Pearson Edexcel Level 1/Level 2 GCSE (9–1)

Time 1 hour 30 minutes

| Paper reference | **1ST0/1F** |

Statistics
PAPER 1
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 ▶

1 The incomplete pictogram gives information about the numbers of cars sold from a garage in August and in September.

Month	Number of cars
August	⬜⬜ ⬜⬜ ⬜⬜ ⬜
September	⬜⬜ ⬜
October	
November	

Key:

⬜⬜ represents 10 cars

In October, 25 cars were sold from the garage.

(a) Complete the pictogram for October.

(1)

(b) Compare the number of cars sold from the garage in August to the number of cars sold from the garage in September.
Give a reason for your answer.

....

....

....

....

(2)

In November, 17 cars were sold from the garage.

(c) Explain why the use of this key may **not** be appropriate for representing 17 cars in the pictogram.

....

....

(1)

(Total for Question 1 is 4 marks)

2 Rahul has a 5-sided spinner.
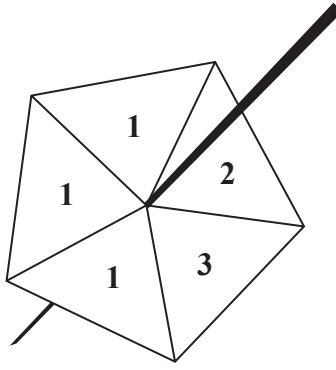
The spinner is a fair spinner.

(a) Explain what the word **fair** means in this description of the spinner.

....

(1)

Rahul is going to spin his spinner once.

(b) Underline the word from the list below that best describes the likelihood that the spinner will land on 3

impossible   unlikely   evens   likely   certain

(1)

(c) Underline the word from the list below that best describes the likelihood that the spinner will land on 4

impossible   unlikely   evens   likely   certain

(1)

Chloe has a 6-faced dice.

She rolls the dice 60 times and records the number on which the dice lands each time.

Here is information about Chloe's results.

Number	Frequency
1	9
2	19
3	11
4	10
5	1
6	10

(d) Explain what the information in the table tells you about whether the dice is fair.

....

....

....

....

....

(2)

(Total for Question 2 is 5 marks)

3 The time series graph shows the percentages of households in the United Kingdom with internet access from 2000 to 2018
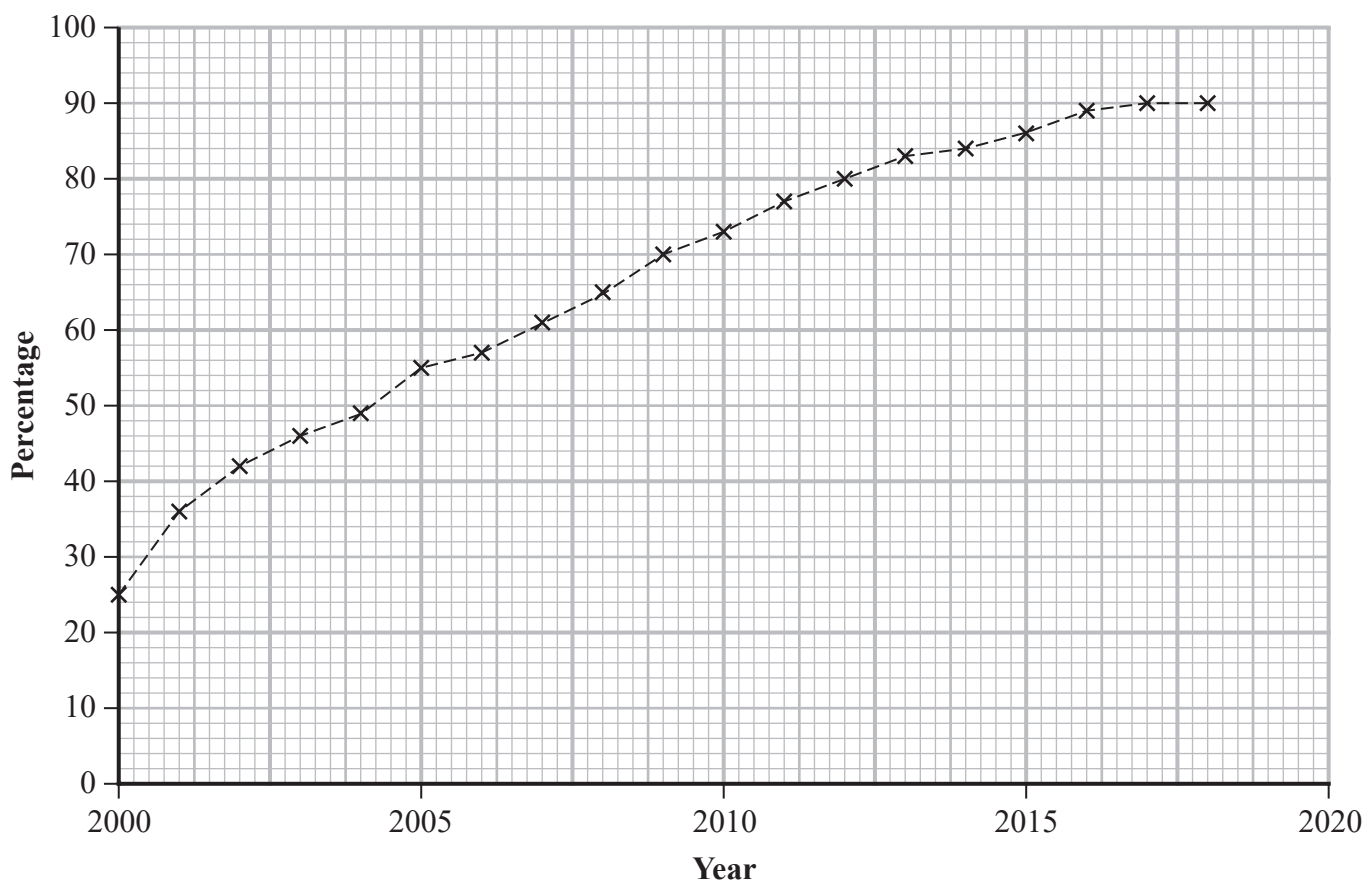
(Source: *Office for National Statistics*)

(a) Write down an estimate for the percentage of households with internet access in

(i) 2000

.... %

(ii) 2018

.... %

(2)

The percentage of households with internet access in 2019 was 93%

(b) Plot this information on the time series graph.

(1)

(c) Write down the first year where the percentage of households with internet access was greater than 50

....

(1)

Using the information given in the time series graph,

(d) compare the change in percentage of households with internet access from 2001 to 2005 with the change in percentage of households with internet access from 2015 to 2019

You must show how you get your answer.

....

....

....

....

(2)

(Total for Question 3 is 6 marks)

4 Alexa needs to know the population of California.

Using the internet, Alexa found that an estimate for the population of California in 2019 is 39.5 million people.

(Source: *www.worldpopulationreview.com*)

(a) Explain why this statistic is an example of secondary data.

....

....

(1)

(b) Give one advantage and one disadvantage of using secondary data.

Advantage ....

....

Disadvantage ....

....

(2)

(Total for Question 4 is 3 marks)

5 Jane is a car mechanic at MPG Autos.

Jane recorded the time, in minutes, that she took to change a tyre on each of 9 cars.

Here are her results.

| 25 | 32 | 19 | 27 | 17 | 24 | 35 | 29 | 15 |

(a) Explain whether or not her results have a mode.

....

....
(1)

(b) Find the median of her results.

.... minutes
(2)

(c) Show that the range of her results is 20 minutes.

(1)

The range of the times that Jane took to change the oil in each of these 9 cars is 28 minutes.

(d) Are the times that Jane took to change a tyre on each of these 9 cars more consistent than the times that she took to change the oil in each of these 9 cars?
Give a reason for your answer.

....

....
**(1)**

(Total for Question 5 is 5 marks)

6 Northern Stationery Supplies has only three branches, one in each of Manchester, Rotherham and York. Employees working at Northern Stationery Supplies work either in the office or they work in the warehouse.

The composite bar chart gives information about the numbers of employees of Northern Stationery Supplies who work in either the Manchester branch or in the Rotherham branch.
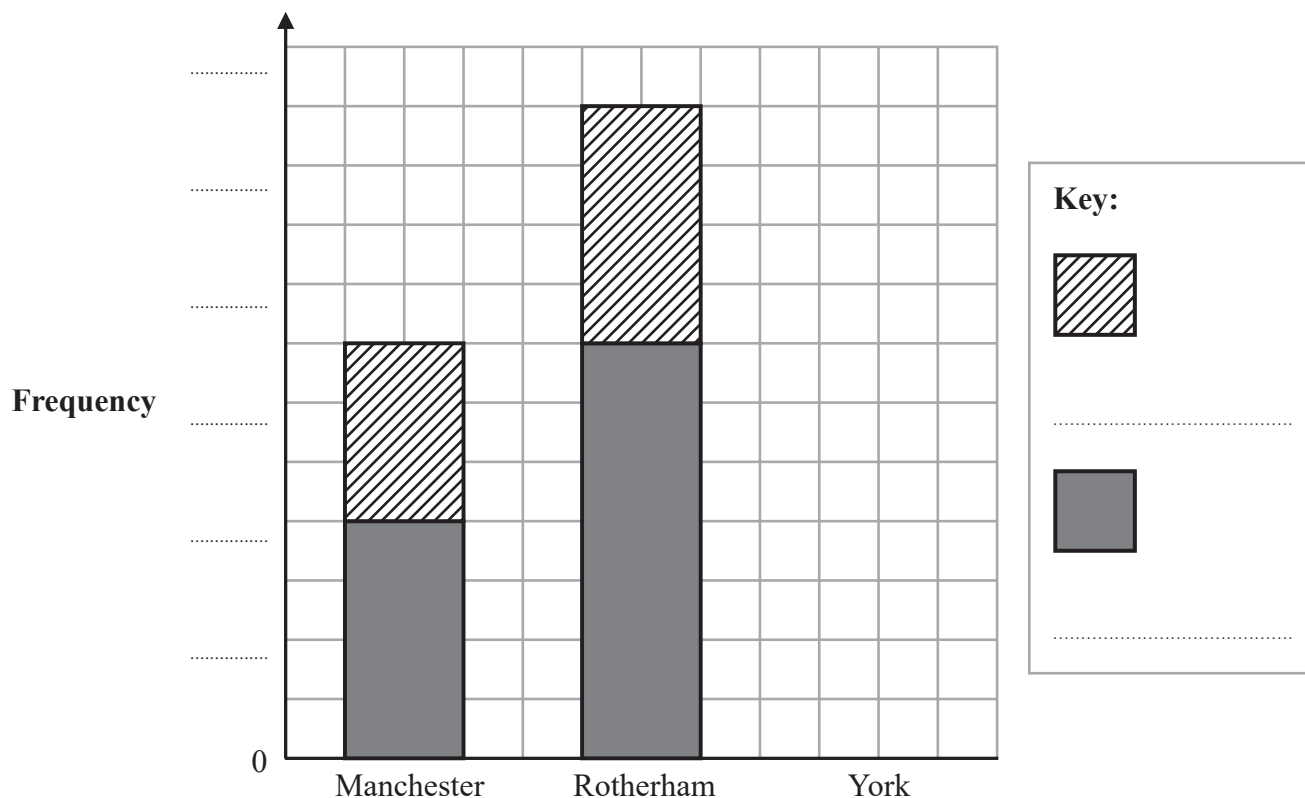
There is a total of 35 employees who work in the Manchester branch.

(a) Complete the labelling on the frequency axis.

(1)

20 employees work in the warehouse in the Rotherham branch.

(b) Complete the key.

(1)

In the York branch, 25 employees work in the office and 10 employees work in the warehouse.

(c) Complete the composite bar chart for the York branch.

(2)

(d) Find the total number of employees who work in the offices of Northern Stationery Supplies.

....

(1)

Northern Stationery Supplies has a total of 125 employees.

One employee from Northern Stationery Supplies is chosen at random.

(e) Find the probability that this employee works in the warehouse in the Manchester branch.

....

(2)

(Total for Question 6 is 7 marks)

7 Mrs Singh is the Headteacher of a school with 1200 students.

She wants to find out about the ways the students at her school use the internet to complete their homework.

She is going to write a plan for her enquiry.

Write down one thing Mrs Singh should include in her plan for each of

- data collection

- choice of sampling method

- a diagram to represent the data collected.

Explain why each of the things you have written down is appropriate.

....

....

....

....

....

....

....

....

....

....

....

....

(Total for Question 7 is 6 marks)

8 Sheffield City Council has more than 8000 employees.
These employees work at different locations in Sheffield.

(Source: *Wikipedia.org*)

A researcher wants to find out about the different methods of transport that these employees use to get to work.

She plans to ask the 9 employees, who work in one library, how they got to work that morning.

Give two reasons why this is **not** a good plan.

Reason 1 ....

....

Reason 2 ....

....

(Total for Question 8 is 2 marks)

9   Dylan investigates the ages of 150 people at a music festival. The youngest person at the festival is 21 years old and the oldest person is 68 years old.

Dylan decides to use a grouped frequency table to show his results.

(a) (i) Give **two** advantages of using grouped data rather than raw data.

....

....

....

....

(2)

(ii) Give **one** disadvantage of using grouped data rather than raw data.

....

....

(1)

Dylan is considering using one of the two possible frequency tables, **Table A** or **Table B**, shown below.

<div>

Table A

Age (a years)	Frequency
$0 < a \leqslant 20$	0
$20 < a \leqslant 40$	73
$40 < a \leqslant 60$	66
$60 < a \leqslant 80$	11
$80 < a \leqslant 100$	0

Table B

Age (a years)	Frequency
$20 < a \leqslant 30$	6
$30 < a \leqslant 40$	67
$40 < a \leqslant 50$	53
$50 < a \leqslant 60$	13
$60 < a \leqslant 70$	11

</div>

Dylan claims that **Table B** gives more detail than **Table A** about his results.

(b) Assess whether or not Dylan's claim is appropriate.

....

....

....

....

(2)

Dylan wants to work out the average age of the 150 people at the music festival.
He decides to use **Table B**.

(c) Calculate an estimate for the mean age of the 150 people at the music festival.
   Give your answer to one decimal place.

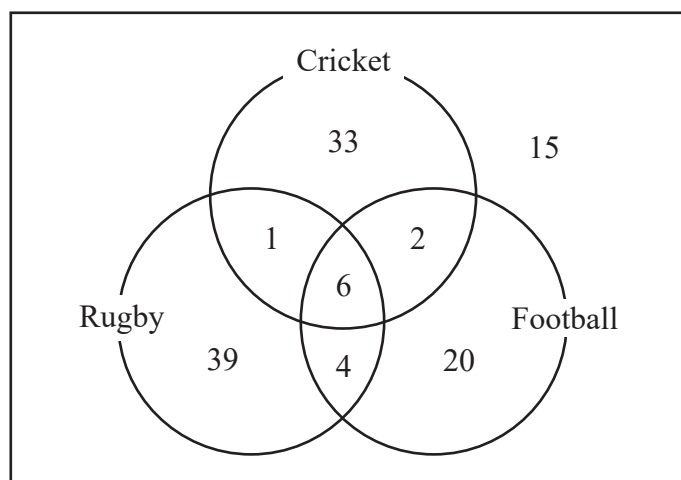... years

(3)

(Total for Question 9 is 8 marks)

P68619A01528

10 120 adults were each asked if they enjoy watching any of the sports Football, Cricket or Rugby.

The Venn diagram gives information about the numbers of these adults who enjoy watching these sports.

(a) How many of these adults enjoy watching all three sports?

....

(1)

One of these adults is chosen at random.

(b) Find the probability that this adult enjoys watching Cricket.

....

(1)

(c) Find the probability that this adult enjoys watching Rugby or Football.

....

(2)

One of the adults who enjoys watching Football is chosen at random.

(d) Find the probability that this adult also enjoys watching Cricket.

....

(2)

Here are five words that can be used to describe different types of data.

qualitative   quantitative   bivariate   discrete   continuous

(e) (i) Use one of the words in the list to describe the type of data needed for **names of sports**.

....

(1)

(ii) Use two of the words in the list to describe the type of data needed for **numbers of adults**.

.... and ....

(2)

(Total for Question 10 is 9 marks)

11 A vet carried out a survey to see if there is a relationship between the weight of a dog and its lifespan.
The vet found the mean weight, in kg, and the median lifespan, in years, of a sample of five dogs in each of ten breeds of dogs.

Here are her results.

Breed	A	B	C	D	E	F	G	H	I	J
Mean weight (kg)	5.5	11.5	20.5	5.0	10.0	50.0	83.0	18.0	7.0	43.0
Median lifespan (years)	15.5	14.8	14.6	14.6	13.5	7.5	6.8	6.3	14.0	6.8

(Source: adapted from *www.instituteofcaninebiology.org*)

The vet used statistical software to draw this scatter diagram for her results.

Mean weight (kg)
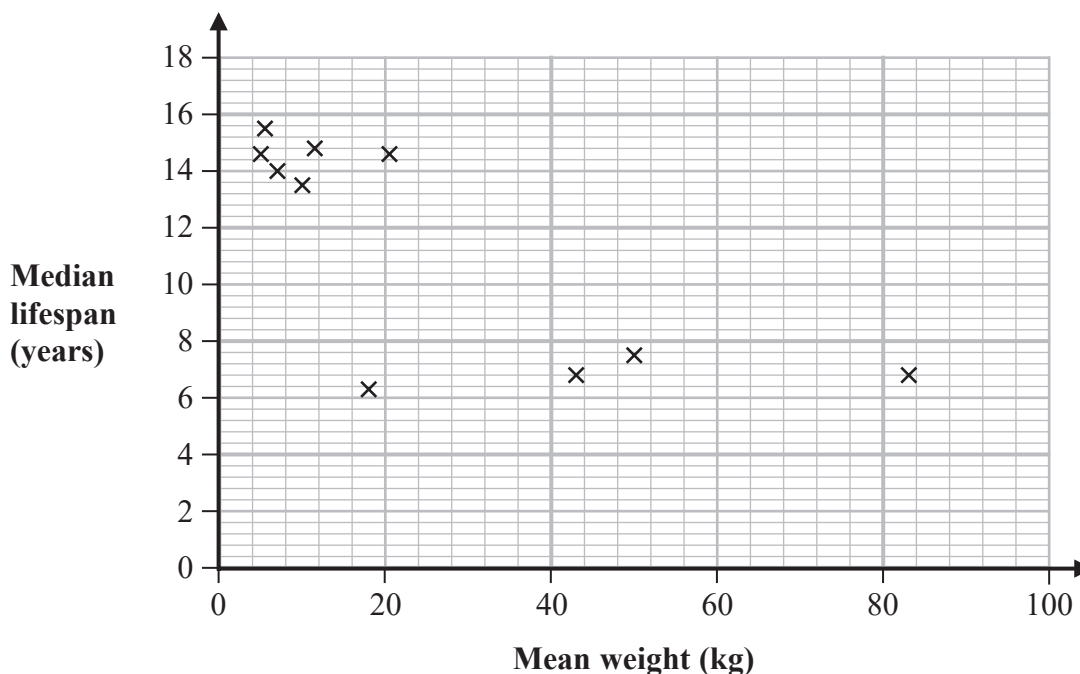
From the scatter diagram the vet concluded that there is a relationship between the weight of a dog and its lifespan.

Using the given statistical information,

(i) assess whether or not the vet's conclusion is appropriate,

....

....

(ii) assess the reliability of the vet's conclusion.

....

....

(Total for Question 11 is 2 marks)

12 Bill is investigating how being grown in the shade and being grown in sunlight affects the heights of tree seedlings.

The following stem and leaf diagrams give information about the heights, in centimetres, of 17 tree seedlings grown in the shade and 17 tree seedlings grown in sunlight.

The seedlings were all planted at the same time.

Shade

1	7 8 9
2	1 2 3 4 5 7 8 9
3	1 1 1 2 4 5
4	
5	

Key:

2 | 1 represents 21 cm

Sunlight

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

Key:

3 | 3 represents 33 cm

Compare the average height of the tree seedlings grown in the shade with the average height of the tree seedlings grown in sunlight.

State clearly the value of each average you use in order to make your comparison.

....

....

....

....

....

....

....

(Total for Question 12 is 3 marks)

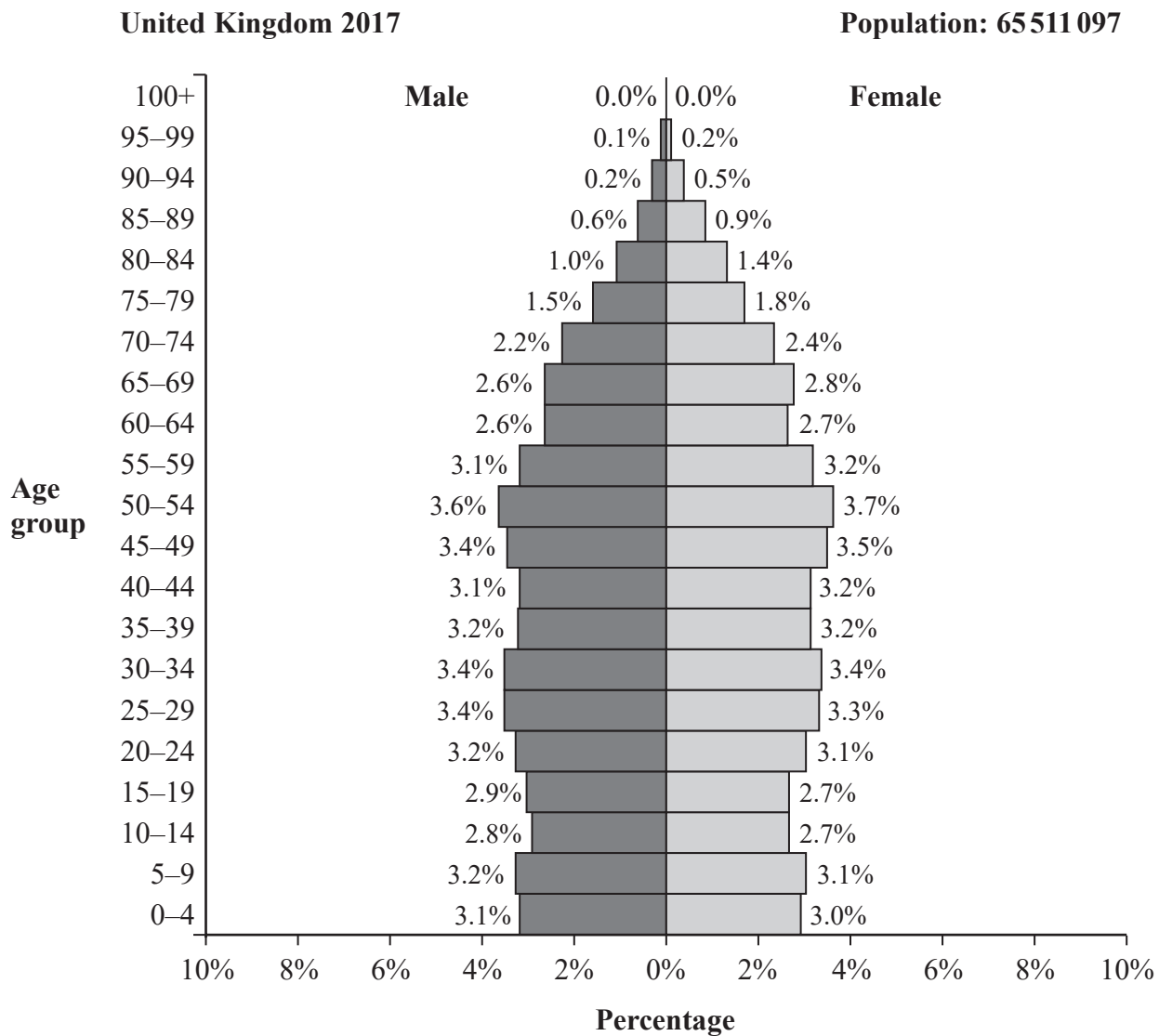**BLANK PAGE**

TURN OVER FOR NEXT QUESTION.

13 The population pyramid gives information about the percentages of the population of the United Kingdom who are male and who are female in each age group for 2017

Each percentage is given correct to one decimal place.

United Kingdom 2017   **Population: 65 511 097**

(Source: *Populationpyramid.net*)

(a) Write down the percentage of the population who are female in the age group 50–54

.... %

(1)

(b) Work out the percentage of the population who are male in the age group 10–19

.... %

(2)

In 2017, the number of people age 100 and older (100+) in the United Kingdom was 13 310

(Source: *Office for National Statistics*)

Using the information above and information from the population pyramid,

(c) explain why the percentage of the population in the age group 100+ is given as 0.0% on the population pyramid.
You must show your working.

....

....

(3)

Jamie is carrying out research into the ages of people in the United Kingdom.
He uses the information in the population pyramid to claim,

"In the United Kingdom in 2017 the number of males who were older than 40 was greater than the number of females who were older than 40"

(d) Explain whether or not Jamie's claim is correct.

....

....

....

....

....

(2)

(Total for Question 13 is 8 marks)

14 Weronika works for a road traffic organisation. One day she is investigating the speeds of cars and the speeds of motorcycles along a motorway.

Here is part of the spreadsheet that Weronika used to record her results.

Speed (s miles per hour)	Percentage of cars	Percentage of motorcycles
$30 \leqslant s < 40$	3	0
$40 \leqslant s < 50$	18	two
$50 \leqslant s < 60$	42	7
$60 \leqslant s < 70$	27	124
$70 \leqslant s < 80$	6	56
$80 \leqslant s < 90$		11
Total	**100**	**100**

(a) Give a reason why Weronika will need to clean the data.

....

(1)

Weronika concludes that the value of 124 in the spreadsheet must be wrong.

(b) Explain why.

....

(1)

Here is the information about motorcycles from the spreadsheet with the data cleaned.

Speed (s miles per hour)	Percentage of motorcycles
$30 \leqslant s < 40$	0
$40 \leqslant s < 50$	2
$50 \leqslant s < 60$	7
$60 \leqslant s < 70$	24
$70 \leqslant s < 80$	56
$80 \leqslant s < 90$	11
Total	**100**

(c) Use linear interpolation to work out an estimate of the median speed of the motorcycles.

.... miles per hour

(3)

The frequency polygon for the speeds of cars has been drawn on the grid.

(d) On the same grid, draw the frequency polygon for the speeds of the motorcycles.

(2)

Speed (miles per hour)
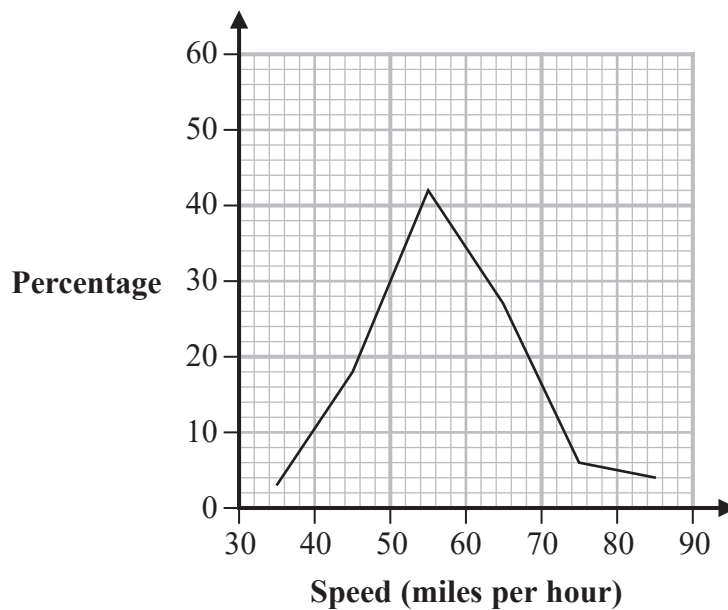
(e) Using the two frequency polygons, compare the skew of the distribution of the speeds of the cars with the skew of the distribution of the speeds of the motorcycles. Explain what your comparison means in context.

....

....

....

....

....

(2)

(Total for Question 14 is 9 marks)

15 The choropleth map gives information about the numbers of orangutans living in the forests of the Malaysian state of Sabah in northern Borneo. Some regions of the state are shaded and labelled with a number.

Key:

☐ 1–100 orangutans

▨ 101–500 orangutans

▧ 501–1500 orangutans

■ More than 1500 orangutans

(Source: *www.researchgate.net*)

(a) Write down the number of a shaded region where there are not more than 100 orangutans.

....

(1)

Regions 15 and 16 are protected forest regions.

Adam claims that in a protected forest region there is a greater number of orangutans.

(b) Does the choropleth map support Adam's claim?
    You must give a reason for your answer.

....

....

....

....

(2)

(Total for Question 15 is 3 marks)

TOTAL FOR PAPER IS 80 MARKS

BLANK PAGE

P 6 8 6 1 9 A 0 2 8 2 8