2 diabetes so these are eight yes and no questions. The simplest way to code these is to code as 1 where the option has been selected and 0 otherwise, giving coding of 0 = no and 1 = yes.

Question 10 (Figure 1.1b) gives scale data. It requires the participant to put their actual height and weight in the two boxes provided (data for this study were collected through an online questionnaire so there were dropdown menus giving possible options). If the questionnaire was completed using a pen and paper, it would be advantageous to put the possible units beside the measure and instruct the participant to select the units they are giving their measurements in. The number of variables equal to the number of types of units used would be entered into SPSS. For example, with height, there may be two variables: height in metres and height in inches. Exact heights and weights would be entered into SPSS. Conversion into one set of units for each variable would take place using SPSS to ensure there are no human errors resulting from the calculations being done by hand.

12. Questions 12–16 ask about your opinion of health promotion programmes in the workplace. Please select the option that best represents your view.

	Believe strongly	Believe slightly	Not sure	Believe a little	Not at all
My company should help create greater awareness about diabetes prevention at work.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My workplace is a suitable environment for promoting diabetes prevention awareness programmes.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	1	2	3	4	5

FIGURE 1.1C EXCERPT FROM A QUESTIONNAIRE – QUESTIONS GIVING ORDINAL DATA

The questions in Figure 1.1c use likert scales in which participants are asked to select the response closest to their opinion. The possible responses are ordered through the spectrum of opinions. The responses range from ‘believe strongly’ to ‘not at all’ with a neutral category in the middle. As these options are ordered, the codes that will be entered into SPSS should also be ordered. Therefore, in this example, the responses are coded from 1 to 5.

Further information

For more information on coding see Chapter 2 of Peacock and Kerry (2007).

OPENING SPSS

Now there are some data to put into SPSS, the program can be opened. To do this, click on Start → Programs → SPSS for Windows → SPSS for Windows. When this has been done, the initial screen will appear. An example of this can be seen in Figure 1.2. A window appears on top of the SPSS Data Editor. If the data file required has been used recently, it will appear in the box containing recently used data files. Otherwise click Cancel → File → Open → Data... → find the data file required from where it was last saved. The file name will end .sav.

The Status Bar is the section at the bottom of the screen that usually says SPSS Processor is ready (Figure 1.2). This can be turned off by clicking on View → Status Bar.

When the dataset is open, the SPSS Data Editor will look similar to Figure 1.3. This is Data View. All variables have been coded to numbers. Each row represents a different participant.

Click on the Variable View tab at the bottom left of the screen (indicated in Figure 1.3), to get a screen like Figure 1.4. This is the screen where category codes and missing data codes can be viewed or declared to SPSS. In Figure 1.4 it can be seen that there is one row for each variable in the dataset. The meaning of each of the columns in Variable View and what should be placed in them will be shown later in this chapter in the section on entering data.

6 INTRODUCTORY STATISTICS FOR HEALTH AND NURSING USING SPSS

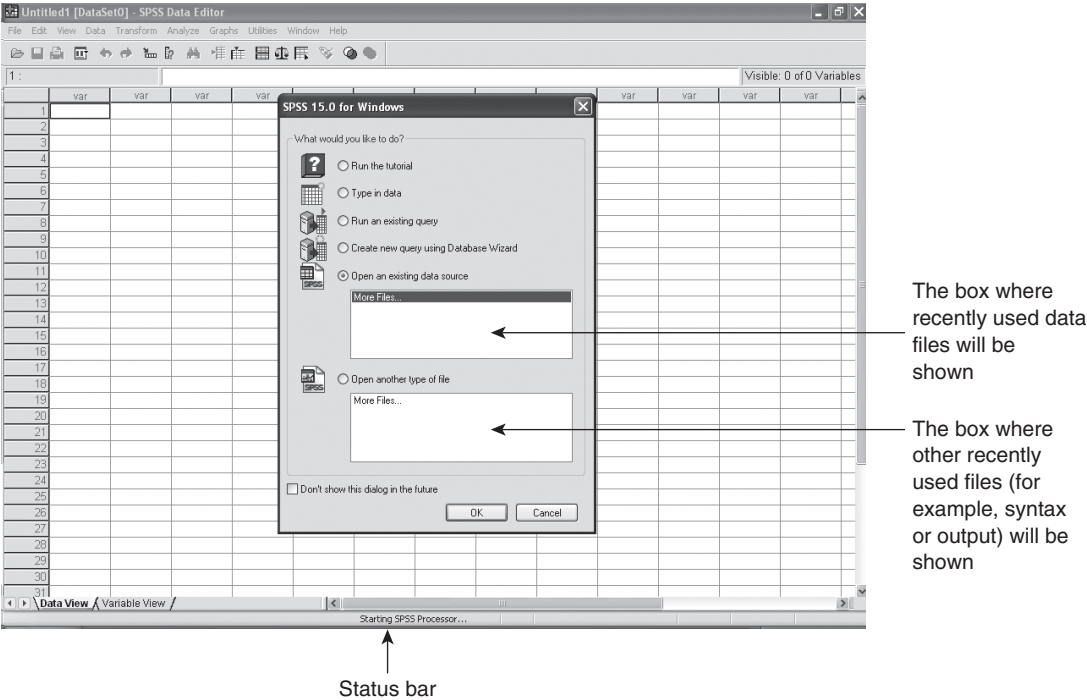


FIGURE 1.2 INITIAL SCREEN WHEN SPSS OPENS

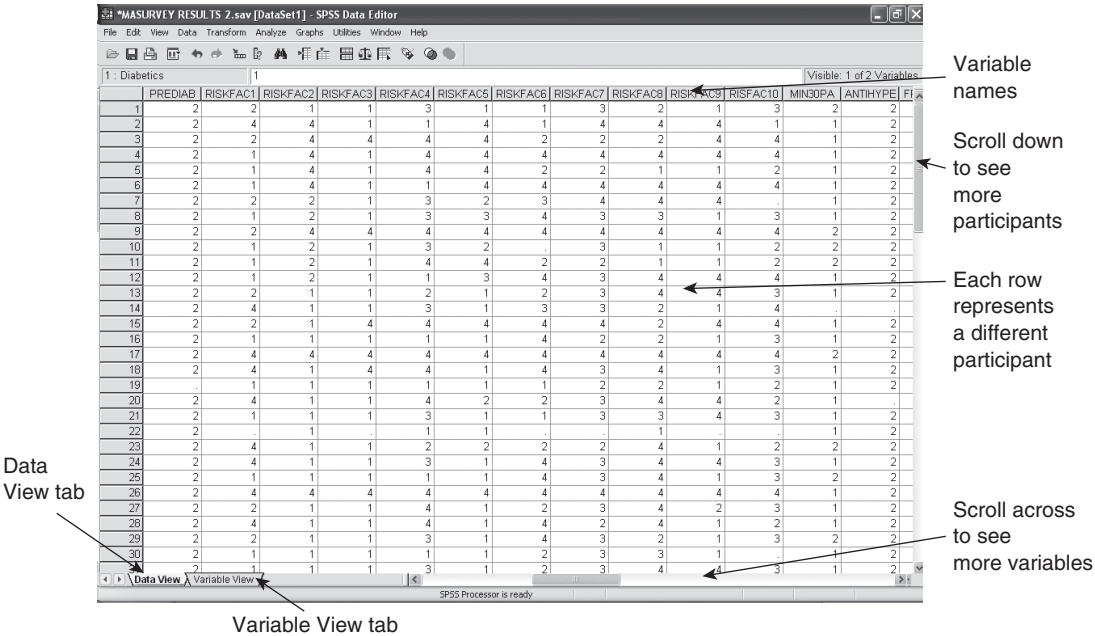


FIGURE 1.3 DATA VIEW

A TOUR OF THE SPSS DATA VIEW MENUS AND TOOLBAR

This section will explain options under the SPSS menus which are likely to be used by health scientists and nurses. If an explanation on using any other options within these menus is needed, the built in Help gives information on all functions and commands within SPSS.

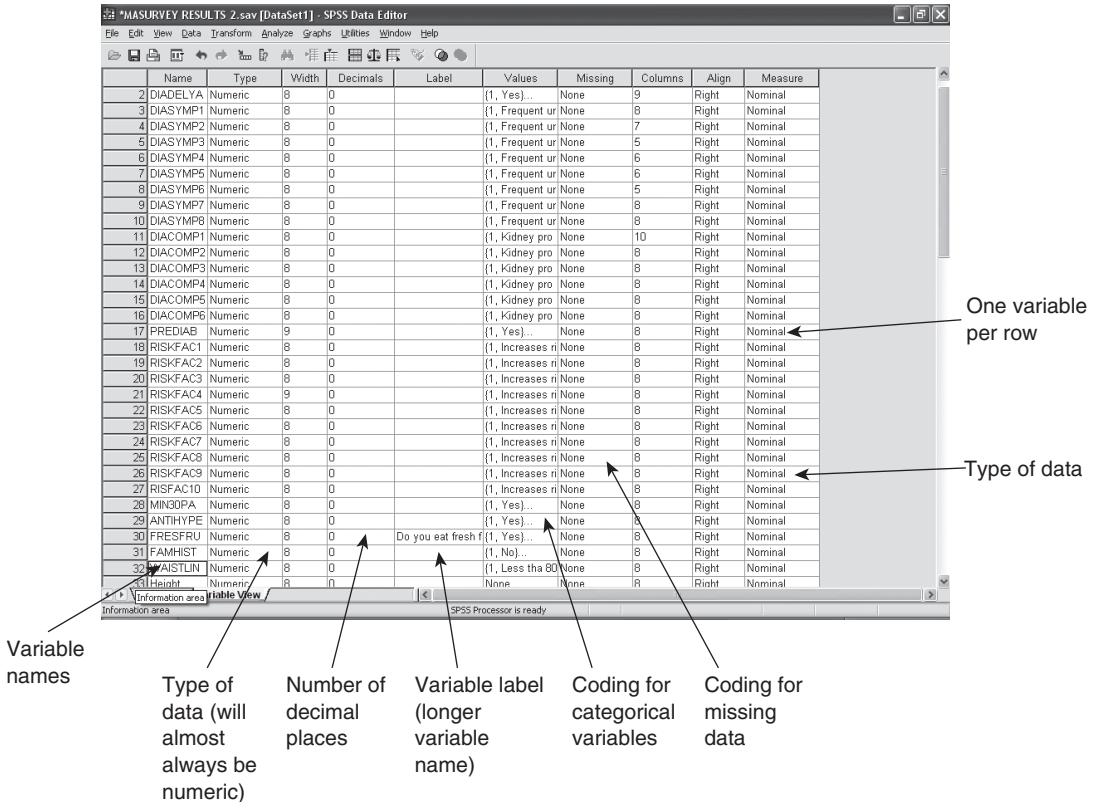


FIGURE 1.4 VARIABLE VIEW

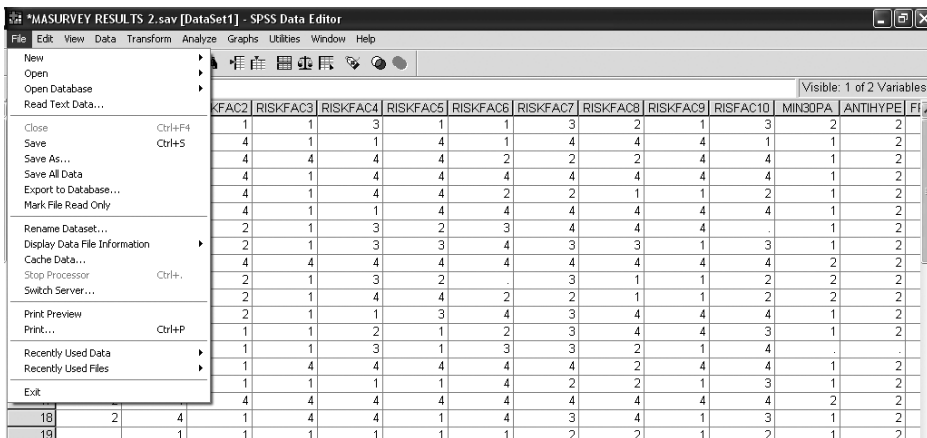


FIGURE 1.5 DATA VIEW SHOWING THE CONTENTS OF THE FILE MENU

File

- Options available under the File menu are shown in Figure 1.5. Many of the options under this menu are similar to those found in other Windows based programs: New, Open, Save, Save As..., Print Preview, Print and Exit.
- Open Database invokes an interface that allows data from a database such as Microsoft Access to be opened directly into SPSS. Likewise, Read Text Data... allows data stored in text formats to be opened.
- Mark File Read Only is used for data protection purposes, so that datasets cannot be modified or deleted. If it is necessary to change the dataset whilst this is in operation, it has to be resaved using a different name. Alternately it is possible to reverse the permissions using Mark file Read Write from the File Menu.
- Display Data File Information shows information about the dataset in the SPSS Viewer (the SPSS Viewer is explained later in this chapter). This includes: the Variable name, Position in the dataset, Label, Measurement Level, Column Width, Alignment, Print Format, Write Format, Missing Values, Value and Label. Most of this information is available in Variable View (Figure 1.4).
- Recently Used Data and Recently Used Files allow datasets or files (output or syntax) that have been opened recently to be opened easily without having to browse directories to find the dataset or file required.

Edit

- This menu (Figure 1.6) largely includes the editing options that are available in other Windows based programs. An additional option is Paste Variables.... Paste Variables... allows exact copies of variables to be created in Variable View. This is useful when a dataset has a number of variables that have similar attributes; variables can be copied and the necessary minor changes made. For example if weight was collected at three time points, the variable attributes can be copied and the variable names and labels changed to make them unique and reflect the three time points.
- Insert Variable and Insert Case allow the user to insert another variable or set of data from a participant into the dataset respectively. These need not necessarily be at the end of the dataset.
- Options... allow the user to change a number of attributes of SPSS, customising to their own preferences.

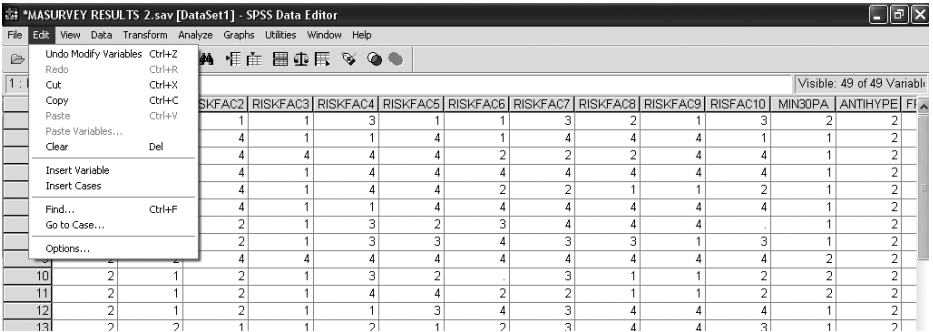


FIGURE 1.6 DATA VIEW SHOWING THE CONTENTS OF THE EDIT MENU

View

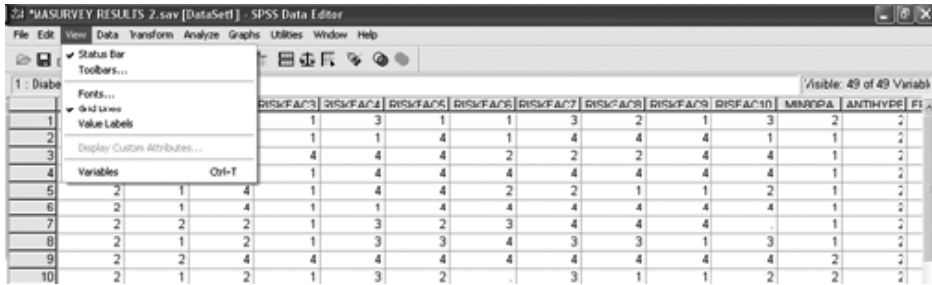


FIGURE 1.7 DATA VIEW SHOWING THE CONTENTS OF THE VIEW MENU

- The icons on the toolbar can be altered using Toolbars....
- The font type and size used in the Data Editor can be changed using Fonts....
- Grid Lines allows the gridlines on Data View to be seen (the default is that they are visible). These can be turned off.
- It is possible to see how the categorical variables have been labelled whilst in Data View by turning Value Labels on.
- When in Data View, clicking on Variables changes the view to Variable View. In Variable View, this menu item is replaced by Data.

Data

- Using the data menu (Figure 1.8), data within a given variable or variables can be sorted using Sort Cases....
- Transpose... creates a new dataset with the rows from the original dataset appearing in the columns and vice versa.

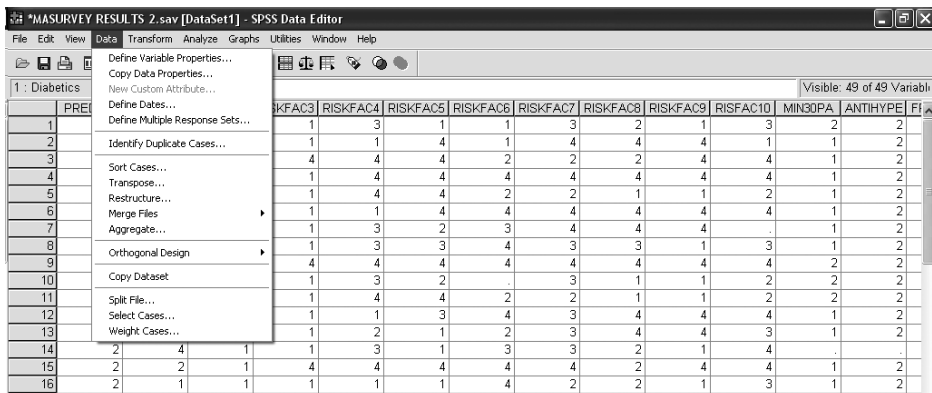


FIGURE 1.8 DATA VIEW SHOWING THE CONTENTS OF THE DATA MENU

- Additional data from another file can be added to the dataset currently open using Merge Files. The data can either be additional participants; in which case at least one of the variable names should be identical (those that are not identical will form new variables in the original dataset). Alternatively, new variables can be added, in which case, the participant identifying variable should be the same in both datasets. This will be explained further in Chapter 2.
- Split File... separates data into groups based on at least one variable for analysis purposes.
- Select Cases... is used if only specific cases (usually defined by specific characteristics) are required. Further explanation on this is in Chapter 2.

Transform

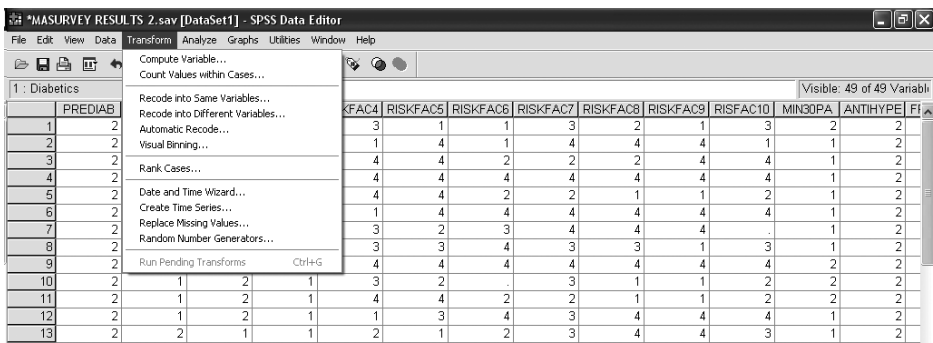


FIGURE 1.9 DATA VIEW SHOWING THE CONTENTS OF THE TRANSFORM MENU

- Compute Variable... is used to generate new variables. These usually utilise SPSS built-in functions such as LN and/or existing variables in the dataset. Details on how to do this will be given in Chapter 2.
- Recode into Same Variables and Recode into Different Variables are used when the coding of variables is altered for analysis, to create coding to transform a continuous variable to a categorical one or to combine responses from more than one variable to create one variable. It may also be used to populate a variable generated using Compute Variable.... Examples of some of these uses are shown in Chapter 2.

Analyze

- All statistical analyses are initiated from the Analyze menu shown in Figure 1.10.
- Reports give elementary case summaries of the dataset. It is more beneficial to summarise the dataset using Descriptive Statistics.
- Descriptive Statistics is used to calculate summary statistics. These will be explained more fully in Chapters 5 and 6.

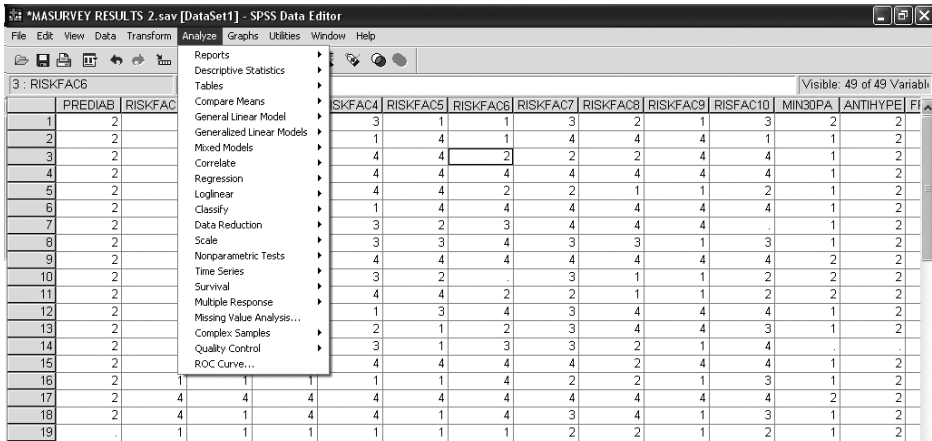


FIGURE 1.10 DATA VIEW SHOWING THE CONTENTS OF THE ANALYZE MENU

- Tables creates tables of summary statistics. Tables of summary statistics can also be created by the levels of categorical variables.
- Compare Means is used to carry out statistical tests that involve the comparison of means, for example, t-tests and one-way ANOVA. These will be discussed in Chapter 9.
- Correlate is used to carry out correlations. These will be explained in Chapter 11.
- Most regression analyses are invoked from the Regression item on the menu. Linear regression and logistic regression will be explained in Chapters 11 and 12.
- All non-parametric tests (for example, Mann-Whitney U test, Wilcoxon Matched Pairs test, Kruskal-Wallis H test) are initiated through the Nonparametric Tests menu item on this menu. An explanation of these will be given in Chapter 10.

Graphs

- The menu shown in Figure 1.11 is the starting point for the production of all stand alone graphics (some graphics can be produced from within specific Analyze menu commands). The easiest way to invoke graphics is to use the Chart Builder.... The use of this will be explained in relation to specific graph types throughout the course of the book.

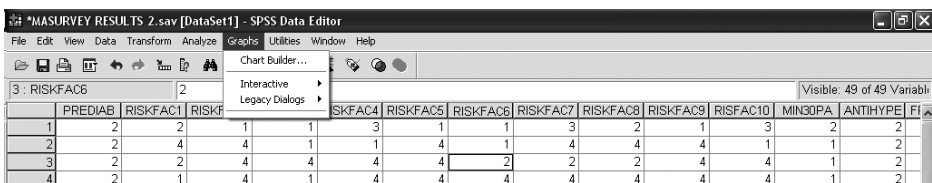


FIGURE 1.11 DATA VIEW SHOWING THE CONTENTS OF THE GRAPHS MENU

Utilities

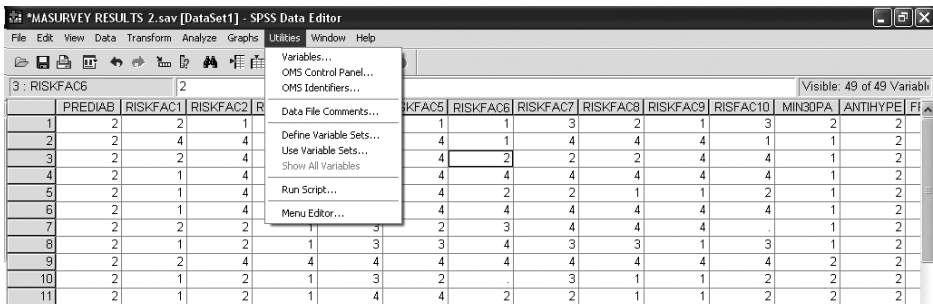


FIGURE 1.12 DATA VIEW SHOWING THE CONTENTS OF THE UTILITIES MENU

- Under the Utilities menu (Figure 1.12), Variables... gives properties, similar to those displayed in Variable View for individual variables. It does not give information for a number of variables simultaneously, so is of limited use.
- Data File Comments... enables the user to add comments about the dataset, which are saved with the dataset. These comments can be shown in the output.

Window

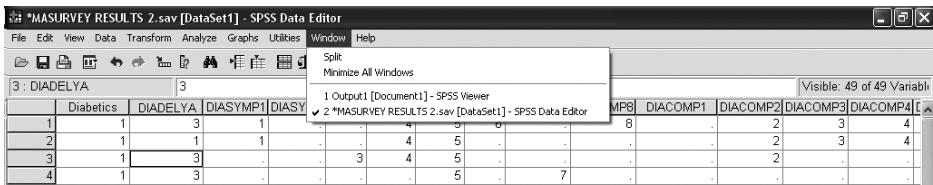


FIGURE 1.13 DATA VIEW SHOWING THE CONTENTS OF THE WINDOW MENU

- Using the Window menu (Figure 1.13), the screen can be divided so that different parts of the dataset can be viewed at the same time using the Split command. The default splits can be moved into positions to suit the user. When the divisions are no longer required, click on Remove Split from the Window menu.
- Minimise All Windows reduces all SPSS windows open to the Taskbar at the bottom of the screen.
- The third item on this menu relates to the window that is currently open within SPSS. When there are a number of windows open it is possible to move from one to another using the list under this menu.

Help

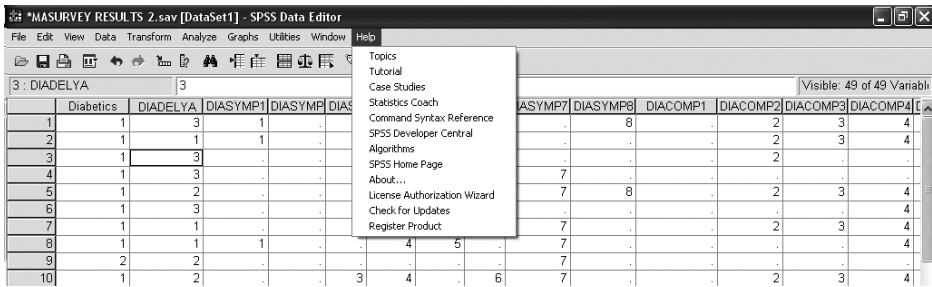


FIGURE 1.14 DATA VIEW SHOWING THE CONTENTS OF THE HELP MENU


- The most useful item on the Help menu (Figure 1.14) is Topics. This performs in a similar way to other Windows based programs.
- Tutorial gives a series of step-by-step guides to a variety of aspects of SPSS including issues related to getting started with SPSS and simple data analysis procedures. Where appropriate the Tutorial is given as a response to the search within Topics.
- Statistics Coach gives basic instructions on what test should be carried out based on the type of data. This function is not a substitute for a good textbook or seeing a statistician.
- Selecting Command Syntax Reference invokes a .pdf document explaining the SPSS syntax.
- Selecting SPSS Home Page invokes an internet browser and then opens the SPSS home page www.spss.com as long as the computer is online.
- About... gives the copyright statement, version number and who the copy of SPSS is licensed to.
- License Authorization Wizard enables the user to give SPSS Inc their authorisation code which locks the copy of SPSS to the hardware.
- Check for Updates connects to www.spss.com to find out whether there are update patches available and if so allows them to be downloaded and installed.
- Register Product enables SPSS to be registered.

Data Editor toolbar

- The data editor toolbar is located underneath the File, Edit, View menus in both Data View and Variable View. This is shown in Figure 1.15.



FIGURE 1.15 DATA EDITOR TOOLBAR

- Starting from the left of Figure 1.15, the first icon is Open File (data).
- The second icon is Save File (greyed out in Figure 1.15). If the data have not previously been saved, Save As is automatically invoked.
- The following icon is Print. However, it is not recommended that the dataset (Data View) is printed because datasets are often very large and inferences about the data will not be able to be made by looking at pages of numbers.
- Dialog Recall displays recently used commands, which can be clicked on to invoke them again.
- Undo and Redo (the left and right pointing arrows, greyed out in Figure 1.15) are used in the same way as with other Microsoft Windows programs. They become usable when an operation has been carried out.
- The icon which resembles a ruler with a red arrow on it is Go To Case. This enables the user to move to a particular case in the dataset. However, this is determined by the grey numbers down the left hand side of Data View rather than any identifying number which may be part of the dataset. Therefore this function has limited use since if the data are sorted, the cases will change in relation to the grey numbers on Data View.
- The following icon is Variables. It enables the user to see much of the information on Variable View variable by variable. This is the same information that can be seen using the Variables command from the Utilities menu.
- The binoculars icon is used for finding particular numbers and/or words within Data View.
- The two icons to the right of Find are Insert Cases and Insert Variable. To use these, highlight a case/variable as required. When the appropriate button is pressed a new case/variable is inserted above/to the left respectively of those already there.
-  is used to split the file, as under the Window menu.
- The scales icon is used to weight cases for statistical analysis. However, for most analyses it will not be necessary to use this function. Weights can also be invoked from the Data menu.
- The first of the final three icons invokes the Value Labels (the same command as that on the View menu), followed by Use Variable Sets and Show All Variables. The final two commands can also be invoked from the Utilities menu.

Opening data saved in formats other than .sav

- Data entered using another program such as Microsoft Excel can be opened in SPSS.
- Click on File → Open → Data... → click on the drop down arrow next to Files of type: to see the types of files that it is possible to open in SPSS (as in Figure 1.16). Choose the format of the file that is required and then locate that file. SPSS will ask for confirmation that the first row contains the variable names and of the length and breadth of the dataset. Then click OK.
- When opening data from some formats (for example, Microsoft Excel), there will be no variable Labels or Values produced. These details can be entered using Variable View.

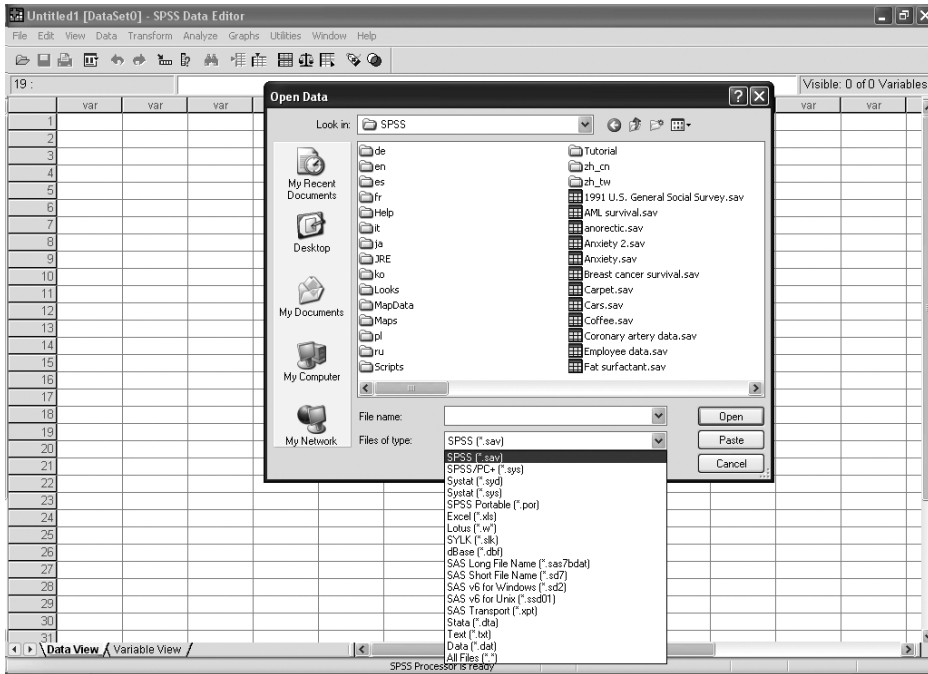


FIGURE 1.16 SCREENSHOT SHOWING THE DATA FORMATS THAT SPSS WILL OPEN

ENTERING DATA

When questionnaires or data extraction sheets have been coded, they are ready to be entered into SPSS.

When SPSS opens, click on the Variable View tab where you will see a screen like Figure 1.17. This screen should be set up to define the variables before data can be entered into Data View.

Name is the variable name, which will appear at the top of columns in Data View. This should be 8 to 10 characters long, and start with an alphanumeric character (letter).

The default for Type is Numeric. This will be appropriate for most data entered into SPSS. Occasionally String (text) or Date (to allow dates to be entered into SPSS) may be needed.

There should be no need to change the Width from the default unless entering text, then this may need to be wider.

Some variables will not need decimal places (for example, the ID number or categorical variables which have been coded). In these circumstances the number of decimal places can be decreased to 0 by either using the arrow in the Decimals cell or typing in the desired number of decimal places, making the data as shown in Data View easier to read. This is shown in Figure 1.18.

Use Label to put in a longer variable name. For example the question as it appears on the questionnaire may be put here. If this is used it will appear in Variables: boxes

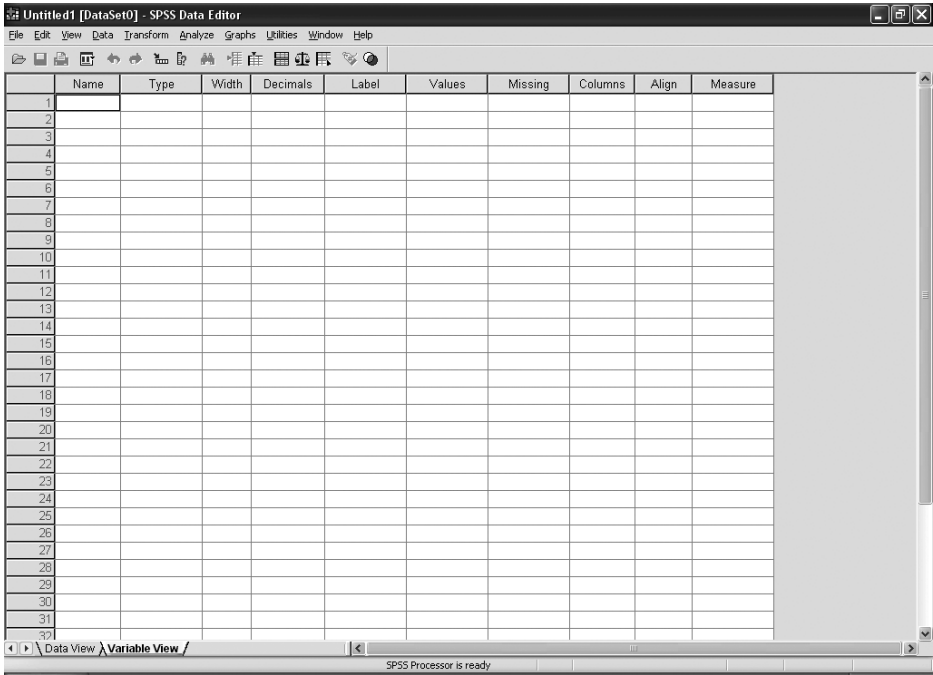
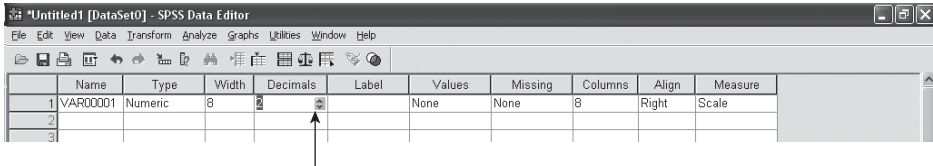


FIGURE 1.17 SCREENSHOT OF VARIABLE VIEW



Click on the Decimals box to show the arrows, press the up arrow to increase the number of decimal places and the down arrow to decrease the number of decimal places

FIGURE 1.18 SETTING UP VARIABLE VIEW'S DECIMALS BEFORE ENTERING DATA

alongside the variable name when choosing variables to analyse. It will appear instead of the variable name in some output.

For questions that give nominal or ordinal data, you can tell SPSS how the responses are coded, so that when analyses are carried out, tables and graphics show the textual coding. Click on the Values cell and ... appears in the right of the cell, click on this to get a dialog box like Figure 1.19. In the Value: box put a code for a category of the variable in question, then in the Label: box beneath it, put the category label. When both pieces of information have been added, click Add. Repeat this until all categories within the variable in question have been declared, at which time click OK to return to Variable View (Figure 1.17).

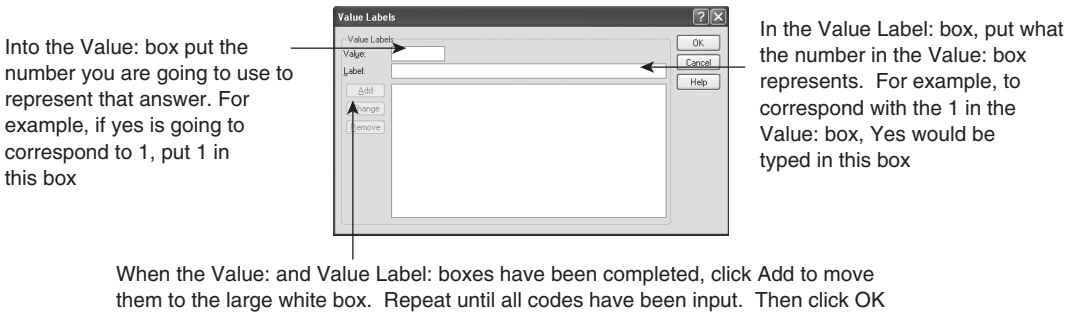


FIGURE 1.19 VALUE LABELS DIALOG BOX

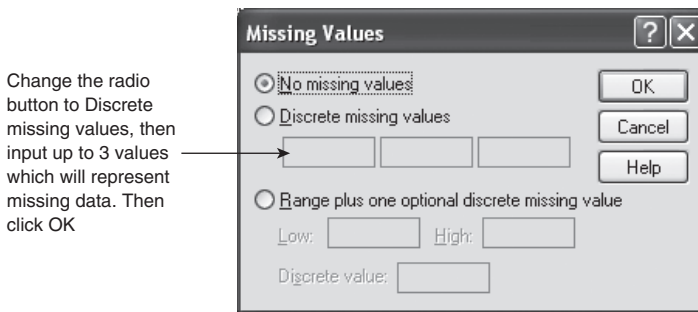


FIGURE 1.20 MISSING VALUES DIALOG BOX

If the dataset is going to have a specified code for missing data, then the Missing cell should be clicked on. This will give ..., click on this. You will get the dialog box shown in Figure 1.20. The advantage of doing this is that it will be obvious that data for that variable have not been omitted to be entered, and that when entering large amounts of data that are present do not get entered into the wrong cell by not moving to the correct cell. Codes for missing data should not be legal values for the variable in question. For example, if the variable was age of respondent, the value indicating missing should be one that is out of range. There are a number of conventions regarding how missing data are coded. Sometimes columns with missing data are filled with 9, 99, 999 or 9999 depending on possible legal values for the variable. In the age example, the codes for missing values could not be 9 or 99 (depending on the study population), so the most sensible coding for this variable would be 999. Some studies such as the Health Survey for England use negative values to indicate data are missing. The easiest way to declare missing values in the Missing Values dialog box (Figure 1.20) is to click the Discrete missing values radio button, then insert up to three values that indicate missing data. When missing value codes have been declared, click OK to return to Variable View (Figure 1.17).

On Variable View, there should be no need to change the Columns or Align boxes. The Measure column should be changed to match the type of data to be inputted. The default is Scale. This should be used for continuous or discrete data, such as

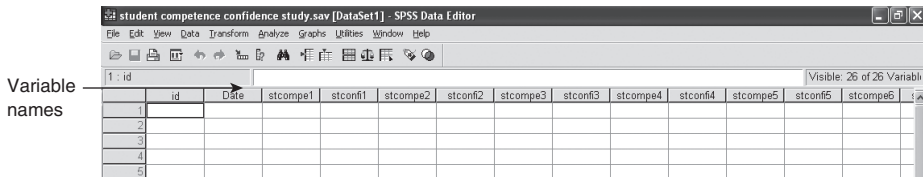


FIGURE 1.21 DATA VIEW WHEN VARIABLE VIEW HAS BEEN SET UP

height or weight. Other options are nominal or ordinal. Variables assigned to these data types should match the descriptions and examples given earlier in this chapter.

The process of setting up the attributes for the variables should be repeated for all variables that will be in the dataset. Once all variables have been defined click on Data View to get a screen like Figure 1.21.

To start entering data, click on the first cell (usually the top left), and begin keying in the data. To move to another cell use the arrow keys on the keyboard (the mouse can also be used, but using the keyboard is quicker). It is easiest to enter all data for each participant before moving on to the next participant.

SAVING DATA

Data will need to be saved. As with other computer programs it is best to do this at regular intervals. In Data View or Variable View click on → File → Save As... and then proceed as for other computer programs.

SPSS VIEWER

All tables, graphics, warnings and information are displayed in the SPSS Viewer (Figure 1.22). The SPSS Viewer opens when SPSS is opened. However, a new SPSS Viewer can be manually opened by clicking on File → New → Output. All subsequent output will be appended to the SPSS Viewer that is open.

Output created in a previous SPSS session can be opened without the relevant dataset being open by clicking on File → Open → Output... → then locating the file required (which will end .spo). If saved output is open and more analysis is carried out, the additional analysis will appear in that SPSS Viewer.

Unwanted output (for example, logs, warnings or results from analyses that are no longer required) can be removed from the SPSS Viewer before it is saved or printed by clicking on the output that is not wanted then pressing delete. Alternately, the item that is no longer required can be located in the Outline Pane and selected. Once selected a box will appear round the item in the main part of the SPSS Viewer; then press delete. Using the Outline Pane is the quickest way of deleting a large number of items from the SPSS Viewer.