7.1 Measures of Center and Variation

Learning Target:	Understand measures of center and variation.
Success Criteria:	• I can find the mean, median, and mode of a data set. • I can find the range and interquartile range of a data set. • I can identify outliers in a data set. • I can describe how outliers affect the measures of center and variation of a data set.

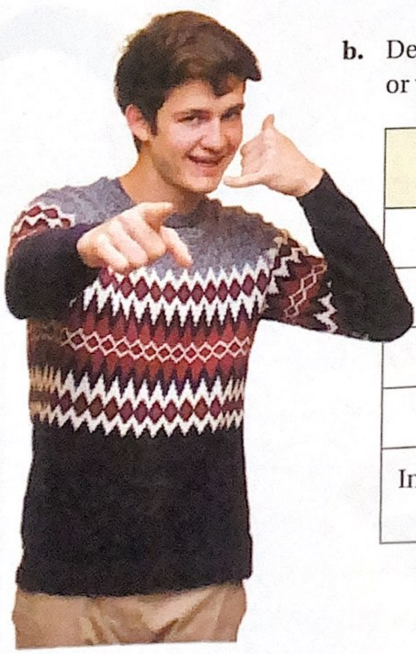
Data Analysis and Probability
preparing for MA.7.DP.1.1 Determine an appropriate measure of center or measure of variation to summarize numerical data, represented numerically or graphically, taking into consideration the context and any outliers.

Exploration 1 Finding Measures of Center and Variation

Work with a partner.

a. In your own words, describe the difference between a measure of *center* and a measure of *variation*.

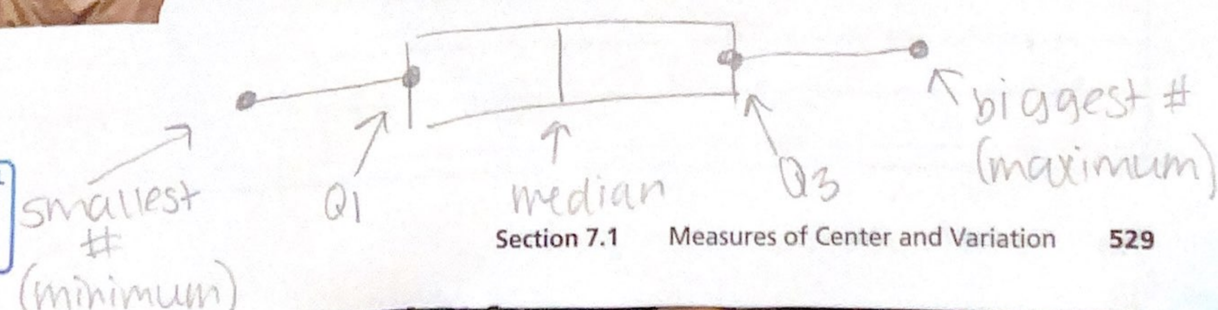
measures of center - describes the typical value
measures of variation - describes the distribution



b. Define each term in the table. Classify each as a measure of center or variation. Explain.

Term	Definition	Measure of *Center* or *Variation*?
Mean	average	center
Median	middle	center
Mode	# that appears most	center
Range	big - smallest	variation
Interquartile Range	Q3 - Q1	variation

Box and whisker



GO DIGITAL

7.1 Lesson

Recall that a *measure of center* is a measure that describes the typical value of a data set. The mean, median, and mode are measures of center.

Key Ideas

average

add all #s divide by how many #s there are

Mean

The mean of a data set is the sum of the data divided by the number of data values.

Data set: $\{4, 7, 7, 12, 15\}$

Mean: $\dfrac{4+7+7+12+15}{5} = \dfrac{45}{5} = 9$

Median middle

Order the data. For a set with an odd number of values, the median is the middle value. For a set with an even number of values, the median is the mean of the two middle values.

*has to be in order from least to greatest

Median: $\{4, 7, 7, 12, 15\}$

The median is 7.

Data can have one mode, more than one mode, or no mode. When all values occur only once, there is no mode.

Mode which # appears the most

The mode is the value or values that occur most often.

Mode: $\{4, 7, 7, 12, 15\}$

The mode is 7.

Example 1 Finding Measures of Center

Daily Numbers of Visitors				
132	318	265	282	211
161	195	304	247	195

Find the mean, median, and mode of the daily numbers of visitors to a beach.

Mean: $\dfrac{132+318+265+282+211+161+195+304+247+195}{10} = \dfrac{2310}{10} = 231$

Median: 132, 161, 195, 195, 211, 247, 265, 282, 304, 318

Order the data.

$$\frac{211+247}{2} = \frac{458}{2} = 229$$

Add the two middle values and divide by 2.

Mode: 132, 161, 195, 195, 211, 247, 265, 282, 304, 318

The value 195 occurs most often.

▶ The mean is 231, the median is 229, and the mode is 195.

GO DIGITAL

Try It Find the mean, median, and mode of the data.

1. 12, 6, 9, 11, 15, 22, 9, 20

mean: 12+6+9+11+15+22+9+20

104 ÷ 8 = 13

median: 6, 9, 9, 11, 12, 15, 20, 22

11.5

mode: 9

2. 26, 8, 12, 31, 14, 25, 34, 7, 23

mean: 180 ÷ 9 = 20

median:

7, 8, 12, 14, 23, 25, 26, 31

mode: none

Recall that a *measure of variation* is a measure that describes the distribution of a data set. The range and the interquartile range (IQR) are measures of variation.

range

subtract biggest − smallest

Key Ideas

Range and Interquartile Range (IQR)

The quartiles of a data set divide the data set into four equal parts.

least value → {3 3 4 5 6 6 8 9 9 12} ← greatest value

$Q_1 = 4$ median (Q_2) = 6 $Q_3 = 9$

The median is also called the second quartile (Q_2). The median divides the data set into two halves. The median of the lower half is the first quartile (Q_1), and the median of the upper half is the third quartile (Q_3).

The range is the difference of the greatest value and the least value.

{3 3 4 5 6 6 8 9 9 12}

range = greatest value − least value

= 12 − 3

= 9

The interquartile range is the difference of the third quartile and the first quartile. The IQR represents the range of the middle half of the data set.

{3 3 4 5 6 6 8 9 9 12}

$\text{IQR} = Q_3 - Q_1$

$= 9 - 4$

$= 5$

IQR

Q3 − Q1

Example 2 Finding Measures of Variation

Find the range and interquartile range of the data in Example 1.

median $= \dfrac{211 + 247}{2} = 229$

least value → 132 161 195 195 211 247 265 282 304 318 ← greatest value

$Q_1 = 195$ $Q_3 = 282$

Range: 318 − 132 = 186 **IQR:** 282 − 195 = 87

▶ The range is 186 and the interquartile range is 87.

Try It **Find the range and interquartile range of the data.**

3. 11, 6, 16, 20, 9, 15, 17

range: 20 − 6 = 14

6, 9, 11, $\boxed{15}$, 16, 17, 20 Q3 Q1

17 − 9 = IQR 8

4. 84, 76, 94, 102, 91, 98, 77, 85

range: 102 − 76 = 26

76, 77, 84, 85, 91, 94, 98, 102

80.5 $\boxed{88}$ 96

96 − 80.5 = IQR 15.5

You can use the interquartile range to identify outliers. Any value in a data set less than $Q_1 - 1.5(\text{IQR})$ or greater than $Q_3 + 1.5(\text{IQR})$ is an outlier.

Example 3 Identifying Outliers

Lovebugs thrive in Florida because of its warm, humid climate.

The numbers of eggs laid by several female lovebugs are represented by the data set {362, 320, 360, 377, 326, 211}. Identify any outliers.

Order the data values from least to greatest. Find the quartiles.

median $= \dfrac{326 + 360}{2} = 343$

211 320 326 360 362 377

$Q_1 = 320$ $Q_3 = 362$

The IQR is 362 − 320 = 42. Use the IQR to find the outlier boundaries.

$$Q_1 - 1.5(\text{IQR}) = 320 - 1.5(42) = 257 \qquad Q_3 + 1.5(\text{IQR}) = 362 + 1.5(42) = 425$$

▶ The only data value less than 257 is 211. There are no data values greater than 425. So, the only outlier is 211.

Outliers → #s that are either really small or really big compared to others

GO DIGITAL