2.1 | Stem-and-Leaf Graphs (Stemplots), Line Graphs, and Bar Graphs

One simple graph, the **stem-and-leaf graph** or **stemplot**, comes from the field of exploratory data analysis. It is a good choice when the data sets are small. To create the plot, divide each observation of data into a stem and a leaf. The leaf consists of a **final significant digit**. For example, 23 has stem two and leaf three. The number 432 has stem 43 and leaf two. Likewise, the number 5,432 has stem 543 and leaf two. The decimal 9.3 has stem nine and leaf three. Write the stems in a vertical line from smallest to largest. Draw a vertical line to the right of the stems. Then write the leaves in increasing order next to their corresponding stem.

Example·2.1¶

For Susan Dean's spring pre-calculus class, scores for the first exam were as follows (smallest to largest):
33; 42; 49; 49; 53; 55; 55; 61; 63; 67; 68; 68; 69; 69; 72; 73; 74; 78; 80; 83; 88; 88; 88; 90; 92; 94; 94; 94; 94; 96; 100

Table 2.1 Stem-and-Leaf Graph

Stem	Leaf
3	3
4	299
5	355
6	1378899
7	2348
8	03888
9	0244446
10	0

The stemplot shows that most scores fell in the 60s, 70s, 80s, and 90s. Eight out of the 31 scores or approximately 26 percent ($\frac{8}{31}$) were in the 90s or 100, a fairly high number of As.

2.1 For the Park City basketball team, scores for the last 30 games were as follows (smallest to largest):
32; 32; 33; 34; 38; 40; 42; 42; 43; 44; 46; 47; 47; 48; 48; 48; 49; 50; 50; 51; 52; 52; 52; 53; 54; 56; 57; 57; 60; 61
for

Solution 2.1

Stem	Leaf
3	22348
4	022346778889
5	00122234677
6	01

The stemplot is a quick way to graph data and gives an exact picture of the data. You want to look for an overall pattern and any outliers. An **outlier** is an observation of data that does not fit the rest of the data. It is sometimes called an **extreme value.** When you graph an outlier, it will appear not to fit the pattern of the graph. Some outliers are due to mistakes (for example, writing down 50 instead of 500) while others may indicate that something unusual is happening. It takes some background information to explain outliers, so we will cover them in more detail later.

Example 2.2

The data are the distances (in kilometers) from a home to local supermarkets. Create a stemplot using the data:
1.1; 1.5; 2.3; 2.5; 2.7; 3.2; 3.3; 3.3; 3.5; 3.8; 4.0; 4.2; 4.5; 4.5; 4.7; 4.8; 5.5; 5.6; 6.5; 6.7; 12.3

Do   of

Hint: The leaves are to the right of the decimal.

Solution 2.2

The value 12.3 may be an outlier. Values appear to concentrate at three and four kilometers.

Stem	Leaf
1	15
2	357
3	23358
4	025578
5	56
6	57
7	
8	
9	
10	

11	
12	3

Try It Σ

2.2 The following data show the distances (in miles) from the homes of off-campus statistics students to the college. Create a stemplot using the data and identify any outliers:

0.5; 0.7; 1.1; 1.2; 1.2; 1.3; 1.3; 1.5; 1.5; 1.7; 1.7; 1.8; 1.9; 2.0; 2.2; 2.5; 2.6; 2.8; 2.8; 2.8; 3.5; 3.8; 4.4; 4.8; 4.9; 5.2; 5.5; 5.7; 5.8; 8.0

Solution 2.2   8.0

Stem	Leaf
0	57
1	12233557789
2	0256888
3	58
4	489
5	2578
6	
7	
8	0

The value 8.0 may be an outlier. Values appear to concentrate at one and two miles.

Example 2.3

A **side-by-side stem-and-leaf plot** allows a comparison of the two data sets in two columns. In a side-by-side stem-and-leaf plot, two sets of leaves share the same stem. The leaves are to the left and the right of the stems. **Table 2.4** and **Table 2.5** show the ages of presidents at their inauguration and at their death. Construct a side- by-side stem-and-leaf plot using this data.

Table 2.4 Presidential Ages at Inauguration

President	Age	President	Age	President	Age
Washington	57	Lincoln	52	Hoover	54
J. Adams	61	A Johnson	56	F Roosevelt	51
Jefferson	57	Grant	46	Truman	60
Madison	57	Hayes	54	Eisenhower	62
Monroe	58	Garfield	49	Kennedy	43
J Q Adams	57	Arthur	51	L Johnson	55
Jackson	61	Cleveland	47	Nixon	56
Van Buren	54	B Harrison	55	Ford	56
W H Harrison	68	Cleveland	55	Carter	52
Tyler	51	McKinley	54	Reagan	69
Polk	49	T Roosevelt	42	G H W Bush	64
Taylor	64	Taft	51	Clinton	47
Fillmore	50	Wilson	56	G W Bush	54
Pierce	48	Harding	55	Obama	47
Buchanan	65	Coolidge	51		

Table 2.3

Table 2.5 Presidential Age at Death

President	Age	President	Age	President	Age
Washington	67	Lincoln	56	Hoover	90

President	Age	President	Age	President	Age
J. Adams	90	A Johnson	66	F. Roosevelt	63
Jefferson	83	Grant	63	Truman	88
Madison	85	Hayes	70	Eisenhower	78
Monroe	73	Garfield	49	Kennedy	46
J. Q. Adams	80	Arthur	56	L. Johnson	64
Jackson	78	Cleveland	71	Nixon	81
Van Buren	79	B. Harrison	67	Ford	93
W. H. Harrison	68	Cleveland	71	Reagan	93
Tyler	71	McKinley	58		
Polk	53	T. Roosevelt	60		
Taylor	65	Taft	72		
Fillmore	74	Wilson	67		
Pierce	64	Harding	57		
Buchanan	77	Coolidge	60		

Try It Σ

2.3 The table shows the number of wins and losses the Atlanta Hawks have had in 42 seasons. Create a side-by-side stem-and-leaf plot of these wins and losses.

Losses	Wins	Year	Losses	Wins	Year
34	48	1968-1969	41	41	1989-1990
34	48	1969-1970	39	43	1990-1991
46	36	1970-1971	44	38	1991-1992
46	36	1971-1972	39	43	1992-1993
36	46	1972-1973	25	57	1993-1994
47	35	1973-1974	40	42	1994-1995
51	31	1974-1975	36	46	1995-1996
53	29	1975-1976	26	56	1996-1997
51	31	1976-1977	32	50	1997-1998
41	41	1977-1978	19	31	1998-1999
36	46	1978-1979	54	28	1999-2000
32	50	1979-1980	57	25	2000-2001
51	31	1980-1981	49	33	2001-2002
40	42	1981-1982	47	35	2002-2003

Table 2.6

Losses	Wins	Year	Losses	Wins	Year
39	43	1982-1983	54	28	2003-2004
42	40	1983-1984	69	13	2004-2005
48	34	1984-1985	56	26	2005-2006
32	50	1985-1986	52	30	2006-2007
25	57	1986-1987	45	37	2008-2009
32	50	1987-1988	35	47	2008-2009
30	52	1988-1989	29	53	2009-2010

Table 2.6

Solution 2.3

Number of Wins		Number of Losses
3	1	9
9 8 8 6 5	2	5 5 9
8 7 6 6 5 5 4 3 1 1 1 1 0	3	0 2 2 2 2 4 4 5 6 6 6 9 9 9
8 8 7 6 6 6 3 3 3 2 2 1 1 0	4	0 0 1 1 2 4 5 6 6 7 7 8 9
7 7 6 3 2 0 0 0 0	5	1 1 1 2 3 4 4 6 7
	6	9

Another type of graph that is useful for specific data values is a **line graph**. In the particular line graph shown in **Example 2.4**, the **x-axis** (horizontal axis) consists of **data values** and the **y-axis** (vertical axis) consists of **frequency points**. The frequency points are connected using line segments.

Table 2.7

Number of times Teenager is Reminded	Frequency
0	2
1	5
2	8
3	14
4	7
5	4

Figure 2.2
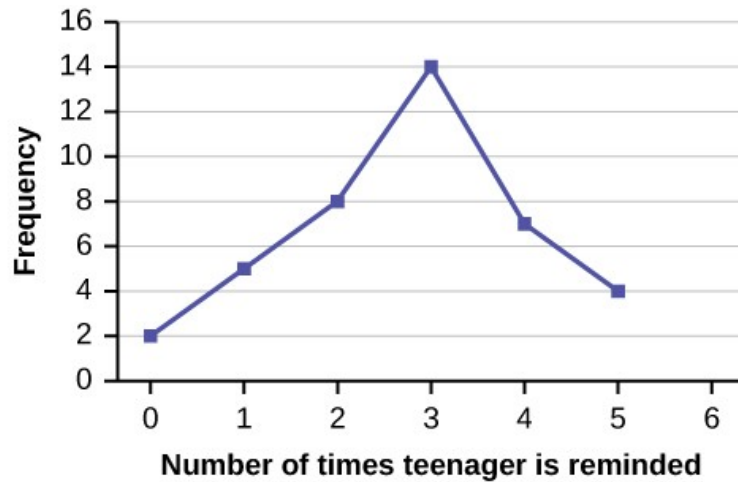
2.4 In a survey, 40 people were asked how many times per year they had their car in the shop for repairs. The results are shown in **Table 2.8**. Construct a line graph.

Number of times in shop	Frequency
0	7
1	10
2	14
3	9

Table 2.8

Solution 2.4

Figure 2.

Bar graphs consist of bars that are separated from each other. The bars can be rectangles or they can be rectangular boxes (used in three-dimensional plots), and they can be vertical or horizontal. The **bar graph** shown in **Example 2.5** has age groups represented on the **x-axis** and proportions on the **y-axis**.

Example 2.5

By the end of 2011, Facebook had over 146 million users in the United States. **Table 2.9** shows three age groups, the number of users in each age group, and the proportion (%) of users in each age group. Construct a bar graph using this data.

Table 2.9

Age Groups	Number of Facebook users	Proportion (percent) of Facebook users
13-25	65,082,280	45 percent
26-44	53,300,200	36 percent
45-64	27,885,100	19 percent

Solution 2.5

Figure 2.3
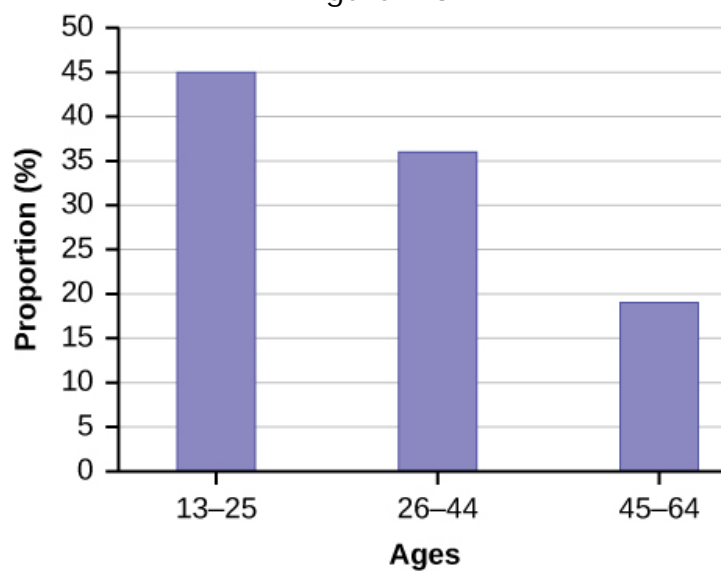
Try It Σ

2.5 The population in Park City is made up of children, working-age adults, and retirees. **Table 2.10** shows the three age groups, the number of people in the town from each age group, and the proportion (percent) of people in each age group. Construct a bar graph showing the proportions.

Table 2.10

Age groups	Number of people	Proportion of population
Children	67,059	19 percent
Working-age adults	152,198	43 percent
Retirees	131,662	38 percent

Solution 2.5


Figure 3.
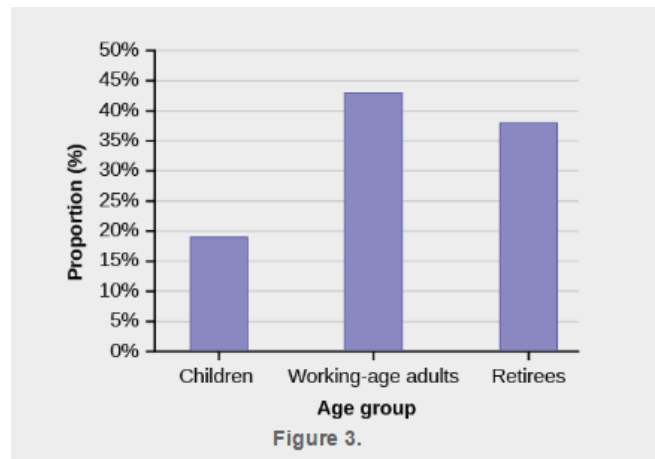
Example 2.6

The columns in **Table 2.11** contain: the race or ethnicity of students in U.S. public schools for the class of 2011, percentages for the Advanced Placement examine population for that class, and percentages for the overall student population. Create a bar graph with the student race or ethnicity (qualitative data) on the *x*-axis, and the Advanced Placement examinee population percentages on the *y*-axis.

Table 2.11

Race/Ethnicity	AP Examinee Population	Overall Student Population
1 = Asian, Asian American or Pacific Islander	10.3 percent	5.7 percent
2 = Black or African American	9.0 percent	14.7 percent
3 = Hispanic or Latino	17.0 percent	17.6 percent
4 = American Indian or Alaska Native	0.6 percent	1.1 percent
5 = White	57.1 percent	59.2 percent
6 = Not reported/other	6.0 percent	1.7 percent

Solution 2.6

Figure 2.4
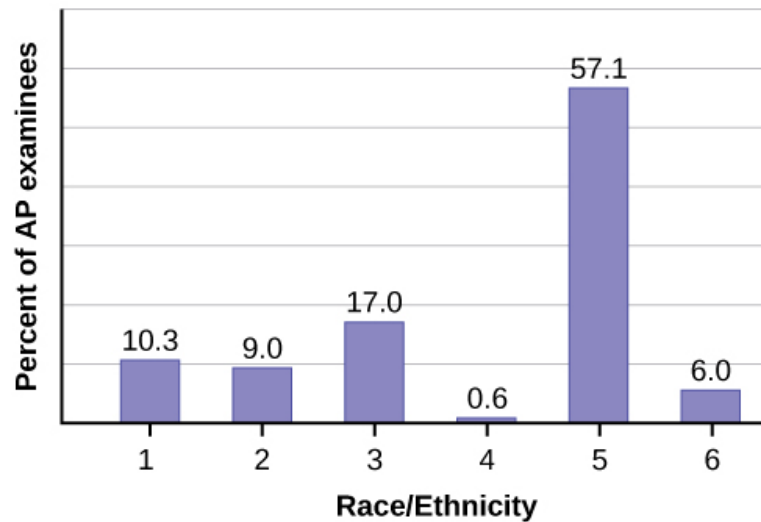
2.6 Park City is broken down into six voting districts. The table shows the percent of the total registered voter population that lives in each district as well as the percent total of the entire population that lives in each district. Construct a bar graph that shows the registered voter population by district.

District	Registered voter population	Overall city population
1	15.5 percent	19.4 percent
2	12.2 percent	15.6 percent
3	9.8 percent	9.0 percent
4	17.4 percent	18.5 percent
5	22.8 percent	20.7 percent
6	22.3 percent	16.8 percent

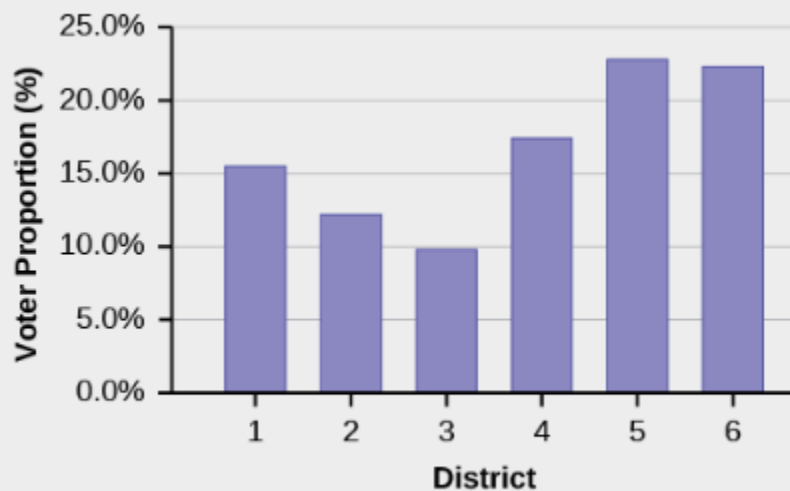Table 2.12

Solution 2.6

Figure 4.

References

Burbary, Ken. *Facebook Demographics Revisited – 2001 Statistics,* 2011. Available online at http://www.kenburbary.com/2011/03/facebook-demographics-revisited-2011-statistics-2/ (accessed August 21, 2013).

"9th Annual AP Report to the Nation." CollegeBoard, 2013. Available online at http://apreport.collegeboard.org/goals-and-findings/promoting-equity (accessed September 13, 2013).

"Overweight and Obesity: Adult Obesity Facts." Centers for Disease Control and Prevention. Available online at http://www.cdc.gov/obesity/data/adult.html (accessed September 13, 2013).

Chapter Review

A **stem-and-leaf plot** is a way to plot data and look at the distribution. In a stem-and-leaf plot, all data values within a class are visible. The advantage in a stem-and-leaf plot is that all values are listed, unlike a histogram, which gives classes of data values. A **line graph** is often used to represent a set of data values in which a quantity varies with time. These graphs are useful for finding trends. That is, finding a general pattern in data sets including temperature, sales, employment, company profit or cost over a period of time. A **bar graph** is a chart that uses either horizontal or vertical bars to show comparisons among categories. One axis of the chart shows the specific categories being compared, and the other axis represents a discrete value. Some bar graphs present bars clustered in groups of more than one (grouped bar graphs), and others show the bars divided into subparts to show cumulative effect (stacked bar graphs). Bar graphs are especially useful when categorical data is being used.

For each of the following data sets, create a stemplot and identify any outliers.
Exercise

The miles per gallon rating for 30 cars are shown below (lowest to highest).

19, 19, 19, 20, 21, 21, 25, 25, 25, 26, 26, 28, 29, 31, 31, 32, 32, 33, 34, 35, 36, 37, 37, 38, 38, 38, 38, 41, 43, 43
Solution

Stem	Leaf
1	9 9 9
2	0 1 1 5 5 5 6 6 8 9
3	1 1 2 2 3 4 5 6 7 7 8 8 8 8
4	1 3 3

Exercise

The height in feet of 25 trees is shown below (lowest to highest).

25, 27, 33, 34, 34, 34, 35, 37, 37, 38, 39, 39, 39, 40, 41, 45, 46, 47, 49, 50, 50, 53, 53, 54, 54

Exercise

The data are the prices of different laptops at an electronics store. Round each value to the nearest ten.

249, 249, 260, 265, 265, 280, 299, 299, 309, 319, 325, 326, 350, 350, 350, 365, 369, 389, 409, 459, 489, 559, 569, 570, 610

Solution

Stem	Leaf
2	5 5 6 7 7 8
3	0 0 1 2 3 3 5 5 5 7 7 9
4	1 6 9
5	6 7 7
6	1

Exercise

The data are daily high temperatures in a town for one month.

61, 61, 62, 64, 66, 67, 67, 67, 68, 69, 70, 70, 70, 71, 71, 72, 74, 74, 74, 75, 75, 75, 76, 76, 77, 78, 78, 79, 79, 95

For the next three exercises, use the data to construct a line graph.

Exercise

In a survey, 40 people were asked how many times they visited a store before making a major purchase. The results are shown in Table.

Number of times in store	Frequency
1	4
2	10
3	16
4	6
5	4

Solution

Figure 5.

Exercise
In a survey, several people were asked how many years it has been since they purchased a
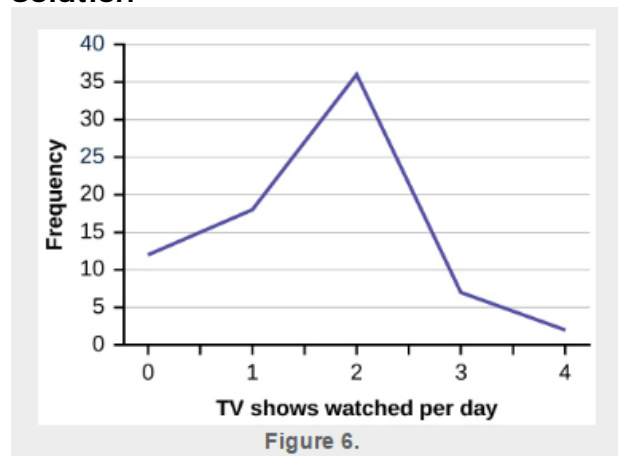
mattress. The results are shown in Table.

Years since last purchase	Frequency
0	2
1	8
2	13
3	22
4	16
5	9

Exercise

Several children were asked how many TV shows they watch each day. The results of the survey are shown in Table.

Number of TV Shows	Frequency
0	12
1	18
2	36
3	7
4	7

Solution

Figure 6.

Exercise

The students in Ms. Ramirez's math class have birthdays in each of the four seasons. The Table shows the four seasons, the number of students who have birthdays in each season, and the percentage (percent) of students in each group. Construct a bar graph showing the number of students.
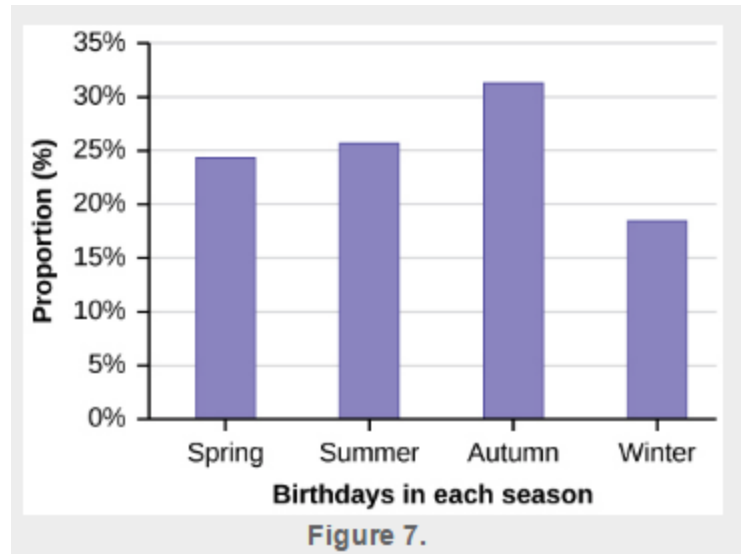
Seasons	Number of students	Proportion of population
Spring	8	24 percent
Summer	9	26 percent
Autumn	11	32 percent

Winter	6	18 percent

Exercise

Using the data from Mrs. Ramirez's math class supplied in Exercise, construct a bar graph showing the percentages.

Solution

Figure 7.

Exercise

David County has six high schools. Each school sent students to participate in a county-wide science competition. Table shows the percentage breakdown of competitors from each school, and the percentage of the entire student population of the county that goes to each school. Construct a bar graph that shows the population percentage of competitors from each school.

High School	Science competition population	Overall student population
Alabaster	28.9 percent	8.6 percent
Concordia	7.6 percent	23.2 percent
Genoa	12.1 percent	15 percent
Mocksville	18.5 percent	14.3 percent
Tynneson	24.2 percent	10.1 percent
West End	8.7 percent	28.8 percent

Exercise

Use the data from the David County science competition supplied in Exercise. Construct a bar graph that shows the county-wide population percentage of students at each school.
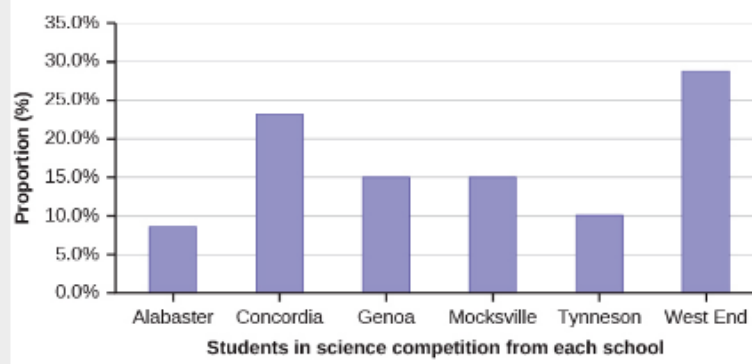
Solution

Figure 8.