Organizing Data

Chapter 2

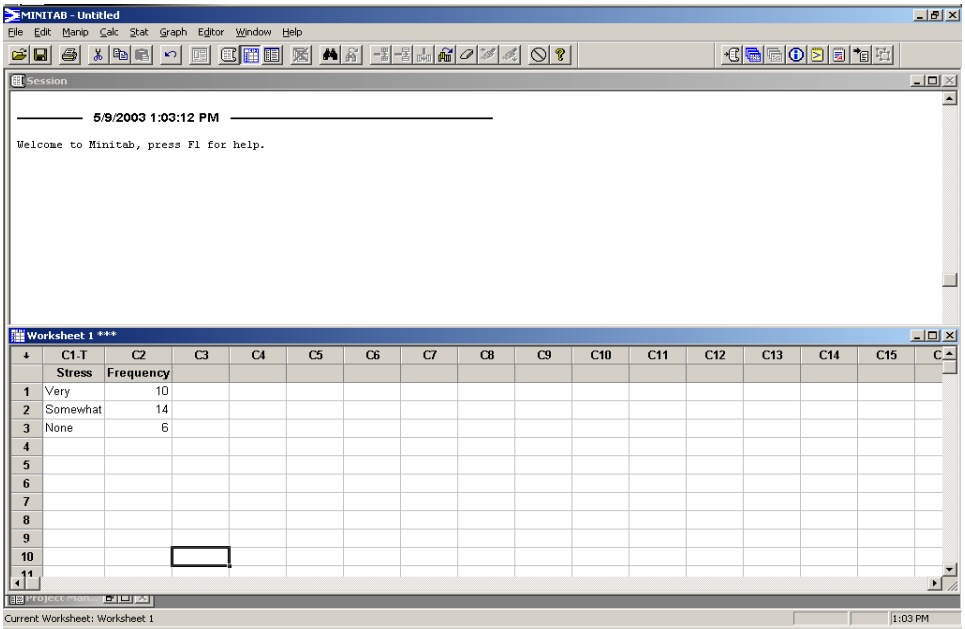
Section 2.2

Example 2-1, pg. 28

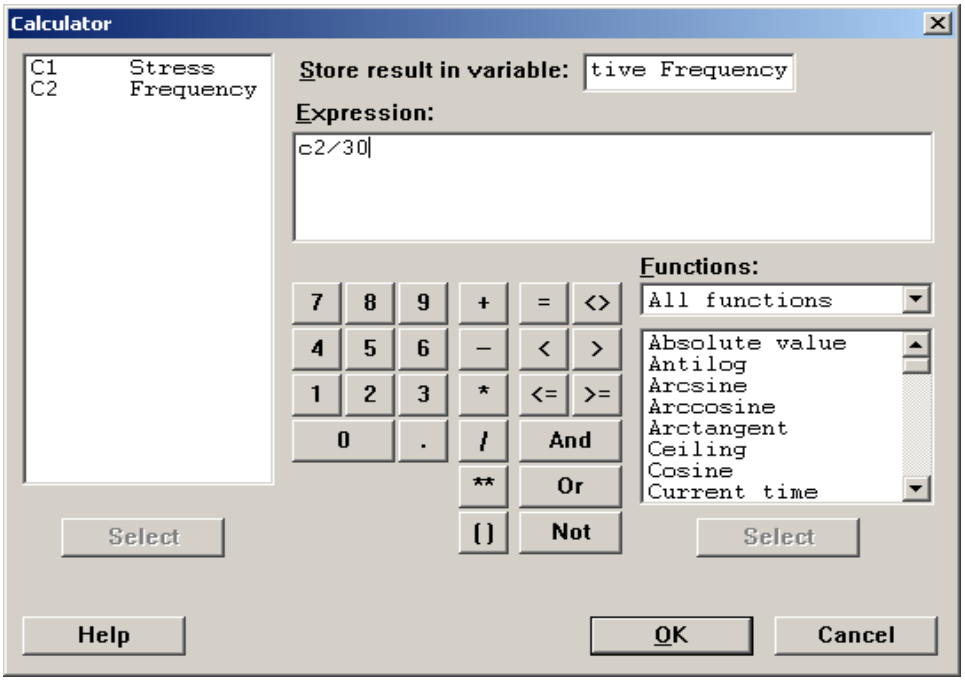
A sample of 30 employees from large companies was selected, and these employees were asked how stressful their jobs were. The data is summarized below, as well as in Table 2.4 on page 29 of the textbook.

Stress on Job	Frequency
Very	10
Somewhat	14
None	6

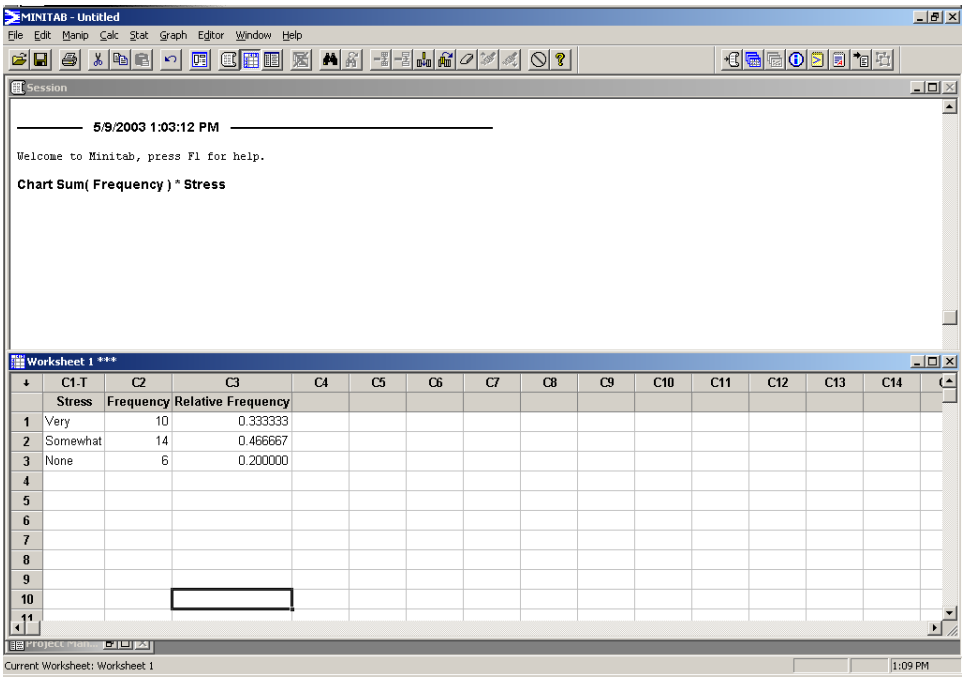
To create a graph of this data, you will need to enter the data into the Data Worksheet. In column 1, type in the three stress levels. Label the column by typing "Stress" in the gray cell in C1. Label column 2 "Frequency", and type in the frequencies.



Next, let Minitab calculate the relative frequencies. Since there are 30 employees in your sample, divide each frequency by 30. To do this, click on **Calc → Calculator.** You are going to create a new column of relative frequencies, so name the column by entering "Relative Frequency" beside **Store result in variable:** as shown below. Next tell Minitab what you want to calculate. Click in the box beneath **Expression**, then type in **C2 / 30**. This tells Minitab to divide each value in C2 by 30.

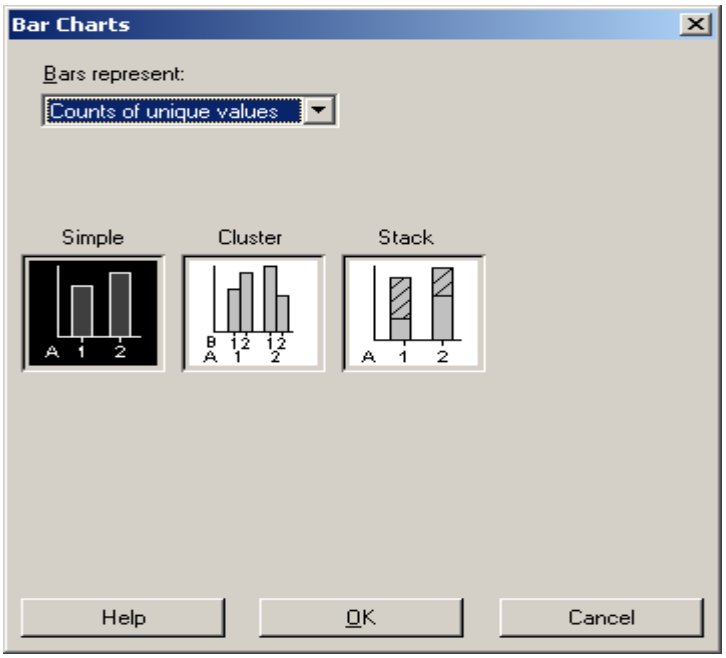


Click on **OK** and the relative frequencies should be in C3.

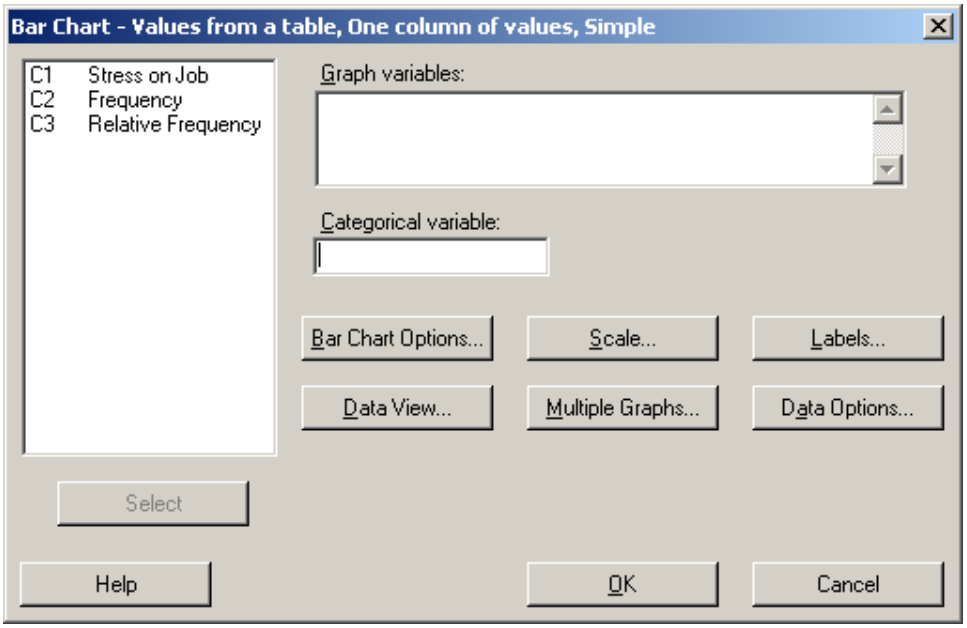


Bar Graphs, pg. 30

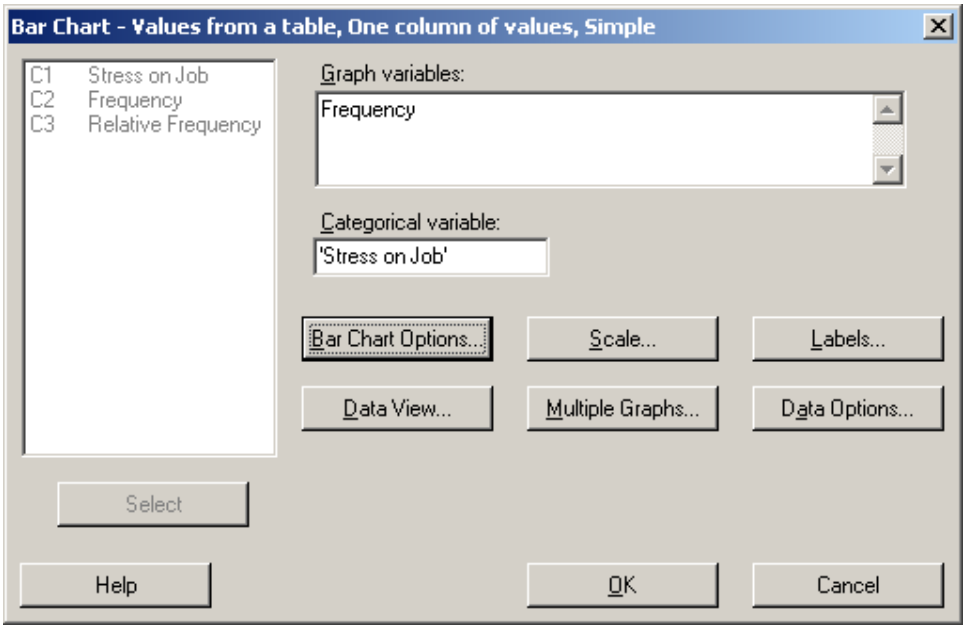
Now that the data is entered, you can proceed to make a bar graph. We will also use this data to make a pie chart. Click on: **Graph → Bar chart**.



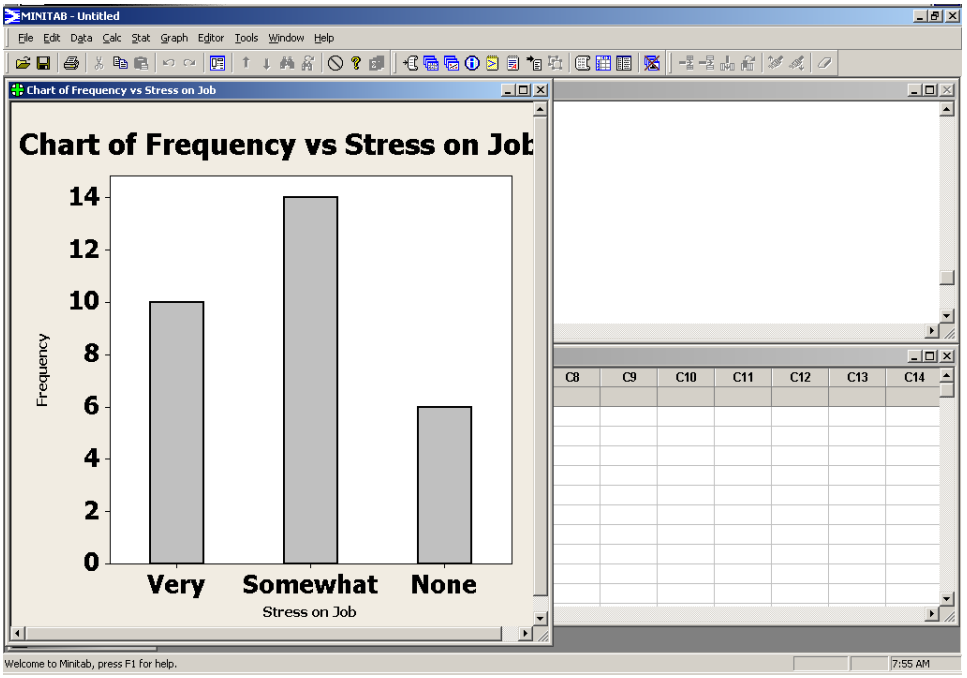
Notice there are 3 types of bar charts to choose from: simple, cluster, or stack. For this example, we will be creating a simple bar chart. To select this option, click on and highlight the picture of the simple bar chart. At the top of this pop-up screen, beside the field **Bars represent**, click on the down arrow to choose **Values from a table**. We are selecting this option because the data is summarized for us already. Click on **OK**.



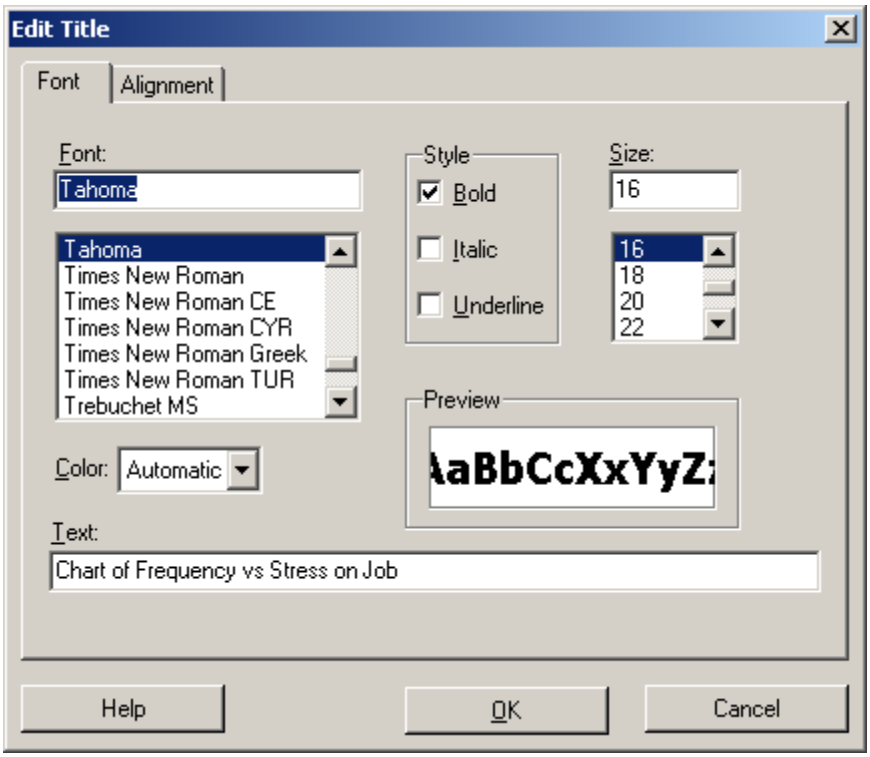
Now you must tell Minitab which variables to graph. Click in the field below **Graph variables:**. This variable must be a numerical measurement. In this case, it is our frequencies. Minitab is expecting a *column number* here. Since C2 contains the frequencies, type in **C2**. (The other way to do this is to notice that when you click on the field below "**Graph variables**", a list of columns appears on the left side of the screen. You can select a column from the list by double-clicking on it, or after a single click highlights the variable, click on **SELECT** at the bottom of the screen.) Next click on the field below **Categorical variables**. Select **C1**, the stress levels. In the screen shot below, you see the column names, and not the column numbers. This is because the double-click method (described above) was used to select the fields.



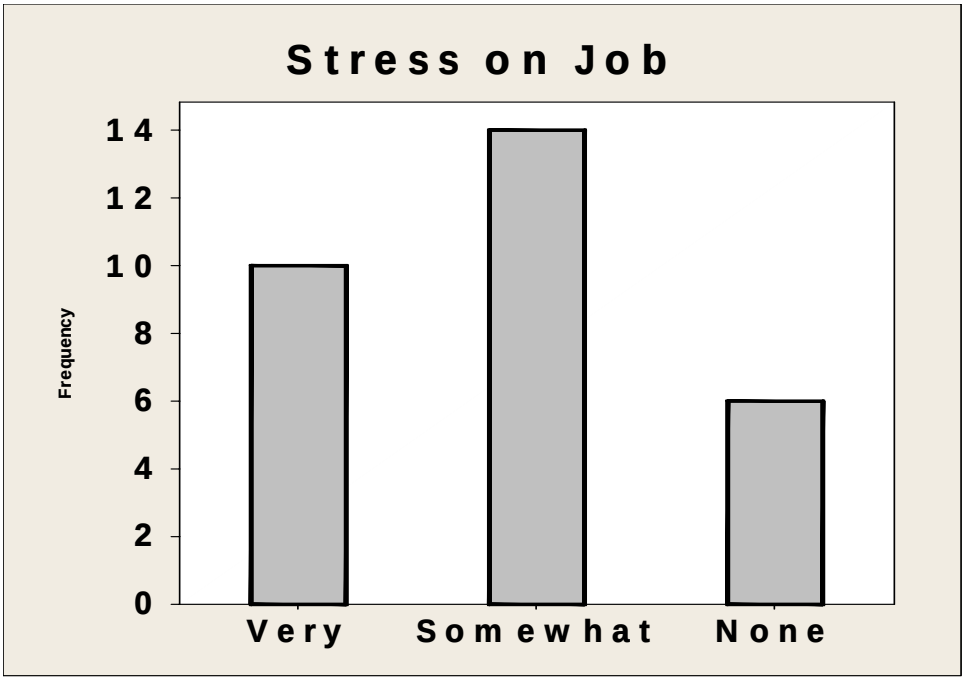
You can click on **Labels** and enter a title for your graph. However, if you click on **OK**, MINITAB will draw a bar graph using the default settings. Your graph will look like the one below.



Using the default settings has produced a good graph in this case. The only thing that needs to be changed is the title. It would be better to have a title "Stress on the Job". To change the title (or other graph options), right click on the title on the graph window. On the pop-up screen that appears, select **Edit title.** The following screen will appear.



From this screen, you can change the font, font size, color, style, and the text. Change the T**ext** to "Stress on the Job". Click on **OK** to view the bar chart now.

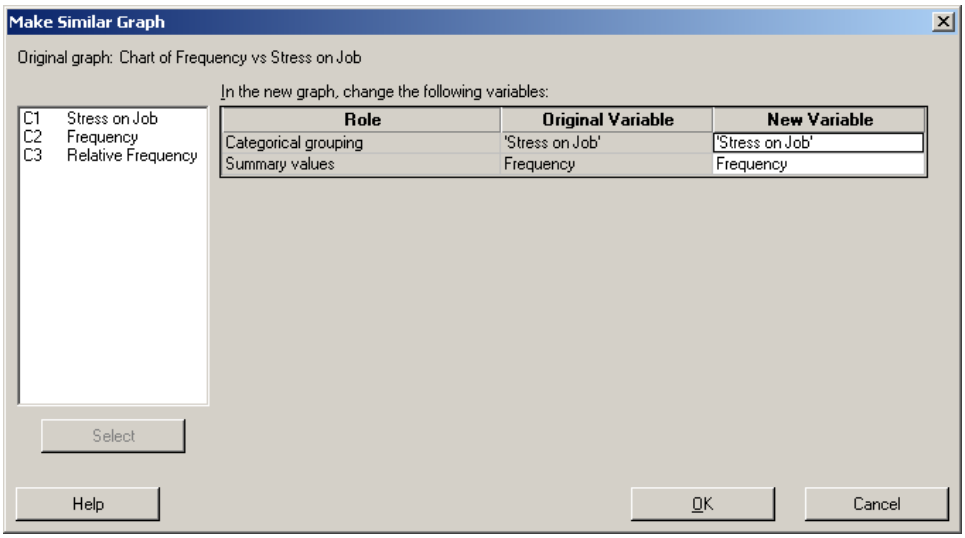


If you want to print the graph, click on **File → Print Graph.** Click **OK** and the graph should print. Once the graph prints, you can close the Graph Window by clicking on the X in the upper right corner of the window.

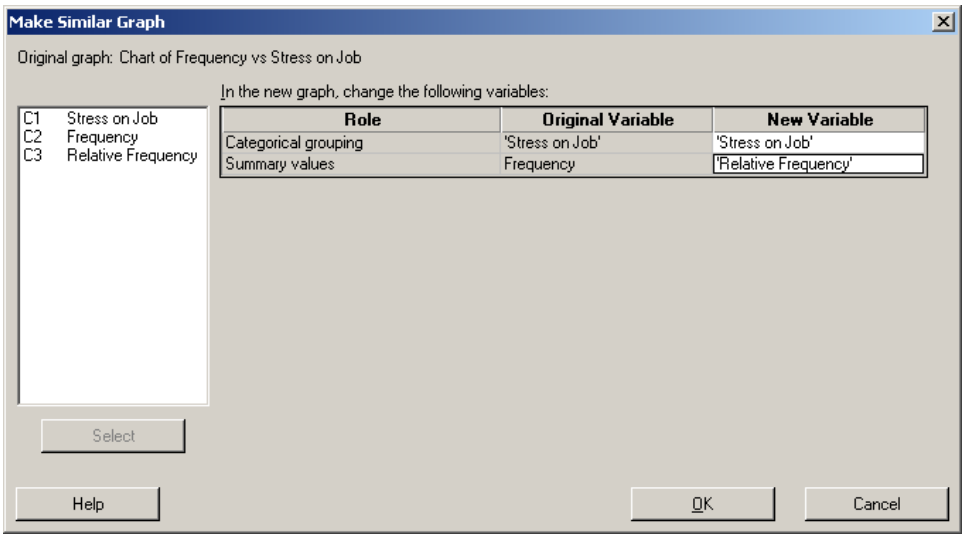
You can change any option on the graph by right clicking on the part that you want to change and selecting it from the pop-up that appears. For example, if you want to change the scale on the x-axis, just click anywhere on the x-axis and choose **Edit X Scale** from the menu.

To create a relative frequency bar chart, you can do the same thing as above. This time you would select the Relative Frequencies (in C3) as the **Graph variable**. However, a much simpler way to graph the relative frequencies is to use the frequency bar chart that you just created as a starting point.

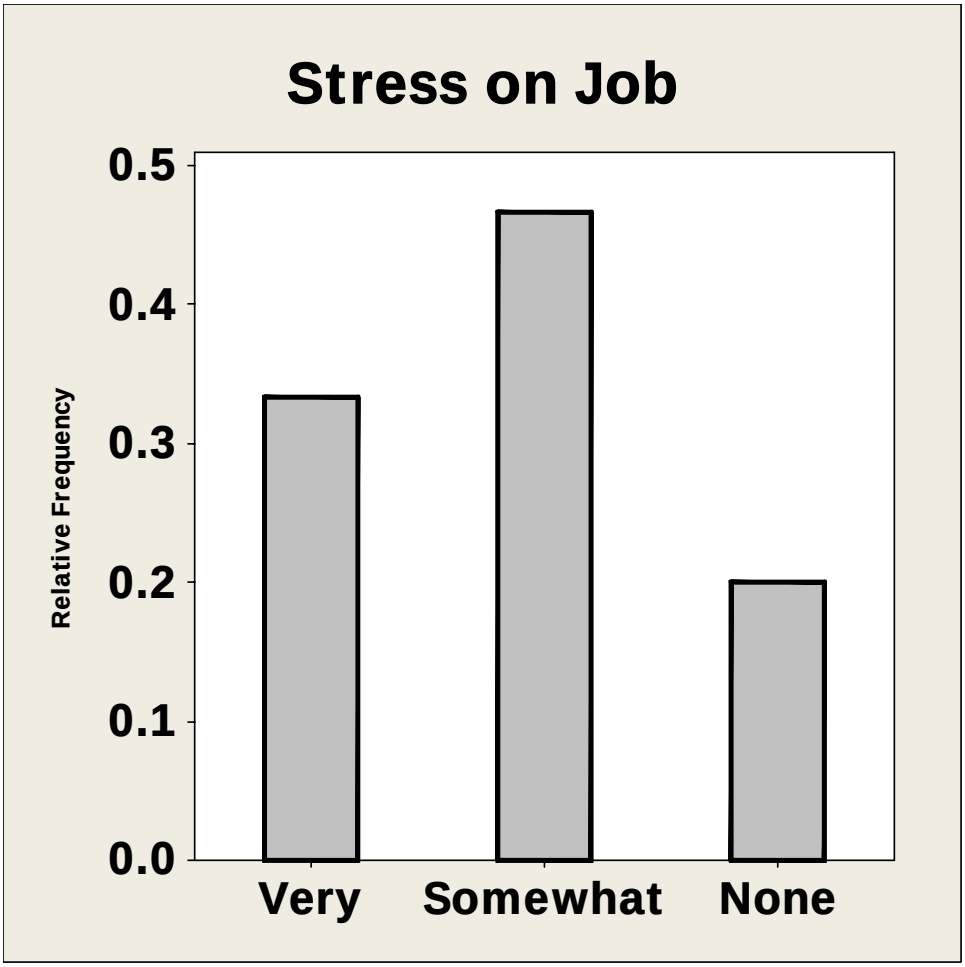
With the Graph Window still open, click on **Editor → Make Similar Graph.** The following window will pop up.



Beneath the column **New Variable**, click on "Frequency" to highlight it. Next, double click on C3 (Relative Frequency) on the far left of the window. Relative Frequency should now have replaced Frequency in the **New Variable** column.



Click on **OK** and the relative frequency bar chart should be displayed.

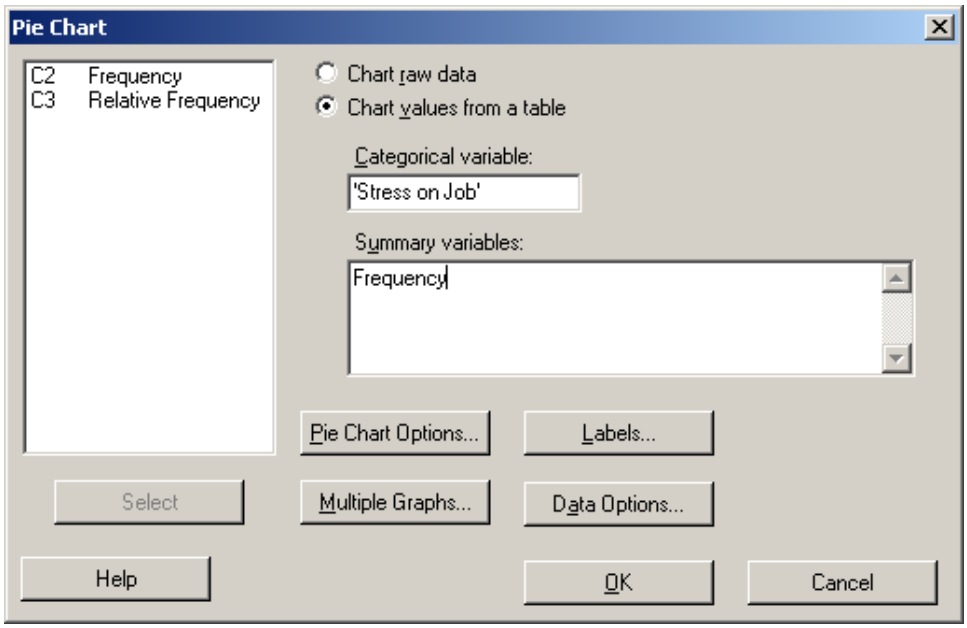

Stress on Job
Relative Frequency
0.5
0.4
0.3
0.2
0.1
0.0
Very
Somewhat
None

*

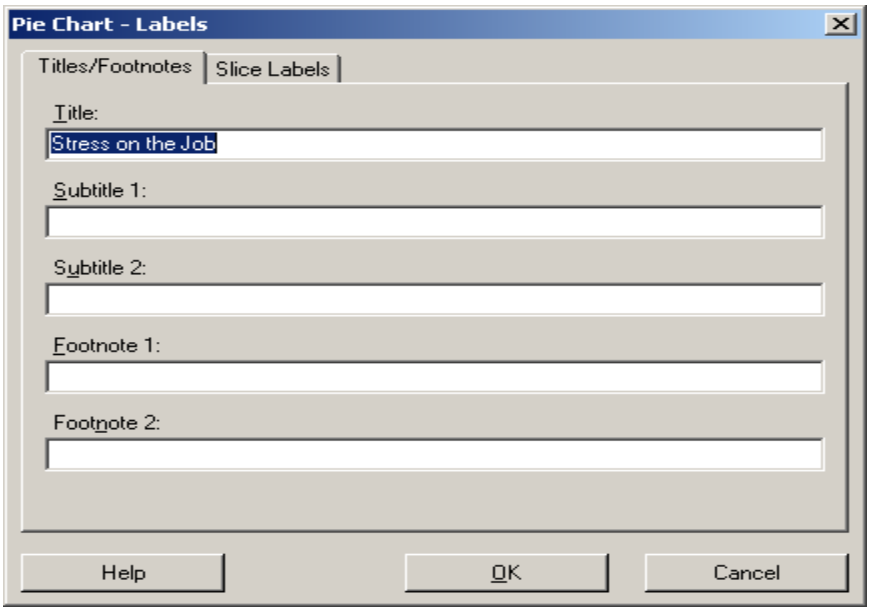
Pie Charts, pg. 30

Use the job stress data to make a pie chart. Pie charts use the frequencies, rather than the relative frequencies, so be sure to use C2.

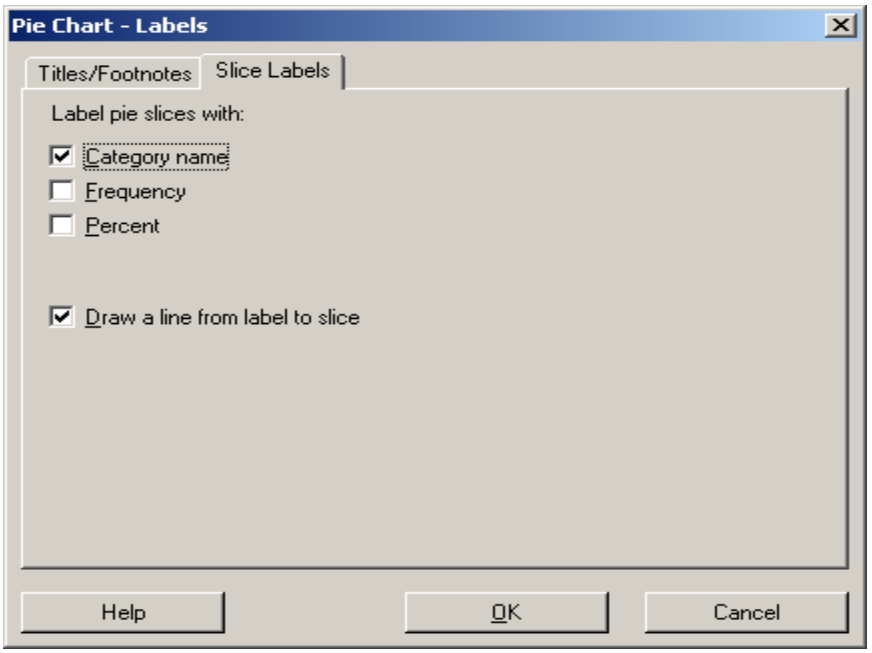
To create a Pie chart, click on **Graph → Pie chart.** On the screen that appears, select **Chart values from a table** by clicking on the small circle to the left of it. Select "Stress on Job" as the **Categorical variable**, and "Frequency" as the **Summary variable.**



Next, click on **Labels** to enter a title for the pie chart.



Click on the tab **Slice Labels.** For this pie chart, the labels for each slice should be the **Category name.** If you'd like a line drawn from the label to the slice, you can select this option also.



Click on **OK** twice and the pie chart should be displayed.


Stress on the Job
Category
Very
Somewhat
None
None
Very
Somewhat

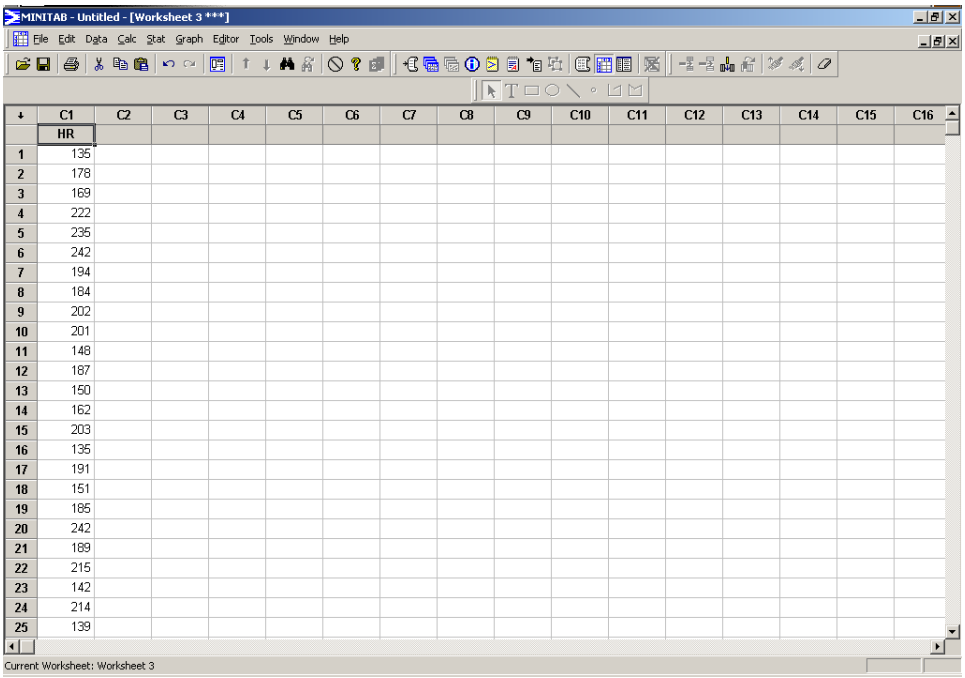
*

Section 2.3

Table 2.9 (page 37 of the textbook) gives the total home runs hit by all the players of each of the 30 Major League Baseball teams during the 2004 season.

Team	HR	Team	HR	Team	HR
Arizona	135	Florida	148	Oakland	189
Atlanta	178	Houston	187	Philadelphia	215
Baltimore	169	Kansas City	150	Pittsburgh	142
Boston	222	Anaheim Angels	162	St. Louis	214
Chicago Cubs	235	LA Dodgers	203	San Diego	139
Chicago White Sox	242	Milwaukee	135	San Francisco	183
Cincinnati	194	Minnesota	191	Seattle	136
Cleveland	184	Montreal (Washington)	151	Tampa Bay	145
Colorado	202	NY Mets	185	Texas	227
Detroit	201	NY Yankees	242	Toronto	145

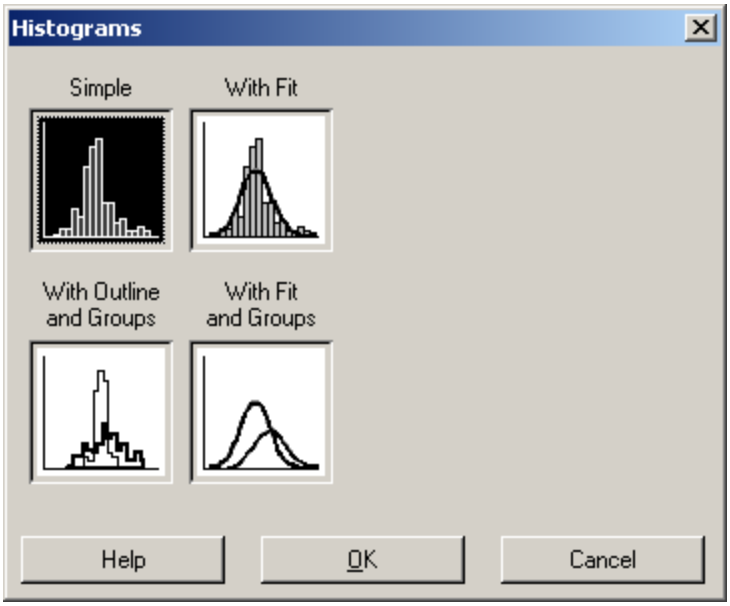
Enter the data into a Minitab worksheet. Enter the number of homeruns into C1 and give C1 an appropriate name. The team name will not be used for these graphs, so you don't need to type these into the Data Window. This data will be used to draw a histogram.



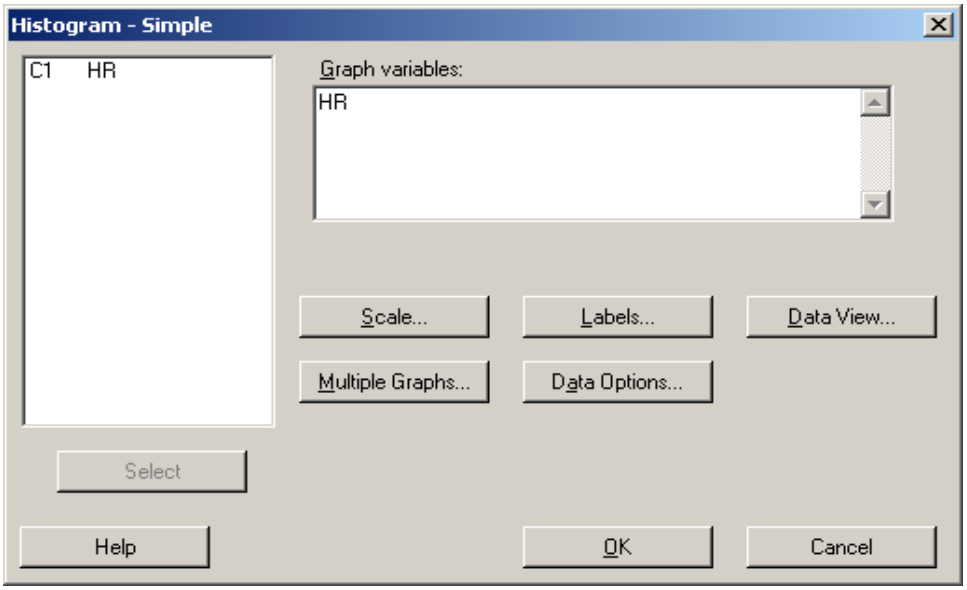
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16
	HR															
1	135															
2	178															
3	169															
4	222															
5	235															
6	242															
7	194															
8	184															
9	202															
10	201															
11	148															
12	187															
13	150															
14	162															
15	203															
16	135															
17	191															
18	151															
19	185															
20	242															
21	189															
22	215															
23	142															
24	214															
25	139															

Histograms, pg. 39

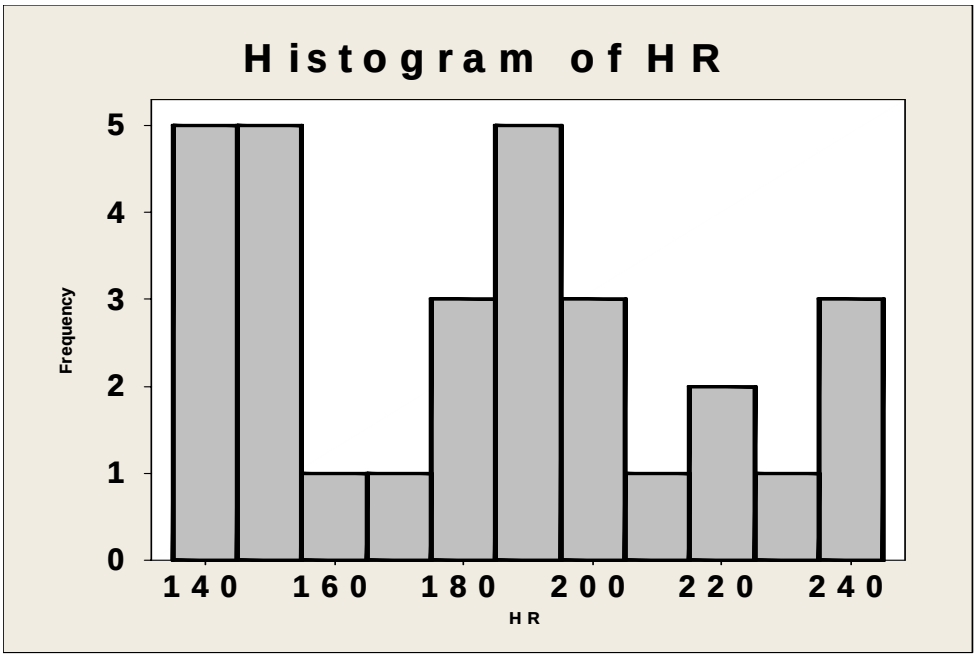
To make a histogram of the home run data, click on: **Graph** → **Histogram**.



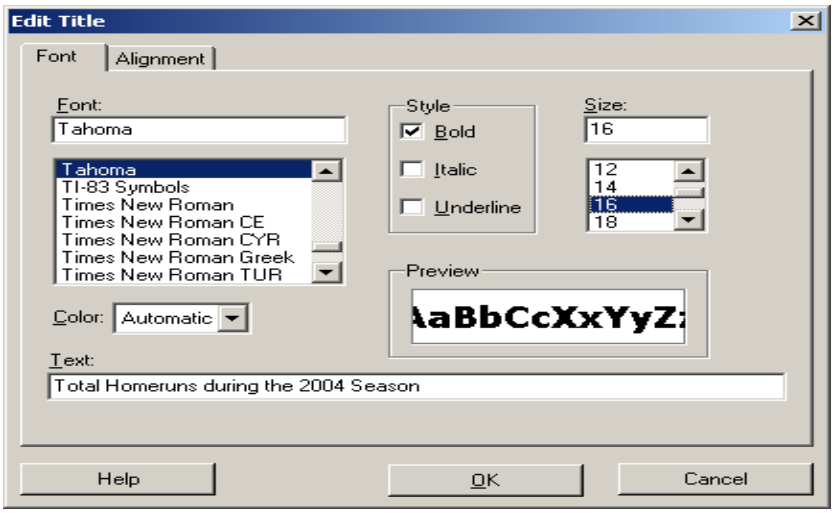
Select a **Simple** histogram and click on **OK**.



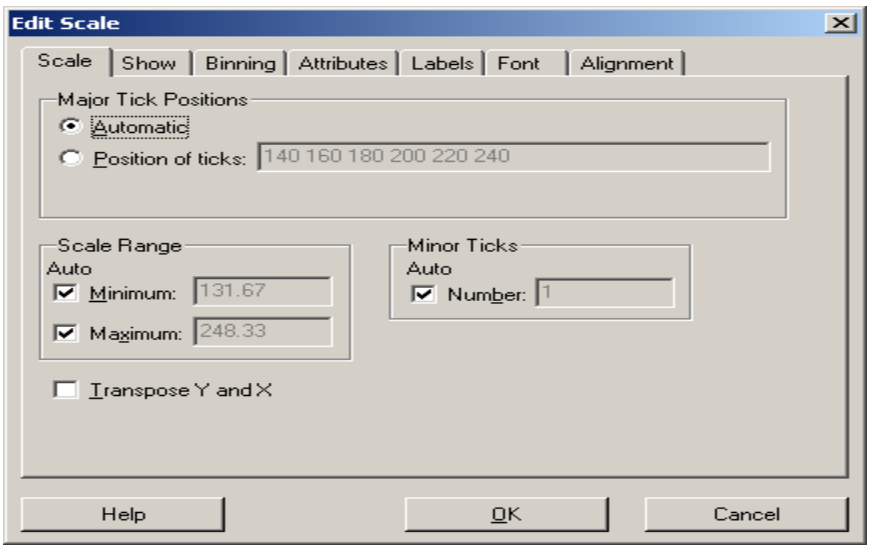
Since “HR” is the variable that you want to graph, double click on C1 in the large box at the left of the screen. "HR" should now be filled in as the **Graph variable**. If you click on **OK** now, MINITAB will draw a histogram using default settings. Your histogram will look like the one below.



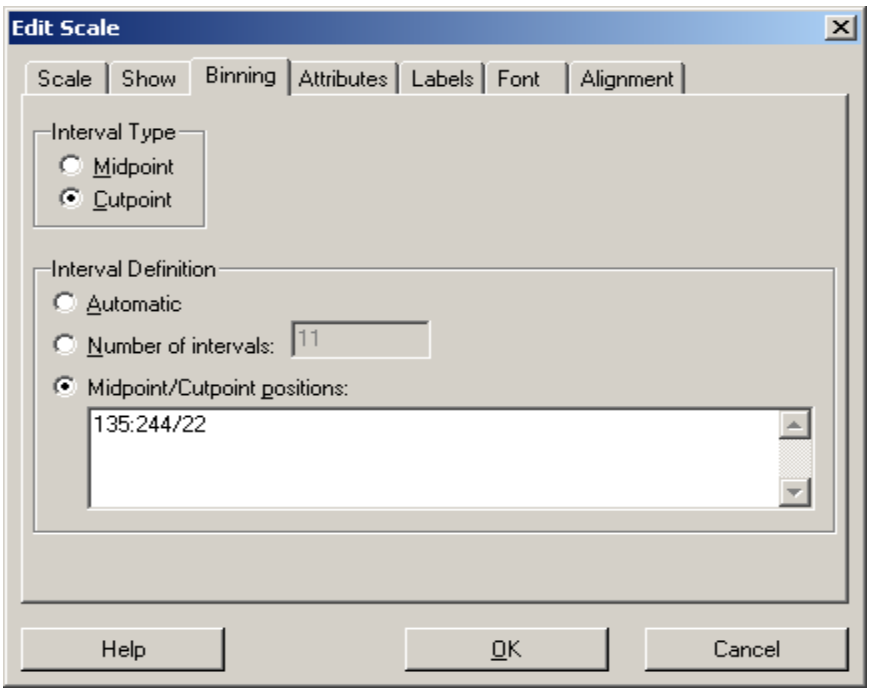
You would probably want to change this title to a more descriptive title. Also, notice that the numbering on the X-axis is not the same as in the textbook. Ordinarily, you would not be trying to make a histogram look exactly like another picture, however, it is worth learning how to edit the settings so that you know how to work with the Minitab program. First, put an appropriate title on your histogram. Right click on the title of the graph. On the drop down list that appears, select **Edit Title.** On the following popup, enter an appropriate title for the graph beneath **Text:.**



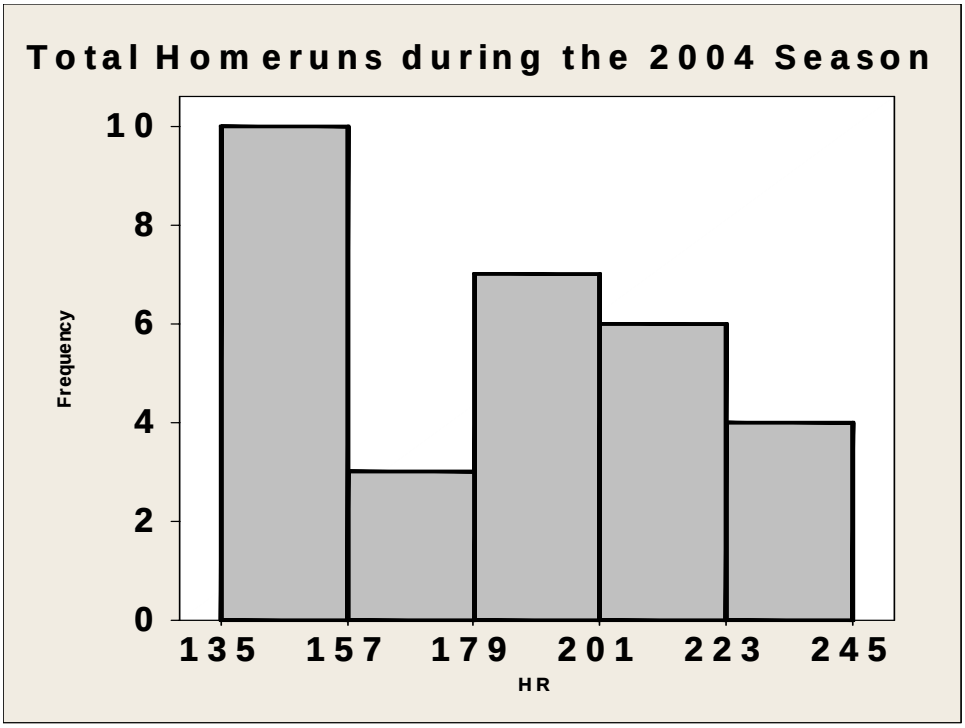
Click on **OK** and the title should be changed on your graph. To change the numbering on the X-axis, right click on the x-axis and select **Edit X Scale.**



Click on the tab labeled “Binning” to change the classes and interval sizes of the histogram.



Look at the classes used in the textbook. The first class begins at 135, the class width is 22, and the largest data point is 244. To use these settings, first select **Cutpoint** for **Interval Type.** A cutpoint locates the numbers for each class at the start of the bar, rather than at the midpoint. Next, for **Interval Definition,** click on **Midpoint/cutpoint positions** and type in **135:244/22.** This tells Minitab you want the numbering to go from 135 to 244 in steps of 22. Click on **OK** to view the histogram.



This time the graph looks good. To print the graph, click on **File → Print Graph. C**lick **OK** and the graph should print.

❋

Polygons pg. 40

To make a polygon of this data, repeat all steps shown above for the histogram of the data. You can edit this histogram to create the polygon. Right click on the graph and move the cursor up to **Add** and then select **Smoother.** The **Degree of smoothing** is 0 and the **Number of steps** is 2. Click on **OK** and the polygon should be superimposed on the bars of the histogram. To remove the bars, right click on the graph, move the cursor up to **Add** and then select **Data display.** On the pop-up that appears, **Un**check **Bars**. Click on **OK** to view the polygon.

Single-Valued Classes, pg. 43

Use Minitab's bar chart to graph single-valued classes. Just click on **Graph → Bar Chart.** Instructions for bar charts were given previously.

Section 2.5

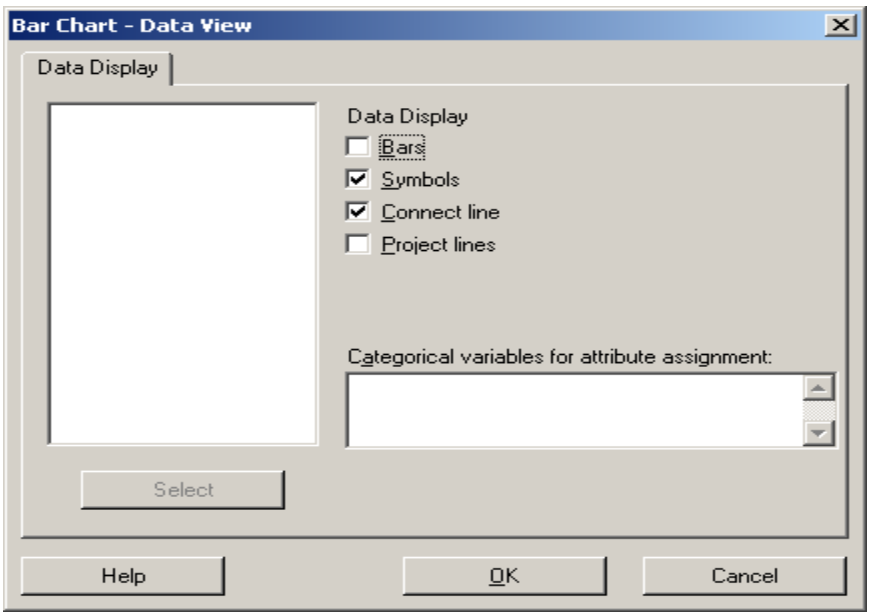
Ogives, pg. 52 (with summarized data)

In Minitab, you must create an ogive from summarized data. You will need to know the lower class boundaries and the cumulative frequency of each class. They are summarized below. Enter them into the Data Window.

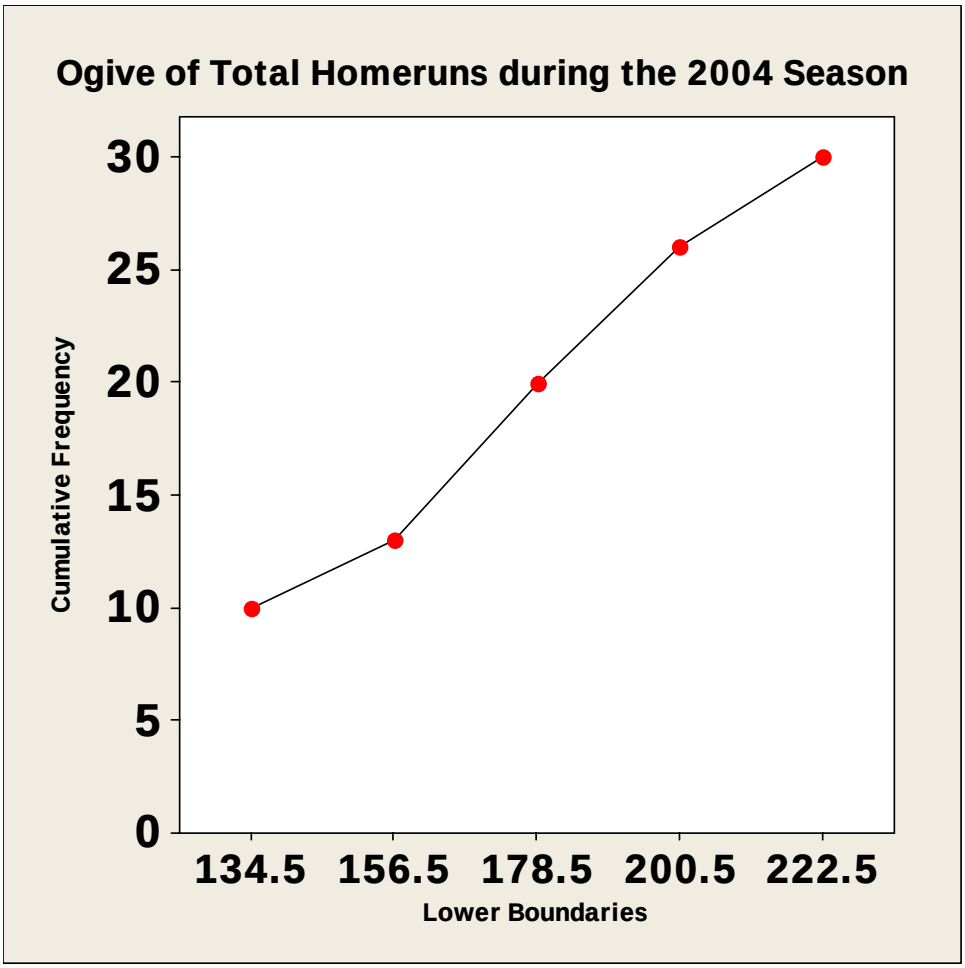
Lower Boundaries	Cumulative Frequency
134.5	10
156.5	13
178.5	20
200.5	26
222.5	30

Click on **Graph → Bar Chart.** The "cumulative frequency" is the **Y** variable and the "lower boundaries" is the **X** variable. Select a **Simple** bar chart and **Bars represent Values from a table.** Enter "Lower Limits" for the **Categorical variable** and "Cumulative Frequency" for the **Graph variable.**

Next, click on the button labeled "**Data View"** select **Symbols** and **Connect line**.



Click on **OK.** **C**lick on the **Labels** button and enter a title for the ogive. Click on **OK** twice to view the ogive.



✱

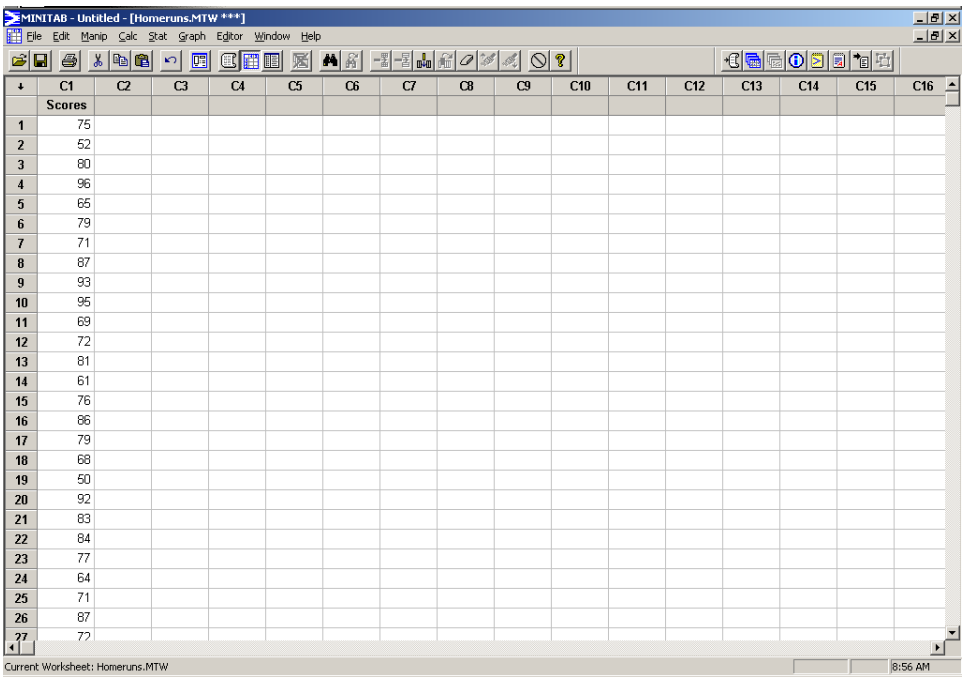
Section 2.6

Example 2-8, pg. 55 Stem-and-leaf displays

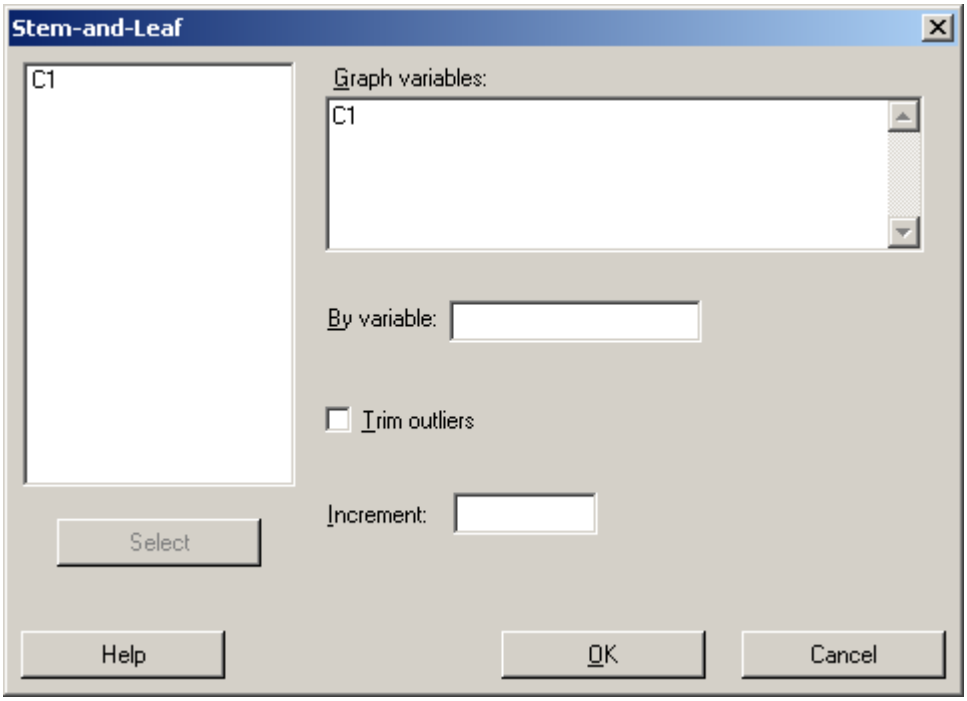
The following data are the scores of 30 students on a statistics test.

75	52	80	96	65	79	71	87	93	95
69	72	81	61	76	86	79	68	50	92
83	84	77	64	71	87	72	92	57	98

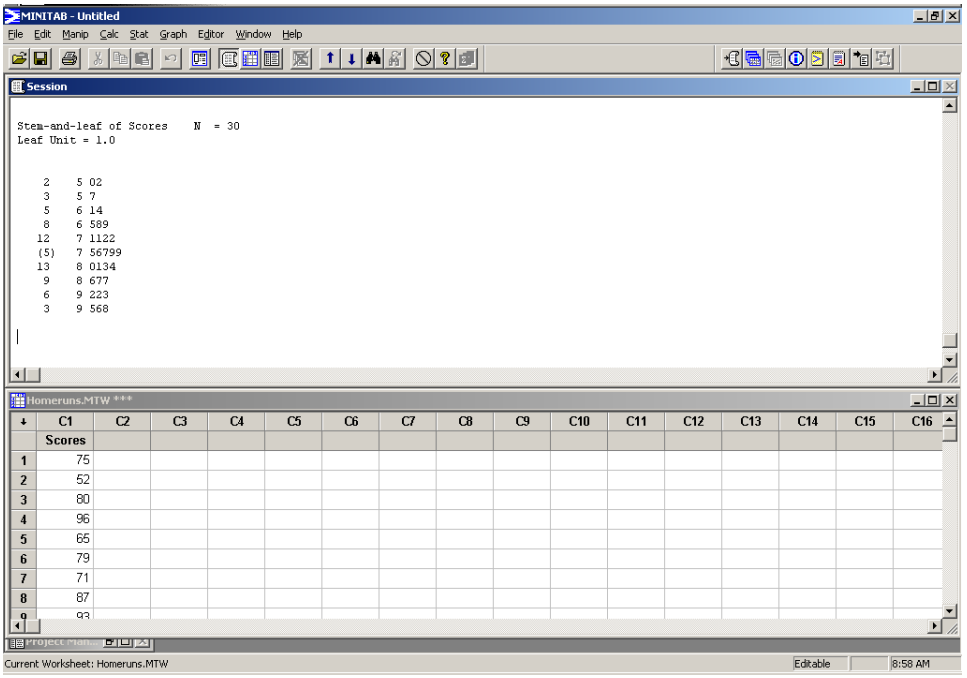
Enter the data into a Minitab worksheet.



To construct a Stem-and-leaf plot, click on **Graph → Stem-and-Leaf.** On the screen that appears, select C1 as your **GraphVariable** by doubling clicking on C1.



Click on **OK.** The stem and leaf plot will be displayed in the Session Window.



In this MINITAB display, the first column on the left is a counter. This column counts the number of data points starting from the smallest value (at the top of the plot) down to the median. It also counts from the largest data value (at the bottom of the plot) up to the median. Notice that there are two data points in the first row of the stem and leaf plot, so the counter is "2". Row 2 has 1 data point so the counter increases to “3” (2+1=3). The row that contains the median has the number “5” in parentheses. This number counts the number of data points that are in the row that contains the median. If there are no data points in a row, the counter on the left remains at the number from the previous row.

The second column in the display is the **Stem**. In this example, the Stem values range from 5 to 9. Notice that this display contains two rows for each of the values. These are called **split-stems**. For each stem value, the first row contains all data points with leaf values from 0 to 4 and the second row contains all data points with leaf values from 5 to 9. Notice that MINITAB constructs an *ordered* stem and leaf.

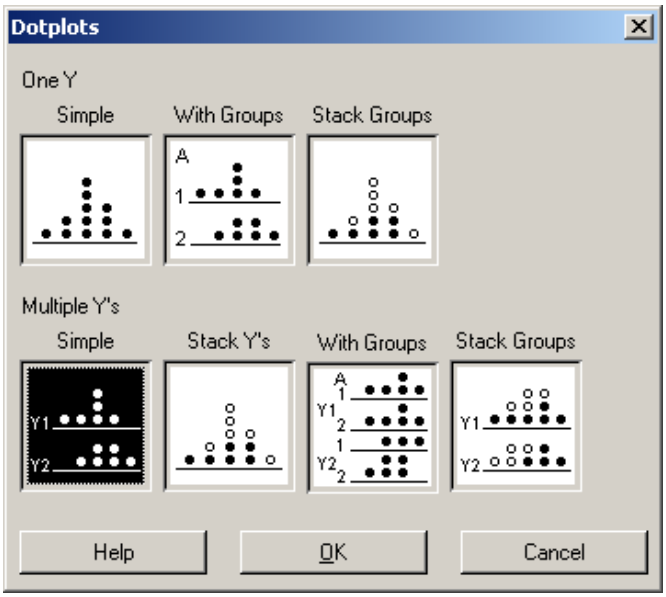
The leaf values are shown to the right of the stem. The leaf values may be the actual data points or they may be the rounded data points. To find the actual values of the data points in the display, use the “Leaf Unit=” statement at the top of the display. The "Leaf Unit" gives you the place value of the leaves. In this stem and leaf plot, the first data point has a stem value of 5 and a leaf value of 0. Since the "Leaf Unit=1.0", the leaf value of 0 is the "ones" place and the stem value of 5 is the "tens" places. Thus the data point is 50. A 50 was the lowest score on the Statistics test.

✱

Section 2.7

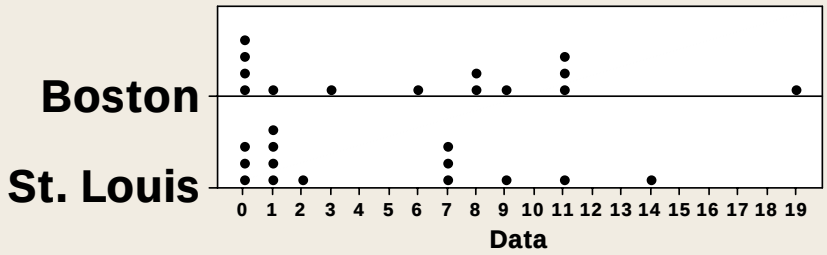
Example 2-12, pg. 60 Dotplots

Enter the Boston RBIs from Table 2.16 on pg. 59 of the text into C1 of a Minitab worksheet, and the St. Louis RBIs from Table 2.17 into C2. (It is not necessary to enter the names of the batters.) To create the dotplots, click on **Graph → Dotplot.** On the screen that appears, select a **Simple** dotplot with **Multiple Y's** and click **OK**.



On the screen that appears, select both C1 and C2 as your **GraphVariables.** Click on **Labels** and enter an appropriate **Title** for the plot. Click on **OK** to view the dotplots.

RBIs During 2004 Playoffs



*

Suggested Exercises

Section 2.2

pp. 32-33: 2.5, 2.6, 2.10

Section 2.3

pp. 48-51: 2.19, 2.20, 2.28, 2.29

Section 2.5

pp. 54: 2.34, 2.35

Section 2.6

pp. 58: 2.48, 2.49

Section 2.7

pp. 61: 2.63, 2.64