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

Total Marks

Statistics
PAPER 2
Higher Tier
(Calculator)

Monday 17 June 2024 – Afternoon

Time: 1 hour 30 minutes

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

Surname						
Other names						
Centre Number						
Candidate Number						

V75439A

Pearson

YOU MUST HAVE

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

YOU WILL BE GIVEN

Data Booklet

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 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 in case you need them.

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. Matthew is investigating average household income for different states in the USA.

(a) Give a reason why it is appropriate to use secondary data for this.
(1 mark)

(continued on the next page)

1. continued.

Look at the diagram for Question 1(b) and 1(c) in the Data Booklet.

Matthew creates the choropleth map in the Data Booklet giving information about the mean household income by state for the USA in 2023

(b) Which THREE states have the lowest mean household income? (1 mark)

(continued on the next page)

1. continued.

Matthew concludes that the mean household incomes are highest on the West coast and the East coast.

(c) Does the choropleth map support this conclusion?

Give a reason for your answer.

(2 marks)

(Total for Question 1 is 4 marks)

2. Look at the diagram for Question 2(a), 2(b) and 2(c) in the Data Booklet.

It shows an incomplete box plot diagram.

Some researchers investigated the hand span, in centimetres, of adult pianists by their level – international, national and amateur.

The box plots in the Data Booklet give information about the hand spans for national level and amateur level pianists.

(continued on the next page)

2. continued.

(a) Circle the word in the list below that describes hand span, in centimetres, as a type of data. (1 mark)

qualitative

ordinal

continuous

bivariate

(continued on the next page)

2. continued.

The table below gives information about the hand spans of the international level pianists.

Greatest hand span	27·5 cm
Median hand span	24·0 cm
Lower quartile	23·0 cm
Range	5·5 cm
Interquartile range	1·5 cm

(continued on the next page)

2. continued.

(b) Using the information in the table, draw on the grid in the Data Booklet a box plot for the hand spans of the international level pianists.

(3 marks)

(continued on the next page)

2. continued.

(c) Compare the three distributions of hand spans.

Give THREE comparisons and interpret TWO of your comparisons.

(5 marks)

Answer lines continue on the next page.

2. (c) continued.

(continued on the next page)

2. continued.

Look at the table for Question 2(d) in the Data Booklet.

Pavel owns a music shop.

He wants to investigate the keyboard sizes used by pianists with different hand spans.

He collects data about the hand spans of the pianists who use his shop.

The table gives information about the number of these pianists with hand spans in each of four size categories.

(continued on the next page)

2. continued.

Pavel plans to sample 20 of these pianists stratified by hand span size.

(d) Explain how Pavel can obtain his stratified sample.

You should include details of any calculations he should use.

(3 marks)

Answer lines continue on the next page.

2. (d) continued.

(Total for Question 2 is 12 marks)

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

Khatia organises two different training courses, Course **A** and Course **B**, to help people to learn to type.

She wants to compare the two different courses to see which is better.

At the end of each course the people are given a skills test.

The table shows the number of participants who passed and failed the skills test for each of the two courses.

(continued on the next page)

3. continued.

(a) Find the relative risk of failing the skills test having taken Course **A** compared to Course **B**.

(3 marks)

(continued on the next page)

3. continued.

(b) Give an interpretation of your answer to part (a).

(1 mark)

(Total for Question 3 is 4 marks)

4. The snow depth, in centimetres, at Whistler Mountain was measured daily in February, in March and in April 2022

(continued on the next page)

4. continued.

The table below gives information about the snow depths for these 89 days.

Snow depth (x cm)	Cumulative frequency
$170 \leq x < 180$	15
$170 \leq x < 190$	40
$170 \leq x < 200$	55
$170 \leq x < 210$	75
$170 \leq x < 220$	80
$170 \leq x < 260$	89

(continued on the next page)

4. continued.

(a) Look at the diagram for Question 4(a) in the Data Booklet.

It shows a grid.

Draw a cumulative frequency diagram on the grid for this information.

(2 marks)

(continued on the next page)

4. continued.

(b) Find an estimate for the number of these days where the snow depth was between 195 cm and 215 cm

(2 marks)

(Total for Question 4 is 4 marks)

5. Look at the diagram for Question 5(a), 5(b) and 5(c) in the Data Booklet.

It shows an incomplete back–to–back stem and leaf diagram.

Zack is comparing the players from the England Rugby Union team with the players from the Welsh Rugby Union team.

The players are listed as Rugby Union Forwards or Rugby Union Backs.

(continued on the next page)

5. continued.

The weights are below, in kilograms, of the England Rugby Union team players who are listed as Backs.

96	112	94	87
93	88	90	92
98	96	96	82
107	111	84	88

(continued on the next page)

5. continued.

The back–to–back stem and leaf diagram gives the weights, in kilograms, of the Welsh Rugby Union Backs.

(a) Complete the back–to–back stem and leaf diagram in the Data Booklet with the weights, in kilograms, of the England Rugby Union Backs.

(3 marks)

(continued on the next page)

5. continued.

(b) Work out the median weight for the Welsh Rugby Union Backs.

(1 mark)

_____ kg

(continued on the next page)

5. continued.

(c) Work out the interquartile range for the Welsh Rugby Union Backs.

(2 marks)

_____ kg

(continued on the next page)

5. continued.

The table below gives information about the weights for the England Rugby Union Backs.

Median	Interquartile range
93·5 kg	9 kg

(continued on the next page)

5. continued.

Zack thinks that the Welsh Rugby Union Backs are heavier and have less consistent weights than the England Rugby Union Backs.

(d) Do the statistics support these conclusions?

You must give reasons for your answer.

(3 marks)

Answer lines continue on the next page.

5. (d) continued.

(continued on the next page)

5. continued.

Amy wants to use the median and interquartile range statistics in a news article for a sports magazine. The article will compare the players on the two teams who are Backs.

(e) Comment on the appropriateness of using the median and the interquartile range in the article. (1 mark)

(continued on the next page)

5. continued.

(f) Give a limitation of using Zack's statistics to compare ALL the players on the two teams.

(1 mark)

(Total for Question 5 is 11 marks)

6. Look at the table for Question 6 in the Data Booklet.

Researchers investigated the effect of changing land usage on the amount of surface runoff water for areas of land around Brandywine Creek.

They recorded the change in forest cover area (km^2) from 1992 to 2000

They also recorded the change in the amount of surface runoff (mm water) from 1992 to 2000

The table gives information about their data.

(continued on the next page)

6. continued.

(a) Calculate Spearman's rank correlation coefficient for the data in the Data Booklet. Interpret this correlation in the context of the question.

(6 marks)

Answer space and lines continue on the next page.

6. (a) continued.

(continued on the next page)

6. continued.

Elizabeth wants to investigate possible drainage solutions that reduce surface runoff.

She wants to establish whether a particular drainage solution reduces surface runoff.

Elizabeth plans to measure surface runoff in 10 areas for a year without the drainage solution.

She will then install the drainage solution in these 10 areas and measure the surface runoff for a year.

(continued on the next page)

6. continued.

(b) Explain how Elizabeth's plan to collect data controls some extraneous variables.

You should include in your answer an example of an extraneous variable that is likely to be controlled in this investigation.

(2 marks)

(continued on the next page)

6. continued.

Elizabeth would like to reduce the time that she is collecting data to one year overall.

(c) Describe how she could do this using a matched pairs approach. (1 mark)

(Total for Question 6 is 9 marks)

7. Look at the table for Question 7(a) and 7(b) in the Data Booklet.

The Consumer Price Index (CPI) is a measure of the rate of change of prices in everyday life.

The table shows the annual average CPI from 2017 to 2021 with 2015 as the base year.

(continued on the next page)

7. continued.

(a) Give an interpretation of the number $108 \cdot 9$ in the table.

(2 marks)

(continued on the next page)

7. continued.

(b) Find the percentage increase in consumer prices from 2017 to 2021

You must show your working.

Give your answer correct to one decimal place.

(2 marks)

(continued on the next page)

7. continued.

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

A representative sample of consumer product prices is included in the calculations for CPI.

In the non–alcoholic beverages category there are two classes of product included.

The table gives the weightings for these two classes and the price index for one of the classes for November 2022

(continued on the next page)

7. continued.

The price index for non–alcoholic beverages in November 2022 was 124·2

(continued on the next page)

7. continued.

(c) Calculate the price index for mineral waters, soft drinks and juices in November 2022

Give your answer correct to one decimal place.

(2 marks)

(continued on the next page)

7. continued.

Look at the table for Question 7(d) in the Data Booklet.

The weights in the non–alcoholic beverage category were different in 2021 and in 2022

The table gives these weights.

(continued on the next page)

7. continued.

(d) Explain what this means for the relative importance in the CPI of 'Coffee, tea and cocoa' and 'Mineral waters, soft drinks and juices' in 2021 compared with 2022

(1 mark)

(continued on the next page)

7. continued.

Look at the table for Question 7(e) in the Data Booklet.

The rate of inflation is the change in prices for goods and services over time.

David wants to compare annual inflation in Germany from 2017 to 2021 with annual inflation in the UK from 2017 to 2021

The table gives the annual inflation for the UK from 2017 to 2021

(continued on the next page)

7. continued.

From 2017 to 2021 the average annual inflation for Germany was 1·66%

(e) By calculating an appropriate geometric mean, compare the average annual inflation for Germany from 2017 to 2021 with the average annual inflation for the UK from 2017 to 2021

You must show your working.

(3 marks)

Answer space and lines are on the next page.

7. (e) continued.

(Total for Question 7 is 10 marks)

8. Look at the information for Question 8 in the Data Booklet.

Roxann wants to investigate the resting heart rates for members of her running club.

She believes that the resting heart rates will be normally distributed.

The information in the Data Booklet shows the plan for her investigation.

(continued on the next page)

8. continued.

Discuss whether Roxann's plans for collecting and presenting data are appropriate.

(6 marks)

Answer lines continue on the next three pages.

8. continued.

8. continued.

8. continued.

(Total for Question 8 is 6 marks)

9. Look at the table for Question 9 in the Data Booklet.

Researchers wanted to find a method to predict the height of ancient Egyptians based upon bones from their skeletons.

They calculated the Pearson's product moment correlation coefficient between the length of particular bones and the height of the skeleton for some male skeletons and for some female skeletons.

(continued on the next page)

9. continued.

They also found regression equations for the relationship between bone length (x centimetres) and height (y centimetres) for each of these bones in male skeletons and in female skeletons.

The table gives information about these product moment correlation coefficients and these regression equations.

(continued on the next page)

9. continued.

(a) Which bone measurement would you recommend using to estimate the height of an ancient Egyptian?

Give a reason for your answer.

(2 marks)

(continued on the next page)

9. continued.

(b) Interpret in context the figure **2·594** in the regression equation for the humerus of male skeletons.

(1 mark)

(continued on the next page)

9. continued.

(c) Use the regression equations to compare the relationships between bone length and height for the different bones.

Include in your answer comparisons between male and female ancient Egyptians.

(3 marks)

Answer lines continue on the next page.

9. (c) continued.

(continued on the next page)

9. continued.

Dina suggests comparing the y–intercepts for the regression equations.

(d) Is Dina's suggestion appropriate? (2 marks)

(continued on the next page)

9. continued.

A museum has some bones that were recovered from an Egyptian pyramid.

The museum wants to predict the height of the ancient Egyptians from whom the bones were recovered.

(continued on the next page)

9. continued.

(e) Is it appropriate to use these regression equations to make this prediction?

(2 marks)

(Total for Question 9 is 10 marks)

10. Researchers used the Petersen capture recapture method to estimate the number of humpback whales off the coasts of California, Oregon and Washington.

They used aerial photographs to identify the whales, recording the whales' unique markings in the first sampling rather than tagging the whales.

(continued on the next page)

10. continued.

In 1995 they recorded the markings of 331 individual whales.

In 1996 they identified 264 individual whales of which 104 had been identified in the previous year.

(continued on the next page)

10. continued.

(a) Calculate an estimate for the number of humpback whales off the coasts of California, Oregon and Washington in 1996 (2 marks)

(continued on the next page)

10. continued.

(b) Discuss the validity and reliability of this estimate.

(3 marks)

(Total for Question 10 is 5 marks)

11. Look at the diagram for Question 11(a), 11(b) and 11(c) in the Data Booklet.

It is a Venn diagram showing information about 50 books on a reading list.

A is the event that the book is a science fiction book.

B is the event that the book is an audio book.

The numbers in the Venn diagram show the number of books.

(continued on the next page)

11. continued.

One of the books is chosen at random.

(a) Find

P(A)

(1 mark)

(continued on the next page)

11. continued.

(b) Find

$$P(A|B)$$

(1 mark)

(continued on the next page)

11. continued.

Mike concludes that **A** and **B** are independent events.

(c) Is Mike correct?

Give a reason for your answer.

(1 mark)

(continued on the next page)

11. continued.

Two different events, C and D, are such that

$$P(C) = 0 \cdot 75$$

$$P(D) = 0 \cdot 4$$

$$P(C \cap D) = 0 \cdot 24$$

(continued on the next page)

11. continued.

(d) Find

$$P(C \cup D)$$

(2 marks)

(Total for Question 11 is 5 marks)

TOTAL FOR PAPER IS 80 MARKS

END OF PAPER

Sources

Question 4
(Source adapted from: aqrt.nrs.gov.bc.ca)

Question 5
(Source: www.englandrugby.com/england/senior–men#squad)

Question 10
(Source: digitalcommons.unl.edu)