

Student Name

Mathematics Stage 5 Core

DATA ANALYSIS A

Book 3 Represent datasets using box plots
and use them to compare datasets

Version: 260207

Report mistakes & suggest improvements:
<https://MrDingMaths.com>

CONTENTS

Overview	2
Box Plots	5
Box Plot Percentages	9
Comparing Datasets	12
Five-Number Summary from Stem & Leaf Plots	19
Five Number Summary from Dot Plots	21
Five Number Summary from Cumulative Histogram	23
Skewness and Box Plots	28
Check Your Understanding	35

OVERVIEW

SYLLABUS CONTENT

MA5-DAT-C-01 compares and analyses datasets using summary statistics and graphical representations

Represent datasets using box plots and use them to compare datasets

- Represent numerical datasets using a box plot to display the median, upper and lower quartiles, and maximum and minimum values
- Compare 2 or more numerical datasets using parallel box plots drawn on the same scale
- Compare and contrast the centres, spreads and shapes of 2 or more numerical datasets, using box plots and numerical statistics, including the 5-number summary
- Determine quartiles from datasets displayed in histograms and dot plots, and represent these as a box plot
- Identify and describe skewness or symmetry of datasets displayed in histograms, dot plots and box plots
- Interpret box plots to draw conclusions and make inferences about the dataset

REVIEW OF PRIOR KNOWLEDGE

1 Calculate the five-number summary and IQR of these datasets.

a 3 18 23 23 27 29 30 34

b 8 2 3 4 2 11 9 12 14 7

A quartile splits the dataset into parts.

2 What percentage of scores are below the third quartile?

.....

3 What percentage of scores does the IQR represent?

4

a Find 25% of 200.

b Find 50% of 470.

c Find 75% of 120.

d Find 25% of 164.

e Find 50% of 96.

f Find 75% of 360.

5 Determine the median from these charts.

a

Tens	Units		
0	0	7	7
1	2	5	9
2	2	3	7
3	1		
4	2	6	

b



c

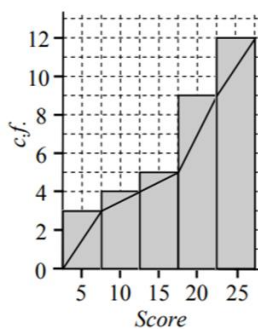
Score	f	$c.f.$
0	2	2
2	3	5
4	1	
6	6	
8	5	
10	2	
$\Sigma f =$		

check

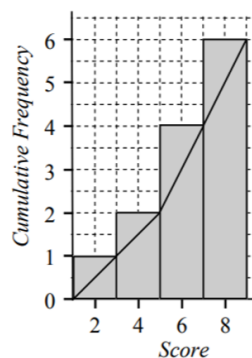
d

x	f	$c.f.$
1	1	
2	4	
3	12	
4	9	
5	12	
6	9	
$\Sigma f =$		

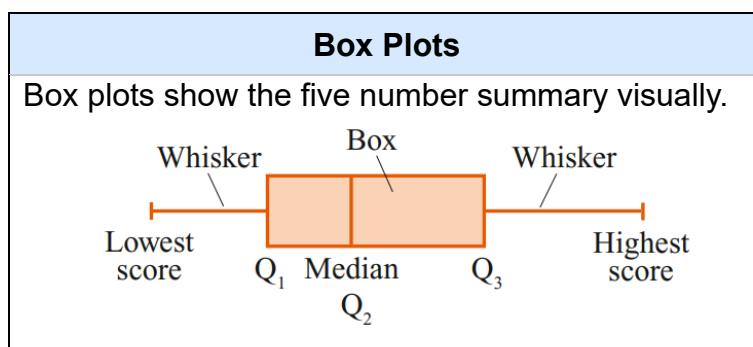
e This is a cumulative frequency histogram



f This is a cumulative frequency histogram

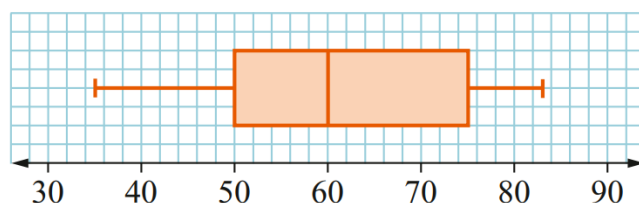


BOX PLOTS



Example Interpret a box plot.

A dataset of student test scores is presented as a box plot. Determine the five-number summary.



Min:	Q_1 :	Q_2 :	Q_3 :	Max:
35	50	60	75	83

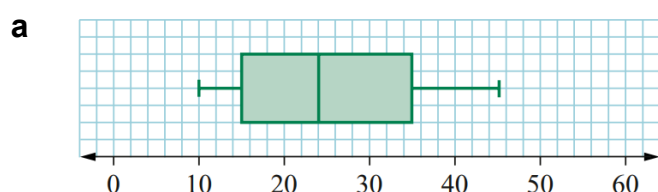
What is the median?

The median is, shown on the box plot as the line in the middle of the

What is the interquartile range?

$IQR = \dots - \dots = \dots$

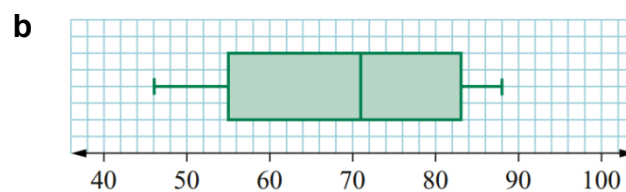
Interpret these box plots.



Min:	Q_1 :	Q_2 :	Q_3 :	Max:
------	---------	---------	---------	------

$IQR = \dots$

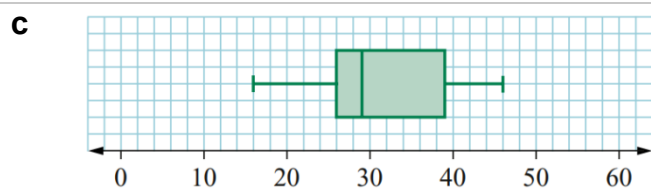
FNS: 10, 15, 24, 35, 45 | $IQR = 20$



Min:	Q_1 :	Q_2 :	Q_3 :	Max:
------	---------	---------	---------	------

$IQR = \dots$

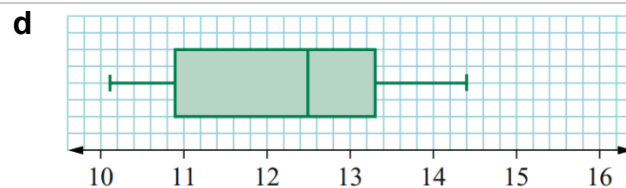
FNS: 46, 55, 71, 83, 88 | $IQR = 28$



Min:	Q_1 :	Q_2 :	Q_3 :	Max:
------	---------	---------	---------	------

$IQR = \dots$

FNS: 16, 26, 29, 39, 56 | $IQR = 13$

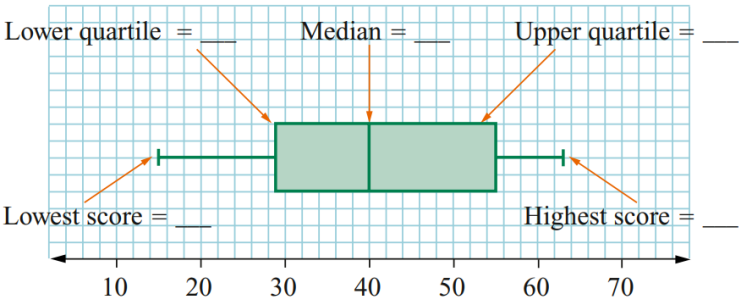


Min:	Q_1 :	Q_2 :	Q_3 :	Max:
------	---------	---------	---------	------

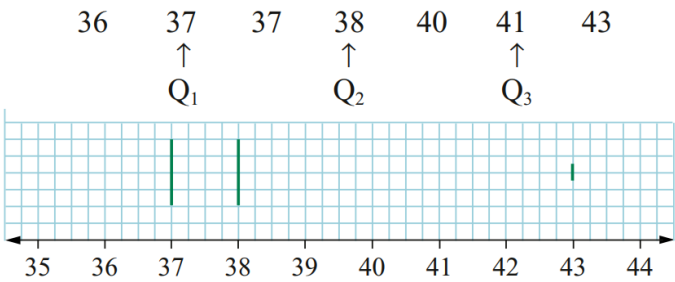
$IQR = \dots$

FNS: 10.1, 10.9, 12.5, 13.3, 14.4 | $IQR = 2.4$

1 Interpret this box plot.

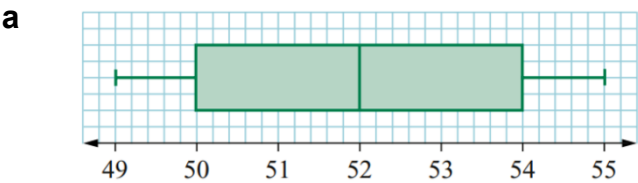


2 Complete the following then draw the box plot.



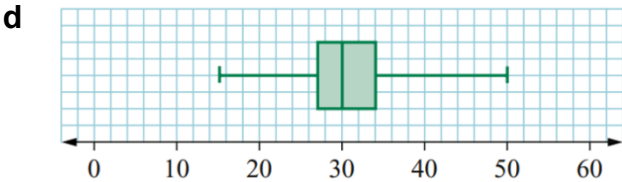
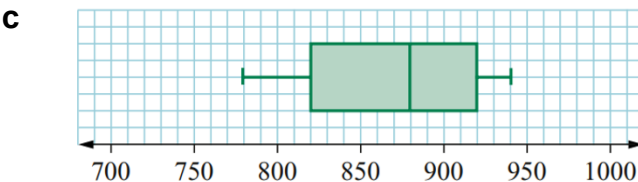
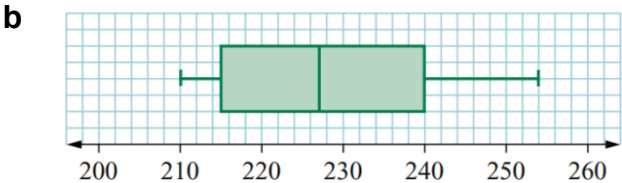
Median = 40
Upper quartile = 41
Lower quartile = 37
36 score = 36
43 score = 43

3 Interpret these box plots.

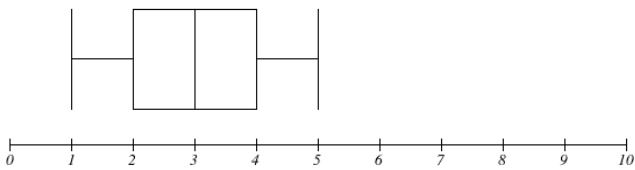
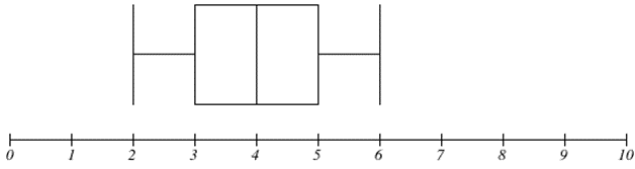
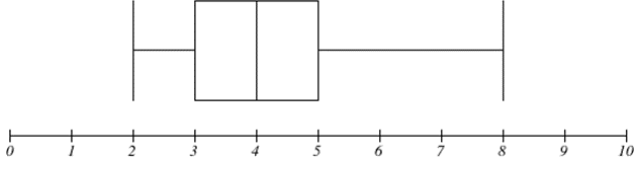
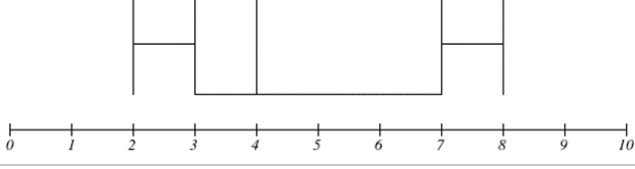
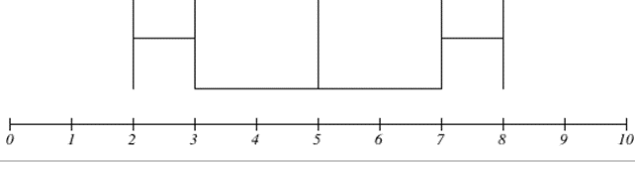
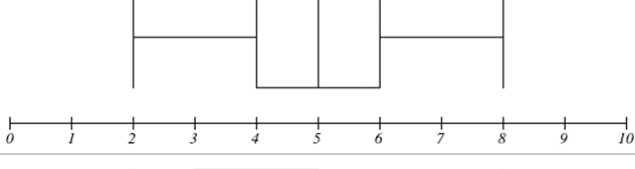
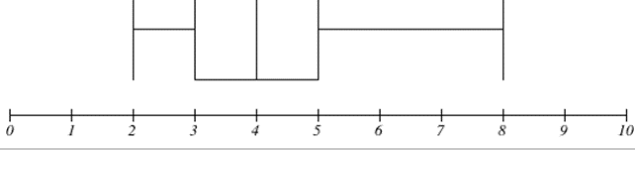


Min: Q_1 : Q_2 : Q_3 : Max:

IQR =



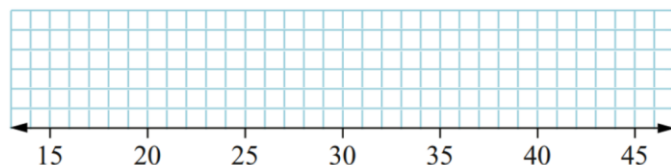
4 Fill in the table based on the diagram.

	Median	Range	IQR
a 			
b 			
c 			
d 			
e 			
f 			
g 			
h 			

5 Draw box plots for these five number summaries.

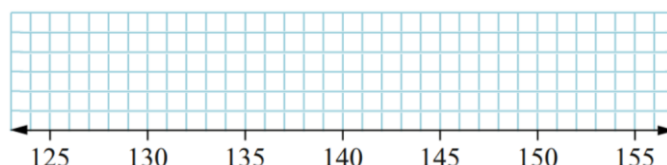
a

Min	Q_1	Q_2	Q_3	Max
15	23	28	32	40



b

Min	Q_1	Q_2	Q_3	Max
130	139	141	148	153



c

Min	Q_1	Q_2	Q_3	Max
6	7	10	18	28



d

Min	Q_1	Q_2	Q_3	Max
71	73	78	80	83



6 State the five number summary of these datasets and draw the box plot.

a 34 35 36 36 37 38 39 39 39 40

Min: Q_1 : Q_2 : Q_3 : Max:



b 4 5 8 8 10 12 12 14 15 19

Min: Q_1 : Q_2 : Q_3 : Max:



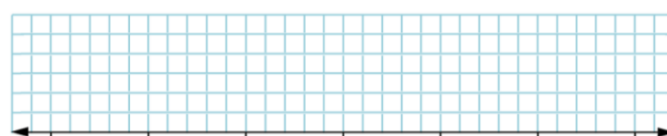
c 21 21 23 24 24 24 26 28 30

Min: Q_1 : Q_2 : Q_3 : Max:

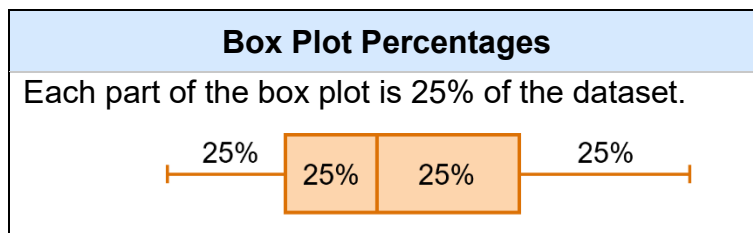


d 89 90 92 95 95 98 102 103

Min: Q_1 : Q_2 : Q_3 : Max:

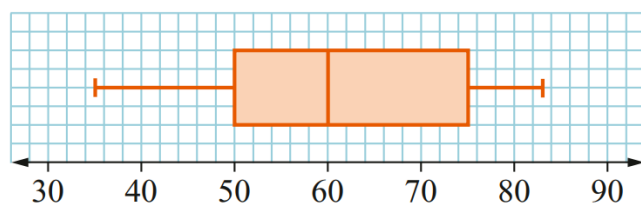


BOX PLOT PERCENTAGES



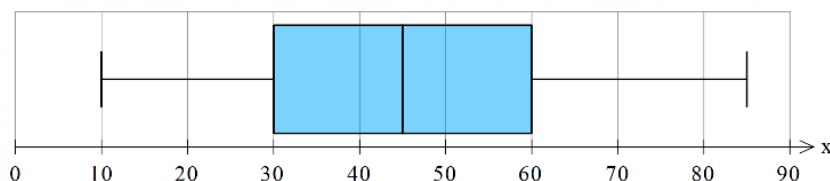
Example Interpret a box plot.

200 students took a test.
Their scores are presented as a box plot.



- a** Between which two scores is the middle 50% of data?
50 and 75
- b** What proportion of people scored greater than 50?
75%
- c** How many people scored greater than 50?
 $200 \times 0.75 = 150$

The box plot shows how long 500 girls spent watching TV last week.



- a** Between which two scores is the middle 50% of data?
- b** What percentage watched between 60 and 85 minutes?
- c** What percentage watched more than 30 minutes?
- d** How many girls watched less than 60 minutes?
- e** How many girls watched more than 45 minutes?
- f** How many girls watched between 30 and 85 minutes?

30 and 60

25%

75%

375

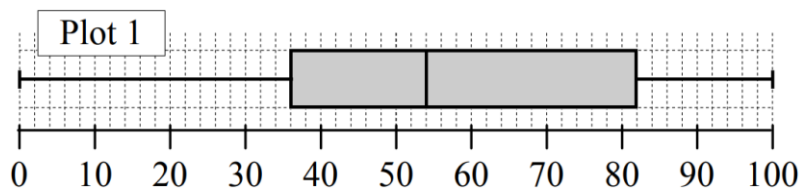
250

375

Key ideas

1. A box plot visually represents the summary of a dataset.
2. Each part of the box plot represents% of the data.
3. The two boxes in the middle represent% of the data.

1 Answer these questions using the box and whisker diagram.



- a The lowest 25% is between 0 and
- b The lower-middle 25% is between and
- c The upper-middle 25% is between and
- d The upper 25% is between and
- e The middle 50% is between and
- f The lower 50% is between and
- g The upper 50% is between and

What percentage of scores lie between:

- | | | | |
|--------------|---------------|-------------|--------------|
| h 54 and 82? | i 36 and 100? | j 0 and 54? | k 36 and 82? |
|--------------|---------------|-------------|--------------|

If the box plot represents 86 000 people, how many are between:

- | | | | |
|-------------|--------------|-------------|--------------|
| l 0 and 36? | m 36 and 82? | n 0 and 82? | o 54 and 82? |
|-------------|--------------|-------------|--------------|

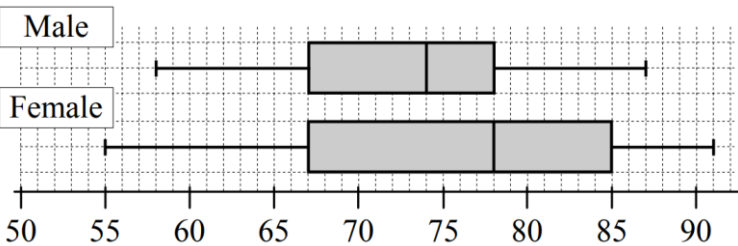
2 This box plot represents people living in a retirement complex.

- a What percentage of male retirees are aged between:

- i 58 and 67?
- ii 74 and 87?

- b What percentage of female retirees are aged between:

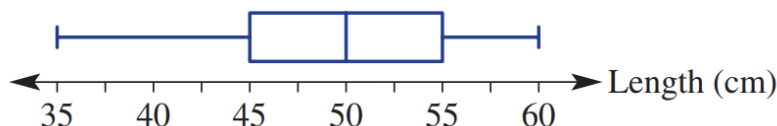
- i 55 and 85?
- ii 67 and 78?



- c If there are 432 males and 584 females, how many:

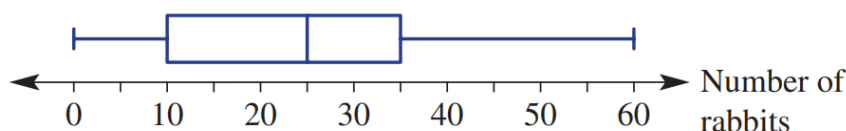
- i Females are aged between 67 and 85?
- ii Males are aged between 67 and 78?
- iii Females are aged between 67 and 78?
- iv Retirees in total aged between 67 and 78?

- 3 This box plot summarises the length of babies born in a particular week in a hospital.



- State the minimum and maximum baby lengths.
- Find the range of the length of the babies.
- State the median baby length.
- Find the interquartile range.
- 50% of the baby lengths are below what amount?
- 25% of the babies are born longer than what amount?
- If there were 80 babies born in the week,
how many would be expected to be less than 45 cm in length?

- 4 This box plot show the number of rabbits spotted per day in a paddock over a 100-day period.

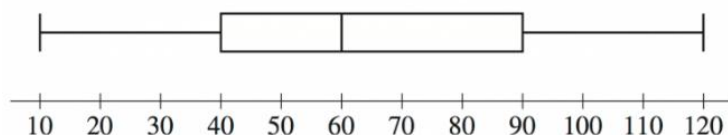


- State the minimum and maximum rabbits spotted.
- Find the range of the number of rabbits spotted.
- State the median number of rabbits spotted.
- Find the interquartile range.
- 75% of the days the number of rabbits spotted is below what amount?
.....
- 50% of the days the number of rabbits spotted is more than what amount?
.....
- How many days was the number of spotted rabbits less than 10?
.....

5 **2015 HSC Standard 2 Band 4**

The times, in minutes, that a large group of students spend on exercise per day are presented in the box-and-whisker plot. What percentage of these students spend between 40 minutes and 60 minutes per day on exercise?

- 17%
- 20%
- 25%
- 50%



COMPARING DATASETS

- **Box plots** placed side by side allow for quick **comparison** of two **datasets**.
- A **smaller IQR** means **more consistent**.

Example Compare datasets using parallel box plots.

Two datasets are shown.

Describe the similarities.

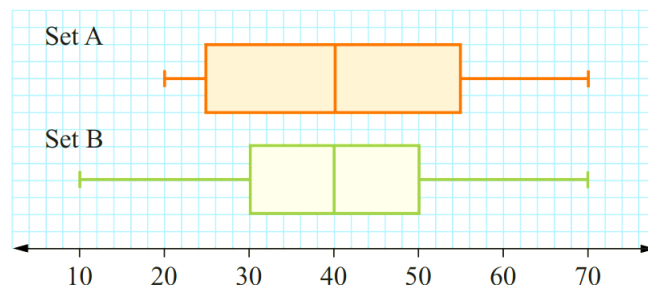
Same median and same maximum value.

Which dataset has a larger range?

Set B has a larger range

Which dataset has a larger interquartile range?

Set A has the larger IQR



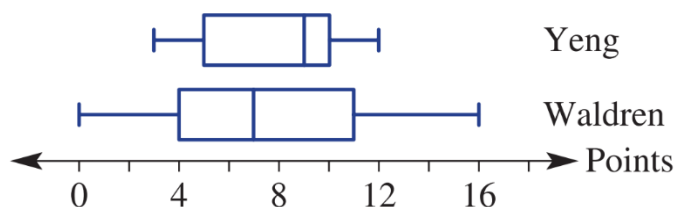
Without doing any calculations, how do we determine which box plot has the larger range?

.....

Without doing any calculations, how do we determine which box plot has the larger IQR?

.....

The number of points per game for two basketball players over a season is summarised in these box plots.



- Which player has the higher maximum?
- Which player has the higher median?
- Which player has the smaller IQR?
- Which player is a more consistent basketball scorer? Give reasons.

Waldren

Yeng

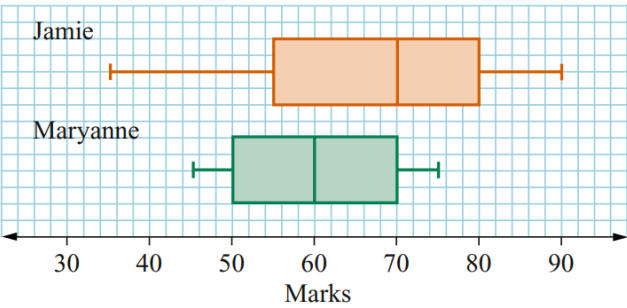
Yeng

Yeng, as he has a smaller range and IQR.

Key ideas

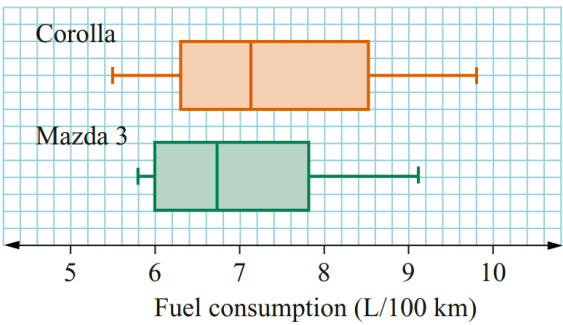
- The dataset with the higher “average” has a higher
- The dataset that is more consistent has a lower

1 A class has eight assessment tasks over a year. The parallel box plots show a summary of the marks for the assessments for two students, Jamie and Maryanne.



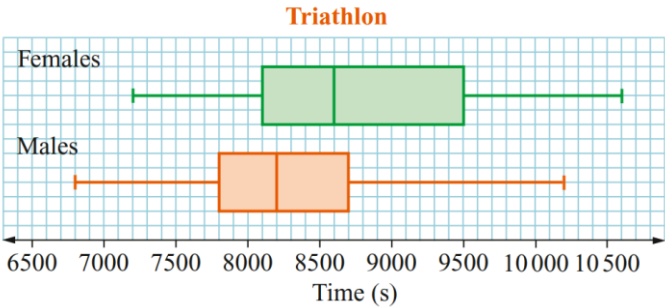
- a Who scored the highest mark?
- b Who scored the lowest mark?
- c What was the range of marks for each student?
- d Who had the greater spread of marks?
- e What was the interquartile range for each student?
- f Whose marks were the more consistent?
- g Who had more marks over 70?

2 A new Toyota Corolla and Mazda 3 were each taken for ten test drives over the same routes. These parallel box plots show a summary of the fuel consumption, in L/100 km, of each vehicle



- a Which car recorded the highest fuel consumption?
- b Which car recorded the lowest fuel consumption?
- c Which car had the greater range of results?
- d Which car demonstrated the more consistent fuel consumption over all routes? Give a reason.
.....
- e Which car used less than 7 L/100 km more often?
.....

3 This parallel box plot shows the times (in seconds) male and female competitors take to complete a triathlon.

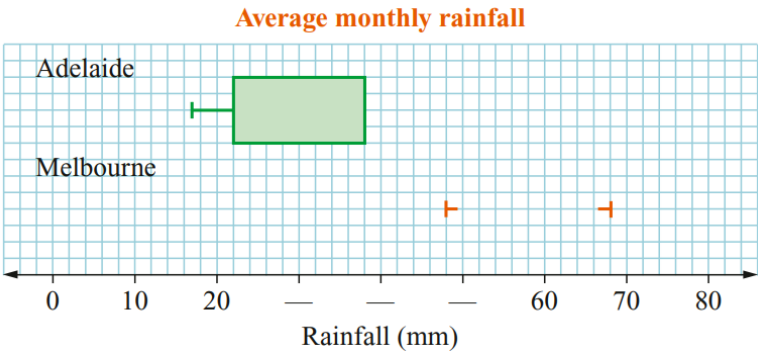


- a What was the fastest time for the event?
- b What was the difference in time between the slowest male and slowest female competitor?
.....
- c How long did it take the first half of the male competitors to finish?
- d How long did it take the top 75% of the males to finish?
- e How long did it take the first 25% of the females to finish?
- f Which group had the greatest range of times?
- g Which group had the smallest interquartile range of times?
- h Which group had the fastest overall times?

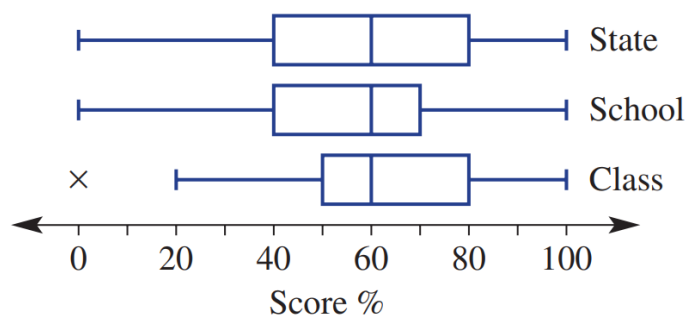
4 The average monthly rainfall, in millimetres, for Adelaide and Melbourne is shown in this table.

Month	J	F	M	A	M	J	J	A	S	O	N	D
Adelaide	17	19	22	38	59	52	65	50	43	38	24	23
Melbourne	48	48	52	58	58	50	49	51	59	68	60	59

- a Order the data in ascending order:
Adelaide:
Melbourne:
- b Calculate the five number summary for each city:
Adelaide:
Melbourne:
- c Complete the two box plots. Make a comment comparing the rainfall of the two cities.



- 5 The results for a Year 12 class are to be compared with the Year 12 results of the school and the State, using the parallel box plots shown.



- a Describe the main differences between the performance of:

i the class against the school

.....

ii the class against the State

.....

iii the school against the State

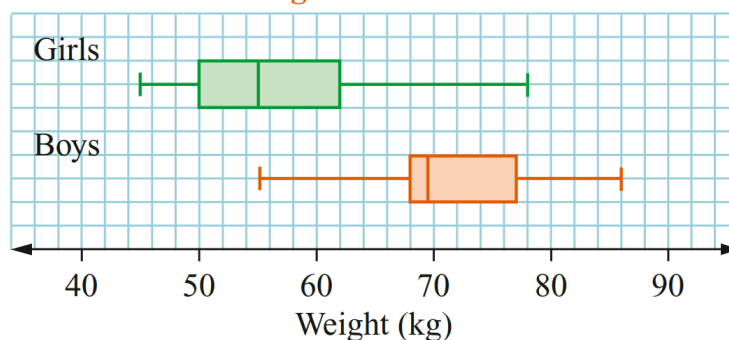
.....

- b The X represents an outlier. Explain how someone could be an outlier in a class but not in a school.

.....

- 6 Write as many things as you can determine from these box plots.

Weights of students

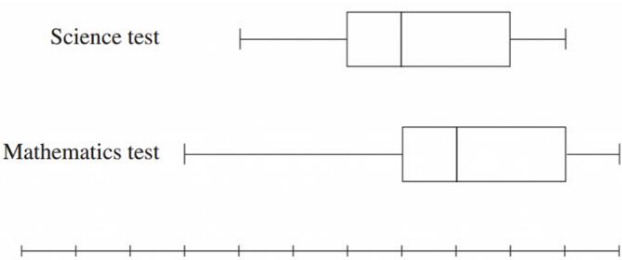


.....

7 2008 HSC Standard 2 Band 4

The marks for a Science test and a Mathematics test are presented in box-and-whisker plots. Which measure must be the same for both tests?

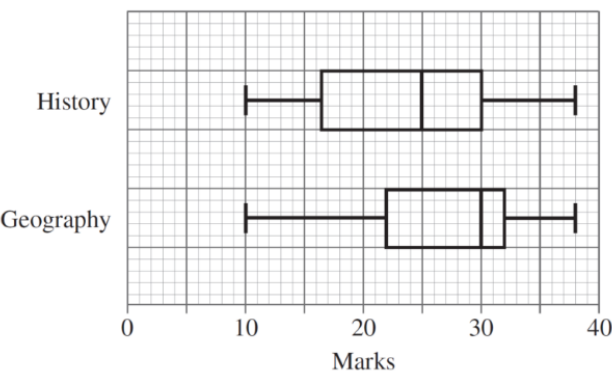
- (A) Mean
- (B) Range
- (C) Median
- (D) Interquartile range



8 2016 HSC Standard 2 Band 4

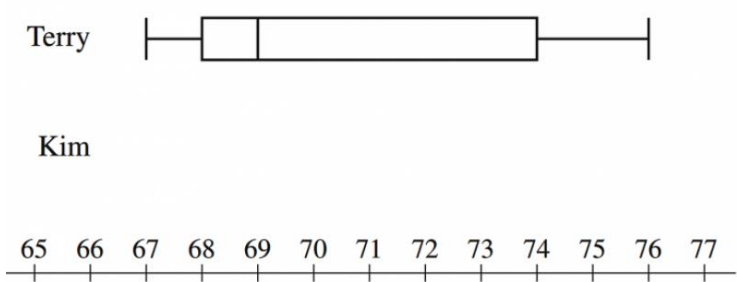
The box-and-whisker plots show the results of a History test and a Geography test. In History, 112 students completed the test. The number of students who scored above 30 marks was the same for the History test and the Geography test. How many students completed the Geography test?

- (A) 8
- (B) 50
- (C) 56
- (D) 112



9 2014 HSC Standard 2 Band 5

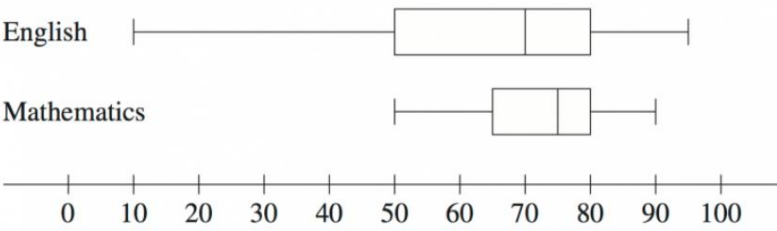
Terry and Kim each sat twenty class tests. Terry's results on the tests are displayed in the box-and-whisker plot shown.



- a Kim's 5-number summary for the tests is 67, 69, 71, 73, 75. Draw a box-and-whisker plot to display Kim's results below that of Terry's results.
- b What percentage of Terry's results were below 69?
.....
- c Terry claims that his results were better than Kim's. Is he correct? Justify your answer by referring to the summary statistics and the skewness of the distributions.
.....
.....
.....

10 2012 HSC Standard 2 Band 5

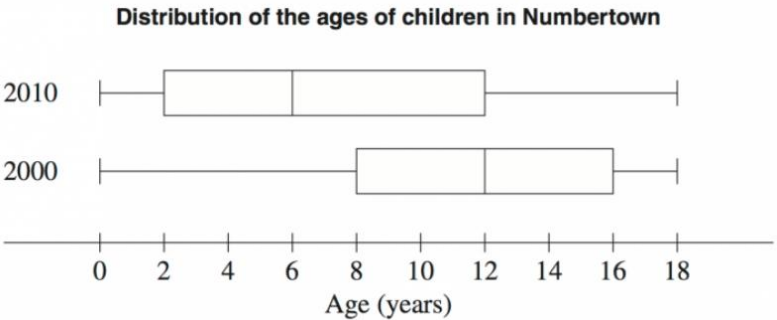
The test results in English and Mathematics for a class were recorded and displayed in the box-and-whisker plots.



- a What is the interquartile range for English?
- b Compare and contrast the two data sets by referring to the skewness of the distributions and the measures of location and spread.
.....
.....
.....

11 2010 HSC Standard 2 Band 5

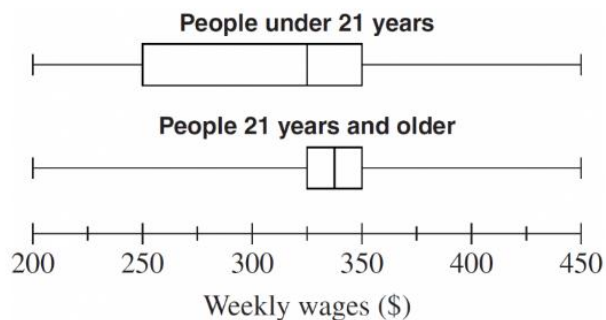
The graphs show the distribution of the ages of children in Numbertown in 2000 and 2010.



- a In 2000 there were 1750 children aged 0–18 years.
How many children were aged 12–18 years in 2000?
.....
- b The number of children aged 12–18 years is the same in both 2000 and 2010.
How many children aged 0–18 years are there in 2010?
.....
- c Identify TWO changes in the distribution of ages between 2000 and 2010.
In your answer, refer to measures of location or spread or the shape of the distributions.
.....
.....
.....
- d What would be ONE possible implication for government planning, because of this change in the distribution of ages?
.....
.....

12 2005 HSC Standard 2 Band 6

Two groups of people were surveyed about their weekly wages. The results are shown in the box-and-whisker plots.



Which of the following statements is true for the people surveyed?

- (A) The same percentage of people in each group earned more than \$325 per week.
- (B) Approximately 75% of people under 21 years earned less than \$350 per week.
- (C) Approximately 75% of people 21 years and older earned more than \$350 per week.
- (D) Approximately 50% of people in each group earned between \$325 and \$350 per week.

FIVE-NUMBER SUMMARY FROM STEM & LEAF PLOTS

Example Determine the five-number summary from a stem-and-leaf plot.

Construct a box plot from this stem and leaf plot.

Stem	Leaf
4	2 4 6
5	0 3 7 7 9 9
6	1 3 4 5 9
7	0 2 4 7
8	1 5 5 6 7

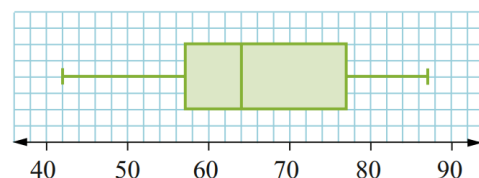
Step 1: Find Q2

Stem	Leaf
4	2 4 6
5	0 3 7 7 9 9
6	1 3 4 5 9
7	0 2 4 7
8	1 5 5 6 7

Step 2: Find Q1 and Q3

Stem	Leaf
4	2 4 6
5	0 3 7 7 9 9
6	1 3 4 5 9
7	0 2 4 7
8	1 5 5 6 7

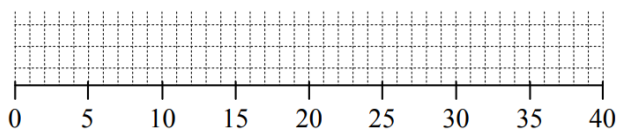
Step 3: Draw box plot.



Draw a box plot from this stem-and-leaf plot.

a

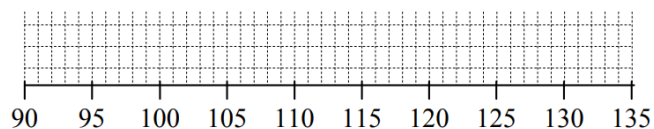
Tens	Units	Min =
0	1 4 9 9 9	Q ₁ =
1	3 8 8 9	Q ₂ =
2	5 9	Q ₃ =
3	2 5	Max =



FNS: 1, 9, 18, 27, 35

b

Stem	Leaf	Min =
9	1 8	Q ₁ =
10	2 6 8	Q ₂ =
11	4 7 8 9	Q ₃ =
12	1 8 8	Max =
13	0 2 2	



FNS: 91, 106, 118, 128, 132

1 Determine the five-number summary from these stem-and-leaf plots.

a

Tens	Units
0	4 4 8
1	3 5
2	5 7 7

Min:
 Q_1 :
 Q_2 :
 Q_3 :
Max:

b

Tens	Units
0	1 7 7
1	2 6
2	0 3 9
3	1

Min:
 Q_1 :
 Q_2 :
 Q_3 :
Max:

c

Stem	Leaf
10	0 4 4
11	6 6 8 9 9
12	1 1 6 6 7

Min:
 Q_1 :
 Q_2 :
 Q_3 :
Max:

</

2 Determine the five-number summary and draw a box plot of the data.

a

Tens	Units
0	1 4 9 9 9
1	3 8 8 9
2	5 9
3	2 5

Min: Q_1 : Q_2 : Q_3 : Max:

b

Tens	Units
0	8 8
1	1 1 2 4 4
2	0 0 1 1 4 6 9

Min: Q_1 : Q_2 : Q_3 : Max:

c

Stem	Leaf
1	7 7 9
2	0 1 1 4
3	2 5 6 6 7
4	1 3 3 8
5	4 5

Min: Q_1 : Q_2 : Q_3 : Max:

d

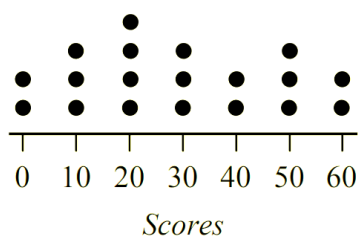
Stem	Leaf
4	2
5	4 6
6	3 3 7
7	1 8 8 9
8	5 6

Min: Q_1 : Q_2 : Q_3 : Max:

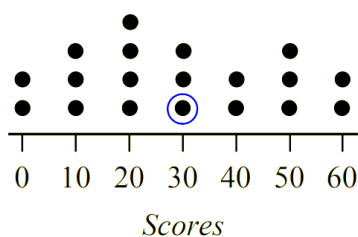
FIVE NUMBER SUMMARY FROM DOT PLOTS

Example Determine the five-number summary from a dot plot.

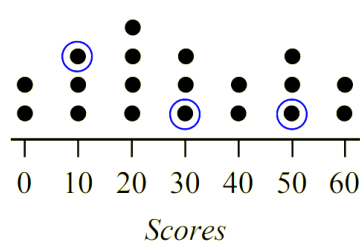
Construct a box plot from this dot plot.



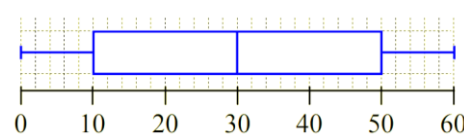
Step 1: Find Q2



Step 2: Find Q1 and Q3

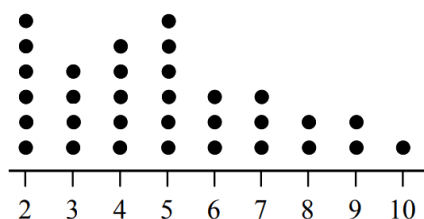


Step 3: Draw box plot.

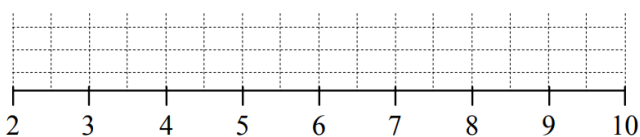


Draw a box plot from this dot plot.

a

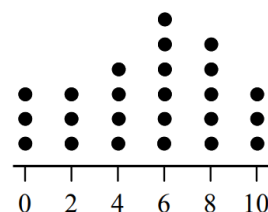


Min: Q_1 : Q_2 : Q_3 : Max:

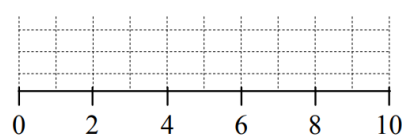


FNS: 2, 3, 5, 6.5, 10

b

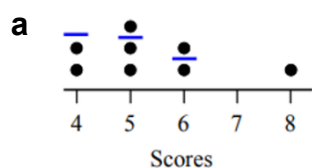


Min: Q_1 : Q_2 : Q_3 : Max:

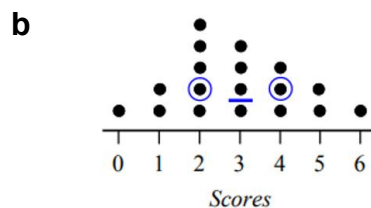


FNS: 0, 3, 6, 8, 10

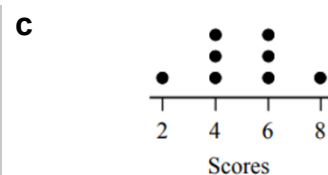
1 Determine the five-number summary from these dot plots.



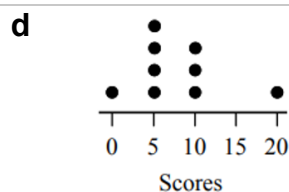
Min: Q_1 : Q_2 : Q_3 : Max:



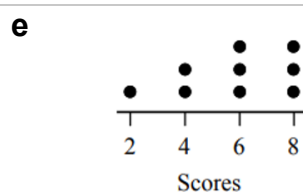
Min: Q_1 : Q_2 : Q_3 : Max:



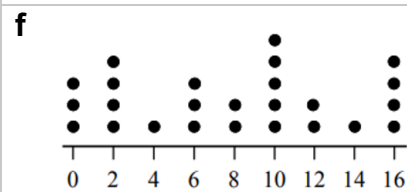
Min: Q_1 : Q_2 : Q_3 : Max:



Min: Q_1 : Q_2 : Q_3 : Max:

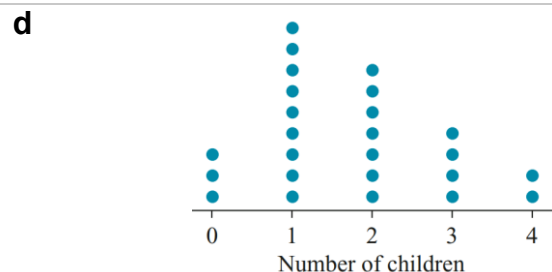
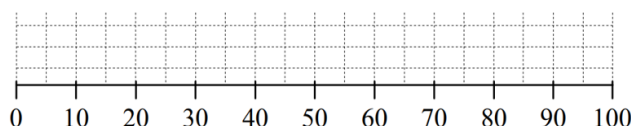
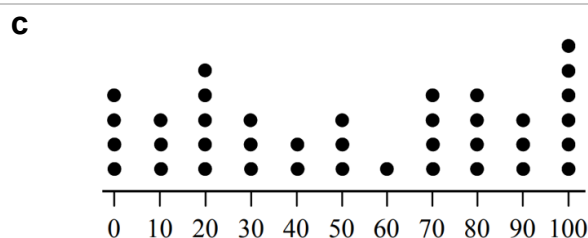
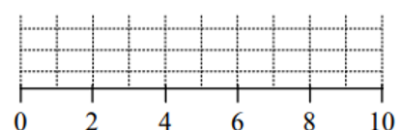
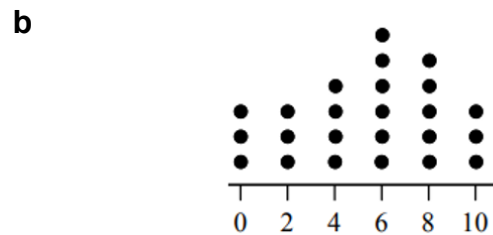
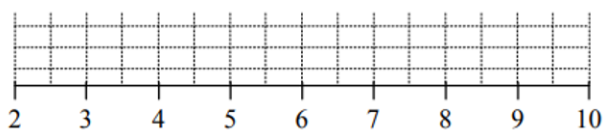
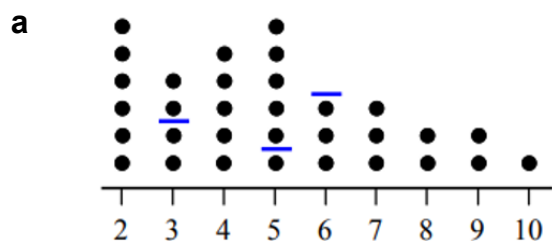


Min: Q_1 : Q_2 : Q_3 : Max:



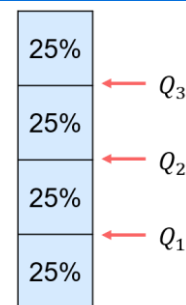
Min: Q_1 : Q_2 : Q_3 : Max:

2 Determine the five-number summary and draw a box plot of the data.



FIVE NUMBER SUMMARY FROM CUMULATIVE HISTOGRAM

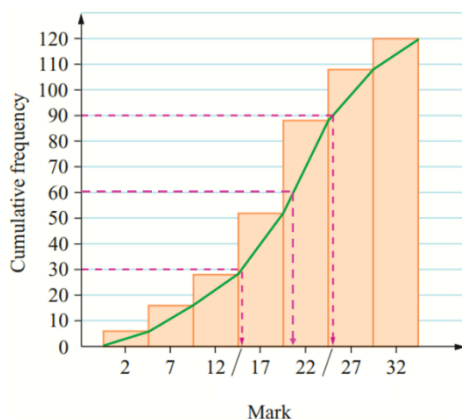
1. Identify the total number of scores. This is the height of the last column.
2. Find Q_1 position: $0.25 \times \text{total}$
3. Find Q_2 position: $0.5 \times \text{total}$
4. Find Q_3 position: $0.75 \times \text{total}$
5. Draw horizontal lines from the vertical axis to the ogive, then estimate the x values.



Example Estimate quartiles from a cumulative histogram.

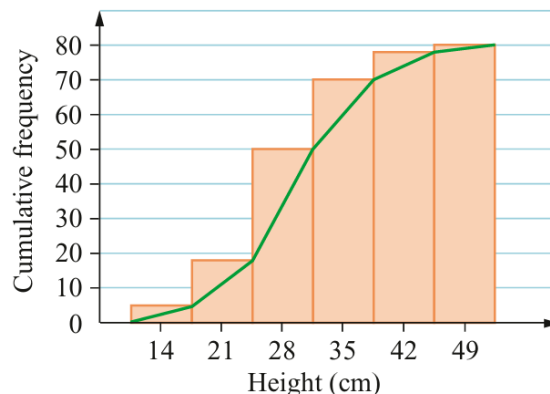
Determine the five number summary from this cumulative histogram.

1. Total scores: 120
2. Q_1 position = $0.25 \times 120 = 30$
3. Q_2 position = $0.5 \times 120 = 60$
4. Q_3 position = $0.5 \times 120 = 90$
- 5.



Min	Q_1	Q_2	Q_3	Max
2	15	21	25	32

1. Total scores:
2. Q_1 position =
3. Q_2 position =
4. Q_3 position =
- 5.

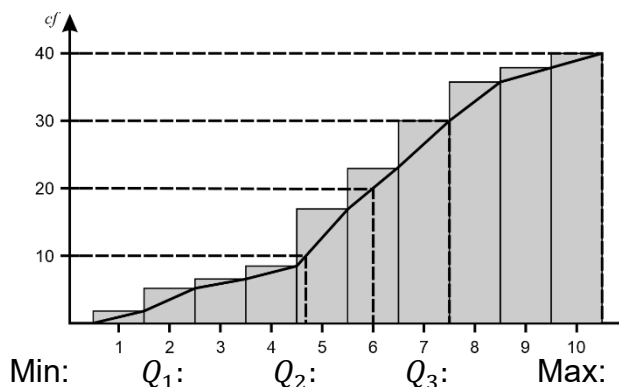


Min	Q_1	Q_2	Q_3	Max

FNS: 14, 25, 30, 35, 49

Determine the five number summary from this cumulative histogram.

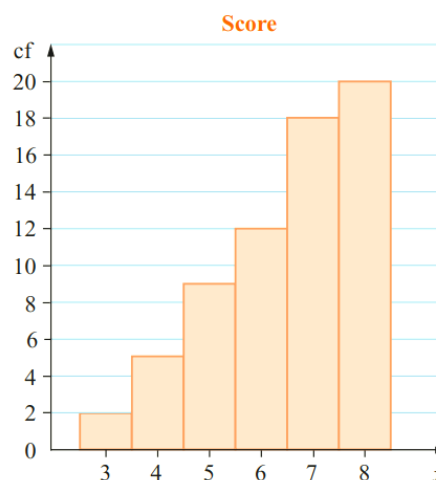
a



Min: Q_1 : Q_2 : Q_3 : Max:

FNS: 1, 4.6, 6, 7.5, 10

b

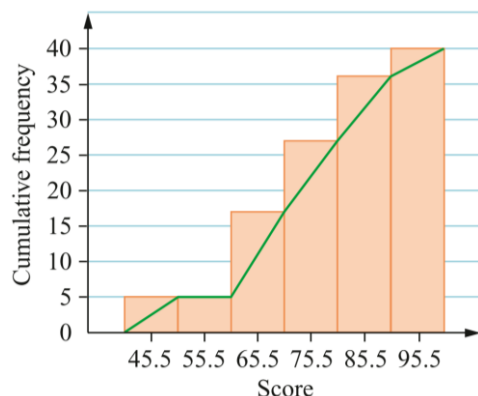


Min: Q_1 : Q_2 : Q_3 : Max:

FNS: 3, 4.5, 5.9, 7, 8

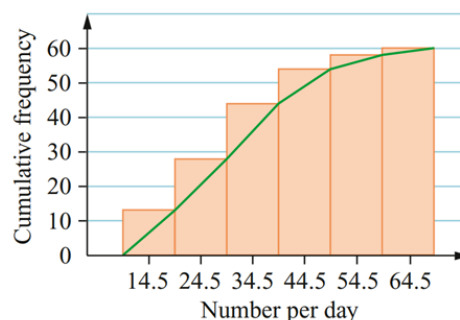
1 Determine the five-number summary from these cumulative histograms.

a



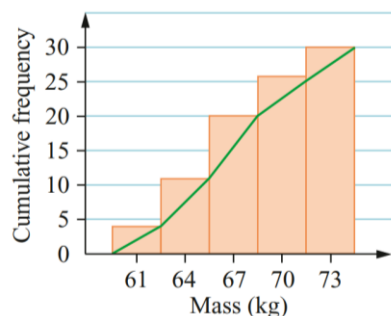
Min: Q_1 : Q_2 : Q_3 : Max:

b



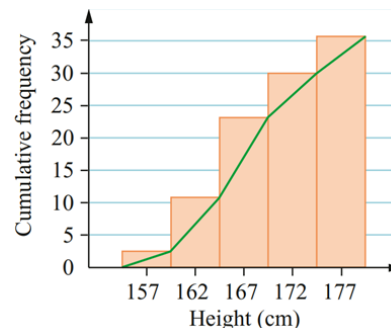
Min: Q_1 : Q_2 : Q_3 : Max:

c



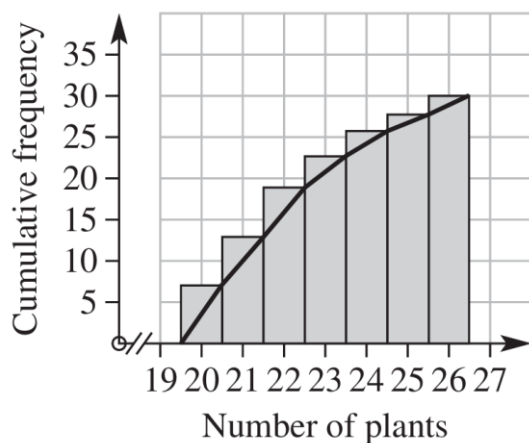
Min: Q_1 : Q_2 : Q_3 : Max:

d



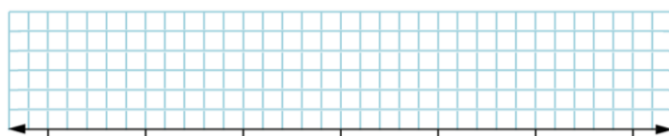
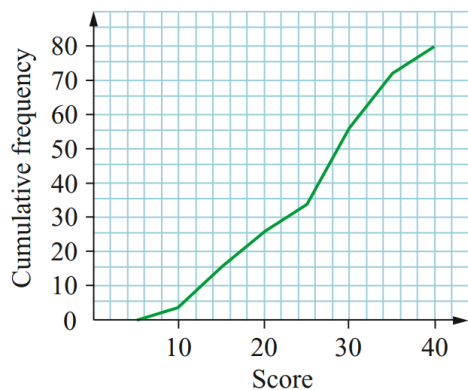
Min: Q_1 : Q_2 : Q_3 : Max:

2 Determine the five-number summary and IQR from this graph.

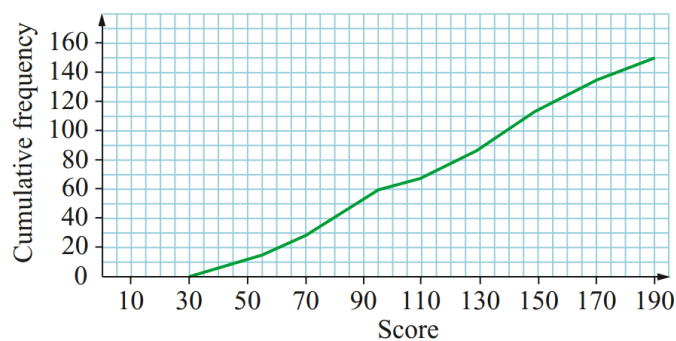


3 Determine the five-number summary from these ogives, then draw the box plot.

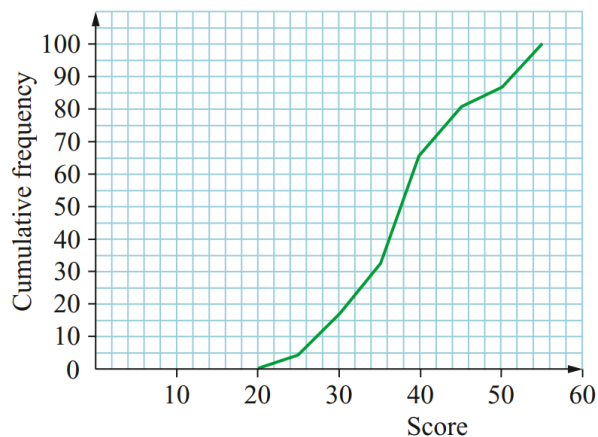
a



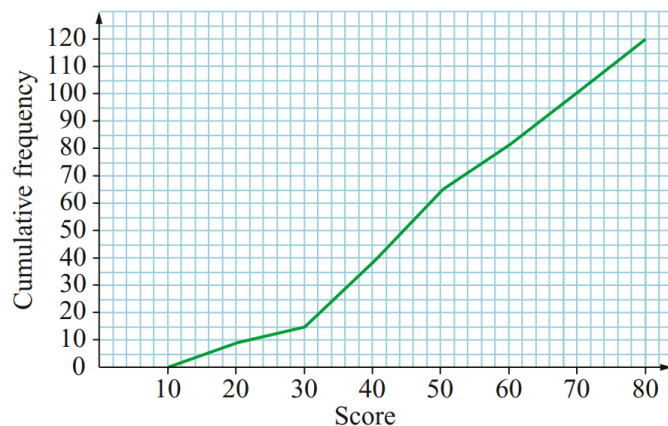
b



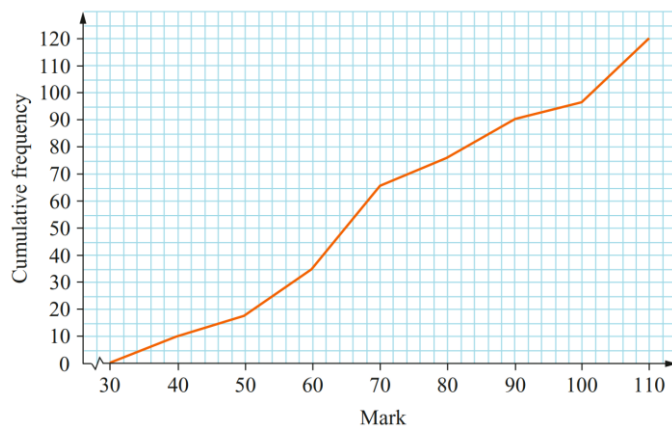
c



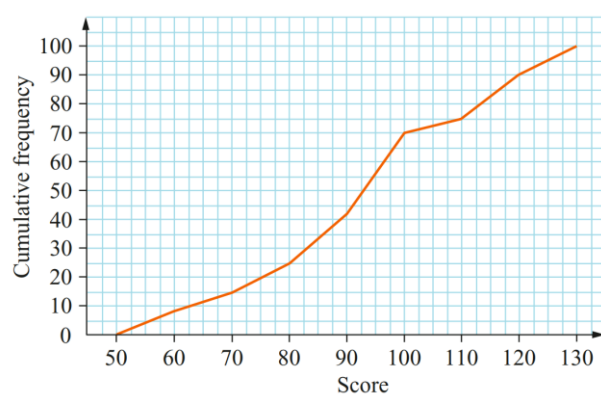
d



e

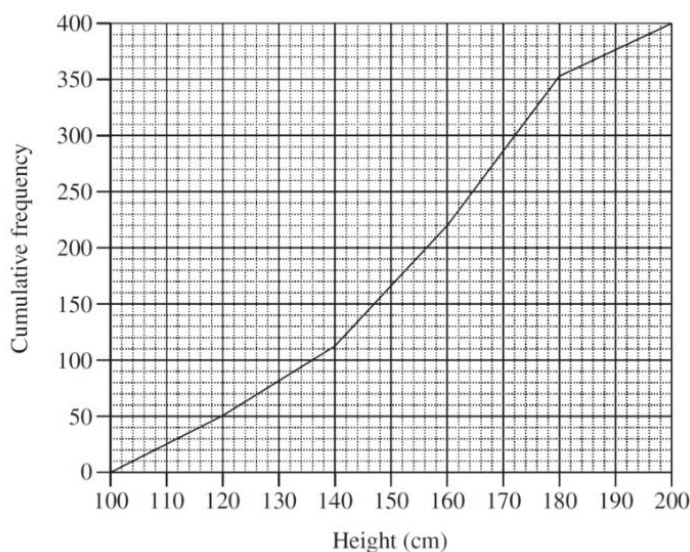


f



4 2016 HSC Standard 2 Band 5

The heights of 400 students were measured. The results are displayed in this cumulative frequency polygon. Use the polygon to estimate the interquartile range.

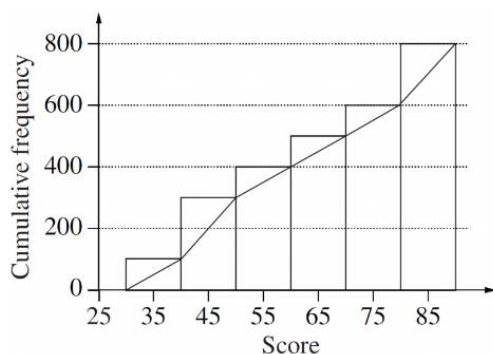
**5 2018 HSC Standard 2 Band 5**

A cumulative frequency table for a data set is shown. What is the interquartile range?

<i>Score</i>	<i>Cumulative frequency</i>
1	5
2	9
3	16
4	20
5	34
6	42

6 2005 HSC Standard 2 Band 5

A set of data is represented by the cumulative frequency histogram and ogive. What is the best approximation for the interquartile range for this set of data?

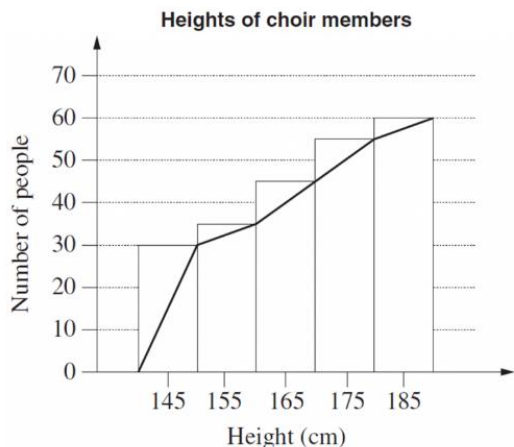


7 2006 HSC Standard 2 Band 5

The heights of the 60 members of a choir were recorded. These results were grouped and then displayed as a cumulative frequency histogram and polygon.

The shortest person in the choir is 140 cm and the tallest is 190 cm.

Draw an accurate box and whisker plot to represent the data.



8 2019 HSC Standard 2 Band 5

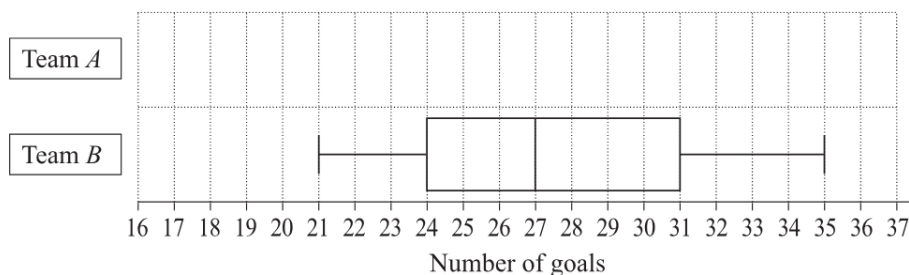
Two netball teams, Team A and Team B, each played 15 games in a tournament.

For each team, the number of goals scored in each game was recorded.

The frequency table shows the data for Team A.

The data for team B was analysed to create the box plot.

<i>Number of goals</i>	<i>Frequency</i>
19	1
20	0
21	1
22	1
23	1
24	3
25	0
26	4
27	3
28	1

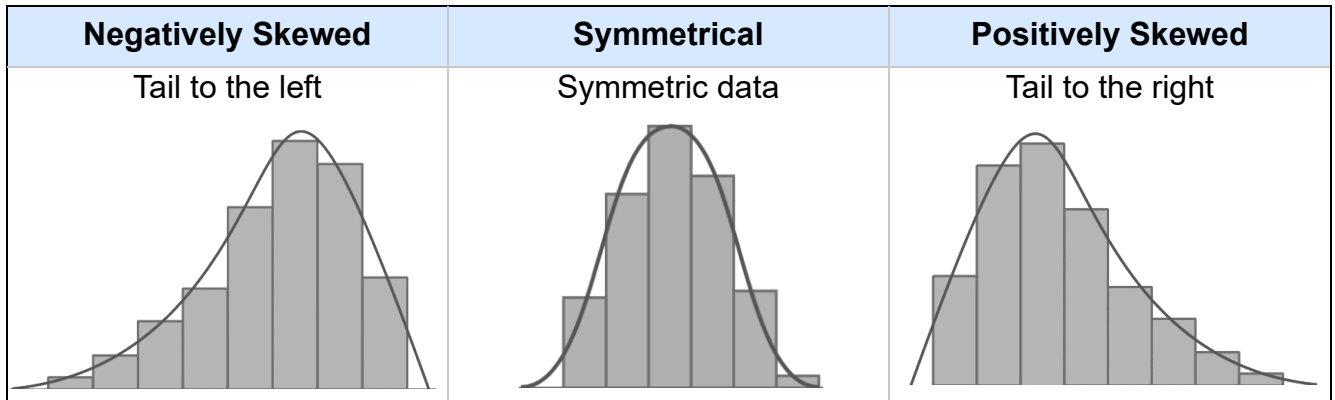


Compare the distributions of the number of goals scored by the two teams.

Support your answer with the construction of a box-plot for the data for Team A.

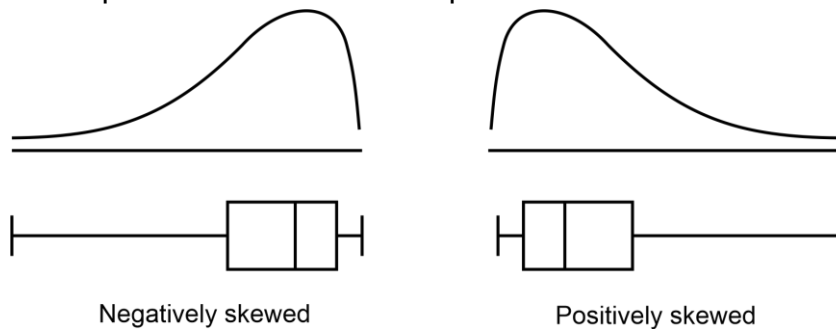
SKEWNESS AND BOX PLOTS

SHAPE OF A DISTRIBUTION



SKEWNESS AND BOX PLOTS

- The whiskers of a box plot tell us about the shape of the distribution.



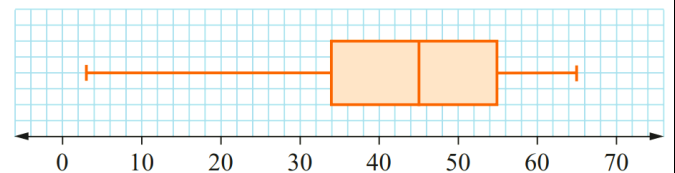
Example Identify skewness of data from a box plot.

Describe the shape of the distribution.

The box plot is not symmetrical.

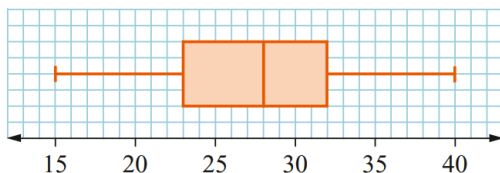
The whisker for the bottom 25% is very long, indicating a tail.

Therefore, the box plot is negatively skewed.



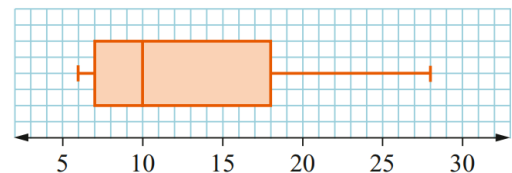
Identify whether these box plots are skewed or symmetrical.

a



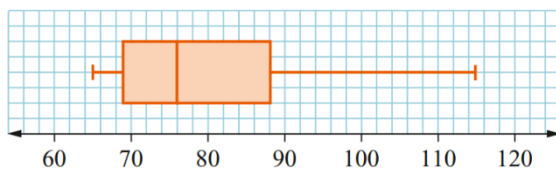
Symmetrical

b



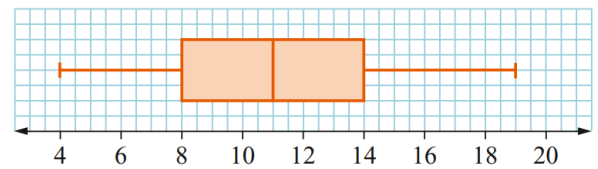
Positively skewed

c



Positively skewed

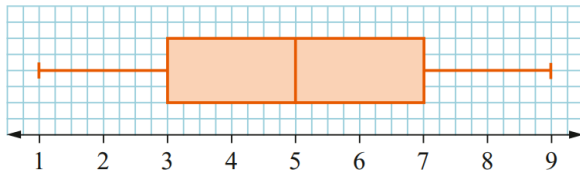
d



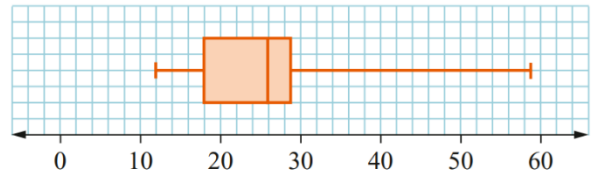
Symmetrical

1 Describe the shape of the data distribution for each box plot.

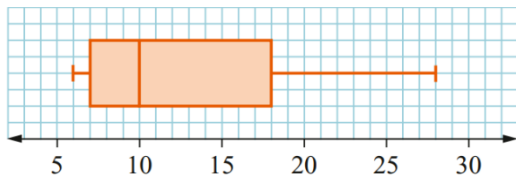
a



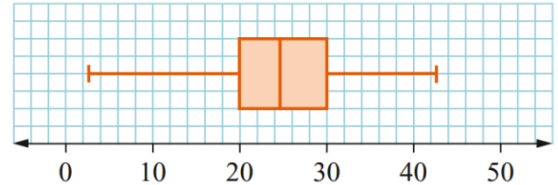
b



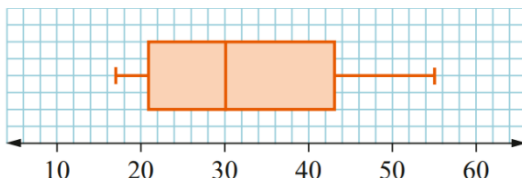
c



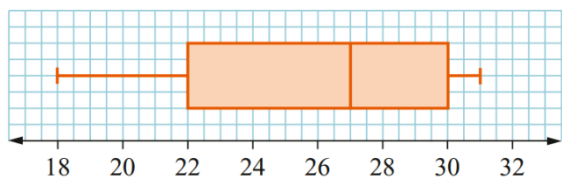
d



e

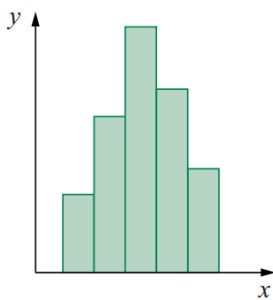


f

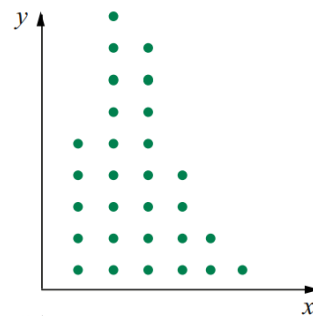


2 Match each histogram or dot plot with its corresponding box plot.

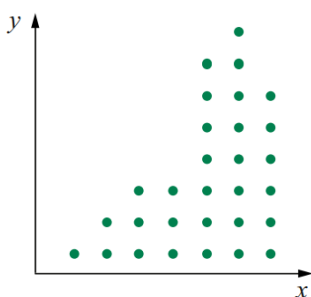
a



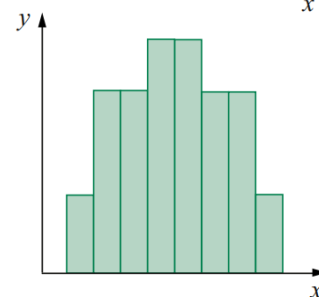
b



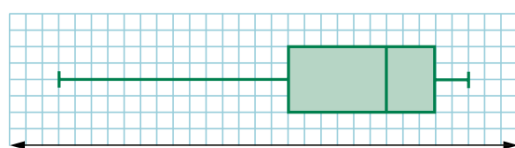
c



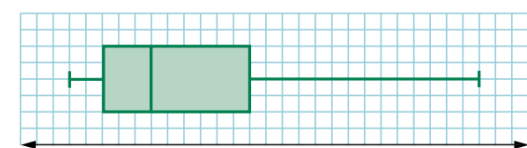
d



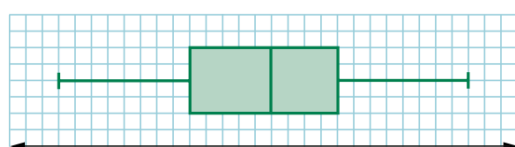
A



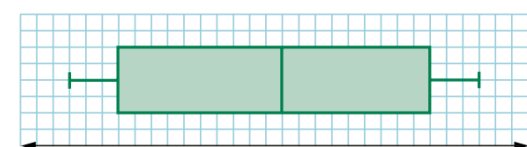
B



C



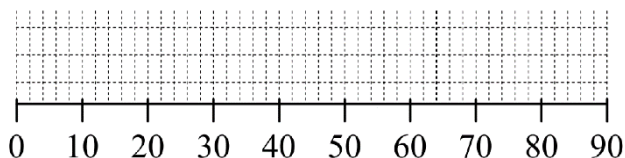
D



3 First find the five-number summary, then sketch the dot plot and decide whether it is skewed or not.

a 4 17 25 27 38 44 55 58 76 76 84

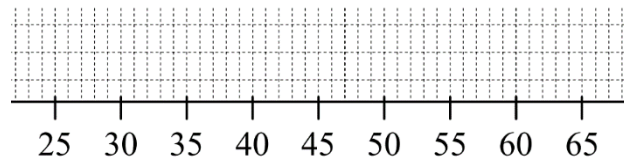
Min: Q_1 : Q_2 : Q_3 : Max:



Positive skew | Negative skew | Symmetric

b 16 21 35 41 47 58 62 72

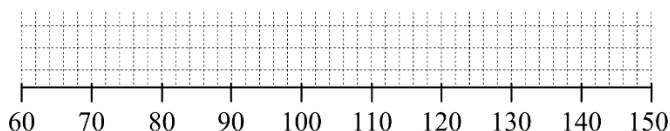
Min: Q_1 : Q_2 : Q_3 : Max:



Positive skew | Negative skew | Symmetric

c 62 64 48 74 86 94 107 111 126 145

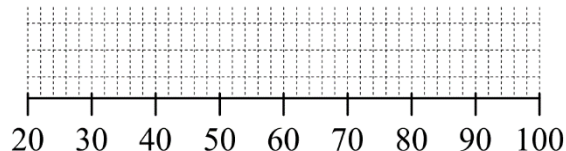
Min: Q_1 : Q_2 : Q_3 : Max:



Positive skew | Negative skew | Symmetric

d 24 39 49 55 63 63 37 72 78 82 86

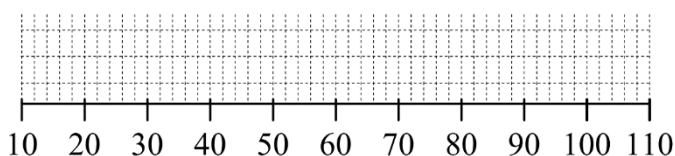
Min: Q_1 : Q_2 : Q_3 : Max:



Positive skew | Negative skew | Symmetric

e 72 24 66 94 11 45 52 78 60 106 18

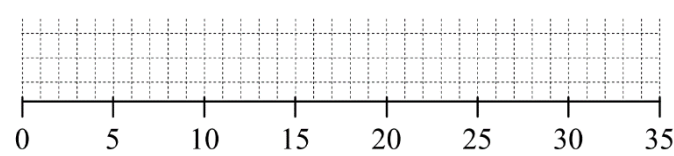
Min: Q_1 : Q_2 : Q_3 : Max:



Positive skew | Negative skew | Symmetric

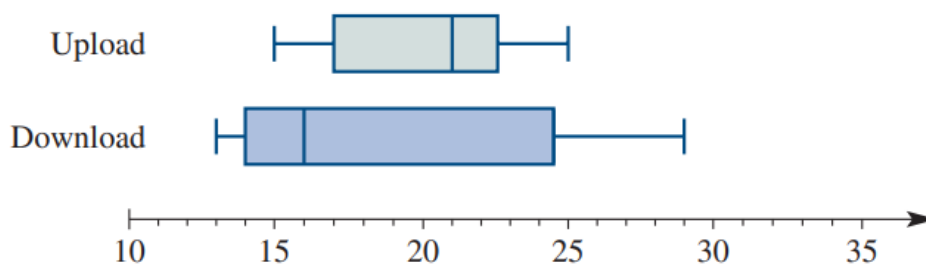
f 14 6 21 19 2 28 11 34 30 16 25 31

Min: Q_1 : Q_2 : Q_3 : Max:



Positive skew | Negative skew | Symmetric

- 4 The parallel box-and-whisker plot shows the amount of data uploaded and downloaded by Anthony each day for the past month.



Does data uploaded or data downloaded have a larger range?

Does data uploaded or downloaded have a larger interquartile range?

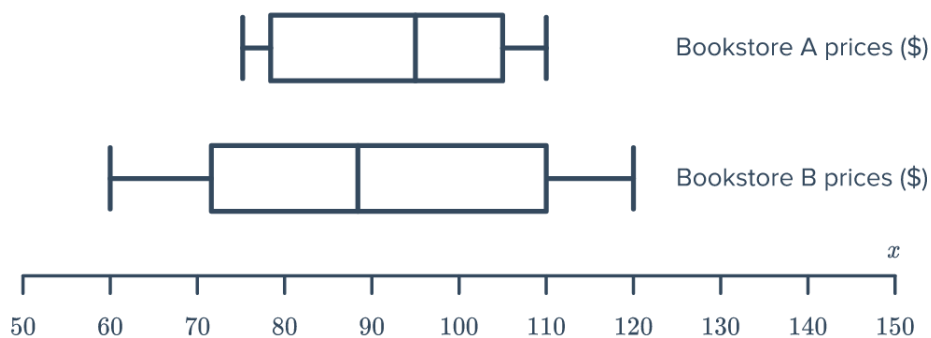
What is the median amount of data downloaded and uploaded?

.....

Describe the shape of the distribution for data uploaded and data downloaded.

.....

- 5 Two bookstores recorded the selling price of all their books. The results are presented in the parallel box plots:



- a Which bookstore had the more consistent prices?
- b Comparing the most expensive books in each store, how much more expensive is the one in store B?

.....

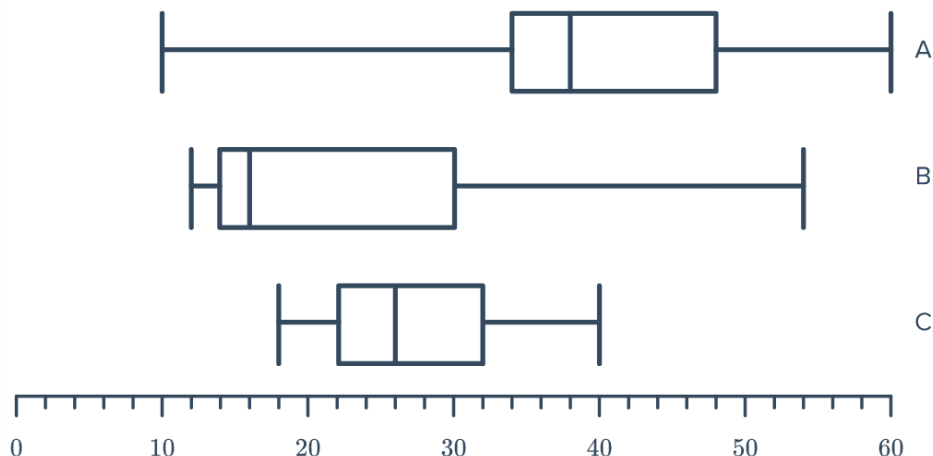
- c True or False: 25% of the books in Bookstore B are at least as expensive than the most expensive book in Bookstore A.

.....

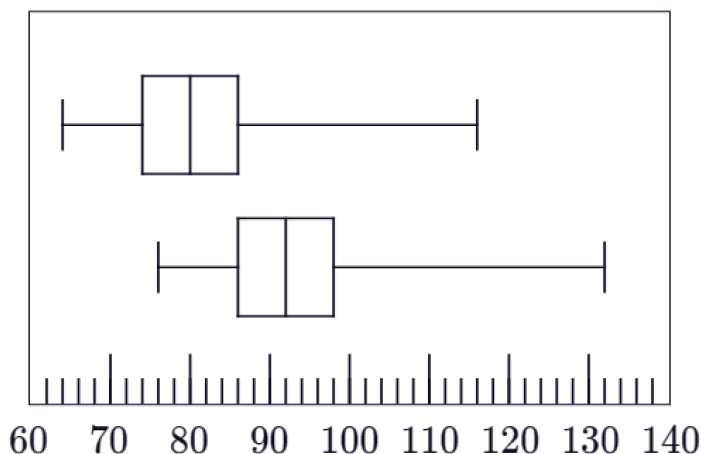
- d True or False: 25% of the books in Bookstore B are cheaper than the cheapest book in Bookstore A.

.....

- 6** A cinema is showing three films, labelled A, B, and C.
The ages of people watching the films are illustrated in the three parallel box plots.



- a** One of the films is very violent and is restricted for viewers under 18.
Which film do you think it is? Explain your answer.
-
-
- b** Which film would you recommend for a group of 15-year-olds to watch?
Explain your answer, making reference to skewness and clustering.
-
-
- c** Which film would you recommend to a family of two parents in their 40's and two teenagers? Explain your answer.
-
-
- 7** A cyclist measures her heartrate in beats per minute (bpm) before and after a ride.
- a** Which box plot do you think is before her ride and which is after?
Explain with reference to clustering.
-
-
-



- b** Compare her IQR before and after a ride. Does her heartrate remain consistent?
-
-
-

3 Draw box plots to compare these two datasets.

Test 1	Tens	Test 2
7 2	4	
6 3 0	5	2 3 8
6 6 5 4 4	6	3 3 3
8 2 2 0	7	1 1 2 2 6 8
9 7 6 6	8	3 4 4 6 8
2 2	9	5 6 6

Test 1

Min =

Q_1 =

Q_2 =

Q_3 =

Max =

IQR = -

IQR =

Test 2

Min =

Q_1 =

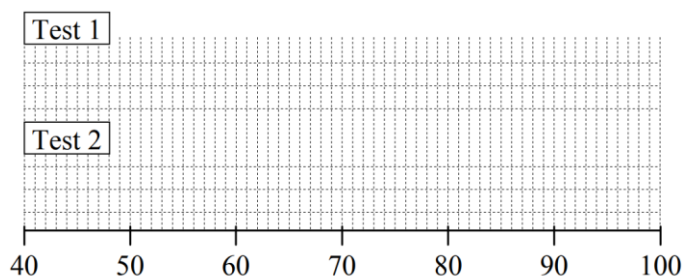
Q_2 =

Q_3 =

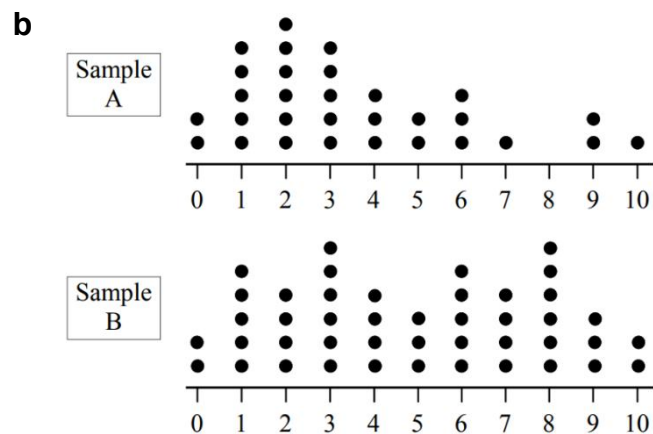
Max =

IQR = -

IQR =



Write a few sentences comparing the measures of centre, spread, and skewness.



Sample A

Min =

Q_1 =

Q_2 =

Q_3 =

Max =

IQR = -

IQR =

Sample B

Min =

Q_1 =

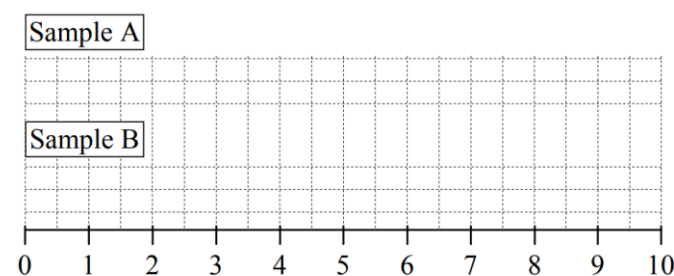
Q_2 =

Q_3 =

Max =

IQR = -

IQR =

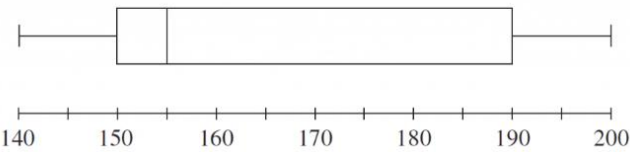


Write a few sentences comparing the measures of centre, spread, and skewness.

8 2011 HSC Standard 2 Band 5

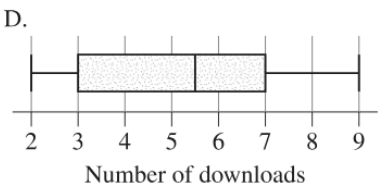
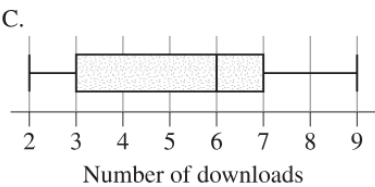
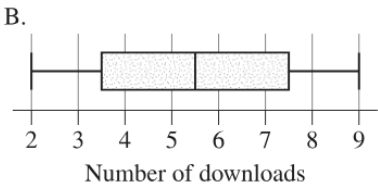
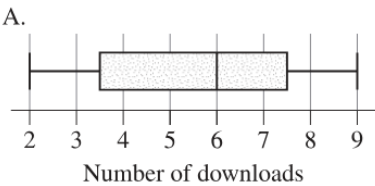
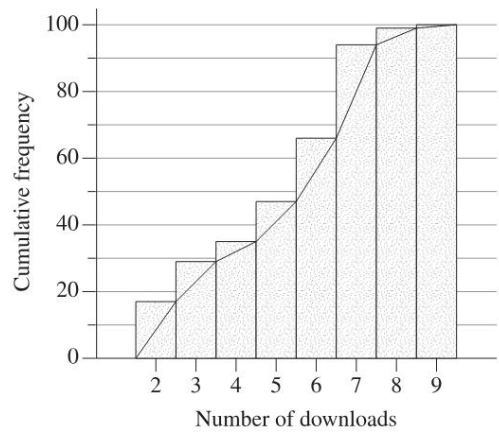
A set of data is displayed in this box-and-whisker plot.
Which of the following best describes this set of data?

- (A) Symmetrical
- (B) Positively skewed
- (C) Negatively skewed
- (D) Normally distributed



9 2022 HSC Standard 2 Band 6

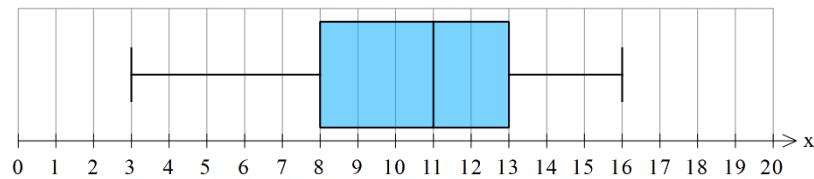
The cumulative frequency graph shows the distribution of the number of movie downloads made by 100 people in one month.
Which box-plot best represents the same data as displayed in the cumulative frequency graph?



CHECK YOUR UNDERSTANDING

Interpret box plots

The box plot shows the scores of a class in a quiz.



1 Identify the median.

- a 10 b 11 c 5 d 13

2 Identify the range.

- a 20 b 11 c 5 d 13

3 Identify the interquartile range.

- a 20 b 11 c 5 d 13

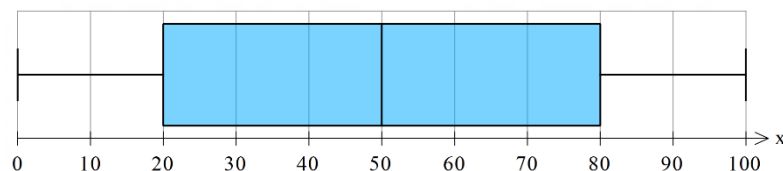
4 How many students took the quiz?

- a 20 b 16 c 5 d Impossible to say

5 What proportion of students scored between 13 and 16?

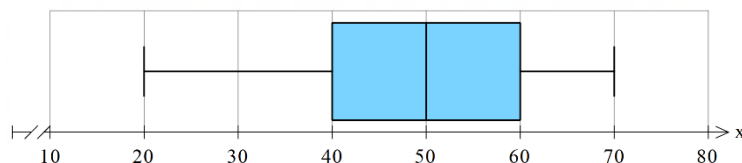
- a $\frac{1}{4}$ b $\frac{3}{13}$ c $\frac{3}{20}$ d Impossible to say

6 What percentage of students scored between 20 and 50 marks?



- a 25% b 30% c 50% d Not enough information.

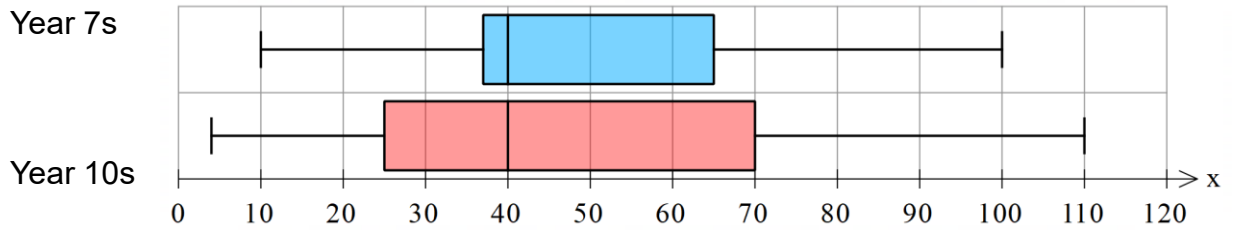
7 The box plot shows how long 300 boys spent watching TV last week. How many boys watched between 40 and 50 minutes of TV?



- a 60 b 10 c 75 d Not enough information.

Compare datasets using box plots

- 1** The box plot shows the results of a survey of how many times students used their mobile phones last week.



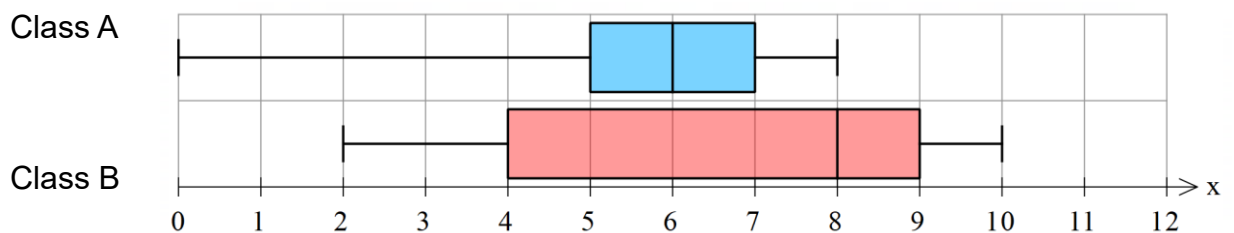
On average, which group used their phones more?

- a** Year 7s **b** Year 10s **c** Both same **d** Impossible to say

Which group had the most consistent phone usage?

- a** Year 7s **b** Year 10s **c** Both same **d** Impossible to say

- 2** The box plot shows the results of two classes' maths exams.



On average, which class performed better?

- a** Class A, their median is higher. **b** Class B, their median is higher. **c** Class B, they had the highest mark **d** Class B, their IQR was greater.