



# Announcements

- ◀ Math MAPs are next week on Monday and Tuesday
- ◀ Flashback Fridays will begin this Friday
  - ◀ 10 questions
  - ◀ On SchoolNet
  - ◀ Graded for accuracy
  - ◀ You MAY consult friends, notes, etc.



## Unit Map - Statistics

- ← ~~Tuesday - Histograms and Box Plots~~
- ▶ **Wednesday - Measures of Central Tendency**
- ▶ Thursday - Line of Best Fit
- ▶ Friday - Line of Best Fit
- ▶ Monday - Math MAP
- ▶ Tuesday - Math MAP (continued)
- ▶ Wednesday - Friday DC Trip
- ▶ Spring Break
- ▶ Monday - Statistics Review
- ▶ Tuesday - Statistics Test



**Tutor.com**

Through OneAccess



# Homework Review

# Measures of Central Tendency

4/3/2019



# Measures of Central Tendency

1. Calculate the mean, median and mode of a set of data.
2. Determine the affect of removing an outlier from a set of data.
3. Describe and support the best measure of central tendency for a given set of data.

**Essential Understanding** You can use different measures to interpret and compare sets of data.

One way to summarize a set of data is to use a *measure of central tendency*. Mean, median, and mode are all **measures of central tendency**.

The measure of central tendency that best describes a data set may depend on whether the data set has an *outlier*. An **outlier** is a data value that is much greater or less than the other values in the set. Below is a review of mean, median, and mode, and when to use each as the measure of central tendency.



take note

## Key Concept Mean, Median, and Mode

### Measure

The **mean** equals  $\frac{\text{sum of the data values}}{\text{total number of data values}}$ . The mean is often referred to as the *average*.

The **median** is the middle value in a data set when the values are arranged in order. For a set containing an even number of data values, the median is the mean of the two middle data values.

The **mode** is the data item that occurs the most times. A data set can have no mode, one mode, or more than one mode.

### When to Use

Use mean to describe the middle of a set of data that *does not* have an outlier.

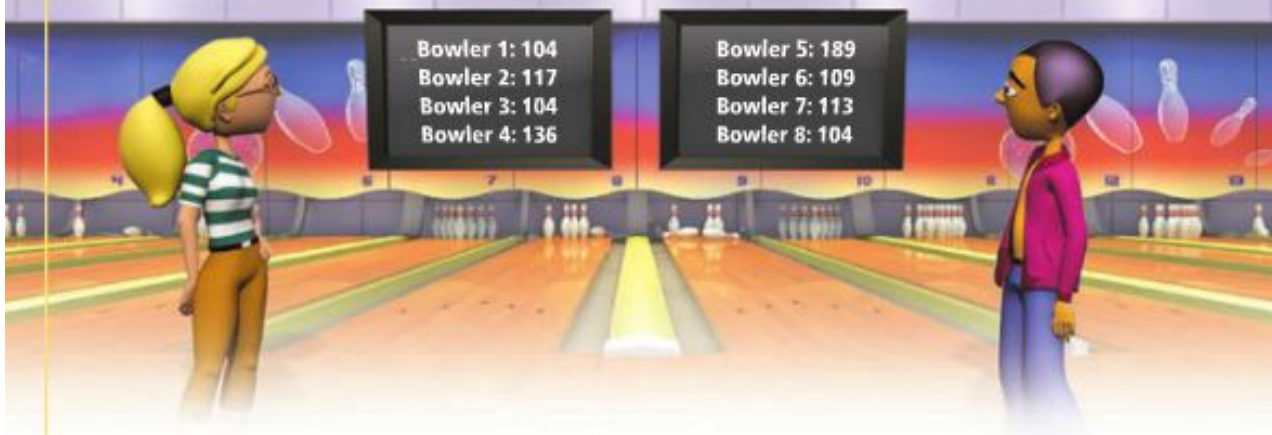
Use median to describe the middle of a set of data that *does* have an outlier.

Use mode when the data are nonnumeric or when choosing the most popular item.



## Problem 1 Finding Measures of Central Tendency

**Bowling** What are the mean, median, and mode of the bowling scores below? Which measure of central tendency best describes the scores?



$$\text{Mean: } \frac{104 + 117 + 104 + 136 + 189 + 109 + 113 + 104}{8} = 122$$

The mean is the sum of the scores divided by the number of scores.

Median: 104 104 104 109 113 117 136 189

List the data in order.

$$\frac{109 + 113}{2} = 111$$

The median of an even number of data values is the mean of the two middle data values.

Mode: 104

The mode is the data item that occurs the most times.

Because there is an outlier, 189, the median is the best measure to describe the scores. The mean, 122, is greater than most of the scores. The mode, 104, is the lowest score. Neither the mean nor the mode describes the data well. The median best describes the data.

Find the measures of central tendency for the following data

Weight in pounds of boxes loaded onto a semi truck: 201, 201, 200, 199, 199

# Find the measures of central tendency for the following data

Weight in pounds of boxes loaded onto a semi truck: 201, 201, 200, 199, 199

## Mean:

add the data values and divide by the number of data values

$$\frac{201+201+200+199+199}{5} = \frac{1000}{5} = 200$$

## Median:

middle value or mean of two middle values of a set of ordered data

199 199 (200) 201 201

middle  
value

## Mode:

value(s) that appear most often in a set of data

(201) (201) 200 (199) (199)

199 and 201 appear most often in the set of data



## Problem 2 Finding a Data Value

**Grades** Your grades on three exams are 80, 93, and 91. What grade do you need on the next exam to have an average of 90 on the four exams?

$$\frac{80 + 93 + 91 + x}{4} = 90 \quad \text{Use the formula for the mean. Let } x = \text{ the grade on the fourth exam.}$$

$$\frac{264 + x}{4} = 90 \quad \text{Simplify the numerator.}$$

$$264 + x = 360 \quad \text{Multiply each side by 4.}$$

$$x = 96 \quad \text{Subtract 264 from each side.}$$

Your grade on the next exam must be 96 for you to have an average of 90.

---

Your grades on three exams are 71, 90, 82.

What grade do you need on the next exam to have an average of 80 on the four exams?

1. Use the formula for mean
2. Let  $x$  represent the fourth exam
3. Solve for  $x$

Find the value of  $x$  such that the data set has the given mean.

1) 3.8, 4.2, 5.3,  $x$ ; mean 4.8

2) 99, 86, 76, 95,  $x$ ; mean 91

A **measure of dispersion** describes how *dispersed*, or spread out, the values in a data set are. One measure of dispersion is *range*. The **range of a set of data** is the difference between the greatest and least data values.

\*\*Measures of dispersion are also called measures of spread or measures of variation





### Problem 3 Finding the Range

**Finance** The closing prices, in dollars, of two stocks for the first five days in February are shown below. What are the range and mean of each set of data? Use the results to compare the data sets.

Stock A: 25 30 30 47 28

$$\text{range: } 47 - 25 = 22$$

$$\begin{aligned}\text{mean: } & \frac{25 + 30 + 30 + 47 + 28}{5} \\ & = \frac{160}{5} = 32\end{aligned}$$

Stock B: 34 28 31 36 31

$$\text{range: } 36 - 28 = 8$$

$$\begin{aligned}\text{mean: } & \frac{34 + 28 + 31 + 36 + 31}{5} \\ & = \frac{160}{5} = 32\end{aligned}$$

Both sets of stock prices have a mean of 32. The range of the prices for Stock A is 22, and the range of the prices for Stock B is 8. Both stocks had the same average price during the 5-day period, but the prices for Stock A were more spread out.

---

Adding the same amount to each value in a set of data has special consequences for the mean, median, mode, and range.

Consider the data set 5, 16, 3, 5, 11.

mean: 8    median: 5    mode: 5    range: 13

If you add 5 to each data value, you get the data set 10, 21, 8, 10, 16.

mean: 13    median: 10    mode: 10    range: 13

Notice that the mean, median, and mode all increased by 5. The range did not change. For any data set, if you add the same amount  $k$  to each item, the mean, median, and mode of the new data set also increase by  $k$ . The range does not change.



## Problem 4 Adding a Constant to Data Values

**Athletics** The table shows the times several athletes spend on a treadmill each day during the first week of training. The athletes add 5 min to their training times during the second week. What are the mean, median, mode, and range of the times for the second week?

**Step 1** Find the mean, median, mode, and range for the first week.

$$\text{mean: } \frac{20 + 20 + 20 + 30 + 41 + 50 + 50}{7} = 33$$

$$\text{median: } 30 \quad \text{mode: } 20 \quad \text{range: } 50 - 20 = 30$$

**Step 2** Find the mean, median, mode, and range for the second week.

$$\text{mean: } 33 + 5 = 38$$

$$\text{median: } 30 + 5 = 35$$

$$\text{mode: } 20 + 5 = 25$$

$$\text{range: } 30$$

Add 5 to each measure of central tendency.

The range does not change.

### Time on Treadmill

Athlete	Time (min)
Bob	50
Carlota	20
Juan	41
Manuel	20
Rosita	30
Sonia	20
Xavier	50



## Problem 4 Adding a Constant to Data Values

**Athletics** The table shows the times several athletes spend on a treadmill each day during the first week of training. The athletes add 5 min to their training times during the second week. What are the mean, median, mode, and range of the times for the second week?

**Step 1** Find the mean, median, mode, and range for the first week.

$$\text{mean: } \frac{20 + 20 + 20 + 30 + 41 + 50 + 50}{7} = 33$$

$$\text{median: } 30 \quad \text{mode: } 20 \quad \text{range: } 50 - 20 = 30$$

**Step 2** Find the mean, median, mode, and range for the second week.

$$\text{mean: } 33 + 5 = 38$$

$$\text{median: } 30 + 5 = 35$$

$$\text{mode: } 20 + 5 = 25$$

$$\text{range: } 30$$

Add 5 to each measure of central tendency.

The range does not change.

**Time on Treadmill**

Athlete	Time (min)
Bob	50
Carlota	20
Juan	41
Manuel	20
Rosita	30
Sonia	20
Xavier	50

In the third week of training the athletes add 10 minutes to their training. What would mean, median, mode and range be?



## Problem 5 Multiplying Data Values by a Constant

**Shopping** A store sells seven models of televisions. The regular prices are \$144, \$479, \$379, \$1299, \$171, \$479, and \$269. This week the store offers a 30% discount on all televisions. What are the mean, median, mode, and range of the discounted prices?

**Step 1** Find the mean, median, mode, and range of the regular prices.

$$\text{mean: } \frac{144 + 171 + 269 + 379 + 479 + 479 + 1299}{7} = 460$$

$$\text{median: } 379 \quad \text{mode: } 479 \quad \text{range: } 1299 - 144 = 1155$$

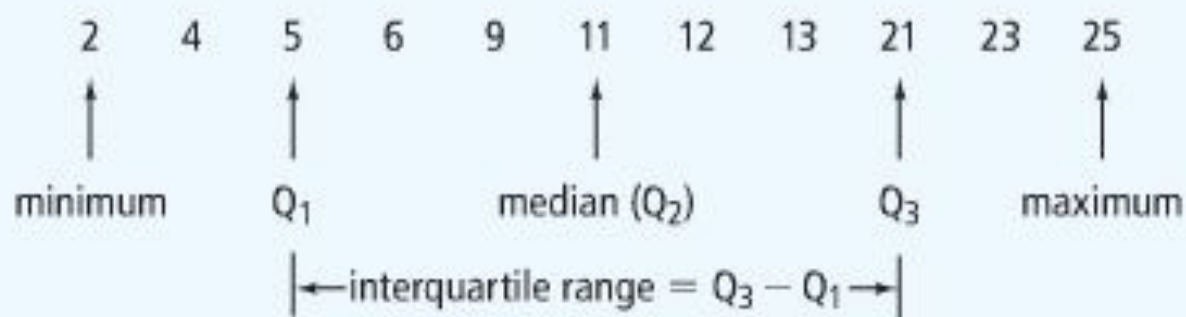
**Step 2** Multiply the mean, median, mode, and range in Step 1 by 0.7 to find the mean, median, mode, and range of the discounted prices.

$$\text{mean: } 460(0.7) = 322 \quad \text{mode: } 479(0.7) = 335.30$$

$$\text{median: } 379(0.7) = 265.30 \quad \text{range: } 1155(0.7) = 808.50$$

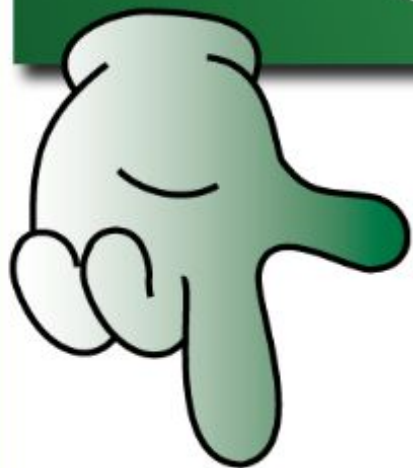
**Essential Understanding** Separating data into subsets is a useful way to summarize and compare data sets.

**Quartiles** are values that divide a data set into four equal parts. The median (or second quartile,  $Q_2$ ) separates the data into upper and lower halves. The first quartile ( $Q_1$ ) is the median of the lower half of the data. The third quartile ( $Q_3$ ) is the median of the upper half of the data. The **interquartile range** is the difference between the third and first quartiles.



For a set of data that has an odd number of values, you do not include the median in either half when finding the first and third quartiles.

# Outlier



A data value that is distinctly separate from the rest of the data.

One definition of outlier is any data value that is more than **1.5 interquartile ranges** (IQR) below the first quartile or above the third quartile.

# Outlier video

Click [HERE](#)



Distance traveled in miles to visit relatives during winter break:

210, 45, 10, 108, 452, 225, 35, 95, 140, 25, 65, 250

1. Find the mean, median, and mode of the data set.
2. Identify if there are any outliers.
3. If there is an outlier, find the mean, median, and mode without the outlier.
4. If there is an outlier, which measure was affected more by removing the outlier?

Time in minutes spent on social networking sites per day:

25, 35, 45, 30, 65, 50, 25, 100, 45, 35, 5, 105, 110, 190, 40, 30, 80

1. Find the mean and median of the data set.
2. Identify if there are any outliers.
3. If there is an outlier, find the mean and median without the outlier.
4. If there is an outlier, which measure was affected more by removing the outlier?

# Standard Deviation



Is a measure of how the values in a data set vary, or deviate, from the mean

Statisticians use several special symbols in the formula for standard deviation.

The Greek letter sigma ( $\sigma$ ) represents standard deviation.

The capital sigma ( $\Sigma$ ) represents the sum of a series of numbers.

$$\sigma = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n}}$$

$x$  is a value in the data set.  
 $\bar{x}$  is the mean of the data set.

$n$  is the number of values in the data set.

Step 1 - Have a data set

Step 2 - Find the mean

Step 3 - Subtract the mean from the first number in the data set, then square the difference

Step 4 - Continue with each number in your data set

Step 5 - Find the mean of the differences

Step 6 - Take the square root of the mean

Find the standard deviation given the data set:

3, 6, 9, 16

# Standard Deviation Video

Click [HERE](#)

## Exercises

Find the mean and standard deviation of each data set. Round to the nearest hundredth.

1. 4 8 5 12 3 9 5 2

2. 102 98 103 86 101 110

3. 8.2 11.6 8.7 10.6 9.4 10.1 9.3

## Exercises

Find the mean and standard deviation of each data set. Round to the nearest hundredth.

1. 4 8 5 12 3 9 5 2

2. 102 98 103 86 101 110

3. 8.2 11.6 8.7 10.6 9.4 10.1 9.3

## Exercises

1. 6, 3.16

2. 100, 7.23

3. 9.7, 1.07



# What does standard deviation tell us?

Let's say the class had an average of 80% and a small standard deviation. That would mean that most people in the class had a score really close to an 80%.

If the class had an average of 80% but a large standard deviation, that would mean that even though the average was 80%, there are lots of people who got scores that were not close to 80%. Some people probably did really well, while others probably struggled.

Standard deviation tells me if my data is close to the mean, or really spread out.

Standard deviation is a measure of spread (like range).

# Released EOC Questions

- 25 The table below shows the area of several states.

<b>State</b>	<b>Area</b> (thousands of square miles)
Connecticut	6
Georgia	59
Maryland	12
Massachusetts	11
New Hampshire	9
New York	54
North Carolina	54
Pennsylvania	46

Delaware has an area of 2,000 square miles. Which is true if Delaware is included in the data set?

- A The mean increases.
- B The range decreases.
- C The interquartile range decreases.
- D The standard deviation increases.

- 36 The number of points scored by a basketball player in the first eight games of a season are shown below.

15, 35, 18, 30, 25, 21, 32, 16

What would happen to the data distribution if she scored 24, 22, 27, and 28 points in her next four games?

- A The data distribution would become less peaked and more widely spread.
- B The data distribution would become less peaked and less widely spread.
- C The data distribution would become more peaked and less widely spread.
- D The data distribution would become more peaked and more widely spread.

A set of nine data points is shown below.

8, 11, 12, 10, 9, 7, 5, 3, 9
------------------------------

Which statement is true if a tenth data point of 45 is added to the data set?

- A The mean and median will both increase.
- B The mean will increase and the median will decrease.
- C The mean will increase and the median will remain the same.
- D The mean and median will both decrease.

33 Abby scored 87, 93, 96, and 89 on her first four history quizzes. What score does Abby need to get on her fifth quiz to have an average of exactly 91 on her history quizzes?

- A 90
- B 94
- C 98
- D 100

42 The table below shows the weights of 8 different bears at a zoo.

Type of Bear	Weight (pounds)
Asiatic Black Bear	225
Black Bear	300
Brown Bear	550
Panda Bear	200
Polar Bear	1,000
Sloth Bear	300
Spectacled Bear	280
Sun Bear	100

If the weight of the polar bear is removed, which statement is true?

- A The mean decreases more than the median because the polar bear is a high outlier.
- B The mean decreases less than the median because the polar bear is a high outlier.
- C The mean decreases more than the median because the high value balances the low value.
- D The mean decreases less than the median because the high value balances the low value.

# Homework

Page 730 #7-10, 13, 15, 19, 28, 31

Page 733 #1, 2, 3