SPECIMEN MATERIAL

Please write clearly in block capitals.

Centre number | | | | | | Candidate number | | | | |

Surname _____

Forename(s) _____

Candidate signature _____

GCSE STATISTICS

Higher tier Paper 1

H

Date of Exam Morning Time allowed: 1 hour 45 minutes

Materials

For this paper you must have:

- a calculator
- mathematical instruments.

Instructions

- Use black ink or black ball-point pen. Draw diagrams in pencil.
- Fill in the boxes at the top of the page.
- Answer **all** questions.
- You must answer the questions in the spaces provided. Do not write outside the box around each page or on blank pages.
- Do all rough work in this book. Cross out any work you do not want to be marked.

Information

- The marks for questions are shown in brackets.
- The maximum mark for this paper is 80.
- You may ask for more answer paper and graph paper. These must be tagged securely to this answer booklet.

Do not write outside the box

Answer **all** questions in the spaces provided.

1 Circle the value of the geometric mean for the numbers 2, 4 and 27

[1 mark]

6 11 14.7 25 216

2 In a normal distribution, approximately what percentage of the data lie with two standard deviations of the mean?

Circle your answer

[1 mark]

50 68 95 99

3 Two normal fair dice are rolled and their scores added.

Circle the probability of scoring a total of 12

[1 mark]

$\frac{1}{6}$ $\frac{1}{12}$ $\frac{1}{18}$ $\frac{1}{36}$

4 Which of these values for Pearson's product moment correlation coefficient shows perfect negative correlation?

Circle the correct answer.

[1 mark]

-1 $\quad$ 0 $\quad$ $-\frac{1}{2}$ $\quad$ $+1$

Turn over for the next question

Do not write outside the box

5 American paint manufacturer DuPont carry out annual surveys about the most popular car colours across the world.

Here is a spreadsheet of the results from 2012.

	A	B	C	D	E
1	Colour	North America (%)	Europe (%)	Asia-Pacific (%)	Worldwide (%)
2					
3	White	24	24	22	23
4	Black	19	23	21	21
5	Silver	16	14	14	18
6	Grey	15	115	20	14
7	Red	10	6	7	8
8	Blue	7	8	5	6
9	Brown	5	6	6	6
10	Other	2	3	4	3
11	Green	2	1	1	1

Source: Wikipedia

5 (a) Give one way you could check whether any data in this spreadsheet needs to be cleaned.

[1 mark]

5 (b) Circle the cell in the spreadsheet where the data needs cleaning.

What value do you think it should be?

[1 mark]

Answer _____

5 (c) The spreadsheet shows the number of cars made in each year from 2008 to 2014, to the nearest 100 thousand.

	A	B
1	**Year**	**Number of cars made (millions)**
2	**2008**	**70.5**
3	**2009**	**61.8**
4	**2010**	**77.9**
5	**2011**	**80.0**
6	**2012**	**84.1**
7	**2013**	**87.3**
8	**2014**	**89.7**

Source: Wikipedia

Describe the pattern in the number of cars made from 2008 to 2014

[1 mark]

5 (d) Use **both** spreadsheets to calculate the approximate number of cars made worldwide in 2012 that were painted Red.

Give your answer to a suitable degree of accuracy.

[4 marks]

Answer _____ million

6 Imran drives, walks or cycles to work depending on the weather.

- If it is raining, he will always drive to work.
- If it is not raining, then he will cycle to work, unless it is windy then he walks.

The probability it is raining on any particular day is 0.3

The probability it is not raining but it is windy is 0.18

6 (a) Write down the probability that Imran drives to work.

[1 mark]

Answer _____

6 (b) Work out the probability that Imran drives to work **two** days in a row.

[2 marks]

_____

Answer _____

6 (c) Work out the probability that Imran cycles to work.

[2 marks]

_____

_____

Answer _____

6 (d) From the information given, is it possible to work out the probability of it being windy on any particular day?

Tick a box.

[1 mark]

Yes ☐ No ☐

Give a reason for your answer.

Reason _____

_____

_____

Turn over for the next question

7 Den and Pete sell games consoles.

The number of consoles sold, x, by Den in each of the 12 months of 2015 is summarised by

$$\sum x = 132 \qquad \sum x^2 = 1560$$

7 (a) Calculate the mean and standard deviation for the number of consoles Den sold per month.

[4 marks]

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

where n is the number of months in 2015

Mean = _____

Standard deviation = _____

7 (b) Den earns £500 each month plus £100 for each console sold.

On average, Den works 16 days per month.

Den says,

"I earned £100 per day on average for my work."

Is Den correct?

Show working to support your answer.

[3 marks]

7 (c) During 2015 Pete earned a mean of £1450 per month with a standard deviation of £275

Compare the earnings each month for Den and Pete.

[2 marks]

7 (d) Do you think it is appropriate for the mean to be used with these data?

[1 mark]

Tick a box.

Yes ☐ No ☐

Give a reason for your answer.

Reason _____

Turn over for the next question

8 A national chain of gyms employs 572 trainers in 30 gyms of different size.

A sample of 50 trainers is to be selected from this total and their views sought on changes to activities offered.

The following are suggested as alternative sampling methods to use.

Method A One trainer is selected from the 10 smallest gyms. Two trainers are selected from each of the remaining 20 gyms.

Management will then select the sample of trainers in any convenient way.

Method B All 572 trainers are numbered from 000 to 571. Start with number 010 and take every 11th trainer to be part of the sample.

Method C All 572 trainers are numbered from 000 to 571. Using random number tables 50 numbers within the range are chosen and the corresponding trainers included in the sample.

Name and compare each sampling method.

Make a reasoned choice of which method should be used.

[7 marks]

Turn over for the next question

9 The table shows, for each of a random sample of 9 books, rankings based on number of pages and retail price, lowest first.

	Rankings	
Book	**Number of pages**	**Retail price**
A	1	2
B	4	3
C	9	7.5
D	5	9
E	3	4
F	6	5
G	8	7.5
H	2	1
I	7	6

9 (a) Explain what the ranking for the retail price of books C and G shows.

[1 mark]

9 (b) (i) Niles uses a spreadsheet to calculate $\sum d^2 = 24.5$

Spearman's rank correlation coefficient $= 1 - \dfrac{6\sum d^2}{n(n^2 - 1)}$

Complete the calculation of the value of Spearman's rank correlation coefficient for the data.

[4 marks]

Answer _____

9 (b) (ii) Explain, in context, what this shows.

[1 mark]

9 (c) The equation of the regression line of y (retail price, £) on x (number of pages) for these 9 books is $y = 0.02x + 1.35$

9 (c) (i) What does the value of 0.02 show in this context?

[2 marks]

9 (c) (ii) The difference in retail price of two other books is £10.30.

The larger book has 765 pages.

Estimate the number of pages in the smaller book.

[3 marks]

Answer £ _____

10 In a town in 2015 the crude death rate was 7.5 and the crude birth rate was 8.5

Quinlan says,

'In 2015 the population of the town will have increased from 2014'

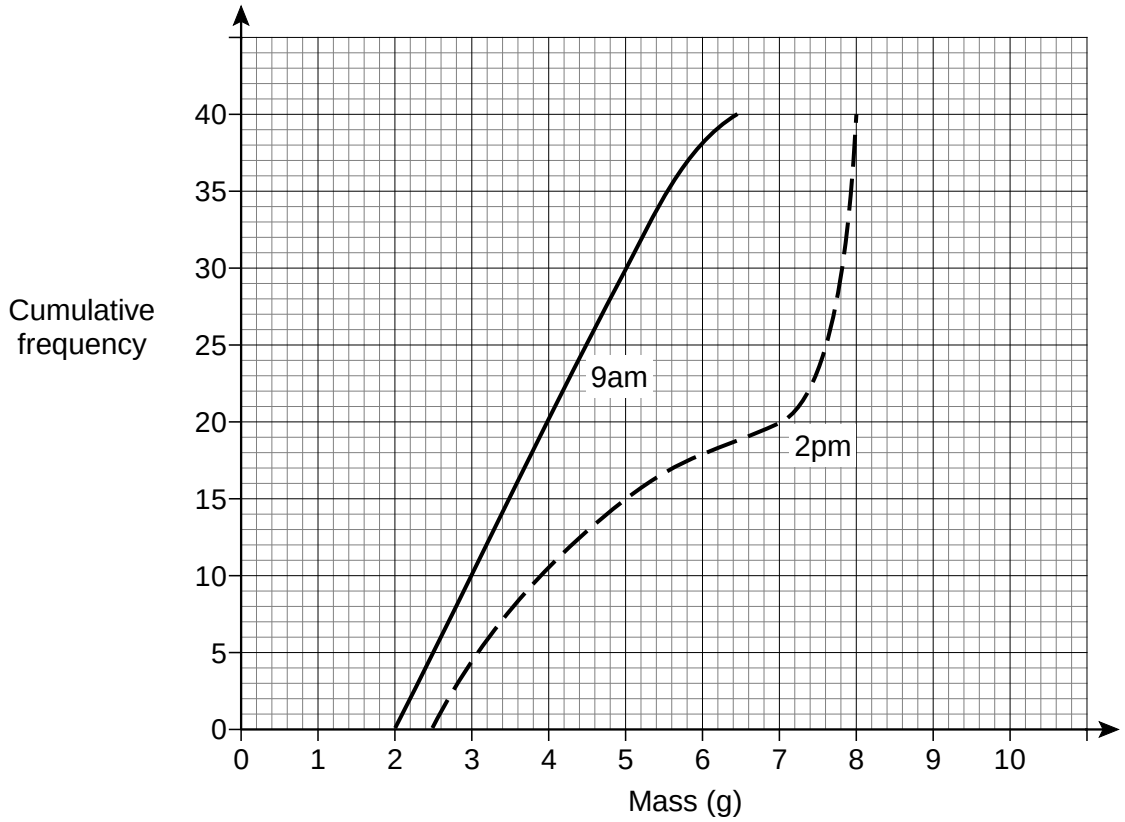
10 (a) Give **one** reason why Quinlan could be correct.

[1 mark]

10 (b) Give **one** reason why Quinlan could be wrong.

[1 mark]

11 The two cumulative frequency curves show the masses (in grams) of 40 goldcrests, Britain's smallest bird, at 9am and 2pm on the same day.



Compare statistically the masses of goldcrests at 9 am and 2 pm.

[6 marks]

12 Kal collected data on the number of votes cast in two constituencies in the 2015 general election.

Constituency	**Northtown**	**Southtown**
Number of votes cast	54 620	76 468

She decided to represent the data using two comparative pie charts.

Constituency	**Northtown**	**Southtown**
Radius of pie chart	4 cm	

12 (a) Calculate the radius for Southtown, to **three** significant figures.

[4 marks]

Answer _____ cm

12 (b) In Northtown, the angle used to represent the Labour share of the total vote was 126°

Calculate the number of people who voted Labour in Northtown.

[3 marks]

Answer _____

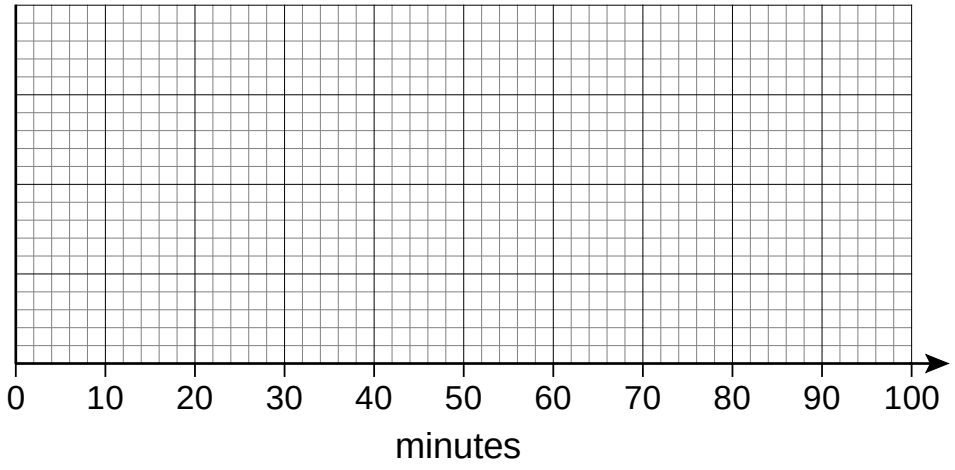
13 Two large groups of students attempt to solve some problems.

Their times to solve the problems are recorded.

13 (a) Group A have results that are approximately normally distributed with a mean of 35 minutes and a standard deviation of 10 minutes.

Sketch the approximate distribution of these results on the grid below.

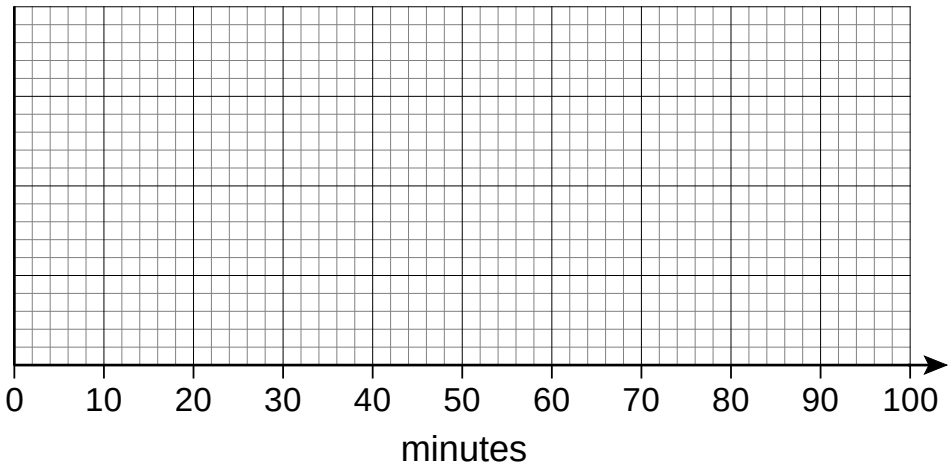
[3 marks]



13 (b) Group B have results that have a mean of 40 minutes but with a positive skew.

Sketch the approximate distribution of these results on the grid below

[2 marks]



14 There is a **Data Sheet** Insert for **Question 14**

Dave is investigating distances thrown in the women's discus competitions in the Olympic Games in different years.

He collects data from the website:

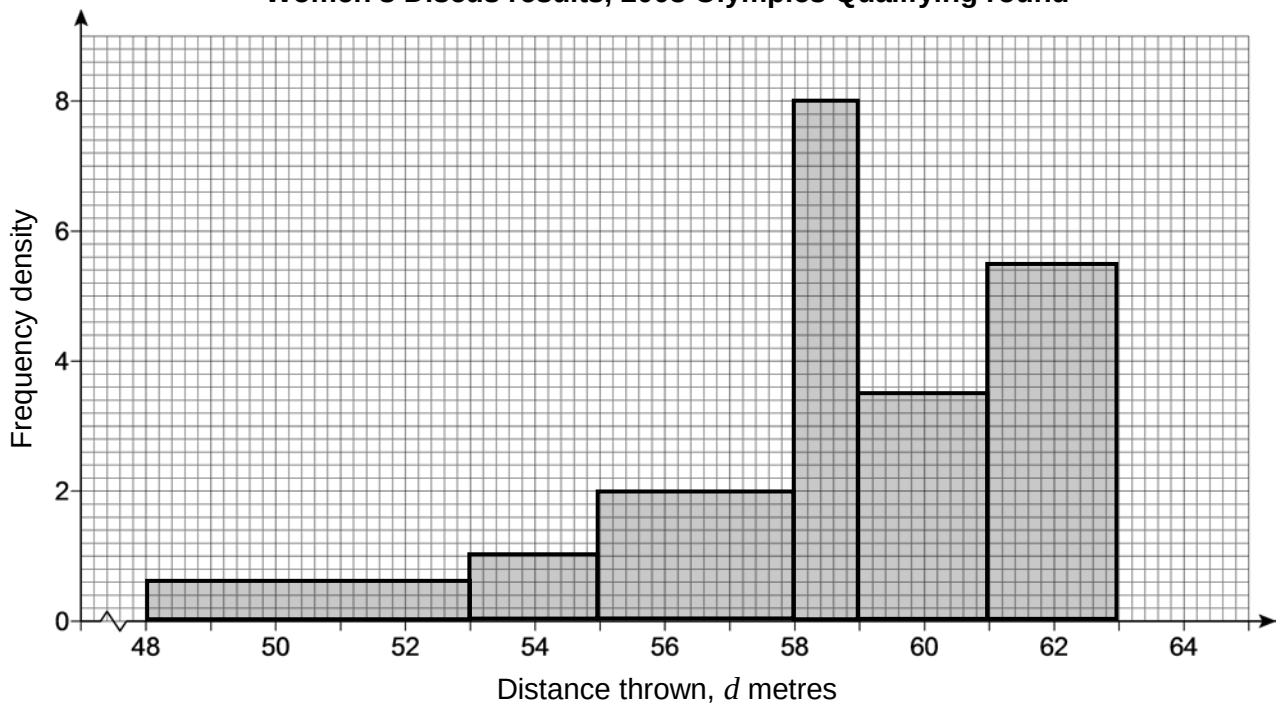
www.olympic.org/olympic-results/london-2012/athlectics/discus-throw-w

14 (a) Suggest a hypothesis he could use.

[1 mark]

The histogram below shows the results for 2008

Women's Discus results, 2008 Olympics Qualifying round



14 (b) State **one** reason why a histogram is appropriate for this type of data.

[1 mark]

14 (c) State **one** reason why unequal widths are useful in this case.

[1 mark]

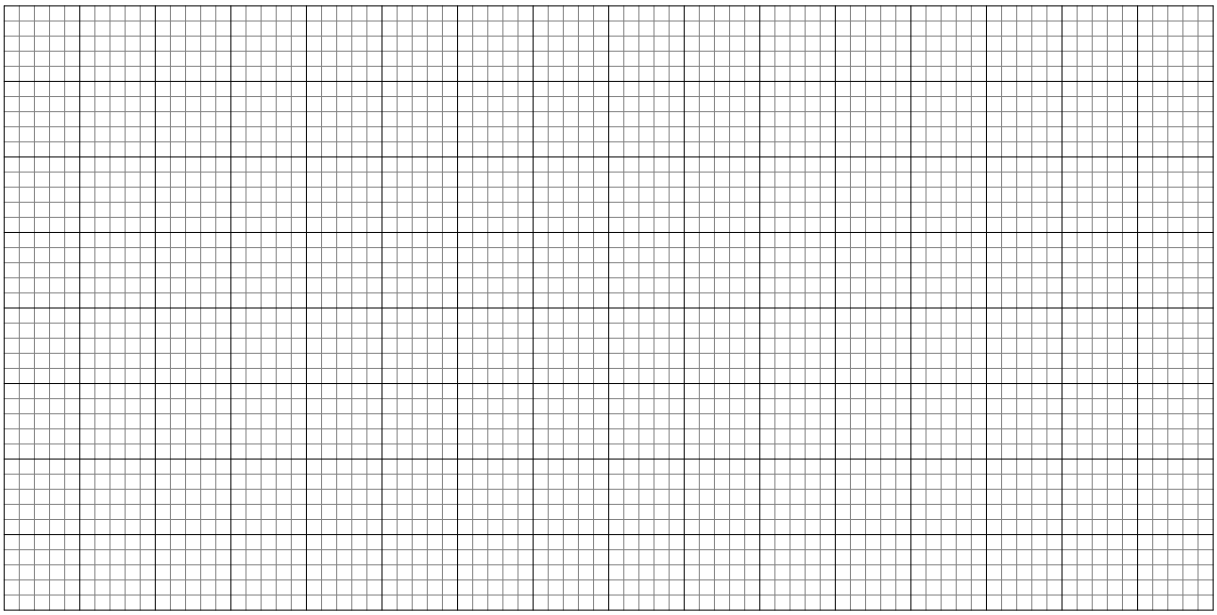
14 (d) On the graph paper below draw a histogram to show the results for the **2012** women's **Qualifying** Round.

You may use the table below.

Give a clear justification for your choice of class widths.

[8 marks]

Class			



14 (e) Interpret both histograms fully.

You should make clear reference to features shown in the histograms and how they either support or do not support the hypothesis you stated in **part (a)**.

[4 marks]

END OF QUESTIONS

There are no questions printed on this page

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED

There are no questions printed on this page

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED