

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

For Day:

$$\text{IQR} = Q_3 - Q_1 = 82.5 - 56 = 26.5$$

$$1.5 * \text{IQR} = 39.75$$

$$\text{Lesser range of outlier} = Q_1 - (1.5 * \text{IQR}) = 56 - 39.75 = 16.25$$

$$\text{Greater range of outlier} = Q_3 + (1.5 * \text{IQR}) = 82.5 + 39.75 = 122.25$$

For Night:

$$\text{IQR} = Q_3 - Q_1 = 89 - 78 = 11$$

$$1.5 * \text{IQR} = 16.5$$

$$\text{Lesser range of outlier} = Q_1 - (1.5 * \text{IQR}) = 78 - 16.5 = 61.5$$

$$\text{Greater range of outlier} = Q_3 + (1.5 * \text{IQR}) = 89 + 16.5 = 105.5$$

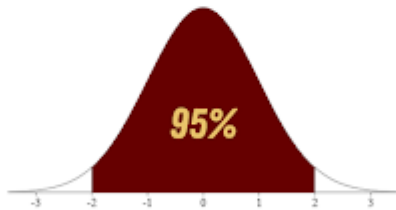
Conclusion: For Night data, Lesser range outliers present since the lesser range outlier value is 61.5 which is greater than minimum value 25.5

Why multiplying by 1.5 ?

As we know most the data follows the Gaussian distribution (Normal distribution – bell curve).

- About 68.26 percent of the data set lies within one standard deviation ($<\sigma$) of the mean (μ).
- About 95.44 percent of the data set lies within two standard deviations (2σ) of the mean (μ), taking both sides into account.
- About 99.72 percent of the data set lies within three standard deviations ($<3\sigma$) of the mean (μ), taking both sides into account.
- The first and the third quartiles, Q_1 and Q_3 , lies at -0.675σ and $+0.675\sigma$ from the mean, respectively.

THE 68, 95, 99 RULE



- If we take scale value 1 then Lower Bound and Higher bound values are -2.025σ to 2.025σ . which is lesser than 3σ . Too many outliers.
- If we take the scale value as 2 then bound values are -3.375σ to 3.375σ which is more than 3σ . Very few outliers.
- If we take the scale value as 1.5 then bound values are -2.7σ to 2.7σ which is almost equal to 3σ . This decision range is the closest to what Gaussian Distribution.

So we are multiplying the IQR with the scale 1.5