Chapter 3

Finding Sums

This chapter covers procedures for obtaining many of the summed values that are commonly used in statistical calculations. These procedures will produce values that are identical to the hand calculations typically required in introductory and advanced statistics courses and these procedures can be used to check your work. Specifically, we demonstrate methods for finding ΣX and ΣY. Also, this section presents examples that use the compute command to create new variables with sums that would reflect the following: ΣX^2, ΣY^2, $\Sigma(X-1)$, $\Sigma(Y-1)^2$, and ΣXY. Even if you have already done so, it may be a good idea for you to review chapter 2, both before you proceed with this section and as a reference as you step through these procedures.

Setting Up the Data

For the following examples, we turn to the cartoon below (Bush-in-the-White House Data Set). Assume that a researcher is interested in the degree to which attitudes about having a "bush" in the white house influence the frequency with which a "bush" is actually used. Based

Table 3.1 Bush-in-the-White House Data Set

Dog	*Desire a "bush" in the White House (X) Independent Variable* (X)	*Number of Times a "bush" was Used (Y) Dependent Variable* (Y)
1	1	2
2	2	2
3	3	3
4	4	5
5	4	7
6	5	9
7	5	11
8	7	11
9	8	12
10	9	13

on self reports obtained from ten dogs, the data are presented in Table 3.1. The first column of data represents the participant number, the second column labeled *X* represents the independent variable (Desire for a "bush" in the White House), and the third column labeled *Y* represents the dependent variable (Number of Times a "bush" was used). As with all of the SPSS examples in this book, our first job is set up the data set in the SPSS data editor. First, we need to define our variables. Figure 3.1 is the variable view of the SPSS data editor. We have created two new variables named *X* and *Y* to correspond with our Independent Variable (Desire for a "bush" in the White House) and our Dependent Variable (Number of Times a "bush" was used), respectively.

Figure 3.1 SPSS Variable View for Bush in the White House Data

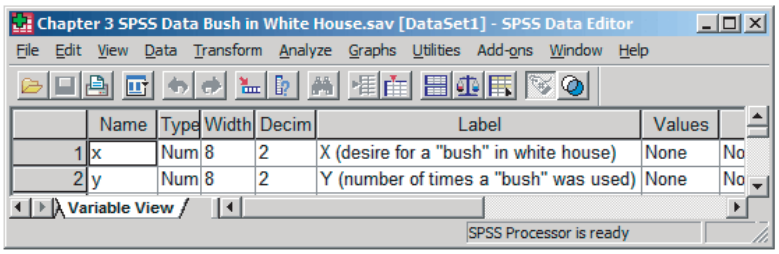


Figure 3.2 SPSS Data View for Bush in The White House Data

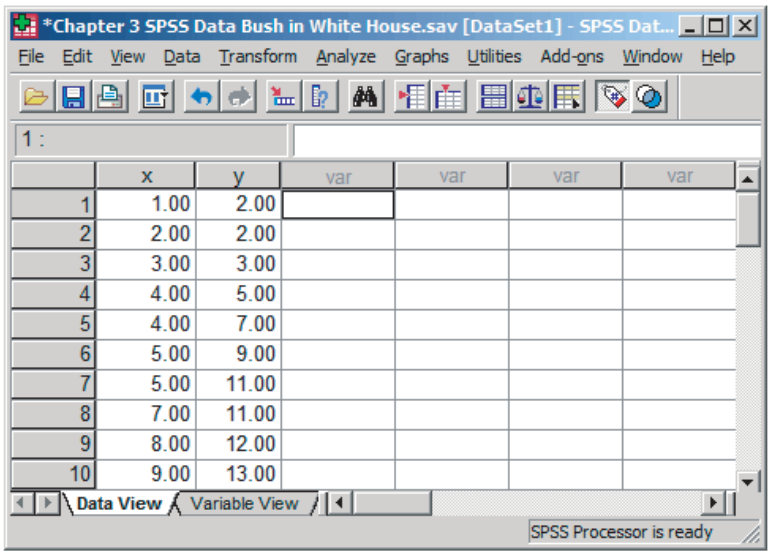


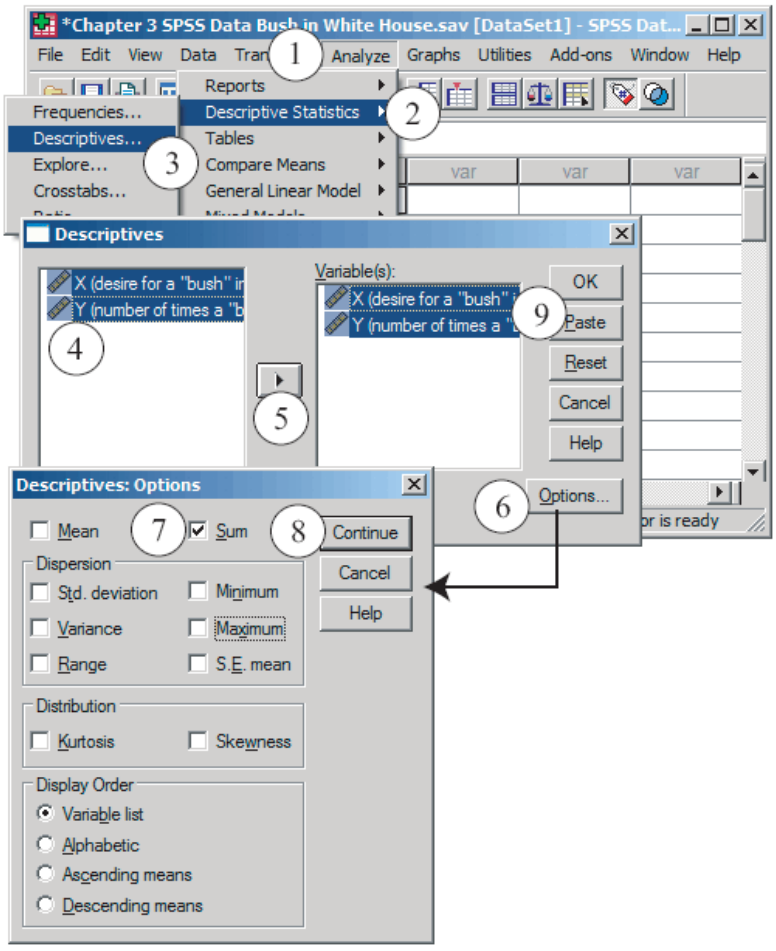
Figure 2.3 presents the data view of the SPSS data editor. Here, we have entered the data

for all 10 the dogs in our data set for the two variables we are investigating. Remember that columns in the data view of the data editor represent the variables and the rows represent each case (e.g., subject, observation, unit, dog, etc.).

Running the Analysis

In this first example we present the steps for running descriptive analyses, which is the simplest way to obtain descriptive information about one or more variables. Note that SPSS calls them Descriptive Statistics, but remember that whether they are statistics or parameters depends on whether your data reflects a population or a sample.

Figure 3.3 SPSS Running Descriptive Analyses



Descriptive Statistics Steps (See Figure 3.3): From the **Analyze** (1) pull down menu, select **Descriptive Statistics** (2), then select **Descriptives...** (3) from the side menu. In the **Descriptives** dialogue box, enter the variables (4) **X** and **Y** in **Variable(s)** field by either double-left-clicking on each

variable or selecting each variable and left-clicking on the boxed arrow pointing to the right, (5). To select the desired descriptive information you want, select **Options** (6). In the **Descritpives: Options** dialogue box you will need to check the options you want. Notice that the Mean, Standard Deviation, Minimum and Maximum options are automatically selected. In our example we have unselected these options and have checked the **Sum** (7) option. This will simplify the resulting output that we obtain. Next, click **continue** (8). Finally, double check your variables and either select **OK** (9) to run, or **Paste** to create syntax to run at a later time.

If you selected the paste option from the procedure above, you should have generated the following syntax:

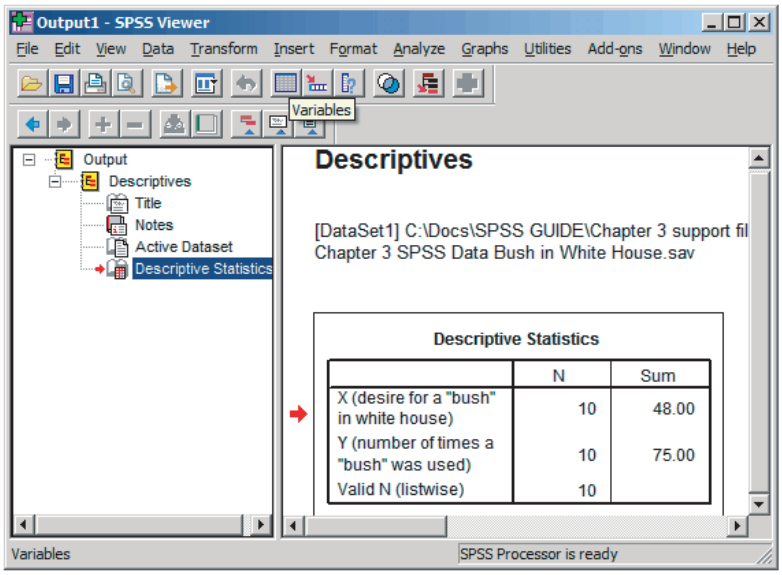
```
DESCRIPTIVES
  VARIABLES=x y
  /STATISTICS=SUM .
```

To run the analyses using the syntax, while in the Syntax Editor, select **All** from the **Run** pull-down menu.

Reading the Output

As presented in Figure 3.4, the requested descriptive information for each of the variables we included (in this case the sums of X and Y) is presented within the columns of the Descriptive Statistics table. Each row of data within the table represents a different variable and is labeled with the Variable Label that you gave each variable (if you did not give the variable a label then the variables will be identified by their less descriptive variable name). In our current example the first column of data reports the number of cases for each variable (N). In this case

Figure 3.4 SPSS Output: Sum of X and Y for Bush in the White House Data



we have data for 10 dogs for both X and Y. The last row of this column [labeled "Valid N (listwise)"] tells reports the number of cases (dogs) that are not missing any data. This information will be presented every time you run the descriptive analyses, regardless of whether you request it or not. In this example, we have data for all 10 dogs for both variables. However, if we did not know the number times a "bush" was used (Y) for one dog, then the Valid N (listwise) would be 9 instead of 10. Similarly, if we were also missing data for a different dog for variable X or Y, then the Valid N would be 8 instead of 10. With respect to the sum of X and Y, we obtain the values 48 and 75, respectively.

Checking Your Hand Calculations

As you can see, running these analyses can be an accurate way of double checking the work that you do by hand. If it happens that the SPSS output and your hand calculations do differ, it means one of them is incorrect. It is probably best to first check the data that you entered in the data editor and ensure that it is correct, before you start redoing the more time consuming hand calculations. If the SPSS data seems correct, then go back and check your calculations and see if you can identify where you made your mistake.

Finding ΣX^2, ΣY^2, and other Complex Summations

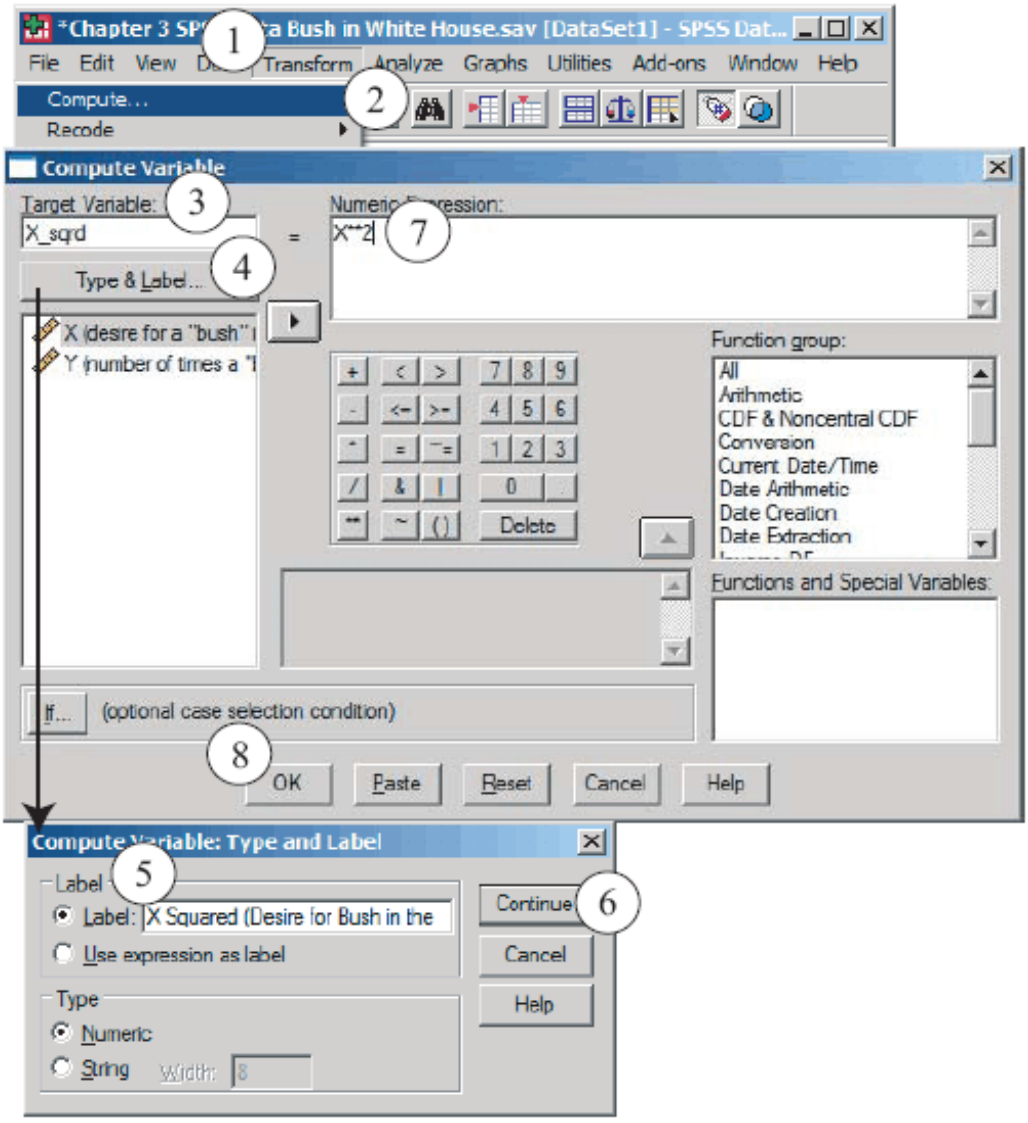
Often we are interested in obtaining sums that require some kind of mathematical operation that is applied to each individual score before we sum them. For example, ΣX^2 tells us to square each score for variable X before we sum it. SPSS allows us to do this by creating a new variable using the Compute Variables procedure. The use of the compute statement is illustrated bellow for ΣX^2.

Running Compute Variable Procedure

Compute Variable Steps (See Figure 3.5): From the **Transform** (1) pull down menu, select **Compute** (2). In the **Compute Variable** dialogue box, enter the desired variable name for the new variable in the **Target Variable** (3) field. In this case we have named our new variable **X_sqrd**. To give the new variable a variable label, click on the **Type & Label** (4) button, and enter the desired variable label [in this case: **X squared (Desire for Bush in White House Squared Before Summing)**] in the **Label** field. Also, note that the mathematical equation used to create the variable can be used as the variable label by selecting the **Use expression as label** option. Click **Continue** to return to the **Compute Variable** dialogue box.

The final steps involve to writing the mathematical equation (numerical expression). In this case we want to square variable X. First, enter the variable name of the variable to be altered in the **Numeric Expression:** field (7). Next, write the desired mathematical equation. In this case we use the following equation: **X**2**. In SPSS ** is the exponential operator. Would could

Figure 3.5 SPSS Using Compute to Create X^2



also compute our desired X^2 variable using this equation: **X*X**. Like most computer programs, SPSS treats * as the multiplication operator. Finally, double check your variables and equations, and either select **OK** (8) to run, or **Paste** to create syntax to run at a later time.

If you selected the paste option from the procedure above, you should have generated the following syntax:

```
COMPUTE X_sqrd = x ** 2 .
VARIABLE LABELS X_sqrd 'X squared (Desire for Bush in White House Squared Before Summing)' .
EXECUTE .
```

To run the analyses using the syntax, while in the Syntax Editor, select **All** from the **Run** pull-down menu.

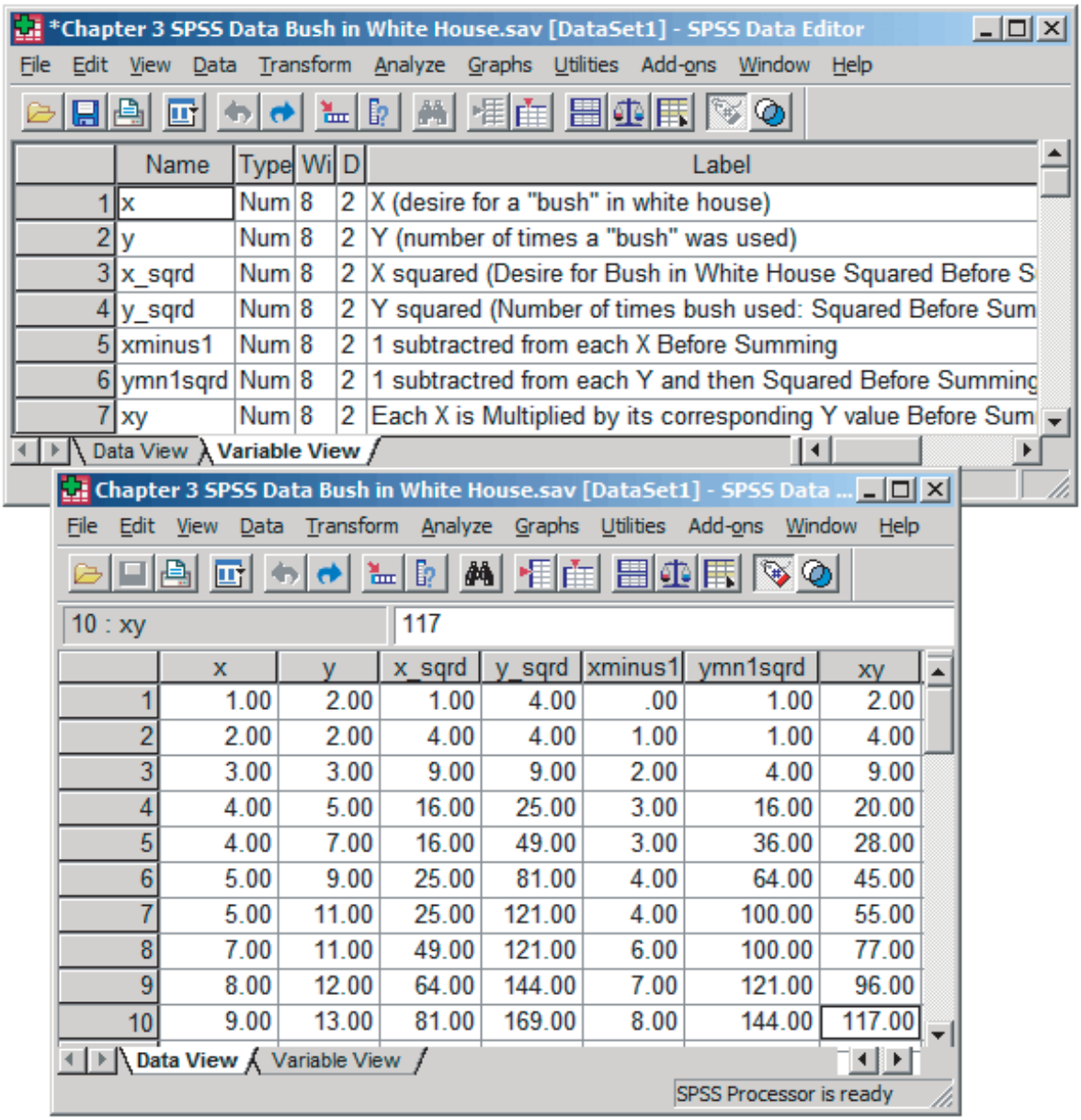
Similarly we can create other complex variables using appropriate numerical expressions, which we have presented in Table 3.2. The first column lists the variable names for the newly created variable. The second column presents the new variable label. The numerical expression used to create the new variable is presented in the last column.

Table 3.2 More Compute Variable Procedures

New Variable Name	New Variable Label	Numerical Expression
Y_sqrd	Y Squared (Number of times bush was used: Squared Before Summing)	Y**2
Xminus1	1 subtracted from each X Before Summing	X-1
Ymn1sqrd	1 subtracted from each Y and then Squared Before Summing	(Y-1)**2
XY	Each X is multiplied by its corresponding Y value Before Summing	X*Y

After running the desired compute statements, the new variables are added to the data editor. Figure 3.6 presents the variable view and data view of the SPSS data editor after running the five compute statements presented above. As each compute statement is run, the newly created variable is inserted into the next available row of the variable view and the next available column of the data view. To keep the new variables for future use, you will need to save the data file before closing.

Figure 3.6 Variable View and Data View with Newly Computed Variables



Find the Sums and Interpret the Output

Once the desired variables have been computed the sums for each new variable can be obtained by repeating the steps for running descriptive statistics presented above (illustrated in Figure 3.3) and adding the newly created variables to the **Variable(s)** field. The analyses can either be run, by selecting **OK**, or saved to run later, by selecting **Paste**. Selecting **Paste** will should provide the syntax which is similar to the following.

```
DESCRIPTIVES
  VARIABLES=x y x_sqrd y_sqrd xminus1 ymn1sqrd xy
  /STATISTICS=SUM .
```

This syntax requests the sums for the original *X* and *Y* variables as well as the sums for the newly created variables.

Figure 3.7 SPSS Output: Sum X, Y, and Newly Created Variables Using Bush in the White House Data

Descriptives

Descriptive Statistics

	N	Sum
X (desire for a "bush" in white house)	10	48.00
Y (number of times a "bush" was used)	10	75.00
X squared (Desire for Bush in White House Squared Before Summing)	10	290.00
Y squared (Number of times bush used: Squared Before Summing)	10	727.00
1 subtractred from each X Before Summing	10	38.00
1 subtractred from each Y and then Squared Before Summing	10	587.00
Each X is Multiplied by its corresponding Y value Before Summing	10	453.00
Valid N (listwise)	10	

Figure 3.7 presents the output obtained by running the descriptive analyses. Again the rows present the data for each variable, which are identified using their respective variable labels.

More Compute Operators

Table 3.3 Compute Variable Mathematical Operators

Operators					
+	addition	<	greater than	~	not
-	subtraction	>	less than	~=	not equal to
*	multiplication	<=	greater than or equal to	&	and
/	division	>=	less than or equal to	|	or
**	exponentiation	=	equal to	( )	Grouping operator e.g. order of operations

Table 3.3 present other mathematical operators that SPSS uses in compute statements. As you become more familiar with SPSS you will find many of these operators to quite useful. These operators can be used separately or in combination to produce a variety of data transformations. The most common use of compute statements by researchers is to combine scores from multiple variables, such as test items or questions on a survey, to form a single score (typically a sum or an average). For example:

var1 + var2 + var3 : will produce a new variable that is the sum of three variables.

(var1 + var2 + var3)/3 : will produce a new variable that is the average of three variables.

SPSS also provides Functions that are operator shortcuts. For example:

var1 + var2 + var3 can also be obtained with the function: SUM(var1, var2, var3)

(var1 + var2 + var3)/3 cam also be obtained with the function: MEAN(var1, var2, var3).

Summary

This chapter presented methods of obtaining sums of simple variables (e.g. *X* & *Y*) and complex variables obtained by transforming existing variables using the Compute Variable operation. These procedures can be used to the hand calculations typically required in

introductory and advanced statistics courses. This chapter also introduced advanced Compute Operators that you may find useful as you become more advanced users of SPSS.