

## Slides from Measures of Center and Dispersion (Tutorial 25)

### Measures of Center and Dispersion

Fred Dillon

This handout contains selected slides to use when reviewing this tutorial topic with or without the video. To access all slides, open thumbnail link on the tutorial interface.

### Measures of Center

The **mode** is the item(s) occurring most often in a data set.

- There may be more than one mode for a data set
- Best for categorical data
- Easily influenced when additional data are added

### Measures of Center

The **median** is the middle value when the item(s) is arranged in numerical order.

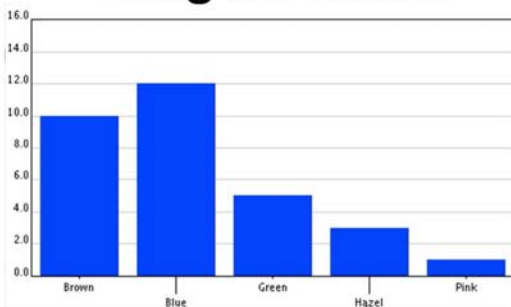
- Useful for describing data with outliers as median is little affected by outliers

### Measures of Center

The **mean**, or arithmetic mean, is the sum of all the data items in the data set divided by the number of values.

- Most commonly used measure of center
- Easily affected by outliers

### Using the Mode



“Blue” is the **mode** of the data set.

### Finding the Median

2, 3, 4, 5, 5, 6, 8, 9, 9, 10, 12

**median = 6**

## Slides from Measures of Center and Dispersion (Tutorial 25)

### Finding the Median

2, 3, 4, 5, 5, 6, 7, 8, 9, 9, 10, 12

$$\text{median} = \frac{6+7}{2} = 6.5$$

There is an even number of values in the data set.

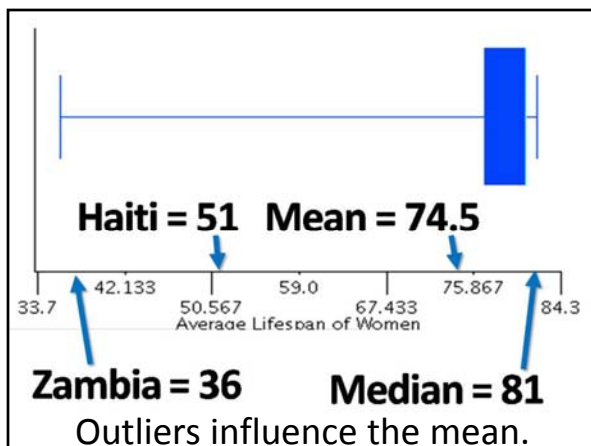
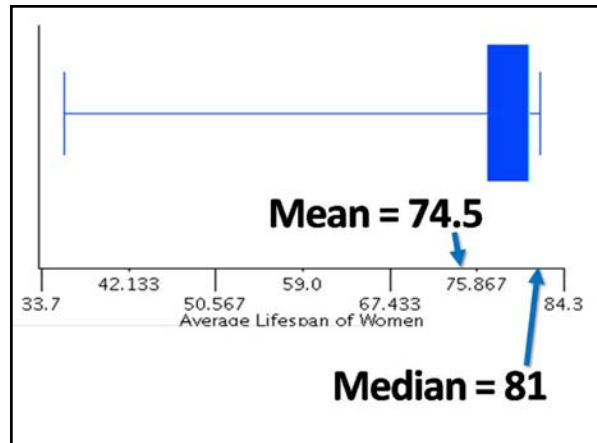
6.5 is the mean of the two middle values

## Median versus Mean

### Average Life Span for Women

Britain	79	Netherlands	81
France	82	Norway	81
Germany	79	Russia	70
Greece	81	Spain	81
Haiti	51	Sweden	82
Italy	81	Switzerland	82
Mexico	77	United States	79
		Zambia	36

Source: <http://www.efmoody.com/longterm/lifespan.html>



## Measures of Dispersion

*describe the spread of the data*

Three commonly used measures:

- Range
- Interquartile Range
- Standard Deviation

## Slides from Measures of Center and Dispersion (Tutorial 25)

### Average Lifespan for Women

Britain	79	Netherlands	81
France	82	Norway	81
Germany	79	Russia	70
Greece	81	Spain	81
Haiti	51	Sweden	82
Italy	81	Switzerland	82
Mexico	77	US	79

Note: Data from Zambia is omitted.

Source: <http://www.efmoody.com/longterm/lifespan.html>

### Measures of Dispersion

The **range** is the difference between the greatest (maximum) and least (minimum) values in a set of data.

### Average Lifespan for Women

Britain	79	Netherlands	81
France	82	Norway	81
Germany	79	Russia	70
Greece	81	Spain	81
Haiti	51	Sweden	82
Italy	81	Switzerland	82
Mexico	77	US	79

maximum: 82      minimum: 51  
**range: 82 - 51 = 31**  
**The range is 31 years.**

Source: <http://www.efmoody.com/longterm/lifespan.html>

### Interquartile Range

The **interquartile range (IQR)** is the difference of the third (upper) and first (lower) quartiles.

### Average Lifespan for Women

Britain	79	Netherlands	81
France	82	Norway	81
Germany	79	Russia	70
Greece	81	Spain	81
Haiti	51	Sweden	82
Italy	81	Switzerland	82
Mexico	77	US	79

51, 70, 77, 79, 79, 79, 79, 81, 81, 81, 81, 81, 82, 82, 82

**Median = 81**

Source: <http://www.efmoody.com/longterm/lifespan.html>

### Finding the Interquartile Range

51, 70, 77, 79, 79, 79, 79, 81, 81, 81, 81, 81, 82, 82, 82

51, 70, 77, 79, 79, 79, 81

median of the lower half of the data

**The first (lower) quartile,  $Q_1$ , is 79**

**$Q_1 = 79$**

Source: <http://www.efmoody.com/longterm/lifespan.html>

# Tutorials for High School Mathematics

## Slides from Measures of Center and Dispersion (Tutorial 25)

*Finding the Interquartile Range*  
 51, 70, 77, 79, 79, 79, 81, 81, 81, 81, 81, 82, 82, 82

81, 81, 81, 81, 82, 82, 82

median of the upper half of the data

The third (upper) quartile,  $Q_3$ , is 81

$Q_3 = 81$

Source: <http://www.efmoody.com/longterm/lifespan.html>

*Finding the Interquartile Range*  
 51, 70, 77, 79, 79, 79, 81, 81, 81, 81, 81, 82, 82, 82

$Q_1$       median       $Q_3$

$IQR = Q_3 - Q_1 = 81 - 79 = 2$

**Interquartile Range = 2**

An outlier does not effect the median or the IQR.

### Interquartile Range

- Identifies the middle 50% of the data
- Not affected by one or two outliers

### Standard Deviation

The *standard deviation* is the “average distance from the mean” for each data point.

### Standard Deviation

The mean of the life span data is approximately 77.6 years.

To find the distance each data point is from the mean, subtract the 77.6 from each data point.

Average Life Span for Women	Distance from mean
Britain 79	$79 - 77.6 = 1.4$
France 82	$82 - 77.6 = 4.4$
Germany 79	$79 - 77.6 = 1.4$
Greece 81	$79 - 77.6 = 1.4$
Haiti 51	$51 - 77.6 = -26.6$
Italy 81	$81 - 77.6 = 3.4$
Mexico 77	$77 - 77.6 = -0.6$
Netherlands 81	$81 - 77.6 = 3.4$
Norway 81	$81 - 77.6 = 3.4$
Russia 70	$70 - 77.6 = -7.6$
Spain 81	$81 - 77.6 = 3.4$
Sweden 82	$82 - 77.6 = 4.4$
Switzerland 82	$82 - 77.6 = 4.4$
United States 79	$79 - 77.6 = 1.4$

## Slides from Measures of Center and Dispersion (Tutorial 25)

Average Life Span for Women		Distance from mean
Britain	79	$79 - 77.6 = 1.4$
France	82	$82 - 77.6 = 4.4$
Germany	79	$79 - 77.6 = 1.4$
Greece	81	$79 - 77.6 = 1.4$
Haiti	51	$51 - 77.6 = -26.6$
Italy	81	$81 - 77.6 = 3.4$
Mexico	77	$77 - 77.6 = 1.4$
Netherlands	81	$81 - 77.6 = 3.4$
Norway	81	$81 - 77.6 = 3.4$
Russia	70	$70 - 77.6 = -7.6$
Spain	81	$81 - 77.6 = 3.4$
Sweden	82	$82 - 77.6 = 4.4$
Switzerland	82	$82 - 77.6 = 3.4$
United States	79	$79 - 77.6 = 1.4$

**Sum = -0.4**

Distance from mean	Square of differences
$79 - 77.6 = 1.4$	1.96
$82 - 77.6 = 4.4$	19.36
$79 - 77.6 = 1.4$	1.96
$79 - 77.6 = 1.4$	1.96
$51 - 77.6 = -26.6$	707.56
$81 - 77.6 = 3.4$	11.56
$77 - 77.6 = 1.4$	1.96
$81 - 77.6 = 3.4$	11.56
$81 - 77.6 = 3.4$	11.56
$70 - 77.6 = -7.6$	57.76
$81 - 77.6 = 3.4$	11.56
$82 - 77.6 = 4.4$	19.36
$82 - 77.6 = 3.4$	11.56
$79 - 77.6 = 1.4$	1.96

**Sum = 887.4286**

### Variance

The *variance* is the arithmetic mean of squared distances.

### Variance

The *variance* is the arithmetic mean of squared distances.

$$\text{Variance} = \frac{887.4286}{14} \approx 63.3878$$

*sum of squared distances from the mean divided by the number of data points*

### Standard Deviation

The *standard deviation* is the square root of the *variance*.

### Standard Deviation

The *standard deviation* is the square root of the *variance*.

$$\text{Standard Deviation} \approx \sqrt{63.3878} \approx 7.962$$

*find the square root to "undo" the squaring of the distances*

## Slides from Measures of Center and Dispersion (Tutorial 25)

### Formula for Standard Deviation

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

symbol for standard deviation (sigma)

### Formula for Standard Deviation

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

find the sum of values obtained when the calculation that follows is performed on each data point

### Formula for Standard Deviation

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

data point

number of data points

mean

### Formula for Standard Deviation

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

## Standard Deviation

### Ways of using standard deviation:

- The percent of the data within one standard deviation of the mean
- The size of the standard deviation relative to the size of the data

A large standard deviation in relation to the data size indicates a large spread of the data.

## Using a Calculator to Find the Measures of Center and Dispersion

# Tutorials for High School Mathematics

## Slides from Measures of Center and Dispersion (Tutorial 25)

Data set: 7, 10, 13, 3, 6

- Enter the data in List 1 (L1)
- Use the following keystrokes to find the summary statistics
  - ✓ Hit STAT button
  - ✓ Arrow over to CALC
  - ✓ Select 1: 1 - Var Stats
  - ✓ Hit ENTER twice
  - ✓ Screen shows  $\bar{x}$  and other statistics
  - ✓ Arrow down

```
1-Var Stats
↑n=5
minX=3
Q1=4.5
Med=7
Q3=11.5
maxX=13
```

```
1-Var Stats
↑n=5
minX=3
Q1=4.5
Med=7
Q3=11.5
maxX=13
```

The **range** is the difference between the “maxX” and “minX”.

The **interquartile range** is found by subtracting  $Q_3 - Q_1$ .

```
1-Var Stats
x̄=7.8
Σx=39
Σx²=363
Sx=3.834057903
σx=3.42928564
↓n=5
```

The **standard deviation** is  $\sigma_x$ .

The **variance** can be found by squaring the standard deviation.

## Summary

What each tells us about the data and how to calculate:

- the measures of center (mode, median and mean)
- the measures of dispersion (range, interquartile range and standard deviation)