Qualification Accredited

OCR
Oxford Cambridge and RSA

Examiners' report

LEVEL 3 CERTIFICATE

CORE MATHS B (MEI)

OCR Level 3 Certificate in Core Maths B (MEI)

H869

For first teaching in 2016

H869/02 Summer 2025 series

Version 1

ocr.org.uk/maths

Contents

Introduction

Our examiners' reports are produced to offer constructive feedback on candidates' performance in the examinations. They provide useful guidance for future candidates.

The reports will include a general commentary on candidates' performance, identify technical aspects examined in the questions and highlight good performance and where performance could be improved. A selection of candidate responses is also provided. The reports will also explain aspects which caused difficulty and why the difficulties arose, whether through a lack of knowledge, poor examination technique, or any other identifiable and explainable reason.

Where overall performance on a question/question part was considered good, with no particular areas to highlight, these questions have not been included in the report.

A full copy of the question paper and the mark scheme can be downloaded from Teach Cambridge.

Paper 2 series overview

This paper aims to present candidates with a wide range of real-life situations in which they can demonstrate statistical problem-solving skills. Candidates engaged well with every question on this paper. Calculations were mostly presented clearly, and a good range of tables, diagrams and situations were analysed and interpreted correctly. Some candidates did not have a secure understanding of using graphs to determine approximate rates of change in context. The most successful responses saw candidates respond confidently to a wide range of scenarios and statements.

Candidates who did well on this paper generally:	Candidates who did less well on this paper generally:
• structured their calculations clearly • applied statistical thinking and modelling skills in a variety of real-life situations • worked accurately and addressed all points in each question • gave clear explanations supported by evidence when asked to critique or justify statements and hypotheses • worked confidently in response to questions based on the pre-release data set.	• only worked confidently in familiar situations • did not present their calculations clearly • did not seem to consider whether their answers were sensible in context • did not use all relevant information when asked to critique or justify statements and hypotheses • did not address all points when responding to questions.

Section A overview

Candidates generally made a strong start to Section A, using relevant information routinely in various contexts. While candidates seemed reasonably familiar with using the Normal distribution for modelling, some were not able to present appropriate calculations for probabilities for the scenario given. Most candidates were confident with the structure of the chi-squared test for independence, but a common error in carrying out the test was the calculation of the degrees of freedom.

Question 1 (a)

1 Emma is investigating the reactions of people of different ages to climate change issues.

Fig. 1.1 shows a questionnaire she designs.

She gives the questionnaire to 90 people. There are 30 people from each of these three age groups: 16 to 24, 25 to 64, 65 and over.

Fig. 1.1

Questionnaire

To what extent do you agree with these statements about climate change?

Circle your answer to each question.

0 = Disagree **1 = Neither agree nor disagree** **2 = Agree**

A All electricity should be generated from renewable sources.
0 1 2

B Diesel and petrol cars should be banned as soon as possible.
0 1 2

C The numbers of all flights should be cut down by at least half.
0 1 2

D Farmers should produce less meat.
0 1 2

Circle your age group: 16 to 24 25 to 64 65 and over

Thank you for participating.

Most people return the questionnaire to Emma, but some do **not**.

In all the returned questionnaires, all four questions have been answered.

For each returned questionnaire, Emma adds the person's answers for questions A to D together to find a score between 0 and 8 for each person.

Fig. 1.2 shows the number of people with each score, split by their age group.

(a) Complete **Fig. 1.2** in the answer space by filling in the empty cells. **[1]**

1(a) Fig. 1.2

Age \ Score	0	1	2	3	4	5	6	7	8	Total
16 to 24	0	0	1	1	2	1	4	9	12	**30**
25 to 64	1	2	3	2	4	4	3	3	3	
65 and over	4	3	3	2	1	3	1	0	3	
Total	**5**	**5**	**7**	**5**	**7**	**8**	**8**			

This was nearly always answered correctly. Occasionally the total was given incorrectly as 150.

Question 1 (b)

(b) How many people did **not** return the questionnaire? **[1]**

This was nearly always answered correctly.

Question 1 (c)

(c) According to the data in **Fig. 1.2**, which age group seems the most concerned about climate change?

How do you know? **[1]**

Nearly all candidates chose the correct age group, and there were many correct justifications. The most successful responses referred clearly to the higher frequency of scores 7 and 8 for this age group.

Question 1 (d) (i)

(d) Emma is concerned about the different numbers of responses from each age group. She describes the questionnaires that were **not** returned as "missing data".

She considers two possible options for dealing with the missing data.

Option 1: Ignore the missing data and work only with the scores from the responses she has.

Option 2: Give all the missing responses a score of zero.

(i) Show that, using **Option 1**, the mean score for the whole sample is 5.04. **[3]**

This question required candidates to show a given answer. The most successful responses showed clear calculations to obtain the sum of the scores multiplied by the frequencies (378) and then divided this by the total number of returned questionnaires (75).

Question 1 (d) (ii)

(ii) Work out the mean score for the whole sample using **Option 2**. **[1]**

Most candidates realised that the process here was the same as in Question 1 (d) (i) apart from the different divisor (90).

Question 1 (d) (iii)

(iii) State which of the two options you think is more appropriate.

Give a reason for your answer. **[1]**

Many good justifications for choosing Option 1 were seen.

Question 2 (a)

2 A forest containing ash trees was planted around the year 1900. A hundred years later, 240 of the ash trees had matured and grown to full height. They supported a great quantity and variety of wildlife.

The heights of these 240 mature ash trees can be modelled by a Normal distribution with mean 32 metres and standard deviation 2.5 m.

(a) Show that over half of these ash trees were over 100 feet tall.

Use the conversion 1 foot is 0.3048 metres. **[2]**

Nearly all candidates scored at least 1 mark, by correctly converting 32 metres into feet, or 100 feet into metres using the given conversion. Many candidates did not link their converted answer to the question by writing a short conclusion in context.

Question 2 (b)

(b) Calculate an estimate of the number of these ash trees over 35 metres tall. **[3]**

An encouraging number of fully correct solutions were seen for the application of the Normal distribution to this context. The most successful responses either calculated a z-score to find the probability associated with the height of a tree over 35 metres tall or correctly stated and used a probability from a calculator menu. Occasionally a much rougher estimate from approximate areas under the Normal curve was used. This gained some credit, but did not give a very accurate approximation as 35 metres was not exactly one standard deviation from the mean as implied by those using this method.

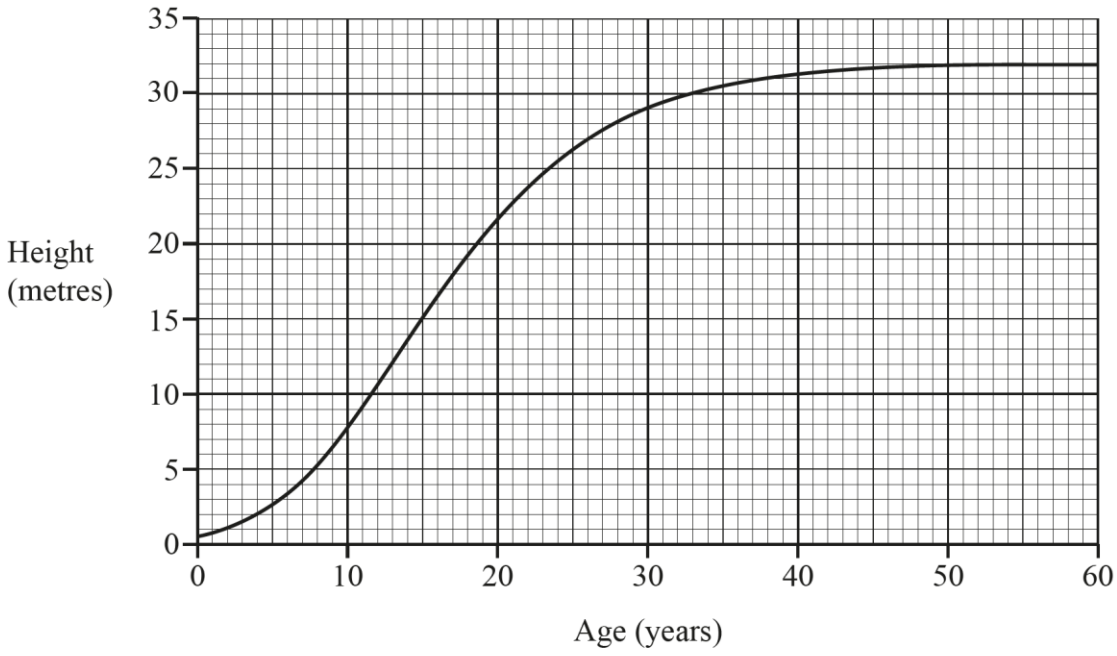
Question 2 (c)

(c) **Fig. 2.1** shows the height of a typical ash tree against its age.

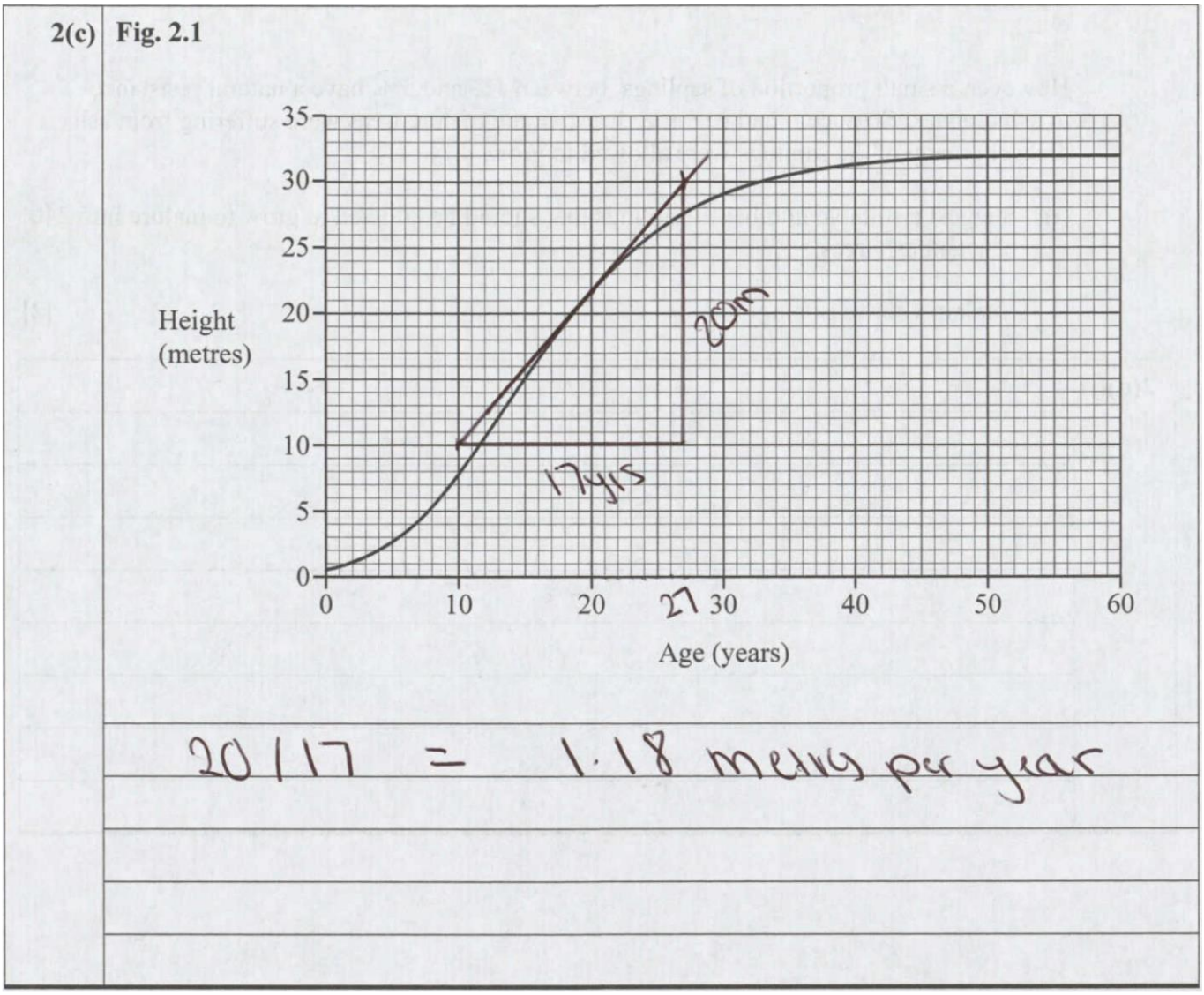
Use **Fig. 2.1** to estimate the rate of growth of a 20-year-old ash tree.

You will need to draw a suitable line on the graph. **[3]**

2(c) **Fig. 2.1**



Exemplar 1



Responses to Question 2 (c) were often incorrect. Where a rate of change is required at a particular point, in this case *at 20 years of age*, candidates should be encouraged to mark that point on the graph, before drawing an appropriate tangent to the curve. In the exemplar shown here, the tangent is drawn, and a triangle has been formed with height and base measurements identified clearly. The candidate has then used these values appropriately to calculate gradient and has given units which relate to the context.

Less successful responses fell into two categories. Either a tangent was not attempted, or the attempt at gradient calculation was upside down.

Misconception

Some candidates incorrectly thought that connecting the point at 20 years to the start of the graph at time zero would represent the growth rate at 20 years. This is not a correct method and would give the *average* rate of growth *over* the first 20 years, rather than the *growth rate at 20 years*.

OCR support

Section check in tests feature questions on a single section of the specification and can be used as formative assessments.

For support with working with graphs and gradients, download the Graphs section check in test from Teach Cambridge.

Question 2 (d) (i)

(d) In 2012, some of the mature ash trees in the forest showed signs of ash dieback, a disease that occurs in ash trees. After a number of years an affected tree will die and fall down. There is no known cure.

Every year a large number of new ash trees (saplings) start to grow in the forest. Until the arrival of ash dieback in 2012, all these saplings had been treated as weeds and were taken out.

However, a small proportion of saplings, between 1% and 5%, have a natural resistance to ash dieback. So once it had become clear that the mature trees were suffering from ash dieback, many of the saplings were allowed to grow.

(i) Suggest a suitable number of saplings that should be allowed to grow to mature into 240 resistant ash trees.

Explain your reasoning. **[2]**

Many candidates made a good estimate, often using 1% or 5%, and sometimes a value in between these two proportions. A short statement was also required to explain their reasoning for the number of saplings that should be allowed to grow.

Question 2 (d) (ii)

(ii) Comment briefly on changes to the forest over time, from 2012 onwards.

Your answer should include what is likely to happen in the future. **[1]**

Most candidates gave a very good answer, that drew on the context of allowing more saplings to grow to replenish those trees dying, leading to recovery of the forest in the long term.

Question 3 (a)

3 Sam drives to work each **weekday**, Monday to Friday.

She is often delayed by having to stop at a level crossing over a railway line.

Sam wants to know whether the delays are related to the time she leaves home.

She tries three different start times which she records as Early, Normal and Late.

She records the delay each day as None, Short, Medium or Long.

Fig. 3.1 shows the number of days in each category.

(a) Fill in the empty cells in **Fig. 3.1**. **[1]**

3(a) **Fig. 3.1**

Observed frequencies

START	DELAY: None	Short	Medium	Long	Total
Early	10	10	9	6	
Normal	7	11	16	11	**45**
Late	13	15	8	4	
Total	**30**				

All values in the table were nearly always correct.

Question 3 (b)

(b) How many weeks did Sam collect the data for? **[1]**

Most candidates realised that the total number of days should be divided by the number of days in the working week. A few missed this detail from the opening stem of the question and divided by seven.

Question 3 (c) (i)

(c) Sam looks at the data in **Fig. 3.1**. She thinks the numbers suggest that changing her start time would reduce her travel time.

Sam wants to check how strong the evidence is before asking if she can arrive at work at a different time.

She carries out a chi squared test at the 10% significance level.

(i) Which of these is a suitable null hypothesis, H_0, for the test?

Tick (✓) **one** box. **[1]**

3(c)(i)

The later the start time the greater the delay.	☐
The length of the delay depends on the start time.	☐
The length of the delay is independent of the start time.	☐
Trains come in groups.	☐

The correct null hypothesis was chosen by most candidates.

Question 3 (c) (ii)

(ii) **Fig. 3.2** shows the Expected frequencies.

Fill in the empty cells in **Fig. 3.2**. **[2]**

Nearly all candidates completed the cells correctly. Mostly this was done very simply by using the earlier answers for the totals in each row and column. Occasionally a correct probability process was used.

Question 3 (c) (iii)

(iii) Fill in the **six** empty boxes in the calculation of the value of X^2 in **Fig. 3.3** on the next page. **[2]**

3(c)(iii) **Fig. 3.3**

$$X^2 = \frac{1.25^2}{8.75} + \frac{(-0.5)^2}{10.5} + \frac{(-0.625)^2}{9.625} + \frac{(-0.125)^2}{6.125}$$
$$+ \frac{(-4.25)^2}{11.25} + \frac{(-2.5)^2}{13.5} + \frac{3.625^2}{12.375} + \frac{3.125^2}{7.875}$$
$$+ \frac{3^2}{10} + \frac{3^2}{12} + \square + \square$$
$$= 0.17857... + 0.02380... + 0.04058... + 0.00255...$$
$$+ 1.60555... + 0.46296... + 1.06186... + 1.24007...$$
$$+ \square + \square + \square + \square$$
$$= 8.37 \text{ (2 decimal places)}$$

This element of the paper was answered correctly by nearly all candidates. Occasionally candidates made a processing error from an incorrect statement of the contributions, e.g. $\frac{-3^2}{11}$ or $\frac{(-3^2)}{11}$. These errors are quite fundamental and were not accepted. Correct responses were $\frac{(-3)^2}{11}$ or just $\frac{3^2}{11}$.

Question 3 (c) (iv)

(iv) Carry out the chi squared test **and** interpret the result in context. **[4]**

A recurring issue with responses to the chi squared test is that candidates often make a mistake in calculating the degrees of freedom for the test. The degrees of freedom for a contingency table should be calculated as $(r - 1) \times (c - 1)$, where r is the number of rows and c is the number of columns.

Degrees of freedom for a contingency table

Fig 3.1 has been adapted below with a border and shading to emphasise that for this contingency table the data is in *3 rows* and *4 columns*, hence DOF = (3 - 1) × (4 - 1) = 2 × 3 = 6. Six degrees of freedom represents the number of cells that would need to be filled, before all cells are effectively known. The use of shading or borders in this way when working with candidates may reinforce this important concept.

	None	Short	Medium	Long	Total
Early	10	10	9	6	**35**
Normal	7	11	16	11	**45**
Late	13	15	8	4	**40**
Total	**30**	**36**	**33**	**21**	**120**

Section B overview

All candidates were able to find and use relevant information from the spreadsheet insert to use in their calculations. Strength of understanding of population growth and how this could be related to graphs and charts was very mixed. Most candidates showed familiarity with the hypothesis test, but many did not appreciate the implication of a negative value for Spearman's correlation coefficient when using table values to make comparisons and draw conclusions. There were many good explanations in context for each question, but some responses lacked detail or overlooked some information given in the question.

Question 4 (a)

4 (a) Write down the **median age** and the **annual population growth rate** for Italy. **[1]**

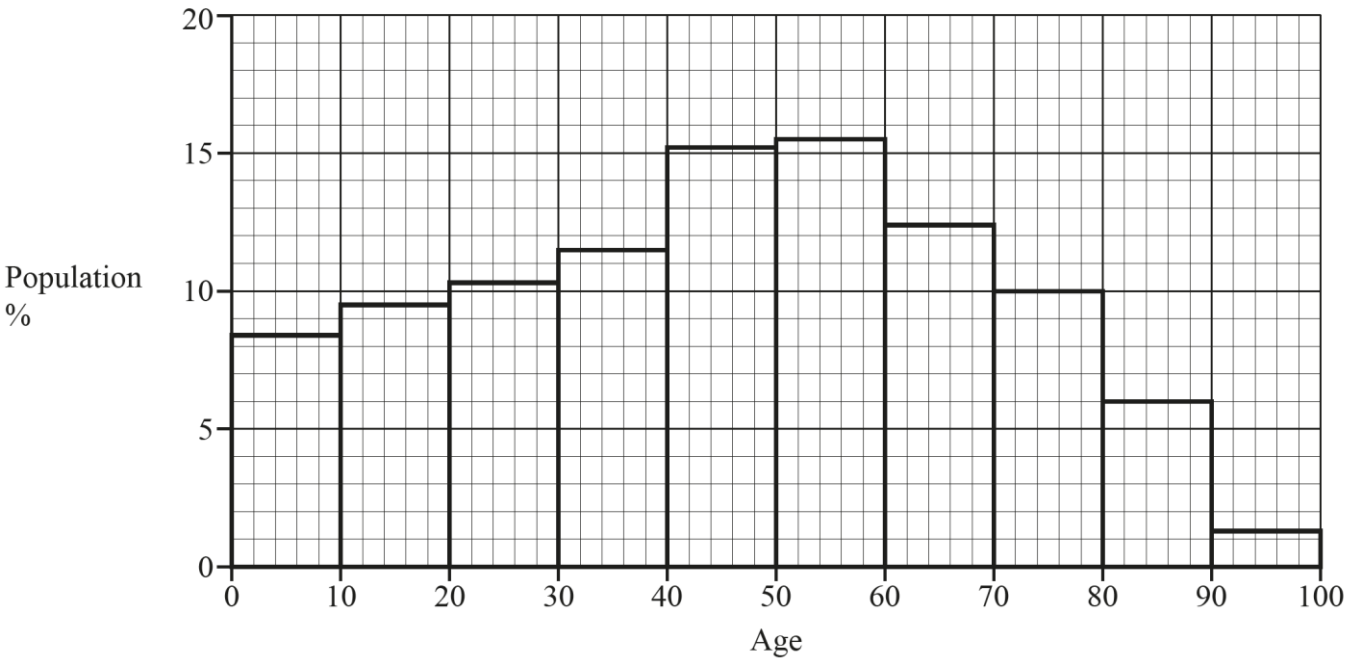
This was nearly always answered correctly.

Question 4 (b)

Fig. 4.1 shows a chart of the population profile for Italy.

(b) Draw a vertical line on **Fig. 4.1** to show the median age. **[1]**

4(b) Fig. 4.1



Nearly all candidates correctly drew a line onto the diagram to indicate the median (46.5) somewhere between 46 and 48 on the horizontal scale as required.

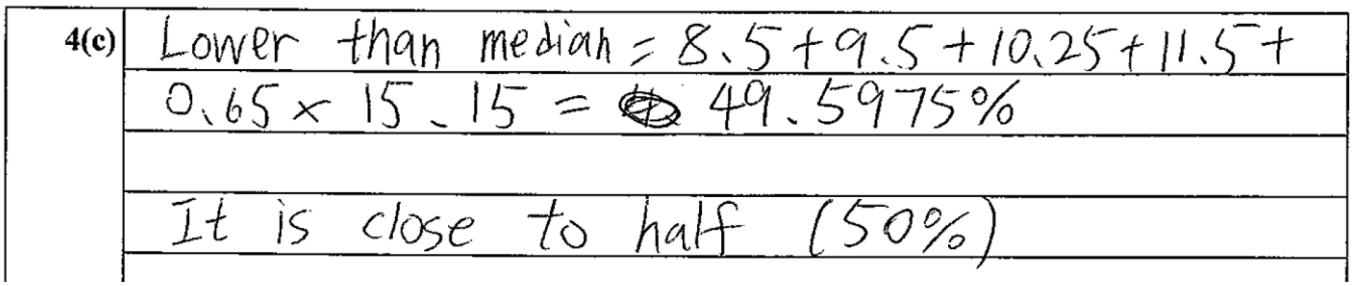
Question 4 (c)

(c) Show that the position of the median line is consistent with the values on the chart.

Use calculations in your answer. **[2]**

Very few candidates were able to demonstrate that the median position could be verified using values from the chart. An important point to reinforce with candidates is that 50% of the dataset should lie either side of the median.

Exemplar 2



The exemplar here shows a good process for summing the values for the first four bars from the population profile for Italy, and then adding to this a proportion of the bar in which the median lies.

Question 4 (d)

(d) State **one** feature of **Fig. 4.1** that is consistent with Italy's annual population growth rate. **[1]**

Successful responses to this question were seen rarely. The best justifications usually referred very clearly to the heights of the bars declining in the younger cohorts/age groups, for example, 'there are fewer people aged 0-10 years than 10-20 years'. Other successful answers justified the negative growth rate in population with simple correct statements such as 'there are fewer young people than middle-aged adults'.

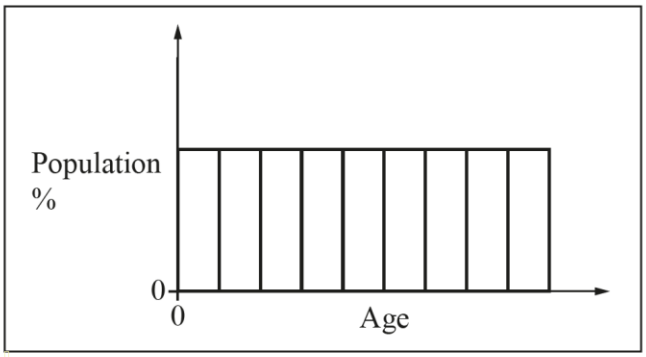
Question 4 (e)

(e) For Pakistan, the median age is 22 and the annual population growth rate is 1.91%.

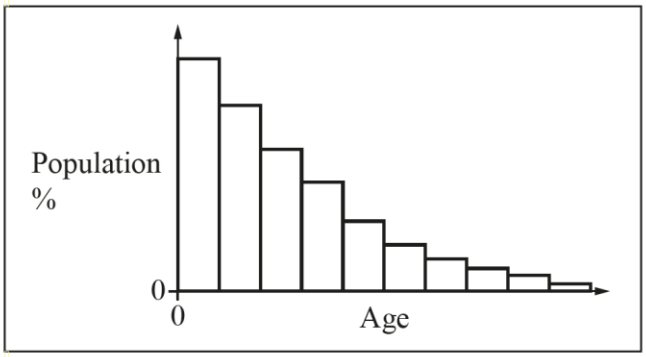
Which of these charts could show the complete population profile of Pakistan?

Tick (✓) **one** box. **[1]**

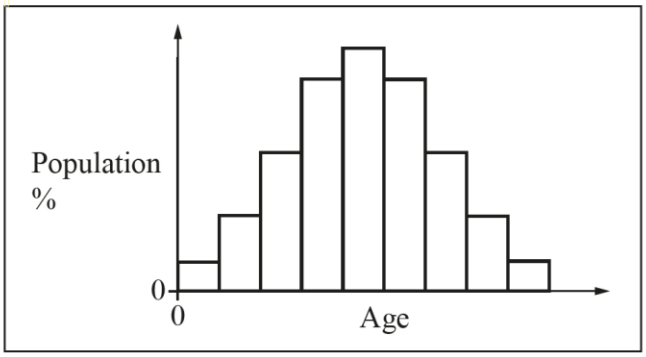
4(e)



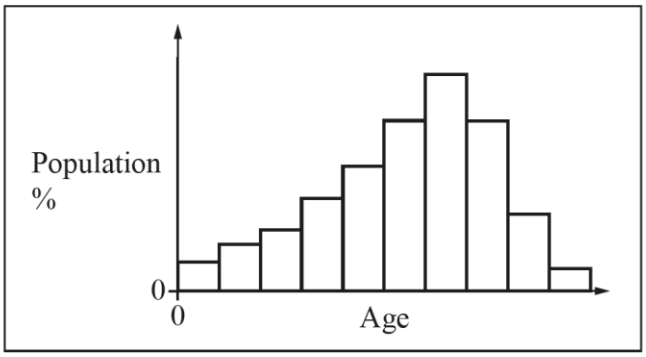
☐



☐



☐



☐

It was clear that conceptually very few candidates could confidently link the increasing population to the correct chart. A common wrong answer was the fourth chart, which like Italy shows a declining population in recent years.

Question 5 (a) (i)

5 (a) Libya is in North Africa.

(i) Explain how you know that the forest percentage in Libya is one of the lowest in Africa.

Suggest why this is the case. **[2]**

Most candidates scored one or other of the marks here, by correctly justifying the given statement or by describing the climate of Libya. Sometimes candidates did not refer to forestation figures from the spreadsheet in the exam insert. Other candidates did not realise that there were two points to be addressed in their answer.

Question 5 (a) (ii)

(ii) State the proportion of Libya's population that live in towns and cities. **[1]**

This was nearly always answered correctly.

Question 5 (b) (i)

Jane makes this statement:

'Countries with a low forest percentage are likely to have a high urban population percentage'.

She selects 10 countries for a pilot investigation of her statement.

Fig. 5.1 shows the countries she selects.

(b) Jane carries out a hypothesis test using Spearman's rank correlation coefficient.

The null and alternative hypotheses are:

H_0: There is no association between forest percentage and urban population percentage.

H_1: There is a negative association between forest percentage and urban population percentage.

(i) Is the test 1-tail or 2-tail? **[1]**

Most candidates worked out correctly that this was a 1-tailed test from Jane's description.

Question 5 (b) (ii)

(ii) Fill in the empty cells in **Fig. 5.1**. **[2]**

5(b)(ii) **Fig. 5.1**

Country	Urban population %	Urban rank, x	Forest %	Forest rank, y	$d = x - y$	d^2
Tunisia						
Botswana	72.9	6	19.8	8	–2	
Belize	46.6	9	60.6	1	8	
Greenland	87.9	3	0.0	10	–7	
Chile	88.0	2	21.9	7	–5	
Indonesia	58.6	8	51.7	3	5	
Russia	75.3	5	49.4	4	1	
Cuba	77.5	4	27.3	5	–1	
Belgium	98.2	1	22.4	6	–5	
Samoa	17.5	10	60.4	2	8	
				Total, $\sum$	**0**	**262**

The cells in the top row were nearly always filled in correctly, using values looked up from the exam insert. The differences in the rankings were nearly always squared correctly to obtain positive values for d^2.

Question 5 (b) (iii)

(iii) Calculate the value of r_s and show that the test is significant at the 5% level. **[4]**

Most candidates seemed familiar with the hypothesis test process, but many did not properly account for the *negative value* of the correlation coefficient when using critical values from the tables to make a valid comparison.

Question 5 (c)

(c) Suggest how Jane selected these countries. **[1]**

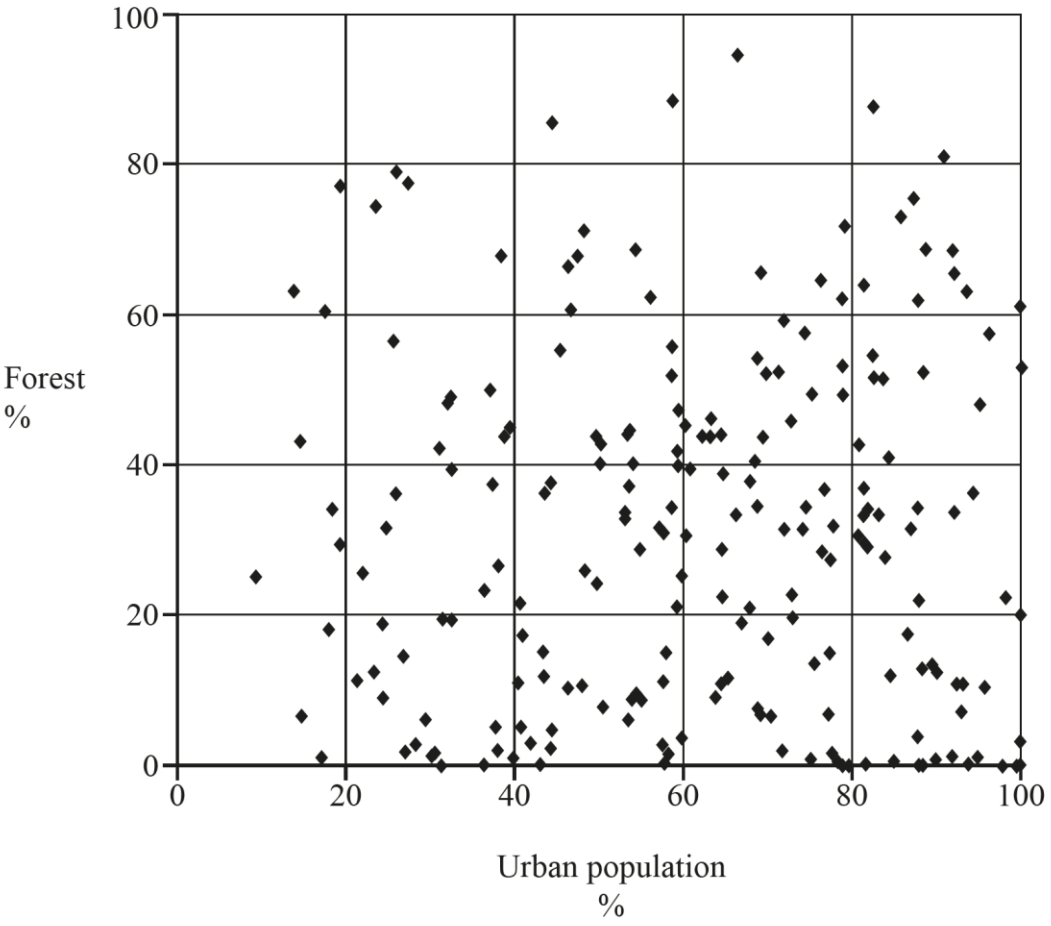
Most candidates demonstrated familiarity with the exam insert by correctly stating that Jane had most likely chosen one from each region.

Question 5 (d) (i)

(d) Jane cleans the data from the whole pre-release spreadsheet appropriately.

She uses it to draw the scatter diagram of forest % against urban population % in **Fig. 5.2**.

Fig. 5.2



(i) Decide whether these statements about the data in **Fig. 5.2** are **true** or **false**.

Tick (✓) **one** box in each row. **[1]**

5(d)(i)

	True	False
The higher the urban population percentage, the lower the forest percentage.		
There are some countries where almost everyone lives in towns or cities.		

It was the first statement in Question 5 (d) (i) that drew the most incorrect answers. If this statement had been true, then there would have been a visible trend (negative correlation) in Fig 5.2.

Question 5 (d) (ii)

(ii) Jane then uses the cleaned pre-release data to work out the Pearson's product moment correlation coefficient for the data in columns D and F.

The answer is -0.0087 ... .

Compare the Pearson's product moment correlation coefficient to the Spearman's rank correlation coefficient that you found in part **(b)(iii)**.

You should state which you consider to be the more reliable answer and suggest a reason for the difference between the two. **[2]**

Most candidates explained correctly that Pearson's product moment correlation would give the more reliable answer in this context, because of the much larger sample size. Candidates' comparisons of the two coefficients were often flawed as their explanations did not correctly refer to the relative size of the two coefficients, given that in this example both coefficients were negative. Some very good answers were seen, however, that clearly referred to the 'strength' of correlation indicated by both coefficients and that the very weak correlation indicated by Pearson's coefficient seemed to be supported by the scatterplot in Fig 5.2.

Question 6 (a) (i)

6 (a) (i) Write down the **population** and **CO_2 emissions** of India. **[1]**

Answers to this were nearly always correct.

Question 6 (a) (ii)

(ii) Calculate the average CO_2 emissions per person in India. **[1]**

Mostly this was calculated correctly, with the correct division seen to support candidates' answers. Occasionally the division was the wrong way up.

Question 6 (b)

(b) Show that the average CO_2 emissions per person is nearly 10 times higher in the United States than it is in India. **[2]**

Most candidates scored full marks here by firstly calculating the CO_2 emissions per person for the United States and then testing the statement by creating a ratio of the two figures using appropriate division.

Question 6 (c) (i)

Sasha uses the spreadsheet of pre-release data to investigate the impact that different countries are having on the environment.

He starts new columns P, Q and R and then:

- copies column C, Population, into column P
- copies column H, CO_2 emissions, into column Q
- enters " = Q2/P2" into cell R2
- copies cell R2 down to R236
- formats column R to give the numbers in the cells to **2** decimal places.

(c) (i) Show that the number in cell R2 is 3.39 and state what it represents. **[2]**

Nearly all candidates correctly justified the 3.39 with an appropriate calculation, but many forgot to describe what the figure represents.

Question 6 (c) (ii)

(ii) The number in cell R209 is 5.84.

What does this tell you? **[1]**

Nearly all candidates correctly described what the value represented.

Question 6 (d)

Before he can continue, Sasha needs to clean the data.

In some rows, column Q contains "#N/A" or the number zero. In those rows he deletes the content of cells Q **and** P so that they are both blank.

(d) State **one** country for which the cells in columns P and Q are now both blank. **[1]**

There were many countries to choose from the exam insert, and nearly all candidates made a correct selection.

Question 6 (e)

After Sasha's data cleaning, the total of the entries in column Q is 35 532 499 700.

The total for column P is 7 978 557 003.

(e) Determine whether people in the United Kingdom have above or below average CO_2 emissions when compared with the rest of the world.

Show your reasoning. **[2]**

Most candidates made a strong finish to the paper with an appropriate calculation to support their reasoning.

Supporting you

Teach Cambridge

Make sure you visit our secure website **Teach Cambridge** to find the full range of resources and support for the subjects you teach. This includes secure materials such as set assignments and exemplars, online and on-demand training.

Don't have access? If your school or college teaches any OCR qualifications, please contact your exams officer. You can forward them this link to help get you started.

Reviews of marking

If any of your students' results are not as expected, you may wish to consider one of our post-results services. For full information about the options available visit the OCR website.

Keep up-to-date

We send a monthly bulletin to tell you about important updates. You can also sign up for your subject specific updates. If you haven't already, sign up here.

OCR Professional Development

Attend one of our popular CPD courses to hear directly from a senior assessor or drop in to a Q&A session. Most of our courses are delivered live via an online platform, so you can attend from any location.

Please find details for all our courses for your subject on **Teach Cambridge**. You'll also find links to our online courses on NEA marking and support.

Signed up for ExamBuilder?

ExamBuilder is the exam paper builder platform for a range of our qualifications. Find out more.

Free for all OCR centres with an Interchange account, it allows unlimited users per centre. We require an Interchange username to verify your centre's users for ExamBuilder.

If you do not have an Interchange account, please contact your centre administrator (usually the Exams Officer) to request a username.

Once you have an Interchange username, use the form on this page to sign up for ExamBuilder, or ask an existing user at your centre to create an ExamBuilder account for you.

Tell us what you think

Your feedback plays an important role in how we develop, market, support and resource qualifications now and into the future. We want you and your students to enjoy and get the best out of our qualifications and resources, but to do that we need your honest opinions to tell us whether we're on the right track or not.

You can email your thoughts to resources.feedback@ocr.org.uk or visit our feedback page to learn more about how you can help us improve our resources.

Designing and testing in collaboration with you and your students

Helping young people develop an ethical view of the world

Equality, diversity, inclusion and belonging (EDIB) are part of everything we do

---

Contact the team at:

01223 553998

ocr.org.uk

To stay up to date with all the relevant news about our qualifications, register for email updates at ocr.org.uk/updates

Visit our Online Support Centre at support.ocr.org.uk

CAMBRIDGE
UNIVERSITY PRESS & ASSESSMENT