## **IOR 1.5 usage rule**

The 1.5 IQR is a rule of thumb, developed by statistician John Tukey for identifying potential outliers in a dataset. The number 1.5 was chosen as a practical compromise to balance sensitivity and stringency in outlier detection

**How the 1.5 IQR rule works**

The interquartile range (IQR) represents the middle 50% of your data. The 1.5 IQR rule sets up a "Outliers" beyond this range to identify data points that are significantly different from the rest.

**Find the quartiles-**

**** Q1 is the first quartile (25th percentile)

Q3 is the third quartile (75th percentile)

**Calculate the IQR-**

IQR=Q3-Q1

**Determine the Outliers-**

Lesser Outlier-Q1-1.5\*IQR

Greater Outlier-Q1-1.5\*IQR

**Identify outliers:** Any data point outside the lower and upper fences is considered an outlier

**The reasoning behind 1.5**

The number 1.5 was not derived from a formal proof but was selected by Tukey as a workable constant. When asked why he chose 1.5, he reportedly replied that 1 was too small, and 2 was too large. This provides a useful, non-parametric method that works well in many scenarios.

**Comparison with normal distribution**

In a normal distribution, the 1.5 IQR rule sets a cutoff roughly equivalent to identifying outliers that are about 2.7 standard deviations from the mean

* A multiplier of 1 would be too sensitive, flagging about 5% of the data as outliers.
* A multiplier of 2 would be too inclusive, potentially missing many anomalies
* A multiplier of 1.5 strikes a balance by flagging about 0.7% of the data, a reasonable cutoff that catches suspicious points without being overly sensitive