Paper Reference 1ST0/1H
Pearson Edexcel
Level 1/Level 2 GCSE (9–1)

Total Marks

Statistics

Paper 1
(Calculator)
Higher Tier

Thursday 13 June 2019 – Afternoon

Time: 1 hour 30 minutes plus your additional time allowance.

In the boxes below, write your name, centre number and candidate number.

Surname					
Other names					
Centre Number					
Candidate Number					

Y60681A

Pearson

YOU MUST HAVE

Ruler, protractor, compasses, writing and drawing equipment, scientific calculator.

YOU WILL BE GIVEN

Data Book
Formulae Pages

INSTRUCTIONS

Answer ALL questions.

Answer the questions in the spaces provided in this Question Paper or on the separate data sheets – there may be more space than you need.

Scientific calculators may be used.

You must NOT write anything on the Formulae Pages. Anything you write on the Formulae Pages will gain NO credit.

You must show all your working out with your answer clearly identified at the end of your solution.

INFORMATION

The total mark for this paper is 80

The marks for EACH question are shown in brackets – use this as a guide as to how much time to spend on each question.

There may be spare copies of some data sheets.

ADVICE

Read each question carefully before you start to answer it.

Try to answer every question.

Check your answers if you have time at the end.

Answer ALL questions.

Write your answers in the spaces provided.

You must write down all the stages in your working.

1. Look at the diagram for Question 1 in the Data Book.

   It shows an incomplete choropleth map.

   The diagram on the next page represents a children's playground that has been divided into 20 squares of equal area.

   In the playground there are some children and some play equipment only.

   The number of children in each square at 11 am one Saturday is shown on the next page.

(continued on the next page)

1. continued.

Key:

9	means 9 children in this square.

Number of children

11	10	7	5	0
9	7	6	3	1
8	4	3	1	1
5	4	2	0	1

(continued on the next page)

1. continued.

(a) Use the information on the previous page to complete the choropleth map in the Data Book. There are eight spaces to fill.
(2 marks)

(continued on the next page)

1. continued.

Grace concludes that there is likely to be more play equipment in that part of the playground represented by the squares in the top left hand corner of the choropleth map than elsewhere in the playground.

(b) Assess the validity of Grace's conclusion with reference to the choropleth map.

(1 mark)

(Total for Question 1 is 3 marks)

2. Look at the diagram for Question 2 in the Data Book.

The cumulative frequency diagram gives information about the heights, in metres, of a sample of 100 oak trees in Camden, London.

(a) Using the cumulative frequency diagram, complete the table below for the heights of these 100 trees.

Lower quartile	Median	Upper quartile

(2 marks)

(continued on the next page)

2. continued.

Look at the diagram for Question 2(b) in the Data Book.

The box plot shows information about the heights, in metres, of a sample of maple trees in Camden, London.

For the sample of oak trees
the least height is $2 \cdot 5$ metres
the greatest height is $22 \cdot 5$ metres

(b) On the grid, draw a box plot for the heights of the sample of oak trees.
(2 marks)

(continued on the next page)

2. continued.

(c) Compare the two distributions of heights.

Give THREE comparisons and interpret one of these comparisons.

(4 marks)

Answer lines continue on the next page.

2. (c) continued.

(Total for Question 2 is 8 marks)

3. Look at the table for Question 3 in the Data Book.

It gives information about the numbers of students from different types of schools who applied to Cambridge University in **2016**

Richard is going to take a sample of **200** of these students stratified by gender.

(continued on the next page)

3. continued.

(a) Work out how many female students there should be in this sample.

(2 marks)

(continued on the next page)

3. continued.

(b) Describe a situation when it would NOT be appropriate to take a sample stratified by gender.

(1 mark)

(continued on the next page)

3. continued.

Richard could have used a different category for his stratified sample.

(c) What is this different category?

(1 mark)

_____

_____

(continued on the next page)

3. continued.

A student is to be chosen at random from the 9963 students.

F is the event that the student chosen is female.

I is the event that the student chosen is from an independent school.

M is the event that the student chosen is from a maintained school.

(continued on the next page)

3. continued.

(d) Explain why the event F and the event I are NOT mutually exclusive.

(1 mark)

(continued on the next page)

3. continued.

(e) Find $P(I \text{ or } M)$

(2 marks)

(Total for Question 3 is 7 marks)

4. Diana is a journalist working for a local newspaper.

She is writing a newspaper article about how house prices in the local area have changed.

Diana has house price data for **1996** and for **2016**

She plans to include in her article the median house price for **1996** and the median house price for **2016**

(continued on the next page)

4. continued.

Mika thinks that Diana should also include in her article the interquartile range of house prices for **1996** and the interquartile range of house prices for **2016**

(a) Give one reason why including the interquartile ranges in the article would be appropriate.
(1 mark)

(continued on the next page)

4. continued.

(b) Give one reason why including the interquartile ranges in the article would NOT be appropriate.
(1 mark)

(Total for Question 4 is 2 marks)

5. Look at the table for Question 5 in the Data Book.

Chris is carrying out an experiment to see if left handed people have a better memory than right handed people.
He has two groups of people with the same number of people in each group.

Group A contains left handed people only and group B contains right handed people only.

Each person is given 15 objects to memorise and then they are tested to see how many objects they remember.

(continued on the next page)

5. continued.

The variables for the experiment are shown in the table.

(a) For each variable, put a mark in the correct column of the table to show the type of variable.
(2 marks)

(continued on the next page)

5. continued.

Chris is not sure whether to work out the mean or the median number of objects each group remembered.

(b) Describe a situation for which it would be more appropriate to work out the median than the mean.

(1 mark)

_____

_____

(continued on the next page)

5. continued.

Chris worked out these summary statistics for the number of objects people in group A remembered.

Lower quartile = 5

Median = 7

Upper quartile = 8

Sanjit is a member of group A

He remembered 14 objects.

(continued on the next page)

5. continued.

(c) Determine whether or not the number of objects Sanjit remembered is an outlier.

(2 marks)

(Total for Question 5 is 5 marks)

6. Look at the two tables for Question 6 in the Data Book.

Table 1 gives the crude birth rates and crude death rates, per **1000** people, for two Caribbean islands in the year **2015**

Jamil concludes that the crude birth rate and the crude death rate for Saint Lucia show that the population of Saint Lucia increased in **2015**

(continued on the next page)

6. continued.

(a) Explain how the data in Table 1 can be used to support Jamil's conclusion.

(1 mark)

(continued on the next page)

6. continued.

(b) Give a reason why the data might NOT support Jamil's conclusion.

(1 mark)

(continued on the next page)

6. continued.

In 2015, the total population of Barbados was 284 217

(c) Using the formula below, work out the number of births in Barbados in 2015

$$\text{crude birth rate} = \frac{\text{number of births} \times 1000}{\text{total population}}$$

(2 marks)

Answer space continues on the next page.

6. (c) continued.

(continued on the next page)

6. continued.

Look at Table 2 below Table 1 for Question 6 in the Data Book.

The standard population of Barbados and of Saint Lucia for three different age groups in **2015** is shown in Table 2

(d) Give an interpretation of the number **64** in the table.

(1 mark)

_____

_____

_____

(continued on the next page)

6. continued.

(e) Compare the crude death rate for Barbados with the crude death rate for Saint Lucia.

By referring to the information in the Standard population table, suggest a reason for your comparison.

(2 marks)

Answer lines continue on the next page.

6. (e) continued.

(Total for Question 6 is 7 marks)

7. Look at the diagram for Question 7 in the Data Book.

The two pie charts show information about the numbers of visitors to the UK for the given reasons in January 2017 and in July 2017

(continued on the next page)

7. continued.

(a) The pie charts do NOT show that there were more visitors to the UK on business in January 2017 than in July 2017

Explain why.

(1 mark)

(continued on the next page)

7. continued.

The number of visitors to the UK in January 2017 in order to visit friends or relatives is 1 080 733

(b) Work out the number of visitors to the UK in January 2017 to have a holiday.

(2 marks)

Answer space continues on the next page.

7. (b) continued.

(continued on the next page)

7. continued.

Look at the table for Question 7(c) in the Data Book.

It gives information about the total number of visitors to the UK in January 2017 and in July 2017

Ruth thinks there is a more appropriate way to draw pie charts now that she knows the information in the table.

(continued on the next page)

7. continued.

(c) Explain, giving reasons, how she should do this.

You must refer to the information in the table in your explanation.

(5 marks)

Answer lines continue on the next page.

7. (c) continued.

(Total for Question 7 is 8 marks)

8. Look at the table for Question 8 in the Data Book.

It shows some of the expected percentile weights (kg) for babies between 3 and 6 months old, based on data collected from the World Health Organisation.

(continued on the next page)

8. continued.

Antonia says,

"An estimate of the expected 30th percentile for a 3 month old boy is 5·8 kg"

(a) Explain why this is NOT a good estimate.

(1 mark)

(continued on the next page)

8. continued.

(b) Compare the expected 9th to the 91st interpercentile range for 5 month old girls with the expected 9th to the 91st interpercentile range for 5 month old boys.

You must show your working.

(3 marks)

The answer space and answer lines continue on the next page.

8. (b) continued.

(continued on the next page)

8. continued.

(c) Give an interpretation of your comparison in part (b)

(1 mark)

(Total for Question 8 is 5 marks)

9. Look at the information for Question 9 in the Data Book.

According to an internet site, an estimate of the number of reindeer in a region of Ontario is **5000**

(Source: www.ontario.ca)

Giovani wants to verify this estimate. His method and conclusion are shown in the Data Book.

(continued on the next page)

9. continued.

Discuss the appropriateness of Giovani's method and of his conclusion.

As part of your discussion you should show your calculations and state any assumptions made.

(5 marks)

Answer lines continue on the next two pages.

9. continued.

9. continued.

(Total for Question 9 is 5 marks)

10. Look at the table for Question 10 in the Data Book.

It shows information about the amount of time that each member of a group of 46 teenagers spent on social media during one day.

(a) (i) Use linear interpolation to find an estimate of the median time spent on social media by the 46 teenagers.

(3 marks)

Answer space continues on the next page.

10. (a) (i) continued.

_____ minutes

(continued on the next page)

10. (a) continued.

The average person will spend approximately **116** minutes on social media each day.
(Source: www.socialmediatoday.com)

(ii) Compare the amount of time spent on social media by the **46** teenagers with the amount of time spent on social media by the average person.
(1 mark)

_____

_____

_____

(continued on the next page)

10. continued.

Look at the table for Question 10(b) in the Data Book.

It gives the mean, the standard deviation and the median for the times spent on social media during one day by a sample of sixty year olds.

(b) Calculate the skew of the times spent on social media by the sample of sixty year olds.

(2 marks)

Answer space continues on the next page.

10. (b) continued.

(continued on the next page)

10. continued.

(c) Interpret your answer to part (b)

(2 marks)

(Total for Question 10 is 8 marks)

11. Some students at a school walk home and some students go home by bus.

The times taken by the students at the school to walk home have a mean of 25 minutes and a standard deviation of 5 minutes.

The times taken to walk home can be modelled by a normal distribution.

(continued on the next page)

11. continued.

(a) Shanaya says,

"More than 80% of the students who walk home take between 20 and 35 minutes"

Use statistical calculations to assess Shanaya's conclusion.

(5 marks)

Answer space continues on the next two pages.

11. (a) continued.

11. (a) continued.

(continued on the next page)

11. continued.

Look at the diagram for Questions 11(b) and (c) in the Data Book.

It shows a sketch of the distribution of the times taken by students to go home by bus.

(continued on the next page)

11. continued.

(b) Explain why it is not appropriate to use a normal distribution to model the times taken by students to go home by bus.

(1 mark)

(c) On the same grid, sketch a diagram showing the distribution of the times taken by students to walk home.

(2 marks)

(continued on the next page)

11. continued.

(d) A is the event that a student walks to school, where $P(A) = 0 \cdot 6$

B is the event that a student is driven to school, where $P(B) = 0 \cdot 3$

Some students are driven some of the way to school and then they walk the rest of the way to school, where $P(A \text{ and } B) = 0 \cdot 15$

(continued on the next page)

11. (d) continued.

A student is picked at random from the students at the school.

Work out $P(A \text{ or } B)$

(2 marks)

(Total for Question 11 is 10 marks)

---

12. Look at the table for Question 12 in the Data Book.

Mr Singh collected data about the number of students in each of three classes in his school and the mean mark of each class in a science test.

The table gives some information about his data.

(continued on the next page)

12. continued.

Mr Singh plans to use one of the following two methods to work out the mean mark of ALL the students in the three classes.

Method 1 Work out the mean of 63, 72 and 55

Method 2 Given that Mr Singh knows the value of n, work out the weighted mean mark for the three classes.

(continued on the next page)

12. continued.

(a) For each of these two methods, assess whether or not the method is an appropriate way to work out the mean mark of all the students in the three classes.
(2 marks)
Answer lines continue on the next page.

Method 1

12. (a) continued.

Method 2

(continued on the next page)

12. continued.

The weighted mean mark for the three classes is $64 \cdot 1$ correct to one decimal place.

(b) Calculate the value of n

(2 marks)

(continued on the next page)

12. continued.

Look at the information for Question 12(c) in the Data Book.

Mr Singh thinks that he has made a mistake with one of the equations.

(c) Compare the equations of the three regression lines and explain which one is most likely to be incorrect.
Explain your answer in context.
(2 marks)
Answer lines continue on the next page.

12. (c) continued.

(continued on the next page)

12. continued.

Look at the diagram for Question 12(d) in the Data Book.

It shows a scatter diagram.

Mr Singh decides to compare the science marks, x and the history marks, y, of class A

He uses statistical software to draw the scatter diagram and the regression line shown in the Data Book.

(continued on the next page)

12. continued.

(d) Find an equation of the regression line of this scatter graph in the form $y = a + bx$

(3 marks)

(Total for Question 12 is 9 marks)

13. Look at the information for Question 13 in the Data Book.

A company wants to investigate the number of sick days its employees have off work.

The company uses a questionnaire.

One of the questions on the questionnaire is shown in the Data Book.

(continued on the next page)

13. continued.

(a) Assess the appropriateness of the method the company uses. (2 marks)

(continued on the next page)

13. continued.

All the company's employees completed the question.

615 ticked box **A**

102 ticked box **B**

(b) Show that an estimate of the number of employees who ticked box **A** because they answered yes to the question is **137**

(1 mark)

Answer space is on the next page.

13. (b) continued.

(Total for Question 13 is 3 marks)

TOTAL FOR PAPER IS 80 MARKS

END OF PAPER