IQR Report

(Outlier,Replacing)

Reason of multiplying 1.5 with IQR

- Multiplying 1.5 with the interquartile range (IQR) is a common techmique used in statistics to identify outliers using the concept of the “Tukey method” or “Tukey’s fences”.
- The interquartile range is a measure of statistical dipersion, represending the range between the first quartile (25^{th} Percentile) and the third quartile (75^{th} Percentile) of the dataset.

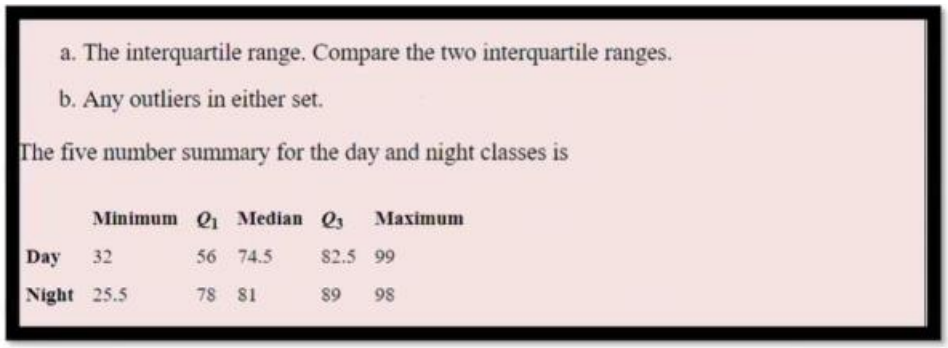
Tukey’s Fences : Tukey’s fences are thresholds used to define the bounds beyond which data points are considered outliers. These fences are calculated as follows:

```
Lower Bound: (Q1 - 1.5 * IQR)
Upper Bound: (Q3 + 1.5 * IQR)
```

- The choice of as the multiplier is somewhat arbitrary but is widely accepted as a balance between detecting genuine outliers and not excessively labelling data points as outliers.
- This value provides a reasonable compromise between sensitivity to potential outliers and the risk of falsely identifying normal data points as outliers.
- John Tukey, a prominent statistician, introduced this method as part of exploratory data analysis. While

other multipliers can be used, 1.5 has become a standard in many statistical analyses due to its effectiveness in identifying potential outliers without overly inflating the outlier count.

Example:



a. The interquartile range. Compare the two interquartile ranges.

b. Any outliers in either set.

The five number summary for the day and night classes is

	Minimum	Q_1	Median	Q_3	Maximum
Day	32	56	74.5	82.5	99
Night	25.5	78	81	89	98

Solution:

IQR = Q3-Q1 = 82.5 – 56 = 26.5

IQR = 26.5

Lessser Range Outlier:

Q1-1.5(IQR) = 16.25

Greater Range of Outlier:

Q3 + 1.5(IQR) = 82.5 + 1.5(26.5)

= 82.5 + 39.75

= 122.75

Q3 + 1.5(IQR) = 122.75

Conclusion:

Day and Night classes lower than 16.25 are lesser range outliers and classes higher than 122.25 are higher range outliers.