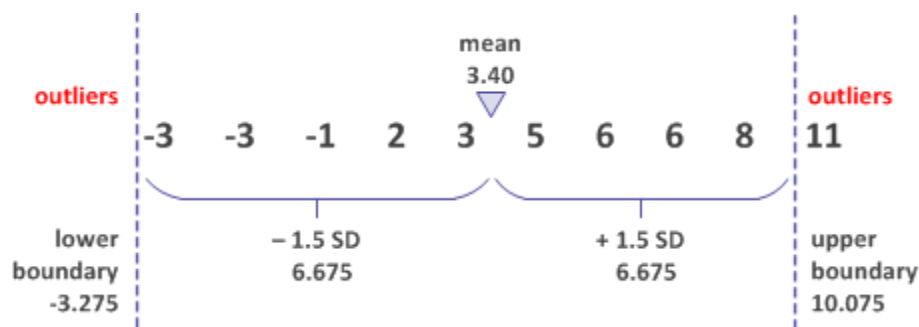


## Assignment About 1.5 Scale in Outlier-25.05.2024

### The mean $\pm$ 1.5 standard deviations

In the second example, you use the mean  $\pm$  1.5 standard deviations to establish the upper and lower outlier boundaries. Now only one value is identified as an outlier.



In **Number of times of S.dev**, specify a multiple of the standard deviation to use for the outlier boundaries.

You can specify any positive integer or decimal numeral (0.5, 1, 1.5, 2 . . . ).

When scale is taken as 1.5, then according to the IQR method any data that lies beyond  $2.7\sigma$  from the mean ( $\mu$ ), on either side, shall be considered an outlier. This decision range is the closest to what Gaussian Distribution tells us, i.e.,  $3\sigma$ . In other words, this makes the decision rule closest to what Gaussian distribution considers for outlier detection, and this is exactly what we wanted.

To get exactly  $3\sigma$ , we'd need to take the **scale = 1.7**. However, 1.5 is more "symmetrical" than 1.7, and we've always been a little more inclined towards symmetry, haven't we?

Also, IQR method of outlier detection is not the only nor best method for outlier detection. Some trade-off is acceptable.

Now you know why we take it **1.5 \* IQR**. But this scale depends on the distribution followed by the data. If my data seem to follow exponential distribution, then this scale would change.