1

SPSS Syntax Usage

1.1 Introduction △

SPSS[1] is a user-friendly program because most options can be found in clear menus. More experienced users will, however, use syntax files, in which commands are stored to modify and/or to analyze the data. Syntax files keep track of every modification and analysis, making it possible to control and to replicate scientific results—two necessary conditions for scientific research. This book provides an overview of the most commonly used commands in terms of SPSS syntax. It is by no means meant to be a complete overview. For specialized applications, we refer to the very extensive manual, which is included in the program as a PDF file (see the help menu) for the exact notation of the commands. Moreover, a set of examples for specialized commands can be found on the Internet pages that support this book (see http://www.ru.nl/mt/syntax/home).

When using SPSS syntax, you will find that there are often several ways to achieve the same results. In such cases, we have striven to explain the least complicated method.

Some familiarity with modifying and analyzing statistical data is required from the reader as this book only deals with the practical use of SPSS syntax and not the underlying statistical techniques. For an approachable introduction to these techniques that does not require prior knowledge of mathematics, we refer the reader to *Discovering Statistics Using SPSS* (2012) by Andy Field.

How to Use SPSS Syntax is set up as a collection of independent sections where information can be found on specific commands. So it is possible to skip parts of the book that deal with topics that one is not familiar with or that are not needed. Where extra knowledge is required for a command, we will refer to the relevant section(s) of the book. However, when describing the commands, we assume that the reader has read Chapter 2 of this book in which general rules for preparing syntax are described, before reading (sections from) Chapters 3 or 4. In the following sections, we will provide some general information about how to use this book.

---

[1]SPSS is now IBM SPSS. SPSS was acquired by IBM in October 2009.

⚠ 1.2 Structure of the Book

To help the reader learn SPSS syntax, we have included a variety of examples from research in the social sciences. These examples are stored in SPSS data files (we use the "sav" format) and are freely available online (see http://www.ru.nl/mt/syntax/home). When downloading these files, it is recommended to store them in a directory named DATA. You may alter this destination directory, but the syntax used in this book refers to the directory DATA. It is assumed that this directory is created on the c-drive when using a Windows computer (*notation*: c:\data). This eventually means that you have to alter this reference in the syntax accordingly when you use a directory other than c:\data. We further programmed the results to be stored in c:\data as well, but again, you may adapt this to your own needs. In case the reader has a Macintosh computer, all notations in this book referring to computer files are without c:, so only \data.

The syntax used throughout this book is placed in a "box" to separate it from normal text so that it can be easily located. The syntax is then summarized again at the end of each chapter. The reader is invited to type the syntax from the book into syntax files to maximize the benefits of learning by doing. As a final reference for the reader, we have also included the relevant outcomes of analyses.

In SPSS, it is also possible to copy the commands from the menus into syntax files (this is called "pasting") and adapt them to one's needs. This is particularly useful for statistical analyses. In Chapter 4, we will elaborate this approach of creating syntax files.

From Chapter 3 onward, the book consists of more or less independent sections. In each section, one common SPSS command is explained by naming its function and the exact structure, followed by one or more examples and the most relevant results that originate from that particular command.

⚠ 1.3 What Is SPSS Syntax?

Up until the 1990s, SPSS did not have any menus or windows. To modify and analyze data, one would write a "job"—a series of commands—and then have SPSS execute these commands. Over the years, software has become more accessible, and SPSS added a graphical user interface with (usually) clear menus. The program, however, is still constructed in the same way. So behind all menus still lie the commands that originally had to be typed manually. These commands become visible again when pushing the "paste" button in the menus. These SPSS commands are nothing but text in which the name of the command is entered, followed by the options. Furthermore, in most commands, you will find one or more variables to which the

command refers. These commands exist for almost all functions of SPSS, and all data modifications and analyses are made up of a series of commands. Managing SPSS using syntax is still a very viable way of working. The reasons for this are explained in the next section.

1.4 The Usefulness of SPSS Syntax △

Saving all actions in a syntax file has its advantages when modifying and analyzing data. In such a file, you can check exactly which commands have been executed and repair any mistakes if something went wrong. For example, when you have recoded a variable and made a mistake, it is possible that all the executed analyses since are also incorrect. Repairing the error and executing all actions again is relatively easy in a syntax file. This, of course, is much more efficient than starting again from scratch. Another advantage is that commands from a syntax file can be copied directly into a new syntax file. By changing only the variable names in the syntax, you can avoid having to find out how the modifications and analyses are done in SPSS again and again. Last, working with syntax files has the advantage that even others can read and know exactly what has been done with respect to data modifications and statistical analyses. The advantages are so great that not only experienced SPSS users but also beginners (e.g., students who are working on a paper or thesis) often work with syntax. As mentioned earlier, SPSS syntax can be created by the "paste" button in almost all SPSS menus. As this book will prove, however, you can, besides being flexible, work faster and more efficiently when typing the syntax directly into a syntax file. Additionally, not all of the features within SPSS are included in the menus.

1.5 Syntax Window △

Besides the data and output windows, SPSS also has a syntax window to type the syntax lines. The largest part of this window contains space for writing text (commands). The window also has a lot of the menus that are present in the other windows. An empty syntax window is opened by choosing File → New → Syntax. To open an existing syntax file, choose File → Open → Syntax. An important menu in the syntax window is Run, which is used to execute the commands in the syntax window. To run selected commands fast, use the *Crtl-r* shortcut (press the control button and "r" at the same time). If nothing is selected in the window, SPSS will execute the part of the syntax where the cursor is positioned. There are different ways to select syntax: If you want to execute a selection of all the commands, you can select the commands you want by "dragging" the mouse (holding the left

mouse button and moving the mouse) or by holding the shift button and moving the cursor with the arrow keys. Selecting all commands can be done using the C*rtl-a* shortcut.

As explained earlier, you can also use the "paste" button in the SPSS menus. This will paste the selected commands in the menu into the syntax window that is active. This command can then be executed using the C*trl-r* shortcut or by the *Run* option in the menu.

Finally, we want to point out that SPSS does not have an option to automatically save syntax, so if the computer crashes, you will lose the syntax! However, it is possible to recover your syntax if it was saved in the session journal. To check this, go to the menu and select Edit → Options (SPSS Version 15 and earlier) or Edit → Options → File Locations, and then check "Record syntax in Journal" and "Append," and click on ok. We recommend saving the syntax files regularly, especially before (blocks of) commands are executed. You can save quickly by using the C*trl-s* shortcut.

⊿ 1.6 SPSS Syntax Files: An Example

Open the SPSS program to begin your session. Do not open any data yet—it does not matter if the data window is empty. Open a new syntax window with File → New → Syntax. In most cases, a data file must be opened first. To do this, type the following command into the syntax file (please note that the period (.) at the end of the command is part of the syntax):

> GET FILE "c:/data/chapter1.sav".
>
> * *Macintosh computers use:* GET FILE "/data/chapter1.sav".

In this command, it is assumed that the file "chapter1.sav" is located in c:/data/. If necessary, replace this with the directory where you have saved it on your own computer. Instead of the double straight quotation marks before and after the file name in SPSS syntax, one may also use straight single quotes, *but* back quotes (`) are not allowed. Furthermore, please note the period (.) after the command. If everything seems in order, execute the command by placing the cursor on the command line and pushing C*trl-r* or choosing Run → Current from the menu. If you did not make any mistakes, the data window should now show the data from the example file. If this is not the case, there will be an error message in the output window indicating the error made in the syntax command. After loading the data into SPSS, you may retrieve specific information from it. To obtain

information about the variables, please type the following command after the GET FILE command and execute the command:

DISPLAY DICTIONARY.

A description of all variables will appear in the output window after executing this command (Table 1.1):

Table 1.1 The Results of Display Dictionary

Variable Information					
Variable	**Label**	**Measurement Level**	**Column Width**	**Alignment**	**Print Format**
respnr	Respondent identification number	Nominal	8	Right	F5
SEX	<None>	Nominal	8	Right	F1
EDUC	Highest completed educational level	Ordinal	8	Right	F1
YBIRTH	Year respondent was born	Scale	8	Right	F4

Note: Variables in the working file.

Variable Values		
Value		**Label**
SEX	1	Male
	2	Female
EDUC	1	Elementary school
	2	Middle school
	3	Junior high school
	4	Senior high school
	5	Vocational school
	6	College
	7	Bachelor degrees
	8	Master degrees

Apparently, a variable called YBIRTH is present in the data file containing the year of birth of the respondents. The data in the file "chapter1.sav" were collected in 1995, so we are able to calculate the respondent's age with the formula "1995 – YBIRTH." The syntax is as follows (please type it in your syntax file and run it):

```
COMPUTE AGE = 1995 – YBIRTH.
```

You have now created a variable AGE containing the age of each respondent in the file. If you want to check whether the new variable is correctly computed, you can execute the following command:

```
DESCRIPTIVES AGE YBIRTH
 /STATISTICS MEAN STDDEV MIN MAX.
```

This will produce an output that displays the mean, standard deviation, and minimum and maximum values of AGE and YBIRTH. After executing the syntax, the following will be displayed in the output window (Table 1.2):

Table 1.2 Descriptive Statistics of AGE and YBIRTH

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
AGE	1974	18.00	70.00	42.2533	13.74781
YBIRTH year respondent was born	1974	1925	1977	1952.75	13.748
Valid N (listwise)	1974				

The computation of AGE is correct, since the minimum value for YBIRTH is 1925 and the maximum value for AGE is 70 (1995–1925). Furthermore, the number of valid scores is equal for both variables (1974). Finally, the sum of the two mean scores is equal to 1995 (42.2533 + 1952.7467).

Another variable present in the data file is SEX. A frequency distribution is useful to get full information about this variable (Table 1.3):

```
FREQUENCIES SEX.
```

Table 1.3 Frequency Distribution of SEX

				Valid	Cumulative
		Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Valid	1 Male	945	47.9	47.9	47.9
	2 Female	1029	52.1	52.1	100.0
	Total	1974	100.0	100.0	

This provides us with a table including the absolute numbers and percentage of men and women in the data file.

As you may have noticed, the data operations and descriptions in this chapter have been executed with relatively simple commands. The next chapter deals with the general structure of such commands. In Chapters 3 and 4, the individual commands are described in more detail.

1.7 SPSS Syntax: An Overview △

We conclude by listing all of the syntax used in this chapter. Notice that syntax comments can be made by starting a sentence with an asterisk (*) and ending it with a period (.). This is a useful way to annotate your work for future reference. The syntax (Windows and Macintosh) is also available at the website (http://www.ru.nl/mt/syntax/home).

```
* Opens the data file used in this chapter.
GET FILE "c:/data/chapter1.sav".
* Macintosh users:.
GET FILE "/data/chapter1.sav".

* Retrieves information about the variables in the data set.
DISPLAY DICTIONARY.
```

```
* Creates a new variable AGE (sample is from 1995).
COMPUTE AGE = 1995 − YBIRTH.

* Checks whether new variable was computed correctly.
DESCRIPTIVES AGE YBIRTH
 /STATISTICS MEAN STDDEV MIN MAX.

* Shows the frequency distribution of the variable SEX.
FREQUENCIES SEX.
```