

Y10 STATISTICS UNIT 4: SUMMARISING DATA

Key words and definitions:

- **Interquartile range** – the difference between Q3 and Q1. It describes the range of half of the data.
- **Percentile** - a value that has a certain percentage of scores below it. For example, if the 70th percentile on an exam was 156 and you scored 156 on the exam, your score was better than 70 percent of test takers.
- **Standard deviation** - a measure of how dispersed the data is in relation to the mean
- **Outlier** – a data point that differs significantly from other observations
- **Skew** - a measure of the asymmetry of the probability distribution of a real-valued random variable about its mean.

Weighted mean

A **weighted mean** is used when different numbers or values have different **weights** i.e. some of the values have more statistical 'importance' than others

$$\text{weighted mean} = \frac{\sum (\text{value} \times \text{weight})}{\sum \text{weights}}$$

Example:

50% of the workers in a company earn £460 per week. 35% of the workers earn £600 per week, and the other 15% earn £820 per week. Work out the mean weekly earnings for workers in the company.

$$\frac{(460 \times 50) + (600 \times 35) + (820 \times 15)}{50 + 35 + 15}$$
$$= \frac{56300}{100} = 563$$

Geometric mean

The geometric mean is found by:

- Multiplying the values together
- Taking the nth root (e.g. square root for 2 numbers, cube root for 3 numbers, etc)

Example:

The geometric mean of 2, 9 and 12 is...

$$\sqrt[3]{2 \times 9 \times 12} = \sqrt[3]{216} = 6$$

The geometric mean gives a better 'average' for numbers that are going to be **multiplied**. For example, in questions involving **percentage changes**:

The geometric mean of the percentage change **multipliers** should be found, not of the percentages on their own e.g. use 1.02 for an increase of 2% or 0.98 for a decrease of 2%

Standard deviation

These are on the formula sheet but you need to know **HOW** to use them!

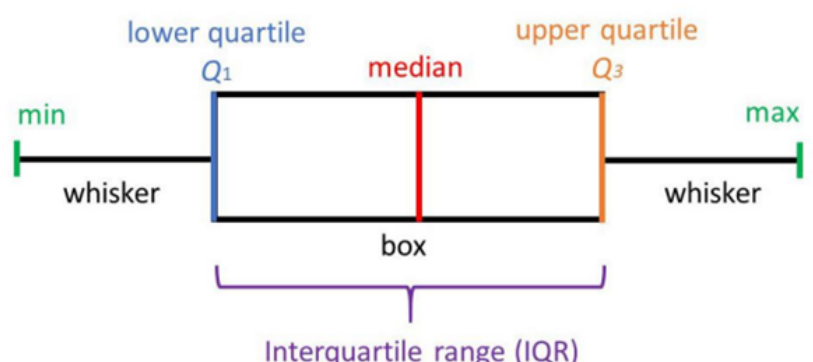
$$\sigma = \sqrt{\frac{1}{n} \sum (x - \bar{x})^2}$$

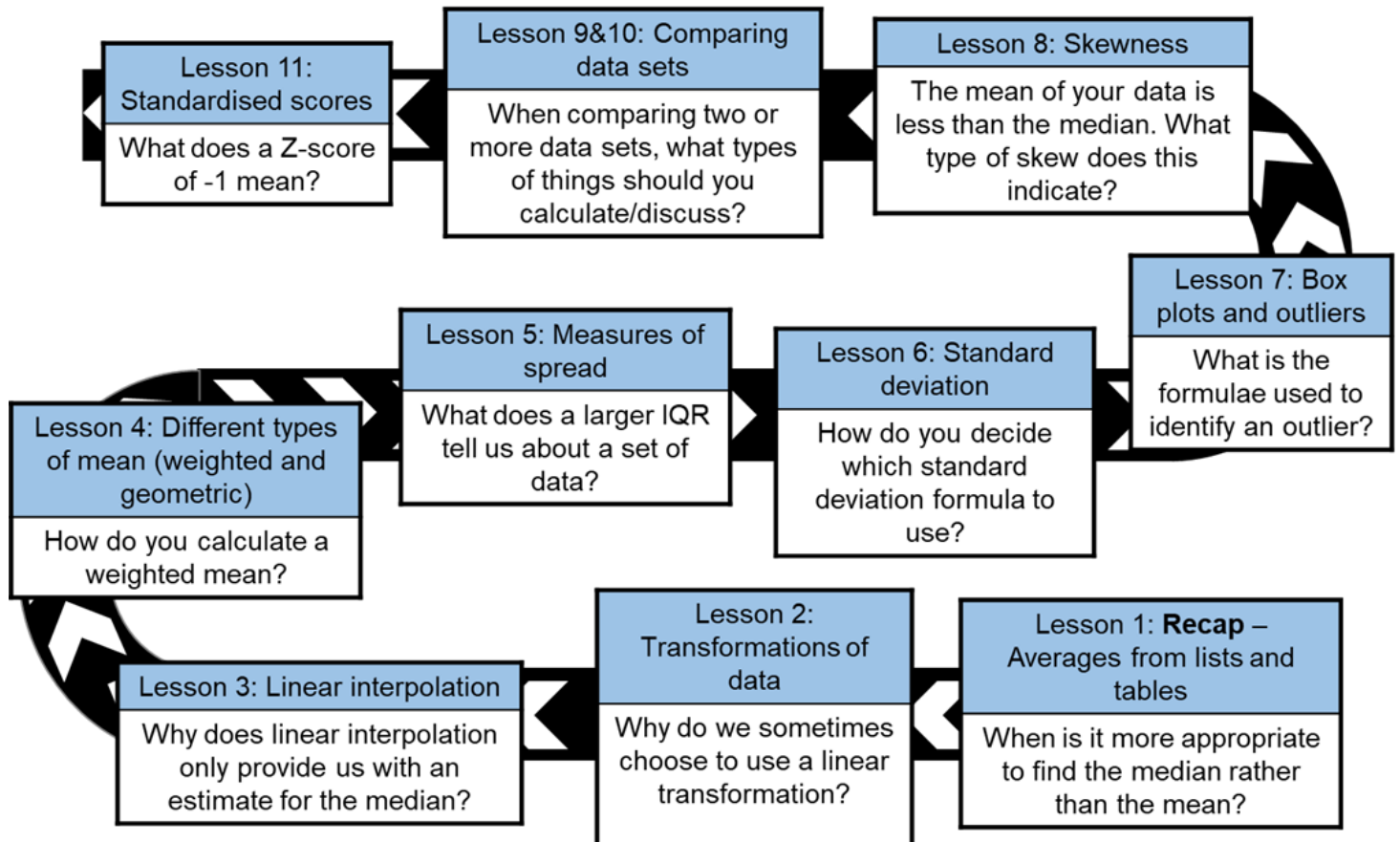
$$\sigma = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

Box plots

Box plots show the distribution of a set of data.

It is important to talk percentages when referring to a box plot, knowing that Q1 is 25% of the data as Q1 represents 1 quarter, Q2 (median) represents half of the data, and so on.





DEFINITION ETYMOLOGY/MORPHOLOGY	CHARACTERISTICS
EXAMPLES 	NON-EXAMPLES