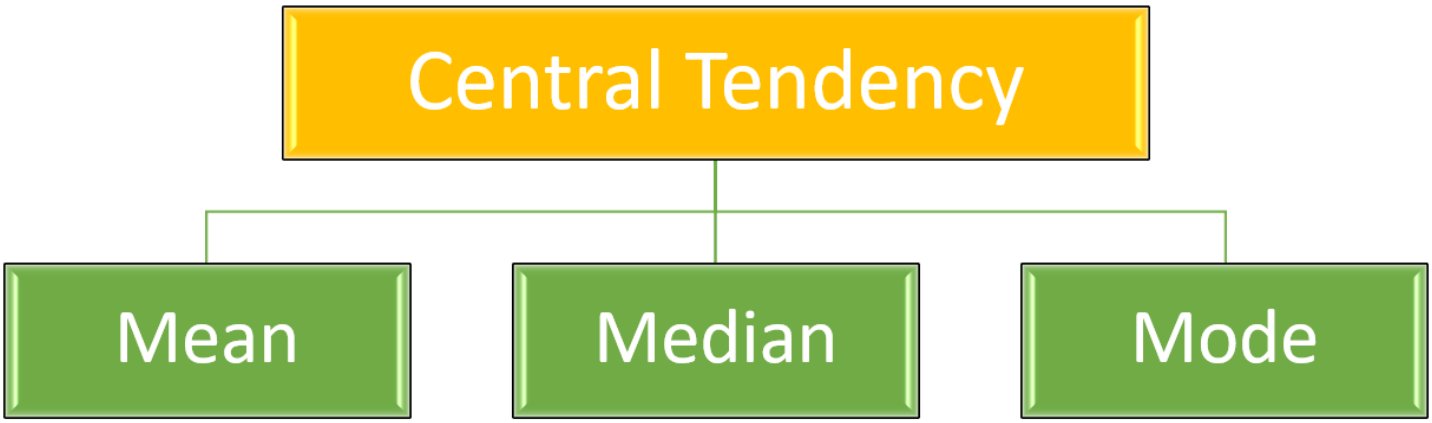
Statistics Basics (DS-11)

Central Tendency (Measure of Centre)

Done: Go through the activity to the end

Description

The central tendency concept is that one single value can best describe the data. Mean, median, and mode are the three important parameters in statistics. Essentially, all three of them refer to a single aspect called the **Central Tendency**. Let's take a closer look at this.



The mean (the average) is most the famous measure of central tendency that you are probably familiar with, but there are also others, such as the median and the mode. The mean, median and mode are all valid measures of central tendency, but under various conditions, one measure of central tendency might become more appropriate than others.

Previous Next

You have completed 100% of the lesson

100%

Jump to...

Statistics Basics (DS-11)

Central Tendency (Measure of Centre)

Done: Go through the activity to the end

Let's Practice

Question
Correct
Mark 1.00 out of 1.00

We understand "mean", "median" and "mode" from the central tendency concept.

Select one:

- (•) True ✔
- ( ) False

Check

Congrats! You are right.

Correct

Marks for this submission: 1.00/1.00.

Previous Next

You have completed 100% of the lesson

100%

Jump to...

Statistics Basics (DS-11)

Central Tendency (Measure of Centre)

Done: Go through the activity to the end

Mean (μ)

The mean is equal to the sum of the values in the dataset divided by the number of values. The number of values in the dataset will be equal to the population or sample size. The table below gives the formula for the population mean and the sample mean.

Population Mean

$$\mu = \frac{\sum x}{N}$$

Sample Mean

$$\bar{X} = \frac{\sum x}{n}$$

N = the number of observations in the population

n = the number of observations in the sample

One of the major disadvantages of using mean rather than using median or mode is, the mean is particularly sensitive to the effect of extreme values. Extreme values are also called **outliers**. We will make a technical description of outliers, however, these are the values that are unusual by being relatively small or large in numerical value compared to the rest of the dataset. For example, consider people's salaries at a factory below.

Staff	Salary (thousand $)
1	102
2	33
3	26
4	27
5	30
6	25
7	33

Staff	Salary (thousand $)
8	33
9	24

We can say $102K dollars is an outlier. The total salary of employees is $333K and the sample size is nine. The mean salary for these nine staff is $37K (333/9=37). Inspecting the raw data, however, suggests that this mean value may not be the best way to accurately reflect a staff's typical salary, as most staff receive salaries between $24K and $33K thousand dollars. We would like to have a better measure of central tendency in this situation. Therefore, taking the median might be a better measure of central tendency.

Tips:

- Under various conditions, one measure of central tendency might become more appropriate than others.

Previous Next

Statistics Basics (DS-11)

Central Tendency (Measure of Centre)

Done: Go through the activity to the end

Let's Practice

Question
Correct
Mark 1.00 out of 1.00

The **mean** is always a more appropriate measure of central tendency method than others.

Select one:

- True
- False ✔

Check

Congrats! You are right.

Correct

Marks for this submission: 1.00/1.00.

Previous Next

You have completed 100% of the lesson

100%

Jump to...

Statistics Basics (DS-11)

Central Tendency (Measure of Centre)

Done: Go through the activity to the end

Median

The median is the middle score for a dataset that has been sorted from small to large. Outliers less affect the median. Suppose we have the same data to calculate the median.

Staff	Salary (thousand $)
1	102
2	33
3	26
4	27
5	30
6	25
7	33
8	33
9	24

Firstly, we need to sort the data from small to large.

Staff	Salary (thousand $)
1	24
2	25
3	26
4	27
5	30
6	33
7	33

Staff	Salary (thousand $)
8	33
9	102

The median is the middle score, in this case, it is $30K. $30K is the middle score because there are 4 scores after it and 4 scores before it. This works well when you have an odd number of scores, but what will happen when you have an even number of sample size? Even if you only had 10 scores? In this case, we simply have to take the middle two scores and average the result. So, if we look at the example below:

Staff	Salary (thousand $)
1	24
2	25
3	26
4	27
5	30
6	33
7	33
8	33
9	50
10	102

Now, we should take the 5th ($30K) and 6th ($33K) scores in our dataset and average them to get a median of $31.5K.

Tips:

- The median is the middle score. If the sample size is 9, the fifth element is the median.

Previous Next

Statistics Basics (DS-11)

Central Tendency (Measure of Centre)

Done: Go through the activity to the end

Let's Practice

Question
Correct
Mark 1.00 out of 1.00

The median is the middle score ✔ for a dataset that has been sorted from small to large.

Check

Your answer is correct.

You have correctly selected 1.

Correct

Marks for this submission: 1.00/1.00.

Previous Next

You have completed 100% of the lesson

100%

Jump to...

Statistics Basics (DS-11)

Central Tendency (Measure of Centre)

Done: Go through the activity to the end

Mean vs. Median

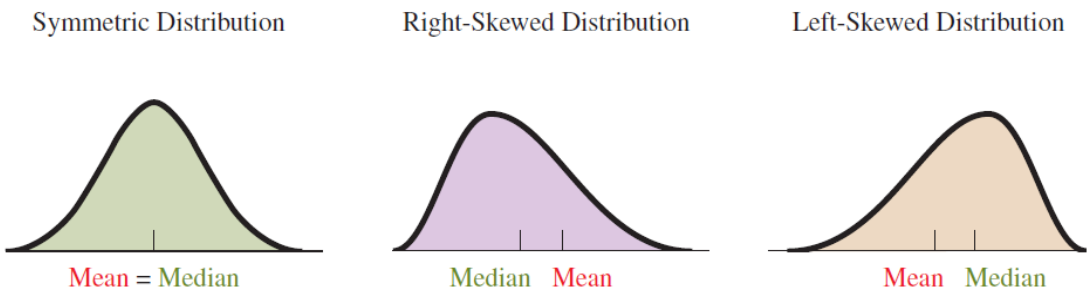
While the mean is the balance point of the data, the median is the middle point.

The mean can be highly influenced by an outlier. The mean is better if a large set of scores does not have an outlier.

The median is not sensitive to outliers. The median is better if a small set of scores has an outlier.

Generally, if the shape is

- Perfectly symmetric, the mean equals the median.
- Skewed to the right, the mean is larger than the median.
- Skewed to the left, the mean is smaller than the median.



Outlier Effect

In 1984 the University of Virginia announced that its Department of Rhetoric and Communication graduates' mean starting salary was $55,000. The outlier, the salary of N.B.A. center Ralph Sampson (his starting salary exceeded $1 million), did not represent the earning power of a B.A. in speech from the University of Virginia. (The median salary wasn't published.)

*Source: **The Cartoon Guide to Statistics** by Larry Gonick*

Previous Next

You have completed 100% of the lesson

100%

Jump to...

Reset user tour on this page

Statistics Basics (DS-11)

Central Tendency (Measure of Centre)

Done: Go through the activity to the end

Let's Practice

Question
Correct
Mark 1.00 out of 1.00

A population of scores has a mean of 26, a median of 23, and a mode of 22.

What is the most likely shape for the population distribution?

- (•) Positively skewed ✔ Generally, if the shape is positively skewed (right skewed), the mean is larger than the median.
- ( ) It cannot be determined from the information given
- ( ) Negatively skewed
- ( ) Symmetrical

Clear my choice

Check

Your answer is correct.

Correct

Marks for this submission: 1.00/1.00.

Previous Next

You have completed 100% of the lesson

100%

Jump to...

Reset user tour on this page

CLARUSWAY
WAY TO REINVENT YOURSELF

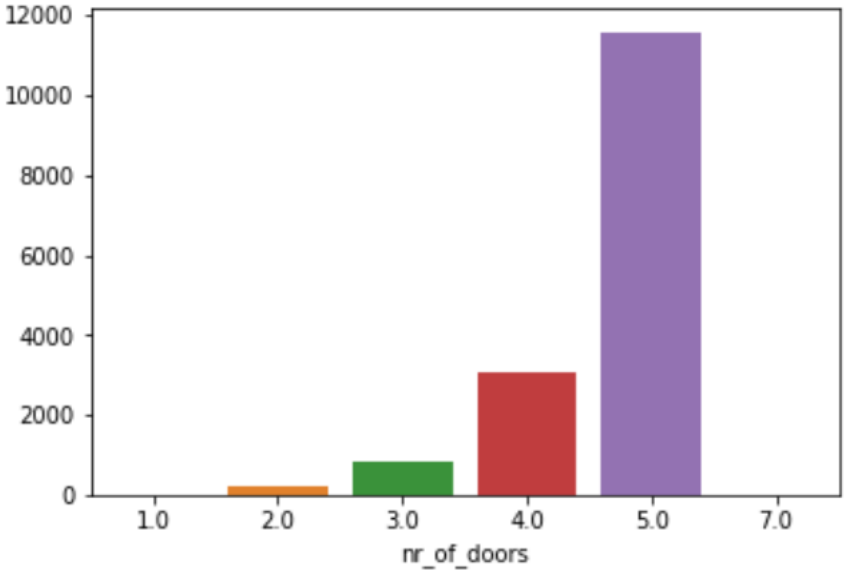
Statistics Basics (DS-11)

Central Tendency (Measure of Centre)

Done: Go through the activity to the end

Mode

The mode is the most frequent score in a dataset. It represents the highest bar in a histogram or bar chart. Therefore, sometimes you can consider the mode as being the most popular option. The mode is normally used for categorical data where we want to know which category is the most common. An example of a mode is presented below.



The histogram shows the number of doors among used cars where the sample size is around 15000. We can say the most popular option is the 5-door cars. Therefore, the mode for this dataset is 5.

To find out the mode, suppose we have the same dataset.

Staff	Salary (thousand $)
1	102
2	33
3	26
4	27
5	30
6	25
7	33

Staff	Salary (thousand $)
8	33
9	24

The mode is the most frequent score in a dataset. So, we can say $33K is the mode in our dataset. Because there are 3 different staff who receives $33K. Let's look at the mean, the median and the mode values for the same dataset:

- mean: $37K
- median: $30K
- mode : $33K

Tips:
- The mode is the most frequent option.
- If all observations are repeated an equal number of times, then the mode does not exist.

Previous Next

Statistics Basics (DS-11)

Central Tendency (Measure of Centre)

Done: Go through the activity to the end

Let's Practice

Question
Correct
Mark 1.00 out of 1.00

The mode is the most frequent ✔ score in a dataset.

Check

Your answer is correct.

You have correctly selected 1.

Correct

Marks for this submission: 1.00/1.00.

Previous Next

You have completed 100% of the lesson

100%

Jump to...

Statistics Basics (DS-11)

Central Tendency (Measure of Centre)

Done: Go through the activity to the end

Calculate Mean, Median and Mode with Python

We can easily calculate mean, median and mode values with Python. We use the "Numpy" library for the mean and median, and the "SciPy" library for the mode. You can compare the values we get by Python with the values we calculated manually.

input :

```
import numpy as np
from scipy import stats

salary = [102, 33, 26, 27, 30, 25, 33, 33, 24]

mean_salary = np.mean(salary)
print("mean:", mean_salary)

median_salary = np.median(salary)
print("median:", median_salary)

mode_salary = stats.mode(salary)
print("mode:", mode_salary)

```

output :

```
mean: 37.0
median: 30.0
mode: ModeResult(mode=array([33]), count=array([3]))

```

Python Playground

Hi **F1607 -**, you can write your code using the editor below.
Once you write the code, click the run button to see the result.

```
1
```

Run

Previous Next

Started on	Monday, 18 April 2022, 10:24 AM
State	Finished
Completed on	Monday, 18 April 2022, 10:30 AM
Time taken	6 mins 21 secs
Marks	6.00/6.00
Grade	**10.00** out of 10.00 (**100**%)

Question **1**

Correct

Mark 1.00 out of 1.00

Which statistical number is the middle number of the data set?

Select one:

- median ✔ Congrats! You are right.
- mode
- range
- mean

Your answer is correct.

The correct answer is: median

Question **2**

Correct

Mark 1.00 out of 1.00

Richard took a sample of 100 pieces of data. He added up all of the pieces of data and then divided by 100. What measure of center is he working on?

Select one:

- range
- mode
- median
- mean ✔ Congrats! You are right.

Your answer is correct.

The correct answer is: mean

Question **3**

Correct

Mark 1.00 out of 1.00

A set of data is given: 3, 7, 10, 10, 16
Jason calculated a measure of center and got 9.2, which measure of center did he just calculate?

Select one:

- range
- mean ✔ Congrats! You are right.
- mode
- median

Your answer is correct.

The correct answer is: mean

Question **4**

Correct

Mark 1.00 out of 1.00

A set of data is given: 3, 7, 10, 10, 16. What is the mode?

Select one:

- 16
- 10 ✔ Congrats! You are right.
- 3
- 7

Your answer is correct.

The correct answer is: 10

Question **5**

Correct

Mark 1.00 out of 1.00

A set of data is given: 100, 55, 95, 150, 101, 99, 53, 57, 70. What is the median?

Select one:

- 57
- 100
- 150
- 95 ✔ Congrats! You are right.

Your answer is correct.

The correct answer is: 95

Question **6**

Correct

Mark 1.00 out of 1.00

Which of the following statements is false?

- The mean can be highly influenced by an outlier.
- The mean is pulled in the direction of the peak in a skewed distribution. ✔ The mean tends to be closer to the longer tail in a skewed distribution.
- In a symmetrical distribution, the mean, median and mode are identical.
- The mean is the balance point of the data.

Your answer is correct.

The correct answer is:
The mean is pulled in the direction of the peak in a skewed distribution.

◄ Central Tendency (Measure of Centre)

Jump to...

Dispersion (Measure of Spread) ►

Statistics Basics (DS-11)

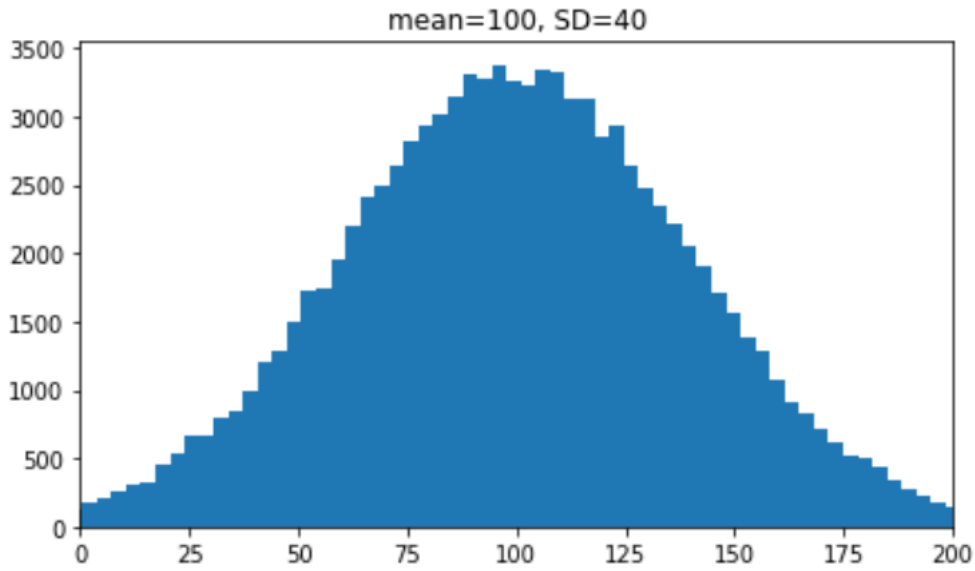
Dispersion (Measure of Spread)

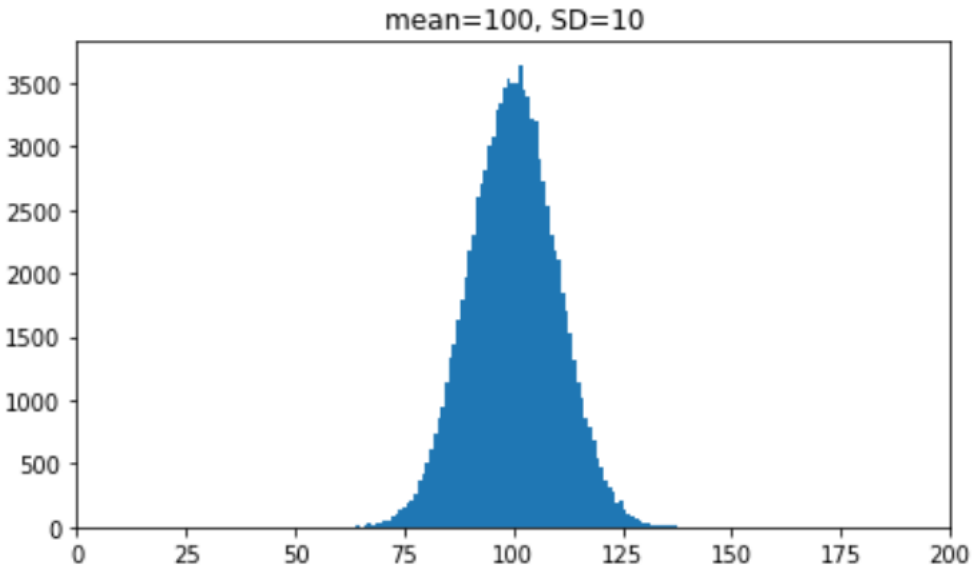
Done: Go through the activity to the end

Introduction

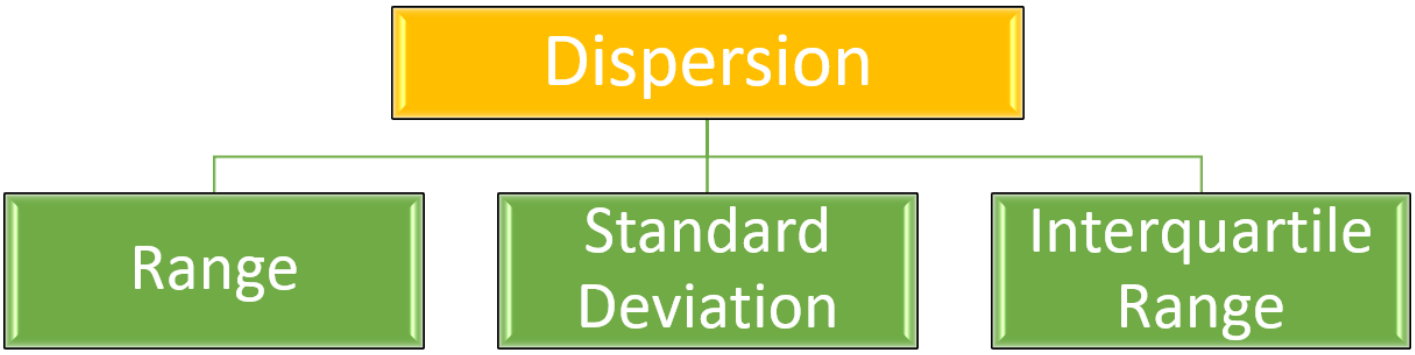
In statistics, the measure of central tendency gives a single value that represents the entire data; however, a single value can not describe the observation exactly. At this point, the dispersion helps us to study the variability of the items. Dispersion is a way to explain how a dataset is distributed. When a dataset has a small value, the values in the dataset are tightly clustered; when it is large the items in the set are widely scattered.

As can be seen in the two histograms below, there may be different distributions with the same mean value (μ =100). The first population is much more dispersed than the second population, however, the mean value for both populations is the same. Therefore, we can say a dispersion explains something more than the central tendency does.





Range, standard deviation and interquartile range are the three widely used measures of dispersion. Now, we are going to consider these concepts.



Tips:

- A dispersion explains something more than the measure of central tendency does.

Previous Next

You have completed 100% of the lesson

100%

Jump to...

Statistics Basics (DS-11)

Dispersion (Measure of Spread)

Done: Go through the activity to the end

Let's Practice

Question
Correct
Mark 1.00 out of 1.00

Central tendency explains something more than a dispersion does.

Select one:

- True
- False ✔

Check

Congrats! You are right.

Correct

Marks for this submission: 1.00/1.00.

Previous Next

You have completed 100% of the lesson

100%

Jump to...

Statistics Basics (DS-11)

Dispersion (Measure of Spread)

Done: Go through the activity to the end

Range

The range is the simple measure of dispersion, which is defined as the difference between the maximum and the minimum values. The main advantage of the range is that it is easy to calculate. There are many disadvantages, on the other hand. It is highly susceptible to extreme values and does not use all the observations in a dataset. In this case, if we look at our salary table again:

Staff	Salary (thousand $)
1	24
2	25
3	26
4	27
5	30
6	33
7	33
8	33
9	102

$$Range = MaximumValue - MinimumValue$$

The difference between the maximum value and the minimum value is 102-24=78. We can say the range for this dataset is 78. As you see in the example it gives the difference between the maximum and minimum two values out of nine values and does not use all the observations. And it is highly susceptible to extreme values because $102K affected the range very badly. If we remove that value, the range of the remaining values will be 9 (33-24= 9).

Previous Next

You have completed 100% of the lesson

100%

Jump to...

Statistics Basics (DS-11)

Dispersion (Measure of Spread)

Done: Go through the activity to the end

Let's Practice

Question
Correct
Mark 1.00 out of 1.00

The range defined as the difference ✔ between the maximum and the minimum values.

Check

Your answer is correct.

You have correctly selected 1.

Correct

Marks for this submission: 1.00/1.00.

Previous Next

You have completed 100% of the lesson

100%

Jump to...

Statistics Basics (DS-11)

Dispersion (Measure of Spread)

Done: Go through the activity to the end

Standard Deviation(σ)

The most commonly used measure of dispersion is the standard deviation (σ). Standard deviation measures the spread around the mean. It is also expressed as the square root of variance. Therefore we must first describe the variance(σ^2). Variance is defined as the average of the squared differences from the mean. The formula for the variance and standard deviation is given below.

Variance=

$$\sigma^2 = \frac{\sum(x-\mu)^2}{N}$$

Standard Deviation=

$$\sigma = \sqrt{\frac{\sum(x-\mu)^2}{N}}$$

μ = population mean
N = number of items in the population

Let's go back to our salary table again and calculate the standard deviation and the variance for these salaries.

Staff	Salary (thousand $)
1	24
2	25
3	26
4	27
5	30
6	33
7	33
8	33
9	102

$$\mu = \frac{24+25+26+27+30+33+33+33+102}{9}$$

$$\mu = \frac{333}{9} = 37$$

$$\sigma = \sqrt{\frac{\sum(x-\mu)^2}{N}}$$

$$\sigma = \sqrt{\frac{(24-37)^2+(25-37)^2+(26-37)^2+(27-37)^2+(30-37)^2+(33-37)^2+(33-37)^2+(33-37)^2+(102-37)^2}{9}}$$

$$\sigma = \sqrt{\frac{(-13)^2+(-12)^2+(-11)^2+(-10)^2+(-7)^2+(-4)^2+(-4)^2+(-4)^2+(65)^2}{9}}$$

$$\sigma = \sqrt{\frac{169+144+121+100+49+16+16+16+4225}{9}}$$

$$\sigma = \sqrt{\frac{4856}{9}}$$

$$\sigma = \sqrt{539,55}$$

$$\sigma = 23,22833518$$

Like the range, also standard deviation is affected by outliers. One value could contribute greatly to the results of the standard deviation. This also means the standard deviation is a good indicator of the existence of outliers.

For example if we remove the outlier from our table, the salary table would be:

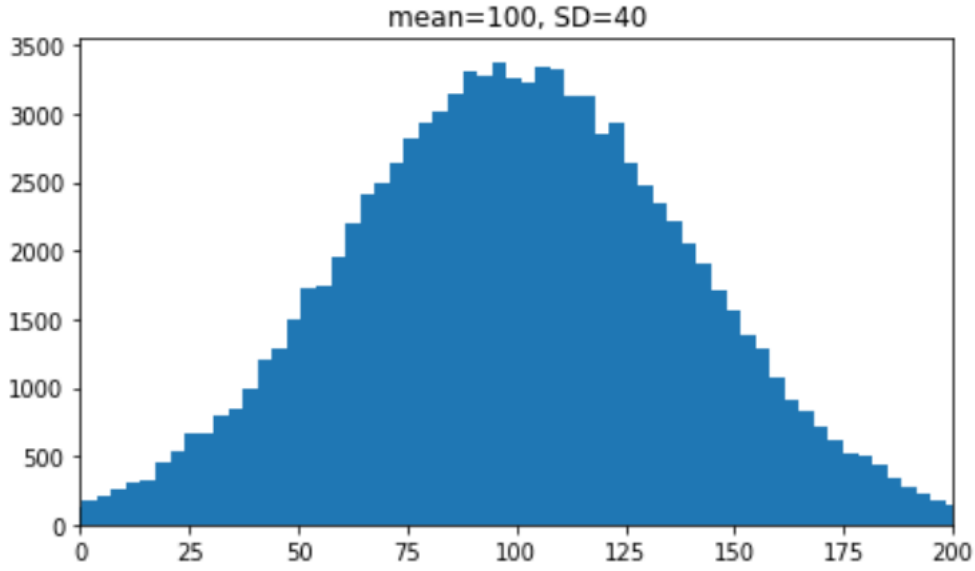
Staff	Salary (thousand $)
1	24
2	25
3	26
4	27
5	30
6	33
7	33
8	33

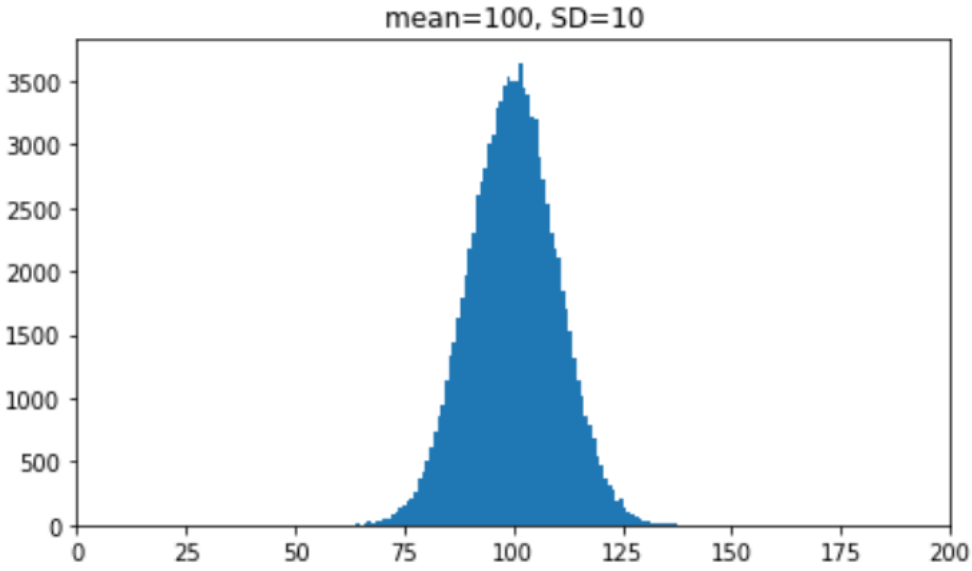
And if we recalculate the standard deviation regarding the new salary table:

$$\sigma = 3,58$$

You can see how only one outlier affects the standard deviation.

The standard deviation is also useful when comparing the spread of two different datasets that have the same mean. The dataset with the smaller standard deviation has a narrower spread of measurements around the mean and therefore usually has relatively less high or low values. In the following example, the standard deviation for the first population is 40, however, the standard deviation for the second one is 10. You see the second population has a narrower spread of measurements around the mean.





Tips:

- The data with the smaller standard deviation has a narrower spread of measurements around the mean.

Previous Next

You have completed 100% of the lesson

100%

Jump to...

Reset user tour on this page

Statistics Basics (DS-11)

Dispersion (Measure of Spread)

Done: Go through the activity to the end

Let's Practice

Question
Correct
Mark 1.00 out of 1.00

Which one is **not** correct about the standard deviation (σ).

Select one:

- (●) It measures the spread around the median. ✔ Congrats! You are right.
- ( ) It can be expressed as the square root of variance.
- ( ) It measures the spread around the mean.

Clear my choice

Check

Your answer is correct.

Correct

Marks for this submission: 1.00/1.00.

Previous Next

You have completed 100% of the lesson

100%

Jump to...

Statistics Basics (DS-11)

Dispersion (Measure of Spread)

Done: Go through the activity to the end

Interquartile Range (IQR)

Quartiles are the values that divide a group of numbers into quarters.

- Q1 or the 25th percentile is the first quartile and defined as the middle number between the smallest number and the median of the dataset.
- Q2 is the second quartile which is the median of the whole dataset.
- Q3 or 75th percentile is the third quartile which is the middle value between the median and the highest value of the dataset.

For example, a dataset consists of those numbers: 0,4,5,7,8,9,10,12,13,14,15,16,20.

The median (Q2) is the value in the middle of the list. In this case, 10 is the median number.

The first quartile (Q1) is the middle number in between the smallest number (0) and the median (10) which is 7. In other words, the middle number between 0 and 10 is 7.

The third quartile (Q3) is the middle value between the median (10) and the highest value (20) in this case that will be 14. In other words, the middle number between 10 and 20 will be 14.

Interquartile Range(IQR) is the **difference between Q3 and Q1**. In this case:

IQR = Q3 - Q1

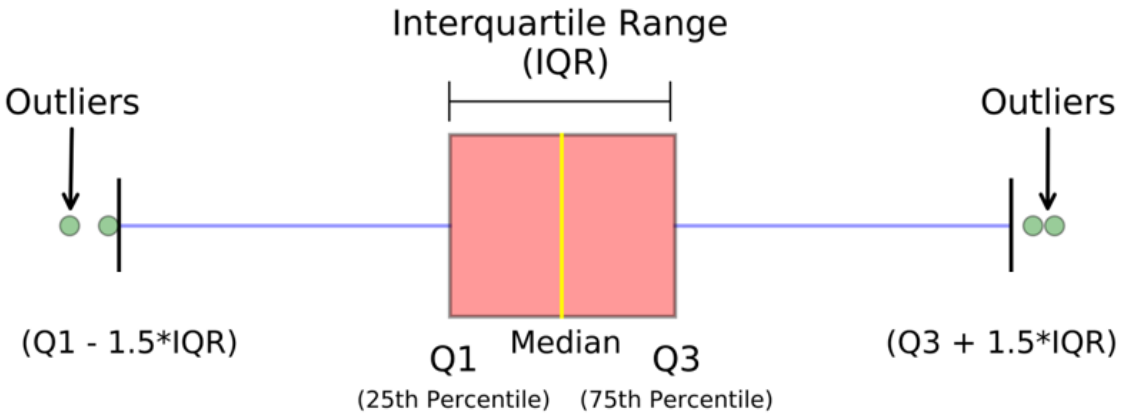
IQR = 14 - 7 = 7

As you remember we have already mentioned the extreme values and named them as outliers. In statistics, an outlier is a data point that differs significantly from other observations. IQR helps us to make a technical description of outliers. A typical definition of the outlier is, any data point more than 1.5 interquartile ranges (IQRs) below the first quartile or above the third quartile.

In this case, we can say:

Outliers are any data point below (Q1 - 1.5 * IQR) or above (Q3 + 1.5 * IQR).

The following picture shows the relationship between IQR and outliers.



Tips:

- Outlier is, any data point more than 1.5 IQR below the Q1 or above the Q3.

Previous Next

You have completed 100% of the lesson

100%

Jump to...

Statistics Basics (DS-11)

Dispersion (Measure of Spread)

Done: Go through the activity to the end

Let's Practice

Question
Correct
Mark 1.00 out of 1.00

A definition of the outlier is, any data point more than 1.5 interquartile ranges (IQRs) below the first quartile or above the third quartile.

Select one:

- (•) True ✔
- ( ) False

Check

Congrats! You are right.

Correct

Marks for this submission: 1.00/1.00.

Previous Next

You have completed 100% of the lesson

100%

Jump to...

Statistics Basics (DS-11)

Dispersion (Measure of Spread)

Done: Go through the activity to the end

Practice IQR

Imagine we have the following number list.

number_list = [1, 5, 10, 15, 40]

Now we will try to find which numbers on our list are the outlier.

We have the following summary:

minimum number = 1

maximum number =40

median=10

Q1 = 5

Q3 = 15

IQR = Q3-Q1

IQR= 15-5 = 10

Therefore, (1.5 * IQR) = 15

To determine if there are any outliers, we must consider the numbers that are 1.5*IQR beyond the quartiles.

Q1 – (1.5 * IQR) = 5-15 = -10

Q3 + (1.5 * IQR) = 15+15 = 30

The last number in our list is 40. And it is outside of the interval from (–10) to (30), therefore 40 is an outlier. The rest of the numbers in the list are not outliers.

Previous Next

You have completed 100% of the lesson

100%

Jump to...

Statistics Basics (DS-11)

Dispersion (Measure of Spread)

Done: Go through the activity to the end

Box Plot

Box plots (also called **box-and-whisker plots** or **box-whisker plots**) give a good graphical image of the concentration of the data. They also show how far the extreme values are from most of the data. A box plot is constructed from five values: *the minimum value, the first quartile, the median, the third quartile,* and *the maximum value.* We use these values to compare how close other data values are to them.

Box Plot
One of the more effective graphical summaries of a data set, the box plot generally shows mean, median, 25th and 75th percentiles, and outliers.

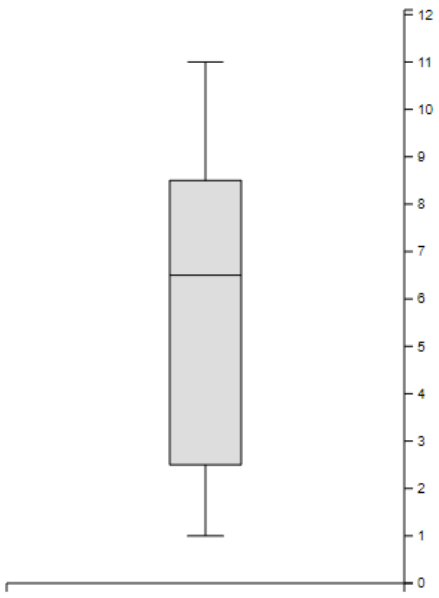
To construct a box plot, use a horizontal or vertical number line and a rectangular box. The smallest and largest data values label the endpoints of the axis. The first quartile marks one end of the box and the third quartile marks the other end of the box. Approximately **the middle 50 percent of the data fall inside the box.** The "whiskers" extend from the ends of the box to the smallest and largest data values. The median or second quartile can be between the first and third quartiles, or it can be one, or the other, or both. The box plot gives a good, quick picture of the data.

Box Plot (Min & Max)

Consider the following dataset.

1	1	2	2	4	6.8	7	8	8.3	9	10	10	11.5

The first quartile is two, the median is seven, and the third quartile is nine. The smallest value is one, and the largest value is 11.5. The following image shows the constructed box plot.



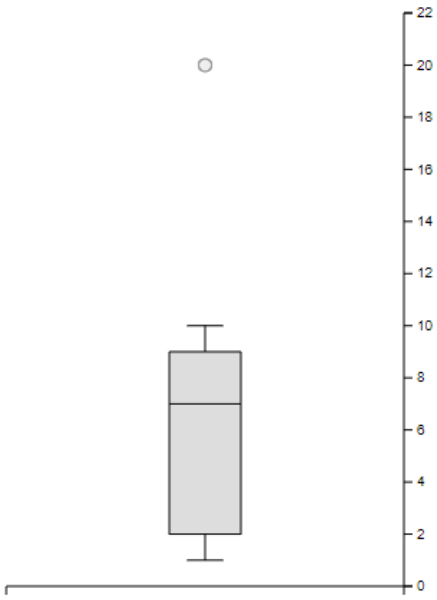
The two whiskers extend from the first quartile to the smallest value and from the third quartile to the largest value. The median is shown in the middle of the box.

Box Plot (1.5xIQR Rule)

Consider the following dataset. (Replace 11.5 with 20)

1	1	2	2	4	6.8	7	8	8.3	9	10	10	20

The first quartile (Q1) is 2, the median (Q2) is 7, and the third quartile (Q3) is 9. If a point is more than 1.5*IQR from an end of the box, it's an outlier. Here the last value 20 is an outlier. The outlier will be drawn individually. Finally, extend "Whiskers" out to the farthest points that are not outliers (within 1.5*IQR of the quartiles). The following image shows the constructed box plot.



Tips:

- You may encounter box-and-whisker plots that have dots marking outlier values. In those cases, the whiskers are not extending to the minimum and maximum values.

Previous Next

You have completed 100% of the lesson

100%

Jump to...

Reset user tour on this page

Statistics Basics (DS-11)

Dispersion (Measure of Spread)

Done: Go through the activity to the end

Let's Practice

Question
Correct
Mark 0.67 out of 1.00

The '*middle*' line drawn inside the box shows the position of the median value ✔ .

The *ends of the 'box'* give the positions of the first and third quartiles ✔ .

The *ends of the 'whiskers'* give the minimum and maximum values ✔ in the data.

Check

Your answer is correct.

You have correctly selected 3.

Correct

Marks for this submission: 1.00/1.00. Accounting for previous tries, this gives **0.67/1.00**.

Previous Next

You have completed 100% of the lesson

100%

Jump to...

Reset user tour on this page

CLARUSWAY
WAY TO REINVENT YOURSELF

Statistics Basics (DS-11)

Dispersion (Measure of Spread)

Done: Go through the activity to the end

Calculate Range, Variance, Std Dev, Quartiles and IQR with Python

We can easily calculate range, variance and standard deviation values with numpy. You can compare the values we get by numpy with the values we calculated manually.

input :

```
import numpy as np
from scipy import stats

salary = [102, 33, 26, 27, 30, 25, 33, 33, 24]

print("Range: ", (np.max(salary)-np.min(salary)))

print("Variance: ", (np.var(salary)))

print("Std: ", (np.std(salary)))

print("Q1:", (np.percentile(salary, 25)))

print("Q2:", (np.percentile(salary, 50)))  #q2 is also called median

print("Q3:", (np.percentile(salary, 75)))

print("IQR:", (stats.iqr(salary)))

```

output :

```
Range:  78
Variance:  539.5555555555555
Std:  23.22833518691246
Q1: 26.0
Q2: 30.0
Q3: 33.0
IQR: 7.0


```

Python Playground

Hi **F1607 -**, you can write your code using the editor below.
Once you write the code, click the run button to see the result.

```
1
```

Run

Previous Next

You have completed 100% of the lesson

100%

Jump to...

Reset user tour on this page

CLARUSWAY
WAY TO REINVENT YOURSELF

Started on	Monday, 18 April 2022, 1:42 PM
State	Finished
Completed on	Monday, 18 April 2022, 1:51 PM
Time taken	9 mins 25 secs
Marks	6.00/7.00
Grade	**8.57** out of 10.00 (**86**%)

Question **1**

Correct

Mark 1.00 out of 1.00

What is the range for the data given:
1, 10, 7, 12, 0, 30, 15, 22, 8, 2

Select one:

- 30 ✔ Congrats! You are right.
- 0
- 1
- 22

Your answer is correct.

The correct answer is: 30

Question **2**

Correct

Mark 1.00 out of 1.00

What is the standard deviation for the data given:
1, 10, 7, 12, 0, 30, 15, 22, 8, 2

Select one:

- 6.089
- 7.089
- 9.089 ✔ Congrats! You are right.
- 8.089

Your answer is correct.

The correct answer is: 9.089

Question **3**

Correct

Mark 1.00 out of 1.00

If a number is inserted into a set that is far away from the mean, how does this affect the standard deviation?

Select one:

- decrease
- remains the same
- increase ✔ Congrats! You are right.
- approaches to zero

Your answer is correct.

The correct answer is: increase

Question **4**

Correct

Mark 1.00 out of 1.00

What is the IQR for the data given:
9, 11, 4, 14, 8, 2, 10, 3, 10, 9, 6, 0, 1

Select one:

- 7 ✔ Congrats! You are right.
- 9
- 11
- 6

Your answer is correct.

The correct answer is: 7

Question **5**

Correct

Mark 1.00 out of 1.00

What is the IQR for the data given:
8, 10, 4, 24, 8, 3, 10, 3, 40, 7, 6, 12, 4

Select one:

- 6 ✔ Congrats! You are right.
- 7
- 11
- 9

Your answer is correct.

The correct answer is: 6

Question **6**

Incorrect

Mark 0.00 out of 1.00

The mean and standard deviation are very good for summarizing the properties of fairly asymmetrical ✖ histograms with outliers ✖ .

Your answer is incorrect.

The correct answer is:
The mean and standard deviation are very good for summarizing the properties of fairly [symmetrical] histograms [without outliers].

Question **7**

Correct

Mark 1.00 out of 1.00

When using IQR to find outliers, which formula will identify the lower limit?

- Q1 + 1.5(IQR)
- Q1 - 1.5(IQR) ✔
- Q3 + 1.5(IQR)
- Q3 - 1.5(IQR)

Your answer is correct.

The correct answer is:
Q1 - 1.5(IQR)

◄ Complementary Lesson About Central Tendency and Dispersion

Jump to...

Interview Questions About Statistics ►