Measures of Position:

Dataset: 10, 12, 15, 16, 17, 19, 21, 23, 25, 30, 32

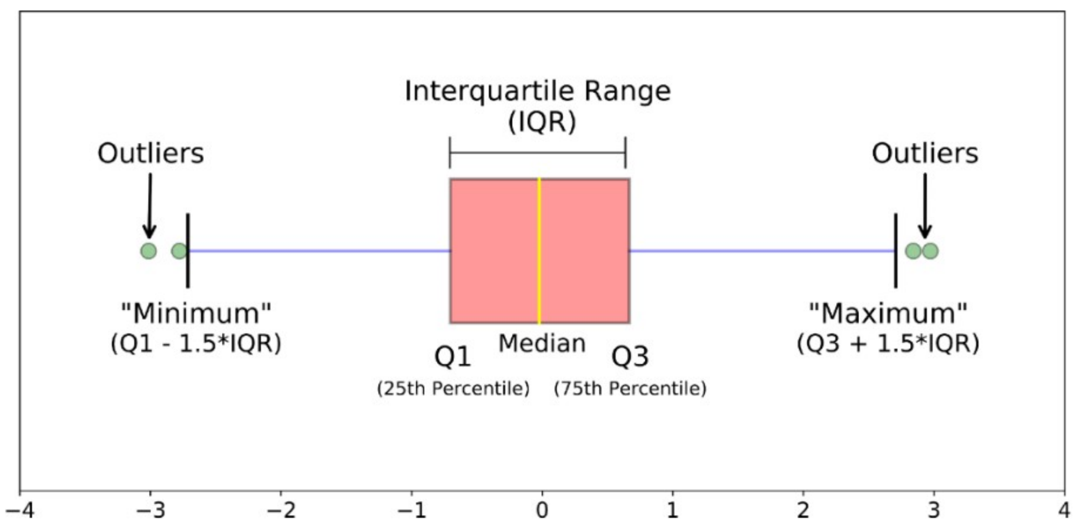
Median = 19, | Q2, 50% or the middle value of the given list of data when arranged in an order.
Q1=15, | Q1 is the value under which 25% of data points are found.
Q3=25, | Q3 the value under which 75% of data points are found.

The Five-Number Summary:

A helpful summary of the data is called the five number summary. The five number summary consists of five values:

The minimum; The lower quartile, Q1; The median (same as Q2); The upper quartile, Q3 and The maximum

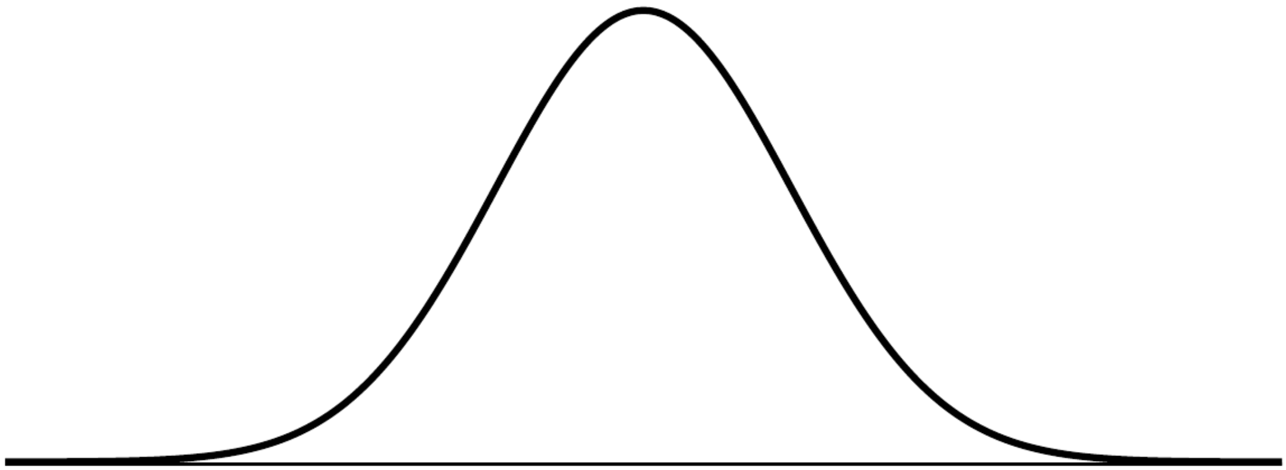
Boxplot —also called box and whisker plot—displays the five-number summary of a set of data.



$$\boldsymbol{IQR = Q3 - Q1}$$

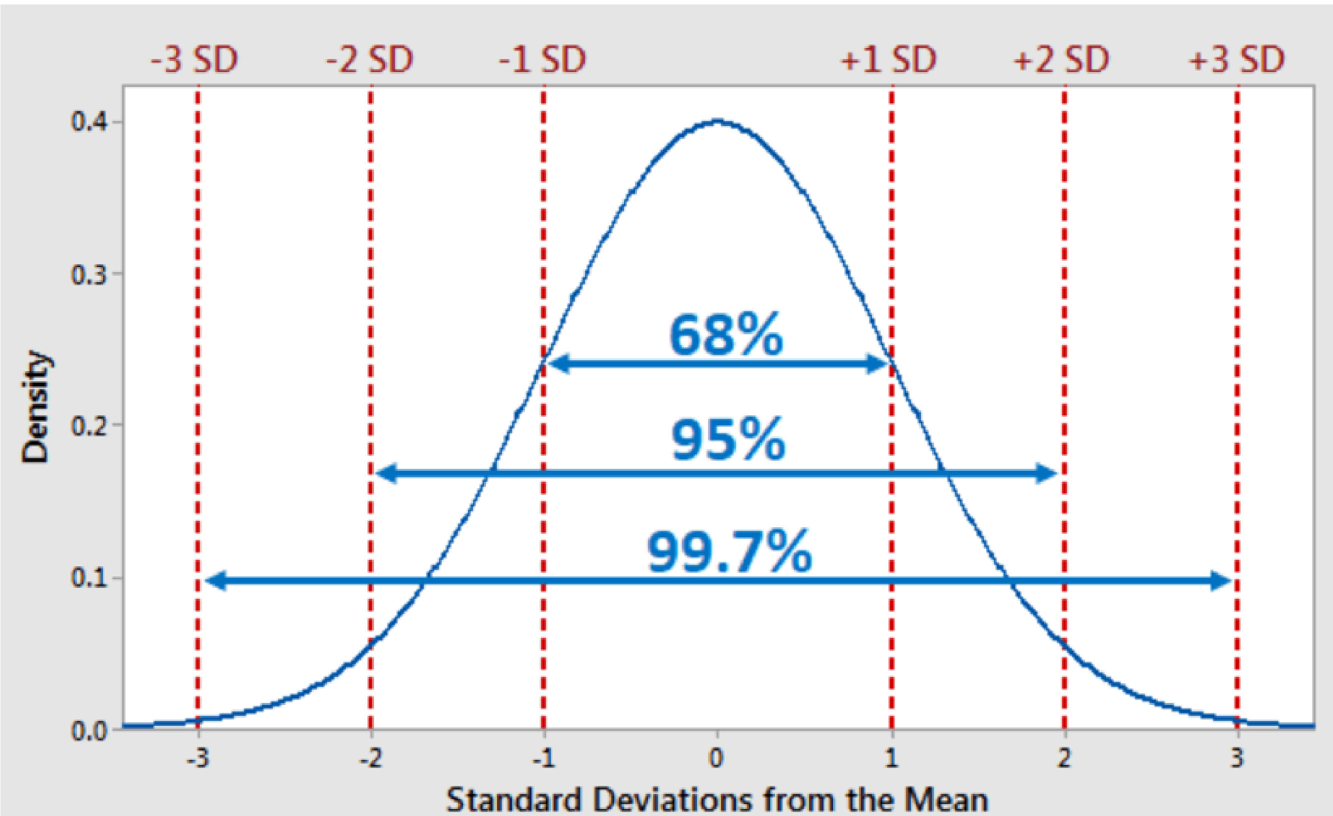
IQR: interquartile range is a measure of variation or dispersion in a set of data.

An outlier is a data point that goes far outside the average value of a group of statistics. A commonly used rule says that a data point is an outlier if it is more than 1.5 IQR above the third quartile or below the first quartile. Said differently, low outliers are below $Q_1 - 1.5$(IQR) and high outliers are above $Q_3 + 1.5(IQR)$



Z score, or number of standard deviations from the mean:

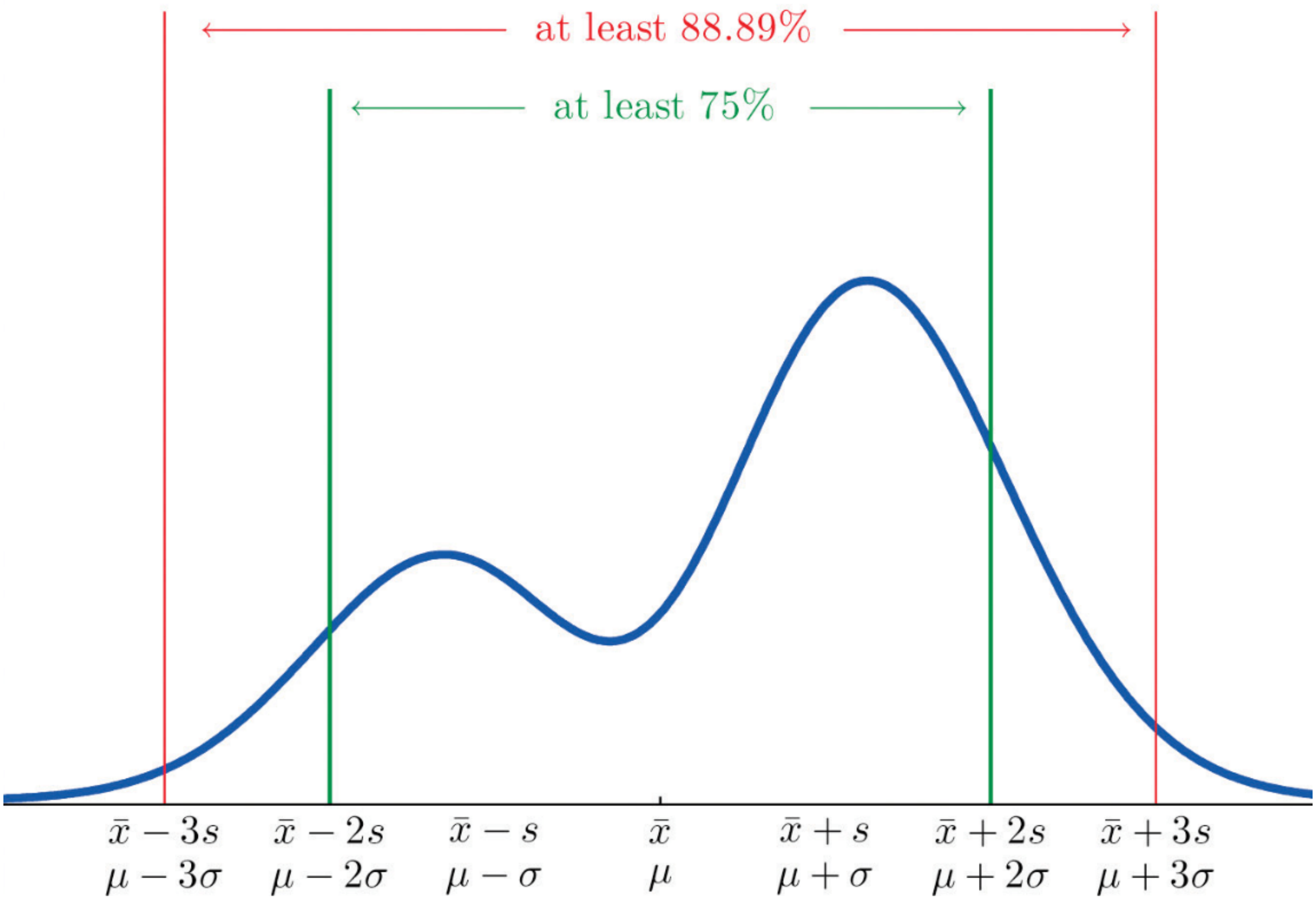
$$z = \frac{x - \mu}{\sigma}$$



Graph credit: https://vitalflux.com/chebyshevs-theorem-concepts-formula-examples/

The Empirical Rule is a rule telling us about where an observation lies in a normal distribution. The Empirical Rule states that approximately 68% of data will be within one standard deviation of the mean, about 95% will be within two standard deviations of the mean, and about 99.7% will be within three standard deviations of the mean. [Credit: https://online.stat.psu.edu/stat800/lesson/2/2.2]

What is the distribution of the date is not symmetrical?



Graph credit: https://vitalflux.com/chebyshevs-theorem-concepts-formula-examples/

Chebyshev's Theorem, also known as Chebyshev's Rule, states that in any probability distribution, the proportion of outcomes that lie within k standard deviations from the mean is at least $1 - 1/k^2$, for any k greater than 1; therefore, the maximum proportion of data values more than k standard deviations from the mean is $\frac{1}{k^2}$

Since at least $\mathbf{1} - \frac{\mathbf{1}}{\boldsymbol{k}^2}$ of the data lie within k standard deviations of the mean ($\boldsymbol{k} \neq \mathbf{1}$):

If $\boldsymbol{k} = \mathbf{2}$, then $\mathbf{1} - \frac{\mathbf{1}}{\mathbf{2}^2} = \mathbf{0.75}$, or $\mathbf{75}\%$ of the data lie within 2 standard deviations of the mean.

If $\boldsymbol{k} = \mathbf{3}$, then $\mathbf{1} - \frac{\mathbf{1}}{\mathbf{3}^2} = \mathbf{0.8888}$…, or $\mathbf{88.89}\%$ of the data lie within 3 standard deviations of the mean.