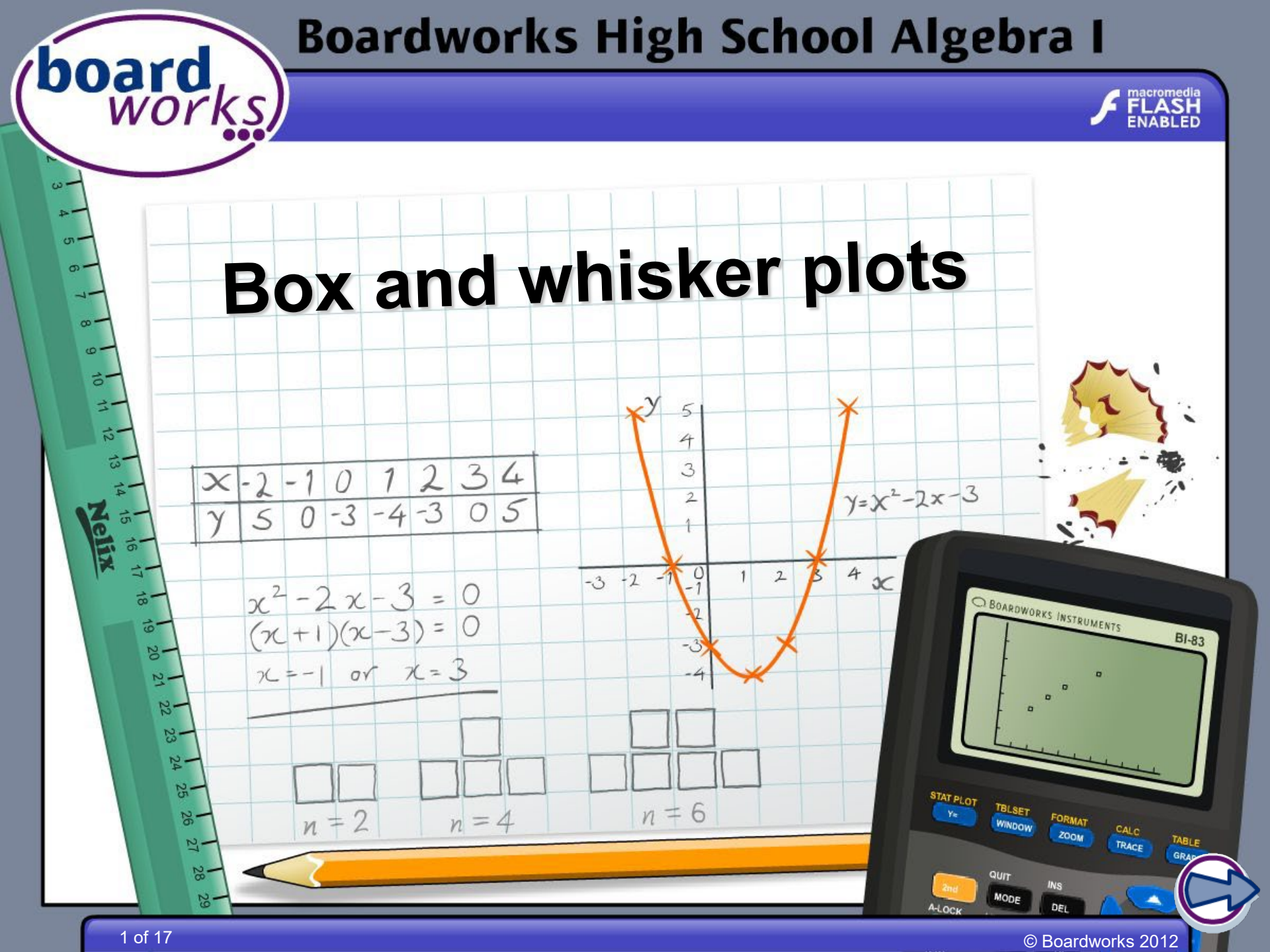
Box and whisker plots



Information

Common core icons

This icon indicates a slide where the Standards for Mathematical Practice are being developed. Details of these are given in the Notes field.

Slides containing examples of mathematical modeling are marked with this stamp.

This icon indicates an opportunity for discussion or group work.

The **Standards for Mathematical Practice** outlined in the Common Core State Standards for Mathematics describe varieties of expertise that mathematics educators at all levels should seek to develop in their students.

They are:

1) Make sense of problems and persevere in solving them.

2) Reason abstractly and quantitatively.

3) Construct viable arguments and critique the reasoning of others.

4) Model with mathematics.

5) Use appropriate tools strategically.

6) Attend to precision.

7) Look for and make use of structure.

8) Look for and express regularity in repeated reasoning.

This icon indicates that the slide contains activities created in Flash. These activities are not editable.

This icon indicates teacher's notes in the Notes field.

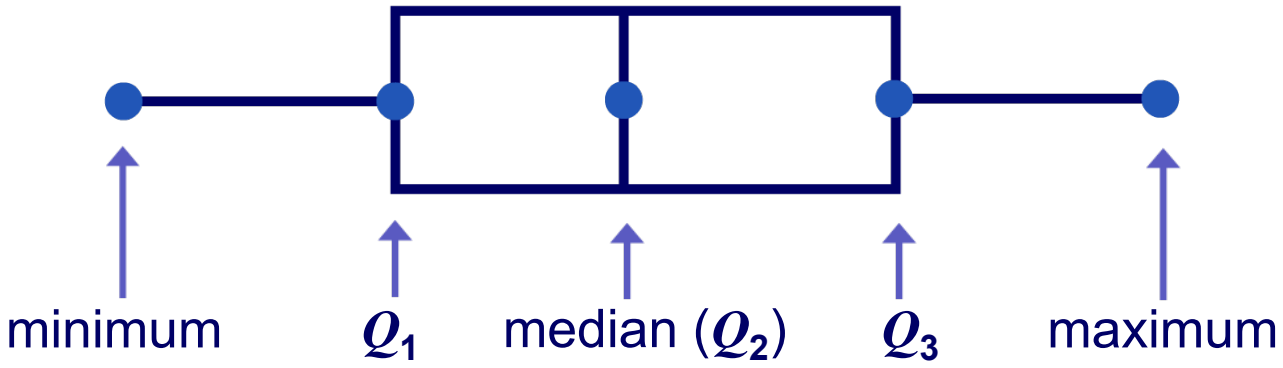
Quartiles and box plots

A set of data can be summarized using five key statistics:

- The **median** (denoted Q_2) – this value is the middle number once the data has been written in order. If there are n numbers in order, the median lies in position $\frac{1}{2}(n + 1)$.
- The **lower quartile** (Q_1) – this value lies one quarter of the way through the ordered data. It is the median of the lower half of values, or the value in position $\frac{1}{4}(n + 1)$.
- The **upper quartile** (Q_3) – this value lies three quarters of the way through the ordered data. It is the median of the upper half of values, or the value in position $\frac{3}{4}(n + 1)$.
- The **minimum** and **maximum** values.

Reading box plots

These five numbers can be shown on a simple diagram known as a **box-and-whisker plot** (or **box plot**):



The box width is the **interquartile range (IQR)**.

interquartile range = $Q_3 - Q_1$

Box plots are drawn along a number line, so that values can be read from them.

Drawing box plots

Median	
Lower quartile	
Upper quartile	
Smallest	
Largest	
Range	
IQR	

plot
plot
plot
plot
plot
plot
plot

Press to reveal values for the boxplot in the table then press 'plot' to see each part drawn in turn.

Draw box plot

Shapes of distributions

roughly symmetrical | positively skewed | negatively skewed

The distribution of a box plot can be described using certain terms.

Press on each of the tabs above to see what shape box plots each term describes.

Comparing box plots

The (ordered) ages of 15 brides marrying at a city hall one month in 2006 were:

18, 20, 20, 22, 23, 23, 25, 26, 29, 30, 32, 34, 38, 44, 53

The **median** is the ½(15 + 1) = 8th number: $Q_2 = 26$

The **lower quartile** is the median of the lower half of values:

$Q_1 = 22$

The **upper quartile** is the median of the upper half of values:

$Q_3 = 34$

The **minimum** and **maximum** values are **18** and **53**.

Draw a box plot of this data.

Comparing box plots

The (ordered) ages of 12 brides marrying at the same city hall in the same month in 2011 were:

21, 24, 25, 25, 27, 28, 31, 34, 37, 43, 47, 61

Q_2 is half-way between the 6th and 7th numbers: $Q_2 = 29.5$

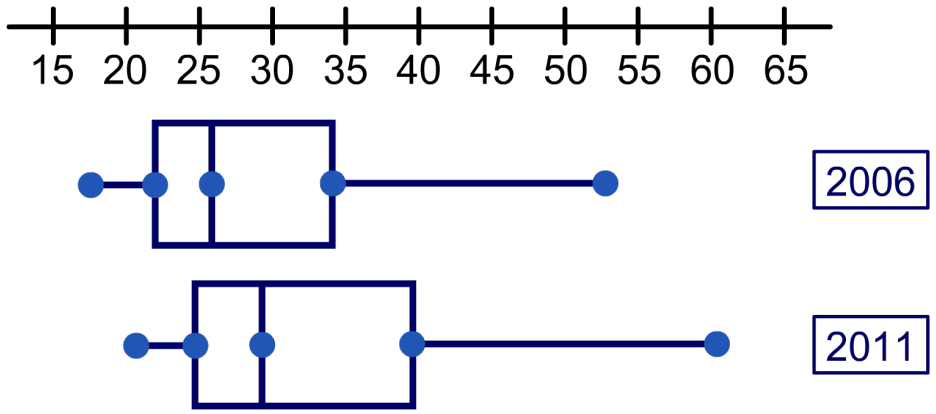
Q_1 is the median of the smallest 6 numbers: $Q_1 = 25$

Q_3 is the median of the highest 6 numbers: $Q_3 = 40$

The minimum and maximum values are **21** and **61**.

Draw a box plot of this data.
Describe the shape of the box plot.

Comparing box plots



These box plots compare the ages of brides in 2006 and 2011.

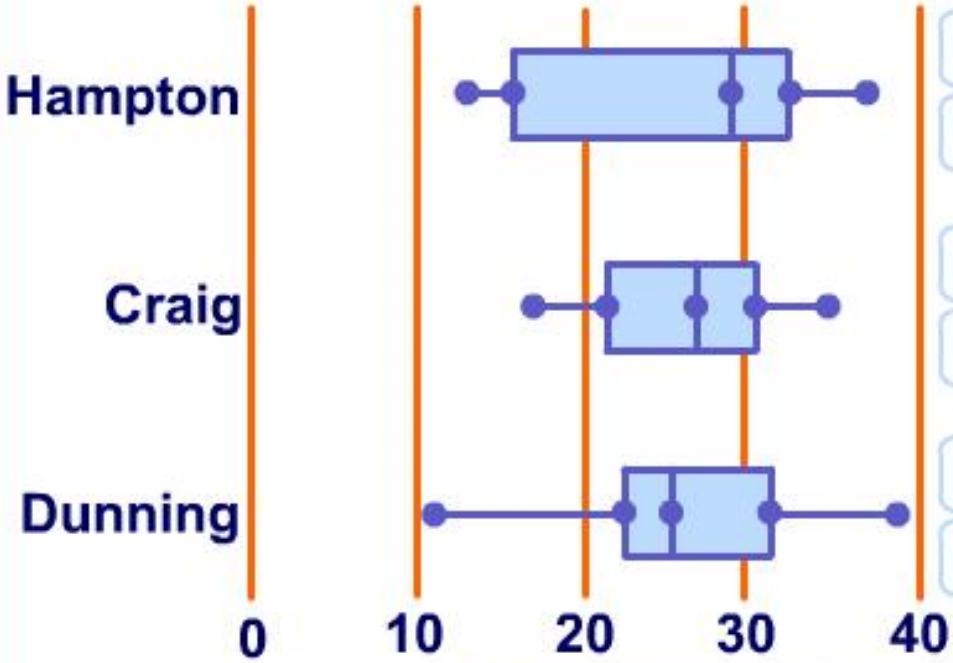
When comparing two box plots, it is important that they are labeled and drawn on the same scale.

The medians show that the brides in 2006 were generally younger than in 2011.

What can you tell from comparing the interquartile ranges?

School exam results

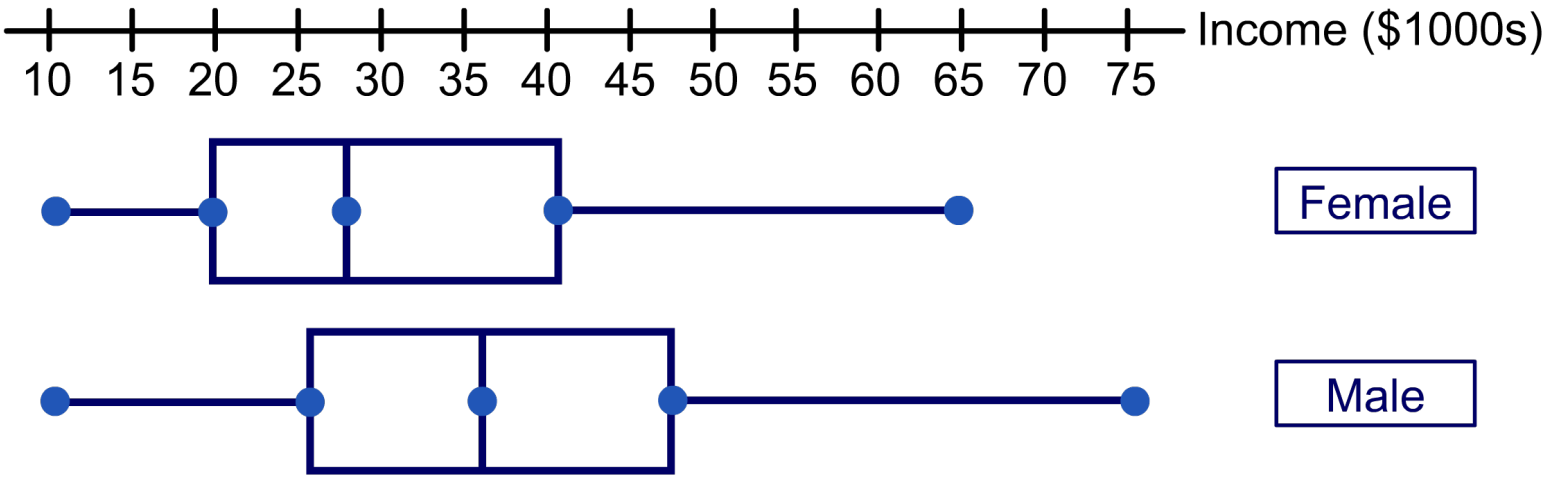
Here are the results for 3 schools in a History exam out of 40 points.
Match the appropriate comments to each school's box plot.
Which school had the best results? Discuss this in groups.



the highest median

Comparing incomes

Income differences between genders for company x



The interquartile range is the same size for males and females. This shows that there is no difference between incomes for different genders.

Is this interpretation correct? Explain your reasoning.

Lap times

James takes part in karting competitions and his mom records his lap times on a spreadsheet.

378 of James' lap times were recorded.

James' fastest time in a race was 51.8 seconds.

In which position in the list is the median lap time?

There are 378 lap times and so the median lap time will be the middle value:

$$\left(\frac{378 + 1}{2}\right)^{\text{th}} \text{ value} \approx \textbf{190}^{\textbf{th}} \textbf{ value}$$

Lap times

In which position in the list is the lower quartile?

There are 378 lap times and so the lower quartile will be the…

$$\left(\frac{378 + 1}{4}\right)^{\text{th}} \text{ value } \approx \textbf{95}^{\textbf{th}} \textbf{ value}$$

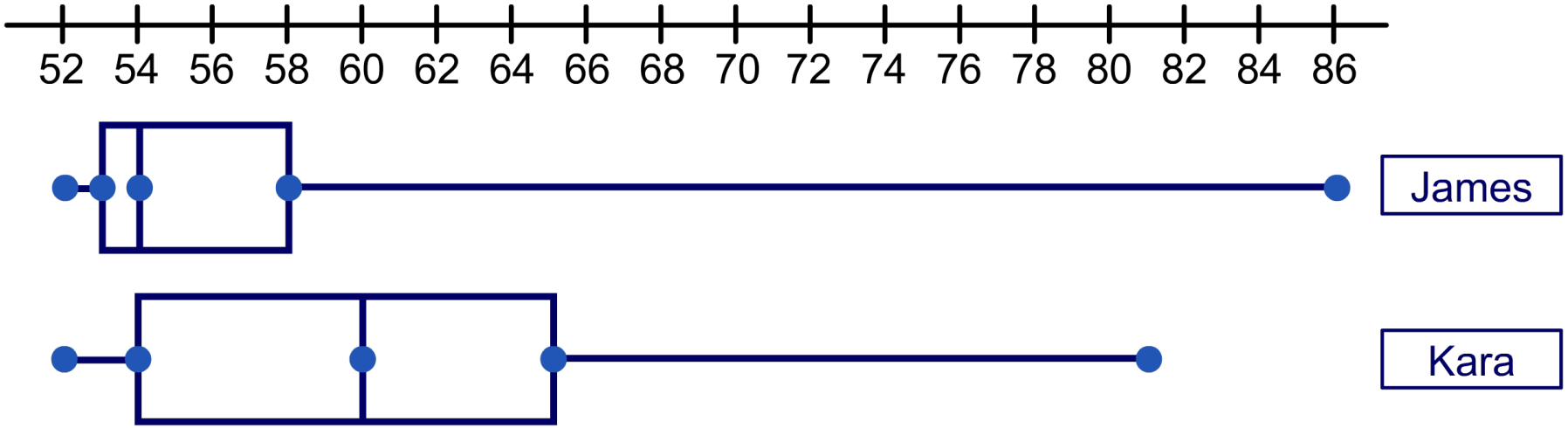
In which position in the list is the upper quartile?

There are 378 lap times and so the upper quartile will be the…

$$\left(3 \times \frac{378 + 1}{4}\right)^{\text{th}} \text{ value } \approx \textbf{284}^{\textbf{th}} \textbf{ value}$$

Comparing lap times

Here are box and whisker plots representing James' lap times and his friend Kara's lap times.



What conclusions can you draw about James' individual performance?

Who is better, James or Kara? Explain your answer.

Cumulative frequency graphs

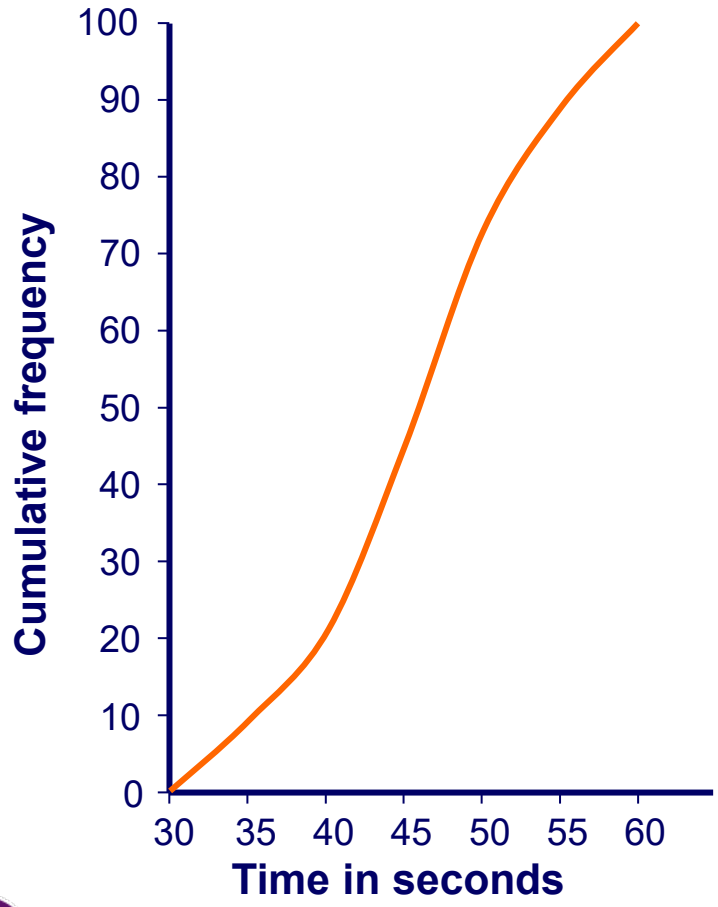
A box plot can be used as an alternative representation of the data displayed in a **cumulative frequency graph**.

Here is the cumulative frequency table showing the number of seconds 100 people can hold their breath.

time in seconds	cumulative frequency
$0 < t \leq 35$	9
$0 < t \leq 40$	21
$0 < t \leq 45$	45
$0 < t \leq 50$	73
$0 < t \leq 55$	89
$0 < t \leq 60$	100

Using graphs

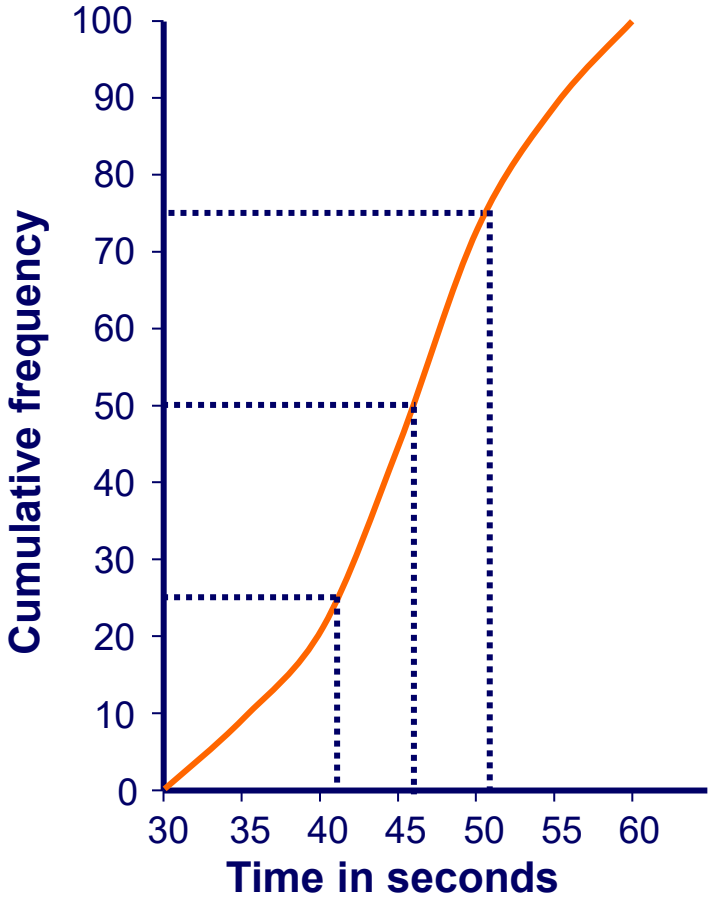
Here is the cumulative frequency graph showing the number of seconds 100 people can hold their breath.



Discuss in groups how you would draw a box plot of this data. Can the median, lower quartile, upper quartile, maximum and minimum values all be found from this graph?

Drawing a box plot from a graph

Once the five required statistics are found from the graph, the corresponding box and whisker plot can be drawn.



Minimum value = **30**

Lower quartile = 25th value = **42**

Median = 50th value = **46**

Upper quartile = 75th value = **51**

Maximum value = **60**