Mr Ronald Maths and Statistics

Predicted Examination Paper

This is a predicted paper for Paper 2 of the
Edexcel GCSE (9–1) Statistics (1ST0) examination on 12 June 2026.

Surname

Forename(s)

GCSE (9–1) Statistics

Paper 2 · Higher Tier

Time allowed: 1 hour 30 minutes
Total marks: 80

You must have: Calculator, ruler graduated in centimetres and millimetres, protractor, pair of compasses, pen, HB pencil, eraser.

Instructions

- Use black ink or ball-point pen.
- Fill in your name in the boxes at the top of this page.
- Answer **all** questions.
- Answer the questions in the spaces provided – *there may be more space than you need.*
- You must show all your working.
- Calculators may be used.
- Diagrams are **not** accurately drawn, unless otherwise indicated.
- Give probabilities as fractions, decimals or percentages.

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.
- A Higher Tier formulae sheet is provided on the next page.

Advice

- Read each question carefully before you start to answer it.
- Keep an eye on the time. Try to answer every question.
- Check your answers if you have time at the end.

Higher Tier Formulae Sheet

You must not write on this formulae sheet. Anything you write on this page will gain no credit.

Standard deviation

$$\sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2} \quad \text{or} \quad \sqrt{\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2}$$

A measure of skewness

$$\frac{3\,(\text{mean} - \text{median})}{\text{standard deviation}}$$

Spearman's rank correlation coefficient

$$r_s = 1 - \frac{6\sum d^2}{n\,(n^2 - 1)}$$

Crude rate (e.g. birth rate)

$$\frac{\text{number of births}}{\text{total population}} \times 1000$$

Crude rate (e.g. death rate)

$$\frac{\text{number of deaths}}{\text{total population}} \times 1000$$

1 A researcher is designing a questionnaire to investigate the financial habits of adults. She writes the following question to find out about each person's annual salary.

How much do you earn?

☐ £0 – £20,000 ☐ £20,000 – £40,000 ☐ £40,000 – £60,000

(a) Write down **two** things that are wrong with this question.

(2)

.....

.....

.....

(b) Design a more suitable question for the researcher to use. You should include responses and response boxes.

(2)

Question: .....

.....

(c) The researcher also records the number of bank accounts each person owns.
Tick all of the boxes below that describe this type of data.

(1)

☐ Discrete ☐ Continuous ☐ Categorical

☐ Qualitative ☐ Quantitative

(d) The researcher decides instead to obtain salary information from a published government report. State whether this data is now primary or secondary, and give a reason for your answer.

(2)

.....

.....

(Total for Question 1 is 7 marks)

2 120 students were asked how they usually travel to school. The two-way table shows some information about their answers.

	Walk	**Bus**	**Car**	**Total**
Boys	18		22	70
Girls		14	12	
Total	42			120

(a) Complete the two-way table.

(3)

(b) One of the boys is chosen at random. Find the probability that this student walks to school.

(2)

.....

.....

.....

Answer .....

(Total for Question 2 is 5 marks)

3 A company has 800 employees. A manager wants to find out what the employees think about a new shift pattern. The manager stands at the main entrance one Monday morning and asks the first 40 employees who arrive.

(a) Write down the name of this sampling method.

(1)

.....

(b) Explain why this method is likely to give biased results.

(2)

.....

.....

.....

(c) The 800 employees consist of 500 full-time staff, 200 part-time staff and 100 contractors. The manager decides instead to take a stratified sample of 40 employees. Work out the number of part-time staff that should be in the sample.

(3)

.....

.....

.....

.....

Answer .....

(Total for Question 3 is 6 marks)

4 The table gives information about the time, t minutes, taken by 50 customers to do their shopping.

Time (t minutes)	**Frequency**
$0 < t \leq 10$	10
$10 < t \leq 20$	20
$20 < t \leq 40$	15
$40 < t \leq 70$	5

(a) Work out an estimate for the mean time taken.

(4)

.....

.....

.....

.....

Answer ..... minutes

(b) Explain why your answer to part (a) is only an estimate.

(1)

.....

.....

(c) Write down the class interval that contains the median.

(2)

.....

.....

Answer .....

(d) When the data was first collected, one of the frequencies was recorded as −3. Give a reason why this value cannot be correct.

(1)

.....

(Total for Question 4 is 8 marks)

5 The summary statistics below describe the number of emails received in one day by a group of office workers.

Minimum = 12 Lower quartile = 28 Median = 35 Upper quartile = 44 Maximum = 70

The mean is 38.2 and the standard deviation is 14.5.

(a) Work out the interquartile range.

(1)

.....

Answer .....

(b) An **outlier** is a value that is either more than $1.5 \times$ IQR below the lower quartile, or more than $1.5 \times$ IQR above the upper quartile.

Determine whether there are any outliers in this data set. You must show your working.

(4)

.....

.....

.....

.....

(c) The value of 70 is found to be a genuine, correctly recorded value. State what should be done with this value before any further analysis is carried out.

(1)

.....

(d) Work out a measure of the skewness of the number of emails received.

(3)

.....

.....

.....

Answer .....

(Total for Question 5 is 9 marks)

6 Maria took a Spanish test and a French test.

- In the Spanish test Maria scored 72 marks. The class mean was 60 marks, with a standard deviation of 8 marks.
- In the French test Maria scored 80 marks. The class mean was 70 marks, with a standard deviation of 12 marks.

(a) Work out the standardised score for Maria's Spanish test.

(2)

.....

.....

.....

Answer .....

(b) Work out the standardised score for Maria's French test.

(2)

.....

.....

.....

Answer .....

(c) In which test did Maria perform better, relative to the rest of her class? You must explain your answer.

(3)

.....

.....

.....

.....

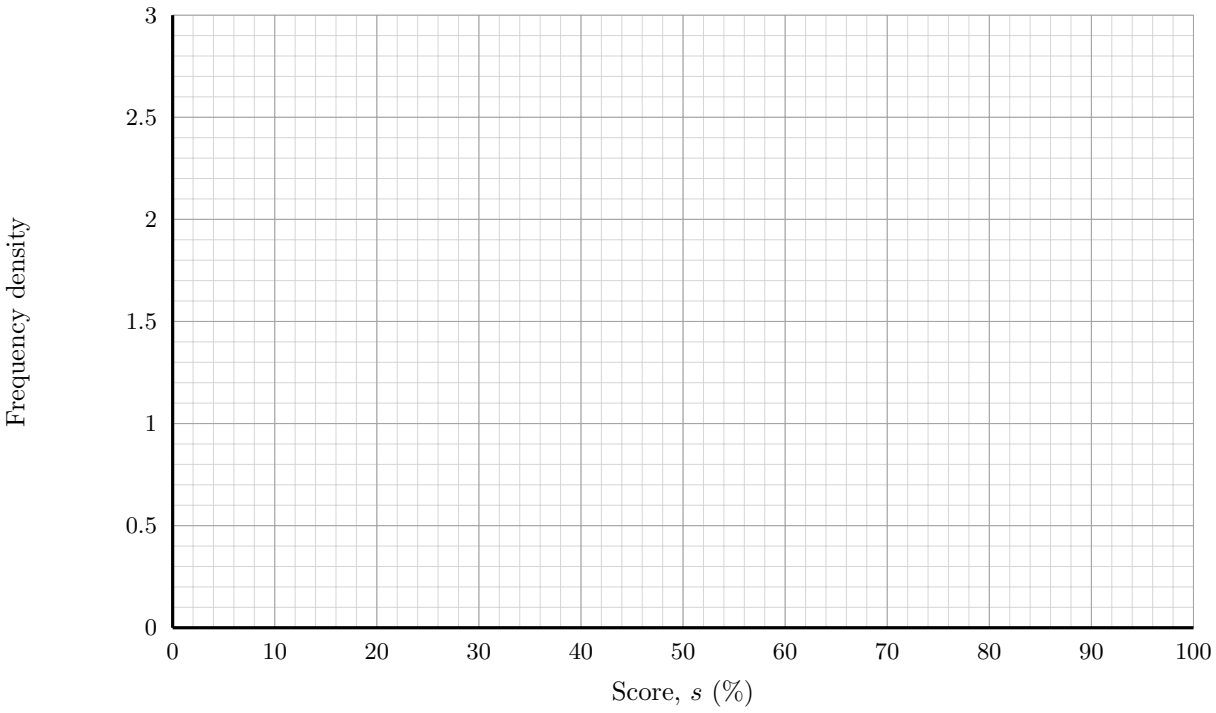
(Total for Question 6 is 7 marks)

7 The table gives information about the test scores, s, of 135 students. Each score is recorded as a percentage.

Score, s (%)	**Frequency**	**Frequency density**
$0 < s \leq 30$	30	
$30 < s \leq 50$	40	
$50 < s \leq 60$	25	
$60 < s \leq 80$	30	
$80 < s \leq 100$	10	

(a) On the grid below, draw a histogram for the information in the table. You may use the final column of the table for your working.

(3)



(b) Use your histogram to estimate the number of students who scored less than 45%.

(3)

.....

.....

.....

.....

Answer .....

(c) Explain why your answer to part (b) is only an estimate, and state the assumption you made in working it out.

(2)

.....

.....

.....

(Total for Question 7 is 8 marks)

8 The table shows the price of an annual train season ticket at the start of each year from 2020 to 2023.

Year	**Price (£)**
2020	1200
2021	1260
2022	1320
2023	1380

(a) Using 2020 as the base year, calculate the price index for 2023.

(2)

.....

.....

.....

Answer .....

(b) Write down what your answer to part (a) tells you about the change in the price of the season ticket between 2020 and 2023.

(1)

.....

.....

(c) Taking 2020 as the base year, the price index for 2024 is 125. Work out the price of the season ticket in 2024.

(2)

.....

.....

.....

Answer ..... £

(d) The Retail Prices Index (RPI) is a weighted index. Give one reason why a weighted index is used rather than a simple index.

(1)

.....

.....

(Total for Question 8 is 6 marks)

9 (a) Two judges each rank 6 competitors in a talent competition. Spearman's rank correlation coefficient for their two sets of rankings is 0.85. Interpret this value in the context of the judges' rankings.

(2)

.....

.....

.....

(b) For two quantitative variables, a researcher calculates a product moment correlation coefficient (PMCC) of -0.92. Describe the correlation between the two variables.

(2)

.....

.....

.....

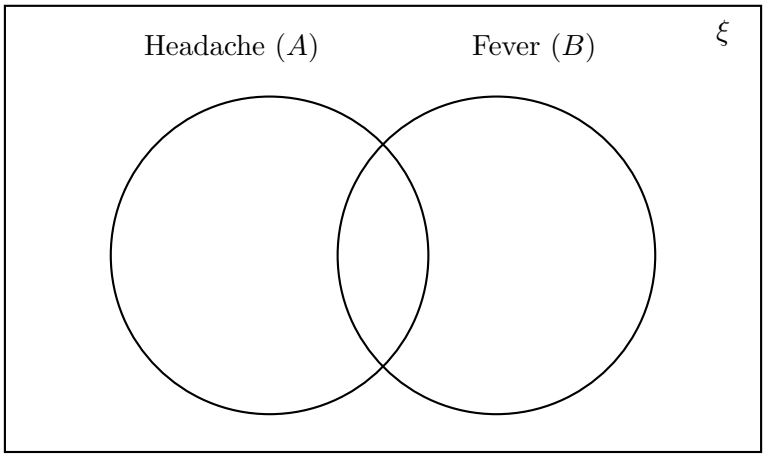
(Total for Question 9 is 4 marks)

10 100 people were asked whether they had experienced a headache or a fever during one winter.

- 58 people reported having had a headache (event A).
- 44 people reported having had a fever (event B).
- 35 people reported having had neither a headache nor a fever.

(a) Use this information to complete the Venn diagram below.

(3)



(b) One of the 100 people is chosen at random. Find $\mathrm{P}(A \cup B)$.

(2)

.....

.....

Answer .....

(c) Write down the absolute risk of a person in this group having a headache.

(1)

.....

(d) Work out the relative risk of having a fever for those who had a headache, compared with those who did not have a headache.

(3)

.....

.....

.....

Answer .....

(Total for Question 10 is 9 marks)

11 Maria has a spinner with the numbers 1 to 6 on it. She spins the spinner 20 times. It lands on the number 6 a total of five times. Maria says,

"This shows the spinner is definitely biased towards the number 6."

Comment on the reliability of Maria's conclusion.

(2)

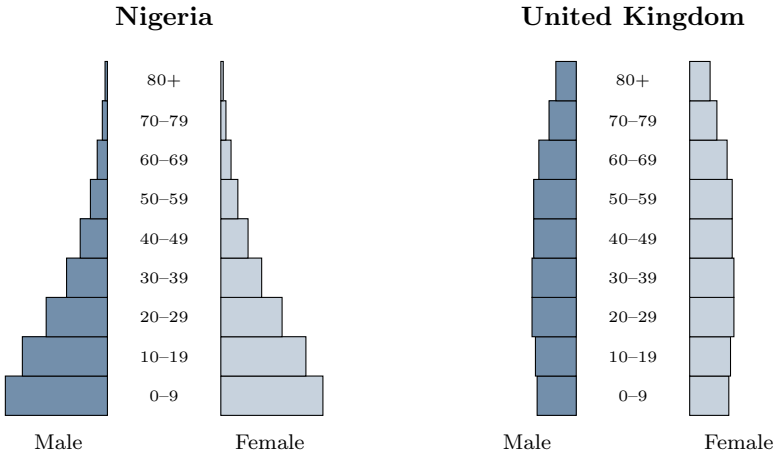
.....

.....

.....

(Total for Question 11 is 2 marks)

12 The two population pyramids below show the age distributions of the populations of Nigeria and the United Kingdom.



Horizontal axis: percentage of the population in each age band

Compare what the shapes of the two pyramids suggest about the birth rate and the life expectancy of the two populations.

(3)

.....

.....

.....

.....

.....

(Total for Question 12 is 3 marks)

13 A teacher records the number of hours, x, that each of eight students spent revising for a test, and the score, y%, that each student achieved in the test. The results are shown in the table.

Revision time, x (hours)	1	2	4	4	5	6	6	8
Test score, y (%)	42	36	64	46	58	74	56	80

(a) Calculate the coordinates of the mean point $(\bar{x}, \bar{y})$.

(2)

.....

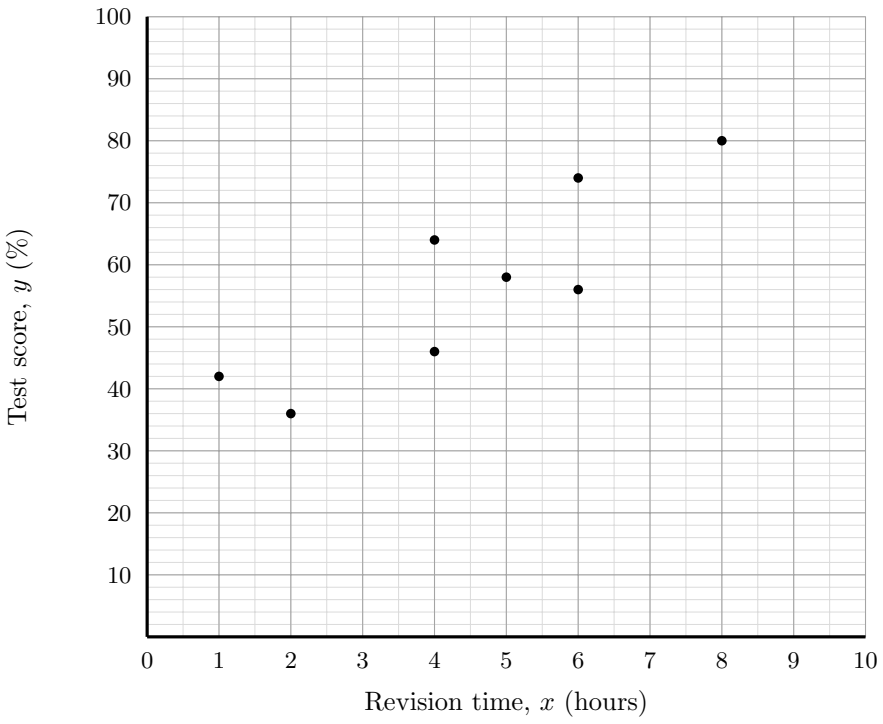
.....

.....

Answer .....

(b) The eight students' results have been plotted on the scatter graph below. Plot your mean point and draw a line of best fit that passes through it.

(1)



(c) Find the equation of your line of best fit. Give your answer in the form $y = a + bx$.

(2)

.....

.....

.....

.....

Answer .....

(d) Mia revised for 12 hours. Explain why using your line of best fit to estimate Mia's test score may not give a reliable result.

(1)

.....

.....

.....

(Total for Question 13 is 6 marks)

TOTAL FOR PAPER = 80 MARKS

END