The median is the median (or centre point), also called **Q2**, of the data (resulting from the fact that the data is ordered).

**Q1** is the first quartile of the data, i.e., to say 25% of the data lies between minimum and Q1.

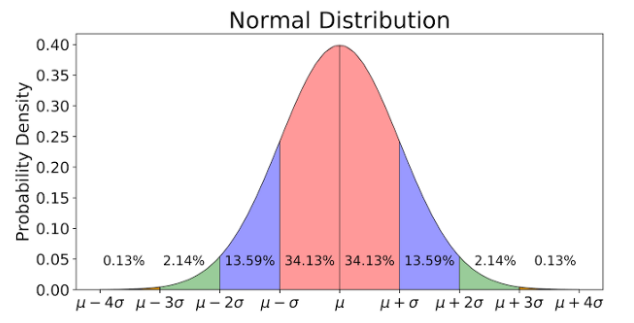
**Q3** is the third quartile of the data, i.e., to say 75% of the data lies between minimum and Q3.

The difference between Q3 and Q1 is called the Inter-Quartile Range or **IQR**.

Lower Bound: (Q1 - 1.5 \* IQR)

Upper Bound: (Q3 + 1.5 \* IQR)

Any data point less than the Lower Bound or more than the Upper Bound is considered as an outlier. This number, or scale, depends on the distribution followed by the data. For example, let’s say our data follows, Gaussian distribution. This is how the distribution looks like:



* About 68.26% of the whole data lies within one standard deviation (<σ) of the mean (μ), taking both sides into account, the pink region in the figure.
* About 95.44% of the whole data lies within two standard deviations (2σ) of the mean (μ), taking both sides into account, the pink+blue region in the figure.
* About 99.72% of the whole data lies within three standard deviations (<3σ) of the mean (μ), taking both sides into account, the pink+blue+green region in the figure.
* And the rest 0.28% of the whole data lies outside three standard deviations (>3σ) of the mean (μ), taking both sides into account, the little red region in the figure. And this part of the data is considered as outliers.
* The first and the third quartiles, Q1 and Q3, lies at -0.675σ and +0.675σ from the mean, respectively.
* This means that for normally distributed data, one-half of the data is within 2/3 of a standard deviation unit of the mean.

One definition of outliers is data that are more than 1.5 times the inter-quartile range before Q1 or after Q3. Since the quartiles for the standard normal distribution are +/-.67, the IQR = 1.34, hence 1.5 times 1.34 = 2.01, and outliers are less than -2.68 or greater than 2.68. Hence for normally distributed data, the probability of being an outlier is 2 times .0037 = .0074. This is less than 1%.

**References**

Grubbs, F. E. (February 1969). "Procedures for detecting outlying observations in samples".

Maddala, G. S. (1992). "Outliers". Introduction to Econometrics (2nd ed.). New York: MacMillan. pp. 89. ISBN 978-0-02-374545-4.

Wikipedia