Unit 2

Summarizing Data

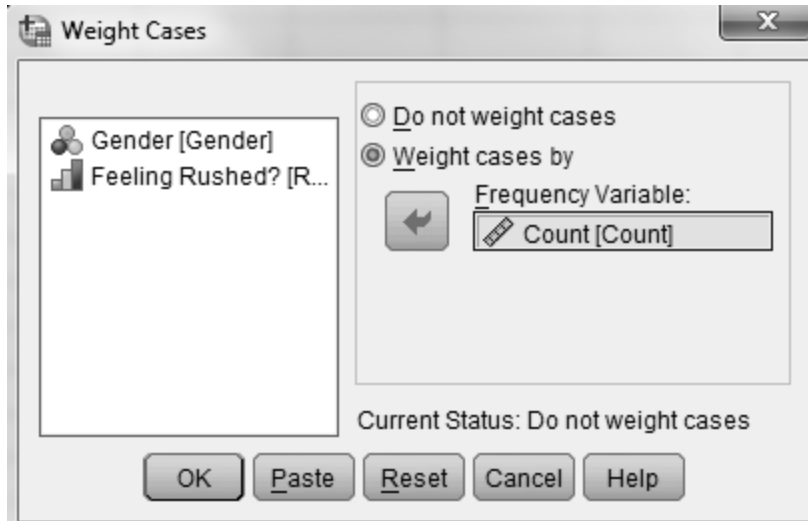# Topic 6

Two-Way Tables

In-Class Activities

Activity 6-1: Feeling Rushed? **6-1**, 6-30, 6-31,16-28,17-11,24-26

b. Use SPSS to produce a bar graph displaying the proportion of respondents in each of the three categories. See Activity 2-1 for instructions on opening a dataset, creating new variables, and creating a bar graph.
i. Use SPSS to create a two-way table and a segmented bar graph of gender and feelings about feeling rushed.
1. Open an empty Data editor by selecting **File > New > Data**. Click the **Variable View** tab. Create the three variables, *Gender, Rushed,* and *Count*. *SPSS Tip:* Make sure that *Gender* has a data type of nominal, *Rushed* has a data type of ordinal, and *Count* has a data type of scale.
2. Assign values to *Gender* and *Rushed*. For *Gender,* Assign a value of 1 with the label "Male" and a value of 2 with the label "Female." For *Rushed,* assign a value of 1 with the label "Always," a value of 2 with the label "Sometimes," and a value of 3 with the label of "Almost Never." The variable *Count* will contain the number observation in the cells of the table. The resulting Data view should look like the following
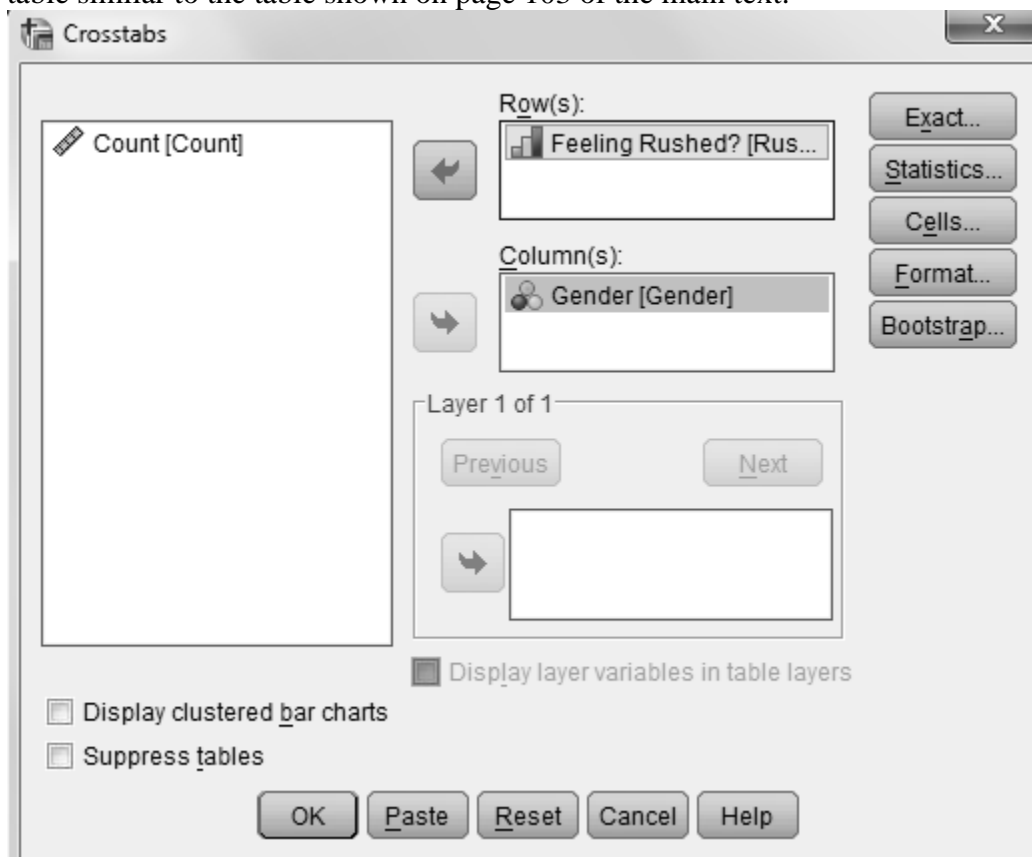
	Gender	Rushed	Count
1	1	1	116
2	1	2	229
3	1	3	82
4	2	1	188
5	2	2	284
6	2	3	78

SPSS Tip: Note that the cells are numbered starting from the upper left corner of the table. The Data view will have six cases, one for each cell in the table. A male who is always feeling rushed would be in the cell in Row 1 and Column 1, a woman who is sometimes feeling rushed would be in Row 2 and Column 2, and so on.
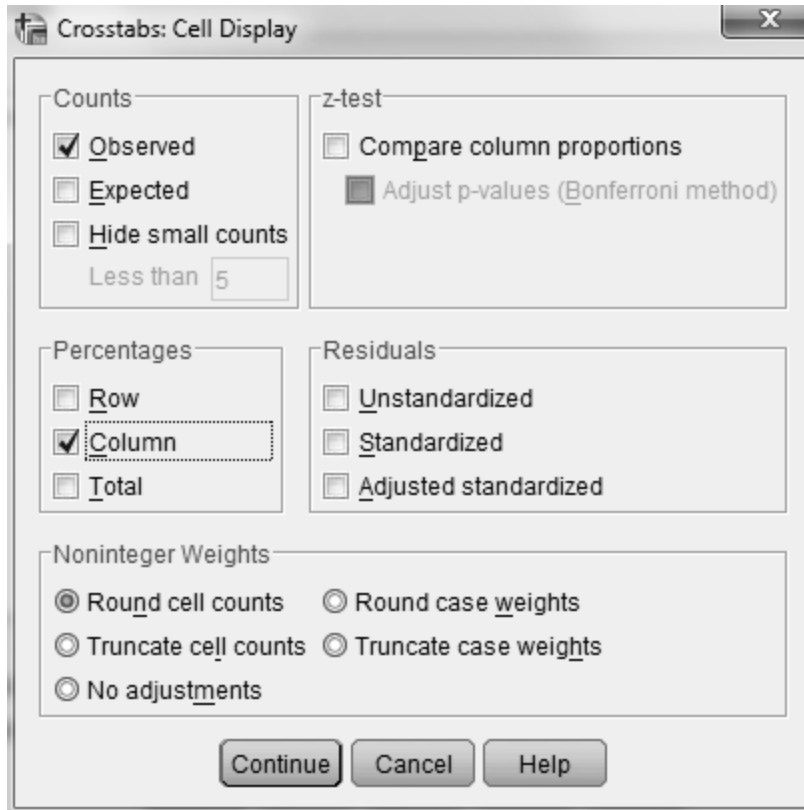3. Select **Data > Weight Cases** to bring up the dialog box shown below. Click **Weight cases by** and select *Count* in the *Frequency Variable* box, and then click **OK**.

4. Select **Analyze > Descriptive Statistics > Crosstabs** to bring up the dialog box shown below. Enter the row variable, *Rushed,* in the *Row(s)* box and the column variable, *Gender,* in the *Column(s)* box. The dialog box is now set up to create a table similar to the table shown on page 103 of the main text.
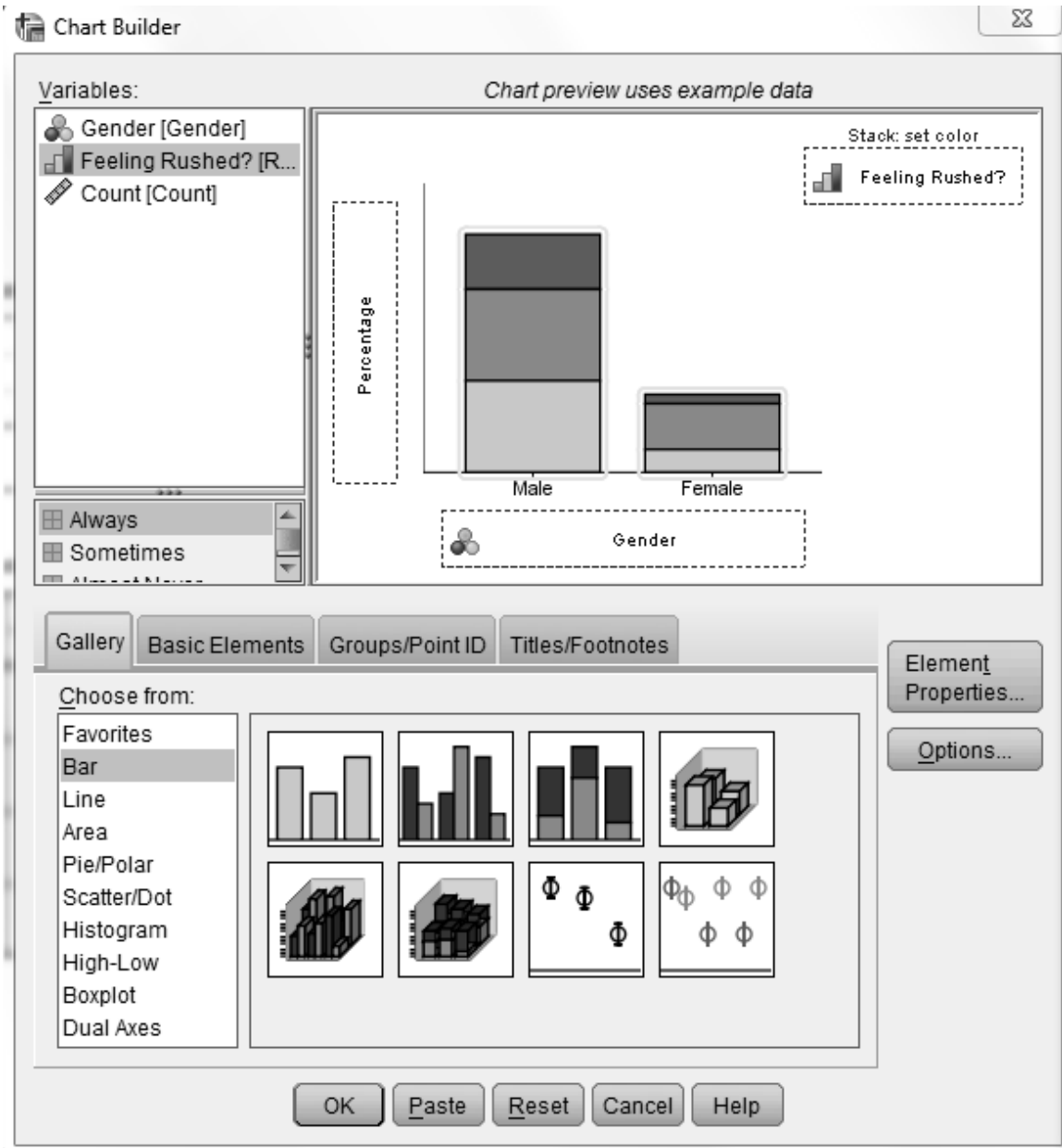
5. Click **Cells**, check *Column* in the *Percentages* area, and click **Continue** to obtain the conditional distribution.

Crosstabs: Cell Display

Counts
- ☑ Observed
- ☐ Expected
- ☐ Hide small counts
  - Less than 5

z-test
- ☐ Compare column proportions
  - ☐ Adjust p-values (Bonferroni method)

Percentages
- ☐ Row
- ☑ Column
- ☐ Total

Residuals
- ☐ Unstandardized
- ☐ Standardized
- ☐ Adjusted standardized

Noninteger Weights
- ◉ Round cell counts   ○ Round case weights
- ○ Truncate cell counts   ○ Truncate case weights
- ○ No adjustments

[Continue] [Cancel] [Help]

6. Click **OK**. The resulting two-way table will appear in an Output window.
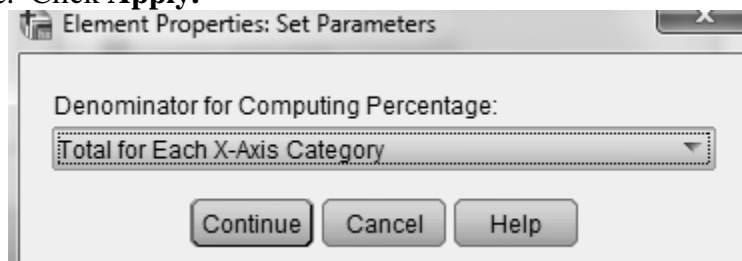
To use SPSS to draw a segmented bar graph:

1. Select **Graphs > Chart Builder** to bring up the dialog box shown below.
2. Select *Bar* from the *Gallery.*
3. Drag the segmented bar chart picture to the *Chart preview uses example data* box. (This is the third picture from the left in the top row. You should see the label *Stacked bar* if you hover your cursor over it.)
4. Drag the column variable, *Gender,* to the *X-Axis?* box.
5. Drag the row variable, *Rushed,* to the *Stack: set color* box.

6. Click **Element Properties** to open the *Element Properties* dialog box if it is not already open.
7. Select *Percentage(?)* from the *Statistic* drop down menu.

8. Click **Set Parameters** and select *Total for Each X-Axis Category*. Click **Continue**. Click **Apply.**

9. Click **OK** to draw the segmented bar chart in an Output window.

Exercises

You will use SPSS whenever possible to create two-way tables, segmented bar graphs, bar graphs, and dotplots.

Exercise 6-6: Lifetime Achievements
6-3, **6-6**, 24-18

Use SPSS to determine the conditional distribution in part a and to construct the requested segmented bar graph in part b of this exercise. Refer to instructions given in Activity 6-1.

Exercise 6-7: Hella Project
6-7, 21-8

Use SPSS to create the two-way table requested in part c, the conditional distribution requested in part d, and create the segmented bar graph requested in part e. Refer to instructions given in Activity 6-1.

Exercise 6-8: Suitability for Politics
6-8. 6-9, 21-15, 25-19, 25-20

Use SPSS to create a cross-tabulation of the relationship between the two variables (follow the steps indicated in Activity 6-1), calculate the conditional distributions, and construct the appropriate displays in the exercise.

Exercise 6-9: Suitability for Politics
6-8, **6-9**, 21-15, 25-19, 25-20

Use SPSS to analyze the data and construct the appropriate graphical displays in this exercise. Refer to Activity 6-1 for instructions on doing this.

Exercise 6-10: A Nurse Accused
1-5, 3-20, **6-10**, 25-23

Use SPSS to create the two-way table requested in part a and calculate the conditional distribution requested in part b. Refer to Activity 6-1 for instructions on doing this.

Exercise 6-11: Children's Television Advertisements
6-11, 25-14

Use SPSS to calculate the conditional distributions requested in part d and to create the segmented bar graph requested in part e.

Exercise 6-13: Weighty Feelings
6-13, 25-25, 26-13

Use SPSS to create the bar graph requested in part b and calculate the conditional distributions and create the segmented bar graph requested in part c.

Exercise 6-14: Preventing Breast Cancer
6-14, 6-15, 21-18, 25-22

Use SPSS to calculate the conditional distributions requested in part d and to create the segmented bar graph requested in part e.

Exercise 6-15: Preventing Breast Cancer
6-14, **6-15**, 21-18, 25-22

Use SPSS to create the two-way table requested in part a, calculate the conditional distributions requested in part b, and create the segmented bar graph requested in part c.

Exercise 6-16: Flu Vaccine
6-16, 21-14

Use SPSS to create the two-way table requested in part b and the numerical and graphical summaries requested in part c.

Exercise 6-17: Watching Films
6-17, 6-18

Use SPSS to create the two-way table requested in part a and the numerical and graphical summaries requested in part b.

Exercise 6-19: Botox for Back Pain
5-4, **6-19**, 6-20, 21-7

Use SPSS to create the two-way table requested in part b and to calculate the conditional proportions requested in part c.

Exercise 6-21: Gender-Stereotypical Toy Advertising
6-21, 6-22

Use SPSS to create the conditional distributions requested in parts b and c and the segmented bar graph requested in part d.

Exercise 6-23: Baldness and Heart Disease
6-23, 25-9

Use SPSS to create the segmented bar graph requested in part d.

Exercise 6-24: Gender and Lung Cancer

Use SPSS to create the two-way table requested in part b, calculate the conditional distributions requested in part c, and create the segmented bar graph requested in part d.

Exercise 6-28: Hypothetical Employee Retention Predictions

Use SPSS to create the segmented bar graph requested in part f.

Exercise 6-33: Government Spending
6-32, **6-33**, 6-34, 25-24, 25-32

Use SPSS to determine the marginal distribution requested in part a, create the bar graph requested in part b, and create the segmented bar graph requested in part c.

Exercise 6-34: Government Spending
6-32, 6-33, **6-34**, 25-24, 25-32

Use SPSS to create the segmented bar graph requested in part a.

Exercise 6-36: Coffee Consumption
2-21, **6-36**, 6-37, 25-29

Use SPSS to determine the marginal distribution requested in part a, create the bar graph requested in part b, determine the conditional distributions requested in part c, and create the segmented bar graph requested in part d.

Exercise 6-37: Coffee Consumption
2-21, 6-37, **6-37**, 25-29

Use SPSS to create the segmented bar graph requested in part b.

Exercise 6-38: Chemotherapy vs. Enzyme Therapy
5-32, **6-38**

Use SPSS to create the two-way table requested in part a, calculate the conditional distributions requested in part b, and create the segmented bar graph requested in part c.

Exercise 6-39: In the News

Use SPSS to create the two-way table requested and to create the graphical and numerical summaries requested.

Topic 7

Displaying and Describing Distributions

In-Class Activities

Activity 7-3: Famous Poets' Poems
7-3, 8-32, 9-20, 10-26

Complete all parts of this activity as directed in your main textbook. Then follow the directions below to use SPSS to create a stemplot.

1. Open and empty Data editor to the Data View tab and store the ages. See Activity 2-1 for instructions on creating a new dataset. Label the variable *Age*. The data type of this variable should be scale.
2. Select **Analyze > Descriptive Statistics > Explore** to bring up the following dialog box.

3. Enter the variable *Age* in the **Dependent List** box.
4. Click either **Both** or **Plots** in the **Display** area.
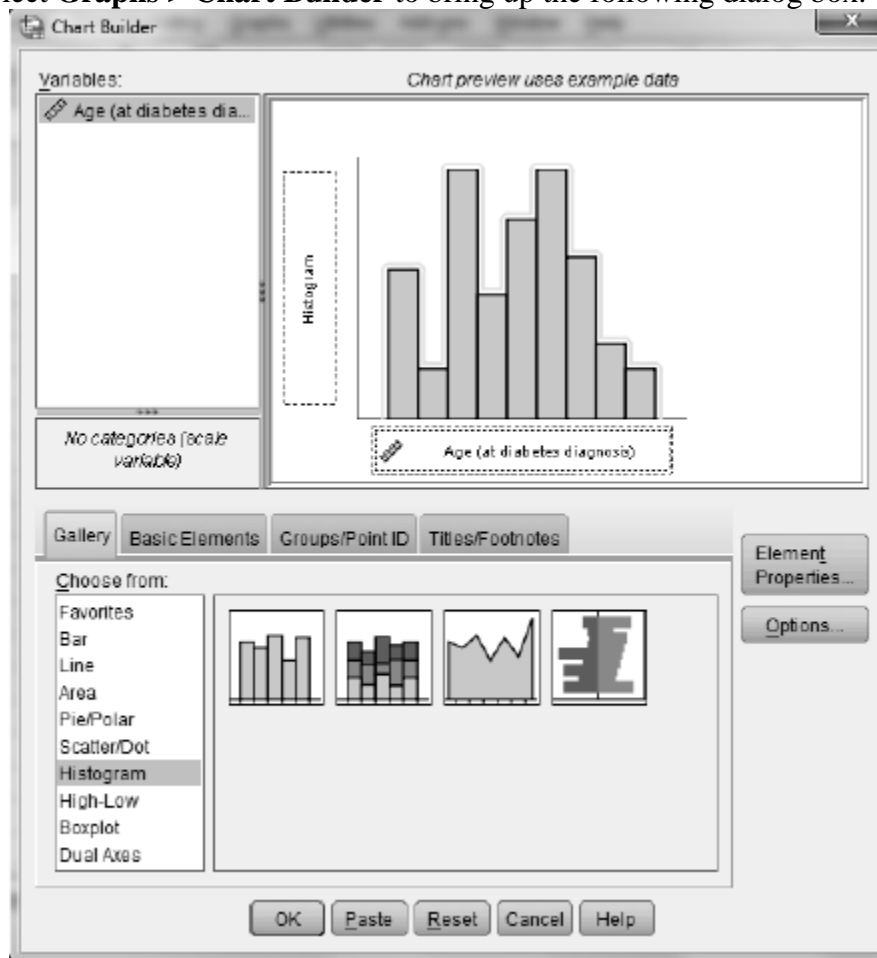5. Click **OK**.

The stemplot is drawn as part of the output in an Output window. Compare SPSS's stemplot and the one in your main textbook on page 132 with the one you drew by hand. Notice that the SPSS output also includes descriptive statistics and other plots.

Activity 7-5: Diabetes Diagnoses
7-5, 10-15

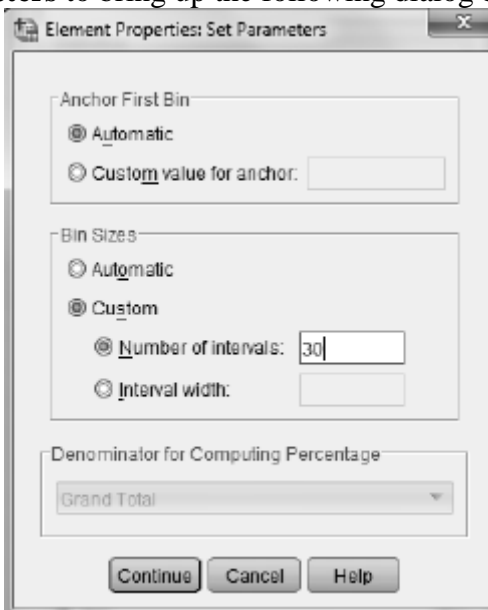You will use SPSS in part f of this activity. Complete all other parts as directed in your main textbook.

f. Follow the directions below to re-create the histogram in SPSS
   1. Open the file `Diabetes.SAV`.
   2. Select **Graphs > Chart Builder** to bring up the following dialog box.

   3. Select **Histogram** in the **Gallery** and drag the histogram to the **Chart preview** box (the leftmost picture in the menu).
   4. Drag the variable to be plotted (*Age*) to the horizontal axis as shown above.
   5. Click **OK**. The histogram will appear in an Output window.
   6. To change the number of subintervals, click **Element Properties**. This dialog box is shown below.

7. Click **Set Parameters** to bring up the following dialog box.

8. In the **BinSizes** area, click **Custom** and enter the desired number of subintervals in the **Number of intervals** box.

9. Click **Continue, Apple, and OK** to draw the histogram.
10. Use SPSS to draw histograms with 10, 5, and 30 subintervals. Comment on the different appearances of the distribution (of the same age data) revealed by these histograms. Which histogram do you think provided the most informative display? Explain.

Exercises

Use SPSS whenever possible to create graphs.

Exercise 7-11: College Football Scores

You will use SPSS in part a of this exercise. Complete the rest of the activity as directed in your main textbook. Refer to SPSS instructions given in Activity 7-3.

Exercise 7-14: Go Take a Hike!
7-6, 7-14

Recall Activity 7-6, which detailed data on day hikes in San Luis Obispo County. Some of the other quantitative variables reported on those hikes include elevation gain and expected time to complete the hike. Use SPSS (the data are stored in the file DayHikes.SAV) to produce graphical displays of these variables. Instructions for creating stemplots and histograms and be found in Activities 7-3 and 7-5. Comment on key features of the distribution of each variable.

Exercise 7-16: Placement Exam Scores
7-16, 9-4

You will uses SPSS in part a of this exercise. The raw scores for this exercise are stored in the file MathPlacement.SAV. Complete part b as directed in your main textbook.

Exercise 7-18: Hitchcock Films

You will use SPSS on part a of this exercise. Complete the rest of the exercise as directed in your main textbook.
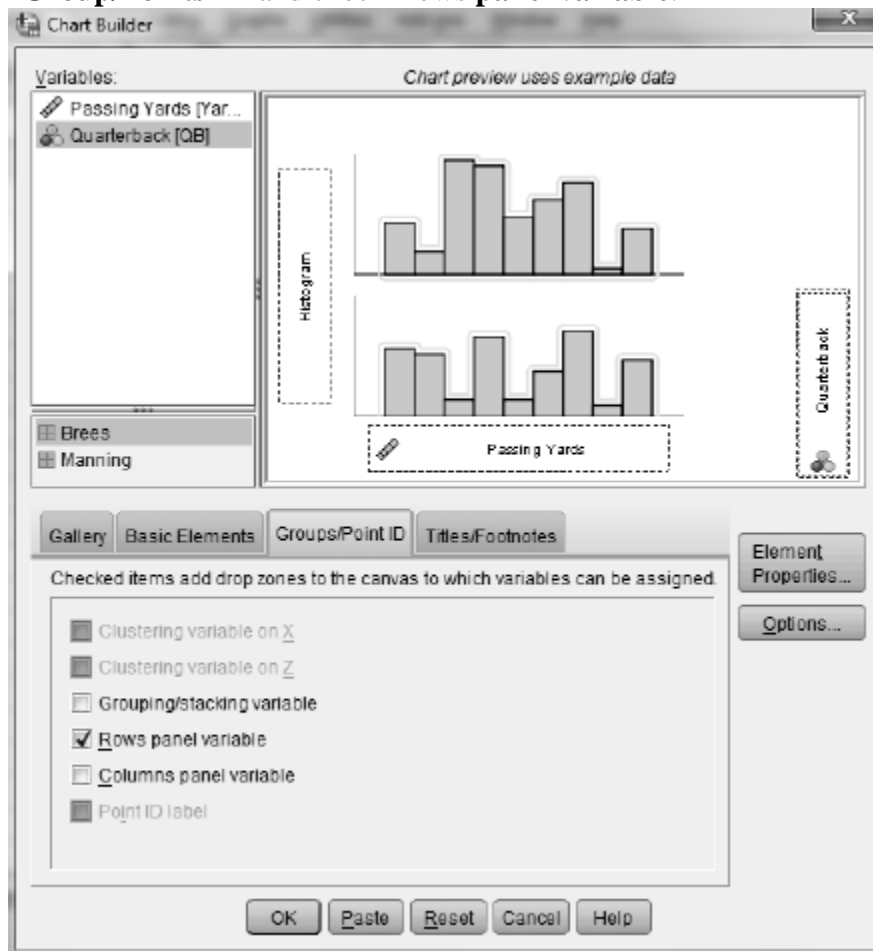
Exercise 7-23: Blood Pressures

The file BloodPressures.SAV contains data from the NHANES study on systolic and diastolic blood pressure measurements. Use SPSS to produce graphical displays of these distributions, and describe their key features. See Activity 7-3 for instructions for creating a stemplot and Activity 7-5 for instructions for creating a histogram.

Use SPSS to create histograms on the same scale as requested in part a. Complete all other parts as directed in the main textbook. To create histograms on the same scale follow the directions given below.

1. Open the file `SuperQBs.SAV`.
2. Select **Graphs > Chart Builder** and set up to draw a histogram of the variable *Yards.*
3. Click **Group/Points ID** and check Rows **panel variable**.

4. Drag *Quarterback* to the **Panel?** box. The resulting dialog box should look like the one shown above.
5. Click **OK** to draw the histogram.
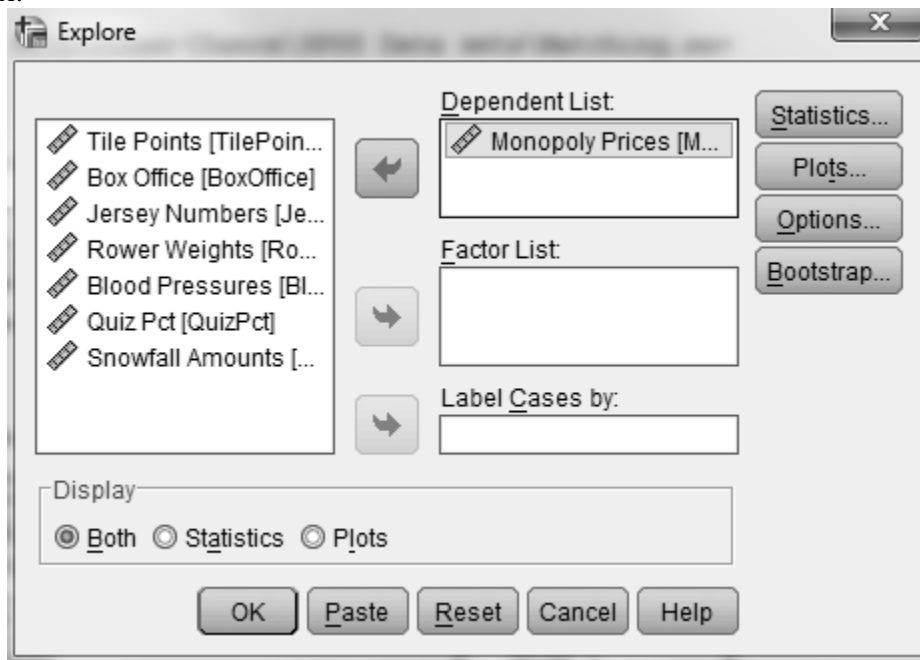
Topic 8

Measures of Center

In-Class Activities

Activity 8-2: Matching Game
7-1, **8-2**

You will use SPSS in parts a and b of this activity. Complete part c as directed in the main textbook.

a. Use SPSS to re-create the dotplots for only these three variables (the data are stored in `Matching.SAV`).   Refer to SPSS instructions for creating dotplots in Activity 2-3.
b. Use SPSS to calculate the mean and median for these variables.
   1. Select **Analyze > Descriptive Statistics > Explore** to open the following dialog box.

   2. Move *Monopoly Prices* to the **Dependent List** box by highlighting it and clicking the right-pointing arrow. Click **OK** and the mean and median along with other statistics will appear in a table in an Output window.
   3. Move the other two variables, one at a time, moving the variable in the **Dependent List** box out. Then click **OK** for each to create a table containing the mean and median. This is necessary because the variables are of different
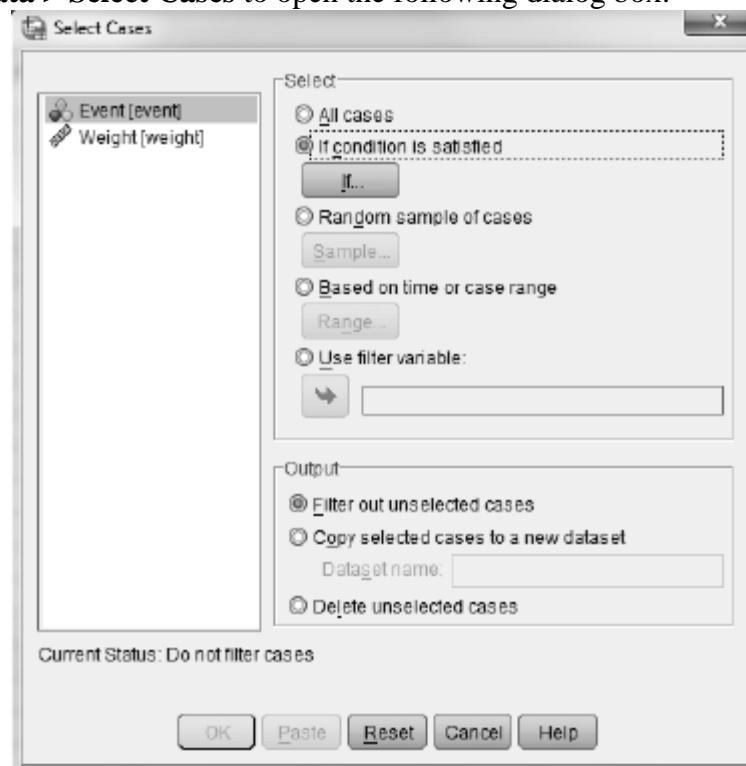
lengths. If you put all three in the **Dependent List** box at once, SPSS will only use the number of values in the shortest variable.
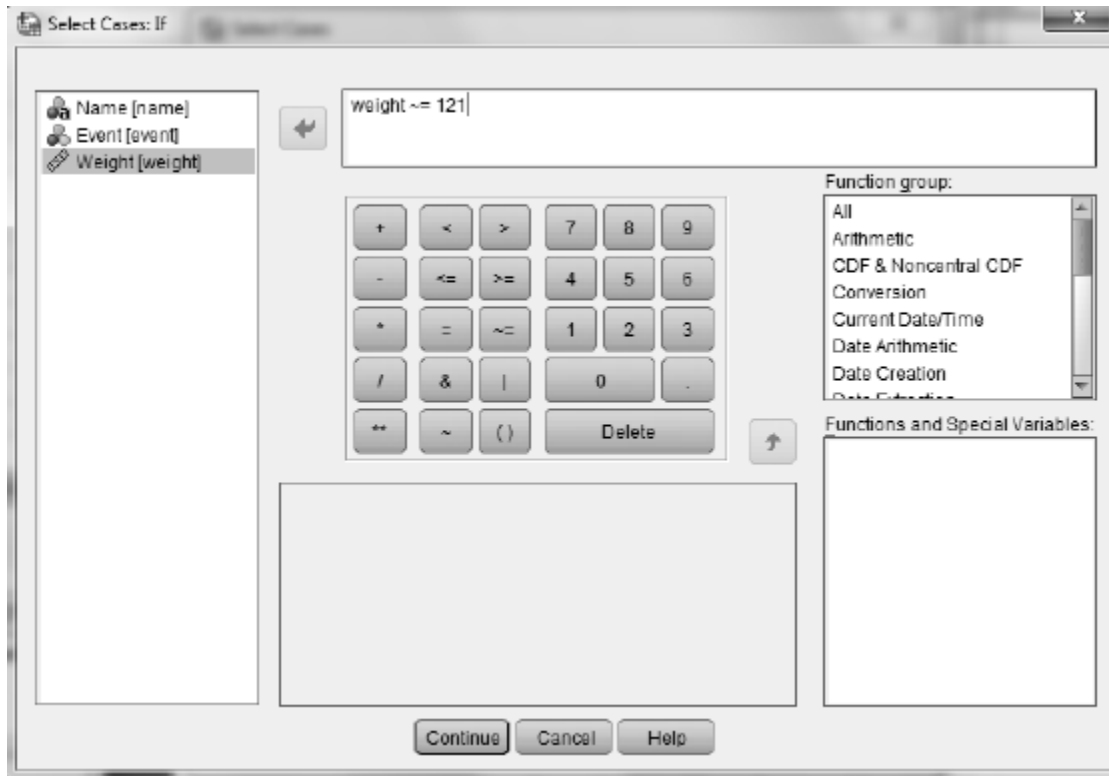
You will use SPSS in parts a, b, d, f, and h of this activity. Record your answers in your main textbook and complete all other parts as directed there.

a. Use SPSS to re-create the dotplot of rowers' weights. (The data are stored in the file `Rowers08.SAV.`) Instructions for creating dotplots are available in Activity 2-3. Based on this dotplot, estimate the values of the mean and median of the rower's weights.
b. Use SPSS to calculate the mean and median as described in Activity 8-2. Recored the mean and median in the first column of the table on page 155 of your main textbook. Why does it make sense that the mean and median differ in this direction? *Hint:* Think about more than the shape of the distribution.
d. Remove the coxswain McElhenney and recalculate the mean and median as described in Activity 8-2. Remove the coxswain as follows.
    1. Select **Data > Select Cases** to open the following dialog box.

    2. Click **If condition is satisfied**, then click **If** to bring up the following dialog box. Enter the condition *weight ~= 121* as shown in the dialog box. You need to use the ~= button (which stands for "not equal to") in the dialog box. It is also good practice the move *weight* by highlighting is and clicking the right pointing arrow. This eliminates the possibility of typographical errors.

3. Click **Continue** followed by **OK**. *SPSS tip:* When selecting cases, always make sure that the unselected cases are filtered, not deleted. Anew variable, *filter_$,* will appear in the Data view. It will contain a 1 for each case that will be used in subsequent analyses and a 0 for each case that will not be used.

Record these new values in the second column of the table on page 155 of your main textbook.
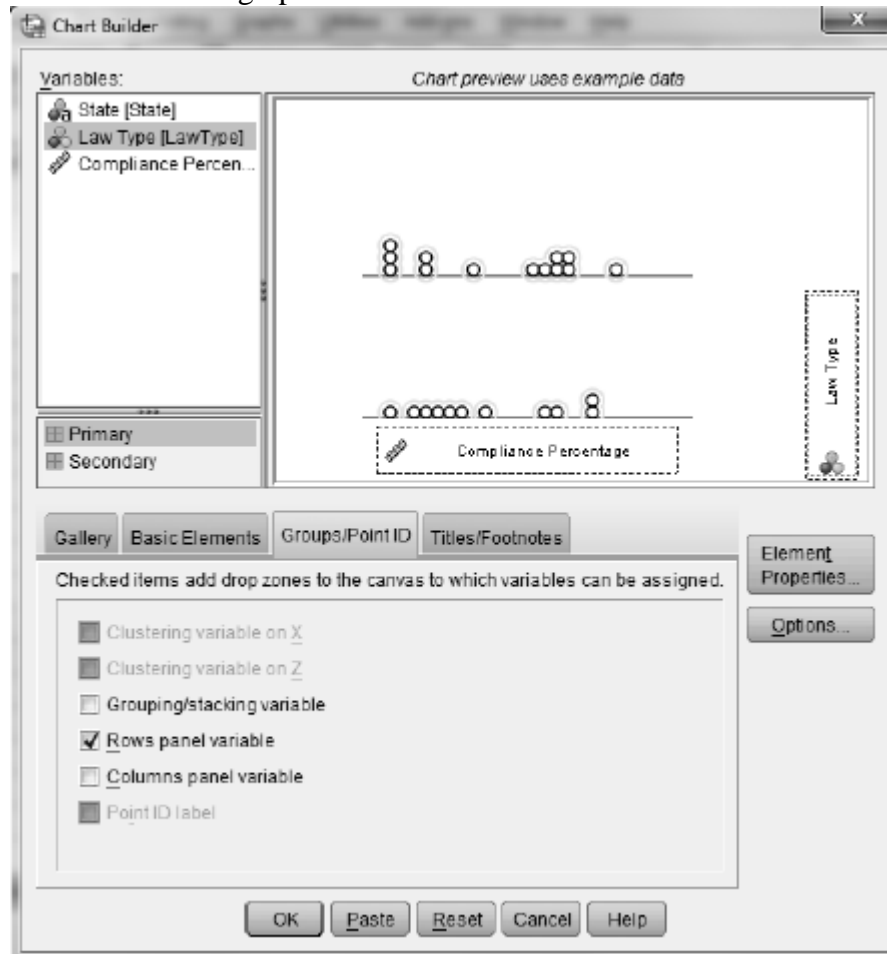
f. Increase Newlin's weight from 229 to 329 by simply typing the new value in the Data vies. Use SPSS to recalculate the mean and median as described in Activity 8-2 and record them in the third column of the table. Did they chance (or not) as you predicted?

h. Increase Newlin's weight to 2229 and use SPSS to recalculate the mean and median as described in Activity 8-2. Record them in the last column of the table.

Activity 8-4: Bucke Up!
2-4, 3-21, **8-4**
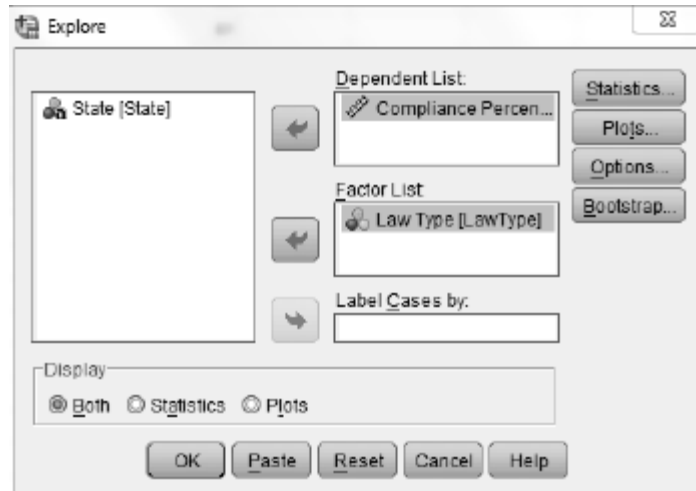
You will use SPSS in parts a and b of this activity. Record your answers in your main textbook and complete part c as directed there.

a. Use SPSS to re-create dotplots for comparing the distributions of seatbelt compliance percentages between the two types of states. (The data are stored in the file `SeatBeltUsage07.SAV.`) To do this you will use SPSS's Chart Builder.

1. Select **Graphs > Chart Builder** to open the following dialog box.
2. From the **Gallery** select **Scatter/Dot** and drag the dotplot picture to the **Chart preview** box.
3. Drag the variable to be plotted (*Compliance Percentage*) to the horizontal axis.
4. To get stacked dotplots with the same scales click the **Groups/Point ID** tab.
5. Check **Rows panel variable**. A *Panel?* box will appear. Drag the group-identifying variable (*Law Type*) to this box.
6. Click **OK** to draw the graphs.

b. Use SPSS to determine the mean and median of the compliance percentages in the two type of states.
   1. Select **Analyze > Descriptive Statistics > Explore** to open the following dialog box.
   2. Move *Compliance Percentage* to the **Dependent List** box.
   3. To get separate analysis for each type of state more *Law Type* to the **Factor List** box.
   4. Click **OK** and a separate table for each type of state will appear in an Output window.

How do the two types of states compare on compliance percentage according to these two measures? By how many percentage points is the compliance percentage greater (on average) in states with strict enforcement laws?

Exercises

Use SPSS whenever possible to create graphs or calculate means and medians.

Exercise 8-12: Planetary Measurements
8-12, 10-20, 19-20, 27-13, 28-20, 28-21

Use SPSS to calculate the requested means in part g. The data are stored in the file `Planets.SAV`.

Exercise 8-14: Sporting Examples
2-6, 3-36, **8-14**, 10-11, 22-26

 Use SPSS to calculate the means and medians requested in part a. The data are stored in the file `SportsExamples.SAV`.

Exercise 8-17: Marriage Ages
8-17, 9-6, 16-19, 16-24, 17-22, 23-1, 23-12, 23-13, 26-4, 29-17, 29-18

Use SPSS to calculate the medians requested in part c and d. The data are stored in the file `MarriageAges.SAV`.
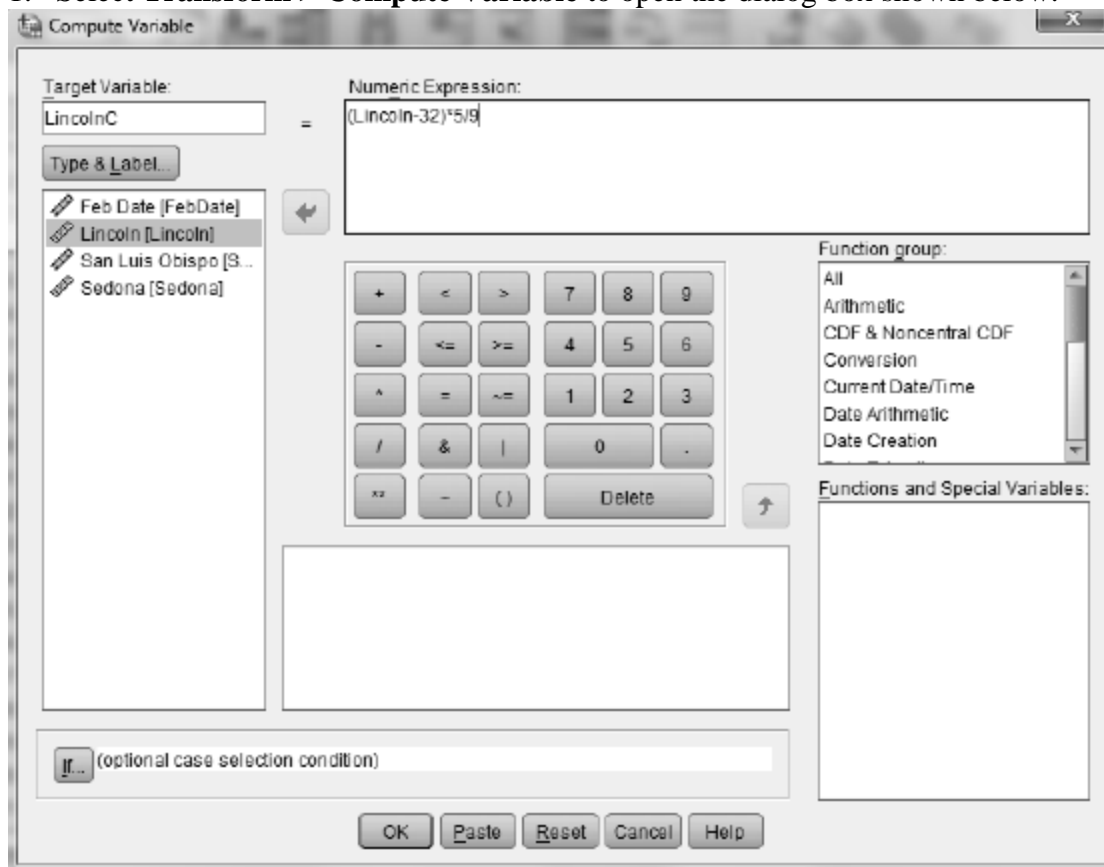
Exercise 8-18: Quiz Percentages

Use SPSS to calculate the mean and median requested in part b and d. The data are stored in the file `QuizPct.SAV`.

Exercise 8-19: February Temperatures
2-5, **8-19**, 9-7

Use SPSS to calculate the mean and median requested in part a and c. Use SPSS to convert the temperatures to Celsius in part b as follows.

1. Select **Transform > Compute Variable** to open the dialog box shown below.

2. For Lincoln, Nebraska, enter *LincolnC* in the **Target Variable** box. Enter *(Lincoln-32)*5/9* in the **Numeric Expression** box.
3. Click **OK** to create the new variable containing Lincoln temperatures in Celsius.
4. Repeat 1-3 for the other cities. Give each target variables names of *SanLuisObispoC* and *SedonaC*.

Use SPSS to calculate the median requested in part a and the median requested in part c.
The data are stored in the file `DealAmounts.SAV`.

Topic 9

Measures of Spread

In-Class Activities

Activity 9-2: Developing Life Expectancies
9-1, **9-2**

h. Use **Analyze > Descriptive Statistics > Explore** to calculate the standard deviations for Botswana and Papua New Guinea. Refer to Activity 8-2 for instructions. The data are stored in the file `LifeExpectancy2008.SAV`. Note that the table that contains the mean and median also contains the standard deviation, interquartile range, the range. Also note that the quartiles for the interquartile range are computed in SPSS using a method that is different from that given in the main textbook.
l. Change the life expectancy for Botswana to 83.5 and use SPSS to recalculate the standard deviation, range, and interquartile range. Record the values in the middle column of the table on page 177 of your main textbook.
m. Change the life expectancy for Botswana to 635 and use SPSS to recalculate the standard deviation, range, and interquartile range. Record the value in the last column of the table on page 177 of your main textbook.

Activity 9-3: Value of Statistics
2-9, 2-10, **9-3**

c. Use SPSS to calculate the standard deviations for each class. Since the number of cases varies from class to class run **Analyze > Descriptive Statistics > Explore** on each class separately. Record the values in the table on page 179 or you main textbook. The data are stored in the file `ValuesFJ.SAV`. Refer to Activity 8-2 for instructions.

Activity 9-6: Marriage Ages
8-17, **9-6**, 16-19, 16-24, 17-22, 23-1, 23-12, 23-13, 26-4, 29-17, 29-18

a. Use SPSS to calculate the mean and median age for each group. Which spouse tends to be older and by how many years on average?
b. Use SPSS to calculate the IQR and standard deviation for each group. Does one group of spouses have more variability in their ages than the other group?

d. Use SPSS to calculate the mean and median of the difference in ages. Do you notice anything about how these compare to the means and medians of husbands and wives individually?
e. Calculate the IQR and standard deviation for difference in ages. Compare these values with those in part b.

Exercises

Use SPSS whenever possible to create graphs or calculate summary statistics.

Exercise 9-7: February Temperatures
2-5, 8-19, **9-17**

Use SPSS to calculate the standard deviations requested in part b. The data are stored in the file `FebTemps.SAV`.

Exercise 9-8: Social Acquaintances
9-8, 9-9, 10-13, 10-14, 19-9, 19-10, 20-12

Use SPSS to calculate the mean and standard deviation requested in part c. The data are stored in the file `AcquaintancesCP.SAV`.

Exercise 9-9: Social Acquaintances
9-8, **9-9**, 10-13, 10-14, 19-9, 19-10, 20-12

Use SPSS to calculate the mean and standard deviation of your class data and compare the results with the data in `AcquaintancesCP.SAV`.

Exercise 9-13: Supreme Court Justices
8-11, **9-13**

Add two years to the tenure of each justice given in the table in Activity 8-11. Enter the data into SPSS and use SPSS to calculate the new mean and median requested in part b and the IQR and standard deviation requested in part d. Double the tenure of each justice given in the table in Activity 8-11. Enter the data into SPSS and use SPSS to calculate the mean, median, IQR, and standard deviation requested in part f.

Exercise 9-16: Tennis Simulations
7-22, 8-21, **9-16**, 22-18

Use SPSS to calculate the standard deviations requested in part c. The data are stored in the file `TennisSim.SAV`.

Exercise 9-24: Hypothetical ATM Withdrawals
9-24, 19-21, 22-25

Use SPSS to create graphical displays for each machine as requested in part a. The data are stored in the file `HypoATM.SAV.`   Use SPSS to calculate the mean and standard deviation for each machine as requested in part b.

Exercise 9-25: Guessing Standard Deviations

Use SPSS to calculate the requested standard deviations in part b. The data are stored in the file `HypoStdDev.SAV.`

Exercise 9-26: Super Bowl Quarterbacks
7-28, **9-26**, 9-27, 10-27

Use SPSS to create the graphical displays requested in part a. Use SPSS to calculate the IQR and standard deviation requested in part b. Delete the last two games for each quarterback and recalculate the IQR and standard deviation requested in part d. The data are stored in the file `SuperQBs.SAV.`

Exercise 9-27: Super Bowl Quarterbacks
7-28, 9-26, **9-27**, 10-27

Use SPSS to calculate the mean and standard deviation for Manning's games. The data are stored in the file `SuperQBs.SAV.`

Exercise 9-28: Rowers' Weights
7-2, 8-3, **9-28**, 10-21

Use SPSS to calculate the mean and standard deviation of the weights as requested in part a. The data are stored in the file `Rowers08.SAV.`
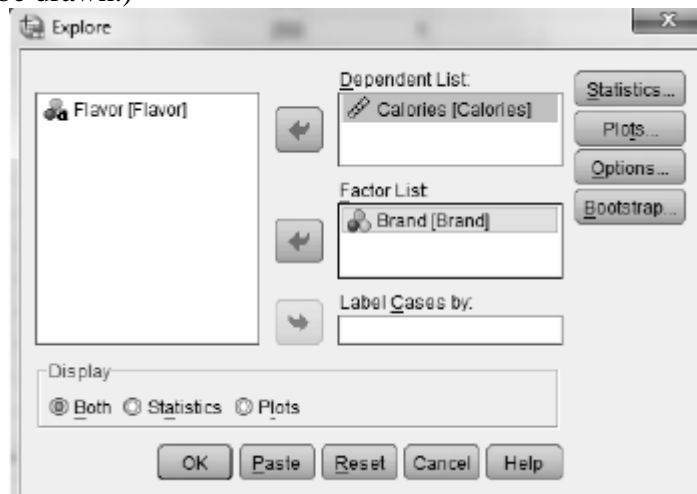
Topic 10

More Summary Measures and Graphs

In-Class Activities

Activity 10-3: Ice Cream Calories

You will use SPSS to learn about and construct modified boxplots in parts a, b and f-h of this activity. Complete all other parts as directed in your main textbook. The data are stored in the file `IceCreamCalories.SAV`. It contains data on calorie amounts per serving for various flavors of Ben & Jerry's, Cold Stone Creamery, and Dryer's ice cream.
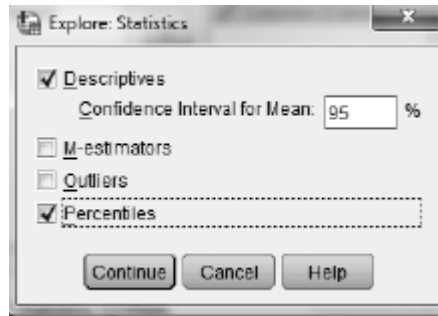
a. Select **Analyze > Descriptive Statistics > Explore** to calculate the five-number summaries of the caloric amounts for all flavors in each brand. Report these values in the table on page 203 of your main textbook. Follow these steps to analyze the ice cream flavors separately by company.
   1. Enter the variable *Brand* in the **Factor List** box and enter the variable *Calories* in the **Dependent List** box.
   2. Click **Both** in the **Display** area. (This will ensure that the boxplots asked for in part b will be drawn.)

   3. Click the **Statistics** button to open the following dialog box.
   4. Click **Percentiles**, click **Continue**, then click **OK**.
       A table of percentiles will appear in an Output window below the table of summary statistics. Percentiles are calculated in two ways. The percentiles referred to as *Tukey's Hinges* calculates quartiles (the 25th and 75th percentiles) using the method you learned in Topic 9. The percentiles referred to as *Weighted Averages* are calculated using a more complicated method not discussed here.
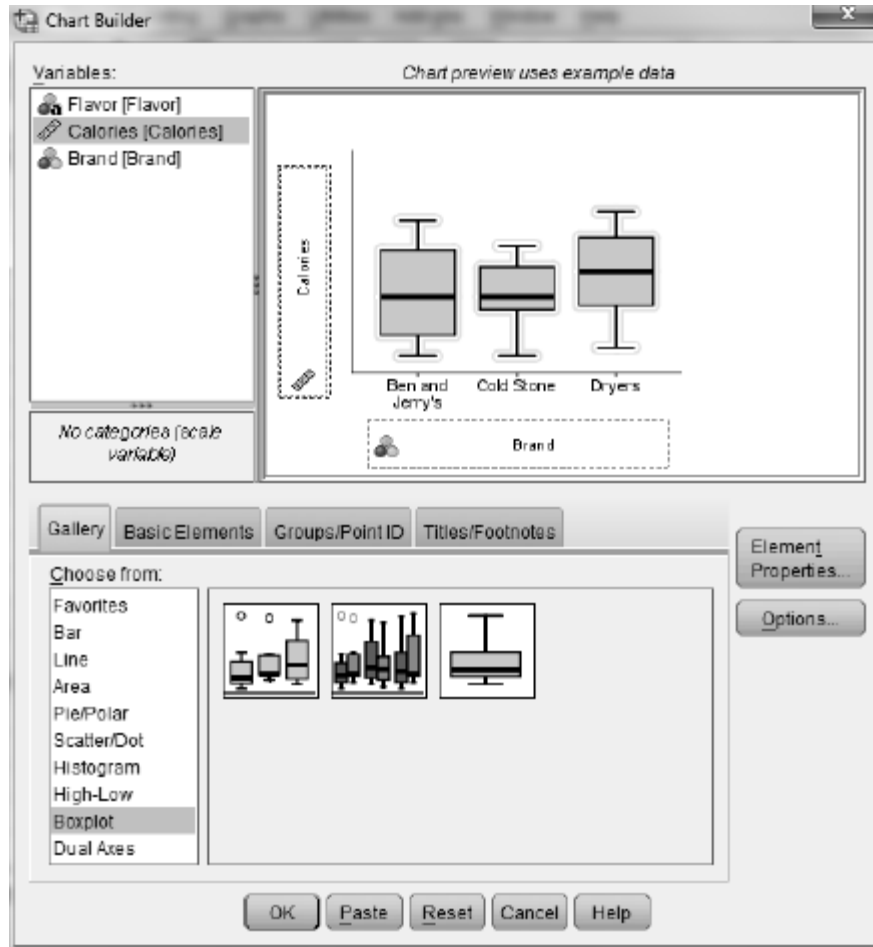
The interquartile range that RPSS reports uses these weighted average quartiles instead of the Tukey's hinges. The percentiles you see in the SPSS output may therefore differ from the ones you calculate by hand, but the differences should not be dramatic.
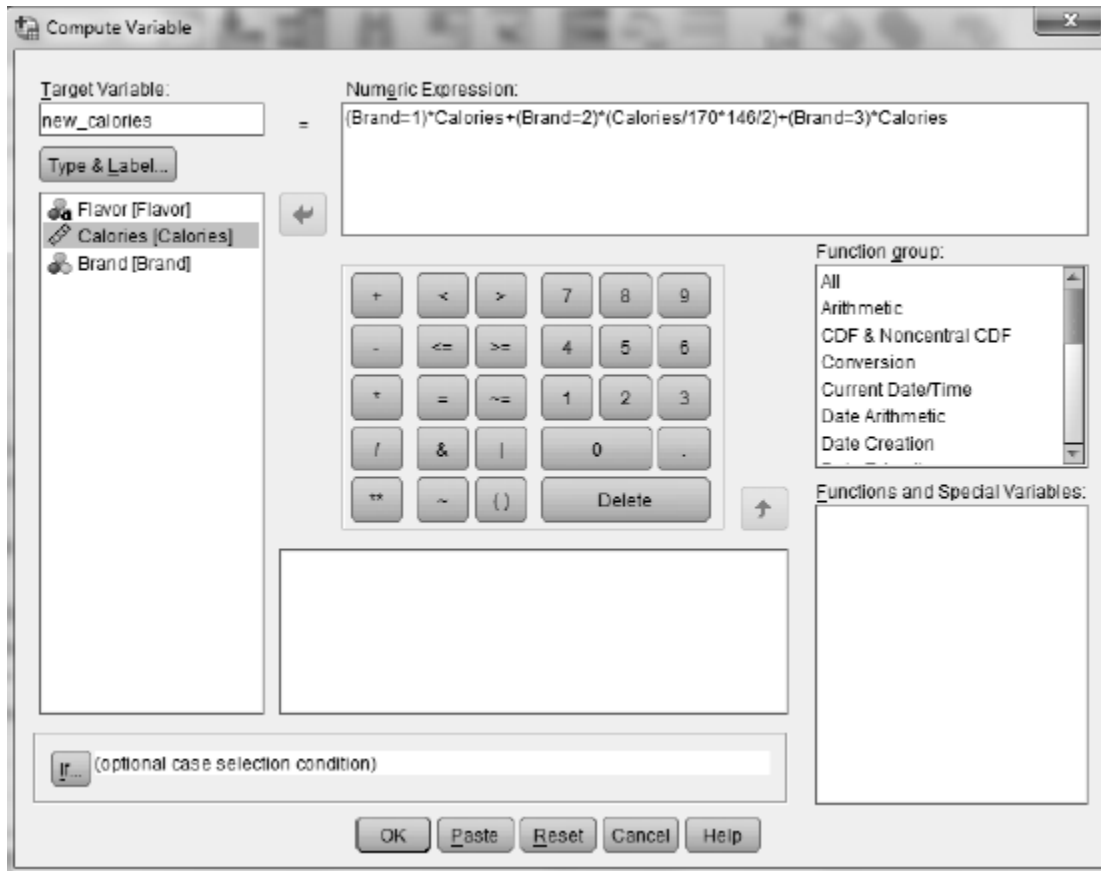

b. The SPSS output from the **Explore** dialog box will contain modified boxplots of the distribution of caloric amounts for the three brands on the same axis. Discuss how the three brands' caloric amounts compare, based on these boxplots. Note that the boxplots produced by SPSS are slightly different from those discussed in this topic. Name, some outliers are denoted by circles and some by asterisks. The ones with circles are between 1.5 and 3 interquartile ranges from the box and are classified as *mild* outliers. Those with astericks are more than 3 interquartile ranges from the box and are classified as *extreme* outliers.

You can use this alternative method for drawing boxplots:
1. Select **Graphs > Chart Builder** to open the following dialog box.
2. From the **Gallery** select **Boxplot** and drag the side-by-side boxplot picture to the **Chart preview** window. This is the leftmost picture in the menu; you will see the label *Simple Boxplot* if you hover your cursor over it.
3. Drag the variable to be graphed (*Calories*) to the **Y-Axis?** box and the group-identifying variable (*Brand*) to the **X-Axis?** box as shown below.
4. Click **OK** to draw the boxplots in an Output window.

f. Select **Transform > Compute Variable** to do this conversion. Let the **Target Variable** be *new_calories*, and let the **Numeric Expression** be as shown here. *Hint:* First, divide Cold Stone's calorie amounts by 170 to convert them to a "per gram" basis, then multiply by 146 to convert the calorie counts to a "per cup" basis, and finally, divide by 2 to convert the caloric counts to a "per half cup" basis. It is also good practice to highlight and move the variables to the **Numeric Expression** box rather than typing them directly. This eliminates the possibility of typographical errors. Finally, click **OK** to create the new variable. *Note:* The *Brand = 1, Brand = 2,* and *Brand = 3* entries evaluate to 1 if the *Brand* is than number and a 0, otherwise. This has the effect of moving Ben & Jerry's and Dryer's to the new variable unchanged and converting only the Cold Stone amounts.
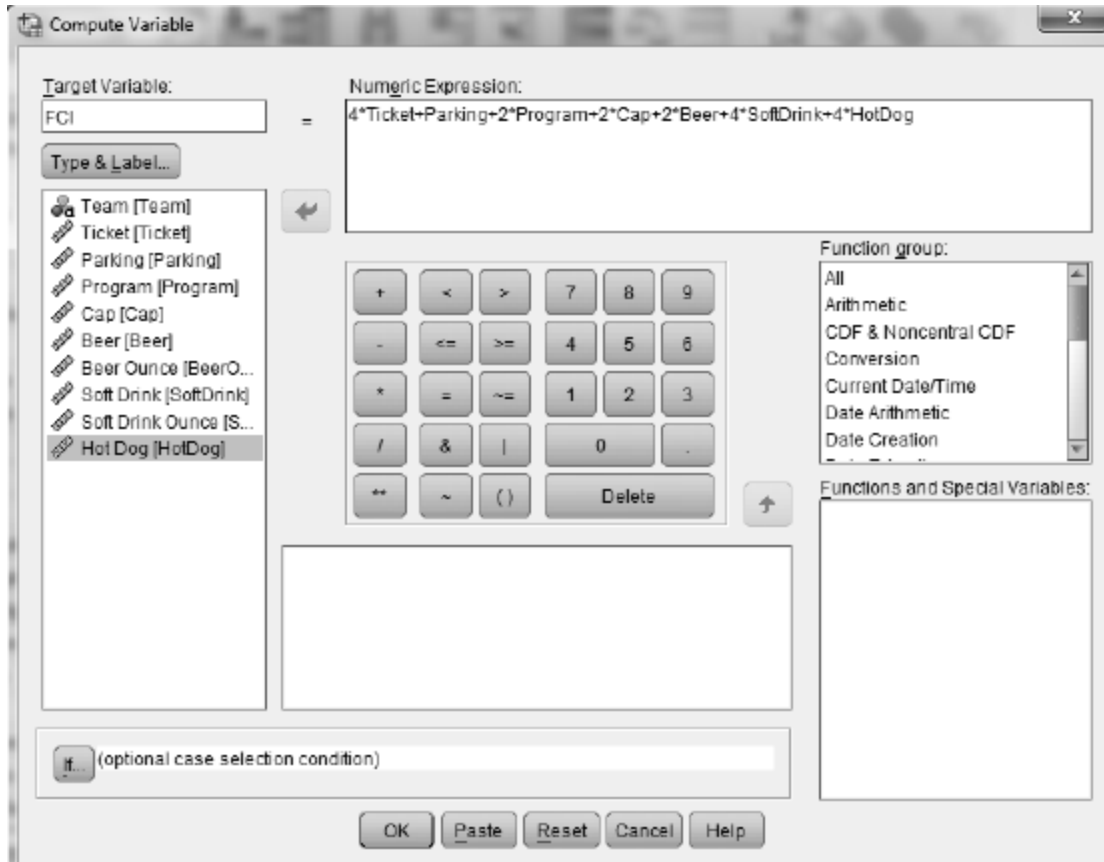
g. Select **Analyze > Descriptive Statistics > Explore** to recalculate the five-number summary for Cold Stone's amounts using the "per half cup" scale. How has the five-number changed?

h. Part of the output from part g is the set of new modified boxplots for the three distributions of caloric amounts. Comment in your main textbook on what the boxplots reveal about how the three ice cream brands compare.

Activity 10-4: Fan Cost Index

You will use SPSS for parts a, b, e, and h. Complete all other parts as directed in your main textbook.

a. The fan cost index (FCI) is defined as the cost of two average (non-premium) tickets, parking, two programs, two baseball team caps, two small beers, four small sodas, and four hot dogs. The data are stored in the file FanCost09.SAV. Compute the fan cost index as follows.
   1. Select **Transform > Compute Variable** to open the dialog box shown below.
   2. To calculate the fan cost index for each team enter *FCI* in the **Target Variable** box. Enter the numeric expression exactly as shown in the **Numeric Expression** box.
   3. Click **OK** to create the new variable. *Hint:* To make sure your formula is entered properly, the FCI for Arizona is $114.24.

b. Use SPSS to create a dotplot of the distribution of the FCI values. Comment on the dotplot's features.

e. Select **Transform > Compute Variable** and create the new variable called **M**e **C**ost **I**ndex (*MCI*) by changing the number of tickets and what is purchased. That is, follow what you did previously, but change the coefficients of the variables *Ticket, Parking*, etc to better reflect what your group would actually spend at a game. Report in your main textbook your new equation and which team has the highest and which has the lowest MCI, along with those values. Also comment on where your team falls in the distribution of MCI values.

h. Select **Transform > Compute Variable** and create a new variable, *price_per_ounce,* for comparing soda prices by converting to a "price pre ounce" variable. Your **Numeric Expression** should be *SoftDrink/SoftDrinkOunce.* Which ballpark has the most expensive soft drink per ounce, and which has the least expensive? What are those values? Record them in your main textbook.

Activity 10-5: Digital Cameras
10-5, 26-18, 27-21

You will use SPSS in parts a and d of this activity. Complete the other parts as directed in your main textbook. The data are stored in the file `DigitalCameras.SAV`. The first few lines are shown in the table on page 207 of your main textbook.

a. Select **Analyze > Descriptive Statistics > Explore** to calculate the five-number summaries and produce comparative boxplots of the camera prices, based on the camera type. *Pricein$* should be in the **Dependent List** box, and *Type* should be in the **Factor List** box. Is there much overlap in prices among the types of cameras? Do some types of cameras tend to cost more than others?

d. Select **Analyze > Descriptive Statistics > Explore** to calculate five-number summaries and to produce comparative boxplots of the camera ratings based on camera type. *RatingScore* should be in the **Dependent List** box, and *Type* should be in the **Factor List** box. Write a paragraph in you main textbook summarizing what these boxplots reveal about the distributions of rating scores across the types of cameras.

Exercises

Use SPSS whenever possible to create graphs or calculate statistics.

Exercise 10-7: Natural Selection
10-1, 10-6, **10-7**, 12-20, 22-21, 23-3

Use SPSS to construct the requested boxplots. The data are stored in the file `Bumpus.SAV.`   Refer to Activity 10-3 for instructions.

Exercise 10-8: Welfare Reform

Complete part a as directed in your main textbook. You will use SPSS in parts b-e of this exercise. The data area stored n the file `WelfareReform.SAV`.

b. Use SPSS to create a new variable, *reduction in number of families receiving welfare assistance.* Choose a short variable name that reflects this measure and label the variable appropriately. *Hint:* A positive value should indicate that the number of families receiving assistance *declined* over this period.

c. Now use SPSS to create another new variable, *percentage reduction.* Identify the three states with the greatest percentage reductions, and the three states with the lowest reductions. *Hint:* To create the percentage reduction variable, select **Transform > Compute Variable**. In the **Numeric Expression** box, divide variable created in part b by the number of families receiving assistance to begin with, and multiply that amount by 100 to convert to a percentage. Then sort the file by percentage reduction. To do this select **Data > Sort Cases** and enter the percentage reduction variable you created in the **Sort by** box.

d. Use SPSS to create a dotplot of the percentage reduction variable, and comment on its key features.

e. The data file also lists the geographic region the stat is in, classified as South, West, Northeast, and Midwest. Use SPSS to create comparative boxplots of the percentage reductions across these four regions. See instructions for creating boxplots in

Activity 10-3.   Comment on what these boxplots reveal about whether some regions of the country tended to see a greater percentage reduction in the number of welfare recipients that did the other regions of the country.

Exercise 10-11: Sporting Examples
2-6, 3-36, 8-14, **10-11**, 22-26

Use SPSS to create the requested boxplots. The data are stored in the file `SportsExamples.SAV.`

Exercise 10-12: Backpack Weights
2-13, **10-12**, 19-6, 20-1, 20-17

Use SPSS to create the boxplots requested in part a. Use SPSS to create the new variable, *ratio of backpack weight to body weight,* and produce the boxplots requested in part b.

Exercise 10-16: Hazardousness of Sports

Use SPSS to create the new variable, *injuries per thousand,* requested in part b. The data are stored in the file `SportsHazards.SAV.`   Do this by dividing the number of injuries by the number of participants and multiplying the result by 1000.

Exercise 10-17: Gender of Physicians

Use SPSS to create the new variable, *percent women,* requested in part b. The data are stored in the file `Physicians.SAV.` Do this by dividing the number of women by the total number of physicians and multiplying that result by 100. Use SPSS to produce the graphical displays requested in part h.

Exercise 10-18: Draft Lottery
10-18, 27-7, 29-9

Use SPSS to calculate the medians requested in part b and create the boxplots requested in part c.

Exercise 10-24: Fantasy Draft Times
10-24, 10-25

Use SPSS to create the boxplots requested in part a. The data are stored in the file `FantasyDraftTimes.SAV.` Use SPSS to calculate the medians requested in part b and c, and to calculate the IQR requested in part d and e.