

# Foundation STATISTICS & PROBABILITY

## Averages

**mode/modal** – most common value or values (modal class).

**median** – the middle number when they are in ascending order.

**mean** – add the numbers up and divide by how many there are.

**range** – the difference between the largest and smallest value.

## Important Terms

**frequency** – the number of elements in a group.

**quantitative data** – information about numbers, e.g. ages or heights (quantities).

**qualitative data** – information about everything else, e.g. eye colour or favourite food.

**random sampling** – every piece of data has the same chance of being chosen.

## Sample Space

A fair coin is flipped and a fair dice is rolled. The sample space diagram below can be used to represent the outcomes.

	1	2	3	4	5	6
H	H, 1	H, 2	H, 3	H, 4	H, 5	H, 6
T	T, 1	T, 2	T, 3	T, 4	T, 5	T, 6

## Pie Charts

To calculate the angle needed, we divide  $360^\circ$  by the total frequency. This tells us the number of degrees needed for 1 person. We can then multiply this by the frequencies to find the angles.

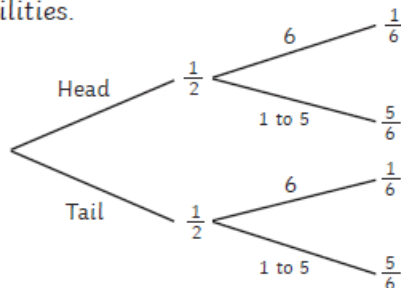
E.g. 10 people were asked their favourite colour.

Favourite Colour	Frequency	Degrees
Red	3	$3 \times 36 = 108^\circ$
Yellow	5	$5 \times 36 = 180^\circ$
Blue	2	$2 \times 36 = 72^\circ$

$$360 \div 10 = 36$$

## Tree Diagrams

A fair coin is flipped and a fair dice is rolled. The tree diagram below can be used to represent some outcomes and their probabilities.



## Probability

Probability is about estimating how likely something is to happen. We use fractions, decimals and percentages to describe probability. Only occasionally do we use words (for example, **likely**, **impossible**, **certain**) and we never use ratios!

Probability of an outcome =

$$\frac{\text{number of ways the outcome can happen}}{\text{total possible outcomes}}$$

The probability of rolling a 5 on a fair dice is  $\frac{1}{6}$

## Scatter Graphs

Easy to spot as the coordinates are scattered. Always draw a **straight line of best fit** (which follows the trend of the data) when you see this type of graph. The line of best fit can be used to make estimates.

These can have **positive correlation** when the line slopes upwards or **negative correlation** when the line slopes downwards.

If you cannot draw a line of best fit, there is **no correlation**.

## Mean from a Frequency Table

where  $f$  is the frequency and  $x$  is the data (e.g. time, number of pets). **Remember**, with continuous data you need to find the midpoint first.