DESCRIPTIVE STATISTICS

Descriptive Statistics

A summary: The Three Key Statistics to Measure of...

Central Tendency

- The extent to which the data values are grouped or clustered around a central value
- Mean, Median, Mode

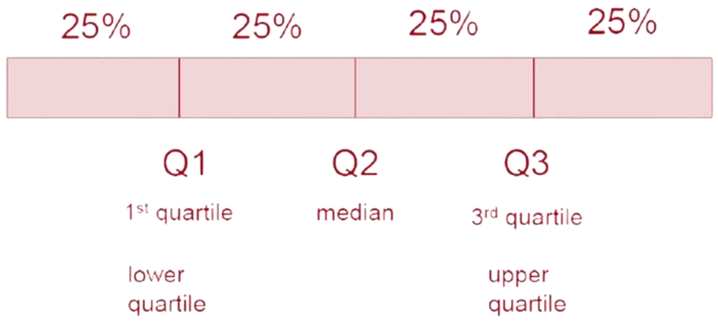
Variation

- The spread, scattering or dispersion of data values, i.e. how the spread out the data is.
- Range, Interquartile-range (IQR), Standard deviation, Variance, Coefficient of variation.

Shape

- The pattern of the distribution of data values from the lowest value to the highest value.
- Skewness, Kurtosis

Quartiles



Measures of Central Tendency / Location

Measures of centre identify the central position of the data set to represent a 'typical' value in the data set.

Formulas and Examples

Mean

The mean gives a 'typical' or central value for a data set.

Mean: average of a set of values

n = total number of data points

Example:

Average Days Late**:**

$\frac{1+2+3+4+5}{5} = 3.0$

$$Mean = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n}$$

Median

The median is the middle value in a set of data that has been ordered from lowest to highest value. 50% of the values are equal to or smaller than the median and 50% of the values are equal to or larger than the median.

Example:

n is odd: 1, 3, 7, 8, 10, 13, 17 — The median is "8"

n is even: 1, 3, 7, 10, 11, 16, 17, 20 — The median is "10.5"

Organise the data from low values to high when determining the median.

Mode

The mode is the value in a data set that appears most frequently. A set of data will have no mode if none of the values is 'most typical' – that is if no data value occurs more than once.

Corresponds to the highest bar in the histogram.

Data set:

45	47	49	51
46	47	49	52
47	47	50	53
47	48	51	54

Mode is: 47