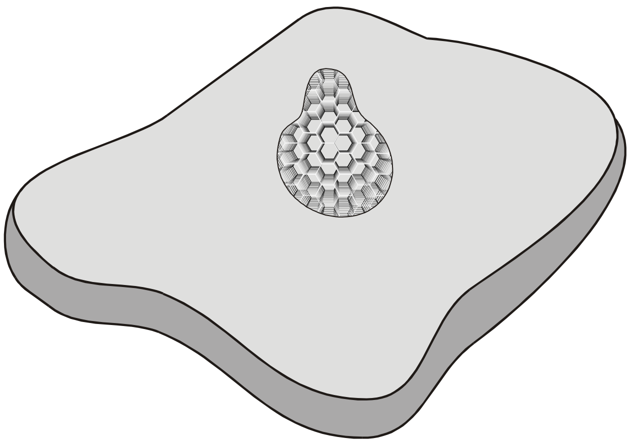
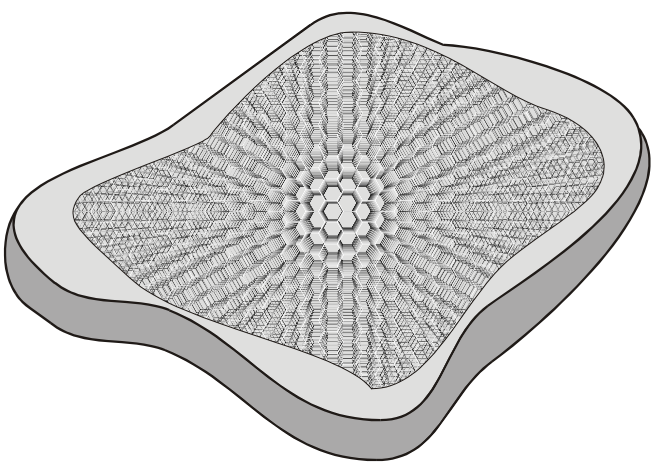
Standard Deviation

Spread

Imagine a slice of bread with a dollop of honeycomb in the centre

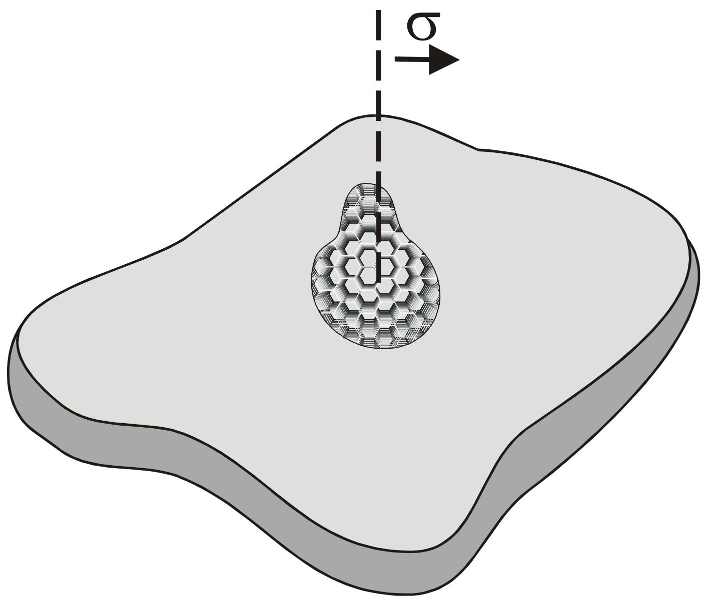


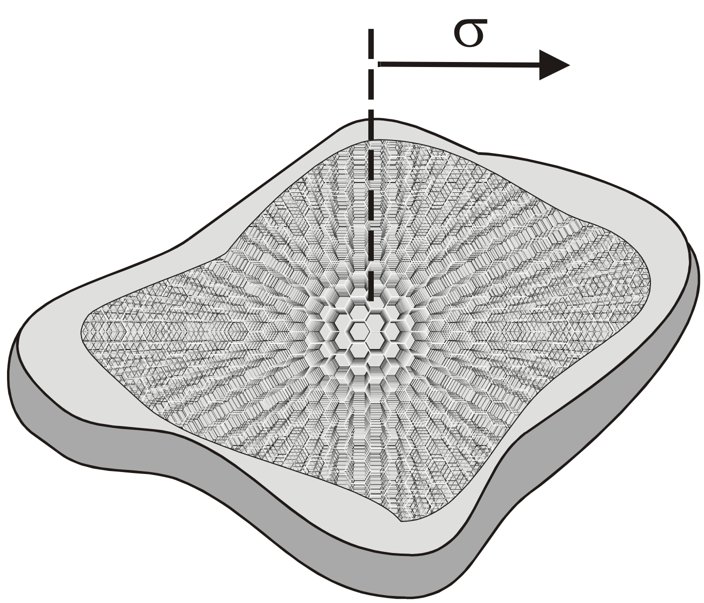
The same quantity of honeycomb could be either concentrated in the middle of the slice of bread, or it could be more evenly spread out, like this



Statisticians also need a measure of spread. They use a number, called the standard deviation, to indicate how spread out data is.

The larger the standard deviation, the more spread out the data is.





Standard deviation

The data we obtain comes in the form of some set of numbers. We need to be able to treat these numbers in some way so as to obtain a measure of their spread – their standard deviation. This is achieved by means of the following formulae.

For a ungrouped simple set of data (that is, without a frequency table), the mean, or measure of central tendency is

$$\bar{x} = \frac{\sum x}{n}$$

The symbol $\sum$ means "add up". The mean is found by adding up all the values and dividing by the total number of values. The standard deviation, or measure of dispersion or spread is

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

Example (1)

Calculate the mean and standard deviation of the set of numbers

2 3 5 7 8 9 10

Solution

$$\bar{x} = \frac{\sum x}{n} = \frac{2+3+5+7+8+9+10}{7} = \frac{44}{7} = 6.2857...$$

$$\frac{\sum x^2}{n} = \frac{2^2+3^2+5^2+7^2+8^2+9^2+10^2}{7}$$

$$= \frac{4+9+25+49+64+81+100}{7} = \frac{332}{7} = 47.4285...$$

$$\sigma = \sqrt{\frac{\sum x^2}{n} - (\bar{x})^2} = \sqrt{47.4285... - (6.2857...)^2} = 2.81 \ (3.S.F.)$$

For grouped frequency set of data, the mean, or measure of central tendency is

$$\bar{x} = \frac{\sum xf}{n}$$

The standard deviation, or measure of dispersion or spread is

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

Example (2)

Calculate the mean and standard deviation of the following grouped frequency distribution.

Width / m	Frequency
10 – 15	3
15 – 20	7
20 – 25	13
25 – 30	27
31 – 35	18
36 – 40	4

Solution

Since this is a grouped frequency table we must take the mid-value of each interval. Also we have to calculate xf and $x^2 f$ for each entry in the table.

Interval	x	f	xf	x^2	$x^2 f$
10 – 15	12.5	3	37.5	156.25	468.75
15 – 20	17.5	7	122.5	306.25	2143.75
20 – 25	22.5	13	292.5	506.25	6581.25
25 – 30	27.5	27	742.5	756.25	20418.75
30 – 35	32.5	18	585	1056.25	19012.5
35 – 40	37.5	4	150	1406.25	5625
		$\sum f = 72$	$\sum xf = 1930$		$\sum x^2 f = 54250$

$$\overline{x} = \frac{\sum xf}{n}$$

$$= \frac{1930}{72} = 26.8(3.S.F.)$$

$$\frac{\sum x^2 f}{n} = \frac{54250}{72} = 753.47222...$$

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

$$= \sqrt{753.47222 - (26.8055...)^2}$$

$$= 5.91(3.S.F.)$$