

The Interquartile Range (IQR) is a powerful statistical tool in data science used to understand the spread and variability of a dataset.

What is IQR?

- IQR is the range between the first quartile (Q1) and the third quartile (Q3).
- Formula: $IQR = Q3 - Q1$
- It captures the middle 50% of the data, excluding the extremes.

Greater Outlier Formula

A data point is considered a **greater outlier** if:

$$\text{Value} > Q3 + 1.5 \times IQR \quad \text{Value} > Q3 + 1.5 \times \text{IQR}$$

Where:

- **Q3** = Third quartile (75th percentile)
- **IQR** = Interquartile Range = $Q3 - Q1$

Explanation

- The IQR measures the spread of the middle 50% of the data.
- By adding $1.5 \times IQR$ to the upper quartile (Q3), we set a threshold beyond which values are unusually high.
- Any value **above this threshold** is flagged as a **greater outlier**.

A data point is considered a **lesser outlier** if:

$$\text{Value} < Q1 - 1.5 \times IQR \quad \text{Value} < Q1 - 1.5 \times \text{IQR}$$

Where:

- **Q1** = First quartile (25th percentile)
- **IQR** = Interquartile Range = $Q3 - Q1$

Explanation

- The IQR measures the spread of the middle 50% of the data.
- By subtracting $1.5 \times IQR$ from the lower quartile (Q1), we set a threshold below which values are unusually low.
- Any value **below this threshold** is flagged as a **lesser outlier**.