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

Monday 17 June 2024

Afternoon (Time: 1 hour 30 minutes)

Paper reference **1ST0/2F**

Statistics

PAPER 2
Foundation Tier

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 ▶

Pearson

1   David has 10 cards each with a single letter on it as shown.

| A | A | A | A | A | B | B | C | D | D |

A card is picked at random.

(a) Underline the word from the list below that best describes the likelihood that the card has a letter A on it.

impossible   certain   likely   evens   unlikely

(1)

(b) Complete this sentence using two different letters.

Cards with the letters .... and .... are equally likely to be picked.

(1)

(c) On the probability scale below, mark with a cross (×) the probability that the card has a letter C on it.

$$0 \qquad\qquad \frac{1}{2} \qquad\qquad 1$$
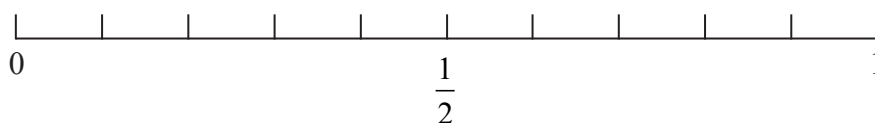
(1)

(d) On the probability scale below, mark with a cross (×) the probability that the card has a letter A or a letter C on it.

$$0 \qquad\qquad \frac{1}{2} \qquad\qquad 1$$

(1)
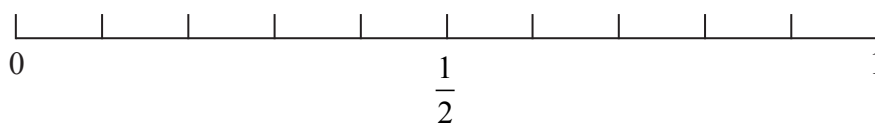
(Total for Question 1 is 4 marks)

2 Jenny is investigating how many days per week people use a gym.

She asks the 40 people in her fitness group how often they use the gym each week.
Jenny draws this bar chart for her data.
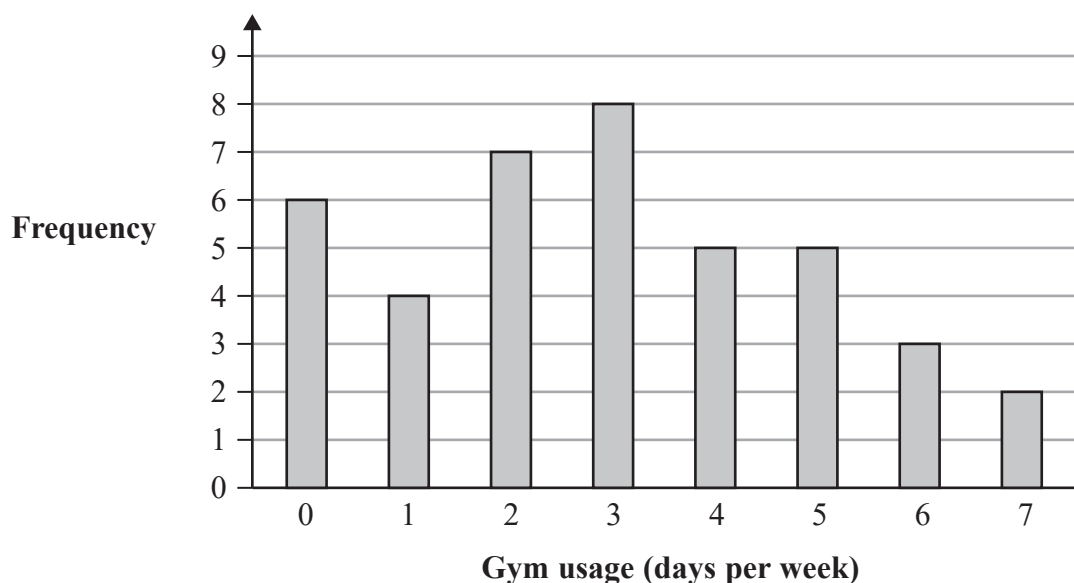
One of these people is chosen at random.

(a) Find the probability that this person uses the gym exactly 2 days per week.

....

(1)

(b) What is the modal number of days to use the gym each week?

....

(1)

Jenny thinks that there are a lot of people in her fitness group who are exercising less than 2 days per week as there is a total of 10 people who used the gym on 0 days or 1 day per week.

(c) Explain why Jenny might **not** be correct.

....

....

(1)

(Total for Question 2 is 3 marks)

3 Ben is researching information about the number of British swimming medals won at the Olympics.

Here are his results, giving the number of British swimming medals won at the Olympics from 1900 to 2016

3	0	7	6	2	4	4	2	0
1	1	2	3	1	1	1	3	5
5	3	1	2	0	2	3	3	6

(Source: *www.teamgb.com*)

(a) Fill in the tally chart for Ben's results **and** complete the frequency column.

Number of Olympic medals won	Tally	Frequency
0		
1		
2		
3		
4		
5		
6		
7		

(2)

(b) Suggest a suitable diagram that could be used for Ben's results.

....

(1)

(c) Write down the mode or modes.

....

(1)

(d) Work out the median.

....

(2)

Ben wants to use an average to summarise the data.

(e) Which of the mode or the median would be more appropriate?
Give a reason for your answer.

....

....

....

<div align="right">(2)</div>

<div align="right">**(Total for Question 3 is 8 marks)**</div>

4 Tachi collects data on the heights, in metres, of a sample of Egyptian pyramids.

Here is her data.

> 136.4   101.1   104   62.6   138.8   65.5   93.5

Tachi picks one of these pyramids at random.

(a) Find the probability that this pyramid will have a height of more than 100 m.

....
(1)

The mean height of a sample of Mexican pyramids is 53.5 m.

Tachi says, "On average these Egyptian pyramids are twice as high as the Mexican pyramids."

(b) Is Tachi correct?
You must show working to support your answer.

....

....
(4)

The range of heights for the Mexican pyramids is 45 m.
The lowest height of the Mexican pyramids is 30 m.

(c) Work out the greatest height of the Mexican pyramids.

.... m
(1)

(Total for Question 4 is 6 marks)

5 Claire is planning an investigation into the length of time that a learner has to wait for a driving test.
She wants to find out about how waiting time varies in different regions of the UK.

Here is her plan for data collection, for calculations and for diagrams.

> **Data collection**
> Visit a random sample of driving test centres in each region to ask for their waiting time in June.
>
> **Calculations**
> Calculate the average waiting time for each region for June.
> Calculate the range of the waiting times for each region for June.
>
> **Diagrams**
> Draw a bar chart showing the average waiting time for each region in June.
> Draw a pie chart showing the range of waiting times for each region in June.

Discuss whether Claire's plans for data collection, for calculations and for diagrams are appropriate.

....

....

....

....

....

....

....

....

....

....

....

....

....

....

(Total for Question 5 is 6 marks)

6 Ana collected information about the Scottish Football Championship results for the 2021/2022 season.

The composite bar chart gives some information about the number of points scored by nine of the teams.

Points are scored from wins or draws.
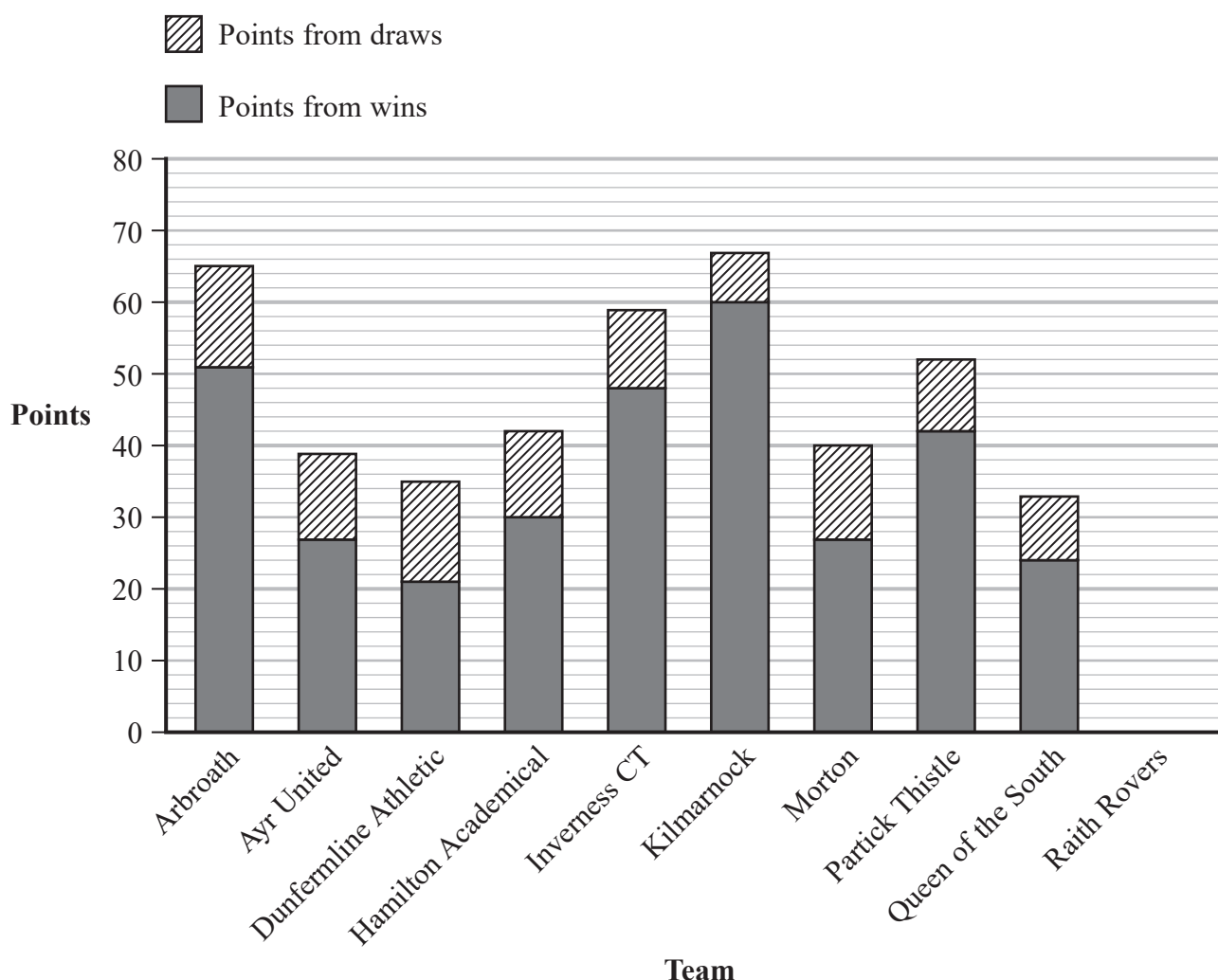
(Source: *www.skysports.com*)

Raith Rovers scored 36 points from wins and 14 points from draws.

(a) Complete the composite bar chart for Raith Rovers.

(2)

(b) What does the overall height of each bar represent?

....

....

(1)

P 7 5 4 3 8 A 0 8 2 4

(c) Compare the points scored by Dunfermline Athletic with the points scored by Queen of the South.

....

....

....

....

....

....

(3)

(Total for Question 6 is 6 marks)

7   Chris is a manager at a theme park.

He wants to find out what food options visitors would like to be able to buy in the theme park.

(a) State the population for this investigation.

....

(1)

Chris decides that he will take a convenience sample of visitors in the section of the park selling food.

(b) Explain what is meant by a convenience sample.

....

....

....

(1)

(c) Give one disadvantage of using a convenience sample.

....

....

(1)

Chris plans to use the data collection sheet below.

Type of food	Tally
Pizza	
Chinese	
Curry	
Fish and chips	

(d) Discuss whether this data collection sheet is appropriate.

You should consider how Chris might use the data and describe any problems he might have when he uses the data collection sheet.

....

....

....

....

....

....

(2)

Chris suggests using a stem and leaf diagram to represent the data that he collects.

(e) Discuss whether or not this would be a suitable diagram to represent his data.

....

....

....

....

(2)

(Total for Question 7 is 7 marks)

8 Timur is investigating the heights and weights of rugby players.
He collected data from the internet about the heights and weights of players from the Wales rugby squad.

The players from the Wales rugby squad are classed as Forwards or Backs.
Timur draws a scatter diagram for the heights and weights of some of the Backs from the Wales rugby squad.
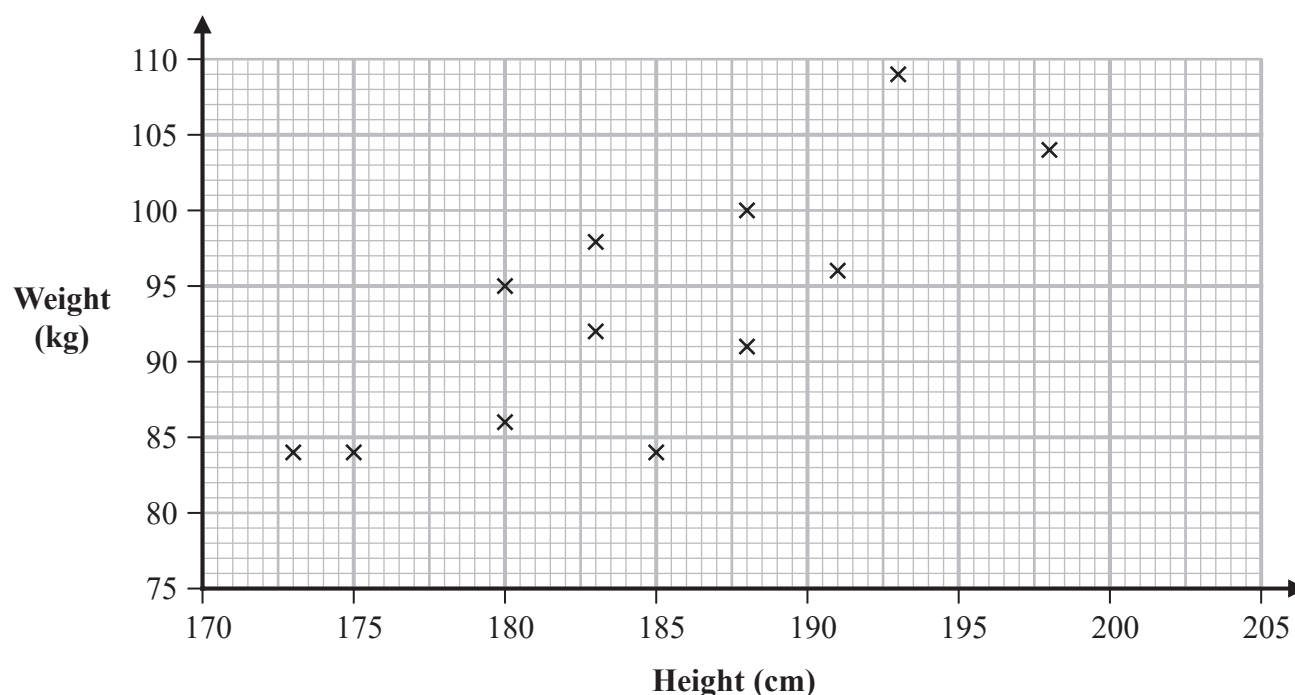
(Source: *www.wru.wales/*)

(a) Explain why a scatter diagram is appropriate for the type of data Timur collected.

....

(1)

The data for three of the players was not plotted on the scatter diagram.
The height and weight of each of these players is given in the table below.

Player	A	B	C
Height (cm)	183	183	173
Weight (kg)	95	87	78

(b) Complete the scatter diagram by plotting the points for players **A**, **B** and **C**.

(2)

(c) Describe and interpret the type of correlation shown by the scatter diagram.

....

....

(2)

The double mean point for the heights and weights of all of the Backs is (184, 92).

(d) Draw a line of best fit through the double mean point.

(2)

Timur also collected the heights and weights of players from the England rugby squad. He found the Spearman's rank correlation coefficient for the heights and weights of the Forwards from the England rugby squad.

The correlation coefficient was 0.00

(e) Interpret this correlation coefficient in context.

....

....

(1)

Timur used statistical software to find the following information about the heights (x cm) and weights (y kg) of the Backs from the Wales rugby squad and the Backs from the England rugby squad.

	Spearman's rank correlation coefficient	Gradient of line of best fit
Wales rugby squad Backs	0.81	0.96
England rugby squad Backs	0.65	1.02

(Source: *www.englandrugby.com*)

(f) Compare the Spearman's rank correlation coefficients and interpret this comparison in the context of the question.

....

....

....

(2)

Timur uses the information in the table to conclude that the weight of the England rugby squad Backs increases faster than the weight of the Wales rugby squad Backs as their height increases.

(g) Assess the validity of Timur's conclusion with reference to the statistical results.

....

....

....

....

(2)

(Total for Question 8 is 12 marks)

9 Mobeen is investigating whether there is a difference in the amount of time spent reading by pupils in Green Park school and pupils at Golden Plains school.

He uses a census of all of the pupils at each school.
Each pupil is asked to record the amount of time spent reading in a week.

Mobeen then collects this information from each student through an online database.

Part of the database is shown below.

	School	**Time spent reading**
1	Green Park	3 hours and 10 minutes
2	Golden	2.5 hours
3	GP	45
4	GREEN PARK	1h30
5	Golden Plains	$3\frac{1}{2}$ h
6	Green park	About 5 hours
7	Green park school	None
8	–	90
9	Golden plains	1.5h

(a) Give **two** reasons why the data should be cleaned before processing.

....

....

....

....

(2)

Mobeen wants to compare the data for Green Park school with the data for Golden Plains school.

Once the data has been cleaned Mobeen plans to use all of the times to draw a single box plot.

(b) Explain why this is **not** an appropriate thing to do.

....

....

....

(1)

(Total for Question 9 is 3 marks)

10 Matthew is investigating average household income for different states in the USA.

(a) Give a reason why it is appropriate to use secondary data for this.

....

....

(1)

Matthew creates a choropleth map giving information about the mean household income by state for the USA in 2023

Mean annual household income in $ thousands.

1. Vermont
2. New Hampshire
3. Massachusetts
4. Rhode Island
5. Connecticut
6. New Jersey
7. Delaware
8. Maryland
9. District of Columbia
10. West Virginia
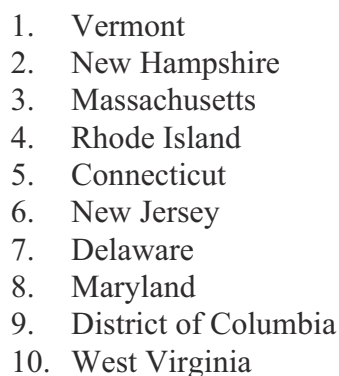
(Source: *worldpopulationreview.com*)

Key

☐ < 70

▦ $70\text{–}80$

▨ $80\text{–}90$

▩ $90\text{–}100$

■ $\geqslant 100$

(b) Which **three** states have the lowest mean household income?

....

(1)

Matthew concludes that the mean household incomes are highest on the West coast and the East coast.

(c) Does the choropleth map support this conclusion?
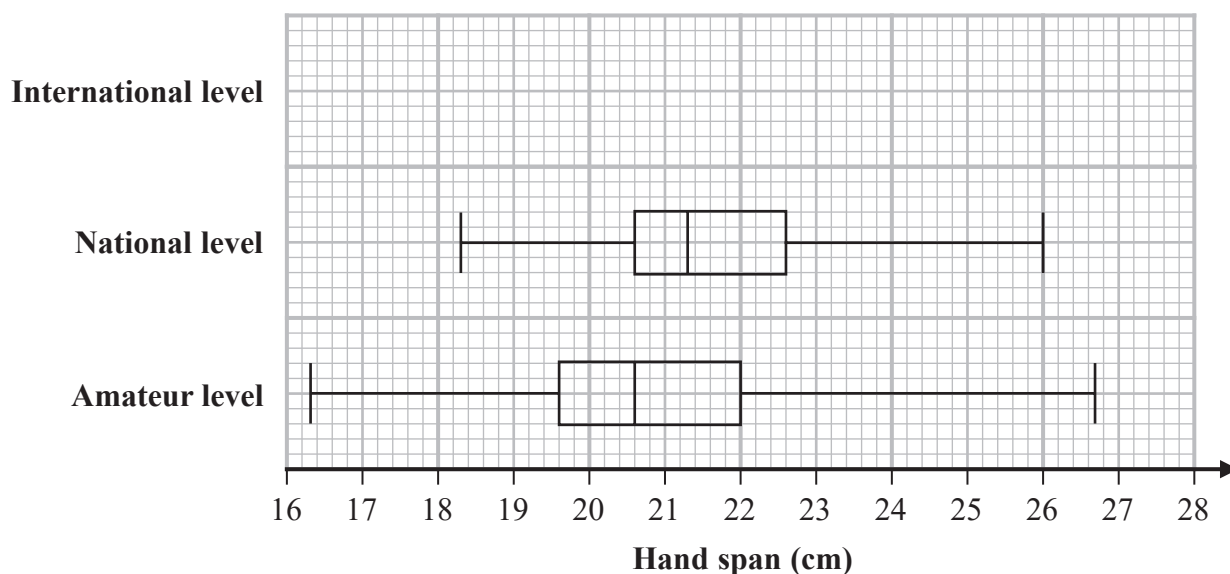   Give a reason for your answer.

....

....

....

(2)

(Total for Question 10 is 4 marks)

11 Some researchers investigated the hand span, in centimetres, of adult pianists by their level – international, national and amateur.

The box plots below give information about the hand spans for national level and amateur level pianists.

(Source: *www.appca.com.au*)

(a) Circle the word in the list below that describes hand span, in centimetres, as a type of data.

qualitative   ordinal   continuous   bivariate

(1)

The table gives information about the hand spans of the international level pianists.

Greatest hand span	27.4 cm
Median hand span	23.9 cm
Lower quartile	23.2 cm
Range	5.1 cm
Interquartile range	1.1 cm

(b) Using the information in the table, draw on the grid above a box plot for the hand spans of the international level pianists.

(3)

P75438A01824

(c) Compare the three distributions of hand spans.
Give **three** comparisons and interpret **two** of your comparisons.

....

....

....

....

....

....

....

....

(5)

Pavel owns a music shop.
He wants to investigate the keyboard sizes used by pianists with different hand spans.
He collects data about the hand spans of the pianists who use his shop.

The table gives information about the number of these pianists with hand spans in each
of four size categories.

Hand span (cm)	A (less than 19)	B ($19 \leqslant$ span < 22)	C ($22 \leqslant$ span < 24)	D (24 or more)
Number of pianists	24	65	57	14

Pavel plans to sample 20 of these pianists stratified by hand span size.

(d) Explain how Pavel can obtain his stratified sample.
You should include details of any calculations he should use.

....

....

....

....

....

....

....

(3)

(Total for Question 11 is 12 marks)

12 Khatia organises two different training courses, Course A and Course B, to help people to learn to type.
She wants to compare the two different courses to see which is better.
At the end of each course the people are given a skills test.

The table shows the number of participants who passed and failed the skills test for each of the two courses.

	Passed	**Failed**	**Total**
Course A	35	15	50
Course B	48	32	80

(a) Find the relative risk of failing the skills test having taken Course A compared to Course B.

....

(3)

(b) Give an interpretation of your answer to part (a).

....

....
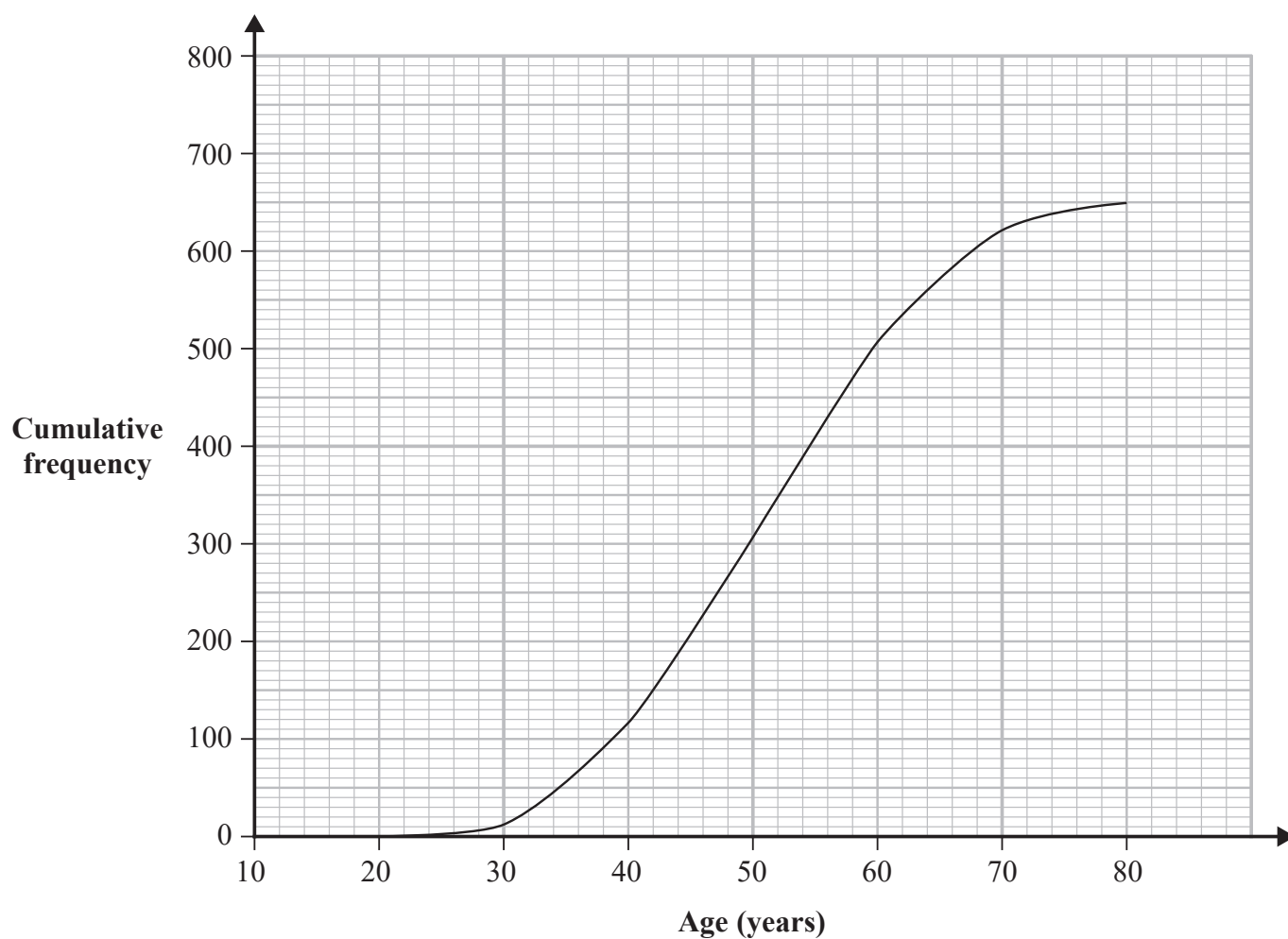
(1)

(Total for Question 12 is 4 marks)

BLANK PAGE

Turn over for next question

Turn over ▶

13 The cumulative frequency graph gives information about the ages, in years, of the 650 members of the UK Parliament in 2017

(Source: *commonslibrary.parliament.uk*)

(a) Using the cumulative frequency diagram, find an estimate of the median age for the members of the UK Parliament in 2017

.... years

(1)

P 7 5 4 3 8 A 0 2 2 2 4

In 2017 the German Parliament had 51.6% of members of Parliament aged between 45 and 60 years old.

(Source: *https://www.dw.com/en/germanys-new-bundestag-who-is-who-in-parliament/a-41082379*)

(b) Compare this figure to the percentage of members of the UK Parliament with ages between 45 and 60 years in 2017

....

....

(4)

(Total for Question 13 is 5 marks)

TOTAL FOR PAPER IS 80 MARKS

BLANK PAGE