Edexcel GCSE

Statistics

Paper 1H

Higher Tier

Wednesday 18 June 2008 – Morning

Time: 2 hours 30 minutes

Materials required for examination	**Items included with question papers**
Ruler graduated in centimetres and millimetres, protractor, compasses, pen, HB pencil, eraser, electronic calculator.	Nil

Instructions to Candidates

In the boxes above, write your centre number, candidate number, your surname, initials and signature.
Check that you have the correct question paper.
Answer ALL the questions. Write your answers in the spaces provided in this question paper.
Some questions must be answered with a cross in a box (☒).
If you change your mind about an answer, put a line through the box (☒) and then mark your new answer with a cross (☒).
You must NOT write on the formulae page or any blank pages. Anything you write on these pages will gain NO credit.
If you need more space to complete your answer to any question, use additional answer sheets.

Information for Candidates

The marks for individual questions and the parts of questions are shown in round brackets: e.g. **(2)**.
This question paper has 8 questions in Section A and 9 questions in Section B. The total mark for this paper is 100.
There are 28 pages in this question paper. Any blank pages are indicated.

Advice to Candidates

Show all stages in any calculations.
Work steadily through the paper. Do not spend too long on one question.
If you cannot answer a question, leave it and attempt the next one.
Return at the end to those you have left out.

Turn over

edexcel
advancing learning, changing lives

GCSE Statistics 1389

Higher Tier Formulae

You must not write on this page.
Anything you write on this page will gain NO credit.

Mean of a frequency distribution $= \dfrac{\sum fx}{\sum f}$

Mean of a grouped frequency distribution $= \dfrac{\sum fx}{\sum f}$, where x is the mid-interval value.

Variance $= \dfrac{\sum (x-\bar{x})^2}{n}$

Standard deviation (set of numbers) $\sqrt{\left[\dfrac{\sum x^2}{n} - \left(\dfrac{\sum x}{n}\right)^2\right]}$

or $\sqrt{\left[\dfrac{\sum (x-\bar{x})^2}{n}\right]}$

where $\bar{x}$ is the mean set of values.

Standard deviation (discrete frequency distribution) $\sqrt{\left[\dfrac{\sum fx^2}{\sum f} - \left(\dfrac{\sum fx}{\sum f}\right)^2\right]}$

or $\sqrt{\left[\dfrac{\sum f(x-\bar{x})^2}{\sum f}\right]}$

Spearman's Rank Correlation Coefficient $1 - \dfrac{6\sum d^2}{n(n^2-1)}$

SECTION A
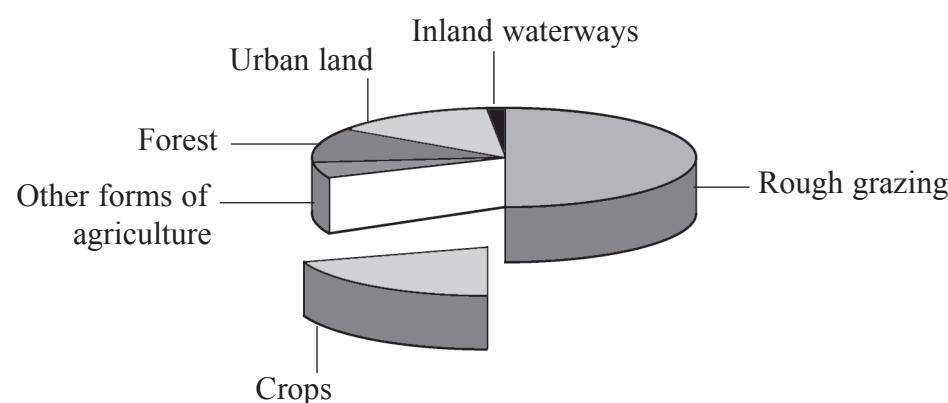
Answer ALL the questions. Write your answers in the spaces provided.

You must write down all stages in your working.

1. The pie chart gives information about land use in the U.K.

Land use in the U.K.

(Source: www.sustainable-development.gov.uk)

Describe **two** features of the diagram that can be misleading.

1 ....

....

....

2 ....

....

....

(Total 2 marks)

M31091A0328

2. The table shows information about births in England and Wales for the years 1995 to 2002.

Births in England and Wales

Year	Number of births	Births per 1000 women 15 - 44	Male births per 1000 female births	Mean age of mothers at childbirth (years)	Percentage of births outside marriage	Percentage of births to mothers not born in the UK
1995	648 138	60.5	1051	28.5	33.9	12.6
1996	649 485	60.6	1055	28.6	35.8	12.8
1997	643 095	60.0	1051	28.8	37.0	13.1
1998	635 901	59.2	1051	28.9	37.8	13.6
1999	621 872	57.8	1055	29.0	38.9	14.3
2000	604 441	55.9	1050	29.1	39.5	15.5
2001	594 634	54.7	1050	29.2	40.0	16.5
2002	596 122	54.7	1055	29.3	40.6	17.7

(Source: National Statistics - Live births)
Reproduced under the terms of the click-use licence C2006010031.

(a) How many births were there in the year 2000?

....

(1)

(b) Describe the trend in the mean age of mothers at childbirth between the years 1995 and 2002.

....

....

....

(1)

For every year between 1995 and 2002 there were more male births than female births.

(c) How does the table show this?

....

....

....

(1)

(d) From the table, what can you infer about the trend in the percentage of births to mothers born in the UK?
Give a reason for your answer.

....

....

....

....

....

(2)

(Total 5 marks)
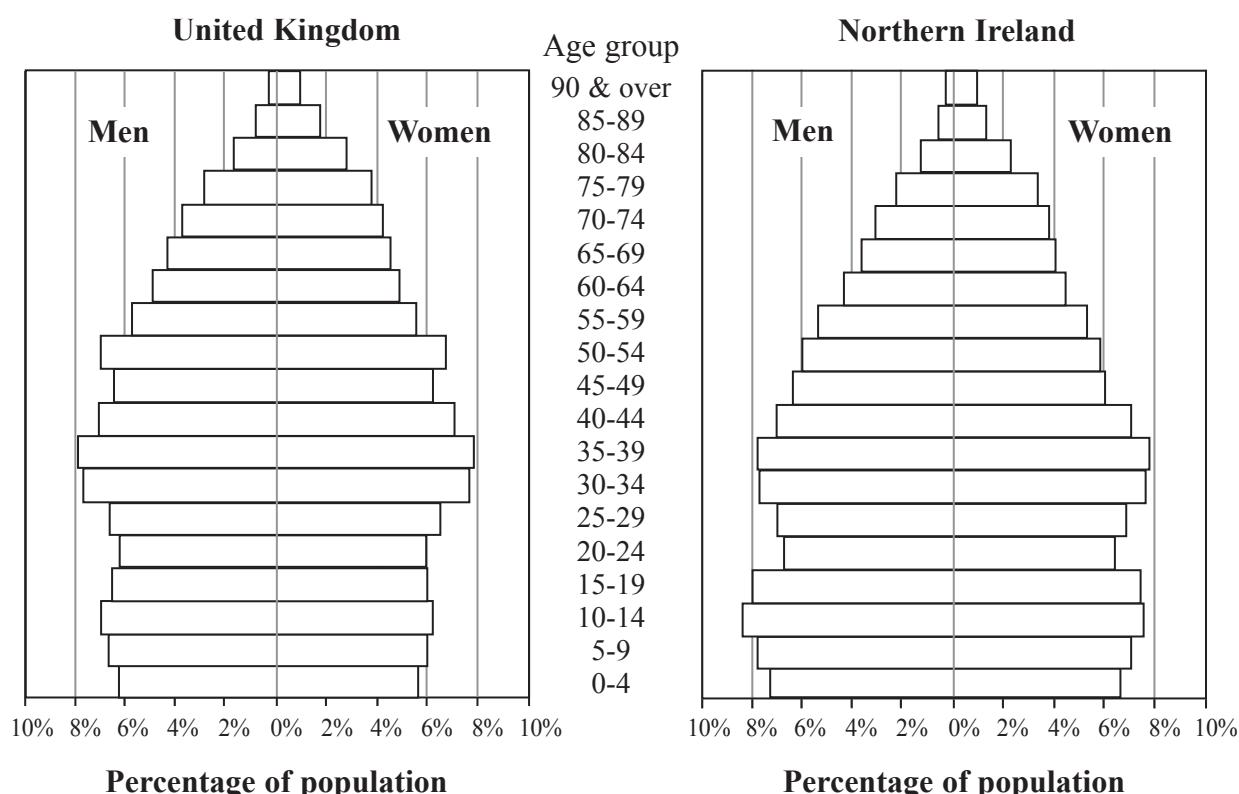
3. The two population pyramids show the percentages of men and women in each age group in the United Kingdom, and in Northern Ireland.

The percentages of men and women in each age group

United Kingdom	Age group	Northern Ireland

(Source: Gov.uk/census 2001)
Reproduced under the terms of the click-use licence C2006010031.

(a) Write down the age group which has the greatest percentage of women in the United Kingdom.

....
(1)

6% of the men in Northern Ireland are in one age group.

(b) Write down this age group.

....
(1)

(c) Compare the percentage of people up to the age of 19 in the United Kingdom with the percentage of people up to the age of 19 in Northern Ireland.

....

....

....
(1)

(Total 3 marks)

4. A survey was done to count the number of eggs in blue-tit nests.
The data is recorded in the table.

Numbers of eggs in blue-tit nests

Number of eggs (*x*)	Frequency (*f*)	*fx*
7	2	14
8	8	64
9	12	108
10	15	150
11	16	
12	9	
Total		

(Source: Internet Survey)

(a) Complete the table.

(1)

(b) Write down the modal number of eggs.

....
(1)

(c) Work out the mean number of eggs.

....
(2)

(d) Find the median number of eggs.

....
(1)

(e) Which average would **best** describe the number of eggs in a blue-tit nest?
Give a reason for your answer.

....

....
(1) **Q4**

(Total 6 marks)

5. A company is planning a survey about the types of transport their workers use to get to the factory.

(a) Give **two** reasons why a pilot study should be carried out.

1 ....

....

2 ....

....

(2)

Megan is going to stand outside the factory on one Monday morning.
As the workers arrive she will ask them how they travelled to work.

(b) Give **one** reason why this will not give a random sample.

....

....

(1)

(c) Describe how to take a random sample from the workers at the factory.

....

....

....

(2)

The company records the distance each worker travels to work and the type of transport they use.

(d) Put a cross in the box to indicate whether each of the following is qualitative, discrete or continuous.

(i) The distance travelled to work.

Qualitative	Discrete	Continuous
☒	☒	☒

(ii) The type of transport.

Qualitative	Discrete	Continuous
☒	☒	☒

(2)

(Total 7 marks)

6. The table shows the Price Index Numbers for the average monthly rental of two-bedroomed properties in Manchester between April and August.
The base month is April.

Average monthly rental of two-bedroomed properties in Manchester

Month	April	May	June	July	August
Price index	100	97	110	70	96

(Source: adapted from Internet Sources)

(a) In which month was the average monthly rental higher than it was in April?

....
(1)

The average monthly rental was £720 in April.

(b) Calculate the average monthly rental in July.

£ ....
(2)

(Total 3 marks)

7. The table shows information about the numbers of fatal road accidents in Great Britain.

Fatal road accidents in Great Britain

Year	2004				2005			
Quarter	1	2	3	4	1	2	3	4
Number of fatal road accidents (x)	703	825	801	892	740	727	818	916

(Source: Department of Transport)

(a) Calculate the mean number of fatal road accidents per quarter.

....

(1)

For the numbers of fatal road accidents in the table $\sum x^2 = 5\,196\,408$

(b) Calculate the standard deviation of the numbers of fatal road accidents.
Give your answer to **1** decimal place.

....

(2)

In the first quarter of 2006 the number of fatal road accidents was 720

A motoring organisation is going to calculate the mean number of fatal road accidents per quarter for 2004, 2005 and the first quarter of 2006.

(c) Without carrying out any more calculations compare this mean with the mean found in part (a).
Give a reason for your answer.

....

....

....

(2)

Q7

(Total 5 marks)

8. The weight of coffee in jars of Groundly coffee is normally distributed with a mean weight of 260 g and a standard deviation of 4.5 g.

A jar of Groundly coffee is chosen at random.

(a) What is the probability that this jar of coffee weighs less than 251 g?

....

(2)

Jars of Groundly coffee have the weight marked as 250 g ℮.

The letter ℮ shows that the weight of the coffee in the jars conforms to the mean weight standard.
The mean weight standard allows no more than $2\frac{1}{2}\%$ of the jars of coffee to contain less than the weight marked on the jar.

(b) Discuss whether or not these jars of Groundly coffee conform to the mean weight standard.

....

....

....

....

....

(2) Q8

(Total 4 marks)

TOTAL FOR SECTION A: 35 MARKS

SECTION B

Answer ALL the questions. Write your answers in the spaces provided.
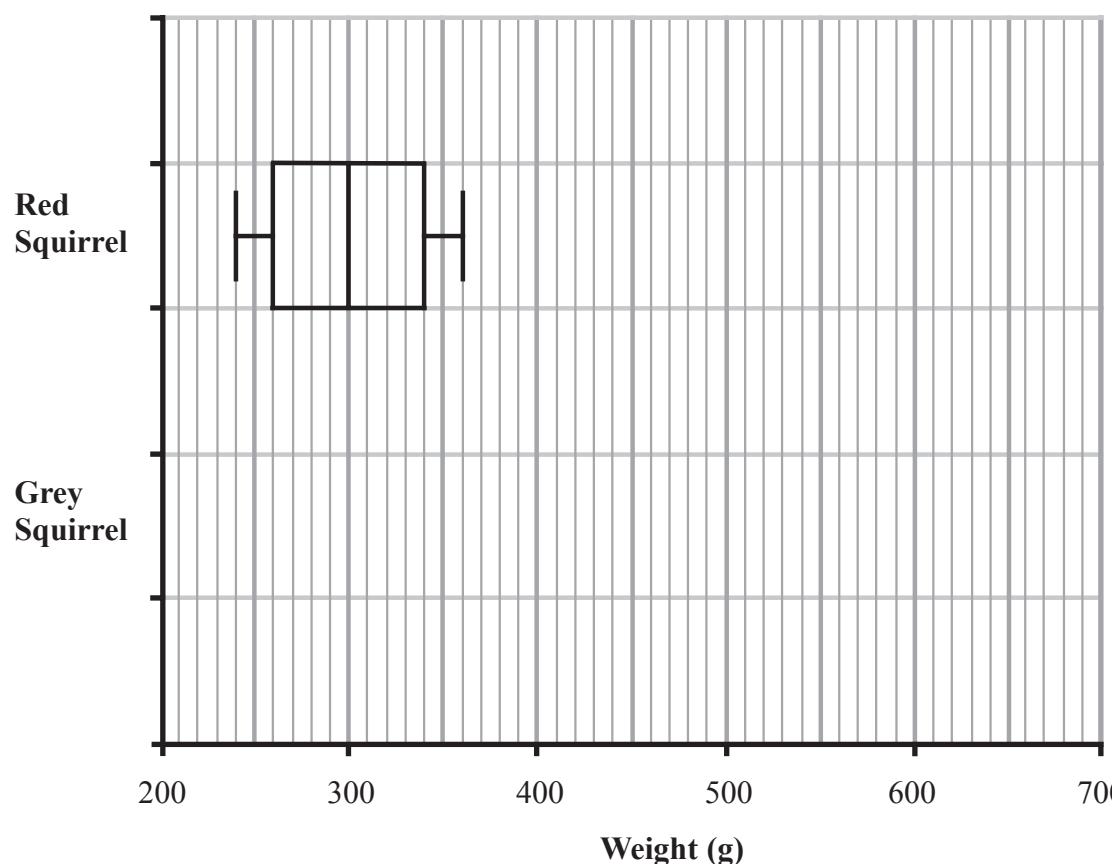
You must write down all stages in your working.

1. There has been research into the survival of Red Squirrels.
   A large number of adult Red Squirrels and a large number of adult Grey Squirrels were weighed.
   The summary statistics of body weights are shown in the table.

Weight in grams	Minimum weight (g)	Lower quartile (g)	Median weight (g)	Upper quartile (g)	Maximum weight (g)
Red Squirrel	240	260	300	340	360
Grey Squirrel	300	450	500	570	620

A box plot has been drawn on the grid to show the distribution of the weights of Red Squirrels.

Distributions of the weights of squirrels

(Source: Wildlife Trust's Squirrel Survey)

(a) On the grid, draw a box plot to show the distribution of the weights of Grey Squirrels.

(3)

(b) Describe the skewness of each of the distributions.

....

....

....

(2)

(c) Compare the distributions of the weights of the Red Squirrels and the weights of the Grey Squirrels.

....

....

....

....

....

(2)

Some Grey Squirrels have a lot of red fur.
Some Red Squirrels have some grey fur.
This means that it is not always possible to tell if a squirrel is red or grey by looking at the colour of its fur.

There are Red Squirrels and Grey Squirrels in a wood.

(d) How could you use the box plots to help you find out whether a certain squirrel is a Red Squirrel or a Grey Squirrel?

....

....

....

....

....

(2)

Q1

(Total 9 marks)
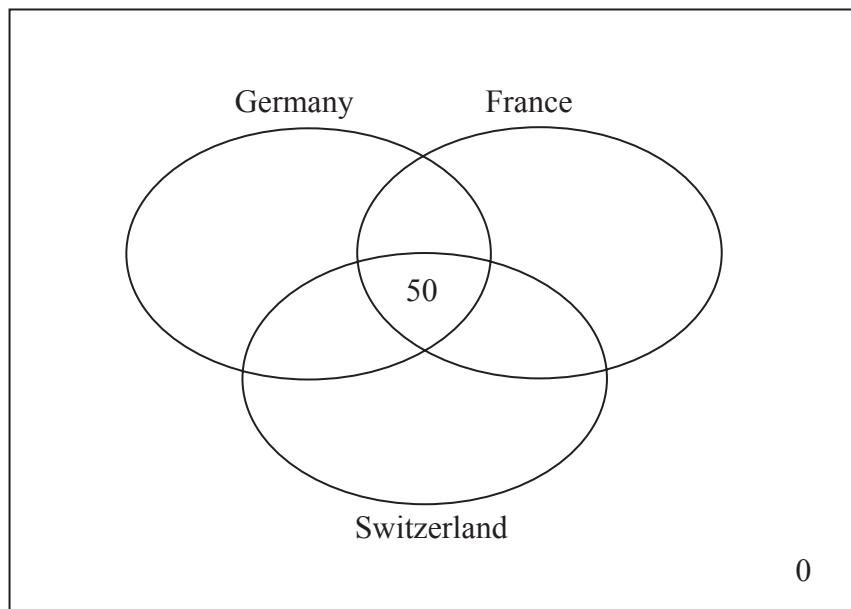
2. A travel agent organises several different tours to Germany, France and Switzerland. Each tour goes to 1, 2 or 3 of these countries.

A total of 200 people went on these tours.

Of these
130 people went on a tour to Germany,
131 people went on a tour to France,
122 people went on a tour to Switzerland,
74 people went on a tour to Switzerland and France,
84 people went on a tour to France and Germany,
75 people went on a tour to Germany and Switzerland,
50 people went on a tour to all three countries.

(a) Complete the Venn diagram for this information.

Germany   France

50

Switzerland

0

(3)

One of the 200 people is chosen at random.

(b) (i) Write down the probability that this person went on a tour to all three countries.

....

(ii) Write down the probability that this person did **not** go on a tour to France.

....

(iii) Work out the probability that this person went on a tour to Switzerland given that the person also went to France.

....

(4)

Q2

(Total 7 marks)
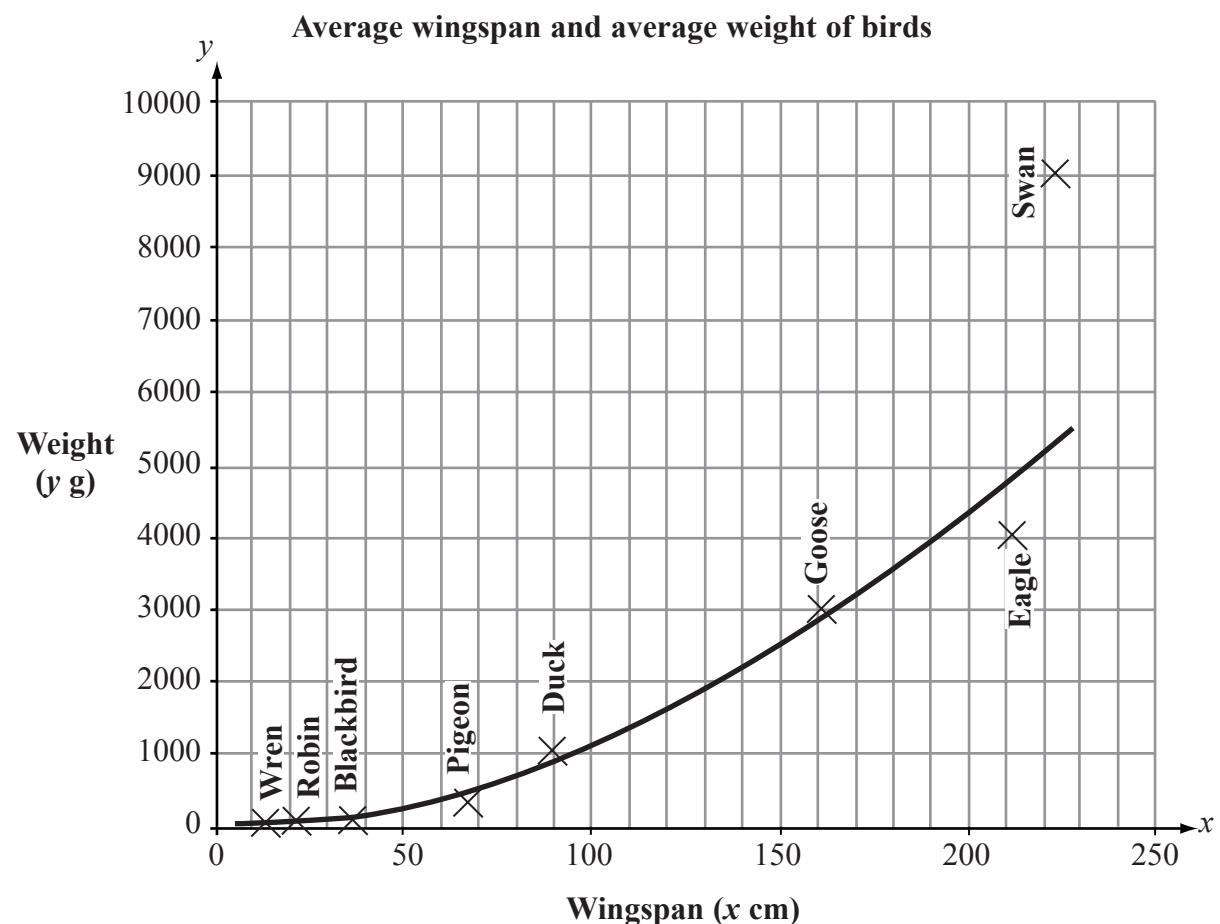
Turn over

BLANK PAGE

3. The scatter diagram shows the average wingspan and the average weight of 8 British birds.

Average wingspan and average weight of birds

(Source: British Trust for Ornithology)

(a) Put a cross in the box to indicate which one of the equations is the best model for these data.

$y = \dfrac{a}{x}$	$y = a\sqrt{x}$	$y = ax^2$
☒	☒	☒

(1)

M31091A01728
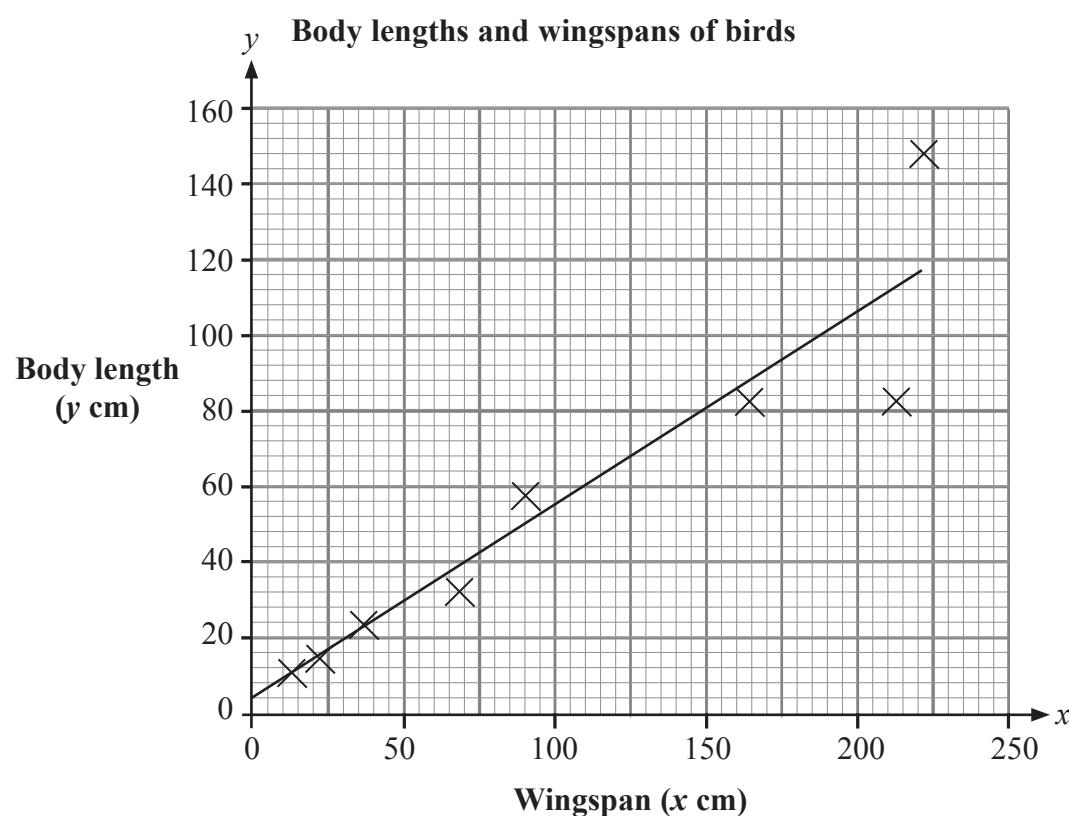
The scatter diagram below shows the body lengths and wingspans of the same birds.

Body lengths and wingspans of birds

The scatter diagram shows positive correlation.

(b) What can you conclude about the relationship between the body length of a bird and its wingspan?

....

....

....

....

(1)

A line of best fit has been drawn on the scatter diagram.

The equation for this line of best fit is of the form $y = ax + b$

(c) Work out the equation for this line of best fit.

....

(3)

A crow has a wingspan of 100 cm.

(d) Estimate the body length of this crow.

.... cm

(2)

Q3

(Total 7 marks)

4. A market research company is going to do a national opinion poll.

They want to find out what people think about the European Union.

(a) Give **two** reasons why it would **not** be a good idea to carry out a census.

1 ....

....

2 ....

....

(2)

The company are going to do a telephone poll.
First they will pick 10 towns at random.
Then they will pick 10 telephone numbers from the telephone book for each town.
They will ring these 100 telephone numbers.
The people who answer will form the sample.

(b) Discuss whether this will form a satisfactory sample for the poll.

....

....

....

....

(2)

Two of the questions suggested for the telephone poll are,

A: Do you think that the European Union is working well? Yes or No?

B: What do you think of the European Union?

(c) Which question is the most suitable?
Discuss the reasons for your choice.

....

....

....

....

(2)

(Total 6 marks)

Q4

5. The probability that an electrical component is faulty is 0.05

(a) Write down the probability that an electrical component is **not** faulty.

....
(1)

(b) Equal sized samples of the electrical components are selected at random. Give the name of a distribution that can be used to model the number of faulty electrical components in the samples.

....
(1)

(c) There are four electrical components in a sample.

(i) Calculate the probability that **exactly three** of the electrical components are **not** faulty.
You may use $(p + q)^4 = p^4 + 4p^3q + 6p^2q^2 + 4pq^3 + q^4$

....

(ii) Calculate the probability that **more than one** of the electrical components is faulty.

....
(5)

(Total 7 marks)

Q5

6. Tyson took a Statistics test and a Maths test.
Both tests were marked out of 100.
The table gives information about Tyson's marks.
It also shows the mean mark and the standard deviation for the group that took the tests.

Test	Tyson's mark	Group mean mark	Group standard deviation
Statistics	55	52	15
Maths	48	45	12

(a) (i) Work out Tyson's standardised score for Statistics.

....

(ii) Work out Tyson's standardised score for Maths.

....

(3)

(b) Write down the subject in which Tyson did better.
Give a reason for your answer.

....

....

....

(2)

(c) Comment on the group's performance in the two tests.

....

....

....

....

....

(2) Q6

(Total 7 marks)

7. At one location on a river, the distance from the bank and the corresponding depth of water were measured.
The table shows the results.

Distance from bank (cm)	0	50	150	200	250	300	350	400	450	500
Depth (cm)	0	10	26	44	59	52	74	104	85	96
Rank of Distance	1	2	3	4	5	6	7	8	9	10
Rank of Depth										
d										
d^2										

(Source: Research Project)

(a) Complete the table.

(2)

(b) Calculate Spearman's rank correlation coefficient for these data.

....
(2)

(c) Interpret your answer to (b).

....

....

....

....

(2)   Q7

(Total 6 marks)
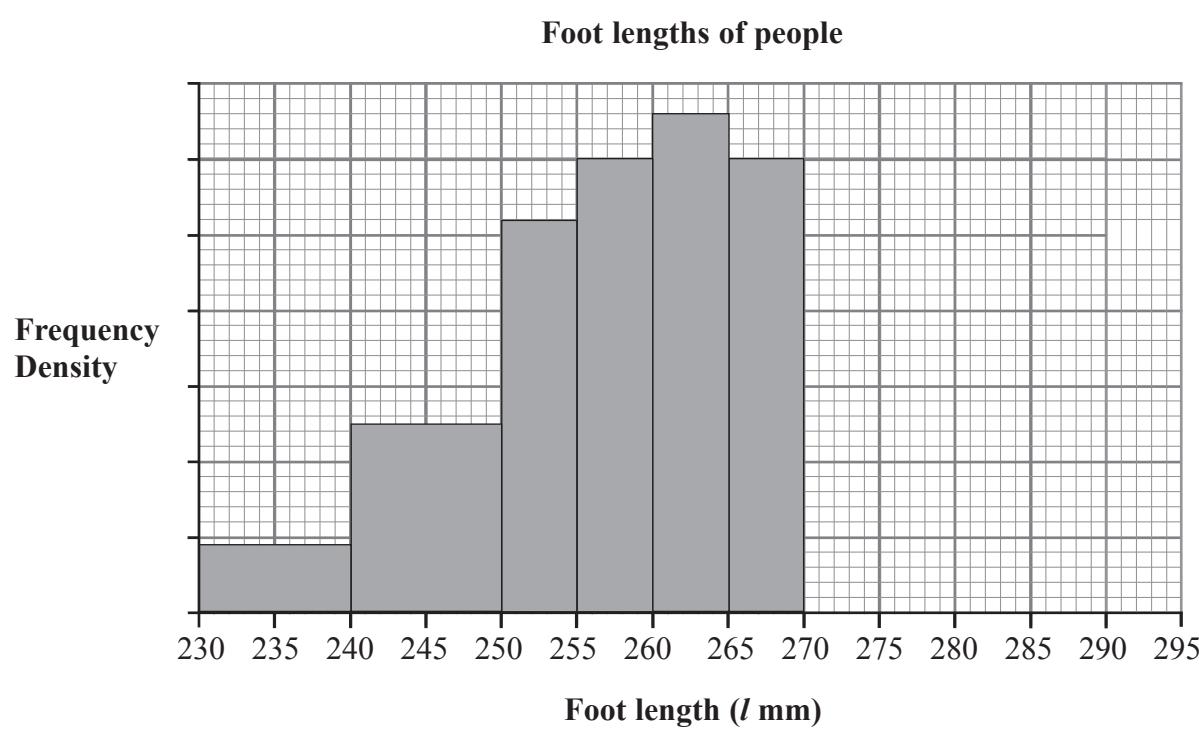
8. A shoe manufacturer measured the length (*l* mm) of 200 people's feet.

The results are summarised in the table below.

Length (*l* mm)	Frequency	
$230 \leqslant l < 240$	9	
$240 \leqslant l < 250$	25	
$250 \leqslant l < 255$	26	
$255 \leqslant l < 260$	30	
$260 \leqslant l < 265$	33	
$265 \leqslant l < 270$	30	
$270 \leqslant l < 280$	29	
$280 \leqslant l < 295$	18	

The incomplete histogram shows information about these data.

Foot lengths of people

Foot length (*l* mm)

(a) Use the information in the table to complete the histogram.

(3)

(b) Discuss how these results could be helpful to the shoe manufacturer.

....

....

....

....

(1)

(c) Calculate an estimate for the median foot length.
Give your answer to 2 decimal places.

.... mm

(3)

(d) What probability distribution could be used to model the lengths of these feet?

....

(1)

Q8

(Total 8 marks)

M 3 1 0 9 1 A 0 2 5 2 8

9. The table shows the numbers of people, in hundreds, going to a small cinema for 12 consecutive quarters.
It also shows some of the four-point moving averages.
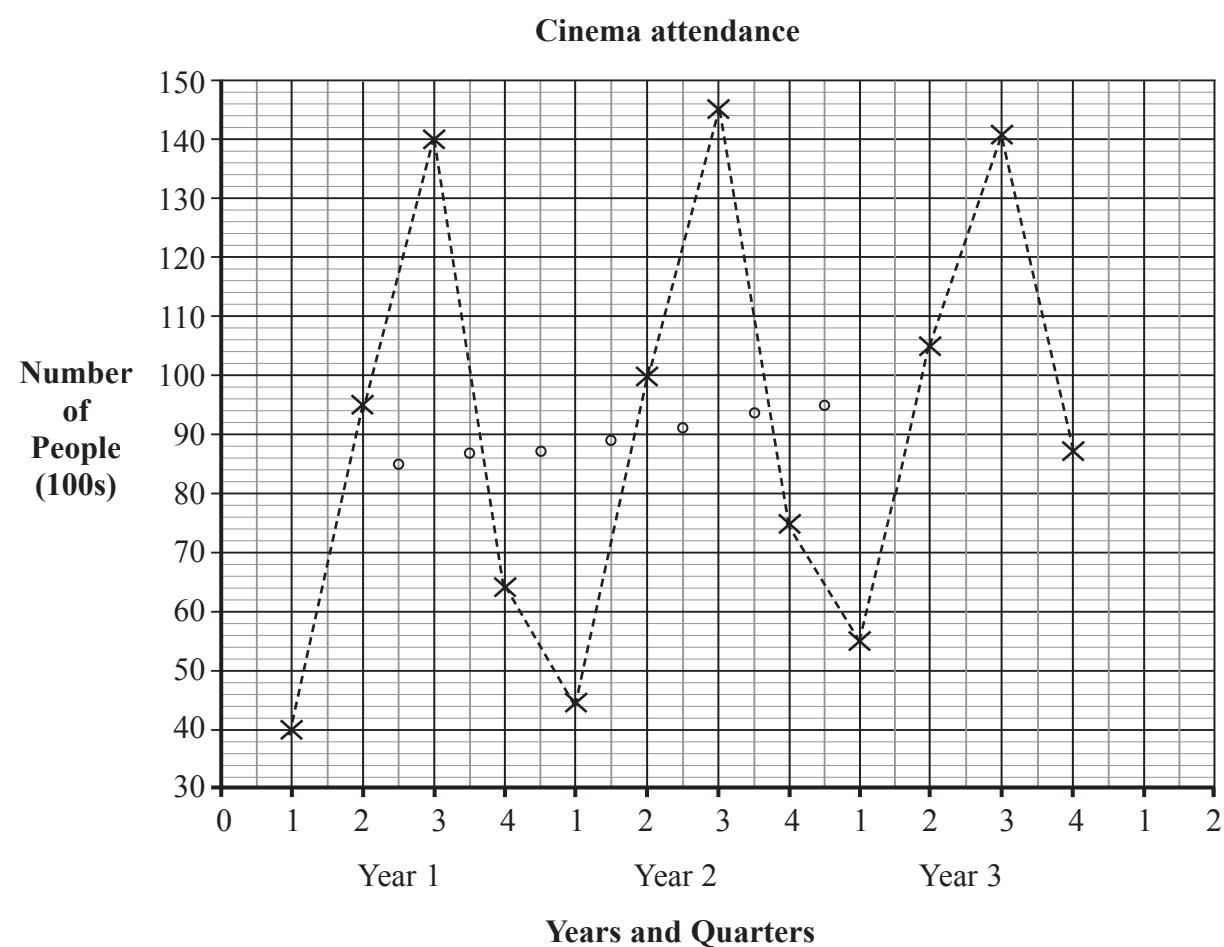
Year	Quarter	Number of people (100s)	Four-point moving averages
1	1	40	
	2	95	85
	3	140	86.25
	4	65	87.5
2	1	45	88.75
	2	100	91.25
	3	145	93.75
	4	75	95
3	1	55	
	2	105	
	3	141	
	4	87	

(a) Complete the table to show the two missing four-point moving averages.

(3)

The data from the table is plotted as a time series graph.

Cinema attendance

Years and Quarters

M31091A02628

(b) Plot the last two moving averages on the time series graph.

(1)

(c) Draw in a trend line for the moving averages.

(1)

The seasonal variations for quarter 1, in hundreds, are – 42, – 44, and – 40 people.

(d) Calculate the average seasonal variation for quarter 1.

.... hundreds

(1)

(e) Using your answer to part (d) and your trend line, find an estimate for the number of people going to the cinema in quarter 1 of year 4.

....

(2) Q9

(Total 8 marks)

TOTAL FOR SECTION B: 65 MARKS
TOTAL FOR PAPER: 100 MARKS

END

BLANK PAGE