What is the 1.5 × IQR Rule?

This is used in the **outlier detection method** for identifying unusual data points in a dataset.

Step-by-Step Reasoning:

1. Understanding the IQR

- The Interquartile Range (IQR) measures the spread of the middle 50% of your data
- IQR = Q3 - Q1 (third quartile minus first quartile)
- It represents where most of your "typical" data lives

2. Why Multiply by 1.5?

The 1.5 multiplier creates **fences** to identify outliers:

- **Lower fence** = Q1 - 1.5 × IQR
- **Upper fence** = Q3 + 1.5 × IQR

Any data point beyond these fences is considered an outlier.

3. Why Specifically 1.5?

This value is somewhat arbitrary but carefully chosen because:

- **It's balanced**: Not too strict (which would flag too many normal points as outliers) and not too loose (which would miss genuine outliers)
- **Statistical reasoning**: For normally distributed data, 1.5 × IQR captures about 99.3% of the data, meaning only about 0.7% would be flagged as outliers
- **Empirical testing**: John Tukey (who developed this rule) found through experience that 1.5 works well across many real-world datasets
- **Practical effectiveness**: It identifies points that are genuinely unusual without being overly sensitive to natural variation

4. The Practical Purpose

This rule helps you:

- Automatically identify data points that deserve closer examination
- Detect data entry errors or measurement problems
- Find genuinely unusual cases in your dataset
- Create box plots (where the whiskers typically extend to 1.5 × IQR)

5. Alternative Multipliers

- **3 × IQR**: Sometimes used for "extreme" outliers (even more unusual)
- **2 × IQR**: A more conservative approach that flags fewer outliers

The 1.5 multiplier has become the standard because it strikes the right balance between sensitivity and specificity in outlier detection across diverse applications.