

5 Data

Representing and interpreting data



Carl Linnaeus was a Swedish scientist who classified all living things into categories. His system organises human beings as:

Kingdom	Animalia
Phylum	Chordata
Class	Mammalia
Order	Primates
Family	Hominidae
Genus	Homo
Species	sapiens

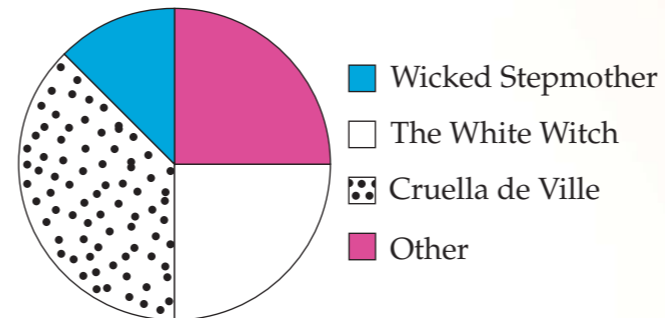
What's the point? Classifying data into groups with similarities makes it easier to use and present the data.

Check in

- Level 3** 1 Order these sets of numbers from smallest to largest.
- a 10, 12, 15, 10, 9, 11, 12, 10
 - b 104, 110, 101, 99, 94, 100, 98, 101, 103, 98
 - c 234, 423, 342, 324, 243, 432

- Level 4** 2 Work these out without using a calculator.
- a $\frac{12 + 14 + 16}{3}$
 - b $\frac{18 + 23 + 25 + 30}{4}$
 - c $\frac{105 + 108 + 110 + 112 + 115}{5}$

- 3 In a survey children chose their favourite storybook villain.
- a What fraction of the children chose the Wicked Stepmother?
 - b What fraction of the children chose Cruella de Ville?
 - c 32 children took part in this survey. How many chose the White Witch?



5b Mean

- Calculate the mean for discrete data

When people talk about the 'average' and don't say which one, they are usually referring to the **mean**.

$$\text{Mean} = \frac{\text{The total of the values}}{\text{The number of values}}$$

Test scores: 3, 4, 4, 8, 8, 9

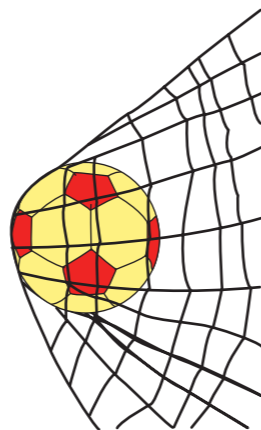
$$\text{Mean} = \frac{3 + 4 + 4 + 8 + 8 + 9}{6} = \frac{36}{6} = 6$$

The mean does not have to be one of the data values, or even a possible value in the context.

For example, in the 2006–2007 season, Manchester United scored a mean of 2.18 goals per match. You can score 2 goals, but how do you score 0.18 goal?

$$\text{The mean was } \frac{83 \text{ goals scored in}}{38 \text{ games played}} = 2.18 \text{ goals per game}$$

Keywords
Average
Mean



example

Find the mean of each set of data.

- a** Resting pulse rates of a group of Year 7 pupils:

67, 71, 70, 66, 63, 71 beats per minute

- b** Cost of Christmas presents Jamie bought:

£3.99, £7, £5.60, £29.99, £6

- a** Total = $67 + 71 + 70 + 66 + 63 + 71 = 408$

$$\text{Mean pulse rate} = \frac{408}{6} = 68$$

- b** Total = $£3.99 + £7 + £5.60 + £29.99 + £6 = £52.58$.

$$\text{Mean cost} = \frac{£52.58}{5} = £10.52 \text{ (to the nearest penny)}$$

example

Martina recorded her test scores for five maths tests during the term. She calculated a mean score of 7.8. Unfortunately, she put a cup of tea on her record book and smudged one of her scores. What was the smudged score?

For a mean of 7.8, the total must be $7.8 \times 5 = 39$.

The four known scores total 31 so the last one must be 8. $39 - 31 = 8$

Maths tests this term	
19 Jan	6
26 Jan	9
2 Feb	7
9 Feb	7
16 Feb	9

Exercise 5b

- Find the mean of each of these data sets.
 - Resting pulse rates of a group of Year 7 students: 61, 71, 80, 66, 68
 - Resting pulse rates of a group of Year 2 students: 93, 78, 81, 81, 71, 95, 82
 - Pulse rates of a group of Year 7 students after running for 5 minutes: 113, 147, 127, 139, 122, 135, 119
- The length of time between eruptions of Old Faithful geyser in Yellowstone National Park is recorded on one day during the park's opening hours. The times are (in minutes) 65, 89, 65, 95, 55, 89, 56, 92. Calculate the mean length of time between eruptions that day.
- A golf instructor asked each of his students to stand 100 metres from the flag. Using a rangefinder, he measured their actual distances, in metres. 107, 95, 102, 110, 96, 98, 107, 93, 101, 91. Calculate
 - the range
 - the mean.
- The number of Year 7 students present each day for two weeks is 159, 159, 153, 149, 152, 158, 161, 160, 161, 154.
 - Calculate the mean number of Year 7 students present.
 - There are 163 Year 7 students on the roll in the school. Calculate the mean number of Year 7 students absent.
- A psychologist measures the reaction times (in hundredths of a second) of different groups of people under different conditions. Calculate the mean reaction time for each of the groups.
 - Normal conditions: 16, 19, 15, 19, 22, 17, 16, 17, 21, 18, 17
 - Using their non-writing hand to react: 19, 22, 23, 19, 17, 24, 27, 21, 18
 - Blindfolded: 20, 21, 18, 19, 23, 21, 24, 18, 17, 22, 19, 23, 20
 - After two alcoholic drinks: 22, 19, 14, 21, 15, 26, 23, 19, 24



puzzle

- Find a set of five numbers which have a mode of 3, a median of 5 and a range of 4.
- How many different sets of five numbers can you find which have a mode of 5, a median of 5 and a range of 4?

5c Frequency tables

- Construct a frequency table for discrete data
- Use a frequency table to find any average (mode, median or mean) and the range

Keywords
Frequency table **Median**
Mean **Mode**
Range

- A large set of data can be summarised in a **frequency table**.
- You can find the **median** from a frequency table by counting a running total.
- You can find the **mean** from a frequency table by multiplying each value by its frequency and adding the results together, then dividing by the total frequency.

example

Class 7C were asked how many pets they had. The frequency table shows their answers.

Number of pets	Frequency
0	8
1	7
2	4
3	1

Calculate

- the range
- the mode
- the median
- the mean of these data.

a The highest value is 3 pets, the lowest is 0 pets. $3 - 0 = 3$

The range is 3 pets.

b 8 students have 0 pets.

8 is the highest number in the frequency column.

The mode is 0 pets.

c There are $8 + 7 + 4 + 1 = 20$ students, so the median is between the 10th and 11th values. The 10th and 11th values are both 1. The median is 1 pet.

$$\begin{aligned} \text{d The mean} &= \frac{\text{Total number of pets}}{\text{Total number of students}} \\ &= \frac{(0 \times 8) + (7 \times 1) + (2 \times 4) + (3 \times 1)}{20} \\ &= \frac{0 + 7 + 8 + 3}{20} = \frac{18}{20} = 0.9 \end{aligned}$$

The mean is 0.9 pets.

The first row on the table tells you that 8 students have 0 pets.

Values 1–8 are 0 pets. Values 9–15 are 1 pet.

The third row of the table tells you that 4 students have 2 pets. Between them, they have 8 pets.

Exercise 5c

- Sally writes down her scores as she throws a dice.
 3, 6, 3, 2, 5, 1, 3, 6, 1, 4, 3, 2, 4, 2, 5, 2, 2
 - Show this information in a frequency table.
 - Write down the modal score.
 - Calculate the median score.
 - Calculate the range of the scores.
- These are the ages of children in a day care centre.
 2, 3, 4, 1, 4, 2, 5, 2, 3, 3, 4, 5, 1, 5, 3
 - Show this information in a frequency table.
 - Calculate the modal age of the children.
 - Calculate the median age of the children.
 - Calculate the range of the ages.
 - Calculate the mean age of the children.
- The number of letters in the words of a child's first reading book are
 3, 3, 2, 4, 3, 1, 3, 3, 4, 2, 3, 3, 3, 2
 - Show this information as a frequency table.
 - Calculate the modal number of letters in a word.
 - Calculate the median number of letters in a word.
 - Calculate the range of the number of letters in the words.
 - Calculate the mean number of letters in a word.
- For a Science project, Simon shells some peas and records the number of peas in each pod.
 3, 4, 3, 8, 2, 3, 6, 7, 2, 7, 5, 5, 3, 3, 7, 8, 3, 4, 5, 6, 2, 8, 3
 - Show this information in a frequency table.
 - Calculate the modal number of peas in each pod.
 - Calculate the median number of peas.
 - Calculate the range of the number of peas.
 - Calculate the mean number of peas.



puzzle

In a rugby team the mean weight of the 8 forwards is 95 kg and the mean weight of the 7 backs is 83 kg.

- Bill says the mean weight of the team is 89 kg. Is he correct? Explain your answer.
- If the median weight of the forwards is 95 kg and the median weight of the backs is 83 kg, what can you say about the median weight of the team?

5d Frequency diagrams

• Understand and use information given on different kinds of bar charts

• You can represent **frequency** using a **bar chart**.

• You can use a **comparative bar chart** to compare different sets of data.

Keywords
Bar chart *Frequency*
Comparative bar chart *Stacked bar chart*

Sometimes a comparative bar chart is called a multiple bar chart.

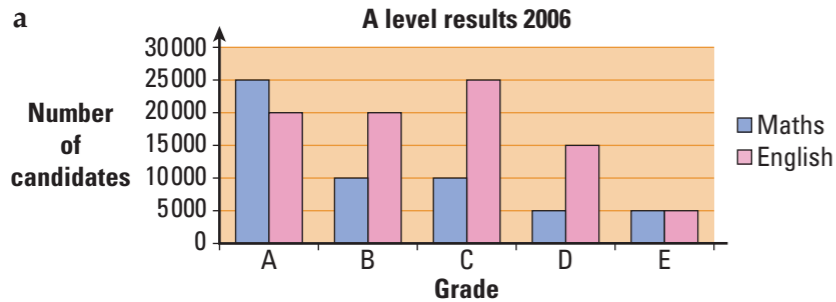
example

The bar chart below shows how many students achieved grades A–E in their maths A level in 2006. The table shows how many students achieved grades A–E in their English A level in the same year.

a Add the English data to the bar chart.

b Comment on what the chart tells you about the results for A level maths and English in 2006.

Grade	Number of candidates
A	20 000
B	20 000
C	25 000
D	15 000
E	5 000



Always include a key.

The height of bar A for English is 20 000 because 20 000 candidates achieved grade A.

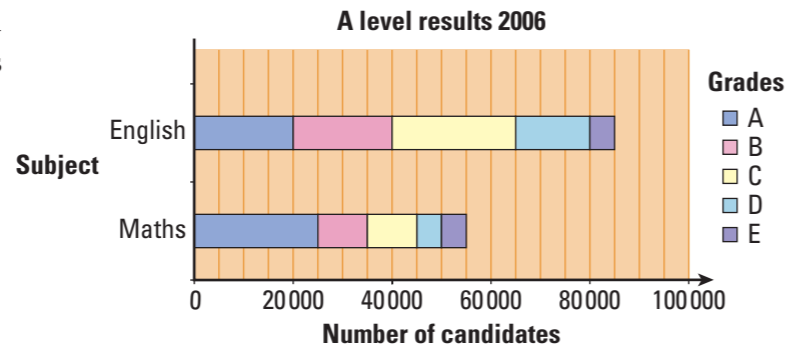
The modal grade has the highest bar.

b The modal grade in maths was A. The modal grade in English was C.

• A **stacked bar chart** can be used to show how much each group contributes to the total.

Sometimes a stacked bar chart is called a component or composite bar chart.

From this graph, it is easy to tell that there were more candidates for English than for maths. You can also see that more candidates got at least a grade B in English than in maths, but more got A grades in maths.



Exercise 5d

1 The table shows the average daily temperatures in °C in London and Sydney.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
London	3.9	4.2	5.7	8.5	11.9	15.2	17	16.6	14.2	10.3	6.6	4.8
Sydney	25.7	25.6	24.6	22.2	19.2	16.7	16	17.5	19.7	21.9	23.5	25.1

a Draw a comparative bar chart to represent this information.
 b Compare the average temperatures in London and Sydney over the year.

2 The table shows the percentages of UK households with access to at least one car, in 1990 and in 2000.

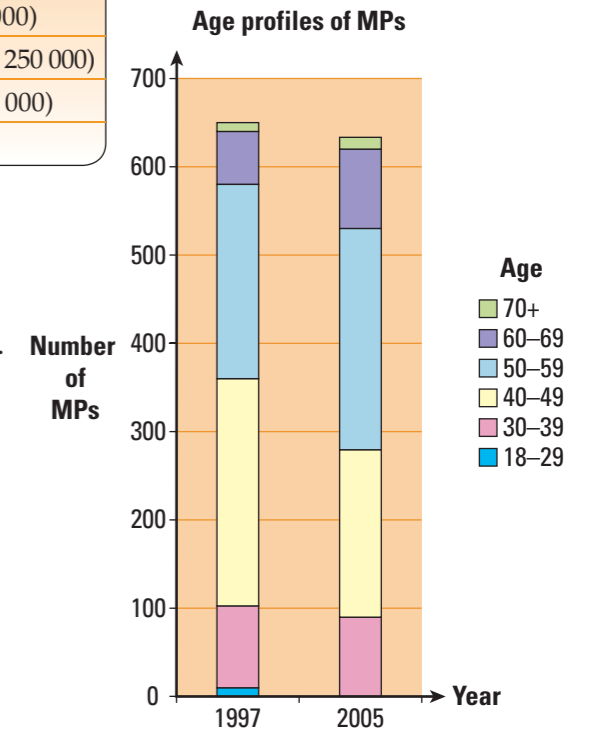
% of households		
1990	2000	Year
60	63	London boroughs
58	62	Metropolitan built-up areas
65	68	Large urban (population over 250 000)
67	73	Medium urban (population 25 000 to 250 000)
72	78	Small urban (population 3 000 to 10 000)
81	85	Rural

a Draw a comparative bar chart to represent this information.

b Describe the change from 1990 to 2000.

3 The chart shows the ages of MPs in 1997 and 2005.

- a i How many MPs were under 30 in 1997?
 ii How many MPs were aged 30–39 in 1997?
 iii How many MPs were aged 40–49 in 1997?
 iv How many MPs were under 50 in 1997?
 b How many MPs were under 50 in 2005?
 c How many more MPs were under 60 in 1997 than in 2005?



challenge

Using the graph in question 3, do you think the average age of MPs was higher in 1997 or in 2005? Give a reason for your answer. Can you be sure from the information provided?

5e Reading and interpreting pie charts

- Understand and use information given on pie charts

A **pie chart** is a circular chart divided into 'slices'.

If you know the total **frequency**, or the frequency of any one group, you can calculate the frequency of each group.

You can use two pie charts to compare the relative size of the groups within each chart.

Keywords
Frequency
Pie chart
Proportion

Pie charts are useful when you want to know the **proportion** of each category, compared to the whole.

example

The pie charts show the GCSE maths and French grades of the Year 11 students at a small school.

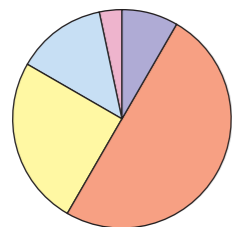
Thirty students achieved an A grade in maths.

a How many Year 11 students are there at the school?

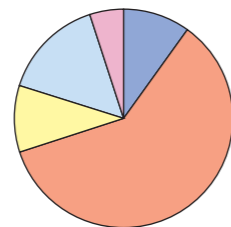
b How many students achieved grades A*, B, C and D in maths?

c Did more students achieve grade A in French than in maths? Explain your answer.

GCSE maths grades



GCSE French grades



■ A*
■ A
■ B
■ C
■ D

a 60 students

You can see that half of the students achieved an A grade. $30 \times 2 = 60$

b A* : 5 students

Measure the A* sector with a protractor.

The angle is 30° . $\frac{30}{360} \times 60 = 5$

B : 15 students

The angle is 90° . $\frac{90}{360} \times 60 = 15$

C : 8 students

The angle is 48° . $\frac{48}{360} \times 60 = 8$

D : 2 students

The angle is 12° . $\frac{12}{360} \times 60 = 2$

c The **proportion** of students achieving an A grade is greater in French than in maths, but you cannot say whether the number of students is greater because you don't know how many students took French GCSE.

There are 360° in a circle.

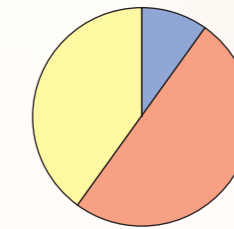
Exercise 5e

1 An airport classifies flight arrivals as **early** if they land more than 5 minutes before the scheduled time, and as **late** if they land more than 5 minutes after it. The pie chart shows information about the arrival times of a sample of 100 flights. 10 planes landed early.

a How many planes were late?

b How many planes landed on time?

Arrivals of flights



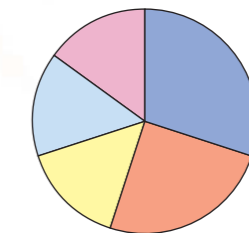
■ Early
■ On tim
■ Late

2 The pie chart shows the number of different types of pets owned by the pupils in an infants' school. There are 30 cats.

a How many pets are there altogether?

b How many dogs are there?

Pet ownership



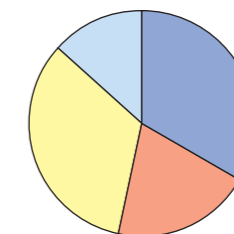
■ Dogs
■ Cats
■ Fish
■ Rabbits & hamsters
■ Others

3 The pie charts show the type of film preferred by the 15 girls and 20 boys in class 7C.

a How many i girls ii boys like science fiction/fantasy best?

b Comment on what else the pie charts tell you about the types of film the girls and boys in class 7C like best.

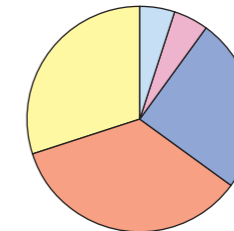
Favourite film type (girls)



Film type

■ Comedy
■ Action/adventure
■ Science fiction/fantasy
■ Costume drama
■ Westerns

Favourite film type (boys)



discussion

Computer spreadsheets make it easy to produce very fancy '3-D' charts. What are the advantages and disadvantages of this kind of chart?



5f Line graphs for time series

- Understand and use information given on line graphs

- A **time series** can be plotted as a **line graph** to show you how something has changed over time.

Keywords
Estimate
Line graph
Time series

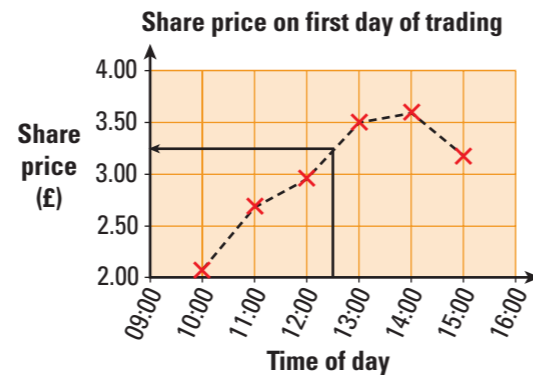
example

A new company making computer games sells shares on the stock market. The table shows the share price at hourly intervals between 10:00 and 15:00 on the first day of trading.

Time of day	10:00	11:00	12:00	13:00	14:00	15:00
Share price	£2.10	£2.70	£2.90	£3.50	£3.60	£3.20

- Plot these data on a line graph.
- Estimate** the share price at 12:30.
- Can you tell from the graph what the highest share price was? Explain your answer.

- Plot a point showing the value at each recorded time. Join them with a broken line.
- £3.20 Read up from 12:30 on the Time axis, then across to the Share price axis.
- No. We don't know the share price at the times between the points. For example, between 14:00 and 15:00 it could have risen beyond £3.60 (the highest recorded price) before falling to £3.20.

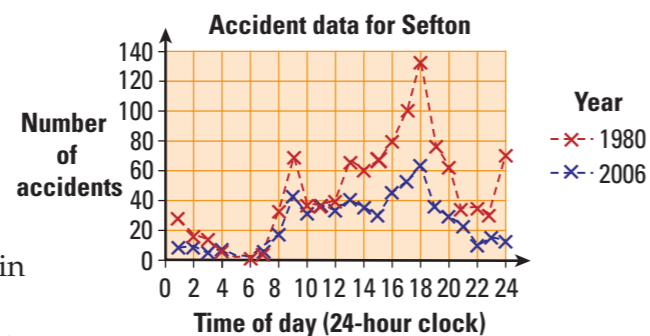


- Two time series can be plotted on the same graph. This helps you to compare the two sets of data.

example

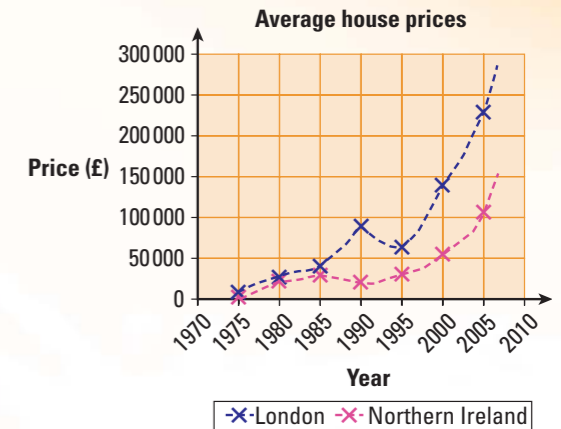
The graph shows the number of traffic accidents at different times of day in Sefton in 1980 and 2006. Comment on the similarities and differences between the two series.

In both series there are more accidents in the morning and evening rush hours. In 2006 there were fewer accidents and there is no rise in the number of accidents around midnight.



Exercise 5f

- The graph shows the average house prices in London and in Northern Ireland between 1975 and 2005.
 - Describe what has happened to the average house price in London over this period.
 - Describe what has happened to the average house price in Northern Ireland over this period.



- The table shows the life expectancy at birth for males and females in the UK between 1981 and 2004.

Year	1981	1986	1991	1996	1997	1998	1999	2000	2001	2002	2003	2004
Males	70.8	71.9	73.2	74.3	74.5	74.8	75.0	75.4	75.7	75.9	76.3	76.6
Females	76.8	77.7	78.7	79.4	79.6	79.7	79.9	80.2	80.4	80.5	80.7	81.0

- Plot two time series graphs on one set of axes to show this information.
- Describe any similarities and differences you see in the two time series.

challenge

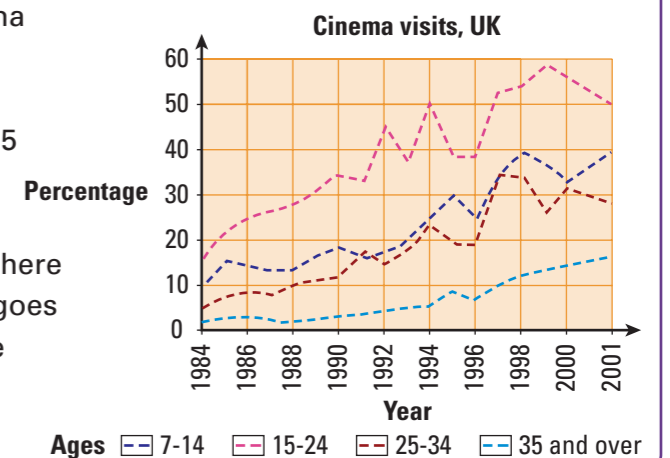
The graph shows the percentage of different age groups who said that they visit the cinema at least once a month.

- Which age group goes to the cinema most often?

Sanjit says that when he goes to the cinema, there are more people over 35 than in any other age group. He says that the graphs must be wrong.

- Assuming Sanjit is right in saying there are more people over 35 when he goes to the cinema, is he correct that the graphs must be wrong?

Explain your answer.



5 Consolidation

5a

- Are the following discrete, continuous or non-numerical data?
 - The heights of the buildings on Trafalgar Square
 - The number of people in Trafalgar Square at different times
 - The cost of entry to exhibitions in the National Gallery
 - The colours of the vehicles in Trafalgar Square at a particular time

- Find the median and range of each of these sets of data.
 - The resting pulse rates of a group of professional athletes:
53, 62, 71, 58, 61, 65, 59, 67, 64, 64
 - The resting pulse rates of a group of accountants:
65, 72, 68, 79, 83, 63, 59, 80, 69
 - The heights (in cm) of a group of professional athletes:
175, 182, 184, 170, 186, 179, 180, 178, 179, 182
 - The heights (in cm) of a group of accountants:
183, 171, 178, 179, 183, 177, 179, 181, 174

5b

- Find the means of the sets of data in question 2 using a calculator.

5c

- The number of text messages sent in an hour by a group of Year 7 students on a school trip is summarised in the table.

Number of text messages	0	1	2	3	4	5
Frequency (Boys)	7	5	2	1	2	0
Frequency (Girls)	3	4	4	2	3	4

Answer these questions for **i** the boys **ii** the girls.

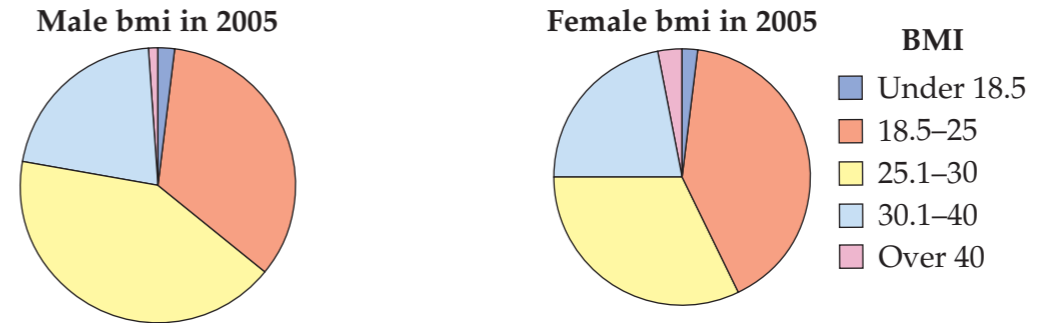
- Write the modal number of text messages.
 - Write the range of the number of text messages.
 - Calculate the mean number of text messages.
- Jamie goes on holiday to Milan in October. He records the maximum temperature (in °C) each day.
17, 17, 20, 21, 20, 18, 18, 19, 21, 21, 22, 19, 18, 17, 17
 - Show this information in a frequency table.
 - Write the modal maximum temperature.
 - Calculate the median maximum temperature.
 - Calculate the range of the maximum temperatures.
 - Calculate the mean maximum temperature.

5d

- Show the data in question 5 on a bar chart.

5e

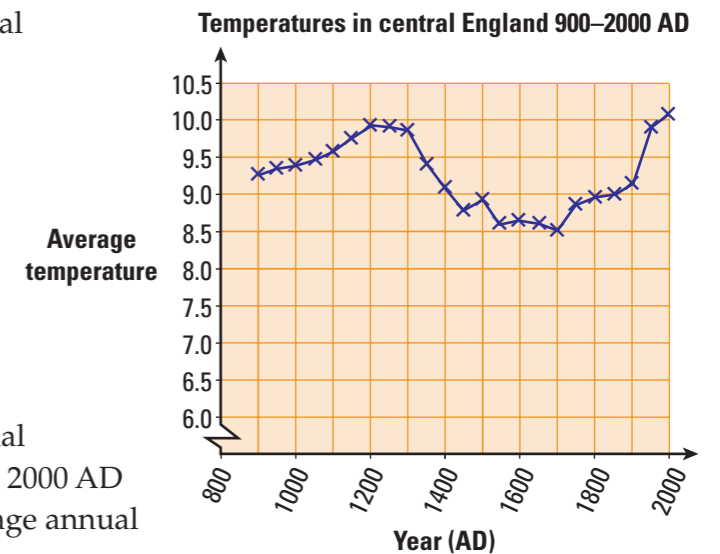
- The pie charts show the bmi (body mass index) of males and females.



- What percentage of male adults have a bmi which is over 18.5 but 25 or under?
- A bmi of over 30 is classed as obese. What percentage of adult males were obese in 2005?
- A bmi of between 25.1 and 30 is classed as overweight. What percentage of adult females were overweight in 2005?

5f

- The graph shows the average annual temperature in central England at 50 year intervals from 900 AD to 2000 AD.



- In which year was the lowest average annual temperature recorded?
- By how much did the average annual temperature drop between 1200 AD and 1400 AD?
- If the same drop in average annual temperature takes place between 2000 AD and 2200 AD, what will the average annual temperature be in 2200 AD?
- Estimate the average annual temperature in 1925.

5 Summary

Key indicators

- Interpret diagrams and graphs (including pie charts) **Level 5**



Level 5

- 1 a Six students in Zeenat's class calculate the distance from their home to school. The distances, in kilometres, are 7, 2, 6, 4, 1, 7. Calculate the median distance.
- b The mode of three numbers is 8. The mean of the three numbers is 7. Find the three numbers.

Mandy's answer ✓

Mandy finds the total of the three numbers

a 1, 2, 4, 6, 7, 7
 Median is $(4 + 6) \div 2 = 5$

b $7 \times 3 = 21$
 $21 - (8 + 8) = 5$
 The three numbers are 5, 8, 8

Mandy puts the numbers in numerical order

Mandy knows two of the numbers are 8



Level 5

- 2 The graph shows at what time the sun rises and sets in the American town of Anchorage, Alaska. The day with the most hours of light is called the longest day. Copy and complete these sentences, using the information from the graph.
- The longest day is in the month of _____. On this day, there are about _____ hours of daylight.
- The shortest day is in the month of _____. On this day, there are about _____ hours of daylight.

