

# STUDY ISLAND FOCUS DAY: STATISTICAL ANALYSIS

# HOMework:

- MAKE SURE 10 questions have been completed in STATISTICAL ANALYSIS
- [www.studyisland.com](http://www.studyisland.com)

TRY 10 TRY 10 TRY 10

TRY 10 TRY 10 TRY 10

TRY 10 TRY 10 TRY 10

# OBJECTIVES

- Students will Describe any overall pattern and any deviations from the overall pattern with reference to the context in which the data were gathered.
- Students will complete 10 questions in Statistical Analysis

**TRY 10 TRY 10 TRY 10**

**TRY 10 TRY 10 TRY 10**

**TRY 10 TRY 10 TRY 10**

# Statistical Analysis

## **Mean**

*The sum of all the values in a data set divided by the number of data values.*

*Also called the average.*

## **Median**

*For an ordered data set with an odd number of values, the median is the middle value.*

*For an ordered set with an even number of values, the median is the average of the two middle values.*

## **Mean Absolute Deviation**

*For a set of data, the mean absolute deviation is the mean of the differences of the values in the data set from the mean.*

## **Interquartile Range**

*The difference between the upper and lower quartiles in the data set.*

## Example

The purpose of a measure of central tendency is to provide a representative value of a set of data. The intended application of that value should also be considered when choosing a measure of center.

The ages of applicants at a driver's license office for one day are shown in the table below.

Ages of Applicants				
17	16	18	16	16
18	18	17	45	16

### Mean

The mean is the average of a set of data.

An outlier in a set of data skews the mean.

Mean is a good measure of center when there are no outliers.

The mean of this data is 19.7. With the outlier 45 removed, the mean of this data is approximately 16.9.

If asked the most representative measure of center for this data, mean might not be the best choice.

### Median

The median is the middle value of a set of data.

Median is a good measure of center when there are outliers, as they tend not to affect the median.

The median of this data is 17.

If asked the most representative measure of center for this data, median would be a good choice.

If there are no outliers, and the mean and median are the same or close to the same, then either is a good measure of center.

If more specific reasoning is given, like wanting to know the average of the data, or wanting to know something about the upper or lower half of the data, then that would help determine if mean or median is the preferred choice.

### Mean Absolute Deviation

The mean absolute deviation is the mean of the differences of the values in the data set from the mean. It measures spread.

The mean absolute deviation of this data is 5.06.

### Interquartile Range

The interquartile range is the difference between the upper and lower quartiles. It measures spread.

The interquartile range of this data is 2.

## **MEAN**

*The mean is the average value of all data in a set.*

*Add up the numbers then divide by the number of values in the set to find the mean.*

### **Example 1:**

Find the mean of the following set of numbers: { 66, 72, 83, 89 }

## **MEDIAN**

*The median is the value that has exactly half the data above it and half below it.  
To find the median, order the numbers from smallest to largest.  
The middle number is the median.*

### **Example 2:**

Find the median of the following set of numbers: { 65, 72, 81, 83, 89 }

### **Solution:**

**Median** = the middle number from smallest to largest

# Mean Absolute Deviation:





**Mean absolute deviation** can be used to measure **variability** in a set of data. It is the mean of the differences of the values in the data set from the mean. It is best used to measure spread when the numbers are farther apart.

To find the mean absolute deviation:

1. Calculate the mean of the data set.
2. Find the absolute values of the differences between each value in the data set and the mean.
3. Find the sum of the absolute values.
4. Divide the sum of the absolute values by the number of values in the data set.

### Example 1:

Find the mean absolute deviation of the data set below.

16, 24, 21, 18, 21

### Solution:

First, find the mean of the data set.

$$\frac{16 + 24 + 21 + 18 + 21}{5} = 20$$

Next, find the sum of the absolute values of the difference between each data point and the mean.

$$|16-20| + |24-20| + |21-20| + |18-20| + |21-20| = 12$$

Finally, divide this sum by the total number of values in the data set.

$$12 \div 5 = 2.4$$

Therefore, the mean absolute deviation of this data set is **2.4**.

To find the mean absolute deviation:

1. Calculate the mean of the data set.
2. Find the absolute values of the differences between each value in the data set and the mean.
3. Find the sum of the absolute values.
4. Divide the sum of the absolute values by the number of values in the data set.

□

20. A group of friends went bowling last Saturday. Their scores are shown below.

**Bowling Scores**

Friend	Score
Brad	121
Alicia	196
Jonas	100
Katy	210
Darius	112
Stephanie	170
Mike	222
Beth	161
Trevor	201
Laura	107

What is the mean absolute deviation of their bowling scores?

- A. 30
- B. 20
- C. 50
- D. 40

To find the mean absolute deviation:

1. Calculate the mean of the data set.
2. Find the absolute values of the differences between each value in the data set and the mean.
3. Find the sum of the absolute values.
4. Divide the sum of the absolute values by the number of values in the data set.

23. The ages of the students in the Children's Theater Company are shown below.

**Children's Theater Company**

<b>Student</b>	<b>Age</b>
Billy	12
Angela	9
Jacob	12
Isabelle	13
Daniel	15
Sarah	8
Marc	11
Amy	14
Todd	10
Lisa	16

What is the mean absolute deviation of their ages?

- A. 12
- B. 1
- C. 8
- D. 2

The lower quartile identifies the 25<sup>th</sup> percentile and the upper quartile identifies the 75<sup>th</sup> percentile.

The difference between the upper and lower quartiles is known as the **interquartile range**.

This is also used to measure **variability**. It is best used to measure spread when the numbers are closer together.

## Example 2:

Find the interquartile range of the data set below.

56, 51, 59, 54, 53, 57, 52

## Solution:

First, find the median. To find the median, order the number from least to greatest and find the middle number.

51, 52, 53, 54, 56, 57, 59

Next, find the upper and lower quartiles. The upper quartile is the median of the numbers *above* the median, and the lower quartile is the median of the numbers *below* the median.

51, 52, 53, 54, 56, 57, 59

The interquartile range is the difference between the upper quartile, 57, and the lower quartile, 52.

$$57 - 52 = 5$$

So, the interquartile range of the data set is **5**.

1. The heights in inches of seven basketball players are listed below.

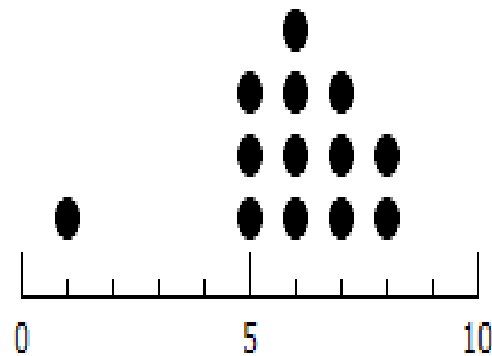
70, 74, 80, 88, 78, 85, 81

What is the interquartile range of the heights?

- A. 11
- B. 79
- C. 12
- D. 10

*Striking deviations, or outliers, are data points that stick out from the rest of the data.*

**Example:**

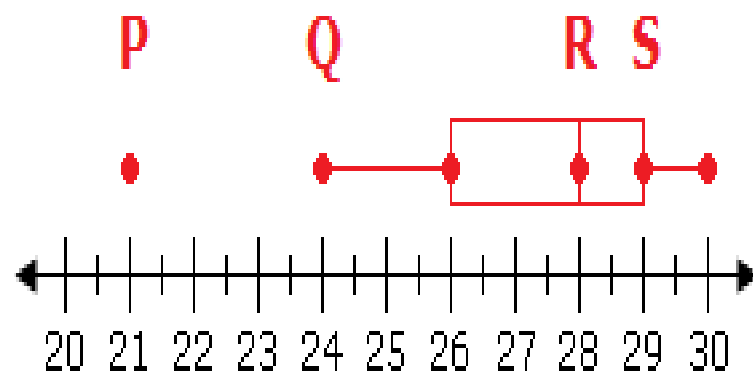


Where is there a striking deviation in the dot plot above?

**Solution:**

Looking at the dot plot, the point that is separated from the majority of the data is at 1.

3.



Which of the following points represents a striking deviation in the box plot above?

- A. P
- B. R
- C. S
- D. Q

# DIRECTIONS (this is your homework if you do not finish in class):

- Sign into [www.studyisland.com](http://www.studyisland.com)
- Click + Sign by ALL SIXTH GRADE (this will show your pathways)
- Click 6.5 Math Pathway
- Click Statistical Analysis
- APPLICATION SHARE AS YOUR WORK

TRY 10 TRY 10 TRY 10

TRY 10 TRY 10 TRY 10

TRY 10 TRY 10 TRY 10