## Measures of Spread

**A measure of spread**

sometimes also called a **measure** of **dispersion(how data is distributed)**

**It**  is used **to** describe **the** variability in a sample or population.

**It** is usually used in conjunction with a **measure** **of central tendency**,

such as the mean or median,

**To** provide an overall description of a set of data.

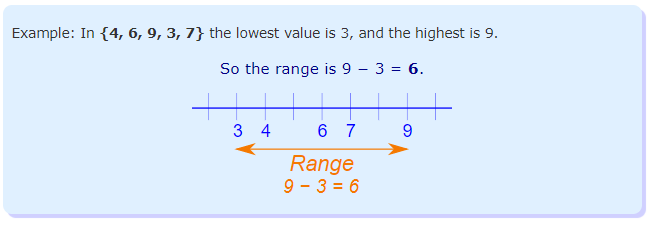
(used in quantitative data)

**Type of MEasure of Spread**

* Range
* Quartiles Range
* Interquartile Range
* Percentile
* Mean Deviation
* Standard deviation and Variance

**The Range (Statistics)**

The Range is the difference between the lowest and highest values.

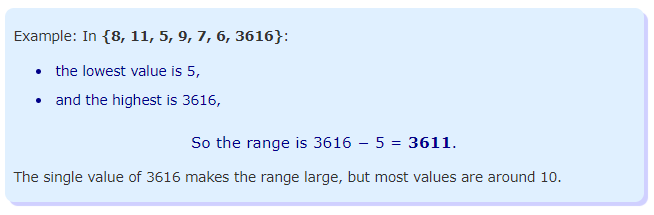


It is that simple!

But perhaps too simple ...

## The Range Can Be Misleading

The range can sometimes be misleading when there are extremely high or low values.



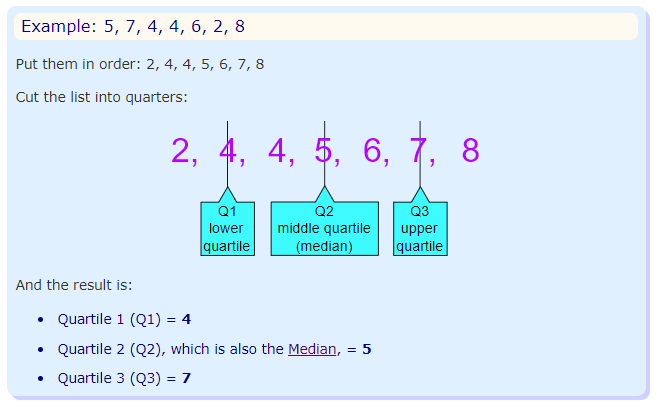
So we may be better off using Interquartile Range or Standard Deviation.

**Quartiles**

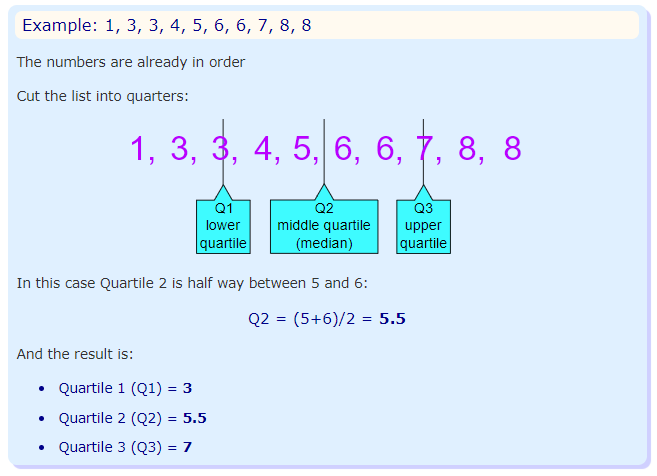
Quartiles are the values that divide a list of numbers into quarters:

* Put the list of numbers **in order**
* Then cut the list into **four equal parts**
* The Quartiles are at the "cuts"

Like this:

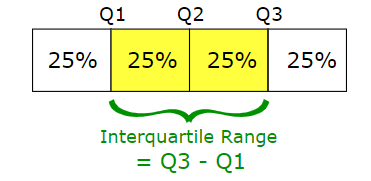


Sometimes a "cut" is between two numbers ... the Quartile is the average of the two numbers.

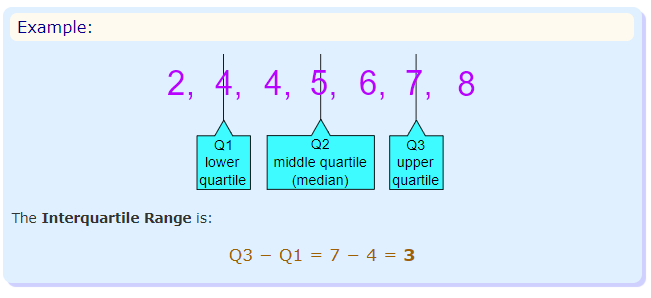


## Interquartile Range

The "Interquartile Range" is from Q1 to Q3:



To calculate it just **subtract Quartile 1 from Quartile 3**, like this:



## Interquartile range gives another measure of variability. It is a better measure of dispersion than range because it leaves out the extreme values. It equally divides the distribution into four equal parts called quartiles. First 25% is 1st quartile (Q1), last one is 3rd quartile (Q3) and middle one is 2nd quartile (Q2).

## 2nd quartile (Q2) divides the distribution into two equal parts of 50%. So, basically it is same as Median.

## 

The interquartile range is the distance between the third and the first quartile, or, in other words, IQR equals Q3 minus Q1

IQR = Q3- Q1

##### **How to calculate IQR**

Step 1: Order from low to high

Step 2: Find the median or in other words Q2

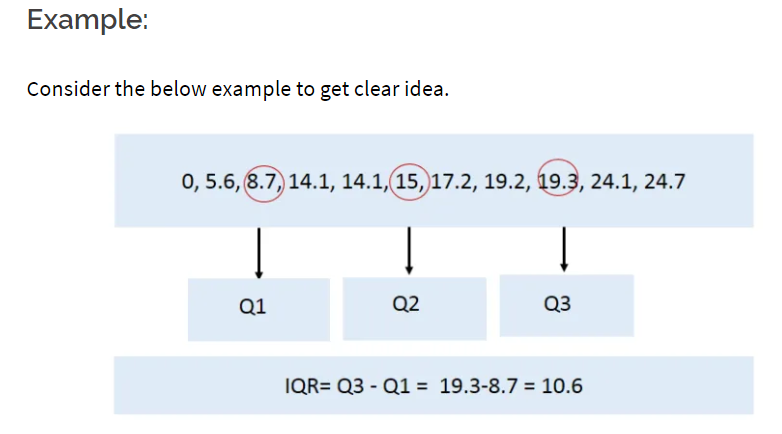
Step 3: Then find Q1 by looking the median of the left side of Q2

Steps 4: Similarly find Q3 by looking the median of the right of Q2

Steps 5: Now subtract Q1 from Q3 to get IQR.

##### **Example:**

Consider the below example to get clear idea.



Consider another example to get better understanding.

Consider the following numbers: 1, 3, 4, 5, 5, 6, 7, 11. Q1 is the middle value in the first half of the data set. Since there are an even number of data points in the first half of the data set, the middle value is the average of the two middle values; that is, Q1 = (3 + 4)/2 or Q1 = 3.5. Q3 is the middle value in the second half of the data set. Again, since the second half of the data set has an even number of observations, the middle value is the average of the two middle values; that is, Q3 = (6 + 7)/2 or Q3 = 6.5. The interquartile range is Q3 minus Q1, so IQR = 6.5 – 3.5 = 3.

##### **Advantage of IQR:**

* The main advantage of the IQR is that it is not affected by outliers because it doesn’t take into account observations below Q1 or above Q3.
* It might still be useful to look for possible outliers in your study.
* As a rule of thumb, observations can be qualified as outliers when they lie more than 1.5 IQR below the first quartile or 1.5 IQR above the third quartile.

Outliers = Q1 – 1.5\* IQR OR

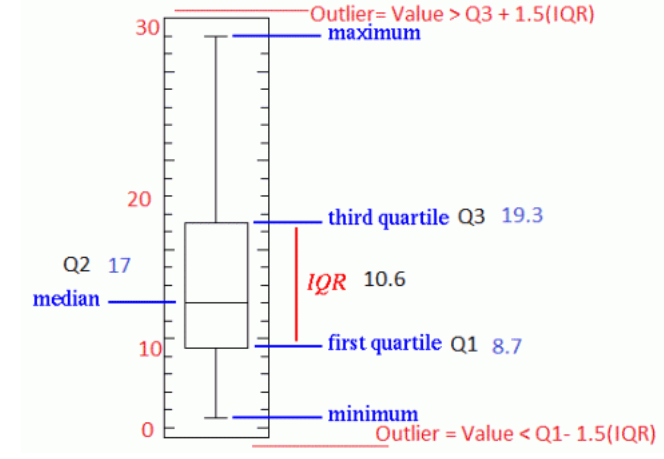
=Q3 + 1.5\*IQR

#### **Box Plot:**

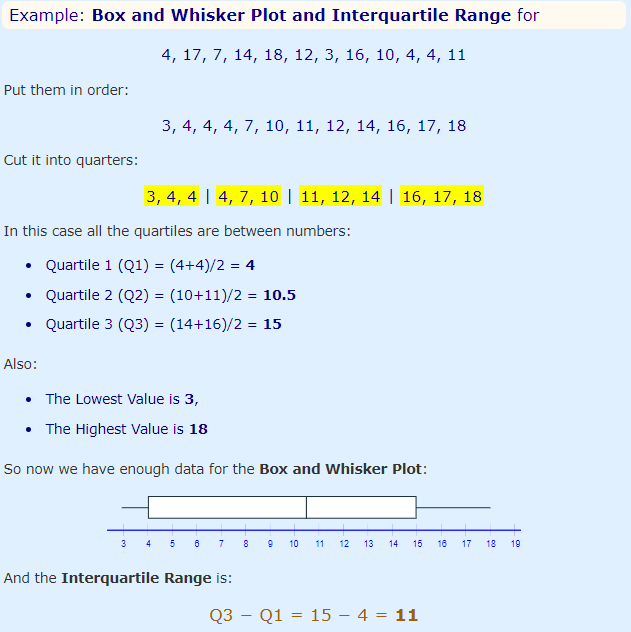
There is one graph that is mainly used when you are describing center and variability of your data.

It is also useful for detecting outliers in the data.

Carefully, observe the above first IQR example when it is plotted in a boxplot.



## 

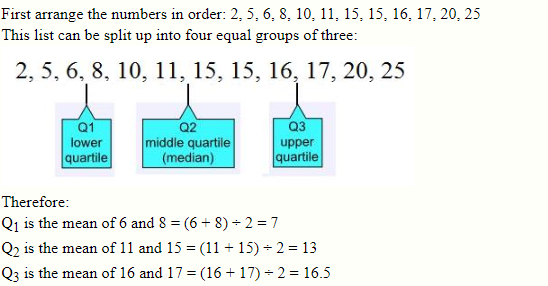


**Question**

What are the quartiles for the following set of numbers?

8, 11, 20, 10, 2, 17, 15, 5, 16, 15, 25, 6

**Answer**

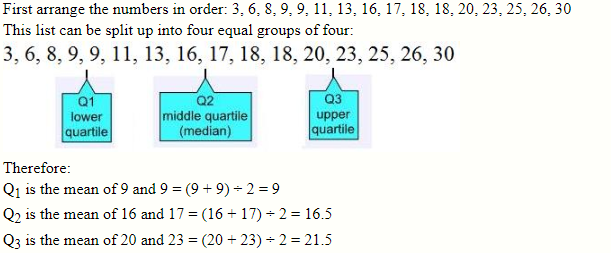


**Question**

What are the quartiles for the following set of numbers?

13, 18, 6, 20, 25, 11, 9, 18, 3, 30, 16, 9, 8, 23, 26, 17

**Answer**

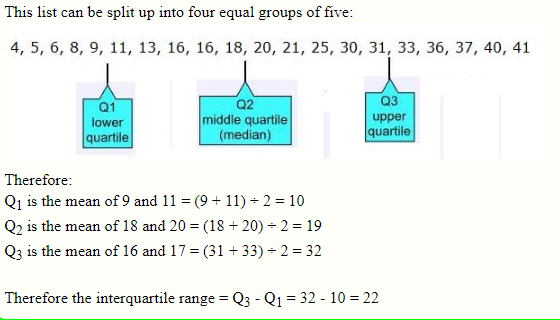


**Question**

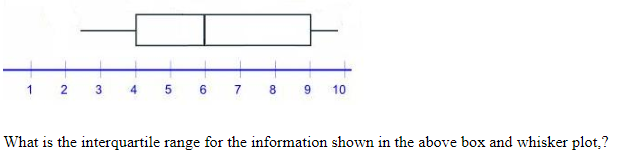
What is the interquartile range for the following set of numbers?

4, 5, 6, 8, 9, 11, 13, 16, 16, 18, 20, 21, 25, 30, 31, 33, 36, 37, 40, 41

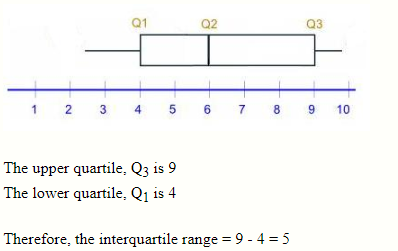
**Answer**

****

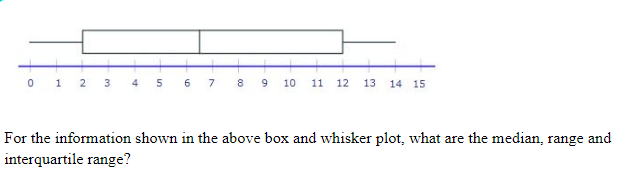
**Question**

****

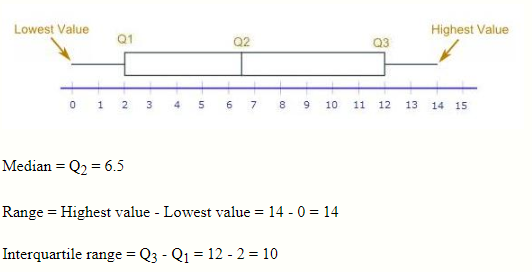
**Answer**

****

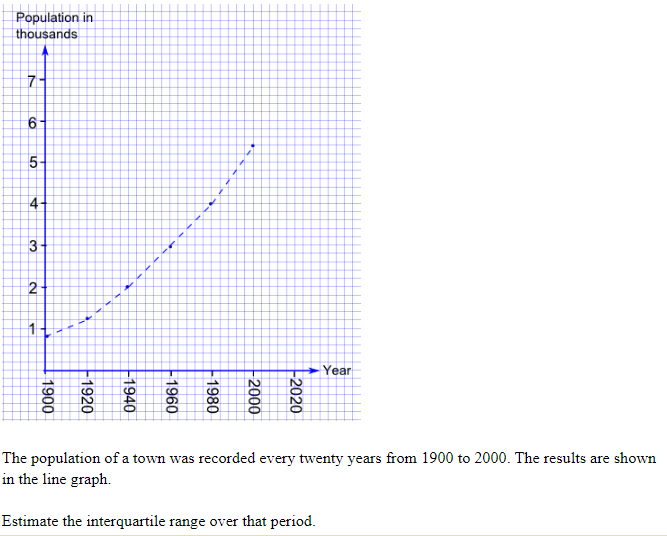
**Question**



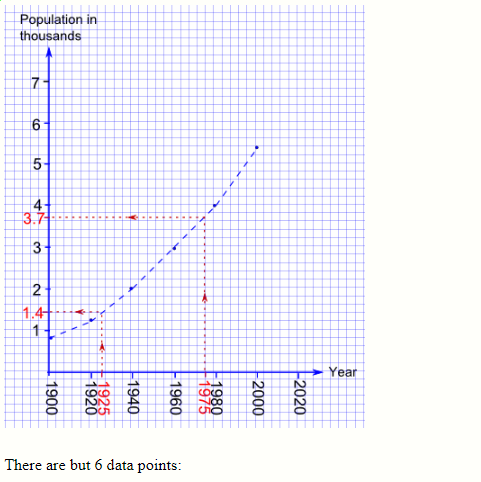
**Answer**

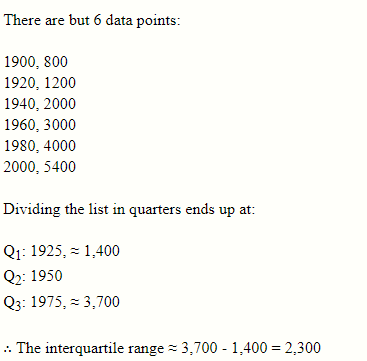


**Question**



**Answer**

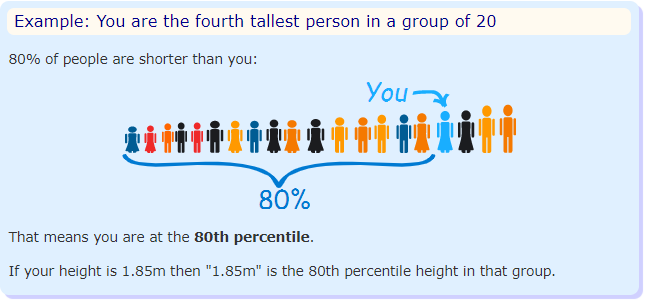




# Percentiles

**Percentile**:

the value below which a percentage of data falls.



## In Order

Have the data **in order**, so you know which values are above and below.

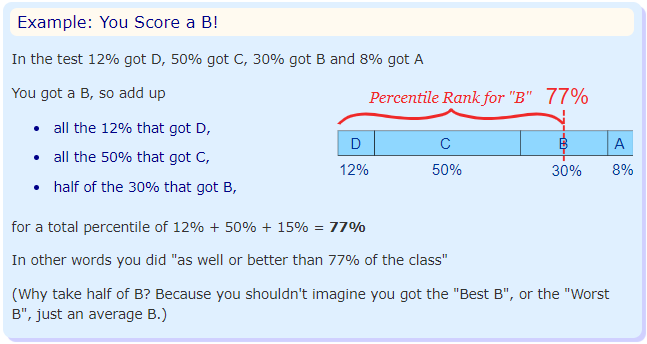
* To calculate percentiles of height: have the data in height order (sorted by height).
* To calculate percentiles of age: have the data in age order.
* And so on.

## 

## Grouped Data

When the data is grouped:

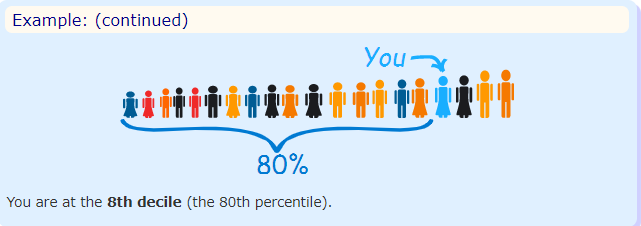
* Add up all percentages **below** the score.
* plus **half** the percentage **at** the score.



## Deciles

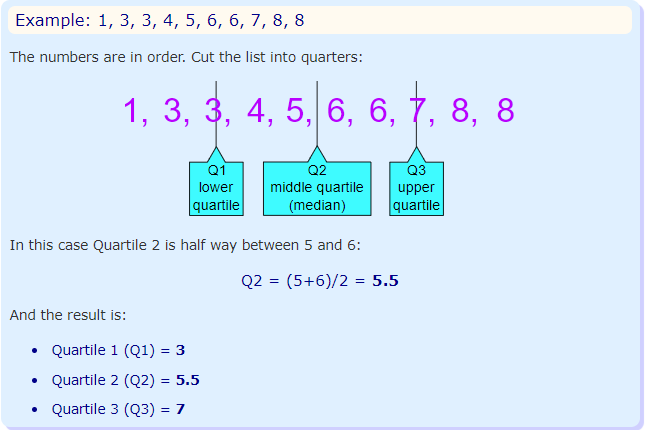
**Deciles** are similar to Percentiles (sounds like decimal and percentile together), as they split the data into **10% groups**:

* The **1st decile** is the **10th percentile** (the value that divides the data so that **10%** is below it)
* The **2nd decile** is the **20th percentile** (the value that divides the data so that **20%** is below it)
* etc!



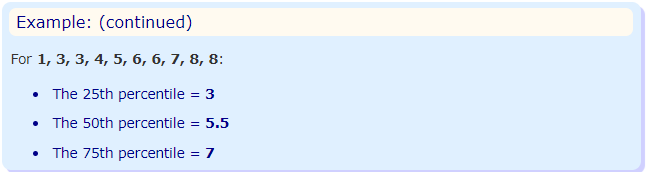
## Quartiles

Another related idea is Quartiles, which splits the data into quarters:



The Quartiles also divide the data into divisions of 25%, so:

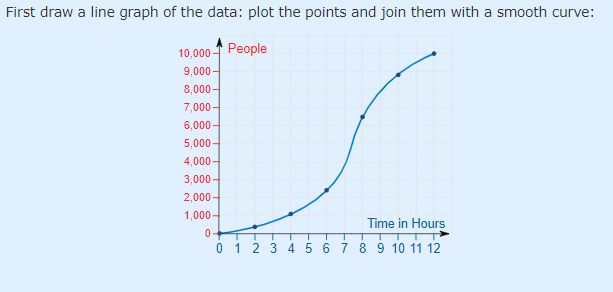
* Quartile 1 (Q1) can be called the **25th percentile**
* Quartile 2 (Q2) can be called the **50th percentile**
* Quartile 3 (Q3) can be called the **75th percentile**

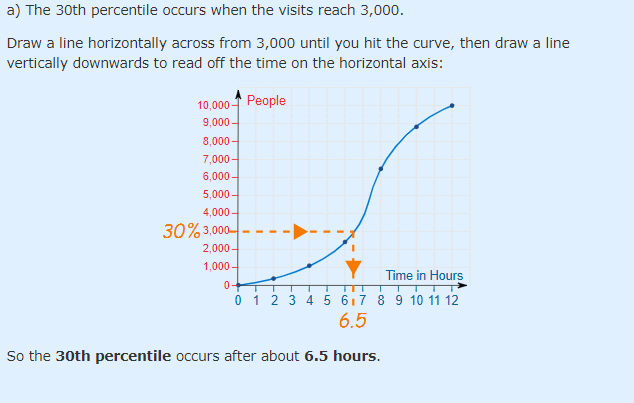
****

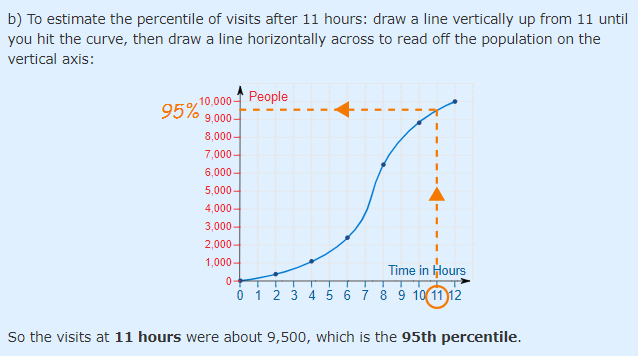
## Estimating Percentiles

**We can estimate percentiles from a line graph.**

****

****

****

****

# Mean Deviation

**How far, on average, all values are from the middle.**

The mean deviation is the first measure of dispersion that we will use that actually uses each data value in its computation.

It is the mean of the distances between each value and the mean.

It gives us an idea of how spread out from the center the set of values is.

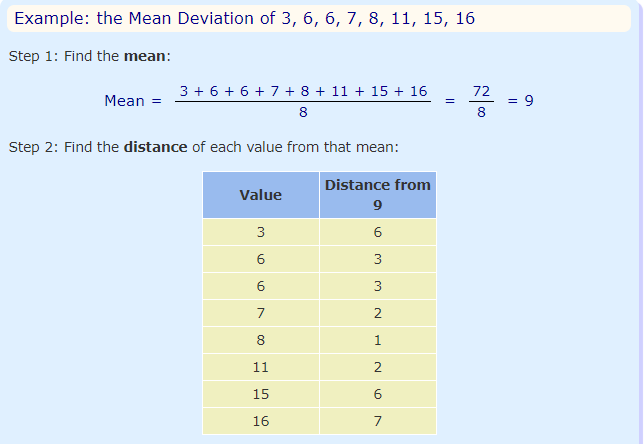
Calculating It

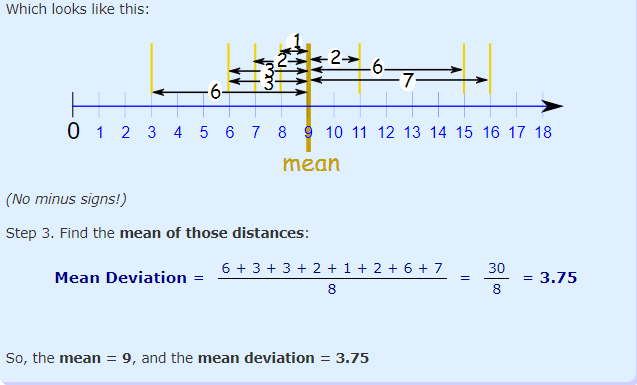
Find the mean of all values ... use it to work out distances ... then find the mean of those distances!

In three steps:

* 1. Find the [mean](https://www.mathsisfun.com/mean.html) of all values
* 2. Find the distance of each value from that mean (subtract the mean from each value, ignore minus signs)
* 3. Then find the mean of those distances

Like this:

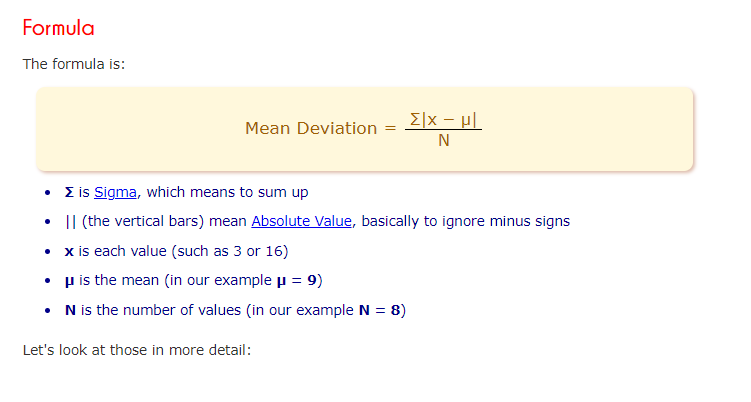


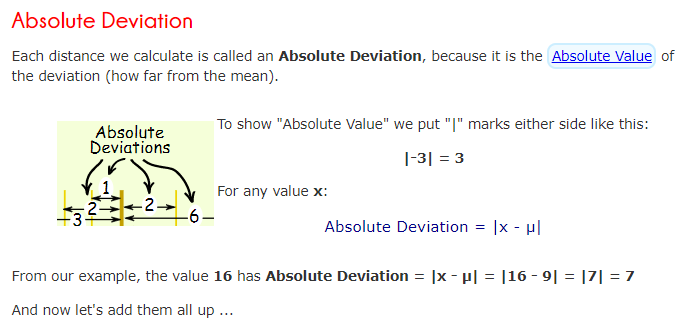


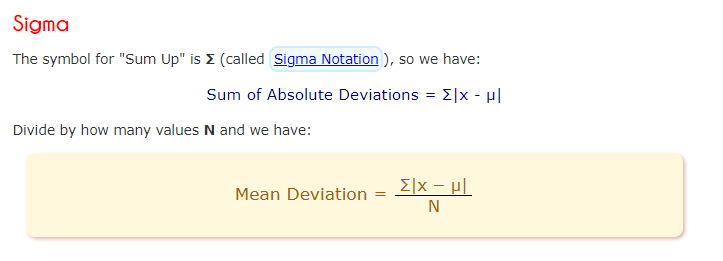
It tells us how far, on average, all values are from the middle.

In that example the values are, on average, 3.75 away from the middle.

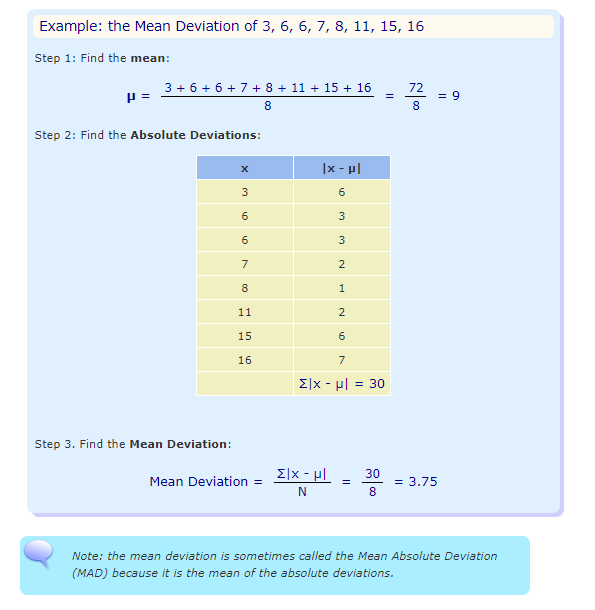
**Deviation = Distance**







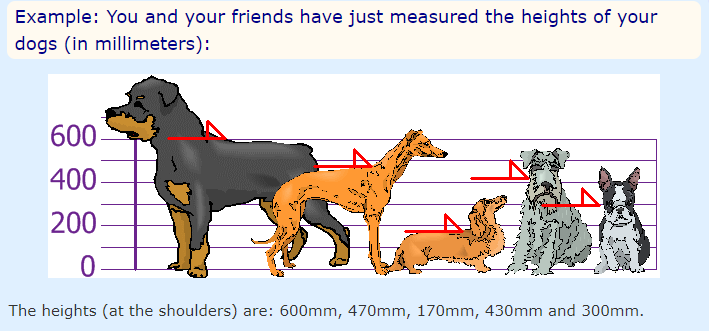
Let's do our example again, using the proper symbols:

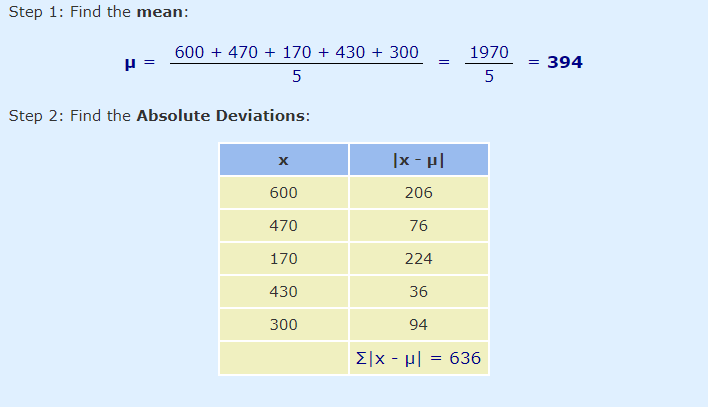


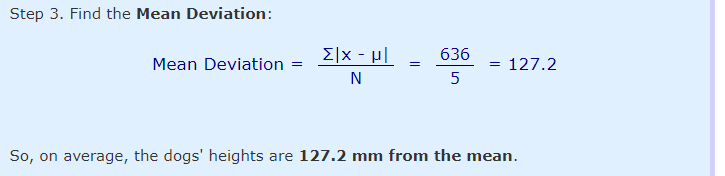
## What Does It "Mean" ?

Mean Deviation tells us how far, on average, all values are from the middle.

Here is an example :



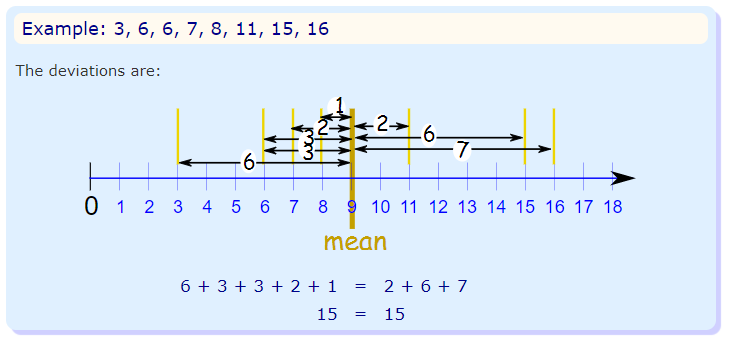


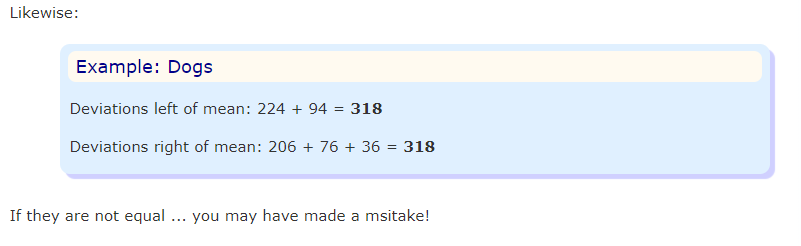


## A Useful Check

The deviations on **one side** of the mean should equal the deviations on the **other side**.

From our first example:

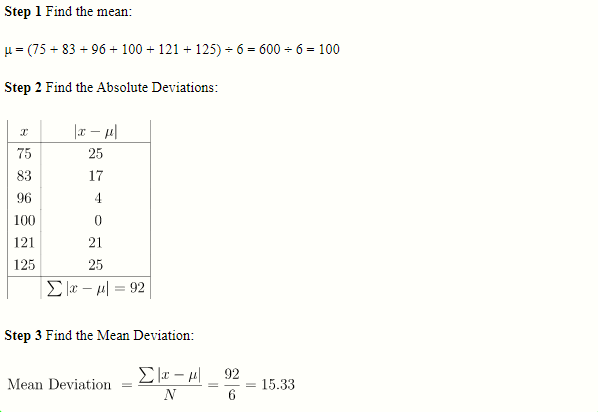




**Question :1**

Calculate the mean deviation for the numbers: 75, 83, 96, 100, 121 and 125

**Answer**



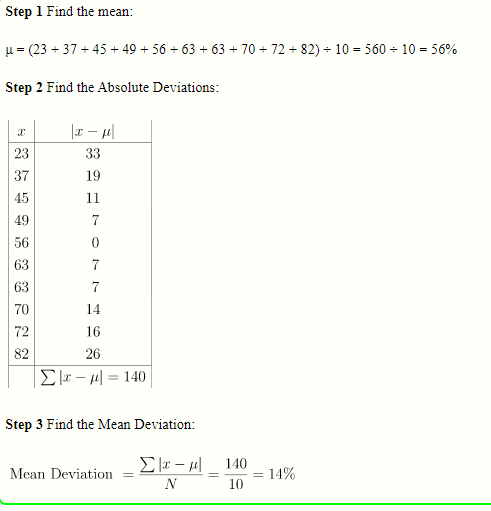
**Question:2**

Ten friends scored the following marks in their end-of-year math exam:

23%, 37%, 45%, 49%, 56%, 63%, 63%, 70%, 72% and 82%

What was the mean deviation of their marks?

**Answer**



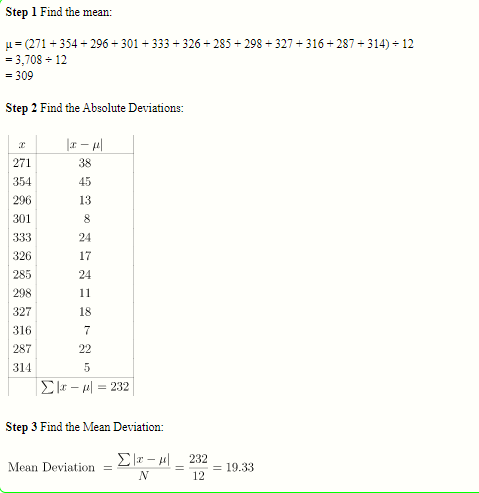
**Question :3**

A booklet has 12 pages with the following numbers of words:

271, 354, 296, 301, 333, 326, 285, 298, 327, 316, 287 and 314

What is the mean deviation of the number of words per page?

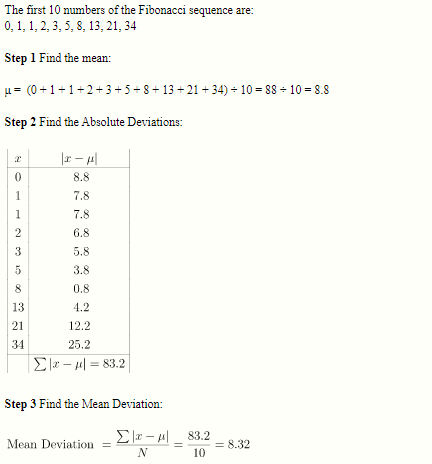
**Answer**



**Question :4**

What is the mean deviation of the first 10 numbers of the Fibonacci sequence?

**Answer**

****

# Standard Deviation and Variance

*Deviation just means how far from the normal*

## *Standard Deviation*

*The Standard Deviation is a measure of how spread out numbers are.*

*Its symbol is* ***σ*** *(the greek letter sigma)*

*The formula is easy: it is the* ***square root*** *of the* ***Variance.*** *So now you ask, "What is the Variance?"*

*Variance*

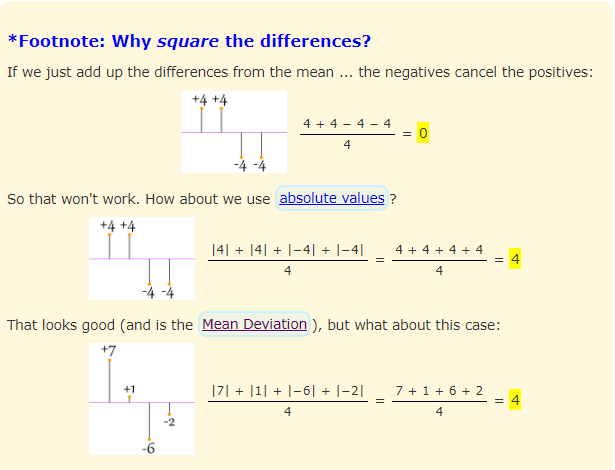
*The Variance is defined as:*

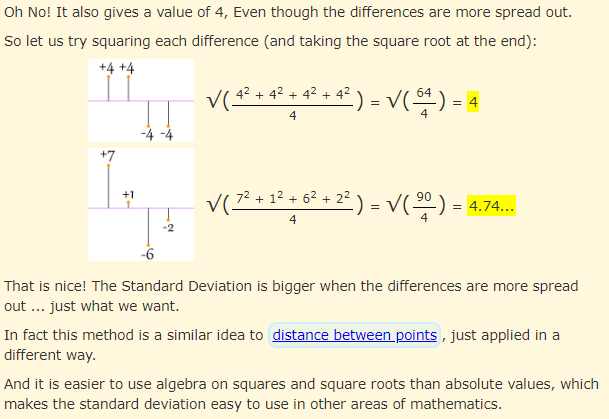
*The average of the* ***squared*** *differences from the Mean.*

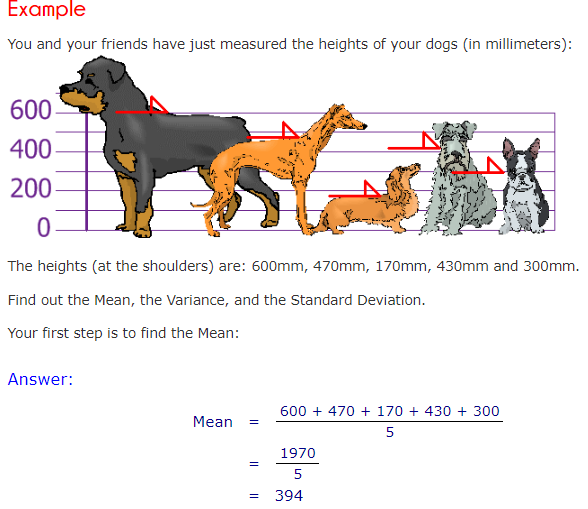
*To calculate the variance follow these steps:*

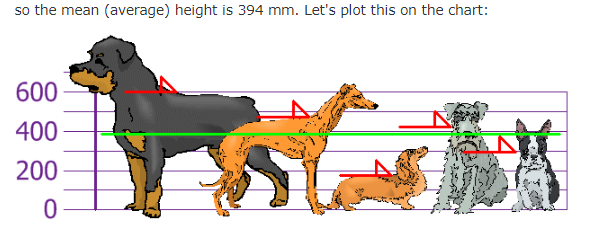
* *Work out the* [*Mean*](https://www.mathsisfun.com/mean.html) *(the simple average of the numbers)*
* *Then for each number: subtract the Mean and square the result (the squared difference).*
* *Then work out the average of those squared differences.*

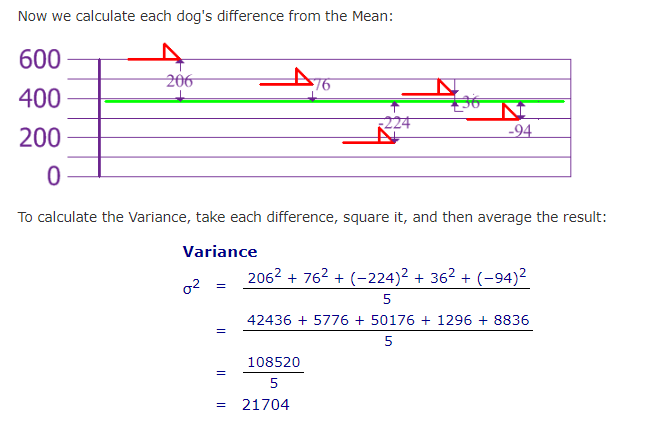
***Why square ?***

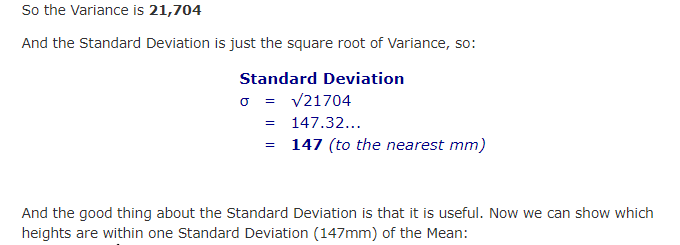
******

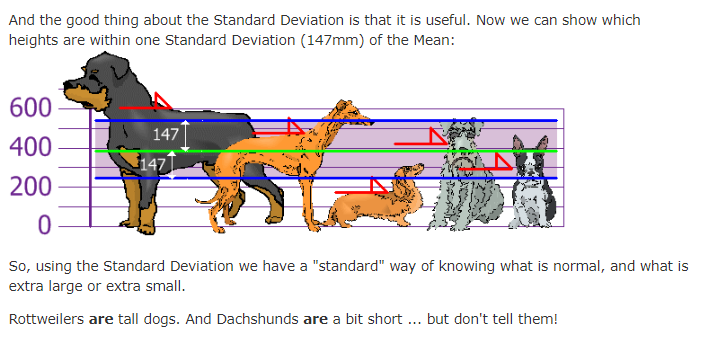
******

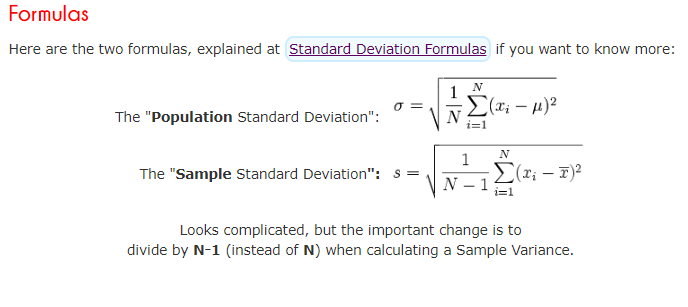
******

******

******

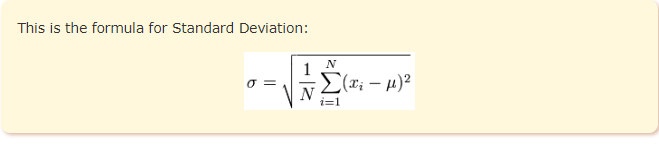
******





here we explain **the formulas**.

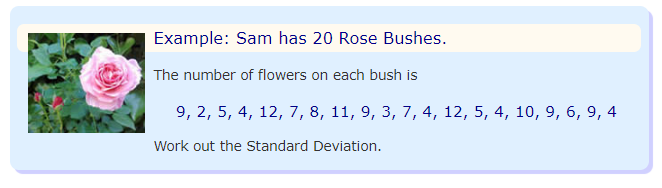
The symbol for Standard Deviation is **σ** (the Greek letter sigma).



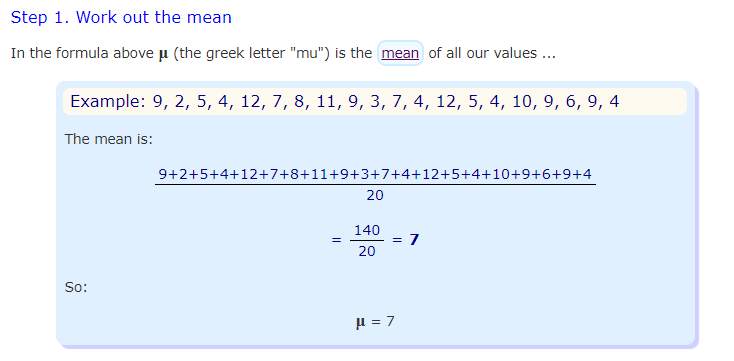
To calculate the standard deviation of those numbers:

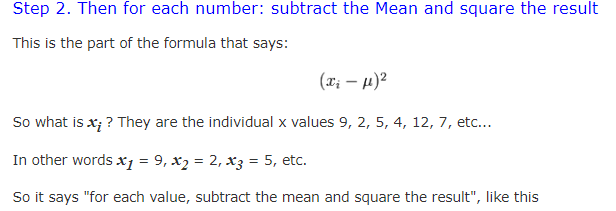
* 1. Work out the [Mean](https://www.mathsisfun.com/mean.html) (the simple average of the numbers)
* 2. Then for each number: subtract the Mean and square the result
* 3. Then work out the mean of **those** squared differences.
* 4. Take the square root of that and we are done!

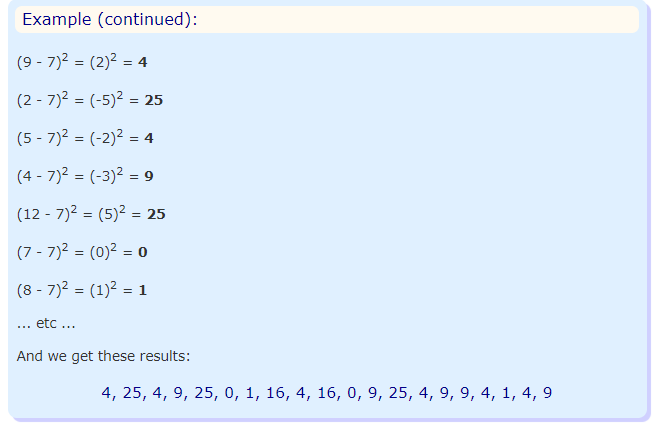
**Example:1**

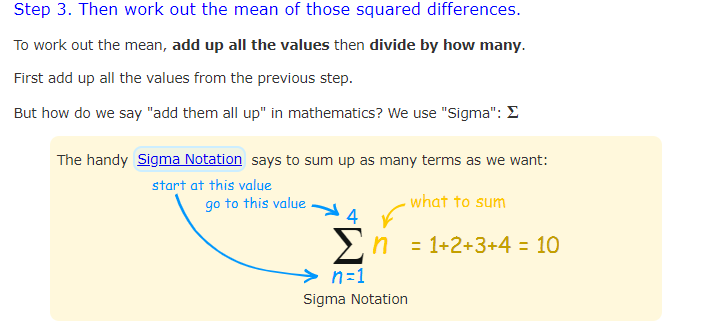
****

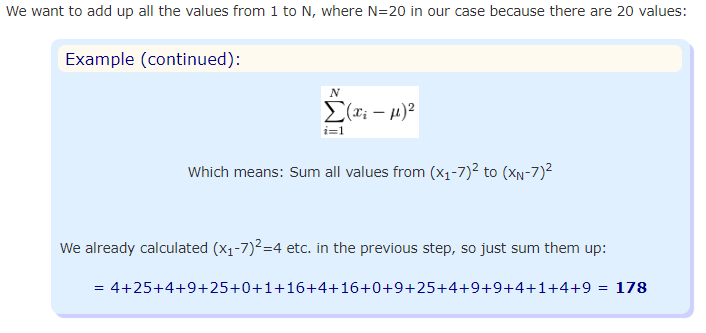
**Answer**

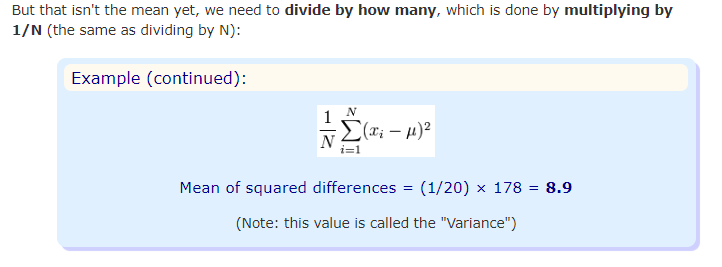


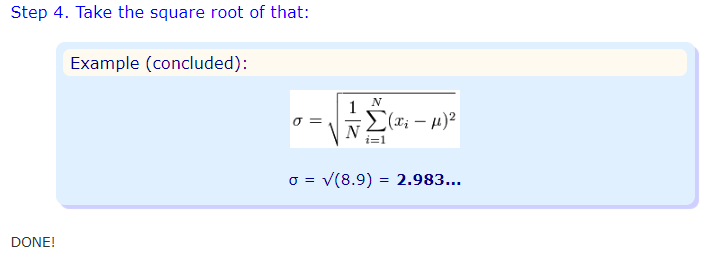


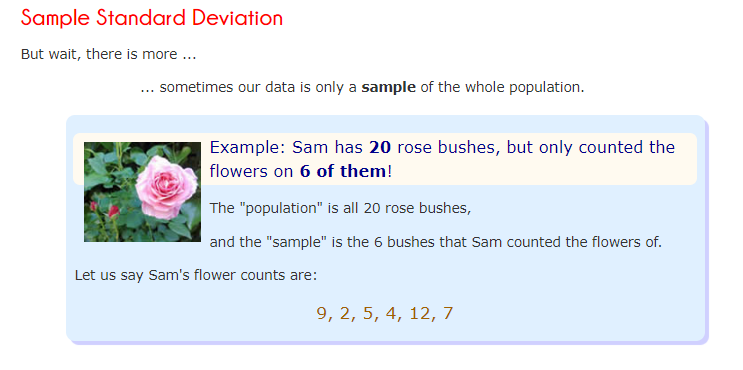


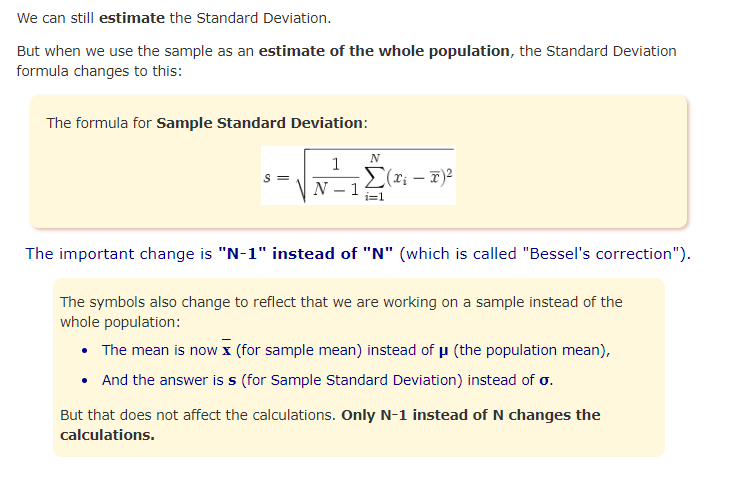


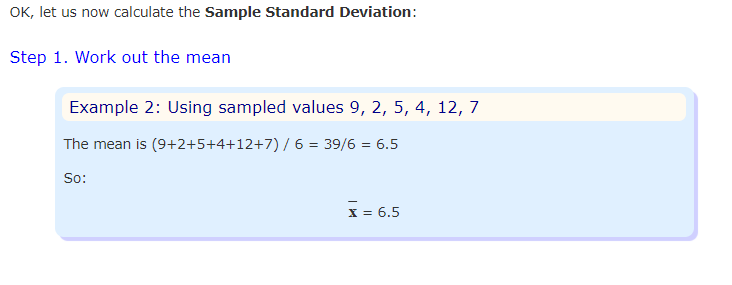


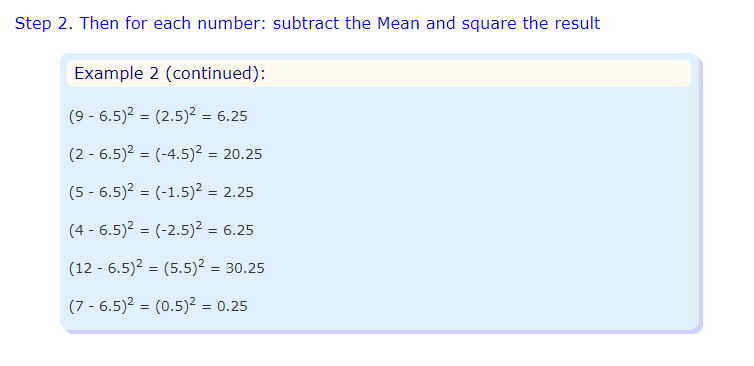


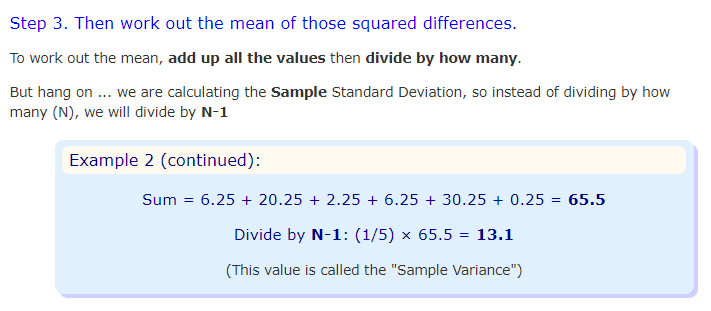


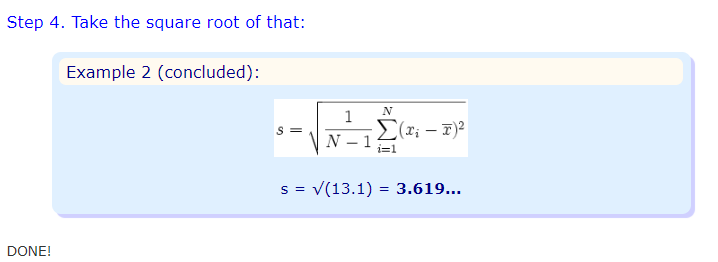


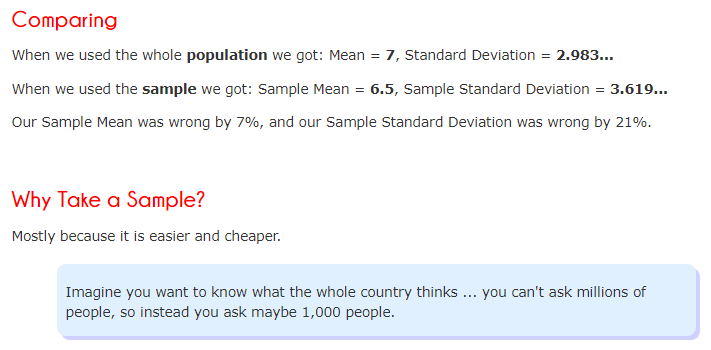


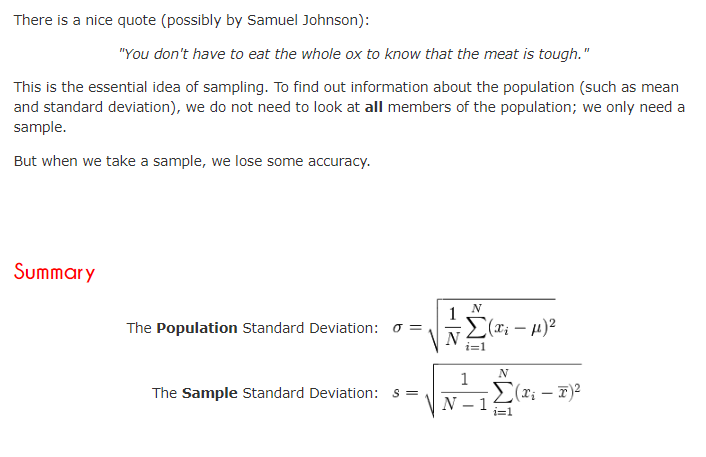












**Example :2**

What is the population standard deviation for the numbers: 75, 83, 96, 100, 121 and 125?

(Try to do this yourself, without using the Standard Deviation calculator.)

**Answer**

1. Firstly find the mean:

Mean = (75 + 83 + 96 + 100 + 121 + 125) ÷ 6 = 600 ÷ 6 = 100

2. Next find the variance. To calculate the Variance, take each difference, square it, and then average the result:

(75 - 100)2 + (83 - 100)2 + (96 - 100)2 + (100 - 100)2 + (121 - 100)2 + (125 - 100)2

= (-25)2 + (-17)2 + (-4)2 + (0)2 + (21)2 + (25)2

= 625 + 289 + 16 + 0 + 441 + 625

= 1,996

So the Variance = 1,996 ÷ 6 = 332.66...

3. The Standard Deviation is just the square root of the Variance

= √(332.66...)

= 18.2 correct to 1 decimal places

**Example :3**

Ten friends scored the following marks in their end-of-year math exam:

23%, 37%, 45%, 49%, 56%, 63%, 63%, 70%, 72% and 82%

What was the standard deviation of their marks?

(Try to do this yourself, without using the Standard Deviation calculator.)

**Answer**

1. Firstly find the mean:

Mean = (23 + 37 + 45 + 49 + 56 + 63 + 63 + 70 + 72 + 82) ÷ 10 = 560 ÷ 10 = 56

2. Next find the variance. To calculate the Variance, take each difference, square it, and then average the result:

(23 - 56)2 + (37 - 56)2 + (45 - 56)2 + (49 - 56)2 + (56 - 56)2 + (63 - 56)2 + (63 - 56)2 + (70 - 56)2 + (72 - 56)2 + (82 - 56)2

= (-33)2 + (-19)2 + (-11)2 + (-7)2 + (0)2 + (7)2 + (7)2 + (14)2 + (16)2 + (26)2

= 1,089 + 361 + 121 + 49 + 0 + 49 + 49 + 196 + 256 + 676

= 2,846

So the Variance = 2,846 ÷ 10 = 284.6

3. The Standard Deviation is just the square root of the Variance

= √(284.6)

= 16.9 correct to 1 decimal place

**Example :4**

A booklet has 12 pages with the following numbers of words:

271, 354, 296, 301, 333, 326, 285, 298, 327, 316, 287 and 314

What is the standard deviation number of words per page?

**Answer**

1. Firstly find the mean number of words per page:

Mean = (271 + 354 + 296 + 301 + 333 + 326 + 285 + 298 + 327 + 316 + 287 + 314) ÷ 12

= 3,708 ÷ 12

= 309

2. Next find the variance. To calculate the Variance, take each difference, square it, and then average the result:

(271 - 309)2 + (354 - 309)2 + (296 - 309)2 + (301 - 309)2 + (333 - 309)2 + (326 - 309)2 + (285 - 309)2 + (298 - 309)2 + (327 - 309)2 + (316 - 309)2 + (287 - 309)2 + (314 - 309)2

= (-38)2 + (45)2 + (-13)2 + (-8)2 + (24)2 + (17)2 + (-24)2 + (-11)2 + (18)2 + (7)2 + (-22)2 + (5)2

= 1,444 + 2,025 + 169 + 64 + 576 + 289 + 576 + 121 + 324 + 49 + 484 + 25

= 6,146

So the Variance = 6,146 ÷ 12 = 512.166...

3. The Standard Deviation is just the square root of the Variance

= √(512.166...)

= 22.6 correct to 1 decimal place