# 6.1 - Measures of Central Tendency and Variation

## Objectives

Find measures of central tendency and measures of variation for statistical data.

Identify any outliers and describe how it affects the data.

## Vocabulary

- variance
- standard deviation
- outlier

## Example 1: Finding Measures of Central Tendency

**Find the mean, median, and mode of the data.**

**deer at a feeder each hour: 3, 0, 2, 0, 1, 2, 4**

**Mean:** $\frac{3+0+2+0+1+2+4}{7} = \frac{12}{7} \approx 1.7$ deer

**Median:** 0 0 1 2 2 3 4 = 2 deer

**Mode:** The most common results are 0 and 2.

**In TI 30XS–**

**DATA**
**Enter in L1**
**$2^{nd}$ DATA**
**1-Var Stats**
**L1**
**ONE**
**CALC → 2: x (mean)**
**→ 9: Med (median)**

## You Try! Example 2

**Find the mean, median, and mode of the data set.**

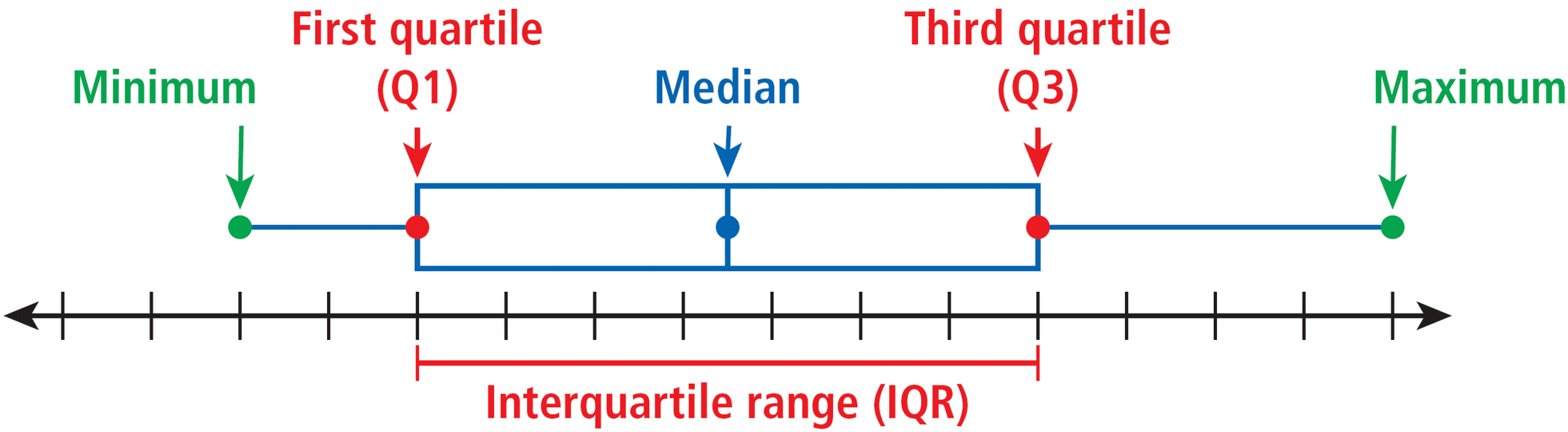
**{6, 9, 3, 8}**

**Mean:** $\dfrac{6+9+3+8}{4} = \dfrac{26}{4} = 6.5$

**Median:** 3 6 ¦ 8 9 $\dfrac{6+8}{2} = 7$

**Mode:** None

A *box-and-whisker plot* shows the spread of a data set. It displays 5 key points: the **minimum** and **maximum** values, the **median**, and the **first** and **third quartiles**.



The quartiles are the medians of the lower and upper halves of the data set. If there are an odd number of data values, do not include the median in either half.

The *interquartile range*, or IQR, is the difference between the 1st and 3rd quartiles, or Q3 – Q1. It represents the middle 50% of the data.

## Example 3: Making a Box-and-Whisker Plot and Finding the Interquartile Range

**Make a box-and-whisker plot of the data. Find the interquartile range.**
**{6, 8, 7, 5, 10, 6, 9, 8, 4}**

**In CALC –**
**DATA**
**Enter in L1**
**$2^{nd}$ DATA**
**1-Var Stats**
**L1**
**ONE**
**CALC**

**Step 1** Enter values in calculator

**Step 2** Find the minimum, maximum, median, and quartiles.

[4, 5, 6, 6] 7, [8, 8, 9, 10]

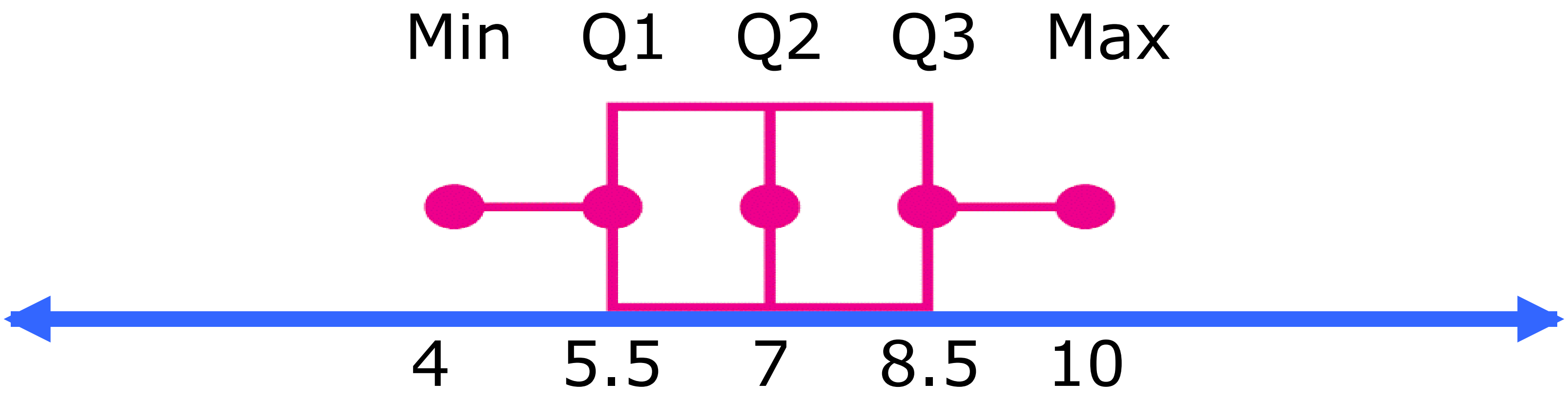
Mimimum

Median

Maximum

First quartile
5.5

Third quartile
8.5




IRQ: 8.5 – 5.5 = 3

The interquartile range is 3, the length of the box in the diagram.

## You Try! Example 4

**Make a box-and-whisker plot of the data. Find the interquartile range. {13, 14, 18, 13, 12, 17, 15, 12, 13, 19, 11, 14, 14, 18, 22, 23}**

**Step 1** Plug values into L1

**Step 2** Find the minimum, maximum, median, and quartiles.

11, 12, 12, 13, 13, 13, 14, 14, 14, 15, 17, 18, 18, 19, 22, 23

Mimimum: 11

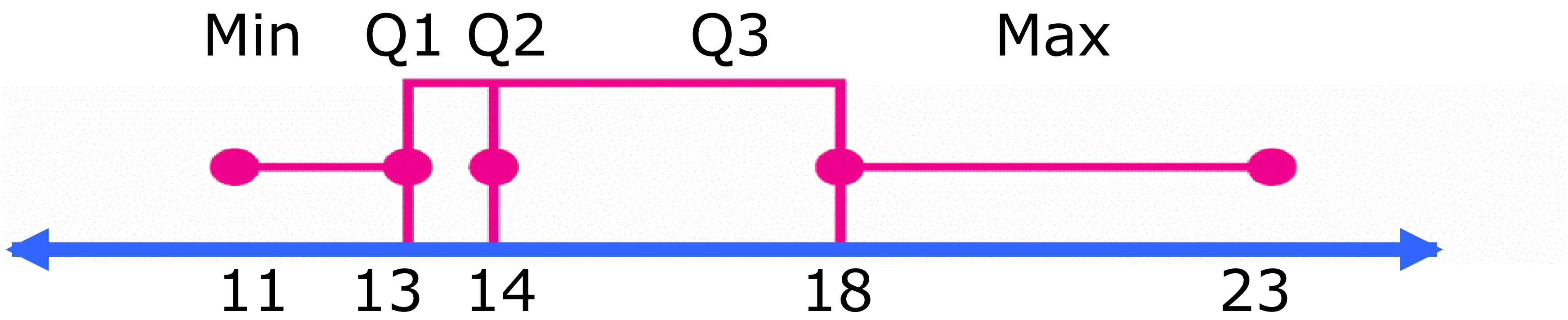
Median: between 14 and 14

Maximum: 23

First quartile: 13

Third quartile: 18

**Step 3** Draw a box-and-whisker plot. Find the IQR




IQR = 18 – 13 = 5

The interquartile range is 5, the length of the box in the diagram.

# Any Questions??

The data sets **{19, 20, 21}** and **{0, 20, 40}** have the same mean and median, but the sets are very different. The way that data are spread out from the mean or median is important in the study of statistics.

A *measure of variation* is a value that describes the spread of a data set. The most commonly used measures of variation are the *range*, the interquartile range, the *variance*, and the *standard deviation*.

## Reading Math

The symbol commonly used to represent the mean is $\overline{x}$, or "$x$ bar." The symbol for standard deviation is the lowercase Greek letter *sigma*, $\sigma$.

The **variance**, denoted by $\sigma^2$, is the average of the squared differences from the mean. **Standard deviation**, denoted by $\sigma$, is the square root of the variance and is one of the most common and useful measures of variation.

Low standard deviations indicate data that are clustered near the measures of central tendency, whereas high standard deviations indicate data that are spread out from the center.

## Example 5: Finding the Mean and Standard Deviation

**Find the mean and standard deviation for the data set of the number of people getting on and off a bus for several stops.**

**{6, 8, 7, 5, 10, 6, 9, 8, 4}**

**Step 1** Enter data in calculator L1

**Step 2** Locate Mean (2:x) and standard deviation (4: σx)

$$\bar{x} = 7$$

$$\sigma = \sqrt{3.3} \approx 1.83$$

**In CALC –**

**DATA**

**Enter in L1**

**2nd DATA**

**1-Var Stats**

**L1**

**ONE**

**CALC**

The mean is 7 people, and the standard deviation is about 1.83 people.

## You Try! Example 6

**Find the mean and standard deviation for the data set of the number of elevator stops for several rides.**

**{0, 3, 1, 1, 0, 5, 1, 0, 3, 0}**

**Step 1** Find the mean.

$$\bar{x} = \frac{0+3+1+1+0+5+1+0+3+0}{10} = 1.4$$

**Step 2** Find the standard deviation.

$$\sigma = \sqrt{2.64} \approx 1.6$$

**In CALC –**

**DATA**
**Enter in L1**
**$2^{nd}$ DATA**
**1-Var Stats**
**L1**
**ONE**
**CALC**

The mean is 1.4 stops and the standard deviation is about 1.6 stops.

An **outlier** is an extreme value that is much less than or much greater than the other data values. Outliers have a strong effect on the mean and standard deviation. If an outlier is the result of measurement error or represents data from the wrong population, it is usually removed.

There are different ways to determine whether a value is an outlier. One is to look for data values that are **more than 3 standard deviations from the mean**.

# Example of Outlier

**In the 2003-2004 American League Championship Series, the New York Yankees scored the following numbers of runs against the Boston Red Sox: 2, 6, 4, 2, 4, 6, 6, 10, 3, 19, 4, 4, 2, 3. Identify the outlier, and describe how it affects the mean and standard deviation.**

Morton, Spring 2014

# Example of Outlier (continued)

**Step 1: Enter the data values into list L1 on a graphing calculator.**

**Step 2: Find the mean and standard deviation.**

```
1-Var Stats
 x̄=5.357142857
 Σx=75
 Σx²=663
 Sx=4.482566965
 σx=4.319509609
↓n=14
```

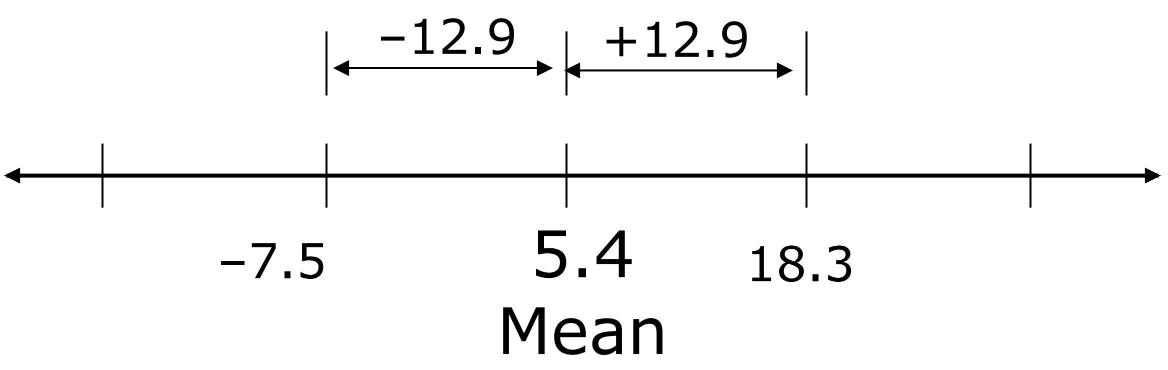
The mean is about 5.4 and the standard deviation is 4.32.

Morton, Spring 2014

# Example of Outlier (continued)

**Step 3: Identify the outliers.** Look for the data values that are more than 3 standard deviations away from the mean in either direction.

Three standard deviations is about 3(4.3) = 12.9.




Values less than -7.5 and greater than 18.3 are outliers, so 19 is an outlier.

# Example of Outlier (continued)

**Step 4: Remove the outlier to see the effect that it has on the mean and standard deviation.**

**All data**

```
1-Var Stats
 x̄=5.357142857
 Σx=75
 Σx²=663
 Sx=4.482566965
 σx=4.319509609
↓n=14
```

**Without outlier**

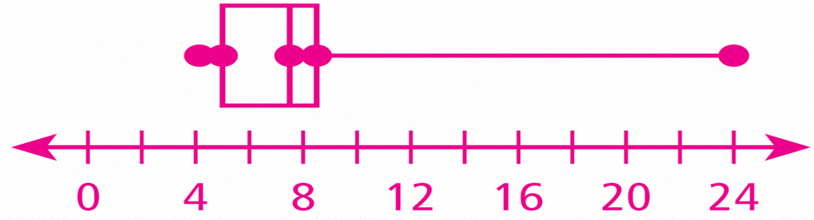
```
1-Var Stats
 x̄=4.307692308
 Σx=56
 Σx²=302
 Sx=2.250356097
 σx=2.162072203
↓n=13
```

The mean decreased from 5.4 to 4.3 and the standard deviation from 4.3 to 2.2.

## PRACTICE PROBLEM!!

**Use the data set for 1-3:**
**{9, 4, 7, 8, 5, 8, 24, 5}**

**1.** Make a box-and-whisker plot of the data. Find the interquartile range.




IQR: 3.5

**2.** Find the variance and the standard deviation of the data set.

var: ≈ 35.94; std. dev: ≈ 5.99

**3.** Are there any outliers in the data set?

No! Outliers less than -9.22, greater than 26.72