Collecting data

- Identifying possible sources of bias and minimising them
- Organising data into grouped frequency tables

Keywords

You should know

explanation 1a | explanation 1b | explanation 1c | online

1 Decide whether you would collect primary or secondary data for these investigations.

a The predicted population in the UK over the next 50 years.

b The hours of sport done each week by pupils in the UK.

c The style of jeans preferred by pupils in your class.

d The number of goals scored by Manchester United last season.

e The number of left-handed pupils in your school.

f The number of taxis waiting at your nearest railway station at 6 p.m. on a weekday.

2 Write whether the data in each set is discrete or continuous and explain how you decided.

a Year 9 examination results.

b The times taken by pupils in your class to run 100 m in a race.

c The heights of sunflowers in a field.

d The weights of books in the school bags of pupils in your class.

e Birthdays of pupils in your class.

f Distances jumped by the long jump team on sports day.

g The time taken by your classmates to get to school.

h The age of each pupil in your class on their last birthday.

3 Say whether the following data is qualitative or quantitative.
If the data is quantitative, state whether it is discrete or continuous and explain how you decided.

a The heights of pupils in your year.

b Favourite sports personalities.

c The foot lengths of pupils in your year.

d The shoe sizes of pupils in your year.

e The colours of the cars in the school car park.

f Different forms of renewable energy.

g The names of pupils in your year.

explanation 2a | explanation 2b | explanation 2c | explanation 2d

4 Write whether each of these methods of collecting data might give biased data and if so why.

a You want to investigate the favourite computer games of pupils in your school. You select a sample of 15 girls and 5 boys in your year.

b You want to investigate the amount of sleep pupils in your school get. You ask pupils in the first two years.

c You want to investigate the music preferred by young people. You ask young people in the High Street between 9 a.m. and 10 a.m. on Saturday morning.

d The government wants to find pupils' opinions of the range of sport offered in schools. It asks pupils from all the schools in Birmingham.

e A local political party want to know the voting intentions of the adults in a town of 50 000 people. It asks 70 people using the shopping mall one Monday morning.

5 A television talent show invites people to vote for their favourite singer.
Do you think unbiased results would be obtained from each of these methods of data collection? Explain your answer for each part.

a Viewers must phone to vote for their favourite singer.

b Viewers must text to vote for their favourite singer.

c Viewers must vote online for their favourite singer.

d Viewers can choose their own method of voting for their favourite singer.

6 Jo wants to test the hypothesis that, in her school, most pupils' favourite type of food is Chinese. Explain how she could select a representative sample of pupils.

7 You plan to survey the pupils in your class to find out about how they spend their free time.

a Write suitable questions to find the following information, trial them on a small number of respondents and refine them if necessary.

i Favourite leisure activity

ii Hours of sport played each week

iii Hours spent on the internet each week

b Write two more questions you could include in the survey.

8 This data collection question could be improved.
Write a better version with response boxes for the answers.

> What kind of cool music do you like: Indie, rock, hard rock or metal?

explanation 3

9 These are the results for 26 pupils in a mathematics examination.

a Copy the grouped frequency table so the teacher can allocate end-of-term grades. Use it to group the marks.

70	57	55	56	69	66	65	74	78	60	86	67	68
76	91	90	87	88	70	63	70	79	85	68	72	66

b Which class interval has the greatest number of pupils?

Mark in examination	Tally	Frequency
50–59		
60–69		
70–79		
80–89		
90–99		

10 These are the heights to the nearest centimetre of 24 boys.

133 150 142 139 143 129 158 144 159 135 134 146

150 168 188 136 172 153 142 172 155 138 176 158

a Copy and complete the grouped frequency table.

Height, h (cm)	Tally	Frequency
$120 \leq h < 130$		
$130 \leq h < 140$		
$140 \leq h < 150$		
$150 \leq h < 160$		
$160 \leq h < 170$		
$170 \leq h < 180$		
$180 \leq h < 190$		

b Which class interval has the greatest frequency?

11 These were the total numbers of goals scored by the 20 teams in the Premier League in one season.

80 65 71 48 74 20 41 46 66 50

36 36 38 67 55 42 45 34 43 45

a Put the information into a grouped frequency table choosing suitable class intervals.

These were the total numbers of goals scored by the 20 teams in the Premier League in the previous season.

83 29 47 63 45 38 57 43 32 34

35 52 44 57 52 29 38 37 52 64

b Put the information into a grouped frequency table using the same class intervals as in part **a**.

c Which class interval has the greatest frequency for each of your grouped frequency tables?

12 Annie is collecting data for a science project.
She measures the lengths of 30 leaves.
These are the lengths of the leaves measured in centimetres correct to the nearest millimetre.

4.3	2.6	3.7	3.2	5.4	3.9
4.2	4.3	5.2	2.8	3.2	3.1
5.3	2.4	2.7	5.3	3.0	3.4
5.3	2.1	2.5	4.7	4.5	3.2
4.9	5.1	5.1	2.9	3.9	4.2

Put the results in a grouped frequency table. Choose suitable class intervals.

13 The data below shows the times, correct to the nearest 0.01 second, taken to run 100 m by 24 world-class athletes.

9.81	9.85	10.01	10.06	10.12	9.90	9.92	10.06
10.04	9.83	9.78	10.07	9.97	10.13	9.94	9.95
10.10	10.02	10.02	9.87	10.02	10.08	9.95	9.91

Put the results in a grouped frequency table with suitable class intervals.

explanation 4

14 Chris surveyed everyone in her year group.
She displayed her findings in a two-way table.

	Right-handed	Left-handed	Total
Girls	52	7	
Boys	68	17	
Total			

a How many pupils did she survey?

b What question do you think she asked?

c How many boys are there in the year group?

d How many left-handed pupils are there in the year group?

e Chris picks a pupil at random.
What is the probability that the pupil will be a left-handed boy?

15 Pupils had to choose between art and music.
This table shows the choices of 52 pupils.
Some of the table has been filled in.

	Art	Music	Total
Boys	5		
Girls	20	7	
Total			

a Copy and complete the table.

b What does the figure in pink represent?

c What does the figure in the blue space represent?

d What percentage of these pupils chose art?
Give your answer correct to one decimal place.

16 This table shows the type of accommodation 100 people stayed in for their summer holiday. Some of the table has been filled in.

	Hotel	Camping	Other	Total
July		11	5	20
August	15		8	
September		14		32
Total	28			100

a What does the figure in orange represent?

b What does the figure in blue represent?

c Copy and complete the table.

d What percentage of people did not stay in a hotel?

17

	Can swim	Cannot swim	Total
Boys			
Girls			
Total			

40 pupils were asked if they can swim or not.

28 of the pupils were girls.

8 boys cannot swim.

23 girls can swim.

Use this information to complete the two-way table.

18 Sahil asked 60 English-speaking adults if they could also speak another language.

6 of the women could speak another language.

28 of the adults were men.

51 adults could not speak another language.

a Design a two-way table to show this information.

b Put the given information into your two-way table.

c Complete your two-way table.

19 Mandy asked 80 pupils in her year if they preferred biology, chemistry or physics.

36 of the pupils were girls.

15 of the boys preferred chemistry.

45 pupils preferred biology.

12 pupils preferred physics.

19 boys preferred biology.

a Design a two-way table to show this information.

b Put the given information into your two-way table.

c Complete your two-way table.