

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

Total Marks

Statistics
PAPER 2
Foundation 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					

YOU MUST HAVE

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

YOU WILL BE GIVEN

Data Booklet

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.

Turn over

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. David has **10** cards each with a single letter on it as shown.

A	A	A	A	A
B	B	C	D	D

A card is picked at random.

- (a) Circle the word from the list below that best describes the likelihood that the card has a letter **A** on it.

(1 mark)

impossible

certain

likely

evens

unlikely

(continued on the next page)

Turn over

1. continued.

Remember the 10 cards:

A A A A A

B B C D D

**(b) Complete the sentence below using
two different letters.**

(1 mark)

**Cards with the letters _____ and
_____ are equally likely to be picked.**

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

It shows a probability scale.

**On the probability scale, mark the probability
that the card has a letter C on it.**

(1 mark)

(continued on the next page)

Turn over

1. continued.

Remember the **10** cards:

A A A A A

B B C D D

(d) Look at the diagram for Question 1(d) in the
Data Booklet.

It shows a probability scale.

On the probability scale, mark the probability
that the card has a letter **A** or a letter **C** on it.

(1 mark)

(Total for Question 1 is 4 marks)

- 2. Look at the diagram for Question 2 in the Data Booklet.**

It shows a bar chart.

Jenny is investigating how many days per week people use a gym.

She asks the 40 people in her fitness group how often they use the gym each week.

Jenny draws the bar chart shown in the Data Booklet for her data.

One of these people is chosen at random.

- (a) Find the probability that this person uses the gym exactly 2 days per week.**
(1 mark)
-

(continued on the next page)

Turn over

2. continued.

(b) What is the modal number of days to use the gym each week?

(1 mark)

(continued on the next page)

Turn over

2. continued.

Jenny thinks that there are a lot of people in her fitness group who are exercising less than 2 days per week as there is a total of 10 people who used the gym on 0 days or 1 day per week.

**(c) Explain why Jenny might NOT be correct.
(1 mark)**

(Total for Question 2 is 3 marks)

3. Ben is researching information about the number of British swimming medals won at the Olympics.

His results are below, giving the number of British swimming medals won at the Olympics from 1900 to 2016

3	0	7	6	2	4	4
2	0	1	1	2	3	1
1	1	3	5	5	3	1
2	0	2	3	3	6	

- (a) Look at the table for Question 3(a) in the Data Booklet.

Fill in the tally chart for Ben's results AND complete the frequency column in the Data Booklet.

(2 marks)

(continued on the next page)

Turn over

3. continued.

(b) Suggest a suitable diagram that could be used for Ben's results.

(1 mark)

(continued on the next page)

3. continued.

Remember Ben's results:

3	0	7	6	2	4	4
2	0	1	1	2	3	1
1	1	3	5	5	3	1
2	0	2	3	3	6	

(c) Write down the mode or modes.
(1 mark)

(d) Work out the median.
(2 marks)

(continued on the next page)

Turn over

3. continued.

Ben wants to use an average to summarise the data.

(e) Which of the mode or the median would be more appropriate?

Give a reason for your answer.

(2 marks)

(Total for Question 3 is 8 marks)

4. Tachi collects data on the heights, in metres, of a sample of Egyptian pyramids.

Her data is shown below.

136.4 101.1 104 62.6 138.8 65.5 93.5

Tachi picks one of these pyramids at random.

- (a) Find the probability that this pyramid will have a height of more than **100** metres.

(1 mark)

(continued on the next page)

4. continued.

Remember Tachi's Egyptian data:

136.4 101.1 104 62.6 138.8 65.5 93.5

The mean height of a sample of Mexican pyramids is **53.5** metres.

Tachi says, "On average these Egyptian pyramids are twice as high as the Mexican pyramids."

(b) Is Tachi correct?

You must show working to support your answer.

(4 marks)

Answer space and answer lines continue on the next page.

Turn over

4. (b) continued.

(continued on the next page)

Turn over

4. continued.

The range of heights for the Mexican pyramids is
45 metres.

The lowest height of the Mexican pyramids is
30 metres.

(c) Work out the greatest height of the
Mexican pyramids.

(1 mark)

_____ metres

(Total for Question 4 is 6 marks)

5. Look at the information for Question 5 in the Data Booklet.

Claire is planning an investigation into the length of time that a learner has to wait for a driving test.

She wants to find out about how waiting time varies in different regions of the UK.

The information in the Data Booklet shows her plan for data collection, for calculations and for diagrams.

Discuss whether Claire's plans for data collection, for calculations and for diagrams are appropriate.

(6 marks)

Answer lines continue on the next page.

5. continued.

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.

(Total for Question 5 is 6 marks)

Turn over

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

It shows an incomplete composite bar chart.

Ana collected information about the Scottish Football Championship results for the **2021/2022** season.

The composite bar chart gives some information about the number of points scored by six of the teams.

Points are scored from wins or draws.

Raith Rovers (**RR**) scored **30** points from wins and **15** points from draws.

- (a) Complete the composite bar chart for Raith Rovers (**RR**).
(2 marks)

(continued on the next page)

6. continued.

(b) What does the overall height of each bar represent?

(1 mark)

(continued on the next page)

6. continued.

- (c) Compare the points scored by
Dunfermline Athletic (**DA**) with the points
scored by Queen of the South (**QS**).
(3 marks)

(Total for Question 6 is 6 marks)

7. Chris is a manager at a theme park.

He wants to find out what food options visitors would like to be able to buy in the theme park.

**(a) State the population for this investigation.
(1 mark)**

(continued on the next page)

7. continued.

Chris decides that he will take a convenience sample of visitors in the section of the park selling food.

(b) Explain what is meant by a convenience sample.

(1 mark)

(continued on the next page)

7. continued.

(c) Give one disadvantage of using a convenience sample.

(1 mark)

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

It shows the data collection sheet Chris plans to use.

Discuss whether the data collection sheet is appropriate.

You should consider how Chris might use the data and describe any problems he might have when he uses the data collection sheet.

(2 marks)

Answer lines continue on the next page.

7. (d) continued.

(continued on the next page)

Turn over

7. continued.

Chris suggests using a stem and leaf diagram to represent the data that he collects.

(e) Discuss whether or not this would be a suitable diagram to represent his data.

(2 marks)

(Total for Question 7 is 7 marks)

8. Look at the diagram for Question 8(a), 8(b), 8(c) and 8(d) in the Data Booklet.

It shows an incomplete scatter diagram.

Timur is investigating the heights and weights of rugby players.

He collected data from the internet about the heights and weights of players from the Wales rugby squad.

The players from the Wales rugby squad are classed as Forwards or Backs.

Timur draws a scatter diagram for the heights and weights of some of the Backs from the Wales rugby squad.

- (a) Explain why a scatter diagram is appropriate for the type of data Timur collected.

(1 mark)

(continued on the next page)

Turn over

8. continued.

The data for three of the players was not plotted on the scatter diagram.

The height and weight of each of these players is given in the table below.

Player	A	B	C
Height (cm)	182	188	173
Weight (kg)	95	100	80

(b) Complete the scatter diagram by plotting the points for players **A**, **B** and **C**.
(2 marks)

(continued on the next page)

8. continued.

(c) Describe and interpret the type of correlation shown by the scatter diagram.

(2 marks)

The double mean point for the heights and weights of all of the Backs is (184, 92)

(d) Draw a line of best fit through the double mean point.

(2 marks)

(continued on the next page)

8. continued.

Timur also collected the heights and weights of players from the England rugby squad.

He found the Spearman's rank correlation coefficient for the heights and weights of the Forwards from the England rugby squad.

The correlation coefficient was 0·00

(e) Interpret this correlation coefficient in context.
(1 mark)

(continued on the next page)

8. continued.

Look at the table for Question 8(f) and 8(g) in the Data Booklet.

Timur used statistical software to find the information shown in the table about the heights (x cm) and weights (y kg) of the Backs from the Wales rugby squad and the Backs from the England rugby squad.

- (f) Compare the Spearman's rank correlation coefficients and interpret this comparison in the context of the question.**
(2 marks)

(continued on the next page)

Turn over

8. continued.

Timur uses the information in the table to conclude that the weight of the England rugby squad Backs increases faster than the weight of the Wales rugby squad Backs as their height increases.

(g) Assess the validity of Timur's conclusion with reference to the statistical results.

(2 marks)

(Total for Question 8 is 12 marks)

Turn over

- 9. Look at the table for Question 9 in the Data Booklet.**
- Mobeen is investigating whether there is a difference in the amount of time spent reading by pupils in Green Park school and pupils at Golden Plains school.**

He uses a census of all of the pupils at each school.

Each pupil is asked to record the amount of time spent reading in a week.

Mobeen then collects this information from each student through an online database.

Part of the database is shown in the Data Booklet.

(continued on the next page)

9. continued.

(a) Give TWO reasons why the data should be cleaned before processing.

(2 marks)

(continued on the next page)

9. continued.

**Mobeen wants to compare the data for
Green Park school with the data for
Golden Plains school.**

**Once the data has been cleaned Mobeen plans to
use all of the times to draw a single box plot.**

**(b) Explain why this is NOT an appropriate thing
to do.**

(1 mark)

(Total for Question 9 is 3 marks)

10. 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)

10. continued.

Look at the diagram for Question 10(b) and 10(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)

10. 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 10 is 4 marks)

- 11. Look at the diagram for Question 11(a), 11(b) and 11(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.

- (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)

Turn over

11. 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

(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)

Turn over

11. continued.

- (c) Compare the three distributions of hand spans.
Give THREE comparisons and interpret TWO of
your comparisons.
(5 marks)**

(continued on the next page)

Turn over

11. continued.

Look at the table for Question 11(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.

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

(continued on the next page)

11. continued.

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

You should include details of any calculations he should use.

(3 marks)

(Total for Question 11 is 12 marks)

12. Look at the table for Question 12 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)

12. 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)

Turn over

12. continued.

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

(1 mark)

(Total for Question 12 is 4 marks)

13. Look at the diagram for Question 13 in the Data Booklet.

It is a cumulative frequency graph giving information about the ages, in years, of the 600 members of the UK Parliament in 2017

- (a) Using the cumulative frequency diagram, find an estimate of the median age for the members of the UK Parliament in 2017**
(1 mark)

_____ years

(continued on the next page)

Turn over

13. continued.

In 2017 the German Parliament had 51·6% of members of Parliament aged between 45 and 60 years old.

(b) Compare this figure to the percentage of members of the UK Parliament with ages between 45 and 60 years in 2017

(4 marks)

Answer space and answer lines continue on the next page.

13. (b) continued.

(Total for Question 13 is 5 marks)

TOTAL FOR PAPER IS 80 MARKS

END OF PAPER

Sources

Question 3

(Source: www.teamgb.com)

Question 4

(Source: www.rankred.com)

Question 13

(Source: <https://www.dw.com/en/germanys-new-bundestag-who-is-who-in-parliament/a-41082379>)