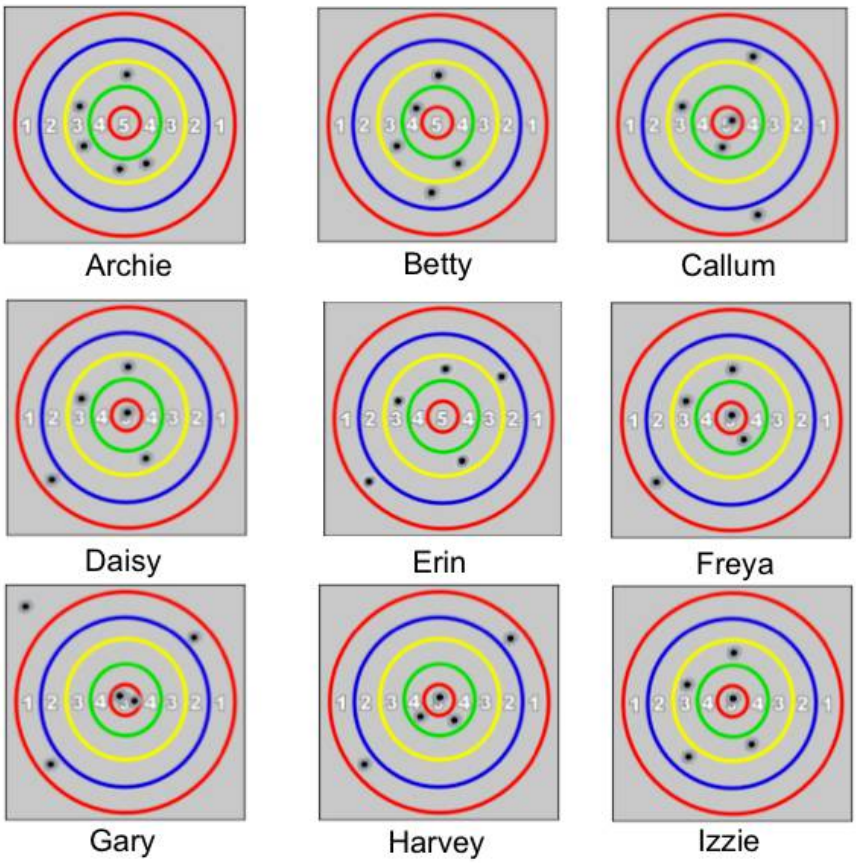
Hotshot

Nine friends took place in shooting competition where each person had 5 shots. Their target maps are shown below. Afterwards there was some disagreement about how the competitors should be ranked.



1 How would you rank to friends in order? Explain your thinking.


Amber

I think Gary should be last as one of his shots missed the target..

But he also had two bull's eyes.


Jon


Matthew

Well I think Archie should win because he was the most consistent.

We need a simple system that takes account of all of their shots. That's only fair.


Alice

In Unit 1, we looked at effective ways of describing distributions with pictures. In this unit we think about how we can describe distributions with numbers."

How many throws does it take to “throw a six”?

Lots of dice games depend on throwing a six



“But on average Watson, how many throws does it take to get a six?”

“I am not too sure actually, what do you think?”

2 How would you answer Sherlock's question?



Miss R

To help make a informed decision, we are going to play a game called “Greedy” in which throwing a six spells disaster and the loss of all of your accumulated points.

Rules of the game

1. Everyone stands up
2. A normal six sided die is rolled twice.
3. Your initial score is the sum of these two throws.
4. You can sit down and record your score on the score sheet.
5. Or you can remain standing and we throw the die again.
6. If the die is a SIX, you lose all of your score, sit down and record your score as ZERO.
7. If the die is not a SIX, you add this to your score.
8. You can now sit down or the game proceeds as before.
9. The aim is to score as many as possible in each round.

You play 5 rounds and your total score is the sum of your round scores

.

Warning

If you are too GREEDY and always choose to remain standing, you will eventually score ZERO.

Section 1 – Summary Statistics



Sarah

Sarah's scores

Round	Score
1	16
2	0
3	23
4	15
5	0
Total	54

Matthew's scores

Round	Score
1	0
2	0
3	0
4	35
5	0
Total	35



Matthew



Alice

As I thought, "Boys are greedier than girls."

I think you need a lot more evidence than that.



Jon

Here are the results for 9Z. We have used a back to back stem and leaf plot to compare the data.

Are Boys greedier than girls?

	Boys		Girls	
(3)	7 0 0	0		
(2)	7 3	1	2 4	(2)
(3)	6 4 3	2	2 8	(2)
(2)	7 5	3	0 2 8	(3)
(3)	6 2 0	4	0 2 3 5 5 5	(6)
		5	4	(1)
		6	1 8	(2)
(2)	8 4	7		
(15)	Count		Count	(16)

Key 3|2 represents 32 points



Evie

I think Alice is right. On average boys are greedier than girls.

On average ...

3

Evie said “On average the boys are greedier than the girls.”
What do you think she means by “On average”?



I think she means that if you find the “average” score for the boys and the “average” score for the girls. The girls will have a higher average.

Sarah

Agreed but there are 3 types of averages. Which one is the best to use?



Matthew

“On average” is a commonly used expression in every day life but as Matthew says there are actually 3 different types of average; the mode, the median and the mean. In this unit we will be considering how to calculate theses 3 measures in different circumstances and, perhaps more importantly, which to decide which measure is the most appropriate to use in each case.

4

In your own words, explain what the three “averages” actual measure and how to find them in each case.



Amber

The “mode” is the most common result. So for the boys the mode is 0 and for the girls it is 45.

Are Boys greedier than girls?

Boys		Girls	
(3)	7 0 0	0	
(2)	7 3	1 | 2 4	(2)
(3)	6 4 3	2 | 2 8	(2)
(2)	7 5	3 | 0 2 8	(3)
(3)	6 2 0	4 | 0 2 3 5 5 5	(6)
		5 | 4	(1)
		6 | 1 8	(2)
(2)	8 4	7	
(15) Count			Count (16)

Key 3|2 represents 32 points

On average ...



Evie

The “median” is the middle result if the results are placed in order. So the median for the boys is 8th result, that is 26.
The median for the girls is between the 8th and 9th results, so between 40 and 42. So the median for the girls is 41.

Are Boys greedier than girls?

Boys		Girls
(3) 7 0 0	0	
(2) 7 3	1 | 2 4	(2)
(3) 6 4 3	2 | 2 8	(2)
(2) 7 5	3 | 0 2 8	(3)
(3) 6 2 0	4 | 0 2 3 5 5 5	(6)
	5 | 4	(1)
	6 | 1 8	(2)
(2) 8 4	7	
(15) Count		Count (16)

Key 3|2 represents 32 points

5

Can you explain why Evie choose the 8th number for the boys but mid way between the 8th and 9th number for the girls?.



Matthew

Matthew's explanation

15 boys results 0, 0, 7, 13, 17, 23, 24, 26, 35, 37, 40, 42, 46, 74, 78

7 results Middle result 7 results

16 girls results 12, 14, 22, 28, 30, 32, 38, 40, 42, 43, 45, 45, 45, 54, 61, 68

8 results Middle results 8 results

On average ...



Jon

Jon's working

The "mean" is a way of sharing out the total number of points equally.
In other words, equalising their scores.

Boys:
0 + 0 + 7 + 13 + 17 + 23 + 24 + 26 + 35 + 37 + 40 + 42 + 46 + 74 + 78
= 462
462 ÷ 15 = 30.8
The boys mean is 30.8.

Girls:
12 + 14 + 22 + 28 + 30 + 32 + 38 + 40 + 42 + 43 + 45 + 45 + 45 + 54 +
61 + 68 = 619
The girls mean is 38.7

So the girls have a higher mean than the boys.



Sarah

So in each case the "average" score for the girls was higher than the boys!
We can say confidently that "Boys are greedier than girls."

6

Now use the results for your group to test the hypothesis that "Boys are greedier than girls"..
Do all of the averages give the same result?



Alice

Adding up the total score from a stem and leaf diagram was quite difficult.
Is there an easier way of adding up the score?

7

Can you think of a way of adding up the numbers that would be faster, easier and more accurate?

Section 1 – Summary Statistics

Fitness check

In the first 3 cases find the mode, median and mean for each dataset. Which average would you say is most sensible to use to describe a "typical" result.



1. Olympic golf tournament 1st round scores.



2. Gym membership by age

Gym members by age

2	2 4 5 8 9	(5)
3	0 1 1 5 7 8 9	(7)
4	0 3 4 5 8 9	(6)
5	2 4 5 6 7 7 8 9	(8)
6	1 4 5 7 9 9 9	(7)
7	0 0 2 3 6 7 8 7 9	(9)
8	0 2 3	(3)
	Count	(45)

Key 2|5 represents 25 years old

3. Reading scores from two reception classes
   a) Class 1: 5.3, 4.9, 5.2, 5.4, 5.6, 5.1, 5.8, 5.3, 4.9, 6.1, 6.2, 5.7, 5.4, 6.9, 4.3, 5.2, 5.6. 5.9, 5.3, 5.8
   b) Class 2: 4.7, 5.0, 5.5, 4.1, 6.8, 5.0, 4.7, 5.6, 4.9, 6.3, 7.8, 3.6, 8.4, 5.4, 4.7, 4.4, 5.6, 3.7, 6.2, 7.5

How many throws does it take to “throw a six”?



Miss R

In the Greedy game, the key to making a high score was anticipating when, on average, you might get a six. Based on your experience so far, how would you answer that question now?
We will collect up your estimates and record the group's thoughts. This we will call the **wisdom of the crowd**.



Evie

Evie's working

Our initial thoughts
4, 6, 5, 3, 7, 4, 9, 5, 6, 5, 8, 10, 8, 6, 9, 6,
6, 9, 6, 5, 4, 7, 6, 8, 5, 5, 3, 8, 6, 7
30 results
Reordering them smallest to largest
3, 3, 4, 4, 4, 5, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6,
6, 6, 7, 7, 7, 8, 8, 8, 8, 9, 9, 9,10
Mode = 6
Median = 6
Mean = 186 ÷ 30 = 6.2



Miss R

Now you going to carry out an experiment of your own using a six sided dice.
In pairs, one of you throw a die repeatedly until your get a six.
The other person count the number of throws needed each time and record the results.
Have 5 goes each and then swap over.
This means we will collectively have a lot of data and that's better shown in a frequency diagram.

9Z – Experimental data.

Number of throws (x)	Frequency (f)	Number of throws (x)	Frequency (f)	Number of throws (x)	Frequency (f)
1	22	9	4	17	2
2	15	10	0	18	1
3	15	11	1	19	0
4	9	12	2	20	0
5	11	13	2	21	1
6	12	14	1	22	0
7	6	15	3	23	0
8	8	16	1	24	1

How many throws does it take to “throw a six”?

8

Can devise an effective method of calculating the three averages from a frequency table?

Number of throws (x)	Frequency (f)
1	22
2	15
3	15
4	9
5	11
6	12
7	6
8	8
9	4
10	0
11	1
12	2
13	2
14	1
15	3
16	1
17	2
18	1
19	0
20	0
21	1
22	0
23	0
24	1
Total	

We need to know how many goes we had altogether.



Alice

And the total number of throws it took to get a six.



Jon



Evie

I've got an idea

How many throws does it take to “throw a six”?

Add an extra column!


Evie

Number of throws (x)	Frequency (f)	fx
1	22	22
2	15	30
3	15	45
4	9	36
5	11	55
6	12	72
7	6	42
8	8	64
9	4	36
10	0	0
11	1	11
12	2	24
13	2	26
14	1	14
15	3	45
16	1	16
17	2	34
18	1	18
19	0	0
20	0	0
21	1	21
22	0	0
23	0	0
24	1	24
Total	117	1055

The most common answer is 1. So the mode 1.


Sarah


Matthew

The mean is 1055 ÷ 117.
So the mean is 9.0

The median is the middle result, so number 59.
That is 4.


Amber


Alice

That's interesting. This time the three averages are quite different.

I think I know why. It's to do with the shape of the distribution.


Jon

9 Can you explain Jon's thinking?

Section 1 – Summary Statistics

Fitness check

In each of these cases summarise the data by finding the mode, median and mean.



1. Julianne is confident that the girls in her form are fitter than the boys and so challenged them to take part in a two minute step up challenge. Each pupil did as many steps ups as possible in two minutes.
   The results are shown in this back to back stem and leaf plot.
   Do your findings for the mode, median and mean support Julianne's claim?

Are girls fitter than boys?

Boys			Girls
(0)		3 | 6	(1)
(3)	6 4 2	4 | 3 7	(2)
(2)	9 5	5 | 2 4 5 9 9	(5)
(5)	5 4 4 4 0	6 | 0 0 0 4	(4)
(1)	2	7 | 2 9	(2)
(2)	7 4	8 |	(0)
(2)	8 2	9 | 6	(1)
(15)	Count		Count (15)

Key 1 \| 2 represents 12 points

2. Here is a back to back stem and leaf diagram of the heights of the pupils in 9Y and 9Z
   Find the mode, median and mean of each dataset.
   Do your results support the claim that 9Y are taller than 9Z?

Are 9Y taller than 9Z?

9Z			9Y	
(3)	8 8 5	13 | 8		(1)
(3)	4 4 3	14 | 1 4		(2)
(2)	9 7	14 | 8 8 8 8		(4)
(5)	4 3 2 1 0	15 | 0 1 1 4 4		(5)
(7)	9 9 7 7 6 6 5	15 | 5 5 6 7 8 8 9 9		(9)
(8)	4 4 4 2 1 1 1 0	16 | 0 0 1 3 4		(5)
(2)	9 5	16 | 6 8		(2)
		17 | 0		(1)
(30)	Count			Count (29)

Key 14 \| 1 represents 141cm

3. The number of goals scored by the Year 8 hockey team in each match last season is listed below.

   0, 1, 2, 2, 3, 1, 3, 2, 0, 1, 5, 2, 4, 0, 2, 1, 2, 3, 5, 5

   Copy and complete the frequency table below.

Score (x)	Frequency (f)	Frequency x score (fx)
0	3	0
1	4	4
2	6	
3	3	
4		
5		
Total		

   From this data, calculate the mean number of goals scored per match last season.

4. The number of goals scored by the Year 8 hockey team in each match this season is given in the frequency table below.

Score (x)	Frequency (f)	Frequency x score (fx)
0	1	
1	5	
2	4	
3	3	
4	4	
5	3	
Total	20	

   From this data, calculate the mean number of goals scored per match last season.

5. By considering the mode, median and mean scores of the Year 8 hockey team's scores over the past two seasons, which team would you consider to be the most attacking? Make sure you give your reasons clearly.

6. 8X and 8Y held a sponsored Geography problem solving competition to raise money for a local charity. There were 10 problems to solve in one hour.

   Their results are shown below.

8X Number of correct solutions (x)	Frequency (f)
0	0
1	0
2	0
3	2
4	7
5	6
6	5
7	5
8	1
9	2
10	0
Total	28

8Y Number of correct solutions (x)	Frequency (f)
0	0
1	0
2	4
3	2
4	4
5	7
6	5
7	2
8	2
9	0
10	1
Total	27

   a) From this data, calculate the mean, mode and median number of correct solutions for each group.

   b) Which average would you prefer to use if you were a member of 8X?

   c) Which average would you prefer to use if you were a member of 8Y?

5. Simon and Pat are keen golfers and played together 36 times last season. Their matches are generally very close. Their scores are shown in the frequency table below.

Score	66	67	68	69	70	71	72	73	74	75
Simon's frequency	1	0	1	3	6	8	7	4	4	2
Pat's frequency	0	0	0	3	7	9	9	5	3	0

Produce a meaningful statistical argument as to who should be described as the better player.

On average, how many points did the rugby team score?





Miss F

The England Ladies rugby team have had an incredible run of victories. They seem to be scoring lots of points. Here is some data to help you but, on average, how many points have they scored per match?

Points scored	Frequency (f)
21 - 25	1
26 - 30	2
31 - 35	3
36 - 40	2
41 - 45	3
46 - 50	0
51- 55	2
56 - 60	4
Total	17



Jon

Well the modal group is 56 – 60.

The median would be the 9th result, so the median is in the 41 – 45 group.



Alice



Taro

I've got an idea how we could "estimate" the mean.

10 What do you think Taro's idea might be?

On average, how many points did the rugby team score?



Sarah

Taro said, “estimate”, so I think he is going to assume everyone in the each group scored the same amount.

In which, perhaps he gave every the middle score in each group?



Matthew



Taro

Taro's working

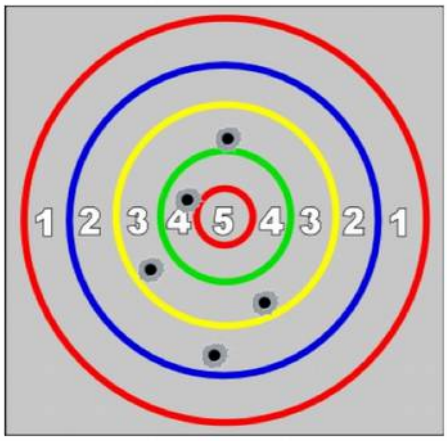
Points scored	Mid point (x)	Frequency (f)	fx
21 - 25	23	1	23
26 - 30	28	2	56
31 - 35	33	3	99
36 - 40	38	2	76
41 - 45	43	3	129
46 - 50	48	0	0
51- 55	53	2	106
56 - 60	58	4	232
Total	n/a	17	675

An approximation for the mean = 675 ÷ 17 = 39.7 points per game.

11 Can you explain why Taro says this is an “estimate” of the mean?

Section 1 – Summary Statistics

Do you get better with practice ?



9X and 9Y went away for an activity weekend. One of the activities was “Target Practice” where each student had 5 shots at their target and the one with the highest total score was awarded the title of “Hotshot of the day”.

On Friday evening everyone had their first go. Their results are shown below in Table 1.

On Saturday morning, they were all given 30 minutes free practice to help them improve. Their best scores are recorded in Table 2.

On Sunday afternoon, they held another competition. Their results are shown in Table 3.

Score	Frequency
1 - 5	4
6 - 10	32
11 - 15	18
16 - 20	1
21 - 25	0
Total	55

Table 1

Score	Frequency
1 - 5	0
6 - 10	15
11 - 15	30
16 - 20	10
21 - 25	0
Total	55

Table 2

Score	Frequency
1 - 5	1
6 - 10	33
11 - 15	17
16 - 20	4
21 - 25	0
Total	55

Table 3

In each case state the modal group, the group in which the median lies and estimate the mean.



Mr T

Now, compare the results from the three days using a statistical argument.
Would you say that the students did get better with practice?
Justify your conclusions.

Measuring spread

Psychology students at a local university were carrying out an investigation into stress. They were interested in how "stressful" students found their university examinations. The students had their pulse measured during a lecture at college two weeks before their exams began and then again immediately after their first exam. The results are shown below.

Pulse rate at rest

4	6 8 8	(3)
5	2 4 7 7	(4)
6	4 8 8 9 9	(5)
7	0 2 2 2 5 6	(6)
8	6	(1)
	Count	(19)

Key 5 | 2 represents 52 beats per minute

Pulse rate after the exam

4	9	(1)
5	8 9	(2)
6	2 5 8 8 8	(5)
7	4 6 8 8 9 9	(6)
8	2 3 5 8 9	(5)
	Count	(19)

Key 6 | 2 represents 62 beats per minute

12 Find the mode, median and mean pulse rate for each group. Comment on your results.



Matthew

Matthew's working

At rest Mode = 72, Median = 68, Mean = 1223 ÷ 19 = 64.4

After exam Mode = 68, Median = 76, Mean = 1388 ÷ 19 = 73.1

The mode went down (that's a surprise) but the median and the mean both went up by about by about 12 – 13%. That suggests the mode isn't very helpful in this case.

Splitting the data into 4

Although the three averages are useful in summarising the data they don't tell the whole story as they reveal nothing about how spread out the data might be.

One way to measure spread is to record the minimum and maximum values. The **range** is then the difference between these two numbers.

In the first case the range = 86 – 46 = 40 beats per minute.

This measure shows the full spread of the data but can be misleading as it is influenced by extreme results or **outliers**. For example, if we ignore the 86, the range falls to 30 beats per minute.

We can improve our description of the spread by also looking at the spread of the middle half of the data and we do this by splitting the data in 4 groups.

The median represents the middle of a distribution. In other words it is where would you split the group so that there an equal number below and above the median.

In that way we have split the original data into 4 quarters. These new splits are called the **lower quartile** (First quartile, Q_1 or 25th percentile) and the **upper quartile** (Third quartile, Q_3 or 75th percentile). The median is sometimes written as Q_2 (the second quartile) or the 50th percentile. The interquartile range (IQR) measures the spread of the middle 50% of the results as is calculated by

$$IQR = Q_3 - Q_1$$

Finding the quartiles

To find the quartiles, the data must be in order.
The at rest pulse rate data from the previous page is listed in order below

46, 48, 48, 52, 54, 57, 57, 64, 68, 68, 69, 69, 70, 72, 72, 72, 75, 76, 86

Step 1 – Find the median in this case 68 beats per minute

Step 2 – Split the distribution into two new distributions, those below the median and those above the median.

Step 3 – Find the median of each of the new distibutions.

Step 4 – Calculate the IQR

13 Can you find the range and interquartile range for the "at rest" pulse data?



Sarah

Sarah's working

46, 48, 48, 52, 54, 57, 57, 64, 68, 68, 69, 69, 70, 72, 72, 72, 75, 76, 86

Q_1 Q_2 Q_3

Range = 86 – 46 = 40 beats per minute

IQR = $Q_3 - Q_1$ = 72 – 54 = 18 beats per minute

14 Can you find the range and interquartile range for the "after the exam" pulse data? Comment on the differences between the two sets of results



Alice

Alice's working

49, 58, 59, 62, 65, 68, 68, 68, 74, 76, 78, 78, 79, 79, 82, 83, 85, 88, 89

Q_1 Q_2 Q_3

Range = 89 – 49 = 40 beats per minute

IQR = $Q_3 - Q_1$ = 82 – 65 = 17 beats per minute

All the results appear to have moved up by roughly the same amount as shown by the range being the same and the IQR being almost the same.
This is in line with our comments about the median and the mean going up by between 12 and 13%.

Comparing distributions

The interquartile range (IQR) is often useful in comparing distributions. It is a measure of the spread of the middle 50% of the data as it measures the gap between the lower and upper quartiles. For example, earlier we looked at the data from Julianne's step up test. She believed that this would show that girls are fitter than boys.
We can now look at this data more closely.
It is easier if the data is listed in ascending order.

Boys: 42, 44, 46, 55, 59, 60, 64, 64, 64, 65, 72, 84, 87, 92, 98

Girls: 36, 43, 47, 52, 54, 55, 59, 59, 60, 60, 60, 64, 72, 79, 96

15 Can you find the interquartile range for each dataset?


Evie

Evie's working

Boys: 42, 44, 46, (55), 59, 60, 64, (64), 64, 65, 72, (84), 87, 92, 98

Q_1 Q_2 Q_3

$IQR = Q_3 - Q_1 = 84 - 55 = 29$ step ups

Girls: 36, 43, 47, (52), 54, 55, 59, (59), 60, 60, 60, (64), 72, 79, 96

Q_1 Q_2 Q_3

$IQR = Q_3 - Q_1 = 64 - 52 = 12$ step ups

16 Can you now interpret your results?


Matthew

It shows that in the middle, the girls' results are less spread out.

It shows that the middle 50% of the girls' results are much closer together than the boys' results.


Sarah


Alice

It shows that the middle 50% of girls are more consistent than the middle 50% of the boys.

Comparing distributions graphically using the five number summary

Because the mean is sensitive to the presence of outliers, data analysts sometimes prefer a less sensitive way of summarising data. The **five number summary** is an alternative method of summarising a dataset and includes the median and the quartiles, which are less sensitive to outliers. The five number summary also includes the minimum and maximum data values which are, of course, affected by outliers.

The **five number summary** consists of the following measures

1. Minimum; the smallest value in the data set Q_0
2. Lower quartile, or first quartile, Q_1
3. Median, Q_2
4. Upper quartile or third quartile, Q_3
5. Maximum; the largets value in the data set, Q_4

Earlier we looked at the data from Julianne's step up test which she believed would show that girls are fitter than boys.

17 Can you find the five number summary for each dataset?



Evie

Evie's working

Boys: 42, 44, 46, 55, 59, 60, 64, 64, 64, 65, 72, 84, 87, 92, 98

Q_0 Q_1 Q_2 Q_3 Q_4

Range $= Q_4 - Q_0 = 98 - 42 = 56$ step ups

IQR $= Q_3 - Q_1 = 84 - 55 = 29$ step ups

Girls: 36, 43, 47, 52, 54, 55, 59, 59, 60, 60, 60, 64, 72, 79, 96

Q_0 Q_1 Q_2 Q_3 Q_4

Range $= Q_4 - Q_0 = 96 - 36 = 60$ step ups

IQR $= Q_3 - Q_1 = 64 - 52 = 12$ step ups

Comparing distributions graphically using a box plot.

As e have emphasised many times over throughout this course, a picture or diagram is often very useful in summarising and comparing distributions.

The **boxplot** (sometimes called a **box-and whisker plot**) is a convenient graphical display of the five-number summary of a data set. The boxplot allows the data analyst to evaluate the **symmetry** or **skewness** of a data set.

Once again, look at the data from Julianne's step up test which she believed would show that girls are fitter than boys.

We can summarise the two data sets, using the five number summaryand by constructing comparative box plots.

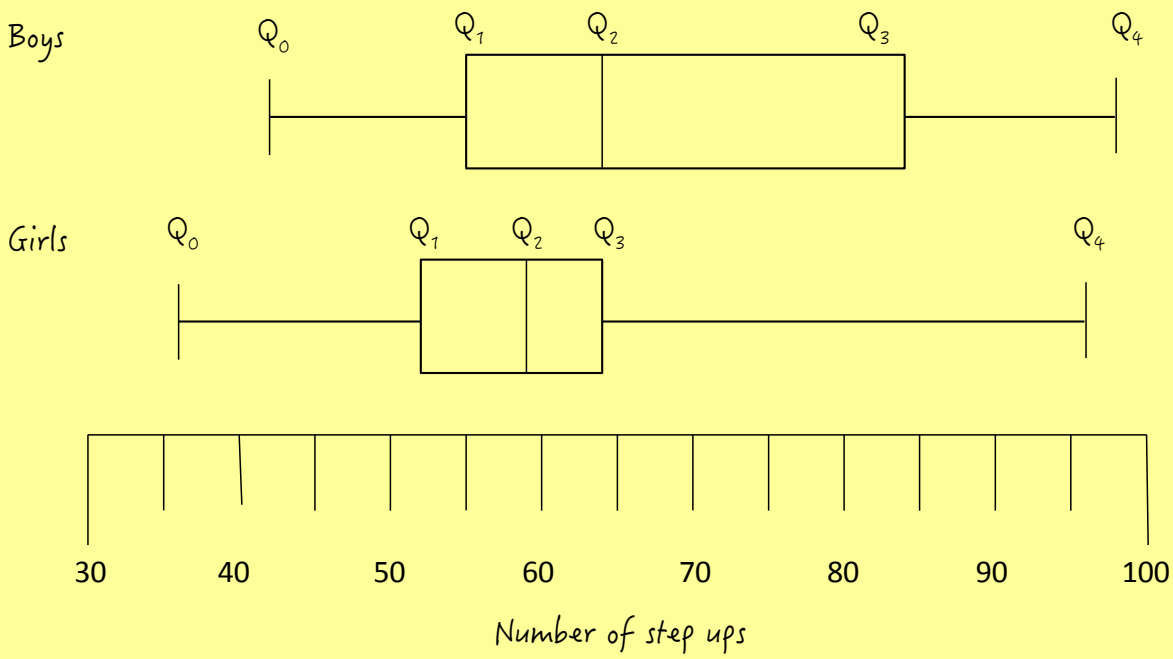


Mr T

Mr T's working

Boys: $Q_0 = 42$, $Q_1 = 55$, $Q_2 = 64$, $Q_3 = 84$, $Q_4 = 98$

Girls: $Q_0 = 36$, $Q_1 = 52$, $Q_2 = 59$, $Q_3 = 64$, $Q_4 = 96$



18 Why do you think it is important to put in the scale?

19 How would you compare these two distributions based on the boxplots?

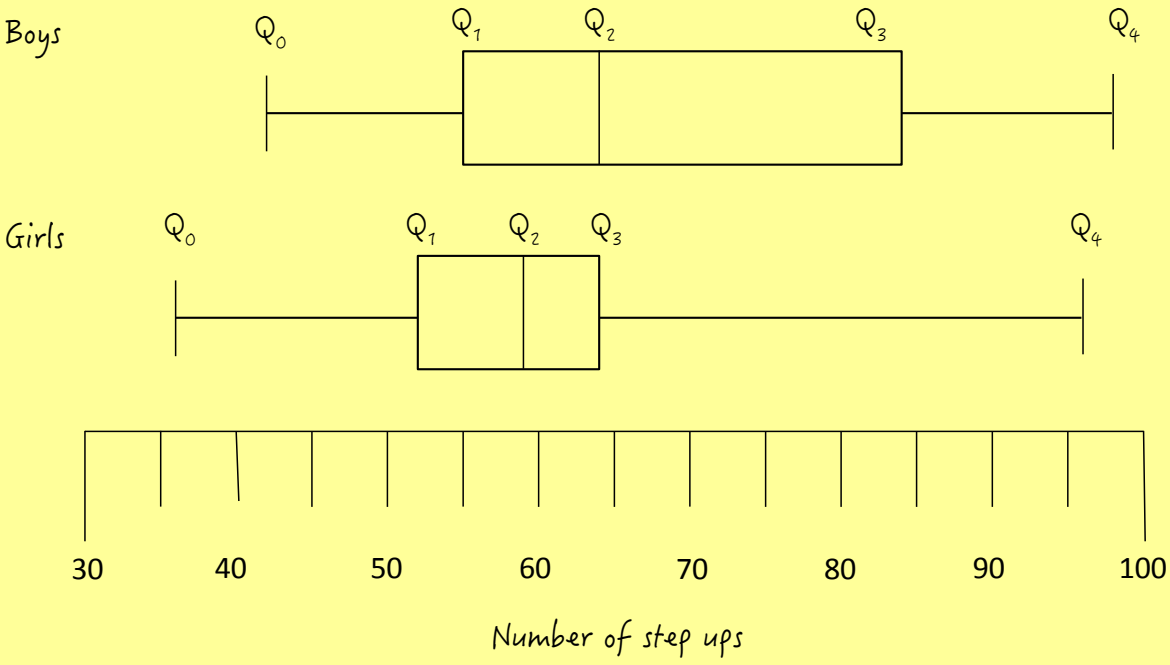


Mr T

Mr T's working

Boys: $Q_0 = 42$, $Q_1 = 55$, $Q_2 = 64$, $Q_3 = 84$, $Q_4 = 98$

Girls: $Q_0 = 36$, $Q_1 = 52$, $Q_2 = 59$, $Q_3 = 64$, $Q_4 = 96$





Sarah

The scale is important because we want to summarise our thoughts by simply looking at the picture – so it needs to represent the data accurately.



Taro

The results of the top 25% of the girls are spread out much more than the top 25% of the boys.



Amber

The results of the middle 50% of the girls' results are much more closely grouped than the middle 50% of the boys' results.



Matthew

The results of the boys in the the third quartile are much more spread out than the results of the quartiles in the third quartile of their results.

Is it really getting hotter?



The summer of 2022 will be remembered as a dry and sunny three months, and for England, the joint warmest summer on record according to the Met Office.

The Met Office say that this means that four of the five warmest summers on record for England have occurred since 2003, as the effects of human-induced climate change are felt on England's summer temperatures.



Mr T

But is this true for all places in England? This data suggests that out weather is quite different at these four seaside locations.

Maximum temperature (°C) in August								
Town	2015	2016	2017	2018	2019	2020	2021	2022
Bognor Regis	26	27	26	34	31	29	22	35
Blackpool	25	25	23	23	28	31	22	31
Whitley Bay	22	23	23	26	26	26	22	30
St Ives	20	22	23	24	22	24	22	30

19 Can you think of some different ways we might compare these distributions?



Mikka

I am going to look at different types of average

I am going to use line graphs



Jasmine

We are going to look at the 5 number summaries ...



Sarah



Meera

… and then we are going to draw boxplots

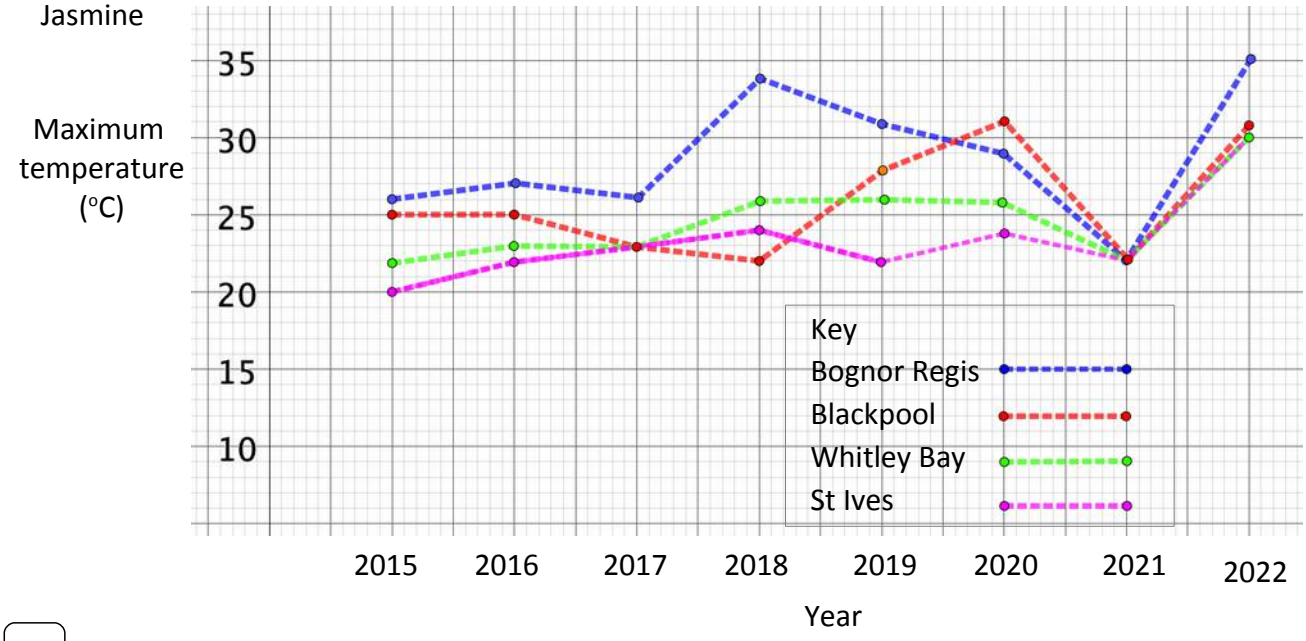
Comparing the data sets

Maximum temperature (°C) in August								
Town	2015	2016	2017	2018	2019	2020	2021	2022
Bognor Regis	26	27	26	34	31	29	22	35
Blackpool	25	25	23	23	28	31	22	31
Whitley Bay	22	23	23	26	26	26	22	30
St Ives	20	22	23	24	22	24	22	30



Jasmine

I am going to use line graphs.



20

Look carefully at Jasmine's line graphs. Talk to a friend and just say what you can see.



Mikka

Bognor Regis is nearly always the warmest place

St Ives is nearly always the coolest place.



Charo



Jon

The widest range of temperatures was in 2018

Whitley Bay seems to be fairly consistent. Except for 2022!



Evie

Comparing the data sets

Maximum temperature (°C) in August								
Town	2015	2016	2017	2018	2019	2020	2021	2022
Bognor Regis	26	27	26	34	31	29	22	35
Blackpool	25	25	23	23	28	31	22	31
Whitley Bay	22	23	23	26	26	26	22	30
St Ives	20	22	23	24	22	24	22	30



Mikka

I am going to look at different types of average

Mikka's working

First, I am going to put the daily temperatures for each town in order

BognorRegis	22, 26, 26, 27, 29, 31, 34, 35	Total 230
Blackpool	22, 23, 23, 25, 25, 28, 31, 31	Total 208
Whitley Bay	22, 22, 22, 23, 26, 26, 26, 30	Total 197
St Ives	20, 22, 22, 22, 23, 24, 24,30	Total 187

	Mode	Median	Mean
Bognor Regis	26	28	28.75
Blackpool	23, 25, 31	25	26.0
Whitley Bay	22, 26	24.5	24.6
St Ives	22	22.5	23.4

21

Look carefully at Mikka' summary. Discuss with a friend what these averages show you? What average would you say sumarises the data most fairly?



Amber

The mode is ceratinly the least helpful average.



Mikka

The mean for St Ives is distorted by one very high temperature.

In that case, I think the median is the most useful statitic.



Taro

Comparing the data sets

Maximum temperature (°C) in August								
Town	2015	2016	2017	2018	2019	2020	2021	2022
Bognor Regis	26	27	26	34	31	29	22	35
Blackpool	25	25	23	23	28	31	22	31
Whitley Bay	22	23	23	26	26	26	22	30
St Ives	20	22	23	24	22	24	22	30



Meera

We are going to look at the 5 number summaries ...

... and then we are going to draw boxplots



Sarah

First, we are going to put the daily temperatures for each town in order

BognorRegis	22,	26,	26,	27,	29,	31,	34,	35
Blackpool	22,	23,	23,	25,	25,	28,	31,	31
Whitley Bay	22,	22,	22,	23,	26,	26,	26,	30
St Ives	20,	22,	22,	22,	23,	24,	24,	30
		Q_0		Q_1		Q_2		Q_3		Q_4

22

Can you decide how you should work out the lower and upper quartiles when it it lies between two numbers?



Jasmine

We could use the same method as we use for finding the median when it's between two numbers! That would make sense.

23

Calculate the 5 number summaries for each of the four towns.

Section 2 – Measuring spread



Meera

We are going to look at the 5 number summaries ...

... and then we are going to draw boxplots



Sarah

First, we are going to put the daily temperatures for each town in order

BognorRegis	22,	26,	26,	27,	29,	31,	34,	35
Blackpool	22,	23,	23,	25,	25,	28,	31,	31
Whitley Bay	22,	22,	22,	23,	26,	26,	26,	30
St Ives	20,	22,	22,	22,	23,	24,	24,	30

	Q_0	Q_1	Q_2	Q_3	Q_4
Bognor Regis	22	26	28	32.5	35
Blackpool	22	23	25	29.5	31
Whitley Bay	22	22	24.5	26	30
St Ives	20	22	22.5	24	30

24

Now plot the 4 boxplots on one graph.



Mr T

Look carefully at your boxplots and explain exactly what they show you.
Try to include as much useful information as possible.

Section 2 – Measuring spread

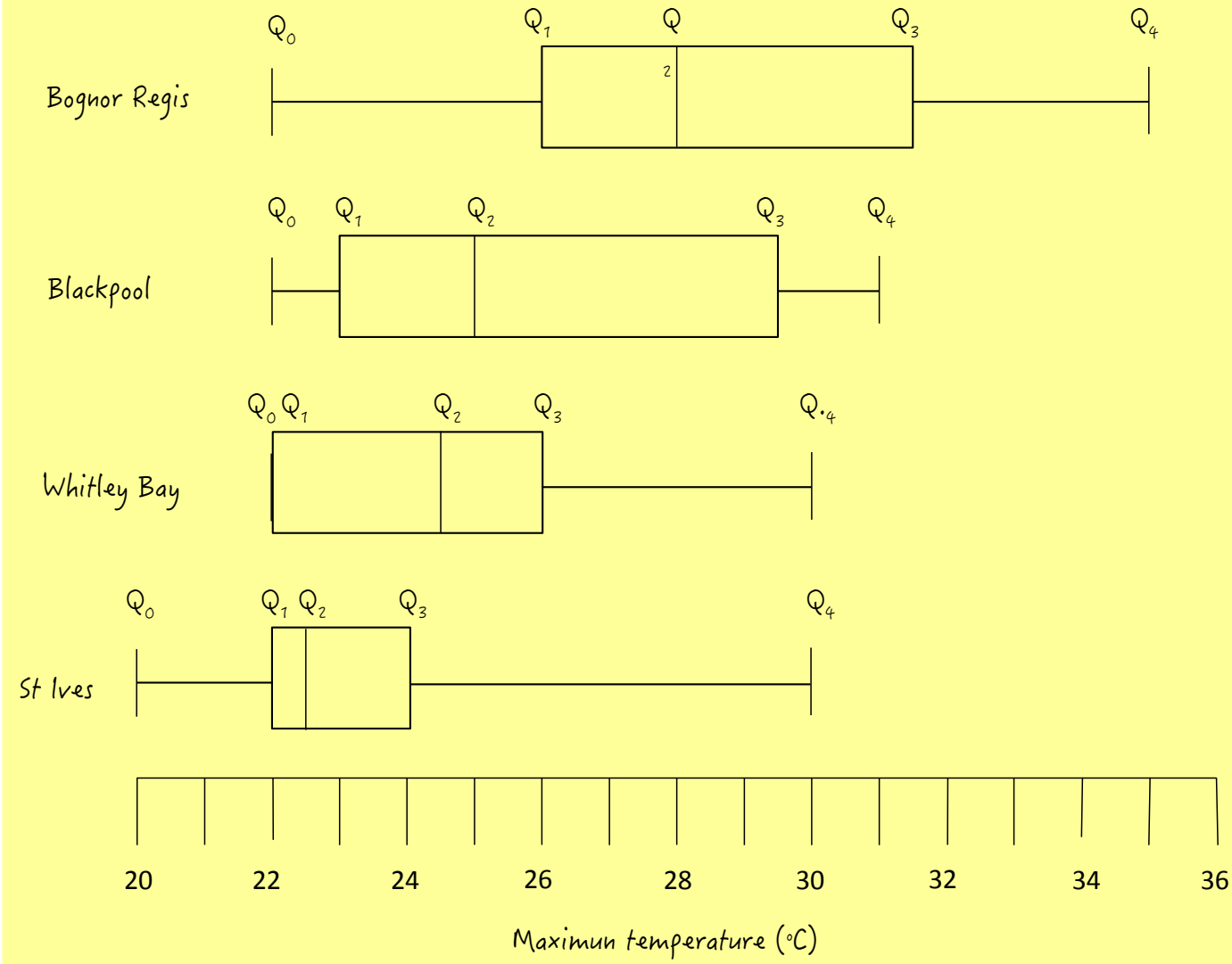


Meera

These are our boxplots.



Sarah





Mikka

For more than 25% of the time Bognor Regis is warmewr than anywhere else

And for 25% of the time, St Ives was colder than anywhere else



Jasmine



Evie

The interquartile range is smallest for St Ives which means ....

The interquartile range is largest for Blackpool which means ....



Charo

Section 2 – Measuring spread

Fitness check



1. Can you explain what the interquartile range actually measures?
2. Can you explain the significance of the interquartile range being smallest for St. Ives and largest for Blackpool?

3. Is it possible to have a large range and a small interquartile range?
   If so, what would it show about the data?

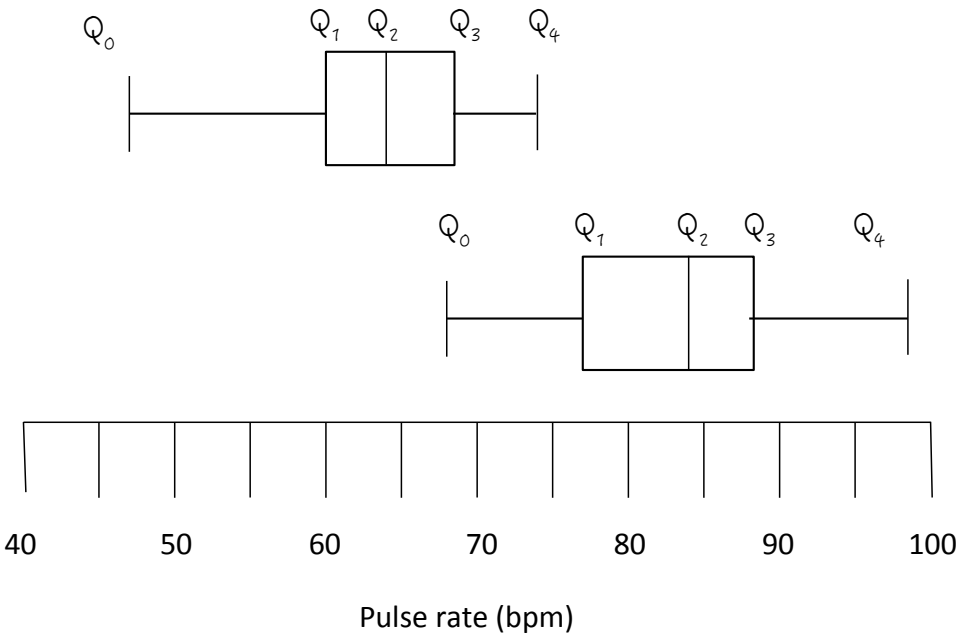
4. Two dog breeders are comparing the number of Newfoundland puppies in the litters they have bred over the past five years.

Breeder 1	0	5	6	6	8	4	6	7	7	6
Breeder 2	8	5	7	5	8	5	7	7	10	0

   a) For each breeder find the five key number summary and draw comparative boxplots of the two sets of data.

   b) From your results, decide which breeder is more successful and give your reasons.

5. These two boxplots show the pulse rates of 12 athletes before and 5 minutes after a race.



Explain what these boxplots show you?

Statistics 4b Mastery

1. 8Z have been investigating whether, "Boys have a better memory than girls". The results are shown in Fig. 1.

   What is the modal score for the boys?

A	4
B	7
C	14
D	17

Do boys have a better memory than girls?

Boys			Girls
(2)	4 4	0 | 4	(1)
(3)	9 6 6	0 | 6 9	(2)
(4)	4 4 3 2	1 | 2 2 2 3 4	(5)
(6)	9 8 7 7 7 5	1 | 5 5 6 7 8	(5)
(0)		2 | 0	(1)
(15)	Count		Count (14)

Key 1 \| 2 represents 12 points

Fig. 1

2. Using the results from Fig. 1, what is the median score for the girls?

A	B	C	D
13	13.5	14	15

3. Using the results from Fig. 1, what is the median score for the boys?

A	B	C	D
13	13.5	14	15

4. Using the results from Fig. 1, what is the range for the girls' scores?

A	B	C	D
16	8	20	4

5. Using the results from Fig. 1, what is the mean score for the boys to 1dp.?

A	B	C	D
5.7	12.3	4.5	13.1

6. Using the results from Fig. 1, what is the mean score for the girls to 1dp.?

A	B	C	D
5.7	12.3	4.5	13.1

7. The frequency table shows the number of goals scored by the Year 8 soccer team this season. What was the modal number of goals scored?

Goals scored	Frequency
0	3
1	5
2	3
3	7
4	2

A	B	C	D
1	2	3	4

8. Using the same data as in Q7, what was the mean number of goals scored per game?

A	B	C	D
1	2	3	4

9. Using the same data as in Q7, what was the median number of goals scored per game?

A	B	C	D
1	2	3	4

9. Using the same data as in Q7, what was the range in the number of goals scored per game?

A	B	C	D
0 - 4	4	5	7

10. This table shows a summary of the spots counted on 25 ladybirds.

Number of spots	Frequency
5	3
6	8
7	9
8	5

Calculate the mean number of spots per ladybird.

A	B	C	D
1	41.5	1.04	6.64

11. This frequency table summaries hourly rate of workers at a local holiday camp.

    How many people earned less than £9.00 per hour?

A	B	C	D
3	10	30	37

Hourly rate	Frequency
£5.50	4
£7.50	14
£8.00	12
£9.00	7
£10.00	3

12. The students in 9Z all participated in a touch typing course to improve their IT skills. They then were assessed by how many words they could type in a minute.
    Their results are shown in this table.

    In which group would the median result lie?

A	B	C	D
6 - 10	11 - 15	16 - 20	21 - 25

Words	Frequency
6 - 10	3
11 - 15	11
16 - 20	12
21 - 25	1
Total	27

13. Using the results from Q12, which is the modal group?

A	B	C	D
6 - 10	11 - 15	16 - 20	21 - 25

14. Using the results from Q12, what is their estimated mean score?

A	B	C	D
15	16	18	27

15. The basketball team have been having some extra shooting practice. Each player had 15 shots and their coach recorded the number of "baskets" they scored and then ranked them in ascending order.

    His notebook is shown in Fig. 2.

    What was the lower quartile of the scores?

Shooting Practice 21/02/18

3, 4, 6, 7, 8, 8, 9, 9, 9, 10,

10, 11, 12, 12, 13, 14, 15, 15

18 players

Ted

Fig. 2

A	B	C	D
7	8	9	10

16. Using the results from Fig. 2, what is the upper quartile of the scores?

A	B	C	D
13	13.5	14	15

17. Using the results from Fig. 2, what is the IQR of the scores?

A	B	C	D
4	5	7	15

In questions 18 – 20, you must show all of your working.

18. The temperature at 2am in February 2018 in Barcelona was measured and recorded. The results are shown in this frequency table.

Temperature °C	0	1	2	3	4	5	6	7	8	9
Number of nights	0	1	4	3	3	3	3	5	4	2

a) Work out the range in the temperatures

b) Calculate the mean temperature.

   Give your answer correct to 1 decimal place.

19. Books for younger children generally have shorter sentences. 50 sentences were taken from each of two book and the number of words in each sentence counted. The results are shown below.

Words	Frequency
1 - 5	12
6 - 10	13
11 - 16	14
17 - 20	5
21 - 25	6
Total	50

Words	Frequency
1 - 5	11
6 - 10	18
11 - 16	15
17 - 20	9
21 - 25	7
Total	50

On average, which book contains the longer sentences? Make sure you use statistical measures to support your decision.

20. Two groups of students were given a short mental arithmetic test. Group A took the test in silence and Group B took the test whist background music played. Their results are shown below.

Group A: 6, 6, 7, 7, 7, 7, 7, 8, 8, 9, 9, 10, 10, 12

Group B: 1, 5, 6, 7, 8, 8, 8, 8, 8, 9, 9, 9, 10, 10, 12

a) Find the 5 key number summary for both sets of data.
b) Find the range of scores for each group.
c) Find the IQR (Inter quartile range) for each group.
d) In your opinion which of these figures best represents the spread of the data? Explain your thinking.
e) Draw two comparative boxplotrs of your results.
f) What conclusions can you draw by looking at these boxplots?