19.3 teacher notes

Objectives

- Analyze stem-and-leaf plots.
- Create and analyze box-and-whisker plots.

Prerequisites

Interpreting a stem-and-leaf plot

Finding the mean and median of a data set

Graphing numbers on a number line

Vocabulary

Stem-and-leaf plot (Lesson 19-1)
Box-and-whisker plot
Range
Median (Lesson 19-1)
Quartile
First quartile (Q_1)
Third quartile (Q_3)
Interquartile range
Five-number summary
Outlier (Lesson 19-1)

Get Started

- Say, "A business magazine, *Widgets Quarterly*, published information about the salaries at two top widget-producing companies."
- On the board or overhead projector, display the following information:

Williams' Widgets	Wayne's Widgets
mean = $36,000	mean = $36,000
median = $35,000	median = $35,000

- Ask the class to discuss the information. Point out that the mean and median salaries are identical for both companies. Ask, "What impression does the information give?" (Students will probably conclude that, on average, employees of the two companies earn about the same amount). Say, "The data published in the magazine gives the impression that salaries at the two companies are very similar."

- On the board or overhead projector, display the following information:

Williams' Widgets
$20,000 $31,000 $35,000 $35,000 $45,000 $50,000

Wayne's Widgets
$33,000 $33,000 $35,000 $35,000 $40,000 $40,000

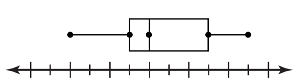
Say, "This is the data used by the magazine writer to report the mean and median salaries."

- Ask students to verify that the mean and median given earlier are accurate. Then, lead the class in a comparison of the two data sets. Point out that although the companies have the same mean and median salary, the employees' salaries at Williams' Widgets are more varied. Encourage students to realize that the range of salaries is $33,000 at Williams' Widgets, but only $7,000 at Wayne's Widgets.
- Say, "Today, we will use a graph called a box-and-whisker plot to display a data set and to compare two data sets. A box-and-whisker plot is a visual representation of the spread or dispersion of the data. It gives information about how the data are distributed. A box-and-whisker plot can be especially useful when comparing two data sets."

Section 1

Expand Their Horizons

In this lesson, students will be introduced to box-and-whisker plots. A box-and-whisker plot is a graph that depicts the distribution of data in a data set. It is based on quartiles—data points that divide the set into four subsets each containing approximately 25% of the data. The name "box-and-whisker plot" comes from its appearance. It consists of a rectangular box and two lines that extend horizontally from each side of the box. A box-and-whisker plot is shown below.



Students will use data displayed in a stem-and-leaf plot to identify five values used to construct a box-and-whisker plot: (1) the minimum value, (2) the first quartile or Q_1, (3) the median or Q_2, (4) the third quartile or Q_3, and (5) the maximum value. These five values are commonly referred to as a five-number summary. A stem-and-leaf plot is useful because it displays data in a manner which makes it easy to determine the values in a five-number summary.

Consider the stem-and-leaf plot used in the DVD to display winning average speeds in Indianapolis 500 auto races.

13	4
14	2 5 6 8
15	4 7
16	1 6 8

14|5 = 145

The data values are arranged in ascending order and displayed in an easy-to-use format. The minimum value, 134, is the first value in the stem-and-leaf plot, and the maximum value, 168, is the last value in the plot. That takes care of two of the five numbers in the summary.

13	**4**
14	2 5 6 8
15	4 7
16	1 6 **8**

14|5 = 145

Minimum value	134
First quartile, or Q_1	
Median, or Q_2	
Third quartile, or Q_3	
Maximum value	168

The median—also referred to as the second quartile (Q_2)—is the middle value. Because this data set contains an even number of values, the median is the average of the two middle values: (148 + 154) ÷ 2 = 151.

13	4
14	2 5 6 **8**
15	**4** 7
16	1 6 8

Minimum value	134
First quartile, or Q_1	
Median, or Q_2	151
Third quartile, or Q_3	
Maximum value	168

14|5 = 145

The first quartile or Q_1 is determined by finding the median of the lower subset of the data, not including the median. In this case, Q_1 = 145.

13	4
14	2 **5** 6 8
15	4 7
16	1 6 8

Minimum value	134
First quartile, or Q_1	145
Median, or Q_2	151
Third quartile, or Q_3	
Maximum value	168

14|5 = 145

The third quartile (Q_3)is determined by finding the median of the upper subset of the data, not including the median. In this case, Q_3 = 161.

13	4
14	2 5 6 8
15	4 7
16	**1** 6 8

Minimum value	134
First quartile, or Q_1	145
Median, or Q_2	151
Third quartile, or Q_3	161
Maximum value	168

14|5 = 145

We now have all five values needed to construct a box-and-whisker plot.

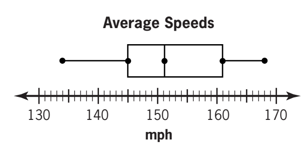
Begin construction of the box-and-whisker plot by drawing a number line that extends past the minimum and maximum values. Use equal intervals on the number line. Then, above the number line, mark the points for the five values in the five-number summary. Create the box by drawing a rectangle with its left side at the first quartile (Q_1) and its right side at the third quartile (Q_3). Next, draw a vertical line segment inside the box at the median. Create the left whisker by drawing a horizontal line segment that extends from the left side of the box to the minimum value and create the right whisker by drawing a horizontal line segment which extends from the right side of the box to the maximum value.

Understanding the structure of a box-and-whisker plot is essential to drawing conclusions about the data it represents. Students should understand the concept of range. The range is the numerical value that represents the spread of data within a set or subset. The range of an entire data set is the difference between the set's maximum and minimum values. In the example above, the range of the data set is 168 − 134 = 34.

In a box-and-whisker plot, the box represents a subset of the data—50% (or approximately 50%) of the data in the middle of the set. Because the box is bounded by the first and third quartiles, we can calculate the interquartile range (IQR) by subtracting Q_1 from Q_3. In the example above, the interquartile range is 161 − 145 = 16. The interquartile range is the length of the box. The longer the box, the more "spread out" is the middle subset of the data.

The left whisker represents the distance between the minimum value and the first quartile, so it represents approximately the lower 25% (or approximately 25%) of the data. Similarly, the right whisker represents approximately the upper 25% (or approximately 25%) of the data. The longer the whisker, the more "spread out" is the subset of the data represented by that whisker. An extremely long whisker may indicate the presence of an outlier—a data value that is much higher or much lower than others in the set.

Consider the box-and-whisker plot created in the DVD to show the distribution of winning average speeds in the Indianapolis 500.



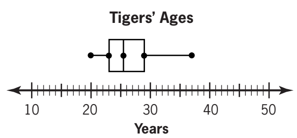
The longer left-hand whisker indicates speeds which were more spread out than the speeds represented by the shorter right-hand whisker. In fact, the left-hand whisker represents a range of 11 mph, while the right-hand whisker represents a range of seven mph.

The median bar's location to the left of the box's center indicates the speeds between the first and second quartiles are more closely clustered than the speeds between the second and third quartiles.

1 **Using the given five-number summary, make a box-and-whisker plot for the ages of the baseball players who play for the Tigers.**

	Age (yrs)
Minimum:	20
First Quartile, Q1:	23
Median, Q2:	25.5
Third Quartile, Q3:	29
Maximum:	37

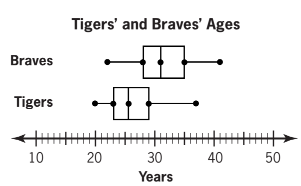
In this exercise, the five-number summary is provided. To construct a box-and-whisker plot, begin by drawing a number line. The five-number summary contains ages ranging from 20 to 37. Be sure the number line includes at least all of the data points. Next, plot all five numbers and draw the box and whiskers.



2 **Using the given five-number summary, make a box-and-whisker plot for the ages of the baseball players who play for the Braves.**

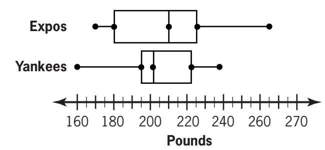
	Age (yrs)
Minimum:	22
First Quartile, Q_1:	28
Median, Q_2:	31
Third Quartile, Q_3:	35
Maximum:	41

This exercise shows how two box-and-whisker plots can be created on the same number line for comparison. To create the additional box-and-whisker plot, graph each number in the five-number summary and then, draw the box and whiskers.



When two box-and-whisker plots are graphed on the same number line, comparisons are often made easier. For example, in the graph created in Guided Notes Questions 1 and 2, it can be seen that the Braves players are, in general, older the Tigers players; that the Braves' players ages are more diverse (that is, the range of ages on the Braves is greater); that the median age of the Braves' players is greater than that of the Tigers. Review the graph carefully with the class, pointing out these conclusions and others that can be drawn.

3 **Compare the box-and-whisker plots for the weights of the Yankees' players and the Expos' players.**



After showing the screen for Guided Notes Question 3, ask students to write statements presenting conclusions which can be drawn from the double box-and-whisker plot. For example, they might write, "Half the Expos weigh 210 lb or more, while half the Yankees weigh 202 lb or more;" "The range of the Yankees' weights is less than the range of the Expos' weights;" or "The Expos have the heaviest player." Consider allowing the class to work in pairs to write as many conclusions as possible. Discuss each conclusion to determine its validity.

Connections

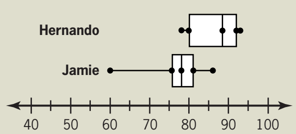
Students are probably familiar with the word *percentile*. The *n*th percentile of a set of numbers is the number in the set such that *n* percent of the numbers in the set are less than or equal to it. A box-and-whisker plot shows the 25th, 50th, 75th and 100th percentiles at a glance.

Additional Examples

1. Determine the five-number summary for each data set and then, make box-and-whisker plots of the data sets. Use the same number line for both graphs.

Jamie's grades:
60, 72, 78, 76, 81, 86, 77, 78, 76, 80, 82

Hernando's grades:
90, 78, 80, 86, 88, 93, 92, 84, 80, 88, 92



2. Use the box-and-whisker plot to compare Jamie's and Hernando's grades. Compare the two sets of data.

Jamie's grades have a greater spread than Hernando's grades. Hernando's median grade is higher than Jamie's highest grade.

Look Beyond

Box-and-whisker plots are used to display the distribution of data. Sometimes, one or more data values are quite different from the rest of the data. Such values might be called outliers. One way to decide whether a value is an outlier is to determine where it lies in relation to number-line boundaries called *fences*.

To find the *lower fence*, multiply the IQR by 1.5, and then, subtract the result from the lower quartile. A value that is less than the lower fence can be considered an outlier. To find the *upper fence*, add 1.5 (IQR) to the upper quartile. A value which is greater than the upper fence can be considered an outlier. Some box-and-whisker plots do not include outliers in the box and whiskers but instead show them as isolated points. In a box-and-whisker plot of this type, the left whisker is drawn to the least value greater than the lower fence, and the right whisker is drawn to the greatest value less than the upper fence. Outliers are then placed on the graph using special notation, such as small circles or stars.