

# Chapter 2

## Organizing and Graphing Data

---

### Section 2.1

**2.1** Data in their original form are often too large and unmanageable. It is easier to make sense of grouped data than ungrouped data and easier to make decisions and draw conclusions using grouped data.

**2.3 a.**

Category	Frequency
Y	23
N	13
D	4

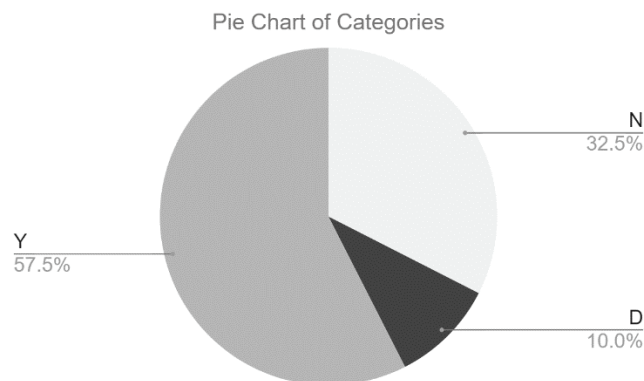
**b.**

Category	Relative Frequency	Percentage
Y	$23/40 = 0.575$	57.5
N	$13/40 = 0.325$	32.5
D	$4/40 = 0.100$	10.0

**c.** 57.5% of the elements belong to category Y.

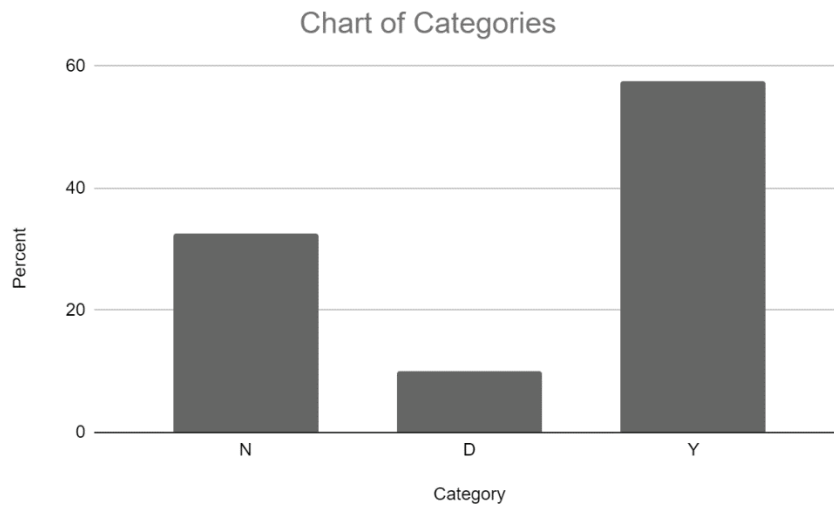
**d.**  $32.5\% + 10.0\% = 42.5\%$  of the elements belong to categories N or D.

**e.**



## 2 Chapter 2

f.



2.5 a.

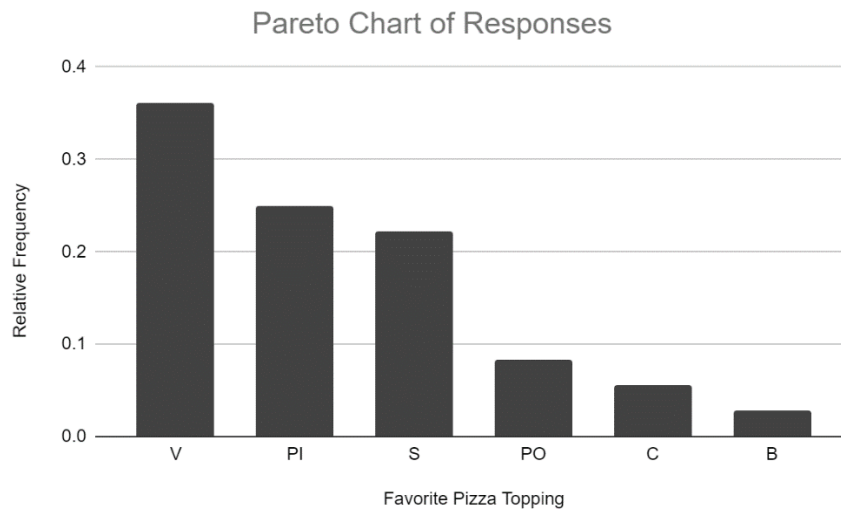
Pizza Topping	Frequency
PI	9
S	8
V	13
PO	3
B	1
C	2

b.

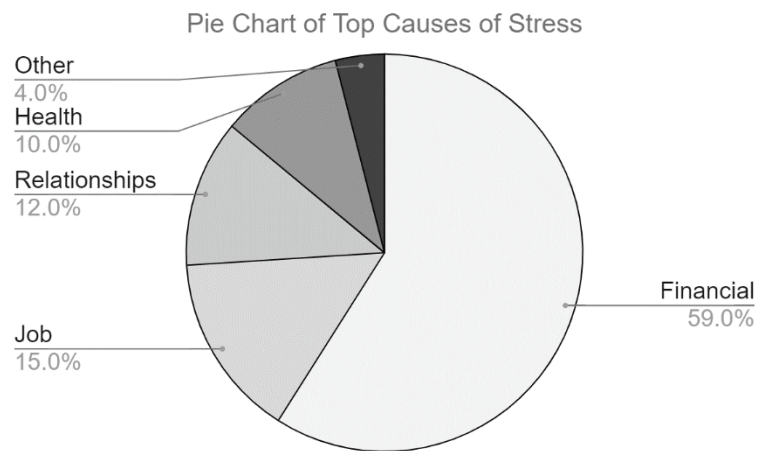
Pizza Topping	Relative Frequency	Percentage
PI	$9/36 = 0.250$	25.0
S	$8/36 = 0.222$	22.2
V	$13/36 = 0.361$	36.1
PO	$3/36 = 0.083$	8.3
B	$1/36 = 0.028$	2.8
C	$2/36 = 0.056$	5.6

- c.  $V + PO + C = 36.1\% + 8.3\% + 5.6\% = 50\%$   
 50% of the respondents mentioned vegetables and fruits, poultry, or cheese.

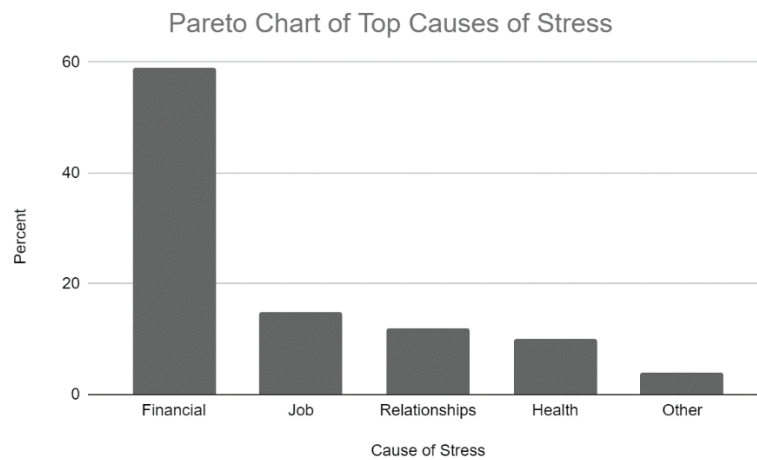
d.



2.7 a.



b.



**Section 2.2**

**2.9** The relative frequency for a class is obtained by dividing the frequency of that class by the sum of frequencies of all classes. The percentage for a class is obtained by multiplying the relative frequency of that class by 100. Example 2–4 is an example that illustrates the calculation of relative frequencies and percentages.

**2.11 a.**  $31 + 78 + 49 + 81 + 117 + 13 = 369$  customers were served.

**b.** Yes, each class has a width of 4.

Gallons of Gas (Class Limits)	Class Width	Class Midpoint
0 to less than 4	4	$\frac{0+4}{2} = 2$
4 to less than 8	4	$\frac{4+8}{2} = 6$
8 to less than 12	4	$\frac{8+12}{2} = 10$
12 to less than 16	4	$\frac{12+16}{2} = 14$
16 to less than 20	4	$\frac{16+20}{2} = 18$
20 to less than 24	4	$\frac{20+24}{2} = 22$

**c.**

Gallons of Gas	Number of Customers	Relative Frequency	Percentage
0 to less than 4	31	$31/369 \approx 0.084$	8.4
4 to less than 8	78	$78/369 \approx 0.211$	21.1
8 to less than 12	49	$49/369 \approx 0.133$	13.3
12 to less than 16	81	$81/369 \approx 0.220$	22.0
16 to less than 20	117	$117/369 \approx 0.317$	31.7
20 to less than 24	13	$13/369 \approx 0.035$	3.5

**d.**  $22.0 + 31.7 + 3.5 = 57.2\%$  of the customers purchased 12 gallons or more.

**e.** The number of customers who purchased 10 gallons or less cannot be determined exactly because 10 is not a boundary value.

f.

Gallons of Gasoline	Cumulative Frequency	Cumulative Relative Frequency	Cumulative Percentage
0 to less than 4	31	$31/369 = 0.084$	8.4
0 to less than 8	$31 + 78 = 109$	$109/369 = 0.295$	29.5
0 to less than 12	$31 + 78 + 49 = 158$	$158/369 = 0.428$	42.8
0 to less than 16	$31 + 78 + 49 + 81 = 239$	$239/369 = 0.648$	64.8
0 to less than 20	$31 + 78 + 49 + 81 + 117 = 356$	$356/369 = 0.965$	96.5
0 to less than 24	$31 + 78 + 49 + 81 + 117 + 13 = 369$	1.000	100.0

2.13 a.

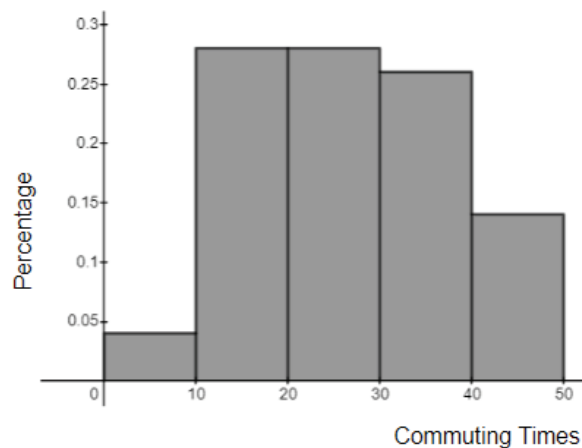
Commuting Times	Frequency
0 to 9	2
10 to 19	14
20 to 29	14
30 to 39	13
40 to 49	7

b.

Commuting Times	Frequency	Relative Frequency	Percentage
0 to 9	2	$2/50 = 0.04$	4
10 to 19	14	$14/50 = 0.28$	28
20 to 29	14	$14/50 = 0.28$	28
30 to 39	13	$13/50 = 0.26$	26
40 to 49	7	$7/50 = 0.14$	14

e.

Histogram of Commuting Times



**6** Chapter 2

**d.**  $26\% + 14\% = 40\%$  of the workers in the sample commute for 30 minutes or more.

**e.**

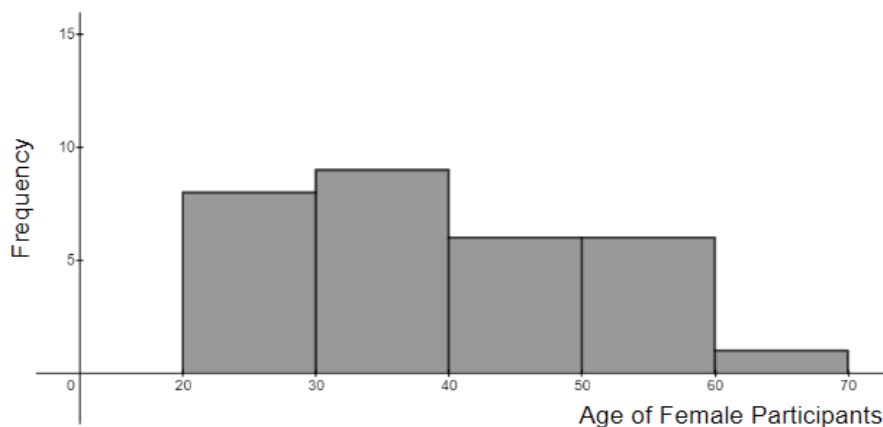
Commuting Times	Cumulative Frequency	Cumulative Relative Frequency	Cumulative Percentage
0 to 9	2	$2/50 = 0.04$	4
10 to 19	$2 + 14 = 16$	$16/50 = 0.32$	32
20 to 29	$2 + 14 + 14 = 30$	$30/50 = 0.60$	60
30 to 39	$2 + 14 + 14 + 13 = 43$	$43/50 = 0.86$	86
40 to 49	$2 + 14 + 14 + 13 + 7 = 50$	$50/50 = 1.00$	100

**2.15 a. and b.**

Age	Frequency	Relative Frequency	Percentage
20 to 29	8	$8/30 = 0.267$	26.7
30 to 39	9	$9/30 = 0.300$	30.0
40 to 49	6	$6/30 = 0.200$	20.0
50 to 59	6	$6/30 = 0.200$	20.0
60 to 69	1	$1/30 = 0.033$	3.30

**c.**

**Frequency Histogram of Age of Female Participants**



**d.**  $26.7\% + 30\% = 56.7\%$  of female participants are younger than 40 years of age.

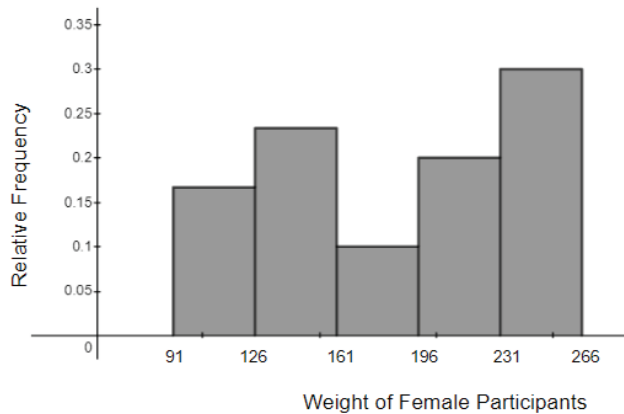
**e.** The male and female age distributions are fairly similar; however, male ages are more spread out and older than the female ages.

2.17 a. and b.

Weight	Frequency	Relative Frequency	Percentage
91 to 125	5	$5/30 = 0.167$	16.7
126 to 160	7	$7/30 = 0.233$	23.3
161 to 195	3	$3/30 = 0.100$	10.0
196 to 230	6	$6/30 = 0.200$	20.0
231 to 265	9	$9/30 = 0.300$	30.0

c.

Relative Frequency Histogram of Weight of Female Participants



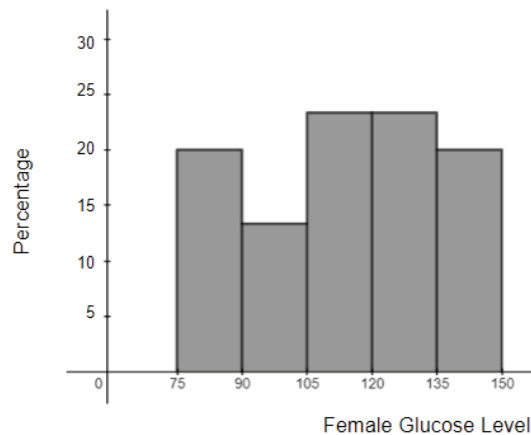
- d.  $16.7\% + 23.3\% = 40\%$  of female participants weighed less than 161 pounds.
- e. The weight distributions were similar, but 10% more of the females had a weight below 161 pounds than did the males. The percentage above 195 pounds was about the same in both distributions, but of these more females fell in the 231 to 265 pounds range than did males.

2.19 a. and b.

Blood Glucose Level	Frequency	Relative Frequency	Percentage
75 to 89	6	$6/30 = 0.200$	20.0
90 to 104	4	$4/30 = 0.133$	13.3
105 to 119	7	$7/30 = 0.233$	23.3
120 to 134	7	$7/30 = 0.233$	23.3
135 to 149	6	$6/30 = 0.200$	20.0

c.

Percentage Distribution Histogram for Female Glucose Level



- d.  $23.3\% + 20.0\% = 43.3\%$  of female participants had a blood glucose level higher than 119.
- e. Taking the center of both distributions to be the class 105 to 119, the shape of the left and right tails between the two distribution is swapped, meaning there are fewer females with blood glucose levels less than 105 than there are males with such glucose levels, and there are more females with blood glucose level greater than 119 than there are males with such blood glucose levels.
- f.

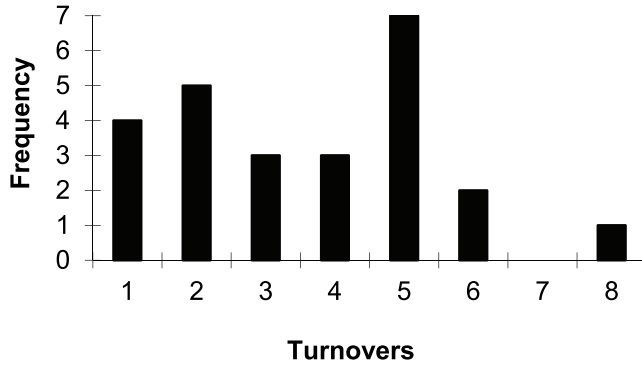
Blood Glucose Level	Cumulative Frequency	Cumulative Relative Frequency	Cumulative Percentage
75 to 89	6	$6/30 = 0.200$	20.0
75 to 104	$6 + 4 = 10$	$10/30 = 0.333$	33.3
75 to 119	$6 + 4 + 7 = 17$	$17/30 = 0.567$	56.7
75 to 134	$6 + 4 + 7 + 7 = 24$	$24/30 = 0.800$	80.0
75 to 149	$6 + 4 + 7 + 7 + 6 = 30$	$30/30 = 1.000$	100.0

2.21 a. and b.

Turnovers	Frequency	Relative Frequency	Percentage
1	4	$4/25 = 0.160$	16.0
2	5	$5/25 = 0.200$	20.0
3	3	$3/25 = 0.120$	12.0
4	3	$3/25 = 0.120$	12.0
5	7	$7/25 = 0.280$	28.0
6	2	$2/25 = 0.080$	8.0
7	0	$0/25 = 0.000$	0.0
8	1	$1/25 = 0.040$	4.0



- d.**



### Section 2.3

- 2.23** To prepare a **stem-and-leaf display** for a data set, each value is divided into two parts; the first part is called the *stem* and the second part is called the *leaf*. The stems are written on the left side of a vertical line and the leaves for each stem are written on the right side of the vertical line next to the corresponding stem. Example 2–9 is an example of a stem-and-leaf display.

- 2.25** The data that were used to make this stem-and-leaf display are: 218, 245, 256, 329, 367, 383, 397, 404, 427, 433, 471, 523, 537, 551, 563, 581, 592, 622, 636, 647, 655, 678, 689, 810, 841

- 2.27 a.**

[illegible][illegible]

**b.**

[illegible]

c. Answers will vary, but the split stem-and-leaf display seems to better discern differences in the data in the range 30 – 70.

2.29

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

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

**b.**

2-4	58	*	20	45	*	30	38	60	90										
5-6	05	30	38	50	60	65	70	75	*	10	17	20	35	38					
7-9	02	05	06	20	21	28	65	87	*	40	45	68	70	90	*	57	68		

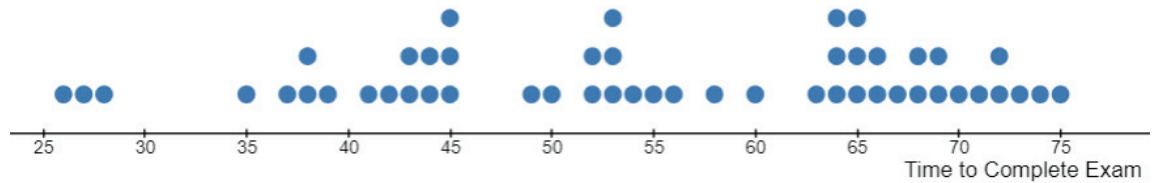
## Section 2.4

**2.31** In order to prepare a **dotplot**, first we draw a horizontal line with numbers that cover the given data set. Then we place a dot above the value on the number line that represents each measurement in the data set. Example 2–12 illustrates this procedure.

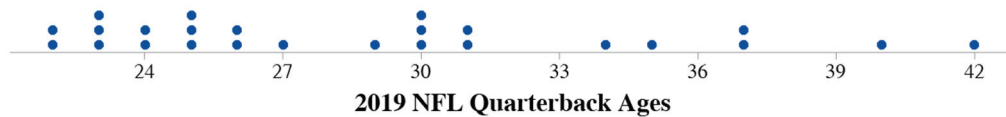
2.33



2.35



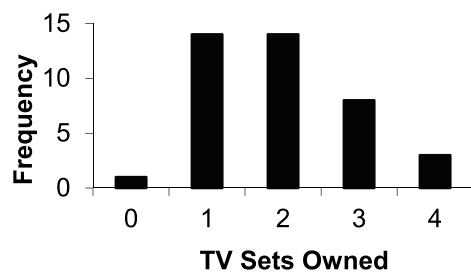
2.37

Supplementary Exercises

2.39 a. and b.

TV sets owned	Frequency	Relative Frequency	Percentage
0	1	$1/40 = 0.025$	2.5
1	14	$14/40 = 0.350$	35.0
2	14	$14/40 = 0.350$	35.0
3	8	$8/40 = 0.200$	20.0
4	3	$3/40 = 0.075$	7.5

c.

d.  $35.0\% + 20.0\% + 7.5\% = 62.5\%$  of the households own two or more television sets.

## 12 Chapter 2

### 2.41 a. and b.

Number of Orders	Frequency	Relative Frequency	Percentage
23 – 29	4	$4/30 = 0.133$	13.3
30 – 36	9	$9/30 = 0.300$	30.0
37 – 43	6	$6/30 = 0.200$	20.0
44 – 50	8	$8/30 = 0.267$	26.7
51 – 57	3	$3/30 = 0.100$	10.0

- c. For  $20.0\% + 26.7\% + 10.0\% = 56.7\%$  of the hours in this sample, the number of orders was more than 36.

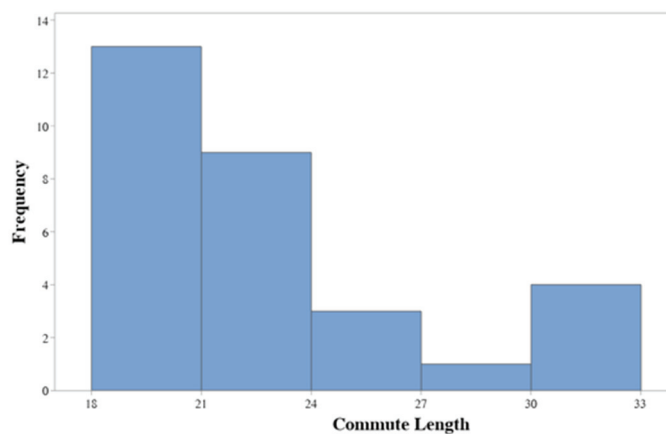
### d.

Number of Orders	Cumulative Frequency	Cumulative Relative Frequency	Cumulative Percentage
23 – 29	4	$4/30 = 0.133$	13.3
30 – 36	13	$13/30 = 0.433$	43.3
37 – 43	19	$19/30 = 0.633$	63.3
44 – 50	27	$27/30 = 0.900$	90.0
51 – 57	30	$3/30 = 1.000$	100.0

### 2.43 a. and b.

Commute Length (in minutes)	Frequency	Relative Frequency	Percentage
18 to less than 21	13	$13/30 = 0.433$	43.3
21 to less than 24	9	$9/30 = 0.300$	30.0
24 to less than 27	3	$3/30 = 0.100$	10.0
27 to less than 30	1	$1/30 = 0.033$	3.3
30 to less than 33	4	$4/30 = 0.133$	13.3

### c.



d.

Commute Length (in minutes)	Cumulative Frequency	Cumulative Relative Frequency	Cumulative Percentage
18 to less than 21	13	$13/30 = 0.433$	43.3
21 to less than 24	22	$22/30 = 0.733$	73.7
24 to less than 27	25	$25/30 = 0.833$	83.3
27 to less than 30	26	$26/30 = 0.867$	86.7
30 to less than 33	30	$30/30 = 1.000$	100

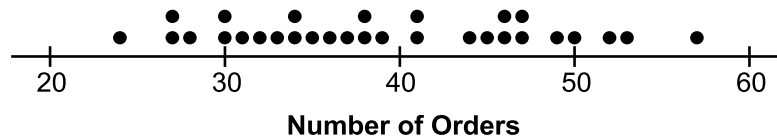
2.45

```

2 | 8 4 7 7
3 | 4 1 8 5 2 9 3 7 0 8 4 6 0
4 | 4 1 7 6 1 9 5 6
5 | 2 3 7 0

```

2.47



2.49 a. Using Sturge's formula:

$$\begin{aligned}
 c &= 1 + 3.3 \log n = 1 + 3.3 \log 135 \\
 &= 1 + 3.3(2.13033377) \\
 &= 1 + 7.03 = 8.03 \approx 8
 \end{aligned}$$

b. Approximate class width =  $\frac{\text{Largest value} - \text{smallest value}}{\text{Number of classes}} = \frac{53 - 20}{8} = 4.125$ .

Use a class width of 5.

2.51 a. Answers will vary.

b. i.

```

9 | 9
10 | 2 8 8
11 | 0 4 5 5 6 9
12 | 3 3 3 5 8 8
13 | 2 3 8
14 | 6 7 7 8
15 | 5 9
16 | 1 2 4 8
17 | 4 4 5 9 9 9
18 | 0 2 3 9
19 | 3 3 5
20 | 2 4

```

- ii. The display shows a bimodal distribution, due to the presence of both females and males in the sample. The males tend to be heavier, so their weights are concentrated in the larger values, while the females' weights are found primarily in the smaller values.

c.

Females						Males					
					9	9					
			8	8	2	10					
9	6	5	5	4	0	11					
8	8	5	3	3	3	12					
					3	13	2	8			
			8	7	6	14	7				
					5	15	9				
					4	16	1	2	8		
						17	4	4	5	9	9
						18	0	2	3	9	
						19	3	3	5		
						20	2	4			

Males are heavier than females. Females concentrate around smaller stems, but males concentrate among higher stems.

### Self-Review Test

1. An **ungrouped data set** contains information on each member of a sample or population individually. The first part of Example 2–1 in the text, listing the responses of each of the 30 employees, is an example of ungrouped data. Data presented in the form of a frequency table are called **grouped data**. Table 2.4 in the solution of Example 2–1 is an example of grouped data.
3. A histogram that is identical on both sides of its central point is called a **symmetric histogram**. A histogram that is **skewed to the right** has a longer tail on the right side, and a histogram that is **skewed to the left** has a longer tail on the left side. Figure 2.9 in the text provides graphs of symmetric histograms, Figure 2.10a displays a histogram skewed to the right, and Figure 2.10b displays a histogram that is skewed to the left.

5 a.

Monthly Expense on Gas (in dollars)	Frequency
50 to 149	9
150 to 249	13
250 to 349	11
350 to 449	9
450 to 549	6

Monthly Expense on Gas (in dollars)	Relative Frequency	Percentage
50 to 149	$9/48 = 0.188$	18.8
150 to 249	$13/48 = 0.271$	27.1
250 to 349	$11/48 = 0.229$	22.9
350 to 449	$9/48 = 0.188$	18.8
450 to 549	$6/48 = 0.125$	12.5

A histogram showing the distribution of monthly gas expenses. The x-axis is labeled 'Monthly Expense on Gas (in \$)' and ranges from 0 to 550 with major ticks every 50 units. The y-axis is labeled 'Percentage' and ranges from 0 to 0.3 with major ticks every 0.1 units. The histogram consists of five bars of equal width (100 units each). The approximate percentages for each bin are: 50-150: 0.19, 150-250: 0.27, 250-350: 0.23, 350-450: 0.19, and 450-550: 0.13.

Monthly Expense Range (in \$)	Percentage
50 - 150	0.19
150 - 250	0.27
250 - 350	0.23
350 - 450	0.19
450 - 550	0.13

Monthly Expense on Gas	Cumulative Frequency	Cumulative Relative Frequency	Cumulative Percentage
50 to 149	9	$9/48 = 0.188$	18.8
150 to 249	$9 + 13 = 22$	$22/48 = 0.458$	45.8
250 to 349	$9 + 13 + 11 = 33$	$33/48 = 0.688$	68.8
350 to 449	$9 + 13 + 11 + 9 = 42$	$42/48 = 0.875$	87.5
450 to 549	$9 + 13 + 11 + 9 + 6 = 48$	$48/48 = 1.000$	100.0

[illegible]

9.

