

Write your name here

Surname

Other names

Centre Number

Candidate Number

**Pearson Edexcel**  
**Level 1/Level 2 GCSE (9–1)**

# Statistics

## Paper 1

**Higher Tier**

Specimen Papers Set 1

**Time: 1 hour 30 minutes**

Paper Reference

**1ST0/1H**

**You must have:**

Ruler graduated in centimetres and millimetres, protractor, pen,  
HB pencil, eraser, scientific calculator.

Total Marks

### Instructions

- Use **black** ink or ball-point pen.
- **Fill in the boxes** at the top of this page with your name, centre number and candidate number.
- Answer **all** questions.
- Answer the questions in the spaces provided – *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.*

### 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.

Turn over ►

S59307A

©2018 Pearson Education Ltd.

6/6/4/6/1/



  
**Pearson**

## Higher Tier Formulae

You must not write on this page.

Anything you write on this page will gain NO credit.

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

$$\text{Standard deviation} = \sqrt{\frac{1}{n} \sum (x - \bar{x})^2}$$

*An alternative formula for standard deviation is*

$$\text{standard deviation} = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

Spearman's rank correlation coefficient

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

Rates of change (e.g. Birth rate =  $\frac{\text{number of births} \times 1000}{\text{total population}}$  )

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA



DO NOT WRITE IN THIS AREA

Answer ALL questions.

Write your answers in the spaces provided.

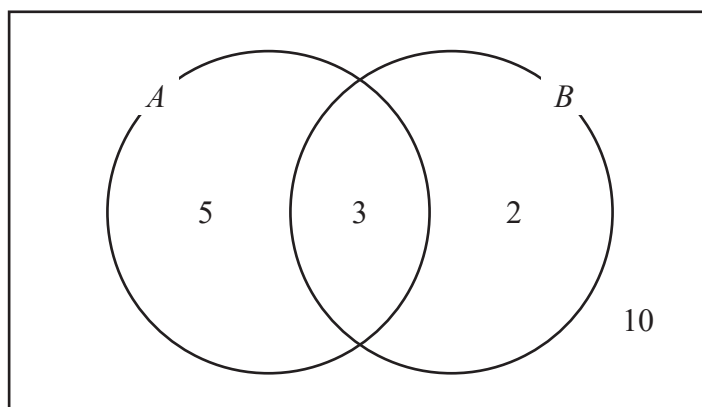
You must write down all the stages in your working.

1 The Venn diagram shows information about 20 films shown in the UK in 2015

$A$  is the event that the film was produced in the UK.

$B$  is the event that the film made more than £40 million.

The numbers in the Venn diagram indicate the number of films.



(Source: BFI Statistical Yearbook)

(a) Explain fully what the number 3 represents in the Venn diagram.

(1)

One of the films is chosen at random.

(b) Find  $P(B)$

(1)

(c) Find  $P(B | A)$

(2)

(d) Using your answers to part (b) and part (c), explain whether or not  $A$  and  $B$  are independent events.

(2)

(Total for Question 1 is 6 marks)



- 2 There are 11 727 students at a university.  
Their nationality is classified as UK, EU or International.

The table shows information about the nationality of these students.

Nationality	UK	EU	International	Total
Number of students	9393	979	1355	11 727

(Source: *www.ox.ac.uk*)

The manager of a book shop wants to carry out a survey into the books read by the students at this university.

She is going to take a sample of 600 of these students.

The manager plans to sample 200 UK students, 200 EU students and 200 International students.

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

.....  
(1)

- (b) Give a reason why this method of sampling might **not** be appropriate.

.....  
.....  
(1)

A shop assistant suggests that it would be better to select a sample of 600 students, stratified by nationality.

- (c) Work out how many EU students there should be in this sample.

.....  
(2)



DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Three students, Amy, Beth and Carlos, bought the same eight books from the book shop.

The manager asked Amy, Beth and Carlos to rank these eight books in increasing order of how much they each enjoyed them.

The manager calculated the Spearman's rank correlation coefficient for the ranks given by Amy and Beth.

She got a result of 1.2

(d) Explain how you know that this result is not correct.

(1)

The manager also calculated the Spearman's rank correlation coefficient for the ranks given by Beth and Carlos.

She got a result of 0.74

(e) (i) What type of correlation is shown by this result?

(ii) Interpret this result.

(2)

**(Total for Question 2 is 7 marks)**



3 The table shows a summary of the results of a long jump competition.

Distance jumped ( $x$ m)	Frequency ( $f$ )		
$7.1 < x \leq 7.3$	1		
$7.3 < x \leq 7.5$	2		
$7.5 < x \leq 7.7$	5		
$7.7 < x \leq 7.9$	9		
$7.9 < x \leq 8.1$	8		
$8.1 < x \leq 8.3$	3		

(Source: IAAF)

- (a) Find an estimate of the mean distance jumped.  
You may use the extra columns in the table.

..... m  
(3)

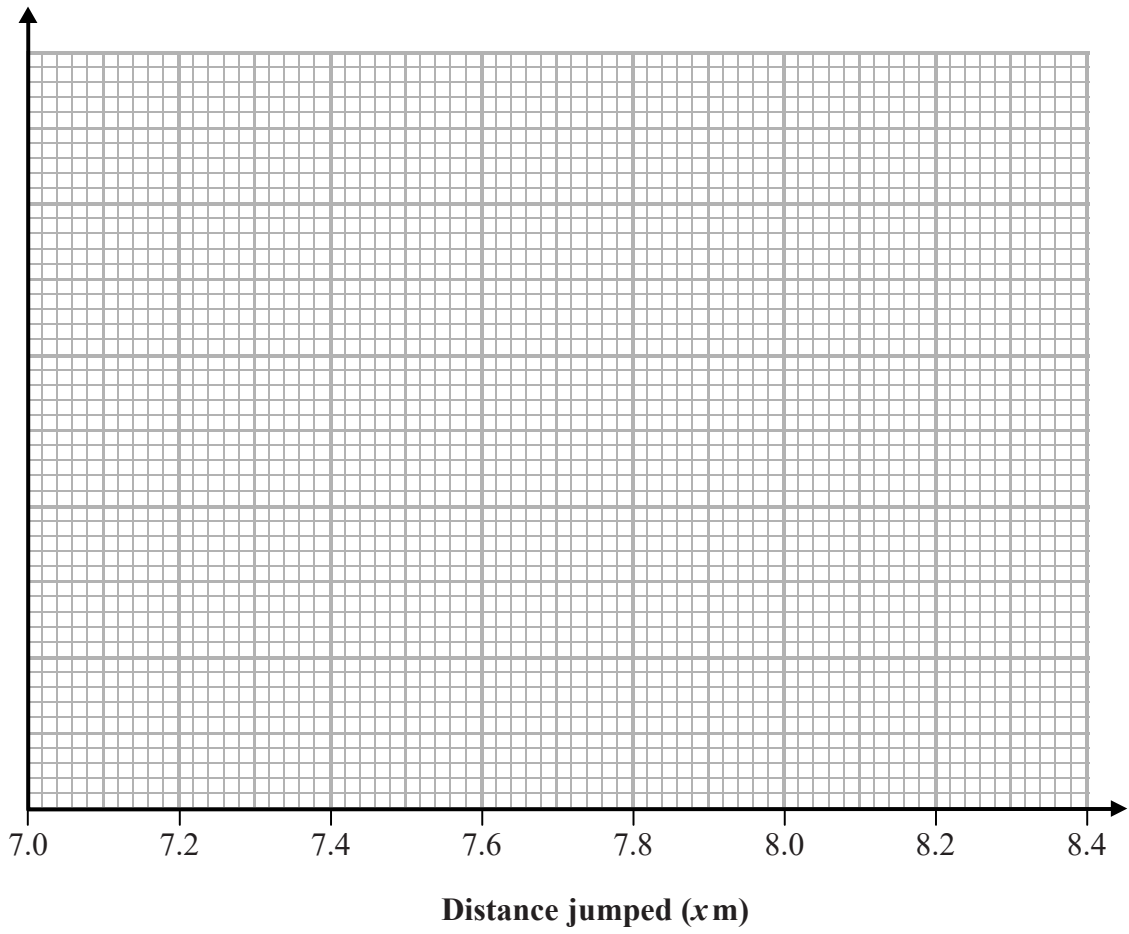


DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

(b) Draw a frequency polygon for the data.



(3)

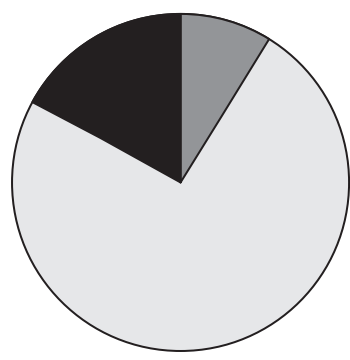
(c) Describe any skew shown by the data.

(1)

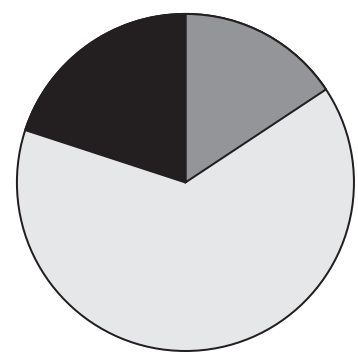
(Total for Question 3 is 7 marks)



4 The pie charts show information about the departure country of the overseas visitors to the UK in January 2016 and in July 2016



January 2016



July 2016

**Key:**

- North America
- Europe
- Other

(Source: *International Passenger Survey*)

Based on the two pie charts, Raul makes these two conclusions.

- 1 There are more visitors from Europe than from North America in both months.
- 2 There are more visitors from Europe in January than in July.

(a) Assess the validity of each of Raul's conclusions.

1 .....

.....

2 .....

.....

(2)

Judith decides to replace the pie charts with comparative pie charts.

(b) Explain how comparative pie charts are more appropriate than pie charts to compare the number of visitors in each month.

.....

.....

(1)





DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

The total number of overseas visitors to the UK in January 2016 was 2 599 000

The total number of overseas visitors to the UK in July 2016 was 3 786 000

Judith's comparative pie chart for January 2016 has radius 3 cm.

- (c) Calculate the radius of Judith's comparative pie chart for July 2016  
Give your answer correct to 1 decimal place.

..... cm

(2)

**(Total for Question 4 is 5 marks)**





DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

6 A large number of runners took part in a 5 km race.

Here are the results of some of the runners.

Runner	Race time
A	16 minutes 24 seconds
B	22:31
C	17.34
D	21:03
E	19.2 minutes
F	30,57
G	24:45

(Source: Parkrun)

Maya wants to represent the results of all of the runners using a stem and leaf diagram.

(a) Discuss two things Maya must do to the data to be able to represent it in a stem and leaf diagram.

.....

.....

.....

.....

(2)

(b) Explain whether or not it would be appropriate to represent the results of all of the runners in the race using

(i) a scatter diagram,

.....

.....

(ii) a histogram,

.....

.....

(iii) a time series graph.

.....

.....

(3)

(Total for Question 6 is 5 marks)



- 7 The table gives the simple index numbers, with 2013 as base year, for the average prices of houses in Scotland and in Wales for the years 2014 to 2017

The table also gives the chain base index numbers for the same information.

	Year	2014	2015	2016	2017
<b>Scotland</b>	Simple index number	103.6	110.1	111.7	113.5
	Chain base index number	103.6	106.3	101.5	101.6
<b>Wales</b>	Simple index number	103.0	106.7	109.7	115.0
	Chain base index number	103.0	103.6	102.8	104.8

(Source: HM Land Registry)

- (a) Give an interpretation of the simple index number 103.6 for Scotland.

(2)

David bought his house in Wales for £165 000 in 2013

- (b) (i) Use an appropriate index number to calculate an estimate for the price of his house in 2017

£.....

- (ii) Comment on the reliability of this estimate.

(3)



(c) By finding appropriate geometric means, compare and interpret the yearly average change in house prices from 2013 to 2017 in Scotland and in Wales.

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

(5)

(Total for Question 7 is 10 marks)



8 Stacey is the manager of a large company.

She wants to investigate the differences between the numbers of days of holiday taken by employees in different age groups.

She obtains the following information about the ages of the employees.

Age group	Number of employees	Number of males	Number of females
20 to 34	106	48	58
35 to 49	135	73	62
50 to 65	84	61	23

Assess the suitability of taking a sample of 20 employees stratified by age and by gender, for her investigation.

.....

.....

.....

.....

.....

.....

.....

.....

**(Total for Question 8 is 3 marks)**



DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

**BLANK PAGE**

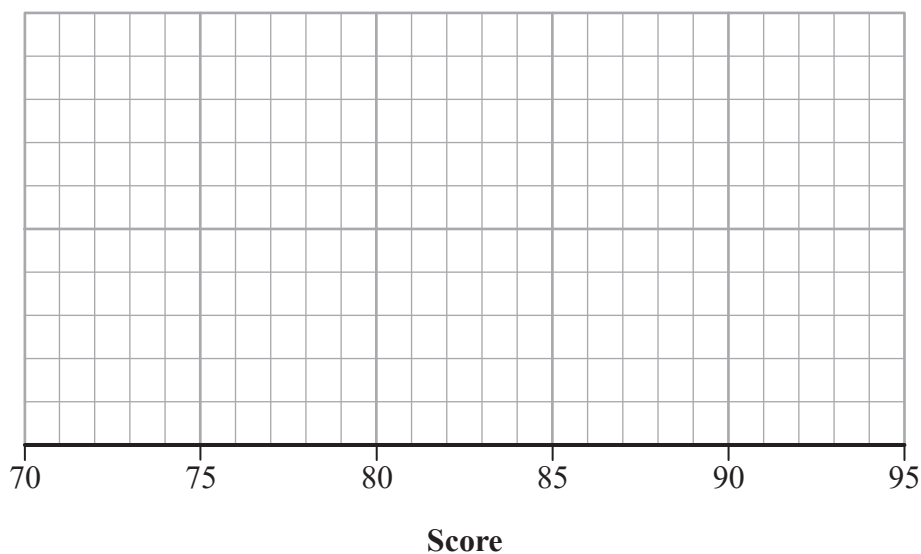


- 9 The table gives information about the scores achieved by students in a French test and in a German test.

The scores in each test are modelled by a normal distribution.

Test	Mean score	Standard deviation
French	78	2.0
German	85	2.5

- (a) On the grid, sketch a diagram of each distribution of scores.  
Label each diagram.



(4)

- (b) Use the information in the table to make **two** comparisons between the scores in the French test and the scores in the German test.  
Interpret your comparisons in context.

.....

.....

.....

.....

.....

.....

(3)





One student who took the French test is selected at random.

- (c) Work out an estimate for the probability that this student scored more than 76 on the French test.

.....  
(3)

Lucy took both the French test and the German test.

Before she took these tests she said that the probability that she will score more than the mean score in both tests is

$$\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

- (d) Comment on the validity of what Lucy said.

.....  
.....  
.....  
.....  
(2)

**(Total for Question 9 is 12 marks)**



10 A company makes bars of chocolate on a production line in a factory.

If a chocolate bar is found to weigh less than 60 g, the company will not be allowed to sell it.

The company uses quality assurance control charts to monitor the weight of the chocolate bars.

Samples of chocolate bars are taken at regular intervals.

The mean weights of the samples are normally distributed.

- (a) Explain why it would not be appropriate for the production line to be set so that 60 g is the target weight.

(1)

In fact the production line is set so that the sample means should be normally distributed with a mean of 62 g and a standard deviation of 0.4 g.

- (b) What can be deduced from this information about the mean and standard deviation of all of the chocolate bars made on the production line?

(2)



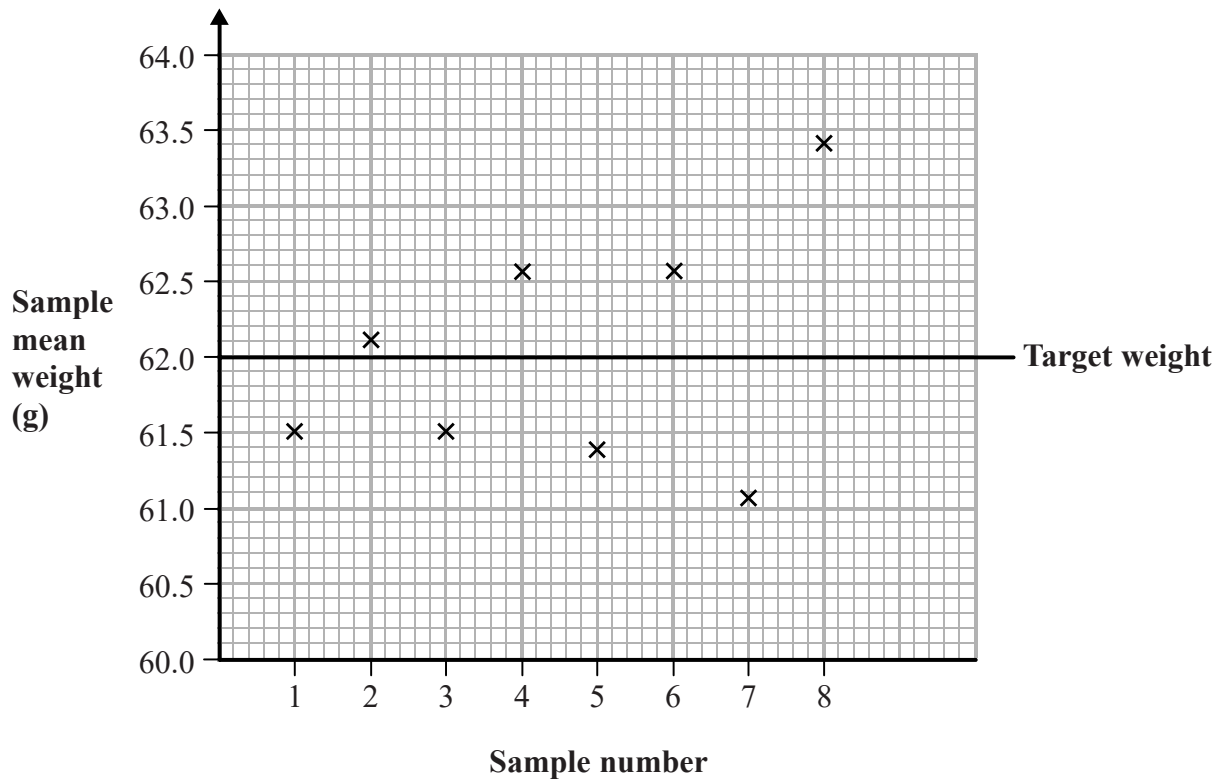
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Alice takes the samples in order to monitor the sample means of the weights.

Here is her incomplete control chart showing 8 sample means.



- (c) By completing the control chart, determine what actions, if any, Alice should have taken based on the information given.

You must justify your answer.

(6)

(Total for Question 10 is 9 marks)



S 5 9 3 0 7 A 0 1 9 2 4

11 A farmer recorded the birth weights, in kg, of a sample of 50 piglets born on his farm.

The grouped frequency table gives information about his results.

Weight ( $w$ kg)	Frequency ( $f$ )
$0.5 \leq w < 1.5$	4
$1.5 \leq w < 2.0$	12
$2.0 \leq w < 2.5$	17
$2.5 \leq w < 3.0$	13
$3.0 \leq w < 4.0$	3
$4.0 \leq w < 6.0$	1

Using the information in the table and the values of  $w$  as the class midpoints, the farmer finds that an estimate for the mean birth weight of these piglets is 2.29 kg.

He also finds that  $\sum fw^2 = 286.875$

- (a) Show that an estimate of the standard deviation of the birth weights of these 50 piglets is 0.7 kg, correct to 1 decimal place.

(2)



(b) Calculate an estimate for the skew of the birth weights of the piglets.

.....  
(4)

(c) Interpret your answer to part (b).

.....  
.....  
(1)

(d) Using only the information given in this question, explain whether it is possible to deduce that any of the birth weights are outliers.

Justify your answer.

.....  
.....  
.....  
.....  
(3)

**(Total for Question 11 is 10 marks)**

---

**TOTAL FOR PAPER IS 80 MARKS**



DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

**BLANK PAGE**



DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

**BLANK PAGE**



DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

**BLANK PAGE**

