Level 1 Experimental Probability Workbook







Name:



Written by Liz Sneddon

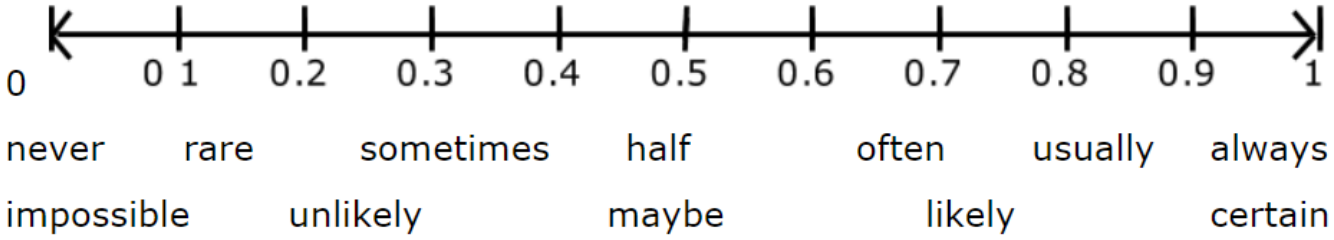
© Liz Sneddon 2024

Language of probability

Probability is a measure of the chance of an event happening.

It can be written as a decimal, fraction, or percentage (decimal or fraction are best)

Probabilities MUST be between 0 and 1 (inclusive)



Exercise 1:

1) Rate the following events as certain, likely, unlikely, impossible, or even chance (circle your answer)

a)	A lion having four legs	Certain / Likely / Impossible / Even / Unlikely
b)	Rolling a 7 with a normal die	Certain / Likely / Impossible / Even / Unlikely
c)	Choosing a black card	Certain / Likely / Impossible / Even / Unlikely
d)	Monday will follow Sunday next week	Certain / Likely / Impossible / Even / Unlikely
e)	Winning Lotto (assuming you have not bought a ticket)	Certain / Likely / Impossible / Even / Unlikely
f)	Throwing a tail with a coin	Certain / Likely / Impossible / Even / Unlikely
g)	Rolling a 3 with a normal die	Certain / Likely / Impossible / Even / Unlikely
h)	A year having 410 days	Certain / Likely / Impossible / Even / Unlikely

2) Select the most appropriate percentage from the probabilities below to describe the chance implied by each of the following words (circle your answer).

a)	Certain	0%, 25%, 50%, 75%, 100%
b)	An even chance	0%, 25%, 50%, 75%, 100%
c)	Probably	0%, 25%, 50%, 75%, 100%
d)	Unlikely	0%, 25%, 50%, 75%, 100%
e)	Definitely	0%, 25%, 50%, 75%, 100%
f)	Likely	0%, 25%, 50%, 75%, 100%
g)	Possibly	0%, 25%, 50%, 75%, 100%

3) Select from the words below to best describe an event that has a probability of the values below (circle your answer).

a)	90%	Certain / Likely / Impossible / Even / Unlikely
b)	1/2	Certain / Likely / Impossible / Even / Unlikely
c)	1	Certain / Likely / Impossible / Even / Unlikely
d)	0	Certain / Likely / Impossible / Even / Unlikely
e)	15%	Certain / Likely / Impossible / Even / Unlikely
f)	100%	Certain / Likely / Impossible / Even / Unlikely
g)	0·007	Certain / Likely / Impossible / Even / Unlikely

Theoretical Probability

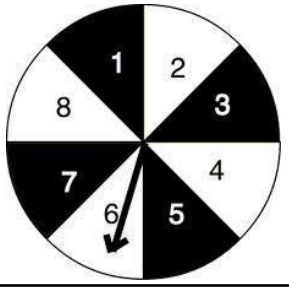
A coin:

A coin has two sides, heads, and tails.



A Spinner:

A spinner is a circle (usually) that has been divided up into any number of pieces.



A die (plural: dice):

A standard die has six sides, with the numbers 1 – 6 on each side. You can get die with 10, 12, or other numbers of sides.



Random number generator:

Your calculator has a random number generator, which can generate a random number between 0 and 1.



A pack of cards:

In a pack of cards there are four suites: Diamonds, Hearts, Clubs, Spades

Each suite has thirteen cards:

Ace, King, Queen, Jack, 10, 9, 8, 7, 6, 5, 4, 3, 2



Terminology

Outcome

One possible result of an experiment.

E.g. rolling a 6

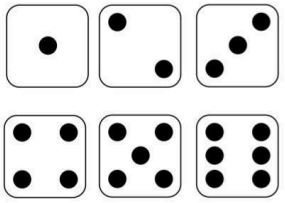


© Liz Sneddon

Sample space

A list of all the possible outcomes of an experiment.

E.g. a dice



© Liz Sneddon

Event

Something which has a number of outcomes.

E.g. rolling a die



© Liz Sneddon

Trial

A single run through of an experiment.

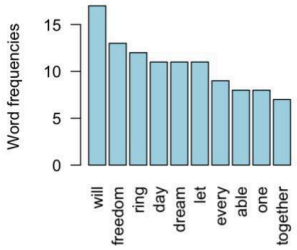
E.g. a roll of the dice



© Liz Sneddon

Frequency

the number of times an event occurs

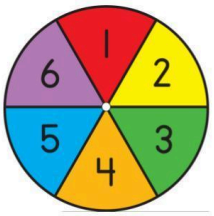


© Liz Sneddon

Equally likely

Outcomes with the same chance.

E.g. outcomes on this spinner



© Liz Sneddon

P(outcome A) =

"The probability of getting outcome A is ...:"

© Liz Sneddon

Independent

Events that do not affect or are not affected by another event.



Each coin toss is not affected by the previous coin toss.

© Liz Sneddon

Relative Frequency

$$P(\text{event}) = \frac{\text{Frequency of outcome}}{\text{Total number of outcomes}}$$

© Liz Sneddon

Example:

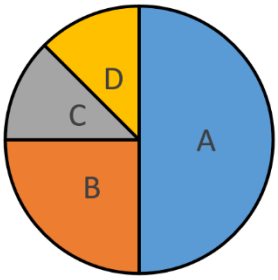
Write the sample space for the following:

1) A single dice



Sample space = 1, 2, 3, 4, 5, 6

2) A spinner

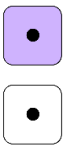
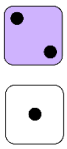
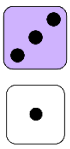
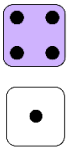
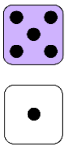
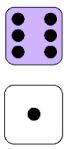
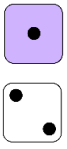
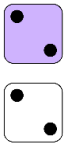
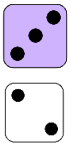
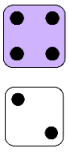
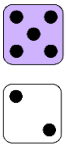
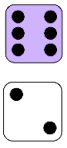
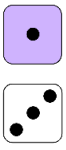
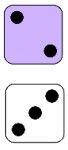
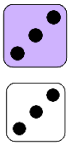
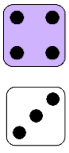
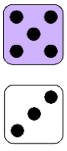
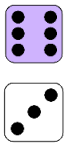
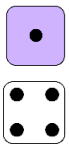
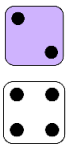
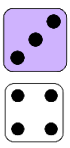
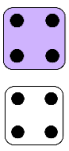
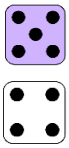
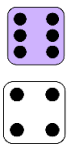


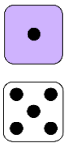
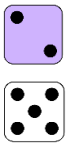
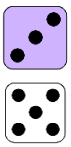
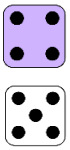
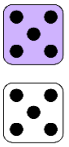
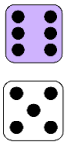
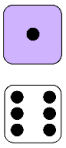
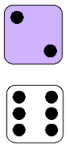
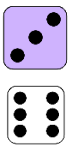
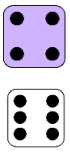
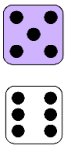
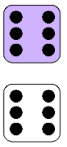
Sample space = A, B, C, D

3) Two dice

Sample space:



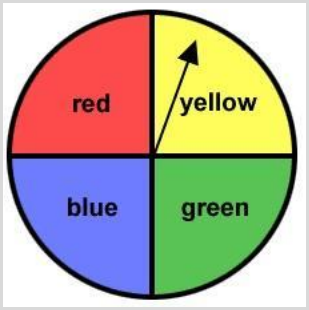
2nd dice \ 1st dice						
						
						
						
						

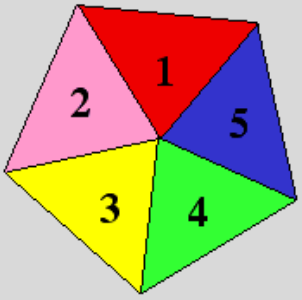
Exercise 2:

1) Write down the sample space for the following objects:

a)



b)



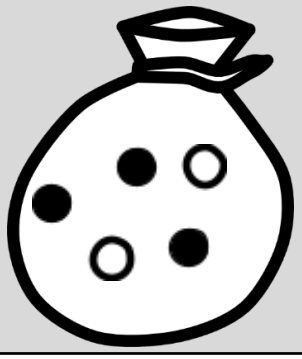
c) A 20-sided dice



d) Lotto balls



e) A bag of black and white marbles

2) Write the sample space for each of the following.

a)	selecting a day of the week	
b)	selecting a month of the year	
c)	rolling a die once	
d)	tossing a coin once	
e)	2 students playing a game of rugby.	

3) The letters of the word MATHEMATICS are written on cards and turned face down. A card is then selected at random.

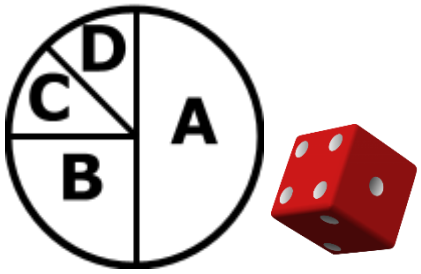
a)	Write the sample space.	
b)	How many elements are in the sample space?	
c)	How many **different** elements are in the sample space?	

4) For each of the following, state whether **each element** of the sample space is **equally likely** to occur (circle your answer).

a)	tossing a coin	Equally likely / not equally likely

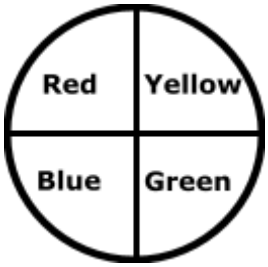
b)	rolling a die	Equally likely / not equally likely
c)	the result of a cricket game between two teams	Equally likely / not equally likely
d)	selecting a card from a normal pack of cards	Equally likely / not equally likely

5) Complete the table below for the sample space for a dice and a spinner.



		Dice					
							
Spinner							

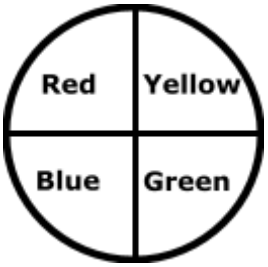
6) Complete the table below for the sample space for a dice and a spinner.





Spinner \ Dice						

7) Complete the table below for the sample space two spins of the spinner.



2nd spin \ 1st spin				

8) Use the following words to complete these sentences.

discrete data	continuous data	mean	range
variable	mode	median	

a)	The average of a set of scores is called the _____
b)	When the scores are arranged in order of size (ascending or descending order) the middle value is called the _____
c)	The score that occurs the most is called the _____
d)	The difference between the highest and the lowest scores is called the _____
e)	Any quantity that varies is called a _____
f)	_____ is count data that has values that are whole numbers.
g)	_____ are measurements that can have any value in a given range.

Probabilities

To find the probability of an event

$$P(event) = \frac{number\ of\ ways\ it\ can\ happen}{total\ number\ of\ outcomes}$$

Example:

1) Rolling a die has the following outcomes: 1, 2, 3, 4, 5, or 6

The probability that the number 5 comes up when a die is rolled is:



$P(5\ on\ an\ die) = \frac{1}{6}$

2) A bag of marbles has 3 blue and 2 white marbles.

The probability that a blue marble is chosen randomly is:



$P(blue) = \frac{3}{5}$

Exercise 3:

1) A die is thrown once. Find the probability that it is:

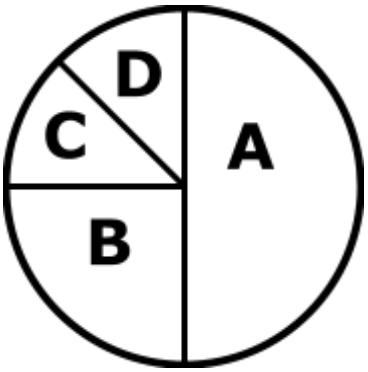
a)	P(a six) =
b)	P(an even number) =
c)	P(a number greater than 4) =
d)	P(a seven) =

2) A bag contains 5 red, 3 blue and 2 white balls. If a ball is drawn at random, find the probability that it is:

a)	P(blue) =
b)	P(red) =
c)	P(not red) =
d)	P(either blue or red) =

3) The spinner is spun once. Find the probability that:

a)	P(A) =
b)	P(B) =



c)	P(C) =
d)	P(D) =
e)	P(A or B) =

4) A number is chosen at random from 1 to 10. Find the probability that the number is:

a)	P(4) =
b)	P(even) =
c)	P(1 or 2) =
d)	P(not 5) =
e)	P(more than 7) =
f)	P(less than 3) =

5) From the letters of the word '**PROBABILITY**', one letter is selected at random. What is the probability that the letter is:

a)	P(a vowel) =
b)	P(the letter P) =
c)	P(the letter P or B) =
d)	P(the letter M) =

6) A card is drawn at random from a normal pack of 52 cards.



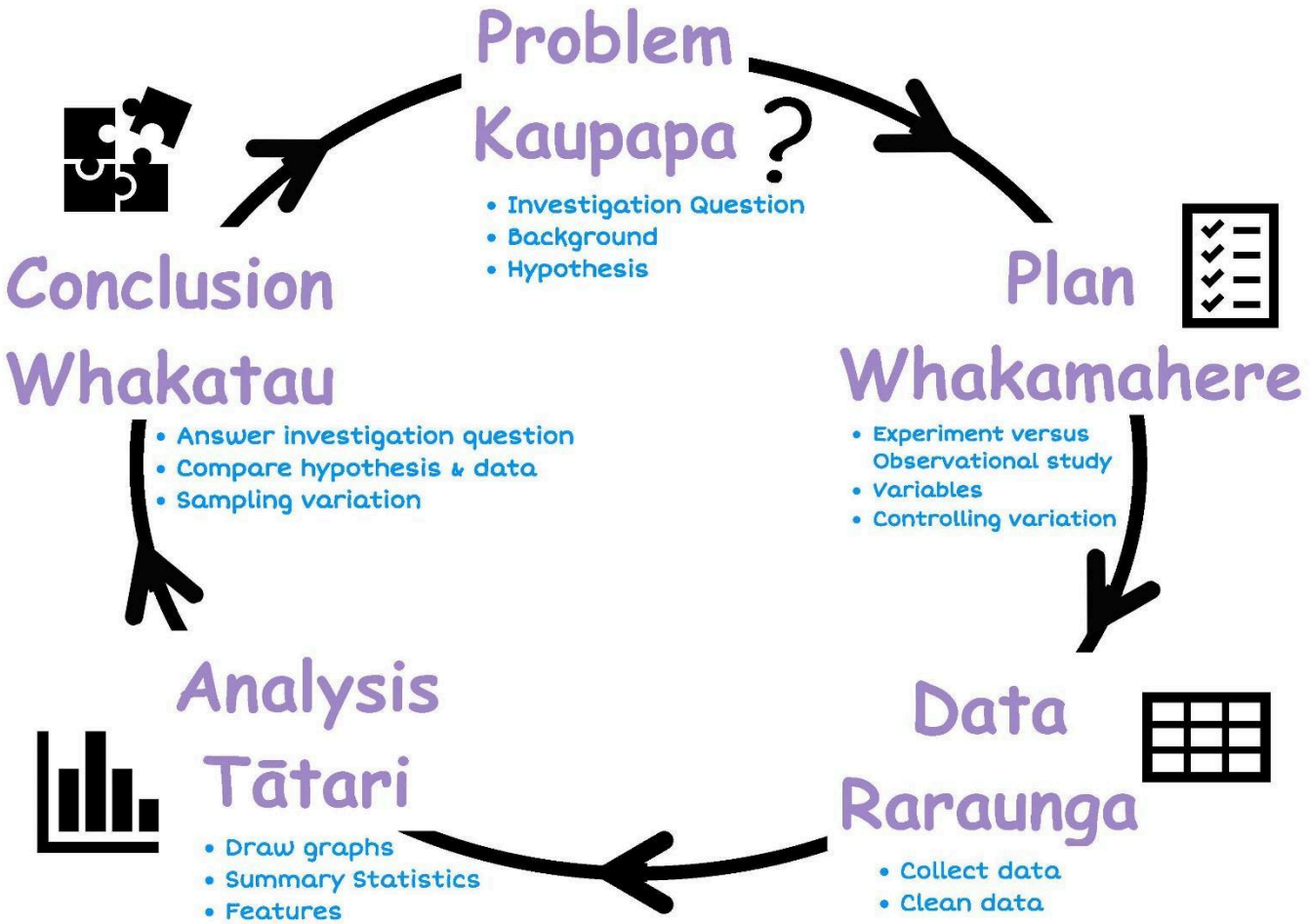
Find the probability that the card is:

a)	P(a diamond) =
b)	P(a red card) =
c)	P(a king) =
d)	P(not a club) =
e)	P(a red ace) =
f)	P(a 5 of clubs) =
g)	P(an 8) =
h)	P(Jack, Queen, or King) =
i)	P(less than 4) =
j)	P(not a 5) =
k)	P(not a red 3) =

Experimental Probability Investigations

The previous section was a reminder of some of the terminology and ideas around probability. We now need to move onto investigating situations where we often do not know what the probability of the outcome is. For example, investigating the chance of me scoring a goal in a game of netball, or the chance that the photocopier breaks down.

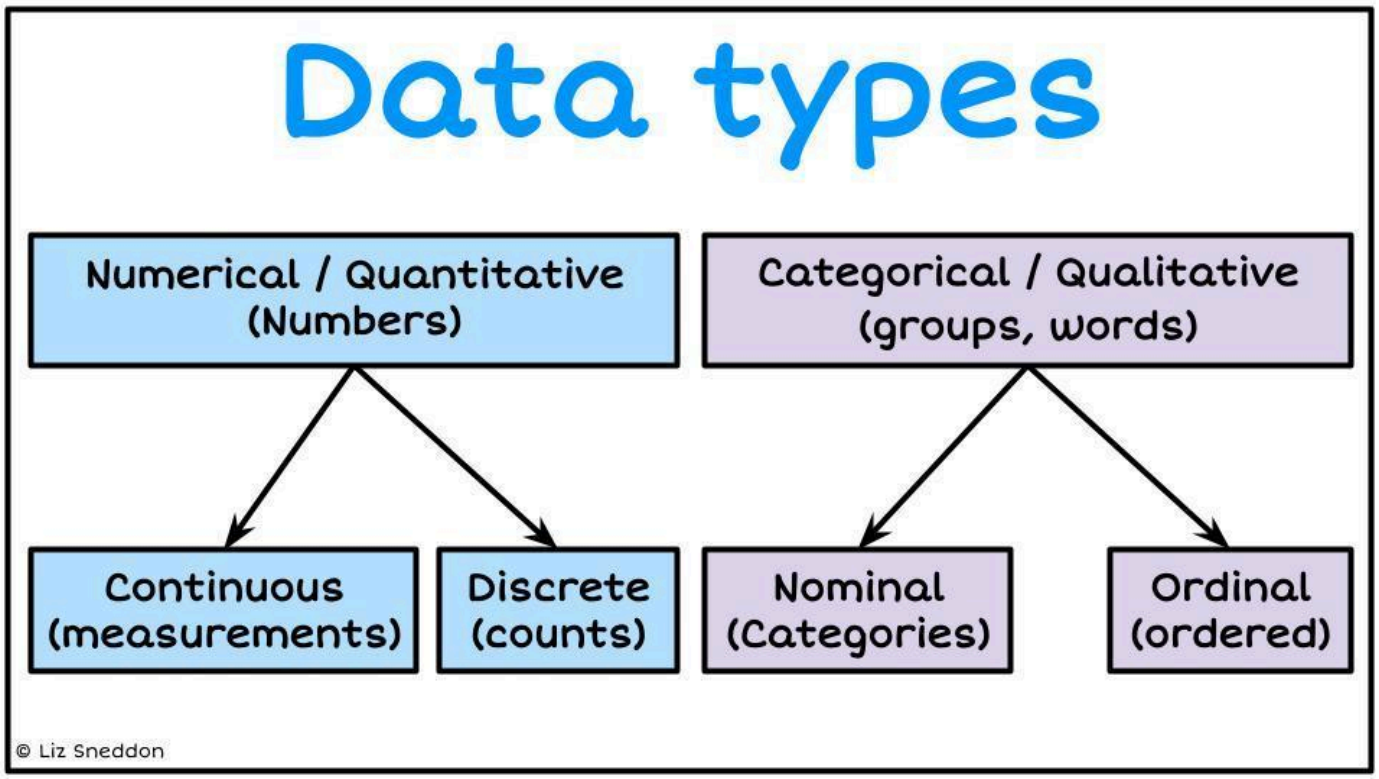
To investigate experimental probabilities, we will use the PPDAC cycle:

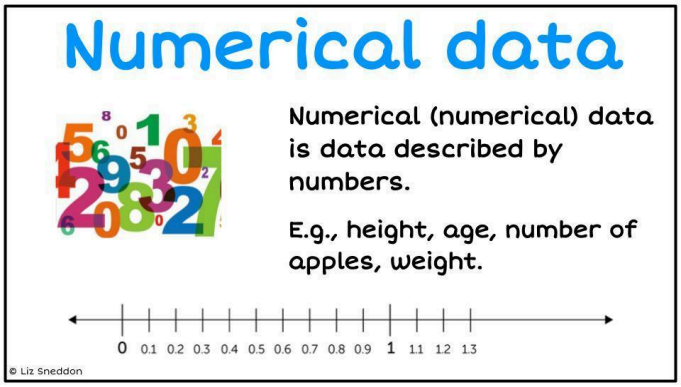


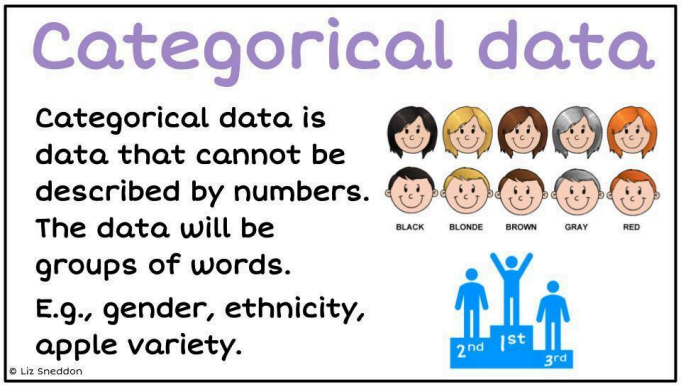
Problem / Kaupapa

Your investigation question will be about the **probability** that an **event** occurs.

A probability question can be asked of either categorical data or numerical data.







First you need to identify the variable and all the events (or sample space). Then you can ask an investigation question about one or more of the events.

You can also investigate the probability of NOT getting something, or of getting several events combined.

Example:

1)

Situation	The situation we will explore is selecting either an apple, pear or peach from a bowl which has 50 pieces of fruit in it.
Variable	Types of fruit in the bowl
Sample space (all events)	Apple, pear, peach
Data type	Categorical
Investigation question(s)	1) I wonder what the probability of getting an **apple** out of the fruit bowl is? 2) I wonder what the probability of getting a **pear** out of the fruit bowl is? 3) I wonder what the probability of getting a **peach** out of the fruit bowl is? 4) I wonder what the probability of **NOT getting an apple** out of the fruit bowl is? 5) etc

2)

Situation	Mrs Sneddon wants to explore how many good night's sleep she gets each week.
Variable	Number of good night's sleep per week
Sample space (all events)	0, 1, 2, 3, 4, 5, 6, 7
Data type	Numerical
Investigation question(s)	1) What is the probability that Mrs Sneddon has **7 good nights sleep** each week? 2) What is the probability that Mrs Sneddon has **0 good nights sleep** each week? 3) What is the probability that Mrs Sneddon has **at least 5 good night sleeps** each week? 4) … This is just a few of the many investigation questions that could be posed.

Exercise 4:

For each situation below, identify the variable, the sample space (all events) whether the data type is numeric or categorical, and write appropriate investigation questions.

1)

Situation	Mrs Sneddon has a kete (basket) with highlighters for students to use, including kōhai (yellow), kākāriki (green), Kikorangi (blue) and māwhero (pink).
Variable	
Sample space (all events)	
Data type	Numerical / Categorical
Investigation question(s)	

2)

Situation	A bag of lollipops has 4 different flavours, grape, cola, orange, and strawberry.
Variable	
Sample space (all events)	
Data type	Numerical / Categorical
Investigation question(s)	

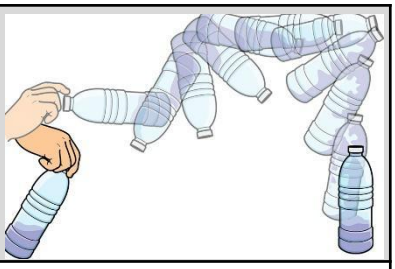
3)

Situation	When playing basketball and doing a penalty shootout, the player gets 3 shots at goal.
Variable	
Sample space (all events)	
Data type	Numerical / Categorical
Investigation question(s)	

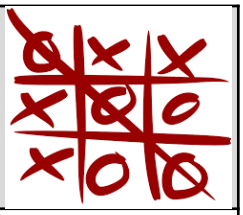
4)

Situation	Apps will often send out notifications when a new message or post is made. Explore notifications received in a one-hour time period.
Variable	
Sample space (all events)	
Data type	Numerical / Categorical
Investigation question(s)	

5)

Situation	Doing a bottle flip challenge where each person tries to flip a bottle 5 times to see how often it lands standing up. 
Variable	
Sample space (all events)	
Data type	Numerical / Categorical
Investigation question(s)	

6)

Situation	Two students want to see who is better at playing tic tac toe. 
Variable	
Sample space (all events)	
Data type	Numerical / Categorical
Investigation question(s)	

Plan / Whakamahere

There are several things that you need to identify once you have written your investigation question.

- Define the variable,
- Identify the sample space (a list of all the outcomes),
- State the number of trials (sample size),
- Any possible sources of variation,
- Write a set of instructions to carry out the experimental investigation.

Example:

Situation:

Mrs Sneddon wants to explore how many good nights sleep she gets each week.

Variable	The number of good nights' sleep each week.
Sample space	0, 1, 2, 3, 4, 5, 6, 7
Number of trials	50 weeks
Sources of variation	• Record whether it is a good night's sleep at the same time each day. • Sleep in the same bed each night. • Try to go to sleep at the same time each night. • Try to have the same routine each night to wind down before going to sleep.
Instructions	1) Starting on a Monday morning, Mrs Sneddon will record whether she had a good nights' sleep. 2) On Sunday of each week, Mrs Sneddon will count up the number of good nights sleep for the week and record this on a table. The table will have one column with the week number (e.g., Week 1, 2 …), one column with tallies recording if it was a good night's sleep, and one column with the number of how many good nights sleep she had each week. 3) Repeat steps one and two for a total of 50 weeks.

Exercise 5:

For each of the situations below, identify the variable, sample space, number of trials, sources of variation, and write a set of instructions to collect the data.

1)

Situation	Mrs Sneddon has a kete (basket) with highlighters for students to use, including kōhai (yellow), kākāriki (green), Kikorangi (blue) and māwhero (pink).
Variable	
Sample space	
Number of trials	
Sources of variation	
Instructions	

2)

Situation	A bag of lollipops has 4 different flavours, grape, cola, orange, and strawberry.
Variable	
Sample space	
Number of trials	
Sources of variation	
Instructions	

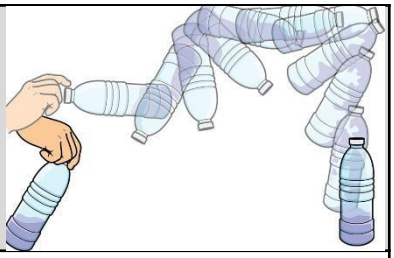
3)

Situation	When playing basketball and doing a penalty shootout, the player gets 3 shots at goal.
Variable	
Sample space	
Number of trials	
Sources of variation	
Instructions	

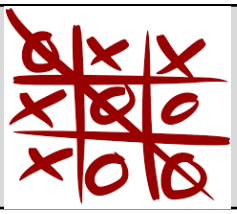
4)

Situation	Apps will often send out notifications when a new message or post is made. Explore notifications received in a one-hour time period.
Variable	
Sample space	
Number of trials	
Sources of variation	
Instructions	

5)

Situation	Doing a bottle flip challenge where each person tries to flip a bottle 5 times to see how often it lands standing up. 
Variable	
Sample space	
Number of trials	
Sources of variation	
Instructions	

6)

Situation	Two students want to see who is better at playing tic tac toe. 
Variable	
Sample space	
Number of trials	
Sources of variation	
Instructions	

Data / Tātari

Once you have your plan, the next step is to collect the data. This requires you to record the data in a table, and then form graphs of the data. We want to be able to do this using Excel or other spreadsheets on our computer.

The first table you want to create is recording the outcomes of each trial. The second table is a summary of the data.

Example:

Situation:

Mrs Sneddon wants to explore how many good night's sleep she gets each week.

Table 1: Recording the data

Trial / Week	Tally	Number of good night's sleep
1		
2		
3		
…		
50		

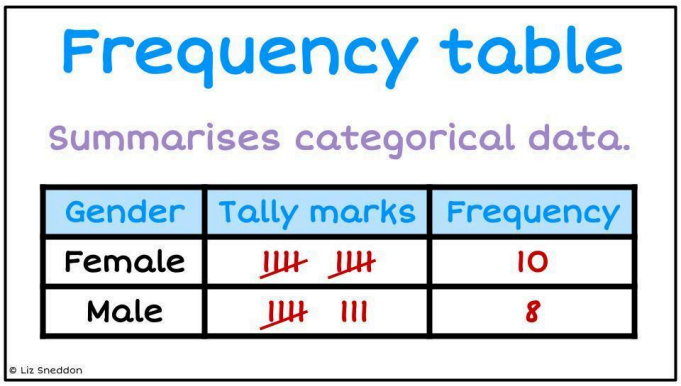
Table 2: Summarising the data

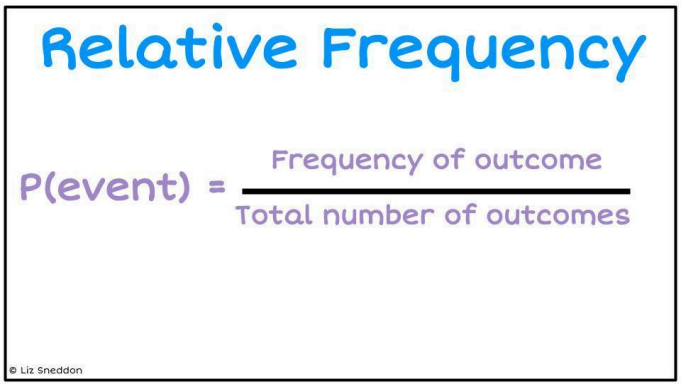
Sample space (number of good night's sleep)	Frequency	Relative frequency
0		
1		
2		
3		
4		
5		
6		
7		

Tally, Frequency and Relative Frequency

A tally is recording a tally mark when you have a certain outcome in your data set.

A frequency is counting up how many times each number occurs.





Example:

The frequency of the following numbers is:

4, 8, 3, 2, 7, 7, 8, 7, 3, 6, 5, 2, 8, 5, 4, 7, 6, 2, 3, 2

To find the relative frequency, divide the frequency by the total number of trials.

Score	Tally	Frequency	Relative Frequency
2	||||	4	4 / 20 = 0.2
3	|||	3	3 /20 = 0.15
4	||	2	2 / 20 = 0.1
5	||	2	2 / 20 = 0.1
6	||	2	2 / 20 = 0.1
7	||||	4	4 / 20 = 0.2
8	|||	3	3 / 20 = 0.15
		Total = 20	

Exercise 6:

1) The test results obtained by a class of 30 students is given. The test was marked out of 10 points. Complete the table below.

Trial / Student	Test mark
1	9
2	7
3	7
4	5
5	6
6	7
7	6
8	7
9	8
10	6
11	6
12	6
13	8
14	9
15	10
16	5
17	6
18	7
19	8
20	6
21	5
22	4
23	3
24	6
25	7
26	9
27	8
28	9
29	7
30	9

Score	Tally	Frequency	Relative frequency
0			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

2) Twenty-five games of tic-tac-toe were played and the following set of data was obtained. Complete the frequency distribution table.

Trial	Winner
1	Liz
2	Draw
3	Jamie
4	Liz
5	Liz
6	Liz
7	Draw
8	Jamie
9	Jamie
10	Liz
11	Draw
12	Jamie
13	Liz
14	Jamie
15	Draw
16	Jamie
17	Liz
18	Liz
19	Jamie
20	Liz
21	Jamie
22	Jamie
23	Liz
24	Draw
25	Liz

Winner	Tally	Frequency	Relative frequency
Liz			
Jamie			
Draw			

Excel Spreadsheet

Entering data

We want to use Excel to record the data. Here are the instructions.

Situation: selecting either an apple, pear or peach from a bowl which has 50 pieces of fruit.

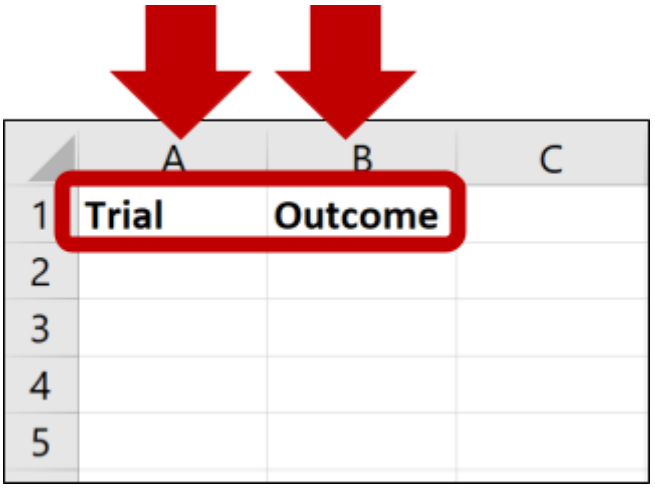
Step 1:

Open up Excel on your laptop and open a blank document.



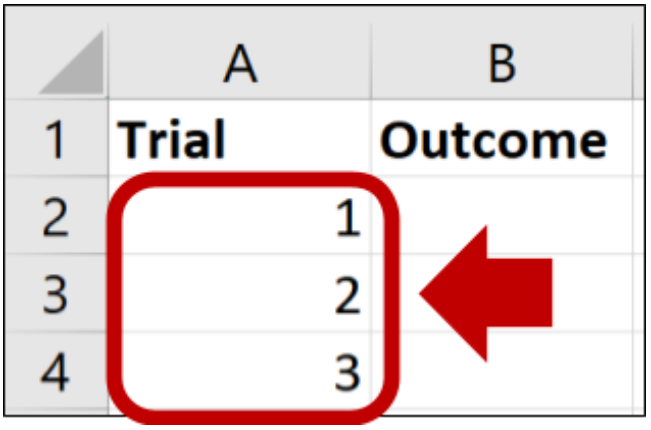
Step 2:

Label the first column **Trial** and label the second column **Outcome.**



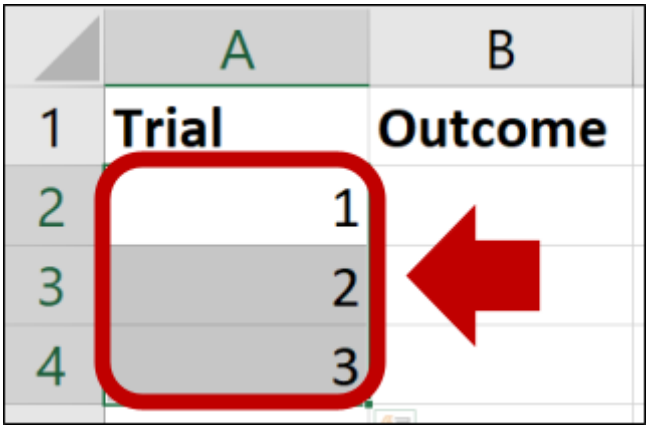
Step 3:

In column A, we want to enter the numbers between 1 and 50 (the number of trials we want to do). Excel can recognise patterns, so type the numbers 1, 2 and 3 into the next 3 rows.



Step 4:

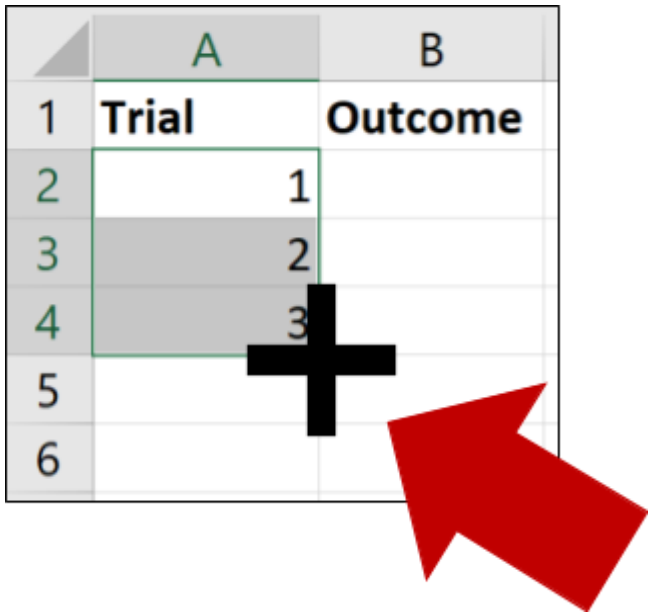
Next, select these cells with the numbers 1, 2 and 3.



Step 5:

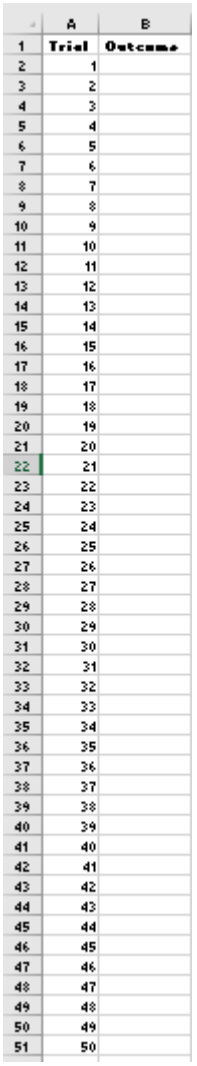
Move your mouse to the bottom right-hand corner of these selected cells,

and a + symbol will appear.



Step 6:

Click and hold your mouse button, and drag it down until the 51st row, and Excel will fill in the numbers up until 50.



Step 7:

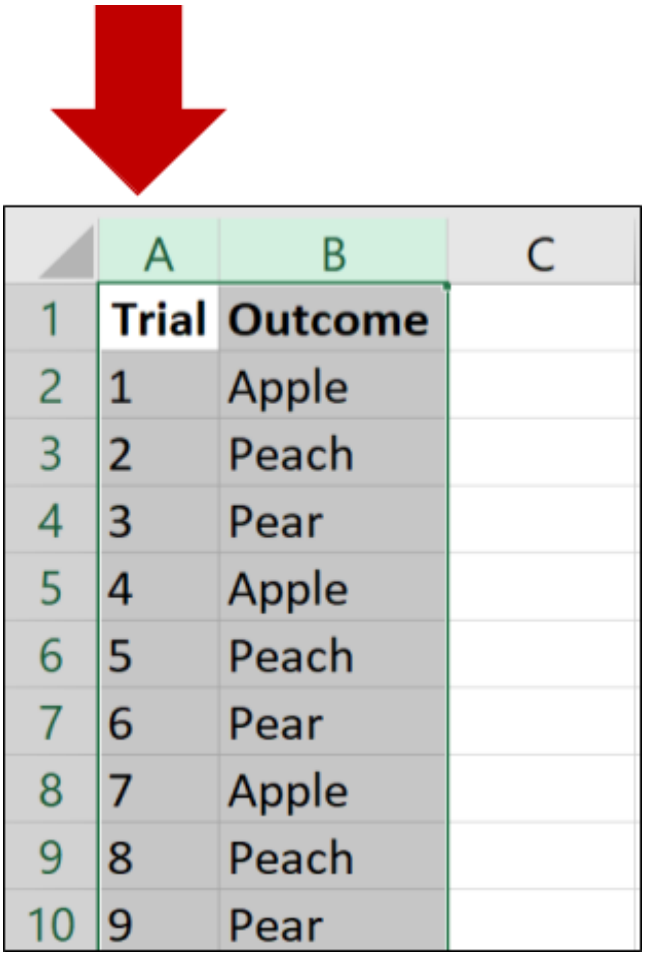
Then enter the outcomes in the second column, and this is your first table with data.

	A	B
1	**Trial**	**Outcome**
2	1	Apple
3	2	Peach
4	3	Pear
5	4	Apple
6	5	Peach
7	6	Pear
8	7	Apple
9	8	Peach
10	9	Pear

Excel Tip to resize the width of the columns

Step 1:

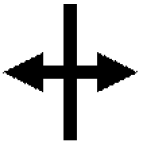
We can get Excel to automatically resize the width of the columns. Select the columns you want to resize, in this case, select columns A and B by going to the letter A, click and hold, and move the mouse across to the letter B. Then let go of the mouse button.



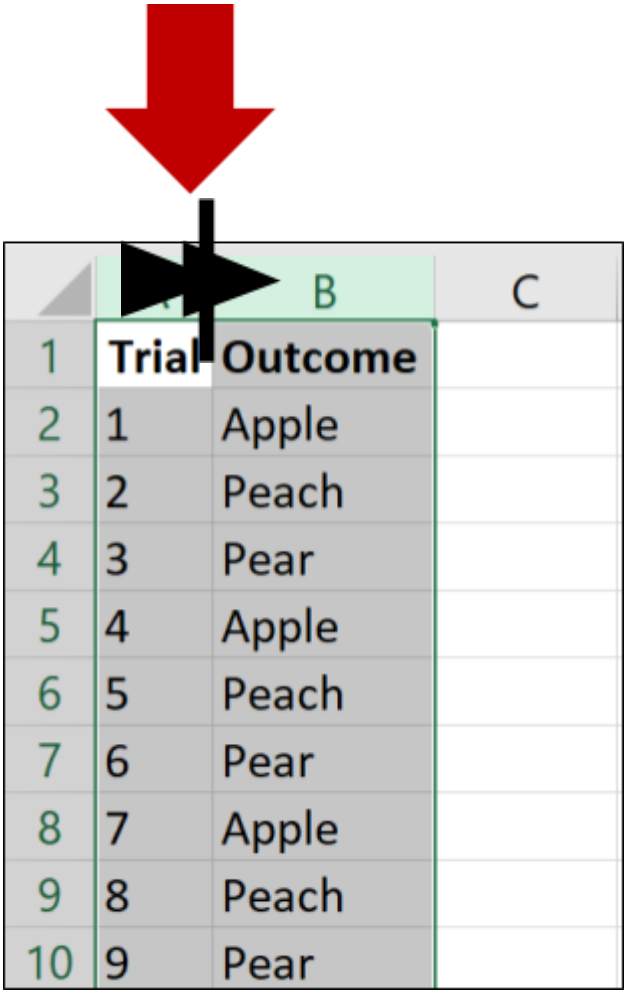
	A	B	C
1	Trial	Outcome	
2	1	Apple	
3	2	Peach	
4	3	Pear	
5	4	Apple	
6	5	Peach	
7	6	Pear	
8	7	Apple	
9	8	Peach	
10	9	Pear	

Step 2:

Move your mouse to be on the line between column A and column B, and a



symbol will appear. Double click on this symbol and the columns will automatically resize to the right width.



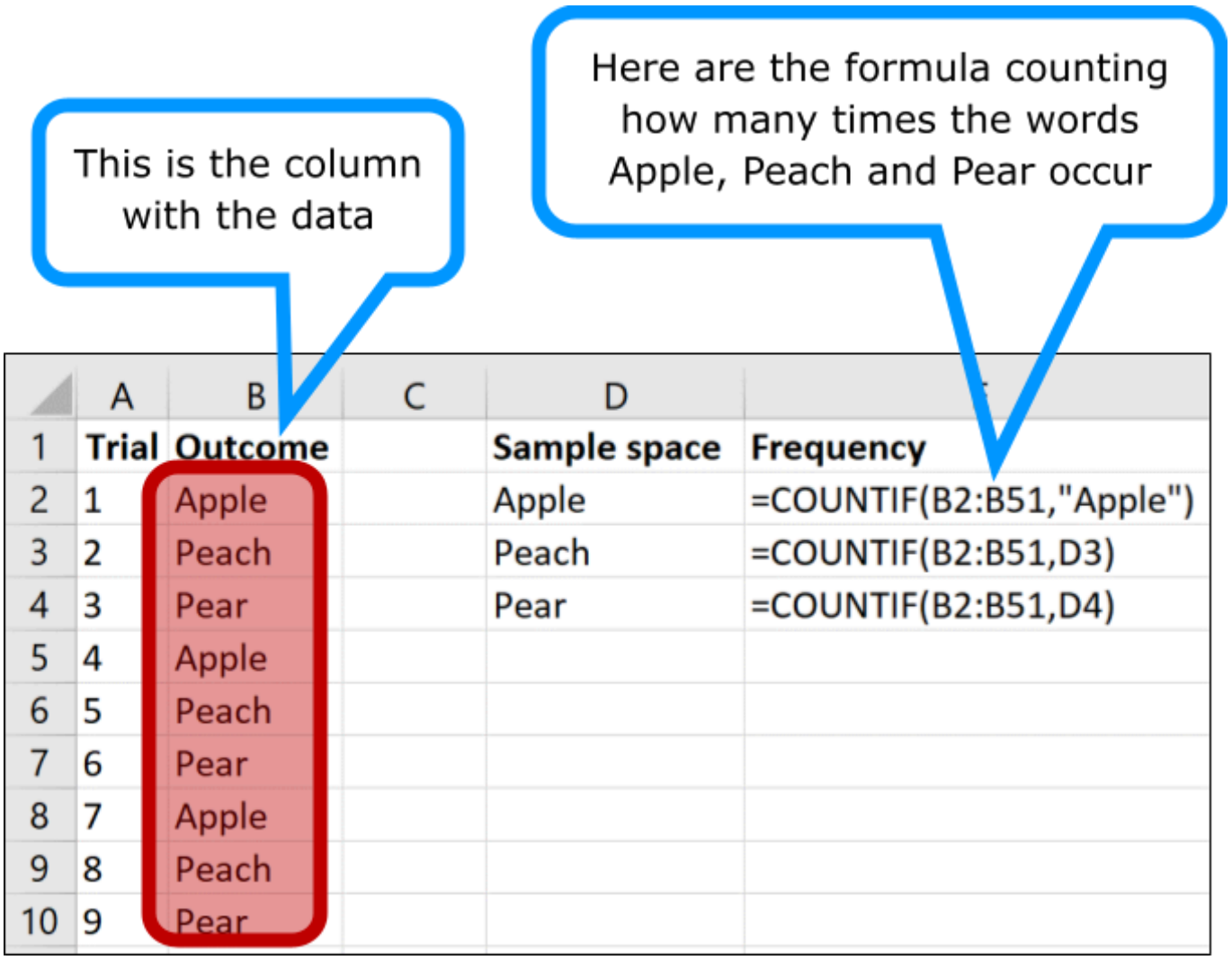
	B	C
1	Trial	Outcome
2	1	Apple
3	2	Peach
4	3	Pear
5	4	Apple
6	5	Peach
7	6	Pear
8	7	Apple
9	8	Peach
10	9	Pear

COUNTIF formula

This is a useful formula to use. It will count the number of times a specific outcome occurs.

= COUNTIF (Range of cells , Criteria)

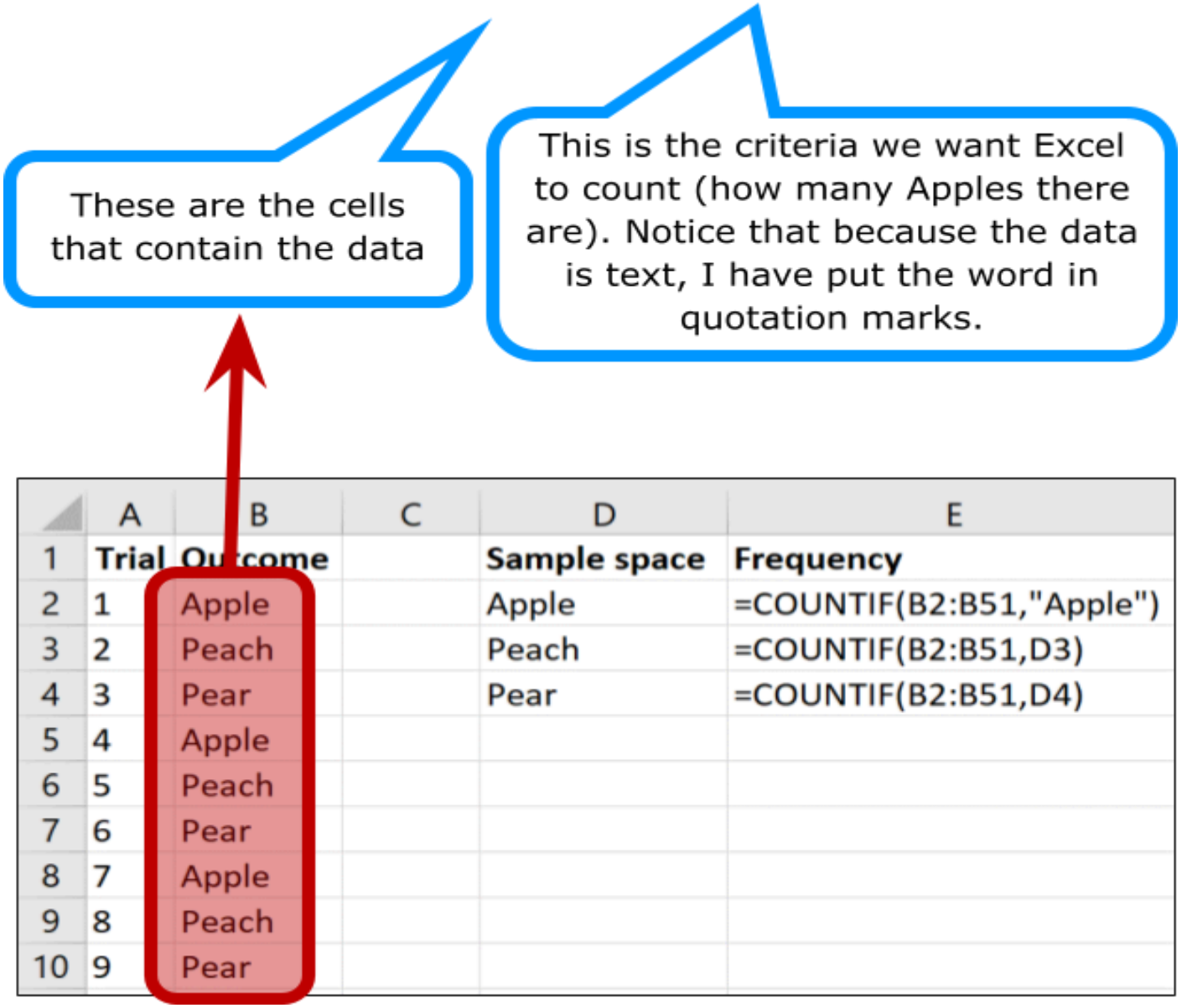
Example:



	A	B	C	D	E
1	**Trial**	**Outcome**		**Sample space**	**Frequency**
2	1	Apple		Apple	=COUNTIF(B2:B51,"Apple")
3	2	Peach		Peach	=COUNTIF(B2:B51,D3)
4	3	Pear		Pear	=COUNTIF(B2:B51,D4)
5	4	Apple			
6	5	Peach			
7	6	Pear			
8	7	Apple			
9	8	Peach			
10	9	Pear			

Let’s look at the different formula to understand how they work.

= COUNTIF (B2:B51 , "Apple")

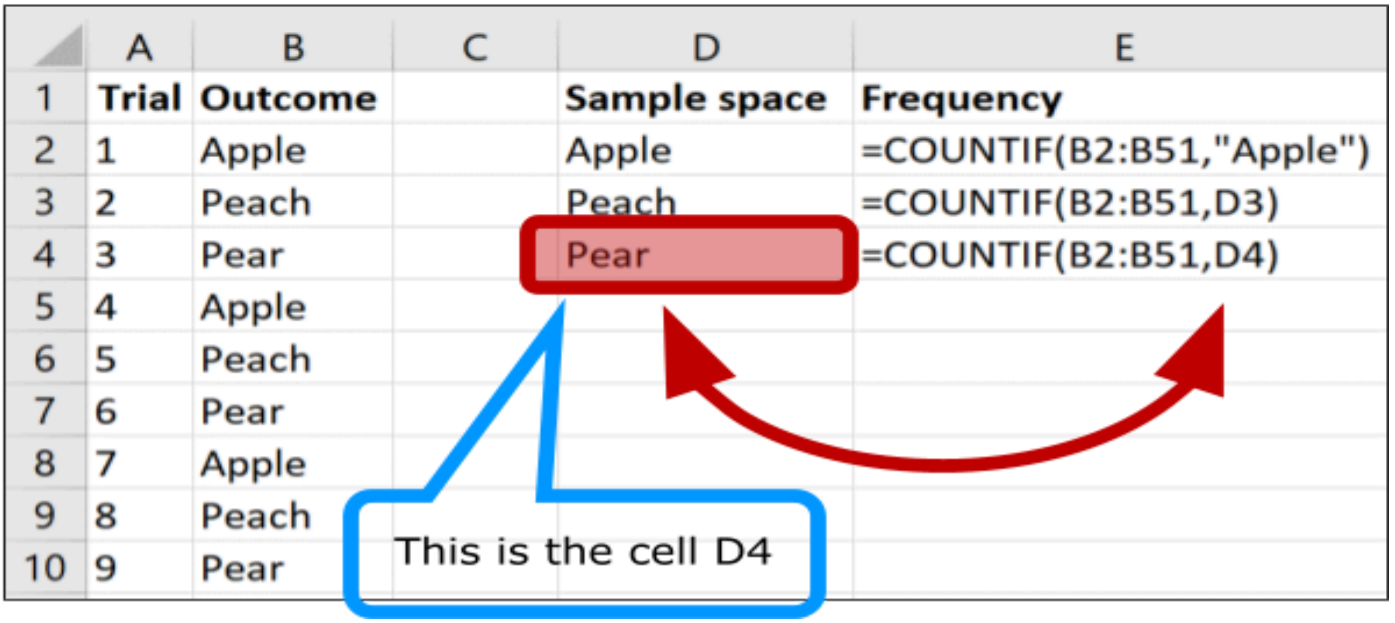


	A	B	C	D	E
1	Trial	Outcome		Sample space	Frequency
2	1	Apple		Apple	=COUNTIF(B2:B51,"Apple")
3	2	Peach		Peach	=COUNTIF(B2:B51,D3)
4	3	Pear		Pear	=COUNTIF(B2:B51,D4)
5	4	Apple			
6	5	Peach			
7	6	Pear			
8	7	Apple			
9	8	Peach			
10	9	Pear			

= COUNTIF (B2:B51 , D4)

These are the cells that contain the data

This is the criteria we want Excel to count. Notice that this is referring to the cell **D4**, where I typed the word **Pear**.



	A	B	C	D	E
1	**Trial**	**Outcome**		**Sample space**	**Frequency**
2	1	Apple		Apple	=COUNTIF(B2:B51,"Apple")
3	2	Peach		Peach	=COUNTIF(B2:B51,D3)
4	3	Pear		Pear	=COUNTIF(B2:B51,D4)
5	4	Apple			
6	5	Peach			
7	6	Pear			
8	7	Apple			
9	8	Peach			
10	9	Pear			

Summary table in Excel

Here are instructions on how to make a summary table in Excel.

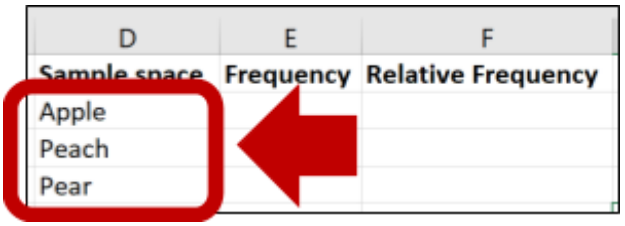
Step 1:

Make headings for the summary table. Label the first column **Sample space**, the second column **Frequency**, the third column **Relative Frequency**. Resize the column widths.

D	E	F
Sample space	**Frequency**	**Relative Frequency**

Step 2:

Under the first column, we want to enter the sample space. E.g., test scores, highlighter colours, etc.



Step 3:

We now want to get Excel to count up the number of times that each outcome occurs.

Type the formula **= COUNTIF (__ , __)** into each cell in the Frequency column.

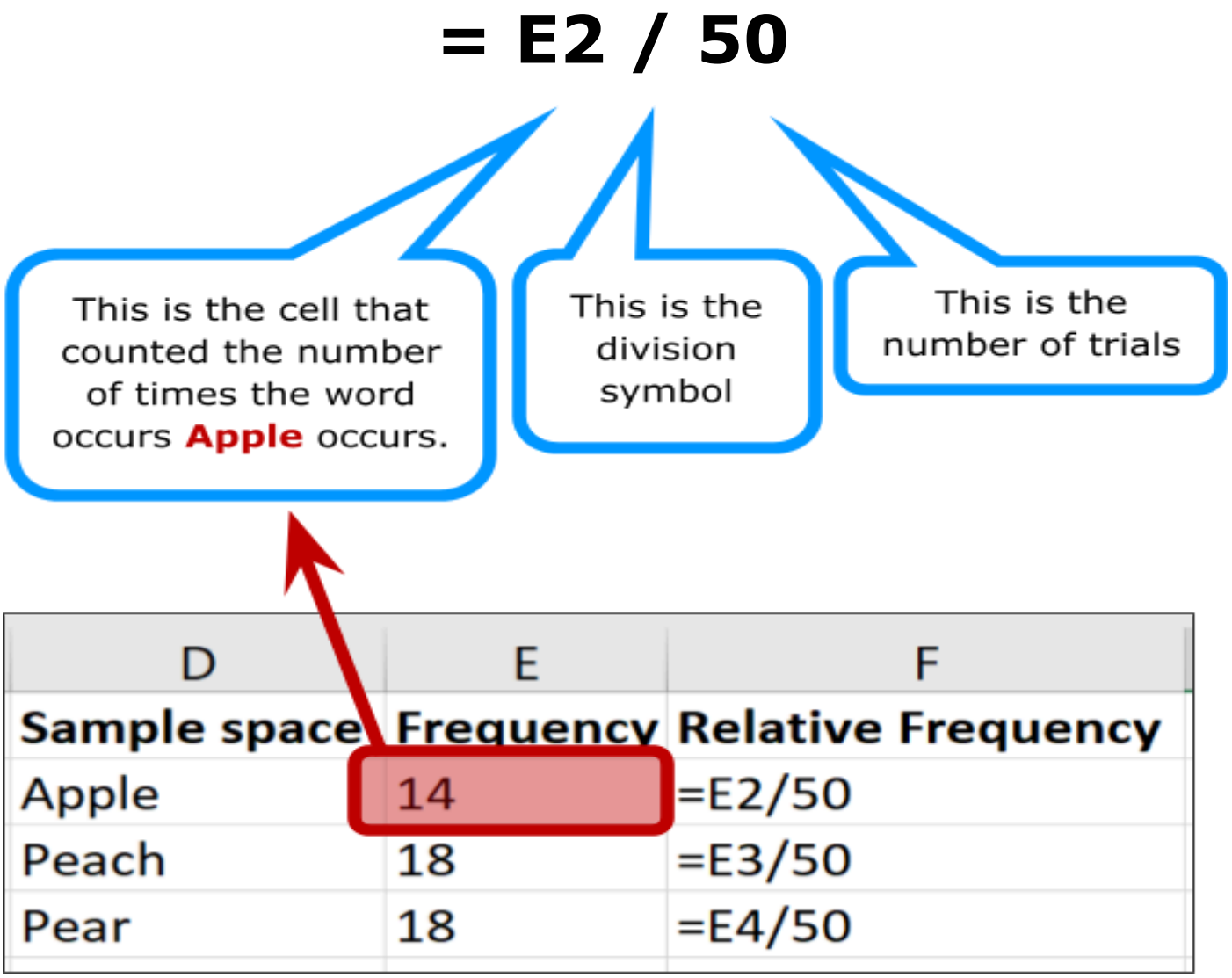
	A	B	C	D	E
1	**Trial**	**Outcome**		**Sample space**	**Frequency**
2	1	Apple		Apple	=COUNTIF(B2:B51,"Apple")
3	2	Peach		Peach	=COUNTIF(B2:B51,D3)
4	3	Pear		Pear	=COUNTIF(B2:B51,D4)
5	4	Apple			
6	5	Peach			

Step 4:

Next, we want to calculate the relative frequency (which is the probability). We just need to divide each frequency by the total number of trials, 50.

D	E	F
Sample space	**Frequency**	**Relative Frequency**
Apple	14	=E2/50
Peach	18	=E3/50
Pear	18	=E4/50

Let’s look at the formula to understand how it works.



Exercise 7:

Enter both of the situations below into Excel spreadsheets. Make a data table and a summary table.

1) The test results obtained by a class of 30 students is given. The test was marked out of 10 points. Complete the table below.

Trial / Student	Test mark
1	9
2	7
3	7
4	5
5	6
6	7
7	6
8	7
9	8
10	6
11	6
12	6
13	8
14	9
15	10
16	5
17	6
18	7
19	8
20	6
21	5
22	4
23	3
24	6
25	7
26	9
27	8
28	9
29	7
30	9

Score	Frequency	Relative frequency
0		
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

2) Twenty-five games of tic-tac-toe were played and the following set of data was obtained. Create both a table of the data and a summary table.

Trial	Winner
1	Liz
2	Draw
3	Jamie
4	Liz
5	Liz
6	Liz
7	Draw
8	Jamie
9	Jamie
10	Liz
11	Draw
12	Jamie
13	Liz
14	Jamie
15	Draw
16	Jamie
17	Liz
18	Liz
19	Jamie
20	Liz
21	Jamie
22	Jamie
23	Liz
24	Draw
25	Liz

Winner	Tally	Frequency	Relative frequency
Liz			
Jamie			
Draw			

Frequency and Relative Frequency bar graphs

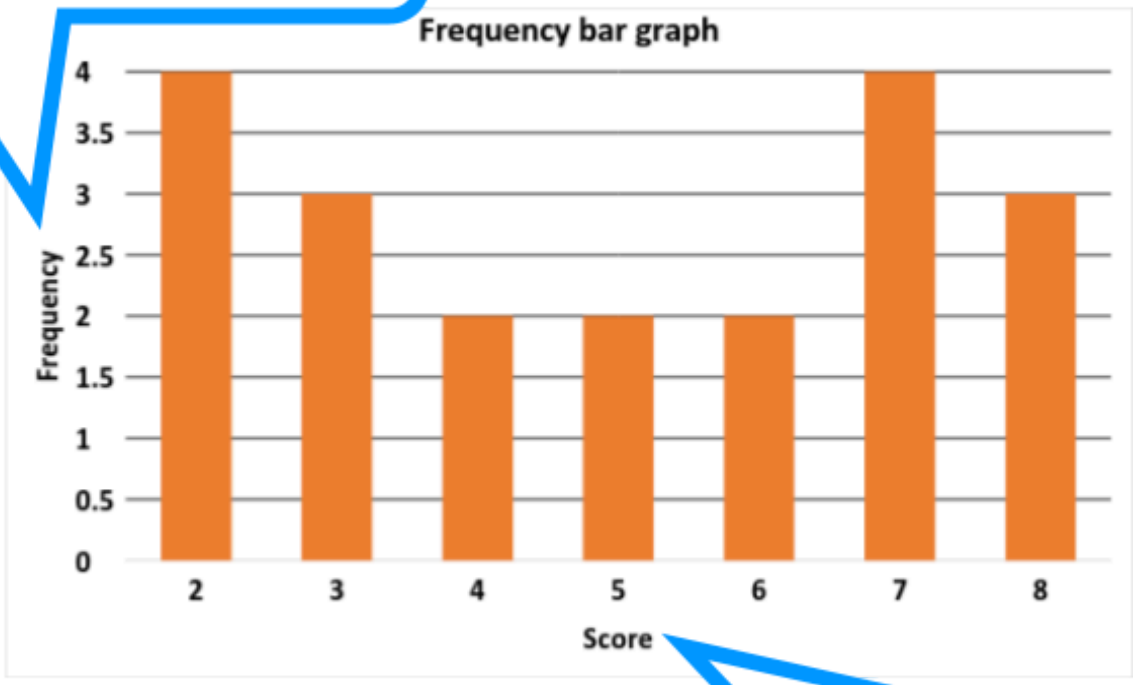
A frequency graph plots the frequency on the vertical axis and the scores or groups on the horizontal axis.

A relative frequency graph plots the relative frequency (as a decimal between 0 and 1) on the vertical axis and the scores or groups on the horizontal axis.

Example:

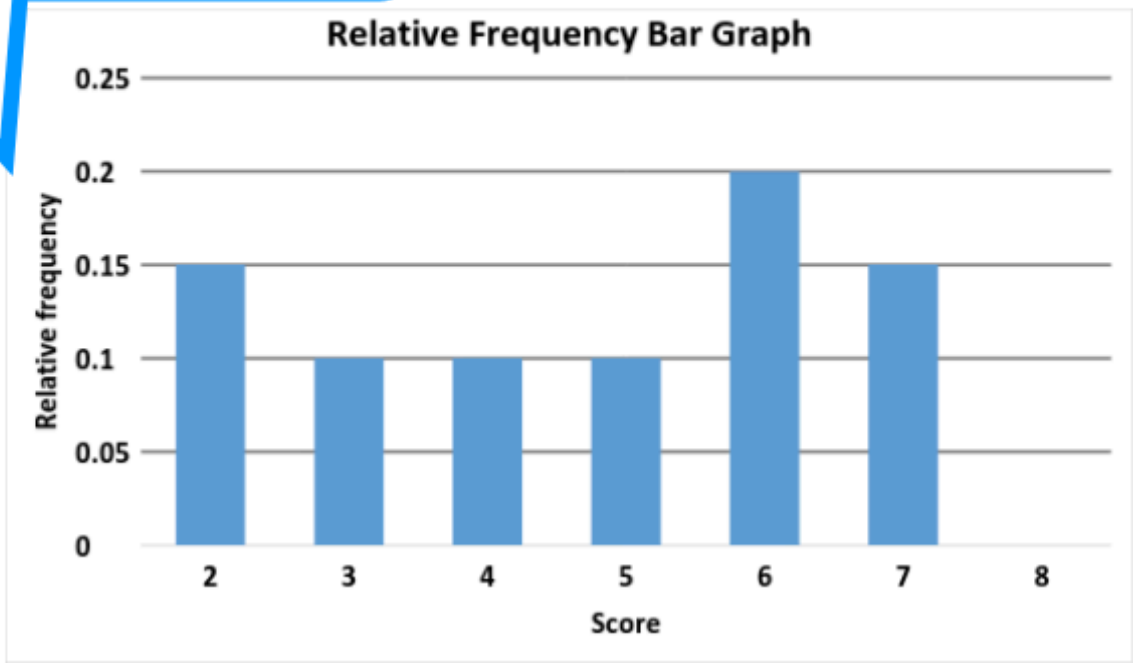
Score	Frequency	Relative Frequency
2	4	4 / 20 = 0.2
3	3	3 /20 = 0.15
4	2	2 / 20 = 0.1
5	2	2 / 20 = 0.1
6	2	2 / 20 = 0.1
7	4	4 / 20 = 0.2
8	3	3 / 20 = 0.15
	Total = 20	

Notice that the vertical axis is the **frequency**



Notice that the horizontal axis is the outcomes (scores)

Notice that the vertical axis is the **relative frequency**



Frequency Bar graphs in Excel

Here are the instructions on how to make a frequency bar graph.
The situation we will explore is selecting either an apple, pear or peach from a bowl which has 50 pieces of fruit in it.

Step 1:

Open up the Excel document with your data and summary table.

	A	B	C	D	E	F
1	Trial	Outcome		Sample space	Frequency	Relative Frequency
2	1	Apple		Apple	14	0.28
3	2	Peach		Peach	18	0.36
4	3	Pear		Pear	18	0.36
5	4	Apple				
6	5	Peach				

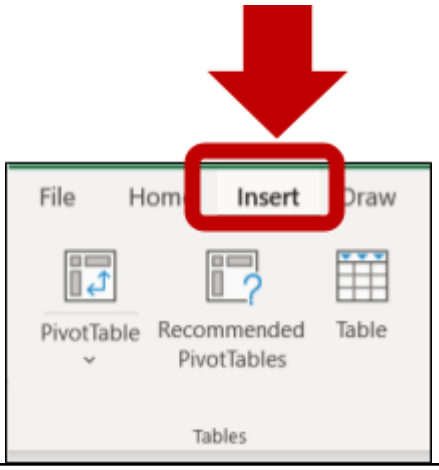
Step 2:

Use the summary table, select the **sample space column** and the **frequency column**.

	A	B	C	D	E	F	G
1	Trial	Outcome		Sample space	Frequency	Relative Frequency	
2	1	Apple		Apple	14	0.28	
3	2	Peach		Peach	18	0.36	
4	3	Pear		Pear	18	0.36	
5	4	Apple					
6	5	Peach					

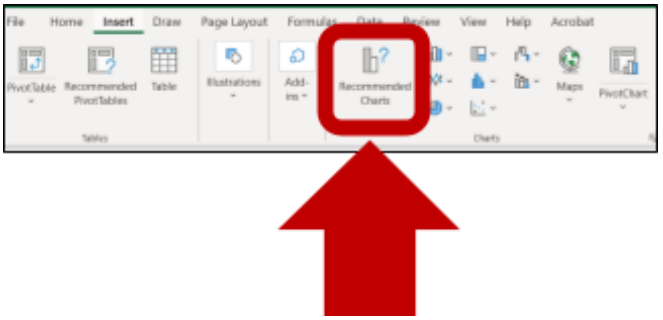
Step 3:

Go to the **Insert** menu



Step 4:

Click on the **Recommended Charts**

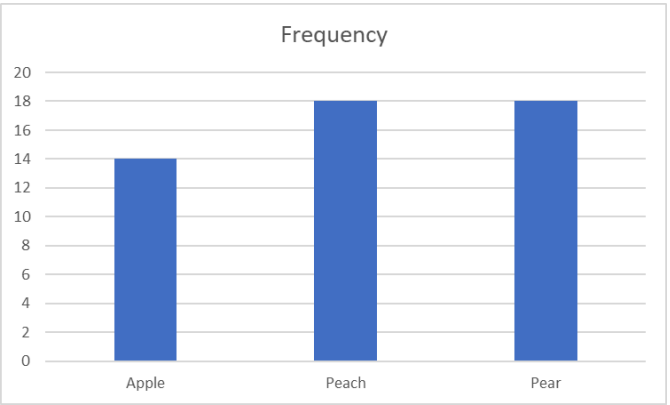


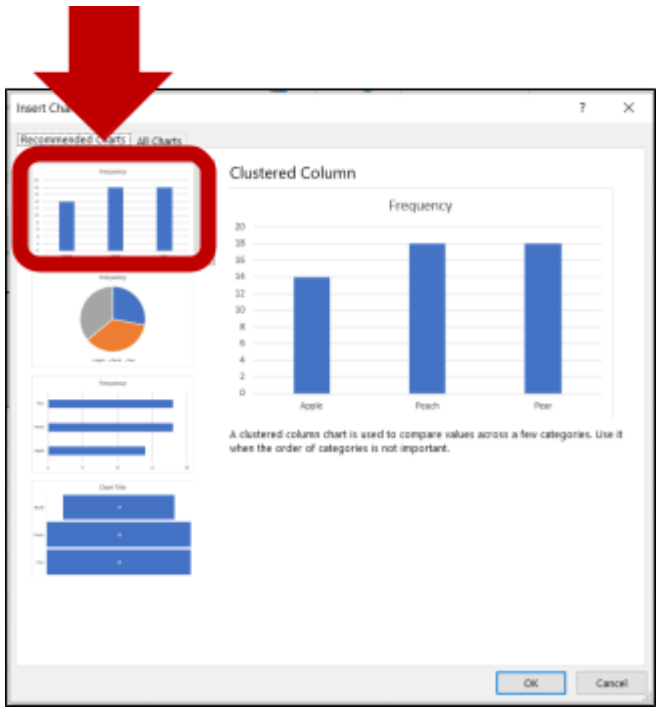
Step 5:

A window will pop up with several graphs to choose from. Select the **bar graph** and click the **OK** button.

Step 6:

This will insert a bar graph in your Excel spreadsheet.

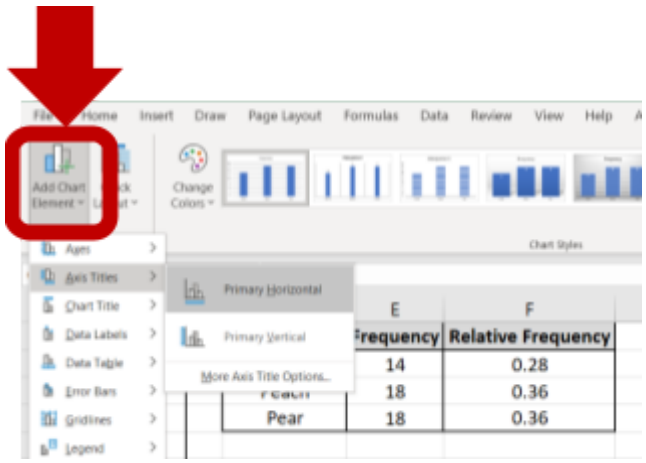




Excel Tip to add Labels and Titles

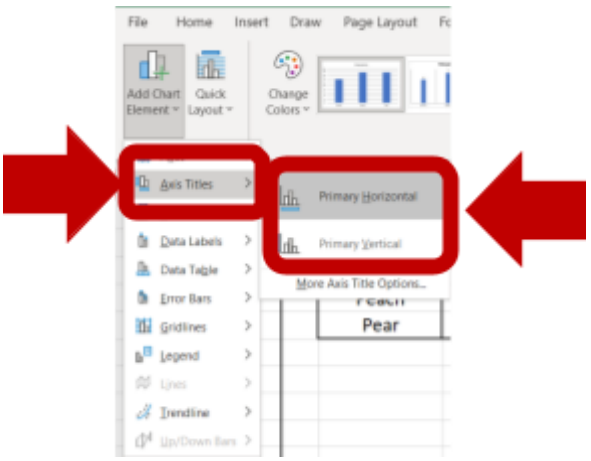
Step 1:

You can now add axis labels to your graph. Click on **Add Chart Element** button.



Step 2:

Then select **Axis Titles** and click on both the **Primary Horizontal** and the **Primary Vertical**.

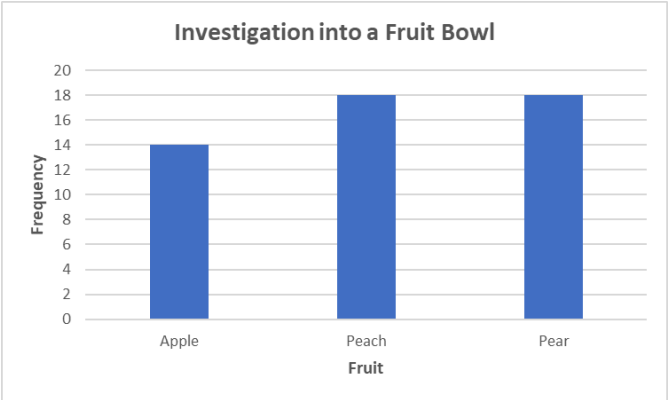


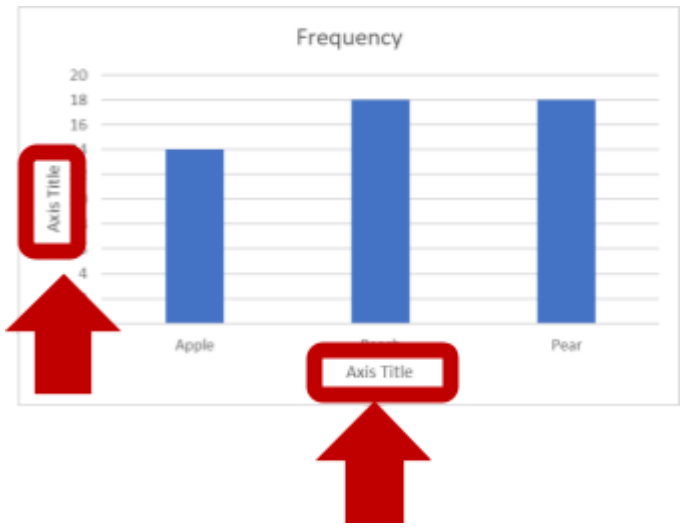
Step 3:

Then go to the graph and change the **Axis Titles**.

Step 4:

You can also change the **Title**.




Frequency
20
18
16
4
Axis Title
Apple
Pear
Axis Title

Relative Frequency Bar graphs in Excel

Here are the instructions on how to make a relative frequency bar graph.

The situation we will explore is selecting either an apple, pear or peach from a bowl which has 50 pieces of fruit in it.

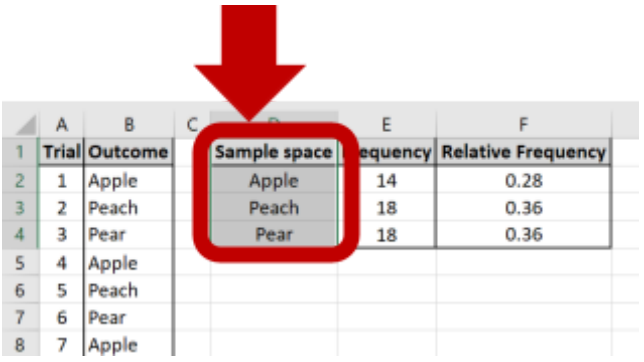
Step 1:

Open up the Excel document with your data and summary table.

	A	B	C	D	E	F
1	Trial	Outcome		Sample space	Frequency	Relative Frequency
2	1	Apple		Apple	14	0.28
3	2	Peach		Peach	18	0.36
4	3	Pear		Pear	18	0.36
5	4	Apple				
6	5	Peach				

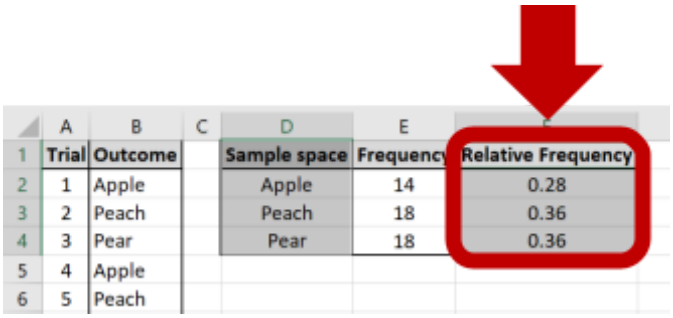
Step 2:

Use the summary table, select the **sample space column**.



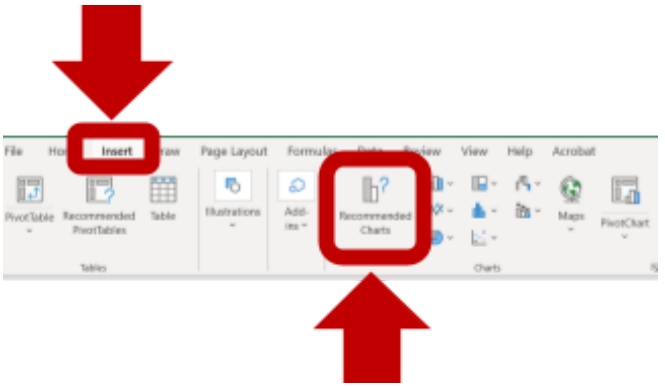
Step 3:

Press the Ctrl key on your keyboard, and then select the relative frequency column.



Step 4:

Go to the **Insert** menu and click on the **Recommended Charts**

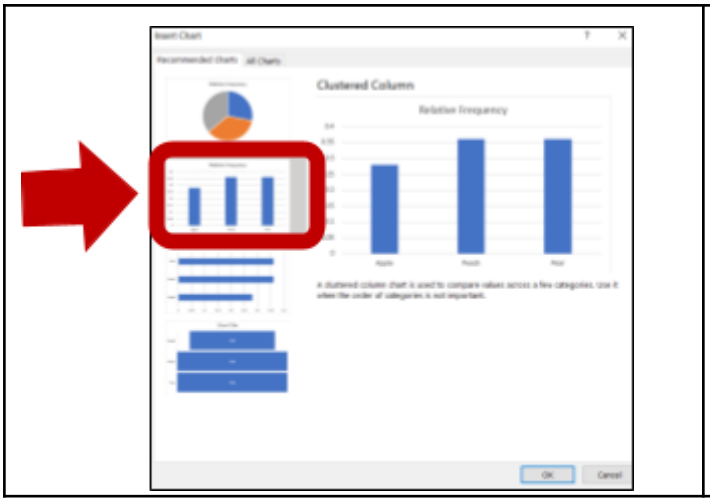


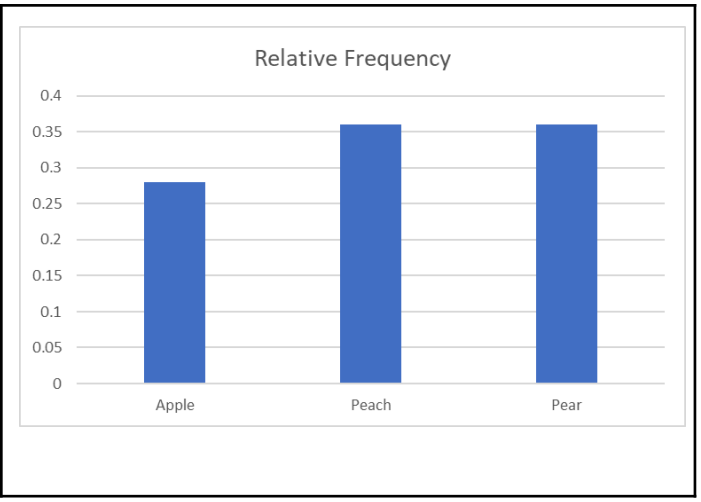
Step 5:

In the window that pops up, select the bar graph, and click the OK button.

Step 6:

This will insert a bar graph in your Excel spreadsheet. Add axis labels and a Title to the graph.





Exercise 8:

For both of the Excel spreadsheets that you made in Exercise 7, create both frequency and relative frequency graphs.

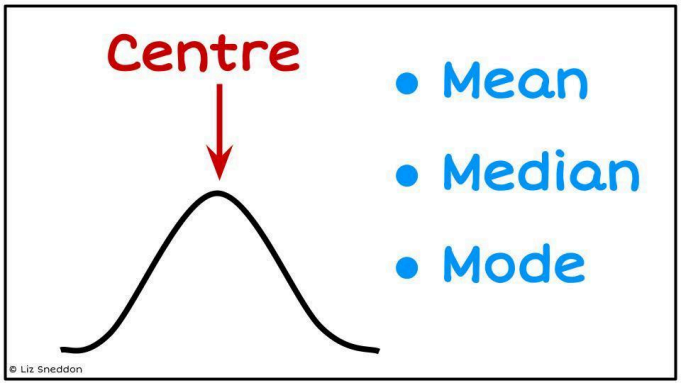
1) The test results obtained by a class of 30 students is given. The test was marked out of 10 points.

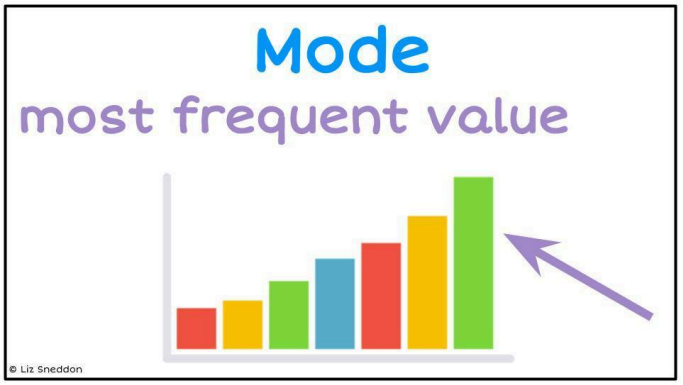
2) Twenty-five games of tic-tac-toe were played.

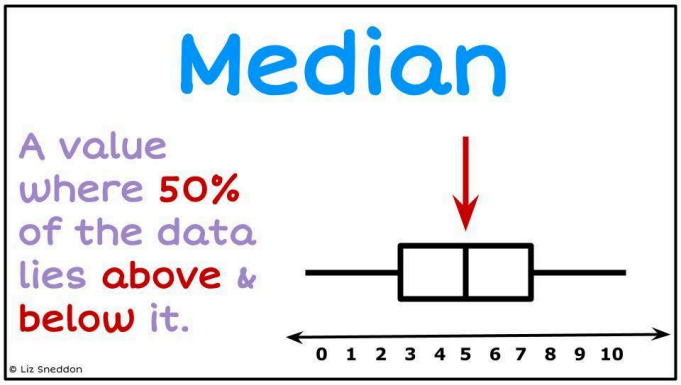
Analysis / Kaupapa

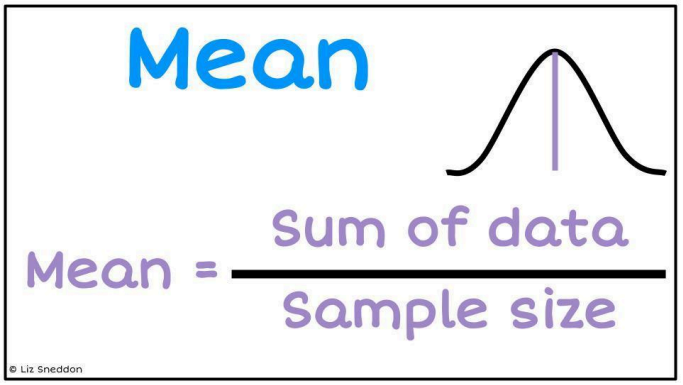
Once we have the data displayed in both tables and graphs, it is then time to look for patterns and discuss the features in the data. The features to look for include:

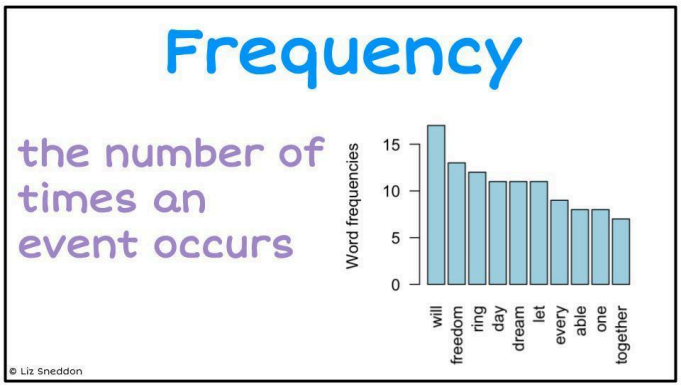
- Centre: mode, median, and/or mean
- Whole distribution: Frequencies or probabilities
- Any other patterns (including shape where appropriate).

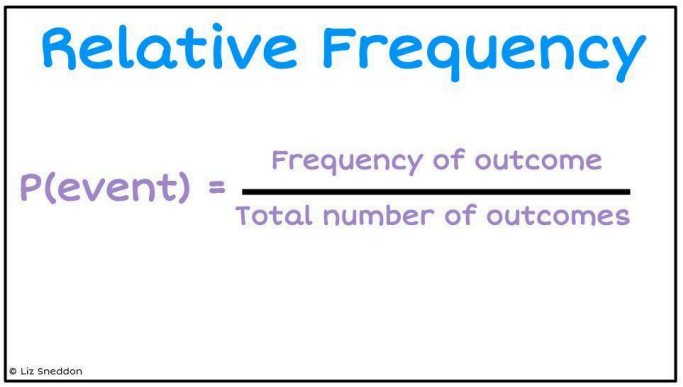










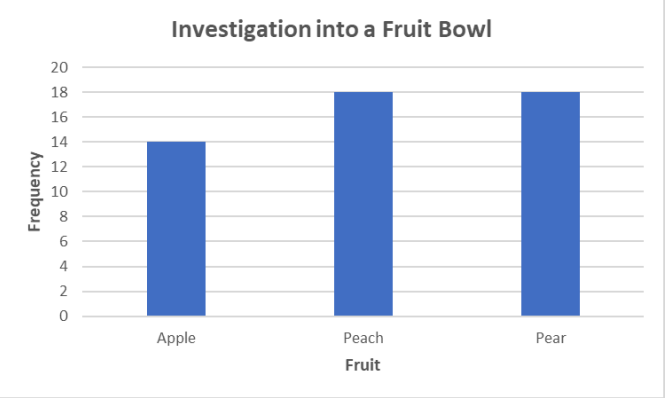


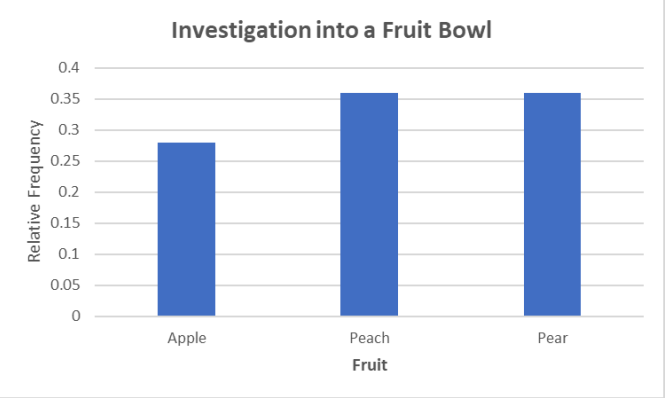
Remember that you can only calculate the mean and median if you have numeric data.

Example 1:

Below are the tables and graphs for the Fruit Bowl investigation.

Sample space	Frequency	Relative Frequency
Apple	14	0.28
Peach	18	0.36
Pear	18	0.36





Centre	The most common fruit is both pear and peach as they both have a relative frequency of 36%
The whole distribution: Frequencies or relative frequencies	The whole distribution could be discussed either by talking about the frequencies or the relative frequencies: • The frequencies of the three fruits are very similar. Pear has the lowest frequency of 14 while pear and peach both have frequencies of 18. • The probabilities of each of the fruits, apples (28%), peach (36%) and pear (36%) are all very similar which suggests that they are close to being equally likely.

Example 2:

Below are the summary table and graphs for an investigation into how many good nights sleeps Mrs Sneddon has each week.

Number of good sleeps	Frequency	Relative frequency
0	0	0
1	1	0.02
2	11	0.22
3	12	0.24
4	16	0.32
5	9	0.18
6	1	0.02
7	0	0

Mean = 3.48 Median = 4





Centre	It was most common for Mrs Sneddon to have 4 good night's sleep each week. The mean number of good nights sleep Mrs Sneddon gets per week is 3.48. The median number of good nights sleep per week is 4.
The whole distribution: Frequencies or relative frequencies	The whole distribution could be discussed either by talking about the frequencies or the relative frequencies: • Mrs Sneddon didn't have any weeks that she had zero or seven good night's sleep (frequency is 0). • The chance of a good night's sleep is lower at the ends than in the middle. For example, the chance of a 5 good nights sleep is 18%, which reduces to 2% for six nights sleep, and 0% for seven nights sleep. There is a similar pattern with zero (0%), one (2%) and two (22%) good night's sleep.
Other patterns	The shape of the distribution of the number of good nights sleep that Mrs Sneddon has each week is roughly a normal distribution

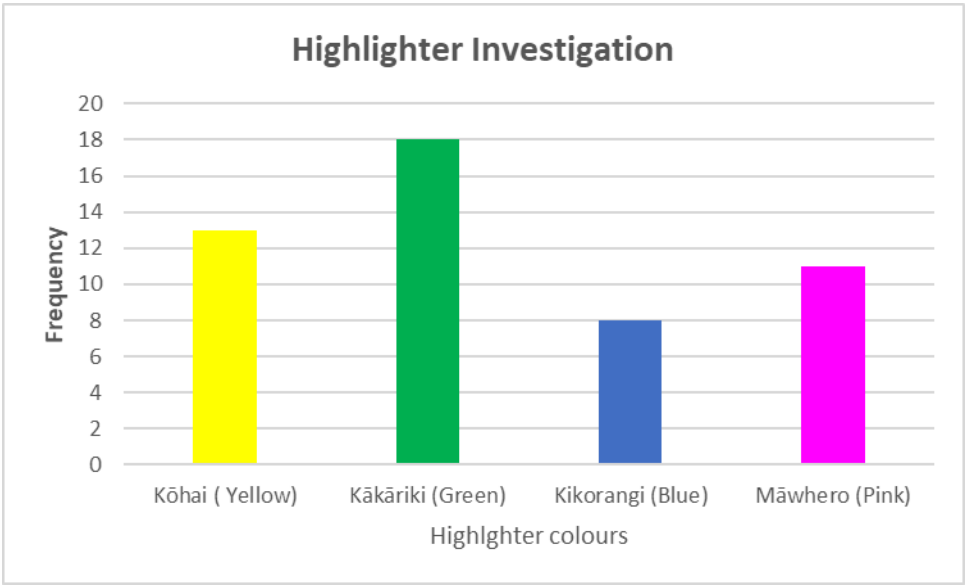
Exercise 9:

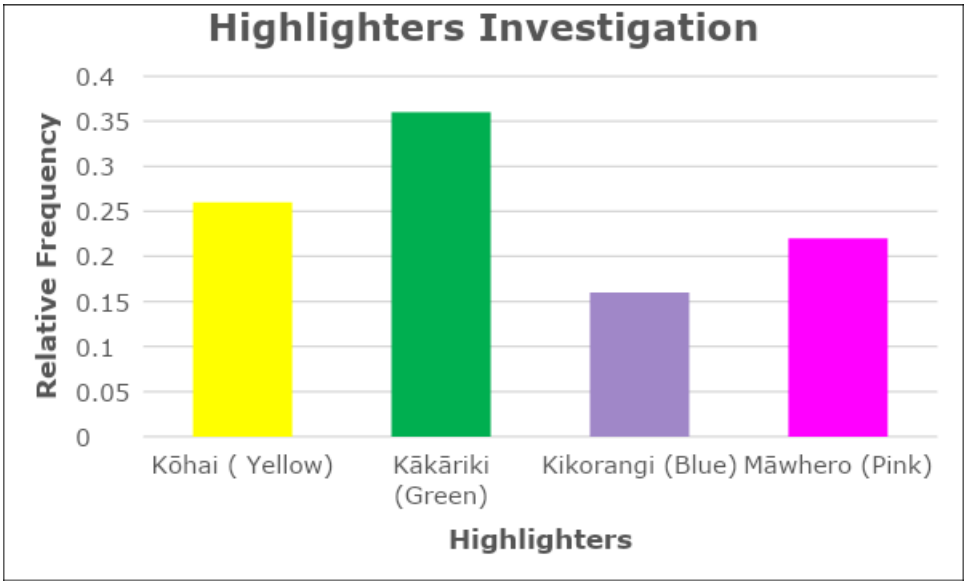
For each of the situations below, analyse any patterns in the tables and graphs.

1)

Situation	Mrs Sneddon has a kete (basket) with highlighters for students to use, including kōhai (yellow), kākāriki (green), Kikorangi (blue) and māwhero (pink).
Summary table and graphs	

Highlighters	Frequency	Relative frequency
Kōhai (Yellow)	13	0.26
Kākāriki (Green)	18	0.36
Kikorangi (Blue)	8	0.16
Māwhero (Pink)	11	0.22

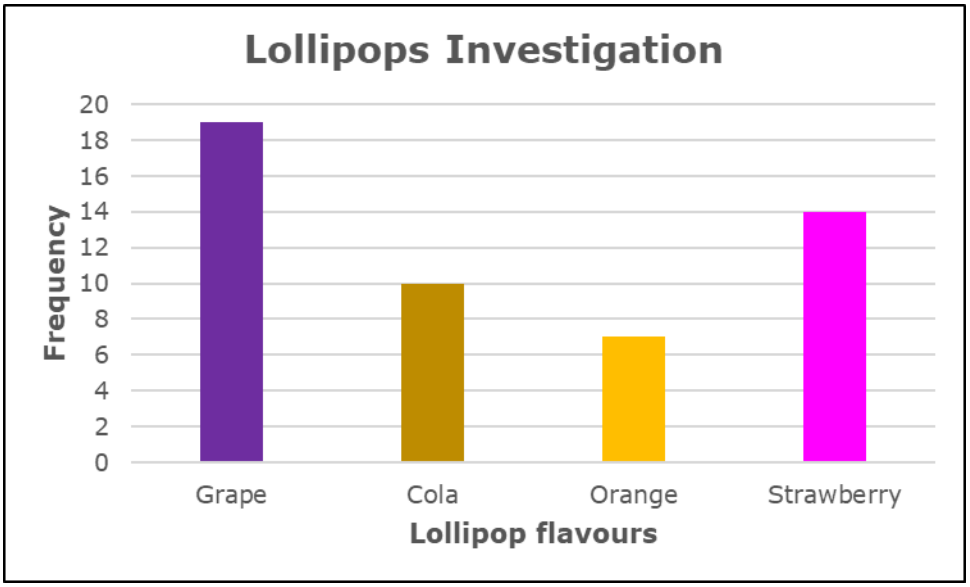


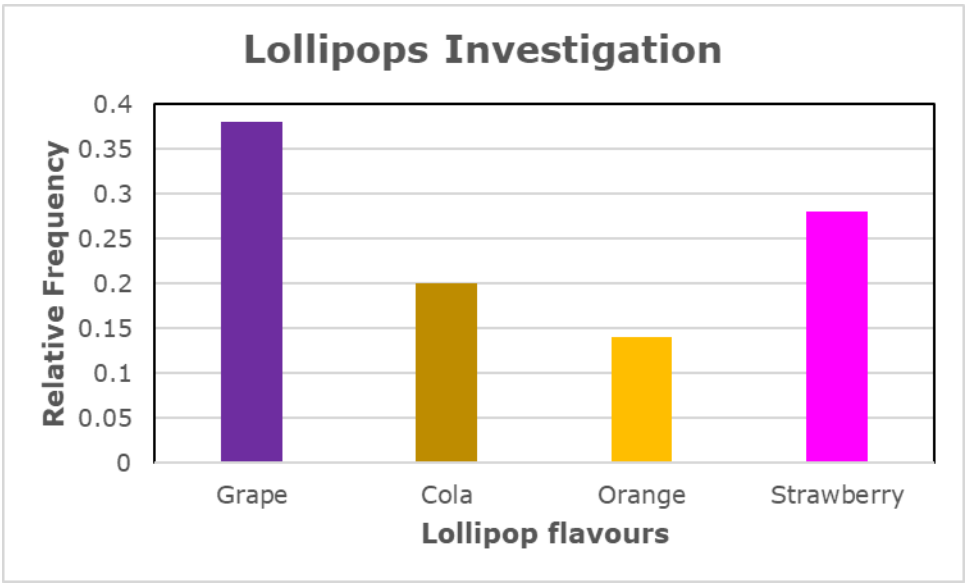
	
Centre	
The whole distribution : Frequencies or relative frequencies	

2)

Situation	A bag of lollipops has 4 different flavours, grape, cola, orange, and strawberry.
Summary table and graphs	

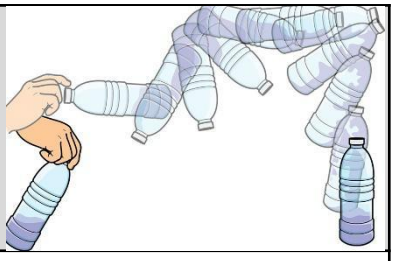
Lollipops	Frequency	Relative frequency
Grape	19	0.38
Cola	10	0.2
Orange	7	0.14
Strawberry	14	0.28




Lollipops Investigation
Relative Frequency
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Grape
Cola
Orange
Strawberry
Lollipop flavours

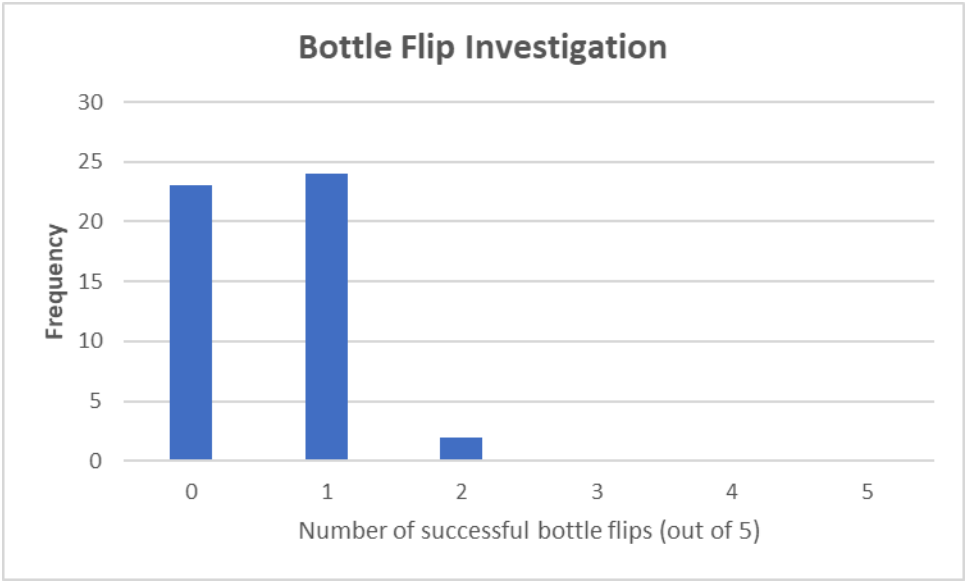
Centre	
The whole distribution : Frequencies or relative frequencies	

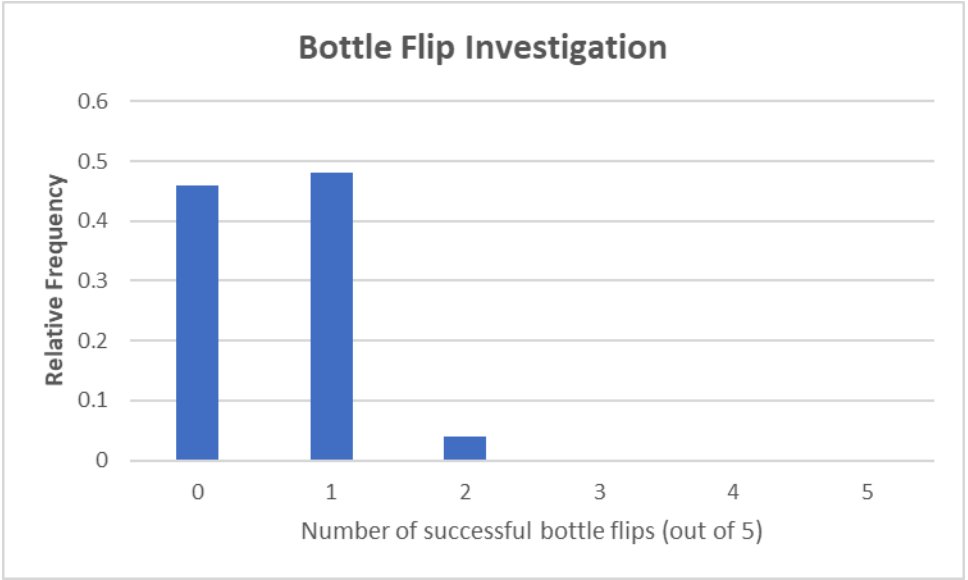
3)

Situation	Doing a bottle flip challenge where each person tries to flip a bottle 5 times to see how often it lands standing up. 
Summary table and graphs	

Sample space	Frequency	Relative Frequency
0	23	0.46
1	24	0.48
2	2	0.04
3	0	0
4	0	0
5	0	0

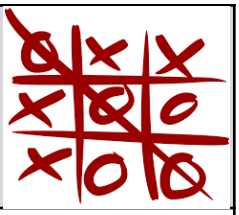
Mean = 0.58 Median = 1



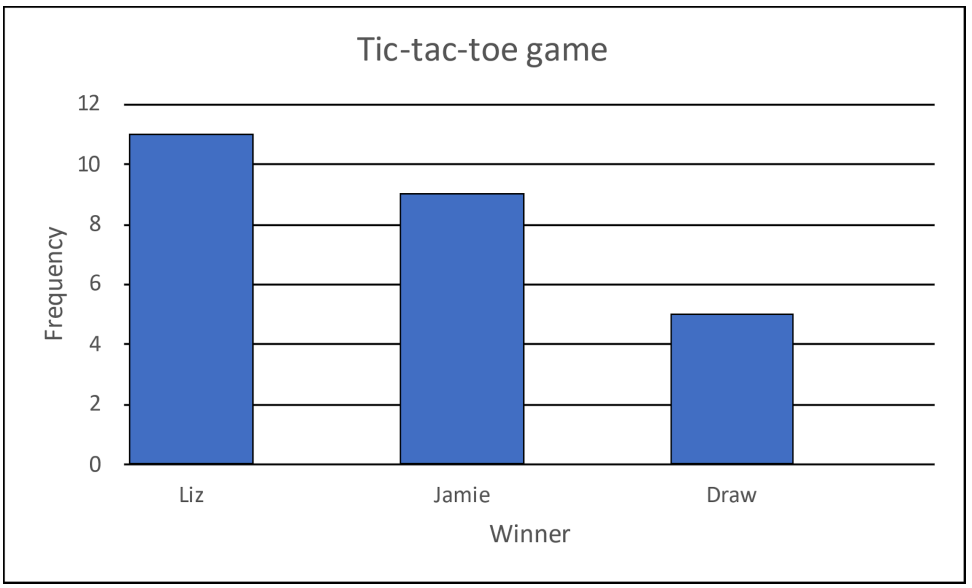


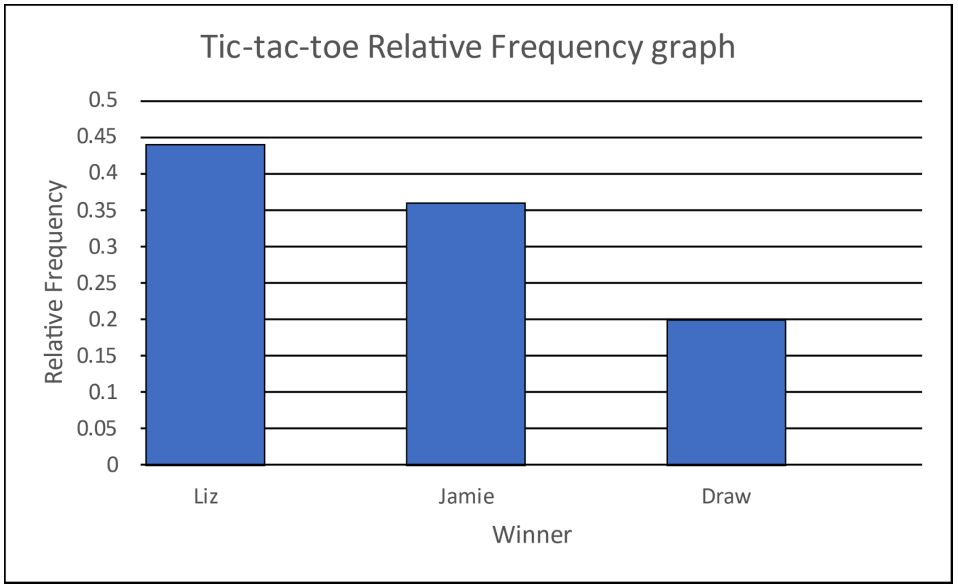
Centre	
The whole distribution : Frequencies or relative frequencies	
Other patterns	

4)

Situation	Two students want to see who is better at playing tic tac toe. 
Summary table and graphs	

Winner	Frequency	Relative Frequency
Liz	11	0.44
Jamie	9	0.36
Draw	5	0.2





Centre	
The whole distribution : Frequencies or relative frequencies	

Conclusion / Whakatau

In the conclusion you need to do the following:

- Answer the investigation question,
- Reflect on the investigation, consider assumptions, limitations or long run relative frequency.

Answer the investigation question

Once you have done your investigation, collected your data, summarised it in tables and graphs, you then need to answer your investigation question, remembering to give the evidence (probability) at the same time, and to put your sentence in context.

Example:

Below are the summary table and graphs for an investigation into how many good nights sleeps Mrs Sneddon has each week.

Number of good sleeps	Frequency	Relative frequency
0	0	0
1	1	0.02
2	11	0.22
3	12	0.24
4	16	0.32
5	9	0.18
6	1	0.02
7	0	0

Mean = 3.48 Median = 4





Question	Answer
What is the **probability** that Mrs Sneddon has at least 5 good night sleeps each week?	The probability that Mrs Sneddon has at least 5 good nights sleeps each week is 20%.

	(0.18+0.02=0.2)

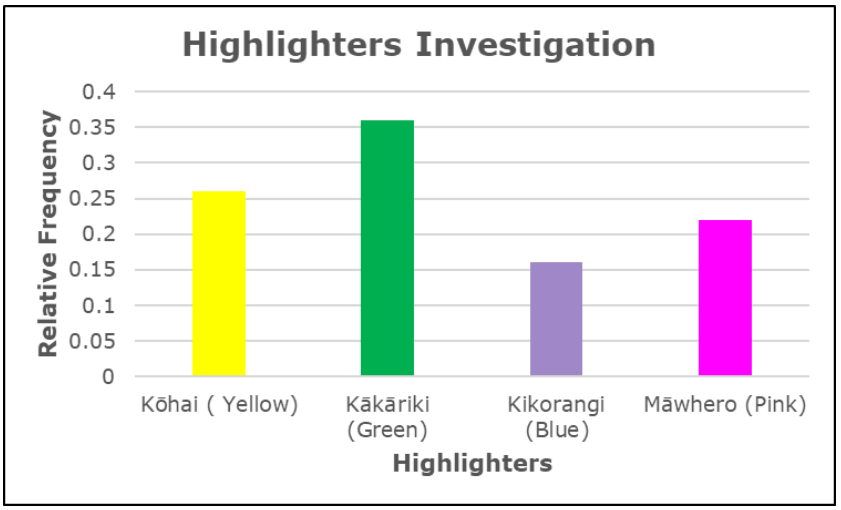
Exercise 10:

Answer the investigation question(s) for each of the situations below.

1)

Investigation question	Which colour has the highest chance of being selected at random?
Summary table and graphs	
Answer	

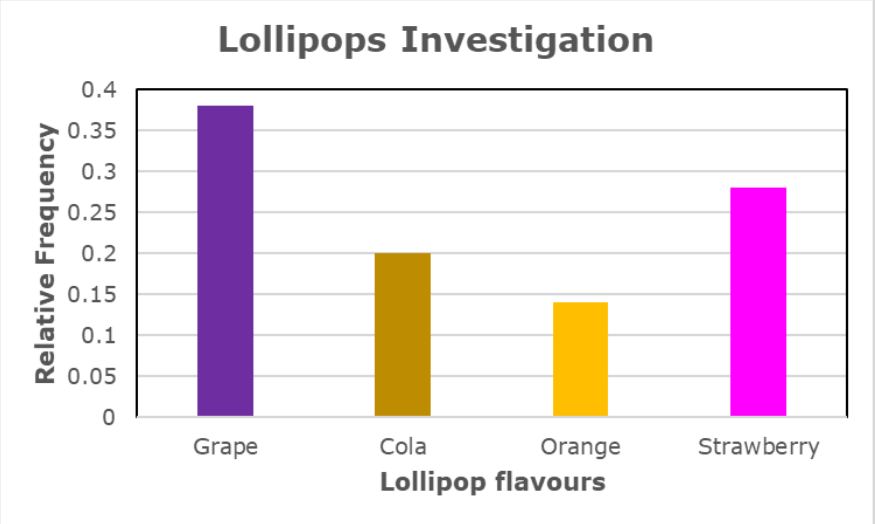
Highlighters	Frequency	Relative frequency
Kōhai (Yellow)	13	0.26
Kākāriki (Green)	18	0.36
Kikorangi (Blue)	8	0.16
Māwhero (Pink)	11	0.22



2)

Investigation question	Which flavour of lollipop is the least likely to be chosen?
Summary table and graphs	
Answer	

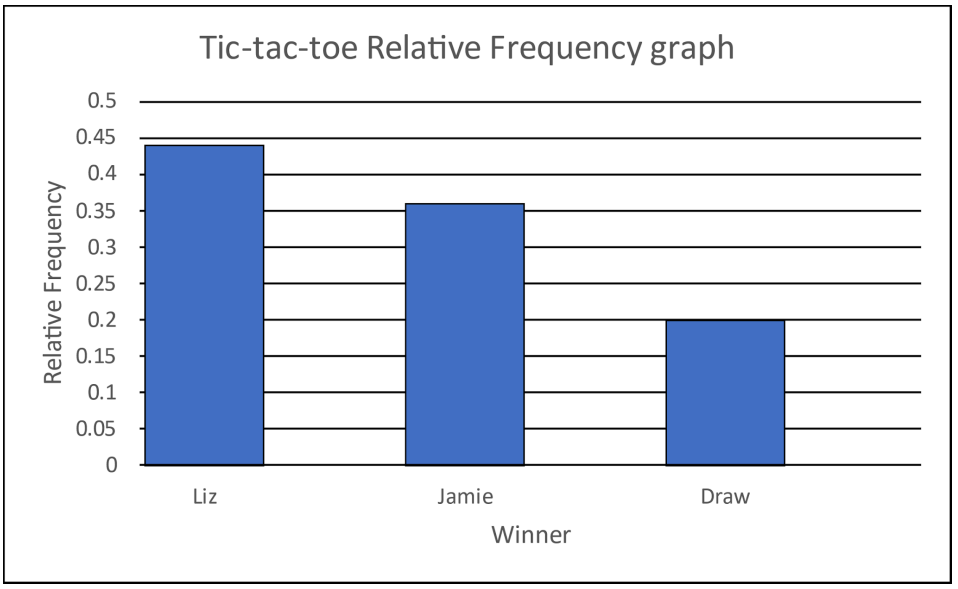
Lollipops	Frequency	Relative frequency
Grape	19	0.38
Cola	10	0.2
Orange	7	0.14
Strawberry	14	0.28



3)

Investigation question	What is the probability of a draw in tic-tac-toe game?
Summary table and graphs	

Winner	Frequency	Relative Frequency
Liz	11	0.44
Jamie	9	0.36
Draw	5	0.2

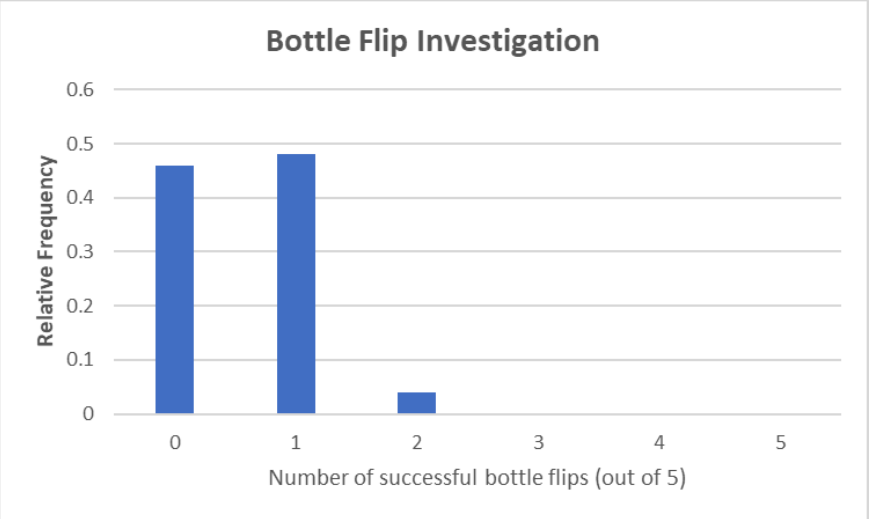


Answer

4)

Investigation question	What is the chance of getting zero bottle flips standing up out of a set of 5 bottle flips?
Summary table and graphs	
Answer	

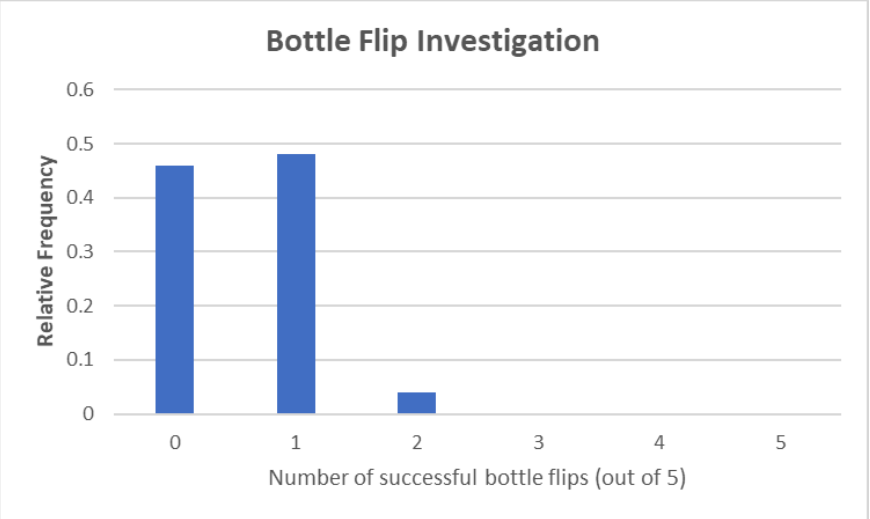
Bottle flips	Frequency	Relative Frequency
0	23	0.46
1	24	0.48
2	2	0.04
3	0	0
4	0	0
5	0	0



5)

Investigation question	What is the chance of getting at least one bottle flip out of a set of 5 bottle flips?
Summary table and graphs	

Bottle flips	Frequency	Relative Frequency
0	23	0.46
1	24	0.48
2	2	0.04
3	0	0
4	0	0
5	0	0

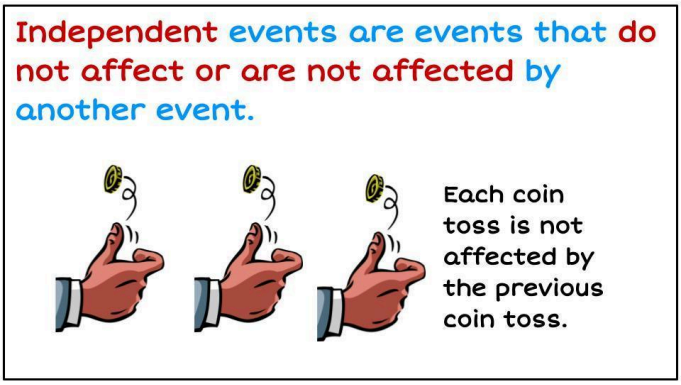


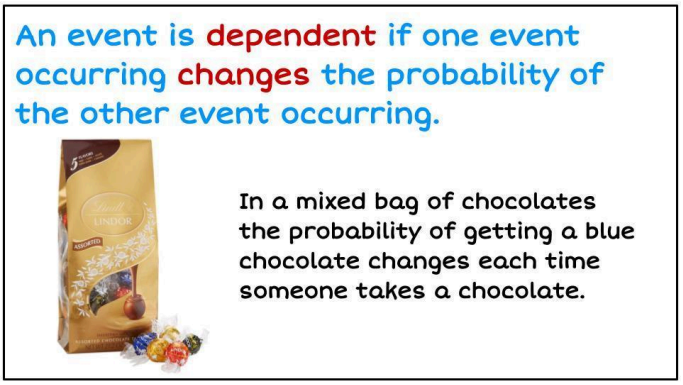
Answer

Assumptions

When we do an experimental probability investigation, we need to think about the situation and what assumptions we make, or what limitations there might be. One common assumption that you can think about is:

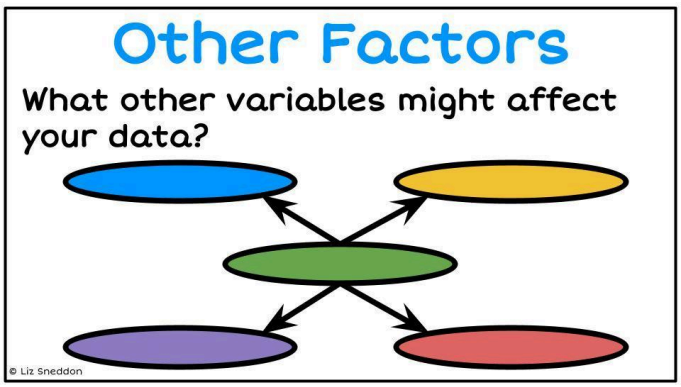
Is each event independent?





Other Factors

In every real-world situation (e.g., not a theoretical situation such as tossing a coin) there are other factors or variables that affect the probability of an event happening.



For example, the chance of selling a house depends on the area the house is in, the size of the house, the number of bedrooms, etc. If the house is in an area of high

demand, there is a greater chance of selling the house than if it is in an area of low demand.

Example 1:

Situation: Selecting either an apple, pear or peach from a bowl which has 50 pieces of fruit in it.

Assumptions:

I assume that selecting one piece of fruit is independent of selecting the next piece of fruit. This might be true for my experiment, but it depends on why people are selecting fruit. For example, if I choose a piece of fruit based on which is on the top, then if the peaches all happen to be together on the top, then the I won't select fruit independently.

Other factors:

One limitation is that this experiment can only work if the fruit stays fresh and doesn't get bruised and rotten. If the fruit was left in the bowl for more than a few days, some of it may go rotten, meaning that the probability of choosing a specific fruit changes. If I had to throw out several pears because they were bruised and rotten, then the chance of choosing a pear would go down.

Another limitation is which season it is during the year, as during the year it will be easier or harder to get apples, pears, and peaches, depending on whether they are available in NZ, or whether they come from overseas. It also depends on the price as to whether I'll buy them.

Example 2:

Situation: Mrs Sneddon wants to explore how many good night's sleep she gets each week.

Assumptions:

I assume that having a good sleep one night is independent of the next night. But if Mrs Sneddon gets sick with the flu for example, then she is more likely to not have a good night's sleep, or if it is the school holidays or weekends, she might be more likely to have a good night's sleep.

Other factors:

One factor that might affect whether Mrs Sneddon gets a good night sleep is how old her children are. For example, if her children are very young and wake up during the night, she isn't likely to have a good night's sleep. Whereas if her children are older, she is more likely to have a good night's sleep.

Another factor that may affect having a good night's sleep, is whether Mrs Sneddon spends these nights sleeping in her normal bed, or whether she goes on holiday or stays with family and sleeps in another bed with a different mattress. The other bed could be more or less comfortable, leading to having a better or worse night's sleep.

Also, depending on Mrs Sneddon's job or other potential stressors in her life (e.g., sick parents, etc.), this may affect whether she has a good night's sleep or not.

Exercise 11:

For each of the situations below, identify and describe any assumptions, limitations, and other factors.

1)

Situation	Mrs Sneddon has a kete (basket) with highlighters for students to use, including kōhai (yellow), kākāriki (green), Kikorangi (blue) and māwhero (pink).
Assumptions and other factors	

2)

Situation	A bag of lollipops has 4 different flavours, grape, cola, orange, and strawberry.
Assumptions and other factors	

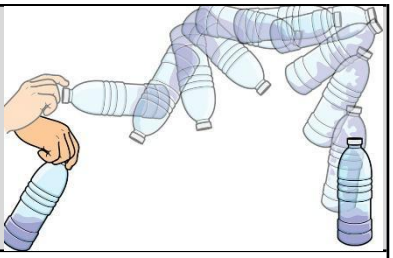
3)

Situation	When playing basketball and doing a penalty shootout, the player gets 3 shots at goal.
Assumptions and other factors	

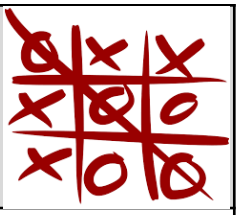
4)

Situation	Apps will often send out notifications when a new message or post is made. Explore notifications received in a one-hour time period.
Assumptions and other factors	

5)

Situation	Doing a bottle flip challenge where each person tries to flip a bottle 5 times to see how often it lands standing up. 
Assumptions and other factors	

6)

Situation	Two students want to see who is better at playing tic tac toe. 
Assumptions and other factors	

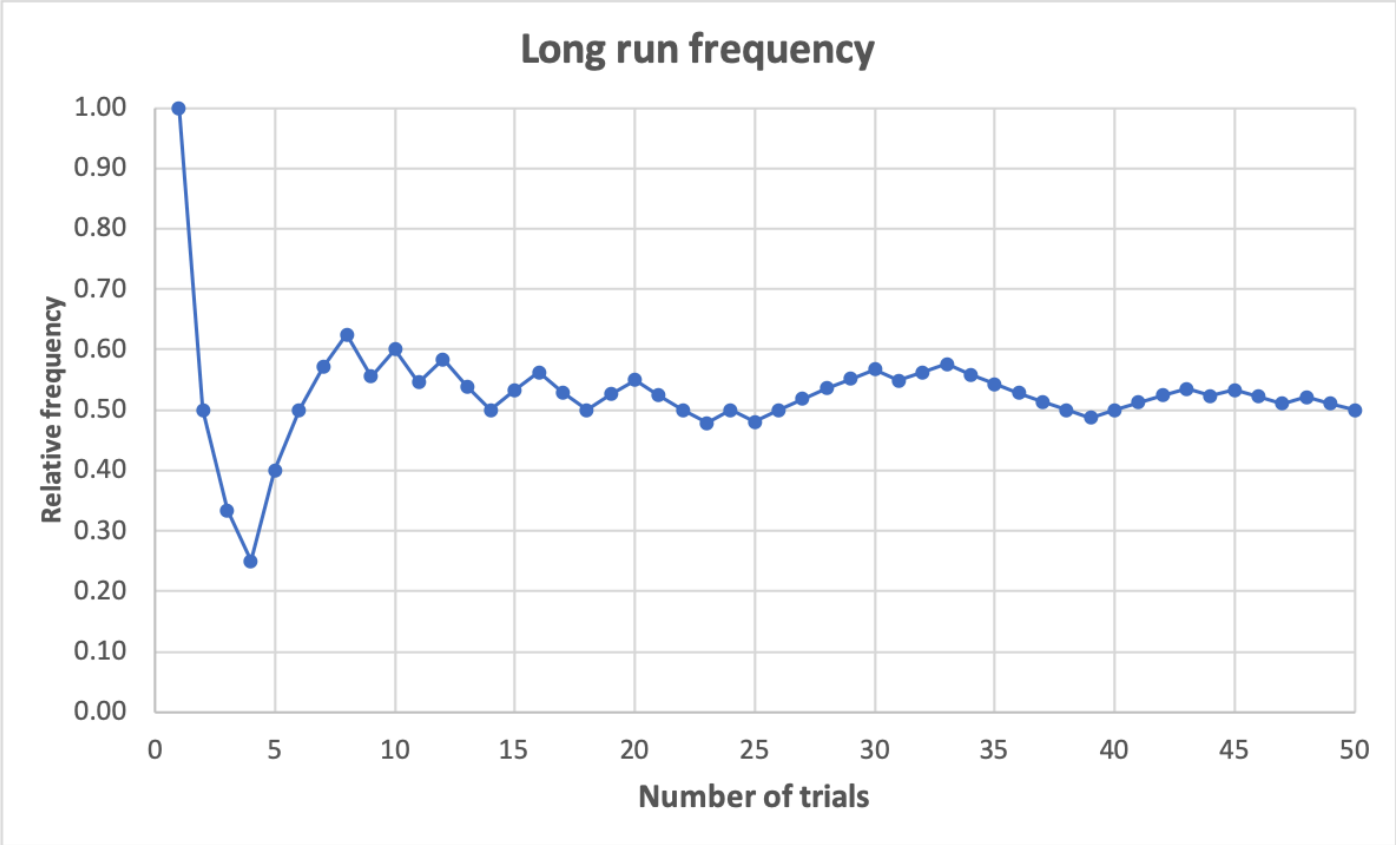
Long run relative frequency

This looks at how the frequency changes as you run more and more simulations. The first few times you run a simulation the probability varies, but as we include more and more data into our estimate of the probability, the accuracy of this increases, and the estimated probability gets closer and closer to the true probability.

$$P(Event) = \frac{Number\ of\ events}{sample\ size\ (or\ number\ of\ trials)}$$

Example:

Here is a long-run frequency graph for tossing a coin and recording the number of heads.



Notice that as the number of trials (our sample size) gets bigger and bigger, the probability of a head gets closer and closer to 0.5 or 50%.

Example:

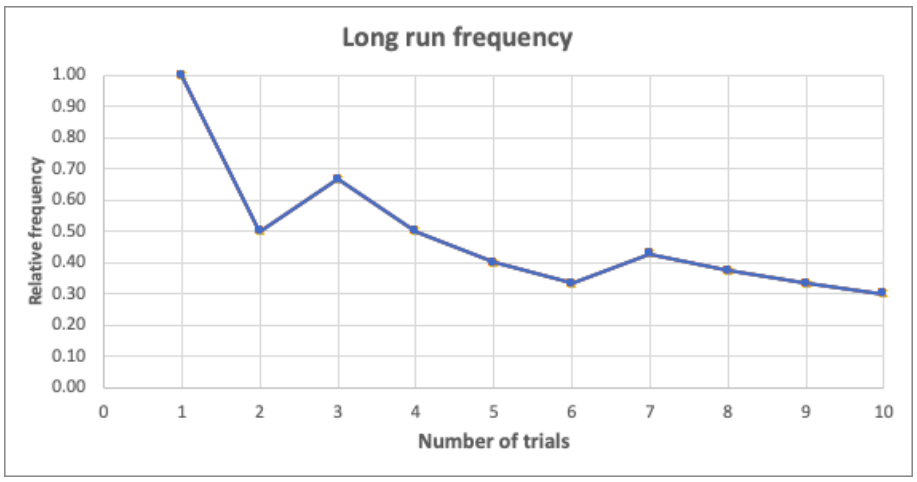
We want to toss a coin and find the long-run frequency of getting a head.

$$P(Heads) = \frac{Number\ of\ heads}{Number\ of\ times\ the\ coin\ tossed}$$

Flip	Result	Heads so far	P(Head)
1	T	0	$\frac{0}{1} = 0.00$
2	H	1	$\frac{1}{2} = 0.50$
3	T	1	$\frac{1}{3} = 0.33$
4	T	1	$\frac{1}{4} = 0.25$
5	H	2	$\frac{2}{5} = 0.40$
6	T	2	$\frac{2}{6} = 0.33$
7	T	2	$\frac{2}{7} = 0.29$
8	T	2	$\frac{2}{8} = 0.25$
9	H	3	$\frac{3}{9} = 0.33$
10	T	3	$\frac{3}{10} = 0.30$
11	H	4	$\frac{4}{11} = 0.36$
12	H	5	$\frac{5}{12} = 0.42$
13	T	5	$\frac{5}{13} = 0.38$
14	H	6	$\frac{6}{14} = 0.43$
15	T	6	$\frac{6}{15} = 0.40$
16	H	7	$\frac{7}{16} = 0.44$

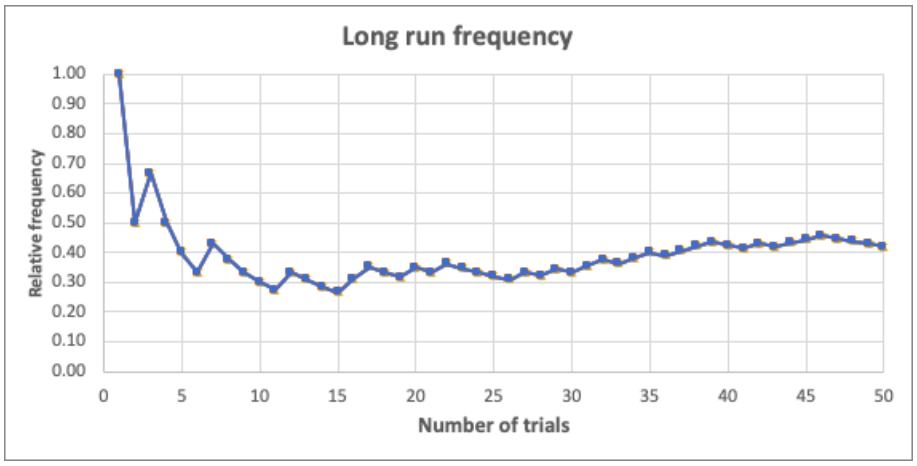
Here is a graph of the first 10 flips.

Notice that the relative frequency is on the y-axis, and the number of flips is on the x-axis. Also notice that there is a lot of variation in the probability.



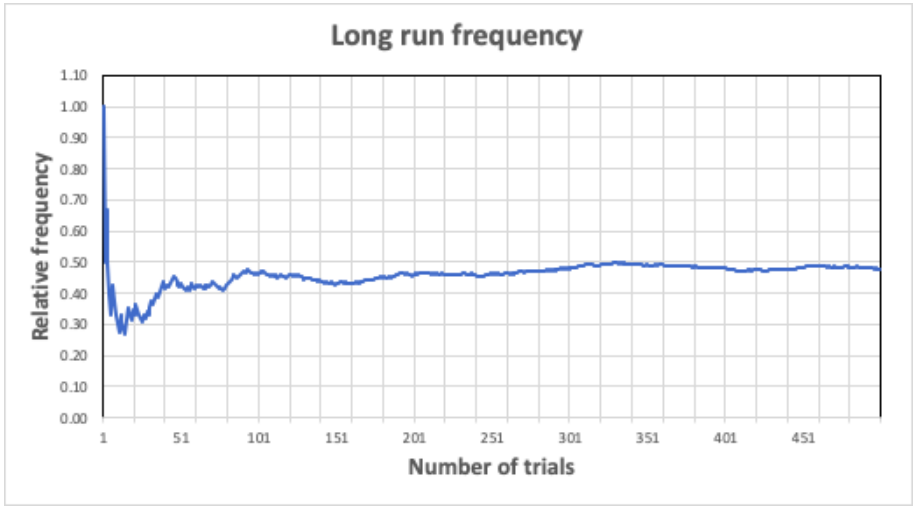
50 trials.

Notice how the probability is starting to flatten out.



500 trials.

Notice how the probability is getting quite stable now?



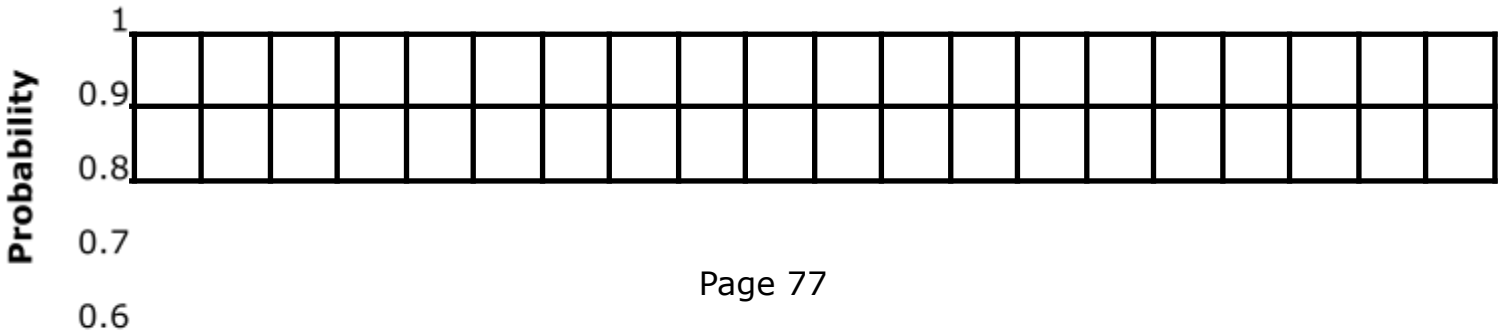
The bigger the sample size (or number of trials), the more accurate our probability estimate is.

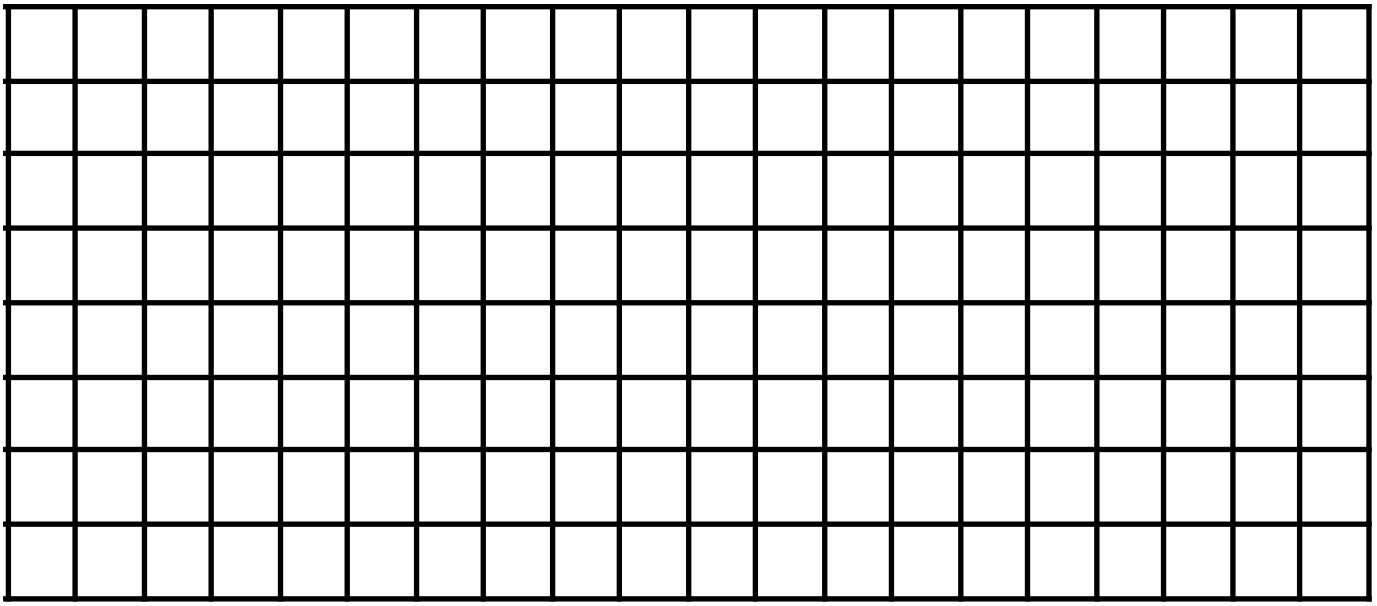
Exercise 12:

We want to estimate the long-run-probability of getting a **1** on a die.

Roll a die 20 times, calculating the long run relative frequency for a **1**. (Click here for an online dice if you do not have one).

Roll	Result	No. of 1's so far	P(1)
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			





Long-run relative frequency graphs in Excel

Here are the instructions on how to make a long-run relative frequency line graph.

Situation: Selecting either an apple, pear or peach from a bowl which has 50 pieces of fruit in it. We are going to find the long-run-relative frequency of getting apples out of the fruit bowl.

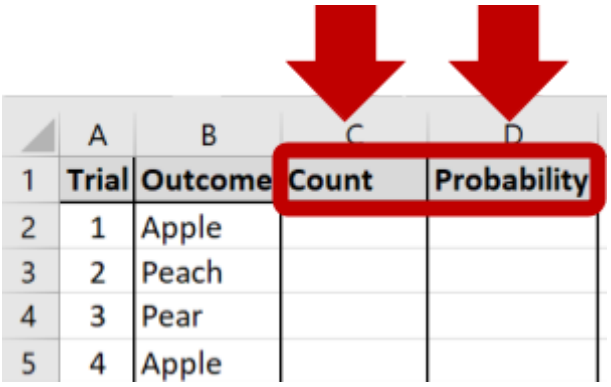
Step 1:

Open up the Excel document with your data and summary table.

	A	B	C	D	E	F
1	Trial	Outcome		Sample space	Frequency	Relative Frequency
2	1	Apple		Apple	14	0.28
3	2	Peach		Peach	18	0.36
4	3	Pear		Pear	18	0.36
5	4	Apple				
6	5	Peach				

Step 2:

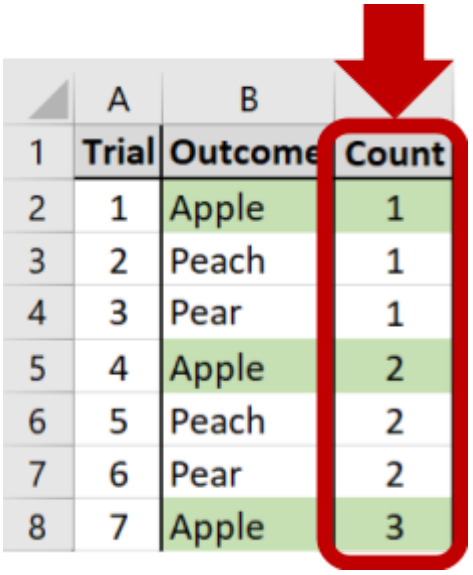
Create two new columns beside the Outcome column, and label them **Count** and **Probability.**



	A	B	C	D
1	Trial	Outcome	Count	Probability
2	1	Apple		
3	2	Peach		
4	3	Pear		
5	4	Apple		

Step 3:

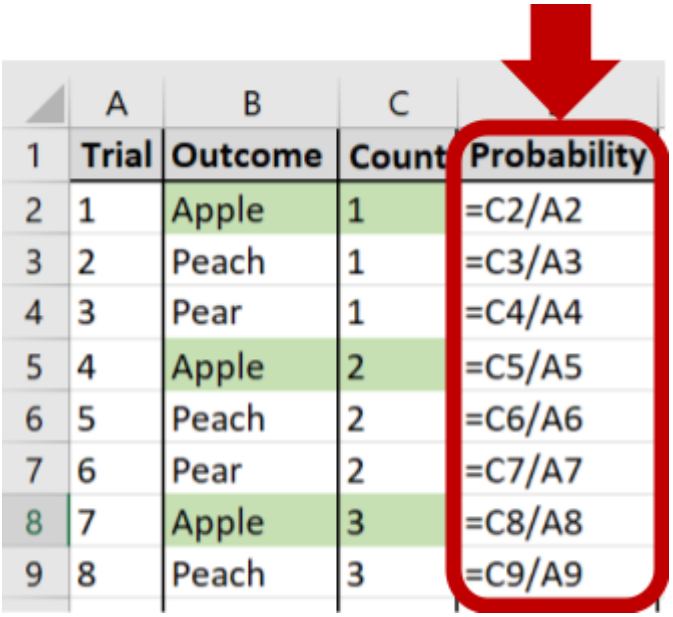
In the **Count** column, enter data for the number of apples.



	A	B	
1	Trial	Outcome	Count
2	1	Apple	1
3	2	Peach	1
4	3	Pear	1
5	4	Apple	2
6	5	Peach	2
7	6	Pear	2
8	7	Apple	3

Step 4:

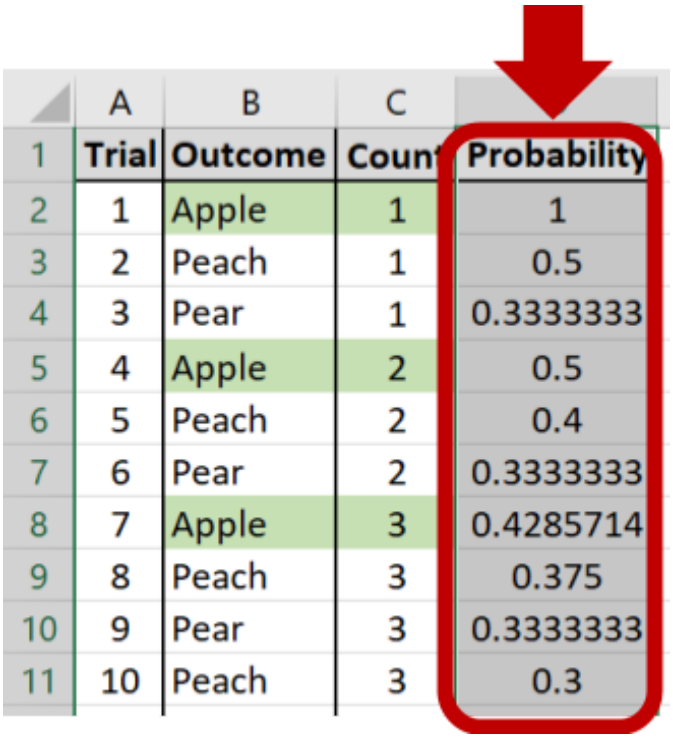
In the **Probability** column, you want to enter a formula which divides the **Count** by the **Probability**.



	A	B	C	
1	Trial	Outcome	Count	Probability
2	1	Apple	1	=C2/A2
3	2	Peach	1	=C3/A3
4	3	Pear	1	=C4/A4
5	4	Apple	2	=C5/A5
6	5	Peach	2	=C6/A6
7	6	Pear	2	=C7/A7
8	7	Apple	3	=C8/A8
9	8	Peach	3	=C9/A9

Step 6:

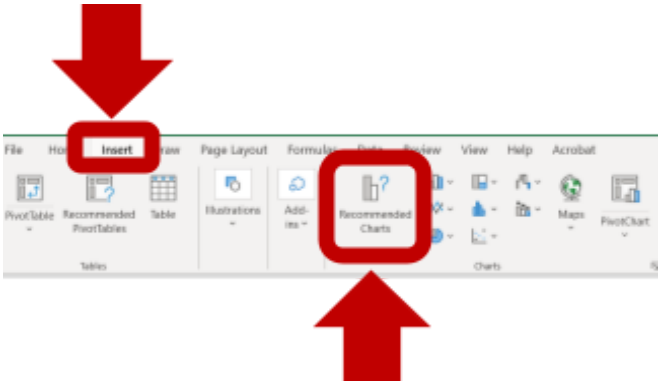
Use the data table, select the **Probability column**.



	A	B	C	
1	Trial	Outcome	Count	Probability
2	1	Apple	1	1
3	2	Peach	1	0.5
4	3	Pear	1	0.3333333
5	4	Apple	2	0.5
6	5	Peach	2	0.4
7	6	Pear	2	0.3333333
8	7	Apple	3	0.4285714
9	8	Peach	3	0.375
10	9	Pear	3	0.3333333
11	10	Peach	3	0.3

Step 7:

Go to the **Insert** menu and click on the **Recommended Charts**



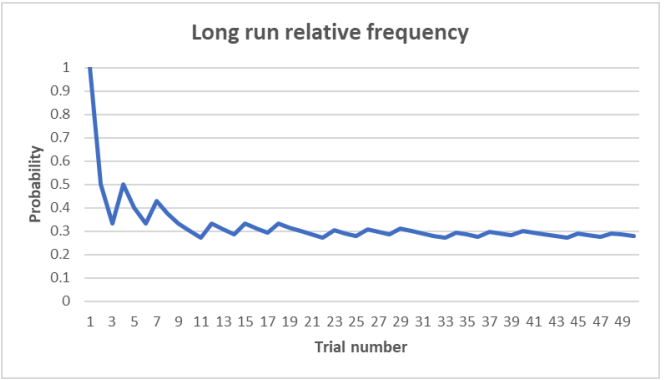
Step 8:

In the window that pops up, select the **line graph**, and click the **OK** button.



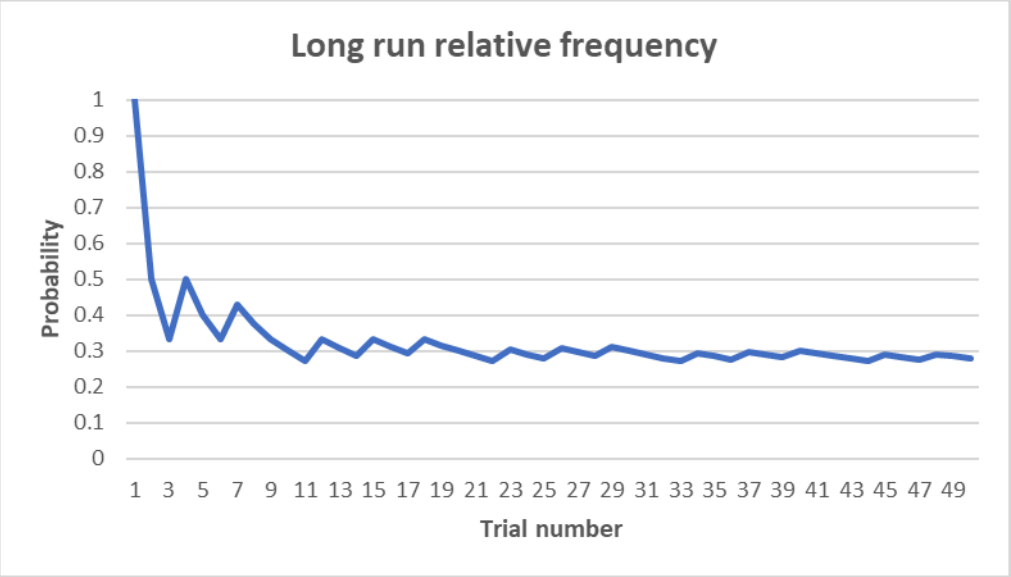
Step 6:

This will insert a line graph in your Excel spreadsheet. Add Axis Labels and Titles.



Example:

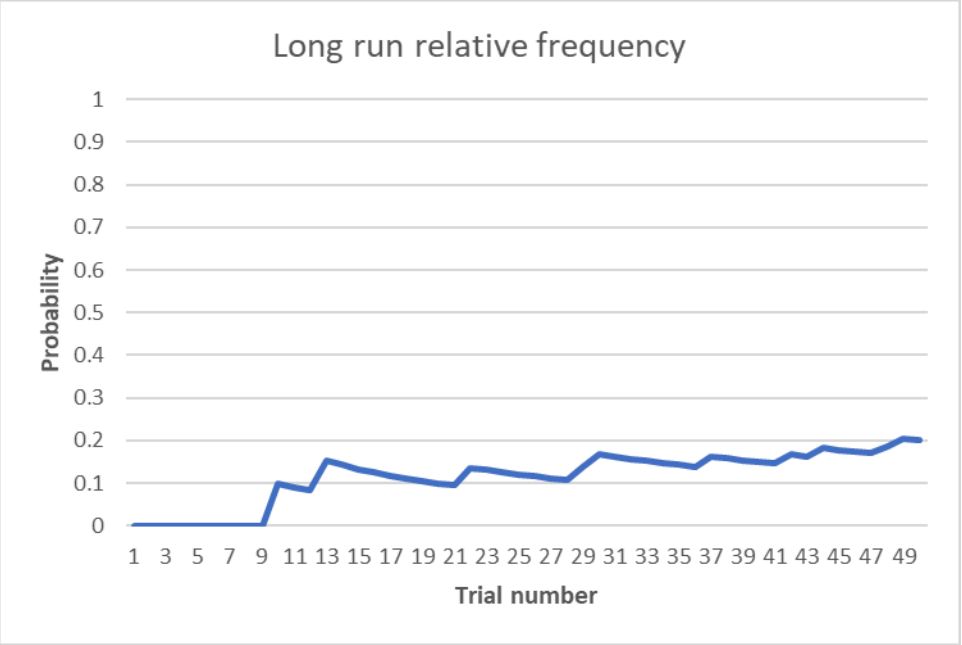
Situation: Selecting either an apple, pear or peach from a bowl which has 50 pieces of fruit in it. We want to explore the long-run-relative frequency of choosing an apple out of the fruit bowl.



Interpretation:
The probability of getting an apple out of the fruit bowl is heading towards 30%.

Example 2:

Situation: Mrs Sneddon wants to explore how many good night's sleep she gets each week. We want to explore the long-run-relative frequency of getting at least 5 night's sleep each week.



Interpretation:

The probability that Mrs Sneddon gets at least 5 good nights sleep each week is heading towards 20%.

Exercise 13:

Here is a link to a Google Spreadsheet with the data. **bit.ly/LRRFExercise**
Open Excel and copy the data from the Google spreadsheet into your Excel file and make a long run relative frequency graph. Then write your interpretation below.

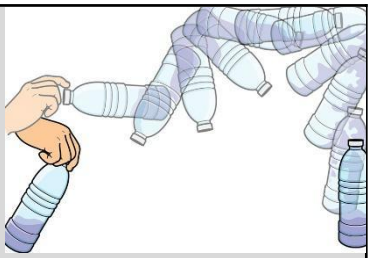
1)

Situation	Mrs Sneddon has a kete (basket) with highlighters for students to use, including kōhai (yellow), kākāriki (green), Kikorangi (blue) and māwhero (pink). Find the long run relative frequency for getting a kōhai (yellow) highlighter.
Interpretation	

2)

Situation	A bag of lollipops has 4 different flavours, grape, cola, orange, and strawberry. Find the long run relative frequency of getting a Grape lollipop at random.
Interpretation	

5)

Situation	Doing a bottle flip challenge where each person tries to flip a bottle 5 times to see how often it lands standing up. Find the long-run relative frequency of getting one bottle landing standing up. 
Interpretation	