Write your name here

Surname	Other names

Pearson Edexcel GCSE

Centre Number | | | | | |

Candidate Number | | | | |

Statistics

Paper 1H

Higher Tier

Monday 23 June 2014 – Afternoon
Time: 2 hours

Paper Reference
5ST1H/01

You must have:
Ruler graduated in centimetres and millimetres, protractor, pen
HB pencil, eraser, electronic 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.*

Information

- The total mark for this paper is 100.
- The marks for **each** question are shown in brackets
  – *use this as a guide as to how much time to spend on each question.*
- Questions labelled with an **asterisk** (*) are ones where the quality of your written communication will be assessed
  – *you should take particular care on these questions with your spelling, punctuation and grammar, as well as the clarity of expression.*

Advice

- Read each question carefully before you start to answer it.
- Keep an eye on the time.
- Try to answer every question.
- Check your answers if you have time at the end.

Turn over ▶

P43362A

5/5/5/3/

P 4 3 3 6 2 A 0 1 2 8

PEARSON

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

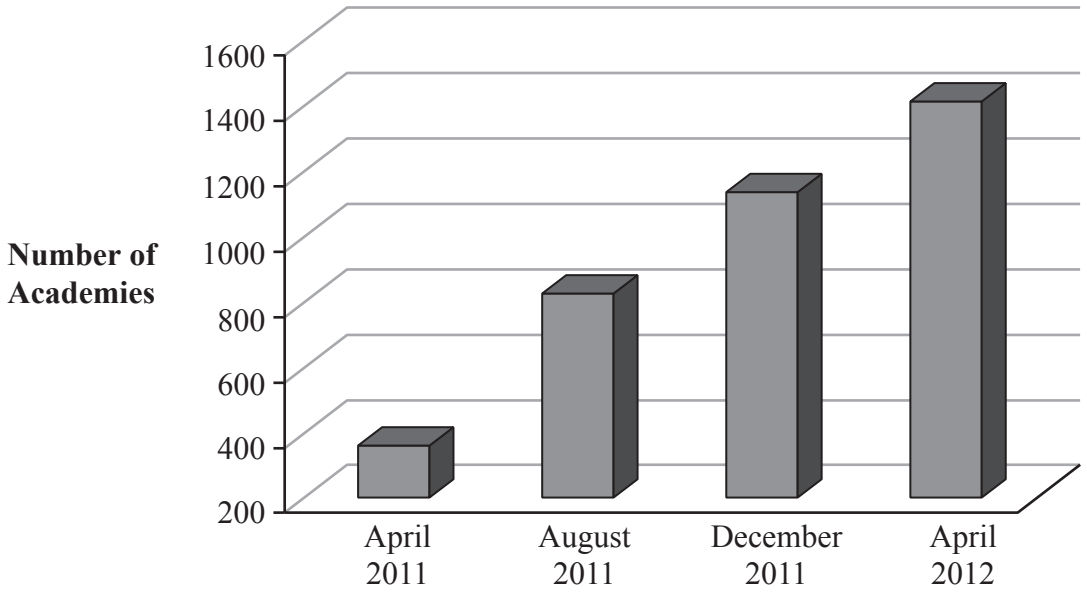
Spearman's Rank Correlation Coefficient $1 - \dfrac{6\sum d^2}{n(n^2-1)}$

Answer ALL the questions.

Write your answers in the spaces provided.

You must write down all stages in your working.

1 The graph shows the number of secondary schools in the UK that became Academies between April 2011 and April 2012



(Data source: Department for Education)

Give **two** reasons why this graph may be seen as misleading.

.....

.....

.....

.....

(Total for Question 1 is 2 marks)

2 Julie and Bevan own a sandwich company.

They deliver sandwiches to customers for lunch in each of 30 offices every day.
There are a number of customers in each office.

Julie wants to make changes to the sandwich menu.
She decides to find out the opinions of the customers.

(a) Describe the population for the survey.

.....

.....

(1)

Bevan wants to use a census to collect the customers' opinions.

(b) Write down **one** advantage of using a census.

.....

.....

(1)

Julie wants to use a sample of the customers, rather than a census.

(c) Give **two** reasons why a sample might be better.

Reason 1 .....

.....

Reason 2 .....

.....

(2)

(d) Explain what is meant by a random sample.

.....

.....

.....

(1)

Julie designs a questionnaire to give to customers.

One question on Julie's questionnaire is

Do you agree that the sandwiches are good value for money?

This is **not** a good question.

(e) Give two reasons why.

Reason 1 .....

.....

.....

Reason 2 .....

.....

.....

(2)

Bevan wants to use face to face interviews with the customers.

(f) Give one advantage and one disadvantage of using face to face interviews rather than a questionnaire given to customers.

Advantage .....

.....

.....

Disadvantage .....

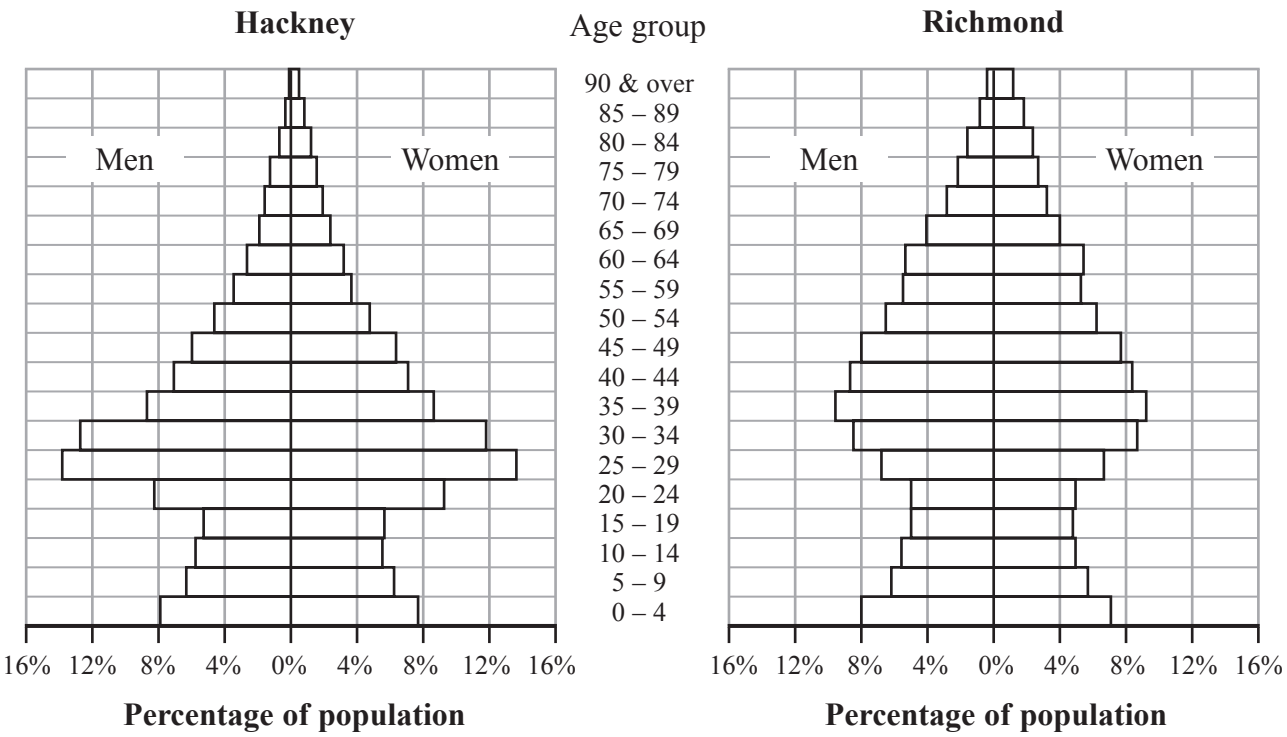
.....

.....

(2)

(Total for Question 2 is 9 marks)

3 The two population pyramids show the percentages of men and women in each age group in Hackney and in Richmond in 2011



(Data source: Office for National Statistics 2011 Census)

(a) Write down the age group that has the greatest percentages of both men and women for

(i) Hackney,

.....

(ii) Richmond.

.....

(2)

In Richmond, 4% of men and 4% of women are in the same age group.

(b) Write down this age group.

.....

(1)

(c) Compare the percentage of people aged 60 and over in Hackney with the percentage of people aged 60 and over in Richmond.

.....

.....

.....

(1)

(Total for Question 3 is 4 marks)

4 The table shows the average monthly rental price for a two-bedroom property in the UK in 2010 and in 2012

Year	2010	2012
Average monthly rental price	£1042	£1187

(Data source: www.rentright.co.uk)

(a) Using 2010 as the base year, calculate the index number for the average monthly rental price in 2012

.....

(2)

The index number for the average monthly rental price for a three-bedroom property in the UK for the same period is 120

*(b) Compare how the monthly rental prices changed from 2010 to 2012 for two-bedroom and three-bedroom properties in the UK.

.....

.....

.....

.....

.....

.....

(2)

(Total for Question 4 is 4 marks)

5 The table gives information about the public spending (£ billions) by Central Government and by Local Government in 2005 and in 2010

Public spending (£ billions)	**2005**			**2010**		
	Central Government	**Local Government**	**Total**	**Central Government**	**Local Government**	**Total**
1. General public services	37.0	5.2	42.2	49.6	5.5	55.1
2. Defence	29.7	0.0	29.8	37.6	0.1	37.7
3. Public order and safety	14.0	14.5	28.5	17.0	17.2	34.2
4. Economic affairs	25.0	8.7	33.7	34.4	11.9	46.3
5. Environmental protection	2.5	4.5	7.0	4.8	6.2	11.0
6. Housing and community	3.7	3.8	7.6	7.2	5.1	12.3
7. Health	82.6	0.4	82.9	118.1	0.1	118.2
8. Recreation and culture	5.5	4.4	10.0	7.5	5.9	13.4
9. Education	23.0	42.1	65.1	33.2	55.1	88.4
10. Social protection	128.7	35.6	164.3	175.7	49.8	225.5
Other spending	7.8	13.0	20.7	11.8	16.0	27.7
Total public spending	359.6	132.2	491.8	497.0	172.8	669.8

(Data source: ukpublicspending.co.uk)

(a) Write down the amount spent on Health by Central Government in 2010

£ ..... billion

(1)

The largest amount spent by Local Government is in the same public spending category in both years.

(b) Write down this category.

.....

(1)

In 2010, the total public spending on Health was more than the total public spending on Education.

(c) How much more?

£ ..... billion

(1)

The figures in the table for Local Government spending in 2010 add to £172.9 billion.
The total given in the table for this spending is £172.8 billion.

(d) Give a reason why these totals are different.

.....

.....

(1)

(e) Find the percentage increase in total public spending by Central Government between 2005 and 2010

..... %

(2)

(Total for Question 5 is 6 marks)

6 Serika did a survey to find out how many online friends people have.

Here are her results.

Number of friends	Frequency (f)
1 to 80	29
81 to 160	16
161 to 240	12
241 to 320	6
321 to 400	5
401 to 480	2

(a) Write down the modal class interval.

.....

(1)

Using x for the midpoint of each class interval, Serika found that $\Sigma fx = 9875$

(b) (i) Calculate an estimate for the mean number of friends.

.....

(2)

(ii) Explain why your answer to (i) is only an estimate.

.....

.....

.....

(1)

Serika considered using a class width of 20 to find her estimate for the mean.

(c) Describe the likely effect of this change on the estimate for the mean.

.....

.....

(1)

Charles thinks that the median is more appropriate than the mean as a measure of central tendency for Serika's data.

(d) Explain why he might think this.

.....

.....

.....

(1)

(Total for Question 6 is 6 marks)

7 Here is a list of statistical techniques which you can use to investigate data.

A Comparative box plots

B Composite percentage bar chart

C Line of best fit

D Pictogram

E Scatter diagram

F Spearman's rank correlation coefficient

G Standard deviation

H Standardised scores

From the list above, choose the best technique to use for each of the following.

(a) To see if there is a relationship between an athlete's height and their time to run 100 m.

.....

(1)

(b) To compare the central tendency, dispersion and skewness of two distributions.

.....

(1)

(c) To compare different groups within two distributions which are not the same size.

.....

(1)

(d) To find how strong the agreement is between two judges of a dance competition.

.....

(1)

(e) To compare the performance of a student in an English exam with their performance in a History exam.

.....

(1)

(Total for Question 7 is 5 marks)

8 A researcher is investigating how safe residents of different ages feel in their community at different times of the day.

(a) Suggest a hypothesis the researcher can use.

.....

.....

(1)

The researcher designs a questionnaire.

She decides to do a pilot study.

(b) Give two reasons why she should do a pilot study.

Reason 1 .....

.....

.....

Reason 2 .....

.....

.....

(2)

The researcher is going to use the questionnaire with a sample of residents in a small town.

She uses the local telephone directory as her sampling frame.

*(c) Explain whether or not this is a good choice of sampling frame.

.....

.....

.....

.....

.....

(2)

(Total for Question 8 is 5 marks)

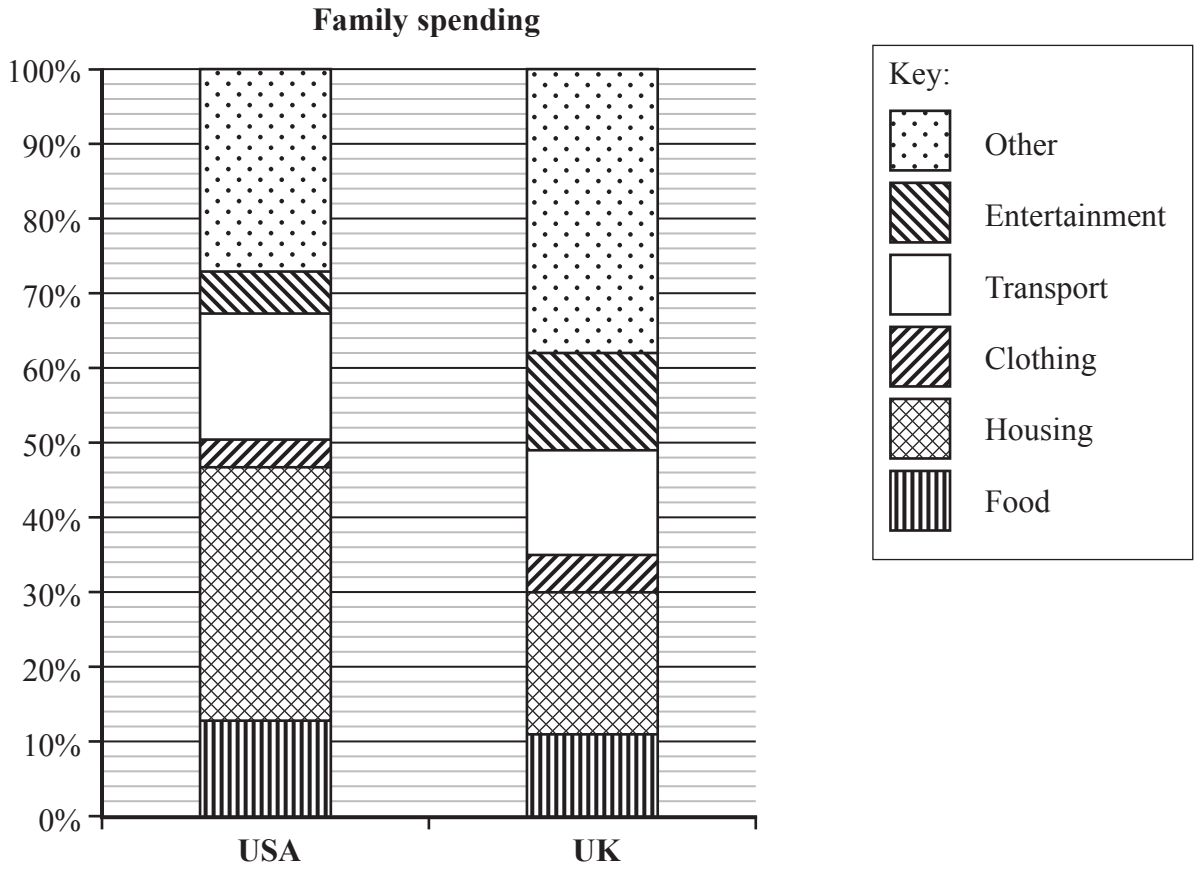
9 Tobi is investigating family spending in the USA and in the UK.

Two of his hypotheses are

Hypothesis 1 The percentage of family spending on housing is greater in the USA than in the UK

Hypothesis 2 The percentage of family spending on entertainment is greater in the USA than in the UK

Tobi collects data from the internet and draws these composite percentage bar charts.



(Data sources: USA – Bureau of Labor Statistics
UK – Office for National Statistics)

(a) Find the percentage of family spending on entertainment in the USA.

..... %

(2)

*(b) Explain, with reasons, whether or not the bar charts support each of Tobi's hypotheses.

Hypothesis 1 .....

.....

.....

.....

.....

Hypothesis 2 .....

.....

.....

.....

.....

(4)

Tobi thinks that the composite bar charts show that family spending on food is more in the USA than in the UK.

(c) Explain why Tobi is not correct.

.....

.....

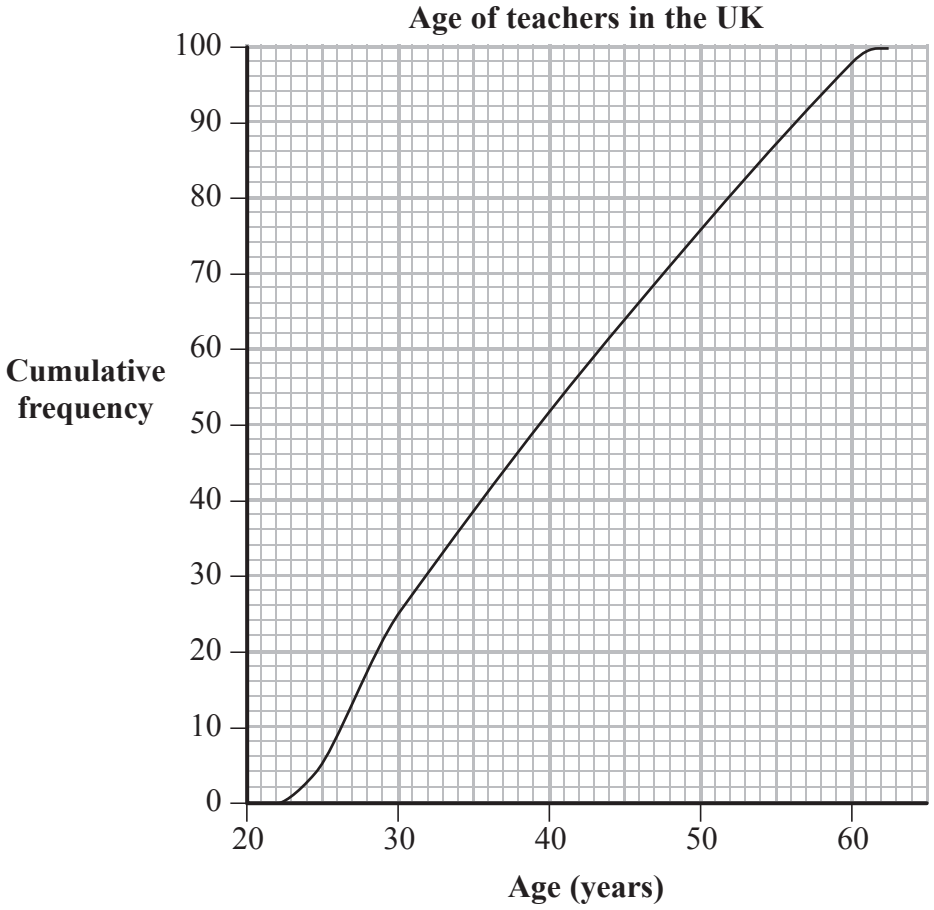
.....

.....

(1)

(Total for Question 9 is 7 marks)

10 The cumulative frequency diagram shows the distribution of ages of a sample of 100 teachers in the UK.

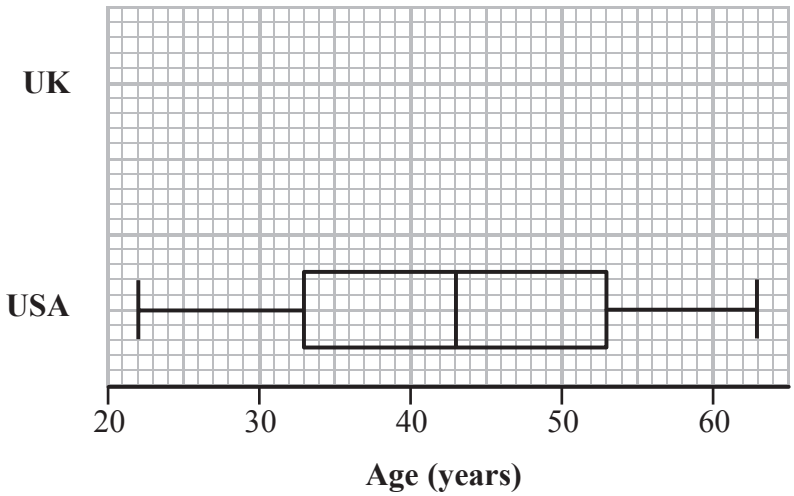


(a) Complete this table for the ages of these 100 teachers.

Lowest	Lower Quartile	Median	Upper Quartile	Highest
22				62

(2)

The box plot shows the distribution of ages of a sample of teachers in the USA.



(b) On the same grid, draw a box plot for the ages of the sample of teachers in the UK.

(2)

A Headteacher thinks that teachers in the UK are younger than teachers in the USA.

(c) Use the box plots to compare the distributions of the ages of the teachers in the UK and the ages of the teachers in the USA.

Is the Headteacher correct?

(4)

(Total for Question 10 is 8 marks)

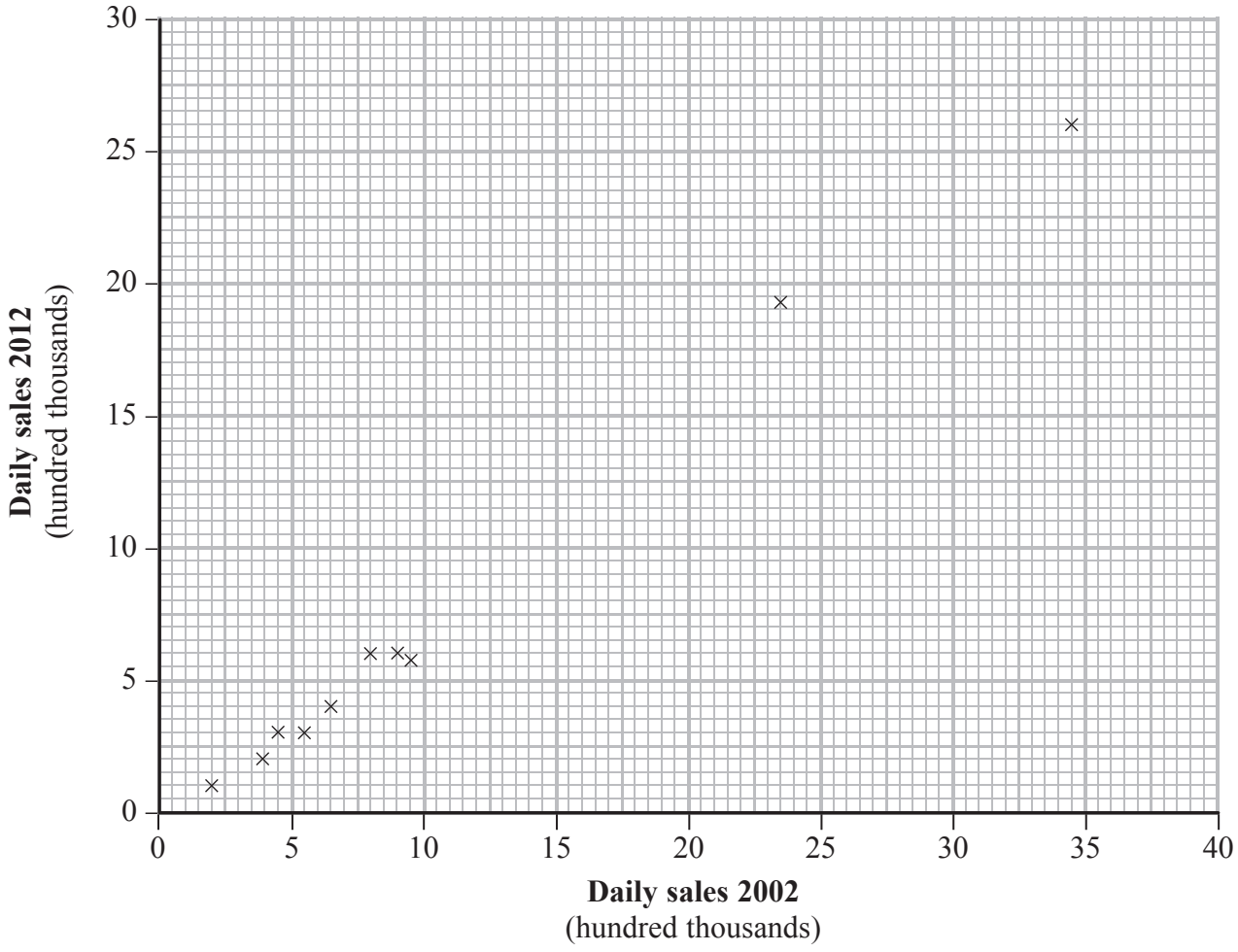
11 Daily newspaper sales fell between 2002 and 2012

The table gives daily sales for some national newspapers in these two years.

Newspaper	**Daily sales 2002** (hundred thousands)	**Daily sales 2012** (hundred thousands)
The Sun	34.6	26.1
Daily Star	7.9	6.1
Daily Record	5.6	2.8
Daily Mail	23.5	19.3
Express	9	6
Telegraph	9.5	5.8
The Times	6.6	4
FT	4.5	3
The Guardian	3.8	2.1
The Independent	1.9	0.9

(Data source: Audit Bureau of Circulations)

Here is a scatter diagram for this information.



Rupert says there is a strong relationship between daily sales in 2002 and 2012 for all these newspapers.

(a) Does the scatter diagram support what Rupert says?
Explain why you think this.

(2)

The daily sales for the Daily Mirror were

2 150 000 in 2002

1 100 000 in 2012

(b) Plot the point for the Daily Mirror on the scatter diagram.

(1)

(c) (i) Describe how the point for the Daily Mirror fits with the other data.

(ii) Discuss the change in sales for the Daily Mirror, between 2002 and 2012, compared with other newspapers.

(3)

(Total for Question 11 is 6 marks)

12 A scientist wants to estimate the number of fish in a disused canal.

He catches a sample of 30 fish from the canal.
He marks each fish with a dye and then puts them back in the canal.

The next day the scientist catches 20 fish from the canal.
He finds that 4 of them are marked with the dye.

(a) Estimate the total number of fish in the canal.

.....

(2)

(b) Write down any assumptions you made.

.....

.....

.....

.....

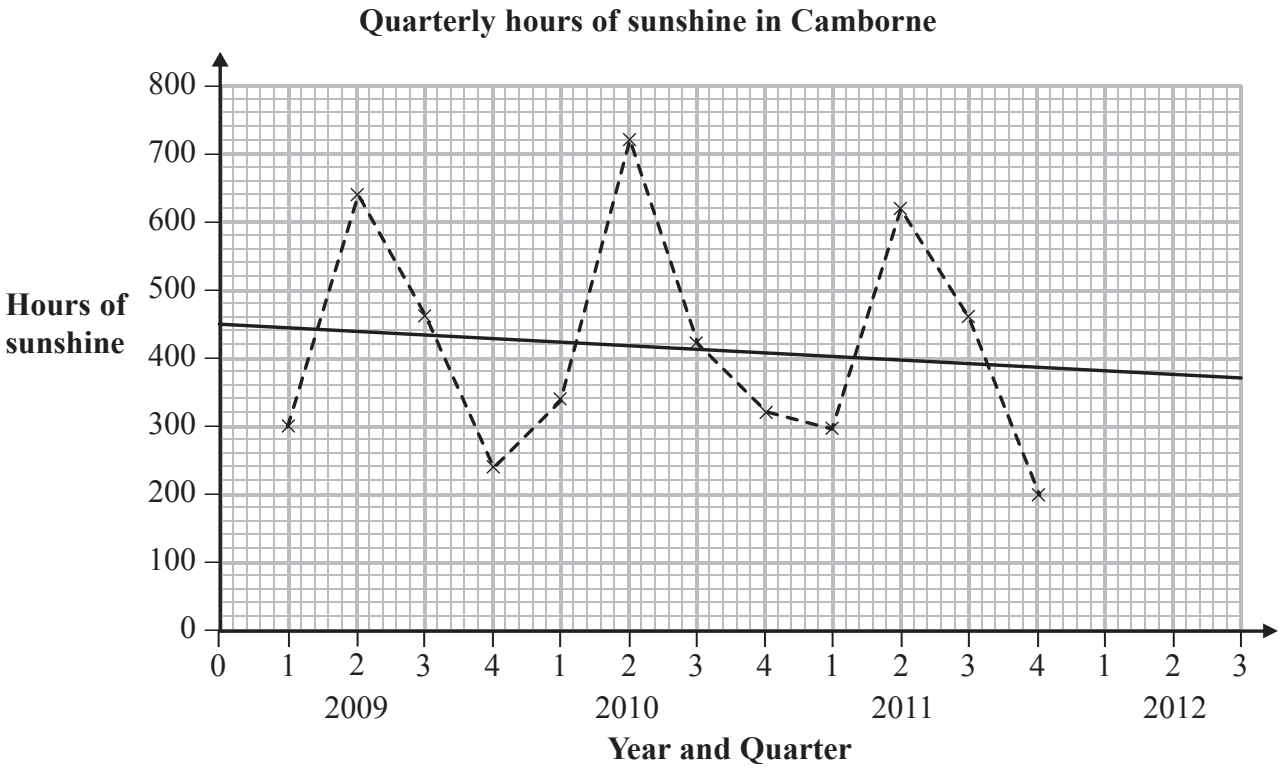
.....

(2)

(Total for Question 12 is 4 marks)

13 The total number of hours of sunshine in Camborne was recorded each quarter over three years.

The time series graph shows the results.



(Data source: www.metoffice.gov.uk)

A trend line has been drawn on the time series graph.

(a) Work out the average seasonal effect for Quarter 2
Give your answer to the nearest whole number.

..... hours

(2)

(b) Use your answer to (a) to predict the total number of hours of sunshine for Quarter 2 of 2012

..... hours

(2)

(Total for Question 13 is 4 marks)

14 SkyLights is a company that makes fireworks.

Sampling is used to test the fireworks when they are set off.

(a) Explain why a census would not be used.

.....

.....

.....

(1)

One day the company makes 3000 of the same type of firework.

A systematic sample of 1% is going to be taken for testing.

*(b) (i) Describe in detail how this sample should be selected.

.....

.....

.....

.....

.....

.....

.....

.....

(ii) Give two disadvantages of using this sampling method.

Disadvantage 1 .....

.....

Disadvantage 2 .....

.....

(4)

(Total for Question 14 is 5 marks)

15 A farmer supplies both free-range eggs and barn eggs.

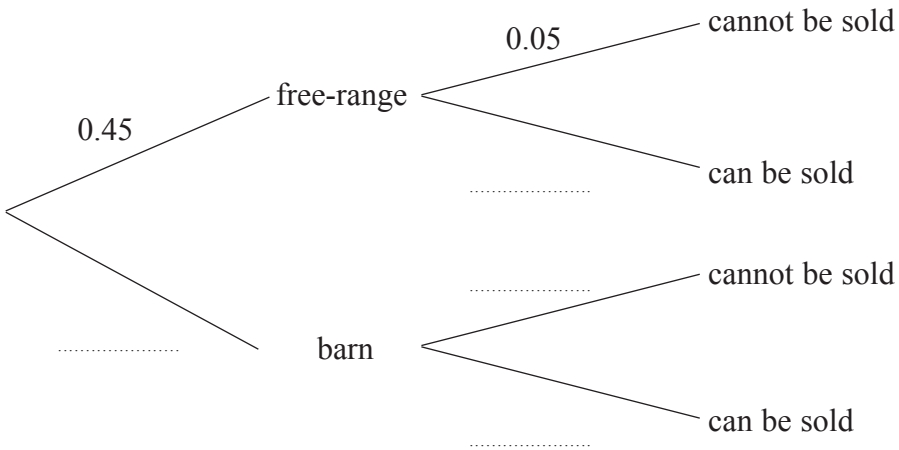
45% of the eggs are free-range. The rest are barn eggs.

An egg cannot be sold when it does not meet a particular standard.

5% of the free-range eggs cannot be sold.

8% of the barn eggs cannot be sold.

(a) Complete the probability tree diagram to show this information.



(2)

One egg is selected at random.

(b) Find the probability that it cannot be sold.

.....

(3)

One of the eggs that cannot be sold is selected at random.

(c) Find the probability that it is a free-range egg.

.....

(2)

(Total for Question 15 is 7 marks)

16 The table gives information about the birth weights (x kg) of 145 elephants.

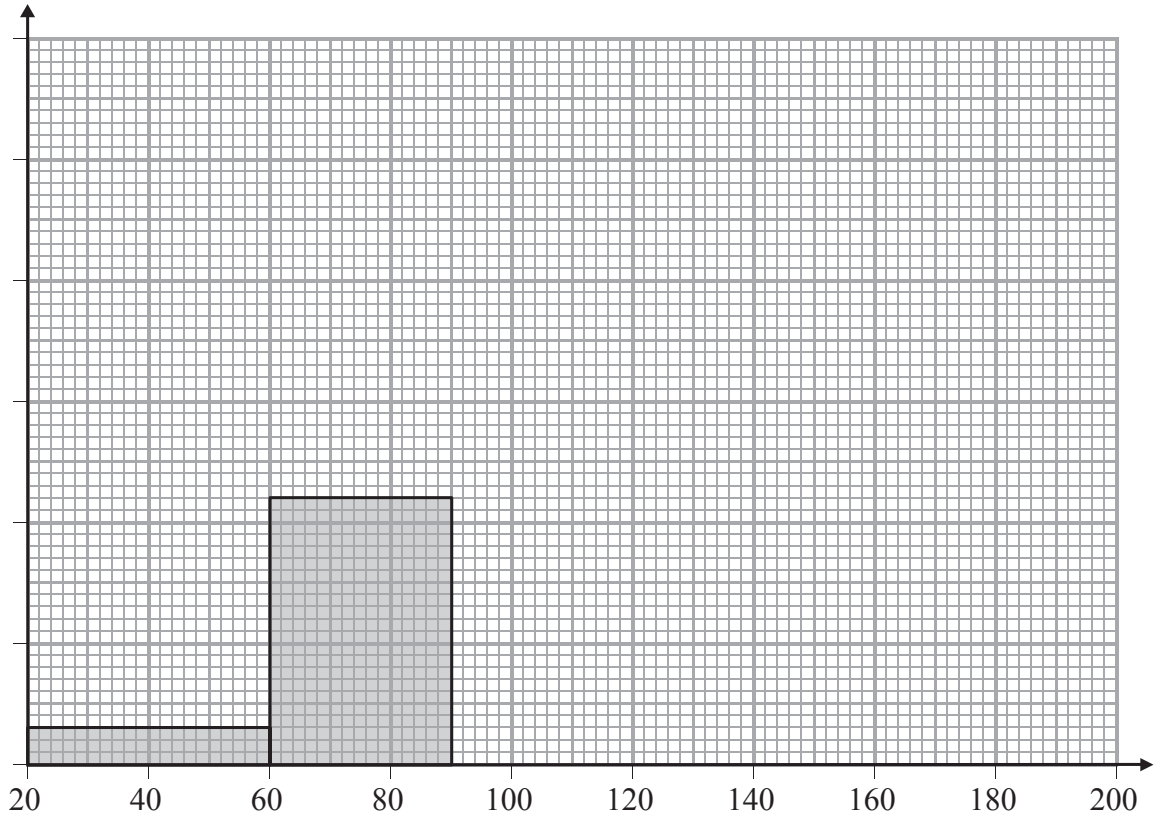
Birth weight (x kg)	**Frequency (f)**		
$20 < x \leqslant 60$	6		
$60 < x \leqslant 90$	33		
$90 < x \leqslant 110$	40		
$110 < x \leqslant 130$	37		
$130 < x \leqslant 160$	24		
$160 < x \leqslant 200$	5		

(Data source: www.elephant.se)

(a) Complete the histogram for this information and label the axes.

You may use the extra columns in the table.

(4)



For these 145 elephants $\Sigma fx = 15\,535$ and $\Sigma fx^2 = 1\,794\,625$

(b) Show that the standard deviation of the birth weights is 30 kg to the nearest kg.

(2)

The mean birth weight for these elephants is 107 kg to the nearest kg.

A zoologist thinks the birth weights are normally distributed.

(c) (i) Between what limits should 95% of the birth weights lie?

..... and .....

(3)

(ii) Comment on whether or not the zoologist is correct.

Give a reason for your answer.

.....

.....

.....

(1)

(Total for Question 16 is 10 marks)

17 Comfi Cabs always has five taxis working at the same time.

Each working taxi is either available or not available.
At any time, the probability that a taxi is **not** available is 0.4

(a) Write down the probability that a particular taxi is available.

.....

(1)

(b) Find the probability that all five taxis are not available at the same time.

.....

(2)

(c) Find the probability that exactly two of the five taxis are available at the same time.

You may use $(p + q)^5 = p^5 + 5p^4q + 10p^3q^2 + 10p^2q^3 + 5pq^4 + q^5$.

.....

(2)

The Smith family need two taxis at the same time to take them all to the airport.

(d) Find the probability that Comfi Cabs cannot take all the Smith family at the same time.

.....

(3)

(Total for Question 17 is 8 marks)

TOTAL FOR PAPER IS 100 MARKS

BLANK PAGE