

Statistics GCSE

Paper 2

2025

Edexcel Higher

Variant 4

1ST0/1H

Mark scheme

Visit our website for tutorials on each question.

www.statsgcse.com

| Question | Mark Scheme | Mark |
|----------|--|------|
| 1 (a) | [1 mark each] for any two from: + There may be repeated random numbers + Random numbers may not correspond to students' numbers + Selected people may not want to participate + Some people may have left the hospital | 2 |

Question 1 (a) model answer

Some people may have left the hospital.

The selected person may not want to participate.

| Question | Mark Scheme | Mark |
|----------|--|------|
| 1 (b) | [1 mark each] for any two from: + Simple, convenient, fast, efficient, or cost-effective + Reflects population or proportion + Enables comparison between in-costume visitors and plain-clothed visitors + Does not require a sampling frame | 2 |

Question 1 (b) model answer

Quota sampling is efficient and allows for comparison between in-costume visitors and plain-clothed visitors.

| Question | Mark Scheme | Mark |
|----------|--|------|
| 1 (c) | [1 mark] One from: + Not all people have an equal chance of being selected. + Only people who go to that entrance are considered. + The people are being picked by Jake | 1 |

Question 1 (c) model answer

Only the people who go to that entrance are considered.

| Question | Mark Scheme | Mark |
|----------|--------------------------------|------|
| 2 (a) | [1 mark] 4.1% to 6.0% increase | 1 |

| Question | Mark Scheme | Mark |
|----------|-------------|------|
| 2 (b) | [1 mark] 1 | 1 |

| Question | Mark Scheme | Mark |
|----------|--|------|
| 2 (c) | [1 mark] Explanation that shows an understanding that we have percentages not numbers. | 1 |

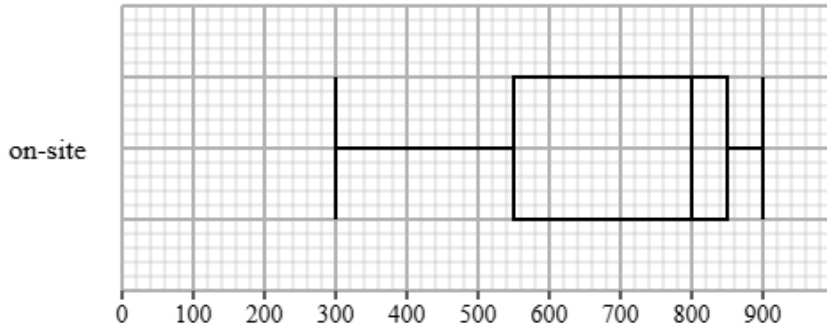
Question 2 (c) model answer

The map does not show the increases and decreases as amounts.

| Question | Mark Scheme | Mark |
|----------|---------------------------|------|
| 2 (d) | [1 mark] choropleth (map) | 1 |

| Question | Mark Scheme | Mark |
|----------|---|------|
| 3 (a) | [1 mark] A box with two whiskers drawn with at least 3 correct values [1 mark] Fully correct | 2 |

Question 3 (a) model answer



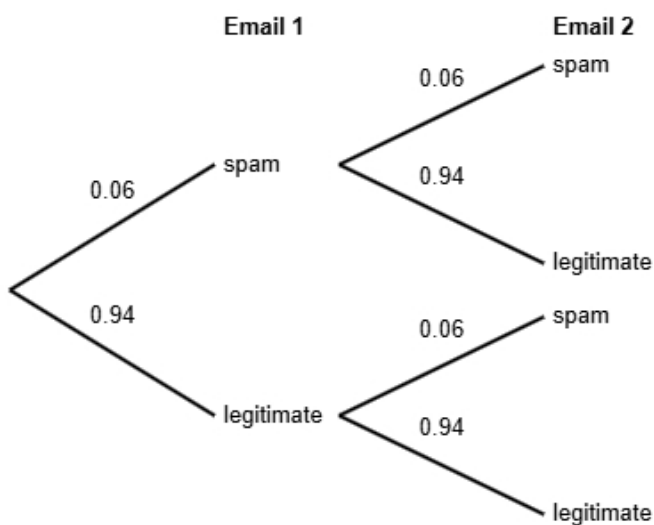
| Question | Mark Scheme | Mark |
|----------|--|------|
| 3 (b) | [1 mark] Correct comparison of the medians [1 mark] Correct comparison of the IQR/range [1 mark] Correct comparison of the skews [1 mark] Correct contextual interpretation comparing medians or IQR/ranges or skew | 4 |

Question 3 (b) model answer

The median steps for remote workers is lower than on-site workers. The IQR for the steps of the remote workers is lower than on-site workers. The skew for the steps of the remote workers is symmetrical and the skew for the on-site workers is negative. The remote workers on average walk less than the on-site workers.

| Question | Mark Scheme | Mark |
|----------|--|------|
| 4 (a) | [1 mark] 0.94 in correct position for Email 1. [1 mark] 0.06, 0.94, 0.06 and 0.94 in correct positions for Email 2. | 2 |

Question 4 (a) model answer



| Question | Mark Scheme | Mark |
|----------|-----------------|------|
| 4 (b) | [1 mark] 0.8836 | 2 |

| Question | Mark Scheme | Mark |
|----------|--|------|
| 4 (c) | [1 mark] for one correct product using their '0.94' or subtracting a not-wanted product from 1 [1 mark] for 0.1128 or 11.28% [1 mark] for 'correct' ft probability and conclusion based on their probability | 3 |

Question 4 (c) model answer

$$\begin{aligned}
 P(\text{spam AND legitimate}) &= 0.06 \times 0.94 \\
 &= 0.0564 \\
 P(\text{exactly one email is spam}) &= 0.0564 \times 2 \\
 &= 0.1128 \\
 0.1128 &= 11.28\% \\
 11.28\% &< 12\%
 \end{aligned}$$

The probability that exactly one email is spam is less than 12%, so Maria is correct.

| Question | Mark Scheme | Mark |
|----------|---|------|
| 5 (a) | [1 mark] for correct comment on the type of data, such as discrete, whole numbers, integers, etc. | 1 |

Question 5 (a) model answer

Because number of traffic accidents reported in a neighbourhood is discrete.

| Question | Mark Scheme | Mark |
|----------|-------------|------|
| 5 (b) | [1 mark] 1 | 1 |

| Question | Mark Scheme | Mark |
|----------|--|------|
| 5 (c) | Part i [1 mark] 0 Part ii [1 mark] Subtracts 50 from 65. [1 mark] 15 | 3 |

| Question | Mark Scheme | Mark |
|----------|-------------|------|
| 5 (d) | [1 mark] 6 | 1 |

| Question | Mark Scheme | Mark |
|----------|-------------------------------|------|
| 5 (e) | [1 mark] for a correct reason | 1 |

Question 5 (e) model answer

The range is 6, so the IQR must be less than 6.

| Question | Mark Scheme | Mark |
|--------------|--|----------|
| 6 (a) | [1 mark] A and C [1 mark] no intersection/overlap | 2 |

Question 6 (a) model answer

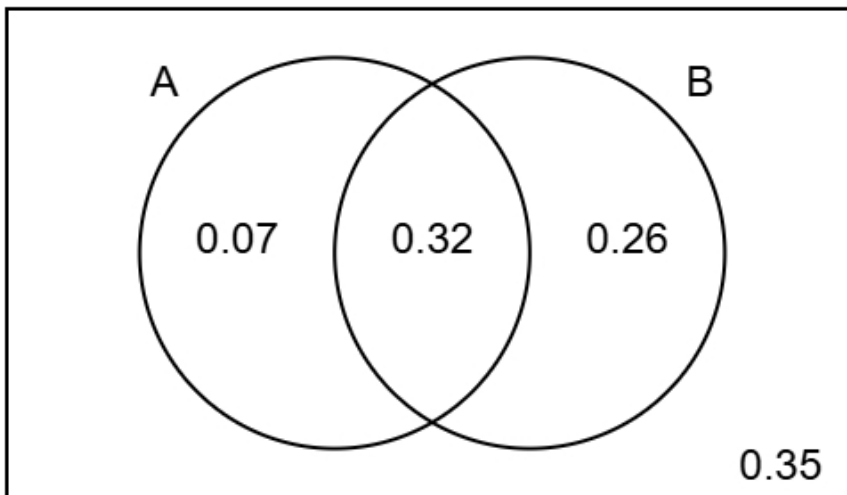
A and C are mutually exclusive because they do not overlap.

| Question | Mark Scheme | Mark |
|--------------|---------------|----------|
| 6 (b) | [1 mark] 0.58 | 1 |

| Question | Mark Scheme | Mark |
|--------------|--|----------|
| 6 (c) | [1 mark] $P(A \text{ or } C) = P(A) + P(C)$ or $(0.07 + 0.32) + (0.08 + 0.29)$ [1 mark] 0.76 | 2 |

| Question | Mark Scheme | Mark |
|----------|--|------|
| 6 (d) | [1 mark] 0.07 and 0.32 in the correct places. [1 mark] 0.26 and 0.35 in the correct places. | 2 |

Question 6 (d) model answer



| Question | Mark Scheme | Mark |
|----------|--------------------------|------|
| 7 (a) | [1 mark] Scatter diagram | 1 |

| Question | Mark Scheme | Mark |
|----------|--|------|
| 7 (b) | [1 mark] Finding d^2 [1 mark] $1 - \frac{6 \times 282}{10(10^2 - 1)}$ [1 mark] -0.71 [1 mark] As the average wait time increases, customer satisfaction decreases. [1 mark] The presence of outliers would decrease the correlation. | 5 |

| Question | Mark Scheme | Mark |
|----------|--|------|
| 7 (c) | [1 mark] PMCC measures linear correlation [1 mark] Spearman's rank correlation is used for ranked/ordinal data [1 mark] Henry should not use PMCC or should use Spearman's rank correlation | 3 |

Question 7 (c) model answer

Henry should not use PMCC because the PMCC only measures linear correlation and Spearman's rank correlation is used for ranked data.

| Question | Mark Scheme | Mark |
|----------|--|------|
| 8 (a) | [1 mark] the pie charts do not show the total budget amount or pie charts only show proportions, not amounts. | 1 |

Question 8 (a) model answer

There may have been more people overall using the village hall in November 1998.

| Question | Mark Scheme | Mark |
|----------|---|------|
| 8 (b) | [1 mark] $\frac{7832}{22} \times 15$ [1 mark] 5340 | 2 |

| Question | Mark Scheme | Mark |
|----------|--|------|
| 8 (c) | [1 mark] Comparative pie charts (or a description of them) [2 marks] A calculation to find the radius of one pie chart from the other. For example, $\sqrt{\frac{43000}{36000}} = 1.09$ x June 1998 will give the radius for November 1998. OR [1 mark] The radii of the pie charts will be in proportion to the square root of the ratio of the totals. [1 mark] The areas will represent the totals. This mark can be award from an example. [1 mark] The proportions/percentages will be the same. | 5 |

Question 8 (c) model answer

Joe can use comparative pie charts.

$$\sqrt{\frac{43000}{36000}} = 1.09$$

The radius of the November 2015 pie chart will be 1.09 larger than the February 2016 pie chart.

Because the totals differ across these pie charts, the overall area of each will reflect that difference, while the proportions remain the same.

| Question | Mark Scheme | Mark |
|----------|--|------|
| 9 (a) | [2 marks] 0.18 OR [1 mark] $0.15 + 3 \times 0.01$ | 2 |

| Question | Mark Scheme | Mark |
|----------|---|------|
| 9 (b) | [1 mark] Stoppage/resets in production process more frequently. | 1 |

Question 9 (b) model answer

The production process would be stopped more frequently.

| Question | Mark Scheme | Mark |
|----------|--|------|
| 9 (c) | [1 mark] The process will need to stop because the range is outside the upper action limit [1 mark] The mean is outside the upper warning limit, but not the action limit | 2 |

Question 9 (c) model answer

The production process should be stopped as the range is outside the upper action limit although the mean is outside the upper warning limit but not the action limit.

| Question | Mark Scheme | Mark |
|----------|--------------|------|
| 10 (a) | [1 mark] 3.7 | 1 |

| Question | Mark Scheme | Mark |
|----------|--|------|
| 10 (b) | <p>[3 marks] 2.04</p> <p>OR</p> <p>[2 marks] $1.2 \times (0.6 - -1.1)$</p> <p>OR</p> <p>[1 mark] $(0.6 \times 1.2) + 3.7$ or 4.42</p> <p>[1 mark] $(-1.1 \times 1.2) + 3.7$ or 2.38</p> | 3 |

Question 10 (b) model answer

$$\begin{aligned} \text{Ben's daily screen time} &= (0.6 \times 1.2) + 3.7 \\ &= 4.42 \\ \text{Charlie's daily screen time} &= (-1.1 \times 1.2) + 3.7 \\ &= 2.38 \end{aligned}$$

Difference

$$4.42 - 2.38 = 2.04$$

| Question | Mark Scheme | Mark |
|----------|--|------|
| 10 (c) | Part i [2 marks] Appropriate and correct reason (such as sample mean will be close to the population mean) OR [1 mark] Correct reason Part ii [2 marks] Not appropriate and correct reason (such as sample mean will be more closely distributed than the individual values) OR [1 mark] Correct reason | 4 |

Question 10 (c) model answer

Using 3.7 hours as the mean is appropriate because the sample mean will be close to the population mean.

Using 1.2 hours as the standard deviation is not appropriate because the sample mean will be more closely distributed than the individual values.

| Question | Mark Scheme | Mark |
|----------|--|------|
| 11 (a) | [1 mark] Correct interpretation, but must be in context. | 1 |

Question 11 (a) model answer

For each 1 hour increase in hours worked the calories burned would decrease by 35 kcal.

| Question | Mark Scheme | Mark |
|----------|---|------|
| 11 (b) | [1 mark] Correct comparison of regression equations of all remote workers [1 mark] Correct comparison of regression equations between Q1 and Q2 [1 mark] Correct comparison of regression equations between junior employees and senior employees | 3 |

Question 11 (b) model answer

For all remote workers, an increase in hours worked led to a decrease in calories burned. As hours worked increased, the calories burned decreased more rapidly in Q1 compared to Q2. The calories burned of senior employees decreased more per hours worked than junior employees.

| Question | Mark Scheme | Mark |
|----------|--|------|
| 11 (c) | Part i [1 mark] Would show a bell curve, symmetry or skew Part ii Max 2 marks [1 mark] All averages equal [1 mark] 68% data within 1 standard deviation, 95% within 2 standard deviations, almost all within 3 standard deviations [1 mark] Can calculate skewness with mean/median and standard deviation | 3 |

Question 11 (c) model answer

If the histogram is symmetrical, a normal distribution could be a suitable model.
 If the skewness is 0 (calculated from the mean, median and standard deviation) and 95% of data is within 2 standard deviations from the mean, a normal distribution could be a suitable model.

| Question | Mark Scheme | Mark |
|----------|---|------|
| 12 (a) | <p>Max 2 marks (only 1 mark if no context)</p> <p>[1 mark] The events are independent using context</p> <p>[1 mark] There are only two outcomes using context</p> <p>[1 mark] The probability remains constant using context</p> | 2 |

Question 12 (a) model answer

There are only two possible outcomes, purchase or not purchase. The chance of a purchase remains constant.

| Question | Mark Scheme | Mark |
|----------|---|------|
| 12 (b) | <p>[2 marks] $\frac{1}{10000}$</p> <p>OR</p> <p>[1 mark] $\left(\frac{10}{100}\right)^3$</p> | 2 |

Question 12 (b) model answer

$$\left(\frac{10}{100}\right)^3 = \frac{1}{10000}$$

| Question | Mark Scheme | Mark |
|----------|--|------|
| 12 (c) | <p>[3 marks] $\frac{523}{10000}$</p> <p>OR</p> <p>[1 mark] Correctly finds a probability e.g. $P(X = 0)$</p> <p>[1 mark] Correct working for $P(X \geq 2)$</p> | 3 |

Question 12 (c) model answer

$$\begin{aligned}
 P(X = 0) &= \left(\frac{90}{100}\right)^4 \\
 &= \frac{6561}{10000} \\
 P(X = 1) &= 4 \times \left(\frac{10}{100}\right) \times \left(\frac{90}{100}\right)^3 \\
 &= \frac{729}{2500} \\
 P(X \geq 2) &= 1 - \left(\frac{6561}{10000} + \frac{729}{2500}\right) \\
 &= \frac{523}{10000}
 \end{aligned}$$