Centre Number						Candidate Number				
Surname										
Other Names										
Candidate Signature										

For Examiner's Use	
Examiner's Initials	
Pages	Mark
3	
4 – 5	
6 – 7	
8 – 9	
10 – 11	
12 – 13	
14 – 15	
16 – 17	
18 – 19	
20 – 21	
22 – 23	
24 – 25	
26 – 27	
TOTAL	

General Certificate of Secondary Education
Higher Tier
June 2013

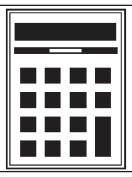
Statistics
Written Paper

43101H

H

Monday 24 June 2013 1.30 pm to 3.30 pm

For this paper you must have:
- a calculator
- mathematical instruments.



Time allowed
- 2 hours

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

Information
- The marks for questions are shown in brackets.
- The maximum mark for this paper is 100.
- You may ask for more answer paper, graph paper and tracing paper. These must be tagged securely to this answer booklet.
- You are expected to use a calculator where appropriate.

Advice
- In all calculations, show clearly how you work out your answer.

You may need to use the following formulae:

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.

Standard deviation for a set of numbers $x_1, x_2, \ldots, x_n$ having a mean value of $\bar{x}$ is given by

$$\sqrt{\frac{\sum (x - \bar{x})^2}{n}} \text{ or } \sqrt{\frac{\sum x^2}{n} - \bar{x}^2}$$

Standard deviation for a frequency distribution

$$\sqrt{\frac{\sum f(x - \bar{x})^2}{\sum f}} \text{ or } \sqrt{\frac{\sum fx^2}{\sum f} - \bar{x}^2}$$

The same formulae apply to the standard deviation of a grouped frequency distribution where x is the mid-interval value.

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

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

1 Josh is investigating how long people spend playing games on their phones.

Here are some of the tasks he does to investigate this.

They are in the wrong order.

A Josh writes his conclusion

B Josh suggests a hypothesis

C Josh calculates some averages

D Josh collects the data

E Josh selects his sample

Put these tasks in the correct order so that Josh has a suitable strategy for his investigation.

First task

Last task

(2 marks)

Turn over for the next question

2

2 Niles is investigating how much sport people play.
He will use a questionnaire.
First he will do a pilot study.

2 (a) What is a pilot study?

.....

.....

(1 mark)

2 (b) Why should Niles carry out a pilot study?

.....

.....

(1 mark)

2 (c) Niles looks at some of the pilot study data about the number of hours each person plays sport per week.

5 4 0 2 6 0 120 4 2 3

2 (c) (i) How many of these people claim to play no sport?

Answer ..... *(1 mark)*

2 (c) (ii) One of the values has been misrecorded.

Which one is it?

Misrecorded value .....

How do you know?

.....

.....

(2 marks)

2 (d) Niles' research question is

"Do men play more hours of sport than women?"

Apart from "hours of sport", what other variable **must** Niles record?

.....

(1 mark)

3 (a) At the beginning of 2012 the town of Brigg had a population of 6000.
In 2012 there were 63 births in the town.

Calculate the crude birth rate for Brigg in 2012.

.....

.....

Answer ..... per thousand *(2 marks)*

3 (b) In 2012 the crude death rate for Brigg was 8 per thousand.

Were there more births or deaths in Brigg in 2012?
Circle your answer.

Births Deaths

Give a reason for your answer.

.....

.....

(1 mark)

9

4 Here is some sample information from a survey about pets in England.

Question: Do you own a cat or dog?	
Cat	28.5%
Dog	24.8%
Neither	53.7%

Source: Ipsos Mori

4 (a) These percentages add up to 107% rather than 100%.

Why is this?

.....

.....

(1 mark)

4 (b) The proportion of people in the **survey** who own a dog is 0.248

Estimate the proportion of the **population** who own a dog.

Answer ..... *(1 mark)*

4 (c) People with cats or dogs were asked how many they owned.

CATS	
Number	Owned by
1	55.9%
2	29.5%
3	8.3%
4 or more	6.3%

DOGS	
Number	Owned by
1	74.8%
2	20.2%
3	3.3%
4 or more	1.7%

Source: Ipsos Mori

What percentage of dog owners own three or more dogs?

.....

Answer ..... % *(2 marks)*

4 (d) Using all three tables in this question, explain why there are probably more cats than dogs in England.

.....

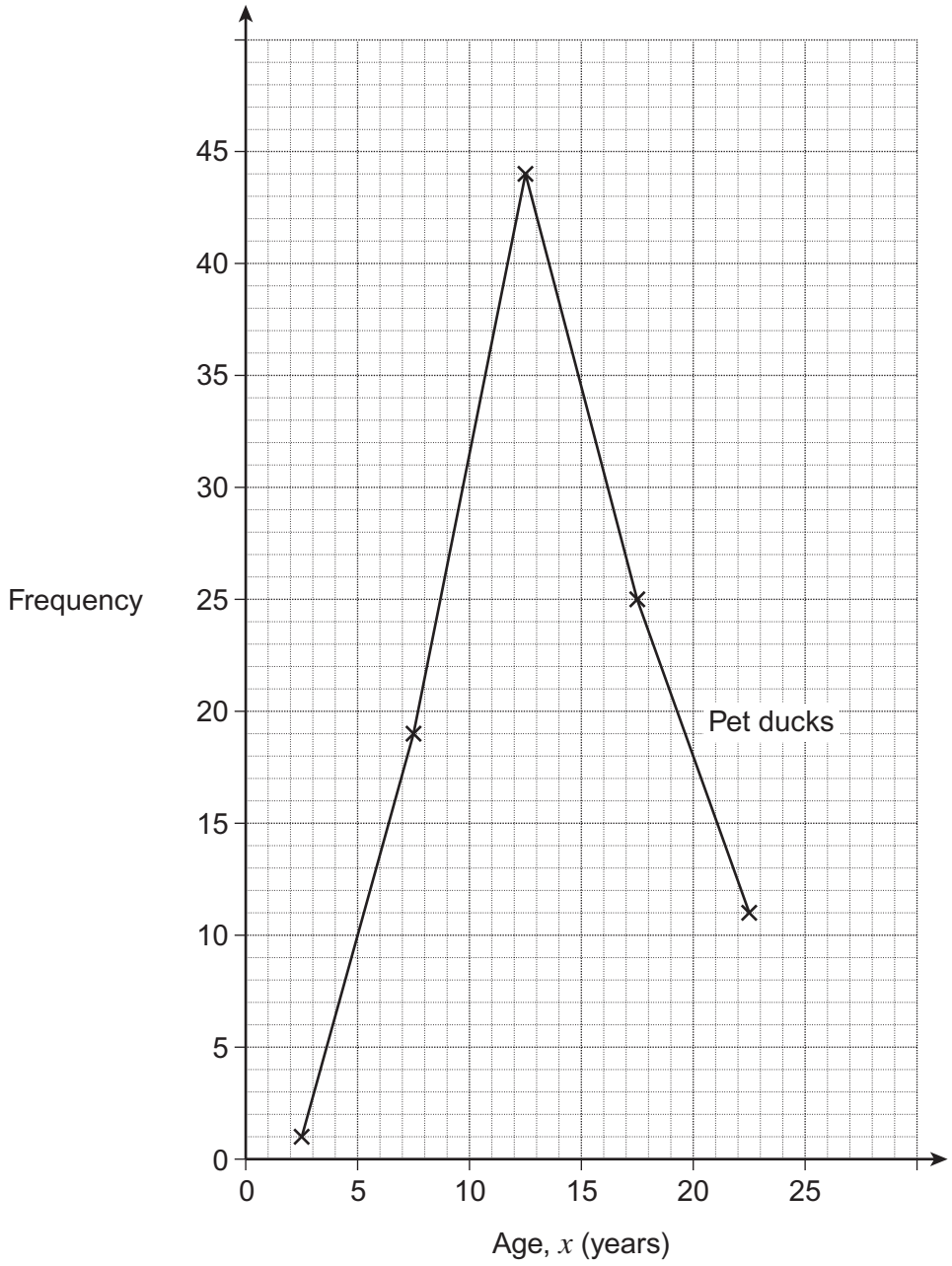
.....

.....

(2 marks)

6

5 The age to which 100 **pet** ducks live is shown on the frequency polygon.



The age to which 100 **wild** ducks live is shown in the table.

Age, x (years)	$0 \leqslant x < 5$	$5 \leqslant x < 10$	$10 \leqslant x < 15$	$15 \leqslant x < 20$
Frequency	18	40	38	4
Midpoint	2.5	7.5		

5 (a) (i) The midpoint of the group $0 \leqslant x < 5$ is 2.5 years.

Complete the table by writing in the **two** missing midpoints.

.....

.....

(1 mark)

5 (a) (ii) Use the midpoints to draw the frequency polygon for wild ducks on the same grid as for pet ducks.

(3 marks)

5 (b) Compare the age to which these pet ducks and wild ducks live.

Comparison 1 .....

.....

Comparison 2 .....

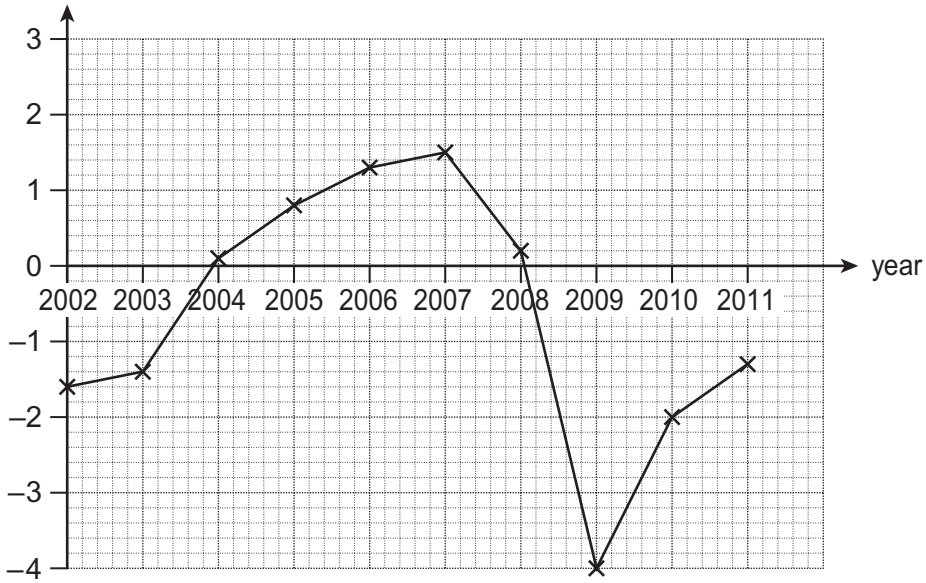
.....

(2 marks)

6

Turn over ▶

6 This is an output gap chart for the USA.



6 (a) The label for the vertical axis is missing.

What should it be?

Answer ..... *(1 mark)*

6 (b) In 2010 the USA was in recession.
How can you see this from the graph?

.....

.....

(1 mark)

6 (c) In a recession, what is likely to happen to unemployment?

Circle your answer.

Decreases Stays the same Increases

(1 mark)

7 A large number of people attend a music festival.

Festival organisers decide to survey a sample of people to find out what they think about buying tickets on-line.

They will interview every 50th person as they leave the festival arena.

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

Answer ..... *(1 mark)*

7 (b) Suggest **two** practical difficulties in conducting the interviews.

Difficulty 1 .....

.....

Difficulty 2 .....

.....

(2 marks)

7 (c) Describe the main difference between the method of sampling used and a quota sample.

.....

.....

.....

(1 mark)

Turn over for the next question

7

8 A clothing store stocks coats in three different sizes (small, medium and large) and three different colours (black, grey and brown).

At the start of last week there were 100 coats in stock as shown in the table.

	Colour		
Size	**Black**	**Grey**	**Brown**
Small	11	25	10
Medium	9	12	8
Large	0	7	18

8 (a) Identify from the table **two** categories of coat that are mutually exclusive.

Answer ..... and ..... *(1 mark)*

8 (b) As part of a quality check, an assistant selects at random one of the 100 coats.

Calculate the probability that the coat chosen is

8 (b) (i) black

.....

Answer ..... *(2 marks)*

8 (b) (ii) medium

.....

Answer ..... *(2 marks)*

8 (b) (iii) brown **or** small

.....

Answer ..... *(3 marks)*

8 (b) (iv) large, given that it is grey.

.....

.....

Answer ..... *(2 marks)*

8 (c) Later, 3 of these 100 coats are found to have faults in the material.

Calculate the probability that none of them is a small, grey coat.

.....

.....

.....

.....

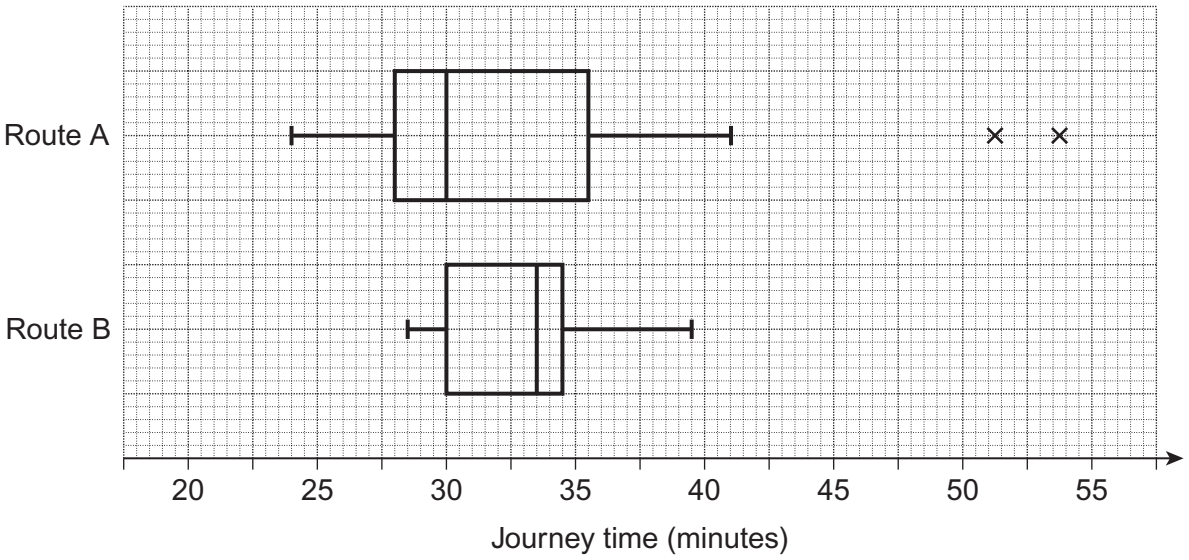
Answer ..... *(3 marks)*

Tun over for the next question

9 Albina drives from work to home each evening.

She takes either route A or route B.

For each route, the times taken, in minutes, for a sample of 40 journeys are summarised in the box and whisker plot.



9 (a) Compare the journey times for both routes.

Comparison 1 .....

.....

Comparison 2 .....

.....

Comparison 3 .....

.....

(3 marks)

9 (b) In each case, state which route Albina should take and give a reason for your choice.

9 (b) (i) She needs to be at home within 30 minutes of leaving work.

Route .....

Reason .....

.....

(1 mark)

9 (b) (ii) She needs to be at home within 40 minutes of leaving work.

Route .....

Reason .....

.....

(1 mark)

Turn over for the next question

5

10 Sarah works as a receptionist at a doctor's surgery.

Each day she records the number of missed appointments.

The table shows the number of missed appointments for a 90-day period.

Number of missed appointments (x)	0	1	2	3	4	5	6	7	8
Number of days (f)	5	7	21	18	15	7	6	7	4

10 (a) Write down

10 (a) (i) the mode

Answer ..... *(1 mark)*

10 (a) (ii) the range of the number of missed appointments.

Answer ..... *(1 mark)*

10 (b) You are given that

$$\sum fx = 315 \qquad \sum fx^2 = 1483$$

10 (b) (i) Calculate the mean number of missed appointments.

.....

.....

Answer ..... *(2 marks)*

10 (b) (ii) Calculate the standard deviation of the number of missed appointments.

.....

.....

.....

Answer ..... *(3 marks)*

10 (c) Is the mean a suitable measure of average in this case?

Tick a box. ☐ Yes ☐ No

Give a reason for your answer.

Reason .....

(1 mark)

10 (d) Later Sarah finds that she has made some mistakes.

All 5 of the days with 'no appointments missed' should have been recorded as 'four appointments missed'.

She corrects the mistakes.

Tick a box to show what happens to the following measures after the corrections have been made.

	Increases	Stays the same	Decreases
Mode	☐	☐	☐
Mean	☐	☐	☐
Range	☐	☐	☐

(3 marks)

11

11 Here is information about migration to the United Kingdom.

Migration to the United Kingdom by age and sex (thousands)									
Year	All ages			Under 45			45 and over		
	Total	Males	Females	Total	Males	Females	Total	Males	Females
2000	479	272	207	440	247	193	39	25	14
2001	519	260	259	484	238	246	35	22	13
2002	532	284	248	499	269	230	33	15	18
2003	586	300	286	546	280	266	40	20	20
2004	563	310	253	521	291	230	42	19	23
2005	589	310	279	554	288	266	35		13
2006	591	315	276	554	293	261	37	22	15

Source: Adapted from ONS 2007

11 (a) What was the total number of females aged under 45 that migrated to the United Kingdom between the years 2001 and 2003 inclusive?

.....

.....

Answer ..... thousand *(2 marks)*

11 (b) Calculate the missing number in the table.

.....

.....

Answer ..... thousand *(2 marks)*

11 (c) (i) Describe the trend in total migration to the United Kingdom over the period for people of all ages.

.....

.....

(1 mark)

11 (c) (ii) Suggest a possible reason for the trend.

.....

.....

(1 mark)

11 (c) (iii) State the year that does not follow this trend.

Answer ..... *(1 mark)*

11 (d) Work out the percentage decrease in total migration to the United Kingdom between 2004 and 2006 for people '45 and over'.

Give your answer to a suitable degree of accuracy.

You must show your working.

.....

.....

Answer ..... % *(4 marks)*

Turn over for the next question

11

12 A park in the Scottish Highlands has an unknown number of deer.

A random sample of 50 deer is caught and marked with a dye before being returned to the park.

Some days later, six samples, each of size 25, are taken from the park.

The number of deer in each sample having a dye mark is as follows:

2 5 3 1 6 0

12 (a) Use these results to estimate the total number of deer in the park.

.....

.....

.....

Answer ..... *(4 marks)*

12 (b) State a condition that must be met to ensure the validity of the estimation method used in part (a).

.....

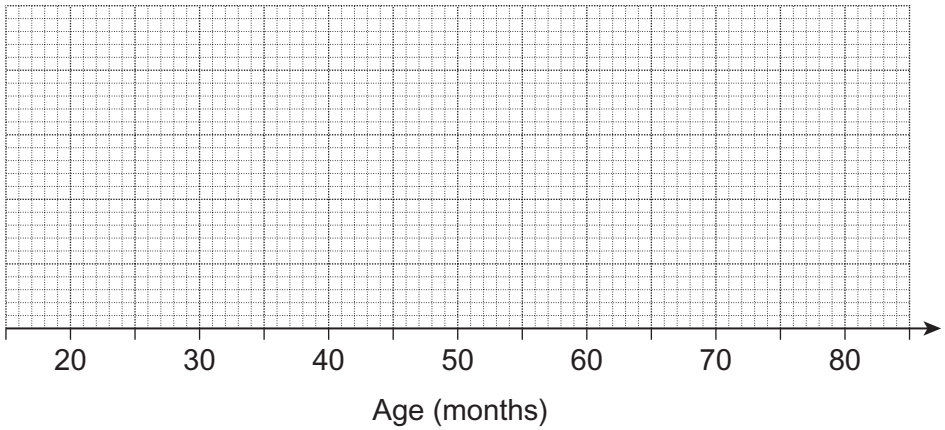
.....

.....

(1 mark)

12 (c) An earlier survey of the ages of deer in the park had shown a mean age of 50 months with standard deviation of 8 months.

Assuming that the distribution of ages is Normal, sketch this distribution on the grid below.



(3 marks)

Turn over for the next question

8

13 A local firm makes cheese.

The table shows the sales of cheese, in tonnes per quarter, from Quarter 1 of 2011 to Quarter 2 of 2013.

The first four moving average values have been calculated and entered in the table.

Year	Quarter (Q)	Sales (tonnes)	Moving Averages
2011	1	8.0	
	2	12.0	
			10.8
	3	14.2	
			11.4
	4	9.0	
			12.3
2012	1	10.4	
			13.5
	2	15.6	
			.....
	3	19.0	
			.....
	4	11.8	
			.....
2013	1	15.2	
	2	19.2	

(Note: The figure of 19.2 tonnes for Quarter 2 of 2013 represents expected sales for that quarter).

13 (a) Calculate the missing moving average values.
Write your answers in the table.

.....

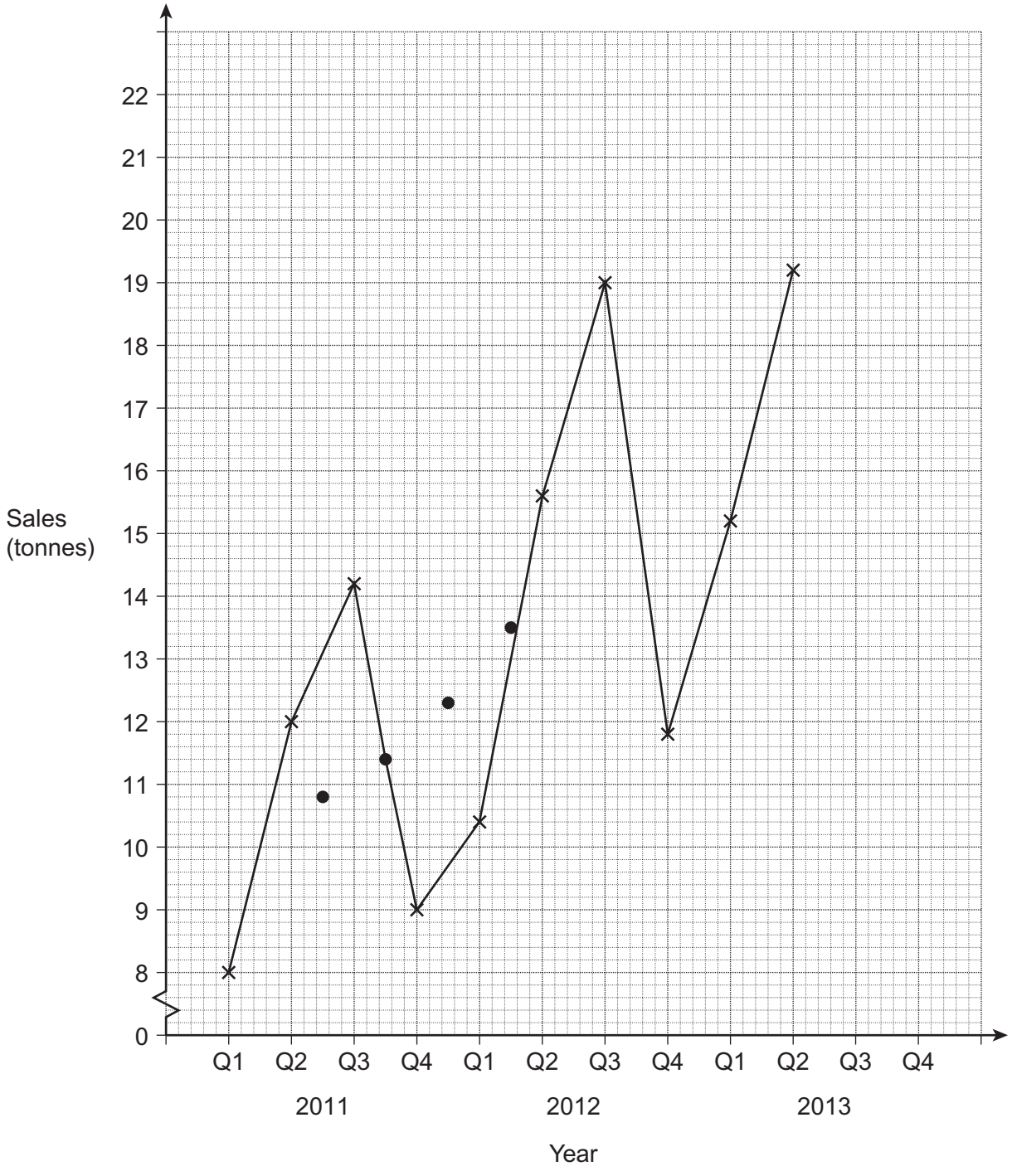
.....

.....

(3 marks)

13 (b) The sales data, together with the first four moving averages, are plotted on the grid.

Plot the remaining moving averages on this grid. *(2 marks)*



13 (c) Draw the trend line. *(1 mark)*

6

Turn over ▶

13 (d) The seasonal effects for Quarter 3 and Quarter 4 are

Quarter 3	Quarter 4
+ 3.65	– 3.49

Use these and your trend line to predict the likely sales for Quarter 3 and Quarter 4 of 2013.

.....

.....

.....

Sales in Quarter 3 ..... tonnes

Sales in Quarter 4 ..... tonnes *(4 marks)*

13 (e) The owners plan to close the business if they do not reach total sales of at least 85 tonnes during 2013.

Advise the owners whether they are likely to achieve this level of total sales.

You **must** show your working.

.....

.....

.....

(2 marks)

14 A sports coach analyses the goal scoring record of a large sample of Premier League Football players over a season.

Part of his analysis involves calculating values for the Product Moment Correlation Coefficient (r).

14 (a) r = +0.65 between number of shots on goal and goals scored.

This is probably correct.

Explain why.

.....

.....

(1 mark)

14 (b) r = – 0.83 between boot size and goals scored.

This is probably incorrect.

Explain why.

.....

.....

(1 mark)

14 (c) r = – 1.14 between age and goals scored.

This is definitely incorrect.

Explain why.

.....

.....

(1 mark)

Turn over for the next question

9

15 The table relates to the Consumer Prices Index for January 2010 (with 2005 as base year).

	Group	Weight	Index
1	Food	108	124.3
2	Alcohol and tobacco	40	119.5
3	Clothing and footwear	56	75.2
4	Housing, water, electricity and gas	129	130.7
5	Furniture	64	107.3
6	Health	22	114.6
7	Transport	164	118.2
8	Communication	25	97.6
9	Recreation	150	99.9
10	Restaurants and hotels	126	114.7
11	Miscellaneous goods and services	116	128.4
		TOTAL = 1000	

Source: Monthly Digest of Statistics, February 2010

15 (a) Explain what the Index value of 75.2 for 'Clothing and footwear' indicates.

.....

.....

(1 mark)

15 (b) A recent survey suggested that at least 25% of household expenditure went on 'Housing, water, electricity and gas'.

Show whether the weighting in the table supports this view.

.....

.....

(1 mark)

15 (c) (i) The 'All Groups' weighted index for January 2010 is 114.8

Recalculate this index when expenditure on 'Food' is excluded.

.....

.....

.....

.....

.....

.....

.....

Answer ..... *(4 marks)*

15 (c) (ii) Give a reason why the value of this index decreases when expenditure on 'Food' is excluded.

.....

.....

.....

(1 mark)

END OF QUESTIONS

7

There are no questions printed on this page

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED