

4

Working with IBM SPSS: Statistical Packages for the Social Sciences

Learning Objectives

The principal goal of this chapter is to introduce you to the SPSS environment. Upon completing this chapter, you should be able to:

- Understand the application of pull-down menus
- Understand how to create a codebook
- Enter and import data into SPSS

Key Terms

Codebook

Data analysis menus

Data definition menus

Data file

Output file

Syntax file

Windows and general
purpose menus

INTRODUCTION TO SPSS

You may be wondering why you need to be familiar with a statistical software package. We are certain that more and more nurses, especially those in advanced practice and leadership positions, will find that familiarity with software programs that facilitate quantitative data analysis will make their evidence-based practice work easier and more efficient. Let us take, for example, a chief nursing officer (CNO) of a large chain of long-term care facilities who must make decisions about staffing ratios and staffing mix to ensure that residents receive good quality care and to meet accrediting requirements. The CNO collects and enters data from each facility on daily resident census and absences among the nursing staff. Over time the CNO, using the descriptive statistics that have calculated, learns that peak absenteeism coincides with night shift duty and weekends (not a big surprise). Here, a software package can be helpful in calculating statistics as the volume of data can be large and some necessary statistics can be very complex. Using these statistical findings, the CNO approaches the board of directors for approval of a significant increase in the night shift and weekend salary differential. Having well-organized statistics allows the CNO to advocate for policy changes to promote improved staffing ratios and avoid costly fees associated with violations of accrediting or inspecting bodies. Nurse managers, advanced practice nurses, quality improvement officers, and bedside nurses may all find occasions in which statistical results generated from data are required to make important policy and patient care decisions or influence those in a decision-making role.

The Statistical Package for the Social Sciences (SPSS) is one of the most popular statistical packages available. Statistical software packages are used to assist researchers and clinicians in answering practice and research questions. SPSS originated in the social sciences as the original acronym suggests, but is utilized extensively in many fields because of its simplicity of use. This chapter will give you a brief overview of how to use this software package and we will use SPSS outputs to explain various statistical techniques in later chapters.

Case Study

Hodgin, R. F., Chandra, A., & Weaver, C. (2010). Correlates to long-term-care nurse turnover: Survey results from the state of West Virginia. *Hospital Topics*, 88(4), 91–97.

Researchers Hodgin, Chandra, and Weaver (2010) conducted a descriptive study that explored turnover with a convenience sample of 275 registered and licensed practical nurses employed in long-term care in West Virginia. They found that turnover was not associated with gender, educational preparation, job title, location, or facility size; instead, their findings indicated that turnover was associated with benefits such as pay, scheduling, and career ladder opportunities, and costs including travel time to work and facility characteristics.

The researchers in this case used a statistical test called the chi square test for independence in which associations between categorical/discrete variables are examined to make inferences from the sample to the population. They also used descriptive statistics to characterize the sample of professional nurses who participated in the study. All of these statistics may be accessed in SPSS.

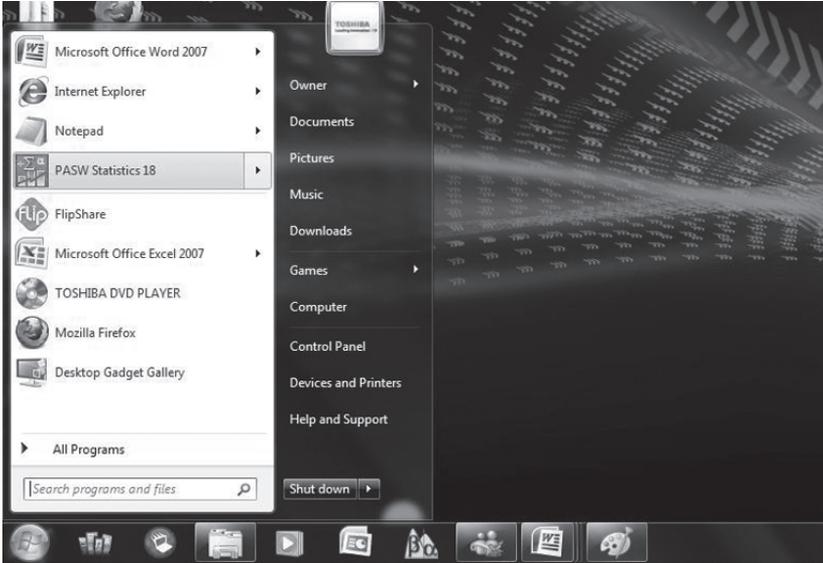
Launching SPSS

You can either double-click the SPSS icon on your desktop, if available, or select SPSS from the Start menu (see **Figure 4-1**). Once you open SPSS, the opening screen should look like **Figure 4-2**.

Major File Types

There are three major file types in SPSS: data files, syntax files, and output files. **Data files** (*.sav) contain the actual data values. Data can

Figure 4-1



be either directly entered into SPSS or be imported from other formats, such as Excel, dBase, and Lotus files. Beginning with version 10.0, data files consist of two views, *data view* and *variable view*. Data view is where the data are inputted and variable view is where the characteristics of

Figure 4-2

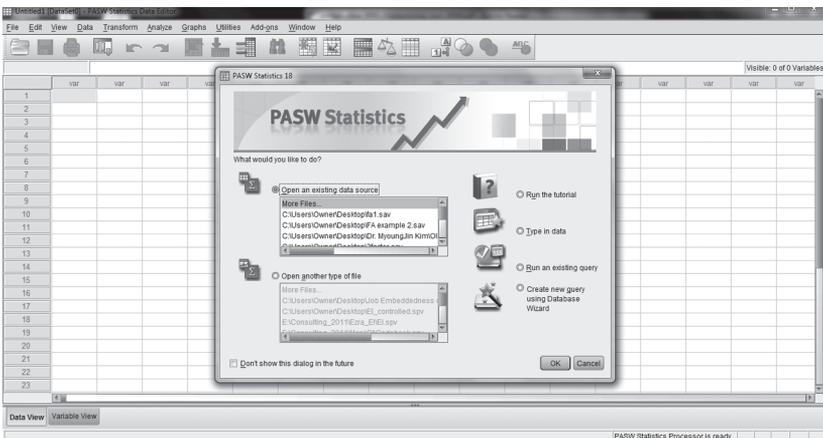
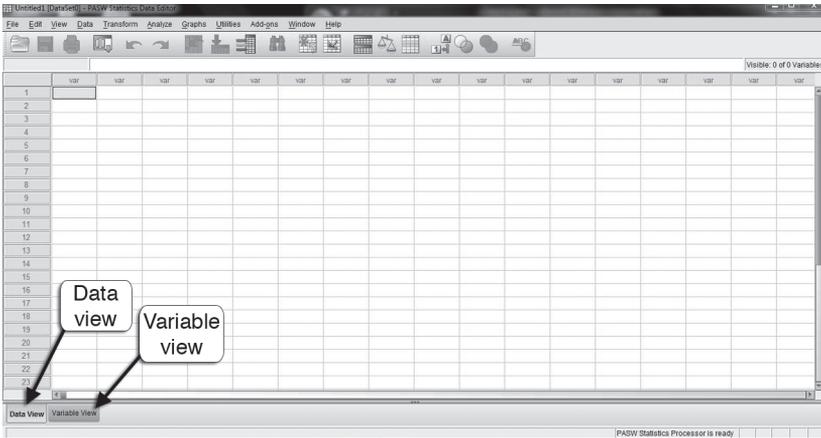


Figure 4-3



the variables, such as variable name, type, label, and values, are defined (See Figure 4-3).

Syntax files contain programmable SPSS commands to conduct analyses (as shown in Figure 4-4), and are a good alternative to using the interactive windows. Although there are certain analyses that can only be performed with syntax files, we will not discuss these in detail

Figure 4-4

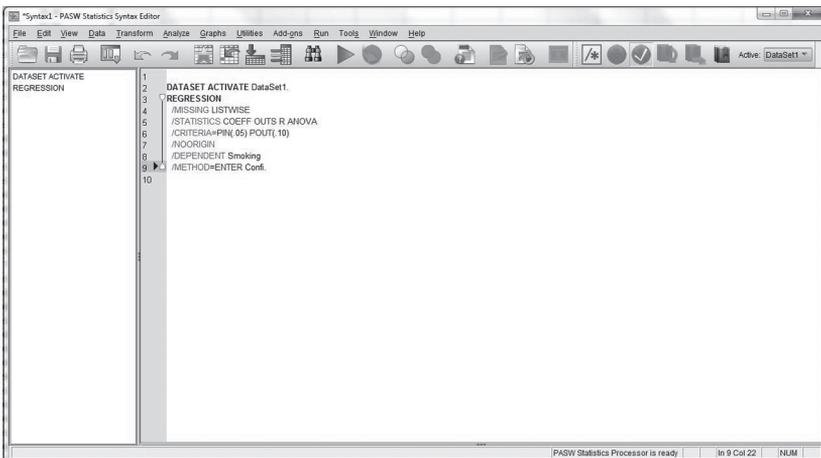


Figure 4-5

Example output files in SPSS.

➔ **Regression**

[DataSet1] J:\Teaching_Spring11\NUR 489.01\Assignments\Assignment 1.sav

Variables Entered/Removed^b

Model	Variables Entered	Variables Removed	Method
1	Self-Confidence ^a	.	Enter

a. All requested variables entered.

b. Dependent Variable: Smoking History

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.856 ^a	.733	.729	10.209

a. Predictors: (Constant), Self-Confidence

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	16334.297	1	16334.297	156.737	.000 ^a
	Residual	5940.246	57	104.215		
	Total	22274.542	58			

a. Predictors: (Constant), Self-Confidence

b. Dependent Variable: Smoking History

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	93.887	5.782		16.238	.000
	Self-Confidence	-.972	.078	-.856	-12.519	.000

a. Dependent Variable: Smoking History

here, as most of the analyses for our purposes can be conducted with interactive windows.

Output files (*.spv) contain the results of your analyses as well as any error messages or warning messages (See **Figure 4-5**). An output file can be saved, edited, printed, or pasted into other applications such as Microsoft Word and PowerPoint.

Types of Pull-Down Menus

As noted above, SPSS is a relatively user-friendly statistical analytic software package due to the interactive windows. On SPSS 18.0, there are a total of 10 pull-down menus, which are categorized into windows and general purpose, data definition, and/or data analyses pull-down menus.

Windows and general purpose menus include FILE, EDIT, VIEW, UTILITIES, and HELP menus. The FILE and EDIT menus should be familiar to most of you as they are similar to those of other software packages. The FILE menu includes options to create new files, open existing files, or save and print the files or outputs. The EDIT menu allows you to find a case or variable, to cut and paste a data value, etc. The HELP menu includes a SPSS tutorial, which we recommend you use as needed as it provides detailed explanations on various topics for a new user. We remember the times when we first used SPSS, and cannot even count how many times the HELP menu saved us! We hope you find this menu as useful as we do.

Data definition menus include DATA and TRANSFORM menus. The DATA menu includes procedures for inserting new variables or cases, sorting cases, merging files, splitting the file, selecting cases, and weighting cases. The TRANSFORM menu includes procedures for recoding variables, replacing missing values, computing new variables using existing variables, and a random number generator. These two menus become useful when a researcher wants to manipulate the file to resolve issues with violated assumptions or extremely unequal sample sizes for different groups. More on this will be discussed in Chapter 7.

Data analysis menus include ANALYZE and GRAPH menus. The ANALYZE menu includes both statistical and psychometric (reliability and validity testing) procedures and the GRAPH menu includes various procedures for creating graphs and plots. One thing to note here is that some of these graph tools are available within statistical procedures under the ANALYZE menu. We will present how to use these functions to perform data analysis and create graphs as we discuss corresponding analytic procedures.

There are many icons on the menu bar that are associated with some of the same procedures as the pull-down menus. However, they will not be discussed as these icons display their function by placing a mouse over them.

Entering Data into SPSS

Completing a Variable View

When entering data into SPSS, the first step is to complete a variable view. This view displays the characteristics of the variables that must be defined so that the data may be entered appropriately into data view. These characteristics include variable name, type, width, decimal, label, values, missing, columns, align, measures, and role. Only the commonly defined characteristics are covered here.

Variable name should begin with a letter and only contain letters and numbers. It should not contain punctuation or spaces and cannot begin with a number. The name can be up to 64 characters in length, but we recommend using short names, approximately 8–10 characters, as variable names longer than this can create wordy outputs that may be difficult to read. For example, a question such, “How many times did you smoke for the past three months?” should be named “Smoke3” instead of using the entire question as a variable name.

Type includes many different formats, but *numeric* and *string* are the most commonly used ones. When the data are quantitative, such as age, the format should be left as the default, numeric. The format should be changed to string when the data are qualitative, such as responses to explain the choice of “other” categories in racial groups.

Decimal characterizes how many decimals are to be shown for variables. The default is set to 2, or the hundredths place, but unless data will be recorded or entered with decimals, we recommend changing to 0 to view the data more clearly.

Label is a description of what a variable represents in a more detail. While we recommend using a short variable name, it may be easy to forget what the abbreviated variable name means. The label field allows up to 255 characters to describe the nature of the variable in more detail. For example, if we are entering data on wound characteristics, the variable name might be “depth,” but the label may be much more detailed, such as “wound depth as measured in millimeters to the nearest whole number.”

Values specify what numbers will be used to represent categories for categorical variables. You will need to indicate how numbers are assigned for categories and this can be accomplished by clicking the gray square box with “...” on the right hand corner of a cell. Let us

consider a case in which we are recording races. We assign a numerical value for each racial group: African American—1, Hispanic—2, Asian/Pacific Islander—3, and White—4. Values allow us to assign a number to a discrete/categorical variable.

Measures represent the level of measurement. Note that interval and ratio levels of measurement are combined in the term *Scale*. **Table 4-1**

Table	
4-1	Characteristics to Define in Variable View
Name	This will be the variable's name that will appear on the top row of data view. Recommend keep name 8 characters or less, as a long variable name will not be shown nicely in data view. Note that you cannot include certain symbols, such as + - % &, nor use spaces.
Type	Define the types of the data. There are different data types such as numeric, comma, date, dollar, and strings; however, you will mostly use numeric (i.e., numbers for data values) or string (i.e., letters/words).
Width	SPSS, by default, sets this at 8 characters. You will be fine with 8 digits for a width of the variable most of the time, but it is subject to change to fit your needs.
Decimals	SPSS, by default, sets this at 2, or the hundredths place. You can change this to other values to fit your needs. For example, you can set it to 0 if your data value will be integers.
Label	Here you can describe the variable in more detail than you can in the short variable name. This will help you keep the variable name short but still remember what it represents.
Values	For categorical variables, you can assign numbers to represent different groups of people. For example, you can assign 1 for male and 2 for female.
Missing	You can assign numbers to missing data. Systematic missing, by default, puts a dot in the cell. You can also assign an impossible value for a certain variable to indicate missing data.
Columns	This is to indicate how many characters will appear in the column. You can change this to fit your needs.
Align	This sets where the data values in data view will appear.
Measure	This represents the level of measurement of a variable. You will choose either nominal, ordinal, or scale.

shows the characteristics that define each variable, and we recommend spending some time in the pull down menu to get the feel of it.

Preparing a Codebook

The next step of data entry is to prepare a **codebook**, which summarizes the characteristics of the defined variables. The codebook guides the creation of the data file and helps minimize errors with the data entry. There are other components that may appear on the codebook, but a typical codebook includes the following:

- Variable name
- Variable labels
- Values for categories of categorical variables
- Values for missing data
- Variable type (i.e., numeric or string)

Go to Analyze > Report > Codebook to create a codebook as shown in Figure 4-6. The variable view of the sample data set is shown in Figure 4-7. Creating a codebook as part of designing a study or evidence-based practice project is often a useful exercise for exposing

Figure 4-6

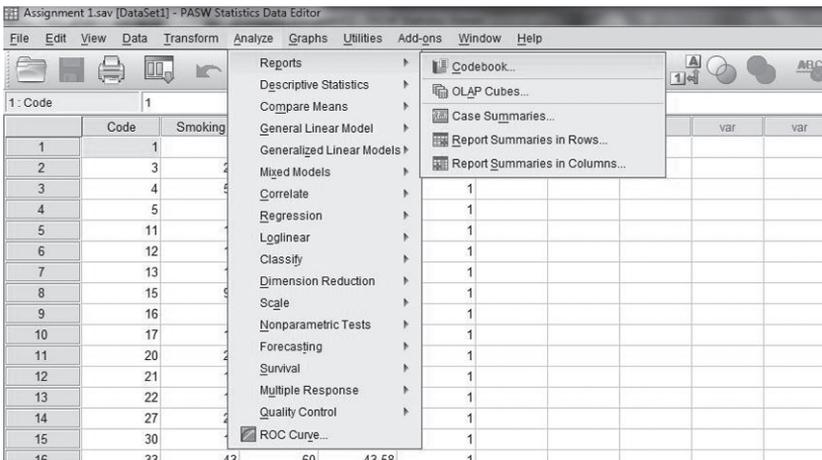


Figure 4-7

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	Smoking	Numeric	8	0	Smoking History	None	None	8	Right	Scale	Input
2	Life	Numeric	8	0	Life Satisfaction	None	None	8	Right	Scale	Input
3	Conf	Numeric	8	2	Self Confidence	None	None	8	Right	Scale	Input
4	Gender	Numeric	8	0	Gender	{1, Male}...	None	8	Right	Nominal	Input
5											
6											

flaws in measurement decisions or prompting questions about how data will be entered and analyzed.

Entering Data

Entering data is simple after the first two steps are complete, as it will be just a matter of typing in the collected data values according to your specifications in the variable view. Table 4-2 contains example data for you to use to practice making a codebook and entering data. Once the above procedures are completed, you should obtain Figure 4-8. Note that Figure 4-8 is not in exactly the same format as you will see in SPSS.

Table				
4-2	Example data to create a codebook.			
Smoking History	Life Satisfaction	Self-Confidence	Gender	
4	40	86.30	1	
27	80	70.18	1	
5	80	61.52	2	
10	67	55.51	2	
9	58	70.10	1	
13	48	95.52	2	

Figure 4-8

Codebook of the sample data set.

Smoking			Life		
		Value			Value
Standard Attributes	Position	1	Standard Attributes	Position	2
	Label	Smoking History		Label	Life Satisfaction
	Type	Numeric		Type	Numeric
	Format	F8		Format	F8
	Measurement	Scale		Measurement	Scale
	Role	Input		Role	Input
N	Valid	6	N	Valid	6
	Missing	0		Missing	0
Central Tendency and Dispersion	Mean	11.33	Central Tendency and Dispersion	Mean	62.17
	Standard Deviation	8.359		Standard Deviation	16.546
	Percentile 25	5.00		Percentile 25	48.00
	Percentile 50	9.50		Percentile 50	62.50
	Percentile 75	13.00		Percentile 75	80.00

Confi			Gender				
		Value		Value	Count	Percent	
Standard Attributes	Position	3	Standard Attributes	Position	4		
	Label	Self Confidence		Label	Gender		
	Type	Numeric		Type	Numeric		
	Format	F8.2		Format	F8		
	Measurement	Scale		Measurement	Nominal		
	Role	Input		Role	Input		
N	Valid	6	Valid Values	1	Male	3 50.0%	
	Missing	0		2	Female	3 50.0%	
Central Tendency and Dispersion	Mean	73.1883					
	Standard Deviation	15.08563					
	Percentile 25	61.5200					
	Percentile 50	70.1400					
	Percentile 75	86.3000					

Importing Data into SPSS

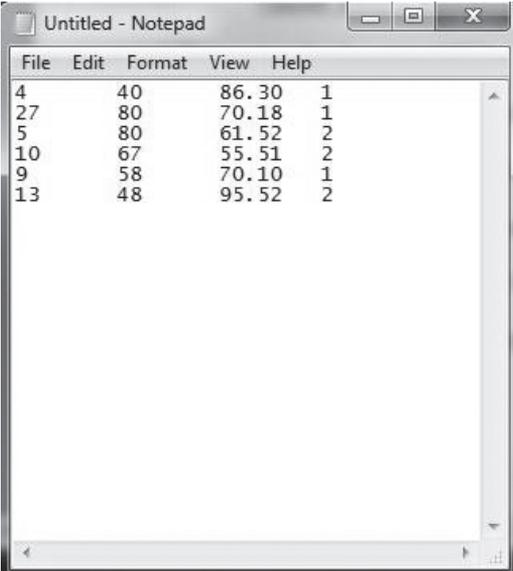
You can also import data from other computer applications/software, such as Microsoft Word or Excel files. While there are other types of formats you can import into SPSS, only text and Excel files are

discussed here as they are the most commonly imported files. Let us consider a situation in which your hospital admissions department keeps patient records using Microsoft Excel, a spreadsheet application used to collect, organize, and conduct arithmetic calculations. As a member of shared governance in the hospital, you have been asked to identify the top 10 admission diagnoses and the length of stay for each primary admission diagnosis. Excel data may be imported into SPSS to calculate a range of statistics that would help identify the most common diagnoses.

Importing a Text File

Text data are composed of contents directly inputted from the computer keyboard and can be easily imported to all computer software. **Figure 4-9** shows the same example data we used in creating a codebook.

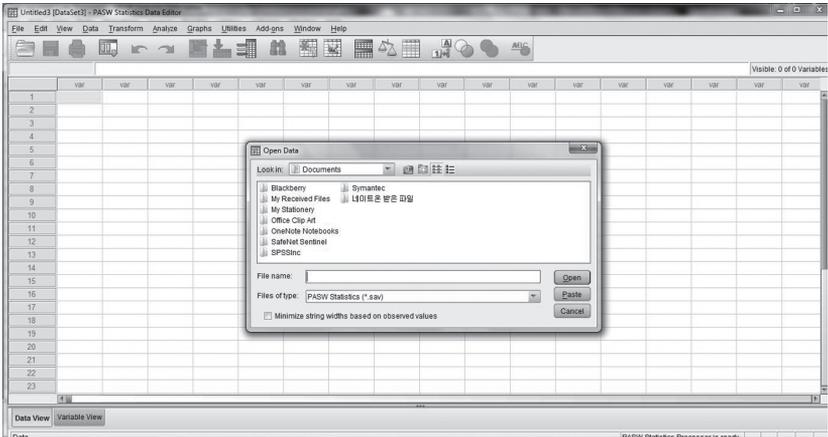
Figure 4-9



The image shows a screenshot of a Notepad window titled "Untitled - Notepad". The window contains a table with four columns of data. The data is as follows:

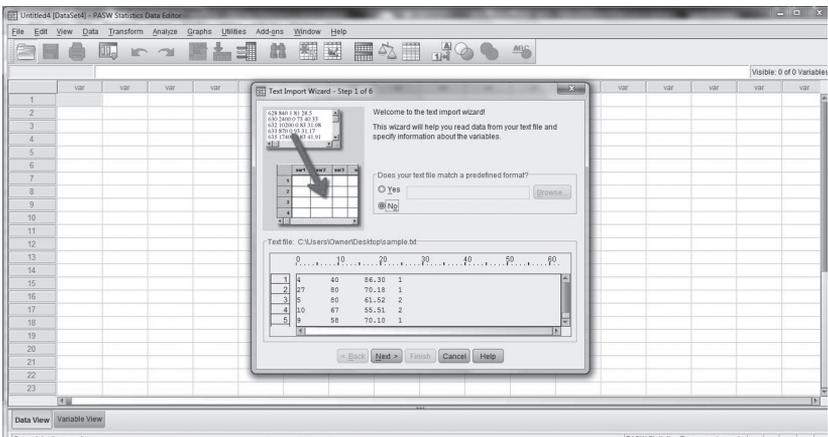
File	Edit	Format	View	Help
4	40	86.30	1	
27	80	70.18	1	
5	80	61.52	2	
10	67	55.51	2	
9	58	70.10	1	
13	48	95.52	2	

Figure 4-10



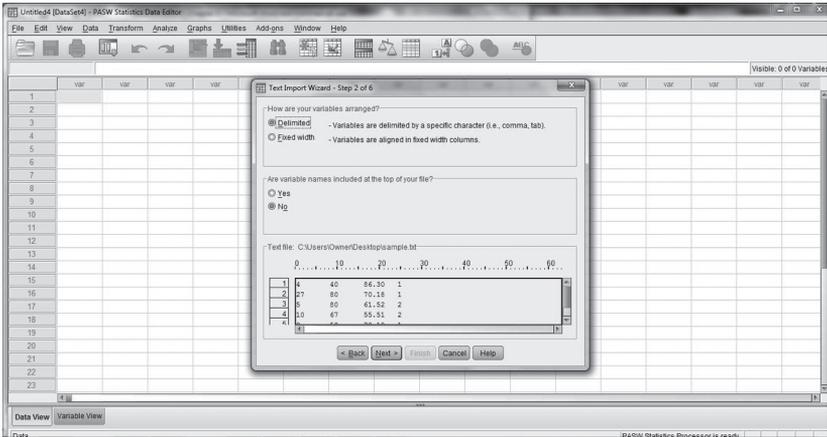
First, go to File > Open > Data. A new window (Figure 4-10) will appear. (Figures 4-11 to 4-16 show the additional screens.) Change “Files of type” from *.sav to *.txt to show your desired text file. Click “open.”

Figure 4-11



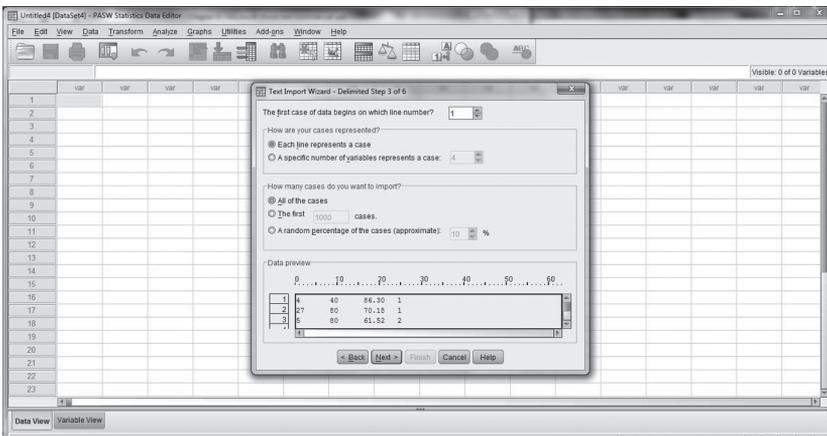
Without special changes, click "Next."

Figure 4-12



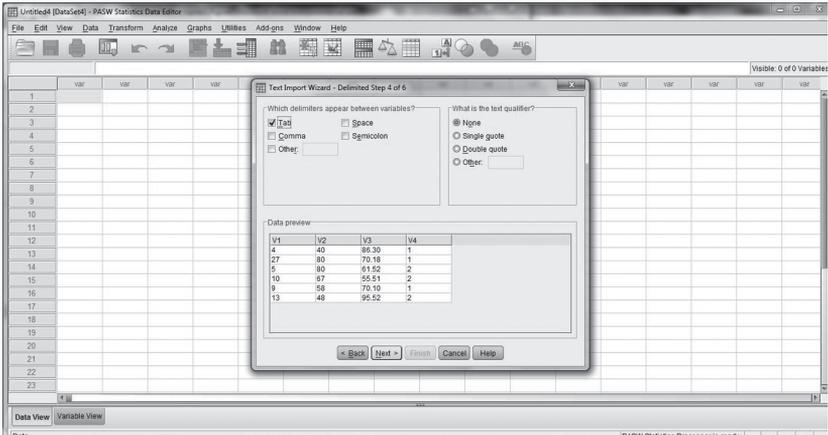
This text file contains the data values separated by tabs, so leave all options in default mode (i.e., delimited and "No" checked) and click "Next."

Figure 4-13



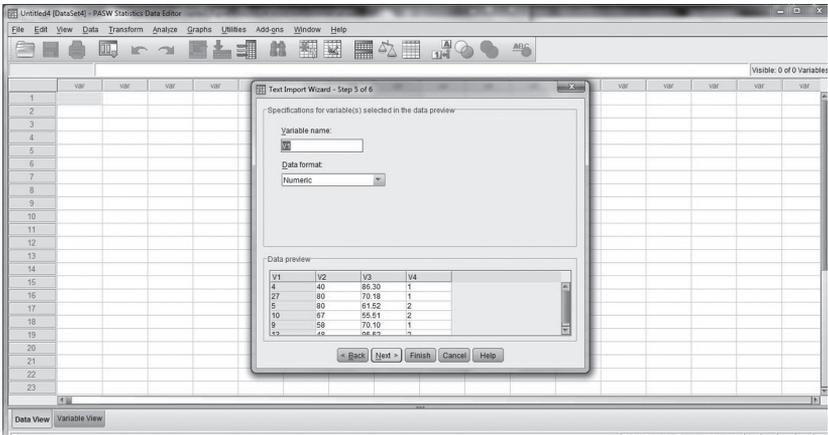
Our data display shows the first case on line number 1 and that each line represents a case. Because we want all of the cases to be imported, do not apply changes and click “Next.”

Figure 4-14



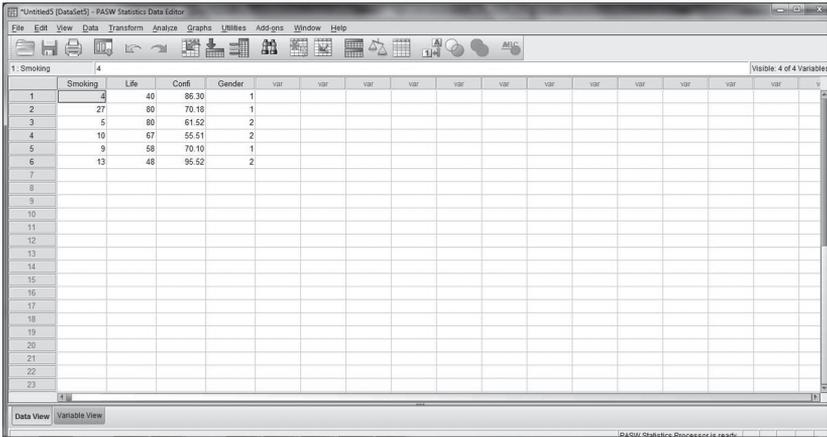
Our data values are separated by tab, so click “Next.”

Figure 4-15



Here, you can specify the variable name and format by clicking each variable at the bottom of the window. Once you are done, click “Next” and then “Finish.” You have now successfully imported a text file into SPSS.

Figure 4-16



Importing an Excel File

Importing an Excel file into SPSS is easier than importing a Text file. Assume the same data file was saved in Excel (**Figure 4-17**).

Figure 4-17

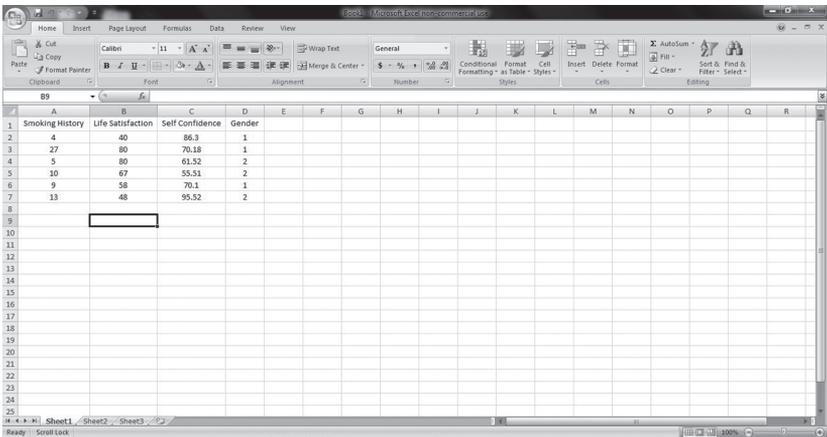
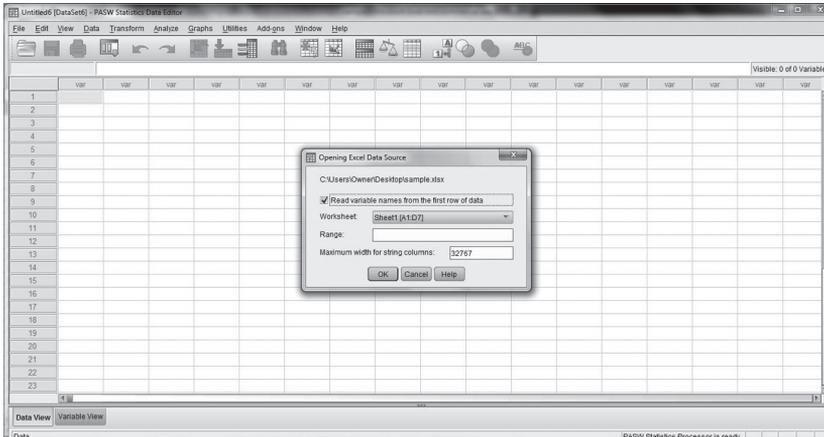


Figure 4-18



Again, go to File > Open > Data. A new window as presented in Figure 4-18 will appear. Change “Files of type” from *.sav to *.xls or *.xlsx to show all Excel files. Open the file as before.

Click OK with the default option if the first row of the Excel file represents the variable name. It should successfully import the Excel file into the SPSS. You may change variable view as needed.

There are many other functions that a researcher can utilize to help with data analysis, but we will discuss them as needed when we explain different types of statistical analyses.

SUMMARY

SPSS is a statistical software package that assists us in efficiently analyzing data to answer evidence-based practice and research questions, and it is one of the most widely used statistical software packages available.

There are three major file types in SPSS: data files, syntax files, and output files. Menus in SPSS are categorized into windows and general purpose, data definition, and/or data analyses pull-down menus.

Entering data begins with defining variables in variable view. Then, a researcher will create a codebook that provides information about the variables in a data set. This should help a researcher with data entry.

Data can be either directly entered into SPSS or different types of data files can be imported into SPSS.

Critical Thinking Questions and Activities

1. What does SPSS stand for?
2. Why is it a good idea to keep variables names short?
3. How will creating a codebook help a researcher in working with SPSS?

Self-Quiz

1. True or False: SPSS only allows you to import the files in *.sav.
2. True or False: A variable name cannot be over 8 characters.

REFERENCE

Hodgin, R. F., Chandra, A., & Weaver, C. (2010). Correlates to long-term-care nurse turnover: Survey results from the state of West Virginia. *Hospital Topics*, 88(4), 91–97.

