In this reading, you will learn how to calculate the standard deviation of a data and from the standard deviation how to find out outliners present in a data. So, please carefully read this article.

**Standard Deviation**

Standard deviation measures the variability of the data set. A smaller standard deviation indicates less variability.

The Standard Deviation is a measure of how spread out numbers are. Its symbol is σ (the Greek letter **sigma**)

Finding standard deviation requires summing the squared difference between each data point and the mean [∑(x-µ)2], adding all the squares, dividing that sum by one less than the number of values (N-1), and finally calculating the square root of the dividend. Mathematically, start with calculating the mean.

Calculate the mean by adding all the data point values, then dividing by the number of data points.

So let's consider a sample data set with 7 values like 20, 24, 25, 36, 25, 22, 23

In the sample data set, 20+24+25+36+25+22+23=175. Divide the sum, 175, by the number of data points, 7, or 175÷7=25. **The mean equals 25**.

Next, subtract the mean from each data point, then square each difference. The formula looks like this: ∑(x-µ)2, where ∑ means sum, x represents each data set value and µ represents the mean value.

Continuing with the example set, the values become: 20-25=-5 and -52=25; 24-25=-1 and -12=1; 25-25=0 and 02=0; 36-25=11 and 112=121; 25-25=0 and 02=0; 22-25=-3 and -32=9; and 23-25=-2 and -22=4.

Adding the squared differences yields: 25+1+0+121+0+9+4=**160**.

Divide the sum of the squared differences by one less than the number of data points. The example data set has 7 values, so N-1 equals 7-1=6. The sum of the squared differences, 160, divided by 6 equals approximately 26.6667.

Calculate the standard deviation by finding the square root of the division by N-1. In the example, the square root of 26.6667 equals approximately 5.164. Therefore, the standard deviation equals approximately **5.164**.

So the mean of the data is 25 and the standard deviation is 5.164. So now let's see how we can use standard deviation to find out the extreme values of the dataset. **Standard deviation helps evaluate data. Numbers in the data set that fall within one standard deviation of the mean are part of the data set. Numbers that fall outside of two standard deviations are extreme values or outliers.**

So to calculate 2 standard deviations just subtract and add the value of standard deviation value 2 times to the mean on either side.

In this example 2 standard deviations are

25 - 5.164 - 5.164 = **14.672**

and

25 + 5.164 +5.164 = **35.328**

So, In the example set, the value 36 lies more than two standard deviations from the mean, so 36 is an outlier. Outliers may represent erroneous data or may suggest unforeseen circumstances and should be carefully considered when interpreting data.