The **sample mean** and **median** are sample statistics that indicate the *middle*. We use $\bar{x}$ (pronounced "x bar") to represent the sample mean. We do not have a standard symbol for the median.

The **standard deviation** and **inter-quartile range** are sample statistics that indicate *spread*. We use s to represent the sample's standard deviation. We abbreviate inter-quartile range as IQR.

Calculating these statistics

Let's start with an example.

participant	score
1	61
2	43
3	48
4	40
5	51
6	60
7	47

To calculate the **mean**, we add up the scores and divide by the sample size.

$$\bar{x} = \frac{61 + 43 + 48 + 40 + 51 + 60 + 47}{7} = 50$$

The general formula for n scores can be expressed in two different ways.

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n}$$

where x_1 is the first score, x_2 is the second score, etc...
We can also use the summation operator.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

where i is a incrementing index that starts at 1 and ends at n.

To determine the **median**, we order the scores and find the middle number.

$$\cancel{40} \quad \cancel{43} \quad \cancel{47} \quad 48 \quad \cancel{51} \quad \cancel{60} \quad \cancel{61}$$

If there were an even number of scores, the median is the average of the two middle scores. For example, the median of $\{1, 3, 4, 9\}$ is 3.5.

To calculate the **standard deviation**, we often use a table.

i	x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1	61	11	121
2	43	−7	49
3	48	−2	4
4	40	−10	100
5	51	1	1
6	60	10	100
7	47	−3	9
		sum →	384

We have found the sum of the squared deviations.

$$\sum_{i=1}^{7} (x_i - \bar{x})^2 = 384$$

We divide this by the sample size minus one, then find the square root of the result.

$$s = \sqrt{\frac{384}{7-1}} = 8$$

The general formula for standard deviation is a little scary.

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

The **IQR** is the difference between the third quartile (Q_3) and the first quartile (Q_1).

$$\text{IQR} = Q_3 - Q_1$$

The first quartile is the median of the low scores. The third quartile is the median of the high scores.

$$\begin{array}{ccc} \underbrace{40 \quad 43 \quad 47}_{\text{low scores}} & \cancel{48} & \underbrace{51 \quad 60 \quad 61}_{\text{high scores}} \\ \cancel{40} \quad \underset{Q_1}{43} \quad \cancel{47} & & \cancel{51} \quad \underset{Q_3}{60} \quad \cancel{61} \end{array}$$

So,

$$\text{IQR} = 60 - 43 = 17$$

And in general,

Step 1 Use the median to divide the ordered data set into two halves.

- If there are an odd number of data points in the original ordered data set, do not include the median in either half.
- If there are an even number of data points in the original ordered data set, split this data set exactly in half.

Step 2 The lower quartile value is the median of the lower half of the data. The upper quartile value is the median of the upper half of the data. Find their difference.

Q1: Determine the mean, median, standard deviation, and IQR of the following scores.

47 53 53 55 55 52 49

Q2: Determine the mean, median, standard deviation, and IQR of the following scores.

50 58 22 53 47 46 52 35 51

Q3: Determine the mean, median, standard deviation, and IQR of the following scores.

$$4 \quad 2 \quad 8 \quad 6 \quad 5 \quad 5$$

Q4: Determine the mean, median, standard deviation, and IQR of the following scores.

$$104 \quad 102 \quad 108 \quad 106 \quad 105 \quad 105$$

Q5: Imagine 19 exams have a mean score of 75. Now, one more student finishes the exam and scores 100. What is the new mean?

Q6: Imagine 10 exams have a mean score of 80. Now, one more student finishes the exam and scores 100. What is the new mean?

Q7: Imagine 4 exams have a mean score of 85. What is the fifth score that can bring the average to 87?

A1: $\bar{x} = 52$
median = 53
$s = 3$
IQR = 6

A2: $\bar{x} = 46$
median = 50
$s = 11$
IQR = 12 (from 52.5 − 40.5)

A3: $\bar{x} = 5$
median = 5
$s = 2$
IQR = 2

A4: $\bar{x} = 105$
median = 105
$s = 2$
IQR = 2

A5: 76.25

A6: 81.82

A7: 95