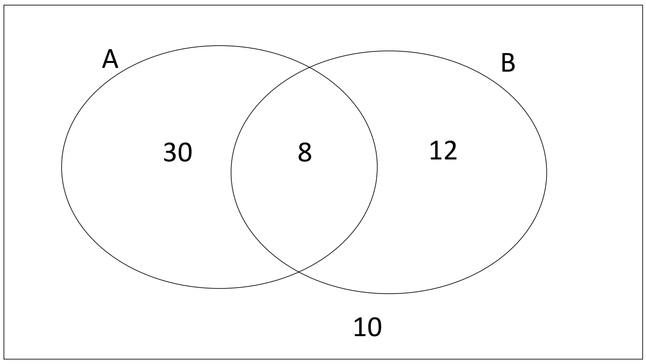
Statistics: Venn diagrams for calculating probabilities

Learning points

Venn diagrams are a convenient way to summarise data where items may belong to two or more different categories or **sets**.

Actual numbers of items may be shown, for example:

$\xi = 60$



The total number of data items included in the Venn diagram is given by the Greek letter ξ, known as xi. The total number of data items represented by the Venn diagram in this case is 60.

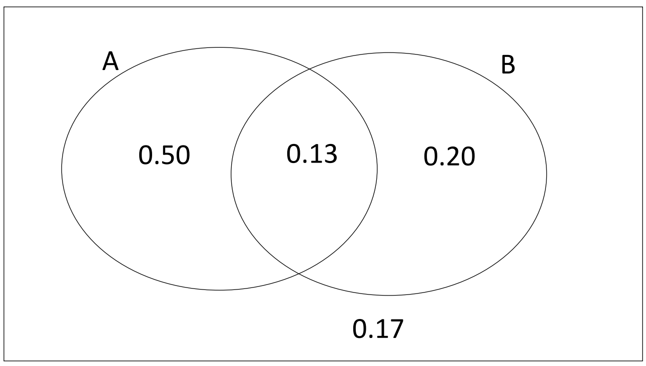
Two sets, A and B are present. 30 items are only in set A, and 12 items are only in set B.

8 items are in both set A and set B. This overlap is known as the **intersection** of the sets, and is written as: $A \cap B$

The total number of items in either set A or B, or both, is 50. This is known as the **union** of the sets, and is written as: $A \cup B$

10 items are not in either of the sets A or B. The area of the Venn diagram outside the circles can be written as: $(A \cup B)'$. The tick mark means **not**, so the area is not in either set A nor set B.

An alternative way of showing data values in a Venn diagram is as **probabilities**:



Each of the figures represents the probability that a data item selected at random will lie within that area of the Venn diagram. For example, if an item is selected at random, there is a 0.5 chance (50%) that it will be in set A but not in set B.

Probabilites are shown by the letter P, for example: $P(A \cap B) = 0.13$

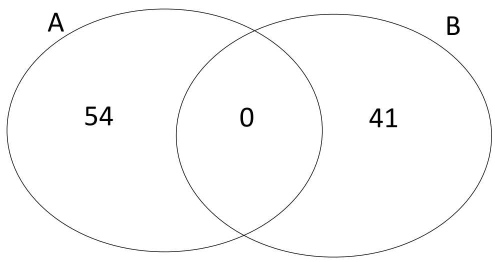
Notice that the total for the probabilities for every separate area of the Venn diagram will add up to 1, representing a certainty that a randomly selected data item will fit into one of the categories.

In some cases, data for every sub-area of the Venn diagram may not be available, and missing values need to be calculated based on various assumptions.

Mutually exclusive sets

It may be the case that data items cannot be in both of a pair of sets.
For example: Set A may represent people aged 21 and under, whilst set B represents people aged over 21. The sets are said to be mutually exclusive.

In this case, we know that the intersection of the sets will be empty:



Statistically independent events

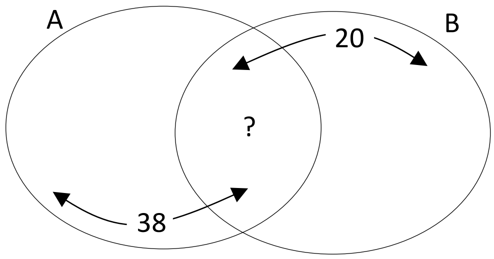
It may be possible to deduce that the event represented by set A has no influence on the event represented by set B.
For example: We might assume that a student's decision to join the college football team and the student's decision to study A-level History are independent events.
In this case, the probability of both events ocurring is found by multiplying the probabilities of the two separate events:

$$P(A \cap B) = P(A) \times P(B)$$

Union and intersection

If the number of data items in either the **union** or **intersection** of two sets is known, it may be possible to calculate the other value.

Consider the data represented by the partially constructed Venn diagram:



Suppose that it is known that there are a total of 38 data items in set A, and 20 data items in set B. It is also known that the two sets together contain 50 data items. We wish to know how many data items are in the intersection of the two sets, shown by the question mark in the diagram above.

The union and intersection of two sets are linked by the formula:

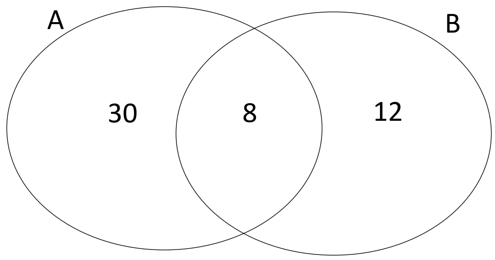
$$A \cup B = A + B - A \cap B$$

This formula makes sense, since the overlap of one of the sets needs to be removed to avoid duplication of the central area of the Venn diagram.

Using the data values given:

$$50 = 38 + 20 - A \cap B$$

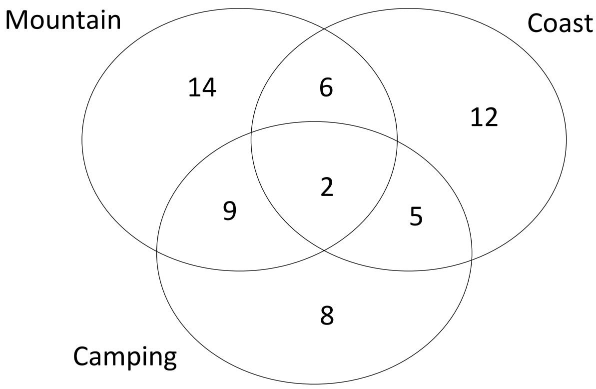
so $A \cap B = 8$. By subtration of this result, we can find the numbers of data items in set A only, or in set B only:



Making deductions from a sub-set of the Venn diagram

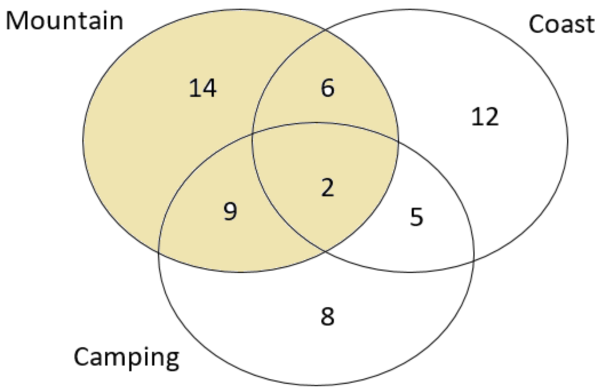
In some cases we are told that some event has already occurred, and are asked to find the probability of some further event occurring.
For example, the Venn diagram below shows the results of a survey of visitors to Wales. Each person was asked:

- Will you be climbing a mountain?
- Will you be walking along the coast path?
- Will you be camping?



Given that the visitor is intending to climb a mountain, we might be asked to calculate the probability that they will also be camping.

To solve this problem, only the area of the Venn diagram representing the given condition needs to be considered:



There are a total of 31 visitors who intend to climb a mountain. Of these, 11 will be camping. This represents a probability of:

$$\frac{11}{31} = 0.35$$

The necessary data can be obtained directly from the Venn diagram, or by using an equivalent formula:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

The vertical line symbol in $P(A|B)$ means "given that...". The probability of A occurring, given that B has already occurred, is equal to the probability of both A and B occurring, divided by the probability of B occurring.

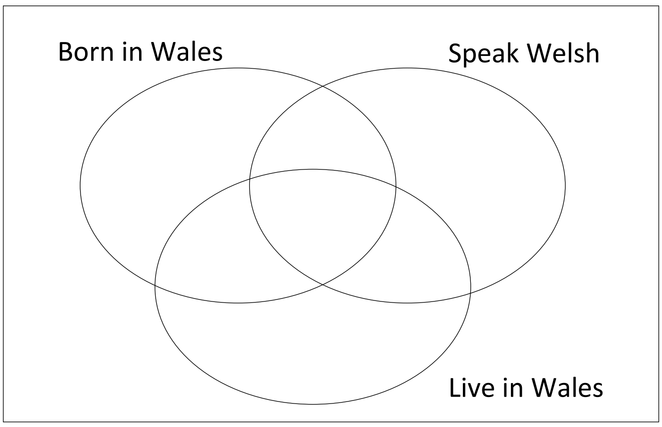
Practical applications

Venn diagrams as a research tool

Compiling a Venn diagram can be a useful method of assembling the data necessary to answer a research question in a subject such as: ecology, social science or economics.

As an example, there is currently an effort to increase the use of the Welsh language. It is of interest to investigate the distribution of Welsh speakers, both in Wales and in other countries. A Venn diagram can be constructed with three sets included in a survey. The questions asked would be:

- Were you born in Wales?
- Do you speak Welsh?
- Do you currently live in Wales?



The necessary data to complete the Venn diagram may be obtained from a number of sources, including the official UK census carried out every ten years, and information from the providers of Welsh language courses.

Once the data has been compiled, the Venn diagram could be used to answer a series of research questions such as:

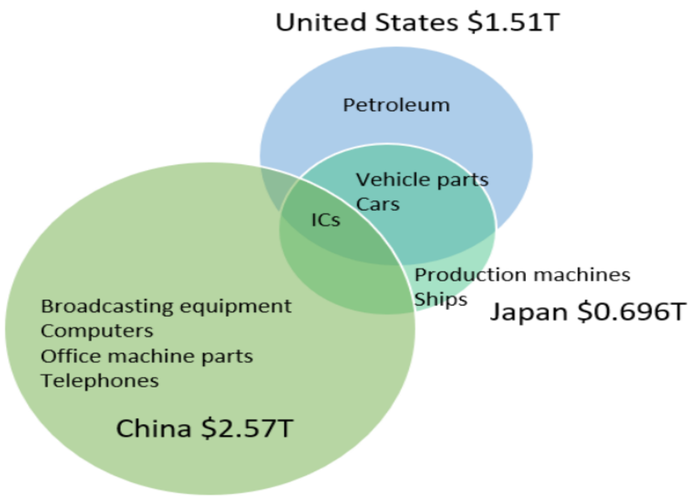
- What proportion of people born in Wales speak Welsh?
- What proportion of people who have moved to live in Wales have learned the language?
- What proportion of Welsh speakers born in Wales now live in other countries?

Venn diagrams in presentations and publications

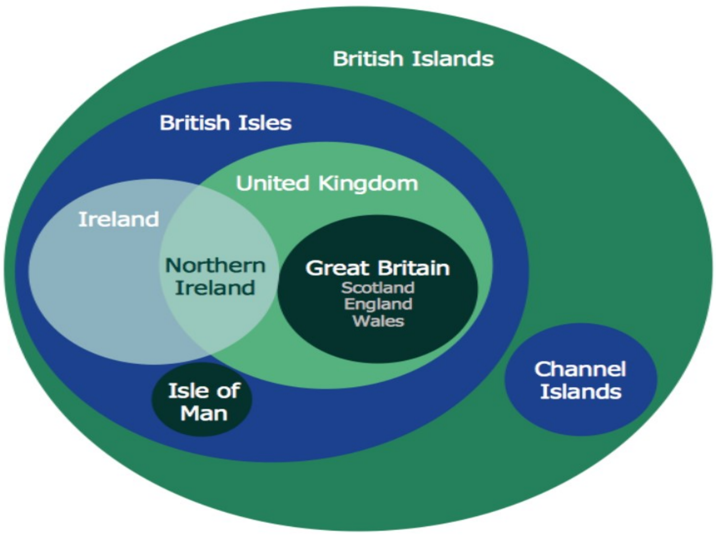
Probably the main use of Venn diagrams is to illustrate Powerpoint presentations or published articles in newspapers or journals. The objective is to demonstrate the often complex relationships which exist between sets of data.

In the Venn diagrams examined so far, the sizes of the circles representing data sets had no special significance. The circles merely provided containers for numerical data. However, it can be clearer for an audience if circles are scaled according to the numerical values they represent.

As an example, the Venn diagram below displays exports from the USA, Japan and China, using proportional circles to represent their relative values:



Many Venn diagrams are only intended to show the relationships between data sets. Numerical values are not included. For example: the Venn diagram below is intended to explain the difference between the terms 'Great Britiain', 'United Kingdon', and 'British Isles', which are often used inaccurately:



Large data set

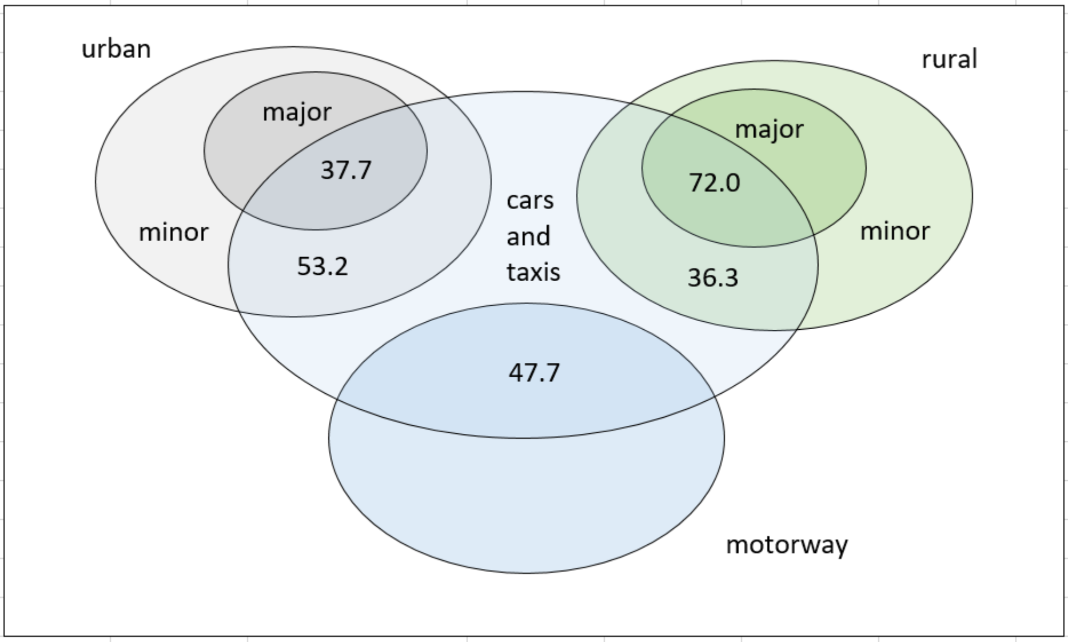
A Venn diagram is a suitable tool to compare various statistics within the large data set.

As an example, we will compare the distances travelled by cars and taxis on diferent types of roads during the year 2023.

Open the large data set and select the spreadsheet giving quarterly rolling annual data. Select the column headings for different types of roads and copy these to a new spreadsheet. Select car travel distances for a row of 2023 data, and add these to the new spreadsheet:

	Major Roads: Motorways	Major Roads: Rural 'A' Roads	Major Roads: Urban 'A' Roads	Minor Roads: Rural	Minor Roads: Urban	All Roads
2023	47.7	72.0	37.7	36.3	53.2	247.0

Use graphics tools provided by the spreadsheet to set up a Venn diagram. This should contain circles to represent major and minor urban and rural roads, and a circle for motorways. Add a circle to represent cars and taxis, so that it intersects the types of road.



Add the mileage figures to complete the Venn diagram.

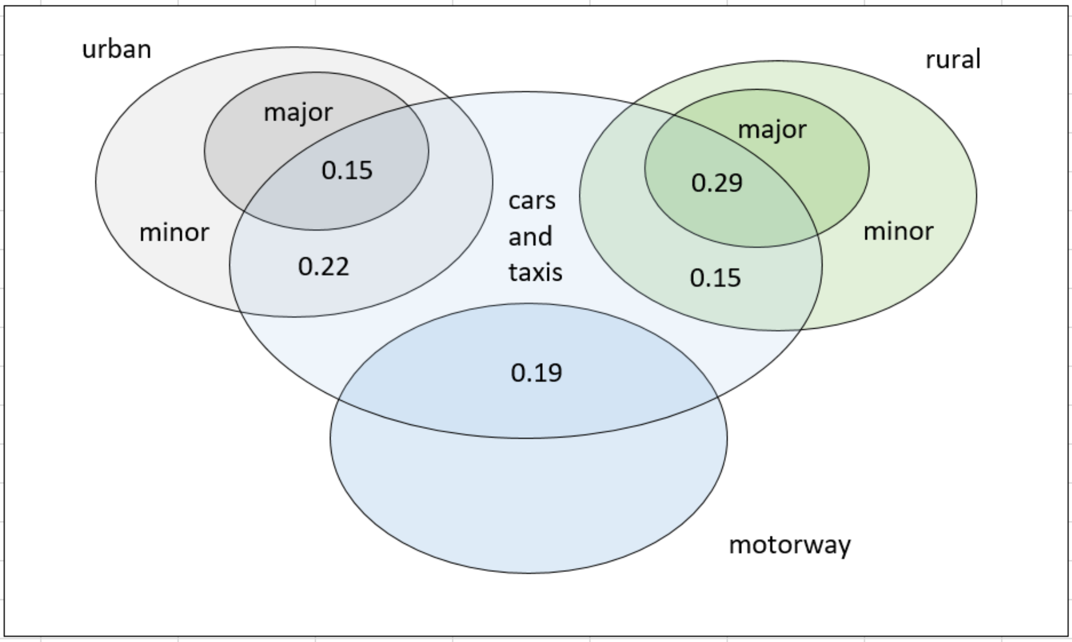
An alternative diagram can be produced to represent probabilites that a car selected at random would be covering its next journey mile on each of the different types of road.

Add a line to the speadsheet which divides each road mileage by the total mileage for all roads.
For example:

$$P(motorway) = \frac{47.7}{247.0} = 0.19$$

	Major Roads: Motorways	Major Roads: Rural 'A' Roads	Major Roads: Urban 'A' Roads	Minor Roads: Rural	Minor Roads: Urban	All Roads
2023	47.7	72.0	37.7	36.3	53.2	247.0
	0.19	0.29	0.15	0.15	0.22	1.00

Make a copy of the Venn diagram, then replace the mileage figures with the calculated probabilities:



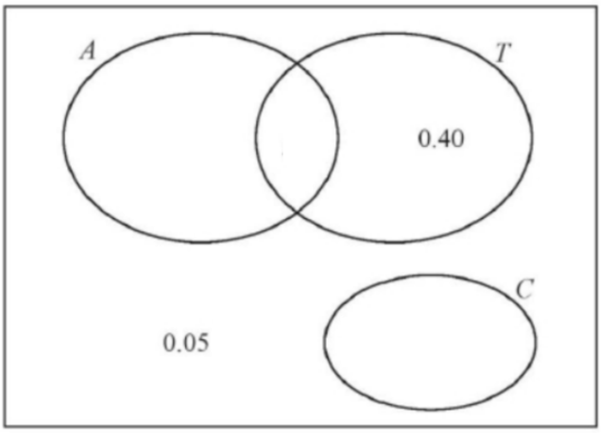
Multiple choice questions

1.

The Venn diagram shows the probabilities for students at a college taking part in various sports.

A represents the event that a student takes part in Athletics.
T represents the event that a student takes part in Tennis.
C represents the event that a student takes part in Cricket.

The probability that a student takes part in both Athletics and Tennis is the same as the probability that the student takes part only in Cricket.



The probability that a student takes part in Athletics or Tennis, or both, is 0.75

Find the probability that a student takes part only in Athletics.

Select an answer

A: 0.10 B: 0.15 C: 0.25 D: 0.20

Correct answer: B

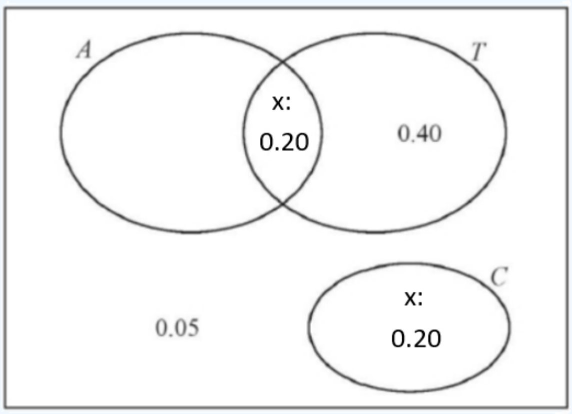
The probability that a student takes part in both Athletics and Tennis is the same as the probability that the student takes part only in Cricket. This can be designated x.

It is stated that the probability that a student takes part in Athletics or Tennis, or both, is 0.75

$$A \cup T = 0.75$$

Cricket and Athletics/Tennis are mutually exclusive, as there is no overlap between Cricket and Athetics/Tennis. The probability x that a student takes part in Cricket is therefore given by:

$$1 - (0.05 + 0.75 + x) \qquad so \quad x = 0.20$$



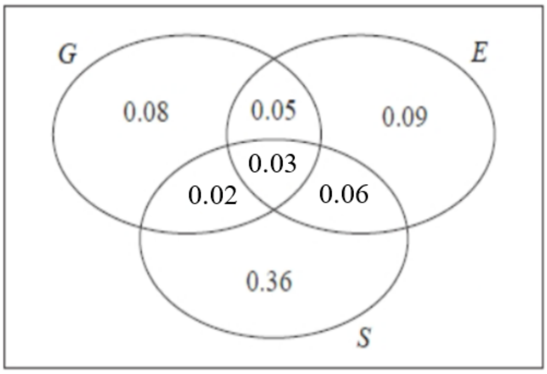
The remaining area of the Venn diagram, representing students who take part only in Athletics, has a probability:

$$1-(0.40+0.20+0.20+0.05) \quad = \quad 0.15$$

---

2. A large college produces three magazines.
One magazine is about green issues, one is about equality and one is about sports.
A student at the college is selected at random and the events *G*, *E* and *S* are defined as follows

G is the event that the student reads the magazine about green issues
E is the event that the student reads the magazine about equality
S is the event that the student reads the magazine about sports



The Venn diagram gives probabilities for sub-sets of readers.

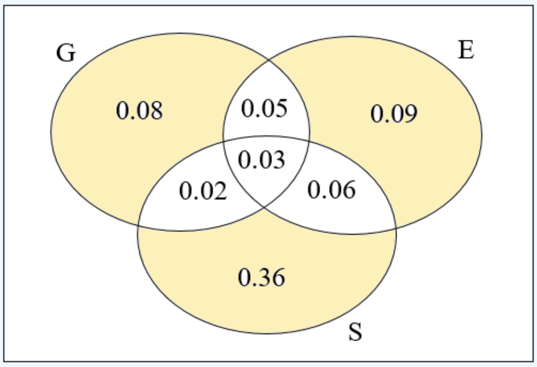
Find the probability that the student will read exactly one of the magazines.

Select an answer

A: 0.46 B: 0.69 C: 0.63 D: 0.53

Correct answer: D

The areas of the venn diagram representing probabilities of reading exactly one magazine are shaded:

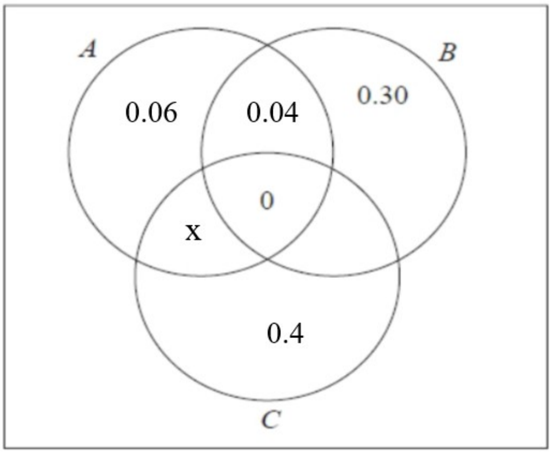


Overlapping circles indicate more than one magazine is read. The area outside the circles represents students who do not read any of the magazines.

The probability of a student reading eactly one magazine is:

$$0.08+0.09+0.36 \quad = \quad 0.53$$

3. The Venn diagram shows three events, *A*, *B* and *C*, and their associated probabilities.



Events *B* and *C* are mutually exclusive.
Events *A* and *C* are independent.

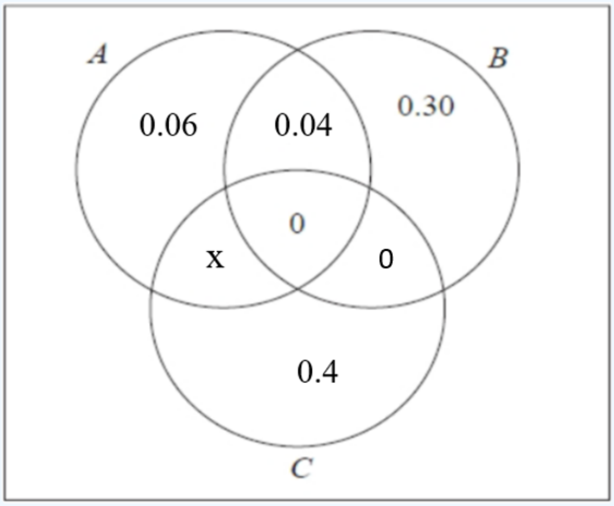
Find the value of x

Select an answer

A: 0.1 B: 0.06 C: 0.2 D: 0.18

Correct answer: A

Events B and C are mutually exclusive, meaning that both events cannot occur together. Probability $B \cap C$ is therefore 0.



Probability of event A is $0.06 + 0.04 + x \;=\; 0.10 + x$

Probability of event B is $0.4 + x$

Events A and C are independent, so $A \cap B = A \times B$

$$A \cap B = (0.10 + x) \times (0.40 + x)$$

$$x^2 + 0.5x + 0.04 = x$$

$$x^2 - 0.5x + 0.04 = 0$$

$$(x - 0.4)(x - 0.1) = 0$$

Possible solutions are: x = 0.4 or x = 0.1

If x = 0.4, the total probability obtained by adding all sub-sets will exceed 1, so this solution is not valid.

The probability of x is therefore 0.1

---

4. A person's blood group is determined by whether or not it contains any of 3 substances *A*, *B* and *C*.

A doctor surveyed 300 patients' blood and produced the table below.

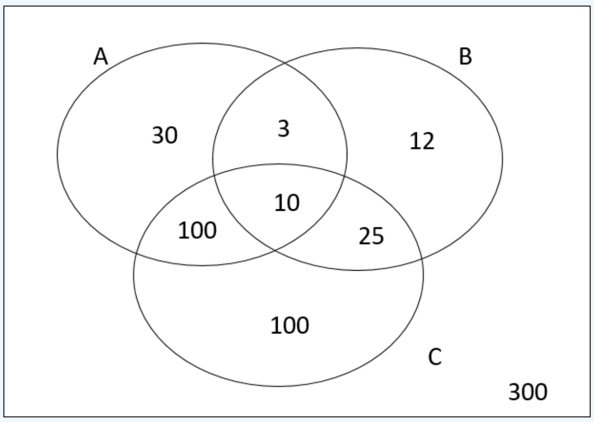
Blood contains	**No. of Patients**
only *C*	100
A and *C* but not *B*	100
only A	30
B and *C* but not *A*	25
only *B*	12
A, *B* and *C*	10
A and *B* but not *C*	3

Draw a Venn diagram to represent this information. Use this to find the probability that a randomly chosen patient's blood contains substance C.

Select an answer

A: 0.69 B: 0.78 C: 0.54 D: 0.82

Correct answer: B



The total number of patients with blood containing substance C is:

$$100 + 100 + 10 + 25 \quad = \quad 235$$

Expressing this as a fraction of the total number of patients in the survey:

$$\frac{235}{300} \quad = \quad 0.78$$

---

5. 40 people were surveyed regarding which games consoles they owned.

8 people said they owned both a Playstation (P) and an Xbox (X),
11 people said they owned both a Playstation (P) and a Nintendo (N),
7 people said they owned both an Xbox (X) and a Nintendo (N),
This data includes 2 people who said they owned all three consoles.

4 people said they owned none of the consoles.

Of those people who owned only one games console, twice as many owned a Nintendo as a Playstation, and half as many owned an Xbox as a Playstation.

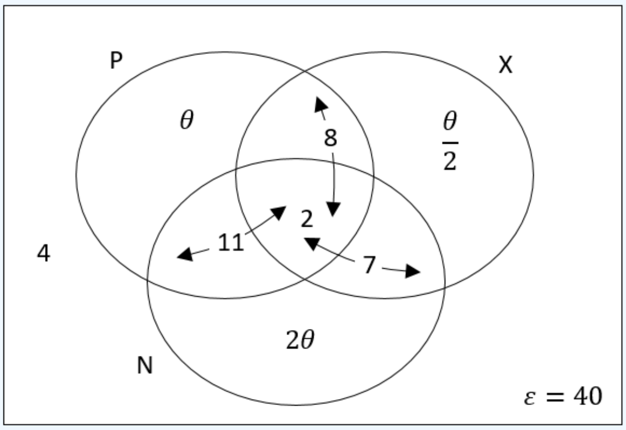
Draw a Venn diagram to illustrate the information given above. Use the diagram to determine the probability that one of the 40 people chosen at random ***does not own a Playstation***.

Select an answer

A: 0.56 B: 0.22 C: 0.39 D: 0.48

Correct answer: D

A diagram can be drawn to show the information as presented in the question. The symbol θ is used to represent the number of people owning only a Playstation.



From this diagram, we can deduce the numbers of people owning exactly two of the consoles are:

Playstation + Xbox: 6 persons
Playstation + Nintendo: 9 persons
Xbox + Nintendo: 5 persons

36 persons own at least one console, and 22 own either two or all three of the consoles.

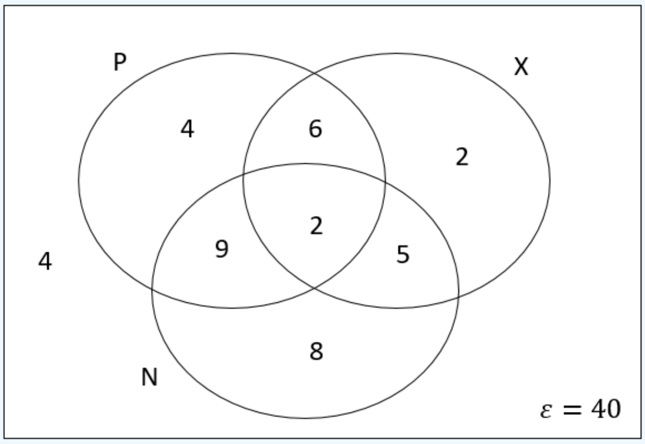
14 persons therefore own exactly one console. This is made up from:

$$\theta + 2\theta + \frac{\theta}{2}$$

where θ is the number of people owning only a Playstation.

$$\frac{7}{2}\theta = 14, \text{ so } \quad \theta = 4$$

A complete Venn diagram can now be drawn:

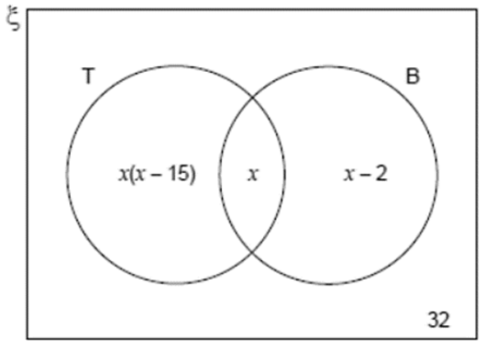


The number of people owning a Playstation is (4 + 6 + 9 + 2) = 21. The number **not** owning a Playstation is therefore 19, giving a probability of $\frac{19}{40} = 0.48$

---

6. The Venn diagram shows information about a coin collection.

ξ = 120 coins in the collection
T = coins from the 20th century
B = British coins



A coin is chosen at random.
It is British.
Work out the probability that it is from the 20th century.

Select an answer

A: 0.49 B: 0.34 C: 0.53 D: 0.61

Correct answer: C

The number of coins in either or both of the sets T and B is (120 – 32) = 88

These coins are represented by three sub-sets of the Venn diagram:

$$[x(x-15)] + [x] + [x-2] = 88$$

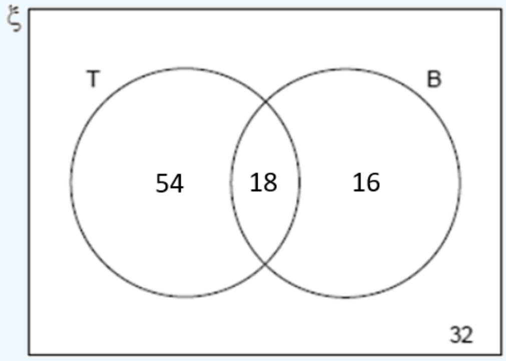
$$x^2 - 13x - 2 = 88$$

$$x^2 - 13x - 90 = 0$$

Factorising: $$(x-18)(x+5) = 0$$

Possible solutions are: x = 18 and x = -5. However, the number of coins cannot be negative, so the only valid solution is x = 18.

A complete Venn diagram can now be drawn:



It is given that the coin chosen is British. Of the 34 British coins, 18 are also 20th century, so the probability is: $\frac{18}{34} = 0.53$

---

7. 100 farms are surveyed in a region well known for growing cereal crops. Most of the farms grow either wheat, barley, or both.

   70 farms grow only wheat or only barley.

   $\frac{4}{5}$ of these 70 farms grow only wheat.

   The number of farms that grow wheat is three times the number that grow barley.

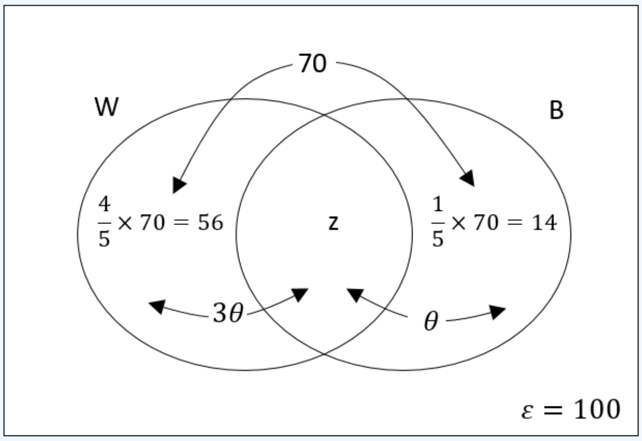
   Produce a Venn diagram to show this information, and use it to determine how many of the farms grow ***barley***, either on its own or with wheat.

   Select an answer

   A: 27 B: 18 C: 34 D: 21

Correct answer: D

A diagram can be produced to summarise the information given in the question.
θ is the number of farms which grow barley. Z is the number of farms which grow both crops:



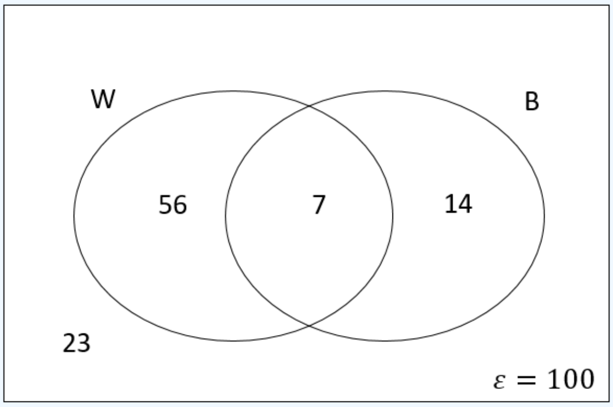
The number of farms growing wheat is three times the number growing barley:

$$(56 + z) = 3 \times (14 + z)$$

$$56 + z = 42 + 3z$$

$$14 = 2z, \quad \text{so } z = 7$$

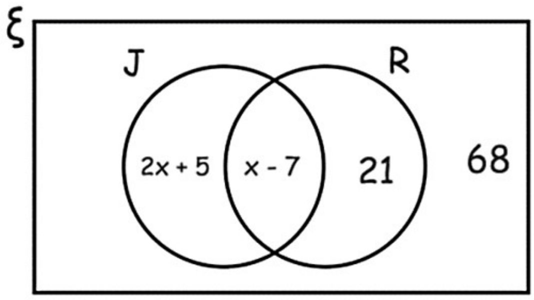
A complete Venn diagram can now be drawn:



The number of farms producing barley, either alone or with wheat, is (14 + 7) = 21

8. The Venn diagram shows information about the cars in a car park.

ξ = 150 cars in the car park
R = red cars
J = cars manufactured in Japan



A car is chosen at random. Find the probability that it is ***red***.

Select an answer

A: 0.15 B: 0.23 C: 0.18 D: 0.32

Correct answer: B

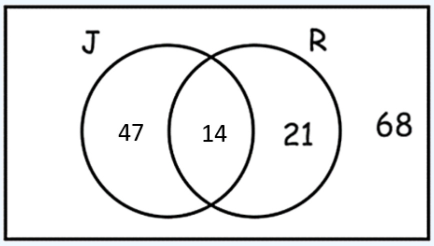
The number of cars which are red, made in Japan, or both, is: (150 – 68) = 82

This number is made up from three sub-sets of the Venn diagram:

$$[2x + 5] + [x - 7] + [21] = 82$$

$$3x = 63, \quad \text{so} \quad x = 21$$

The complete Venn diagram can now be drawn:

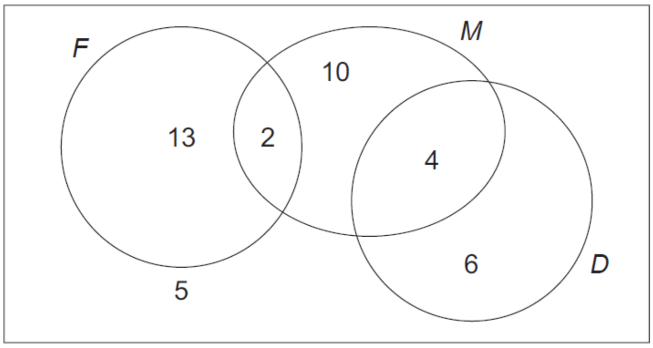


The probability of a car selected randomly from the car park being red is:

$$\frac{35}{150} = 0.23$$

Longer examination questions

1. The Venn diagram shows the subjects studied by 40 sixth form students. F represents the set of students who study French, M represents the set of students who study Mathematics and D represents the set of students who study Drama.
The diagram shows the number of students in each set.



a) Explain what $M \cap D'$ means in this context. [1]

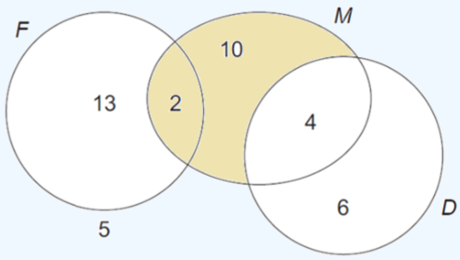
b) One of these students is chosen at random. Find the probability that this student studies

i) exactly two of these subjects,

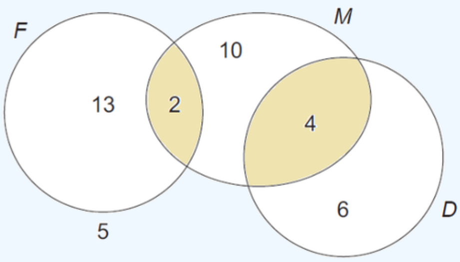
ii) Mathematics or French or both. [3]

c) Determine whether studying Mathematics and studying Drama are statistically independent for these students. [3]

(a) $M \cap D'$ is the intersection or overlap of the set M with the area of the Venn diagram which is <u>not</u> in the set D.

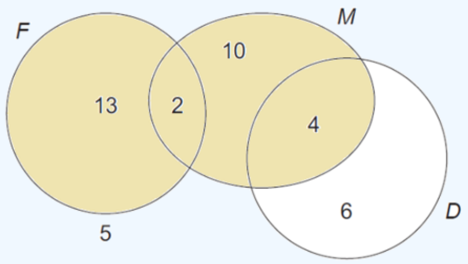


(b) The area of the Venn diagram representing students studying exactly two of the subjects is shaded:



This represents 6 out of 40 students, giving a probability of $\frac{6}{40} = 0.15$

The area of the Venn diagram representing students studying Mathematics or French or both is shaded:



This represents 29 out of 40 students, giving a probability of $\frac{29}{40} = 0.725$

(c) If studying Mathematics and studying Drama are independent events, then:

$$M \cap D = M \times D$$

16 students study Mathematics, giving a probability of $\frac{16}{40} = 0.4$

10 students study Drama, giving a probability of $\frac{10}{40} = 0.25$

The calculated probability of both occurring if the events are independent is $0.4 \times 0.25 = 0.1$
This represents 4 students.

The actual number of students studying both Mathematics and Drama is 4, so it is true that the events are statistically independent.

---

2. Given that

$$P(A) = 0.35\,, \quad P(B) = 0.45 \quad \text{and} \quad P(A \cap B) = 0.13$$

find

(a) $P(A \cup B)$ **(2)**

(b) $P(A' \mid B')$ **(2)**

The event C has $P(C) = 0.20$

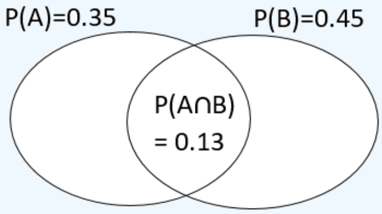
The events A and C are mutually exclusive and the events B and C are independent.

(c) Find $P(B \cap C)$ **(2)**

(d) Draw a Venn diagram to illustrate the events A, B and C and the probabilities for each region. **(4)**

(e) Find $P([B \cup C]')$ **(2)**

(a)

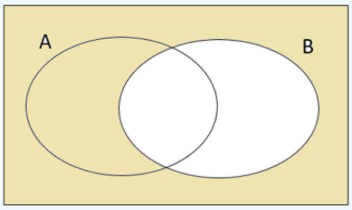


The union of the sets A and B is equal to the sum of the two sets, minus the overlapping intersection:

$$P(A \cup B) = A + B - A \cap B$$

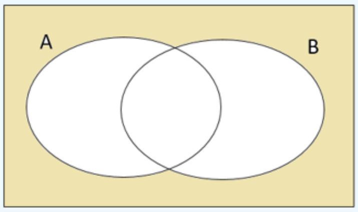
$$= 0.35 + 0.45 - 0.13 \quad = \quad 0.67$$

(b) The probability of not being in set B is represented by the shaded area:



This probability is equal to: $1 - 0.45 = \quad 0.55$

Given that an element in not in set B, the probability of also not being in set A is represented by the shaded area:



From the result calculated in part (a) above, this probability is equal to: $1 - 0.67 = \quad 0.33$

Given that an element is not is set B, the probability of not being in set A is:

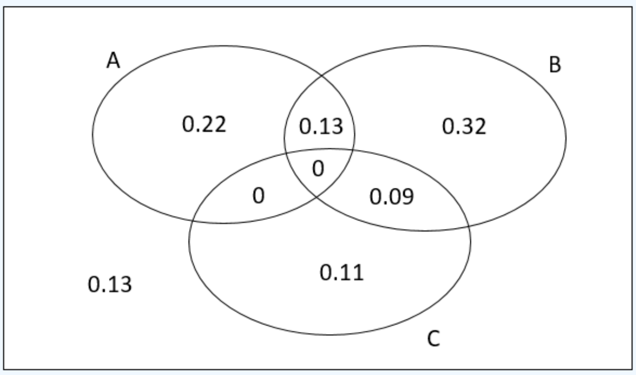
$$\frac{0.33}{0.55} = 0.6$$

(c) P(B) = 0.45, P(C) = 0.20. The events are independent, therefore:

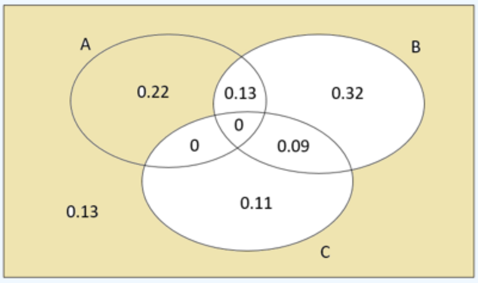
$$P(B \cap C) = A \times B$$

$$= 0.45 \times 0.2 \quad = 0.09$$

(d)



(e) $(B \cup C)'$ is the area of the Venn diagram which is in neither set B nor set C:



This has a probability of:

$$0.13 + 0.22 = 0.35$$

3.

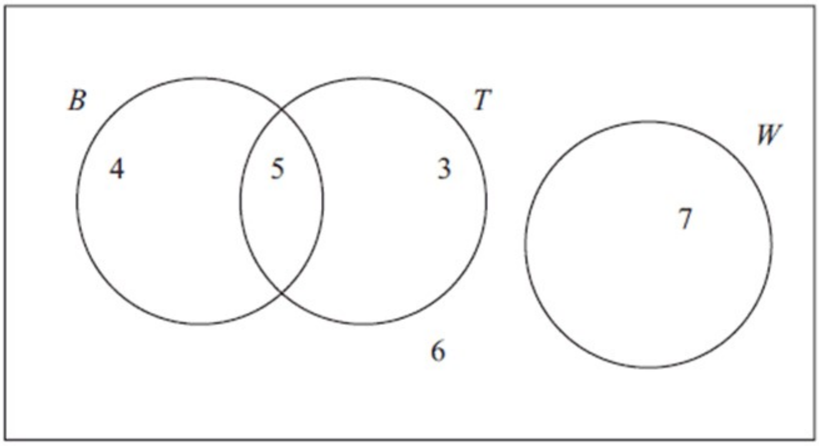


Figure 1

Figure 1 shows how 25 people travelled to work.

Their travel to work is represented by the events

B bicycle

T train

W walk

(a) Write down 2 of these events that are mutually exclusive. Give a reason for your answer. **(2)**

(b) Determine whether or not B and T are independent events. **(3)**

One person is chosen at random.

Find the probability that this person

(c) walks to work, **(1)**

(d) travels to work by bicycle and train.

(1)

(e) Given that this person travels to work by bicycle, find the probability that they will also take the train.

(2)

(a) Travelling by Train and Walking are mutually exclusive, as the circles on the Venn diagram do not intersect. There are no people represented who travel to work by both methods.

(b) If travelling by Bicycle and travelling by Train are independent events, then:

$$P(B) \times P(T) = P(B \cap T)$$

The probability of travelling by Bicycle is: $\frac{9}{25} = 0.36$

The probability of travelling by Train is: $\frac{8}{25} = 0.32$

If the events are independent, then the probability of both occurring is:

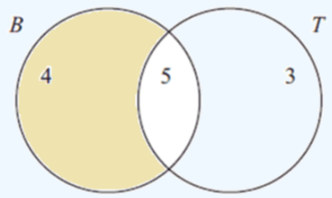
$$0.36 \times 0.32 = 0.115$$

This represents approximately 3 people. The actual number travelling by both Bicycle and Train is 5, so the events are not statistically independent.

(c) The probability of Walking to work is: $\frac{7}{25} = 0.28$

(d) The probability of travelling by Bicycle and Train is: $\frac{5}{25} = 0.20$

(e) Given that a person travels by Bicycle, they may then take the Train:



Of the 9 people travelling by Bicycle, 5 also take the Train.

The probability of taking the Train is therefore: $\frac{5}{9} = 0.56$

---

4. The following shows the results of a survey on the types of exercise taken by a group of 100 people.

65 run
48 swim
60 cycle
40 run and swim
30 swim and cycle
35 run and cycle
25 do all three

(a) Draw a Venn Diagram to represent these data.

(4)

Find the probability that a randomly selected person from the survey

(b) takes none of these types of exercise,

(2)

(c) swims but does not run,

(2)

(d) takes at least two of these types of exercise.

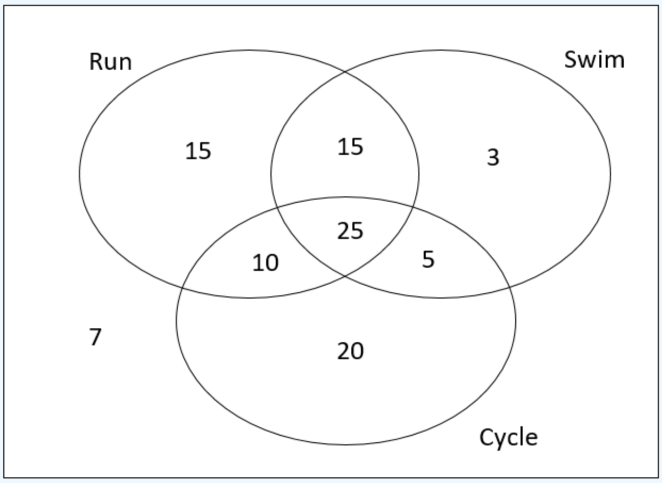
(2)

Jason is one of the above group.
Given that Jason runs,

(e) find the probability that he swims but does not cycle.

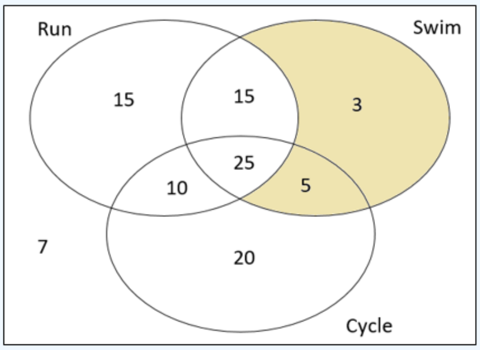
(3)

(a) $\varepsilon = 100$



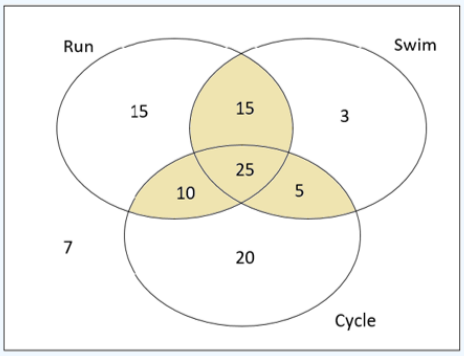
(b) The probability of a person taking none of the types of exercise is: $\frac{7}{100} = 0.07$

(c) The probability of swimming but not running is shown by the shaded area:



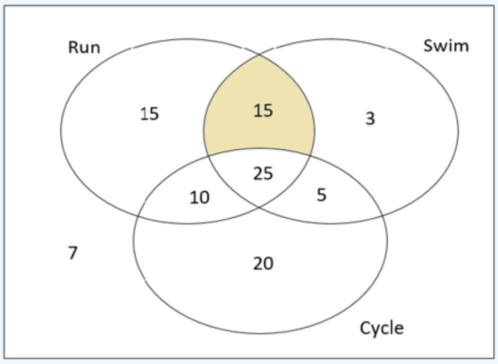
The probability is: $\frac{8}{100} = 0.08$

(d) The probability of taking at least two types of exercise is shown by the shaded area:



The probability is: $\frac{55}{100} = 0.55$

(e) Given that Jason runs, the probability of swimming but not cycling is shown by the shaded area:



Out of 65 people who run, 15 also swim but do not cycle. The probability is given by:

$$\frac{15}{65} = 0.23$$

---

5. The Venn diagram in Figure 1 shows the number of students in a class who read any of 3 popular magazines A, B and C.

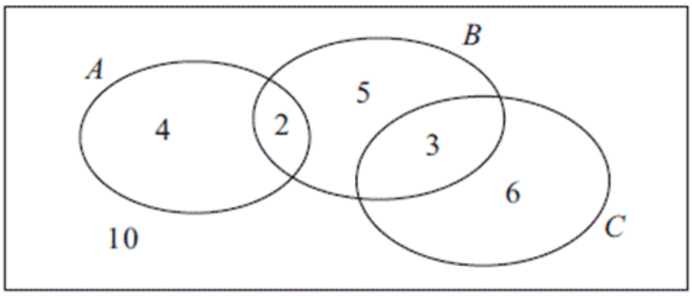


Figure 1

One of these students is selected at random.

(a) Show that the probability that the student reads more than one magazine is $\frac{1}{6}$.

(2)

(b) Find the probability that the student reads A or B (or both).

(2)

(c) Write down the probability that the student reads both A and C.

(1)

Given that the student reads at least one of the magazines,

(d) find the probability that the student reads C.

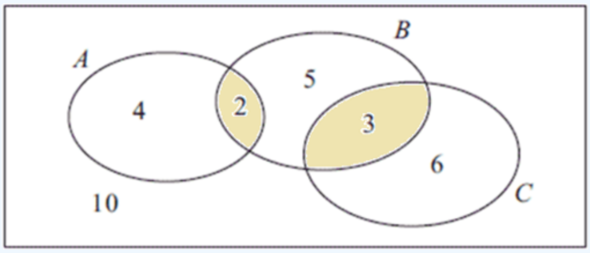
(2)

(e) Determine whether or not reading magazine B and reading magazine C are statistically independent.

(3)

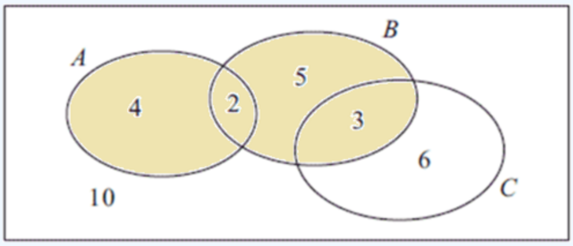
(a) The total number of students represented in the Venn diagram is 30.

The number of students reading more than one magazine is 5.



The probability of a student reading more than one magazine is therefore $\frac{5}{30} = \frac{1}{6}$

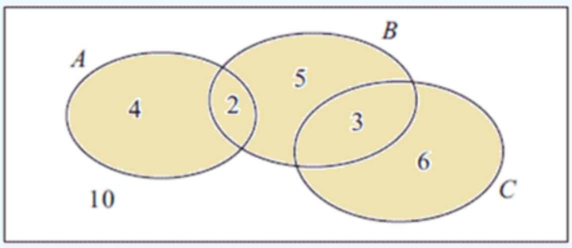
(b) The number of students reading magazine A or B, or both, is 14:



The probability of a student reading A or B or both is $\frac{14}{30} = 0.47$

(c) There is no overlap between circles A and C on the Venn diagram.
No student reads both magazine A and magazine C, so the probability is 0.

(d) Given that a student reads at least one magazine, we are concerned only with the area of the Venn diagram shaded below:



20 students read at least one magazine. Of these, 9 students read magazine C.

The probability is $\frac{9}{20} = 0.45$

(e) If reading magazine B and reading magazine C are independent events, then the probability of reading both, $P(B \cap C)$, is equal to the two individual probabilities P(B) and P(C) multiplied together.

The probability of reading B is $\frac{10}{30}$

The probability of reading C is $\frac{9}{30}$

The probabilities multiply to give: $\frac{10}{30} \times \frac{9}{30} = \frac{90}{900} = 0.10$ This is equivalent to 3 students

The actual number of students reading both magazines is 3, so it is true that the events are statistically independent.

---

6. There are 180 students at a college following a general course in computing. Students on this course can choose to take up to three extra options.

112 take systems support,
70 take developing software,
81 take networking,
35 take developing software and systems support,
28 take networking and developing software,
40 take systems support and networking,
4 take all three extra options.

(a) draw a Venn diagram to represent this information.

(5)

A student from the course is chosen at random.

Find the probability that this student takes

(b) none of the three extra options,

(1)

(c) networking only.

(1)

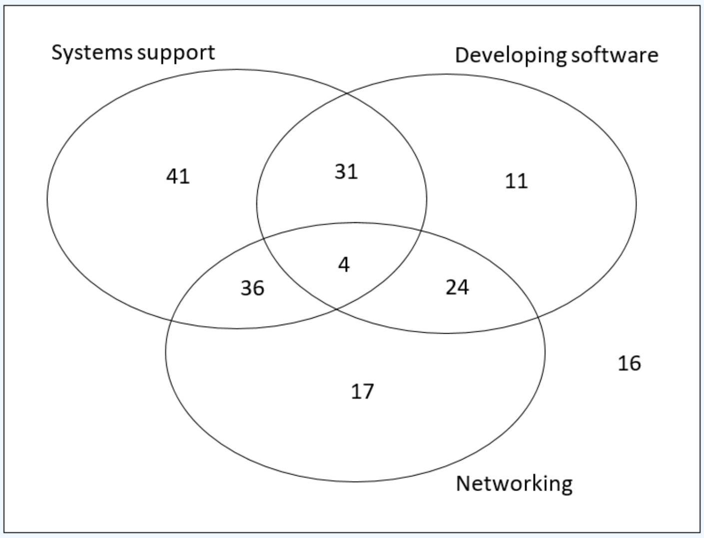
Students who want to become technicians take systems support and networking. Given that a randomly chosen student wants to become a technician,

(d) find the probability that this student takes all three extra options.

(2)

(a)

$\xi = 180$



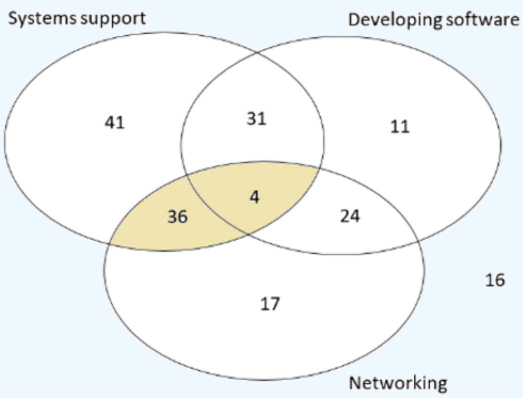
(b) 16 students take none of the options.

This reperesents a probability of $\frac{16}{180} = 0.09$

(c) 17 students take only the Networking option.

This reperesents a probability of $\frac{17}{180} = 0.09$

(d) Given that a student is taking Systems support and Networking, we are concerned only with the shaded area of the Venn diagram:



Out of 40 students taking Systems support and Networking, 4 take all three options.

The probability is given by: $\frac{4}{40} = 0.10$

7. The following shows the results of a wine tasting survey of 100 people.

96 like wine A,
93 like wine B,
96 like wine C,
92 like A and B,
91 like B and C,
93 like A and C,
90 like all three wines.

(a) Draw a Venn Diagram to represent these data.

(6)

Find the probability that a randomly selected person from the survey likes

(b) none of the three wines,

(1)

(c) wine A but not wine B,

(2)

(d) any wine in the survey except wine C,

(2)

(e) exactly two of the three kinds of wine.

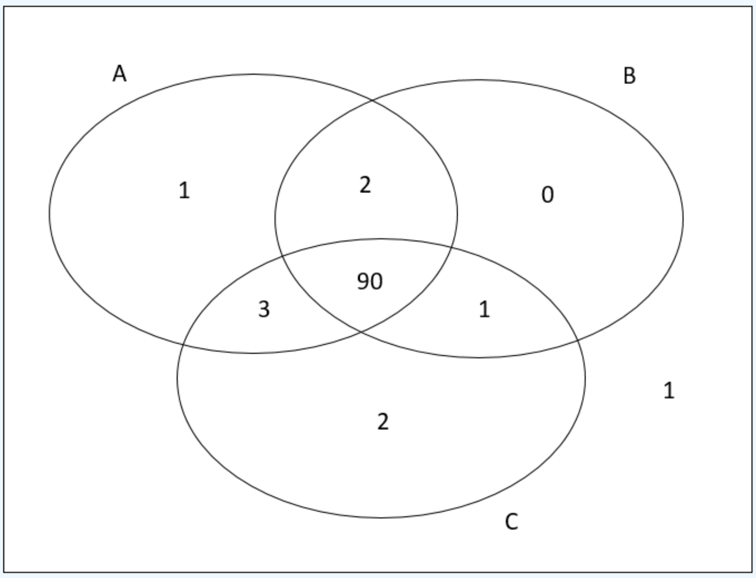
(2)

Given that a person from the survey likes wine A,

(f) find the probability that the person likes wine C.

(3)

(a) $\xi = 100$



(b) 1 person liked none of the wines.

This reperesents a probability of $\frac{1}{100} = 0.01$

(c) 4 people liked wine A but not wine B.

This reperesents a probability of $\frac{4}{100} = 0.04$

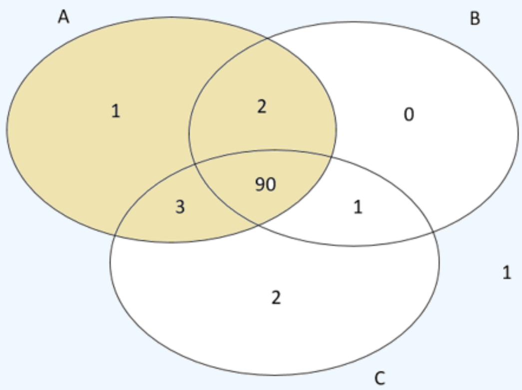
(d) Ignoring wine C, 97 people liked either of the other two wines A and B.

This reperesents a probability of $\frac{97}{100} = 0.97$

(e) 6 people liked exactly two of the wines.

This reperesents a probability of $\frac{6}{100} = 0.06$

(f) Given that a person likes wine A, we are dealing with only the shaded area of the Venn diagram:



96 people like wine A. Of this group, 93 also like wine C.

This reperesents a probability of $\frac{93}{96} = 0.97$

---

8. A survey of the reading habits of some students revealed that, on a regular basis, 25% read quality newspapers, 45% read tabloid newspapers and 40% do not read newspapers at all.

(a) Find the proportion of students who read both quality and tabloid newspapers.

(3)

(b) draw a Venn diagram to represent this information.

(3)

A student is selected at random. Given that this student reads newspapers on a regular basis,

(c) find the probability that this student only reads quality newspapers.

(3)

(a) If 100 students were surveyed, 60 would read at least one type of newspaper.

Out of these 60 students, 25 would read quality newspapers and 45 would read tabloids.

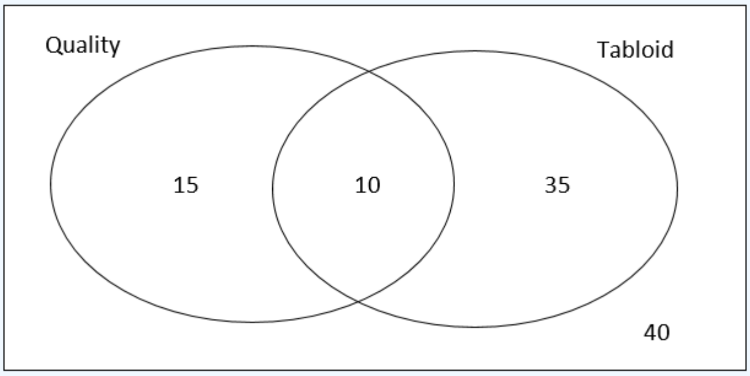
The number of students reading both types, $P(Q \cap T)$, is given by the equation:

$$P(Q) + P(T) - P(Q \cap T) = 60$$

$$P(Q \cap T) = (25 + 45) - 60 = 10\%$$

(b)

$\xi = 100$



(c) 60 students read at least one newspaper. Of these students, 15 read only a quality newspaper.

This reperesents a probability of $\frac{15}{60} = 0.25$