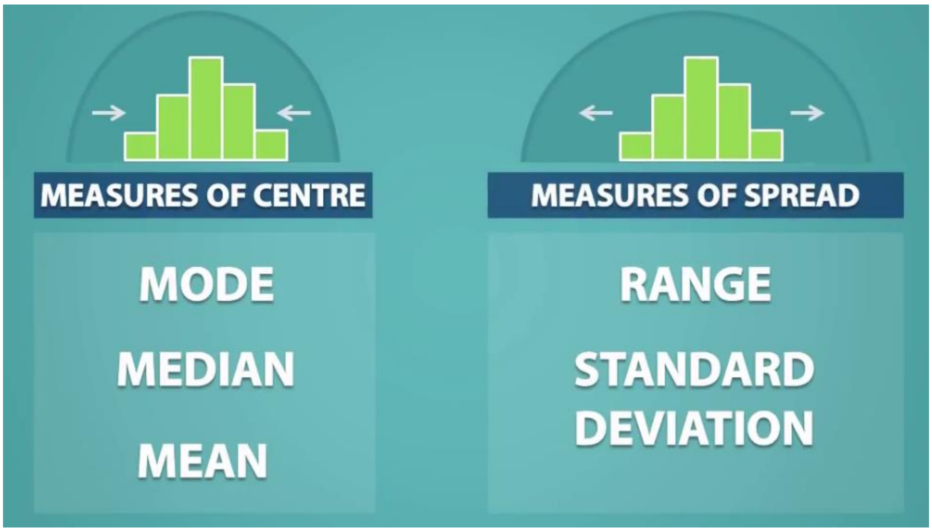
2. Maths - Statistics – PART – 2

Contents

2. Maths - Statistics – PART – 2

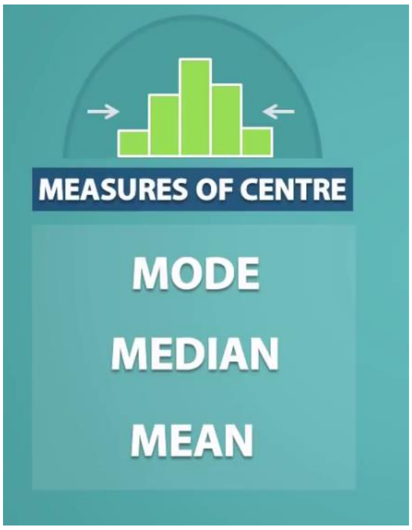
1. Usage of mode, median, mean, range & standard deviation?

- ✓ Identifying and describing like how the dataset got distributed
- ✓ These mode, median, mean, range and standard deviation gives the numerical information and distribution about the dataset
- ✓ These also explains about
  - Measures of centre
  - Measures of spread



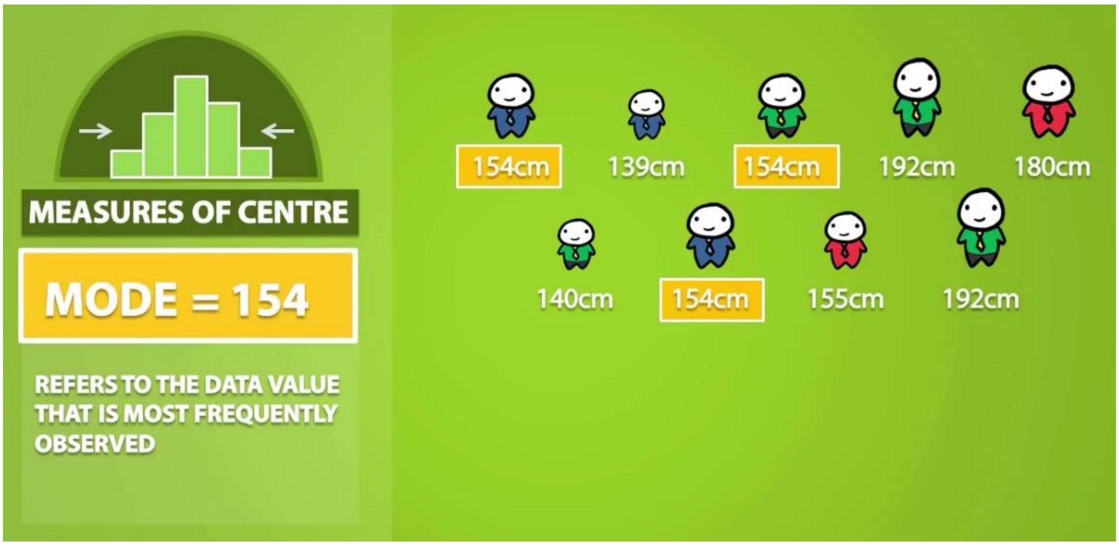
2. Measures of centre

- ✓ Mode
- ✓ Median
- ✓ Mean



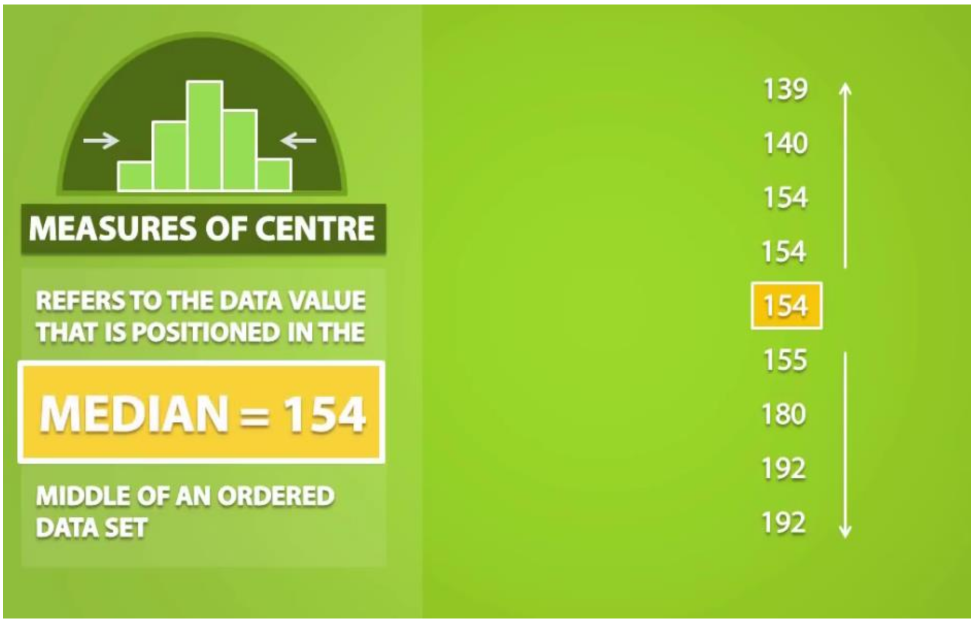
3. Mode

- ✓ Value which is most frequently observed
- ✓ Suppose we have taken random people heights and displayed as below
- ✓ Here 153 value is repeated in 3 times



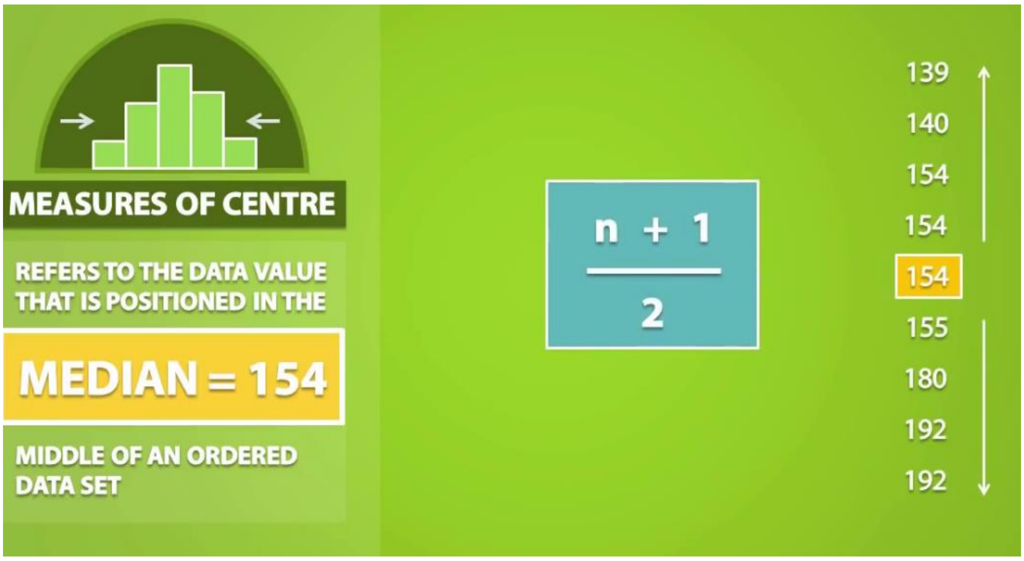
4. Median

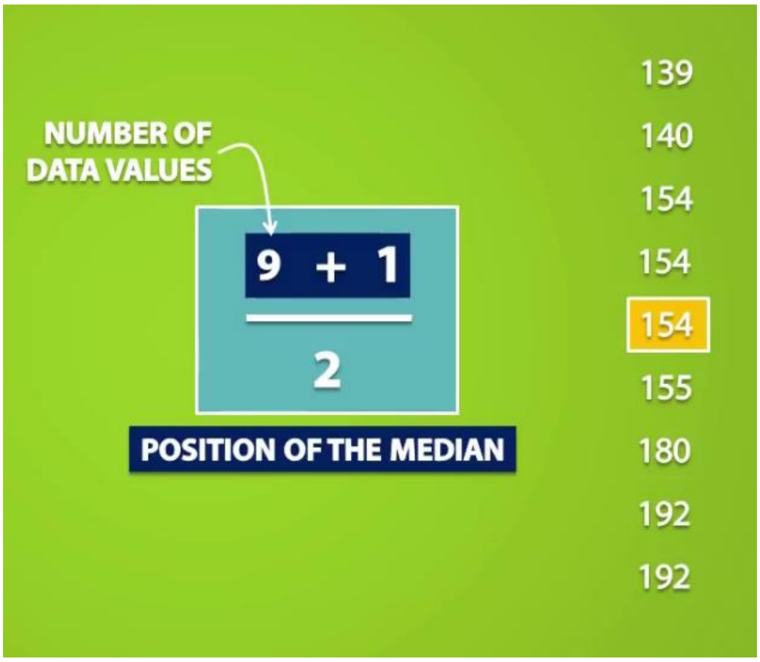
- ✓ Value that is positioned in the middle of an ordered dataset
- ✓ First we need to keep the data into an order
- ✓ We usually order the dataset into smallest to largest

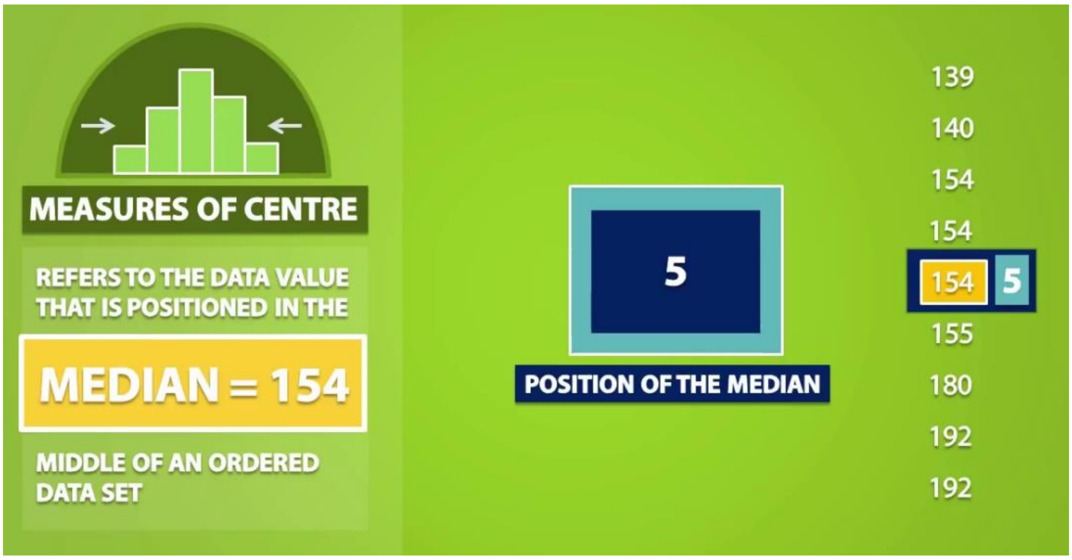


Special cases

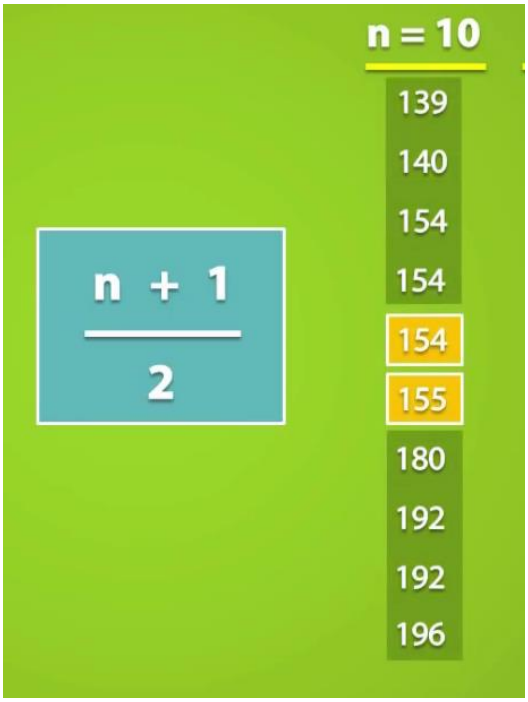
- ✓ If the dataset is extremely large then it might helpful use the below formula



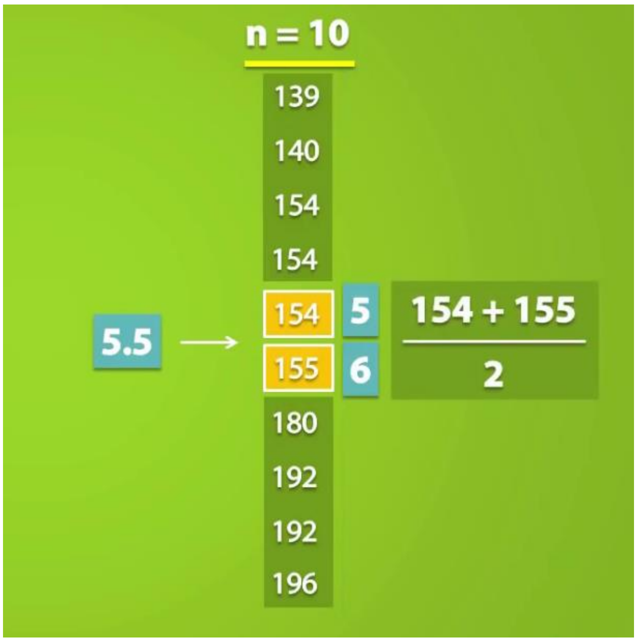



MEASURES OF CENTRE
REFERS TO THE DATA VALUE THAT IS POSITIONED IN THE
MEDIAN = 154
MIDDLE OF AN ORDERED DATA SET
5
POSITION OF THE MEDIAN
139
140
154
154
154
5
155
180
192
192

- ✓ If number of values are in odd or even number then we do have some special scenarios to find out the median value



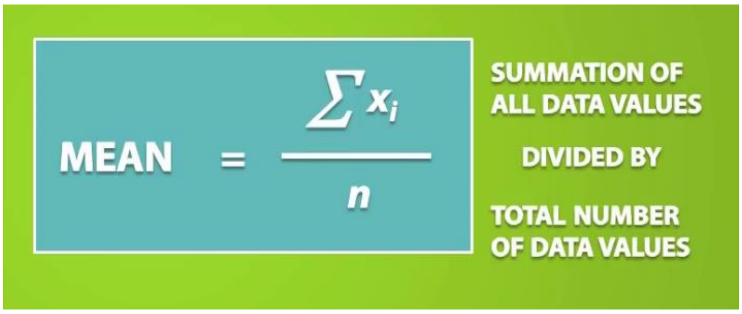

n = 10
139
140
154
154
154
155
180
192
192
196
1
2
3
4
5
6
7
8
9
10
5.5


n = 10
139
140
154
154
154
155
180
192
192
196
5
6
5.5
154 + 155
2


n = 10
139
140
154
154
154
155
180
192
192
196
5.5
5
6
MEDIAN
154.5

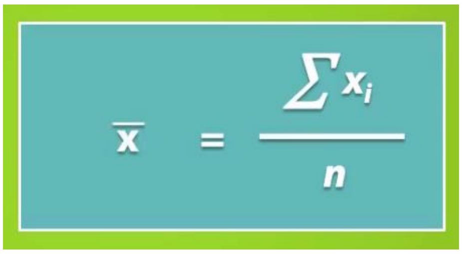
6. Mean

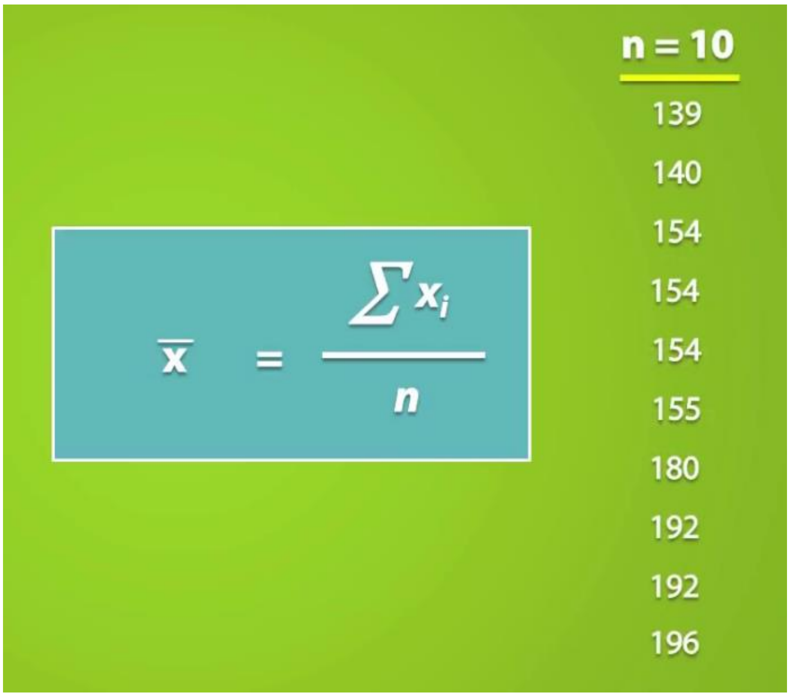
- ✓ The mean is just another name of average
- ✓ Below is the formula which indicates mean of total values

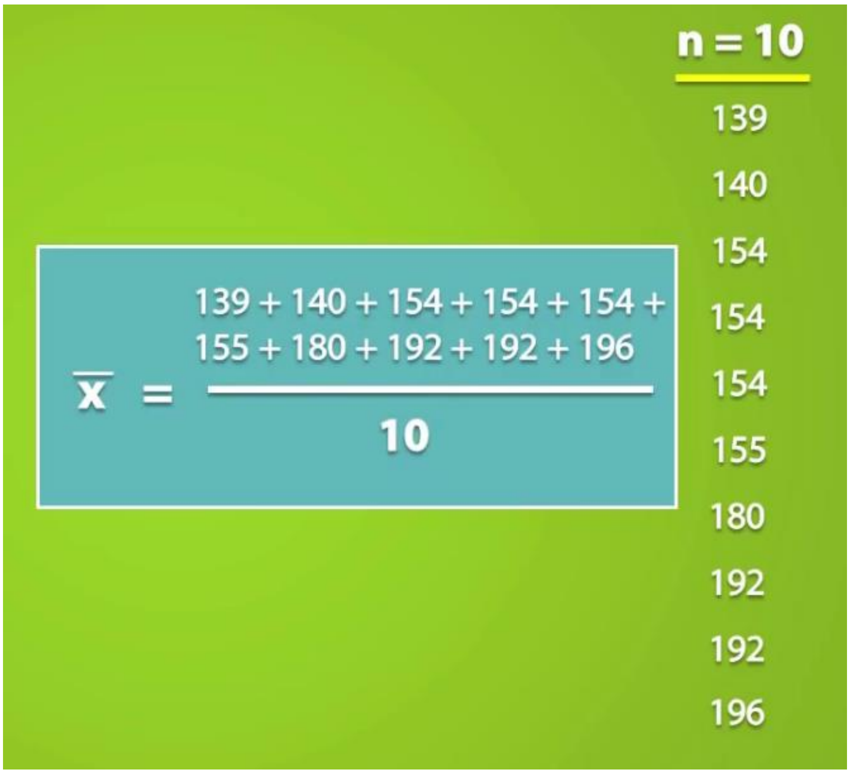


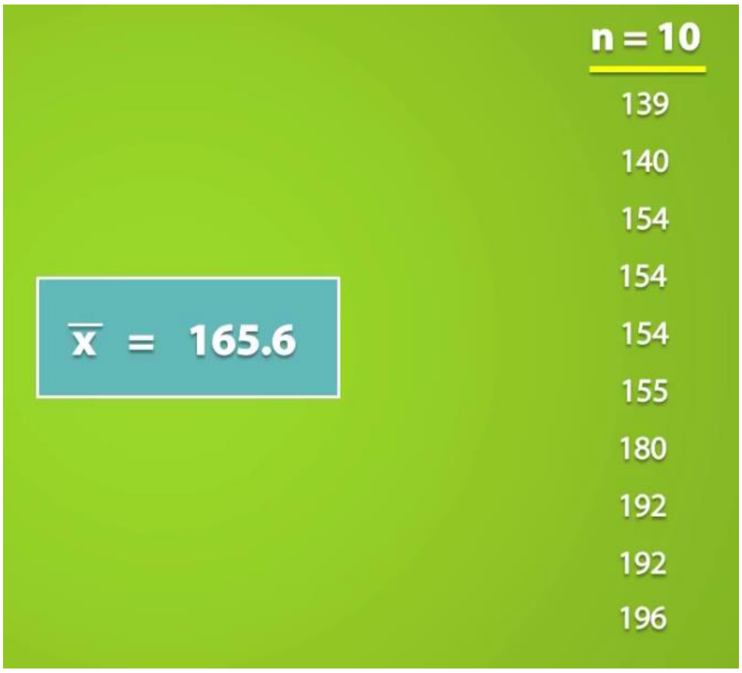
Sample mean

- ✓ Below is the formula which indicates mean of sample values




n = 10
139
140
154
154
154
155
180
192
192
196
$\bar{x} = \frac{\sum x_i}{n}$


n = 10
139
140
154
154
154
155
180
192
192
196
$\bar{x} = \frac{139 + 140 + 154 + 154 + 154 + 155 + 180 + 192 + 192 + 196}{10}$

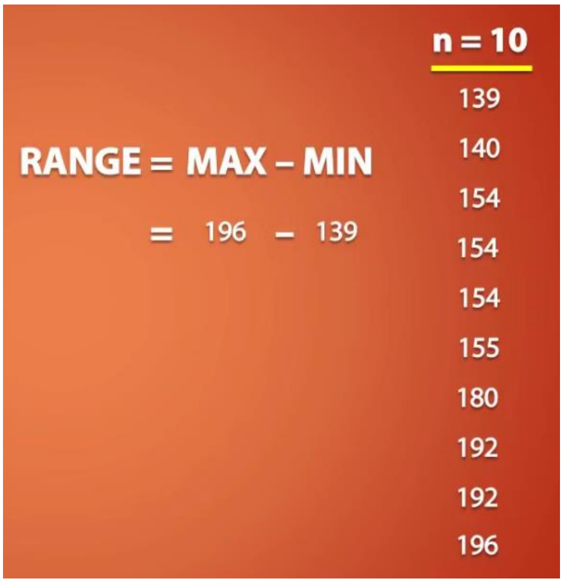

n = 10
139
140
154
154
154
155
180
192
192
196
x̄ = 165.6

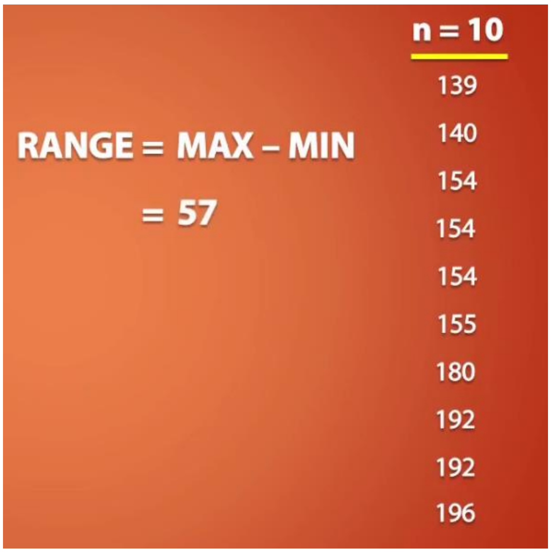
7. Measures of spread

- ✓ Range
- ✓ Standard deviation

8. Range

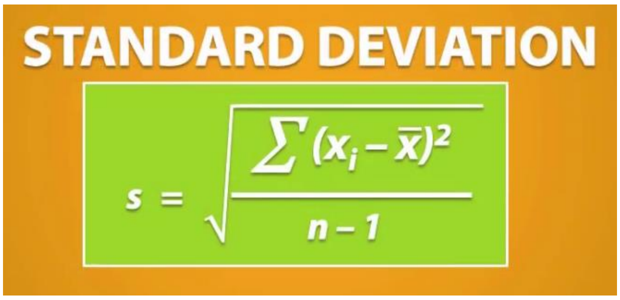
- ✓ Range means difference in between minimum value and maximum value
- ✓ It explains about the data is in between min and max values

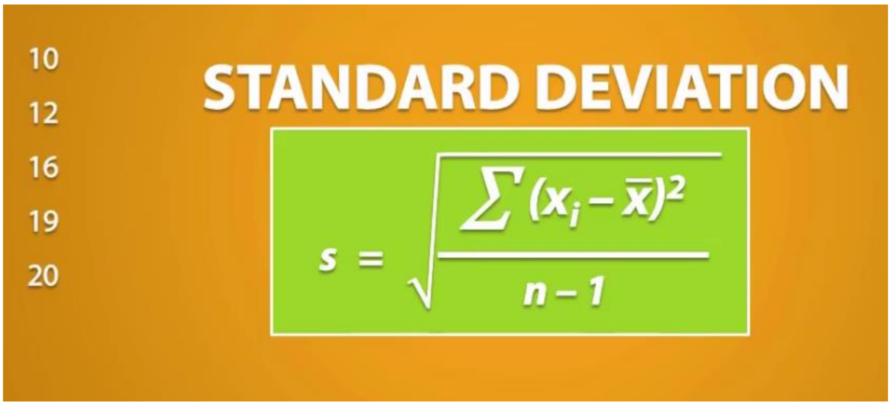


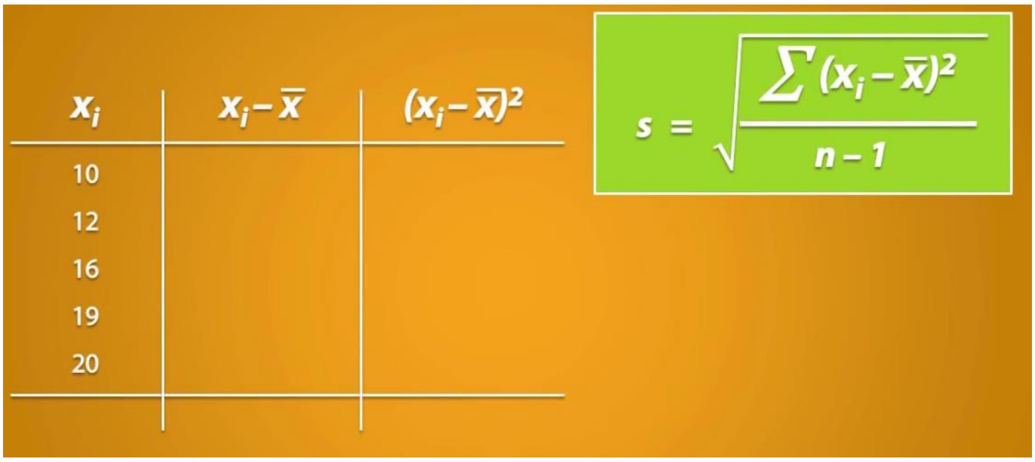


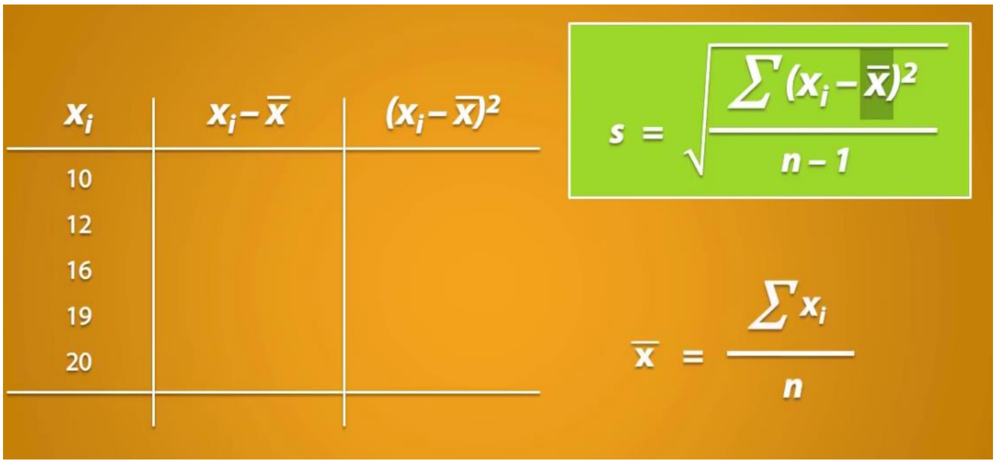
9. Standard Deviation

- ✓ The Standard Deviation is a measure of how spread out numbers.
- ✓ Formula is very simple,   It is the square root of the Variance

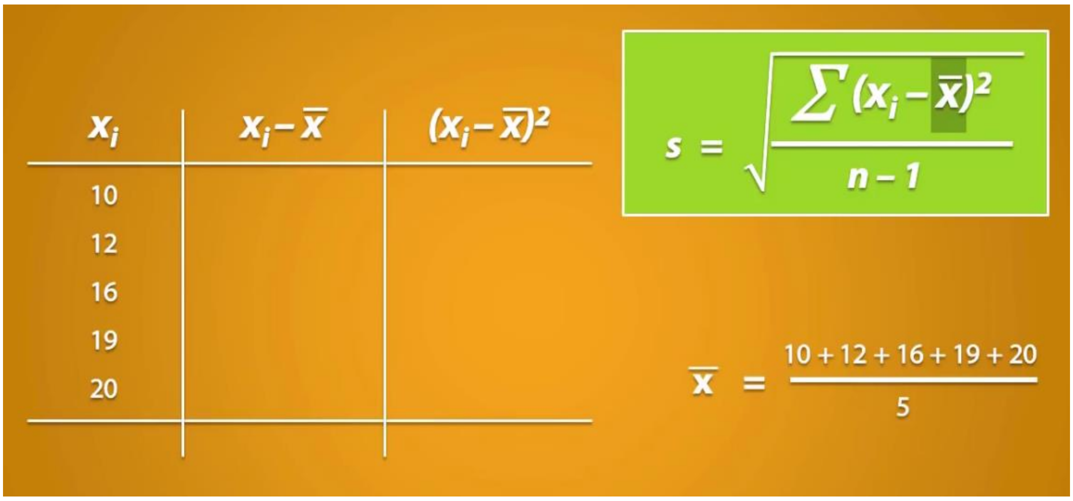




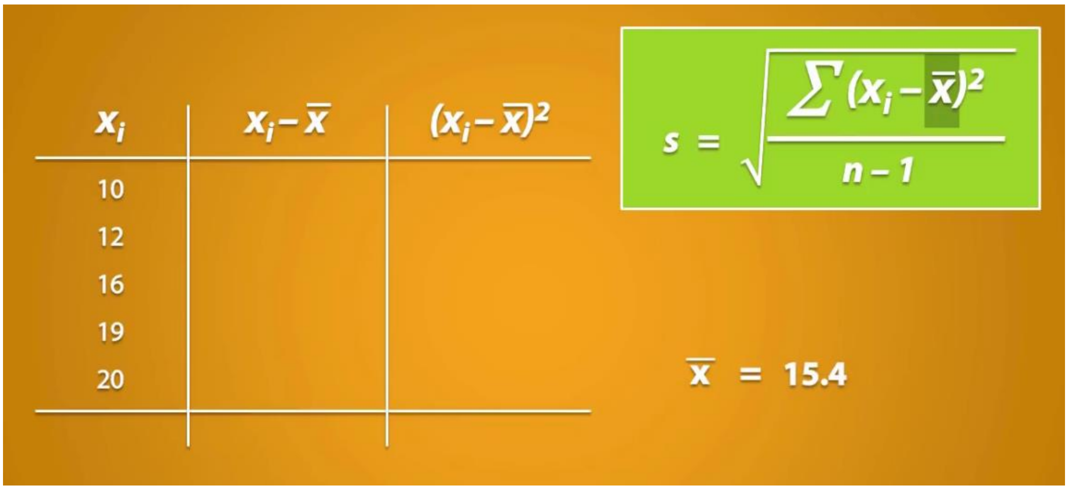


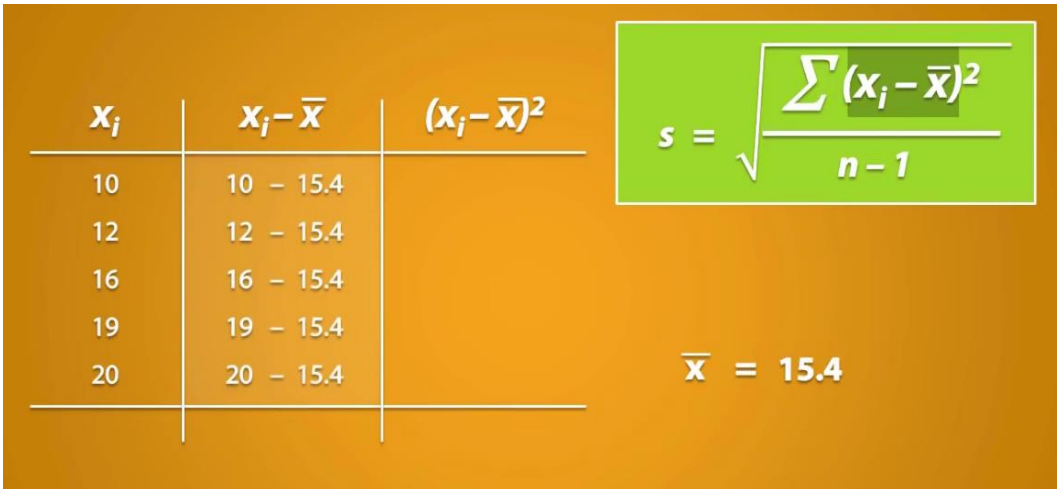


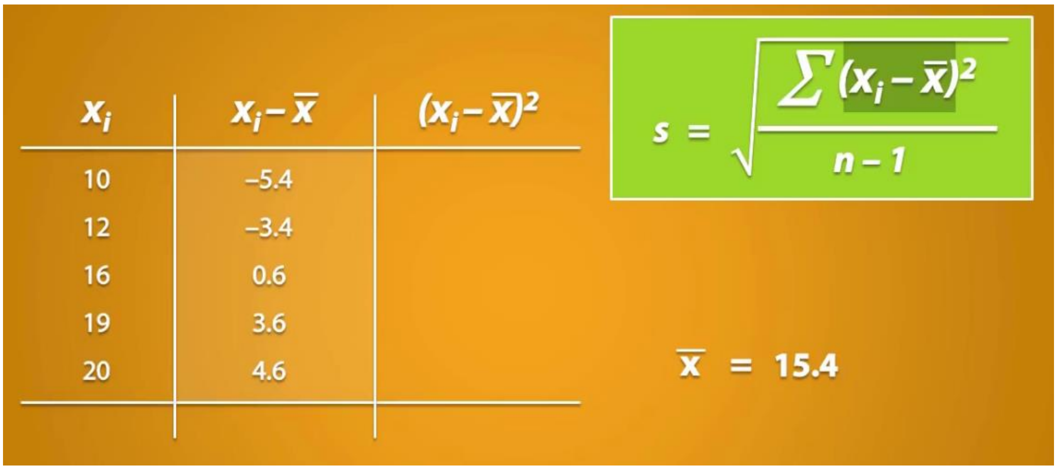
x_i
$x_i - \bar{x}$
$(x_i - \bar{x})^2$
10
12
16
19
20
$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$
$\bar{x} = \frac{\sum x_i}{n}$



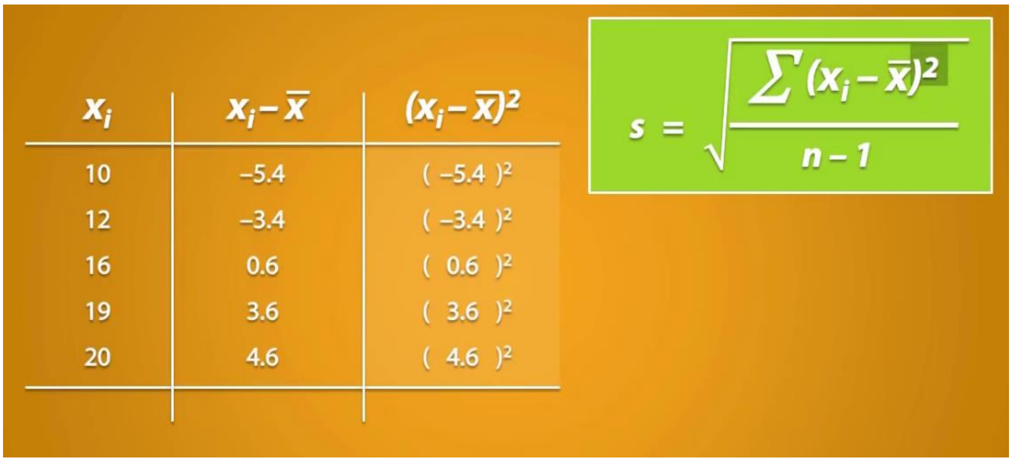
x_i
$x_i - \bar{x}$
$(x_i - \bar{x})^2$
10
12
16
19
20
$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$
$\bar{x} = \frac{10 + 12 + 16 + 19 + 20}{5}$


x_i
$x_i - \bar{x}$
$(x_i - \bar{x})^2$
10
12
16
19
20
$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$
$\bar{x} = 15.4$

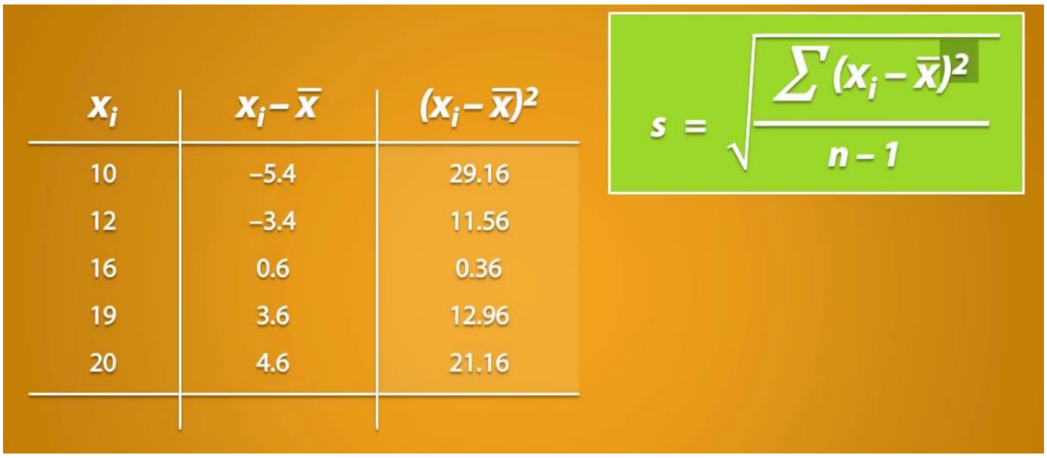

x_i
$x_i - \bar{x}$
$(x_i - \bar{x})^2$
10 10 – 15.4
12 12 – 15.4
16 16 – 15.4
19 19 – 15.4
20 20 – 15.4
$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$
$\bar{x} = 15.4$

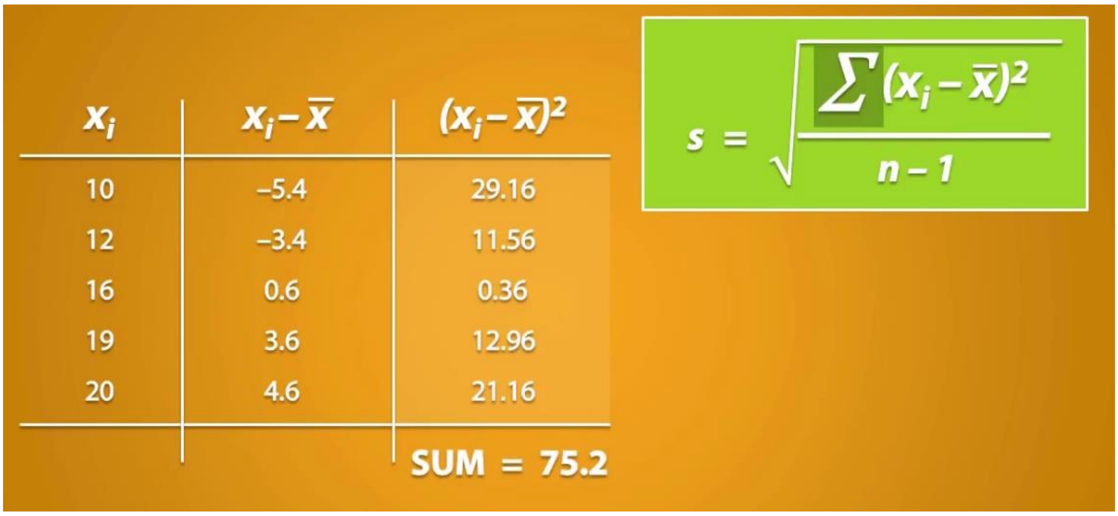


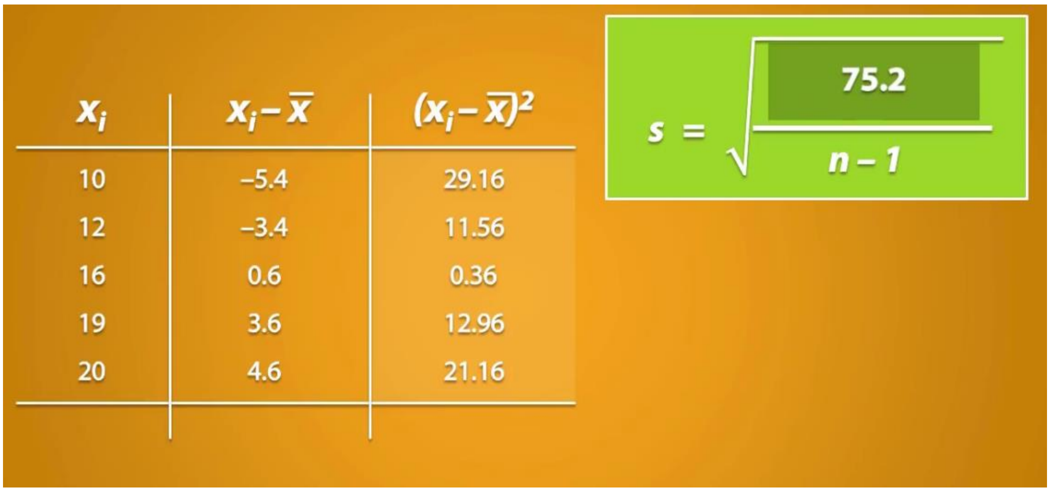
x_i | $x_i - \bar{x}$ | $(x_i - \bar{x})^2$
10 | –5.4
12 | –3.4
16 | 0.6
19 | 3.6
20 | 4.6
$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$
$\bar{x} = 15.4$

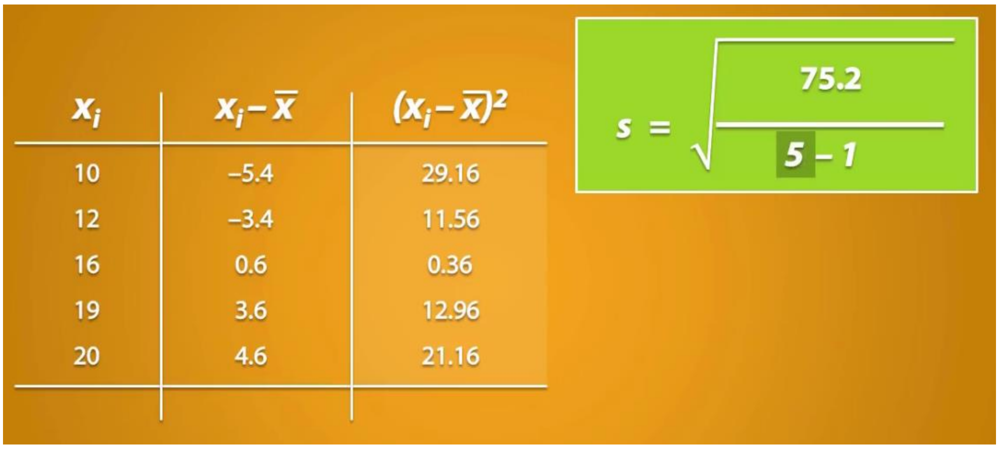


x_i | $x_i - \bar{x}$ | $(x_i - \bar{x})^2$
10 | –5.4 | $(-5.4)^2$
12 | –3.4 | $(-3.4)^2$
16 | 0.6 | $(0.6)^2$
19 | 3.6 | $(3.6)^2$
20 | 4.6 | $(4.6)^2$
$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$


x_i | $x_i - \bar{x}$ | $(x_i - \bar{x})^2$
10 | –5.4 | 29.16
12 | –3.4 | 11.56
16 | 0.6 | 0.36
19 | 3.6 | 12.96
20 | 4.6 | 21.16
$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$


x_i | $x_i - \bar{x}$ | $(x_i - \bar{x})^2$
10 | –5.4 | 29.16
12 | –3.4 | 11.56
16 | 0.6 | 0.36
19 | 3.6 | 12.96
20 | 4.6 | 21.16
SUM = 75.2
$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$


x_i | $x_i - \bar{x}$ | $(x_i - \bar{x})^2$
10 | –5.4 | 29.16
12 | –3.4 | 11.56
16 | 0.6 | 0.36
19 | 3.6 | 12.96
20 | 4.6 | 21.16
$s = \sqrt{\frac{75.2}{n-1}}$

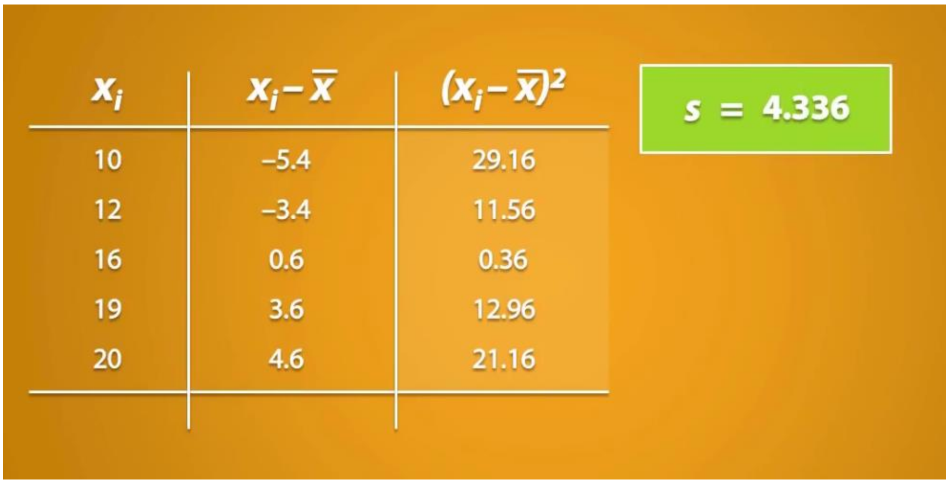

x_i | $x_i - \bar{x}$ | $(x_i - \bar{x})^2$
10 | –5.4 | 29.16
12 | –3.4 | 11.56
16 | 0.6 | 0.36
19 | 3.6 | 12.96
20 | 4.6 | 21.16
$s = \sqrt{\frac{75.2}{5-1}}$

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
10	–5.4	29.16
12	–3.4	11.56
16	0.6	0.36
19	3.6	12.96
20	4.6	21.16

$$s = \sqrt{\frac{75.2}{4}}$$

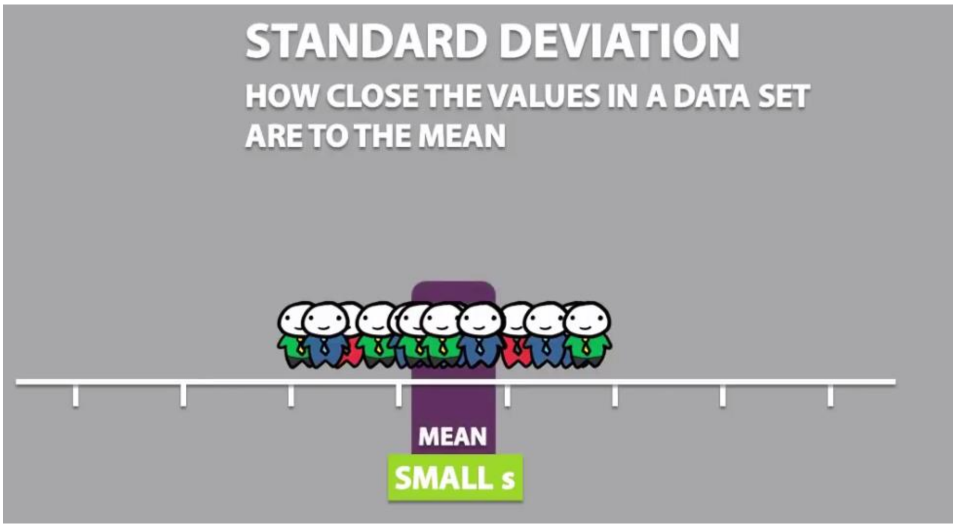
x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
10	–5.4	29.16
12	–3.4	11.56
16	0.6	0.36
19	3.6	12.96
20	4.6	21.16

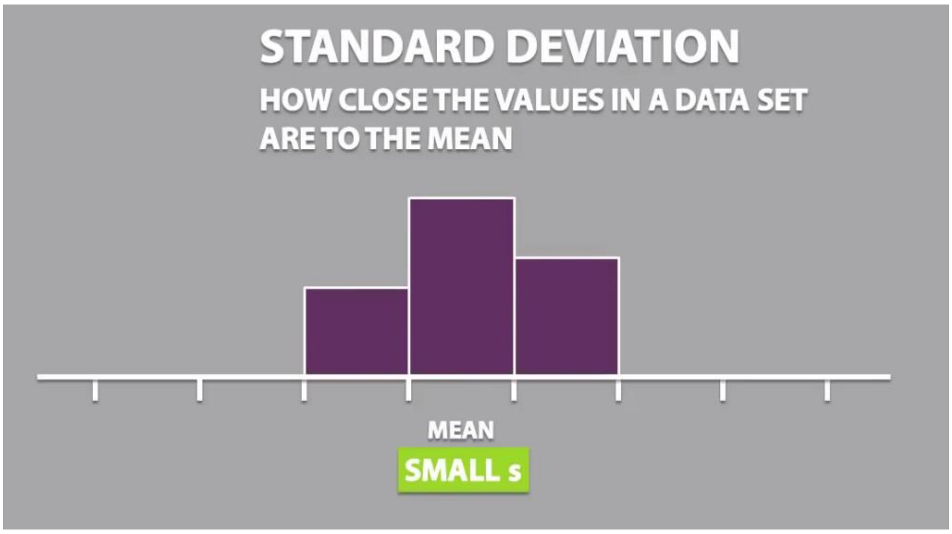
$$s = \sqrt{18.8}$$

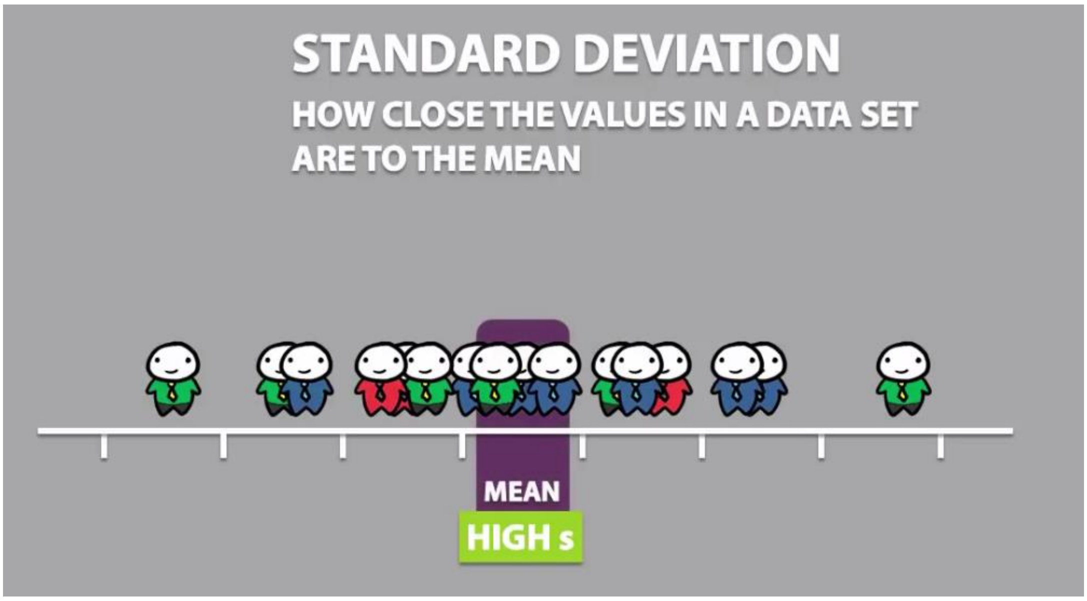


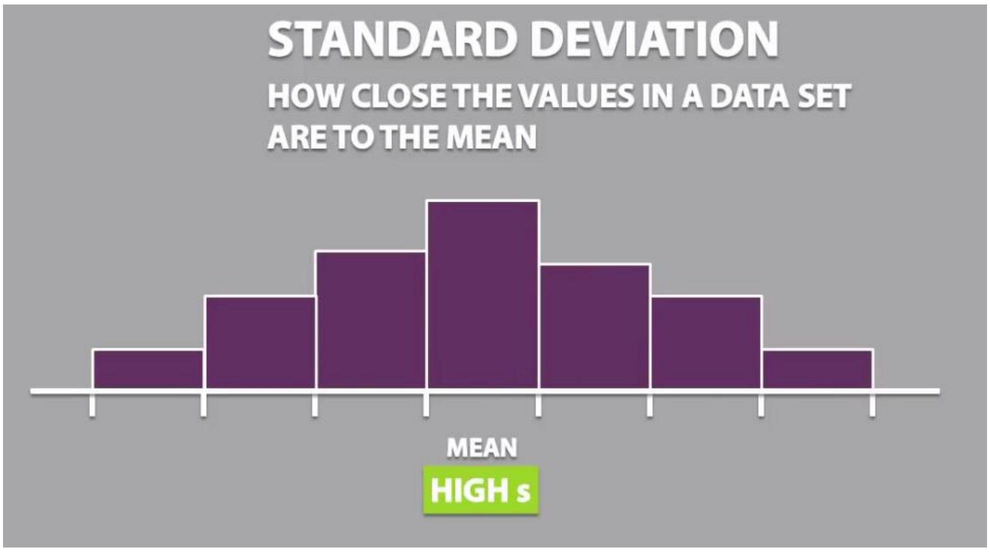
x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
10	–5.4	29.16
12	–3.4	11.56
16	0.6	0.36
19	3.6	12.96
20	4.6	21.16

$s = 4.336$


STANDARD DEVIATION
HOW CLOSE THE VALUES IN A DATA SET ARE TO THE MEAN
MEAN
SMALL s


STANDARD DEVIATION
HOW CLOSE THE VALUES IN A DATA SET ARE TO THE MEAN
MEAN
SMALL s


STANDARD DEVIATION
HOW CLOSE THE VALUES IN A DATA SET
ARE TO THE MEAN
MEAN
HIGH s


STANDARD DEVIATION
HOW CLOSE THE VALUES IN A DATA SET
ARE TO THE MEAN
MEAN
HIGH s

10. Variance

- ✓ The average of squared differences from the mean.
- ✓ Variance is the average of squared differences from the mean
- ✓ By using this we can find how far the data points in a population are from the population mean.