Part 3: Two Way Tables

Often in statistics we look at data where we know information about two (or more variables). We often represent this using a two way table like this:

	Number of Students		
Eye Colour	**Male**	**Female**	**Total**
Brown	6	7	13
Blue	4	3	7
Total	10	10	20

We can show the same data as percentages or decimals

	Number of Students		
Eye Colour	**Male**	**Female**	**Total**
Brown	30%	35%	65%
Blue	20%	15%	35%
Total	50%	50%	100%

	Number of Students		
Eye Colour	**Male**	**Female**	**Total**
Brown	0.3	0.35	0.65
Blue	0.2	0.15	0.35
Total	0.5	0.5	1

All three of these tables have the same information, just written in a slightly different way.

This means we can then ask some questions like:

a. What percentage of the students are male?
   We can see the total for the male column is 50%. We could also work this out as 10/20.

b. What is the probability if choose a student at random they have blue eyes?
   We can see the total for the blue eyes column is 0.35. We could also work this out as 7/20.

c. How many female students are there with brown eyes?
   We can see the box that is in the brown row and the female column has 7 in it.

You will notice that for all of the rows, and all of the columns, the total row is just what the different parts add up to. This means if we are given a table with missing values we can often work out the missing values. For example if I was given this table:

	Number of Students		
Eye Colour	**Male**	**Female**	**Total**
Brown	0.3	0.35	*d*
Blue	0.2	*f*	*e*
Total	*b*	*c*	*a*

I could work out all the missing values:

a. This is 1, as probabilities always add up to 1.
b. This is 0.5, as 0.3 + 0.2 = 0.5
c. This is 0.5 as well, as b and c need to add to make 1.
d. This is 0.65 as 0.3 + 0.35 = 0.65
e. This is 0.35, as d and e need to add up to 1.
f. This is 0.15, as 0.2 and e need to make 0.35.

Occasionally you are given tables without a total column. The first thing you should do if this is the case is add the total column on.