

Please check the examination details below before entering your candidate information

Candidate surname

Other names

Centre Number

Candidate Number

Pearson Edexcel

Level 1/Level 2 GCSE (9–1)

Specimen Paper (Set 2)

(Time: 1 hour 30 minutes)

Paper Reference **1ST0/2F**

Statistics

Paper 2

Foundation Tier

You must have:

Ruler graduated in centimetres and millimetres, protractor,
pair of compasses, pen, HB pencil, eraser, scientific calculator.

Total Marks

Instructions

- Use **black** ink or ball-point pen.
- **Fill in the boxes** at the top of this page with your name, centre number and candidate number.
- Answer **all** questions.
- Answer the questions in the spaces provided
– *there may be more space than you need.*
- Scientific calculators may be used.
- You must **show all your working out** with **your answer clearly identified** at the **end of your solution**.



Information

- The total mark for this paper is 80
- The marks for **each** question are shown in brackets
– *use this as a guide as to how much time to spend on each question.*

Advice

- Read each question carefully before you start to answer it.
- Try to answer every question.
- Check your answers if you have time at the end.

Turn over ►

S61446A

©2018 Pearson Education Ltd.

1/1/



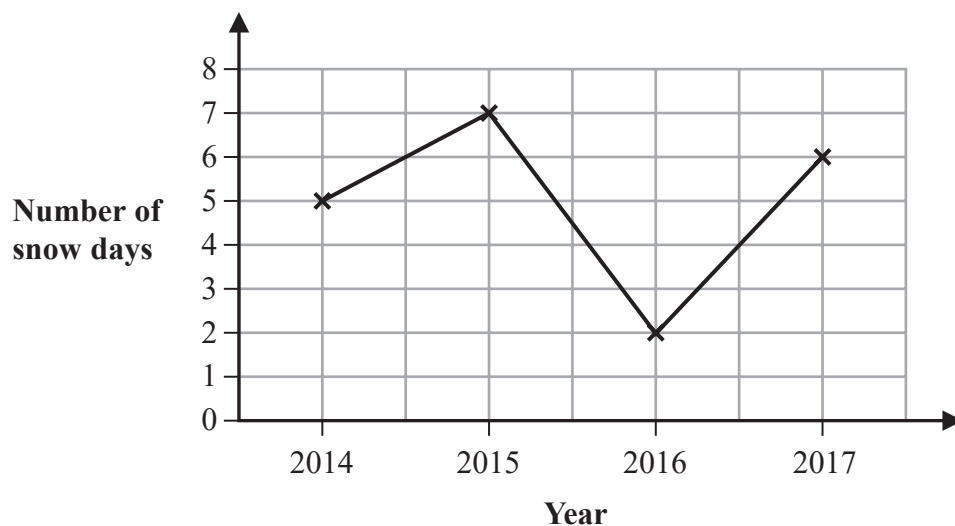

Pearson

Answer ALL questions.

Write your answers in the spaces provided.

You must write down all the stages in your working.

- 1 The graph shows information about the number of snow days in Flurrytown each year from 2014 to 2017



- (a) Write down the year with the greatest number of snow days.

(1)

- (b) Work out the total number of snow days in Flurrytown from 2014 to 2017

(2)

- (c) Explain why it would not be appropriate to start the year axis of the graph at 0

(1)

(Total for Question 1 is 4 marks)



- 2 Antonia is collecting information about the runners in a race.

Here are some types of data.

categorical continuous ordinal discrete

- (a) Use a word from the list above to complete correctly each of the following sentences.

The time taken to run the race is data.

The finishing position (1st, 2nd, 3rd, etc.) of the runner is data.

The number of runners in the race is data.

(2)

- (b) Write down an example of **qualitative** data that Antonia could collect about the runners.

.....
(1)

There are 600 runners in the race.

- (c) State what Antonia could use to help her process quickly the data she collects about the runners.

.....
(1)

(Total for Question 2 is 4 marks)



- 3 The table shows the likelihood of each of five outcomes of an experiment.

Outcome	Likelihood
A	impossible
B	likely
C	certain
D	evens
E	very unlikely

- (a) Write down the outcome with the greatest probability.

.....
(1)

Two outcomes are less likely to occur than outcome D.

- (b) Write down these two outcomes.

..... and
(1)

- (c) Write down the probability that outcome D will occur.

.....
(1)

- (d) Write down the probability that outcome A will occur.

.....
(1)

- (e) Write down the probability that outcome A will **not** occur.

.....
(1)

(Total for Question 3 is 5 marks)



- 4 The incomplete two-way table gives some information about the numbers of tennis players by age and by gender for each of the top 30 male and top 30 female tennis players in January 2018

Age	male	female	Total
18 to 22	2	4	
23 to 29	15		
30 to 36			21
Total	30	30	

(Source: www.wtatennis.com and www.atpworldtour.com)

- (a) Complete the two-way table.

(2)

- (b) Compare the ages of the male tennis players with the ages of the female tennis players. You should make **two** comparisons.

(2)

(Total for Question 4 is 4 marks)



- 5 Sadie is investigating the number of pets that each of the students in her school has.

Sadie's report on her investigation is given below.
The report is not complete.

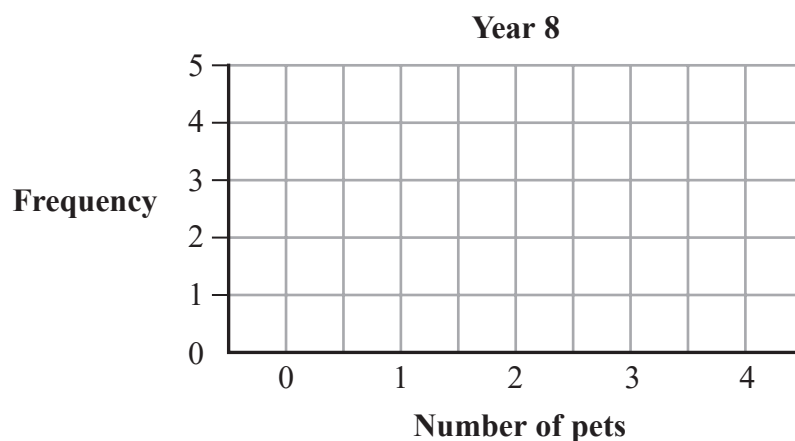
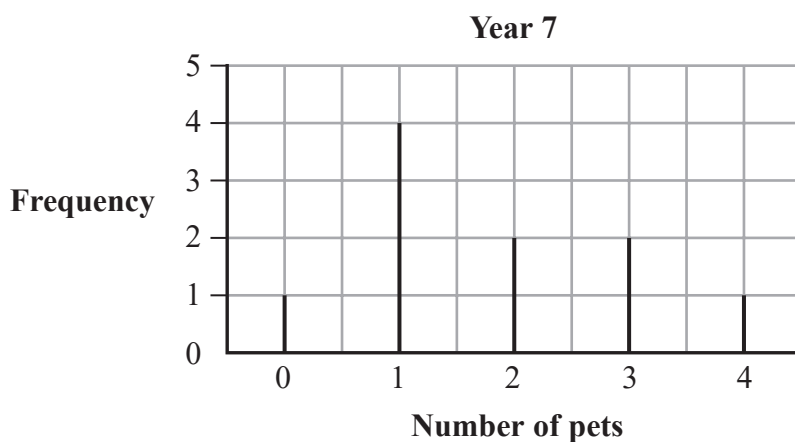
- (a) Complete Sadie's report.

Hypothesis: I believe that students in Year 7 have more pets than students in Year 8

Data collection: I am going to ask 10 students in Year 7 and 10 students in Year 8 how many pets they each have.

Data representation: I am going to represent the data using a bar line chart for each year group.

Number of pets	Year 7	Year 8
0	1	4
1	4	1
2	2	0
3	2	3
4	1	2



The mode of the number of pets for Year 7 is

The mode of the number of pets for Year 8 is

The mode for Year 7 is than the mode for Year 8

Conclusion:

Comparing the modes suggests that.....

.....

.....

(6)

(b) Suggest an improvement that could be made to Sadie's method for data collection.

.....

.....

(1)

(c) Suggest an improvement that could be made to Sadie's method for data representation.

.....

.....

(1)

(d) Comment on the validity of the conclusion to Sadie's investigation.

.....

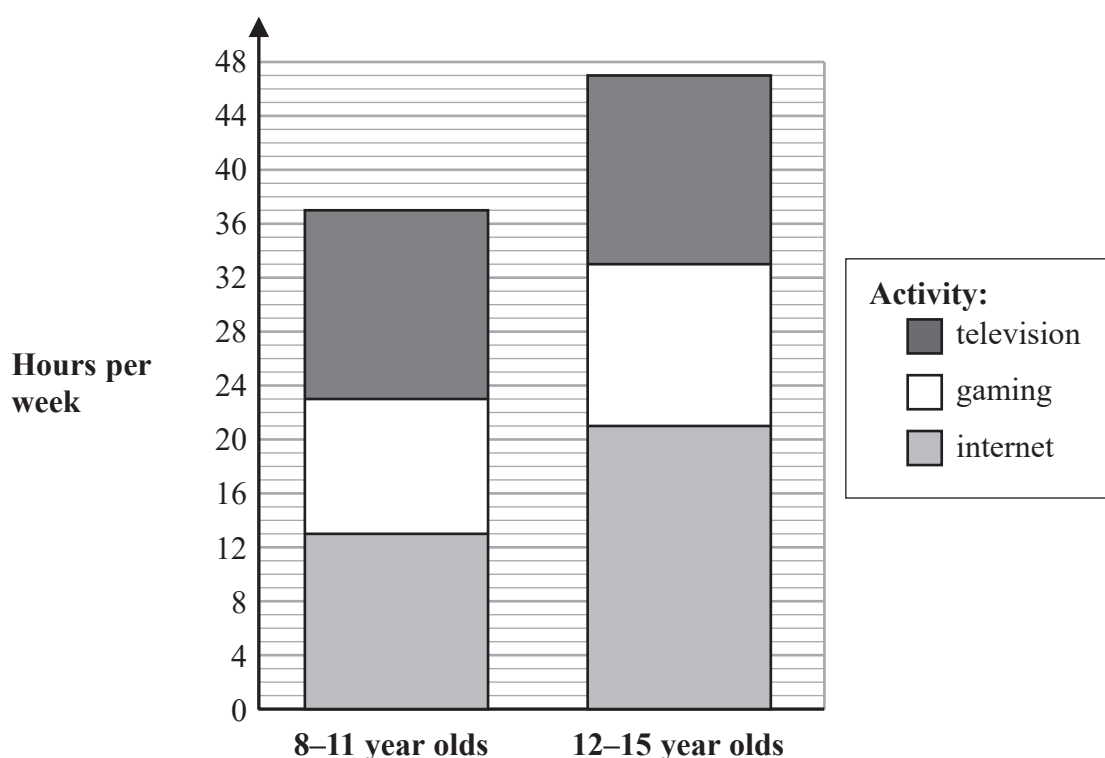
.....

(1)

(Total for Question 5 is 9 marks)



- 6 The composite bar charts give information about the amounts of times children spend each week using the internet, gaming and watching television.



(Source: www.ofcom.org.uk)

The total time spent on these three activities per week is less for 8–11 year old children than it is for 12–15 year old children.

- (a) Explain how the composite bar charts show this.

(1)

- (b) Compare the time spent watching television per week by 8–11 year old children and by 12–15 year old children.

(1)

The information was obtained by asking children how much time they spent doing these activities each week.

- (c) Comment on the reliability of this information.

(1)

(Total for Question 6 is 3 marks)



7 The mayor of a town wants a new library built.

He wants to survey the residents of the town to find out where each resident wants the library to be built.

The mayor has to decide between three different sampling methods for his survey.

Method 1: Opportunity sampling by sampling the first 100 residents who are available.

Method 2: Cluster sampling by dividing the town into 8 districts, selecting one of the districts at random and surveying all the residents in that district.

Method 3: Stratified sampling by dividing the residents into groups by age and randomly selecting a sample of residents from each age group in proportion to the number of residents in that group.

(a) Give one **advantage** of Method 1

(1)

(b) Give one **disadvantage** of Method 2

(1)

There are 6200 residents in the town of which 930 are aged 60 or over.

A sample of 150 residents of the town is taken, stratified by age.

(c) Calculate the number of residents aged 60 or over who should be included in the sample.

(2)

The mayor wants the sample to be representative of the residents of the town.

(d) Which of the three methods of sampling should the mayor use?

Give a reason for your answer.

(2)

(Total for Question 7 is 6 marks)



8 A vet wants to estimate the average weight of adult female shorthair cats.

She took a sample of 8 adult female shorthair cats that were brought to her clinic.

Here are the cats' weights in kilograms.

4.0 4.5 9.8 5.1 4.1 4.5 4.2 3.9

(a) Identify the outlier.

.....kg
(1)

The vet plans to estimate the average weight of adult female shorthair cats using this data.

(b) Give her advice.

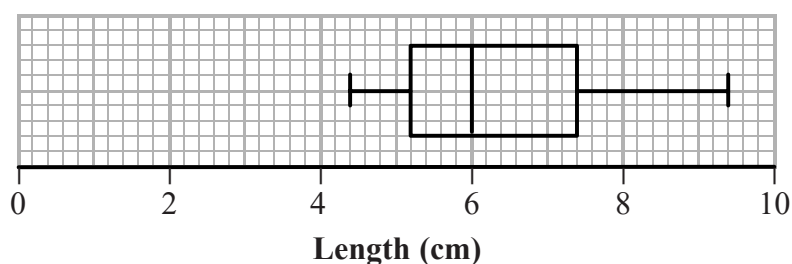
You should consider the data she plans to use and which average she should choose.
Give reasons for your answer.

(5)

(Total for Question 8 is 6 marks)



- 9 Some students were asked to draw a 5 cm long line without using a ruler. The actual length of each line was then measured. The box plot shows information about the distribution of the actual lengths of the lines that were drawn by the students.



- (a) Using the information given by the box plot, comment on the accuracy of the students in drawing a 5 cm long line.

(2)

- (b) Describe the skew of the distribution.

(1)

- (c) Describe what information is given about this distribution by its skewness.

(1)

(Total for Question 9 is 4 marks)



- 10 The table gives information about the times taken by each of 117 swimmers to complete their 50m freestyle race at the World Championships in 2017

Time taken (t seconds)	Frequency
$20 \leq t < 22$	10
$22 \leq t < 24$	68
$24 \leq t < 26$	24
$26 \leq t < 28$	12
$28 \leq t < 30$	3

(Source: www.fina.org)

Caleb suggests using a histogram to represent this information.

- (a) Discuss whether or not a histogram is a suitable diagram to use.

(1)

Caleb also suggests using the mean to summarise the data.

- (b) Discuss whether or not the mean is a suitable average to use.

(1)

Caleb uses the information in the table and linear interpolation to find an estimate of the median. He gets an estimate of 23.4 seconds.

Caleb says that exactly 50% of the swimmers took more than 23.4 seconds to complete the 50m freestyle race.

- (c) Comment on whether or not Caleb's conclusion is appropriate.

(1)

(Total for Question 10 is 3 marks)



DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

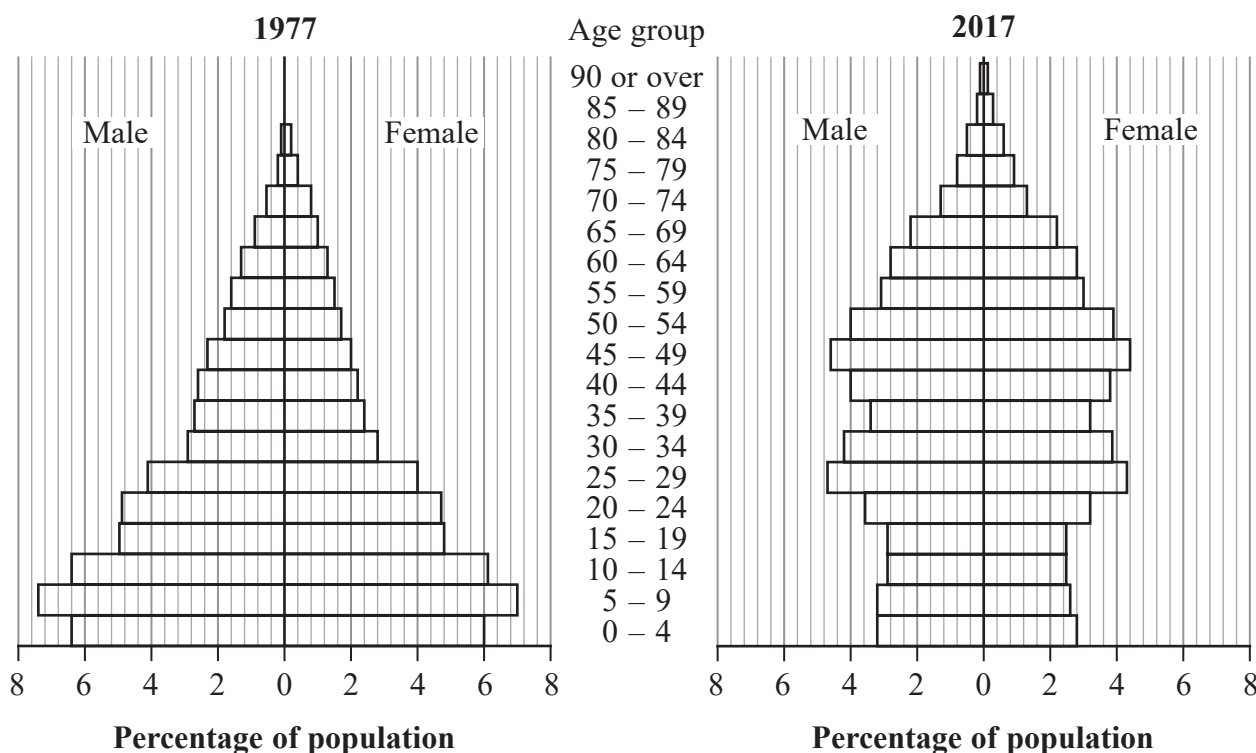
BLANK PAGE



S 6 1 4 4 6 A 0 1 3 2 0

- 11 The two population pyramids show the percentages of males and females in each age group in China for the years 1977 and 2017

Each percentage is based on the total population of China for that year.



(Source: populationpyramid.net/China)

- (a) For the year 1977, write down the percentage of the population who are female in the age group 0 – 4 years.

.....%

(1)

- (b) For the year 2017, write down the age group that has the greatest percentage of females.

.....

(1)

Eric says that more than 25% of the population of China is aged 9 or under in 1977

- (c) Determine whether or not Eric is correct.
You must show your working.

(3)



- (d) Compare the percentages of people in China in the 60 – 64 age group in 1977 with the percentages of people in China in the 60 – 64 age group in 2017

(2)

(Total for Question 11 is 7 marks)



- 12 The table shows information about the number of car journeys per person in a year that are a distance of between 0 and 50 miles.

The information is based on a sample of 382 people from the 2016 National Travel Survey.

Distance (x miles)	$0 < x \leq 1$	$1 < x \leq 2$	$2 < x \leq 5$	$5 < x \leq 10$	$10 < x \leq 25$	$25 < x \leq 50$
Frequency	24	65	131	83	62	17

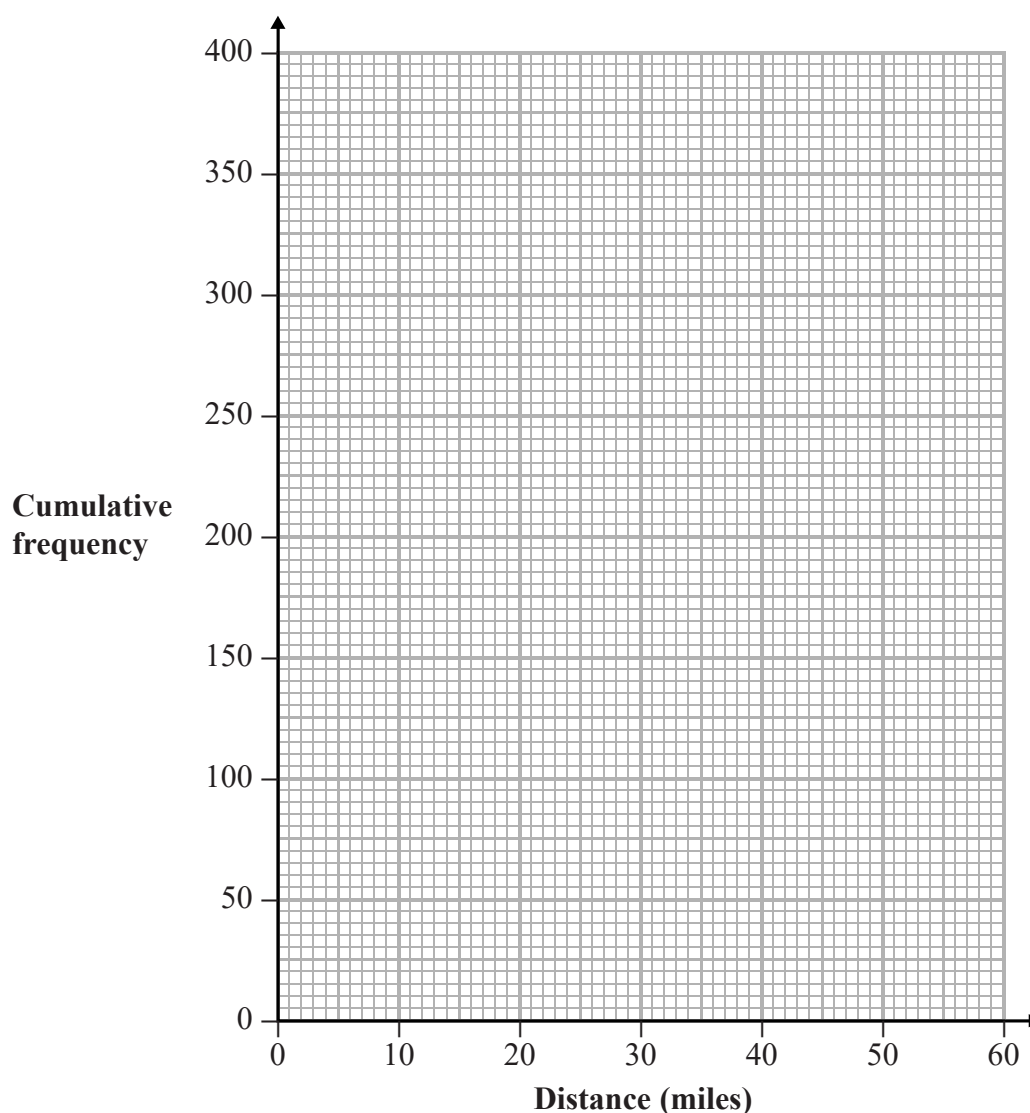
(Source: www.gov.uk)

- (a) Complete the cumulative frequency table for the information above.

Distance (x miles)	$0 < x \leq 1$	$0 < x \leq 2$	$0 < x \leq 5$	$0 < x \leq 10$	$0 < x \leq 25$	$0 < x \leq 50$
Cumulative frequency	24					

(1)

- (b) Draw a cumulative frequency diagram for this information.



(3)



In the same survey, the median distance travelled by train for a sample of people was 11.3 miles.

Hamish says,

“The information collected in the survey shows that people in these samples travel on average further by train than by car.”

(c) Assess whether or not Hamish’s conclusion is appropriate.

(2)

Hamish recorded the number of passengers travelling on 250 buses.

Information about his results is shown in the cumulative frequency table below.

Number of passengers	35	36	37	38	39	40	41	42	43	44	45
Cumulative frequency	5	12	24	39	65	98	138	176	207	234	250

Hamish plans to draw a cumulative frequency step polygon for his results rather than the type of cumulative frequency diagram drawn in part (b).

(d) Explain why Hamish’s plan is appropriate.

(1)

Hamish wants to take a sample of the bus passengers.

Hamish plans to take a sample of 50 men and 50 women.

(e) Name this sampling method and explain why this might be an appropriate method to use.

(2)

(Total for Question 12 is 9 marks)



- 13 The back-to-back stem and leaf diagram gives information about the ages of a random sample of members of parliament in Canada and in the UK.

Canada		UK
7 3 2 2	3	5 7
8 6 4 4 4 3	4	0 0 1 3 3 6 8 8
7 5 4 4 3 1 0	5	2 3 4 5 6 7
9 5 1 1 1	6	0 2 2 2 3 8 9
4 3 1	7	7 9

Key:

2|3|5 represents an age of 32 for a member of parliament in Canada and an age of 35 for a member of parliament in the UK

(Source: *en.wikipedia.org*)

- (a) Give a reason to support the use of a back-to-back stem and leaf diagram to represent this information.

(1)

Some information about the quartiles of these two distributions is given in the table below.

	Canada	UK
lower quartile	44	b
median	a	54
upper quartile	61	c

- (b) Find the value of a , the value of b and the value of c

$a =$

$b =$

$c =$

(3)



- (c) Write down the proportion of members of parliament in the UK that are likely to be older than 54 years old.
Give a reason for your answer.

(1)

- (d) Compare the spread of ages for members of parliament in Canada with the spread of ages for members of parliament in the UK.

State clearly the values of the statistic you use to make your comparison.
Interpret your comparison.

(3)

One member of parliament in the UK wants to investigate the ages of the people living in her constituency.

She suggests using the electoral register as a sample frame for her investigation.

- (e) State one use of a sample frame in an investigation.

(1)

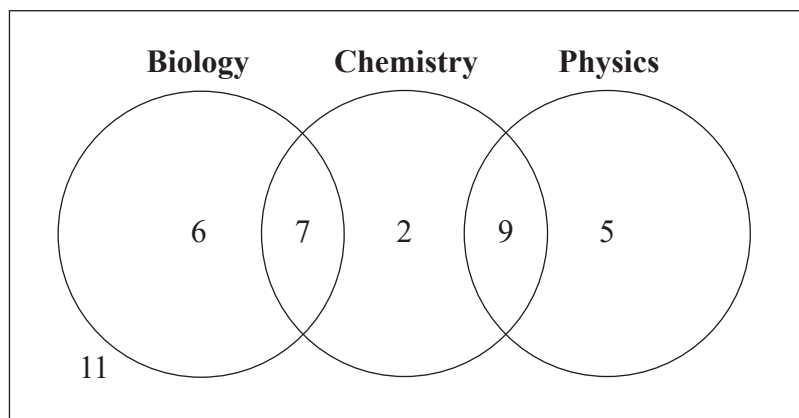
- (f) Assess the suitability of using the electoral register as a sample frame for this investigation.

(2)

(Total for Question 13 is 11 marks)



- 14 There are 40 students in Year 12 at a sixth form college. The Venn diagram gives information about the numbers of students studying Biology, Chemistry and Physics.



One of the 40 students is selected at random.

- (a) Write down the probability that this student

(i) studies Biology,

.....
(1)

(ii) studies Chemistry and Biology.

.....
(1)

X is the event that the student selected studies Chemistry.
 Y is the event that the student selected studies Physics.

- (b) Find

(i) $P(X)$

.....
(1)

(ii) $P(X \text{ and } Y)$

.....
(1)

(iii) $P(Y | X)$

.....
(1)

(Total for Question 14 is 5 marks)

TOTAL FOR PAPER IS 80 MARKS

