1. Mode

The **mode** of a set of values is the value with the highest frequency. There may be several **modal** values.

Examples :

A set of values is composed of 55 ; 56 ; 57 ; 56 ; 55 ; 56 ; 57
In this example, the frequency of 56 is 3, that of 55 and 57 is 2, hence the value 56 for the mode.

A frequency distribution is shown in the table below :

values	11	13	17	19	23
frequency	1	33	1	33	7

Both values 13 and 19 have a frequency 33, which is the highest frequency : hence two modal values that are 13 and 19.

2. Quartiles

Quartiles, as their name suggests, are the quarter-way divisions of the data.

The **lower quartile**, Q_1 , is the value one-quarter of the way through the distribution.

The value half-way through the distribution is not known as a quartile but as the _____ , often referred to as Q_2 .

The **upper quartile**, Q_3 , is the value three-quarters of the way through the distribution.

Therefore, half the values of the distribution must lie between the upper and lower quartiles.

The **interquartile range** is a single number and gives a measure of the **spread** of the values **about** the median : its value is $Q_3 - Q_1$

Exercise 1

A group of people form a club and over a period of time contribute to a fund to purchase equipment. The records showed the contributions as folows :

Contribution, €	0 - 10	10 - 20	20 - 30	30 - 40	40 - 60
Number of customers	5	10	25	40	20

Note : 0 – 10 means 0 is included but 10 is not, etc...

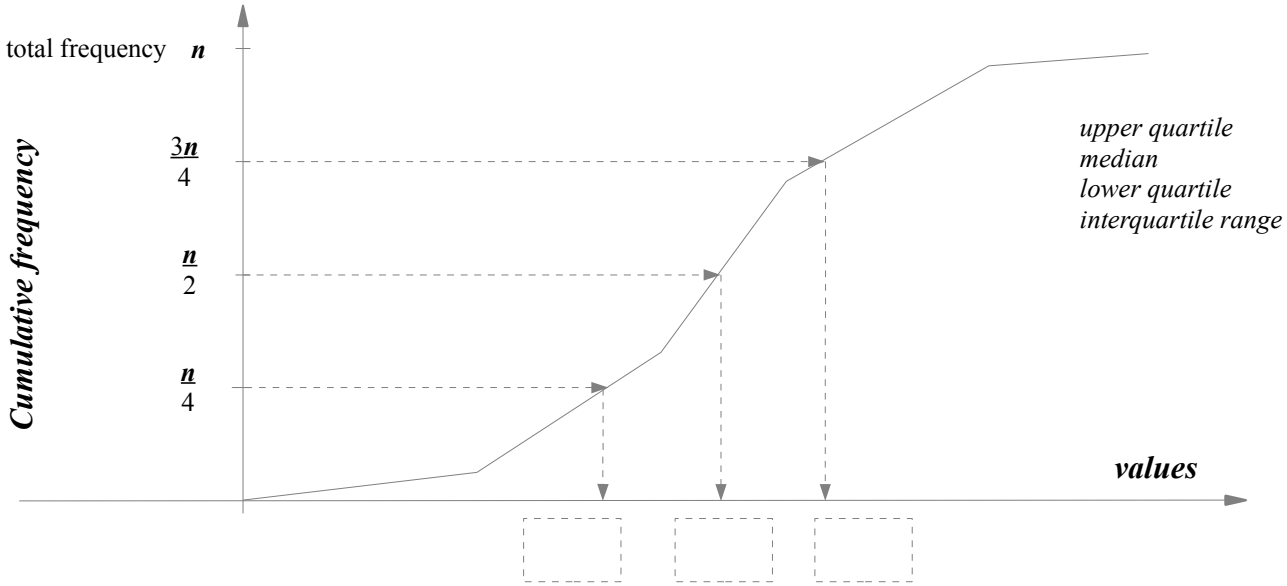
Taking the contribution at the mid-interval values, calculate the mean contribution, and the interquartile range.

Vocabulary interquartile range – modal – mode – (lower / upper) quartile – spread ... about ...

3. Cumulative frequency

In a **cumulative frequency** diagram the frequencies are accumulated. Each accumulated frequency is the combined total of all the previous frequencies up to that particular value. If we fill in the accumulated frequencies in tabular form, we have what is called a **cumulative frequency table**. The graph of a cumulative frequency is called a **cumulative frequency curve** or **ogive**. It has a distinctively lopsided S-shape.

The basic model is a linear function in each class interval : this figures that values are regularly spread within their class, therefore the points of the graph (restricted to one interval) are aligned. Remember though that a line segment is drawn, and not all individual points for values.



Note :

If drawing a cumulative frequency curve is required, use the basic model.
If an ogive is already drawn and is not composed of line segments, just use it to read out quartiles.

Exercise 2

A shopkeeper recorded the amount of money spent by 50 customers on a certain day in the following cumulative frequency table :

Amount spent in €	< 10	< 20	< 30	< 40	< 50	< 60	< 80
Number of customers	2	5	9	17	29	43	50

i. *Draw a cumulative frequency curve to illustrate the data*

Use your cumulative frequency curve to estimate :

ii. *the number of customers who spent less than € 36*

iii. *the percentage of customers who spent between € 47 and € 65*

iv. *complete the corresponding grouped frequency table :*

Amount spent in €	[0 ; 10[	[10 ; 20[	[20 ; 30[	[30 ; 40[	[40 ; 50[	[50 ; 60[	[60 ; 80]
Number of customers							

Vocabulary cumulative frequency – curve – ogive

4. Standard deviation

The **standard deviation** σ is a measure of the **spread** of the values **about** the mean.
In other words, it gives us an indication of how the values we have are spread out.
The higher the standard deviation, the more spread out around the mean is the data, and vice versa.

4.1. Set of values

Let $(x_i)_{1 \le i \le n}$ be a set of values, with total frequency n.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{i=n} d_i^2}{n}} \qquad \text{where } d_i = x_i - \bar{x} \text{ is the deviation}$$

example : Let 1; 3; 6; 7; 8 be an array of numbers. Calculate the standard deviation σ of that array.

first calculate the mean $\bar{x} = \frac{1+3+6+7+8}{5} = 5$

second make out a table

	x_i	d_i	d_i^2
	1	4	16
	3	2	4
	6	1	1
	7	2	4
	8	3	9
sum	**25**		**34**

last calculate the standard deviation $\sigma = \sqrt{\frac{34}{5}} \approx 2.61$

Exercise 3

The grouped frequency table below refers to the ages of 28 people at a party :

Age (in years)	0 - 4	4 - 8	8 - 12	12 - 16	16 - 20
Number of people	2	6	12	6	2

Note : 0 – 4 means greater than 0 but less than 4, etc...

Assuming the data can be taken at the mid-interval values, calculate :

i. *the mean* $\bar{x}$

ii. *the standard deviation* σ

iii. *evaluate the expression* $\sqrt{\bar{x} + 4\sigma - 1}$

iv. *complete the corresponding cumulative frequency table :*

Age (in years)	< 4	< 8	< 12	< 16	< 20
Number of people					

and draw the cumulative frequency curve

v. *Use the curve to estimate the percentage of people in the range* $\bar{x} - \sigma$ *to* $\bar{x} + \sigma$

Vocabulary the spread of ... about ... – standard deviation

4.2. Frequency distribution

Let the following set of couples describe a frequency distribution : $(x_i ; f_i)_{1 \le i \le n}$

where x_i are the values, f_i are the frequencies and n is the total number of values.

To calculate the standard deviation of the frequency distribution, we use the following formula :

$$\sigma = \sqrt{\frac{\sum_{i=1}^{i=n} f_i \, d_i^2}{\sum_{i=1}^{i=n} f_i}} \qquad \text{where } d_i = x_i - \bar{x} \text{ is the deviation}$$

example : Calculate the mean and standard deviation σ of the following frequency distribution : $1 \le i \le 6$

Value x_i	2	4	6	8	10	12
Frequency f_i	7	8	4	4	3	4

first calculate the mean $\bar{x} = \dfrac{7\times2 + 8\times4 + 4\times6 + 4\times8 + 3\times10 + 4\times12}{7+8+4+4+3+4} = \dfrac{180}{30} = 6$

second make out a table

	f_i	x_i	d_i	d_i^2	$f_i d_i^2$
	7	2	4	16	112
	8	4	2	4	
	4	6	0		
	4	8			
	3				
sum					

last calculate the standard deviation σ =

4.3. Grouped frequency distribution

For grouped frequency distribution, the standard deviation is calculated in exactly the same way, except that x_i stands for the mid-interval value.

Exercise 4

Find the mean and standard deviation, correct to 2 d.p., of the following grouped frequency distribution :

value	[0 ; 20[	[20 ; 40[	[40 ; 60[	[60 ; 80]
frequency	11	14	9	6