Determining the probability of a good day for sowing cotton

Level

10

Curriculum Area

Mathematics

Print Resource

Resource Description

This resource encourages critical thinking and an understanding of conditional probability when used to determine the probability of a farmer being able to sow a cotton crop within an optimum time frame, based on real historical soil and air temperature data. This resource also teaches the use of spreadsheet commands for conditionally formatting data.

Use the language of ‘if ....then, ‘given’, ‘of’, ‘knowing that’ to investigate conditional statements and identify common mistakes in interpreting such language.

back to top

Audience

This resource consists of 3 tasks that have been written for students to be able to access, read and complete with teacher guidance. The solutions to these tasks are at the end of the resource. Teachers are provided with the context, background information and Australian curriculum linkages.

back to top

The Mathematics Tasks and Curriculum Content Descriptors

There are three tasks. These are ordered sequentially, however each task has sufficient information within it to be able to stand alone.

Task 1 – Conditional formatting of Microsoft Excel spreadsheets

In this task, students learn to manipulate data in spreadsheets. The conditional formatting function in Excel is used to identify good days for sowing and spreadsheet formulae are used to count these days. This task also uses scatter graphs to determine linear trends in data **(ACMSP279)**

Task 2 – Constructing and interpreting Venn diagrams and two-way tables
This task revises the use of Venn diagrams and the language and notation used to describe sets of data **(ACMSP204 and ACMSP205)**. It also relates data in Venn diagrams to data expressed in two-way tables **(ACMSP292)**.

Task 3- Calculating probability and conditional probability

This task investigates the concept of relative frequency (experimental probability) **(ACMSP226)** and the concept of dependence and independence **(ACMSP246)** using Venn diagrams, two-way tables and tree diagrams. During the investigative process, emphasis is given to understanding the meaning of dependent events and the correct use of conditional language **(ACMSP247)**.

Detailed plans for the delivery of each task and solutions to the questions posed are given below. Following the launch of each lesson, you will need to decide whether students work individually, in pairs or as small groups to answer the given questions, prior to discussing the solutions as a class.

Elaborations

back to top

- using two-way tables and Venn diagrams to understand conditional statements
- using arrays and tree diagrams to determine probabilities

back to top

Aims

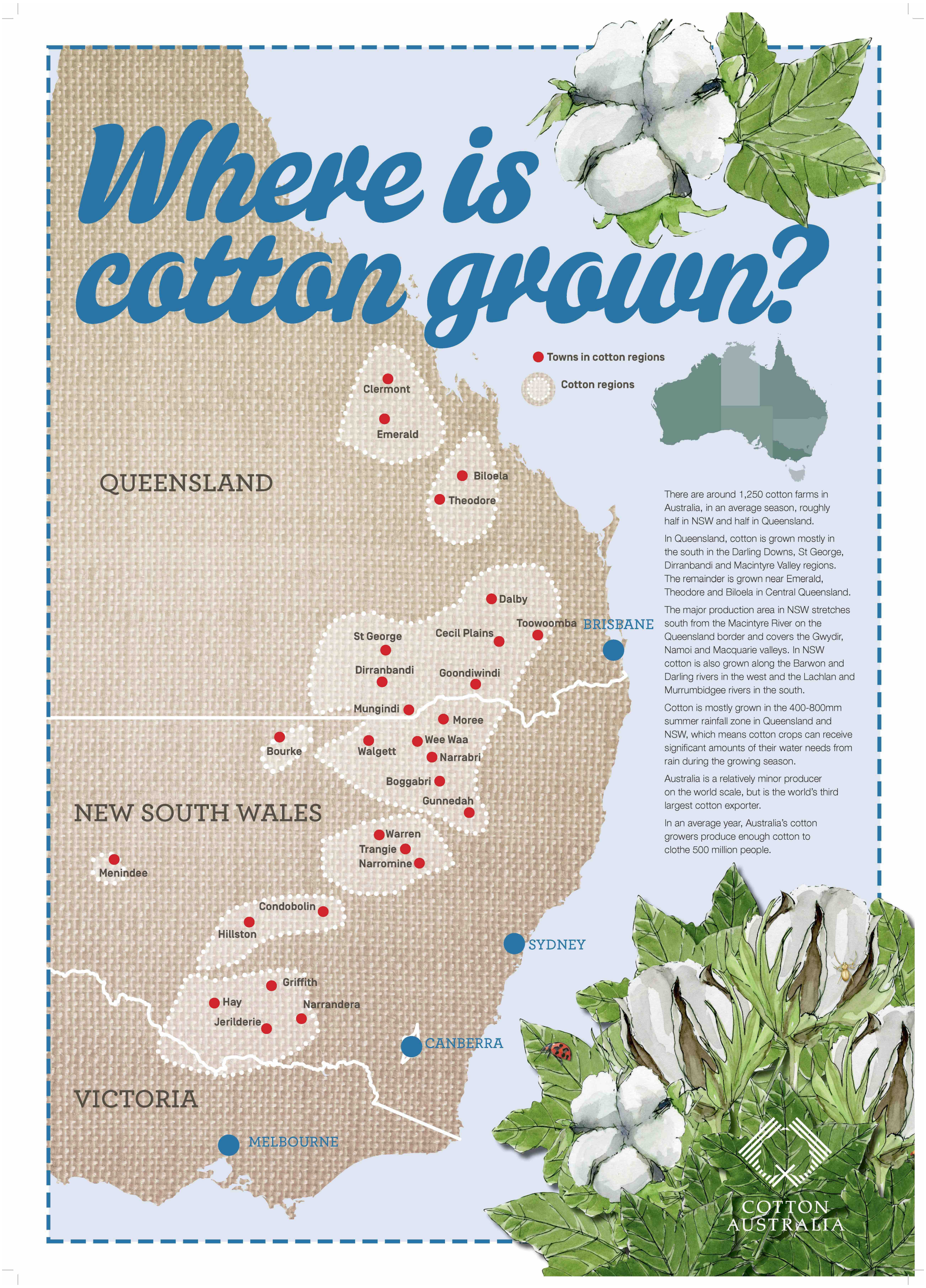
- To consolidate the language and notation used to describe sets of data and reinforce connections between different representations of frequency and relative frequency (experimental probability) though their application to a real life context
- To introduce the language of conditional probability
- To instruct students in the use of excel spreadsheet commands for conditionally formatting data
- To encourage critical thinking regarding the analysis and application of statistical data

back to top

Background Information about the Cotton Industry and the Cotton Plant

Cotton is a natural fibre produced by the cotton plant. Cotton is the most popular fabric in the world because it's soft and strong and doesn't irritate skin or cause allergies. It's also comfortable to wear because its fibre structure provides ventilation. In other words, it breathes. Cotton keeps the body cool in summer and warm in winter.

The production of cotton yarn from cotton plants originated at least 7,000 years ago in the hot, dry regions of India and Central America. Cotton plants were brought to Australia in 1788 with the First Fleet, however Australia's modern cotton industry began in north-western NSW in the 1960s. Today about two thirds of Australia's cotton is grown in NSW and the remainder is grown in Queensland.



Source: http://cottonaustralia.com.au/uploads/resources/Cotton_Map_poster-ONLINE.jpg (http://cottonaustralia.com.au/uploads/resources/Cotton_Map_poster-ONLINE.jpg) © Cotton Australia 2015, accessed 25th May, 2016.

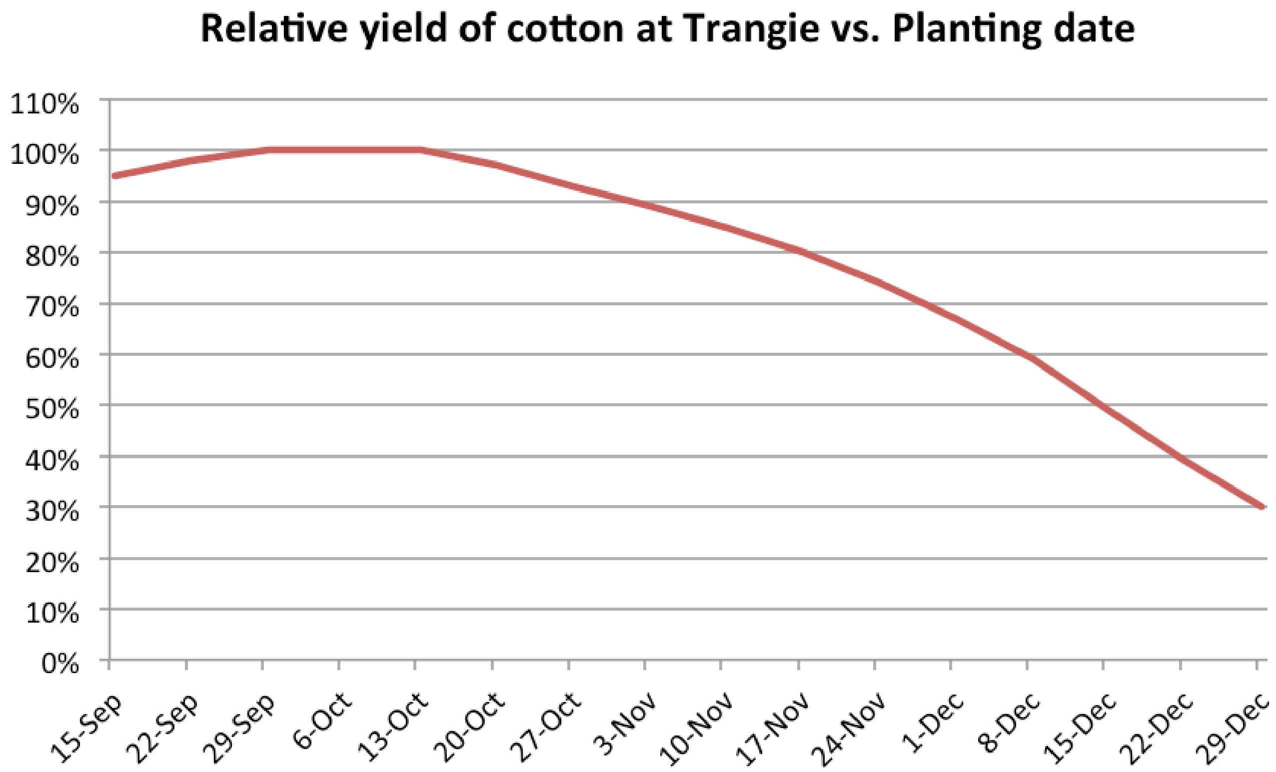
Thanks to cotton research and development, Australian cotton growers produce yields almost three times the world average, and compared to 10–15 years ago they are using 30 per cent less land and 40 per cent less water to produce a tonne of cotton lint (Australian Cotton Production Manual, 2015).

The success of a cotton crop depends on the farmer's management strategies. A good strategy relies on an understanding of the plants' development. Cotton plants follow a specific pattern of development depending on temperature. Temperature records can therefore be used to reliably predict the time from when a crop is sown (i.e. planted) to when it will produce the first flower bud (called 'first square'), when it will flower, when the first cotton boll (i.e. the fruit) will reach its maximum size, and when the boll will open, revealing the cotton fibres.

Cotton is sown in spring. It is wise to plant it not much earlier than the last expected frost because a frost may kill the young cotton seedlings. However after this time, the crop needs to be sown as early as possible to allow the maximum time for growth and development before autumn frosts occur. Timely sowing of the cotton seeds into warm soil, guarantees strong, vigorous seedlings. The later a crop is sown, the lower its potential cotton yield and quality, particularly if the season is cooler than usual. A late sown crop is also more susceptible to insect pests.

The ideal planting time varies between seasons and between cotton-producing districts. Districts such as Bourke, have a longer growing season for cotton than districts further east or south, such as Hillston or Trangie. This is because the more easterly and the more southerly districts are cooler (Australian Cotton Production Manual, 2015).

At Trangie, the average date of the last frost is mid-September, so crops should be sown after this time. The graph below shows how the relative yield of a cotton crop in the Trangie district declines when it is sown later in the spring, assuming that the crop has sufficient water.



The sowing of a cotton crop at Trangie needs to be no later than 20 October for the crop to be profitable. Early crop maturity is essential in this location because of the district's relatively cool temperatures. If the crop doesn't mature in time, yield and cotton quality are sacrificed and in addition to this, rain may further delay harvesting and consequently the planting of the following winter crop, such as wheat.

back to top

Information about Cotton in Australia

Cotton Australia - the website of the peak representative body for the Australian cotton growing industry
http://cottonaustralia.com.au/ (http://cottonaustralia.com.au/)

CottonInfo - a website designed to **connect cotton growers with research**
http://cottoninfo.com.au/search/node/cotton%20production
(http://cottoninfo.com.au/search/node/cotton%20production)

CottASSIST - a website developed by the CSIRO with data and data management tools to assist with cotton management decisions.
https://www.cottassist.com.au/Climate/About.aspx
(https://www.cottassist.com.au/Climate/About.aspx)

Cotton Seed Distributors Ltd - the website of the seed company that sells cotton varieties bred by the CSIRO to cotton growers. All cotton growers in Australia grow varieties bred by the CSIRO. http://www.csd.net.au/ (http://www.csd.net.au/)

CSIRO - the CSIRO website about their research in cotton. http://www.csiro.au/en/Research/AF/Areas/Plant-Science/Cotton (http://www.csiro.au/en/Research/AF/Areas/Plant-Science/Cotton)

back to top

Conditional formatting in Excel

https://www.ablebits.com/office-addins-blog/2014/06/10/excel-conditional-formatting-formulas/#conditional-formatting-using-formulas (https://www.ablebits.com/office-addins-blog/2014/06/10/excel-conditional-formatting-formulas/#conditional-formatting-using-formulas)

back to top

Student Tasks

The following material is written for student to complete the 3 tasks outlined. The teacher notes and solutions are provided at the end of the resource.

back to top

Student Task 1 – Your Consulting Job

Download Task 1 (pdf/maths-cotton-task1.pdf)

A farmer at Trangie is considering investing in the development of part of his small property for cotton production. He has learnt from agronomists that based the mathematical modelling of cotton growth and development in his district, the optimal period for sowing cotton (the planting season) at Trangie is from 15 September until 20 October. If the crop has not been sown by 20 October, it risks losing its profitability.

The farmer has also read information about the temperature conditions necessary for sowing cotton. This information, supplied by Cotton Seed Distributors on their website (http://www.csd.net.au/greenlight (http://www.csd.net.au/greenlight)), is shown below.

Have you got the green light for planting this season?

Planting the cotton crop is one of the most important operations on the farm. It sets the standard for the entire season. There are some key considerations that will help ensure that it is a once only task.

	red light	orange light	green light
Soil temperature at 10 cm depth above 14°C at 9am (AEST)	✘	✔ ✘	✔
Forecast average temps for the week following planting on a rising plane	✘	✘ ✔	✔
	STOP!	**STEADY**	**GO!**

IMPORTANT

- If you cannot give a green tick next to at least one of these statements, then planting conditions are definitely unsuitable - **STOP!**
- If you can give a green tick to only one of these statements - **BE CAUTIOUS**. Adjustments may need to be made.
- If you can give both statements a green tick - **LET'S GO!**

The farmer has decided that he will develop his property for cotton production if he can expect there to be at least five **green light** days each year during the planting season at Trangie.

Your job is to analyse data from climatic records and then use your analysis to **advise the**

farmer

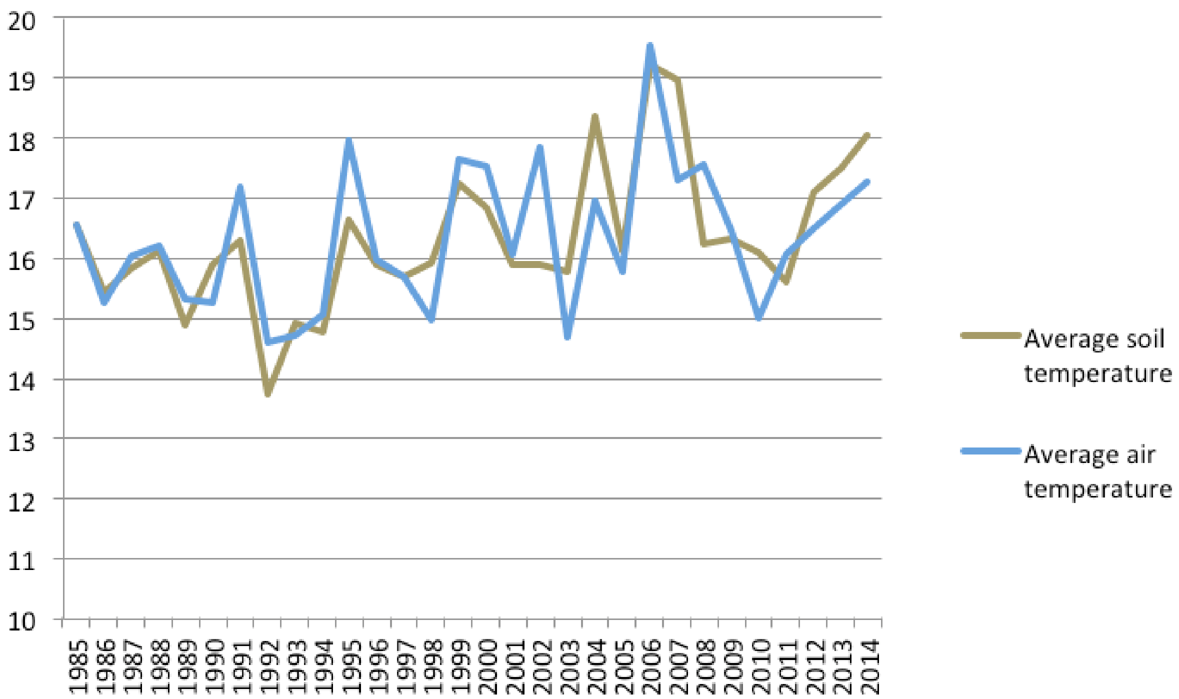
Task 1 - Conditional formatting of Microsoft Excel spreadsheets

This task requires you to conditionally format climatic data on an Excel spreadsheet so the farmer can easily see and count the number of **red light** days, **orange light** days and **green light** days in a planting season at Trangie.

The farmer has obtained 30 years of temperatures of the soil at a 10 cm depth, measured at Trangie Agricultural Research Station at 9 am AEST (1985 – 2014). He has also found daily maximum and minimum air temperature data for Trangie online at the Bureau of Meteorology website (http://www.bom.gov.au (http://www.bom.gov.au)). From these two sets of historical climatic data, he wants to find out how many suitable sowing days for cotton are likely to occur at Trangie during the planting season (15 September and 20 October). The number of days during this period when he can sow cotton will impact on his decision as to whether he will produce cotton, and if he does, how much of his farm he will commit to cotton production.

Average soil temperatures at a depth of 10 cm and average daily temperatures at Trangie during the optimum sowing period of the years 1985 – 2014, are shown in the graph below. Average daily soil temperatures range from 13.7°C (in 1992) up to 19.2 °C (in 2006) and average daily air temperatures range from 14.6°C (in 1992) up to 19.5°C (in 2006). The farmer decides to examine the data for 1992 and 2006 more carefully to find out how many **red light** days, **orange light** days and **green light** days occurred during the planting season in these two extreme years, and also in a more typical season such as that of the year 2000.

Average daily temperatures at Trangie, 15 September – 20 October



This graph was drawn using data in the Microsoft Excel spreadsheet Trangie Planting Season Data.

In this spreadsheet, the data on the worksheets are:

- Daily soil temperature at 10 cm depth read at 9am AEST from 15 September – 20 October
- Daily maximum air temperature from 15 September – 27 October
- Daily mean air temperature (calculated by averaging the maximum and minimum temperatures for each day) from 15 September – 27 October.

Note: Averages calculated at the bottom of each worksheet are over the period from 15 September – 20 October. Air temperature data continues for another week to make it possible to detect whether the forecast temperatures were on a "rising plane" for the week following 20 October.

Open the spreadsheet Trangie Planting Season Data.

Trangie Planting Season Data (pdf/trangie-planting-season-data.xlsx)

The first worksheet shows soil temperature at 10 cm depth during the planting season each year from 1985-2014. These data are measurements made at Trangie Agricultural Research Station before this job was taken over by the Bureau of Meteorology in 2015.

Some of the data in this worksheet are missing. This missing data forms a pattern.

- Why do you think this pattern happens?

Open the spreadsheet Testing for Good Sowing Days.

Testing for Good Sowing Days (pdf/testing-good-sowing-days.xlsx)

This spreadsheet contains three worksheets, each with a year of daily data selected from the spreadsheet Trangie Planting Season Data. These years are 1992 (the year with the coldest planting season) , 2006 (the year with the warmest planting season) and 2000 (a year with a fairly average planting season.)

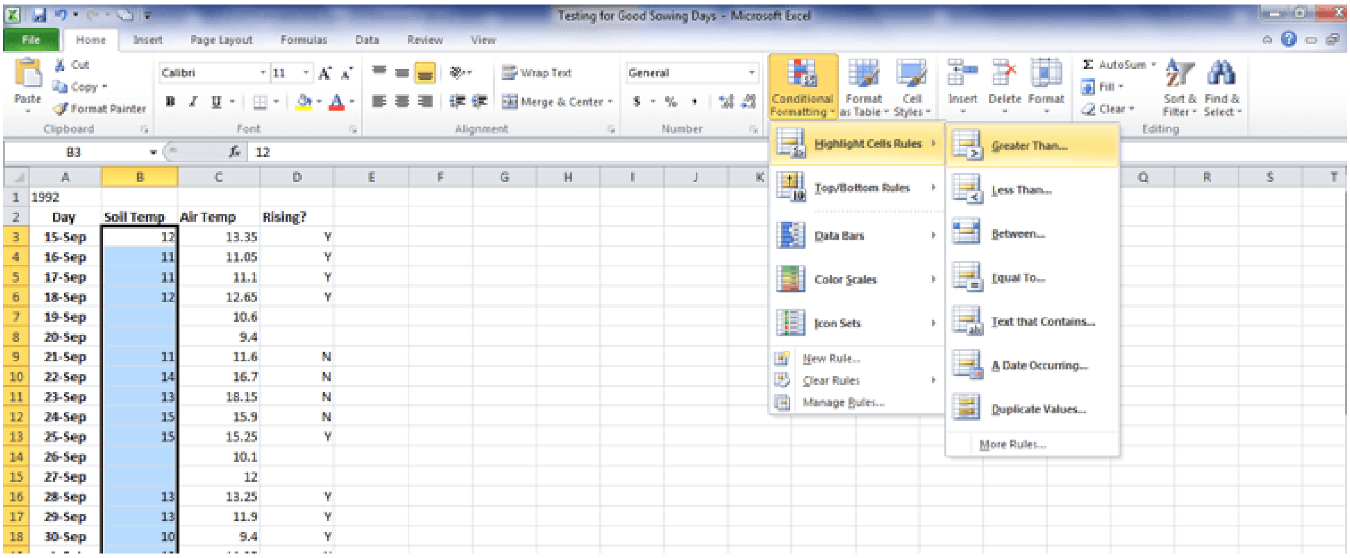
You will be using these data to determine which days in the planting season satisfy one, both or neither of two criteria given by Cotton Seed Distributors as being important when deciding whether to plant cotton on a particular day, namely:

a) Soil temperature at 10 cm depth above 14°C at 9am (AEST)

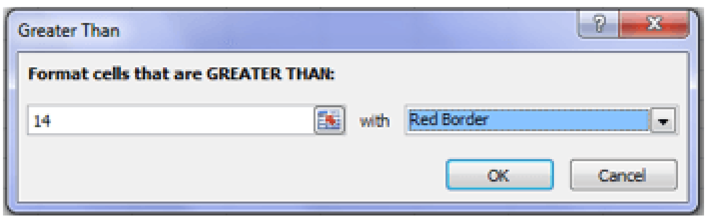
b) Forecast average temps for the week following planting on a rising plane.

In Microsoft Excel spreadsheets you can specify rules for formatting cells according to what is written in the cell.

In the worksheet for 1992, select the soil temperature data in Column B (Cells B3:B38), then select Conditional Formatting >Highlight Cells Rules >Greater Than…



In the resulting box, type **14** and select Red Border from the dropdown menu (shown below).



To count the number of days in the planting season for which there is soil temperature data, go to Cell E47. In the formula box, type **=COUNT(B3:B38)**.

- On how many days in the 1992 planting season, was soil temperature measured?

To count the number of days in the planting season for which the soil temperature is greater than 14°C, go to Cell E48.

In the formula box, type **=COUNTIF(B3:B38,">14")**.

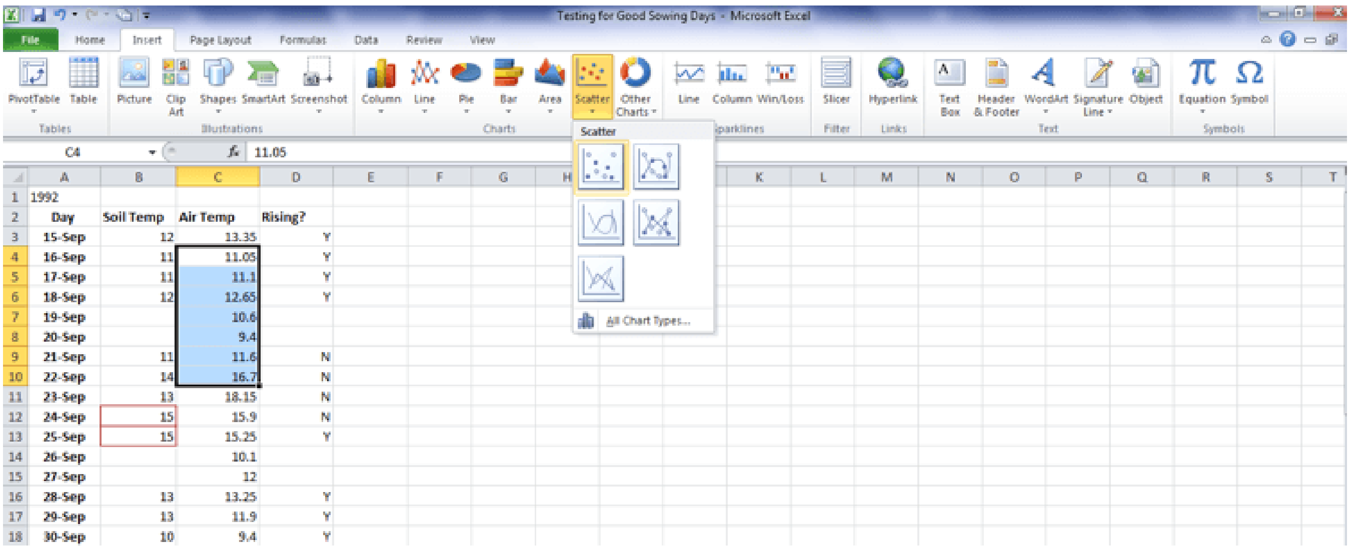
This formula consists of an array of cells to search (B3:B38) followed by a condition for a cell to be included in the count (">14").

- How many of these days have a soil temperature greater than 14°C?
- What fraction of the days when soil temperature was measured, have a soil temperature greater than 14°C? Write your answer as a decimal correct to 2 decimal places.

The second condition for a good sowing date is a forecast of rising air temperatures for the following week. To determine this, a scatter plot can be drawn using the mean air temperature data for the 7 days after the date of the soil temperature measurement.

- By using data in this spreadsheet as "a forecast of average temperatures", what assumption has been made?

To make a scatterplot of air temperature data for the week following 15 September (i.e. the week of 16 – 22 September), select the cells C4:C10. Then select the tab Insert, the icon Scatter, and the first of the types of scatter plot (as shown below).



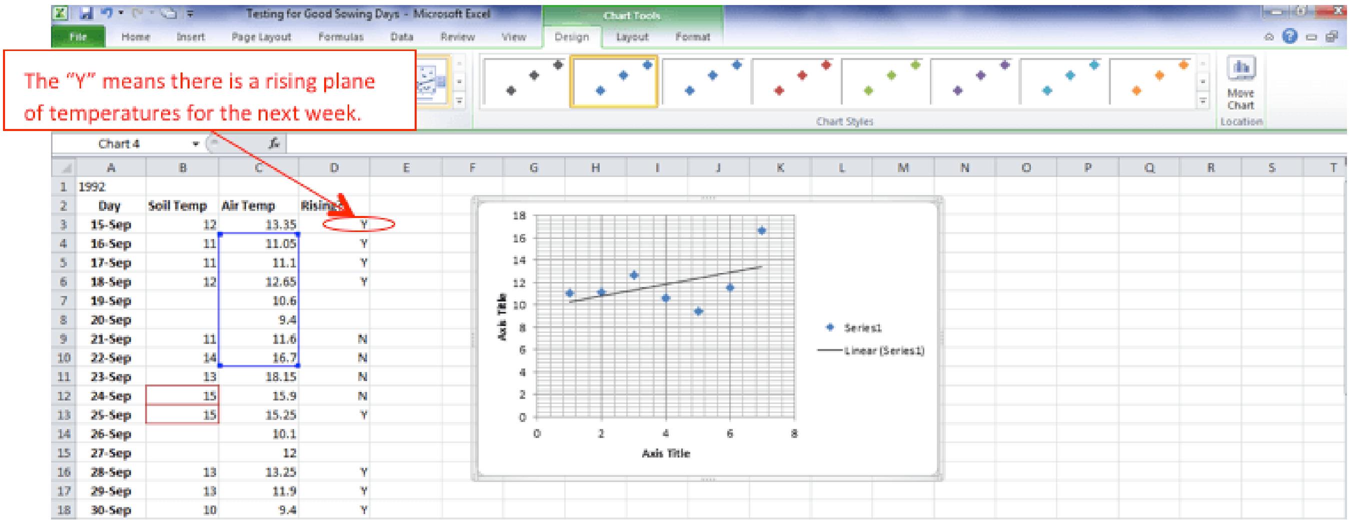
You should now have a scatterplot with 7 data points on it.

The numbers on the horizontal axis represent the order of the 7 dates.

The numbers on the vertical axis represent the mean air temperature for each date (in °C).

Draw a trend line through these 7 data points by selecting Chart Layouts (Under the Design tab) then Layout 3 (as shown below).

You can quickly see by the direction of the slope of the line, whether the week's temperatures are rising or falling.



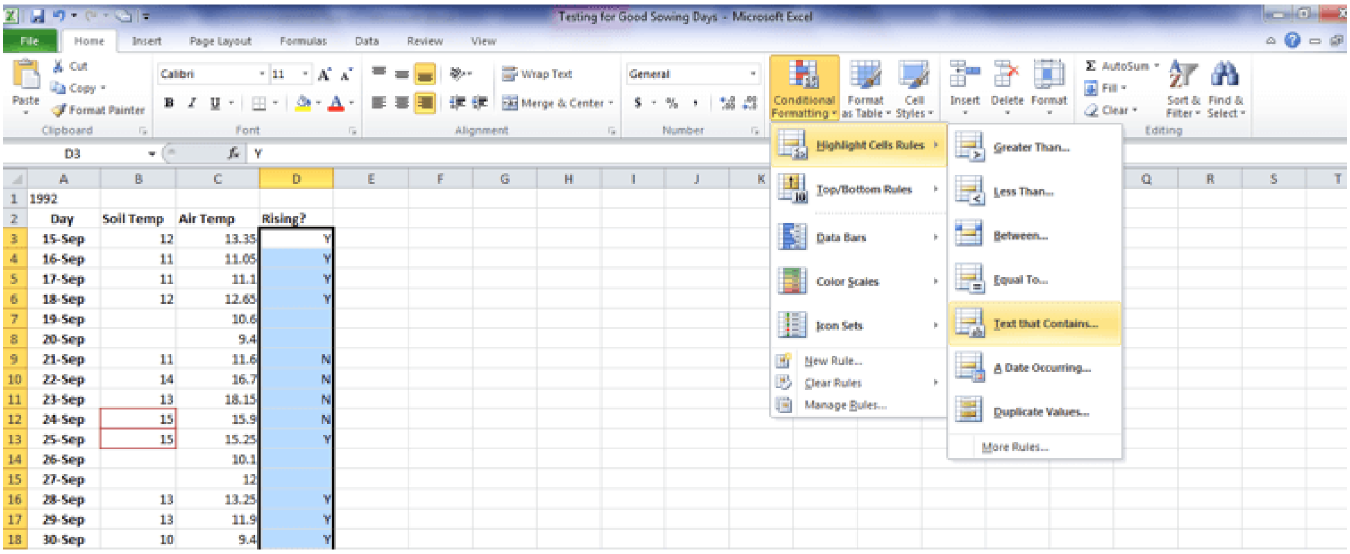
The scatterplot for the week following (and not including) 15 September shows that temperatures are "on a rising plane" (i.e. the trend line is going up). To record this trend in your table for 15 September, the letter "Y" (for Yes) has been placed in Cell D3.

Check that Column D is correct for other days ("Y" for a rising trend and "N" if it is not rising) by using the same method. You can change the data used in an existing scatter plot by right clicking on the plot, selecting Select Data, highlighting a different string of 7 temperatures and clicking OK.

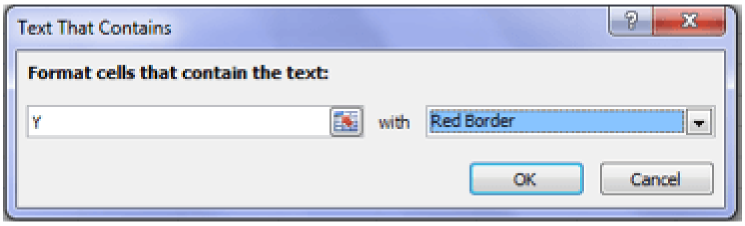
Note: Temperature trends do not need to be found for the weeks when the prior day has no soil temperature data. This is because if a day has no soil temperature measurement it cannot be used to decide whether it is good for sowing cotton or not.

The column indicating whether the temperatures are rising (Column D) can now be conditionally formatted to highlight only those cells that have a "Y" in them.

Select the data in Column D (Cells D3:D38), then select Conditional Formatting >Highlight Cells Rules >Text that Contains…



In the resulting box, type **Y** and select Red Border from the dropdown menu (shown below).



In Cell E49, make a count of the number of days with a forecast of rising temperatures, by typing the formula **=COUNTIF(D3:D38, "Y")** into the formula bar.

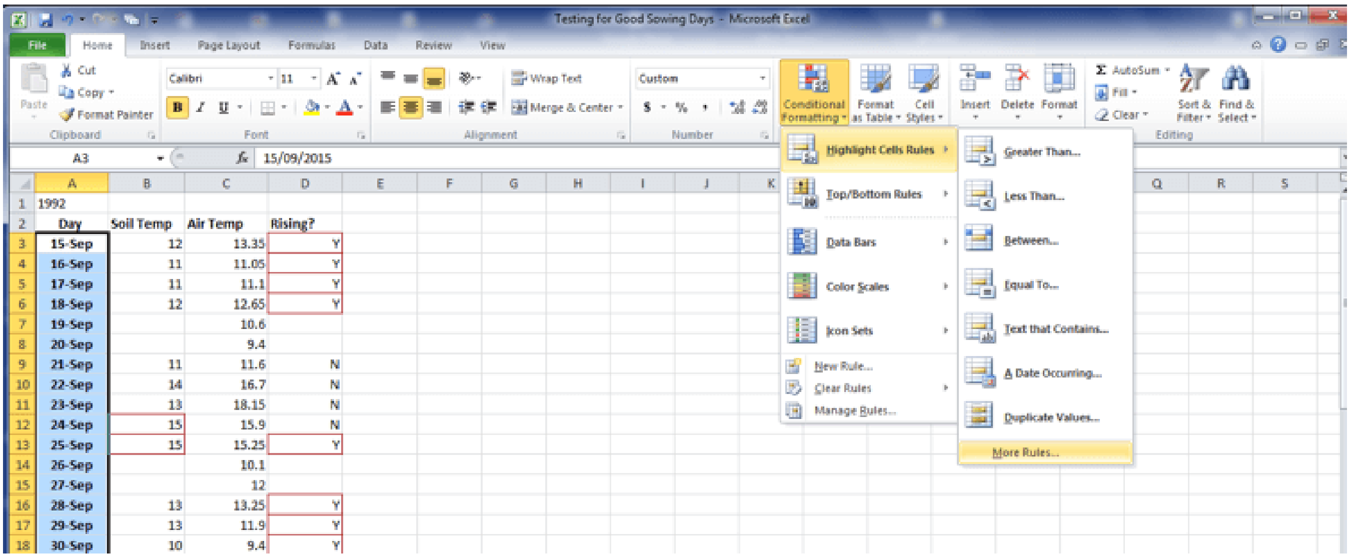
- How many of the days with soil temperature data, have mean temperatures for the following week on a rising plane?
- What fraction of the days with soil temperature data, have average temperatures for the following week on a rising plane? Write your answer as a decimal correct to 2 decimal places.

Look again at the decision-making criteria supplied by Cotton Seed Distributors.

- From conditionally formatting your spreadsheet:
- how could you tell which days are **green light** days?
- how could you tell which days are **red light** days?
- how could you tell which days are **orange light** days?

In Excel, it is possible to format a cell based on the value or text in one or more other cells. Highlight the data in Column A from 15-Sep to 20-Oct (Cells A3:A38).

Then select Conditional Formatting >Highlight Cell Rules >More Rules…as shown below.

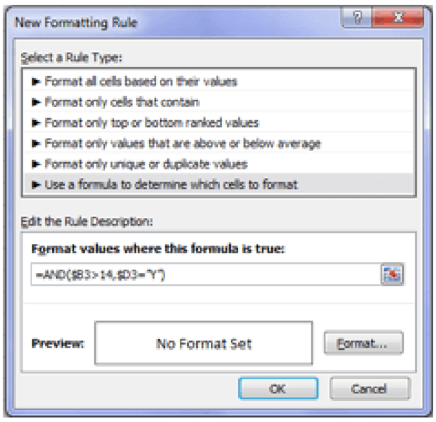


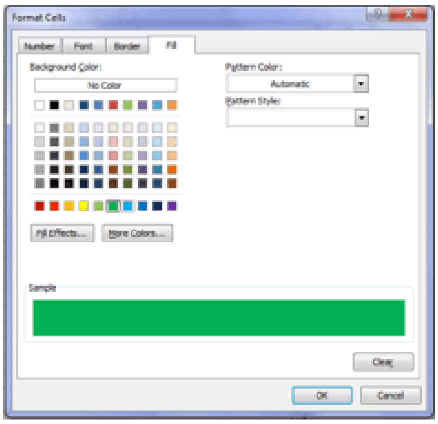
Select the Rule Type **>Use a formula to determine which cells to forma**t

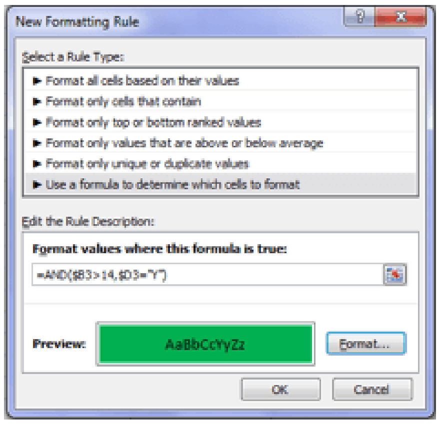
Type the Rule Description as: **=AND($B3>14,$D3="Y")**

Click on Format. Format so the selected cells are filled with green if the rule is obeyed.

Click OK for the formatting, then OK to have the new formatting rule applied.

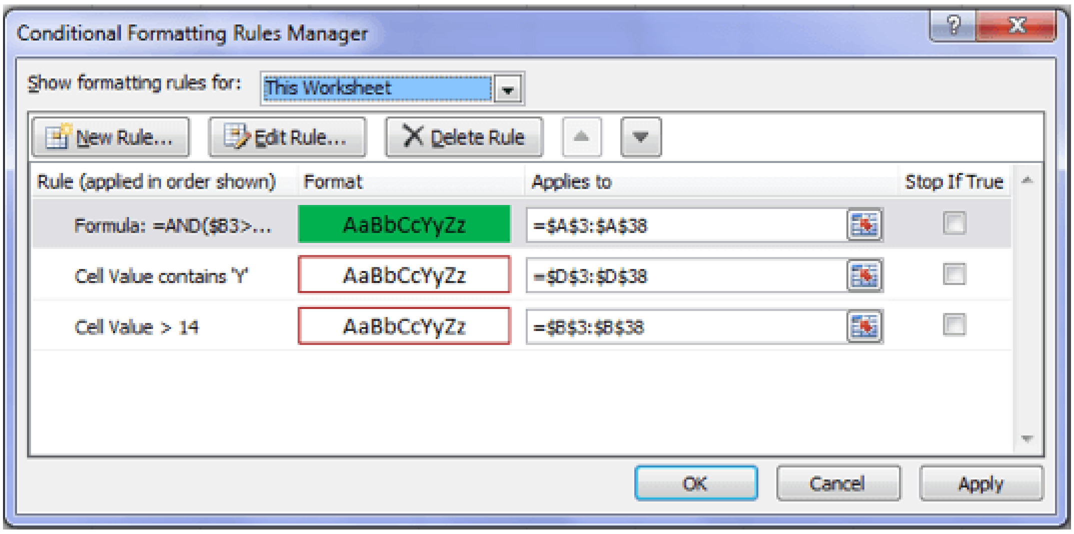






You can see all the formatting rules you have used for your worksheet so far, by selecting Conditional Formatting >Manage Rules…

Select This Worksheet from the dropdown menu at the drop and you should see the rules below.



The rule **=AND($B3>14,$D3="Y")** selects cells where:

- the soil temperature is greater than 14°C **AND**

- temperatures are rising for the following week.

These are **green light** days.

- What are the two conditions for **red light** days?
- Write this rule as an Excel formula to conditionally format these days.

(Note that the symbols <= written together mean "less than or equal to".)

Use your new formatting rule to colour these days in Column A, red.

The conditions for **orange light** days are:

- the soil temperature is greater than 14°C **OR**

- temperatures are rising for the following week.

but **NOT BOTH** of these conditions.

So there are two types of **orange light** days.

To separate them, we will call them **brown light** days when <u>only the soil condition</u> is met, and **orange light** days when <u>only the rising temperature condition</u> is met.

- What are the two conditions for **brown light** days?
- Write this rule as an Excel formula to conditonally format these days.

Use your new formatting rule to colour these days in Column A, brown.

- What are the two conditions for **orange light** days?
- Write this rule as an Excel formula to conditonally format these days.
- Use your new formatting rule to colour these days in Column A, orange.

To count the number of **green light** days, go to Cell E51 and type the formula **=COUNTIFS(B3:B38,">14",D3:D38,"=Y")**.

- Explain the meaning of the symbols in this formula.
- What formula would you type to get a count of the **red light** days into Cell E52?
- What formula would you type to get a count of the **brown light** days into Cell E53?
- What formula would you type to get a count of the **orange light** days into Cell E54?

Type in these formulae and check that the counts made are correct by counting the coloured cells.

In the first column of the table below, write your results for 1992 planting season.

Repeat the conditional formatting exercise for 2000 and 2006 using the other worksheets in the spreadsheet Testing for Good Sowing Days. Write these results into the table below.

	1992 Planting Season	**2000 Planting Season**	**2006 Planting Season**
Number of **green light** days			
Number of **red light** days			
Number of **brown light** days			
Number of **orange light** days			
TOTAL number of days with data	**25**	**24**	**23**

- Why is it better to compare the proportion of days in each category rather than compare the number of days in each category?

For each year, calculate the proportion of days in each category.

Write your answers as decimal fractions (to 2 decimal places) in the table below.

For each season, the proportions should add to 1.

	1992 Planting Season	2000 Planting Season	2006 Planting Season
Proportion of **green light** days			
Proportion of **red light** days			
Proportion of **brown light** days			
Proportion of **orange light** days			

From this data, the minimum proportion of **green light** days in a planting season at Trangie is expected to be the proportion of **green light** days in 1992 (the coolest planting season).

- How many days are there in a planting season (15 September – 20 October)?
- What is the proportion of **green light** days in the 1992 season?
- In the coolest planting seasons, how many **green light** days are expected?

The number of **orange light** days is the sum of the **brown light** days and the **orange light** days.

- What is the proportion of **orange light** days in the 1992 season?
- In the coolest planting seasons, how many **orange light** are expected?

If a farmer has a smaller property, he is usually not prepared to take the risk and sow on an orange light day. He would rather wait for a green light day.
Farmers with larger properties are likely to sow on all days except red light days.

In the coolest planting seasons, on how many days during the coolest planting seasons are farmers with larger properties likely to sow cotton?

back to top

Student Task 2 – Constructing and interpreting Venn diagrams and two-way tables

Download Task 2 (pdf/maths-cotton-task2.pdf)

To plant cotton on a particular day, there
are two important conditions:

Condition A: Soil temperature at 10 cm depth above 14°C at 9am (AEST)
Condition B: Forecast average temperatures for the week following planting on a rising plane.

A day in the cotton planting season may or may not meet these conditions.
When a condition is met, it is said to be **satisfied**.

In Task 1, a farmer decided to look at historical climate data to see how frequently these conditions have been satisfied at Trangie. He decided to look at data for the planting season in 1992 because it was the coolest planting season in the record. He also looked at data for the planting season in 2006 because it was the warmest planting season in the record.

In a planting season, the days when Condition A was satisfied form a **set** of days which will be called Set A. The days when Condition B was satisfied form another **set** of days which will be called Set B.

The Cotton Seed Distributers' website (http://www.csd.net.au/greenlight (http://www.csd.net.au/greenlight)) says:

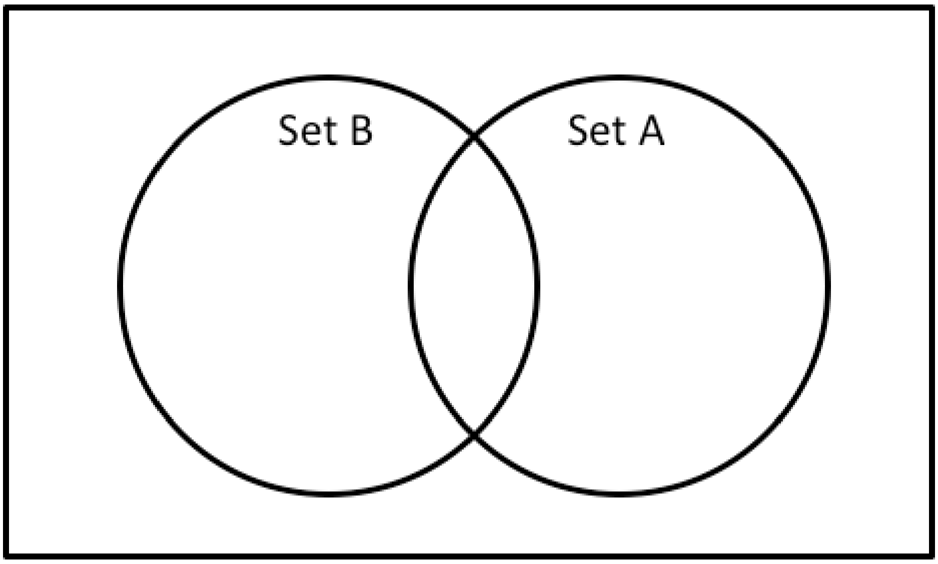
- If both conditions are satisfied, it is a **green light** day, meaning that the farmer should go ahead and sow cotton that day.
- If neither condition is satisfied, it is a **red light** day, meaning that the farmer should stop and not sow cotton that day.
- If only one of the two conditions is satisfied, it is an orange light day, meaning that the farmer should be cautious and only sow cotton that day if there is unlikely to be enough green light days remaining in the planting season.

Days might be orange light days because only condition A is satisfied (**brown light** days) or because only condition B is satisfied (**orange light** days).

Venn Diagrams
The information for each planting season can be shown in a Venn diagram. A Venn diagram shows the relationship between two or more sets, each set being represented by a circle.

The circle labelled Set A represents the number of days that satisfy Condition A.

The circle labelled Set B represents the number of days that satisfy Condition B.



The two circles in the Venn diagram intersect (i.e. overlap) because some days satisfy Condition A and Condition B and are therefore in both Set A and Set B.

- What colour light was given to these days?

On the Venn diagram above, shade the region (i.e. the area) where the circles overlap, with this colour of pencil.

On some days, neither Condition A nor Condition B was satisfied.

- What colour light was given to these days?

On the Venn diagram above, shade the region outside the circles (but inside the rectangle) with this colour.

On some days, Condition A was satisfied but Condition B was not satisfied.

- What colour light was given to these days?

On the Venn diagram above, shade the part of the circle labelled Set A that does not overlap with the circle labelled Set B.

On some days, Condition A was not satisfied but Condition B was satisfied.

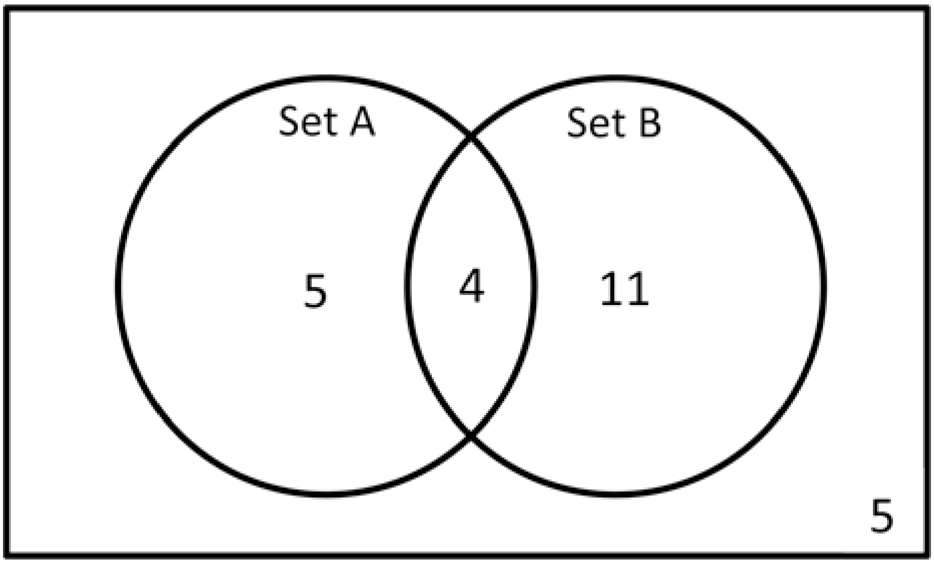
- What colour light was given to these days?

On the Venn diagram above, shade the part of the circle labelled Set B that does not overlap with the circle labelled Set A.

The table below shows the number of the different types of days found in the 1992 planting season.

	1992 Planting Season
Number of **green light** days	4
Number of **red light** days	5
Number of **brown light** days	5
Number of **orange light** days	11

Each number of days in the table matches a colour you used to shade a region of the previous Venn diagram. The Venn diagram below is the same diagram except a number (the number of days of each type) has been written in each region.



Set Language and Notation

The items in a set are called the **elements** of the set.

The elements of the planting season at Trangie, are days.

- What is the total number of days represented in the Venn diagram for the 1992 season?

This number of elements is said to be the number in the **universal set**.

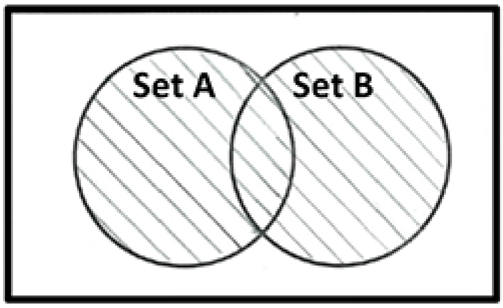
Look at the Venn diagram for the 1992 planting season.

The **brown light** days and **green light** days are in Set A.

The notation **n(A)** means the number of elements in Set A. n(A)= 5 + 4 = 9

- What colour days are in Set B?

- What does n(B) equal?

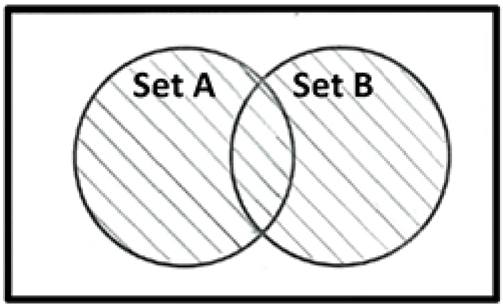


The **union** of two sets, A and B (symbolised as A ∪ B) refers to all the elements from both sets.

It is when either of the two conditions is satisfied.

It can be found on a Venn diagram by shading both circles.

It then includes every region of the diagram that's shaded.

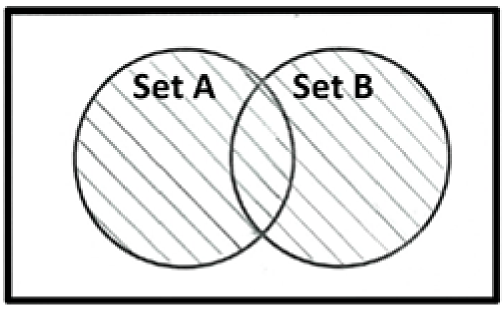


The **intersection** of two sets, A and B (symbolised as A ∩ B) refers to the elements that are common to both sets.

It is when both of the two conditions are satisfied.

It can be found on a Venn diagram by shading one set in one direction and one set in the other direction (as shown).

It is then the region shaded twice (hatched lines in both directions).



If an element is not within Set A, it is in the **complement** of Set A.

The notation used for the complement of A, is **A'**.

The complement of Set A, is Set A'.

Set A' is the shaded region of this Venn diagram.

The Venn diagram below shows the data for the 2000 planting season.

- Shade it with a pencil to help you find the values of the following:

n(A ∪ B)=

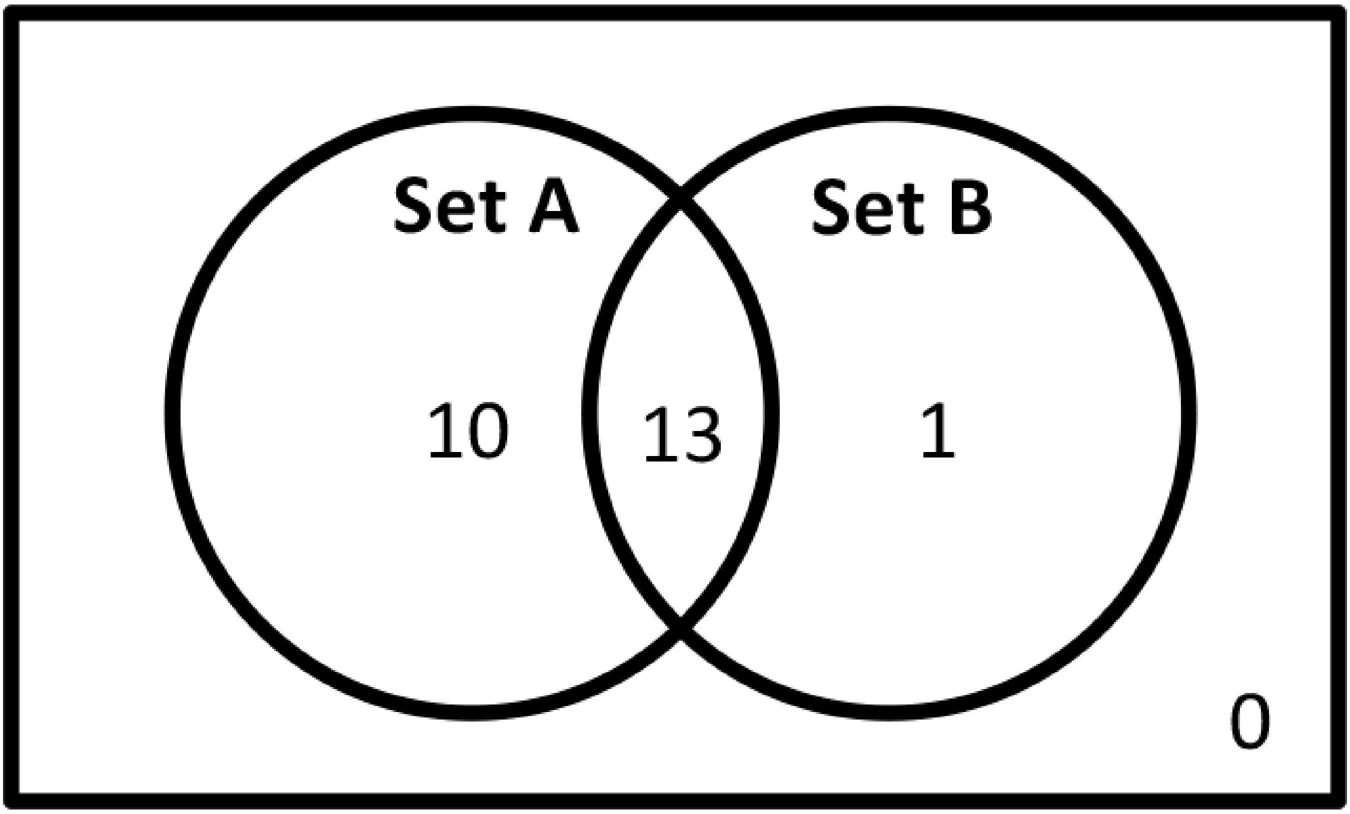
n(A ∩ B)=

n(A' ∪ B)=

n(A' ∩ B)=

n(A' ∪ B')=

n(A' ∩ B')=



- Is the rule n(A ∪ B) = n(A) + n(B) - n(A ∩ B) true for the 2000 planting season?

Is this rule true for the 1996 planting season?

Is this rule true for any number of elements in Set A and Set B? Explain why or why not.

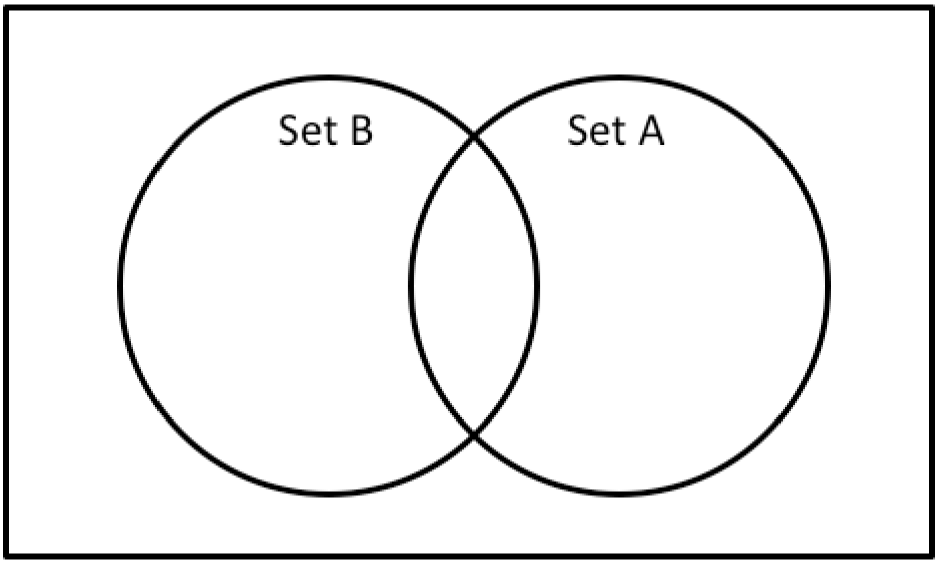
Would this rule apply if there were no **green light** days (i.e. no overlap of the two sets)?

The table below shows the number of the different types of days found in the 2006 planting season at Trangie. This was the warmest planting season in the data record.

	2006 Planting Season

Number of **green light** days	12
Number of **red light** days	0
Number of **brown light** days	11
Number of **orange light** days	0

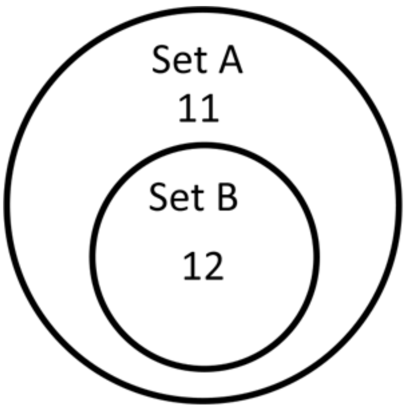
Represent this data by writing a number of days into each region of the Venn diagram below.



- How many days are in Set A?
- How many days are in Set B?

Because all the elements of Set B are in Set A,

the diagram could be drawn as one circle inside the other.



- Colour each region according to the type of day it represents.
- Why do you think it is not necessary to place the

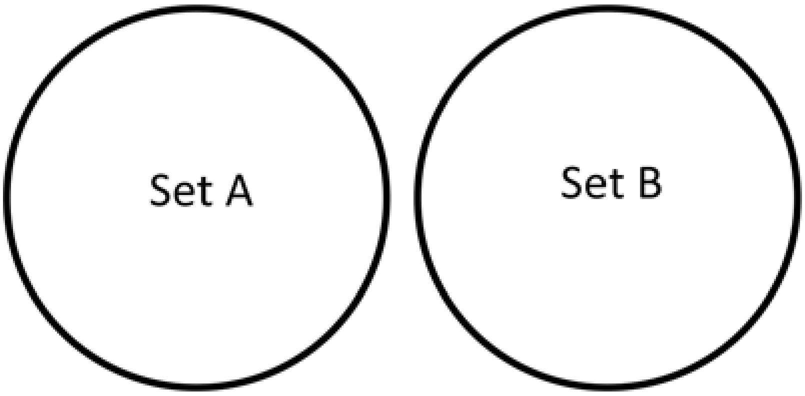
circles inside a rectangle?

Sometimes sets have no elements that overlap.

These sets are described as being **mutually exclusive**.

This means that every element is either in Set A or in Set B, not both.

There is no intersection, as shown in the Venn diagram below.



<u>Two-way Tables</u>

Complement notation is often used as column and row headings in two-way tables.

The columns divide the days into the number of them that satisfy Condition A (Set A) and the number of them that don't satisfy Condition A (Set A').

The rows divide the days into the number of them that satisfy Condition B (Set B) and the number of them that don't satisfy Condition B (Set B').

	Set A	**Set A'**
Set B		
Set B'		

Colour the 4 empty cells of the table above as follows.

green light days satisfy Condition A and Condition B, so colour the cell that is in the Set A column and in the Set B row, green.

Colour the remaining 3 cells to indicate the sets in which there are **red light** days, **brown light** days and **orange light** days.

The table below shows the number of days in each cell.

1992	**Set A**	**Set A'**	**Marginal Totals**
Set B	4	11	
Set B'	5	5	

Marginal Totals			

In the table, write the total of each column and each row. These are called **marginal totals**.

- What does the total of Set B mean in this context?
- What does the total of Set A' mean in this context?
- What is the sum of the two column totals?
- What is the sum of the two row totals?
- Why should the sum of the columns totals, be the same as the sum of the row totals?

This sum is the number in the universal set.

- Using the Venn diagrams for 2000 and 2006, make a two-way table for each of these seasons.
- Which of the three seasons (1992, 2000 or 2006) has the greatest proportion of days in Set B?
- Did you expect all the seasons to have more than half of their days in Set B? Why or why not?

back to top

Student Task 3 – Calculating probability and conditional probability

Calculating Probabilities using a Venn Diagram and a Two-way Table

Download Task 3 (pdf/maths-cotton-task3.pdf)

Probabilities are proportions (i.e. fractions). They indicate what proportion of times a particular outcome is expected to occur.

Venn diagrams and two-way tables can be used to calculate a probability and/or a conditional probability of something occurring.

The thing that occurs is called the **outcome**.

The number of times that an outcome occurs is its **frequency.**

The **relative frequency** of an outcome is its frequency expressed as a fraction of the largest possible number of times it could have occurred.

The **probability** of an outcome occurring in the future can be considered to be its relative frequency (the proportion of times it occurred in the past), assuming that nothing changes.

The relative frequency of an outcome is also called its **experimental probability**.
It is calculated as:

$$Experimental\ Probability\ of\ an\ outcome = \frac{The\ frequency\ of the\ outcome}{The\ sample\ space}$$

The **sample space** is the number of times the “experiment” was done (i.e. the greatest number of times that the outcome could have occurred).

A farmer at Trangie wants to know the relative frequency (probability) of days in a planting season being good for sowing cotton.

There are two conditions which, if satisfied, result in days with a favourable outcome for sowing.
<u>Condition A</u>: Soil temperature at 10 cm depth above 14°C at 9am (AEST)
<u>Condition B</u>: Forecast average temperatures for the week following planting on a rising plane.
A day in the cotton planting season may or may not meet these conditions.

In the language of probability, each of these conditions is an **event**.
Each event has 2 possible **outcomes** so altogether, the 2 events result in 4 possible outcomes:

- If conditions A and B are both satisfied, it is a **green light** day, meaning that the farmer should go ahead and sow cotton that day.
- If neither condition is satisfied, it is a **red light** day, meaning that the farmer should stop and not sow cotton that day.
- If Condition A is satisfied and Condition B is not satisfied it is a **brown light** day, meaning that the farmer should be cautious about sowing cotton that day.
- If Condition B is satisfied and Condition A is not satisfied it is an **orange light** day, meaning that the farmer should be cautious about sowing cotton that day.

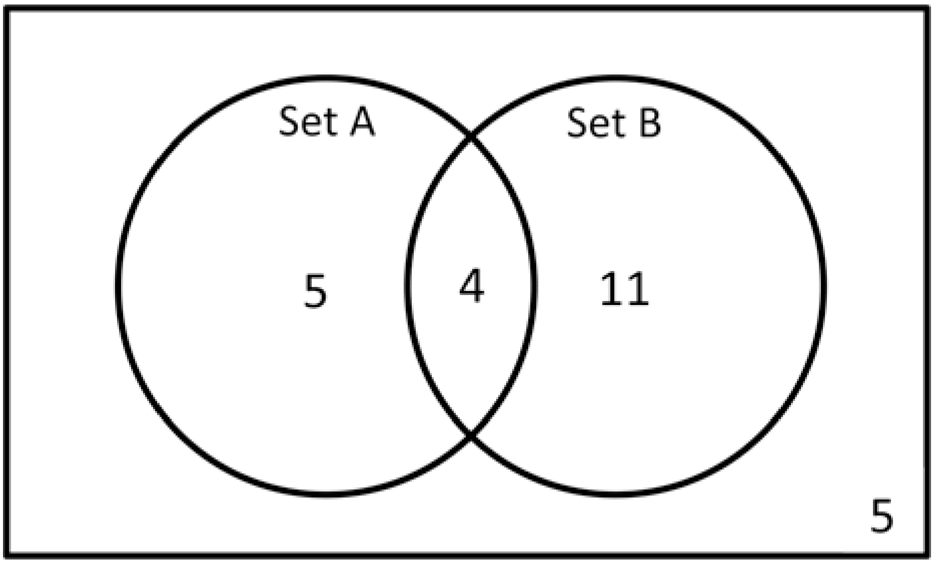
<u>Probabilities</u>
Events are **mutually exclusive** if when one event occurs, it is impossible for another event to also occur. They cannot occur together (mutually).

- Are the two events (a day when soil temperature is greater than 14°C and a day when average temperatures for the following week are on a rising plane) mutually exclusive? Why or why not?

Outcomes are mutually exclusive if when one outcome of an event occurs, it is impossible for another outcome of that event to also occur.

- For Event A (whether soil temperature is greater than 14°C) what are the two possible outcomes?

Are these outcomes, mutually exclusive? Why or why not?
Below is the Venn diagram from Task 2 showing the numbers of the different types of days in the 1992 season (the coldest planting season of the climatic data record).
Set A is the group of days when Condition A was satisfied and Set B is the group of days when Condition B was satisfied.



For each day (a day being an element in the universal set), one of 4 possible **outcomes** occurs.
The number written in each region of the Venn diagram is the frequency of each outcome.
The number of elements in the universal set (the total of the frequencies) is the sample space.

- How many days are in the sample space?

The notation for the frequency (i.e. the number) of elements in Set A is n(A).

The notation for the probability of an element being in Set A is P(A)

$$P(A) = \frac{n(A)}{\textit{The sample space}} = \frac{9}{25}$$

The probability of a **green light** day is P(A ∩ B)

$$P(A \cap B) = \frac{n(A \cap B)}{\textit{The sample space}} = \frac{4}{25}$$

- Shade the Venn diagram above to help you find out which colour of day is represented by each of the following probabilities. Then calculate the probabilities.
- P(A' ∩ B')
- P(A ∩ B')
- P(A' ∩ B)

Conditional Probabilities

The same data that is in the Venn diagram, is in the two-way table below.

1992	**Set A**	**Set A'**	**Marginal Totals**
Set B	4	11	**15**
Set B'	5	5	**10**
Marginal Totals	**9**	**16**	**25**

The shaded cell in the table gives the number of elements in the sample space (the universal set).

Look at the column of the table headed Set A.
The marginal total for Set A tells you that there were 9 days for which Condition A was satisfied.
Of these 9 days, 4 also satisfied Condition B. This is 4/9 of the days in Set A.
The denominator of the fraction (the sample space) is now restricted to the **marginal total**, 9.

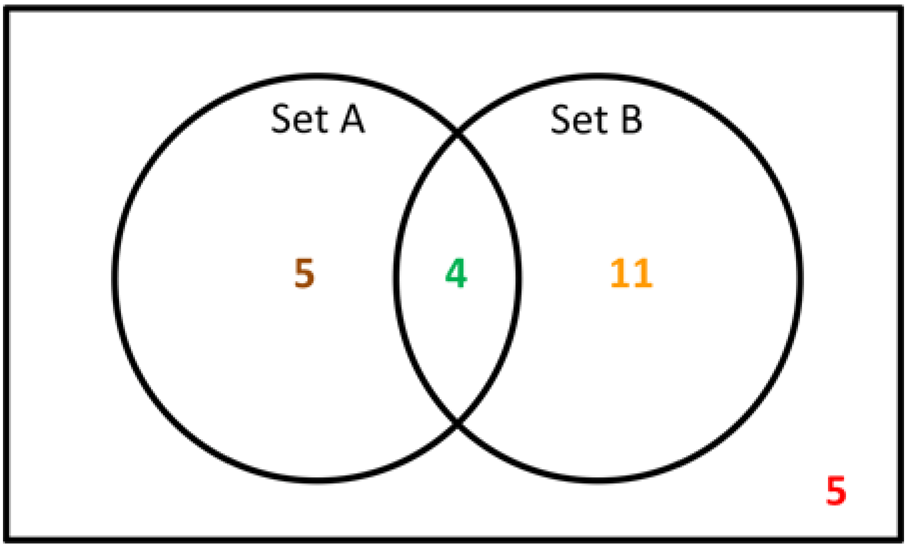
Look at the column of the table for Set A'.
The marginal total for Set A' tells you that there were 16 days for which Condition A was not satisfied. Of these 16 days, 11 did not satisfy Condition B. This is 11/16 of the days in Set A'.

Event A and Event B are **dependent events** because the outcome of Event A (i.e. whether the day is in Set A or Set A') affects the probabilities of the outcomes of Event B (whether the temperatures for the following week are on a rising plane).
If the day is in Set A, P(B)=4/9, whereas if the day is in Set A', P(B)=11/16.

The probability that a day is in Set B, given that day is in Set A is called a **conditional probability**, where the day being in Set A is the **initial condition**.

The notation used for P(B given A) is **P(B|A).**

Looking at Venn diagram below, P(B|A)=(P(A ∩ B))/(P(B))=**4**/(**4**+**5**)=4/9



The probability that a day is in Set B, given that day is in Set A, is **4/9**.

This can be expressed in different ways.

- With a partner, decide whether the following sentences are true or false, i.e. whether they mean the same as The probability that a day is in Set B, given that day is in Set A, is **4/9**.

Then share your answers with another pair of people. Convince them that your answers are correct, or be convinced by them that you are incorrect and change your answers.

1) If a day is in Set B, then the probability that it is in Set A is 4/9

2) If a day is in Set A, then the probability that it is in Set B is 4/9

3) Given that a day is in Set B, the probability that it is in Set A is 4/9

4) Knowing that a day is in Set B, the probability that it is also in Set A is 4/9

5) If you already know that a day is in Set A, the probability that it is in Set B is 4/9

- What is meant by **P(B|A^')=11/16**?
- What is meant by **P(A|B)?**What is the value of **P(A|B)** in the 1992 season?
- What is meant by **P(A|B')?** What is the value of **P(A|B')** in the 1992 season?

If events A and B are **independent**, whether an element is in Set A or not in Set A, has no effect on the probability of it being in Set B or not in Set B.

In this case, P(B|A)=P(B|A')

so P(B|A) is simply the same as P(B). i.e. P(B|A)=P(B)

Also, P(A | B)=P(A | B^') so P(A|B)=P(A)

Independent events are not the same as mutually exclusive events.
Two events are mutually exclusive if the occurrence of one event excludes the occurrence of the other. Mutually exclusive events cannot happen at the same time. In other words an element cannot be in both sets, so for mutually exclusive events P(A ∩ B)=0

When events are independent, the conditional probability of Event B occurring given that Event A has occurred, is simply the probability of Event B occurring i.e. P(B|A) = P(B).

- With a partner, decide whether the pairs of events below are dependent or independent.

Read information at http://cottonaustralia.com.au/australian-cotton/basics/cotton-facts (http://cottonaustralia.com.au/australian-cotton/basics/cotton-facts).
Then share your answers with others so you can convince them or be convinced by them.
(i) Whether a cotton crop is genetically modified, and the quantity of insecticide used on it.
(ii) Whether an area is sown to cotton or left fallow, and the emission of greenhouse gases.
(iii) The weight of a bale of cotton, and what that bale of cotton is used to make.
(iv) Rain occurring on a day at Trangie, and rain occurring on the same day at Narromine.
Probability Trees
Probabilities can be calculated from three different representations:

- Venn diagrams
- Two-way tables
- Probability trees

Probability trees show the possible outcomes by creating a **branch** for each one.
For planting cotton, there are 2 events: whether Condition A is satisfied (Event A) and whether Condition B is satisfied (Event B).

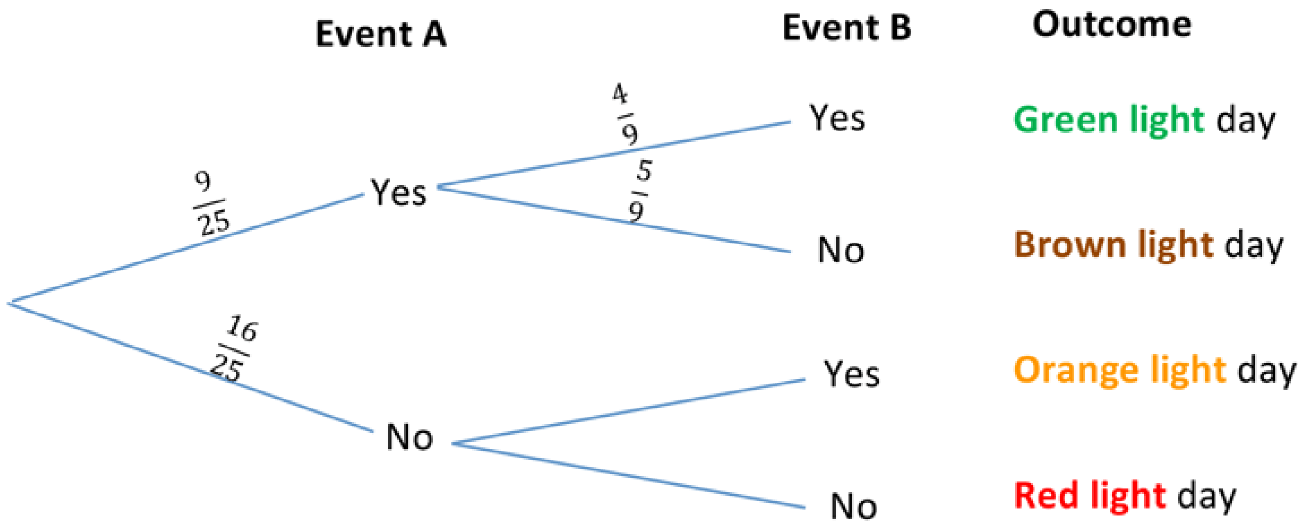
To begin with, the tree diagram for the 1992 season has 2 branches representing the 2 possible outcomes (Yes and No) for Event A.
P(A)=9/25 P(A')=16/25
These probabilities are written on the initial two branches of the probability tree.

Each of these branches then splits into 2, representing the 2 possible outcomes (Yes and No) for Event B, given the outcome for Event A.
P(B | A)=4/9 P(B' | A)=5/9
These probabilities are written on the two branches that split from the initial top branch.

- What probabilities should be written on the branches that split from the initial bottom branch?

Calculate their values. Then write them on these two branches.

The probabilities on the bottom two branches are not the same as the probabilities on the top two branches because Event A and Event B are not independent events.



The probability of the final outcome (the colour of the day) depends on all the probabilities on the branches that lead to that outcome.

For independent events, $P(A \cap B) = P(A) \times P(B)$

For dependent events, $P(A \cap B) = P(A) \times P(B \mid A)$ so $P(B|A) = (P(A \cap B))/(P(A))$ (as seen on page 3)

- The branches of the **green light** day outcome are $P(A) = 9/25$ and $P(B \mid A) = 4/9$

Calculate the probability of a **green light** day using these two values.

- Using the tree diagram , write the calculations needed to find:
- the probability of a **brown light** day in the 1992 season
- the probability of an **orange light** day in the 1992 season
- the probability of a **red light** day in the 1992 season.

The two-way table of frequencies in 2000 (a fairly typical planting season) is shown below.

- Draw a probability tree and calculate the probability of each of the four outcomes.
- For this season, could you have drawn a probability tree with fewer branches?

2000	**Set A**	**Set A'**	**Marginal Totals**
Set B	13	1	**14**
Set B'	10	0	**10**
Marginal Totals	**23**	**1**	**24**

The two-way table of frequencies in 2006 (the warmest of planting season) is shown below.

- Draw a probability tree and calculate the probability of each of the four outcomes.
- For this season, could you have drawn a probability tree with fewer branches?

2006	**Set A**	**Set A'**	**Marginal Totals**
Set B	12	0	**12**
Set B'	11	0	**11**
Marginal Totals	**23**	**0**	**23**

- A farmer with a small property is interested in the probability of a **green light** day occurring.

Does this probability vary much between seasons?

- A farmer with a large property is interested in the probability of a **green light** or a **brown light** or an **orange light** day occurring as he will sow on any of these days.

Calculate the probability of this occurring on a day in the 1992 season, the 2000 season and the 2006 season. Does this probability vary much between seasons?

- If the average soil temperature and the average air temperature over a planting season at Trangie increase in the future,

- would you expect the probability of **green light** days to increase or decrease? Why?
- would you expect the probability of **red light** days to increase or decrease? Why?

back to top

Conclusion: Before you give advice to the farmer

The farmer has employed you to advise him as to whether, in the future, he can expect there to be at least five **green light** days each year during the planting season at Trangie. Given that other conditions for sowing are suitable (eg. the planting equipment is available and it's not so wet on the **green light** days that the cotton planter could get bogged) five **green light** days should give him time to sow his crop without a major yield loss due to temperature conditions.

The farmer would prefer the 5 **green light** days to be consecutive days because of the difficulty of stopping then re-starting the planting operation. In fact, the farmer would rather plant on a **brown light** day or an **orange light** day than stop planting altogether.

Taking the farmer's preferences into consideration, reflect on the three tasks you have completed and advise the farmer. Give the farmer a "Yes" or a "No" answer but also explain how you reached your answer and what assumptions you have made (eg. about temperature trends and missing data). Also let the farmer know whether the planting days are likely to be consecutive.

Note: Because there is missing data in the record (eg. no data for weekends), you will need to work with percentages. For example, there are 43 days in the planting season. If measurements were only taken on 25 of these days and 10 were **green light** days, the percentage of days in the season being **green light** days could be estimated as 10/25 = 40% of the days. 40% of 43 days is about 17 days.

back to top

Teacher Notes & Solutions

The Context

The profitability of growing cotton depends on the prevailing economic conditions and the yield of the crop. Although soils and rainfall influence cotton yields, temperature is the dominant environmental factor.

Temperature conditions vary between locations. In the Trangie district, a farmer needs to be able to sow (i.e. plant) the crop on or between 15 September and 20 October for the yield to be profitable. This period is referred to as the planting season. There are two temperature conditions that need to be considered to determine whether a day in the planting season is suitable for sowing cotton.

Students are given the scenario of farmers in the Trangie district faced with the decision of whether or not they should invest in cotton production. Some of the farmers are more risk-averse than others. The risk-averse farmers will only plant cotton on a day when both temperature conditions are satisfied. The less risk-averse farmers will plant cotton provided that at least one of the two temperature conditions is satisfied. Students analyse historical temperature data at Trangie to determine the relative frequency these conditions having been satisfied in past years.

Before beginning this series of lessons, read the background information provided. You could give this information to your students and/or you could introduce the context using a video or a slide presentation such as the one at https://prezi.com/tzrjuk06rqih/visual-story-of-cotton/ (https://prezi.com/tzrjuk06rqih/visual-story-of-cotton/) which shows how the cotton plant grows.

Cotton is sown in the spring and harvested in the autumn. Emphasise the importance of crop being grown when it will not be damaged by frosts, and that temperature determines the rate at which the cotton plant develops.

Find Trangie on the map of cotton growing districts. Ask them where this district is (north, south, east or west) in relation to other cotton growing districts.

Show the line graph “Relative yield of cotton at Trangie vs. Planting date” in the introduction. From this graph, ask them to find the planting dates for which the cotton grower can expect to get a relative yield of at least 95%. This means that based on previous yields, the yield on average is expected to be within the top 5% of yields.

---

back to top

The Mathematics Tasks and Curriculum Content Descriptors

There are three tasks. These are ordered sequentially, however each task has sufficient information within it to be able to stand alone.

Task 1 – Conditional formatting of Microsoft Excel spreadsheets

In this task, students learn to manipulate data in spreadsheets. The conditional formatting function in Excel is used to identify good days for sowing and spreadsheet formulae are used to count these days. This task also uses scatter graphs to determine linear trends in data **(ACMSP279)**

Task 2 – Constructing and interpreting Venn diagrams and two-way tables

This task revises the use of Venn diagrams and the language and notation used to describe sets of data **(ACMSP204 and ACMSP205)**. It also relates data in Venn diagrams to data expressed in two-way tables **(ACMSP292)**.

Task 3- Calculating probability and conditional probability

This task investigates the concept of relative frequency (experimental probability) **(ACMSP226)** and the concept of dependence and independence **(ACMSP246)** using Venn diagrams, two-way tables and tree diagrams. During the investigative process, emphasis is given to understanding the meaning of dependent events and the correct use of conditional language **(ACMSP247)**.

Detailed plans for the delivery of each task and solutions to the questions posed are given below. Following the launch of each lesson, you will need to decide whether students work individually, in pairs or as small groups to answer the given questions, prior to discussing the solutions as a class.

---

back to top

Task 1 – Conditional formatting of Microsoft Excel spreadsheets

Testing good sowing day - Answers (pdf/testing-good-sowing-days-answers.xlsx)

1) Check that all students understand the language on the webpage http://www.csd.net.au/greenlight (http://www.csd.net.au/greenlight) “Have you got the greenlight for planting this season”

- AEST - Australian Eastern Standard Time
- Forecast average temps – predicted average daily temperatures (found by averaging the maximum and minimum temperature of each day)

- The week following – Note that this doesn't include the day in question
- On a rising plane - trending upwards
- Adjustments - These are measures or actions taken to enable better sowing conditions eg. the planting rate (the amount of seeds sown per hectare) may need to be increased to allow for a potential loss of seedlings if the temperature drops

2) Check that all students understand what is meant by conditionally formatting a spreadsheet.

- Conditionally - subject to, or depending on one or more requirements

3) Show the line graph of "Average daily temperatures at Trangie, 15 September – 20 October".

- Check that all students know the difference in meaning between daily average temperatures and average daily temperature.

The average daily temperature points in the graph are daily average temperatures that have been averaged over the days in the period from 15 September – 20 October (36 days) each year.
From the graph, ask them to find the maximum and minimum values for each line and the year in which these values occurred.

4) Show the spreadsheet Trangie Planting Season Data, or ask students to open it.

- Why do you think this pattern happens?

The pattern happens because the staff at Trangie Agricultural Research Station who recorded the measurements, were not employed on weekends.
In October there is missing data for 3 days in a row because of the October long weekend.
The diagonal pattern is formed because in consecutive years, the days of the week go forward by one (a year being 52 weeks and 1 day). When it is a leap year (52 weeks and 2 days) the day of the week goes forward by 2 days.

5) Ask students to open the spreadsheet Testing for Good Sowing Days.
Point out that there are 3 worksheets, one for each of 3 years.
Revise spreadsheet skills.

6) Give students the Task 1 sheet and ask them to follow the instructions (beginning on page 3) to conditionally format the data so they can see which days are good days for sowing cotton.
In the sheet, bolded numbers or letters are those to be typed into a box.
Questions to answer are marked with a dot.

- On how many days in the 1992 planting season, was the soil temperature measured? 25 days

- How many of these days have a soil temperature greater than 14°C? 9 days
- What fraction of the days when soil temperature was measured, have a soil temperature greater than 14°C? Write your answer as a decimal correct to 2 decimal places. 0.36

- By using data in this spreadsheet as “a forecast of average temperatures”, what assumption has been made?

It is assumed that the temperatures forecast were those that actually happened.

- How many of the days with soil temperature data, have mean temperatures for the following week on a rising plane? 15 days
- What fraction of the days with soil temperature data, have average temperatures for the following week on a rising plane? Write your answer as a decimal correct to 2 decimal places. 0.6

- From conditionally formatting your spreadsheet:
- how could you tell which days are **green light** days?

In their row, they have a red border for the cells in both Column B and Column D.

- how could you tell which days are **red light** days?

In their row, they have a no cells with a red border.

- how could you tell which days are **orange light** days?

In their row, they have only one cell with a red border.

- What are the two conditions for **red light** days?

A soil temperature of 14°C or less, and average temperatures for the following week not rising.

- Write this rule as an Excel formula to conditionally format these days.

(Note that the symbols <= written together mean ”less than or equal to”.)

=AND($B3<=14,$D3=”N”)

- What are the two conditions for **brown light** days?

A soil temperature greater than 14°C and avearge temperatures for the following week not rising.

- Write this rule as an Excel formula to conditonally format these days.

=AND($B3>14,$D3=”N”)

- What are the two conditions for **orange light** days?

A soil temperature of 14°C or less and average temperatures for the following week rising.

- Write this rule as an Excel formula to conditonally format these days.

AND($B3<=14,$D3="Y")

To count the number of **green light** days, go to Cell E51 and type the formula **=COUNTIFS(B3:B38,">14",D3:D38,"=Y")**.

- Explain the meaning of the symbols in this formula.

There is more than one condition, so the formula is =COUNTIFS (not =COUNTIF) and the two conditions are separated by a comma.
Each condition is written in inverted commas after the location of the array of cells to which it is to be applied and a comma.
The formula says to count the number of days (rows) from row 3 to row 38 if they have a value of greater than 14 in Column B and if they have a Y in Column D.

- What formula would you type to get a count of the **red light** days into Cell E52?

=COUNTIFS(B3:B38,"<=14",D3:D38,"=N")

- What formula would you type to get a count of the **brown light** days into Cell E53?

=COUNTIFS(B3:B38,">14",D3:D38,"=N")

- What formula would you type to get a count of the **orange light** days into Cell E54?

=COUNTIFS(B3:B38,"<=14",D3:D38,"=Y")

- Why is it better to compare the proportion of days in each category rather than compare the number of days in each category?

Using proportions makes possible a comparison between different seasons, even when the number of days in each season for which there is data, differs.

- How many days are there in a planting season (15 September – 20 October)? 36 days
- What is the proportion of **green light** days in the 1992 season? 4/25 = 0.16 of the days.
- In the coolest planting seasons, how many **green light** days are expected?

$0.16 \times 36 = 6$ days (to the nearest day)

- What is the proportion of **orange light** days in the 1992 season? 16/25=0.64 of the days.
- In the coolest planting seasons, how many **orange light** are expected?

0.64×36=23 days (to the nearest day)

- In the coolest planting seasons, on how many days during the coolest planting seasons are farmers with larger properties likely to sow cotton?

This is the number of green light days plus the number of orange light days, which together is
0.64+0.16=0.8 of the days
0.8×36=29 days (to the nearest day)

Task 2 – Constructing and interpreting Venn diagrams and two-way tables

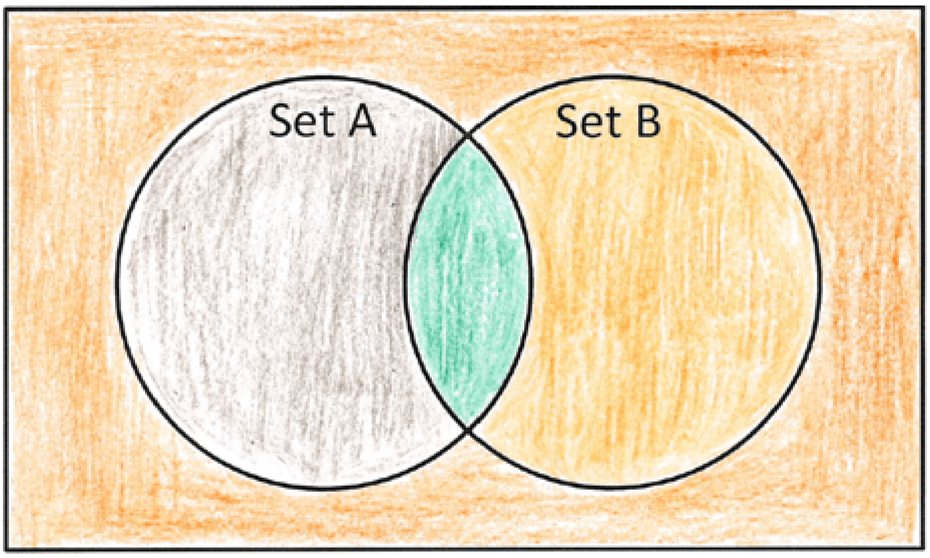
1) Remind students of the context and what is meant by the various colours of day.
Emphasise the words in italics (eg. both, neither).

2) Check that all students understand what is meant by

- The condition is satisfied – the requirements have been fulfilled
- A set – a collection of distinct “things” (eg. numbers or objects) that have something in common

3) Give them the Task 2 sheet and ask them to follow the instructions and answer the questions.

<u>Venn Diagrams</u>



The two circles in the Venn diagram intersect (i.e. overlap) because some days satisfy Condition A and Condition B and are therefore in both Set A and Set B.

- What colour light was given to these days? Green

On some days, neither Condition A nor Condition B was satisfied.

- What colour light was given to these days? Red

On some days, Condition A was satisfied but Condition B was not satisfied.

- What colour light was given to these days? Brown

On some days, Condition A was not satisfied but Condition B was satisfied.

- What colour light was given to these days? Orange

<u>Set Language and Notation</u>

- What is the total number of days represented in the Venn diagram for the 1992 season? 25

- What colour days are in Set B? orange and green
- What does n(B) equal? n(B)=11+4=15

The Venn diagram below shows the data for the 2000 planting season.

- Shade it with a pencil to help you find the values of the following:

$n(A \cup B) = 24$

$n(A \cap B) = 13$

$n(A' \cup B) = 14$

$n(A' \cap B) = 1$

$n(A' \cup B') = 11$

$n(A' \cap B') = 0$

- Is the rule $n(A \cup B) = n(A) + n(B) - n(A \cap B)$ true for the 2000 planting season?

Yes, because 24 = 23 + 14 – 13

Is this rule true for the 1996 planting season?

For the 1996 season: 20 = 9 + 15 – 4

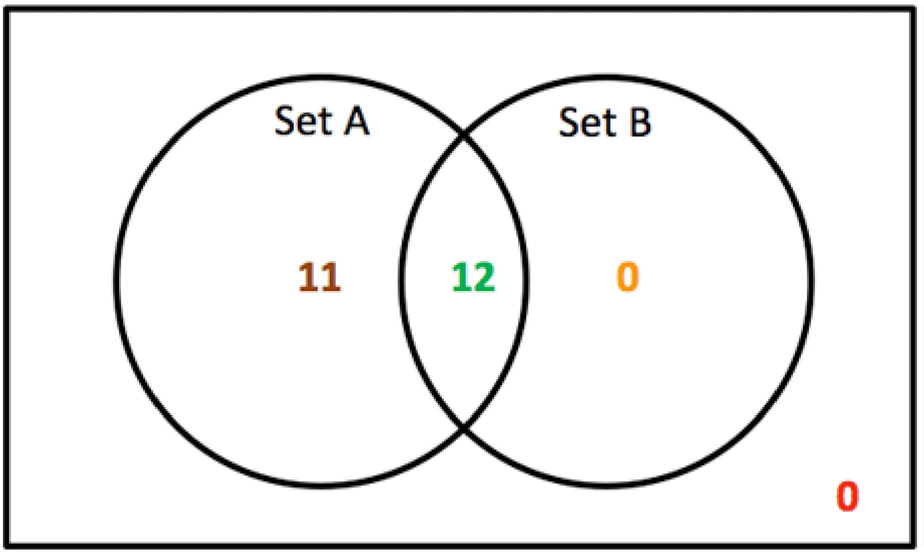
Is this rule true for any number of elements in Set A and Set B? Explain why or why not.

Yes. The intersection is doubled when the two sets overlap so if its value is subtracted from the total of the two sets, it becomes the same as just each of the three regions added together.

Would this rule apply if there were no **green light** days (i.e. no overlap of the two sets)?

Yes. In this case n(A ∩ B) = 0

2006 planting season:



- How many days are in Set A? 12
- How many days are in Set B? 12

Two-way Tables

Colour the 4 empty cells of the table above as follows.

green light days satisfy Condition A and Condition B, so colour the cell that is in the Set A column and in the Set B row, green.

	Set A	**Set A'**
Set B		
Set B'		

Colour the remaining 3 cells to indicate the sets in which there are **red light** days, **brown light** days and **orange light** days.

1992	**Set A**	**Set A'**	**Marginal Totals**
Set B	4	11	15

Set B'	5	5	10
Marginal Totals	9	16	25

- What does the total of Set B mean in this context?

There are 15 days when temperatures were on a rising plane.

- What does the total of Set A' mean in this context?

There are 16 days when temperatures are not above 14°C.

- What is the sum of the two column totals? 25
- What is the sum of the two row totals? 25
- Why should the sum of the columns totals, be the same as the sum of the row totals?

Because every day has to be either in Set A or not in Set A, and every day has to be in Set B or not in Set B. Both these values are total of all the days (the universal set).

- Using the Venn diagrams for 2000 and 2006, make a two-way table for each of these seasons.

2000	**Set A**	**Set A'**	**Marginal Totals**
Set B	13	1	**14**
Set B'	10	0	**10**
Marginal Totals	**23**	**1**	**24**

2006	**Set A**	**Set A'**	**Marginal Totals**
Set B	12	0	**12**

Set B'	11	0	**11**
Marginal Totals	**23**	**0**	**23**

- Which of the three seasons (1992, 2000 or 2006) has the greatest proportion (i.e. fraction) of days in Set B?

1992: 15/25=3/5=0.6

2000: 14/24=7/12≈0.58

2006: 12/23≈0.52

1992 has the greatest proportion of days when the following week had temperatures on a rising plane.

- Did you expect all the seasons to have more than half of their days in Set B? Why or why not?

Yes. Because September and October are months of spring, temperatures overall are expected to be rising.

Task 3 – Calculating probability and conditional probability

1) Remind students of the context and what is meant by the various colours of day.
Emphasise the words in italics (eg. both, neither).

2) Check that all students understand the language of probability.

- Event
- Outcome
- Frequency
- Sample space
- Relative frequency
- Experimental probability

3) Introduce the words mutually exclusive using examples of mutually exclusive events and mutually exclusive outcomes.
eg. If throwing a dice is one event and drawing a card is another event, these events have no common elements (eg. you can't draw an ace by throwing a die) .
eg. For the event of throwing a die once, the outcomes of throwing an odd number and the

outcome of throwing an even number are mutually exclusive (because no number is both odd and even).
Ask students to find their own examples of mutually exclusive events and mutually exclusive outcomes.

4) Give students the sheet for Task 3. Ask them to read the first page and then answer the questions about probabilities on page 2.
Probabilities

- Are the two events (a day when soil temperature is greater than 14°C and a day when average temperatures for the following week are on a rising plane) mutually exclusive? Why or why not?

No. These events are not mutually exclusive because on the the same day there can be a soil temperature is greater than 14°C and average temperatures for the following week on a rising plane

- For Event A (whether soil temperature is greater than 14°C) what are the two possible outcomes?

The two outcomes are:

- A soil temperature greater than 14°C
- A soil temperature not greater than 14°C (i.e. less than or equal to 14°C)

Are these outcomes, mutually exclusive? Why or why not?

Yes, because a temperature cannot be both greater than 14°C and less than or equal to 14°C.

- How many days are in the sample space? 25
- Calculate the probabilities.

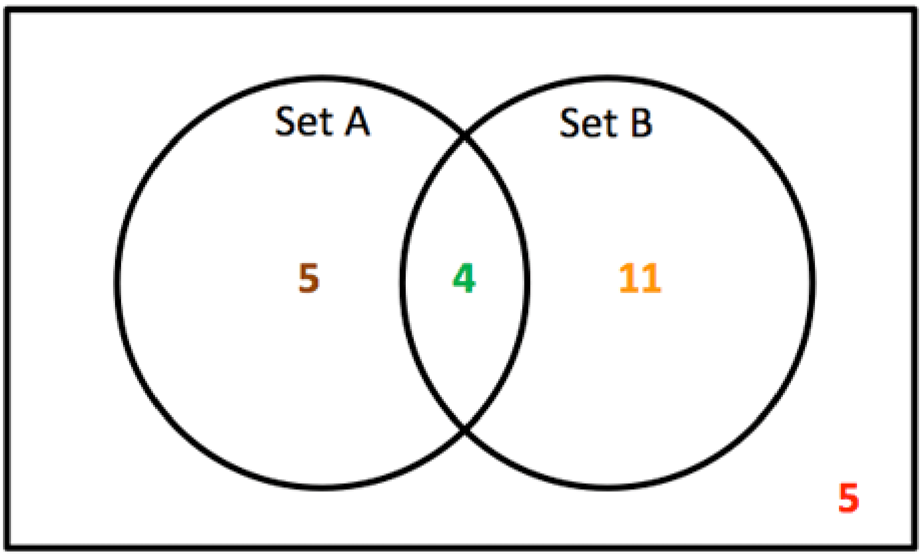
- $P(A' \cap B') = 5/25 = 1/5$
- $P(A \cap B') = 5/25 = 1/5$
- $P(A' \cap B) = 11/25$

4) Ask them to read page 3 which is about conditional probabilities.
Work through this page with them using the table below and the Venn diagram for the 1992 season, so they understand the meaning of:

- The sample space and a restricted sample space
- Marginal total

- Dependent and independent events
- Conditional probability
- Initial condition

1992	**Set A**	**Set A'**	**Marginal Totals**
Set B	4	11	**15**
Set B'	5	5	**10**
Marginal Totals	**9**	**16**	**25**



5) Get students working in pairs to answer the questions on page 4.

Conditional Probabilities

- With a partner, decide whether the following sentences are true or false, i.e. whether they mean the same as The probability that a day is in Set B, given that day is in Set A is **4/9**.

1) If a day is in Set B, then the probability that it is in Set A is 4/9 True

2) If a day is in Set A, then the probability that it is in Set B is 4/9 False

3) Given that a day is in Set B, the probability that it is in Set A is 4/9 True

4) Knowing that a day is in Set B, the probability that it is also in Set A is 4/9 False

5) If you already know that a day is in Set A, the probability that it is in Set B is 4/9 True

- What is meant by **P(B|A') = 11/16**?

This is the probability that a day will have a rising plane of temperatures the following week, if you already know that the soil temperature that day is less than or equal to 14°C.

- What is meant by P(A|B)? What is the value of P(A|B) in the 1992 season?

This is the probability that the soil temperature on a day is greater than 14°C , if you already know that there will a rising plane of temperatures over the following week.
P(A | B) = 4/15

- What is meant by P(A|B')? What is the value of P(A|B') in the 1992 season?

This is the probability that the soil temperature on a day is greater than 14°C , if you already know that there will a rising plane of temperatures over the following week.
P(A | B') = 5/10 = 1/2

- With a partner, decide whether the pairs of events below are dependent or independent.

Read information at http://cottonaustralia.com.au/australian-cotton/basics/cotton-facts (http://cottonaustralia.com.au/australian-cotton/basics/cotton-facts).
Then share your answers with others so you can convince them or be convinced by them.
(i) Whether a cotton crop is genetically modified, and the quantity of insecticide used on it.
On the webpage, if students click the link Biotechnology and Cotton, they will find that:

- The use of biotechnology in cotton has made a significant contribution in the dramatic reduction in insecticides applied to Australian cotton crops.
- Australian cotton growers have reduced their insecticide use by 89% over the last decade, with some crops not sprayed for insects at all.

Genetically modified cotton crops are crops where the cotton variety has been given a gene that reduces its susceptibility to insects, thereby reducing the amount of insecticide needed to control the insects that attack it.
Therefore the two events are dependent.

(ii) Whether an area is sown to cotton or left fallow, and the emission of greenhouse gases.
On the webpage, if students click the link Climate Challenges and Cotton, they will find that:

- Cotton growing has a better-than-neutral carbon footprint. Net on-farm emissions of greenhouse gases on cotton farms are negative because the cotton plants store more carbon than is released from production inputs used during growth.

“Left fallow” means that nothing is grown on the land.
When a crop grows, it uses the process of photosynthesis to take in carbon dioxide from the atmosphere and produce plant material. This is why an area sown to cotton takes in more carbon than is emitted by the processes of growing it (such as the energy used to plant and fertilise it).
Therefore the two events are dependent.

(iii) The weight of a bale of cotton, and what that bale of cotton is used to make.
On the webpage, if students click the link Properties and Cotton Products, they will find

- The cotton lint from one 227 kg bale can produce 215 pairs of denim jeans, 250 single bed sheets, 750 shirts, 1,200 t-shirts, 3000 nappies, 4,300 pairs of socks, 680,000 cotton balls, or 2,100 pairs of boxer shorts

227 kg is the statistical average weight of a bale of cotton in Australia. In other countries, the weight is different eg. In South Africa, cotton bales weigh about 200 kg. The weight of the bale has no effect on what the cotton is used to make.
Therefore the two events are independent.

(iv) Rain occurring on a day at Trangie, and rain occurring on the same day at Narromine.

If you find these two places on a map of NSW you will notice that they are close to each other. They are only 34 km apart by road, so if it is raining at one of these places it is likely to be raining at the other.
Therefore the two events are dependent.

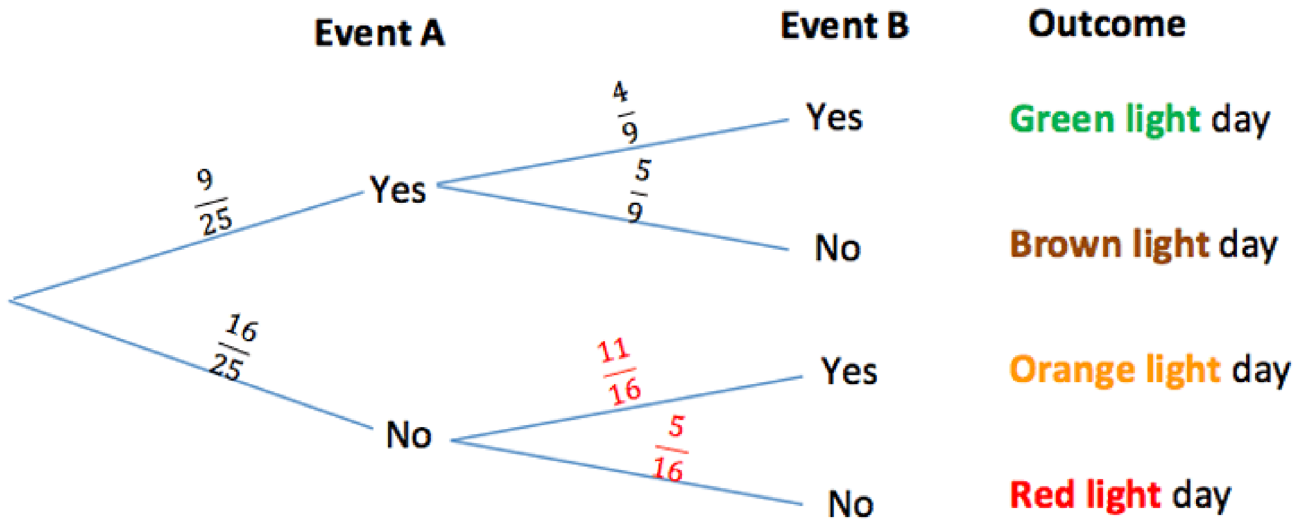
6) Revise the topic of probability trees. Then ask students to answer the questions on Pages 5 & 6.

Probability Trees

- What probabilities should be written on the branches that split from the initial bottom branch?

Calculate their values. Then write them on these two branches.

The probabilities on the bottom two branches are not the same as the probabilities on the top two branches because Event A and Event B are not independent events.



- The branches of the **green light** day outcome are P(A) = 9/25 and P(B | A) = 4/9.

Calculate the probability of a **green light** day using these two values.<.p>

P(green light day) = P(A)×P(B) = 9/25×4/9 = 4/25

- Using the tree diagram , write the calculations needed to find:
- the probability of a **brown light** day in the 1992 season

P(brown light day) = P(A)×P(B') = 9/25 × 5/9 = 5/25 = 1/5

- the probability of an **orange light** day in the 1992 season.

P(orange light day) = P(A')×P(B) = 16/25×11/16 = 11/25

- the probability of a **red light** day in the 1992 season.

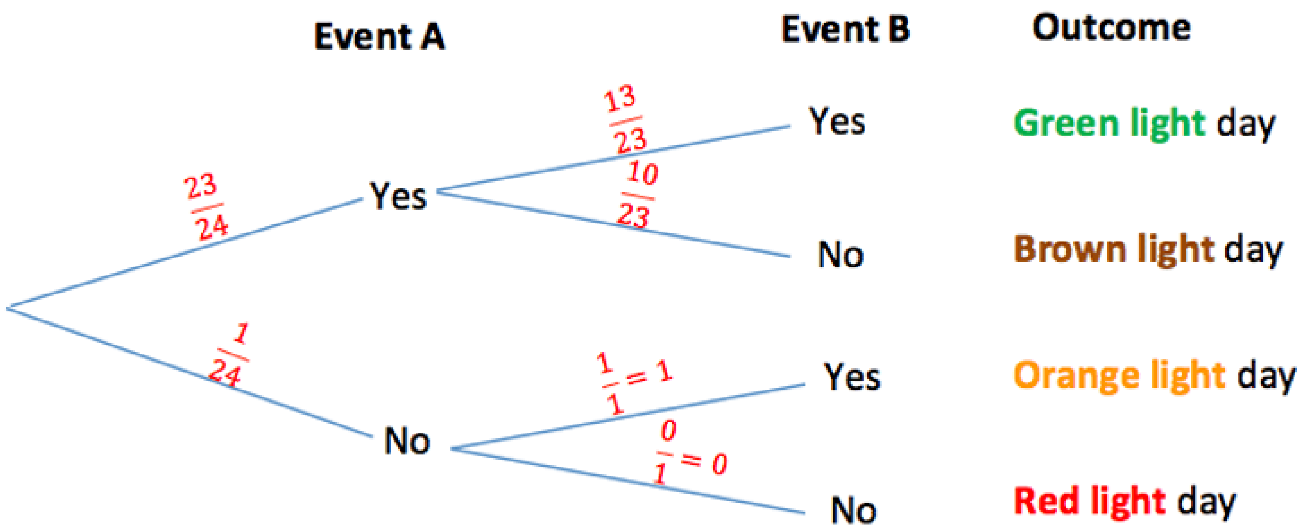
P(red light day) = P(A')×P(B') = 16/25×5/16=5/25=1/5

The two-way table of frequencies in 2000 (a fairly typical planting season) is shown below.

- Draw a probability tree and calculate the probability of each of the four outcomes.

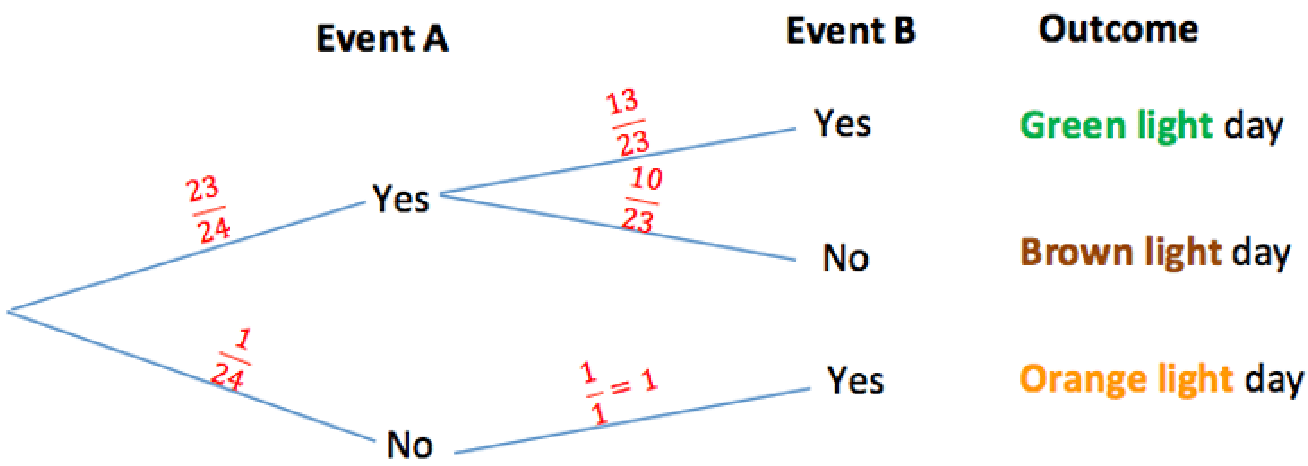
2000	**Set A**	**Set A'**	**Marginal Totals**
Set B	13	1	**14**
Set B'	10	0	**10**

Marginal Totals	**23**	**1**	**24**



- For this season, could you have drawn a probability tree with fewer branches?

Yes. The last of the 4 branches isn't needed because it has a zero probability.



The two-way table of frequencies in 2006 (the warmest of planting season) is shown below.

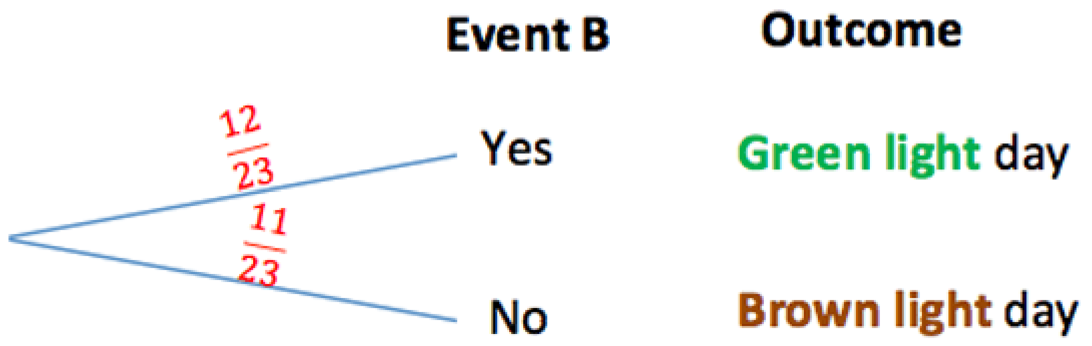
- Draw a probability tree and calculate the probability of each of the four outcomes.

2006	**Set A**	**Set A'**	**Marginal Totals**

Set B	12	0	**12**
Set B'	11	0	**11**
Marginal Totals	**23**	**0**	**23**

- For this season, could you have drawn a probability tree with fewer branches?

Yes. Because Event A is certain i.e. P(A)= 1, only Event B needs to be considered.



- A farmer with a small property is interested in the probability of a **green light** day occurring.

Does this probability vary much between seasons?

1992: 4/25 = 0.16

2000: 13/24 ≈ 0.54

2006: 12/23 ≈ 0.52

There is a large variation due to the number of days with a temperature above 14°C increasing greatly from the coolest season (1992) to an average type of season (2000). For warmer than average seasons, little variation would be expected.

- A farmer with a large property is interested in the probability of a **green light** or a **brown light** or an **orange light** day occurring as he will sow on any of these days.

Calculate the probability of this occurring on a day in the 1992 season, the 2000 season and the 2006 season. Does this probability vary much between seasons?

This is the same as the complement of the probability of a red light day occurring.

1992: 5/25 = 0.2

2000: 24/24 = 1

2006: 23/23 = 1

This probability varies a lot from the coolest season to the average season. It is unlike to vary much once when seasons are warmer than average because in these seasons Condition A (soil temperature above 14°C) is always (or nearly always) satisfied.

- If the average soil temperature and the average air temperature over a planting season at Trangie increase in the future,
- would you expect the probability of **green light** days to increase or decrease? Why?
- would you expect the probability of **red light** days to increase or decrease? Why?

Soil temperatures increasing will increase the probability of green light days and brown light days. However there could still be orange light days because whether a temperature is on a rising plane or not over the next week, is not greatly affected by the temperature of the day. There can still be a very hot day followed by a week of decreasing temperatures.