Chapter

Organizing and Displaying Data

Learning Objectives

The main goal of this chapter is to ensure that you understand the most common approaches to organizing and displaying data and statistical analyses so that these may efficiently and effectively conveyed to an audience. This chapter will prepare you to:

- Employ techniques for displaying data and statistical analyses in tables and graphs
- Choose the best format to display data and statistical analyses
- Accurately interpret data and statistical analyses presented in a graph or table

Key Terms

Percentile Bar chart

Pie chart **Boxplot** Frequency distribution

Stem and leaf plot Histogram

Scatterplot

Line chart

INTRODUCTION

Nurses in practice, administration, and research must present data or statistical analyses to accomplish a variety of purposes—you may need to make a case for changes in practice or promote the use of resources for a particular initiative, or perhaps you need to convince a patient that their behavior is leading to poor health outcomes. Visual representations of data and analyses are an essential component of dissemination whether you are preparing a written report for your institution, giving an oral presentation with visual aids, or writing a manuscript for publication. Following data collection and analysis, what you have are a bunch of numbers and characters; it is important to be able to communicate the most salient aspects of the data to others. Similarly, as you are reading reports, you need to determine if the researcher made a defensible choice in the data analysis and presentation, and be able to understand a table or graph in the report.

If the number of data values is small, it may be easy to interpret the data set using language only; that is, we simply explain in writing or a verbal report what our findings are. However, if the data set is large or complex, it can be very challenging to figure out what the data have to say. Imagine that you have two data sets: one with measurements on infant mortality from 20 countries, and one with measurements from 200 countries. Which one will be easier to understand?

So, how can the nurse represent data and analyses concisely, accurately, and clearly? One important approach is to use tables, graphs, and charts. The old adage "a picture is worth a thousand words" is also true for data points and statistics. A graph or table depicting the data often tells the story in a more compelling fashion than words alone. However, data presentations can be clean and clear or muddy and misleading depending upon the quality of our choices for display or the decisions of the researcher. In deciding what techniques to use for data display, we must ask ourselves two important questions. The first question is, "When should we use a graph or table?" Graphs and tables are likely to be useful when there is a large amount of, or complex, information to report. The second question is, "What is the best type of graph or table to display our data?" Sometimes, a simple bar chart may be adequate to present data. In other cases, displays that are more complex are needed to communicate precisely with the audience. Data

displays should fit with the level of measurement and variable type and account for the audience characteristics.

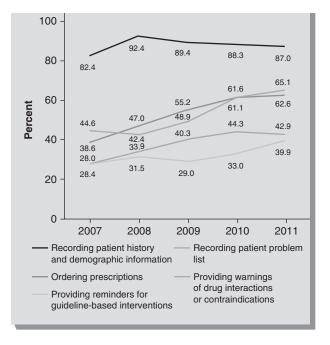
Do you recall the different levels of measurement and types of variables we discussed in Chapter 3? Whether a variable is measured at the nominal, ordinal, interval, or ratio level will, in part, determine what presentation methods you will choose. Suppose we collect data on the gender of subjects. Gender is measured at the nominal level of measurement. A simple bar chart or pie chart may be a good way to convey this information since there are likely to be a limited number of response categories. Now consider data on income of nurses. Income is measured at the ratio level of measurement and the data values are not limited to preset categories. A simple bar chart or pie chart will not be an appropriate display of these data because each income level reported by an individual subject will appear on the chart. Note that a bar or pie chart may be used if income measurements are categorized into a limited number of groups (e.g., 0-\$25,000, \$25,001-\$50,000, and $\ge \$50,000$).

As a nurse in a leadership role, you are expected to accurately interpret and decide on the best approach for displays of data. The goal of this chapter is to provide guidance on what factors to consider when choosing how to present your data and statistical analyses.

Case Study

A well-thought-out analysis and chart can make all the difference on the impact your work has with an audience. In 2009, the United States began an incentive program to promote the use of electronic health records (EHR) that complied with criteria for meaningful use. In 2015, Eric Jamoom and Esther Hing at the National Center for Health Statistics reported trends in health care systems using electronic health records in emergency departments, finding that emergency departments around the country were increasingly able to meet the objectives for meaningful use of EHR systems. The line chart below displays the impact of the incentive program on the use of EHR systems over 4 years.

How might you use such a chart to display data from work that you doing right now? For example, tracking the influence of a new approach to fall prevention on number of falls before and after the intervention; exploring trends in postoperative infections in a same-day surgery center; or following a diabetic client's hemoglobin A_{1C} over time to illustrate their success with diet and exercise?



Jamoom, E., & Hing, E. (2015). Progress with electronic health record adoption among emergency and outpatient departments: United States, 2006–2011. *National Center for Health Statistics Data Brief*, 187.

GROUPING DATA

One of the most common ways of presenting data is a **frequency distribution**, which shows the possible values of a variable and the corresponding frequency of those values. The simplest frequency distribution has two columns, one for data values and the other for

Table								
5-1	Frequency Distribution of Number of Calls at an Emergency Room							
Number of Calls		Frequency						
0		2						
	1	5						
	2	7						
	3	16						
	4	1						

corresponding frequencies. TABLE 5-1 is an example frequency distribution displaying the number of calls per night received at an emergency room in a single month.

The first column shows the number of calls received per night at an emergency room, ranging from zero to four calls. The second column shows how frequently an emergency room (ER) received the corresponding number of calls in one month. There were two nights when the ER received no calls and 16 nights when the ER received three calls.

A frequency distribution table can be extended by adding additional columns such as cumulative frequency and cumulative percentage. Cumulative frequency is the sum of frequency of the current category with that of previous categories, and cumulative percentage is the ratio of cumulative frequency of the category of interest to the total number of subjects. TABLE 5-2 shows the extended frequency distribution table of Table 5-1.

A frequency distribution can be either ungrouped or grouped. Our previous example was an ungrouped frequency distribution. If the data are measured at the categorical level, either nominal or ordinal, an ungrouped frequency distribution is the usual choice as there will be a limited number of category responses. TABLE 5-3 shows a frequency distribution table for gender.

If the variables are measured at the interval or ratio level of measurement, the choice of ungrouped or grouped frequency distribution depends on the range of the data values. If the range of data values is small, such as found in Tables 5-1 and 5-2, an ungrouped frequency

Table				
5-2	Extended Emergenc	- /	ribution of Number	r of Calls at an
Number	of Calls	Frequency	Cumulative Frequency	Cumulative Percentage
	0	2	2	0.06
	1	5	7	0.23
	2	7	14	0.45
	3	16	30	0.97
	4	1	31	1.00

Table	
5-3	Frequency Distribution Table for Gender
Gender	Frequency
Male Female Trans Total	20 70 5 95

distribution table may be still appropriate. If not, a grouped frequency distribution may be a more efficient way of displaying the data.

Suppose that you want to create a frequency distribution of mortality rates of 200 countries. The data will range over many different values and an ungrouped frequency distribution table would have to be quite large to display the data. In this case, a grouped frequency distribution table will show the data more concisely as the distinct intervals of data values will be grouped to simplify the information about a variable. TABLE 5-4 shows a grouped frequency distribution table of mortality rates of 200 countries.

A grouped frequency distribution table is a much easier way to convey information when the data set is large or complex, but we will

Table							
5-4	Grouped Frequency Distribution Table of Mortality Rates						
	ty Rates (Death 000 Live Births)	Frequency					
	0-10	9					
	11-20	27					
	21-30	42					
	31-40	111					
41	and above	11					

lose individual data. For example, how would you answer the question, "Which of these countries have an infant mortality rate of 5 or 6 deaths per 1,000 live births?" It is impossible to answer that question by looking at the grouped frequency distribution table. While a grouped frequency distribution provides good summative information on the data set, we lose information on individual data values.

To create a frequency distribution in IBM SPSS Statistics software ("SPSS"), you will open Frequency.sav and go to Analyze > Descriptive Statistics > Frequencies as shown in FIGURE 5-1. In the Frequencies box, you will then move a variable of interest—in this case, Frequency—into "Variable(s)" as shown in FIGURE 5-2. Clicking "Continue" and then "OK" will then produce the output, shown in FIGURE 5-3.

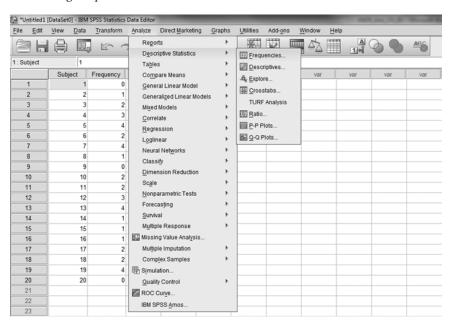
Which frequency distribution table is the best choice depends again on how the data are being measured and the range of values. If the range of data values is small, choosing an ungrouped frequency distribution table and retaining as much of the information as possible is probably best. If the data range is large or very complex, a grouped frequency distribution table will convey information in a more concise format even though some details of the information will be lost.

GRAPHS AND CHARTS

Nurses using data and statistics for practice, quality/process improvement, research, and evidence-based practice may choose to convey

Figure 5-1

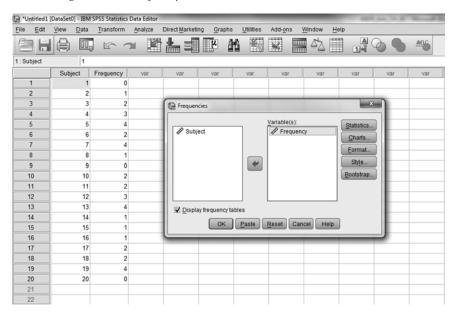
Selecting Frequencies in SPSS.



information about data in graphs and charts that would be difficult or cumbersome to examine in text format. In fact, graphs and charts are the best method of describing data when the data set is large. Graphs and charts are visually impactful and, when well designed, can be easily understood. There are many different types of graphs and charts available, and we will now move to a discussion of how to determine what graphs and charts is the best choice.

A useful chart or graph should show the data or statistics in a meaningful, clear, and efficient manner. Poor choices may lead to ambiguous or misleading interpretations of the data. Suppose we needed to present the collected measurements of the weights of 100 patients. Since each patient will have different measurements in weight, choosing a pie chart as a description method will not display the data clearly, as shown in FIGURE 5-4. When there are many data values to display, seeing each data value as a separate category does not help

Defining a variable frequency in SPSS.



us understand the data. In the case of patient weights, the use of a histogram is a much better way to understand the data (FIGURE 5-5).

There are many types of graphs and charts in SPSS, and we will explain how to create each with examples of variables. We also discuss the appropriate levels of measurement for each graph and chart. Each data set used here is also included in the online resources for this text, accessed using the code found in the front of this text.

Discrete or Categorical Data

Bar charts and pie charts are two useful ways to represent discrete data (i.e., data that are categorical with fixed number of categories measured at the nominal or ordinal level). They are commonly used charts as they are easy to create, use, and interpret.

Figure 5-3

Example output of frequency distribution table in SPSS.

Example output of frequency distribution table in SPSS.

⇒ Frequencies

[DataSetl] E:\2011\StatBook_2011\Chapters\chapter 5\Frequency.sav

Statistics

Frequency

N	Valid	20
	Missing	0

Frequency

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	3	15.0	15.0	15.0
	1	5	25.0	25.0	40.0
	2	6	30.0	30.0	70.0
	3	2	10.0	10.0	80.0
	4	4	20.0	20.0	100.0
	Total	20	100.0	100.0	

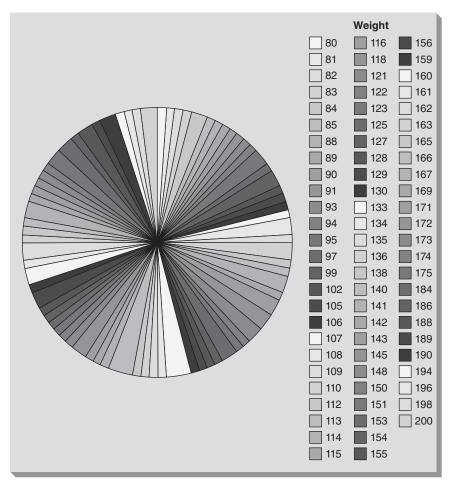
Bar Chart

The bar chart is the most appropriate choice for variables measured at the nominal and ordinal level of measurement and can be used to display one or more variables. If a bar chart is used with the ordinal level of measurement, ordering the rank helps the reader in interpreting the chart (e.g., if discussing pressure ulcers, listing the measurements in order—stage I, stage II, stage III, stage VI). A typical bar chart has the response categories on the horizontal axis and the frequencies of each category on the vertical axis; this helps you discern much about the data, such as the most/least common category, the difference of one bar relative to the other, and changes in frequency over time.

To create almost any chart or graph for a variable in SPSS, you will open Location.sav, go to Graph > Legacy Dialogues, and select the

Figure 5-4

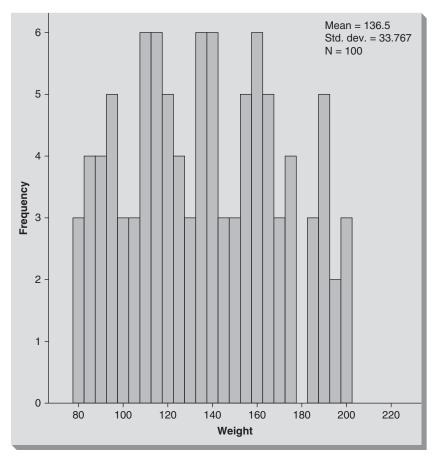
Mistakenly used pie chart.



type of graph/chart desired—in this case, Bar, as shown in FIGURE 5-6. In the Bar Charts box, you will leave "Simple" and "Summaries for groups of cases" selected as default, and click "Define" as shown in FIGURE 5-7. In the Define Simple Bar box, you will then move the variable of interest—in this example, Location—into "Category Axis" as shown in FIGURE 5-8. Click "OK" to produce the output (FIGURE 5-9).

Figure 5-5

Histogram using the same data set as Figure 5-3.



Note that a bar chart can also be created horizontally (i.e., the response categories on the vertical axis and the frequencies of each category on the horizontal axis). Additional examples where a bar chart can be handy to present the data include ethnicity, insurance category (Medicare/Medicaid), marital status, patient acuity, and gender.

Pie Chart

A **pie chart** is a circular chart where pieces within the chart represent a corresponding proportion of each category, and it is an appropriate

Selecting a bar chart in SPSS.

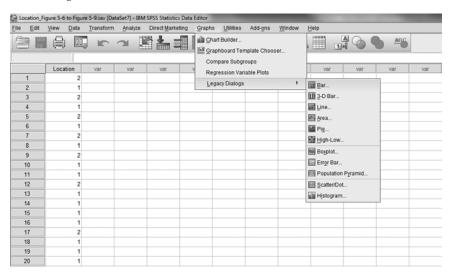


Figure 5-7

Selecting a simple bar chart in SPSS.

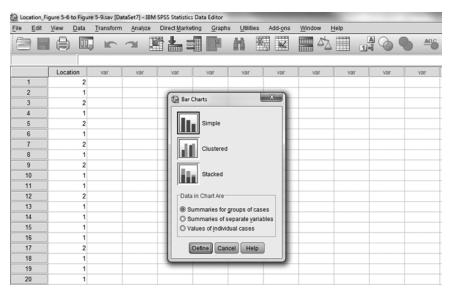
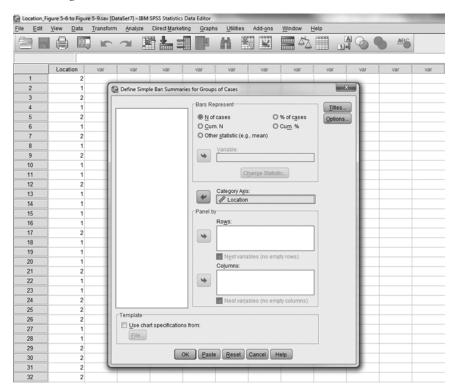


Figure 5-8

Defining a variable in a bar chart in SPSS.

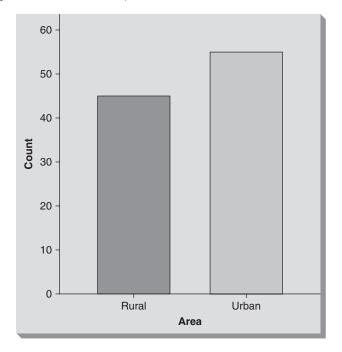


choice for nominal and ordinal level of measurement. A pie chart is simple to create, use, and understand, much like a bar chart when it is created for a single variable. Pie charts are very useful for visualizing the most commonly occurring class compared to the whole and the relative size of different classes. However, it may be difficult to compare data across different pie charts.

To create a simple pie chart for a variable in SPSS, you will open Ethnicity.sav and go to Graph > Legacy Dialogues > Pie, as shown in FIGURE 5-10. In the Pie Charts box, you will leave "Summaries for groups of cases" checked as default and click "Define" as shown in FIGURE 5-11. In Define Pie box, you will then move a variable of

Figure 5-9

Example bar chart for nurses' job location.



interest, Ethnicity, into "Define Slices by" as shown in FIGURE 5-12. Clicking "OK" will then produce the output. An example output is shown in FIGURE 5-13.

Continuous Data

When the variable is continuous (i.e., interval and ratio), both bar charts and pie charts become inefficient in displaying the collected data. It is important to show how the data are distributed with continuous data, since the data values can range differently, unlike with categorical data. Better choices for continuous variables are histograms, stem and leaf plots, and boxplots.

Figure 5-10

Selecting a pie chart in SPSS.

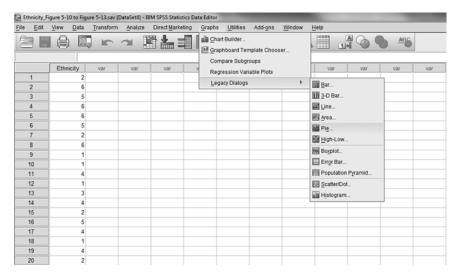
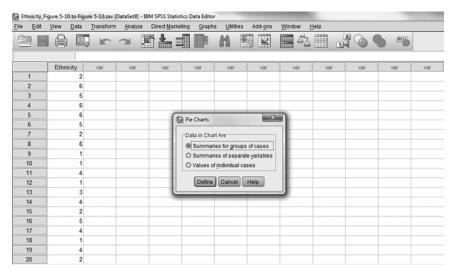
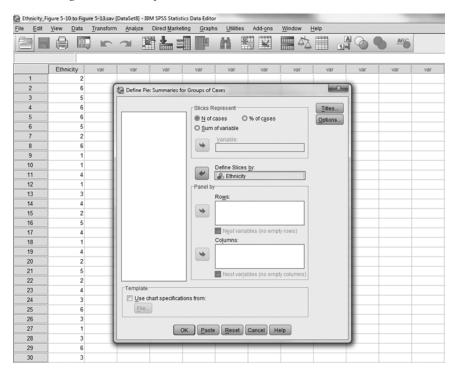


Figure 5-11

Selecting a simple pie chart in SPSS.



Defining a variable in a pie chart in SPSS.



Histogram

A histogram is similar to a bar chart in structure, which explains why histograms are often mistaken for bar charts. However, you can think of it as a graphical way of presenting information from a frequency distribution. It organizes a group of data points into a number of intervals, and the bar in a histogram represents the frequency in corresponding intervals, not in predefined limited numbers of categories as with a bar chart. With a histogram, you can understand the general shape of the data distribution (i.e., what are the data trends) and the most commonly occurring data values.

To create a histogram in SPSS, you will open Age.sav and go to Graph > Legacy Dialogues > Histogram as shown in **FIGURE 5-14**. In

Figure 5-13

Example pie chart for nurses' ethnicity.

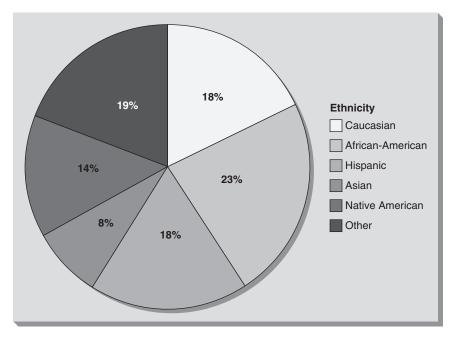
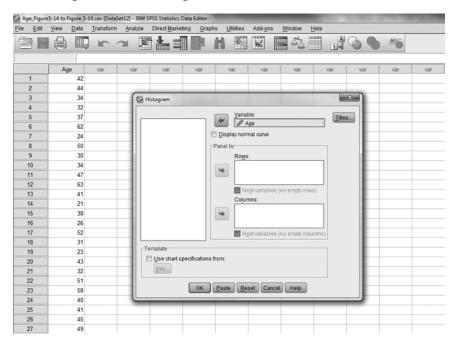


Figure 5-14

Selecting a histogram chart in SPSS.

ile	<u>E</u> dit	View	<u>D</u> ata	Transform	Analyze	Direct Marketing	Graphs	<u>U</u> tilities	Add-ons	Window	<u>H</u> elp	
			00,		21 E		<u> </u>	t Builder hboard Ten pare Subgr	mplate Choo	ser		
		Ag		var	var	var	Regression Variable Plots				var var var	var
	1		42				Legacy Dialogs				Bar	
	2		44				200	,				
	3		34								11 3-D Bar	
	4		32								Line	
	5		37								Area	
	6		62								₽i <u>e</u>	
	7		24								High-Low	
	8		50								Boxplot	
	9		30								Error Bar	
	10		34									
	11		47								Population Pyramid	
	12		63								Scatter/Dot	
	13		41								Histogram	
	14		21									
	15		38									
	16		26									
	17		52									
	18		31							_		
	19		23									
	20		43									

Defining a variable in a histogram in SPSS.



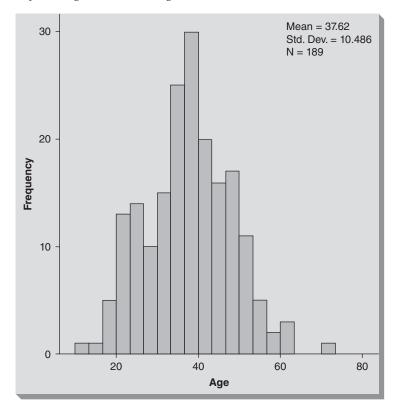
the Histogram box, you will then move a variable of interest, Age, into "Variable" as shown in FIGURE 5-15. Clicking "OK" will then produce the output. An example output is shown in FIGURE 5-16. Note that a histogram can be obtained in other places, such as "Explore."

Stem and Leaf Plot

Stem and leaf plots are good charts for showing the distribution of continuous data. These are similar to histograms, but have greater flexibility and display more information. In addition to the overall shape of distribution, a stem and leaf plot also shows information regarding individual data values. To create a stem and leaf plot in SPSS, you will open Satisfaction.sav and go to Analyze > Descriptive Statistics >

Figure 5-16

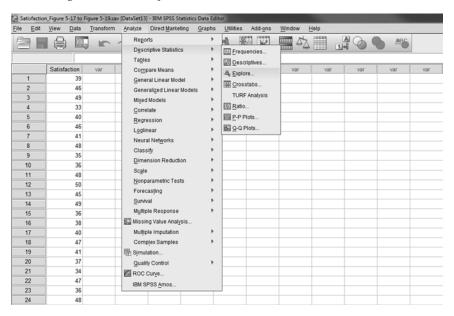
Example histogram for nurses' age.



Explore as shown in FIGURE 5-17. In the Explore box, you will move a variable of interest, Job Satisfaction, into "Dependent List" as shown in FIGURE 5-18. Note that you can move the categorical variable into "Factor List" if you want to create separate stem and leaf plots for a categorical variable; this will create separate plots for different categories of the variable. Stem and leaf plot is checked as the default in "Plots" button, so clicking "OK" will then produce the output. FIGURE 5-19 is an example of a stem and leaf plot.

As you can see in Figure 5-19, the variable, Satisfaction, looks similar on both sides from the center and the actual data values are shown.

Selecting a stem and leaf plot in SPSS.



In this way, stem and leaf plots not only show the distribution, but also information about individual data values.

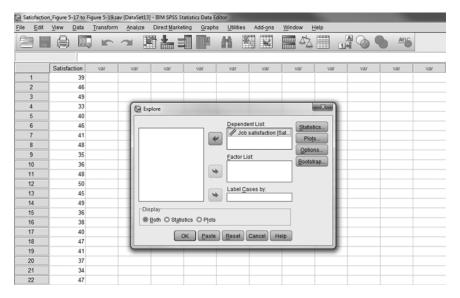
Boxplot

A **boxplot** can be used to display more information than any other chart discussed so far in this chapter. It is a good choice for variables measured on the continuous scale and allows for comparisons across groups. A boxplot does not show individual data values like a stem and leaf plot, but does display other information, such as the overall distribution, the center of the distribution, the quartile, and possible outliers.

To create a boxplot in SPSS, you will open Heartrate.sav and go to Graph > Legacy Dialogues > Boxplot as shown in **FIGURE 5-20**. In the Boxplot box, you will leave "Simple" and "Summaries for groups of

Figure 5-18

Defining a variable in a stem and leaf plot in SPSS.



cases" selected as default, and click "Define" as shown in FIGURE 5-21. In the Define Simple Boxplot box, you will move a variable of interest, Heart Rate, into "Variable" and Gender into "Category Axis" as shown in FIGURE 5-22. Clicking "OK" will then produce the output. An example output is shown in FIGURE 5-23.

A boxplot can be drawn both vertically, as shown in Figure 5-23, and horizontally. It is one of the important charts used to describe the data in exploratory data analysis. Interpretation of a boxplot is as follows.

- The box in the plot contains the middle 50% of the data set. The middle line in the box represents the 50th percentile, the exact middle number of the entire data set, while the upper edge represents the 75th percentile and the lower edge represents the 25th percentile.
- If the middle line is not exactly in the middle of the box, it is an
 indication that the data is not equally distributed in both sides
 from the center.

Example stem and leaf plot for nurses' job satisfaction.

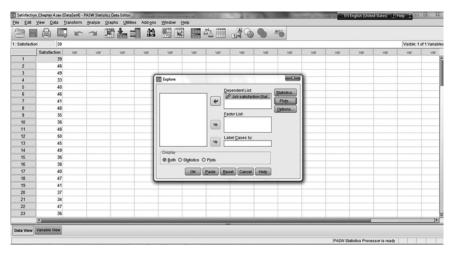


Figure 5-20

Selecting boxplot in SPSS.

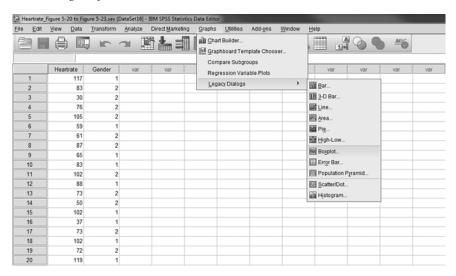
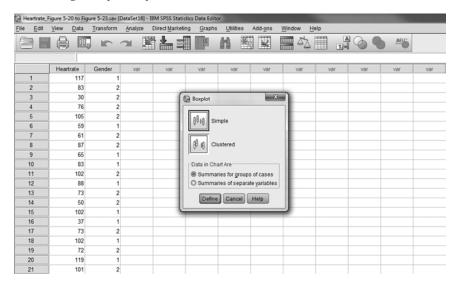


Figure 5-21

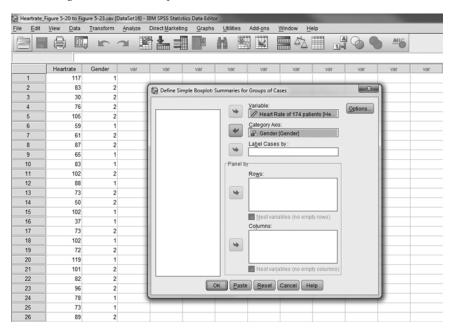
Selecting a simple boxplot in SPSS.



- The ends of the vertical lines, which are called "whiskers," represent the minimum and maximum data values. The lower whisker equals 1.5 times the interquartile range (IQR) below the first quartile (25th percentile), and the upper whisker equals 1.5 times the IQR above the third quartile (75th percentile).
- Any data value outside of the whiskers is considered to be a possible outlier, which is defined as an unusual data value in the current data set.

We need to give you an explanation on percentile before we finish this discussion on boxplots. **Percentile** is a measure of location and tells us how many data values fall below a certain percentage of observations. For example, if you have were in the 75th percentile on an exam, then you did better than 75% of the other people taking that exam

Defining variables in a boxplot in SPSS.



OTHER GRAPHS AND CHARTS

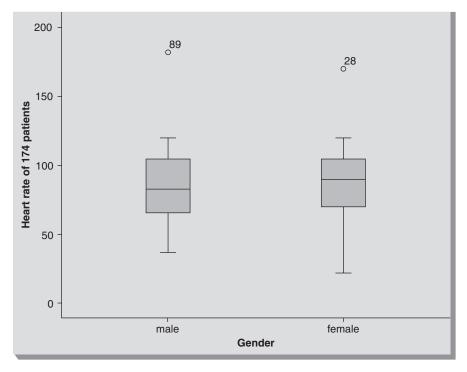
There are other graphs and charts that may be useful for displaying data. They include line chart and scatterplot. Line charts are used to examine trends of variable over time, and scatterplots are used to examine the relationship between variables.

Line Chart

Like the bar chart and pie chart, the **line chart** is good choice for displaying the frequency of categories. It is created by connecting dots, representing the data values of each category as shown in **FIGURE 5-24**. In this example, the horizontal axis represents "Month," a categorical variable, and the vertical axis represents "Systolic Blood Pressure," a continuous variable.

Figure 5-23

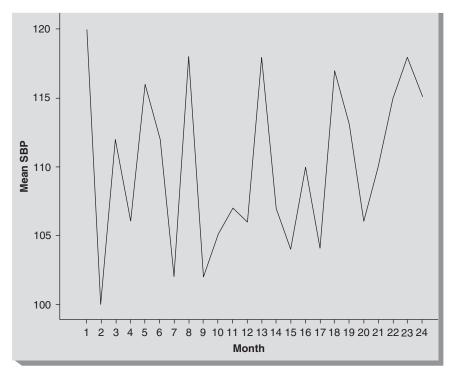
Example boxplot for 174 patients' heart rate.



A line chart is useful when trying to find and compare changes over time or when trying to define meaningful patterns of variables. To create a line chart in SPSS, you will open SBP.sav and go to Graph > Legacy Dialogues > Line as shown in FIGURE 5-25. In the Line Charts box, you will leave "Simple" and "Summaries for groups of cases" selected as default, and click "Define" as shown in FIGURE 5-26. In the Define Simple Line box, you will then move a variable of interest, Systolic Blood Pressure, into "Variable" after clicking a radio button for "Other statistics" (e.g., mean) and Month into "Category Axis" as shown in FIGURE 5-27. Clicking "OK" will then produce the output.

Figure 5-24

Example line chart for systolic blood pressure.



Scatterplot

Scatterplots are used when a researcher wishes to examine the relationship between two continuous variables—for example, age and systolic blood pressure. Relationships can be in either positive or negative directions. A positive relationship means that both variables move in the same direction, such as increased smoking and the increased probability of getting lung cancer, while a negative relationship means that the variables move in opposite directions, such as lower self-esteem and increased depression. FIGURE 5-28 shows an example of a scatterplot.

To create a scatterplot in SPSS, you will open WeightHeight.sav and go to Graph > Legacy Dialogues > Scatter/Dot as shown in FIGURE 5-29.

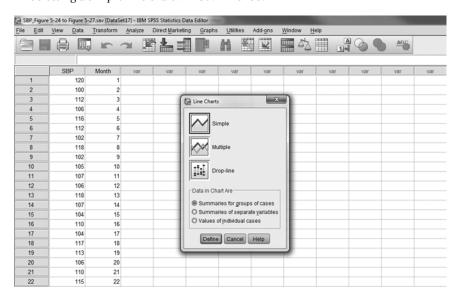
Figure 5-25

Selecting a line chart window in SPSS.

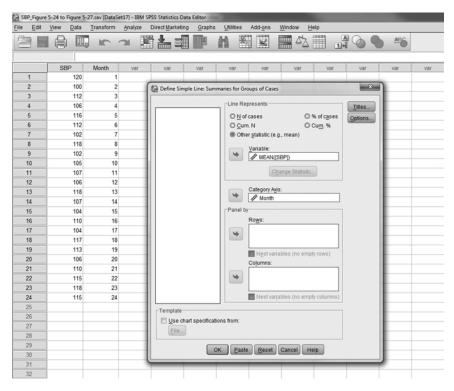
₽ SBP_F	igure 5-24 to Figu	re 5-27.sav [Data	Set17] - IBM	SPSS Statistics Data	a Editor					
<u>F</u> ile <u>E</u>	dit <u>V</u> iew <u>D</u> a	ta <u>T</u> ransform	<u>A</u> nalyze	Direct Marketing	Graphs	Utilities	Add- <u>o</u> ns	Window	<u>H</u> elp	
			71			rt Builder hboard Ten	nplate Choos	er		6
		V	·		Con	pare Subgr	oups			_
	SBP	Month	var	var	Reg	ression Vari	able Plots		var var var	
1		20			Lea	acy Dialogs		-	Bar	
2		00 2							11 3-D Bar	
3		12 3	3					-	Line	
5			5							
6		12						-	Area	_
7		02							Pi <u>e</u>	_
8		18 8							High-Low	_
9		02 9							Boxplot	
10	1	05 10)						III Error Bar	
11	1	07 1	1						Population Pyramid	
12	1	06 12	2						Scatter/Dot	
13	1	18 13	3						Histogram	
14		07 14								
15		04 15								
16		10 16								
17		04 17								
18	1	17 18	3							

Figure 5-26

Selecting a simple line chart window in SPSS.



Defining a variable in a line chart in SPSS.



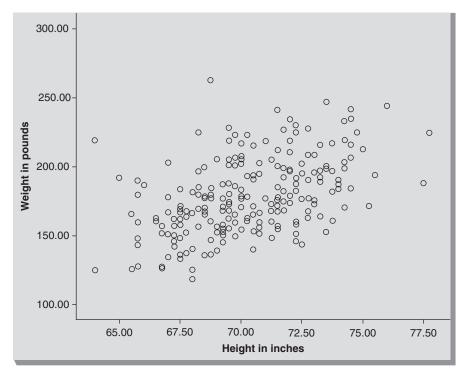
In the Scatter/Dot box, you will leave "Simple Scatter" selected as default, and click "Define" as shown in **FIGURE 5-30**. In the Simple Scatterplot box, you will move an independent variable, Height, into "X Axis" and a dependent variable, Weight, into "Y Axis" as shown in **FIGURE 5-31**. Clicking "OK" will then produce the output.

PRESENTING THE DATA IN THE BEST FORMAT

We have discussed many different methods of displaying data and statistical analyses, and most of these tables and charts are easy to

Figure 5-28

Example scatterplot for a relationship between height and weight.



create, use, and understand. However, if any of these charts are not carefully designed they can present false or misleading information. You should consider four elements selecting the best type of graph or table for your data.

First, you need to ask yourself whether you have chosen the most appropriate type of graph for the data. This means considering level of measurement and size and complexity of the data set. For example, you would not want to generate a histogram to display the data values on an ethnicity variable or a bar chart for sodium content level of 100 patients. Second, you should make sure that you have provided enough information on each component of the graph. The title for the graphic should be descriptive and the variable(s) should be clearly

Selecting a scatterplot in SPSS.

le	Edit	<u>V</u> iew <u>D</u> ata	Transform	Analyze	Direct Marketing	Graphs	<u>U</u> tilities	Add-ons	Window	<u>H</u> elp			
			E,	⊃ı [<u> </u>	Builder hboard Ten	nplate Choos	ser	1		ARC	
		weight	height	var	var		ession Vari			var	var	var	var
1		180.50	69.50			_	cy Dialogs			69 a			
- 2	2	205.25	71.25			Fe24	icy Dialoga	_	-	Bar			
3	3	187.75	69.50							11 3-D Bar			
4	1	182.00	73.75							Line			
ŧ	5	160.25	71.25							Area			
(6	168.00	71.25							Pie			
7	7	218.50	71.00							High-Low			
8	3	247.25	73.50										
(9	191.75	65.00							Boxplot			
1	0	202.25	70.00							Error Bar			
1	1	196.75	68.25							Population	Pyramid		
1	2	203.00	67.00							Scatter/Dot			
1	3	262.75	68.75							Histogram.			
1	4	205.00	69.50							mm - natogram.			
1	5	217.00	70.00										
1	6	212.00	71.50										
1	7	125.25	68.00										
1	8	164.25	73.25										

Figure 5-30

Selecting a simple scatterplot in SPSS.

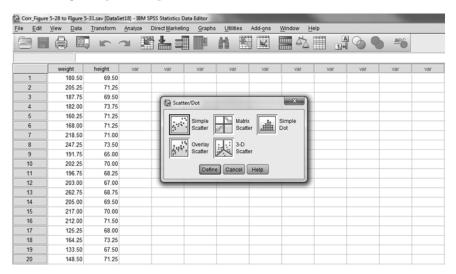
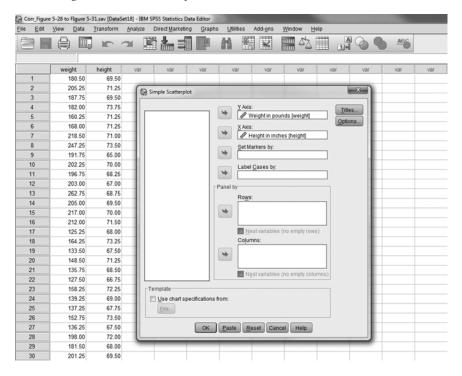


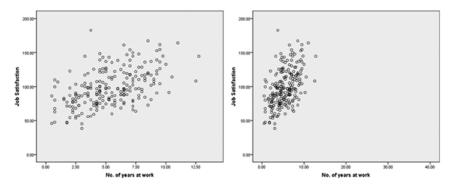
Figure 5-31

Defining variables in a scatterplot in SPSS.



identified. Third, make sure that the independent variable is placed on the horizontal axis while the dependent variable is placed on the vertical axis. Reversed, the graph may illustrate a different picture than what you want to show. Finally, you also have to ensure that the graph has been drawn on the proper scale, as it may show a different scenario than what it actually is occurring if it is incorrectly drawn. **FIGURE 5-32** is a perfect example of how an inappropriately scaled graph can be misleading. While the data do not show a strong relationship between the number of years worked at the current job and job satisfaction (as shown in the graph on the left), the graph on the right shows a very strong positive relationship due to the inappropriately defined scale.

Example of data distortion.



SUMMARY

Graphs and charts are a useful and efficient way of displaying data. When the table, graph, or plot is created carefully and appropriately, your data will be more clearly understood and more meaningful. The careful nurse researcher, clinician, and leader should keep in mind what each chart is good for and choose a chart accordingly.

Bar charts and pie charts are good for displaying the frequency or percentage of given categories. Line charts are also good when the variable for the horizontal axis is categorical.

Histograms, stem and leaf plots, and boxplots are good for displaying the distribution of continuously measured variables. A histogram is similar to a bar chart in structure, but it is used to show the distribution of data values in a user-defined range. Stem and leaf plots show the same information as a histogram, but they show the actual data. A boxplot is probably the chart with the most information, as it shows information on potential outliers, the center values, and the 25th and 75th percentiles.

However, all of these charts can be easily manipulated to produce false information if they are not carefully designed. Therefore, we should think carefully about which type of graphs or charts will best

fit the data, be certain to include enough information on each component of the graph so that readers can accurately interpret the display.

Critical Thinking Questions

- 1. What are the purposes of constructing a graph or table to display information about a variable?
- 2. Levels of measurement are an important factor in determining which chart to use. Explain why this is the case.

Refer to the graph on the next page for questions 3-5.

- 3. Does this chart seem to be appropriate for this data? Why or why not? Explain your answer.
- 4. Is the title appropriately worded?
- 5. Is there any component of this chart that you think it is not complete or is confusing? Explain.

Self-Quiz

- 1. True or false: A histogram is useful when a researcher is trying to display information about a categorical variable.
- 2. True or false: Bar charts and pie charts can convey similar types of information.
- 3. True or false: There is no chart that allows an investigator to identify possible outliers.
- 4. Which of the following is a good example of data that can be appropriately displayed with a line chart?
 - a. Ethnicity
 - b. Systolic blood pressure
 - c. Income
 - d. Age
- 5. Which of these charts allows a researcher to examine a possible relationship between two continuous variables?
 - a. Histogram
 - b. Bar chart
 - c. Scatterplot
 - d. Line chart

Reference 109

REFERENCE

Jamoom, E., & Hing, E. (2015). Progress with electronic health record adoption among emergency and outpatient department: United States, 2006–2011. *National Center for Health Statistics Data Brief*, 187.

 $\hbox{@ Jones \& Bartlett Learning LLC, an Ascend Learning Company. NOT FOR SALE OR DISTRIBUTION.}\\$