Mathematics Enhancement Programme
Primary Demonstration Project

6B Handling Data

Help Booklet

Support for Primary Teachers in Mathematics

Mathematics Enhancement Programme

Help Module 6

HANDLING DATA

Part B

Contents of Part B

Preface

Activities

Tests

Answers

Contents of Part A

Preface

Introductory Notes

Worked Examples and Exercises

Answers

CIMT
University of Exeter

Sponsored by

ESSO

Gatsby Technical
Education Project

PREFACE

This is one of a series of *Help Modules* designed to help you gain confidence in mathematics. It has been developed particularly for primary teachers (or student teachers) but it might also be helpful for non-specialists who teach mathematics in the lower secondary years. It is based on material which is already being used in the *Mathematics Enhancement Programme: Secondary Demonstration Project*.

The complete module list comprises:

1.	ALGEBRA	6.	HANDLING DATA
2.	DECIMALS	7.	MENSURATION
3.	EQUATIONS	8.	NUMBERS IN CONTEXT
4.	FRACTIONS	9.	PERCENTAGES
5.	GEOMETRY	10.	PROBABILITY

Notes for overall guidance:

- Each of the 10 modules listed above is divided into 2 parts. This is simply to help in the downloading and handling of the material.

- Though referred to as 'modules' it may not be necessary to study (or print out) each one in its entirely. As with any self-study material you must be aware of your own needs and assess each section to see whether it is relevant to those needs.

- The difficulty of the material in **Part A** varies quite widely: if you have problems with a particular section do try the one following, and then the next, as the content is not necessarily arranged in order of difficulty. Learning is not a simple linear process, and later studies can often illuminate and make clear something which seemed impenetrable at an earlier attempt.

- In **Part B, Activities** are offered as backup, reinforcement and extension to the work covered in Part A. **Tests** are also provided, and you are strongly urged to take these (at the end of your studies) as a check on your understanding of the topic.

- The marking scheme for the revision test includes B, M and A marks. Note that:

M *marks*	are for method;
A *marks*	are for accuracy (awarded only following a correct M mark);
B *marks*	are independent, stand-alone marks.

We hope that you find this module helpful. Comments should be sent to:

Professor D. N. Burghes
CIMT, School of Education
University of Exeter
EXETER EX1 2LU

The full range of Help Modules can be found at

www.ex.ac.uk/cimt/help/menu.htm

ACTIVITIES

ACTIVITY 6.1   *Misuse of Statistics*

The diagrams below illustrate various data from a variety of statistics. Look carefully at the presentations and suggest ways in which each could be improved.

1.   *Sales of 'singles', 1984-88*
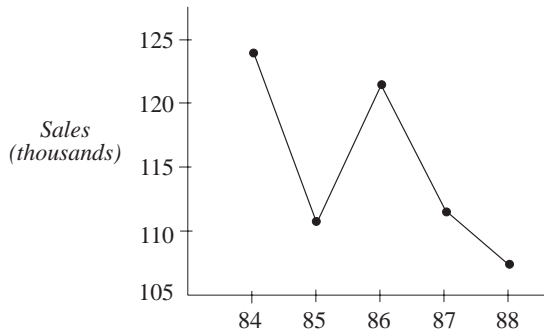
- What does this graph show?
- Why is it misleading?

2.   *Road deaths in Scotland*
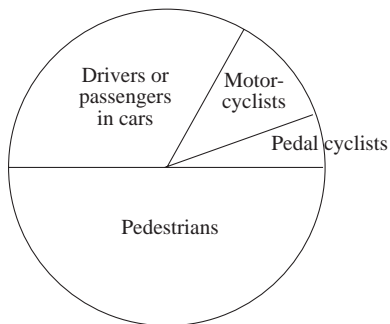
- What conclusion can be drawn?
- What other information is needed?

3.   *Number of TV sets per household*

The results from a survey of 50 houses is shown below.
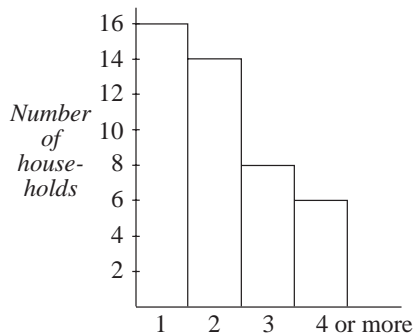
- What is missing from this illustration?
- What other data would be relevant?

4.   *Chocolate bar sales* (average/week)
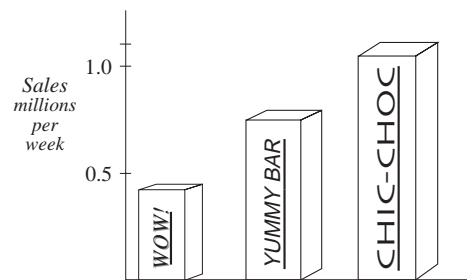
- Why is this illustration misleading?

5.   *Methods of travelling to work*
A survey of how 1000 people travel to work gave these results.

Car:	430	Bus:	356
Train:	136	Other:	78

This is illustrated in the pictogram below.
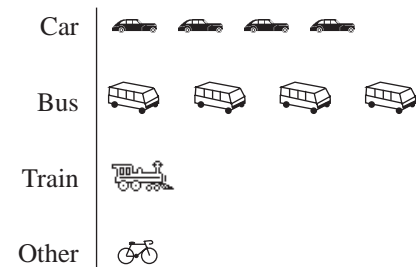
- What is wrong with this pictogram?

6.   *Voting intentions*
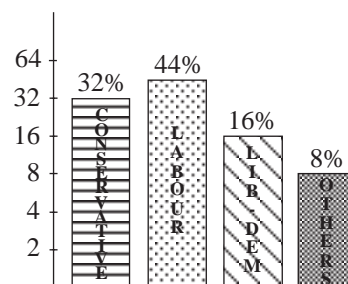
- Why is this misleading?

ACTIVITY 6.2 *Stem and Leaf Plots*

There are many ways of representing data. For example, you are probably familiar with

histograms and bar charts

but there is another very simple way which quickly gives an overall view of the general characteristics of the data. This is called a

Stem and Leaf Plot

The following example illustrates how it works.

Example

The marks gained out of 50 by 15 pupils in a Biology test are given below.

27	36	24	17	35	18	23	25
34	25	41	18	22	24	42	

We form a *Stem and Leaf Plot* by recording the marks with the tens as the 'stem' and the units as the 'leaf', as shown opposite.

Stem	Leaf
0	
1	7 8 8
2	7 4 3 5 5 2 4
3	6 5 4
4	1 2

The leaf part is then reordered to give a final plot as shown.

This gives at a glance both an impression of the spread of the numbers and an indication of the average.

Stem	Leaf
0	
1	7 8 8
2	2 3 4 4 5 5 7
3	4 5 6
4	1 2

1. Form a Stem and Leaf Plot for the following data.

21	7	9	22	17	15	31	5	17	22	19	18	23
10	17	18	21	5	9	16	22	17	19	21	20	

There is no *one* correct way of representing data. The best representation depends on the purpose for which the data is being used.

2. The ages of drivers involved in fatal road accidents in England during one week are given below.

17	82	40	48	21	35	23	24	18	57	62	45
20	21	33	27	24	37	58	69	65	19	15	21
28	71	43	31	73	26	18	21	34	35	51	63
23	65	22	45	23	27	18	19	32	25	61	36

Illustrate the data using   (a)   Stem and Leaf Plot

(b)   Histogram

(c)   Pie Chart

Which do you think is the most informative way of representing the data?

ACTIVITY 6.3   *Averages*

Sometimes data sets have to be summarised by a *single v*alue, usually called an *average*.

There are three types of average measures commonly used:

$$\boxed{Mean} \qquad \boxed{Median} \qquad \boxed{Mode}$$

A   *Calculating each average measure*

Times (in minutes) taken to solve 'logic' question

19	14	15	9	18	16	10	11	16
4	20	10	14	11	9	13	15	13
12	2	17	15	14	10	11	10	12

27 members of a class were set a 'logic' question and the times (in minutes) each pupil took to solve it were noted.

1.   The **MEAN** value of a set of data is $\dfrac{\text{sum of values}}{\text{number of values}}$.

   What is the *mean* (to 2 d.p.) of the times given in the table?

2.   The **MEDIAN** is the *middle* value of an ordered set of data.

   (a)   Write down the times in the table above in *ascending* order, i.e. smallest first.

   (b)   How many values are there?   (c)   What is the *median* ?

3.   The **MODE** is the value which occurs most often, i.e. the most popular.
   What is the *mode* of the times in the table above?

4.   Which of the three measures do you think is most representative of the average time?
   Give your reasons.

B   *Choosing which measure to use*

5.   In a clothes shop, the sizes of a particular dress sold during one week were noted and are shown in the table opposite.

   (a)   Find the *mean*, *mode* and *median* for this data.

   (b)   Which measure is of most use to the sales staff?

Dress sizes sold in one week

10	14	12	16	18
16	12	10	14	16
16	14	18	8	14
12	16	10	10	16
16	18	14	16	8

6.   The wages of factory employees are shown in the table.

   (a)   Find the *mean*, *mod**e** and *median* of the weekly wages.

   (b)   Which of these measures is the most useful?

Factory wages paid per week

10	are paid	£120
35	are paid	£140
25	are paid	£160
30	are paid	£180

ACTIVITY 6.4   *Correlation*

The bodies of most people are in proportion. If you are particularly tall, then you will probably also have long arms and legs and large hands, etc. The purpose of this activity is to attempt to see how clear-cut these relationships are.

We say that two variables have *positive correlation* if they increase in proportion.

Different types of correlation are sketched below in scatter diagrams.

| *Strong negative correlation* | *Weak negative correlation* | *No correlation* | *Weak postive correlation* | *Strong positive correlation* |
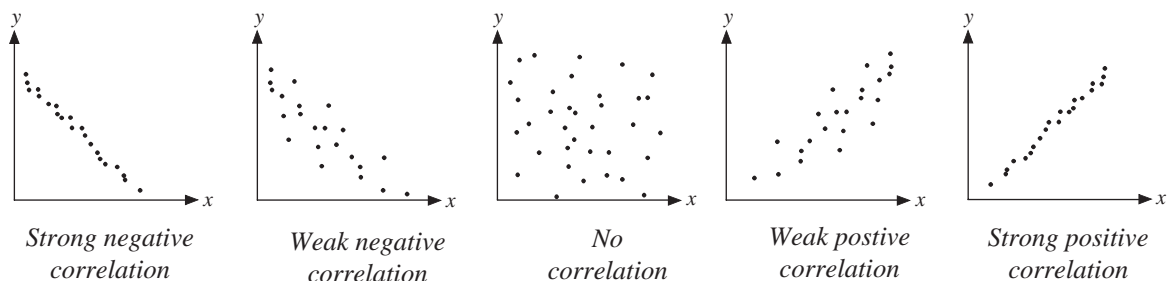
1.   For your class members, or a group of 20 to 30 people, find out their

    (i)   height   (ii)   feet size

    (iii) arm length   (iv) hand size

    (iv)   waist size   (v)   circumference of head.

2.   (a)   Taking the *y* variable as height and *x* as one of the other variables, draw scatter diagrams for each *x* variable.

    (b)   Determine the type of correlation in each case.

ACTIVITY 6.1   *Notes for Solutions*

Notes and solutions given only where appropriate.

6.1   1.   Although the number of singles is on a downward trend, it is not as pronounced as shown in the graph, where the vertical scale starts at 105 (thousands).

A better representation is shown below.
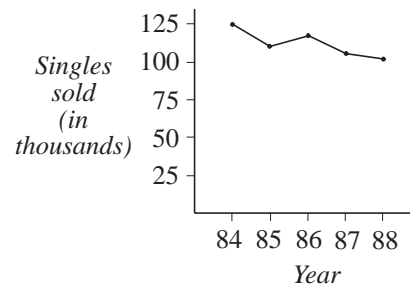
2.   It is difficult to draw any conclusions, except the ratios between the various categories. It would be more helpful to know

(a)   the time length of the survey,

(b)   the total number of road deaths and as a proportion of the population,

(c)   the proportion of cars / motorcycles / cyclists on the roads.

3.   The zero television sets per household is missing.
This gives another column of length 6.

4.   The illustration is 3-dimensional and not in proportion, making the Chic-Choc Bar sales look even more dominant. It would be better to use a bar chart (with equal widths).

5.   This is a complete mess! The scale is missing (e.g. each symbol $\equiv$ 100 people).
When using a pictogram each 'shape' must have equal width – here the 'bus' is much larger than the 'car'. The data has been rounded to the nearest hundred – this is a poor approximation when the sample is only 1000.

It is probably better to use a bar chart or a pie chart.

ACTIVITIES 6.2 - 6.3   *Notes for Solutions*

Notes and solutions given only where appropriate.

6.2 1.

Stem	Leaf
0	5 5 7 9 9
1	0 5 6 7 7 7 7 8 8 9 9
2	0 1 1 1 2 2 2 3
3	1

2. (a)

Stem	Leaf
0	
1	5 7 8 8 8 9 9
2	0 1 1 1 1 2 3 3 3 4 4 5 6 7 7 8
3	1 2 3 4 5 5 6 7
4	0 3 5 5 8
5	1 7 8
6	1 2 3 5 5 9
7	1 3
8	2

6. The non-linear scale exaggerates the performances
of *Liberals* and *Others* and makes the *Labour* lead
look very small.
It would be better to use a linear scale.

6.3 1. 12.59

2. (c) 13

3. 10

4. Mean or median (but *not* the mode)

5. (a) mean = 13.76, median = 14, mode = 16

   (b) The mode is probably of most use.

6. (a) mean = £155, median = £160, mode = £140   (b) the mean

TESTS

Test 6.1

Mental Practice

Answer these questions as quickly as you can, but without the use of a calculator.

1. For the data

    2   1   3   2   4   3   2   6   4

    what is:

    (a)   the mean,

    (b)   the median,

    (c)   the mode,

    (d)   the range?

2. For the data

    4   6   7   11

    what is:

    (a)   the median,

    (b)   the mean?

3. A football team's scoring record for 10 matches is

No. of goals	0	1	2	3	4	5 or more
Frequency	2	1	4	1	2	0

    What is the mean number of goals per match?

4. The mean of three numbers is 2. A fourth number, 6, is added.
    What is the new mean value?

5. When the number 5 is added to a set of 3 numbers, the mean value is 6.5.
    What is the mean of the original three numbers?

6. The mean of a set of 4 numbers is 5. When a fifth number is added to the set,
    the mean increases to 6. What is the value of the fifth number?

Test 6.2   Mathematics Enhancement

Answer these questions as quickly as you can, but without the use of a calculator.

1. For the data

        2   4   1   8   1   9   3

    what is:

    (a)   the mean,

    (b)   the median,

    (c)   the mode,

    (d)   the range?

2.   For the data

        13   8   2   1

    what is:

    (a)   the median,

    (b)   the mean?

3.   The number of shots taken by 10 golfers at a particular hole is summarised below

No. of shots	0	1	2	3	4	5	6	7	8 or more
Frequency	0	0	1	2	5	1	0	1	0

    What is the mean number of shots taken at this hole?

4.   The mean of three numbers is 4. Another number, 8, is added.
      What is the new mean value?

5.   The mean of a set of 5 numbers is 3. When another number is added, the mean
      increases to 3.5. What is the number added?

6.   When the number 2 is added to a set of 5 numbers, the mean value is 4.5.
      What is the mean value of the original set of numbers?

Test 6.3

Revision

40 minutes are allowed

1. A sports shop keeps information about sports shoes on a database.
    Part of this database is shown below.

Model	Manufacturer	Cost
Flyer	Tiger	£39.99
Racer	Cheetah	£37.29
Runner	Cheetah	£35.99
Strider	Tiger	£48.99
Blinder	Lion	£33.49
Sprinter	Leopard	£49.99

    (a) Write down the name of the manufacturer of the cheapest shoe.   (*1 mark*)

    (b) How much dearer is the Strider than the Racer?   (*2 marks*)

    (*LON*)

2. The bar chart below shows the proportions of blood donors in 14 different countries.

    Blood
    donations

    per 1000 total population

Switzerland	97
Denmark	83
Finland	62
Belgium	60
France	58
Sweden	52
Scotland	52
W Germany	51
US	50
Norway	47
England	46
Netherlands	45
Spain	23
Portugal	19
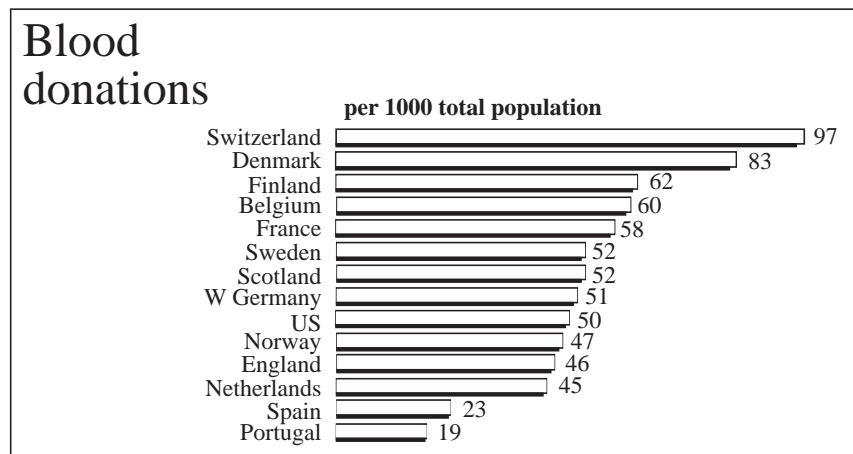
    (a) In Belgium, 60 people in every 1000 are blood donors.
         How many people in every 100 is this?   (*1 mark*)

    (b) Which country has 5.8% of its population giving blood?   (*1 mark*)

    (c) What is wrong with the following statement?

         England has twice as many blood donors as Spain.   (*2 marks*)
         (*NEAB*)

Test 6.3 Revision

3.
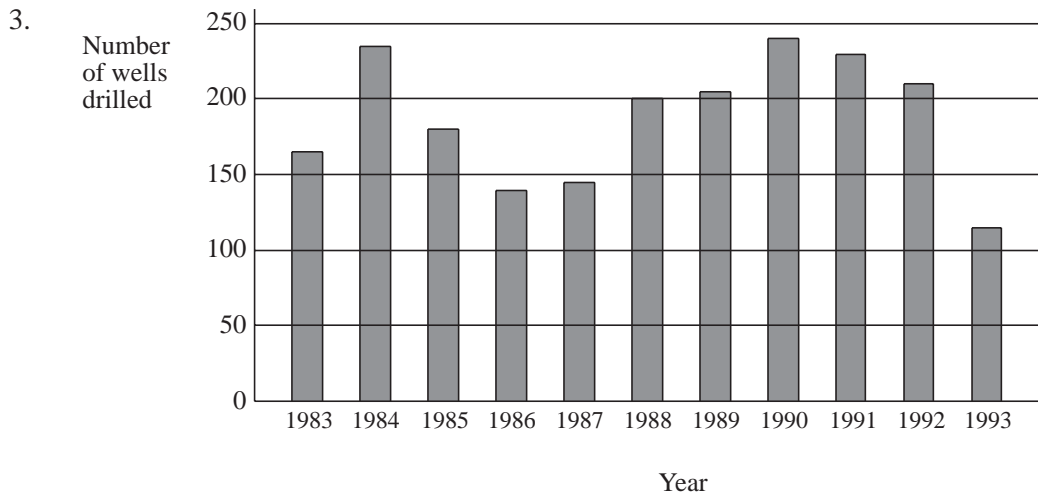
The bar chart shows the number of oil wells drilled in the North Sea between 1983 and 1993.

(a)   In which year were the most wells drilled?   (*1 mark*)

(b)   How many wells were drilled in 1988?   (*1 mark*)

(c)   How many wells were drilled in 1991?   (*1 mark*)

(*LON*)

4.   The table shows the number of different types of homes in a village.

Type of home	Detached house	Bungalow	Terraced house	Semi-detached house
Number of homes	35	52	33	60

(a)   Draw a pie chart to show this information.

You may find it helpful to complete a table like the one below first.

Type of home	Detached house	Bungalow	Terraced house	Semi-detached house
Angle of sector				

(*6 marks*)

(b)   Which type of home is the mode?   (*1 mark*)

(*SEG*)

Test 6.3 Revision

5.   The Severn Trent Water Company published a report on household water usage.

   The pie chart shows the average amounts of water used per day by one person.
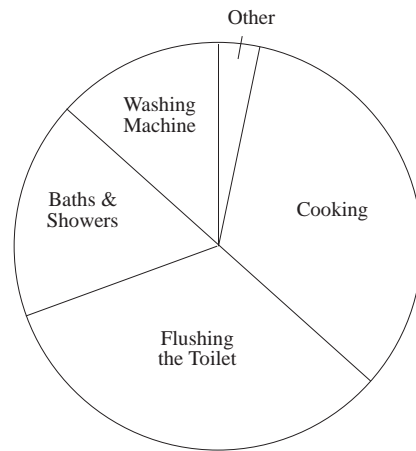
AVERAGE WATER USAGE PER DAY BY ONE PERSON

   The total amount used per day by one person is 30 gallons.

   (a)   Measure the angle for the washing machine.   (*1 mark*)

   (b)   Calculate the number of gallons used per day for the washing machine.   (*3 marks*)

   (c)   Calculate the fraction of the total used for the washing machine.   (*2 marks*)

                                                                                       (*SEG*)

6.   The waiting times, rounded down to the nearest minute, for 60 patients at a certain clinic are as follows:

25	12	53	8	26	5	19	73	67	18
87	42	6	21	14	19	12	15	13	36
36	16	72	36	13	37	11	51	39	32
30	47	6	22	68	25	98	23	45	22
7	9	26	35	27	48	58	56	29	20
32	62	80	41	58	17	54	15	14	74

   (a)   Construct a frequency table using class intervals

                     0–9, 10–19, 20–29 and so on.   (*2 marks*)

   (b)   Draw a frequency diagram to represent this data.   (*2 marks*)

   (c)   Which interval has the largest frequency.   (*1 mark*)

Test 6.3 Revision

7. Boxes of eggs containing broken eggs cannot be sold in shops.

    The table shows how many boxes of Grade *A* eggs and Grade *B* eggs could not be sold in a week at one shop.

Grade	Mon	Tue	Wed	Thur	Fri
A	14	16	16	14	16
B	14	12	9	11	15

    (a) For the Grade *A* eggs calculate:

        (i) the range;

        (ii) the mean.   (*3 marks*)

    (b) For the Grade *B* eggs the range is 6 boxes and the mean is 12.2 boxes.

        Use this information to compare the two Grades.

        Which Grade would you recommend the shopkeeper to stock?

        Give a reason.   (*2 marks*)

                              (*NEAB*)

8. The table shows the time spent on homework in one week by the students in classes 9A, 9B and 9C.

Time spent (to nearest hour)	Number of students 9A	9B	9C
1	1	0	1
2	4	5	0
3	4	6	2
4	5	5	1
5	1	0	0
6	3	2	4
7	0	0	6
8	0	2	3
9	2	0	1
10	0	0	2
Total	20	20	20

    (a) Write down the class which had the highest mean time.   (*1 mark*)

    (b) Which class had the lowest modal time?   (*1 mark*)

    (c) Which class's times were less spread out than the others?   (*1 mark*)

    (d) What was the modal time for all 60 students?   (*2 marks*)

    (e) What was the median time for class 9C?   (*2 marks*)

                          (*SEG*)

Test 6.3 Revision

9. At the end of a typing course all the students are tested. The time taken to type 60 words is recorded.

The table shows the distribution of their times.

Time (w seconds)	Number of students
$0 \le w < 20$	0
$20 \le w < 40$	7
$40 \le w < 60$	19
$60 \le w < 80$	11
$80 \le w < 100$	3

(a) Draw a frequency polygon for these data. (*2 marks*)

(b) Calculate an estimate of the mean of their times. (*3 marks*)

(*SEG*)

10. The number of children per family in a recent survey of 21 families is shown.

$$1 \quad 2 \quad 3 \quad 2 \quad 2 \quad 4 \quad 2 \quad 2$$

$$3 \quad 2 \quad 2 \quad 2 \quad 3 \quad 2 \quad 2 \quad 2$$

$$4 \quad 1 \quad 2 \quad 3 \quad 2$$

(a) What is the range in the number of children per family. (*1 mark*)

(b) Calculate the mean number of children per family.
Show your working. (*2 marks*)

A similar survey was taken in 1960.

In 1960 the range in the number of children per family was 7 and the mean was 2.7.

(c) Describe **two** changes that have occurred in the number of children per family since 1960.

(*2 marks*)

(*SEG*)

Tests 6.1 and 6.2

Answers

Test 6.1

Test 6.2

1. 3

2. 3

3. 2

4. 5

5. $6\frac{1}{2}$

6. 7

7. 2

8. 3

9. 7

10. 10

1. 4

2. 3

3. 1

4. 8

5. 5

6. 6

7. 4

8. 5

9. 6

10. 5

Test 6.3

Answers

1. (a) Lion (also allow Blinder)   B1

    (b) £48.99 – £37.29 = £11.70   M1   A1

    (3 marks)

2. (a) 6   B1

    (b) France   B1

    (c) This is only true if the total populations are the same   B2

    (4 marks)

3. (a) 1990   B1

    (b) 200   B1

    (c) 230 (allow 227 to 233)   B1

    (3 marks)

4. (a)
Type	Detached	Bungalow	Terraced	Semi-detached
Angle	70	104	66	120

    M1 A1 A1 A1 A1

    Pie chart   B1

    (b)   Bungalow   B1

    (7 marks)

5. (a) 48°   B1
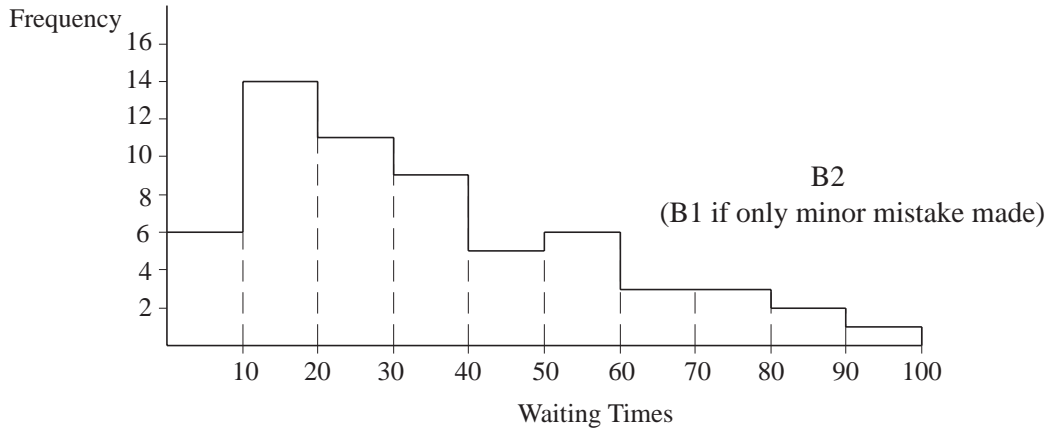
    (b) $\dfrac{48}{360} \times 30 = 4$ M2   A1

    (c) $\dfrac{48}{360} = \dfrac{2}{15}$ M1   A1

    (6 marks)

6. (a)

0–9	6
10–19	14
20–29	11
30–39	9
40–49	5
50–59	6
60–69	3
70–79	3
80–89	2
90–99	1

    B2

Test 6.3 Answers

(b)

B2
(B1 if only minor mistake made)

(c)   10–19   B1

(5 marks)

7.   (a) (i) $16 - 14 = 2$ B1

         (ii) mean = $\dfrac{76}{5} = 15.2$ M1   A1

      (b) B has greater range but lower mean - hence choose B.   B1   B1

(5 marks)

8.   (a) 9C   B1
      (b) 9B   B1
      (c) 9B   B1
      (d) 3   B2
      (e) 7   B2

(7 marks)

9.   (a) Frequency polygon   B2

      (b) $\left[(30 \times 7) + (50 \times 19) + (70 \times 11) + (90 \times 3)\right] \div 40$ M2
           $= 55$ A1

(5 marks)

10.   (a) Range 1 to 4, ie. 3   B1

      (b) mean value = $\dfrac{48}{21} = 2.286$ (allow 2.3)   M1   A1

      (c) Both the range (7 to 3) and mean value (2.7 to 2.3) have decreased

                                                                               B1   B1

(5 marks)

(TOTAL MARKS 50)