## Collecting, Displaying, and Analyzing Data

## Solutions Key

## ARE YOU READY?

1. $\frac{3+5+4+7}{4}=\frac{19}{4}$
$=4 \frac{3}{4}$
$=4.75$
2. $\frac{83+88}{2}=\frac{171}{2}$
$=85 \frac{1}{2}$
$=85.5$
3. $\frac{2.2+1.7+1.8}{3}=\frac{5.7}{3}$
4. $15,18,23,42,45,65$
$=1.9$
5. $87,102,103,104,105,118$
6. 19, 24, 33, 56, 65, 76, 82
7. $2,2,3,3,4,5,6,8,9$
8. $18+26=44$
9. $23+17=40$
10. $75+37=112$
11. $98+64=162$
12. $133-35=98$
13. $54-29=25$
14. $200-88=112$
15. $1,055-899=156$
16. Iion
17. zebra

## LESSON 1

## Think and Discuss

1. Possible answer: A mean can be used when determining grades from several scored assignments.
2. mode
3. An outlier usually has the most effect on the mean. The median may not be affected as much because it is the data value in the middle position. The mode will not be affected at all.

## Exercises

1. Data set: $5,30,35,20,5,25,20$
$5+30+35+20+5+25+20=140$
$140 \div 7=20$
The mean is 20 .
Arrange the numbers in order: 5, 5, 20, 20, 25, 30, 35
The median is 20 .
The numbers that appear most often are 5 and 20.
The modes are 5 and 20.
greatest value - least value $=35-5=30$
The range is 30 .
2. Data set: $44,68,48,61,59,48,63,49$
$44+68+48+61+59+48+63+49=440$
$440 \div 8=55$
The mean is 55 .
Arrange the numbers in order: 44, 48, 48, 49, 59, 61, 63, 68
Determine the average of the middle two numbers.
$49+59=108$
$108 \div 2=54$
The median is 54 .
The number that appears most often is 48.
The mode is 48 .
greatest value - least value $=68-44=24$
The range is 24 .
3. Median; There is an outlier that skews the data.
4. Data set: 4, 12, 3, 1, 4, 7, 5, 4

The outlier is 12 .
Without the outlier:
$4+3+1+4+7+5+4=28$
$28 \div 7=4$
The mean is 4 .
Arrange the numbers in order: 1, 3, 4, 4, 4, 5, 7
The median is 4 .
The number that appears most often is 4 .
The mode is 4.
With the outlier:
$4+12+3+1+4+7+5+4=40$
$40 \div 8=5$
The mean is 5 .
Arrange the numbers in order: 1, 3, 4, 4, 4, 5, 7, 12
Determine the average of the middle two numbers.
$4+4=8$
$8 \div 2=4$
The median is 4 .
The number that appears most often is 4 .
The mode is 4.
12; Adding the outlier increased the mean by 1 . The median and mode did not change.
The mean best describes the data with outlier. The median best describes the data without the outlier.
5. Data set: 92, 88, 65, 68, 76, 90, 84, 88, 93, 89
$92+88+65+68+76+90+84+88+$
$93+89=833$
$833 \div 10=83.3$
The mean is 83.3.
Arrange the numbers in order: 65, 68, 76, 84, 88,
88, 89, 90, 92, 93
Determine the average of the middle two numbers.
$88+88=176$
$176 \div 2=88$
The median is 88 .
The number that appears most often is 88 .
The mode is 88 .
greatest value - least value $=93-65=28$
The range is 28 .
6. Data set: $23,43,5,3,4,14,24,15,15,13$
$23+43+5+3+4+14+24+15+15+$
$13=159$
$159 \div 10=15.9$
The mean is 15.9 .
Arrange the numbers in order: $3,4,5,13,14,15$, 15, 23, 24, 43
Determine the average of the middle two numbers.
$14+15=29$
$29 \div 2=14.5$
The median is 14.5 .
The number that appears most often is 15 .
The mode is 15 .
$43-3=40$ is the difference between the greatest and least values. The range is 40 .
7. Data set: 2.0, 4.4, 6.2, 3.2, 4.4, 6.2, 3.7
$2.0+4.4+6.2+3.2+4.4+6.2+3.7=30.1$
$30.1 \div 7=4.3$
The mean is 4.3.
Arrange the numbers in order: 2.0, 3.2, 3.7, 4.4,
4.4, 6.2, 6.2

The median is 4.4.
The numbers that appear most often are 4.4 and 6.2.
The modes are 4.4 and 6.2.
greatest value - least value $=6.2-2.0=4.2$
The range is 4.2 .
8. Data set: $13.1,7.5,3.9,4.8,17.1,14.6,8.3,3.9$
$13.1+7.5+3.9+4.8+17.1+14.6+8.3+$ $3.9=73.2$
$73.2 \div 8=9.15$
The mean is 9.15 .
Arrange the numbers in order: 3.9, 3.9, 4.8, 7.5, 8.3, 13.1, 14.6, 17.1

Determine the average of the middle two numbers.
$7.5+8.3=15.8$
$15.8 \div 2=7.9$
The median is 7.9.
The number that appears most often is 3.9.
The mode is 3.9 .
greatest value - least value $=17.1-3.9=13.2$
The range is 13.2 .
9. Mean and median; The mode focuses on one data value and does not reflect the data set as a whole.
10. Data set: $13,18,20,5,15,20,13,20$

The outlier is 5 .

## Without the outlier

$13+18+20+15+20+13+20=119$
$119 \div 7=17$
The mean is 17 .
Arrange the numbers in order: $13,13,15,18,20$, 20, 20
The median is 18 .
The number that appears most often is 20 .
The mode is 20 .
With the outlier
$13+18+20+5+15+20+13+20=124$
$124 \div 8=15.5$
The mean is 15.5 .
Arrange the numbers in order: 5, 13, 13, 15, 18, 20, 20, 20
Determine the average of the middle two numbers.
$15+18=33$
$33 \div 2=16.5$
The median is 16.5 .
The number that appears most often is 20.
The mode is 20 .
5 ; Adding the outlier decreased the mean and the median by 1.5. The mode did not change; mean; median.
11. Data set: $45,48,63,85,151,47,88,44,68$

The outlier is 151.
Without the outlier
$45+48+63+85+47+88+44+68=488$
$488 \div 8=61$
The mean is 61 .

Arrange the numbers in order: 44, 45, 47, 48, 63, 68, 85, 88
Determine the average of the middle two numbers.
$48+63=111$
$111 \div 2=55.5$
The median is 55.5 .
No number appears more often than the others.
There is no mode.
With the outlier
$45+48+63+85+151+47+88+44+68$
$=639$
$639 \div 9=71$
The mean is 71 .
Arrange the numbers in order: 44, 45, 47, 48, 63, 68, 85, 88, 151
Determine the average of the middle two numbers.
The median is 63 .
No number appears more often than the others.
There is no mode.
151; Adding the outlier increased the mean by
10 and the median by 7.5 . The mode did not change because there was no mode; median; mean.
12. Because the mean of three data items is 62 , the sum of the three data items must be 186.
$3 \cdot 62=186$.
Two of the data items are 58 and 61. Subtract the sum of the two known data items from 186 to find the missing data item.
$58+61=119$
$186-119=67$
Jon's height at the third checkup was 67 in.
13. Data set: $2,4,5,5,7,8,8,10,11,12,12,12,21$
$2+4+5+5+7+8+8+10+11+12+$
$12+12+21=117$
$117 \div 13=9$
The mean is 9 .
Arrange the numbers in order: $2,4,5,5,7,8,8,10$, 11, 12, 12, 12, 21
The median is 8 .
The number that appears most often is 12 .
The mode is 12 .
The outlier is 21 .
Without the outlier
$2+4+5+5+7+8+8+10+11+12+$
$12+12=96$
$96 \div 12=8$
The mean is 8 .
With the outlier
$2+4+5+5+7+8+8+10+11+12+$
$12+12+21=117$
$117 \div 13=9$
The mean is 9 .
$9 ; 8 ; 12$; adding the outlier increased the mean by 1.
14. Data set: $95,93,91,95,100,99,92$
$95+93+91+95+100+99+92=665$
$665 \div 7=95$
The mean is 95 .
Arrange the numbers in order: 91, 92, 93, 95, 95, 99, 100
The median is 95 .

The number that appears most often is 95 .
The mode is 95 .
The mean, median, and mode of the original data set is 95 . Since all three measures of central tendencies are the same value, adding that value to the data set will not change any of the original central tendencies.
The added value is 95 .
15. Possible answer: The mean or the median, since most of the participants were in their twenties.
16. Possible answers: When estimating this data you can use the whole number values given in the table.
$0+1+1+2+3+4=11$
$11 \div 6 \approx 2$
The mean is approximately 2.
The middle two numbers when the numbers are arranged in order are 1 and 2.
The median is approximately 2.
greatest value - least value $=4-1=3$
The range is approximately 3 .
17. Data set: $10,7,9,5,13,10,7,14,8,11$
$10+7+9+5+13+10+7+14+8+11=94$
$94 \div 10=9.4$
The mean is 9.4.
Arrange the numbers in order: 5, 7, 7, 8, 9, 10, 10, 11, 13, 14
Determine the average of the middle two numbers.
$9+10=19$
$19 \div 2=9.5$
The median is 9.5 .
The numbers that appear most often are 7 and 10. The modes are 7 and 10.
The only measure of central tendency that has a value of 9.5 is the median.
Answer: What is the median of the data set?
18. Possible answer: The mean is most often affected by including an outlier. Often the median and mode will not change, but the mean will always change.
19. a. Median; Use when there are outliers that may distort the data or to describe the middle value.
b. Mean; use when there are no outliers to distort the data.
20. B; Data set: $272,276,281,279,276$
$272+276+281+279+276=1,384$
$1,384 \div 5=276.8$
The mean is 276.8.
21. J; Data set: 4, 3, 4, 3, 4, 6, 4
$4+3+4+3+4+6+4=28$
$28 \div 7=4$
The mean is 4 .
Arrange the numbers in order: 3, 3, 4, 4, 4, 4, 6
The median is 4 .
The number that appears most often is 4 .
The mode is 4.
Data set J has the mean, median, and mode of 4 .

## LESSON 2

## Think and Discuss

1. Possible answer: You can tell the range and median of the data set, as well as how the data is distributed around the median.
2. Possible answer: The range of a set of data is the difference between the greatest and least values. The interquartile range is the difference between the upper and lower quartiles. The interquartile range tells how large the spread of data around the median is.

## Exercises

## 

Data set: $46,35,46,38,37,33,49,42,35,40,37$
Step 1: Find the lower and upper extremes, the median, and the first and third quartiles. Order the data from least to greatest. $33,35,35,37,37,38,40,42,46,46,49$
Find the lower and upper extremes.
(33.) $35,35,37,37,38,40,42,46,46$, 49)

Find the median.
$33,35,35,37,37,38 ., 40,42,46,46,49$
Find the first and third quartiles.

$$
33,35,35,37,37,38,40,42,46,46,49
$$

Step 2: Draw a number line. Above the number line, plot points representing the lower and upper extremes, the median, and the first and third quartiles.
The range is 16 , the interquartile range is 11 , the lower quartile is 35 , and the upper quartile is 46 .
2. airplane $A$
3. airplane $B$
4. airplane A
5.


Data set: $81,73,88,85,81,72,86,72,79,75,76$
Step 1: Find the lower and upper extremes, the median, and the first and third quartiles.
Order the data from least to greatest.
$72,72,73,75,76,79,81,81,85,86,88$
Find the lower and upper extremes.
72.) $72,73,75,76,79,81,81,85,86,88$

Find the median.
$72,72,73,75,76,79.81,81,85,86,88$
Find the first and third quartiles.

$$
33,35,35,37,37,38,40,42,46,46,49
$$

Step 2: Draw a number line. Above the number line, Plot points representing the lower and upper extremes, the median, and the first and third quartiles.
The range is 16 , the interquartile range is 12 , the lower quartile is 73 , and the upper quartile is 85 .
6. city A
7. city A
8. city $B$
9.


Data set: $12,7,15,23,10,18,39,15,20,8,13$
Step 1: Find the lower and upper extremes, the median, and the first and third quartiles.
Order the data from least to greatest.
$7,8,10,12,13,15,15,18,20,23,39$
Find the lower and upper extremes.
7, 8. $10,12,13,15,15,18,20,23$, 39)
Find the median.
$7,8,10,12,13,15.) 15,18,20,23,39$
Find the first and third quartiles.
$7,8,10,12.13,15,15,18,20$, 23., 39
Step 2: Draw a number line. Above the number line, plot points representing the lower and upper extremes, the median, and the first and third quartiles.
The outlier is 39. In the following steps, 39 is not included in the data.
Step 1: Find the lower and upper extremes, the median, and the first and third quartiles.
Order the data from least to greatest.
$7,8,10,12,13,15,15,18,20,23$
Find the lower and upper extremes.
(7, 8. 10, 12, 13, 15, 15, 18, 20, (23)
Find the median.
$7,8,10,12,13,15,15,18,20,23$
Median $=\frac{13+15}{2}=14$
Find the first and third quartiles.
$7,8,10.212,13,15,15,18,20,23$
Step 2: Above the same number line, plot points representing the lower and upper extremes, the median, and the first and third quartiles.
10. The interquartile range increases with the outlier present.
11. The range because it is either the lower or upper extreme.
12.


Data set: 42, 42, 44, 45, 45, 45, 47, 47, 48, 48, 48, $49,49,50,50,50,50,51,51,52,54,56$
Step 1: Find the lower and upper extremes, the median, and the first and third quartiles.
Order the data from least to greatest.
$42,42,44,45,45,45,47,47,48,48,48,49,49$, 50, 50, 50, 50, 51, 51, 52, 54, 56
Find the lower and upper extremes.
42. $42,44,45,45,45,47,47,48,48,48,49,49$, $50,50,50,50,51,51,52,54,56$
Find the median.
42, 42, 44, 45, 45, 45, 47, 47, 48, 48, 48, 49. 49, $50,50,50,50,51,51,52,54,56$
The median is 48.5

Find the first and third quartiles.
$42,42,44,45,45,45,47,47,48,48,48$,
$49,49,50,50,50,50,51,51,52,54,56$
Step 2: Draw a number line. Above the number line, plot points representing the lower and upper extremes, the median, and the first and third quartiles.
13. a.


2030405060708090100
Data set: 103, 92, 63, 49, 48, 37, 33, 32, 30, 30, $27,23,22,19,19$
Step 1: Find the lower and upper extremes, the median, and the first and third quartiles.
Order the data from least to greatest.
$19,19,22,23,27,30,30,32,33,37,48,49,63$, 92, 103
Find the lower and upper extremes.
(19.) 19, 22, 23, 27, 30, 30, 32, 33, 37, 48, 49, 63, 92, 103
Find the median.
$19,19,22,23,27,30,30,32.33,37,48,49,63$, 92, 103
Find the first and third quartiles.
19, 19, 22, 23., 27, 30, 30, 32,
33, 37, 48, 49), 63, 92, 103
Step 2: Draw a number line. Above the number line, plot points representing the lower and upper extremes, the median, and the first and third quartiles.
b. Possible answer: Most countries won between 23 and 49 medals.
14. a.


Data set: $53,55,56,56,58,58,58,59,59,60$, $60,61,61,61,61,61,62,62,62,64$
Step 1: Find the lower and upper extremes, the median, and the first and third quartiles.
Order the data from least to greatest.
$53,55,56,56,58,58,58,59,59,60,60,61,61$, 61, 61, 61, 62, 62, 62, 64
Find the lower and upper extremes.
53. $55,56,56,58,58,58,59,59,60,60,61,61$, $61,61,61,62,62,62,64$
Find the median.
$53,55,56,56,58,58,58,59,59,60,60,61,61$, 61, 61, 61, 62, 62, 62, 64
The median is 60
Find the first and third quartiles.
$53,55,56,56,58,58,58,59,59,60$,
60, 61, 61, 61, 61, 61, 62, 62, 62, 64

Step 2: Draw a number line. Above the number line, plot points representing the lower and upper extremes, the median, and the first and third quartiles.
b. 58 in.
c. 61 in.
15. The student made an error in finding the median of the upper half of the data. The upper quartile is 11 .
16. The plot with the larger box represents a greater range of numbers. To see why, draw two box-andwhisker plots with the same median, equally long whiskers, and boxes of different lengths.

17. Data set: $1,2,4,2,1,0,6,8,1,6,2$

Step 1: Find the lower and upper extremes, the median, and the first and third quartiles.
Order the data from least to greatest.
$0,1,1,1,2,2,2,4,6,6,8$
Find the lower and upper extremes.
(0.) $1,1,1,2,2,2,4,6,6,8)$

Find the median.
$0,1,1,1,2$, (2.) $2,4,6,6,8$
Find the first and third quartiles.
0, 1, (1.) 1, 2, 2, 2, 4, 6. 6, 8
Step 2: Draw a number line. Above the number line, plot points representing the lower and upper extremes, the median, and the first and third quartiles.
The interquartile range is 5 , so an outlier is
5 times 1.5.
$5 \cdot 1.5=7.5$
Since 8 is larger than $7.5,8$ is considered an outlier.
18. D ; The interquartile range for the top box-and-whisker plot is $24-15=9$. The interquartlie range for the bottom box-and-whisker plot is $21-12=9$.
19. 15; The top box-and-whisker plot has the greatest range, so its lower quartile is 15.

## LESSON 3

## Think and Discuss

1. Possible answer: You would use a sample to survey voters before a national election since it is not possible to survey the entire population.
2. Possible answer: It may not be possible for every member of the population to have an equal chance of being chosen.

## Exercises

1. Daria's method is best; It uses a random sample, and Nadia's method uses a convenience sample.
2. The sample is not biased; It is a random sample.
3. The sample is biased; All city residents do not eat at a single restaurant.
4. The claim is not true because the data shows that about 1,200 defective light bulbs are produced each day.
5. Vonetta's method is best; It uses a random sample, and Suzanne's method uses a convenience sample.
6. The sample is biased; Listeners who call the station are more likely to enjoy the music the station plays.
7. The sample is not biased; It is a random sample.
8. Let $x$ represent the students who speak three or more languages.

$$
\begin{aligned}
\frac{30,600}{x} & =\frac{240}{20} \\
240 x & =612,000 \\
\frac{240 x}{240} & =\frac{612,000}{240} \\
x & =2,550
\end{aligned}
$$

2,550 students speak three or more languages.
9. Survey the entire population because the population is relatively small.
10. Use a sample because the population is too large to survey.
11. Survey the entire population because the population is relatively small.
12. Agree. Based on data from the sample, there are 30 fruit flies with deformed wings.
13. The question only asks about green; so, it is biased. An unbiased question is: What is your favorite color?
14. People with unlisted numbers cannot be surveyed.
15. Possible answer: Survey 25 students whose names are randomly chosen from a list of all seventh graders.
16. No; If the manager had chosen a random sample of 200 employees, only 6 or 7 of them would walk to work.
17. B; let $x$ represent the students who have pet dogs.

| $\frac{580}{x}$ | $=\frac{30}{12}$ |
| ---: | :--- |
| $30 x$ | $=6,960$ |
| $\frac{30 x}{30}$ | $=\frac{6,960}{30}$ |
| $x$ | $=232$ |

18. A survey of the listeners of a sports radio program asking whether they read the sports page. Listeners to a sports radio program are likely to read the sports pages.

## READY TO GO ON?

1. $20,000+18,000+14,000+24,000+10,000$

$$
=86,000
$$

$86,000 \div 5=17,200$
The mean is $\$ 17,200$.
2. The number that appears most often is 30 .

The numbers in the table are written in thousands.
The mode is 30,000 miles.
3. Data set: $18,22,5,21,19,21,17,3,19,20,29$, 18, 17
$18+22+5+21+19+21+17+3+19+$ $20+29+18+17=229$
$229 \div 13 \approx 17.6$
The mean is about 17.6.
Arrange the numbers in order: $3,5,17,17,18,18$, 19, 19, 20, 21, 21, 22, 29
Determine the average of the middle two numbers.
The median is 19.
The numbers that appear most often are 17, 18, 19, and 21.
The modes are 17, 18, 19, and 21.
greatest value - least value $=29-3=26$
The range is 26 .
4. Median; The outliers affect the mean, so the median is best.
5. Data set: $14,8,13,20,15,17,1,12,18,10$

Step 1: Find the lower and upper extremes, the median, and the first and third quartiles.
Order the data from least to greatest.
$1,8,10,12,13,14,15,17,18,20$
Find the lower and upper extremes.
(1.) $8,10,12,13,14,15,17,18$, (20)

Find the median.
$1,8,10,12,13,14.15,17,18,20$
Median $=\frac{13+14}{2}=13.5$
Find the first and third quartiles.
$1,8,10,12,13,14,15,17.18,20$
Step 2: Above the same number line, plot points representing the lower and upper extremes, the median, and the first and third quartiles.

6. Data set: $3,8,5,12,6,18,14,8,15,11$

Step 1: Find the lower and upper extremes, the median, and the first and third quartiles.
Order the data from least to greatest.
$3,5,6,8,8,11,12,14,15,18$
Find the lower and upper extremes.
(3.) $5,6,8,8,11,12,14,15,18$

Find the median.
$3,5,6,8,8,11.12,14,15,18$
Median $=\frac{8+11}{2}=9.5$
Find the first and third quartiles.
3, 5, 6.) 8, 8, 11, 12, 14. 15, 18
Step 2: Above the same number line, plot points representing the lower and upper extremes, the median, and the first and third quartiles.

7. The plot of the second data set has a greater interquartile range.
8. Biased; It may not be likely that families with grown children or no children would attend an amusement park with their immediate families.
9. Not biased; The sample is random.
10. The claim is likely to be true; The data shows that about 1,875 fish are in the quarry.

## STUDY GUIDE: REVIEW

1. population; sample
2. outlier
3. mean
4. lower quartile
5. Mean: $324+233+324+399+233+299=$
$1812 \div 6=302$
Median: 233, 233, 299, 324, 324, 399

$$
(299+324) \div 2=311.5
$$

Mode: 233 and 324
Range: $399-233=166$
6. Mean: $48+39+27+52+45+47+49+37=$ $344 \div 8=43$
Median: 27, 37, 39, 45, 47, 48, 49, 52

$$
(45+47) \div 2=46
$$

Mode: none
Range: 25
7. The median is the most useful when the data set has an outlier.
8. Arrange the numbers in order: $272,272,274$, 278, 290
The middle number is 274 .
The median is 274 .
The number that appears most often is 272.
The mode is 272 .
9. $280+278+276+280+279=1,393$ $1,393 \div 5=278.6$
The mean winning score at the U.S. Open for men is 278.6.
10. greatest value - least value $=280-276=4$ The range winning scores for men is 4 . greatest value - least value $=290-272=18$ The range winning scores for women is 18.
11.

12. $42-29=13$
13. 2,500 is a reasonable estimate based on the data.
14. This is not biased because the people were walking down the street and chosen randomly.
15. This is biased because teenagers coming out of a clothing store likely like to buy the type of clothes sold by that store and teenagers who like to buy other types of clothes would not be surveyed.
16. You would use a sample to survey the seniors in your state since it is not possible to survey the entire population.
17. You would use the population to survey the members of the tennis club since it is possible to survey the entire population.
18. You would use a sample to survey some customers since it is not possible to survey the entire population.

## CHAPTER TEST

1. Mean: 22.2; median: 22.5; mode: 12; range: 26
2. The outlier would increase the mean and the median.
3. 

|  | Frequency | Cumulative <br> Frequency |
| :---: | :---: | :---: |
| $10-19$ | 4 | 4 |
| $20-29$ | 4 | 8 |
| $30-39$ | 2 | 10 |

4. 

| Stem | Leaves |
| ---: | :--- |
| 1 | 0228 |
| 2 | 2389 |
| 3 | 26 |

Key: $3 \mid 2$ means 32
5.

6.

7.

8. 18.5
9. Divide the number of bus riders by the number of students surveyed, then multiply by the total number of students in the school, 314.
10. $\frac{4}{5}=\frac{x}{314}$ $1,256=5 x$ $251.2=x$
About 251 students participate in after-school activities.
11. $21.0+20.7+21.2+21.6=84.5$ $84.5 \div 4=21.125$
The mean fuel rate is 21.125 miles per gallon.
12. $21.2+21.6+42.8$ $42.8 \div 2=21.4$
The mean fuel rate for 2005 is 21.4 miles per gallon.
© Houghton Mifflin Harcourt Publishing Company

