GCSE Mathematics Higher

AQA Specification 8300

Part 6

Chapter Twelve: Probability

12.1 Probability Experiments

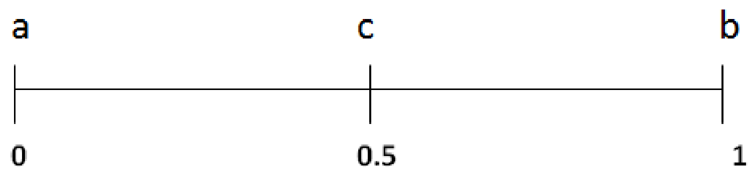
Probability Scales and Notation

Probability refers to how likely an event is to happen. It is measured as a decimal or a fraction or a percentage. It is advised that you convert to a decimal or a fraction if you're given a probability as a percentage.

The value is always between 0 and 1. A probability of 0 means that the event is impossible whereas a probability of 1 means that the event is certain.
These values can be thought of as lying on a probability scale ranging from 0 to 1.

Example:
Draw a probability scale illustrating the following
a. The probability of obtaining an 8 when rolling a standard die.
b. The probability of obtaining a number between 1 and 6 when rolling a standard die.
c. The probability of obtaining an odd number when rolling a standard die.

a. This is impossible so the probability is 0
b. This is certain so the probability is 1
c. There is an even chance of getting an odd number as there are three even numbers and three odd numbers. You will often hear this referred to colloquially as "a 50-50 chance". The probability is 0.5 or ½



These scales can also be used to represent estimated probabilities. For example, if asked to represent the probability that it will be sunny on a July day in Florida you would mark this near to the 1 on the scale.

Writing "the probability of.... is.." gets a bit tedious so there is a shorthand we use instead. The capital letter P stands for "probability of" which is then followed by the event in brackets and what it equals.
For example the probability of getting a head when tossing a fair coin is ½ or 0.5 so we write P(Head) = 0.5 or we could write P(H)=0.5
This can be used any event. As another example we could say that the probability of getting a 2 when rolling a fair coin is $\frac{1}{6}$ as there are 6 sides and only one side contains a 2.
In this case we write P(2) = $\frac{1}{6}$

As seen above the value of the probability is calculated by dividing the number of favourable trials – this is situations in which the required event arises – by the total number of possible outcomes.

Example:
In a bag there are 3 blue balls, 5 red balls and 2 yellow balls. Find the probability of choosing each of these colours.

We can see that there are 3 + 5 + 2 = 10 balls altogether so we have
P(B) = $\frac{3}{10}$ P(R) = $\frac{5}{10} = \frac{1}{2}$ P(Y) = $\frac{2}{10} = \frac{1}{5}$

As you can see, probability is a topic that requires a good understanding of fractions and decimals. Make sure you're confident with adding and multiplying fractions and decimals before continuing.

Activity 12.1 A

1. Draw a probability scale. Mark the probability of each of these events happening on the scale – use the question letter to denote each event.
   a. Getting a head when tossing a fair coin
   b. Getting a diamond from a standard pack of playing cards (jokers removed)
   c. Picking a 15 from a standard pack of playing cards (jokers removed)
   d. It will be Friday before Saturday next week
   e. Tuesday will immediately follow Sunday
   f. It will snow in California in August
   g. It will snow in Chicago in December

2. Using appropriate notation write down the probability of each of these events happening
   a. Getting a tail when tossing a fair coin
   b. Getting an even number when rolling a fair die
   c. Getting a 3 when rolling a fair die
   d. Picking a heart from a standard set of playing cards (jokers removed)

3. A bag contains 26 pieces of paper each with a different letter of the alphabet on
   a. What is the probability of picking a B?
   b. Write down the value of P(Vowel)

4. A bead is drawn randomly from a jar that contains 4 brown beads, 2 pink beads, and 3 blue beads. Find the probability of selecting
   a. A brown bead
   b. A pink bead
   c. A blue bead
   d. A red bead

5. A ball is drawn randomly from a jar that contains 5 pink balls, 2 purple balls, and 13 white balls. Find the probability of selecting
   a. A pink ball
   b. A ball that isn't white
   c. A ball that is pink, purple or white

6. A fair six sided dice is rolled once. Write down the value of P(less than 3).

7. Daisy chooses one letter at random from the word MATHEMATICS. Write down the value of
   a. P(M)

b. P(A)
c. P(D)

<u>Expected Outcomes</u>

If we know a probability of an outcome occurring we can calculate the amount of times we expect this outcome to occur within a given number of trials. This is found by multiplying the number of trials by the probability. The higher the number of trials the more accurate the theoretical outcome.

Example:
A fair die is rolled 50 times. Calculate the number of times you would expect to roll a three.

We know that P(3) = $\frac{1}{6}$ so the number of times we would expect to roll a three when the die is rolled 50 times is
$50 \times \frac{1}{6} = 8.33\dot{3}$
Therefore we would expect to roll a 3 eight times when a fair die is rolled fifty times.

<u>Mutually Exclusive Events</u>

Events are mutually exclusive if they cannot happen at the same time. For example getting an odd number or an even number with a single roll of a die are mutually exclusive events as it is impossible for both of these things to happen at the same time; the number rolled is either odd or even, it cannot be both.

If a set of mutually exclusive events are also exhaustive – that is they encompass all possibilities – their probabilities will sum to 1. For example, when tossing a coin the outcomes are either head or tail, this list of outcomes is exhaustive as there are no other options therefore the probabilities sum to 1.
(Which you knew already because P(H) = ½ and P(T) = ½ so we have P(H) + P(T) = ½ + ½ = 1)

Example:
The probability of winning a game is 0.4. The probability of drawing in the same game is 0.3. What is the probability of losing?

Since win, draw and lose are mutually exclusive and form a complete set of all the possible outcomes we know that the probabilities sum to 1.

P(Win) + P(Draw) + P(Lose) = 1
0.4 + 0.3 + P(Lose) = 1
0.7 + P(Lose) = 1

∴ P(Lose) = 0.3

The above example illustrates an important rule with regards to mutually exclusive events:
The probability of an event happening = 1 – probability of the event not happening

This is also true in reverse:
The probability of an event not happening = 1 – probability of the event happening

Example:
The probability that a child in a class has a pet dog is $\frac{5}{8}$. What is the probability that a child in the same class does not have a dog?

These scenarios are mutually exclusive and exhaustive events as a child either has a dog or does not.
P(no dog) = 1 – P(dog)
P(no dog) = $1 - \frac{5}{8} = \frac{3}{8}$

Using this rule can make it easier to find larger probabilities by finding the complementary event.

Example:
A fair coin is tossed four times, find the probability of getting at least one head.

There are $2^4 = 16$ possible outcomes. You could write all of these out as follows:
HHHH, THHH, HTHH, HHTH, HHHT, TTHH, HTTH, HHTT, HTHT, THTH, THHT, HTTT, THTT, TTHT, TTTH, TTTT (this complete list of all possible outcomes is called the **sample space**)
Then count the number of outcomes that include at least one head which gives $\frac{15}{16}$
Or you can think that P(at least one head) = 1 – P(no heads)
Only one option gives no heads so we have $1 - \frac{1}{16} = \frac{15}{16}$

Activity 12.1 B

1. A fair die is rolled 20 times, find the amount of times you would expect to get:
    a. The number 6
    b. A number less than 5
    c. An even number
    d. The number 7

2. A spinner is divided into four equal sections coloured red, blue, green and yellow. If it is spun 20 times how many times would you expect it to land on red?

3. Jo drives down the same road every day. She has had to stop at the lights 20 times in the last 24 journeys. What is the probability that she'll have to stop at the lights on her next journey down the same road?

4. A coin is tossed 20 times. It lands on heads 15 times and tails 5 times. It's decided that the coin must be biased. Do you agree with this decision?

5. A biased coin is tossed 50 times – the ratio of heads to tails is 1:4. How many times would you expect to see tails?

6. For each of these pair of events decide whether they are mutually exclusive or not. If they are, decide whether they are exhaustive
    a. A child has a dog / a child has a cat
    b. A coin is tossed twice: you get two heads / you get two tails
    c. A fair die is rolled twice: the score is the same both times / the total is odd
    d. A fair die is rolled: you get a prime number / you get an even number
    e. A fair die is rolled: you get an odd number / you get an even number

7. The probability of choosing each coloured ball is shown in the table below. Find the value of x

Blue	Red	Yellow	Green
0.2	0.1	x	0.6

8. A group contains men, women and children. If a person is called at random the probability of calling a man is 0.4 and the probability of calling a woman is 0.35.

   a. What is the probability of calling a child?

   b. If there are 30 people in the group, how many would you expect to be men?

9. A fair coin is tossed 3 times.

   a. Write down all the possible outcomes

   b. Find the probability of at least one tail

10. Two fair dice are rolled together and their scores are added

   a. Draw a table illustrating all of the possible outcomes

   b. Find P(7)

   c. Find P(not 7)

11. A bag contains only red, white and blue marbles. There are three times more white than red and twice as many blue as there are white. When a marble is chosen at random find the probability of choosing each of these colours.

12.2 Sets

Notation

A set is, quite simply, a collection of numbers or objects. These objects that are in a set are called members or, more commonly, elements. A set is denoted with "curly brackets".
For example the set of possible outcomes when rolling a single fair die is{1,2,3,4,5,6}. As another example, if X is the set of all even numbers we have X = {2, 4, 6, 8....}

If we want to denote that a number or object is an element of a set we use the symbol ∈
For example, since 34 is an element of the set X from above we would write 34 ∈ X or
34 ∈ {2, 4, 6, 8...}

The **universal set** that contains all elements is denoted by ε
The **empty set** that contains no elements is denoted by Ø

The **intersection** between two sets is denoted by ∩. This is usually read as "and".
For example if we have A = {1,2,3,4} and B = {2,4,6,8}, A∩B is the set that contains elements in both A and B then we have A∩B = {2,4}

The **union** of two sets is denoted by ∪. This is usually read as "or".
For example, if we have A and B as above, A∪B is the set that contains elements that are in A or B so A∪B = {1,2,3,4,6,8}

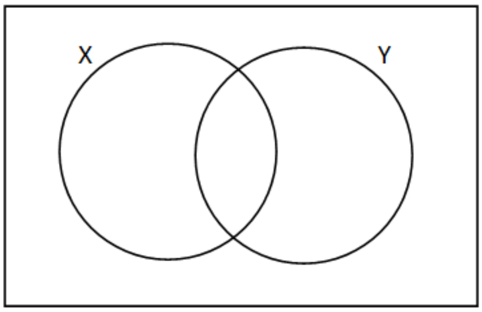
The **complement** of a set is the elements that are not in the set, this is denoted by '.
For example, if A is the set of all even numbers A = {2,4,6,8,10...} then A' is all the elements that are not in A so we have A'= {1,3,5,7,9...}

Venn Diagrams

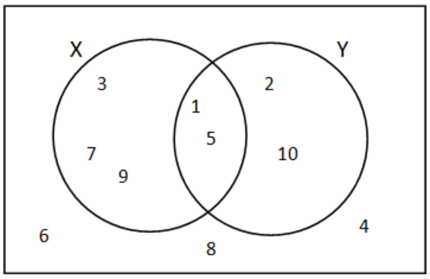
Sets can be illustrated on Venn diagrams. These comprise of a box surrounding two or more circles, depending on the number of sets involved. Each circle is labelled as a set, elements in each set are placed in the respective circles. Any elements that belong to both sets go in the overlap between the circles and any elements that don't belong to either set go inside the box but outside the circles.

Example:
Given the sets below draw a Venn diagram
ε = {1,2,3,...8,9,10}
X = {odd numbers}
Y = {factors of 10}

The first thing we need to do is draw the empty Venn diagram – this is a box containing two circles since we have two sets X and Y (remember that ε is the universal set).



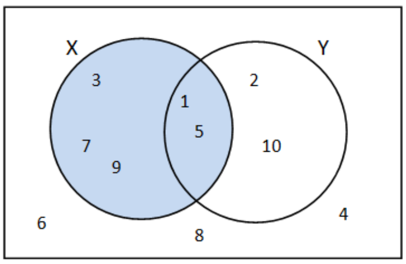
Now all you have to do is fill the numbers in – have a go yourself before looking at the finished diagram below.
Any numbers that are only in X go in the X circle, similarly for Y. Any numbers that are in both go in the intersection and any numbers that are in the universal set but are not elements of either X or Y go outside the circles.



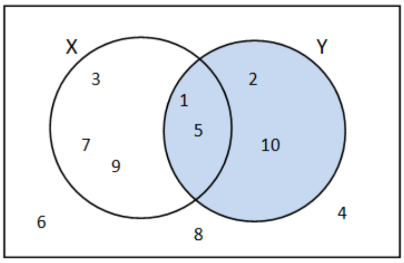
You can use Venn diagrams to work out probabilities. The number of elements in a set is divided by the total number of elements in the universal set to give the value of the probability.

The following probabilities have been calculated from the above Venn diagram

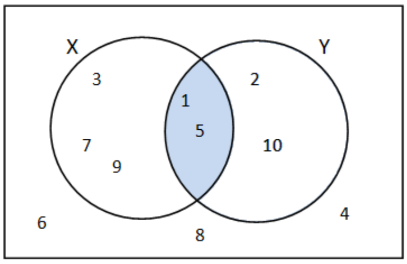
P(X) = $\frac{5}{10} = \frac{1}{2}$ This includes all of the values in the X circle – the ones only in X and in the overlap.



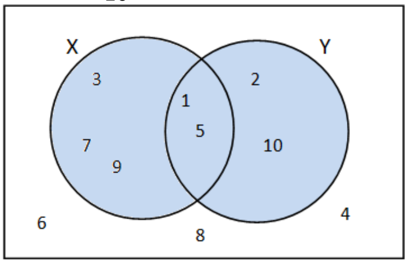
P(Y) = $\frac{4}{10} = \frac{2}{5}$



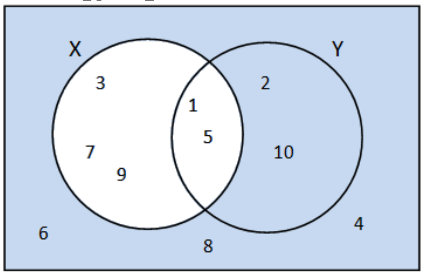
P(X∩Y) = $\frac{2}{10} = \frac{1}{5}$ Remember that ∩ is the intersection between the sets



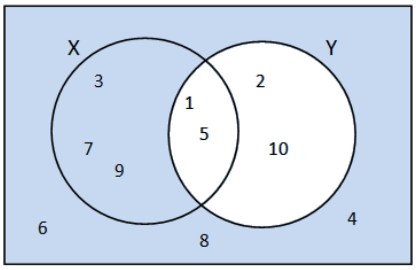
P(X∪Y) = $\frac{7}{10}$ Remember that ∪ is the union of the sets



P(X') = $\frac{5}{10} = \frac{1}{2}$ X' is all the elements not in X

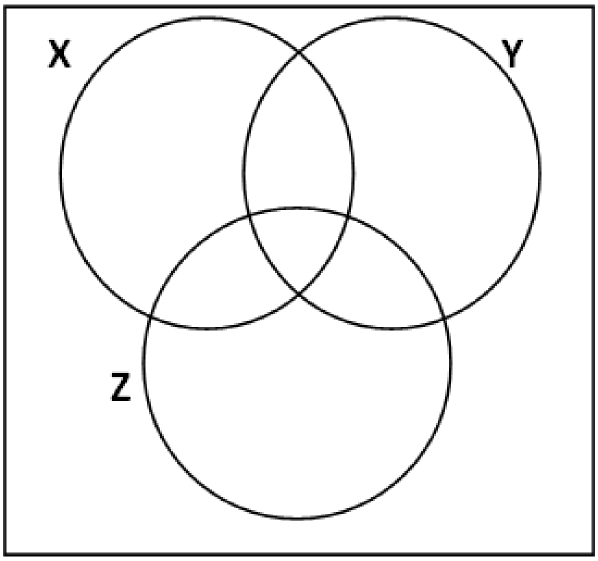


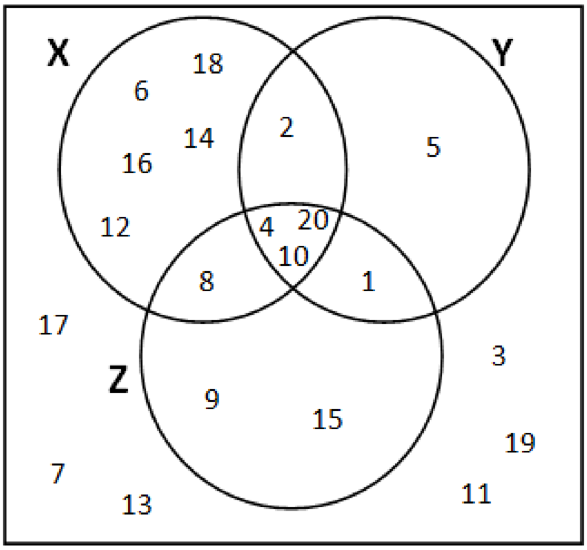
P(Y') = $\frac{6}{10} = \frac{3}{5}$



Example:
Draw a Venn diagram for the sets
ε = {1,2,3,...,19,20}
X = {Even numbers}
Y = {Factors of 20}
Z = {1,4,8,9,10,15,20}

The empty Venn diagram is shown below. Have a go at filling it in yourself before looking at the completed diagram beneath it.

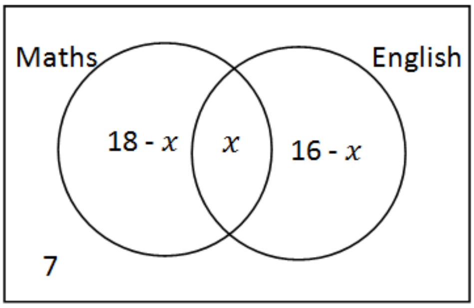




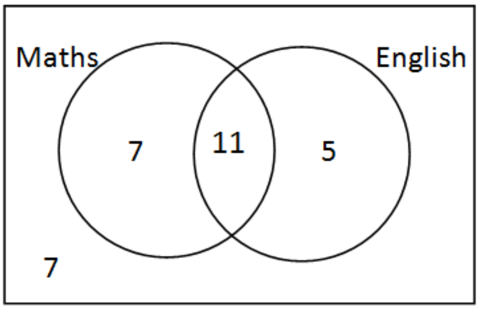
Sometimes, with Venn diagrams, you are given incomplete information and you have to use what you've been given to find the missing values. This is done with algebraic logic as is illustrated in the following example.

Example:
In a group of 30 sixth form students 18 study maths and 16 study English. There are 7 in this group that study neither maths nor English. Find the number of students that study both subjects.

First draw a Venn diagram, denoting the missing value with x



As you can see the numbers inside each circle are the totals with the centre value subtracted from them.
Now we know that there are 30 students in total so we have
$(18 - x) + x + (16 - x) + 7 = 30$
$18 - x + x + 16 - x + 7 = 30$
$41 - x = 30$
$x = 41 - 30 = 11$
So the number of students that study both subjects is 11 and the completed Venn diagram is



Activity 12.2

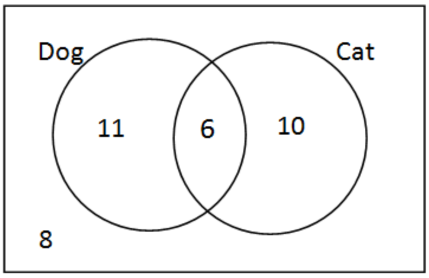
1. X = {Square numbers up to 100}, Y = {Even numbers less than 20}, Z = {Factors of 100}. By listing the elements of each set or otherwise find

a. X∩Y b. X∩Z c. Y∩Z d. X∪Y e. X∪Z f. Y∪Z

2. Given that X = {Even numbers}, which of the following are subsets of X?

a. A = {Square numbers}
b. B = {2,4,6,8}
c. C = {1,2,4,6}
d. D = {5,10,15,20}
e. E = {Multiples of 10}

3. The Venn diagram below shows information about pets owned by children in a class.



a. How many children own both a dog and a cat?
b. How many children own a dog?
c. How many children are there altogether?
d. How many children own a cat?
e. How many children don't own a dog or a cat?
f. How many children have either a dog or a cat?

4. ε = {1,2....12,13} A = {Factors of 12} B = {1,3,5,7,10}
a. Draw a Venn diagram
b. Find P(A)
c. Find P(A∩B)
d. Find P(B')

5. There are 50 people in a gymnastics club. 22 of these people have blue eyes. 25 of the gymnasts have blonde hair. 14 people have neither blue eyes nor blonde hair. Draw a Venn diagram.

6. In a class of 32, 23 students play football, 18 play tennis and 4 do no sports. How many play both football and tennis?

7. 100 students are asked about the subjects they study. All the students asked study one or more of the subjects listed.

42 study food tech	12 study both food tech and psychology
36 study health and social	10 study both health and social and food tech
47 study psychology	9 study both health and social and psychology

If a student is chosen at random, what is the probability that they study all three of the given subjects?

12.3 Tree Diagrams

Tree diagrams are used to solve probability problems involving two or more events. They can be used for a set of mutually exclusive exhaustive events.

When constructing a tree diagram you should write the outcomes at the end of each branch and the value of the probability on each branch. It's important to remember that the probabilities on each set of branches should add to 1.

To find the probabilities of the two or more events you multiply along the branches and where appropriate add the values at the end. These are often thought of as the "and" and "or" rules – you multiply for "and" and you add for "or".
For independent events P(A and B) = P(A) x P(B) and P(A or B) = P(A) + P(B)

Example:
A fair die is rolled twice, find the probability of rolling a 6 then an even number.

Here we're looking for the probability of a 6 *and* an even number – recall that ∩ is used to represent 'and'.
P(6 ∩ even) = $\frac{1}{6} \times \frac{1}{2} = \frac{1}{12}$

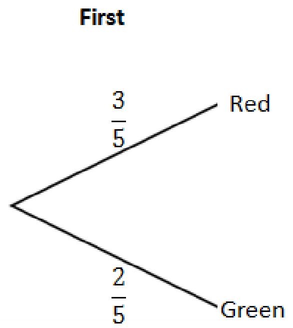
Example:
A fair die is rolled, find the probability of rolling a 2 or a 3.

Here we're looking for the probability of a 2 *or* a 3 so we have
P(2 ∪ 3) = $\frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$

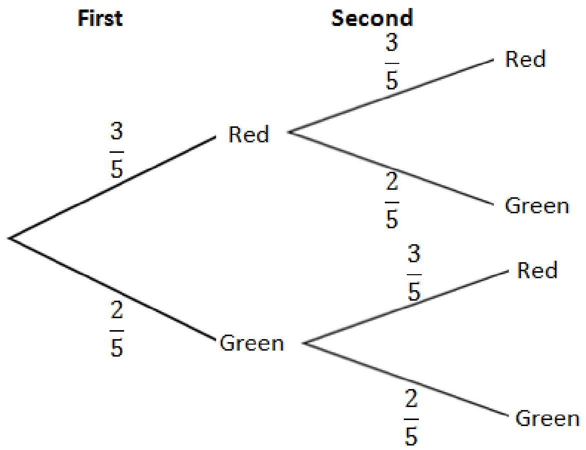
Generally, if you're given a situation in which you have more than one event, the first thing you should do is construct a tree diagram. The steps for construction are illustrated in the following example.

Example:
A bag contains 10 balls; 6 red balls and 4 green balls. One is chosen at random then replaced before a second one is chosen.
Draw a tree diagram and use it to find the probability of
a. Two red balls
b. Red then green
c. At least one red
d. Two of the same colour

The first thing to do is draw the first set of branches, this corresponds with the first ball chosen. The outcomes are written at the end and the probabilities are drawn on the branches.
P(R) = $\frac{6}{10} = \frac{3}{5}$ P(G) = $\frac{4}{10} = \frac{2}{5}$

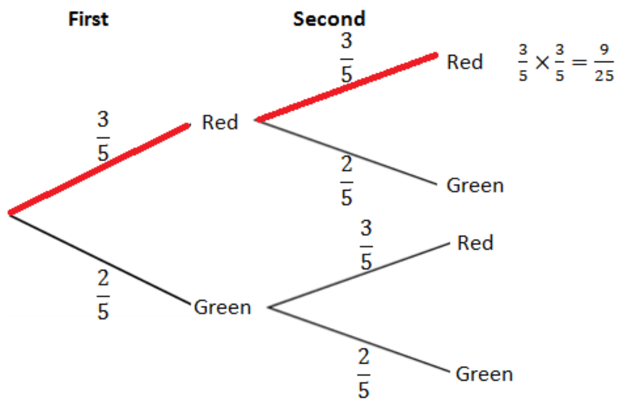


Once these are drawn the second choice of ball is added as shown. Since the ball was replaced the probabilities remain the same.



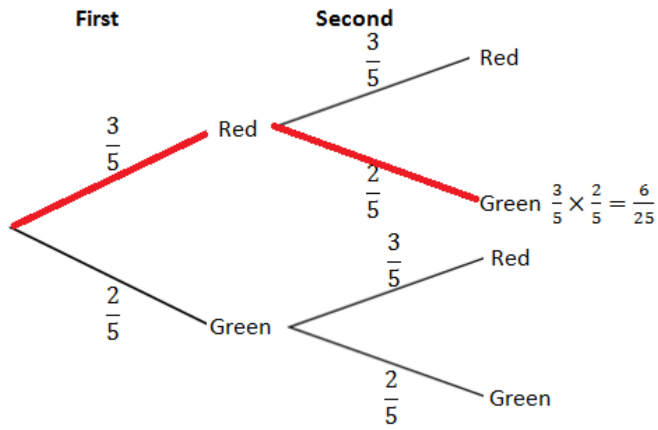
We now have all the events represented so the tree diagram is complete. As you can see they grow very quickly so, if you have more events or options be careful to give yourself a lot of space!

a. Probability of two red balls
We need to follow the branches along, multiplying as we go.



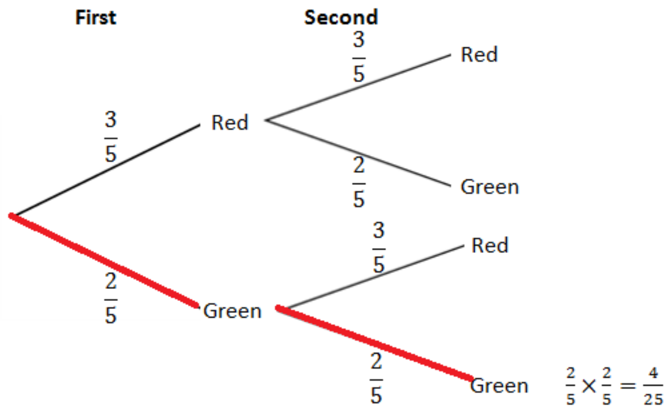
P(RR) = $\frac{9}{25}$

b. Red then green
Again we need to multiply along the branches of these choices



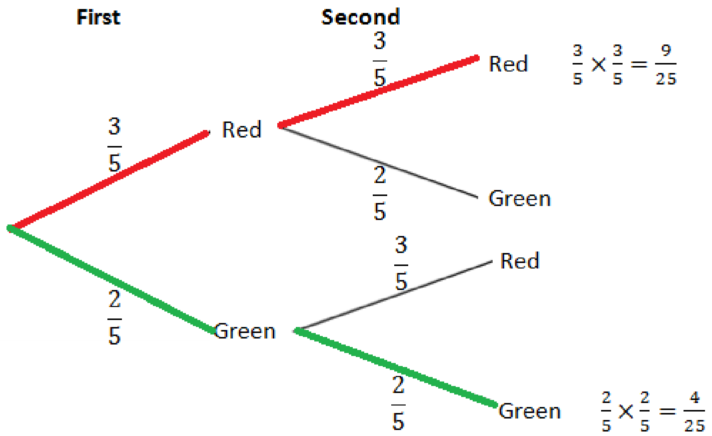
P(RG) = $\frac{6}{25}$

c. At least one red
Here we use the rule we learnt previously: P(at least one red) = 1 – P(no red)
The only option that gives no red is two greens.



P(at least one red) = 1 – P(GG) = $1 - \frac{4}{25} = \frac{21}{25}$

d. Two of the same colour
With this option we have either RR or GG. In this case we follow both of these outcomes along the branches, multiplying as we go like before. However since we want P(RR *or* GG) we must add the results together.

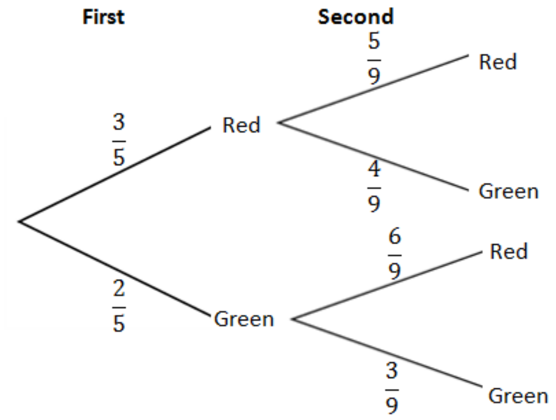


P(Two of the same colour) = P(RR ∪ GG) = $\frac{9}{25} + \frac{4}{25} = \frac{13}{25}$

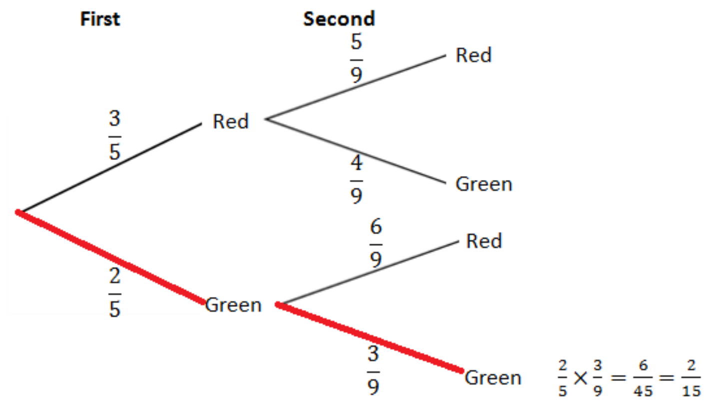
Probability tree diagrams are especially useful in what we call "non replacement" problems. These involve events, such as the one above, where items are removed and not returned causing the probability to change at each stage. A question of this sort is illustrated below.

Example:
As before a bag of 10 balls contains 6 red and 4 green. Two balls are chosen at random – the first is not replaced before the second is chosen.
Draw a tree diagram and find the following probabilities
a. At least one red
b. Two of the same colour

The tree diagram is constructed as before. In this case the first set of branches, corresponding to the first ball chosen, remain the same. However the values of the probabilities change as we move onto the second set of branches.
Once one ball has been removed there are 9 balls in total. If a red one has been removed there is one less red so there are now 5 red and 4 green. If a green one has been removed there are 6 red and 3 green. The probabilities on the tree diagram show this.

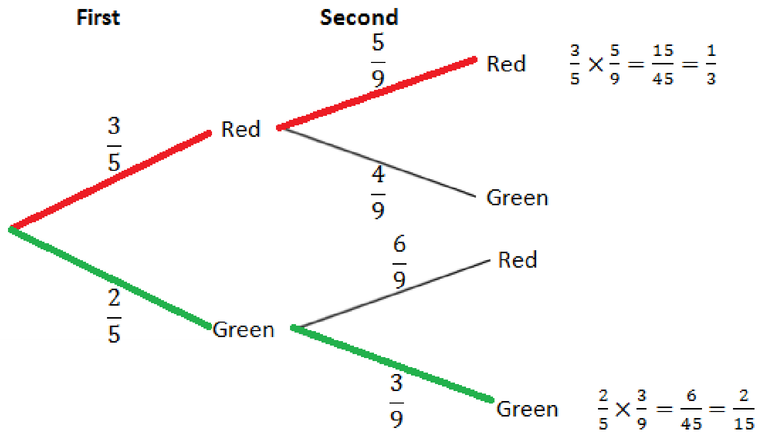


a. P(At least one red) = 1 – P(no red)



P(at least one red) = $1 - \frac{2}{15} = \frac{13}{15}$

b. P(Two of the same colour) = P(RR) + P(GG)



As you can see the fractions were easier to add before they were simplified as they had the same denominator, this is often the case with probability problems and why it is usually easier to not simplify the fractions until the last step.

P(two of the same colour) = P(RR) + P(GG) = $\frac{15}{45} + \frac{6}{45} = \frac{21}{45} = \frac{7}{15}$

Activity 12.3

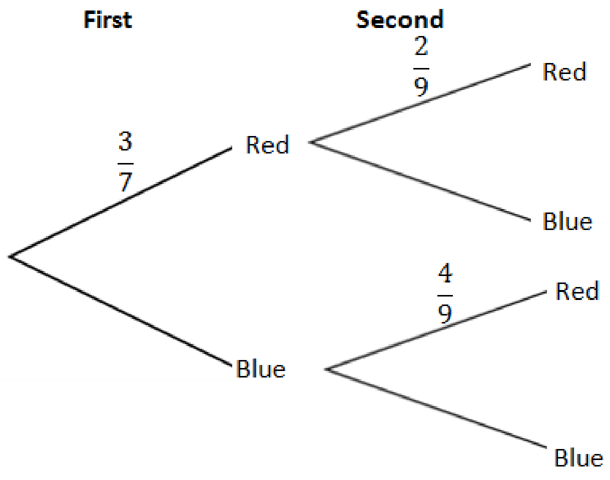
1. A bag has 5 red and 3 blue balls in it. A ball is taken at random and not replaced before a second ball is taken out.
   a. What is the probability of choosing red on the second attempt?
   b. What is the probability of getting no blue balls?

2. Susan works five days a week. She travels to work by car two days a week and by train on the other three. She is late 10% of the time when she travels by car and 20% of the time when she travels by train.
   a. Draw a probability tree using the information above.
   b. On any given work day what is the probability that she travels by car and is late?
   c. On any given work day what is the probability that she is not late?

3. A driving theory test has a pass rate of 75%. Students that pass this test are eligible to take the practical test which has a pass rate of 60%. Draw a probability tree diagram and use your diagram to find the probability of passing both tests.

4. A shop sells lunches consisting of a sandwich and a packet of crisps, the probability of choosing each sandwich filling and crisp flavour are shown below.
Sandwich: P(Cheese) =0.4 P(Ham) =0.1 P(Egg) = 0.5
Crisps: P(Plain) = 0.6 P(Salt and Vinegar) = 0.4
Draw a tree diagram and find the probability of choosing:
a. Cheese sandwich with plain crisps
b. Ham sandwich with salt and vinegar crisps

5. a. Complete this tree diagram



b. Find the probability of choosing two the same.

12.4 Conditional Probability

A conditional probability is where the probability of later events happening depends on earlier events occurring. This was touched upon in questions 2 and 3 in the previous activity.

The probability of event B happening, given that A has already happened, is written as P(B given A). This is calculated as followed

$$\text{P(B given A)} = \frac{\text{P(A and B)}}{\text{P(A)}}$$

Example:
Given that P(A)=0.3, P(B)=0.2 and P(A and B) = 0.1
a. Find P(B given A)
b. Are A and B independent?

a. $\text{P(B given A)} = \frac{0.1}{0.3} = \frac{1}{3} = 0.33\dot{3}$
b. $\text{P(B given A)} \neq \text{P(B)}$ so, no, the events are not independent

---

Activity 12.4

1. P(A)=0.4 and P(B)=0.5 where A and B are independent events. Find P(B given A).

2. P(A)=0.2 and P(B)=0.7 where A and B are independent events. Find P(A given B).

3. A survey of drivers that had been in an accident showed that 97% of drivers wear seatbelts and that, if there is an accident, 34% of those not wearing a seatbelt die whereas 93% of those that do wear a seatbelt survive. Draw a tree diagram showing this information and find the probability of picking a driver at random that did not wear a seatbelt and survived.

ASSIGNMENT NINE

Answers to these questions are not provided. You should send your work to your tutor for marking. Show all of your working.

1. A recipe uses flour, sugar and butter in the ratio 25:20:15, simplify the ratio and find out how much butter and flour there is when 240g of sugar is used. (4)

2. The population of a village decreases by 0.5% each year. There are 40,000 people in 2001.

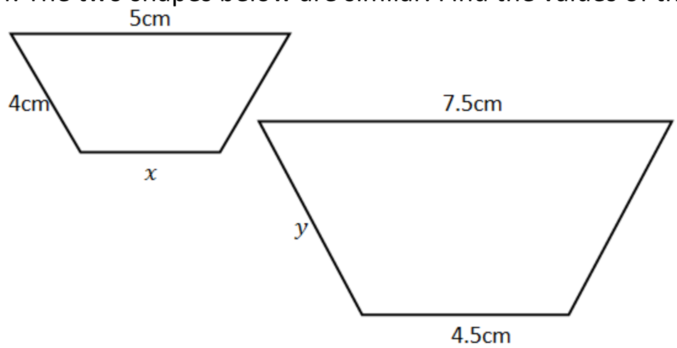
a. How many people would there be in 2002? (2)

b. How many people would you expect there to be in 2011? (2)

c. How many people would you have expected there to be in 2000? (2)

3. The time taken to complete a journey is directly proportional to the distance travelled and inversely proportional to the average speed throughout the journey. It takes a driver 2 hours to complete a journey of 60 miles. If the average speed is doubled would the same driver be able to complete a journey of 180 miles in 3 ½ hours? (3)

4. The two shapes below are similar. Find the values of the missing sides.



(3)

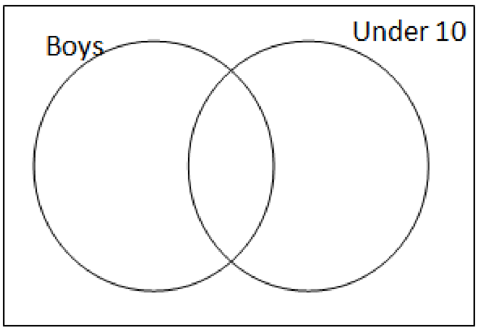
5. Three people flip a bias coin. Andy flips the coin 10 times and gets 7 heads, Bethany flips the coin 20 times and gets 15 heads and Carol flips the coin 30 times and gets 23 heads. Each of them uses their results to calculate the probability of getting heads or tails.

a. Write down the value of P(H) and P(T) that each of them would calculate (3)

b. Who do you think has the most accurate estimate for the probability? Why? (1)

c. Another coin has P(H) = 0.8. If the coin was flipped 50 times how many times would you expect to get tails? (1)

6. There are 50 people in a hockey club. 15 of these are boys under ten years old. 10 are girls over ten years old. There are 23 people under ten years old altogether.
Complete this Venn diagram.



(3)

7. X = {First 10 prime numbers}, Y = {First 10 odd numbers}
Write down the elements of a. $X \cup Y$ b. $X \cap Y$ (2)

8. Annie works on Tuesdays and Thursdays. Each day she is either early, on time or late. On Tuesday P(early) = 0.2, P(on time) = 0.5. On Thursday P(on time) = 0.4, P(late) = 0.3.

a. Draw a tree diagram showing the probabilities of her being early, late or on time on these two days. (3)

b. Find the probability that she is on time both days. (2)

c. Find the probability that she is early at least once (4)

Total 35 marks

Chapter Thirteen: Statistics

13.1 Collecting Data

Sampling

A **sample** is a collection of units or data values drawn from a **population**. The population is the set of units that is being studied. Although the word population is widely used to describe people it is not limited to this and can refer to a variety of things such as objects, drugs or procedures to name but a few. Looking at people as an example the population does not always refer to every person on the planet, it can be limited to those we are interested in for the study. For example the population may be all competitive gymnasts or all people of Irish descent.

The larger the sample the more accurate and representative the data is. For this reason, from a mathematical point of view, it would be much preferable to use the entire population in any study. A sample is used instead because it is usually highly impractical to gather data from every member of a given population. Because of this it is likely that the values gathered from the sample differ slightly from the values that would be gathered from the population. This is called a **sampling error**.

When selecting a sample there are a number of important points to consider:

- Is the sample going to be biased?
  Bias is a systematic error whereby the data differs from the population parameter. For example, if you wanted to find out if people liked football and you collected your data by asking a sample of people coming out of a football stadium your result would be biased.

- Is the person collecting the data going to affect the results?
  Another factor that can introduce bias is the person collecting the data. For example, if a police force want to find out about underage drinking and they send out a police officer to ask a group of underage teenagers if they've consumed alcohol they will all say no, regardless of whether or not that is true.

- Does the data collection method affect the results?
  Using a similar logic to the point above the method used to collect the data can affect the results of the study. As mentioned before people may give the answer they think is correct or the one they think they should rather than a true answer. A questionnaire that has, for example, a question asking if someone is a generous person they are likely to say yes regardless.

- Is the data relevant to the study?
  It is often the case that people will collect the data that is easier to find rather than what they actually want to measure. You need to be careful that questions asked or data collection techniques used measure the variable that you intend to.

- Is the sample large enough?
  We have already noted that a larger sample gives more accurate data but becomes less practical to deal with. With a well chosen sample, the bigger it is the smaller the sampling error so it is a difficult balancing act to find the correct sample size for your study. For example, if a study asked 10 people who they intended to vote for in an upcoming election the data would not be accurate enough to represent the entire population of the country.

Frequency Tables

When you're collecting data you need a quick and concise way of recording it. It is commonplace to use a data collection sheet, also known as a **tally chart**.
For each occurrence of a given value a vertical mark | is used, for every fifth value a horizontal or diagonal line is drawn across the vertical marks ~~||||~~

A frequency table is drawn up by counting the total number of tally marks in each row.

Example:
Fifteen students are asked what their favourite colour is, the results are below.

Yellow	Red	Purple	Red	Yellow
Pink	Blue	Purple	Pink	Blue
Yellow	Red	Pink	Red	Red

Complete a tally chart.

First we need to construct the tally chart. An empty chart is shown below, have a go at filling it in yourself before looking at the completed chart beneath it.

Colour	**Tally**	**Total**
Yellow		
Red		
Purple		
Pink		
Blue		

Colour	**Tally**	**Total**
Yellow	|||	3
Red	~~||||~~	5
Purple	||	2
Pink	|||	3
Blue	||	2

Data can also be shown in a **two way table**. These are used to link two different types of information. An example of a two way table is shown below.

	Male	**Female**
Red	8	2
Green	7	5
Blue	5	7

Cumulative Frequency

A cumulative frequency table is easily constructed from a grouped frequency table by simply calculating a running total of the frequency column.

This is best described by working through an example.

Example:
The grouped frequency table below shows the number of marks gained by students on a test. Draw up a cumulative frequency table.

Score	Frequency
$0 < s \leq 10$	5
$10 < s \leq 20$	9
$20 < s \leq 30$	7
$30 < s \leq 40$	8

A cumulative frequency table is constructed as follows.

Score	Cumulative Freq.
$0 < s \leq 10$	5
$0 < s \leq 20$	$5 + 9 = 14$
$0 < s \leq 30$	$14 + 7 = 21$
$0 < s \leq 40$	$21 + 8 = 29$

Activity 13.1

1. Using the two way table in the previous example, answer the following questions.
    a. How many males are there altogether?
    b. How many people are there altogether?
    c. How many females picked blue?
    d. How many people picked green?

2. Lucy conducts a survey asking people how many animals they have. Draw a tally chart for her data which is shown below.

3	4	1	0	3	1	2	0	3	3	1	1	0
2	5	0	1	2	1	2	2	3	4	0	1	3

3. Sam wants to find out how many people read newspapers, to do this he has to ask a sample of people. Write down a problem with using each of the following samples.
    a. Asking people as they come out of a newsagents
    b. Asking a year 10 class at a local school
    c. Asking every person in town

4. A group of 50 sixth form students are asked about whether they study maths or English or both. Of these students 23 are female. The results show that 12 girls and 14 boys study maths. Fill in the two way table below.

	Maths	**English**	**Total**
Male			
Female			
Total			

5. The number of pets in each house are shown below.

Number of Pets	Number of Houses
0	4
1	6
2	5
3	8
4	2

a. Work out the total number of houses.
b. Work out the total number of pets.
c. Draw a cumulative frequency table.

6. The table below shows some information about favourite fruit. Fill in the gaps.

	Banana	Apple	Orange	Total
Male	5		8	22
Female				
Total	13	21		52

<u>13.2 Statistical Diagrams 1</u>

Various different diagrams are often used to represent data. Different ones are used in different situations – in part the choice depends on the type of data collected and in part it is down to the preference of the person constructing them and what "looks best".

<u>Pictograms</u>

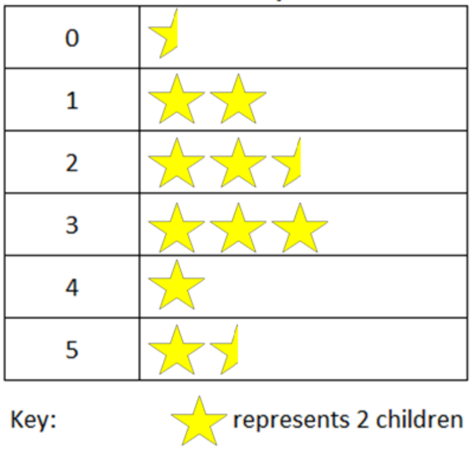
Pictograms use symbols to represent quantities. The symbols can be anything and are often chosen if they relate to the study in some way. For example, if the data was about cars the symbol used may be a picture of a car. It is often the case that part of a symbol, rather than the whole thing, must be used to represent quantities.

Example:
The number of gold stars a group of primary school children receive in a day is displayed in the chart below. Draw a pictogram to represent the data.

Number of Stars	Number of Children
0	1
1	4
2	5
3	6
4	2
5	3

The pictogram is shown below. Here we have chosen to use a star as the symbol as it is appropriate for the situation however the choice of symbol is actually irrelevant. It is, however, important to remember to include a key.
You can also choose how many objects the symbol represents. Here we have chosen two.

Number of Stars in a Day



<u>Bar Charts</u>

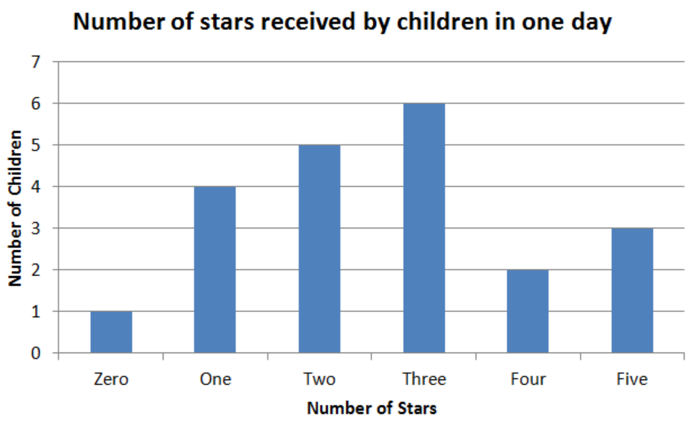
Bar charts are most commonly seen as comprising of vertical bars but they can also be horizontal. Like pictograms they give a good visual representation of the size of each category and are therefore a good way to display discrete data.

There are a few things that are important to remember when drawing bar charts. You need:
- Titles on both axes
- A title for the diagram
- Bars of equal width and spacing
- Values on the y axis evenly spaced starting at 0

Example:
The number of gold stars a group of primary school children receive in a day is displayed in the chart below. Draw a bar chart to represent the data.

Number of Stars	Number of Children
0	1
1	4
2	5
3	6
4	2
5	3

The bar chart for the data is shown below

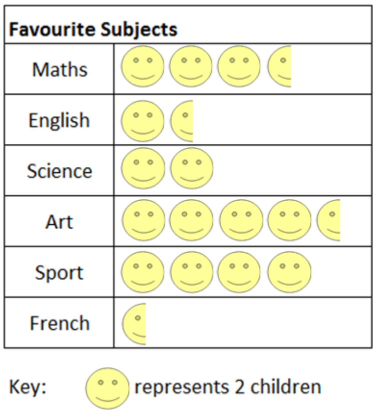


Activity 13.2 A

1. A group of children are asked their favourite subject. The data collected is shown in the table below. Draw a pictogram and a bar chart to represent the data.

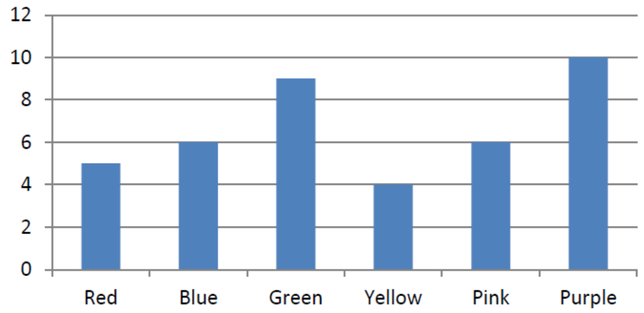
Subject	**Number of Children**
Maths	10
English	3
Science	8
Art	2
Sport	6
French	3

2. A different group of children were asked about their favourite subjects. The data collected is shown on the pictogram below.



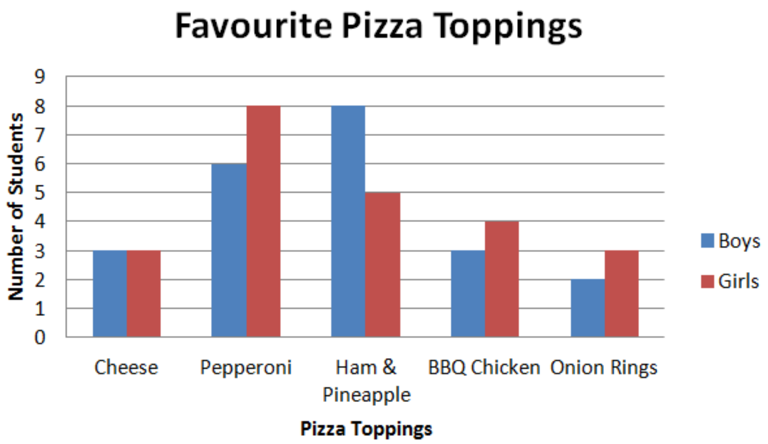
a. Which subject was most popular?
b. Which subject was least popular?
c. How many students chose maths?
d. How many children were asked altogether?

3. A bar chart showing the favourite colours of children is shown below.



a. What is wrong with this bar chart?
b. Which colour was the most popular?
c. How many children picked yellow?
d. How many children picked brown?
e. How many children were asked altogether?

4. The dual bar chart below shows favourite pizza toppings of a group of students.



a. Which topping was most popular with the girls?
b. Which topping was most popular with the boys?
c. Which topping was the favourite for the same number of girls and boys?
d. Which topping was the most popular over all?
e. How many girls were asked?

Pie Charts

Pie charts are another way of representing data, they are often favoured as a good visual representation.
The angle of a section shows the size of each category.

As with bar charts there should always be a title. Alongside this there needs to be labels for each category. These can either be written inside or beside each section or they can be written as a key.

In order to calculate the size of the angle you must divide 360° by the total number of data values and then multiply this answer by the number in each category.

Example:
The favourite colours of 30 children are shown in the table below. Draw a pie chart to represent the data.

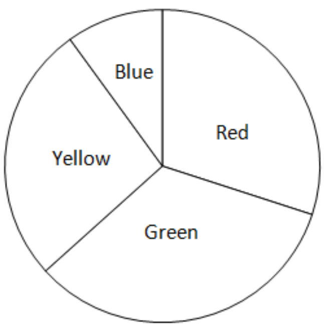
Colour	Number of Children
Red	9
Green	10
Yellow	8
Blue	3

There are 30 data values altogether so the first thing we need to do is
$360° \div 30 = 12°$

Now we need to multiply this value by the number in each category to find the size of the angle needed
Red: $9 \times 12° = 108°$
Green: $10 \times 12° = 120°$
Yellow: $8 \times 12° = 96°$
Blue: $3 \times 12° = 36°$

To construct the pie chart you should draw a circle – of any size – and mark the centre. Next draw a line from the centre to the circumference – this can be anywhere. Use this line as a base to measure the first angle, mark where the angle goes from the protractor and draw a line in. Now use this line as a base to draw the next angle, continue like this until all sections have been drawn.

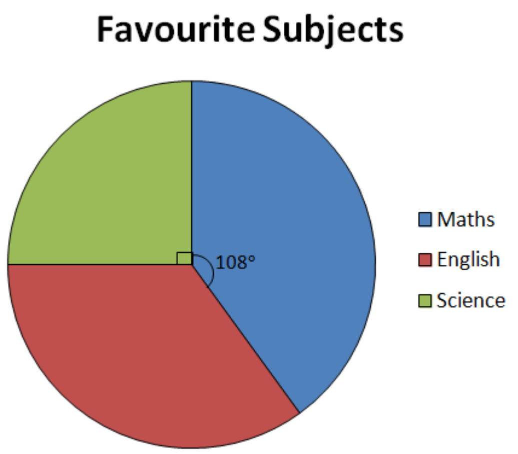
Favourite Colours



You also need to be familiar with how to read pie charts to find the size of each category. Sometimes the pie chart will be drawn accurately, when this is the case you measure each angle. Sometimes the pie chart will say "not to scale" in this case you will be given some of the angles and you use your knowledge of angles about a point adding up to 360 to find the size of each of them.

As before the first thing to do is divide 360 by the total number of data values. Once you have this value, and the size of each angle, you should simply divide the size of the angle by this value to find the size of the category.

Example:
The pie chart below shows the favourite subjects of 20 students. It is not drawn to scale.
Find the number of students that favour each subject.



There are 20 data values altogether so the first step is
$360° \div 20 = 18°$

Now we need to know the size of all of the angles.
Science: 90°
Maths: 108°
English: $360° - 90° - 108° = 162°$

The last step is to divide each of these angles by the value found at the beginning to find the number of students in each category
Science: $90 \div 18 = 5$
Maths: $108 \div 18 = 6$
English: $162 \div 18 = 9$

Activity 13.2 B

1. Given the data in the tables below, draw a pie chart.

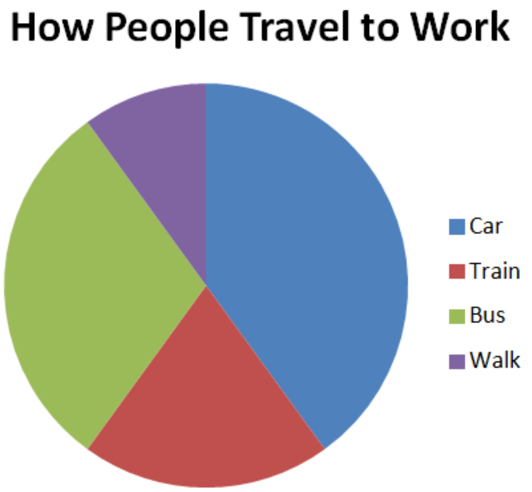
a.

Fruit	**Number of People**
Apple	10
Banana	12
Orange	8
Other	10

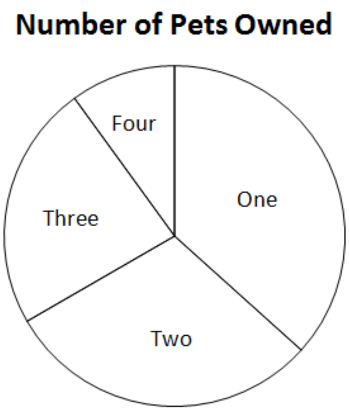
b.

UK Holiday Destination	**Number of People**
Cornwall	16
Blackpool	27
Isle of Wight	38
Skegness	14
Other	25

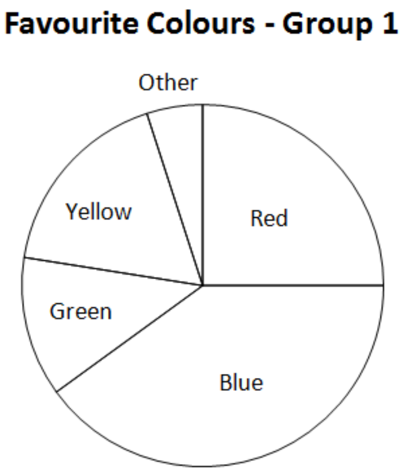
2. Twenty people were asked how they travel to work, the pie chart is shown below. It has been drawn accurately. Find the number of people in each category.

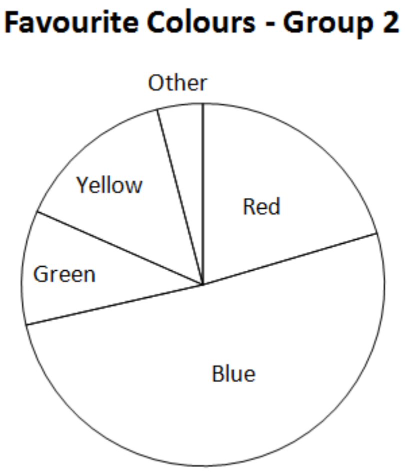


3. Thirty people were asked how many pets they have, the pie chart is shown below. The pie chart is drawn accurately. Find the total number of pets owned by the group.



4. Two different groups of people were asked about their favourite colours. Amanda looks at the pie charts and says that there are more people in the second group than the first group that prefer blue. Do you think she is correct? Explain why.





Histograms

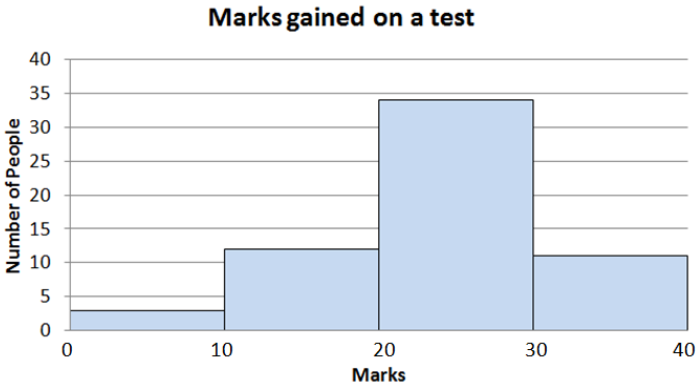
Histograms look a lot like bar charts but there are some important differences. Firstly there can be no gaps between the bars and secondly the width of the bars may vary rather than remaining constant.
As we have seen bar charts can be used to display discrete data, histograms, on the other hand, are used to display continuous grouped data.

Example:
The table below shows the marks gained in a test by students. Draw a histogram to represent the data.

Marks	Number of People
0-10	3
11-20	12
21-30	34
31-40	11

The first thing we need to do is check the class widths and make sure they do not overlap. In this case there is no overlap and they are all the same width.

Now we draw the histogram in much the same way as we would draw a bar chart, the only difference is the labelling along the x axis.



When the class width are different histograms need to be drawn up carefully. The height of the bars represent the frequency density which is calculated as follows

$$\text{Frequency Density} = \frac{\text{Frequency}}{\text{Class Width}}$$

Therefore, if looking at a histogram, you find the area of the bar to find the frequency
Frequency = class width x frequency density = width x height

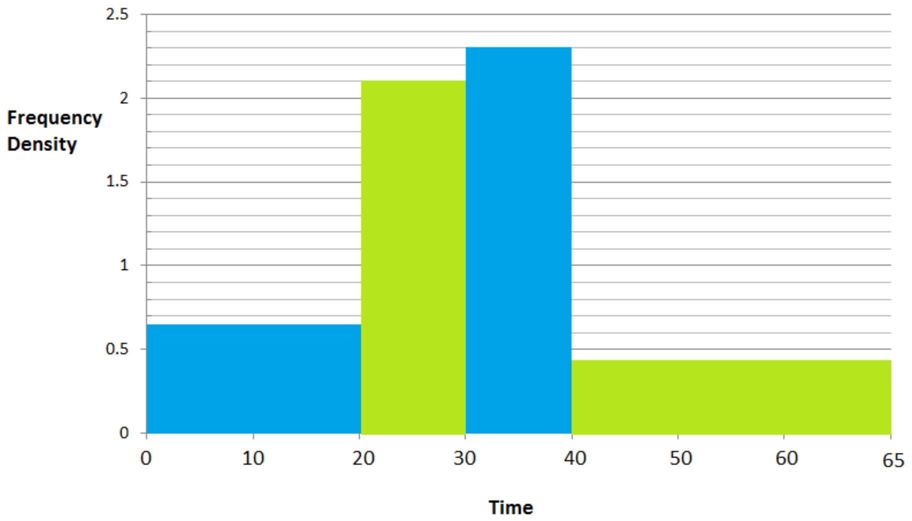
Example:
Draw a histogram for this set of data

Time	Frequency
$0 < t \leq 20$	13
$20 < t \leq 30$	21
$30 < t \leq 40$	23
$40 < t \leq 65$	11

The first thing we need to do is add two columns, one to record the class width and one to calculate the frequency density.

Time	Class Width	Frequency	Freq. Density
$0 < t \leq 20$	20	13	$13 \div 20 = 0.65$
$20 < t \leq 30$	10	21	$21 \div 10 = 2.1$
$30 < t \leq 40$	10	23	$23 \div 10 = 2.3$
$40 < t \leq 65$	25	11	$11 \div 25 = 0.44$

Now all that remains to be done is to plot the graph.



Scatter Graphs and Correlation

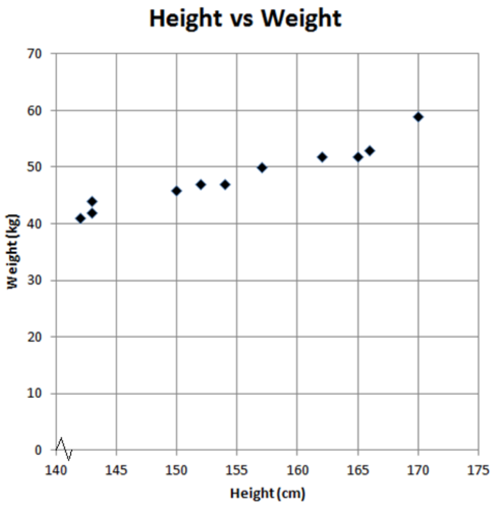
Scatter graphs are particularly useful for comparing two sets of data, for example, height and weight. Data is generally collected in pairs – such as recording someone's height and weight at the same time – and coordinates are plotted.

If the points appear to lie roughly in a straight line the line is drawn in and is called the **line of best fit**. This is an estimate.
If the line has a positive gradient the data is said to have a **positive correlation** and if the line has a negative gradient the data is said to have a **negative correlation**.
The correlation coefficient is a value between -1 and +1. If the points all lie close to a line the correlation is strong and the value is nearing 1 or -1 depending on whether the gradient is positive or negative. If the points are scattered around the correlation is weak and the value is nearing 0 – if we cannot see a line at all we say the value is 0 and there is no correlation.

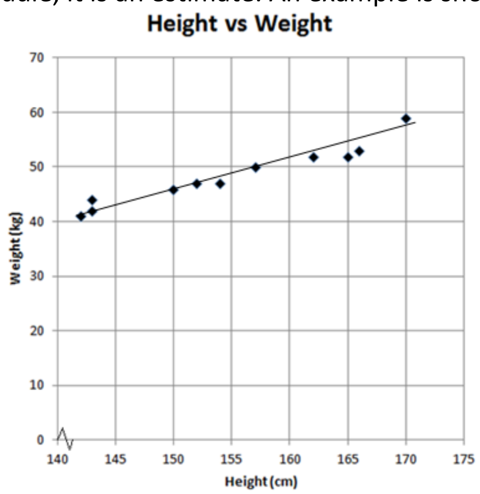
Example:
A random sample of people have their height and weight recorded, the results are shown below.
Draw a scatter graph and describe the results.

Height (cm)	165	154	143	142	157	166	162	170	152	143	150
Weight (kg)	52	47	42	41	50	53	52	59	47	44	46

The first thing we need to do is plot the graph taking height and weight as the axes.



We can easily see that the data appears to be in a line. A line of best fit attempts to go through the middle, it is an estimate. An example is shown below.



The question also asked us to describe the results. With this data we would say something along the lines of “there is a positive correlation between the height and the weight, as one increases so does the other”.

It is important to remember:

Correlation does not imply causation.

A correlation between two variables does not mean that one causes the other. For example, looking at the previous question we can see that, as a rule, as weight increased so did height but it would be incorrect to say that an increase in weight causes an increase in height!

Activity 13.2 C

1. Given the data below draw a histogram.

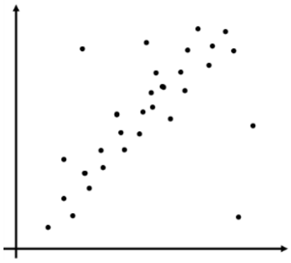
Time	Frequency
$0 < t \leq 10$	21
$10 < t \leq 20$	19
$20 < t \leq 25$	10
$25 < t \leq 35$	11

2. Given the data on height and weight below draw a scatter graph, a line of best fit and describe the correlation.

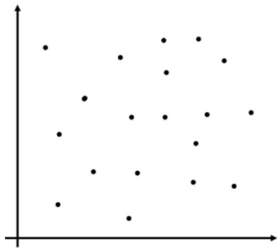
Height (cm)	142	165	174	143	156	167	164	152	146	170	165	159	146	148	165
Weight (kg)	43	52	66	42	46	54	55	47	41	58	55	52	44	50	52

3. Describe the correlation in each of these graphs

a.

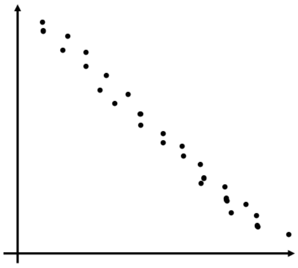


b.

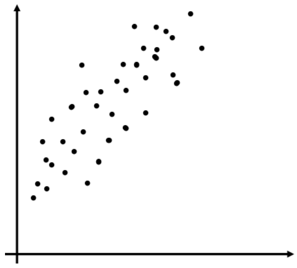


c.

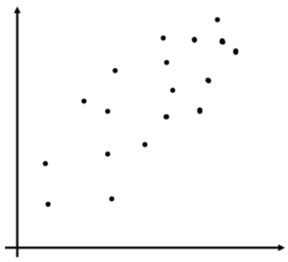
d.



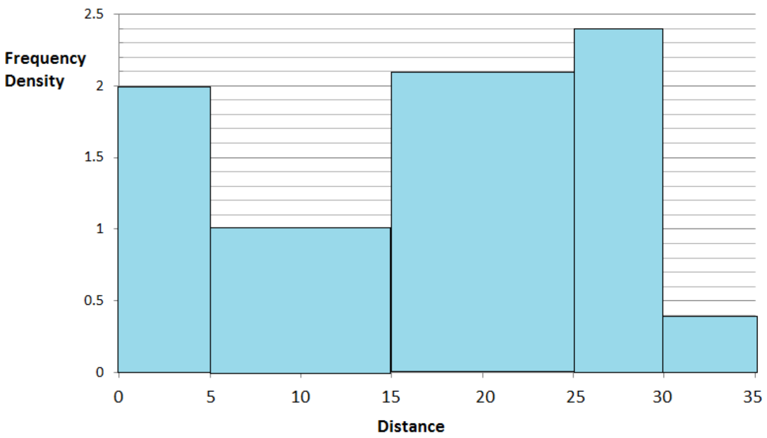
e.



f.



4. Draw a frequency table for this histogram



13.3 Averages

There are four different ways of calculating an average: mean, median, mode and range. Each of these will be described in turn.

Mean

The mean is the most commonly used average. It is often joked that it is called the mean because it is the most difficult one to calculate.

In order to calculate the mean you need to add all the data values together then divide this by the number of values.

This has the advantage of taking into account all the values however the answer is not always a value that occurs within the data set and it can be affected by outliers.

Example:
Find the mean of this list of numbers
1 7 9 4 3 9 3 1 1 3 3

The first thing we need to do is add all of the numbers together
1+7+9+4+3+9+3+1+1+3+3=44

Now we count them to find that there are 11 numbers, therefore we do
$44 \div 11 = 4$

So the mean of this list of numbers is 4

Mode

The mode of a set of data is, quite simply, the value that occurs the most frequently. It is possible for there to be more than one mode.
To remember this think: mode = most common

The mode has the advantage of not being affected by outliers, being easy to calculate and being an actual value within the data set. There's also the fact that you are able to find it when the data values are not numbers. However there isn't always a mode – there can be more than one.

Example:
Find the mode of this list of numbers
1 7 9 4 3 9 3 1 1 3 3

It is a case of simple observation to see which number occurs most often but, as a general rule, it is easier to see when the data values are rearranged into ascending order
1,1,1,3,3,3,3,4,7,9,9
It is now clear to see that the mode is 3

Median

The median is the middle value when the data is arranged in order. If there are an even number of data values meaning there is not a single middle value the median is calculated by adding the two middle values together and dividing the answer by 2.

This average, like the mode, has the advantage of not being affected by outliers and being easy to calculate however it isn't always an actual value in the data set.

Example:
Find the median of this list of numbers
1 7 9 4 3 9 3 1 1 3 3

The first thing we need to do is to rearrange the set so that it is in order

1,1,1,3,3,3,3,4,7,9,9

Now we just need to work out which is the middle value

1 1 1 3 3 3 3 4 7 9 9

So the median is 3

Example:
Find the median of this list of numbers
1 3 7 9 8 9 5 7 3 6

Again the first thing to do is put the data values in order and then find the middle value

1 3 3 5 6 7 7 8 9 9

Here there is not just one middle value so we circle the two middle values instead.
To find the median we have to add them together and divide by 2
(6+7)÷2 = 13÷2 = 6.5
So the median is 6.5

Range

The range is found by subtracting the lowest value in the data set by the highest value.

Example:
Find the range of this list of numbers
1 7 9 4 3 9 3 1 1 3 3

The highest and lowest values can be found by observation but, in order to be sure, it is best to put the set in order.
1,1,1,3,3,3,3,4,7,9,9

Now we can clearly see that the highest value is 9 and the lowest value is 1, therefore the range is 9 – 1 = 8

Questions 3-5 in the activity below illustrates how to calculate averages with a table. Have a go at the question 3 then look in the answers to see the worked answer if you're unsure before continuing with questions 4 and 5.

Interquartile Range

The interquartile range (IQR) is found by subtracting the lower quartile from the upper quartiles.

Quartiles are found in a similar way to the median. The lower quartile is the middle value of the bottom half of the ordered data and the upper quartile is the middle value of the top half.

It is up to you whether you count to find them or use the formulae
Lower quartile = ¼ (n+1)th value
Upper quartile = ¾ (n+1)th value
When n is the number of data points

When calculating the median you can also use the ½ (n+1)th value

Example:
Find the interquartile range for this set of data
1 7 9 4 3 9 3 1 1 3 3

The first thing we need to do is to rearrange the set so that it is in order
1,1,1,3,3,3,3,4,7,9,9

Now we find the median:

1 1 1 3 3 3 3 4 7 9 9

Therefore the upper and lower quartiles are shown in red and blue respectively.

1 1 1 3 3 3 3 4 7 9 9

So we have
IQR = 7 – 1 = 6

Example:
Find the IQR for this list of numbers
1 3 7 9 8 9 5 7 3 6

Again the first thing to do is put the data values in order and then find the middle value

1 3 3 5 6 7 7 8 9 9

So the UQ and LQ are

1 3 3 5 6 7 7 8 9 9

IQR = 8 – 3 = 5

Activity 13.3 A

1. Find the mean, mode, median, range and IQR of each of these lists of numbers

a. 2 8 9 7 7 8 5 9 7 7 6 5

b. 8 9 7 5 7 0 7 6 4 2 8

c. 78 65 87 98 98 65 76 78 99 91 65 71

2. Find the mean, mode and median of this list of numbers. Which do you think gives the best average in this case? Why?
89 101 98 92 100 105 92 91 101 101 5 90

3. The scores a group of students got on a test are shown in the table below. Find the mean, median, mode, range and IQR of the test scores.

Score	Frequency
45	2
46	3
47	6
48	1
49	3
50	2

4. The heights of a group of people were measured, the results are shown in the table below. Find the mean, median, mode, range and IQR of the heights.

Height (cm)	Frequency
153	2
154	3
157	7
159	3
161	8
164	3

5. The number of sweets in boxes were counted, the results are shown in the table below. Find the mean, median, mode and range of the number of sweets.

Number of Sweets	Frequency
11	5
13	2
14	1
16	3
18	2
20	1

6. Find a set of numbers that satisfy each of the conditions. There is more than one possible answer in each case.
   a. Four numbers with a mean of 3
   b. Five numbers with a mode of 4 and a median of 5
   c. Six numbers with a range of 5 and a median of 6
   d. Four numbers with a mean of 4, a median of 3 and a range of 6
   e. Six numbers with a mean of 11, a mode of 7, a median of 10 and a range of 10

Grouped Data

With grouped continuous data – as seen in the histograms section – we can calculate an estimated mean, the modal class and the class interval in which the median lies. The logic used to calculate each of these is very similar to that used with discrete or non-grouped data.

As before the modal class can be seen by observation, as can the median.

The estimated mean is found by following the steps below:
- Multiply the mid-point of the class by the frequency
- Add these values together
- Divide this value by the total frequency

Example:
The table below shows the marks gained by students on a test.
Find the modal class, the class interval in which the median lies and the estimated mean.

Marks	Number of People
0-10	3
11-20	12
21-30	34
31-40	11

By observation we can see that the modal class is 21-30.
The class containing the median is also 21-30.

In order to calculate the estimated mean it is easiest to add two extra columns and one extra row to the table.

Marks	Number of People	*Mid point*	*Midpoint x Frequency*
0-10	3	*5*	*5 x 3 = 15*
11-20	12	*15.5*	*12 x 15.5 = 186*
21-30	34	*25.5*	*34 x 25.5 = 867*
31-40	11	*35.5*	*11 x 35.5 = 390.5*
Total	*60*		*1458.5*

So, to calculate the estimated mean we do 1458.5 ÷ 60 = $24.80\dot{3}$

Activity 13.3 B

For each of the tables below calculate the class interval in which the median lies, the modal class and the estimated mean.

1.

Time Taken	Frequency
$0 < t \leq 5$	10
$5 < t \leq 10$	24
$10 < t \leq 15$	12
$15 < t \leq 20$	21

2.

Age	Frequency
$5 < a \leq 10$	43
$10 < a \leq 15$	12
$15 < a \leq 20$	45
$20 < a \leq 25$	32

3.

Height	Frequency
$125 < h \leq 135$	21
$135 < h \leq 145$	45
$145 < h \leq 155$	36
$155 < a \leq 165$	13

4.

Time Taken	Frequency
$10 < t \leq 15$	23
$15 < t \leq 20$	34
$20 < t \leq 25$	13
$25 < t \leq 30$	21
$30 < t \leq 35$	45

13.4 Statistical Diagrams 2

Cumulative Frequency Graphs

Plotting a cumulative frequency graph is done in the same way as plotting a line graph.
The upper bound of the group forms the y coordinate and the value in the cumulative frequency column forms the x coordinate. Once all of these points are plotted they are joined up.

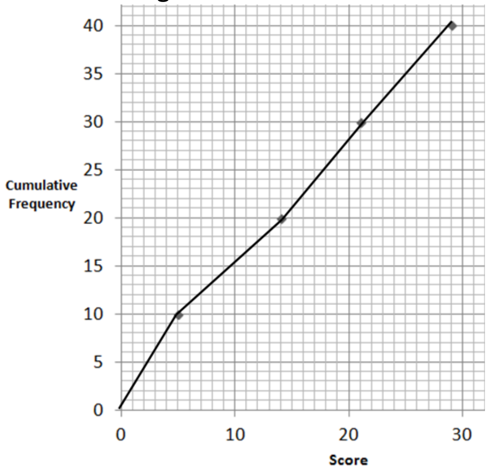
Example:
Draw a cumulative frequency graph for this set of data

Score	Frequency
$0 < s \leq 10$	5
$10 < s \leq 20$	9
$20 < s \leq 30$	7
$30 < s \leq 40$	8

The first thing to do is to construct a cumulative frequency table, this was done in a previous example.

Score	Cumulative Freq.
$0 < s \leq 10$	5
$0 < s \leq 20$	$5 + 9 = 14$
$0 < s \leq 30$	$14 + 7 = 21$
$0 < s \leq 40$	$21 + 8 = 29$

Now we need to plot the points (5,10), (14,20), (21,30) and (29,40) before joining them up with a series of straight lines.



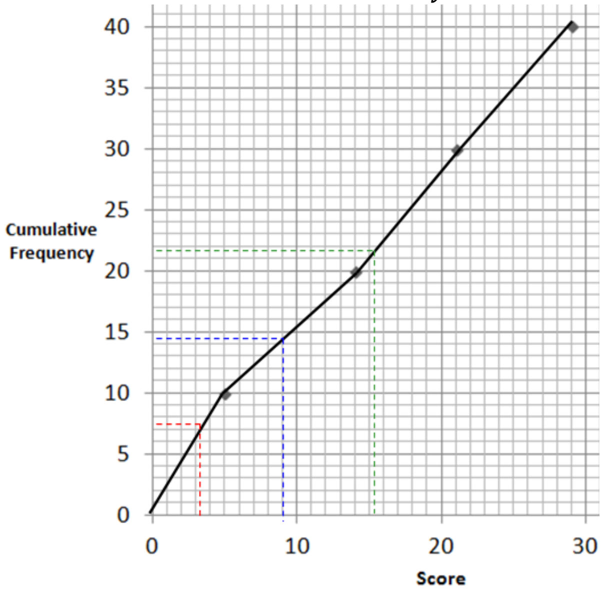
Cumulative frequency graphs can be used to estimate the median and the quartiles.
To find the median first calculate half of total cumulative frequency. Draw a line from the y axis at this value across to the graph then down to the x axis. The value at the x axis gives you the median.
Use a similar technique to find the quartiles, finding ¼ of the total for the lower quartile and ¾ for the upper quartile.

Example:
Using the graph drawn in the previous example find an estimate for the median and the quartiles.

The cumulative total is 29.

Median (Blue): 29 x ½ = 14.5
LQ (Red): 29 x ¼ = 7.25
UQ (Green): 29 x ¾ = 21.75

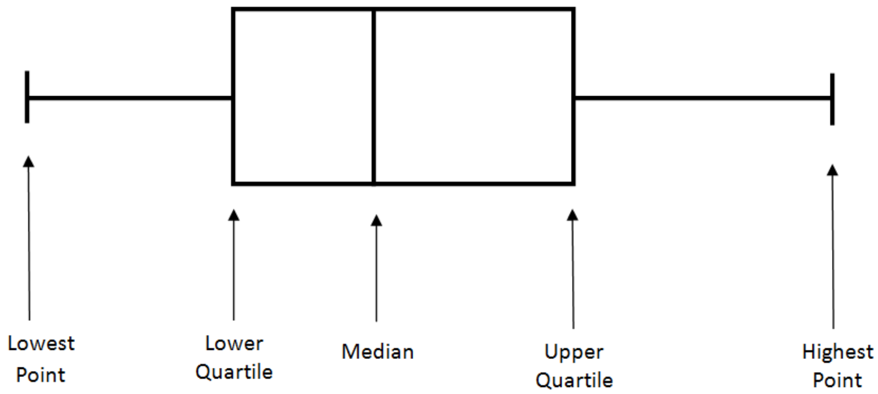
The lines are drawn across from the y axis then down to the x axis.



So our estimates would be
Median = 9
LQ = 3.25
UQ = 15.25

Box Plots

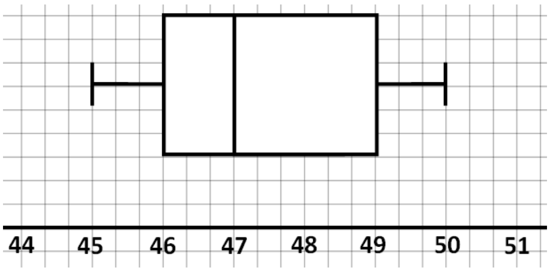
Box plots are used to illustrate the spread of the data.
A box is drawn with the lower quartile at one end and the upper quartile at the other. A line inside the box is drawn at the median. The box then has "whiskers" coming out of either side of the box, the end of these whiskers lie at the lowest value and the highest value in the data set.



Example:
Given this set of data, draw a box plot

Score	Frequency
45	2
46	3
47	6
48	1
49	3
50	2

Lowest Point = 45
Highest Point = 50
Median = 47
LQ = 46
UQ = 49



When asked to compare box plots there are a number of things you could comment on. These include the range, the IQR and the median to name but a few. When two box plots are placed together, on the same scale, they clearly illustrate any differences in range and spread between the two sets of data.

Activity 13.4

For each of the sets of data below
a. Draw a cumulative frequency graph
b. Use the graph to estimate the median and quartiles
c. Draw a box plot (for 1 and 2 use exact values for the median and quartiles)

1.

Height (cm)	Frequency
153	2
154	3
157	7
159	3
161	8
164	3

2.

Number of Sweets	Frequency
11	5
13	2
14	1
16	3
18	2
20	1

3. Youngest = 6, Oldest = 25

Age	Frequency
$5 < a \leq 10$	43
$10 < a \leq 15$	12
$15 < a \leq 20$	45
$20 < a \leq 25$	32

4. Quickest = 12s, Longest = 34s

Time Taken	Frequency
$10 < t \leq 15$	23
$15 < t \leq 20$	34
$20 < t \leq 25$	13
$25 < t \leq 30$	21
$30 < t \leq 35$	45

ASSIGNMENT TEN

Answers to these questions are not provided. You should send your work to your tutor for marking. Show all of your working.

1. You are given the list of data below

5 4 3 2 4 5 4 3 1 4 5 4 3
2 4 1 2 4 3 2 3 5 5 5 3

Using this data draw

a. A frequency table (including tally marks) (1)
b. A bar chart (3)
c. A pie chart (5)

Find

e. The mode (1)
f. The median (2)
g. The mean (2)
h. The range (1)

2. The marks students gained on their maths and English tests are shown below. Draw a scatter graph with a line of best fit and comment on the correlation. (5)

Maths	52	34	46	58	44	42	51	38	53	59
English	46	38	47	49	48	47	45	41	56	51

3. The time taken, in seconds, for a group of students to complete a race are shown below

Time	Frequency
$10 < t \leq 30$	4
$30 < t \leq 40$	8
$40 < t \leq 50$	12
$50 < t \leq 60$	10
$60 < t \leq 75$	6

a. Draw a histogram (5)
b. Draw a cumulative frequency graph (5)
c. Use your graph to estimate the median and quartiles (3)
d. Given that the quickest time was 15s and the longest was 75s, draw a box plot for the data (4)

Total 37 marks

End of Section Revision Quiz

These questions are based upon everything you have covered in the number section of the course. It is not compulsory to complete this quiz but it is recommended that you use it in order to make sure you have a firm understanding of all of the topics covered.
Answers are provided to these questions – if you're finding a question or group of questions particularly difficult you should go back to that section in your notes and go over it to make sure you understand it fully.
The majority of these questions are exam style questions so you should familiarise yourself with them now.

1. Find 51% of 350. Do not use a calculator.

2. Lucy wants to buy a tablet that costs £650. She has £122 already. Every week she gets paid £150, she saves 20% of her pay. How many weeks will she have to save for to be able to afford the tablet?

3. Lisa runs 200 metres in 1 minute. Using this data she says she can run 2,000 metres in 10 minutes. Do you think she's correct?

4. In order to take a group of students out there must be at least enough adults to satisfy the adult to students ratio of 1:8.

   a. How many adults are needed to take 64 students out?
   b. How many adults are needed for 36 students?
   c. If there are 9 adults, what is the maximum number of students that can go out?

5. Given the sets below draw a Venn diagram and write down the value of $P(S \cap E)$.
ε = {1,2,3,4,5,6,7,8,9,10,11,12,13,14,15}
S = {Square numbers}
E = {Even numbers}

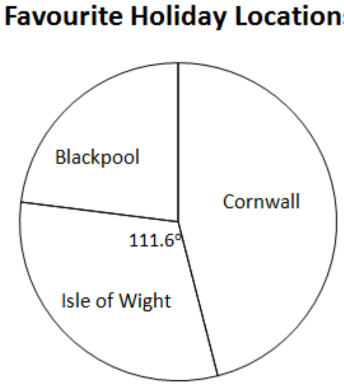
6. Here is a list of numbers. Find the mean, median, mode and range.

8	9	7	6	8	9	5	7	3	9	0	6	4

7. Three bags contain coloured marbles.
Bag A contains 5 blue, 3 red and 2 yellow.
Bag B contains 2 blue, 2 red and 1 yellow.
Bag C contains 8 blue, 7 red and 5 yellow.
You have to select your ball at random without looking, but you are allowed to decide which bag to pick from. If you wanted a blue ball, which bag should you choose?

8. Blue paint costs £2.50 per litre. Yellow paint costs £3.20 per litre. Blue and yellow paint is mixed in the ratio 3:2. Find the cost of 15 litres of mixture.

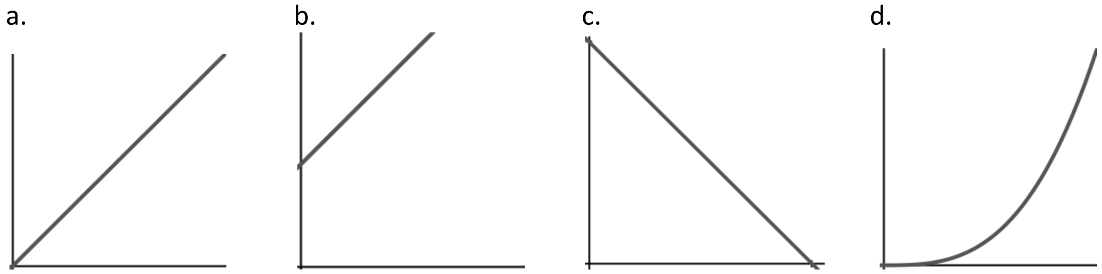
9. The pie chart shows information about favourite holiday locations in the UK of 500 people. There were twice as many people who preferred Cornwall compared to Blackpool. Find the number of people who chose each location.



10. Draw a bar chart to illustrate the numbers found in question 9

11. In a school the ratio of girls to boys is 8:5. There are 63 more girls than boys. Find the total number of students.

12. y is directly proportional to x. Which graph shows this?

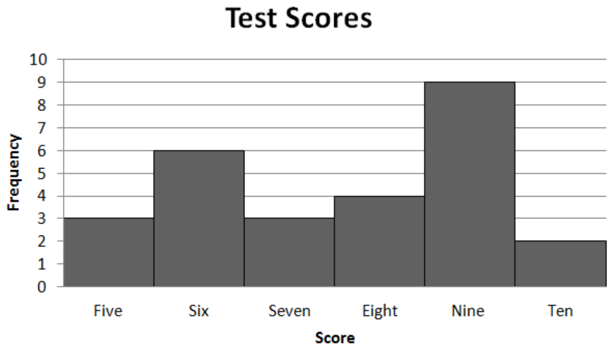


13. A trampoline club with 30 members enter a competition. 20% achieved a bronze medal, 10% got a silver medal and 17% got a gold medal. The rest didn't finish in a podium place – how many was this?

14. A pair of jeans costs £12 in a 30% off sale. What did they cost originally?

15. y is inversely proportional to x. When $y = 2, x = 4$.
   a. What does y equal when x is 3?
   b. What does x equal when y is 8?

16. The bar chart below shows the test scores of a class.



a. A child is picked at random. Find the probability that their test score is the median test score for this class.

b. Abbie's score is the same as the mode for this class. What score did she get?

17. £2500 is placed into a bank account. There are two different interest schemes: simple interest at 2.8% per year or compound interest at 1.1% each year. If the money is to be left in the bank for 5 years, which is the best interest scheme?

18. Rectangle A has sides of 4cm and 6cm. It is congruent to rectangle B which has a longer side of 15cm, find the length of the shorter side in rectangle B.

19. A bias coin is tossed then a bias die is thrown. The probability of getting heads is 0.6. The probability of getting an even number is 0.8.

Find the probability of

a. Heads and an even number

b. Heads and an odd number

c. Not getting tails then an odd number

20. Comment on the correlation of this set of data.

Test 1	2	4	3	9	1	2	4	9	8	10	3	6
Test 2	3	4	2	8	3	3	4	10	7	8	4	5

21. Draw a cumulative frequency graph and box plot for the set of data given in the table below.

The quickest time was 6s, the longest time was 48s

Time	Frequency
$5< t\leq 10$	3
$10< t\leq 15$	12
$15< t\leq 25$	15
$25< t\leq 35$	8
$35< t\leq 50$	4

Congratulations!

You have reached the end of the GCSE Higher course.
It's time to get revising for those exams – best of luck!