

Surname						Other Names					
Centre Number						Candidate Number					
Candidate Signature											

For Examiner's Use
--------------------

General Certificate of Secondary Education  
June 2008



**STATISTICS**  
**Higher Tier**

**3311/H**  
**H**

Wednesday 18 June 2008 9.00 am to 11.30 am

<p><b>For this paper you must have:</b></p> <ul style="list-style-type: none"> <li>• a calculator</li> <li>• mathematical instruments.</li> </ul>	
---	--

For Examiner's Use	
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	
Examiner's Initials	

Time allowed: 2 hours 30 minutes

**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 in the spaces provided.
- You must answer the questions in the spaces provided. Answers written in margins or on blank pages will not be marked.
- Do all rough work in this book.

**Information**

- The maximum mark for this paper is 120.
- The marks for questions are shown in brackets.
- Additional answer paper and graph paper will be issued on request and must be tagged securely to this answer book.
- You are expected to use a calculator where appropriate.

**Advice**

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



J U N 0 8 3 3 1 1 H 0 1

You may need to use the following formulae:

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

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

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

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

Standard deviation for a frequency distribution

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

The same formula applies to the standard deviation of a grouped frequency distribution where  $x$  is the mid-interval value.

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



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

- 1 The table shows information about the length of time patients had been waiting for treatment in National Health Service hospitals (NHS), by region, at the end of March 2005.

**Time spent waiting for NHS treatment**

	Less than 6 months (%)	6 to 12 months (%)	Longer than 12 months (%)	Number of patients (in thousands)
<b>England</b>	75	23	2	1012
North Yorkshire	77	22	1	122
North West	78	20	2	160
Trent	82	16	2	100
West Midlands	83	16	1	85
Eastern	75	22	3	124
London	75	22	3	129
South East	73	23	4	189
South West	75	21	3	103
<b>Wales</b>	68	18	14	71
<b>Scotland</b>	80	17	3	72

Source: Adapted from Social trends No. 35

- 1 (a) What percentage of patients in the South East region had been waiting 6 to 12 months for treatment?

Answer ..... % (1 mark)

- 1 (b) Using the percentage figures for regions of England, select the region which appears to have the shortest recorded waiting times at the end of March 2005.

Answer ..... (1 mark)

- 1 (c) Give a reason why the total for the South West Region does not equal 100%.

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

**Question 1 continues on the next page**



1 (d) How many patients had been waiting for less than 6 months for treatment in the North West region?

.....  
.....

Answer ..... (3 marks)

The multiple bar chart opposite shows the percentage waiting times for all NHS patients in England and Wales.

1 (e) Complete the multiple bar chart to show the percentage waiting times for NHS patients in Scotland.

(3 marks)

1 (f) State **two** differences in the waiting times shown in the multiple bar chart.

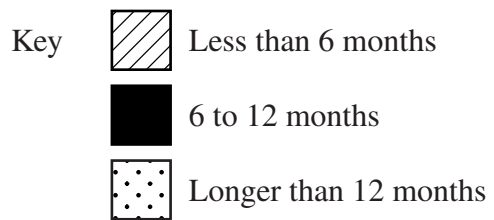
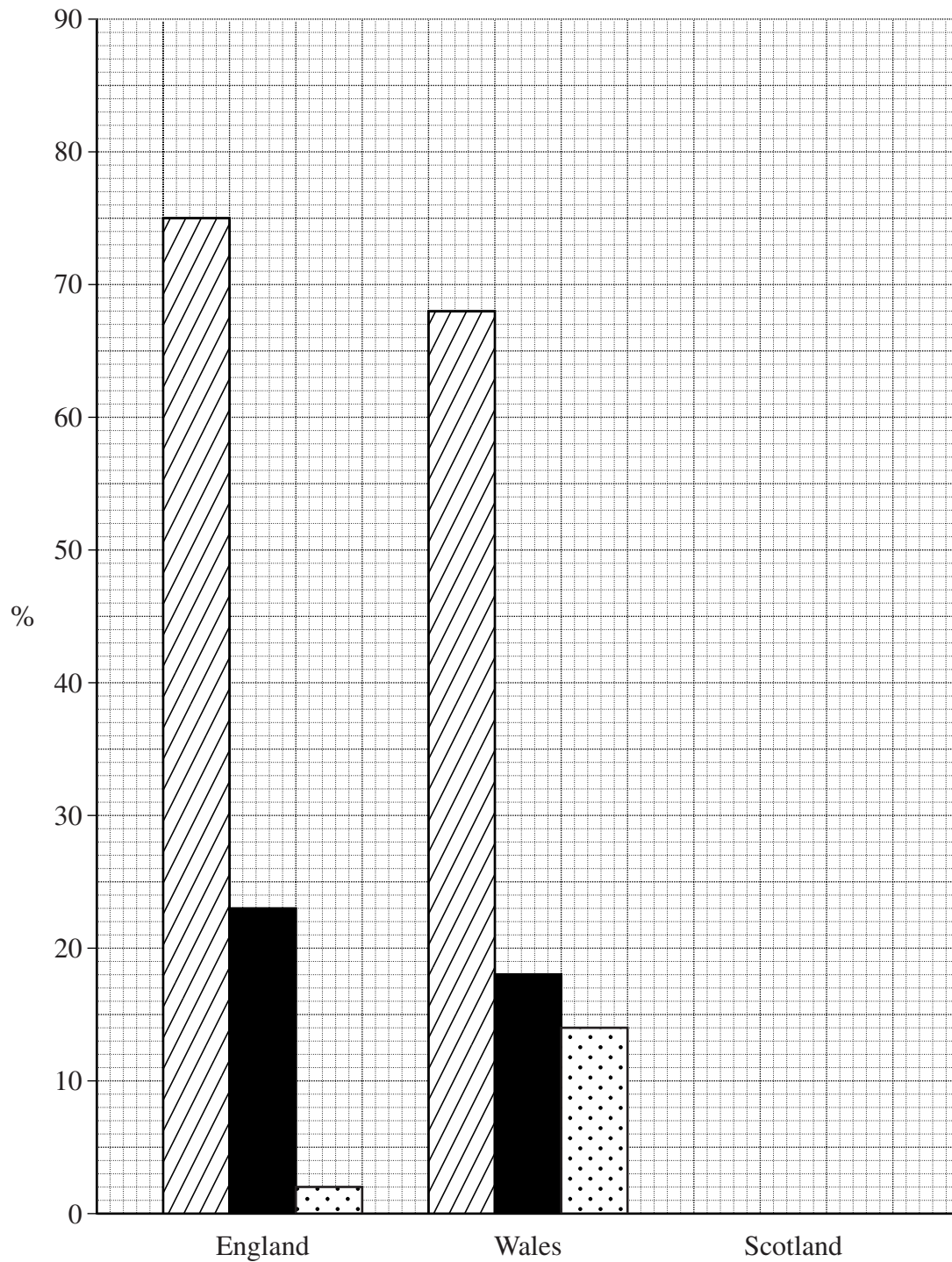
Difference 1 .....  
.....

Difference 2 .....  
.....

(2 marks)



### National Health Service Waiting Times

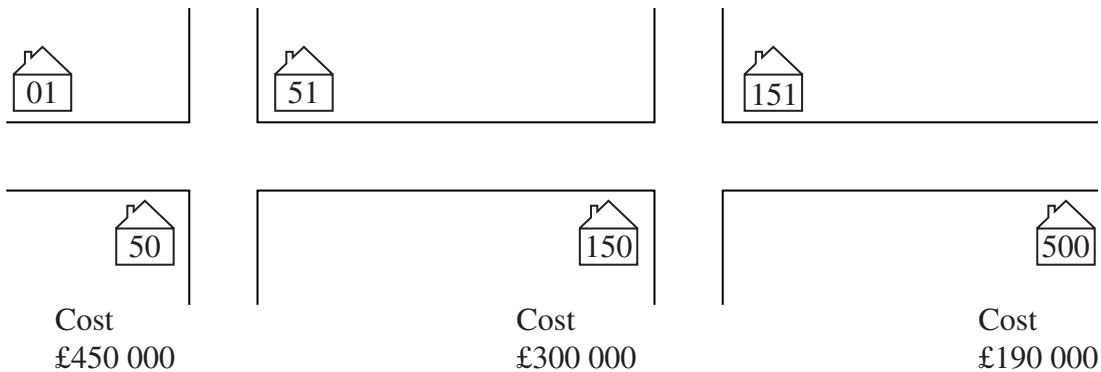


Turn over ►



2 The diagram shows the layout of a new housing estate in Stokeville.

Houses numbered 01 to 50	cost £450 000 each
Houses numbered 51 to 150	cost £300 000 each
Houses numbered 151 to 500	cost £190 000 each



The house builder agrees to undertake a sample survey to find out residents' views on introducing parking charges on the estate.

He decides to question one female from each of the houses numbered 31 to 50 on the estate.

2 (a) Give **two** reasons why this method of sampling is unsuitable.

Reason 1 .....

.....

Reason 2 .....

.....

(2 marks)



2 (b) At a later date the builder extends the survey by selecting a simple random sample of 50 houses from the 500 on the estate.

2 (b) (i) Explain how the sample could be chosen in this case.

.....  
.....  
.....  
.....

(2 marks)

2 (b) (ii) Why might a stratified sample be more appropriate in this case?

.....  
.....  
.....

(1 mark)

2 (b) (iii) In a stratified sample, how many houses costing £300 000 would be included in a sample of 50?

.....  
.....

Answer ..... (2 marks)

**Question 2 continues on the next page**



2 (c) One of the classification questions used in the survey was

‘How much do you pay each month on your mortgage?’

under £200

£200 - £300

£300 - £700

over £700

2 (c) (i) Give **one** criticism of the response section.

.....

.....

*(1 mark)*

2 (c) (ii) Give **one** criticism of the question asked.

.....

.....

*(1 mark)*





**Turn over for the next question**

**DO NOT WRITE ON THIS PAGE  
ANSWER IN THE SPACES PROVIDED**

**2**

**Turn over ►**

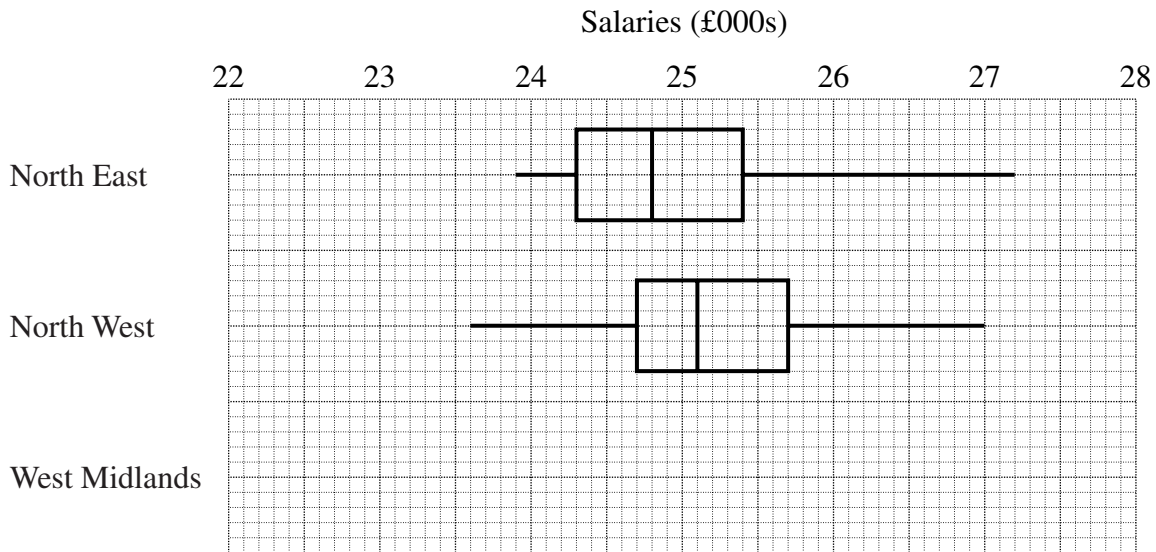


3 Sunita is an unemployed IT consultant.

She is planning to move to the North East, North West or West Midlands region of the UK to find work.

She downloads from the Internet details of twenty-five vacancies for IT staff in each region.

A summary of the salaries in the North East and North West regions is illustrated in the box plots below.



3 (a) What is the range of salaries in the North East region?

.....

.....

Answer £ ..... (1 mark)

3 (b) Find the interquartile range of salaries in the North West Region.

.....

.....

Answer £ ..... (2 marks)



3 (c) Sunita summarised her results from the West Midlands region.

The median salary on offer was £25 000, the lower quartile salary was £23 400, and the interquartile range was £3100.

The highest salary was £26 900 and the lowest £22 300.

Use these results to draw a box plot on the grid.

(3 marks)

.....  
.....

3 (d) Sunita does not want to move to a region where the range in salaries exceeds £4500 or where the median salary is below £25 000.

3 (d) (i) State which region she should choose.

Answer ..... (1 mark)

3 (d) (ii) Explain why the other two regions would not be selected.

Region .....

Reason .....

.....

Region .....

Reason .....

.....

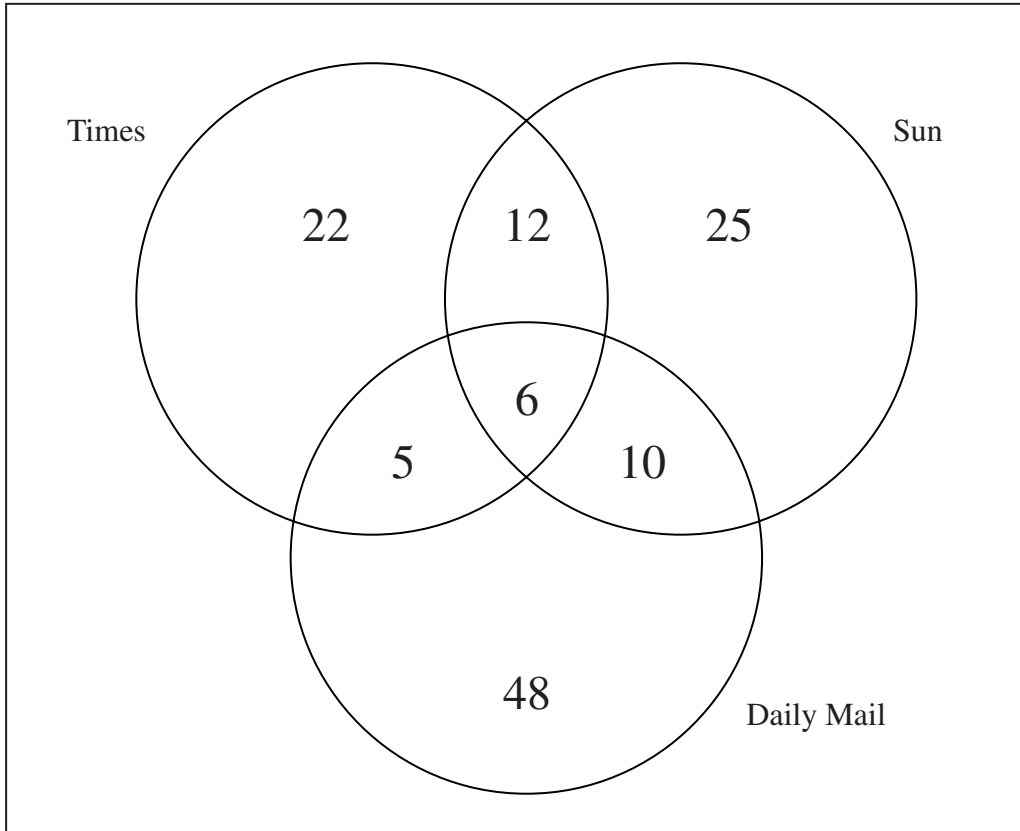
(2 marks)

**Turn over for the next question**



4 150 students were asked which daily newspaper(s) they read.

The results are shown in the diagram.



4 (a) Find the probability that a student chosen at random reads

4 (a) (i) the Times

.....

Answer ..... (1 mark)

4 (a) (ii) only one of the papers

.....

.....

Answer ..... (2 marks)



4 (a) (iii) none of the three papers

.....  
.....  
.....

Answer ..... (3 marks)

4 (a) (iv) the Sun or the Times or both but not the Daily Mail.

.....  
.....

Answer ..... (2 marks)

4 (b) One of the students, Jean, reads the Daily Mail.

Find the probability that she also reads the Times.

.....  
.....  
.....

Answer ..... (3 marks)

4 (c) State whether each of the following variables is qualitative, discrete or continuous.

4 (c) (i) The number of pages in each newspaper.

Answer ..... (1 mark)

4 (c) (ii) The length of time spent reading each newspaper.

Answer ..... (1 mark)



- 5 The table shows some information about the number of schools by type in one Local Education Authority (LEA) in the UK in 2006.

Type of School	Number	Angle on Pie Chart
Primary	63	
Comprehensive	45	120
Grammar		48
Others		
Total	135	360

- 5 (a) Complete the table.

.....

.....

.....

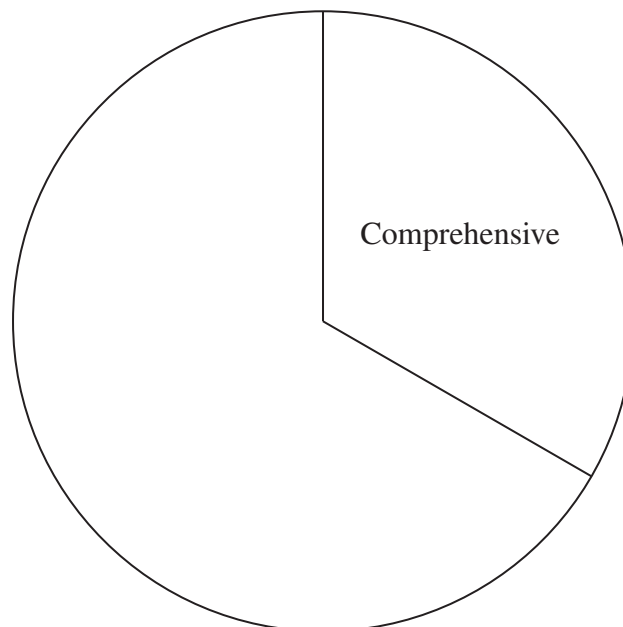
.....

.....

(4 marks)

- 5 (b) A pie chart of radius 4.1 cm is to be drawn to illustrate these data.

Complete the pie chart shown.



(2 marks)



5 (c) A pie chart showing the number of schools by type, in another LEA is to be drawn for comparison.

The radius for this pie chart will be 4.8 cm.

Calculate the difference between the total number of schools in each of the two LEAs.

.....  
.....  
.....  
.....

Answer ..... (3 marks)

**Turn over for the next question**



6 The table shows the time, in minutes, for which a sample of cars were parked in a short stay city car park.

Time, $t$ (minutes)	Frequency
$0 < t \leq 20$	6
$20 < t \leq 60$	18
$60 < t \leq 80$	30
$80 < t \leq 100$	9
$100 < t \leq 160$	12

6 (a) Calculate an estimate of the mean time.

.....  
.....

Answer ..... minutes (2 marks)

6 (b) Show that the standard deviation for the times is approximately 33.6 minutes.

.....  
.....  
.....  
.....  
.....

(3 marks)

6 (c) Draw a histogram to represent these data on the grid opposite.

.....  
.....  
.....

(4 marks)

6 (d) (i) On your histogram shade the region which is within 2 standard deviations of the mean.

.....  
.....

(3 marks)





6 (d) (ii) Approximately what percentage of the total number of cars are included in this region?

.....

Answer ..... % (2 marks)

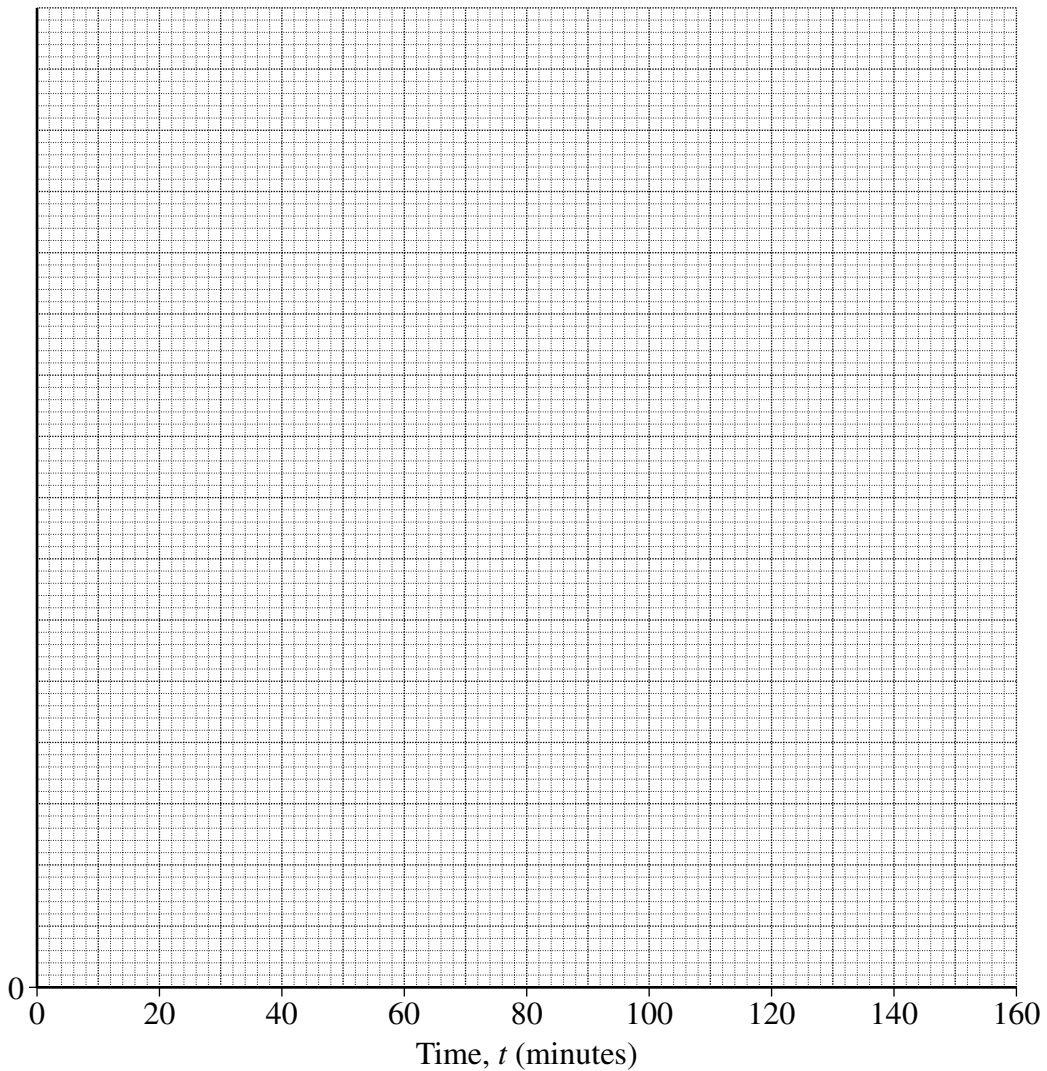
6 (e) Comment on the shape of the distribution.

.....

.....

(1 mark)

Time parked



7 The table shows information collected from the wages department of an engineering company.

Type of work	2003		2006
	Number of employees	Weekly Wage (£)	Weekly Wage (£)
Management	10	525	615
Clerical	18	210	255
Skilled	48	410	528
Semi-skilled	15	190	330

7 (a) The total weekly wage bill for 2003 is £31 560.

Using 2003 as the base year, calculate the index for the total weekly wage bill in 2006.

.....

.....

.....

.....

.....

(5 marks)

7 (b) (i) Calculate the percentage increase in the weekly wage for semi-skilled workers from 2003 to 2006.

.....

.....

Answer ..... % (2 marks)

7 (b) (ii) Explain why this value shows a higher percentage increase than the index calculated in part (a).

.....

.....

.....

(1 mark)



8 (a) Members of a fishing club wish to estimate the number of fish in a nearby lake.

A random sample of 80 fish is taken from the lake and each fish is marked before being released back into the water.

At a later date, a second sample of 280 fish are randomly chosen from the lake.

In this sample there are 33 marked fish.

Use the data to get the best estimate of the total number of fish in the lake.

.....  
.....  
.....  
.....  
.....

Answer ..... (4 marks)

8 (b) Jim and Shaun are members of the club and regularly fish the lake together.

On each occasion they record how long it takes for one of them to catch the first fish.

The times taken (in minutes) are normally distributed with mean and standard deviation as shown.

	Mean time (minutes)	Standard deviation of times (minutes)
Jim	35	6
Shaun	40	10

Records show that on Friday of last week the first fish was caught after 25 minutes.

Which of the two fishermen is more likely to have caught this particular fish?

Explain your answer fully.

.....  
.....  
.....  
.....

(4 marks)

Turn over ►



9 Paula collects data on the distance travelled by a sample of 160 shoppers to their nearest supermarket.

The results are shown in the table.

Distance, $m$ (miles)	Frequency
$0 < m \leq 5$	10
$5 < m \leq 8$	29
$8 < m \leq 12$	29
$12 < m \leq 16$	44
$16 < m \leq 21$	27
$21 < m \leq 30$	21

9 (a) Calculate an estimate of the percentage of shoppers who travel less than 15 miles to the supermarket.

.....

.....

.....

.....

Answer ..... % (3 marks)

9 (b) Two shoppers are selected at random from the sample.

What is the probability that they both travel between 8 and 21 miles to the supermarket?

.....

.....

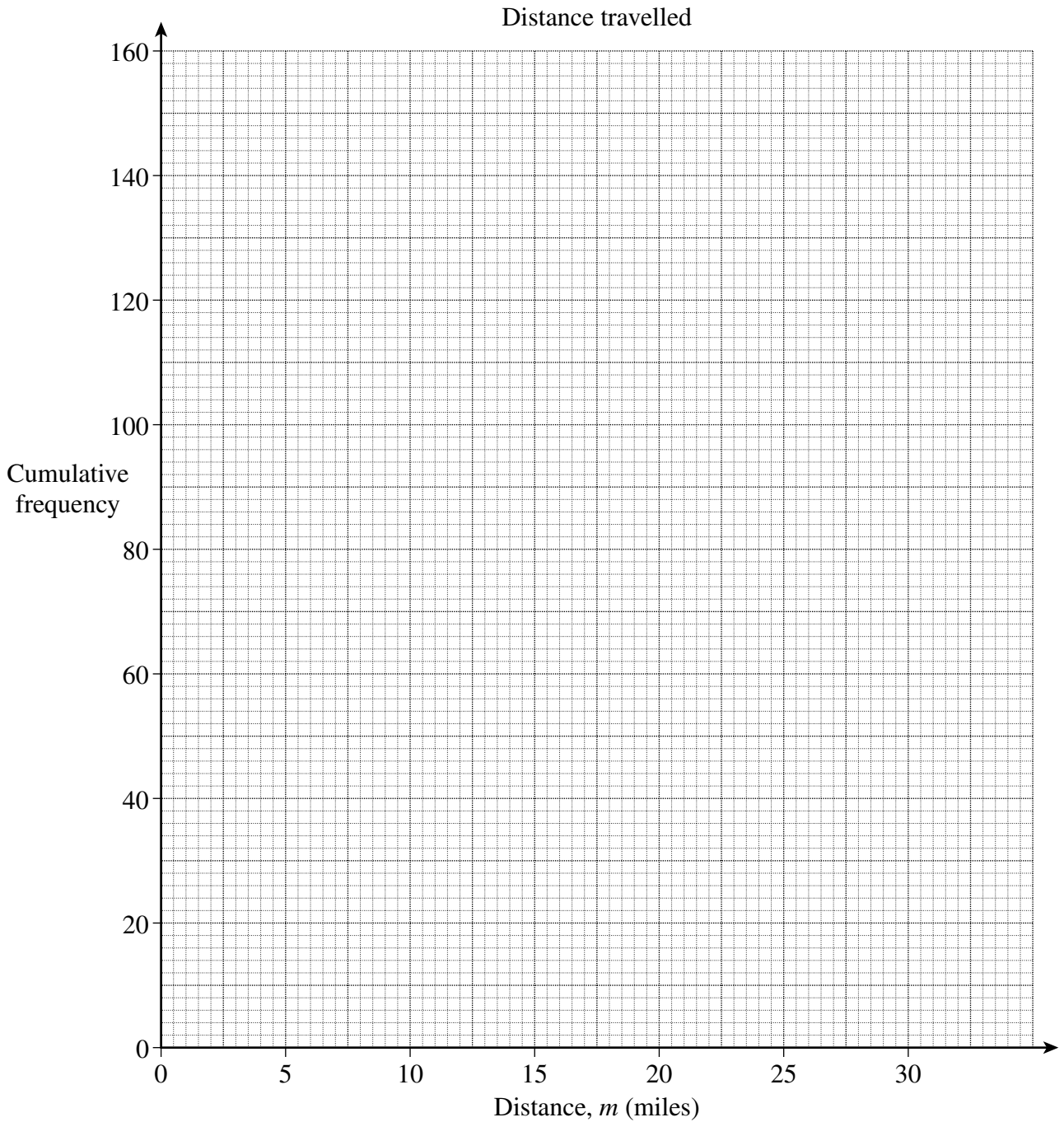
.....

.....

Answer ..... (4 marks)

9 (c) Draw a cumulative frequency polygon to illustrate these data on the grid opposite. (4 marks)





9 (d) Use your graph to estimate

9 (d) (i) the median distance

Answer ..... miles (1 mark)

9 (d) (ii) the range between the 1st and 9th decile.

.....

Answer ..... miles (3 marks)

**Question 9 continues on the next page**

**Turn over ►**



- 9 (e) Give **one** advantage of using the interdecile range rather than the interquartile range.

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

- 9 (f) Road works near to the supermarket will result in all shoppers adding one mile to their journey.

What effect will this have on

- 9 (f) (i) the median distance

.....  
(1 mark)

- 9 (f) (ii) the interdecile range?

.....  
(1 mark)



**Turn over for the next question**

**DO NOT WRITE ON THIS PAGE  
ANSWER IN THE SPACES PROVIDED**

**3**

**Turn over ►**

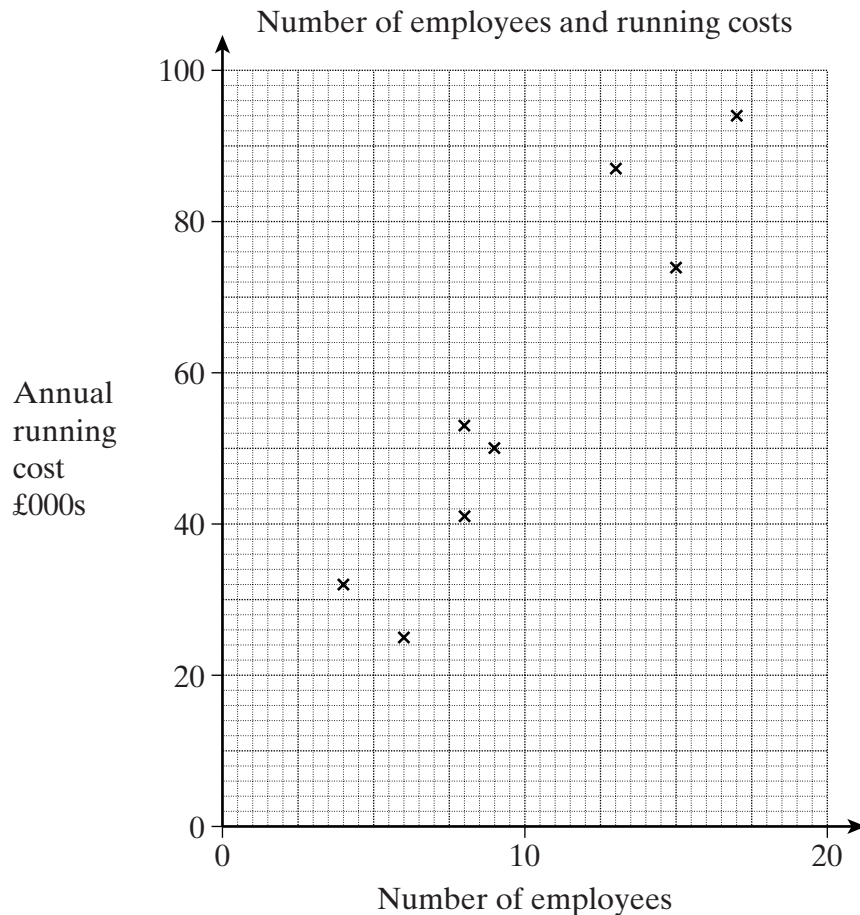


- 10** An insurance company has eight branch offices which vary in size and number of employees.

The table shows the annual cost of running each branch office and the number of employees at each office.

Branch office	A	B	C	D	E	F	G	H
Number of employees	15	13	17	8	9	8	6	4
Annual running cost in thousands of pounds	74	87	94	41	50	53	25	32

These data are shown on the scatter diagram.



For the data:  
the mean number of employees is 10  
the mean annual cost to run the branches is £57 000.

The line of best fit passes through the point (5, 30 500).





**10** (a) Draw the line of best fit on the scatter diagram. (2 marks)

**10** (b) (i) Calculate the gradient of the line of best fit.

.....  
.....  
.....

Answer ..... (2 marks)

**10** (b) (ii) What does the gradient of the line represent in this context?

.....  
.....

(1 mark)

**10** (c) Explain why office A, employing 15 staff, appears good value for money.

.....  
.....

(1 mark)

**10** (d) Calculate Spearman's coefficient of rank correlation between the number of branch employees and branch annual running costs.

.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....

Answer ..... (6 marks)

**Turn over** ►



**11** A food processing plant manufactures pots of organic yoghurt.

The label on each pot reads 'weight 150 g'.

**11** (a) Give **one** reason why the company does not check the weight of all pots produced.

.....  
 .....

(1 mark)

**11** (b) To check the weights of the pots, samples of five are selected at random from the production process.  
 The mean weight of each sample is recorded.

The company requires that the sample mean weights must fall within the range  
 $150\text{ g} \pm 0.8\%$

The mean weight of the first 10 samples taken on Monday is shown below.

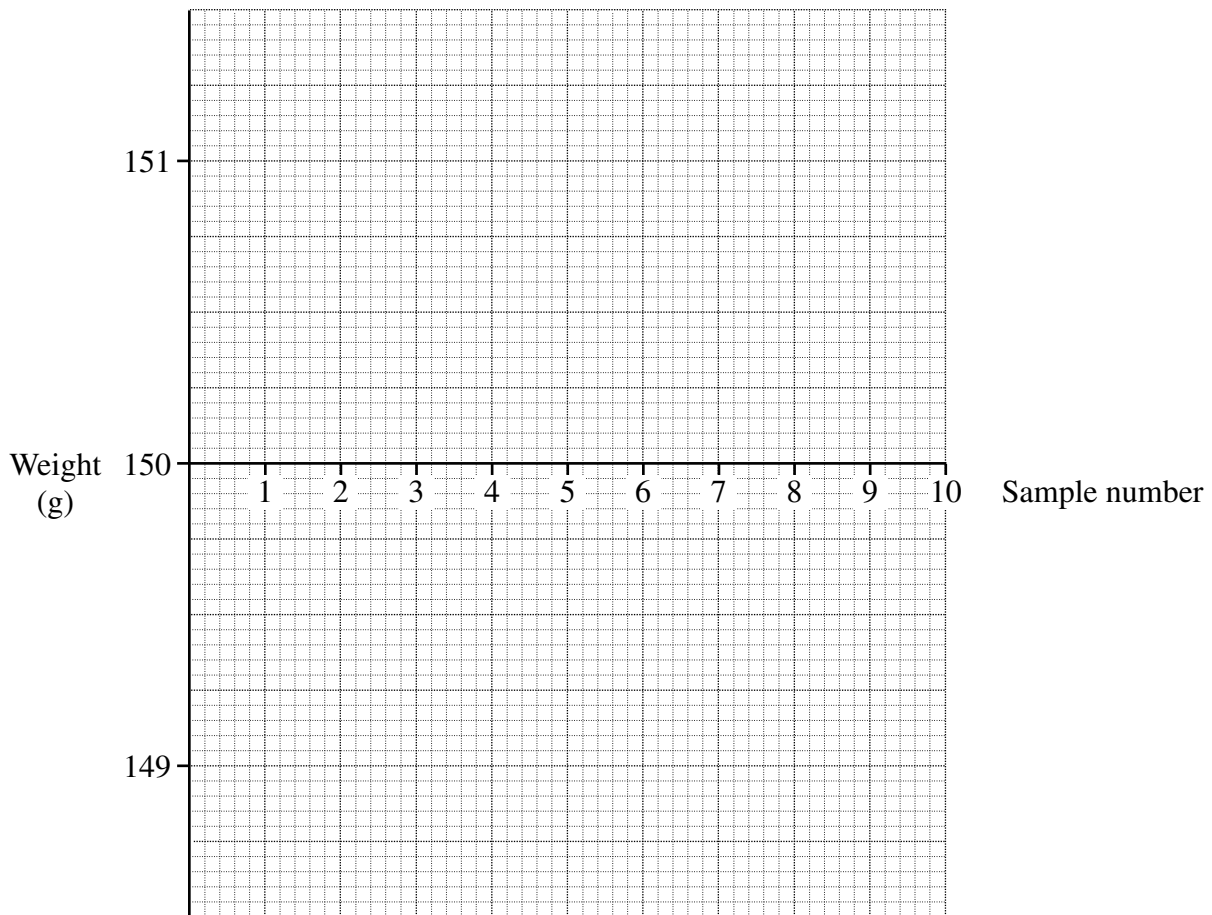
Sample number	1	2	3	4	5	6	7	8	9	10
Sample mean (g)	150.1	149.6	148.9	151.4	151.1	150.0	149.7	148.5	149.1	150.6

Plot the sample means on the chart opposite showing clearly the quality control limits.

.....  
 .....

(5 marks)





11 (c) Comment on any action that should have been taken during the period.

.....  
.....

(2 marks)

**END OF QUESTIONS**



**There are no questions printed on this page**

**DO NOT WRITE ON THIS PAGE  
ANSWER IN THE SPACES PROVIDED**

Copyright © 2008 AQA and its licensors. All rights reserved.

