1.5 RULE IN IQR

Name: S. Rayen Subhiksha

1. Why are we using 1.5 times of IQR?

- The interquartile (IQR) method of outlier detection uses 1.5 as its scale to detect outliers because it most closely follows Gaussian distribution. As a result, the method dictates that any data point that's 1.5 points below the lower bound quartile or above the upper bound quartile is an outlier.
- In practical applications, using 1.5 times the IQR to define outliers helps in detecting values that are unusually high or low compared to the rest of the data. This can be particularly useful in various fields such as finance, quality control, and research, where identifying outliers can highlight important anomalies or errors.

- The choice of 1.5 strikes a balance between detecting outliers and not being overly sensitive to minor deviations in the data. Using a smaller factor would result in many data points being falsely identified as outliers, while a larger factor might miss significant outliers.
- Over time, the use of 1.5 as a multiplier has become a convention in statistics due to its effectiveness and reliability in various applications. This convention is widely taught

and used, further reinforcing its acceptance and usage.

- The 1.5 factor is simple and practical, making it easy to apply and understand across different datasets. It provides a standardized way to identify outliers, which is useful in various statistical analyses and data cleaning processes.
- The factor of 1.5 is derived from empirical observations and statistical theory. In a normally distributed dataset, this factor ensures that approximately 99.3% of data points lie within the bounds of Q1 - 1.5 * IQR and Q3 + 1.5 * IQR, assuming no outliers. This is based on the properties of the normal distribution and the spread of data.