LEARNING MADE EASY

SPSS® Statistics

WORKBOOK

for dummies®

A Wiley Brand

Fine-tune your interpretation skills

Rehearse the most important statistical techniques

Produce more effective tables and charts

Jesus Salcedo, PhD
Keith McCormick

SPSS® Statistics Workbook

by Jesus Salcedo, PhD
and Keith McCormick

for dummies®
A Wiley Brand

SPSS® Statistics Workbook For Dummies®

Published by: John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030-5774, www.wiley.com

ISBN 978-1-394-15630-6 (pbk); ISBN 978-1-394-15632-0 (ebk); ISBN 978-1-394-15631-3 (ebk)

Contents at a Glance

Table of Contents

Introduction

Do you need to use IBM SPSS Statistics for school or work and you feel that you need more practice to become proficient? Maybe you eventually get the results you need when you work in SPSS, but it feels like you aren't doing so efficiently. Maybe you're about to embark on an important assignment or research project, and you feel like some more practice (with carefully presented solutions) is just what you need.

Not to worry. Help has arrived. *SPSS Statistics Workbook For Dummies* is a collection of exercises and solutions to move you towards mastery with SPSS. It is a companion book to *SPSS Statistics For Dummies,* 4th Edition, but you don't need to read both. (You see more about how the books work together later in this introduction.)

We (Keith and Jesus) have trained tens of thousands of students during in-person workshops for SPSS Inc. (before IBM acquired SPSS). That experience, over many years, means we've looked over the shoulders of attendees as they've attempted exercises just like these. We can anticipate possible mistakes before you make them — and if you practice the examples in this workbook, you won't make them at all. We've used that experience, along with the topics that IBM has identified as important for certification, to keep you focused on the critical skills for using SPSS efficiently.

About This Book

This workbook is designed to develop your ability to perform the most common and important tasks in the efficient use of SPSS Statistics. Like all books in the *Dummies* series, it features easy-access organization and doesn't bog you down in unnecessary details. We carefully listed, discussed, debated, and narrowed the topics to maximize an efficient investment of your time.

We understand and value academic pursuits, but this is not an academic book. This is an eminently practical book written by two authors who have spent decades leading software training workshops. However, if you're a student, a researcher, or an academic, we have you covered. We explain critical statistical concepts, choosing the right technique, and interpreting your results. All that and more is explained from a practical point of view.

So why choose this workbook?

- Dozens of practical examples with step by step solutions
- Advice from two lifelong users of SPSS Statistics

- Topics carefully chosen and prioritized for the tasks you will encounter most often
- Tips and tricks to avoid the most common mistakes from two authors who have observed (and made) every mistake that is possible to make in SPSS

We follow a few conventions in this workbook:

- Most chapters in this workbook will have the following repeating structure: explanation, example, questions, and answers. The answers are found at the end of each chapter. While skipping around chapters and sections to find what you need is encouraged, you'll probably have the best luck reading each section, as indicated with a major heading, in its entirety.
- Most of the exercises use the menus and dialogs, so when we show you a series of mouse clicks to access a specific area of the menus, we list the steps like this: choose File ➪ Open ➪ Data and load the GSS2018.sav file.
- Two chapters, 15 and 16, focus on using SPSS Syntax. SPSS Syntax programming code appears in `monofont` to help differentiate it from other text.
- You are encouraged to write directly in this workbook when trying interpretive questions. When the questions prompt for an interpretation like "Describe your result," space is provided for you to write your answer.

Foolish Assumptions

This book is for you if you have IBM SPSS Statistics and want to practice. Because it's a workbook, you get dozens of examples to help you prepare data, create variables, produce statistical analyses, and even write SPSS Syntax. You can even check out practice questions for IBM's certification exam in the final section of the book.

This workbook has been written as a companion to the book *SPSS Statistics For Dummies*, 4th Edition. That book, which is written for first-time users of SPSS, has more theory, overview, and initial setup. You aren't required to have both books, but they make a perfect pair. *SPSS Statistics For Dummies* is the best place to start if you are at ground zero. You can start there, and then try the exercises in this one. Or, and we like this option even better, use them side by side.

This book is about IBM SPSS Statistics. We sometimes call it SPSS Statistics, or even just SPSS. You may encounter something called IBM SPSS Modeler. That is different software, and this book will not discuss SPSS Modeler at all (except just now).

Icons Used in This Book

The tip icon marks tips and shortcuts to make working in SPSS Statistics quicker and easier.

Remember icons highlight information that's especially important to know. These reminders represent best practices in using SPSS.

The technical stuff icon indicates information of a highly technical nature that provides helpful context. You can skip this info to complete a practice example, but it will help you more fully understand SPSS.

The warning icon tells you to watch out! It notifies you of important information that may save you headaches when using SPSS.

The example icon presents a fully worked-out solution as a reference to assist you with the practice problems.

Beyond the Book

This workbook is just one part of the support we provide for mastering SPSS. For details about significant updates or changes that occur between editions of this workbook, go to `www.dummies.com`, type **SPSS Statistics Workbook For Dummies** in the Search box, and open the Download tab on this book's dedicated page.

That page is also where you'll find the book's cheat sheet, which contains advice for exploring a new data set, identifying which menu to use to perform various important data preparation operations, and finding the correct graph type for your data by utilizing level of measurement.

You can find a wealth of related information at the companion book's Dummies page by going to `www.dummies.com` and typing **SPSS Statistics For Dummies** in the Search box. Videos describing the topics covered in *SPSS Statistics For Dummies*, 4th Edition and ten SPSS gotchas can be found at `https://keithmccormick.com/SPSS4Dummies`.

Where to Go from Here

We wrote this book in a nonlinear way. You can skip around. It's okay. We skipped around when we wrote it, too. But it does have a structure. If you have a deep interest in a broad topic, you might want to complete an entire part of the book.

If you need practice getting your data in and set up, check out Part 1. Part 2 is all about creating variables with the Compute dialog and by using other transformations. If you need to apply a formula or modify a variable, check out the dozens of examples in this part.

If you need practice producing statistical analyses and interpreting them, you need Part 3. Part 4 is all about graphing, providing lots of practice making nearly every graph type that SPSS Statistics supports. You also learn about editing output.

If you are intrigued with Syntax but have been afraid to dive in, see Part 5. You get a chance to try Syntax and then check your work. Many of the tasks are similar to those in Part 2, except you work in the Syntax window in Part 5.

In the final section, Part 6, we warn you about skills that are particularly tricky and then offer some practice questions to help you decide if the IBM SPSS Statistics certification exam is something you'd like to try.

1 Getting Data into and out of SPSS

IN THIS PART . . .

Bring data into SPSS and show results in other applications.

Define data properties.

IN THIS CHAPTER

- Resolving trouble with delimiters on import
- Exporting output

Chapter 1
Working Through Import and Export Challenges

In this chapter you learn about ways to transfer data into and out of SPSS. Getting data into SPSS is the first step before any analysis can be done. If the data is available in an SPSS data file (.sav file extension), bringing that data into SPSS is easy. If your data comes from another program such as Excel or is in the txt, CSV, or SAS format, you can import that data into SPSS with just a little more work. The examples in this chapter demonstrate some complications that arise when importing these types of files.

After SPSS analyzes your data and displays results in easy-to-understand tables and graphs, you might want to use the results in another application to share your findings with others. This might involve

- Formatting tables as cleanly as possible for clarity of presentation and ease of viewing
- Exporting output to a file format that can be read by any user, such as Portable Document Format (pdf)
- Doing post-processing on pivot tables in another application, such as Microsoft Excel

Applications such as Microsoft's PowerPoint or Word can display the results of your analyses as plain text, as rich text, or as a metafile, which is a graphical representation of the output. Pivot tables can be pasted or exported to Microsoft Excel with each cell of the pivot table in a separate Excel cell.

In this chapter, you both copy and paste as well as export your SPSS output to another application.

Importing Data

This section contains an example of a procedure you can follow to read data from Excel files into SPSS. Along the way, SPSS keeps you informed about what's going on so there won't be any big surprises at the end. Here are a few things to consider to make the import process a little easier:

- When reading the Excel file into SPSS, the file must not be open in Excel.
- Variable names are read from the first row of the Excel spreadsheet. However, the blank spaces in the variable names are removed because blanks are not allowed in SPSS variable names.
- The measurement levels of variables are assigned based on the criterion defined in Data options. Variables with a small number of values are set to nominal. Variables with many values are set to scale.
- SPSS reads only the values in the cells of the spreadsheet. Formulas in the spreadsheet will be computed and these computed values imported to SPSS.
- The formulas and other spreadsheet characteristics associated with the cells are not imported to SPSS.

TIP

We strongly recommend opening the original file in the software where it is stored so you can see how the file is structured.

In the following example, the Excel workbook has two worksheets: The first is the title page and the second contains the data. Do the following to read this data into SPSS:

1. **Choose File➪Import Data➪Excel.**
2. **Select the GSS2018 Title.xlsx file and then click Open.**

   You can download the file from the book's companion website at `www.dummies.com/go/spssstatisticsworkbookfd`.
3. **In the Worksheet drop-down list, select the GSS2018 worksheet.**

   An Excel file can contain more than one worksheet, and you can choose the worksheet you want from the drop-down list, as shown in Figure 1-1.

Also, if you've elected to read only part of the data, use the Range drop-down list to specify the range of Excel cells that you want to import. Use the five check boxes, as needed, to specify whether the names of variables appear in the first row, the percentage of data to use to determine the variable type, and how to handle leading or trailing spaces.

TIP

Inspect the data preview to make sure the variables and data will be read properly.

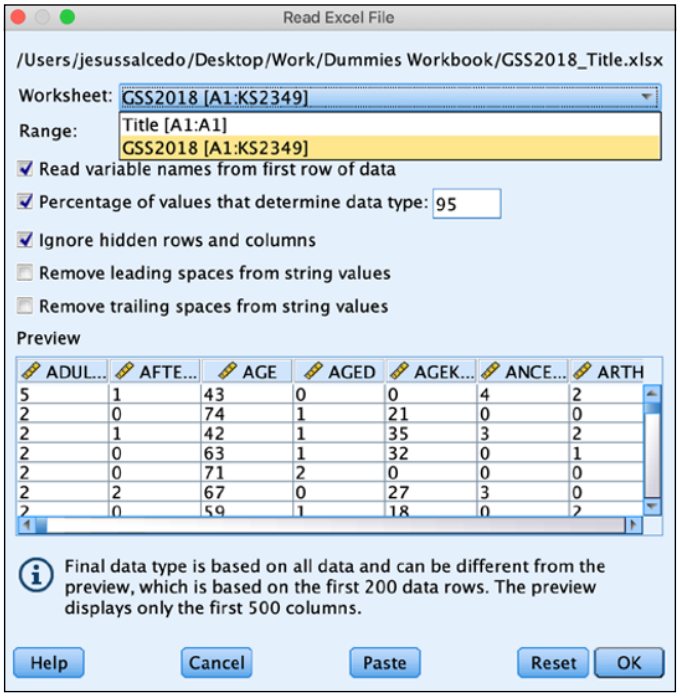


FIGURE 1-1: Select which data in the spreadsheet to include.

4. **Click the Read Variable Names from the First Row of Data check box.**
5. **Click OK.**

   Your data appears in the SPSS window.
6. **Switch to the Variable View tab to examine the variable definitions and make any changes.**

   SPSS makes a bunch of assumptions about your data, and some of those assumptions are probably wrong.
7. **Save the file using your chosen SPSS name, and you're off and running.**

See the following for an example of importing data into SPSS.

Q. Text files are another common source of data. Many spreadsheet programs and databases can save their contents in a text file format. Commonly used delimiter characters are tabs or commas.

Import the file GSS2018.csv to SPSS. Note that the comma-separated file has variable names in the first row. The variable names are in the first row and the data begins in row 2.

A. Do the following:

a. Choose File➪ Import Data➪ CSV Data.

b. Select the GSS2018.csv file, and then click Open.

The Text Import Wizard appears, allowing you to load and format your data. Examine the input data. The screen lets you peek at the contents of the input file so you can verify that you've chosen the right file.

c. Click Advanced Options (Text Wizard). Examine the input data.

If your file uses a predefined format (it doesn't in this example), you can select it here and skip some of the later steps.

d. Click Continue.

e. Specify that the data is delimited and the names are included.

SPSS takes a guess, but you can also specify how your data is organized. It can be divided using commas (as in this example), spaces, tabs, semicolons, or some combination. Or your data may not be divided — it may be that all the data items are jammed together and each has a fixed width. If your text file includes the names of the variables, you need to tell SPSS.

f. Click Continue.

g. Specify how SPSS should interpret the text.

You can tell SPSS something about the file and which data you want to read.

Perhaps some lines at the top of the file should be ignored — this happens when you're reading data from text intended for printing and header information is at the top. By telling SPSS about it, those first lines can be skipped.

Also, you can have one line of text represent one case (one row of data in SPSS), or you can have SPSS count the variables to determine where each row starts.

And you don't have to read the entire file — you can select a maximum number of lines to read starting at the beginning of the file, or you can select a percentage of the total and have lines of text randomly selected throughout the file.

h. Click Continue.

i. Specify tab as the delimiter.

SPSS can use commas, spaces, tabs, and semicolons as delimiting characters. You can even use some other character as a delimiter by selecting Other and then typing the character. You can also specify whether your text is formatted with quotes (which is common) and whether you use single or double quotes.


REMEMBER

Strings must be surrounded in quotes if they contain any of the characters being used as delimiters.

j. Click Continue.

k. If necessary, change the variable name and data format.

SPSS makes a guess for the type of each variable. To change a name, select it in the column heading at the bottom of the window, and then type the new name in the Variable Name field at the top. If you need to change the format, use the Data Format drop-down list.


REMEMBER

You can also change the data types later in the Variable View tab of the Data Editor window.

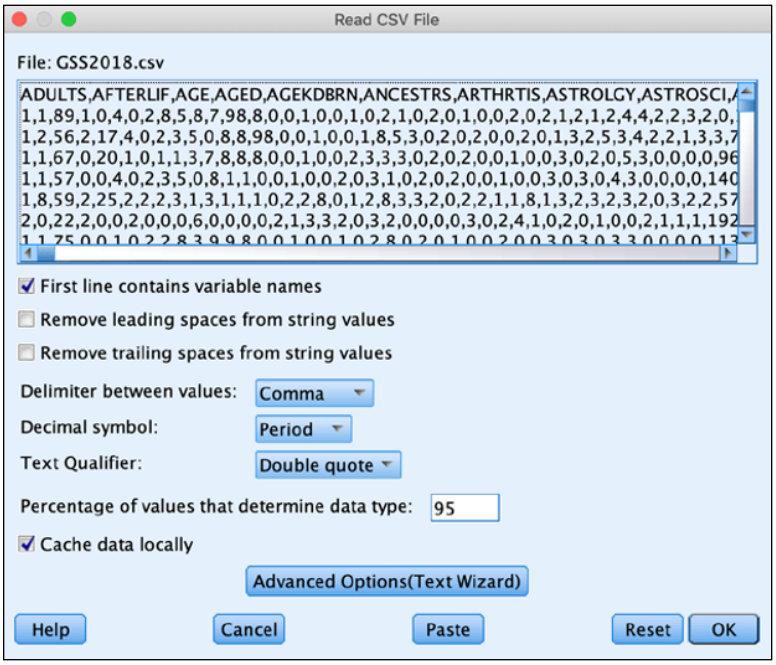
l. Click Continue.

m. In the Would You Like to Save This File Format for Future Use? section, click No.

Saving the file format for future use is something you would do if you were loading more files of this same format into SPSS — it reduces the number of questions to answer and the amount of formatting to do next time.

n. Click the Done button.

o. Look at the data, and correct the data types and formats, if necessary. Then save it all to a file by choosing File➪ Save As.



1. Import the GSS2018.xlsx file to SPSS. This example has one worksheet, the variable names are in the first row, and the data begins in row 2.

2. Import the GSS2018 extra title.xlsx file to SPSS. This example has two worksheets in the Excel workbook. The first worksheet is the title page. The second worksheet contains the data, with the variable names in the first row and the data beginning in row 5.

3. Import the GSS2018.dat file to SPSS. This file has a DAT format. The variable names are in the first row and the data begins in row 2.

4. Import the GSS2018 lines.txt file to SPSS. This file has a TXT format. The variable names are in the first row and the data begins in row 3.

Exporting Results

If you have a single table or a small number of tables, you can copy and paste these directly into a file opened in another application. Alternatively, SPSS provides an export facility to export large numbers of tables and charts into a file in a variety of common formats: Excel, Portable Document Format (PDF), HTML, text, Microsoft Word, and PowerPoint files.

1. **Choose File ➪ Open ➪ Output and load the Chapter 1 Output.spv file.**

   Download the file at `www.dummies.com/go/spssstatisticsworkbookfd`.

2. **Select the R's highest degree frequency table.**
3. **Choose Edit ➪ Copy.**
4. **Switch to Word or another word-processing application.**
5. **Choose Edit ➪ Paste Special.**


TIP

When copying and pasting SPSS pivot tables, always use Paste Special because it provides various options for displaying the table.

6. **Choose Formatted Text (RFT).**

   The table is pasted and looks like a Word table that can be edited, as shown in Figure 1-2.

R's highest degree

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	LT HIGH SCHOOL	262	11.2	11.2	11.2
	HIGH SCHOOL	1178	50.2	50.2	61.3
	JUNIOR COLLEGE	196	8.3	8.3	69.7
	BACHELOR	465	19.8	19.8	89.5
	GRADUATE	247	10.5	10.5	100.0
	Total	2348	100.0	100.0	

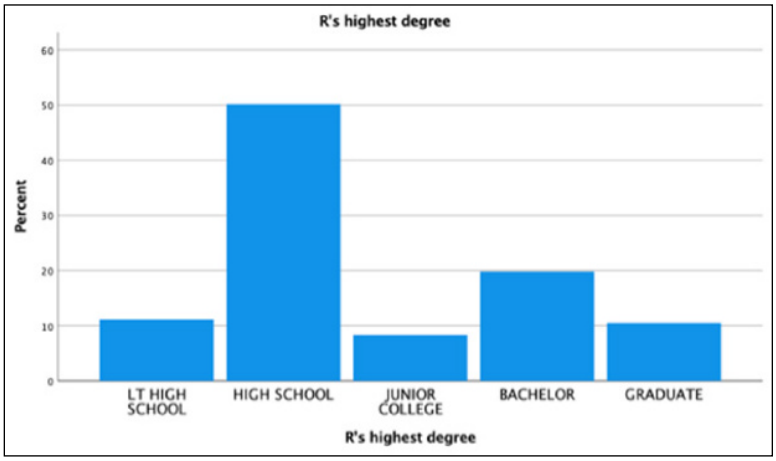
FIGURE 1-2: Pasting a table in Microsoft Word.

See the following for an example of taking SPSS output and bringing into another application.

EXAMPLE

Q. Using the Chapter 1 Output.spv file, copy the graph for R's highest degree and paste it into Word.

A. Select and copy the graph. Then go to Word and choose Edit➪ Paste (or Paste Special).



REMEMBER

Charts are always pasted as images that cannot be edited.

5. Using the Chapter 1 Output.spv file, copy the R's highest degree frequency table and paste in Word as a picture.

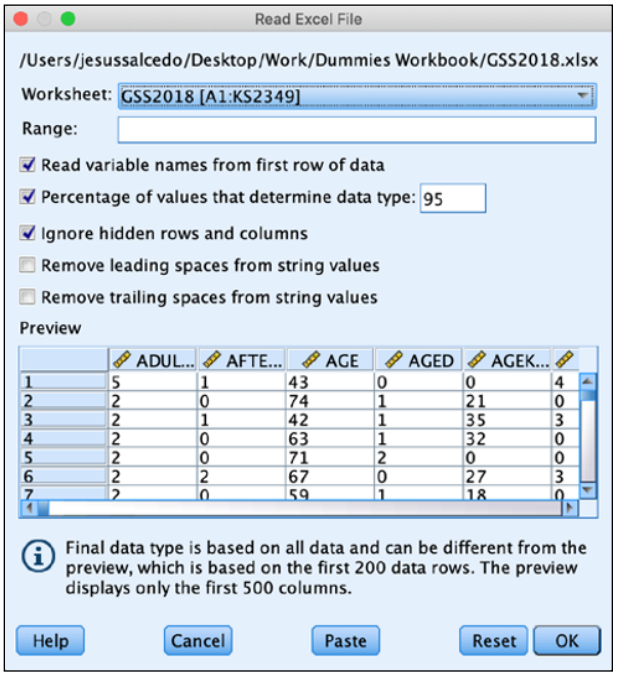
6. Using the Chapter 1 Output.spv file, copy the R's highest degree frequency table and paste it in Excel as a table using Unicode Text.

7. Using the Chapter 1 Output.spv file, copy the R's highest degree frequency table and paste it in Excel as a picture.

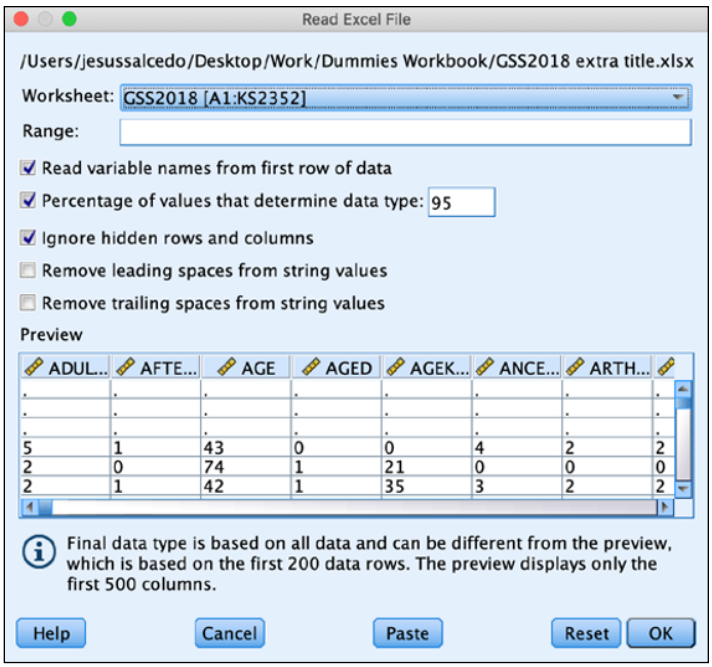
8. Export the Chapter 1 Output.spv file to Word.

Answers to Problems in Working Through Import and Export Challenges

1 Choose File➪ Import Data➪ Excel. Click the Read Variable Names from the First Row of Data check box.



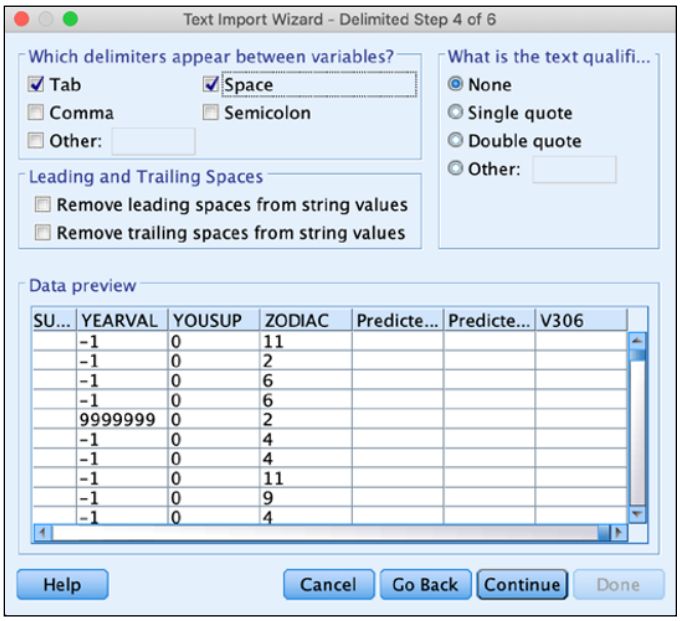
2 Choose File➪ Import Data➪ Excel. Select the GSS2018 worksheet. Click the Read Variable Names from the First Row of Data check box. After the data has been imported, delete the extra rows that appear at the beginning of the file.



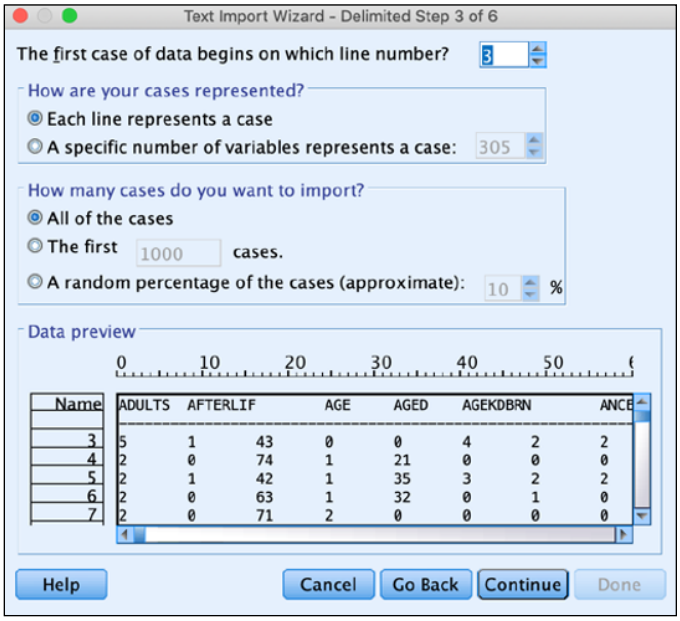
3 Choose File➪ Import Data➪ Text Data. Use the defaults but specify that tab is the only delimiter. Deselecting space as a delimiter will remove the extra column.


TIP

Always check the data preview to make sure you do not have any columns that do not have variable names. If this occurs, you did not select the correct delimiter.



4 Choose File➪ Import Data➪ Text Data. Use the defaults but specify that the first case of data begins on line 3. Also specify that tab is the only delimiter.



5 Select and copy the R's highest degree frequency table. Then go to Word and choose Edit➪ Paste Special and select Picture. You would choose to copy SPSS output as a picture if you did all your editing in SPSS and would like the output to appear exactly as it appeared in SPSS.

R's highest degree

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	LT HIGH SCHOOL	262	11.2	11.2	11.2
	HIGH SCHOOL	1178	50.2	50.2	61.3
	JUNIOR COLLEGE	196	8.3	8.3	69.7
	BACHELOR	465	19.8	19.8	89.5
	GRADUATE	247	10.5	10.5	100.0
	Total	2348	100.0	100.0	

6. Select and copy the R's highest degree frequency table. Then go to Excel, choose Edit➪ Paste Special, and select Unicode Text. You would use the Unicode Text option so that you can further edit your tables in Excel.

	A	B	C	D	E	F	G
1	R's highest degree						
2			Frequency	Percent	Valid Percen	Cumulative Percent	
3	Valid	LT HIGH SCH	262	11.2	11.2	11.2	
4		HIGH SCHOO	1178	50.2	50.2	61.3	
5		JUNIOR COLL	196	8.3	8.3	69.7	
6		BACHELOR	465	19.8	19.8	89.5	
7		GRADUATE	247	10.5	10.5	100	
8		Total	2348	100	100		
9							

7. Select and copy the R's highest degree frequency table. Then go to Excel, choose Edit➪ Paste Special, and select Picture. You would choose to copy SPSS output as a picture if you did all your editing in SPSS and would like the output to appear exactly as it appeared in SPSS.

R's highest degree

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	LT HIGH SCHOOL	262	11.2	11.2	11.2
	HIGH SCHOOL	1178	50.2	50.2	61.3
	JUNIOR COLLEGE	196	8.3	8.3	69.7
	BACHELOR	465	19.8	19.8	89.5
	GRADUATE	247	10.5	10.5	100.0
	Total	2348	100.0	100.0	

8. To export your output from SPSS to Word:

   a. Choose File➪ Export.

   b. In the Objects to Export section, select the items to include in the output. For the example, select All Visible.

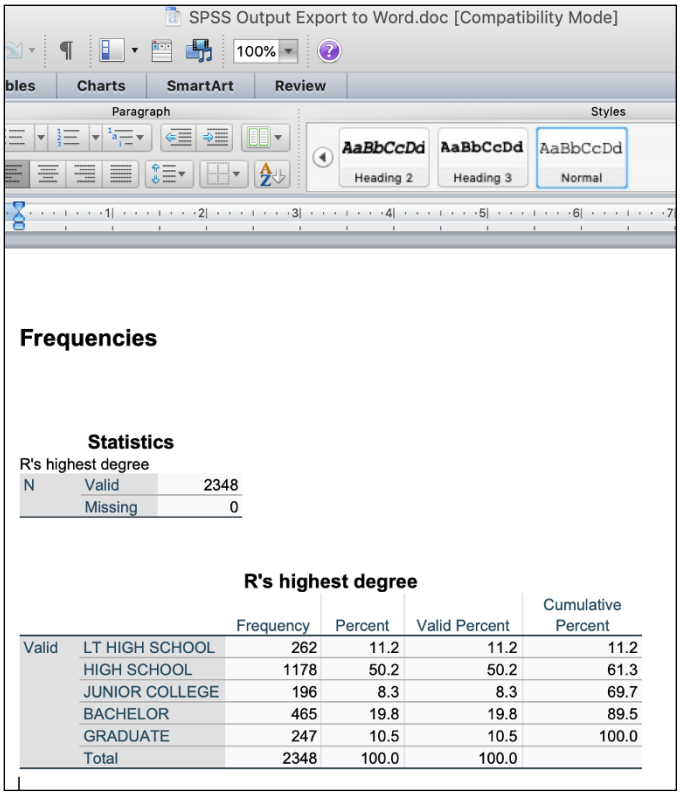
   c. In the Document section's Type drop-down list, choose Word/RTF.

   d. Click the Browse button, select the directory and name of the exported file (SPSS Output Export to Word), and then click Save.

   e. Click OK.

f. To view the exported file, open Microsoft Word.

The exported Word file contains the tables and graph that make up the Chapter 1 Output.spv file. The tables can be edited with the Microsoft Word table editor because the tables were exported as text, not as a picture.



IN THIS CHAPTER

- Defining variables
- Seeing how to work with dates and times
- Finding out how to copy data properties

Chapter 2
Defining Data

SPSS data has three major components: cases (Chapter 3), variables (Chapters 4 and 5), and metadata (Chapter 2). *Metadata* consists of your variable attributes, or definitions. Metadata tells SPSS how a variable is defined and how it can be used. Without a definition, a number serves no purpose. For example, the number 34 could be the number of tickets purchased or the number of sales in the last hour; therefore, a variable's definition is important.

In this chapter, you learn about variable definitions, or attributes. You also work with date and time variables. Finally, because defining data can be time-consuming, we demonstrate a special shortcut menu for copying your data and variable definitions from one dataset to another.

Defining Metadata

Entering data into SPSS is a two-step process. First, you define what sort of data you'll be entering. Then you enter the values. In this section, we briefly describe some of the most important variable attributes in the Variable View tab of Data Editor.

Each variable characteristic has a default, so if you don't specify a characteristic, SPSS uses the default:

- **Name:** Every variable must have a unique name. Following are some handy hints about names:
  - You can use characters in a name, such as @, #, and $, as well as the underscore character (_) and numbers. However, variable names can't start with these special characters.

- Be sure to start every name with a letter, not a number.
- You can't include spaces anywhere in a name, but an underscore is a good substitute.

- **Type:** Most data you enter will be regular numbers. Data such as currency must be displayed in a special format, and data such as dates require special formats so they can be used in calculations. For this type of data, you simply specify what type of data you have, and SPSS takes care of the details for you.
- **Width:** The Width column in the definition of a variable determines the number of characters used to display the value. If the value to be displayed is not large enough to fill the space, the output will be padded with blanks. If it's larger than you specify, it either will be reformatted to fit or asterisks will be displayed.
- **Decimals:** The Decimals column contains the number of digits that appear to the right of the decimal point when the value appears onscreen.
- **Label:** The name and the label serve the same basic purpose: They're descriptors that identify the variable. The difference is that the *name* is the short identifier and the *label* is the long one.
- **Values:** The Values column is where you assign labels to all possible values of a variable. Normally, you make one entry for each possible value that a variable can assume. For example, for a variable named Sex, you could assign the value 1 to the Male label and the value 2 to the Female label.
- **Missing:** You can specify codes for missing data and SPSS will ignore this value when performing calculations.
- **Measure:** Your value for the Measure attribute specifies the level of measurement of your variable. Following are the level of measurement options in SPSS:
  - *Nominal:* A value that specifies a category or type of thing. You can have 0 represent No and 1 represent Yes.
  - *Ordinal:* A value that specifies the position, or order, of something in a list. For example, first, second, and third are ordinal numbers.
  - *Scale:* A number that specifies a magnitude. The scale can be distance, weight, age, or a count of something.

Follow these steps to define a label for a value:

1. **Choose File ➪ Open ➪ Data and load the Happy.sav file.**

   You can download the file from the book's companion website at `www.dummies.com/go/spssstatisticsworkbookfd`.

2. **Click the Variable View tab of Data Editor.**
3. **Click in the Value cell for the Happy variable to open Value Labels dialog.**