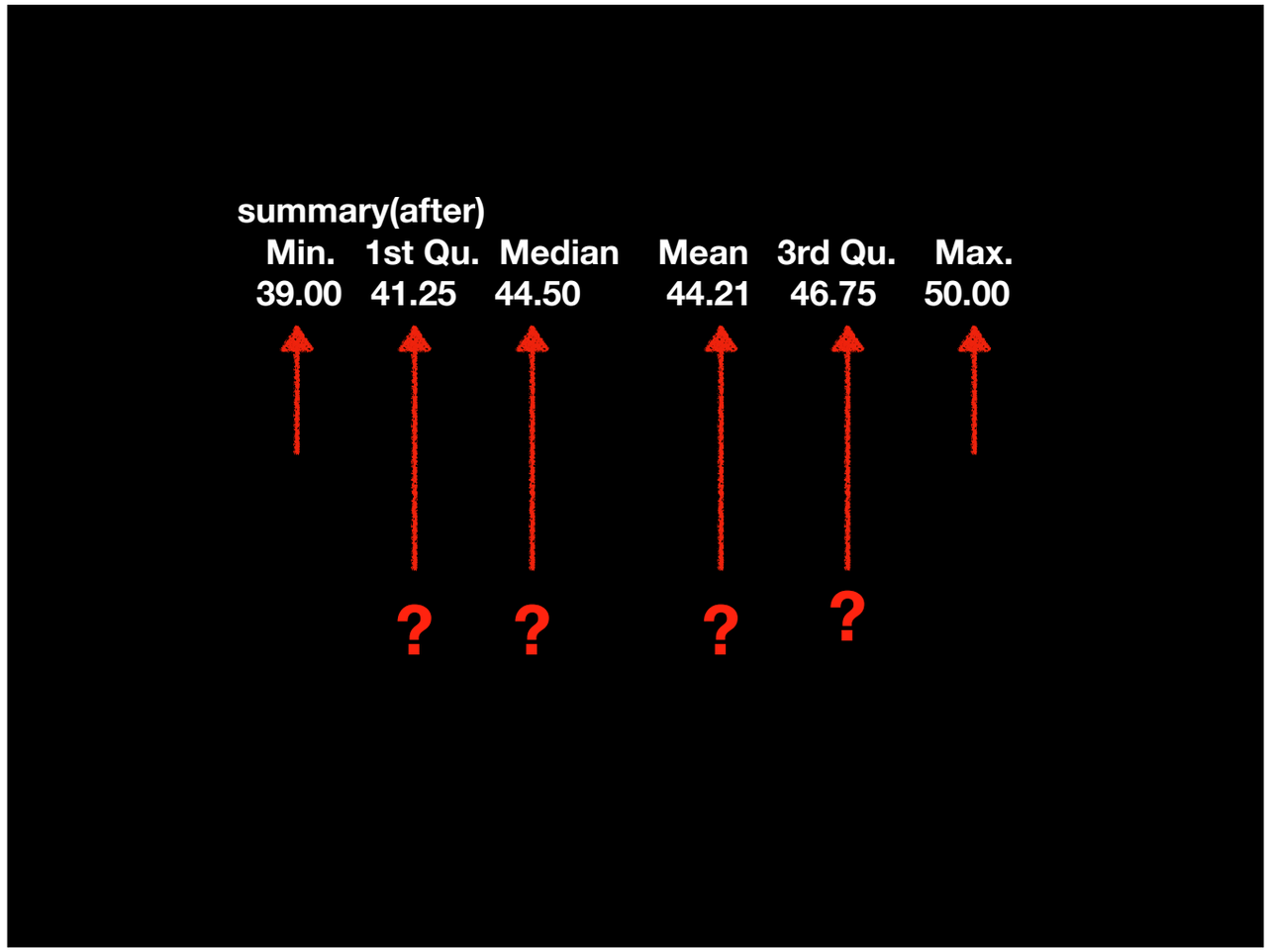

Some of these may look familar but some might be new -- the mean, median and quartiles might be new.

Summary Statistic Definitions!

Mean (Sample) = sum of all data values divided by number of data points

$$\text{Mean} = \frac{\text{Sum of all values}}{\text{Number of values}}$$

$$\text{Symbolically, } \bar{x} = \frac{\sum x}{n}$$

where $\bar{x}$ (read as 'x bar') is the mean of the set of x values,
$\sum x$ is the sum of all the x values, and
n is the number of x values.

(note - only works with “numerical” data types… more about data types later)

Median = if we order the data from smallest to largest, this is the observation in the middle (splits the data in 2 halves)

First/Third Quartiles = where 25% of the data falls below/above

Standard Deviation = this is the square root of the variance, where the variance is roughly the average distance of data values from the mean

$$\text{Standard Deviation (sample)} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \text{mean})^2}{n-1}}$$

Mean and median give a sense of the “typical value” in your dataset.

Summary Statistic Definitions!

Mean (Sample) = sum of all data values divided by number of data points

$$\text{Mean} = \frac{\text{Sum of all values}}{\text{Number of values}}$$

$$\text{Symbolically, } \bar{x} = \frac{\sum x}{n}$$

where $\bar{x}$ (read as 'x bar') is the mean of the set of x values,
$\sum x$ is the sum of all the x values, and
n is the number of x values.

(note - only works with "numerical" data types… more about data types later)

Median = if we order the data from smallest to largest, this is the observation in the middle (splits the data in 2 halves)

First/Third Quartiles = where 25% of the data falls below/above

Standard Deviation = this is the square root of the variance, where the variance is roughly the average distance of data values from the mean

$$\text{Standard Deviation (sample)} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \text{mean})^2}{n-1}}$$

Standard deviation and quartiles give measures of the variability in the data.