10.5 Interpret Box-and-Whisker Plots

Before	You made stem-and-leaf plots and histograms.
Now	You will make and interpret box-and-whisker plots.
Why?	So you can compare sets of scientific data, as in Ex. 19.

Key Vocabulary
- box-and-whisker plot
- quartile
- interquartile range
- outlier

COMMON CORE

CC.9-12.S.ID.1 Represent data with plots on the real number line (dot plots, histograms, and box plots).*

A **box-and-whisker plot** organizes data values into four groups. Ordered data are divided into lower and upper halves by the median. The median of the lower half is the **lower quartile**. The median of the upper half is the **upper quartile**.

EXAMPLE 1 Make a box-and-whisker plot

SONG LENGTHS The lengths of songs (in seconds) on a CD are listed below. Make a box-and-whisker plot of the song lengths.

173, 206, 179, 257, 198, 251, 239, 246, 295, 181, 261

Solution

STEP 1 **Order** the data. Then find the median and the quartiles.

Lower half   Median   Upper half

173, 179, **181**, 198, 206, **239**, 246, 251, **257**, 261, 295

Lower quartile   Upper quartile

STEP 2 **Plot** the median, the quartiles, the maximum value, and the minimum value below a number line.

STEP 3 **Draw** a box from the lower quartile to the upper quartile. Draw a vertical line through the median. Draw a line segment (a "whisker") from the box to the maximum and another from the box to the minimum.

Animated Algebra at my.hrw.com

✓ **GUIDED PRACTICE** for Example 1

1. Make a box-and-whisker plot of the ages of eight family members:
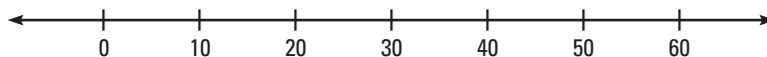   60, 15, 25, 20, 55, 70, 40, 30.

DLR/NASA

INTERPRET A BOX-AND-WHISKER PLOT A box-and-whisker plot separates data into four groups: the two parts of the box and the two whiskers. Each part contains approximately the same number of data values.
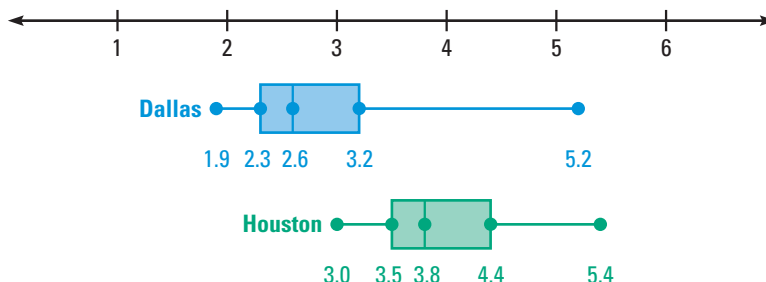
Each whisker represents about 25% of the data.

The box on each side of the median represents about 25% of the data.

You know that the range of a data set is the difference of the maximum value and the minimum value. The **interquartile range** of a data set is the difference of the upper quartile and the lower quartile.

EXAMPLE 2 **Interpret a box-and-whisker plot**

PRECIPITATION The box-and-whisker plots below show the normal precipitation (in inches) each month in Dallas and in Houston, Texas.

a. For how many months is Houston's precipitation less than 3.5 inches?

b. Compare the precipitation in Dallas with the precipitation in Houston.

Solution

a. For Houston, the lower quartile is 3.5. A whisker represents 25% of the data, so for 25% of 12 months, or 3 months, Houston has less than 3.5 inches of precipitation.

b. The median precipitation for a month in Dallas is 2.6 inches. The median for Houston is 3.8 inches. In general, Houston has more precipitation.

For Dallas, the interquartile range is $3.2 - 2.3$, or 0.9 inch. For Houston, the interquartile range is $4.4 - 3.5 = 0.9$ inch. So, the cities have the same variation in the middle 50% of the data. The range for Dallas is greater than the range for Houson. When all the data are considered, Dallas has more variation in precipitation.

✓ **GUIDED PRACTICE** **for Example 2**

2. PRECIPITATION In Example 2, for how many months was the precipitation in Dallas more than 2.6 inches?

OUTLIERS A value that is widely separated from the rest of the data in a data set is called an **outlier**. Typically, a data value is considered to be an outlier if it is greater than the upper quartile by more than 1.5 times the interquartile range or if it is less than the lower quartile by more than 1.5 times the interquartile range.

★ **EXAMPLE 3** **Standardized Test Practice**

> The normal monthly amounts of precipitation (in inches) in Dallas are: 1.9, 2.4, 3.1, 3.2, 5.2, 3.2, 2.1, 2.0, 2.4, 4.1, 2.6, 2.6. These data were used to create the box-and-whisker plot in Example 2. Which value, if any, is an outlier?
>
> **A** 1.9 **B** 5.2 **C** 1.9 and 5.2 **D** No outlier

Solution

From Example 2, you know the interquartile range of the data is 0.9 inch. Find 1.5 times the interquartile range: $1.5(0.9) = 1.35$.

From Example 2, you also know that the lower quartile is 2.3 and the upper quartile is 3.2. A value less than $2.3 - 1.35 = 0.95$ is an outlier. A value greater than $3.2 + 1.35 = 4.55$, is an outlier. Notice that $5.2 > 4.55$.

▶ The correct answer is B. **A** **B** **C** **D**

✓ **GUIDED PRACTICE** for Example 3

3. Which value, if any, is an outlier in the data set?

3.7, 3.0, 3.4, 3.6, 5.2, 5.4, 3.2, 3.8, 4.3, 4.5, 4.2, 3.7

A 3.0 **B** 5.4 **C** 3.0 and 5.4 **D** No outlier

10.5 EXERCISES

HOMEWORK KEY
○ = See **WORKED-OUT SOLUTIONS**
   Exs. 3 and 17
★ = **STANDARDIZED TEST PRACTICE**
   Exs. 2, 8, 9, 18, and 19

SKILL PRACTICE

1. VOCABULARY What is the interquartile range of a data set?

2. ★ **WRITING** *Explain* how you can identify an outlier in a data set.
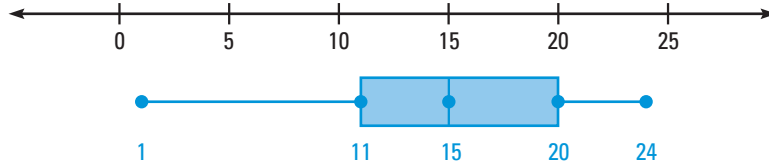
EXAMPLE 1
for Exs. 3–7

BOX-AND-WHISKER PLOTS Make a box-and-whisker plot of the data.

3. 1, 7, 0, 7, 2, 6, 3, 6, 0, 7, 8

4. 10, 1, 7, 5, 1, 8, 5, 4, 6, 5, 9, 12

5. 52, 20, 24, 45, 35, 32, 39, 42, 23, 64

6. 0.8, 0.4, 0.3, 0.6, 0.7, 0.2, 0.7, 0.9

7. ERROR ANALYSIS *Describe* and correct the error in creating a box-and-whisker plot of the data 0, 2, 4, 0, 6, 10, 8, 12, 5.
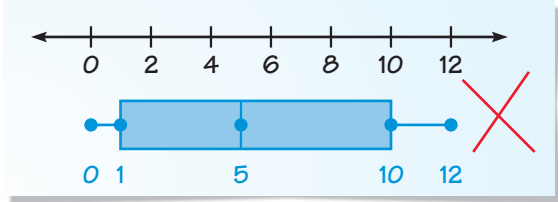
BOX-AND-WHISKER PLOT In Exercises 8–10, use the box-and-whisker plot.

8. ★ MULTIPLE CHOICE About what percent of the data are greater than 20?

    A 25%   **B** 50%   **C** 75%   **D** 100%

9. ★ MULTIPLE CHOICE About what percent of the data are less than 15?

    A 25%   **B** 50%   **C** 75%   **D** 100%

10. ERROR ANALYSIS *Describe* and correct the error in interpreting the box-and-whisker plot.

> About 25% of the data values lie between 11 and 20.

OUTLIERS Make a box-and-whisker plot of the data. Identify any outliers.

11. Hours worked per week: 15, 15, 10, 12, 22, 10, 8, 14, 18, 22, 18, 15, 12, 11, 10

12. Prices of MP3 players: $124, $95, $105, $110, $95, $124, $300, $190, $114

13. Annual salaries: $30,000, $35,000, $48,000, $68,500, $32,000, $38,000

14. CHALLENGE Two data sets have the same mean, the same interquartile range, and the same range. Is it possible for the box-and-whisker plots of such data sets to be different? *Justify* your answer by creating data sets that fit the situation.

PROBLEM SOLVING

15. SEAWAY The average sailing times to the Atlantic Ocean from several ports on the St. Lawrence Seaway are shown on the map. Make a box-and-whisker plot of the sailing times.

○ = See **WORKED-OUT SOLUTIONS** in Student Resources   ★ = **STANDARDIZED TEST PRACTICE**

16. BASEBALL STATISTICS In 2004, Ichiro Suzuki scored 101 runs. The numbers of runs he scored against different opposing teams are listed below. Make a box-and-whisker plot of the numbers of runs scored.

Runs scored: 18, 8, 4, 8, 2, 8, 0, 9, 0, 4, 2, 5, 9, 1, 2, 1, 2, 11, 7

(17.) RETAIL SALES The retail sales (in billions of dollars) of the nine U.S. states with the highest retail sales in 2002 are listed below.

California: $153.1   Florida: $118.2   Georgia: $38.4

Illinois: $52.4   New Jersey: $35.8   New York: $54.7

Ohio: $50.7   Pennsylvania: $49.9   Texas: $107.0

a. Make a box-and-whisker plot of the retail sales.

b. Which states, if any, had retail sales in 2002 that can be considered outliers?

18. ★ **SHORT RESPONSE** The stem-and-leaf plot shows the ages of the first 43 presidents of the United States when they first took the oath of office.

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

Key: 4 | 2 = 42 years

a. Make a box-and-whisker plot of the ages.

b. Ronald Reagan was the oldest United States president, and Theodore Roosevelt was the youngest. Can either of these presidents' ages be considered outliers? *Explain* why or why not.

19. ★ **EXTENDED RESPONSE** The box-and-whisker plots show the diameters (in kilometers) of craters on Jupiter's moons Callisto and Ganymede.
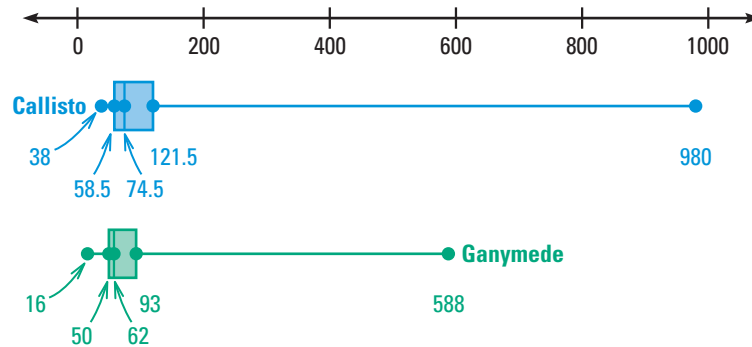
Callisto

a. *Compare* the diameters of craters on Callisto with the diameters of craters on Ganymede.

b. The largest crater in the United States is the Chesapeake Bay in Virginia, with a diameter of 90 kilometers. *Compare* the diameter of the Chesapeake Bay with diameters of craters on Callisto and Ganymede.

c. The largest crater on Earth is Vredefort in South Africa, with a diameter of 300 kilometers. *Compare* the diameter of Vredefort with the diameter of craters on Callisto and Ganymede.

20. CHALLENGE The box-and-whisker plots show the heights (in inches) of singers in a chorus, according to their voice parts. A soprano part has the highest pitch, followed by alto, tenor, and bass, respectively. Draw a conclusion about voice parts and heights. *Justify* your conclusion.
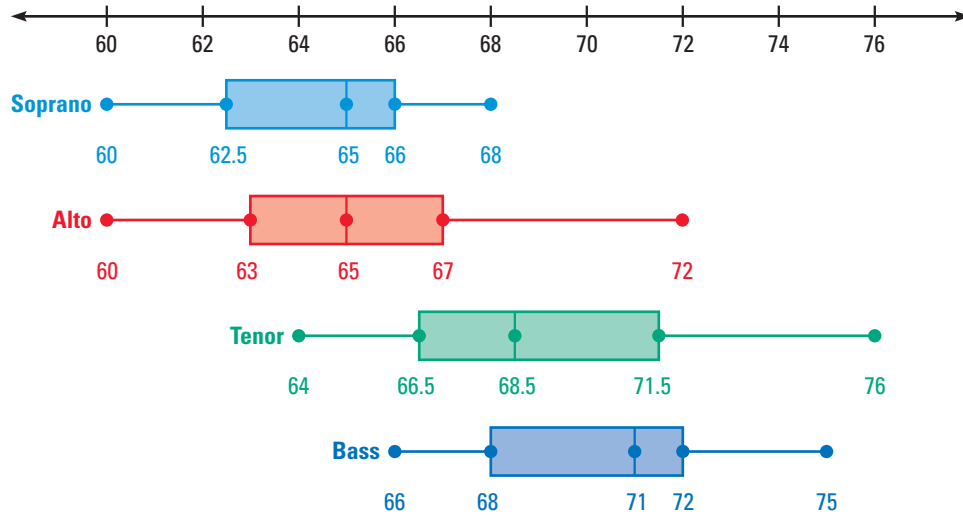
QUIZ

The table shows the number of girls and boys in the ninth grade at Jefferson High School who expressed interest in playing on a soccer team or on a lacrosse team. Use the table for Exercises 1 and 2.

	Soccer	Lacrosse	Total
Girls	64	52	116
Boys	58	60	118
Total	122	112	234

1. Are more ninth graders interested in playing on a soccer team or on a lacrosse team?

2. Did more boys or more girls express interest in playing on a lacrosse team?

3. Make a histogram of the data: 44, 52, 60, 47, 65, 40, 49, 45, 32, 68, 39.

4. Make a stem-and-leaf plot of the data: 1.8, 2.2, 1.2, 2.8, 3.6, 3.3, 1.8, 2.2.

5. **TEST SCORES** The scores on a math exam are given below. Make a box-and-whisker plot of the data. Identify any outliers.

  76, 55, 88, 92, 79, 85, 90, 88, 85, 92, 100, 91, 90, 86, 88

Using ALTERNATIVE METHODS

Another Way to Solve Example 1

MATHEMATICAL PRACTICES

Use appropriate tools strategically.

MULTIPLE REPRESENTATIONS In Example 1, you saw how to analyze the lengths of songs on a CD by drawing a box-and-whisker plot. You can also solve the problem by using a graphing calculator.

PROBLEM

> **SONG LENGTH** The lengths of the songs (in seconds) on a CD are listed below. Make a box-and-whisker plot of the song lengths.
>
> 173, 206, 179, 257, 198, 251, 239, 246, 295, 181, 261

METHOD 1 **Using a Graphing Calculator** An alternative approach is to use a graphing calculator.

STEP 1 **Enter** the data into List 1.

STEP 2 **Go to** the STAT PLOT screen and select the box-and-whisker plot for Plot 1. The Xlist for Plot 1 should be L_1, so that it displays the data from List 1. Make sure Plot 1 is on.
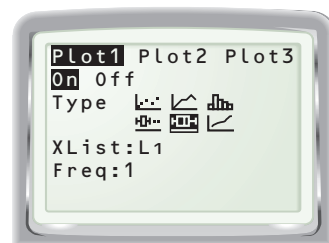
STEP 3 **Press** ZOOM 9 to set the window so that it shows all of the data.

STEP 4 **Press** GRAPH . Use the trace feature to examine the box-and-whisker plot more closely. Notice that the graphing calculator refers to the lower quartile as Q_1 and the upper quartile as Q_3.
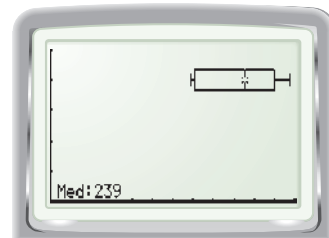
PRACTICE

In Exercises 1 and 2, use a graphing calculator to make a box-and-whisker plot of the data.

1. **REPTILE SPECIES** The number of known reptile species per 10,000 square kilometers in the countries of Asia (excluding the Middle East): 36, 26, 49, 11, 32, 35, 27, 58, 91, 26, 8, 8, 12, 12, 23, 110, 4, 51, 41, 41, 62, 350, 77, 18, 81, 23, 18, 59

2. **BIRD SPECIES** The number of threatened bird species per 10,000 square kilometers in the countries of North and South America: 5, 50, 41, 27, 103, 18, 64, 53, 3, 26, 64, 2, 11, 22

Analyze Data Distribution

GOAL Choose an appropriate display, measure of central tendency, and measure of spread based on the shape of a data distribution.

When you are presenting a set of data, you should consider the distribution of the data before deciding what type of measure of central tendency and graph to use for the data.

DATA THAT ARE CLOSELY GROUPED Use a histogram to display the data. Use the mean as a measure of central tendency. Use standard deviation as a measure of the spread.
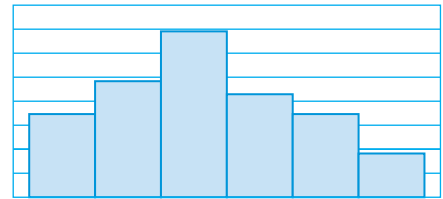
DATA VALUES THAT ARE SPREAD OUT Use a box-and-whisker plot to display the data. Use the median as a measure of central tendency. Use the interquartile range as a measure of the spread.
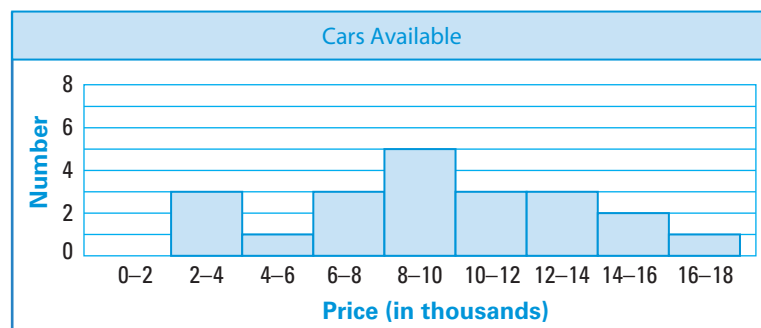
EXAMPLE 1 Choose a display for data

A used car dealer has 21 cars for sale at the prices shown in the table. Choose an appropriate display, measure of central tendency, and measure of spread for this data set.

$2150	$2800	$3500	$5100	$6050	$7100	$7250
$8000	$8850	$9100	$9225	$9900	$10,200	$10,800
$11,750	$12,200	$12,640	$13,020	$14,700	$15,500	$16,400

Solution

The data are close together with no outliers. Use a histogram. The center of the data can be represented by the mean, which is $9,345. The spread can be represented by the standard deviation, which is about $3946.


Cars Available — histogram of Number vs. Price (in thousands), with intervals 0–2, 2–4, 4–6, 6–8, 8–10, 10–12, 12–14, 14–16, 16–18.

EXAMPLE 2 **Choose a display for data**

Another used car dealer has 24 cars for sale at the prices shown in the table. Choose an appropriate display, measure of central tendency, and measure of spread for this data set.

$3,800	$5,100	$7,100	$7,250	$8,850	$9,225	$9,900	$10,200
$10,500	$10,800	$11,400	$11,750	$12,200	$12,350	$12,640	$13,020
$13,890	$14,700	$15,500	$15,990	$17,000	$17,800	$22,900	$38,775

Solution

The data value $38,775 appears to be an outlier. Use a box-and-whisker plot to display the data. The outlier will affect the mean and standard deviation, so they do not represent the data well. The median is $11,975. The interquartile range is $5537.50.

Used Cars

Price (in thousands)
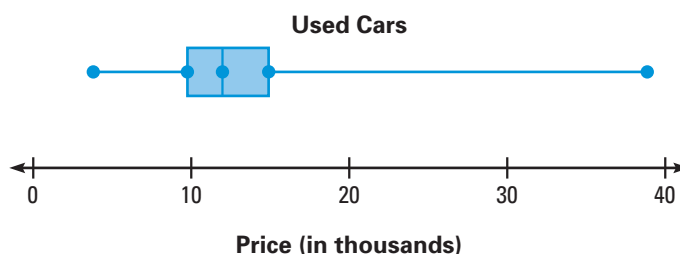
PRACTICE

For Exercises 1–6, choose an appropriate display, measure of central tendency, and measure of spread for the data set. Explain your reasoning.

1. **QUIZ SCORES** The scores on the first quiz in Mr. Stuart's math class were 6, 9, 10, 12, 12, 13, 14, 14, 15, 15, 15, 16, 16, 17, 17, 17, 17, 18, 18, 18, 19, 19, 19, 20, and 20.

2. **FOOTBALL** The points scored by twenty of the top 25 college football teams on Saturday, September 25, 2010 were 24, 73, 37, 42, 17, 31, 70, 35, 10, 20, 37, 65, 22, 31, 20, 24, 12, 27, 14, and 34.

3. **RUNNING** The time (in minutes) it took twenty freshmen to run the mile in physical education class were 7, 7.5, 8, 8, 8.2, 8.4, 8.5, 9, 9, 9, 9.6, 9.8, 10, 10.5, 10.5, 10.8, 11.2, 11.5, 11.7, and 12 minutes.

4. **HOMEWORK** The numbers of hours that twenty-five students spent doing homework last week were 1, 8, 8, 8.5, 9, 9.5, 9.5, 10, 10, 10, 10, 10, 10.5, 10.5, 10.5, 11, 11, 11, 11, 11.5, 11.5, 12, 12, 12, and 12.

5. **COOKIES** The numbers of cookies in 20 boxes at a bake sale are 16, 16, 18, 18, 20, 20, 24, 24, 24, 24, 26, 28, 28, 30, 30, 30, 30, 36, 36, and 36.

6. **BASEBALL** The attendance at a professional baseball team's home games during September are shown in the table.

39,555	31,424	40,788	31,647	31,596	33,623
36,364	37,285	34,481	36,553	39,316	38,057

MIXED REVIEW *of Problem Solving*

MATHEMATICAL PRACTICES Make sense of problems and persevere in solving them.

1. MULTI-STEP PROBLEM The ages of people who attended an opening reception for a theater production are listed below.

54, 25, 28, 64, 30, 42, 33, 50, 27, 35, 40, 39, 41, 52, 49, 48, 56, 60, 58, 37, 56, 45, 57, 62

a. Make a frequency table of the data.

b. Make a histogram of the data.

2. SHORT RESPONSE Students collected fish of two species, blue gill and largemouth bass, from the same pond for a science fair project. The lengths (in millimeters) of the fish they collected are listed below.

Blue gill: 186, 171, 171, 176, 183, 182, 172, 172, 173, 184

Largemouth bass: 354, 297, 300, 344, 317, 360, 432, 457, 392, 395

a. Make a box-and-whisker plot for the data.

b. Based on the fish the students collected, which of the species has more variation in length? *Explain* your reasoning.

**3. MULTI-STEP PROBLEM** Jo and her friend Abe sold tickets for the Homecoming football game. The table shows the number of student tickets and adult tickets each person sold.

	Student	Adult	Total
Jo	178	215	393
Abe	201	188	389
Total	379	403	782

a. How many adult tickets did Jo sell?

b. How many tickets did Jo and Abe sell altogether?

c. Who sold more student tickets?

4. SHORT RESPONSE
The back-to-back stem-and-leaf plot below shows the lengths (in meters) of the eight best men's and women's final long jump results from the 2004 Olympics. *Compare* the lengths of the jumps by men with those by women.
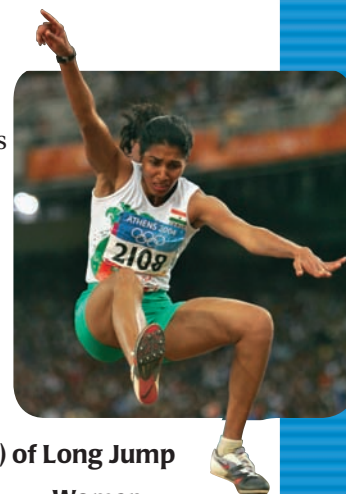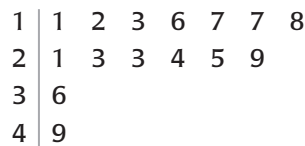
Lengths (in meters) of Long Jump

Men		Women
	6	7 8 8 9
	7	0 0 0 1
6 5 3 3 2 2 2 0	8	

Key: 0 | 7 | 1 = 7.0 m, 7.1 m

5. SHORT RESPONSE The stem-and-leaf plot shows the number of games lost by 15 NCAA football coaches with the greatest career winning percentages after at least 10 years of coaching.

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

Key: 2 | 1 = 21 games

a. Make a box-and-whisker plot of the data.

b. Tom Osborne had a winning percentage of 83.6% over his career and lost 49 games. Can the number of games lost by Tom Osborne be considered an outlier? *Explain* your reasoning.