Qualification Accredited

OCR
Oxford Cambridge and RSA

A LEVEL

Examiners' report

MATHEMATICS B (MEI)

H640

For first teaching in 2017

H640/02 Summer 2023 series

Version 1

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 answers 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 OCR.

Paper 2 series overview

H640/02 is the second of three compulsory components in the A Level assessment. It contributes 36.4% of the total A Level and assesses content from pure mathematics and statistics.

Candidates are expected to have studied statistics using the Large Data Set and to have routinely used spreadsheets, graphing and statistical software when studying this course. They are expected to be able to extract information from software output such as a table of statistics generated when conducting a hypothesis test.

To do well in this component, candidates need to be familiar with the command words detailed in the specification and to use their calculators efficiently in a variety of contexts, such as calculating binomial probabilities or using the inverse Normal function. When interpreting statistical diagrams and tables, they need to make sure their comments relate to the question asked, and to support these comments with numerical arguments where appropriate.

Candidates who did well on this paper generally:	Candidates who did less well on this paper generally:
• understood the need to present sufficient working when completing a 'Detailed reasoning' question • understood that a request to 'Determine' entails providing some justification for their answer • made efficient and appropriate use of their calculators • understood how to extract, process, analyse and make inferences from graphs, tables and software output • were familiar with the pre-release material and were able to apply their understanding in a variety of contexts • expressed their solutions clearly and concisely using correct mathematical notation.	• showed insufficient working when completing their answers to a 'Detailed reasoning' question • did not provide any justification for their answers when faced with a request to 'Determine' • did not always understand when an answer derived from their calculator was clearly wrong – for example, presenting a probability greater than 1 as a final answer • were not able to extract the relevant information from graphs, tables and software output provided in the stem of a question • did not appear to be familiar with the pre-release material • did not express their solutions well, making errors with notation of various sorts • did not set their work out in a clear and orderly way; in some cases very unclear presentation made it virtually impossible for examiners to read the response, particularly in the cases where a sentence or two of explanation was requested.

OCR support

Full details of the command words used on questions can be found in section 2d of the specification. A summary poster and student A4 version can be found on Teach Cambridge.

Section A overview

The questions in Section A are designed to be fairly routine procedural and many of the higher scoring candidates earned close to full marks in this section. There is a steady gradient of difficulty through the section, but most candidates could expect to earn at least half marks in this section.

Question 1

1 Determine the sum of the infinite geometric series $9-3+1-\frac{1}{3}+\frac{1}{9}+\ldots$ **[3]**

Candidates who did well in this question successfully used the formula for the sum of an infinite geometric progression with $a = 9$ and $r = -⅓$ to obtain the correct answer.

Candidates who did less well worked with $r = ⅓$ or simply wrote down the correct answer, showing no working.

Candidates who did not do well worked with $r = -3$ or $+3$.

A small number of candidates did not work with the appropriate formula at all, and nearly always did not earn any marks.

Question 2 (a) and (b)

2 The equation of a circle is

$$x^2-12x+y^2+8y+3=0.$$

(a) Find the radius of the circle. **[2]**

(b) State the coordinates of the centre of the circle. **[1]**

Candidates who did well completed the square twice, successfully, and found the value of r without making any slips. They were then able to write down the coordinates of the centre of the circle.

Candidates who did less well either made a sign error in one of the brackets when completing the square, or slipped up in the arithmetic when finding r. They usually obtained the mark in part (b) as a follow through.

Question 3

3 **In this question you must show detailed reasoning.**

Find the smallest possible positive integers m and n such that $\left(\frac{64}{49}\right)^{-\frac{3}{2}} = \frac{m}{n}$. **[3]**

Candidates who did well in this question showed three distinct steps to obtain the correct final answer.

Candidates who did less well ran two of the steps together, and therefore didn't show sufficient detail of their working to earn full marks.

Some candidates did less well because they made one or slips notation or made bracket errors.

Candidates who did not do well used their calculator to find the answer.

Question 4 (a), (b) and (c)

4 A biased octagonal dice has faces numbered from 1 to 8. The discrete random variable X is the score obtained when the dice is rolled once. The probability distribution of X is shown in the table below.

x	1	2	3	4	5	6	7	8
$\mathrm{P}(X = x)$	p	p	p	p	p	p	p	$3p$

(a) Determine the value of p. **[2]**

(b) Find the probability that a score of at least 4 is obtained when the dice is rolled once. **[1]**

The dice is rolled 30 times.

(c) Determine the probability that a score of 8 occurs exactly twice. **[2]**

The majority of candidates did very well in this question; full marks was more common than not.

Candidates who did less well usually lost marks in part (c), either because they showed no justification for their (correct) answer, because they missed the binomial coefficient, or because they used the wrong value of p. A very few candidates did not recognise that the binomial model was appropriate and gave the solution as 0.3^2

Question 5

5 You are given that $\overrightarrow{OA} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$ and $\overrightarrow{OB} = \begin{pmatrix} 5 \\ -3 \end{pmatrix}$.

Determine the exact length of AB. **[3]**

Candidates who did well in this question showed the subtraction of vectors, the use of Pythagoras and presented their final answer in one of the allowed exact forms.

Candidates who did less well made an arithmetic slip when finding $\overrightarrow{AB}$ or simply wrote down the (correct) answer with no justification. A few candidates added the two vectors instead of finding their difference; a minority of these candidates went on to earn SC1 for a complete response.

Question 6

6 The parametric equations of a circle are

$x = 2\cos\theta - 3$ and $y = 2\sin\theta + 1$.

Determine the cartesian equation of the circle in the form $(x-a)^2 + (y-b)^2 = k$, where a, b and k are integers. **[4]**

Candidates who did well usually expressed x and y or $2x$ and $2y$ in terms of $\cos\theta$ and $\sin\theta$ respectively, before using Pythagoras to find the Cartesian equation. Sometimes they lost the final A mark by forgetting to multiply both sides by 4, or by leaving k as 2^2. Some candidates adopted the second approach detailed in the mark scheme, and were usually successful.

Candidates who did less well wrote down the correct answer with no supporting reasoning to earn SC2. Some candidates began with $x^2 = (2\cos\theta - 3)^2$ and $y^2 = (2\sin\theta + 1)^2$ and then tried to eliminate θ. Occasionally this was done successfully, but more often than not there were slips in the algebra or the equation was left in terms of x, y, $\cos\theta$ and $\sin\theta$, thus earning no marks.

Question 7

7 The coefficient of x^8 in the expansion of $(2x+k)^{12}$, where k is a positive integer, is 79 200 000.

Determine the value of k. **[4]**

Candidates who did well identified the correct binomial coefficient of x^8, recognised that 2^8 was also a component of the term in x^8 and set up a correct equation to obtain the correct answer. Some candidates did so having made a bracket error with an initial statement of $2x^8$ instead of $(2x)^8$.

Candidates who did less well did not recover from an initial step of $2x^8$ but still earned the method mark, or (rarely) worked with ${}^{12}C_5$ instead of ${}^{12}C_4$. Some candidates lost the final mark due to a slip in arithmetic.

Section B overview

In Section B, there are more stretching questions involving more problem solving and interpretation. This section proved accessible to the majority of candidates, with most able to earn some marks on every question. Some candidates were able to obtain full marks on most of the questions.

Question 8 (a)

8 A garden centre stocks coniferous hedging plants. These are displayed in 10 rows, each of 120 plants. An employee collects a sample of the heights of these plants by recording the height of each plant on the front row of the display.

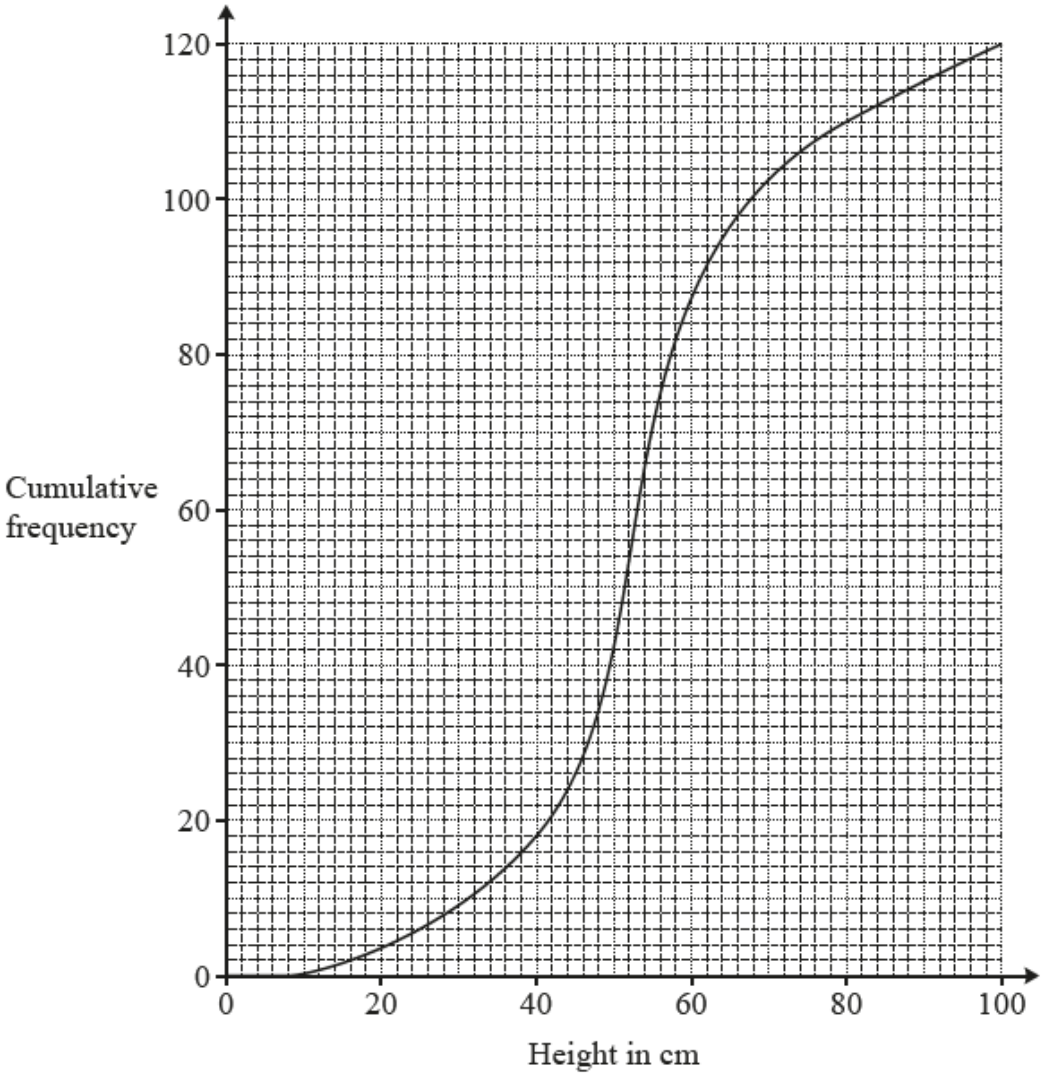
(a) Explain whether the data collected by the employee is a simple random sample. **[1]**

Candidates who did well explained that each possible sample did not have an equal probability of being selected, so the data collected was not a simple random sample. Candidates were also given credit for reasoning that each plant didn't have an equal probability of being selected.

Candidates who did not do well usually tried to describe the sampling method – often incorrectly as, for example, being systematic.

Question 8 (b)

The data are shown in the cumulative frequency curve below.



The owner states that at least 75% of the plants are between 40 cm and 80 cm tall.

(b) Show that the data collected by the employee supports this statement. **[4]**

Candidates who did well found the difference between the appropriate cumulative frequencies, having first stated them by reading them off from the diagram. They then compared this result with 90 (75% of 120) or found 92/120 × 100 and compared the result with 75. They completed the argument by commenting that the calculation supports the owner's statement.

Candidates who did less well gave an incomplete argument by not comparing, for example, 92 with 90, or obtained 110 – 19 = 91 instead of 110 – 18 = 92.

A small minority of candidates did not show sufficient detail for a 'Show that' question.

Question 8 (c)

(c) Explain whether all samples of 120 plants would necessarily support the owner's statement. **[1]**

Candidates who did well reasoned – in a variety of ways, that different samples can produce different results, so another sample would not necessarily support the owner's statement.

Candidates who did not do well often showed some understanding of what was required, but framed their response in terms of rows rather than samples, or went off at a tangent and discussed, for example, growing conditions.

Question 9 (a)

9 The pre-release material contains information concerning the median income of taxpayers in different areas of London. Some of the data for Camden is shown in the table below. The years quoted in this question refer to the end of the financial years used in the pre-release material. For example, the year 2004 in the table refers to the year 2003/04 in the pre-release material.

Year	2004	2005	2006	2007	2008	2009	2010	2011
Median Income in £	21 300	23 200	24 200	25 900	26 900	#N/A	28 400	29 400

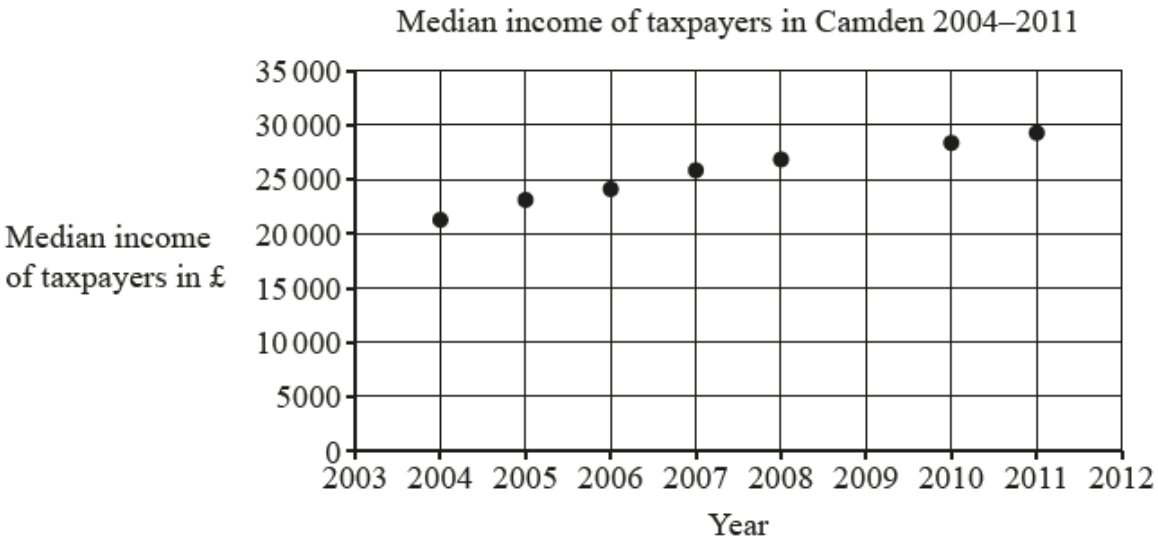
(a) Explain whether these data are a sample or a population of Camden taxpayers. **[1]**

Candidates who did well quoted either the start of the second sentence in the stem or the entry for median income for 2009 to justify the assertion that this is a sample of Camden taxpayers.

Candidates who did not do well either spoiled their response with incorrect reasoning, or stated that the data are a population.

Question 9 (b)

A time series for the data is shown below.



The LINEST function on a spreadsheet is used to formulate the following model for the data:

$I = 1115Y - 2\,212\,950$, where I = median income of taxpayers in £ and Y = year.

(b) Use this model to find an estimate of the median income of taxpayers in Camden in 2009. **[1]**

Nearly all candidates calculated the correct value.

Question 9 (c)

(c) Give **two** reasons why this estimate is likely to be close to the true value. **[2]**

Candidates who did well recognised that the relationship between median income and time appears to be approximately linear, and that the estimate is found from interpolation.

Candidates who did less well either gave only one correct reason, or were unable to express their reasoning sufficiently well. Some candidates reasoned that the estimate is likely to be reliable because 27 085 is between 26 900 and 28 400, which isn't quite the same as interpolation, and didn't earn the mark.

Question 9 (d)

The median income of taxpayers in Croydon in 2009 is also not available.

(d) Use your knowledge of the pre-release material to explain whether the model used in part **(b)** would give a reasonable estimate of the missing value for Croydon. **[1]**

Candidates who did well based their answer on median income being different in different regions, or, in some cases, they knew that median incomes were generally lower in Croydon than in Camden.

Candidates who did not do well usually asserted that the estimate would be reasonable.

Exemplar 1

9(a)

A population, as it includes all of the taxpayer's' incomes in London.

9(b)

$J = 1115Y - 2212950$

$Y = 2009$

$J = 1115 \times 2009 - 2212950$

$= 2240035 - 2212950$

$= 27085$

$= £27085$

9(c)

Follows the pattern of positive correlation ~~and does~~ is between the values from 2008 and 2010.

It uses existy data

9(d)

No, the two areas have different median incomes, so the formula would not apply, as inner and outer london differ in terms of ~~wages~~ income, with inner london having higher income.

There were two clues in the question about the answer to part (a). The second sentence begins 'Some of the data…' and in the table the entry for median income in 2009 is # N/A. In spite of this, many candidates opted for 'population' supported by some spurious reasoning, as in this case.

Nearly all candidates successfully answered part (b), as in this case.

This candidate seems to have grasped the two features of the data which indicate that the estimate is likely to be close to the true value. The comment on positive correlation was sufficient for B1, being taken as equivalent to a suggesting a close linear relationship between the variables, so the use of a linear model is probably sensible.

However, in this case, the second mark is not earned. It is not clear whether the candidate thinks that it is because the value calculated in part (b) is between 26 900 and 28 400 that the estimate is likely to be reliable, or whether it's because 2009 lies between 2008 and 2010 which means we are interpolating.

In part (d) B1 is earned because the candidate states that the estimate is unlikely to be reasonable, and the reasoning presented is based on incomes being different in different areas of London.

Parts (a) and (d) were designated as LDS advantage questions. In part (b), familiarity with using spreadsheets would have been advantageous to the candidates.

Question 10

10 Determine the exact value of $\int_0^{\frac{\pi}{4}} 4x\cos 2x \, dx$. **[5]**

Most candidates identified integration by parts as the correct strategy for this question. Candidates who did well (and most did do well) applied the method successfully and obtained the correct exact answer.

Candidates who did less well made a sign error at either the first or second stage of the application but were still eligible for the second method mark. In some cases they did not show the substitution of the limits and thereby lost this mark. A few candidates started with a simple substitution before going on to apply integration by parts. This was sometimes done completely successfully, but more often than not proved to be an opportunity for a minor slip.

Candidates who did not do well often recognised that integration by parts was the correct approach, but identified u as $\cos 2x$ and v' as $4x$.

Question 11 (a)

11 In this question you must show detailed reasoning.

The variables x and y are such that $\frac{dy}{dx}$ is directly proportional to the square root of x.

When $x = 4$, $\frac{dy}{dx} = 3$.

(a) Find $\frac{dy}{dx}$ in terms of x. **[3]**

Candidates who did well set up the correct differential equation in terms of a constant of proportionality and made the correct substitution to find *k*.

Candidates who did not do well either just wrote down the correct answer, set up an incorrect differential equation, or did not attempt to find a numerical value for *k*.

Question 11 (b)

When $x = 4$, $y = 10$.

(b) Find y in terms of x. **[3]**

Candidates who did well integrated successfully and made the correct substitution to find the value of the constant of integration.

Candidates who did not do well either integrated incorrectly or neglected to include a constant of integration.

Question 12 (a)

12 It is given that

- $f(x) = \pm\frac{1}{\sqrt{x}}, x > 0$
- $g(x) = \frac{x}{x-3}, x > 3$
- $h(x) = x^2 + 2, x \in \mathbb{R}$.

(a) Explain why $f(x)$ is **not** a function. **[1]**

Candidates who did well usually spotted that this is a one-to-many mapping and therefore not a function. Some candidates noted that the mapping is neither a many-to-one nor a one-to-one mapping, and therefore not a function.

Candidates who did not do well spoiled an otherwise correct answer by referring to, for example, a one-to-many function. A significant minority of candidates presented answers which indicated they did not understand the definition of a function.

Question 12 (b)

(b) Find $gh(x)$. **[2]**

The overwhelming majority of candidates earned both marks on this question.

Candidates who did less well lost the accuracy mark by cancelling out x^2 in the fraction and gave an answer of –2.

Question 12 (c)

(c) State the domain of $gh(x)$. **[1]**

Candidates who did well presented their answer as a pair of inequalities.

Candidates who did not do well often recognised that 1 and –1 were crucial but were unable to complete the answer; $x \neq 1, x \neq -1$ was a common wrong answer.

Question 13 (a)

13 A large supermarket chain advertises that the mean mass of apples of a certain variety on sale in their stores is 0.14 kg.

Following a poor growing season, the head of quality control believes that the mean mass of these apples is less than 0.14 kg and she decides to carry out a hypothesis test at the 5% level of significance.

She collects a random sample of this variety of apple from the supermarket chain and records the mass, in kg, of each apple. She uses software to analyse the data. The results are summarised in the output below.

n	80
Mean	0.1316
σ	0.0198
s	0.0199
Σx	10.525
Σx^2	1.4161
Min	0.1
Q1	0.12
Median	0.132
Q3	0.1435
Max	0.19

(a) State the null hypothesis and the alternative hypothesis for the test, defining the parameter used. **[2]**

Candidates who did well stated both hypotheses in terms of μ and went on to define μ as the population mean mass.

Candidates who did less well stated the hypotheses correctly, but did not give a complete definition of μ.

Candidates who used symbols other than μ rarely scored both marks, and often did not earn any marks.

Assessment for learning

When setting up the hypothesis test, μ needs to be defined as the population mean mass in order to earn full credit.

Question 13 (b)

(b) Write down the distribution of the sample mean for this hypothesis test. **[2]**

Candidates who did well were able to identify the mean and variance correctly, and used correct notation when stating the distribution.

Candidates who did less well spoiled an otherwise correct answer by using X instead of $\bar{X}$, or by stating the variance as $\frac{0.0198^2}{80}$ instead of $\frac{0.0199^2}{80}$.

Candidates who did not do well stated the mean as 0.1316 and/or the variance as 0.0198^2 or 0.0199^2.

Misconception

Many candidates mistook the root mean squared deviation for the standard deviation in the software printout when identifying the distribution of the sample mean. It might help to emphasise that the standard deviation will always be the slightly larger quantity, by a scale factor of $\sqrt{\frac{n}{n-1}}$, where n is the sample size.

Question 13 (c)

(c) Determine the critical region for the test. **[2]**

Candidates who did well identified the correct region and stated the answer using correct notation. They may have used $\frac{0.0198^2}{80}$ instead of $\frac{0.0199^2}{80}$.

Candidates who did less well simply stated the critical value, or spoiled a correct answer by stating the region in terms of X or μ instead of $\bar{X}$.

Question 13 (d)

(d) Carry out the test, giving your conclusion in context. **[3]**

Candidates who did well had found the correct critical region in part (c) and noted that 0.1316 is in the critical region. They went on to reject the null hypothesis and give a non-assertive conclusion in context. Some candidates started again and correctly calculated a p-value, which they successfully compared with 0.05. They may have used $\frac{0.0198^2}{80}$ instead of $\frac{0.0199^2}{80}$.

Candidates who did less well spoiled otherwise correct solutions with a response that was too assertive, or contradictory.

Candidates who did not do well did not earn M1, although they often had made well-rehearsed general statements which would have been creditworthy with the correct supporting numerical work.

Question 14 (a)

14 The pre-release material contains information concerning the median income of taxpayers in £ and the percentage of all pupils at the end of KS4 achieving 5 or more GCSEs at grade A*–C, including English and Maths, for different areas of London.

Some of the data for 2014/15 is shown in **Fig. 14.1**.

Fig. 14.1

	Median Income of Taxpayers in £	Percentage of Pupils Achieving 5 or more A*–C, including English and Maths
City of London	61 100	#N/A
Barking and Dagenham	21 800	54.0
Barnet	27 100	70.1
Bexley	24 400	55.0
Brent	22 700	60.0
Bromley	28 100	68.0

A student investigated whether there is any relationship between median income of taxpayers and percentage of pupils achieving 5 or more GCSEs at grade A*–C, including English and Maths.

(a) With reference to **Fig. 14.1**, explain how the data should be cleaned before any analysis can take place. **[1]**

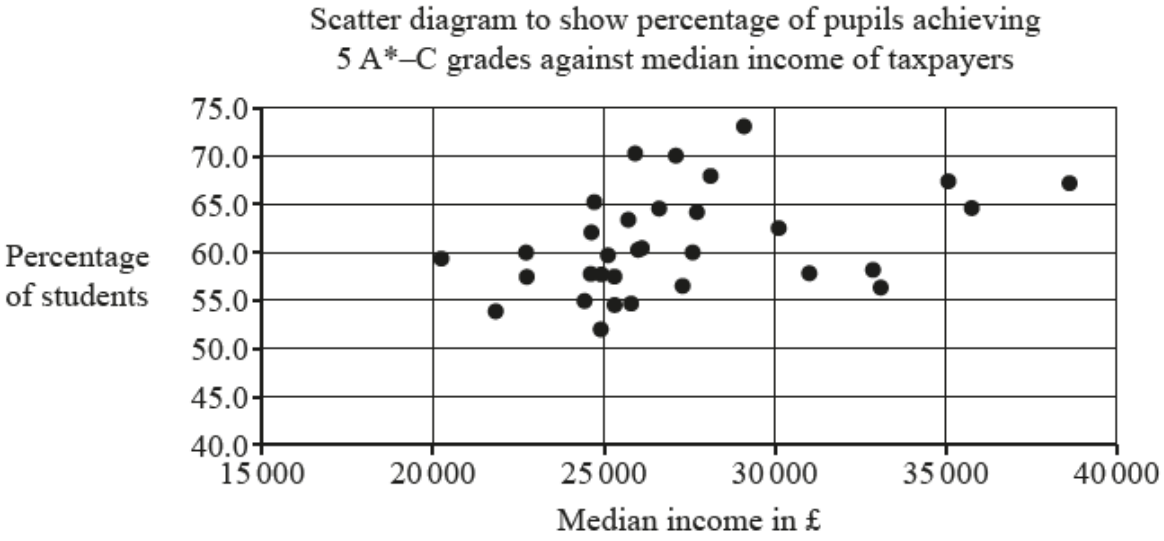
Candidates who did well recognised that regions where some of the data was unavailable would need to be removed.

Candidates who did not do well thought that only part of the data for these regions needed to be removed, or went down the wrong route altogether and tried to identify outliers.

Question 14 (b)

After the data was cleaned, the student used software to draw the scatter diagram shown in **Fig. 14.2**.

Fig. 14.2



The student calculated that the product moment correlation coefficient for these data is 0.3743.

(b) Give **two** reasons why it may not be appropriate to use a linear model for the relationship between median income of taxpayers in £ and the percentage of all pupils at the end of KS4 achieving 5 or more GCSEs at grade A*–C. **[2]**

Candidates who did well compared the value of r with either 0 or 1 and made an appropriate comment. They also referred to the scatter diagram and whether the relationship between median income and percentage of students appears to be linear.

Candidates who did less well only identified one reason correctly.

Candidates who did not do well often made relevant comments, but their responses were too vague – for example 'low correlation' was insufficient to earn either mark.

Question 14 (c)

The student carried out some further analysis. The results are shown in **Fig. 14.3**.

Fig. 14.3

	median income of taxpayers in £	percentage of pupils achieving 5+ A*–C
mean	27216	61.0
standard deviation	4177.5	5.32

The student identified **three** outliers in total.

(c)

- Use the information in **Fig. 14.3** to determine the range of values of the median income of taxpayers in £ which are outliers.
- Use the information in **Fig. 14.3** to determine the range of values of the percentage of all pupils at the end of KS4 achieving 5 or more GCSEs at grade A*–C which are outliers.
- On the copy of **Fig. 14.2** in the **Printed Answer Booklet**, circle the **three** outliers identified by the student. **[4]**

Candidates who did well used the two standard deviation check successfully, and gave their answer as four inequalities. They were able to identify the outliers on the diagram.

Candidates who did less well either only made partial use of the two standard deviation check – just calculating values for median income, for example, or just stated the four values.

Candidates who did not do well used a 1, 1.5 or 3 standard deviation check, or attempted to work with the interquartile range.

Question 14 (d)

The student decided to remove these outliers and recalculate the product moment correlation coefficient.

(d) Explain whether the new value of the product moment correlation coefficient would be between 0.3743 and 1 or between 0 and 0.3743. **[1]**

The relatively small number of candidates who did well on this question identified the correct range of values and reasoned this was so because the scatter was now less linear – although this was expressed in a variety of ways.

Candidates who did not do well gave the correct region but were not able to provide an adequate explanation. A significant minority identified the wrong region, reasoning that removing outliers makes the data more reliable.

Exemplar 2

14(a)

Remove the City of London as it has no data for percentage of pupils achieving 5 or more GCSEs at grade A* - C, including English and Maths.

14(b)

The product moment correlation coefficient is 0.37413, which shows a very weak positive correlation.

The scatter graph does not show a linear pattern.

14(c)

Median income outliers: $27216 \pm 2(4177.5) = 35571$ ~~$27216 + 2(4177.5) = 88571$~~ $= 18861$

~~$27216 - 2(4177.5) = 18864$~~ ~~$23038.5$~~

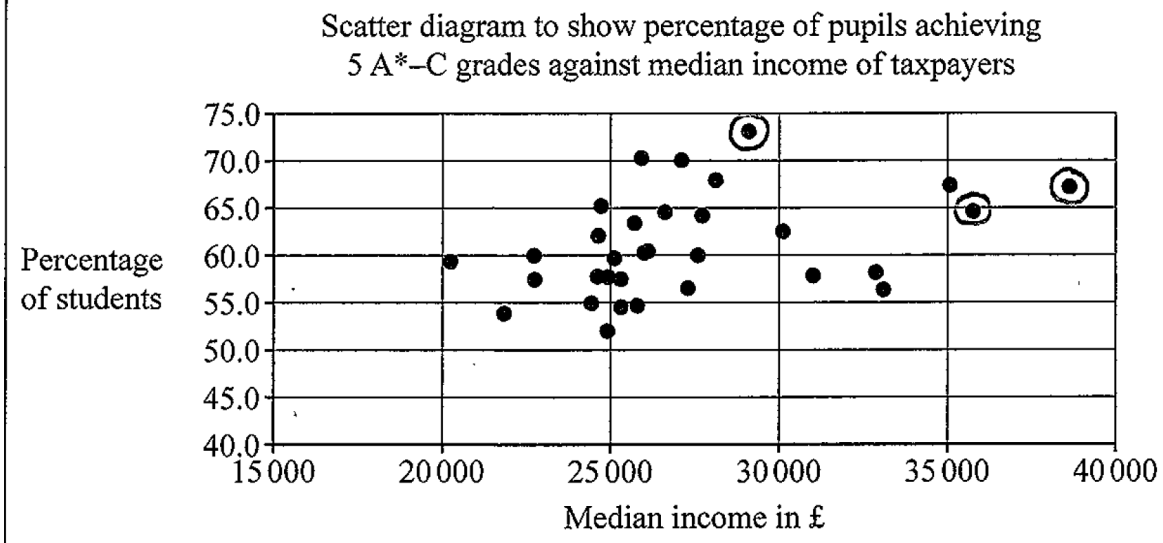
~~Brent < 23038.5 Barking and Dagenham < 23038.5~~

Grade Outliers:

~~$61 + 2(5.32) = 71.64$~~ $61 + 2(5.32) = 71.64$

~~$61 - 2(5.32) = 50.36$~~ $61 - 2(5.32) = 50.36$

~~Bexley~~



14(d) It would be $between 0 and 0.3743 as in Fig. 14.2 all the points cluster in a concentrated area.

In part (a) the candidate noted that all the data relating to City of London needs to be removed, and not just the cell containing #N/A, thereby earning B1. This is an example of an LDS advantage question.

In part (b), although the value of the correlation coefficient is referred to, it is not compared to 0 or to 1, so there is insufficient for B1. The second B1 is earned because of the reference to non-linearity of the scatter.

In part (c) the candidate performs the calculations correctly, but does not go on to identify the ranges of values of outliers in either case, thereby earning M1SC1. The correct points are then identified on the graph for the final A mark.

In part (d) the correct range of values is identified, but the justification is insufficient for the award of the mark. 'All the points cluster in a constant area' doesn't quite convey the idea that the scatter is less linear in nature.

Question 15

15 In this question you must show detailed reasoning.

The equation of a curve is

$\ln y + x^3 y = 8$.

Find the equation of the normal to the curve at the point where $y = 1$, giving your answer in the form $ax + by + c = 0$, where a, b and c are constants to be found. **[7]**

Candidates who did well usually differentiated implicitly with respect to *x* before (unnecessarily) expressing the derivative in terms of *x* and *y* in order to find the gradient of the tangent by substitution of *x* = 2 and *y* = 1. Many went on to obtain full marks, but some slipped up in the initial rearrangement and lost the final accuracy mark. Candidates who differentiated implicitly with respect to *y* were generally equally as successful, but some were unable to obtain the gradient of the normal from working with $^{dx}/_{dy}$

Candidates who did less well made errors with differentiating and then lost the method marks by not showing their substitutions. Those candidates who made *x* the subject (or, occasionally, *y*) more often than not lost marks when differentiating and often didn't recover.

Question 16 (a)

16 Research conducted by social scientists has shown that 16% of young adults smoke cigarettes.

Two young adults are selected at random.

(a) Determine the probability that one smokes cigarettes and the other doesn't. **[2]**

Candidates who did well used the binomial distribution or a tree diagram to obtain the correct answer.

Candidates who did less well simply wrote down the correct answer without regard of the 'determine' command word that signifies that answers need to be justified. Others found 0.16×0.84 but missed out the binomial coefficient.

Question 16 (b)

The same research has also shown that

- 75% of young adults drink alcohol.
- 66% of young adults drink alcohol, but do **not** smoke cigarettes.

(b) Determine the probability that a young adult selected at random **does** smoke cigarettes, but **does not** drink alcohol. **[2]**

Candidates who did well usually worked with a Venn diagram or a contingency table.

Candidates who did not do well usually multiplied probabilities, possibly having worked with a tree diagram.

Question 16 (c)

(c) A young adult who drinks alcohol is selected at random. Determine the probability that this young adult smokes cigarettes. **[2]**

Candidates who did well recognised this as a conditional probability question and extracted the correct numerator from part (b). A small number of candidates successfully worked with $1-\frac{66}{75}$ or equivalent.

Candidates who did not do well generally mistook the request for $P(S \cap A)$ instead of $P(S|A)$.

Question 16 (d)

(d) Using your answer to part **(c)**, explain whether the event that a young adult selected at random smokes cigarettes is independent of the event that a young adult selected at random drinks alcohol. **[2]**

Candidates who did well recognised that $P(S|A) = P(A)$ if the events are independent, and tested accordingly.

Candidates who did less well re-started and showed that $P(S \cap A) \neq P(S) \times P(A)$, to earn SC1.

Question 17

17 In this question you must show detailed reasoning.

Solve the equation $2\sin x + \sec x = 4\cos x$, where $-\pi < x < \pi$. **[6]**

Candidates who did well generally adopted the first or alternative approach detailed in the mark scheme, but there were some successful attempts based on squaring both sides first and obtaining a quartic in $\sin x$ or $\cos x$.

Candidates who did less well, made sign errors following the award of B1M1 and obtained an incorrect quadratic. Candidates who did not adopt the first approach sometimes did less well because they did not appreciate that their method introduced extra solutions.

Candidates who did not do well often made no further progress following the award of B1.

Question 18 (a), (b), (c), (d), (e) and (f)

18 Riley is investigating the daily water consumption, in litres, of his household.
He records the amount used for a random sample of 120 days from the previous twelve-month period.

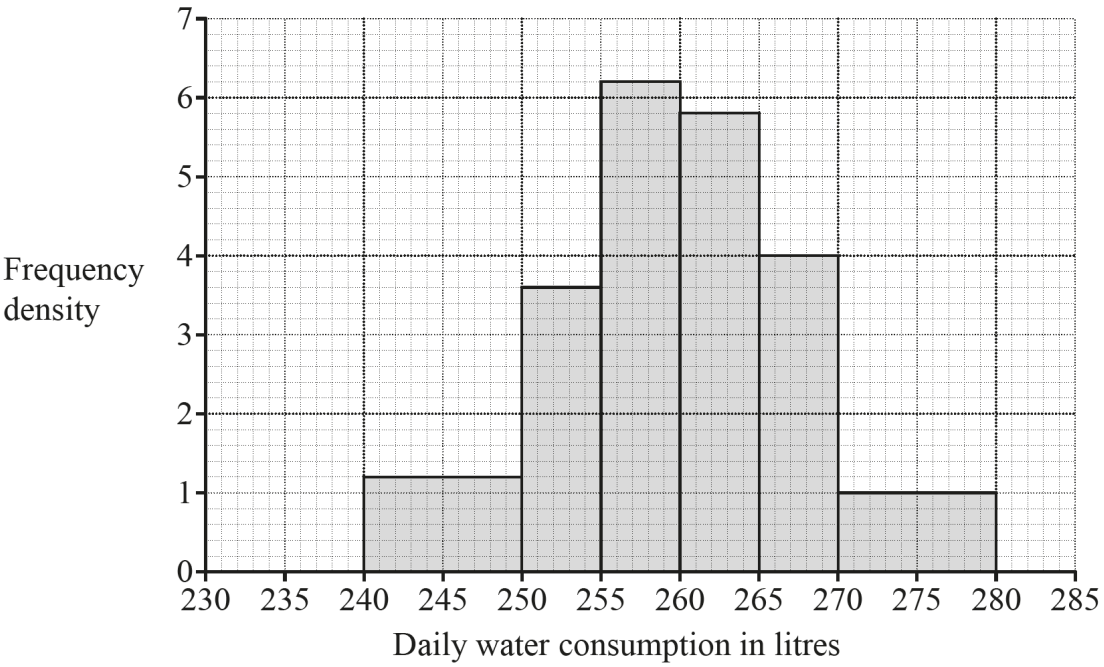
The daily water consumption, in litres, is denoted by x.

Summary statistics for Riley's sample are given below.

$\sum x = 31164.7 \quad \sum x^2 = 8101050.91 \quad n = 120$

(a) Calculate the sample mean giving your answer correct to **3** significant figures. **[1]**

Riley displays the data in a histogram.



(b) Find the number of days on which between 255 and 260 litres were used. **[1]**

(c) Give **two** reasons why a Normal distribution may be an appropriate model for the daily consumption of water. **[2]**

Riley uses the sample mean and the sample variance, both correct to **3** significant figures, as parameters of a Normal distribution to model the daily consumption of water.

(d) Use Riley's model to calculate the probability that on a randomly chosen day the household uses less than 255 litres of water. **[2]**

(e) Calculate the probability that the household uses less than 255 litres of water on **at least** 5 days out of a random sample of 28 days. **[2]**

The company which supplies the water makes charges relating to water consumption which are shown in the table below.

Standing charge per day in pence	7.8
Charge per litre in pence	0.18

(f) Adapt Riley's model for daily water consumption to model the daily **charges** for water consumption. **[3]**

Candidates who did well answered the first three parts completely correctly. They found the correct variance or standard deviation in part (d) to calculate the requested probability. In part (e) they recognised that this is a binomial situation, and used their value from part (d) to calculate the probability. In part (f) they knew how to find the mean and variance for the new model, possibly slipping up with the arithmetic for the variance.

Candidates who did less well ignored the request for 3 significant figures in part (a) and may have only identified one valid reason in part (c). 'The sample size is large' was a common response which did not score. In part (d) the mean squared deviation instead of the variance may have been found, so only the FT mark would be available in part (e). Numerical errors in previous parts may have proved costly in part (f), due to no FT marks being available, but in practice very few candidates fell foul of this, as those who went astray earlier in the question usually did not know how to tackle this question.

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.

Access to Scripts

For the June 2023 series, Exams Officers will be able to download copies of your candidates' completed papers or 'scripts' for all of our General Qualifications including Entry Level, GCSE and AS/A Level. Your centre can use these scripts to decide whether to request a review of marking and to support teaching and learning.

Our free, on-demand service, Access to Scripts is available via our single sign-on service, My Cambridge. Step-by-step instructions are on our 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 question builder platform for a range of our GCSE, A Level, Cambridge Nationals and Cambridge Technicals qualifications. Find out more.

ExamBuilder is **free for all OCR centres** with an Interchange account and gives you unlimited users per centre. We need an Interchange username to validate the identity of your centre's first user account for ExamBuilder.

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

Active Results

Review students' exam performance with our free online results analysis tool. It is available for all GCSEs, AS and A Levels and Cambridge Nationals.

Find out more.

Need to get in touch?

If you ever have any questions about OCR qualifications or services (including administration, logistics and teaching) please feel free to get in touch with our customer support centre.

Call us on
01223 553998

Alternatively, you can email us on
support@ocr.org.uk

For more information visit
ocr.org.uk/qualifications/resource-finder
ocr.org.uk
facebook.com/ocrexams
twitter.com/ocrexams
instagram.com/ocrexaminations
linkedin.com/company/ocr
youtube.com/ocrexams

We really value your feedback

Click to send us an autogenerated email about this resource. Add comments if you want to. Let us know how we can improve this resource or what else you need. Your email address will not be used or shared for any marketing purposes.

I like this **I dislike this**

Please note – web links are correct at date of publication but other websites may change over time. If you have any problems with a link you may want to navigate to that organisation's website for a direct search.

CAMBRIDGE
UNIVERSITY PRESS & ASSESSMENT