

Match each data collection method to one set of data.

Survey

The colours of cars passing a school at lunchtime

Controlled experiment

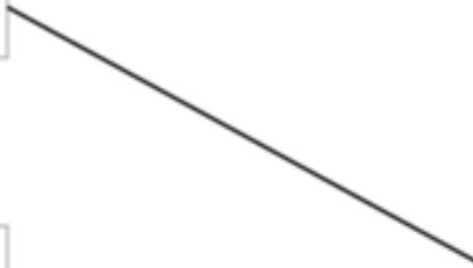
The number of people in each house of a street

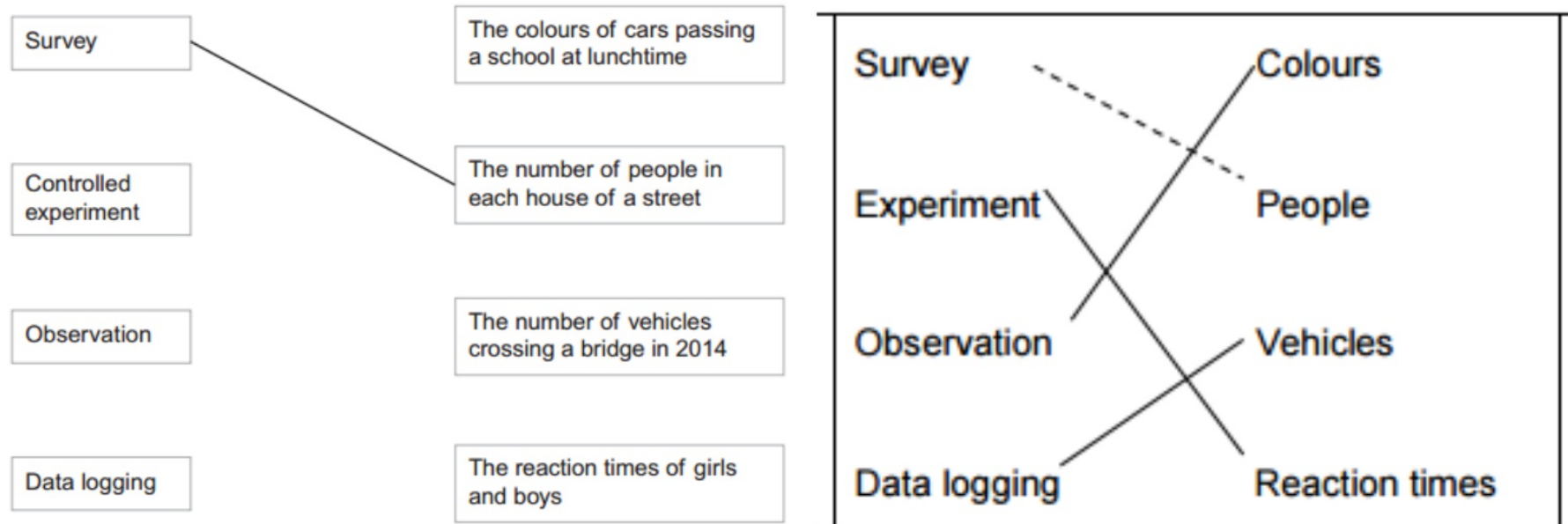
Observation

The number of vehicles crossing a bridge in 2014

Data logging

The reaction times of girls and boys





Data logging is the collection of data over a period of time, and is something often used in scientific experiments. Data logging systems typically monitor a process using sensors linked to a computer. Most data logging can be done automatically under computer control.

Jess wants to know the number of people who live in her street.
She carries out a survey.

Which **two** words describe the data she collects?
Circle your answers.

Primary

Secondary

Discrete

Continuous

Jess wants to know the number of people who live in her street.
She carries out a survey.

Which **two** words describe the data she collects?
Circle your answers.

Primary

Secondary

Discrete

Continuous

Primary selected and Secondary not selected	B1
Discrete selected and Continuous not selected	B1

**A grocer has 100 boxes of strawberries.
He weighs 10 of the boxes.**

Which three words describe the data he collects?

Circle your answers.

continuous discrete sample primary secondary

Name a suitable sampling method to obtain 10 boxes to represent the 100 boxes.

Briefly describe how to carry out your method.

A grocer has 100 boxes of strawberries.
He weighs 10 of the boxes.

Which three words describe the data he collects?

Circle your answers.

continuous discrete sample primary secondary

Name a suitable sampling method to obtain 10 boxes to represent the 100 boxes.

Briefly describe how to carry out your method.

Random (sampling)	Q1	oe
Number the 100 boxes	B1	oe
Obtain random numbers and select them	B1	oe

Four numbers have a mean of 10
The median is 8

Two of the numbers are 1 and 5

Work out the other two numbers.

Four numbers have a mean of 10 so the four numbers add to 40
The median is 8 average of the two middle numbers must be 8

Two of the numbers are 1 and 5

Work out the other two numbers.

1 5 11 23

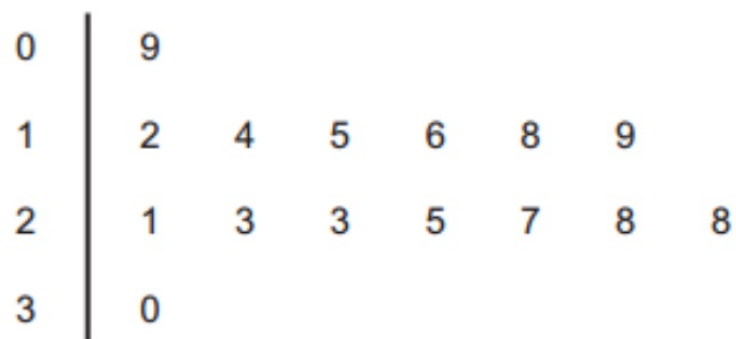
median 8, check $11+5 = 16$
and $16/2 = 8$

fourth number 23 to make the total 40

11 chosen with no other number less than 11 chosen	B1
4×10 or 40	M1
23	A1

Amy and Ben each played a game 15 times.
The stem-and-leaf diagram shows the points scored by Amy.

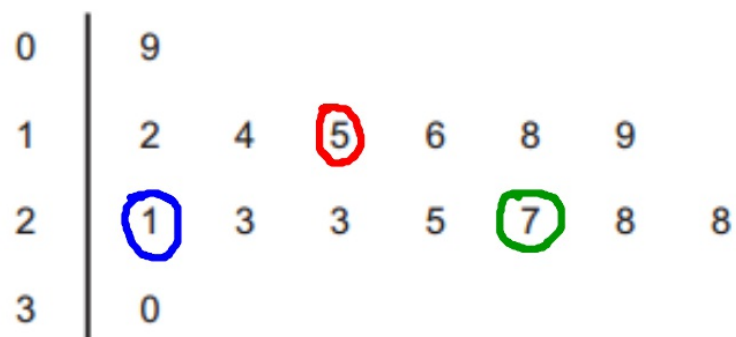
Key: 3 | 0 represents 30 points



- (a) Find the range of Amy's scores.
- (b) Find the median and quartiles

Amy and Ben each played a game 15 times.
The stem-and-leaf diagram shows the points scored by Amy.

Key: 3 | 0 represents 30 points



(a) Find the range of Amy's scores. $30 - 9 = 21$

(b) Find the median and quartiles

median 8th value 21

LQ 4th value 15

UQ 12th value 27

Compare the scores for Amy and Ben

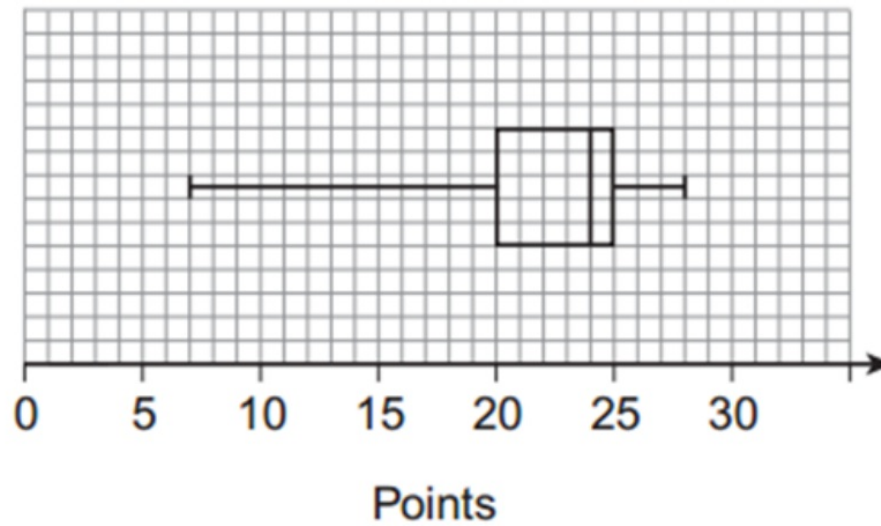
Amy

median 8th value 21

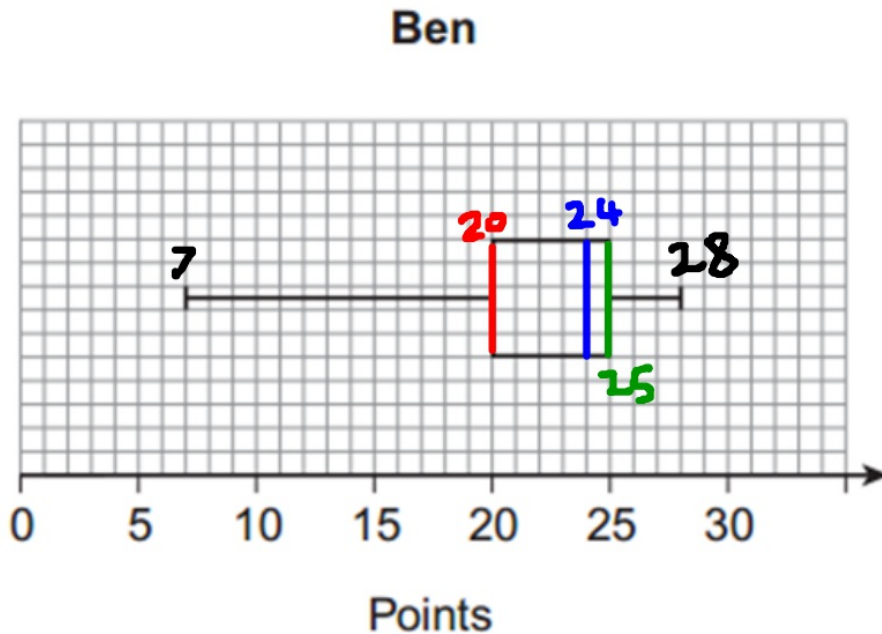
LQ 4th value 15

UQ 12th value 27

Ben



Compare the scores for Amy and Ben



Amy

median 8th value 21

LQ 4th value 15

UQ 12th value 27

Ben's median (24) is bigger than Amy's median (21) so on average Ben has better scores than Amy.

Ben IQR is 5 (25-20), Amy IQR 12 (27-15), so Ben has more consistent scores than Amy

What is a stratified sample?

Here is some information about the age groups of people in a sports club.

Junior	Adult	Senior
35	220	45

A sample of size 60, stratified by age group, is taken.
How many juniors in the sample?

Two people are chosen at random from the sample. Work out the probability that they are both juniors.

What is a stratified sample?

Each group in the sample is in the same proportion (as the population)

Must refer to the idea of proportion

Here is some information about the age groups of people in a sports club.

Junior	Adult	Senior
35	220	45

Total 300

A sample of size 60, stratified by age group, is taken.
How many juniors in the sample?

**35 Juniors in the sports club of 300 members
require $\frac{35}{300}$ of the sample of 60 to be juniors**

$$\frac{35}{300} \times 60 = \frac{35}{5} = 7$$

$\frac{35}{35+220+45} \times 60$	M1
7	A1

Here is some information about the age groups of people in a sports club.

Junior	Adult	Senior
35	220	45

A sample of size 60, stratified by age group, is taken.

How many juniors in the sample? 7

Two people are chosen at random from the sample. Work out the probability that they are both juniors.

$$\frac{7}{60} \times \frac{6}{59}$$

$$\frac{42}{3540} \text{ or } \frac{7}{590}$$

or 0.0118(...) or 0.0119 or 0.012

Chen records his journey times to college.

Time, t (minutes)	Frequency
$25 < t \leq 30$	12
$30 < t \leq 35$	18
$35 < t \leq 40$	24
$40 < t \leq 45$	6
	Total = 60

Midpoint x Frequency

$$27.5 \times 12$$

$$32.5 \times 18$$

$$37.5 \times 24$$

$$42.5 \times 6$$

$$\frac{2070}{60}$$

$$34.5$$

Write down a calculation to calculate the mean mark?
Explain why your answer is an estimate.

Because the data is grouped

The grouped frequency table represents the speeds of the 1000 cars.

Speed, s (mph)	Frequency
$18 \leq s < 20$	80
$20 \leq s < 25$	440
$25 \leq s < 30$	360
$30 \leq s < 40$	120

Describe how you would show the data on a histogram.

The grouped frequency table represents the speeds of the 1000 cars.

Speed, s (mph)	Frequency
$18 \leq s < 20$	80
$20 \leq s < 25$	440
$25 \leq s < 30$	360
$30 \leq s < 40$	120

Frequency Density

40

88

72

12

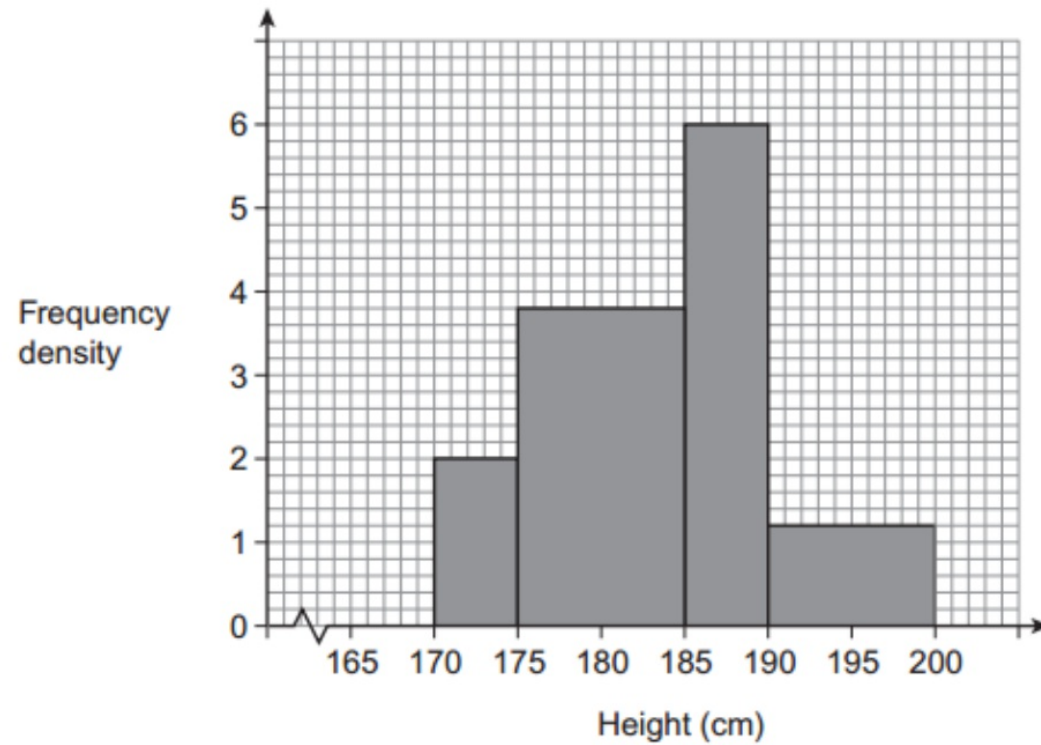
Describe how you would show the data on a histogram.

Calculate frequency density

Care with axes and scales

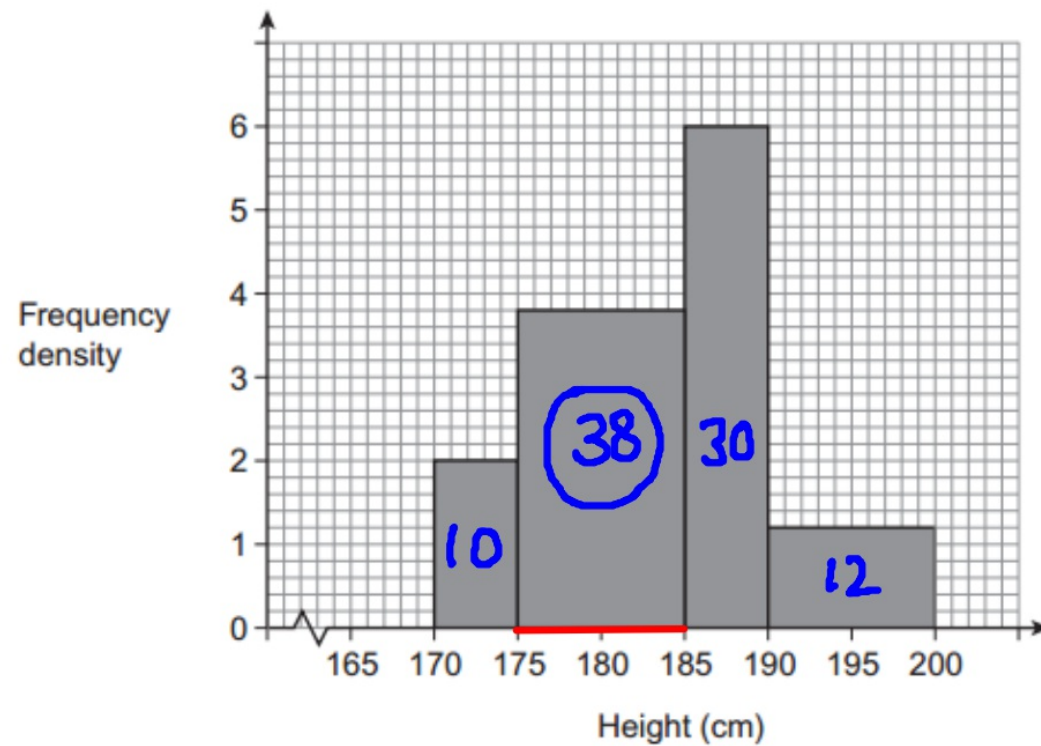
**plot bars of correct width (class width)
and height (frequency density) (y axis)**

The histogram represents the heights of 90 firefighters.



Which of the four bars represents the greatest number of firefighters?

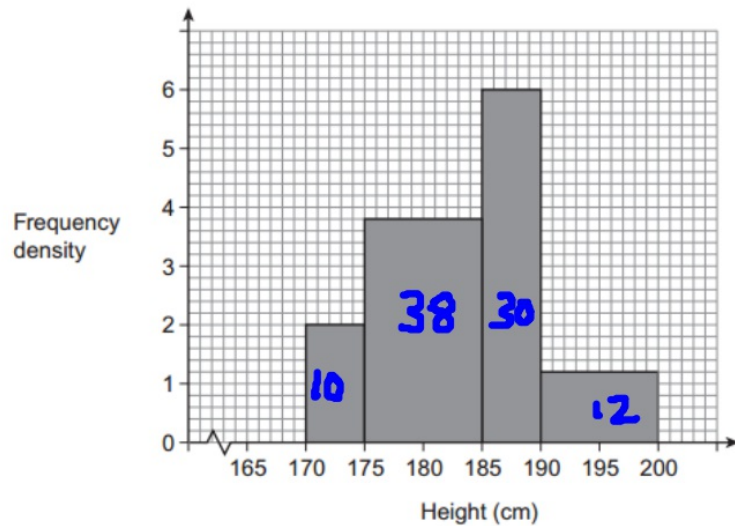
The histogram represents the heights of 90 firefighters.



Which of the four bars represents the greatest number of firefighters?

175 ≤ height < 185 and 38 seen

The histogram represents the heights of 90 firefighters.



How would you estimate the mean height?

$$172.5 \times 10$$

$$180 \times 38$$

$$187.5 \times 30$$

$$195 \times 12$$

$$\frac{16530}{90}$$

Midpoints seen or implied 172.5, 180, 187.5, 195	B1
their Σfx $10 \times 172.5 + 38 \times 180 + 30 \times 187.5 + 12 \times 195$ or $1725 + 6840 + 5625 + 2340$ or 16 530	M1
their $\Sigma fx \div 90$	M1 dep
184 or 183.7 or 183.66... or 183.67	A1

The tallest firefighter was 195.6 cm The shortest firefighter was 170.4 cm Both heights are given to 1 decimal place.

Work out the maximum possible difference in their heights.

$$195.65 - 170.35 = 25.3\text{cm}$$

The table shows information about the marks of 500 students.

Mark, m	Frequency	Cumulative frequency
$15 < m \leq 40$	80	80
$40 < m \leq 60$	220	
$60 < m \leq 80$	125	
$80 < m \leq 100$	75	

Complete the cumulative frequency column.

What are the coordinates of the points you would plot on the cumulative frequency curve?

What is the greatest possible value for the range of the marks?

What is the least possible value for the range of the marks?

The table shows information about the marks of 500 students.

Mark, m	Frequency	Cumulative frequency
$15 < m \leq 40$	80	80
$40 < m \leq 60$	220	300
$60 < m \leq 80$	125	425
$80 < m \leq 100$	75	500

Complete the cumulative frequency column.

What are the coordinates of the points you would plot on the cumulative frequency curve? $(40, 80)$ $(60, 300)$

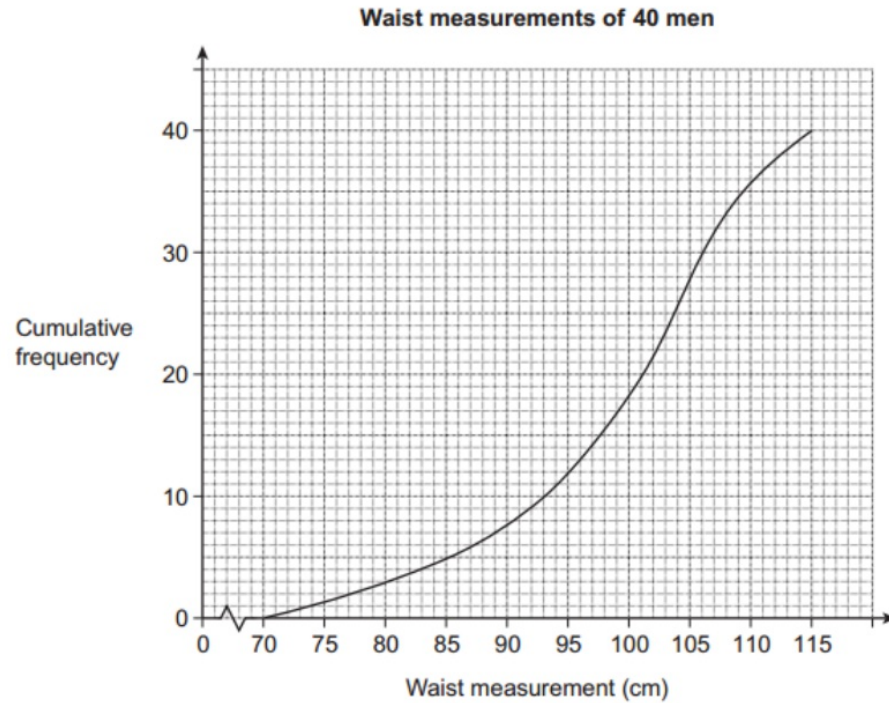
$(80, 425)$ $(100, 500)$

What is the greatest possible value for the range of the marks?

$$100 - 15 = 85$$

What is the least possible value for the range of the marks?

$$81 - 40 = 41$$



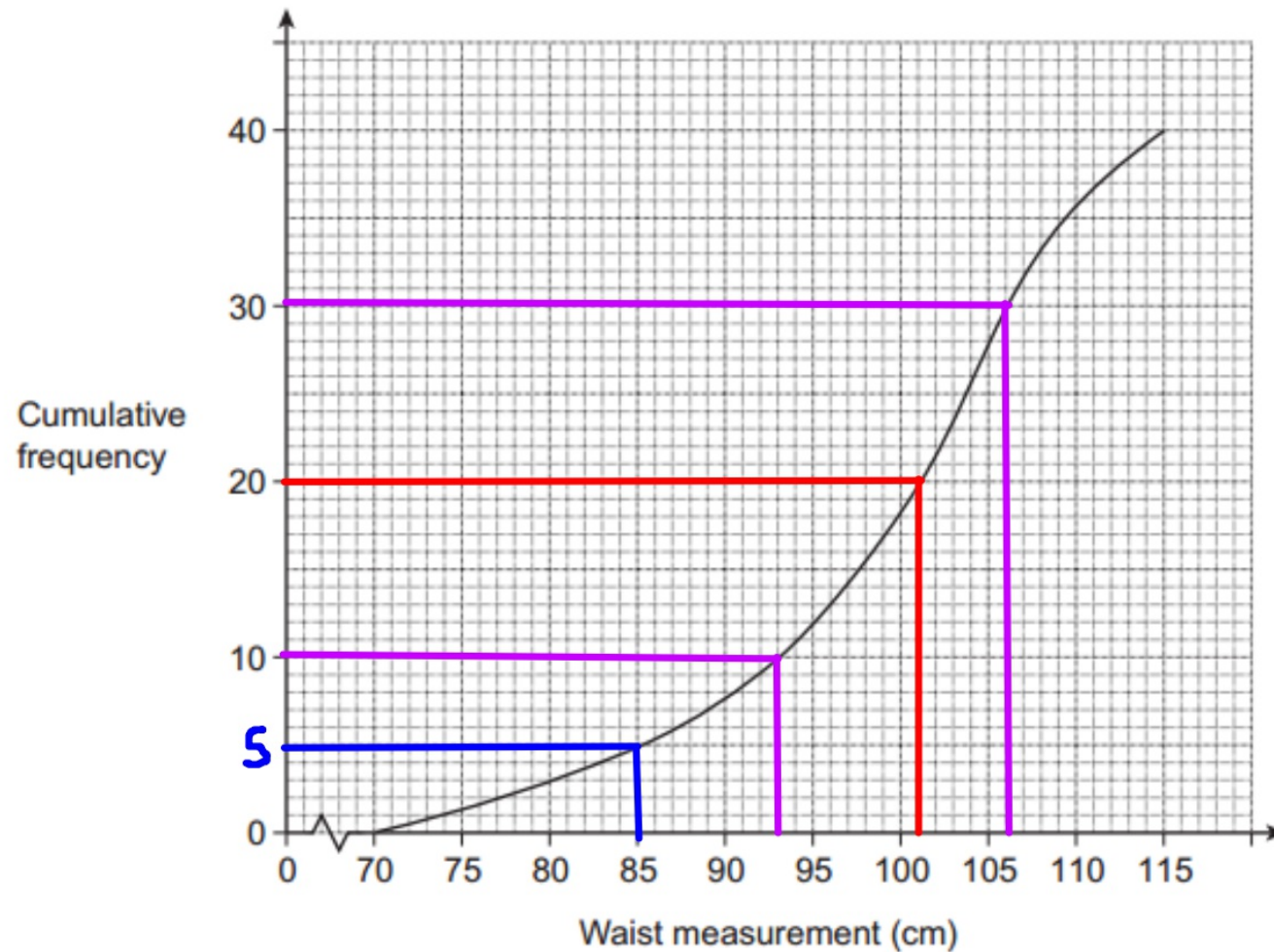
How many men have a waist measurement of 85 cm or less?

How many men have a waist measurement of 85 cm or **more**?

What is the median waist measurement?

What is the interquartile range of the waist measurements?

Waist measurements of 40 men



How many men have a waist measurement of 85 cm or less? **5**

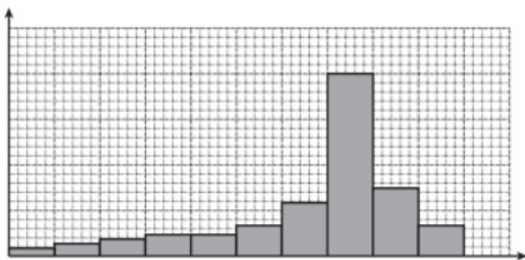
How many men have a waist measurement of 85 cm or more? $40 - 5 = 35$

What is the median waist measurement? **101**

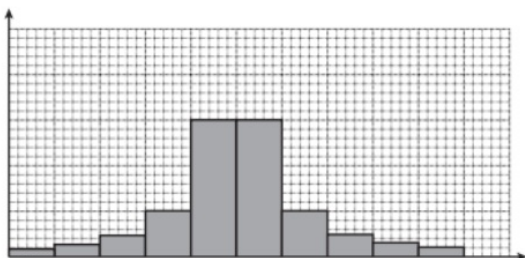
What is the interquartile range of the waist measurements? $106 - 93 = 13$

Here are the histograms for four different sets of data.
Each set of data has the same number of values.

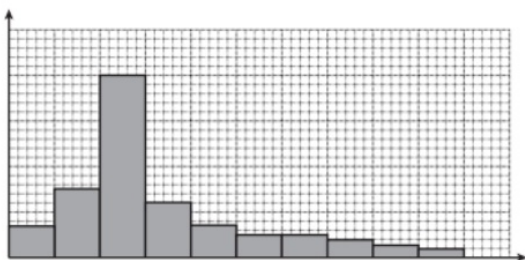
Match each box plot to a histogram



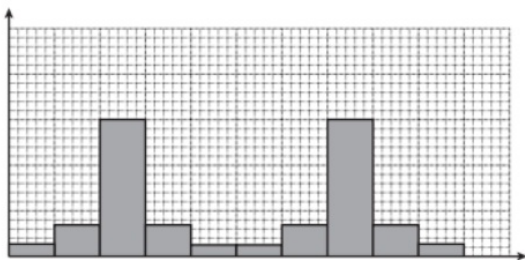
Histogram 1



Histogram 2

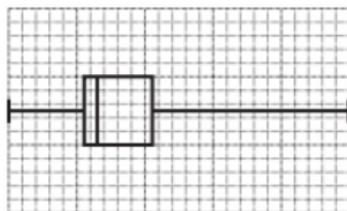


Histogram 3

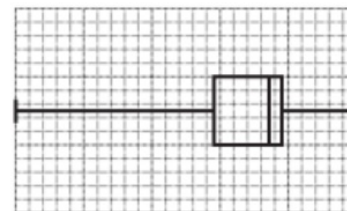


Histogram 4

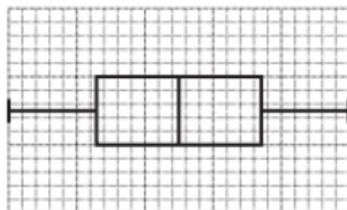
Box plot A



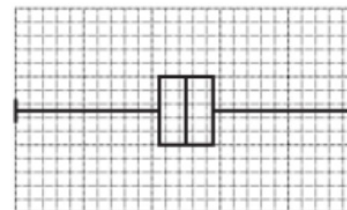
Box plot B



Box plot C

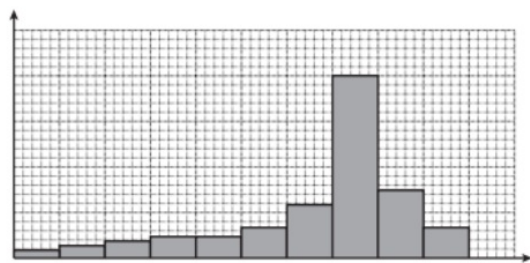


Box plot D



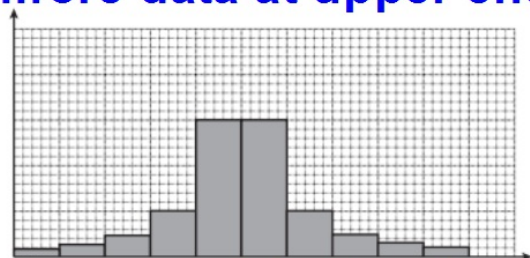
Here are the histograms for four different sets of data.
Each set of data has the same number of values.

Match each box plot to a histogram



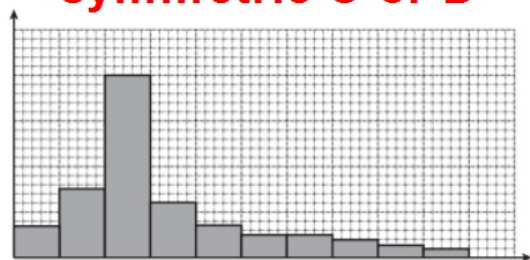
Histogram 1

more data at upper end



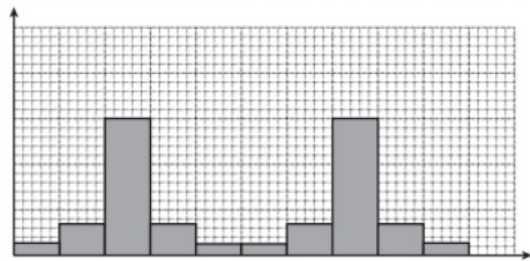
Histogram 2

symmetric C or D



Histogram 3

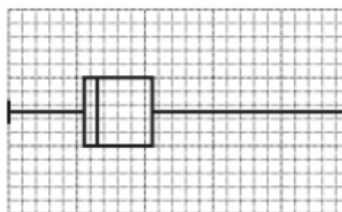
more data at lower end



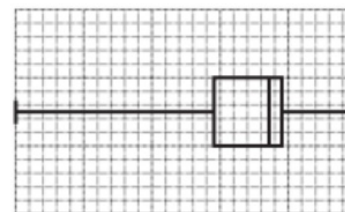
Histogram 4

symmetric C or D

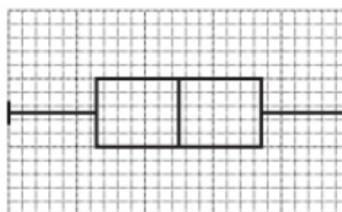
Box plot A



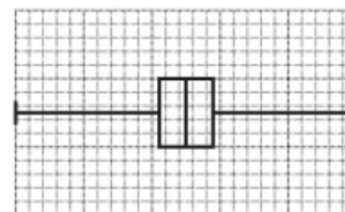
Box plot B



Box plot C



Box plot D



Compare histograms 2 and 4, histogram 2 has less data in the middle of the distribution

1	B
2	D
3	A
4	C

A researcher wants to compare the ages of viewers of BBC 1 and Sky 1.
Write a suitable hypothesis.

He writes a plan to investigate the hypothesis.

Use the Data Handling Cycle to put his plan in the correct order.

- A Work out the mean age for each channel.
- B Select some television viewers to ask.
- C Compare the results and comment on the hypothesis.
- D Collect data about the ages of the television viewers.

A researcher wants to compare the ages of viewers of BBC 1 and Sky 1.

Write a suitable hypothesis.

eg Sky 1 viewers are younger than BBC1 viewers

He writes a plan to investigate the hypothesis.

Use the Data Handling Cycle to put his plan in the correct order.

- 3 A Work out the mean age for each channel.
- 1 B Select some television viewers to ask.
- 4 C Compare the results and comment on the hypothesis.
- 2 D Collect data about the ages of the television viewers.

The sections of a fair spinner are red, white or blue.

The spinner is spun 40 times.

Red	White	Blue	Total
28	9	3	40

Write down the relative frequency of the spinner landing on red.

The spinner has 10 equal sections. Work out the most likely number of sections for each colour.

The sections of a fair spinner are red, white or blue.

The spinner is spun 40 times.

Red	White	Blue	Total
28	9	3	40

$$\begin{array}{r} \text{red} \\ 28 \\ \hline 40 \end{array} = \frac{7}{10}$$

Write down the relative frequency of the spinner landing on red.

The spinner has 10 equal sections. Work out the most likely number of sections for each colour.

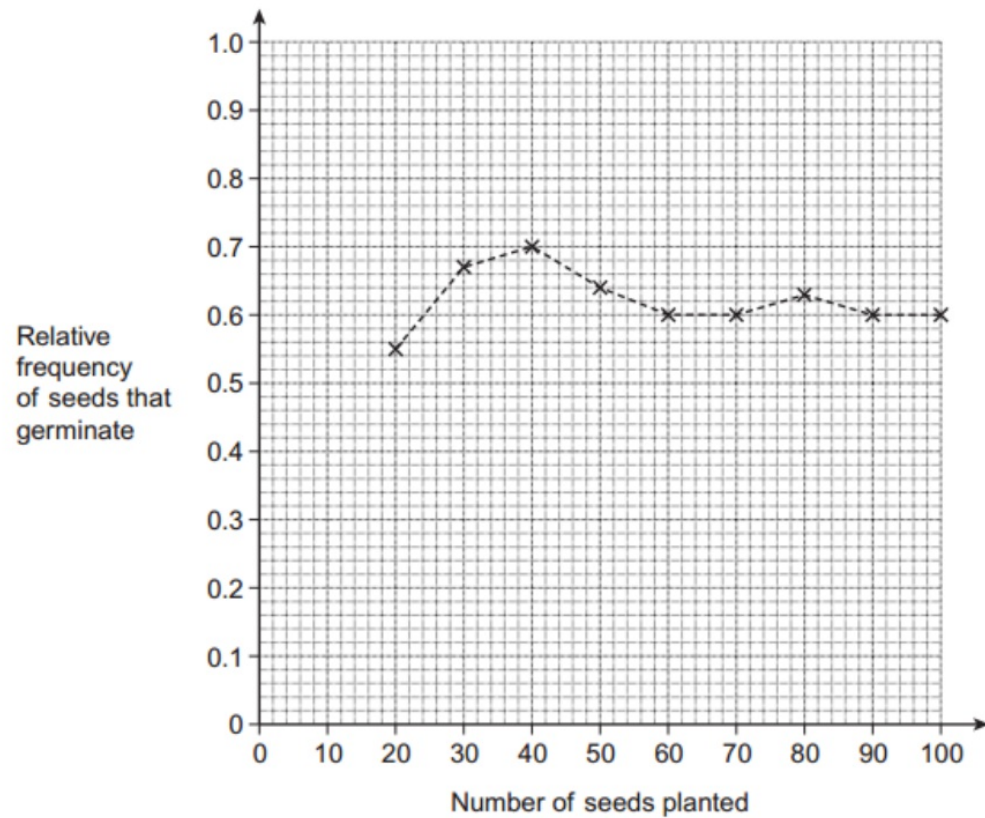
Red	White	Blue	Total
7	2	1	10

Need $\frac{7}{10}$ of the sections red so 7

Need $\frac{9}{40}$ blue (2.25)

and $\frac{3}{40}$ white (0.75) also need whole numbers

A gardener plants ten seeds each week from the same seed packet.
The graph shows the relative frequency of seeds that germinate.

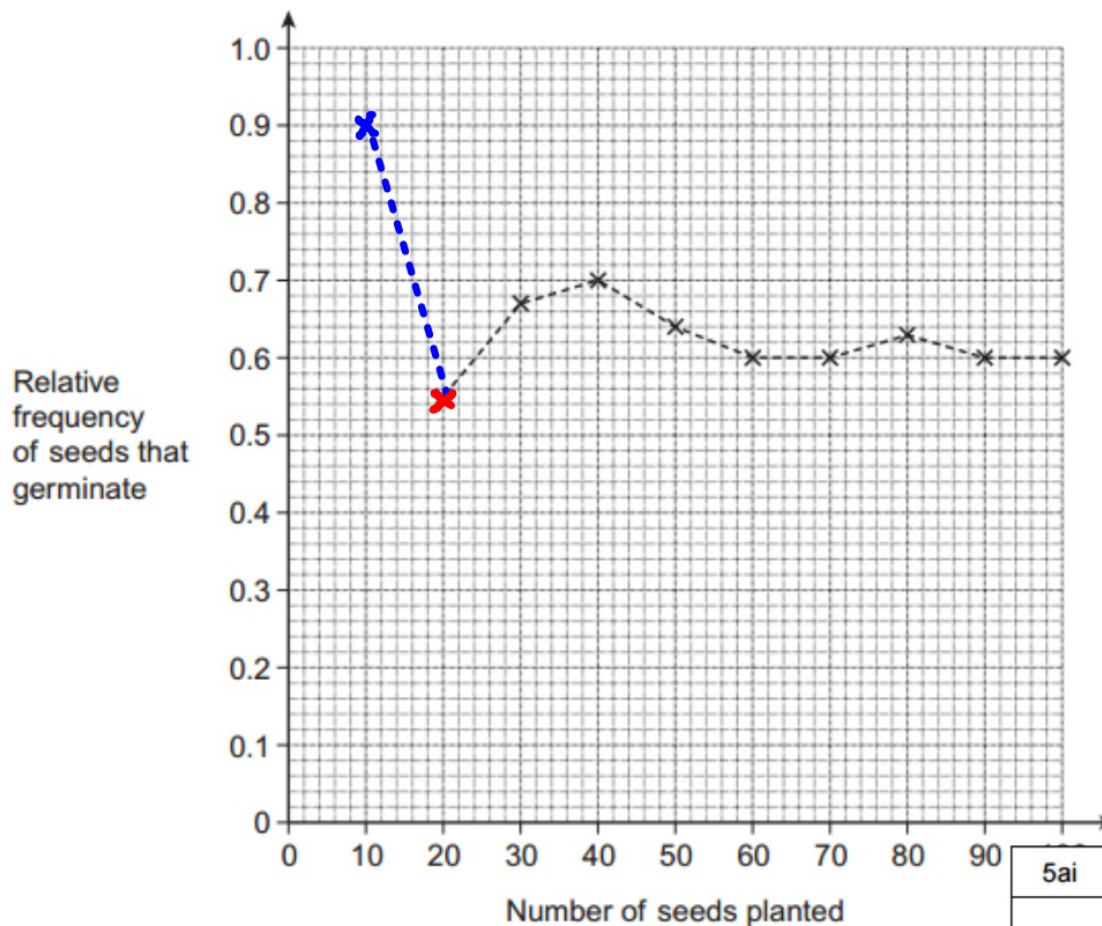


Given that nine seeds out of the ten planted in the first week germinate.

(a)(i) Write down the relative frequency of seeds planted in the first week that germinate.

(a) (ii) Plot your relative frequency on the graph.

(b) How many of the seeds planted in week 2 germinate?



**(b) by the end of week 2
20 seeds planted,
relative frequency is 0.55
so expect $0.55 \times 20 = 11$ to
have germinated by now
and 9 of these germinated
in week 1.**

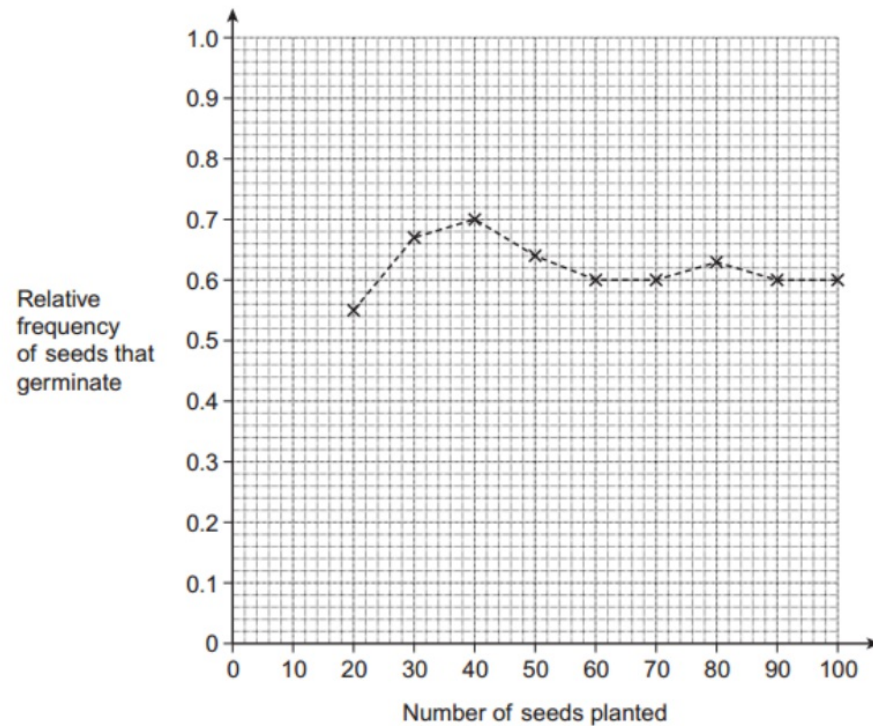
Given that nine seeds out of the ten planted in the
(a)(i) Write down the relative frequency of seeds
that germinate. **0.9**

(a) (ii) Plot your relative frequency on the graph.

(b) How many of the seeds planted in week 2 germinate?

5ai	0.9	B1
5aii	(10, 0.9) plotted	B1 ft
5b	$0.55 \times 20 (-9)$ or 11 (-9)	M1
	2	A1

A gardener plants ten seeds each week from the same seed packet.
The graph shows the relative frequency of seeds that germinate.



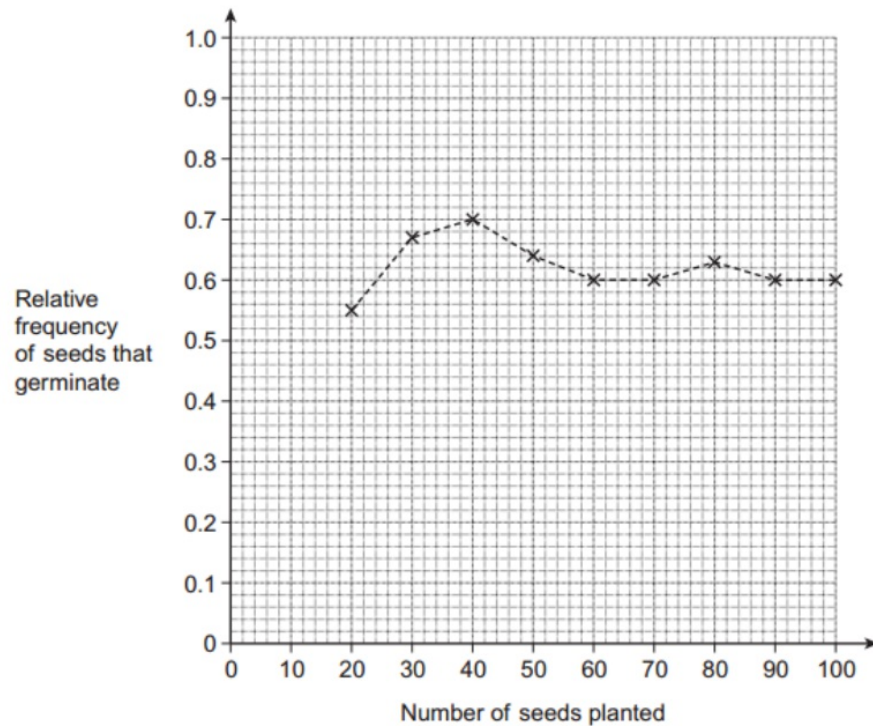
(c) How many of the 100 seeds are expected to germinate

There are 130 seeds in the seed packet.
The label on the packet states:

On average 80 of the seeds will germinate.

Is this statement fair?
Show how you decide.

A gardener plants ten seeds each week from the same seed packet.
The graph shows the relative frequency of seeds that germinate.



(c) How many of the 100 seeds are expected to germinate **60**

There are 130 seeds in the seed packet.
The label on the packet states:

On average 80 of the seeds will germinate.

Is this statement fair?
Show how you decide.

$0.6 \times 130 (= 78)$
78 and no