

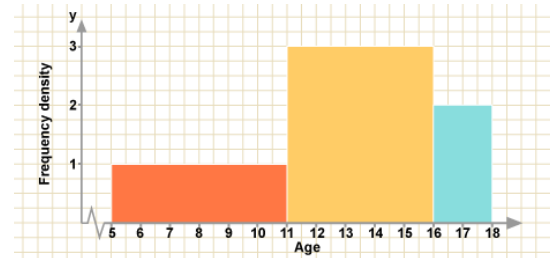
17 - Sampling and More Complex diagrams

Frequency Density = Frequency \div Class Width

Frequency = Frequency Density \times Class width

Histograms

1. There is continuous data (no gaps between classes)
2. When each class width/interval is not equal
3. The area ($l \times b$) represents the frequency, not height



Box plots

When data is ordered:

$$LQ = \frac{1}{4}(n+1)$$

$$\text{Median} = \frac{1}{2}(n+1)$$

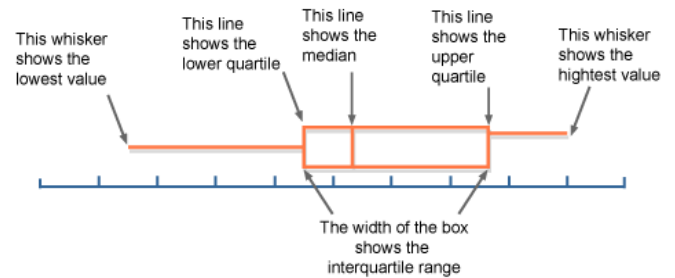
$$UQ = \frac{3}{4}(n+1)$$

Interquartile range = $UQ - LQ$

Positive skew = when median is closer to LQ than UQ

Negative skew = when median is closer to UQ than LQ

Symmetrical distribution = even spread from LQ to UQ



Outlier = does not fit the trend, or more than 1.5x the interquartile range below LQ/above UQ