19.1 teacher notes

Objectives

- Find the mean, median, and mode of a set.
- Create a stem-and-leaf plot.
- Analyze a stem-and-leaf plot.

Vocabulary

Mean
Average
Set (Lesson 1-1)
Median
Mode
Outlier
Statistics
Stem-and-leaf plot

Prerequisite

Finding the average

Get Started

- Prepare pieces of string by cutting them to the following lengths: two 1 in. pieces, three 3 in. pieces, and four 5 in. pieces. Place them, straightened, on a table for all students to view. Do not put them in order of size.
- Ask students what they notice about the pieces. Use leading questions such as, “Which size is the least common?” and “Which size is the most common?”
- Students should notice there are more of the 5-in. pieces than of other-length pieces. Explain that, in this set of string pieces, 5 in. is the mode because that length occurs most often.
- Instruct a student to place the pieces in order from shortest to longest. Ask, “Which piece is in the middle of this arrangement?”
- Students should notice the third 3-in. piece is the middle piece, with four pieces of string on each side of it. Explain that, in this set of string pieces, 3 in. is the median because it is in the middle of the set of strings.
- If desired, complete a manipulative exercise to demonstrate mean. See the “Manipulatives Section” at the end of these Teacher Notes.

Section 1

Expand Their Horizons

Statistics is a branch of mathematics that deals with the collection, analysis, interpretation, and presentation of information referred to as "data." In Section 1, students will learn to find three types of statistical measures: the mean, the median, and the mode. The mean, median, and mode are measures of the "central tendency" of a set of numbers. That is, they are used to describe a typical or representative value of the data set. A data set consists of an unordered group of numbers.

There are different types of means used in statistics; however, the mean used in this lesson is the arithmetic mean. The arithmetic mean is an average—the sum of the values in a data set divided by the number of values in that set. The values in the data set do not need to be placed in any particular order for one to be able to calculate the arithmetic mean. Sometimes, the mean is rounded to the same place value as that of the individual items in the given data set. However, in this lesson, the mean is stated exactly as computed.

Although the arithmetic mean is the measure of central tendency with which people tend to be most familiar, it is not a good representation of the data if there is a value in the data set that is much higher or much lower than the other values. Such a value is referred to as an "outlier" and will skew the results. For example, the mean of the data set {5, 7, 8, 6} is $(5 + 7 + 8 + 6) \div 4 = 26 \div 4 = 6.5$, which is a good representation of the values in the set. Adding one outlier, such as 25, to the set would result in a mean of $(5 + 7 + 8 + 6 + 25) \div 5 = 51 \div 5 = 10.2$, which gives a less significant measure of central tendency.

The median is the middle value in a data set when the values (or members) of the data set are arranged in order. Whether that order is increasing or decreasing does not matter; either order will yield the same median. If the data set contains an odd number of data values, the median is the middle value listed. Some statistics textbooks give the formula $\frac{n+1}{2}$ for obtaining the placement for the median, where n is the number of values in the set. For example, if there are 13 values in the set, $\frac{13+1}{2} = \frac{14}{2} = 7$; the seventh value, when arranged in order, is the median. However, in most cases, students will have no trouble finding the middle value without this formula.

If the data set contains an even number of data values, there are two middle values. In such a situation, the two middle values are averaged to find the median. This may produce a median with a value that was not in the original set of values. For example, in the set {3, 4, 5, 6}, the median is the average of four and five or 4.5. Unlike the mean, the median is not usually significantly affected by an outlier and may be considered a better representation of the central tendency when an outlier is present.

The mode is the value in a data set that occurs most often. A set of values may have one mode, more than one mode, or no mode. A set is said to have no mode if each value of the set occurs equally. The data do not have to be arranged in order to determine the mode, but for larger sets it may be helpful. Also, the mode is not affected by an outlier.

The data in a set can be presented by a stem-and-leaf plot. In the following example, eight, nine, and 10 are stems:

8	
9	
10	

The last digits are the leaves. They are arranged horizontally, usually in increasing order, in the row of their corresponding stem and to the right of a vertical line. In the example below, one, six, five, seven, nine, and zero are leaves.

8	1 6
9	5 7 9
10	0

A stem-and-leaf plot should include a key that demonstrates how the digits are to be interpreted. For example, 10|0 = 100.

The number to the left of the vertical bar is concatenated with the number to the right of the bar for each respective row.

Basically, a stem-and-leaf plot displays the data arranged in order using a short-hand notation. It may be easier for students to place the data in order before creating this plot for the first few examples. A stem-and-leaf plot can help students determine the median and the mode. For the median, the data in the plot is already placed in order and the middle value can easily be found. The mode can quickly be identified by looking for the stem with the greatest number of same-digit leaves.

The following examples demonstrate how to find the mean, median, and mode of a data set and how to construct a stem-and-leaf plot.

Consider the data set {7, 5, 6, 8, 4, 6, 5, 6, 7, 6, 5, 10}. The mean is the average, found by adding the values and dividing by the number of values in the set. The sum of the values in this set is 75, which is divided by 12 because there are 12 values listed. The quotient, which is the mean of the set, is 6.25.

In order to find the set's median, the values must be listed in order: {4, 5, 5, 5, 6, 6, 6, 6, 7, 7, 8, 10}. Because there are 12 values, and 12 is an even number, the median is found by averaging the two middle values. The sixth term from each side is the value six, and the average of six and six (found by adding these to get 12 and dividing by two) is six. Thus, the median for the set is six. The mode is the most frequent value, which in this case is also six.

The mean, median, and mode will not always be so close in value to one another. However, in cases where these measures of central tendency are the same or are very close, any of the measures of central tendency would be a good representation of the data.

Next, consider the data set {95, 100, 86, 97, 99, 81}. Because a stem-and-leaf plot is not very helpful in finding the mean, we will determine the mean from the original values as written: $(95 + 100 + 86 + 97 + 99 + 81) \div 6 = 558 \div 6 = 93$. The mean of the set is 93.

To find the median and the mode, we will use a stem-and-leaf plot. Place every first digit in numerical order, to the left of the line; this is the stem. The last digits, called the leaves, are placed to the right of the line and next to their corresponding stem.

The stems in this example will be eight, nine, and 10. The leaves corresponding to stem eight are six and one (for 86 and 81); the leaves corresponding to stem nine are five, seven, and nine (for 95, 97, and 99); and the leaf corresponding to stem 10 is zero (for 100). In a stem-and-leaf plot, the stems and leaves should be placed in order. In this lesson, they are placed in increasing order. A key is necessary to illustrate how the digits in the plot should be interpreted. Any value in the set may be used. For example, 8|1 = 81, 9|7 = 97, or 10|0 = 100. Choosing the latter, the stem-and-leaf plot becomes this plot.

8	1 6	
9	5 7 9	10|0 = 100
10	0	

Because the stem-and-leaf plot places the values in order, it can be used to easily find the median. This data set contains an even number of values, so the median will be the average of the two middle values, 95 and 97. Averaging these values gives $(95 + 97) \div 2 = 192 \div 2 = 96$. Thus, the median of the set is 96. Note that 96 is not a value in the original set.

The mode is the value that occurs most frequently in the set. Because each value is listed only once, the set has no mode.

Common Error Alert

Students may mistakenly list zero as the mode for a set having no mode at all. Consider the set of values {–3, –1, –1, 0, 0, 0, 5}. In this set, zero occurs most frequently and thus, is the mode. Explain that the value zero can be the mode, but zero should not be written as the mode for a set that has no mode.

Consider this stem-and-leaf plot.

2	4 8 8	
3	8 8	2|4 = 2.4
4	0 0 0 8	

Notice that in this example, the stem is the one's digit, and the leaf is the tenth's digit. The mean is the average of the values. If the set of numbers, 2.4 + 2.8 + 2.8 + 3.8 + 3.8 + 4.0 + 4.0 + 4.0 + 4.8 is added, the sum will be 32.4.

Because there are nine values in this set, divide 32.4 by nine. The result is 3.6. Therefore, the mean is 3.6.

To find the median, the values must be placed in order, as has already been done in the stem-and-leaf plot. The middle value is the fifth value from either side. The middle value is the eight-leaf on the three-stem, indicating that 3.8 is the median.

The mode is the value that appears most often in the set—in this case the zero-leaf on the four-stem. Thus, 4.0 is the mode for this set of values.

Common Error Alert

When finding the mode from a stem-and-leaf plot, the student should consider both the stem and the leaf; that is, the student should consider the entire number. Students should not list as the mode a number corresponding to the leaf which appears most frequently in the *table*, but rather a number corresponding to the leaf that appears most frequently in a single row. In the previous example, a leaf of eight was listed five times. However, it represented three different numbers: 2.8 (twice), 3.8 (twice), and 4.8. Because 4.0 occurred three times and a leaf of eight did not represent any number more than twice, 4.0 was the mode.

Use the following data set for Problems 1 through 3. Children's Pets: 1, 8, 2, 5, 4.

1 **Find the mean.** The mean is the sum of the values, divided by the number of values. Here, the sum is 20, to be divided by the number of values, which is five. Because $20 \div 5 = 4$, the mean number of children's pets is four pets.

2 **Find the median.** To find the median, the data must be arranged in order: 1, 2, 4, 5, 8. Because the data set contains an odd number of values, the median is the middle value, four. Two children have more than four pets, and two children have fewer than four pets.

3 **Find the mode.** Because each value occurs only once, there is no mode. Each child has a different number of pets.

Use the following data set for Problem 4: Newt's Times (in minutes) to Complete Obstacle Course: 31, 22, 15, 19, 17, 19, 20.

4 **Make a stem-and-leaf plot.** The stems will be the ten's digits of one, two, and three, when placed in increasing order. For the stem one, the leaves are five, nine, seven, and nine (for the numbers 15, 19, 17, and 19, respectively). For the stem two, the leaves are two and zero (for the numbers 22 and 20, respectively). For the stem three, there is only the leaf one (for the number 31). The leaves for each stem are arranged horizontally in increasing order. Any of the values can be used as a key. For example, 1|7 represents 17; thus, it can be listed as the key.

1	5 7 9 9	
2	0 2	1|7 = 17
3	1	

Connections

Statistics is a branch of mathematics, which can be used in many fields. A store owner, who wants to determine the number of salespeople needed on a given day of the week, can use one of the measures of central tendency to estimate the number of customers who will visit his store on that day. A firm can collect scores on a college exit exam for those students who turn out to be successful employees and decide what scores to use as criteria for hiring. A farmer can analyze the yield in ears of corn per stalk to determine future planting. People can use statistics to their advantage in many "non-math" fields.

Additional Examples

1. Create a stem-and-leaf plot for the data set {203, 150, 180, 168, 175, 152, 196, 180, 159, 203}.

In a stem-and-leaf plot, every digit except the last digit is arranged vertically in order as a stem. The stems will be 15, 16, 17, 18, 19, and 20. The last digits are the leaves, which are placed, in order, in the row of their corresponding stems and to the right of the vertical line. The leaves for the stem 15 are zero, two, and nine (for 150, 152, and 159). The leaf for 16 is eight (for 168). The leaf for the stem 17 is five (for 175). The leaves for the stem 18 are zero and zero (both for 180). The leaf for 19 is six (for 196). Finally, the leaves for the stem 20 are three and three (both for 203). Any one of these can be used as the key.

15	0 2 9
16	8
17	5
18	0 0
19	6
20	3 3

20|3 = 203

2. Find the mean, median, and mode of the data set in Additional Example Problem 1.

The mean is found by adding the values in a data set and dividing by the number of values in the set. The sum of the values given is 1,766, and the data set contains 10 values. Thus, 176.6 is the mean for this set.

The median is the middle value in the set. Because this set has an even number of values, the two middle values are averaged to find the median. The two middle values are 175 and the first 180. Adding these values and dividing by two yields 177.5 as the median.

The mode is the value listed most frequently in the set. In this set, two values are listed twice; thus, there are two modes for this set, 180 and 203.

Mean: 176.6
Median: 177.5
Modes: 180, 203

Look Beyond

The study of statistics is a mathematics course in its own right and covers many topics. One of the topics studied in a statistics course is the various types of means, such as a weighted mean, which can be used to find the grade average in a course where the final exam is a more significant grade than the other grades. Another topic is measures of dispersion, such as range and standard deviation, which analyze how far apart data values are. These statistical methods can be used to analyze current measures and make predictions, which are the basis of probability—a concept studied in the next module.

Manipulatives

To demonstrate the mean of a set, use cups and beans as manipulatives. There should be one cup for each value in the set, with the values represented by the number of beans in the cups. To find the mean, place the beans from all of the cups in a pile and redistribute the beans so that each cup contains the same number of beans. The number of beans in each cup is the mean.

For example, to find the mean of set {6, 3, 5, 2}, four cups will be needed. These cups will hold six, three, five, and two beans, respectively, to represent the original set. If all the beans from all of the cups are placed in a pile, the pile will have 16 beans. To redistribute the beans equally, place one bean in each cup and then, repeat for a second bean in each cup, etc., until all of the beans are placed evenly. Each cup will then hold four beans, indicating four is the mean of this set.

For a set that will not distribute equally, parts of beans may be used. Consider the set {4, 0, 1, 5, 3, 2}. Use six cups holding four beans, no beans, one bean, five beans, three beans, and two beans, respectively. If all the beans are placed in a pile, the pile will have 15 beans. Place one bean in each of the six cups. Then, place a second bean in each cup. There will be three beans remaining, but three beans cannot be distributed evenly among six cups. However, each of the three beans can be cut in half, and the resulting six halves can be evenly distributed among the cups. Each cup now holds $2\frac{1}{2}$ beans, which is the mean of the set.