

MATHSPACE

CUSTOM TASKS

YEAR: 9MF

TOPIC:

DATA ANALYSIS A

BOOKLET

C.T 12.01C Centre and spread (22Q'S)

1. For each of the following decide whether the statements are true or false.

- a Two sets of data have the same highest and lowest values. This means they must have the same mode.

True

A

False

B

- b Two sets of data that have the same highest and lowest values must have the same mean.

True

A

False

B

2. For each of the following statements, decide whether they are true or false:

- a If two sets of data have the same median then the data sets must themselves be the same

True

A

False

B

- b If two sets of data have very different modes then the highest values cannot be the same

True

A

False

B

3. Assess how various changes to data sets alter their characteristics.

- a Consider the set of data:

1, 2, 2, 4, 4, 5, 6, 6, 8, 8, 8, 9, 9

If one score of 8 is changed to a 9, which two of the following would be altered?

Median

A

Mean

B

Range

C

Mode

D

- b Consider this set of data that represents the number of apps on six people's phones.

11, 12, 15, 17, 19, 19

If each person downloads another 7 apps, which one of the following would *not* change?

Mode

A

Mean

B

Range

C

Median

D

4. Given the stem-and-leaf plot:

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

Stem	Leaf
8	0 9 9
9	4

Key: 1 | 2 = 12

- a Find the mean from the stem-and-leaf plot. Round your answer to two decimal places.
- b Find the mode for the data provided in the stem-and-leaf plot.
- c Find the median.
- d What is the range of the scores?
5. The stem and leaf plot shows the number of hours students spent studying for a science exam.
- From the data in the stem and leaf plot, find (to two decimal places if necessary) the:

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

- a mean.

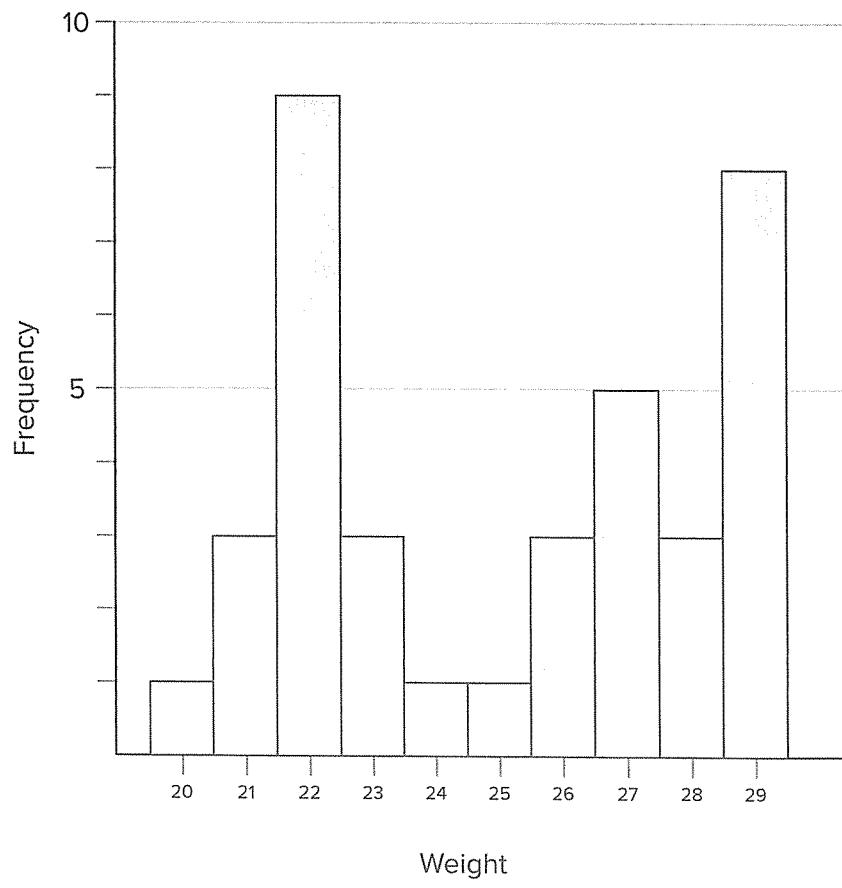
b median.

c mode.

d range.

6. The histogram shows the weights of preschool students in kilograms.

From the data in the bar graph, find the:



a mean.

Round your answer to two decimal places, if necessary

b median.

c mode.

d range.

7. Consider the following set of data:

27, 50, 24, 37, 47, 41, 27, 126, 44, 27

a Fill in this table of summary statistics.

Mean

Median

Mode

Range

b Which data value is an outlier?

c Fill in this table of summary statistics after removing the outlier.

Mean

Median

Mode

Range

- d Let A be the original data set and B be the data set without the outlier.

Fill in this table using the symbols $>$, $<$ and $=$ to compare the statistics before and after removing the outlier.

With outlier Without outlier

Mean: A B

Median: A B

Mode: A B

Range: A B

8. The range of a set of scores is 5, and the highest score is 18.
What is the lowest score in the set?
9. A group of students had a range in marks of 11 and the lowest score was 5.
What was the highest score in the group?
10. Consider the data shown in the table:

Score	Frequency
68	16
69	41
70	30
71	31
72	50
73	29

- a Calculate the range of the scores.
- b What is the mode?
11. What is the modal class of this data set?

Class	Frequency
30 – 39	3
40 – 49	3
50 – 59	4
60 – 69	2
70 – 79	5
80 – 89	8

80 – 89 A 40 – 49 B

60 – 69 C 30 – 39 D

50 – 59

E

70 – 79

F

12. Consider the table below.

Score	Frequency
1 - 4	1
5 - 8	5
9 - 12	10
13 - 16	5
17 - 20	3

- a Use the midpoint of each class interval to determine the mean of the following sample distribution, correct to one decimal place.
- b Which is the modal group?

5 - 8

A

13 - 16

B

1 - 4

C

9 - 12

D

17 - 20

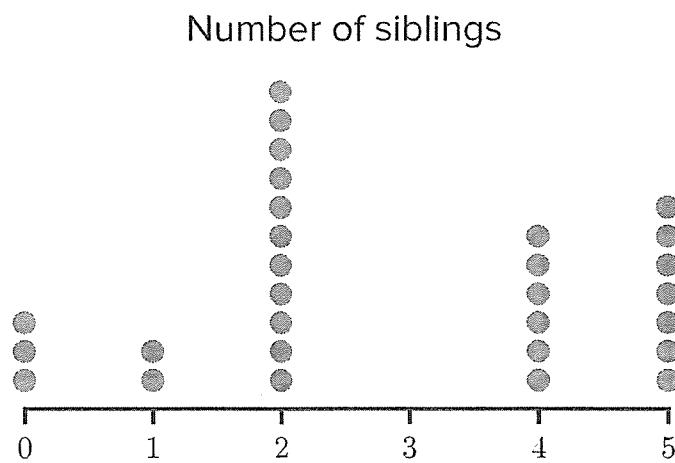
E

13. A real estate agent wanted to determine a typical house price in a certain area. He gathered the selling price of some houses (in dollars):
 327 000, 376 000, 424 000, 439 000, 444 000, 469 000, 472 000, 475 000



- a Calculate the mean house price.

- b What percentage of the house prices exceeded the mean?
- c Determine the median house price.
- d What percentage of house prices exceeded the median?
14. Each student in the class was asked to write down the number of siblings they had. The teacher recorded the results in the dot plot.

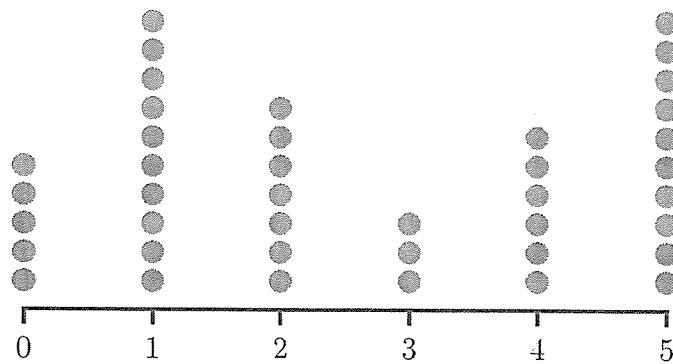


- a How many students were there in the class?
- b If none of the students share the same siblings, then how many siblings are there in total?
- c What is the average number of siblings for a student in this class? Round your answer to one decimal place.

- d What is the average number of children in a family for a student in this class?

15. Here is a dot plot of the number of goals scored across each of Rosey's soccer games.

Number of goals scored



- a How many games were played in total?
- b How many goals were scored in total?
- c What was the average number of goals per game? Round your answer to one decimal place.
16. A cyclist measured his heart rate immediately after finishing each event in which he competed. The results are recorded in a stem plot.

Stem	Leaf
16	2
17	3 8
18	4 5 6 9

Stem	Leaf
19	5 5

Key: 12 | 3 = 123

- a How many events did the cyclist compete in?
- b What is his mean post event heart rate?
17. The size of each earthquake that occurred in a region over a three year period, measured from 0 to 9.9, is recorded in a stem-and-leaf plot.

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

Key: 5 | 2 = 5.2

- a How many earthquakes in total were recorded?
- b What is the mean number of earthquakes per year in the region?
- c It was found that the combined total of all earthquake sizes was 87.

What was the mean size of an earthquake that occurred during the period?

18. Consider this stem and leaf plot:

Stem	Leaf
4	3
5	
6	
7	
8	
9	1 1 2 3
10	1 1 3 5 7
11	2 8

Key: $1 | 2 = 12$

- a What is the range of the data set?

- b Which score is an outlier?

- c If the outlier is removed, what is the new range of the data set?

19. Consider this stem plot containing 10 scores.

Stem	Leaf
3	4 4 9
4	6 6 8 9

Stem	Leaf
5	1 4
6	
7	
8	4

Key: $1 | 2 = 12$

- a What is the mean?

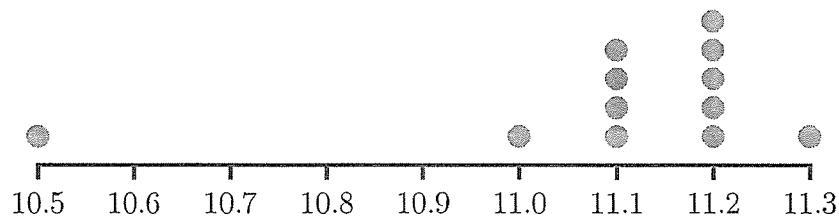
Write your answer as a decimal.

- b If the outlier is removed, what is the new mean?

Round your answer to one decimal place.

20. The glass windows for an airplane are cut to a certain thickness, but machine production means there is some variation. The thickness of 12 panes of glass produced is measured (in millimetres), and the dot plot shows the results.

Glass thickness (mm)



- a What is the median thickness?

- b If the outlier is removed what is the new median?

21. The selling price of recently sold houses is given below.

\$760000, \$650000, \$810000, \$780000, \$760000, \$590000, \$1360000

- a. What is the mean selling price?

Round your answer to the nearest thousand dollars.

- b. What is the median selling price in dollars?

- c. Recalculate the mean selling price in dollars excluding the outlier.

- d. Recalculate the median selling price in dollars excluding the outlier.

- e. Which measure of centre best identifies the typical selling price of recently sold houses?

Mean

A

Median

B

22. The weight of fish caught in a 'weigh and release' fishing competition, in kilograms are given below.

12.5, 15.1, 13, 14.2, 14.5, 14.9, 12.5, 14.3, 1.5

- a. What is the mean weight in kilograms?

- b. What is the median weight in kilograms?

- c Recalculate the mean weight in kilograms excluding the outlier.
- d Recalculate the median weight in kilograms excluding the outlier.
- e Which measure of centre best identifies the typical fish weight?

Mean

A

Median

B

C.T 12.04C Cumulative frequency (11Q'S)

1. Complete the frequency table based on the data set:

3, 3, 3.19, 4, 7, 8, 5.73, 3, 2, 5, 2, 3, 4, 6.35, 3, 6, 2, 3.3, 6.75, 3.4



Score	Frequency
-------	-----------

$2 \leq x < 3$

$3 \leq x < 4$

$4 \leq x < 5$

$5 \leq x < 6$

$6 \leq x < 7$

$7 \leq x < 8$

$8 \leq x < 9$

2. Complete the cumulative frequency column:

Score (x)	Frequency (f)	Cumulative Frequency (cf)
---------------	-------------------	-------------------------------

$1 \leq x < 5$	9	
----------------	---	--

Score (x)	Frequency (f)	Cumulative Frequency (cf)
$5 \leq x < 9$	6	
$9 \leq x < 13$	7	
$13 \leq x < 17$	5	
$17 \leq x < 21$	7	

3. Fill in the missing values in the table below.

Score (x)	Frequency (f)	Cumulative Frequency (cf)
12	8	8
13	5	
14	8	21
15		26
Total = 26		

4. For the given frequency table:

- a Complete the cumulative frequency column.

Score (x)	Frequency (f)	Cumulative Frequency (cf)
1 – 5	25	

Score (x)	Frequency (f)	Cumulative Frequency (cf)
6 – 10	33	
11 – 15	26	
16 – 20	21	
21 – 25	8	
26 – 30	3	

- b Calculate the total frequency.
- c Approximately half of the recorded scores are greater than which value?

20 A 16 B
 C 15 D

5. Consider the cumulative frequency table below:

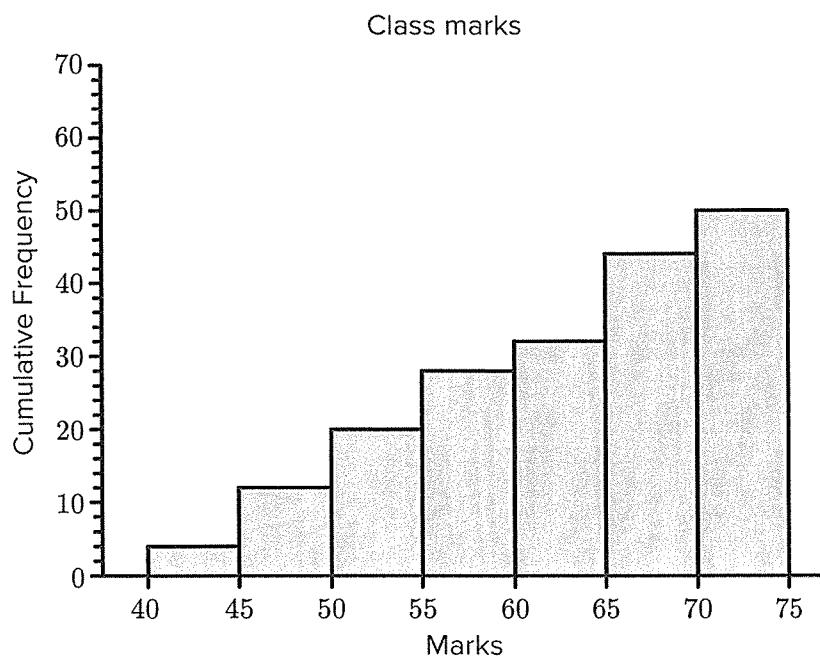
Score (x)	Cumulative frequency (cf)
22	8
23	11
24	15
25	24
26	32

- a Find the total number of scores recorded.

- b Find the number of times a score of 23 occurred.

- c Find the number of times a score less than 25 occurred.

6. For the cumulative frequency histogram below:



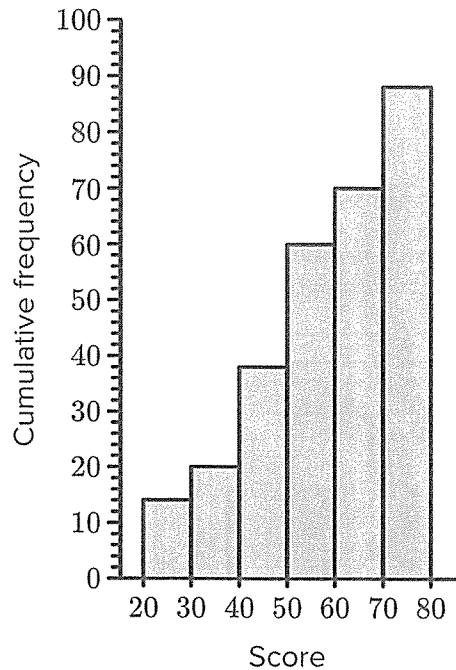
- a Find the total number of scores recorded.

- b Find the number of times a score between 45 and 50 occurred.

- c Find the number of times a score between 60 and 65 occurred.

- d Find the percentage of scores that were 54 or less.

7. For this cumulative frequency histogram:



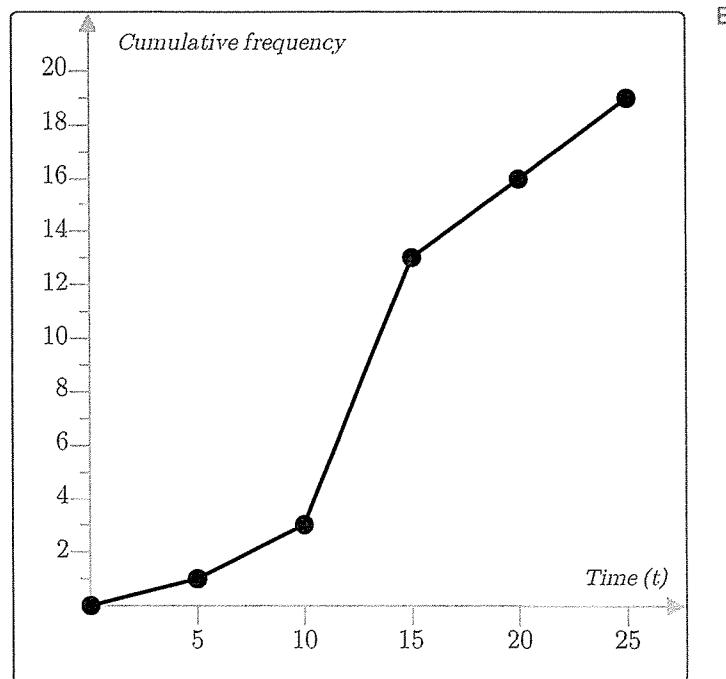
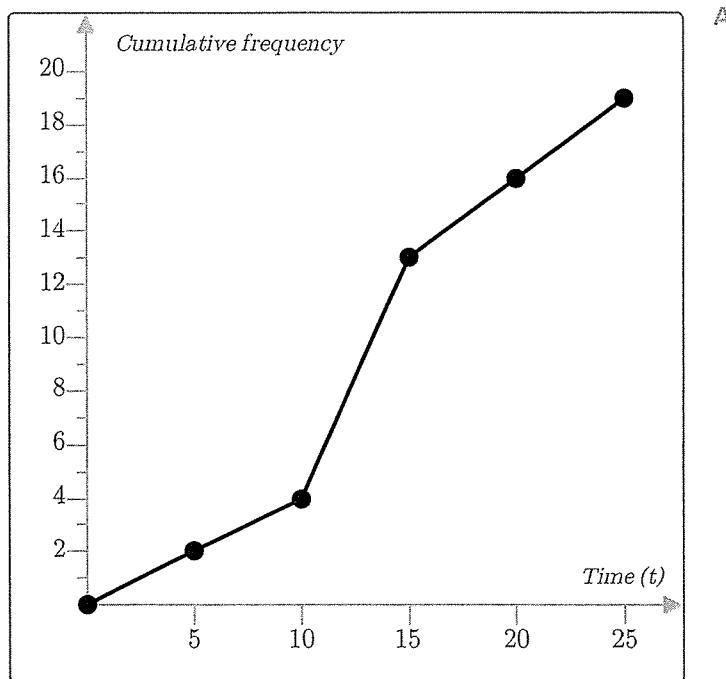
- a How many scores are there in total?
- b How many scores between 40 and 50 occurred?
8. The frequency table shows the wait time (t), in minutes, for customers at a small restaurant during peak hours.
- a Complete the cumulative frequency column.

Time (t)	Frequency (f)	Cumulative frequency (cf)
$0 \leq t < 5$	1	
$5 \leq t < 10$	2	
$10 \leq t < 15$	10	
$15 \leq t < 20$	3	

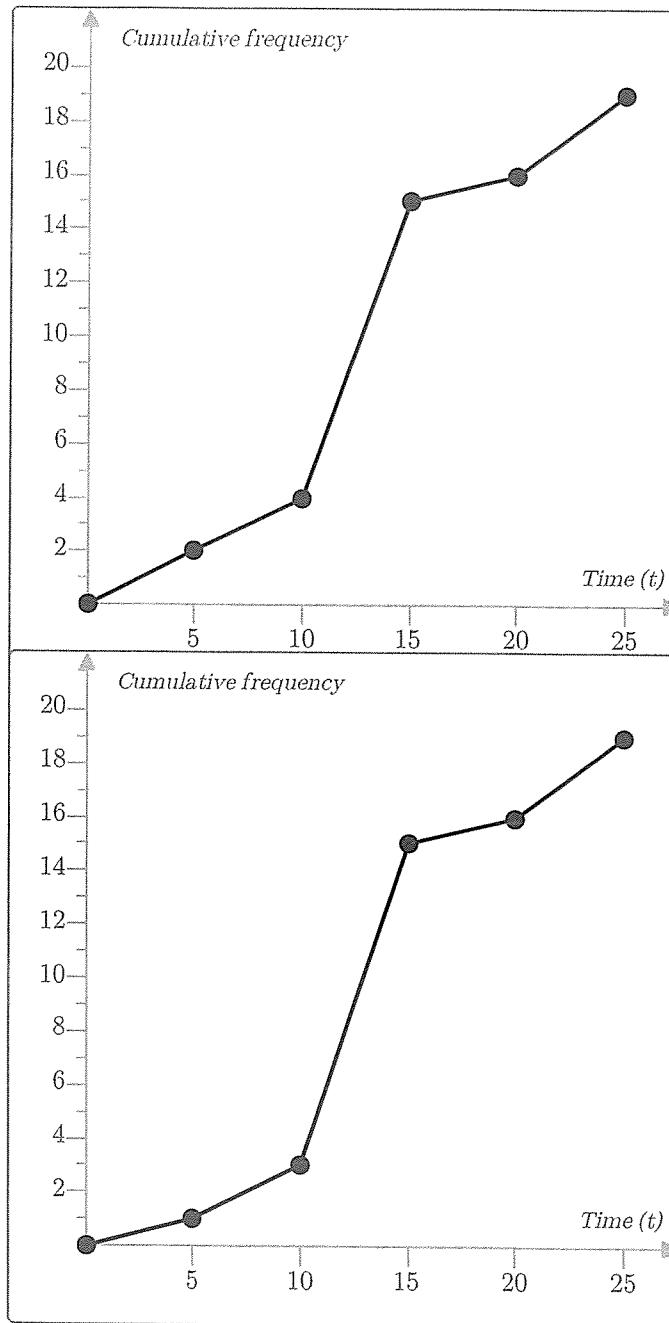
Time (t)	Frequency (f)	Cumulative frequency (cf)
--------------	-------------------	-------------------------------

$20 \leq t < 25$	3	
------------------	---	--

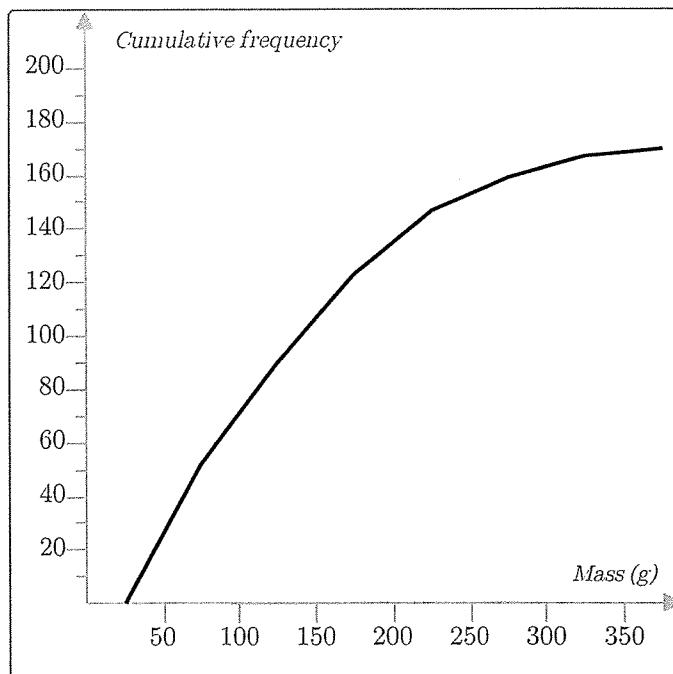
- b Determine which cumulative frequency polygon correctly represents the data.



C

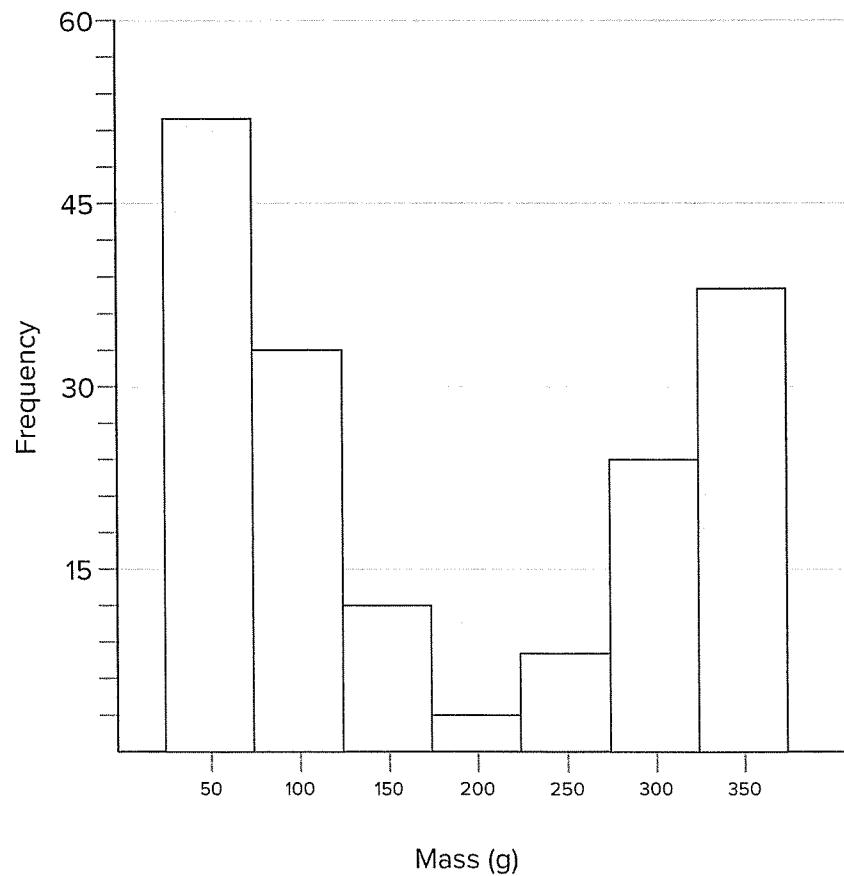


9. The graph below shows the cumulative frequency of the masses of fish caught in a fishing competition.

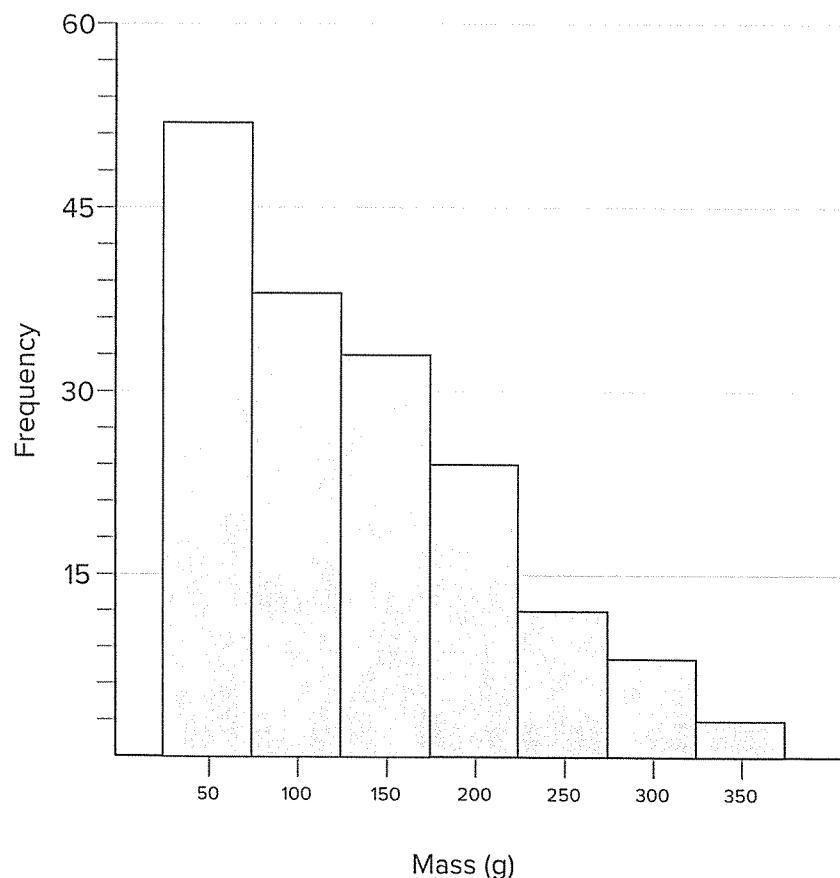


Which of the frequency histograms best depicts the distribution of the masses of the fish?

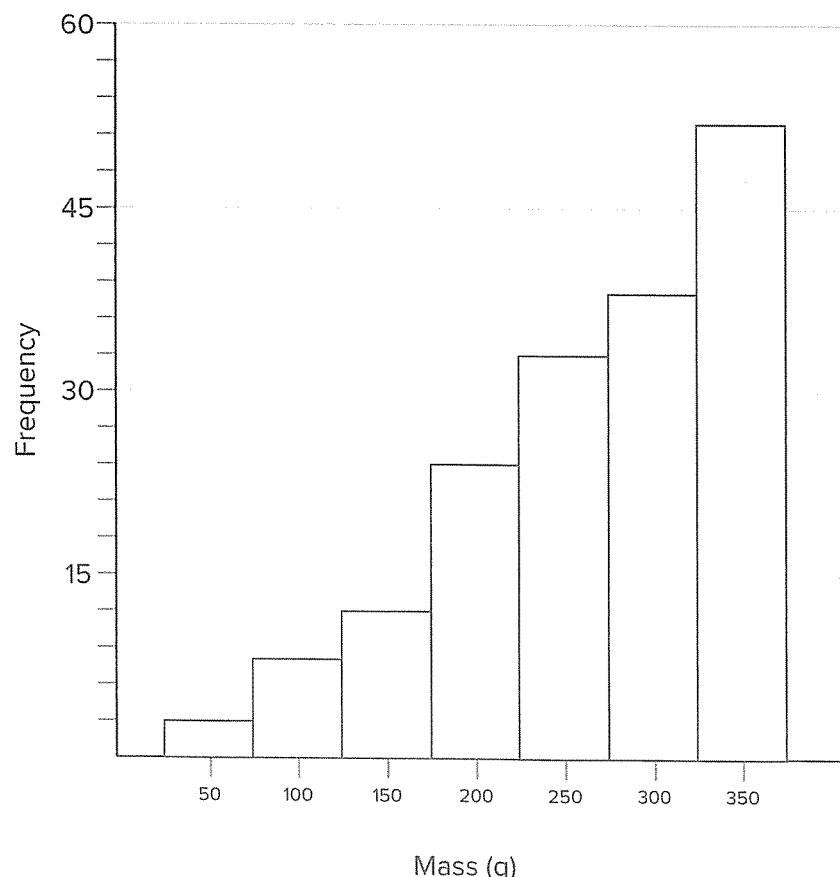
A

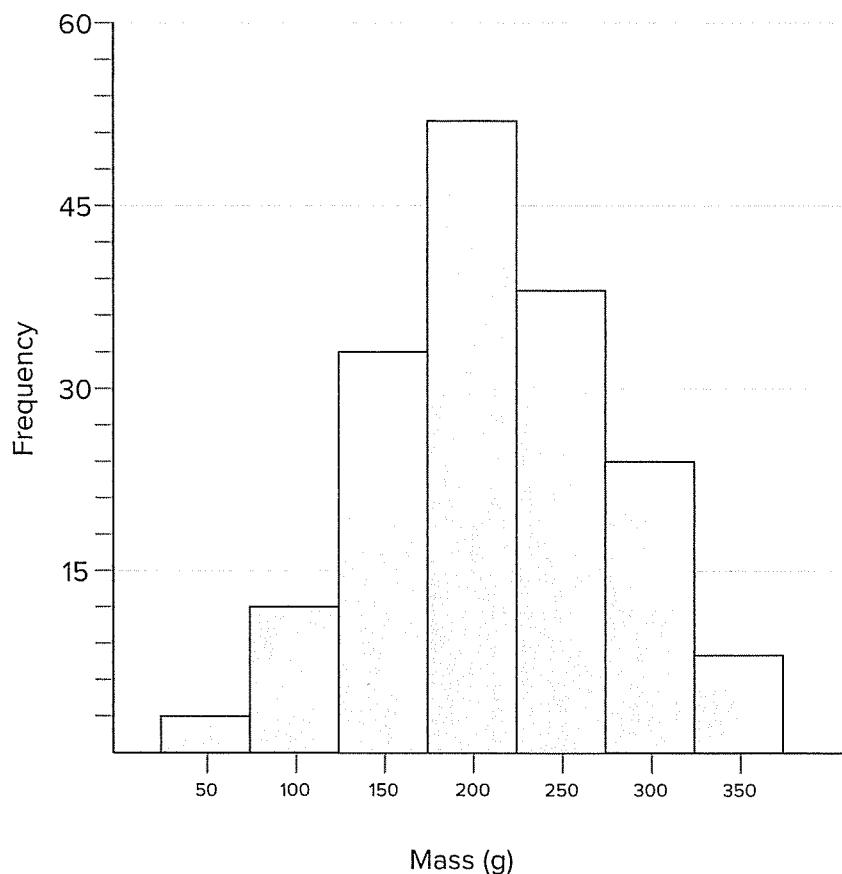


B

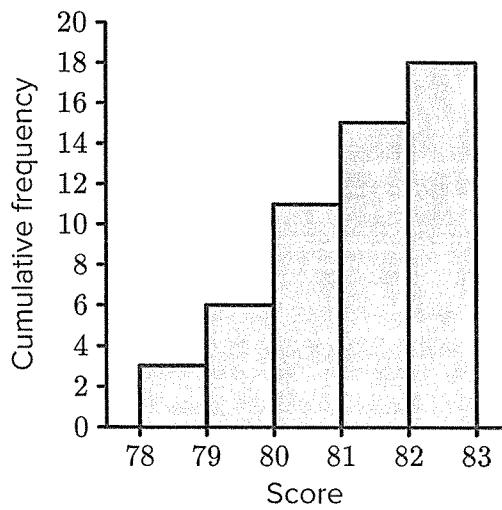


C





10. Use the cumulative frequency histogram to complete the frequency table.



Score	Frequency
-------	-----------

$$78 \leq x < 79$$

$$79 \leq x < 80$$

$$80 \leq x < 81$$

$$81 \leq x < 82$$

$$82 \leq x < 83$$

- 11.** The number of sightings of the Northern Lights were recorded across various Canadian locations over a period of 1 month. The numbers below represent the number of sightings at each location.

11, 10, 10, 9, 7, 8, 8, 12, 12, 12, 12, 12, 12, 9, 9, 12, 9, 9, 8, 8

- a Complete the table.

Number of Sightings	Number of Locations (f)	Cumulative Frequency (cf)
---------------------	-----------------------------	-------------------------------

7

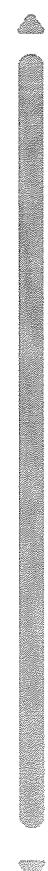
8

9

10

11

12



- b In how many locations were there at least 8 sightings?

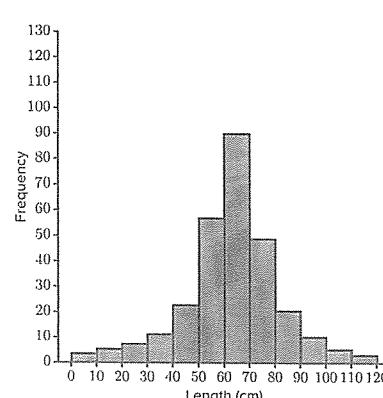
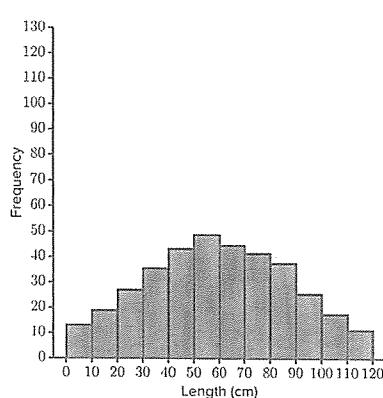
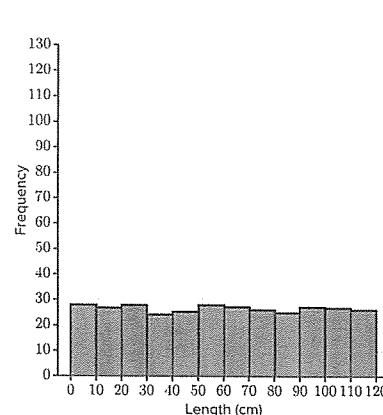
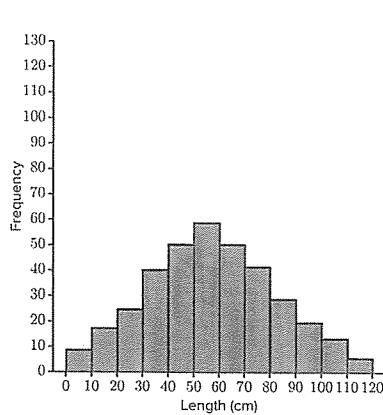
- c In how many locations were there less than 11 sightings?

C.T 12.02C Standard deviation (11Q'S)

1. A higher distribution means that the scores are more spread out from the mean.

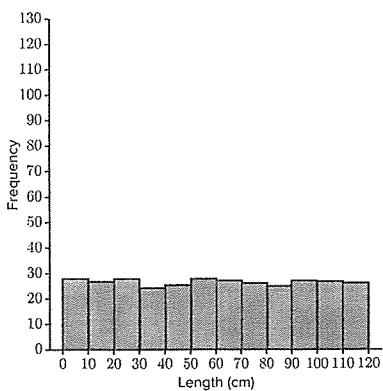
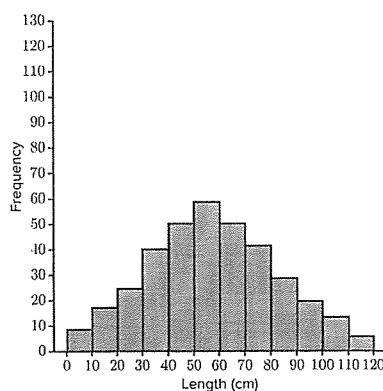
Consider the following graphs.

- a Select the graph with the highest standard deviation.

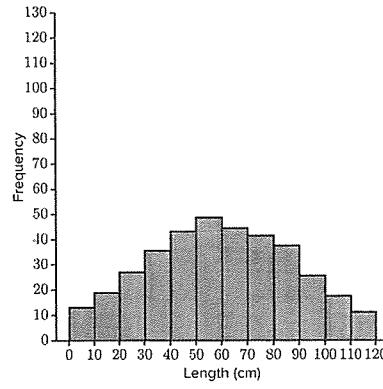


- b Select the graph with the lowest standard deviation.

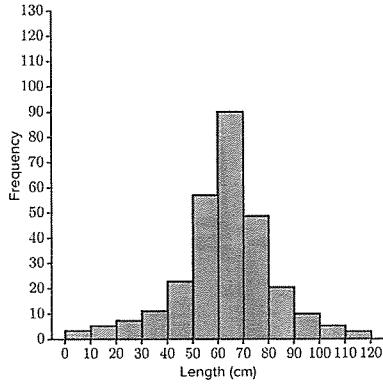
A



C



D



2. Find the standard deviation of the following set of scores by using the statistics mode on a calculator:
- 81, 90, 90, 88, 73, 80, 86, 87, 75, 82, 70, 81, 71, 81, 79, 81, 80, 86



Round your answer to two decimal places.

3. Find the standard deviation of the following set of scores by using the statistics mode on a calculator:
- 20, 44, 27, 25, 21, 28, 41, 24, 27, 39, 35, 43, 30, 17, 40

Round your answer to two decimal places.

4. Use technology to determine the standard deviation for the data represented by the frequency table.

Score	Frequency
15	13
16	9
17	23
18	19
19	8
20	13

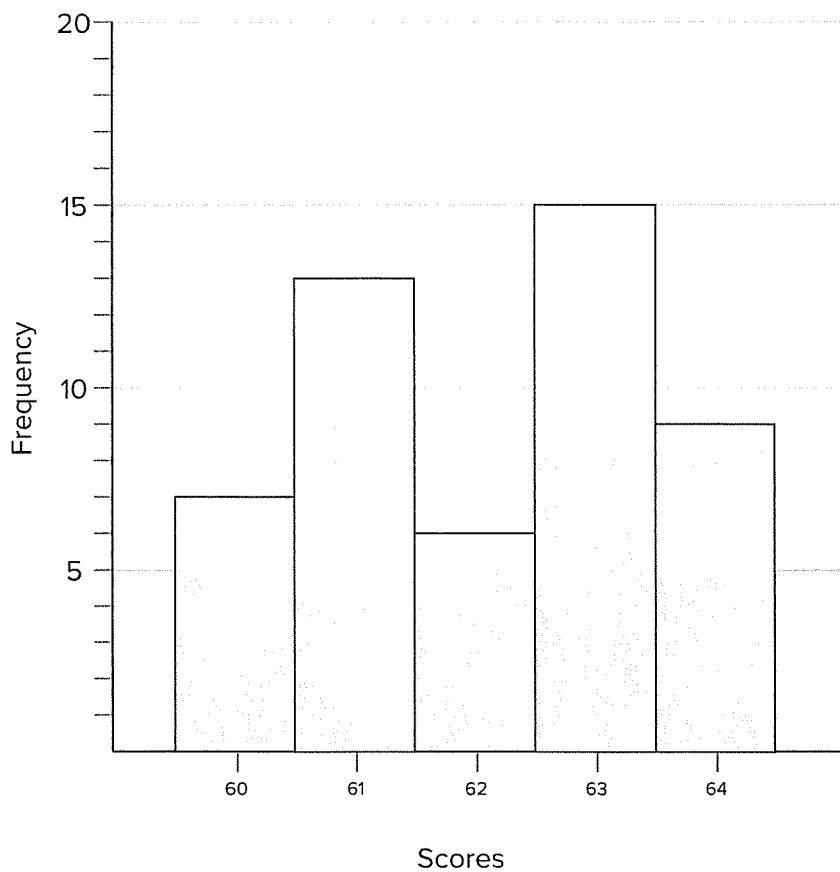
Round your answer to two decimal places.

5. Find the standard deviation of the following set of scores by using the statistics mode on the calculator:

3, 14, 11, 17, 3, 18, 15, 6, 17, 15

Round your answer to two decimal places.

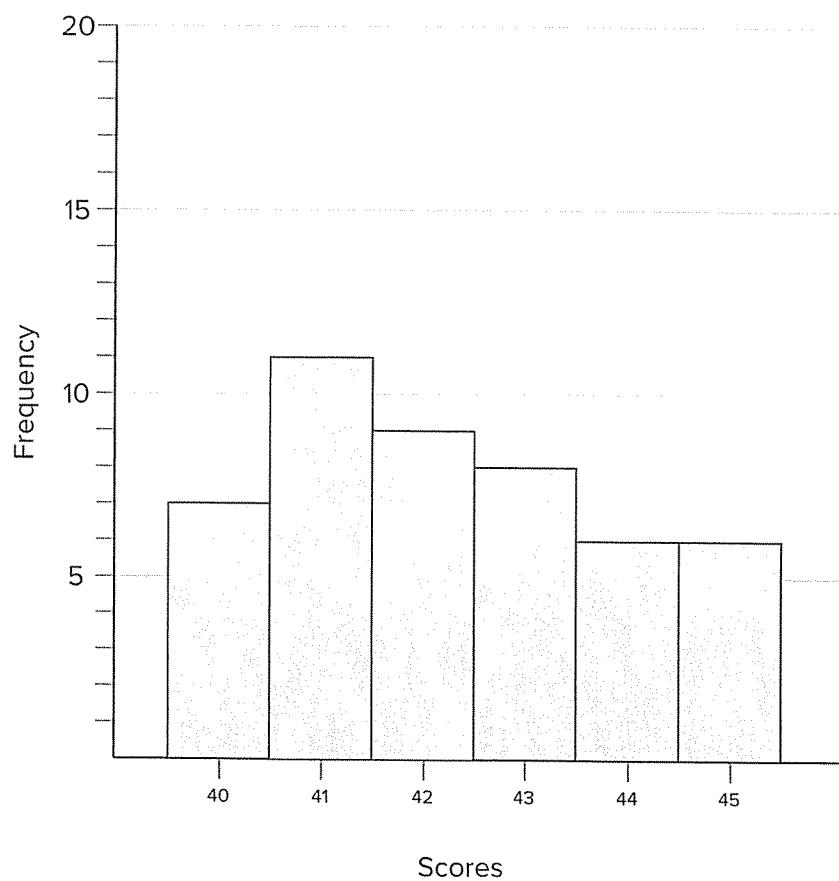
6. Using the data in the histogram below, find each of the following.



- a Find the range of the data set.

 - b Find the mean of the data set.

 - c Find the standard deviation.
Round your answer to two decimal places.
7. Using the data in the histogram below, find each of the following:



- a The range of the data set.
- b The mean of the data set.
Round your answer to two decimal places.
- c The standard deviation of the data set.
Round your answer to two decimal places.
8. The scores of five diving attempts by a professional diver are recorded below.
6.2, 5.0, 5.0, 5.3, 6.3

a Calculate the standard deviation of his scores. Round your answer to two decimal places.

b On his sixth attempt, the diver scores 9.7. This score will:

Increase the mean and
decrease the standard
deviation.

A

Decrease the mean and
increase the standard
deviation.

B

Increase the mean and
increase the standard
deviation.

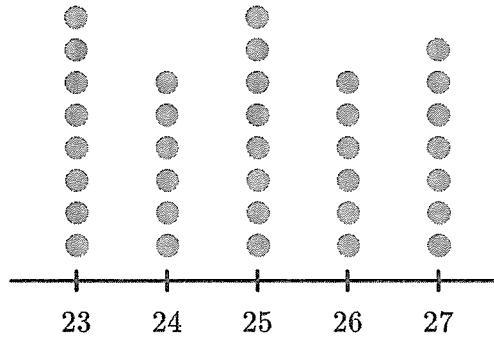
C

Decrease the mean and
decrease the standard
deviation.

D

c If each judge gave the diver the same score, what would be the standard deviation of the judges' scores?

9. Consider the dot plot shown below.



a Find the mean of the data represented by the dot plot, to two decimal places.

b Now find the MAD, rounded to two decimal places.

10. Consider the stem-and-leaf plot below.

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

Key: 1|2 = 12

- a Calculate the mean of the data set, to two decimal places.
- b Now calculate the standard deviation of the data set, to two decimal places.

11. Consider the stem-and-leaf plot below.

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

Key: 1|2 = 12

- a Calculate the mean of the data set, to two decimal places.

- b Now calculate the standard deviation of the data set, to two decimal places.

C.T 12.03C Comparisons using standard deviation (12Q'S)

1. The mean income of people in Canada is \$43000. This is the same as the mean income of people in Germany. The standard deviation of Canada is greater than the standard deviation of Germany. In which country is there likely to be the greatest difference between the incomes of the rich and poor?

Canada

A

Germany

B

2. Two cricketers compare the mean and standard deviation of their runs made per match.

They conclude that Ivan is a more consistent batter but Bianca generally scores more runs per match.

Which of the following is likely to be the case?

Ivan has a lower mean and a lower standard deviation than Bianca.

A

Ivan has a lower mean but a higher standard deviation than Bianca.

B

Ivan has a higher mean but lower standard deviation than Bianca.

C

Ivan has a higher mean and a higher standard deviation than Bianca.

D

3. Two companies record the wait time for calls to their customer hotlines over 10 calls. The recorded values are given below in minutes.

- Company X : 3.1, 2.1, 3.1, 3.2, 3.0, 2.6, 3.8, 2.7, 2.5, 3.5
- Company Y : 2.2, 3.1, 3.0, 2.5, 3.0, 3.2, 3.2, 2.7, 2.5, 2.5

- a Calculate the mean wait time for each company.

Company X : minutes

Company Y : minutes

- b Calculate the standard deviation for the wait times for company X .

Round your answers to two decimal places.

- c Calculate the standard deviation for the wait times for company Y .

Round your answers to two decimal places.

- d Which company generally has better response times?

Company X

A

Company Y

B

- e Which company has more consistent response times?

Company X

A

Company Y

B

4. The life of two brands of batteries are tested using a sample of 10 batteries from each brand. Their battery lives (in hours) are shown below.

- Brand X : 23.3, 19.7, 20.7, 25.3, 22.5, 19.1, 20.0, 20.7, 20.7, 20.9
- Brand Y : 23.2, 27.5, 25.0, 24.5, 22.7, 29.8, 22.9, 26.0, 26.4, 22.6

- a Calculate the mean battery life of each brand.

Brand X : hours

Brand Y : hours

- b Calculate the standard deviation for the battery life of Brand X .

Round your answer to one decimal place.

- c Calculate the standard deviation for the battery life of Brand Y .

Round your answer to one decimal place.

- d Which brand produces batteries that generally last longer?

A Brand X B Brand Y

- e Which brand produces batteries that are more consistent?

A Brand X B Brand Y

5. Two friends compete in triple jump and the distance of 20 jumps were recorded. The mean and standard deviation for the jumps are shown below.

Jumper	Mean (m)	Standard deviation (m)
Fred	12.4	0.8
Tracy	11.6	0.5

- a What does a comparison of the mean of the two friends tell us?

Tracy generally jumps further. A Tracy jumps more consistently. B

Fred generally jumps further. C Fred jumps more consistently. D

b What does a comparison of the standard deviation of the two friends tell us?

Tracy jumps more consistently. A Fred jumps more consistently. B

Fred generally jumps further. C Tracy generally jumps further. D

6. Two friends compete in 100 m sprints and the time to complete 50 sprints were recorded. The mean and standard deviation for the sprints are shown below.

Runner	Mean (s)	Standard deviation (s)
Quentin	13.2	1.2
Hannah	14.6	0.75

a What does a comparison of the mean of the two friends tell us?

Quentin's sprint times are more variable. A Hannah's sprint times are more variable. B

Quentin generally runs faster. C Hannah generally runs faster. D

- b What does a comparison of the standard deviation of the two friends tell us?

Quentin generally runs faster. A Hannah generally runs faster. B

Quentin's sprint times are more variable. C Hannah's sprint times are more variable. D

7. Consider the following pair of data sets below.

- Set A: 2, 4, 5, 8, 8, 9, 9, 9, 10, 10, 10, 11, 12, 15
- Set B: 2, 2, 3, 4, 5, 5, 6, 7, 9, 11, 13, 13, 14, 15

- a Calculate the range of each data set.

Set A:

Set B:

- b Calculate the interquartile range of each data set.

Set A:

Set B:

- c Calculate the standard deviation for set A.

Round your answers to one decimal place.

- d Calculate the standard deviation for set B.

Round your answers to one decimal place.

- e Which data set has more variability?

Set A

A

Set B

B

- f Is range a useful measure to compare variability for these two sets?

Yes

A

No

B

8. Use the statistics mode on your calculator to answer this question.

The scores obtained by two classes are given below:

Red Class: 55, 57, 49, 58, 68, 57, 60, 53, 56, 51

Blue Class: 53, 59, 52, 49, 59, 49, 58, 57, 48, 54

- a Enter the mean of the scores obtained by the red class, correct to two decimal places.
- b Enter the standard deviation of the scores obtained by the red class, correct to two decimal places.
- c Enter the mean of the scores obtained by the blue class, correct to two decimal places.
- d Enter the standard deviation of the scores obtained by the blue class, correct to two decimal places.

- e Which class performed better on average?

Red Class

A

Blue Class

B

- f Which class produced more consistent results?

Red Class

A

Blue Class

B

9. The following table shows the heart rate data of a group of people after exercise.

Height of step	Stepping rate	Heart rate
Short step	Slow	92
Short step	Slow	96
Short step	Medium	105
Short step	Medium	106
Short step	Fast	121
Short step	Fast	125
Tall step	Slow	101
Tall step	Slow	103
Tall step	Medium	119
Tall step	Medium	124
Tall step	Fast	127
Tall step	Fast	130

- a Complete the table. Give all answers to one decimal place.

Stepping rate

Height of
step

Data

Slow

Medium

Fast

Stepping rate				
Short step	Average of heart rate	94.0	105.5	123.0
	Standard deviation of heart rate	2.0	0.5	2.0
Tall step	Average of heart rate			
	Standard deviation of heart rate			

- b Which of the following combinations of step height and stepping rate generated the higher heart rate?

a tall step at a slow stepping rate

A

a short step at a fast stepping rate

B

- c Which of the following combinations of step height and stepping rate showed the least variability?

a tall step at a fast stepping rate

A

a tall step at a slow stepping rate

B

a short step at a medium stepping rate

C

a short step at a fast stepping rate

D

10. Quiana grows two different types of bean plants. She records the number of beans that she picks from each plant for 10 days. Her records are shown below:

Plant A: 1, 9, 10, 4, 9, 6, 7, 7, 9, 5

Plant B: 7, 7, 5, 6, 4, 7, 8, 5, 5, 5

- a Which plant produces more beans on average?

Plant A

A

Plant B

B

- b Which plant has a more consistent yield of beans?

Plant A

A

Plant B

B

11. Kathleen is training for the 100-meter freestyle event. Over a week, she records her best times (in seconds) each day:

$$51.7, 52.7, 51.8, 51.5, 52.6, 52.2, 52.1$$

On the following Monday, she swims the 100-meter in 55 seconds.

- a Will her time on Monday increase or decrease the mean?

Increase

A

Decrease

B

- b Will her time on Monday increase or decrease the MAD (mean absolute deviation)?

Increase

A

Decrease

B

12. Mr. Maximilian's art classes are working on sketching, and he is tracking the time (in minutes) it takes for students in two of his Year 10 classes to complete their sketches. The completion times for the students in his classes are shown in the table:

Class 1	43	47	52	49	55	46	51	54	45	44	48	53
Class 2	49	43	46	47	40	52	41	50	44	45	48	42

- a Calculate the MAD for each class, rounded to three decimal places.

Class 1 =

Class 2 =

- b Is it reasonable to compare the variability of these two classes?

No, because the times are in different units.

A

Yes, because the times are exactly the same.

B

Yes, because both classes have similar tasks and data can be compared.

C

No, because the students are from different classes.

D

- c Which class of students completed their sketches with less variability in time?

Class 1

A

Class 2

B

C.T 12.06C Quartiles (29Q'S)

1. Find the median of the following data set:

10, 11, 14, 16, 17, 23, 24

2. Find the median of the following data set:

22, 23, 31, 36, 41, 51, 45, 43

3. Find the median of the following data set:

28, 30, 33, 39, 43, 53

4. Find the median of the following data set:

44, 46, 48, 53, 53, 54, 63, 64

5. Consider the data set shown below:

0, 1, 3, **5**, 7, 9, 12

Median

a Find the lower quartile.

b Find the upper quartile.

6. Consider the data set shown below:

24, 25, 27, 29, 30, 33

↑
Median: 28

a Find the lower quartile.

b Find the upper quartile.

7. Consider the data set shown below:

11, 12, 16, 17, **20**, 23, 24, 27, 27
Median

- a Find the lower quartile.
 - b Find the upper quartile.
8. Consider the data set shown below:

12, 15 ,17, 17, 19, 20, 21, 23, 24, 27, 31, 32
↑
Median: 20.5

- a Find the lower quartile.
 - b Find the upper quartile.
9. Determine which term is defined by each description:

- a Lowest value

Maximum

A

Median

B

Minimum

C

Lower quartile

D

Upper quartile

E

- b 50% of the data lies on either side of this value

Lower quartile

A

Median

B

Minimum

C

Maximum

D

Upper quartile

E

- c At most 25% of the data are above this value

Median

A

Lower quartile

B

Upper quartile

C

Minimum

D

Maximum

E

- d Highest value

Maximum

A

Minimum

B

Lower quartile

C

Upper quartile

D

Median

E

- e At most 25% of the data are below this value

Maximum

A

Upper quartile

B

Minimum	C	Median	D
Lower quartile	E		

10. Here are Lucy's scores from her last 13 rounds of golf played:

64, 64, 68, 68, 70, 78, 84, 86, 116, 120, 128, 132, 134

- a Find her median score.

- b Find the lower quartile.

- c Find the upper quartile.

- d Find the interquartile range.

11. Here are Roald's scores from his last 17 exams:

38, 42, 48, 51, 56, 62, 64, 70, 74, 76, 80, 82, 87, 88, 96, 98, 100

- a Find his median score.

- b Find the lower quartile.

- c Find the upper quartile.

d Find the interquartile range.

12. Answer the following, given this set of scores:

8, 20, 19, 4, 15, 14, 10

a Sort the scores in ascending order.

b Find the number of scores.

c Find the median.

d Find the lower quartile of the set of scores.

e Find the upper quartile of the set of scores.

f Find the interquartile range.

13. Below is the luggage weight of 30 passengers.

Weight	Frequency
--------	-----------

16	4
----	---

17	5
----	---

18	5
----	---

19	3
----	---

Weight	Frequency
20	4
21	6
22	3

- a What is the mean check in weight?

Give your answer correct to two decimal places.

- b Determine the:

Median: kilograms

Lower Quartile: kilograms

Upper Quartile: kilograms

- c In which quartile does the mean lie?

3

A

1

B

4

C

2

D

14. A sample of boxes of matches were selected for quality control and the number of matches in each box recorded in the following frequency table.

Matches	Frequency
45	1
46	2
47	5

Matches	Frequency
48	3
49	7
50	25
51	6
52	2

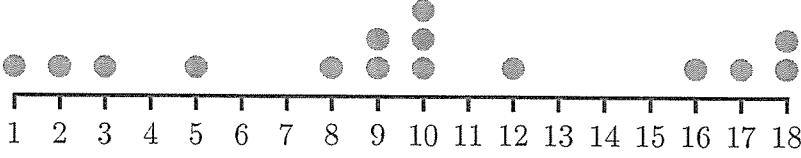
a How many matchboxes were sampled?

b Find the range of the data.

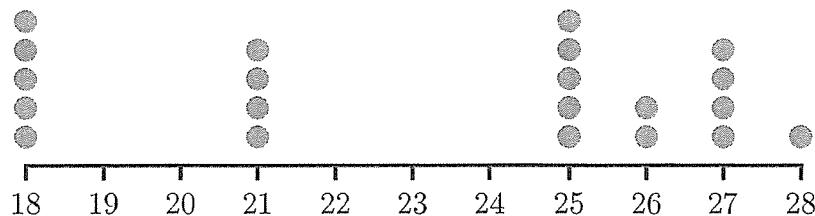
c Find the interquartile range.

15. A group of students were asked how many phone calls they had made the previous day. The information was collected in the following frequency table.

Phone calls	Frequency
0	8
1	5
2	10
3	6
4	6
5	7
6	3
7	2

- a How many students were surveyed?
- b Find the range of the data.
- c Find the interquartile range.
16. Consider the dot plot below and then answer the following questions.
- 
- | Age | Frequency |
|-----|-----------|
| 1 | 1 |
| 2 | 1 |
| 3 | 1 |
| 5 | 1 |
| 8 | 1 |
| 9 | 3 |
| 10 | 4 |
| 12 | 1 |
| 16 | 1 |
| 17 | 2 |
| 18 | 2 |
- a Find the lower quartile of the set of scores.
- b Find the upper quartile of the set of scores.
- c Find the interquartile range.
- d What is the range of this data set?
17. The dot plot shows the ages of customers in a mobile phone store in one day.

Age of customers



- a Find the lower quartile of the set of scores.
- b Find the upper quartile of the set of scores.
- c Find the interquartile range.
18. The stem plot shows the number of hours students spent studying during an entire semester.

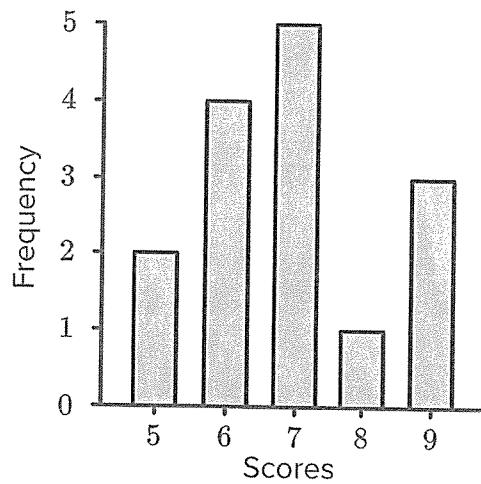
Stem	Leaf
6	2 7
7	1 2 2 4 7 9
8	0 1 2 5 7
9	0 1

Key: 5 | 2 = 52

- a Find the lower quartile of the set of scores.
- b Find the upper quartile of the set of scores.

- c Find the interquartile range.

19. For this bar chart:



- a Organise the data into the frequency table.

Score (x)

Freq (f)

5

6

7

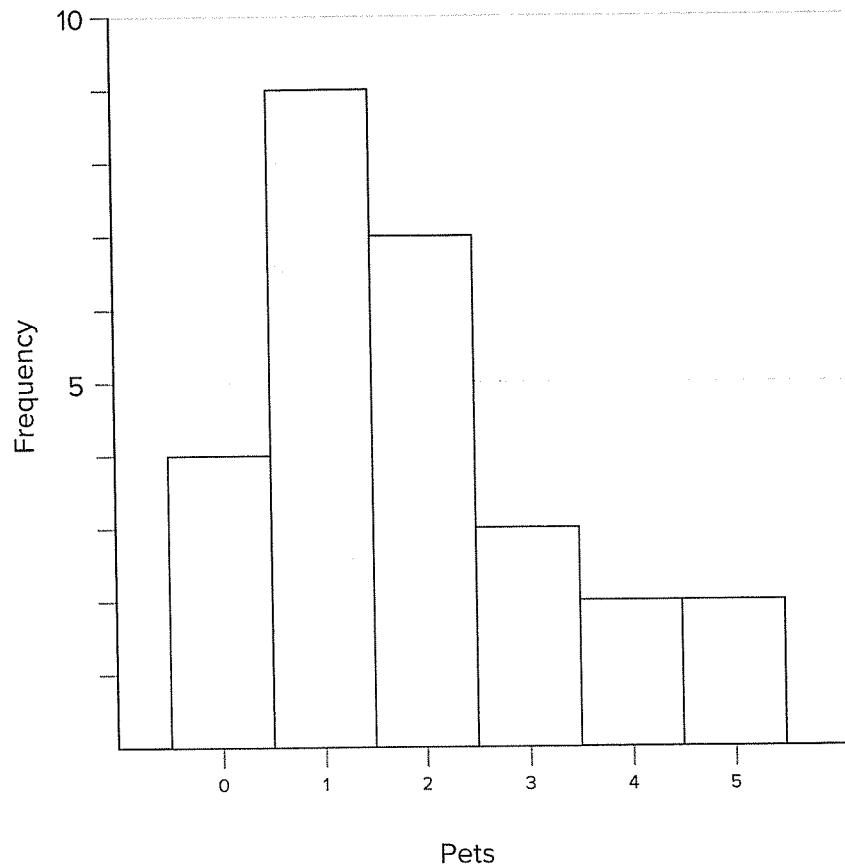
8

9

Totals

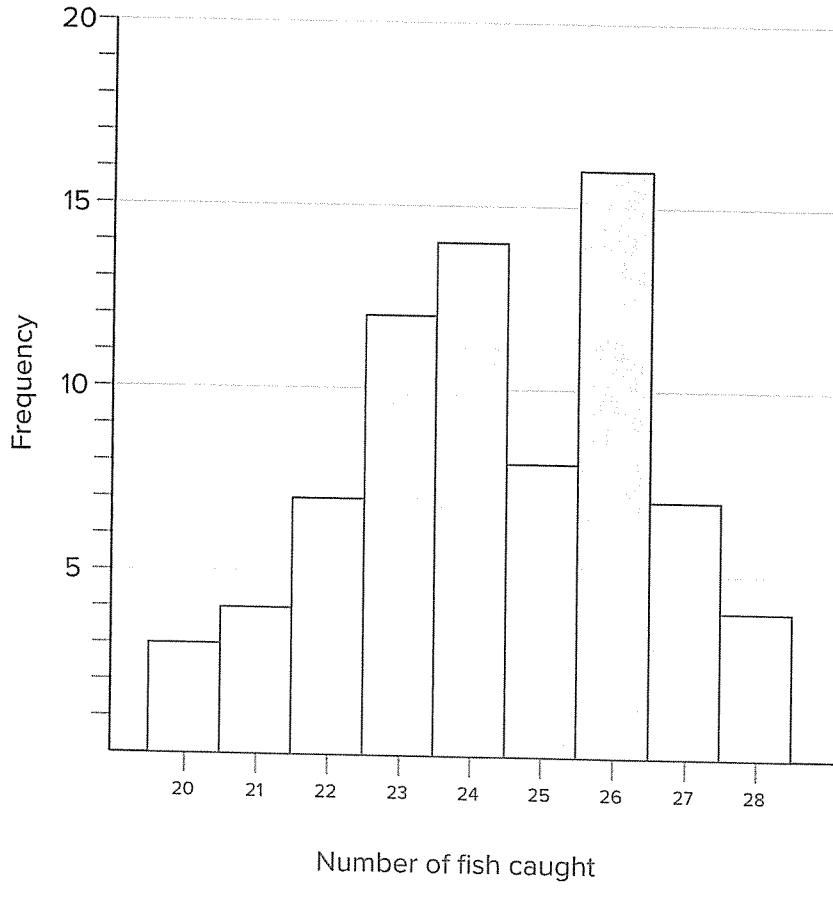
- b Find the median score.
- c Find the lower quartile score.
- d Find the upper quartile score.
- e Find the interquartile range.

20. There are 27 students in a class. The column graph shows the number of pets that each student in the class owns.



- a Find the lower quartile of the set of scores.
- b Find the upper quartile of the set of scores.
- c Find the interquartile range.

21. Dylan records the number of fish he catches for each fishing trip over a period of time.



- a How many fishing trips did he go on?

- b Find the range in the number of fish caught.
- c Find the interquartile range.
- 22.** The table shows the number of points scored by a basketball team in each game of their previous season.
- 75, 86, 90, 80, 70, 79, 65, 93, 62, 82, 79, 81, 80, 59, 73
- a Find the median value.
- b Find the lower quartile.
- c Find the upper quartile.
- d In what percentage of games did the team score more than the upper quartile?
- e What percentage of games did the team score less in than the lower quartile?
- 23.** The stem-and-leaf plot shows the batting scores of two cricket teams, the Ruby Raptors and the Golden Griffins.

Ruby Raptors	Stem	Golden Griffins
9 7 6	6	3 4 8
7 6 3 0	7	2 6 7
1 0	8	1 2 6 8

Ruby Raptors Stem Golden Griffins

9 0 9

Key: 6 | 1 | 2 = 12 and 16

- a Determine the median score for each team.

Ruby Raptors median =

Golden Griffins median =

- b Determine the range for each team.

Ruby Raptors range =

Golden Griffins range =

- c Find the interquartile range for the Ruby Raptors.

- d Find the interquartile range for the Golden Griffins.

- e Determine which team has a more symmetrical distribution.

Golden Griffins

A

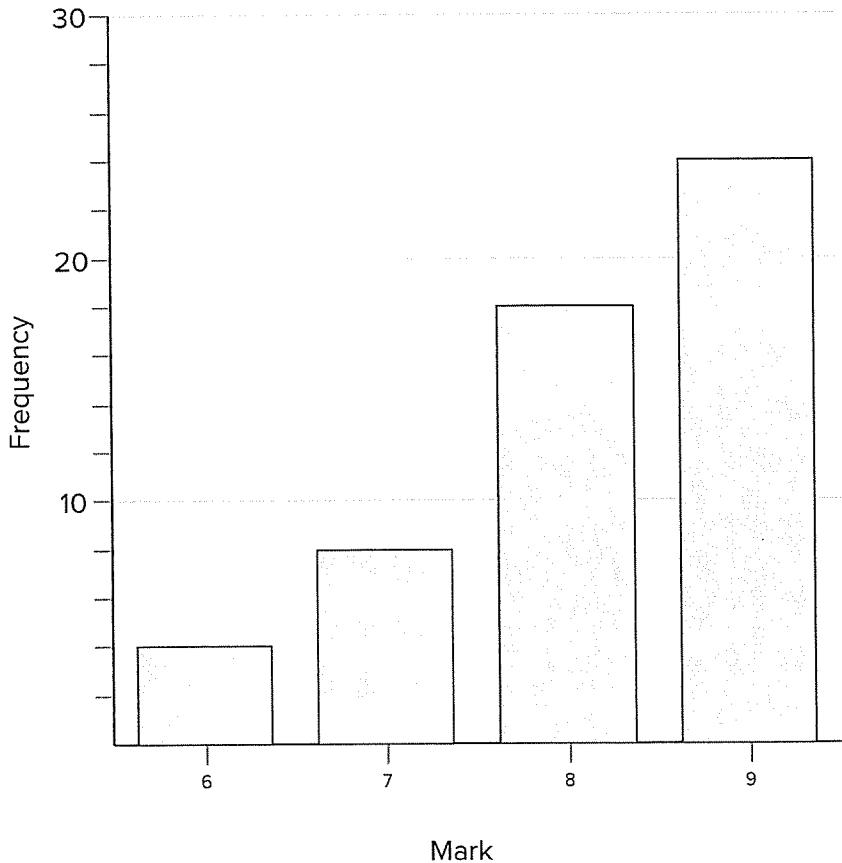
Ruby Raptors

B

Both Ruby Raptors and Golden Griffins

C

24. The bar graph shows the marks (out of 10) that students received on a spelling test.



- a Find the lower quartile of the set of scores.

- b Find the upper quartile of the set of scores.

- c Find the interquartile range.

25. These data sets each contain 10 scores:

Set 1: 13, 14, 17, 18, 20, 22, 24, 25, 28, 29
Set 2: 13, 14, 17, 18, 20, 22, 40, 45, 48, 54
Set 3: 13, 41, 42, 43, 47, 50, 52, 53, 59, 60

- a Determine which data set has the largest interquartile range.

Set 1

A

Set 2

B

Set 3

C

- b Identify the type of distribution for Set 1.

Skewed-left

A

Skewed-right

B

Symmetrical

C

- c Identify the type of distribution for Set 2.

Symmetrical

A

Skewed-right

B

Skewed-left

C

- d Identify the type of distribution for Set 3.

Skewed-left

A

Symmetrical

B

Skewed-right

C

26. There is a test to measure the Emotional Quotient (EQ) of an individual. Here are the EQ results for 21 people listed in ascending order.

93, 94, 94, 95, 95, 96, 96, 98, 101, 104, 105, 106, 111, 112, 113, 118, 118,



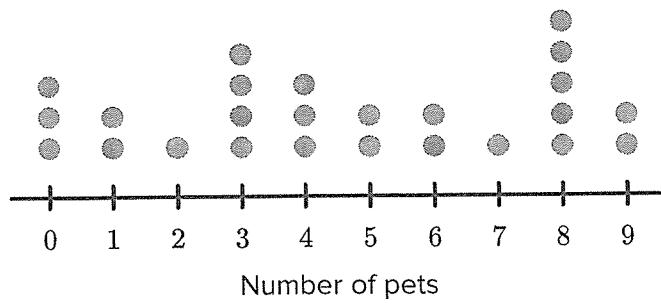
- a Calculate the range.

- b Calculate the interquartile range.

- c Determine the median EQ score.

- d Calculate the mean, rounded to two decimal places.

27. The dot plot shows the number of pets that 25 students have.



- a Find the lower quartile.

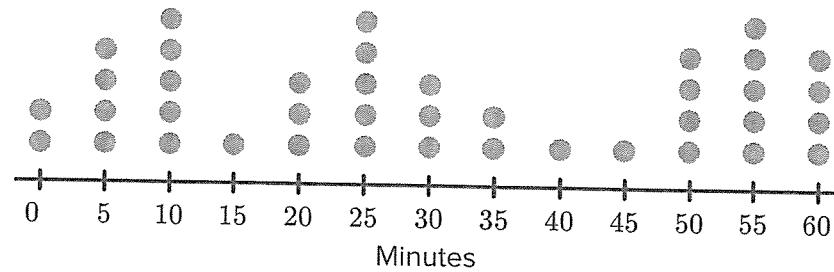
- b Find the upper quartile.

- c Find the proportion of students who have more pets than the upper quartile.

28. In a competition, a contestant must complete 12 challenges earning as many points as possible. Her scores for the first 11 challenges are:
46, 48, 49, 60, 78, 79, 84, 91, 94, 95, 114

Determine her score, x , in the 12th round if the lower quartile of all of her 12 scores is 54.

29. The airline *Flo Air* decided to keep track of flight delay times (the number of minutes after the scheduled time when the plane takes off) over a week. The 40 results are shown in the dot plot.



- a Determine the median delay time of the flights, in minutes.

- b Determine the upper quartile for the data.

- c Determine the lower quartile for the data.

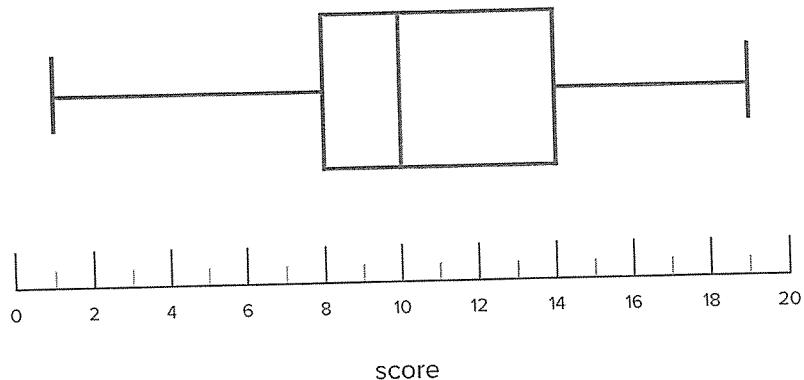
- d Determine the interquartile range.

- e If a flight is delayed for 10 minutes or more, the airline incurs a fee. According to the given dot plot, for what percentage of flights did the airline incur a fee? Round the percentage to one decimal place.

- f A rival airline, *Fly Air*, had a mean delay time during the same week of 50 minutes.
What percentage of *Flo Air*'s flights had delay times that were longer than *Fly Air*'s mean delay time?

C.T 12.07C Box plots (25Q'S)

1. For the box plot shown below, find each of the following:



Complete the table for the given data:

Minimum

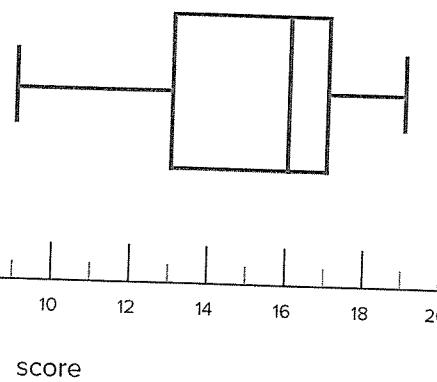
Lower Quartile

Median

Upper Quartile

Maximum

2. For the box plot shown below, find each of the following:



Complete the table for the given data:

Minimum

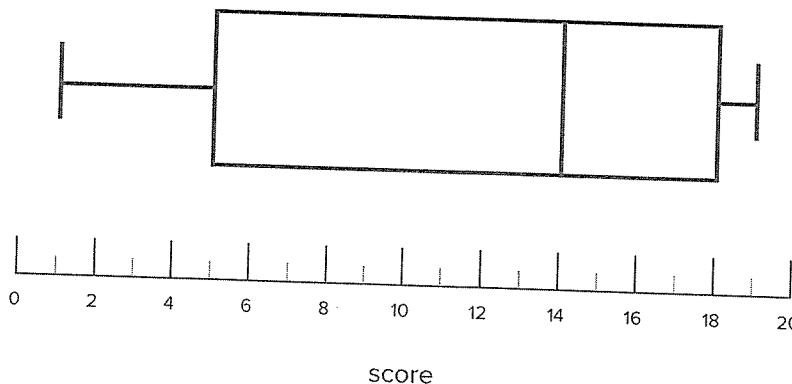
Lower Quartile

Median

Upper Quartile

Maximum

3. For the box plot shown below, find each of the following:



Complete the table for the given data:

Minimum

Lower Quartile

Median

Upper Quartile

Maximum

4. Using the information in the table, create a box plot to represent this data:

Minimum	5
Lower Quartile	25
Median	40
Upper Quartile	55
Maximum	65

5. A geography teacher has marked a set of tests. She wants to represent the results in a box plot. She has already sorted her data and created the table shown. Adjust the box plot to match the data in the table.

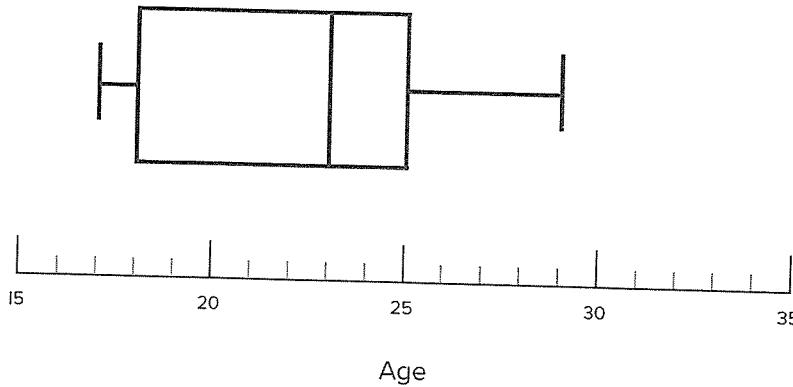
Minimum	4
Lower Quartile	10
Median	22
Upper Quartile	26

Maximum	38
---------	----

6. Create a box plot to represent the data in the table below.

Minimum	10
Lower quartile	20
Median	40
Upper quartile	55
Maximum	75

7. The box plot below shows the age at which a group of people got their driving licences.



- a What is the oldest age at which someone got their licence?

- b What is the youngest age at which someone got their licence?

c What percentage of people were aged from 18 to 23?

10%

A

50%

B

25%

C

d The middle 50% of responders were within how many years of one another?

6

A

8

B

9

C

7

D

e In which quartile are the ages least spread out?

2nd

A

3rd

B

4th

C

1st

D

f The bottom 50% of responders were within how many years of one another?

6

A

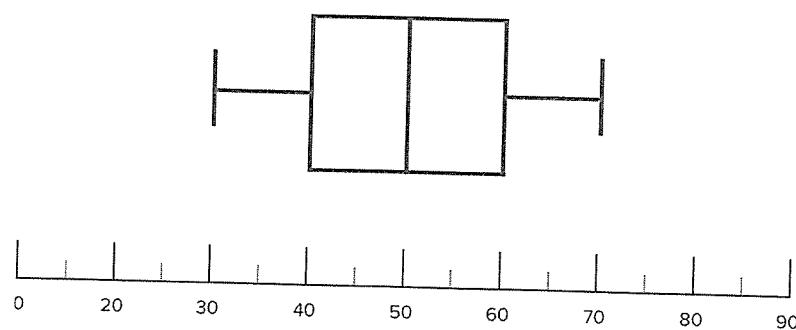
5

B

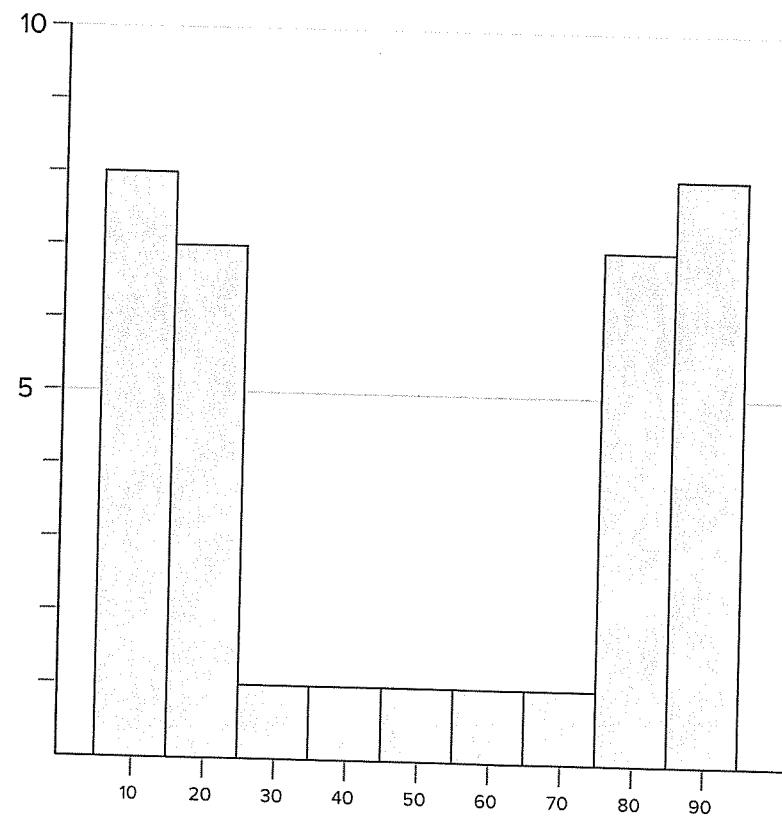
7

C

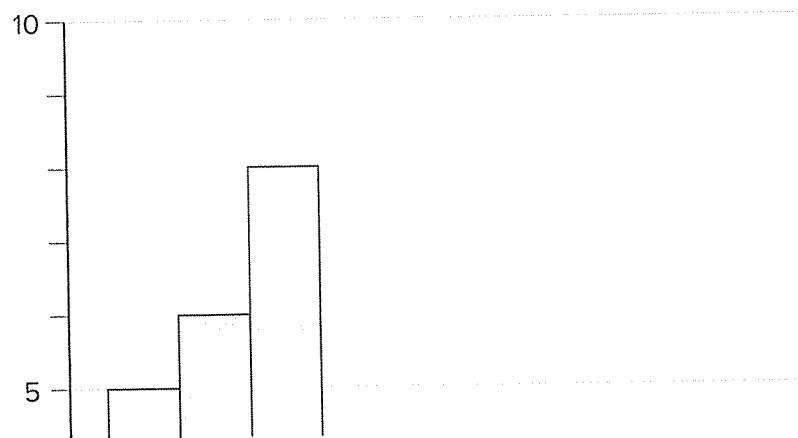
8. Match the box plot shown to the correct column graph.



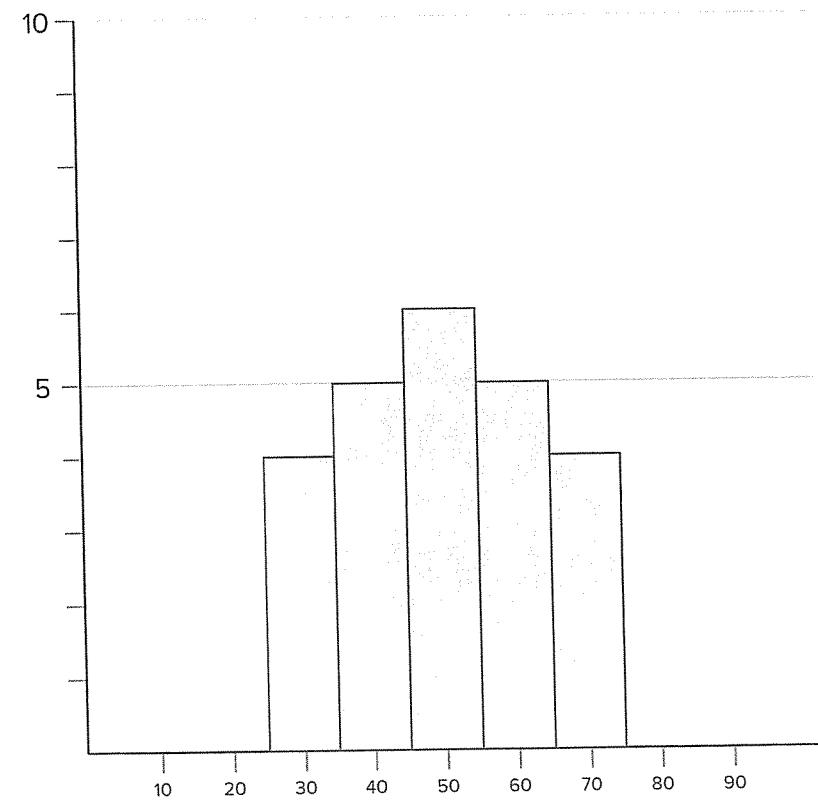
A



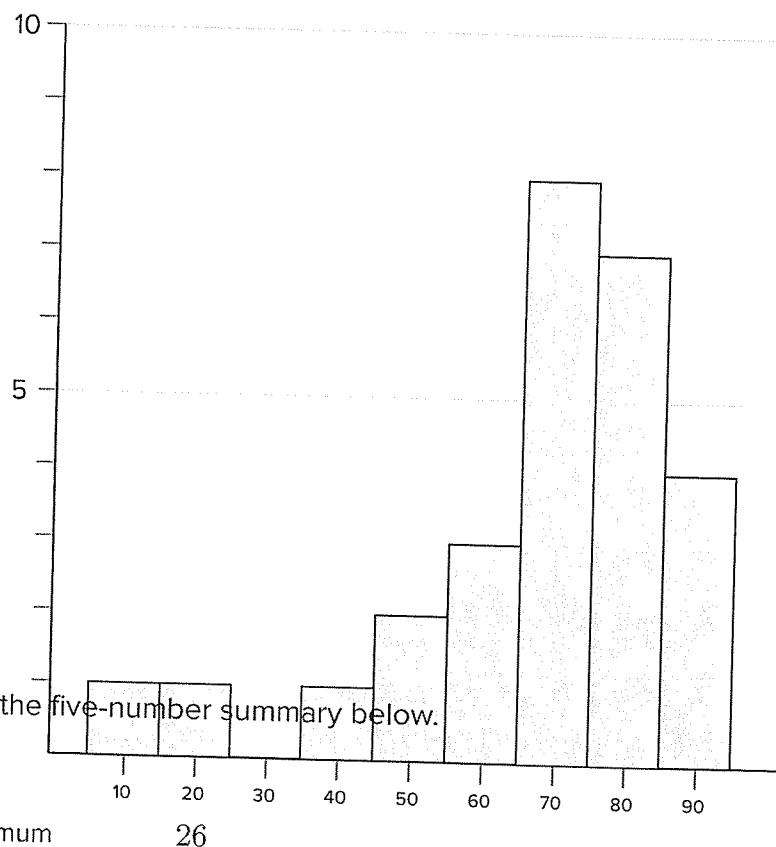
B



C



D



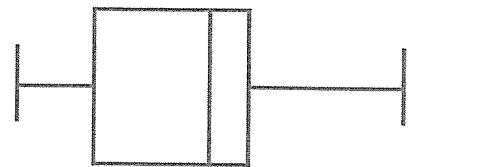
9. Consider the five-number summary below.

Minimum	26
Lower quartile	34
Median	38
Upper quartile	44
Maximum	45

a Find the range.

b Find the interquartile range.

10. Determine if each statement true or false for the data set shown in the boxplot:



- a The maximum is 18.

True

A

False

B

- b The minimum is 7.

True

A

False

B

- c The median is 13.

True

A

False

B

- d There is an error as the median is not exactly in the middle of Q_1 and Q_3 .

True

A

False

B

- e The interquartile range is 4.

True

A

False

B

- f If an additional value of 1 was added to the data set, it would be an outlier.

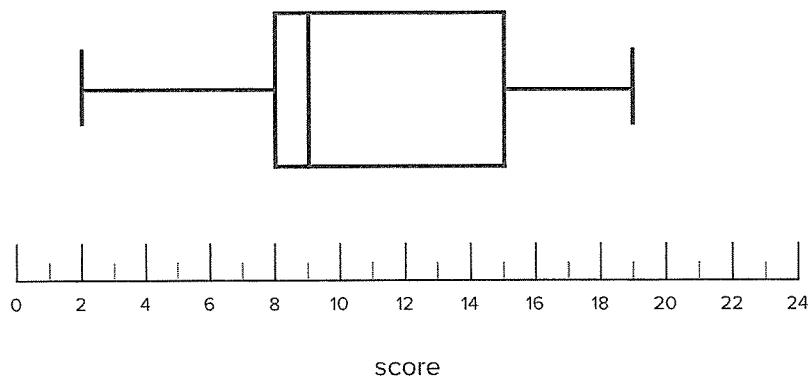
True

A

False

B

11. Consider the box plot shown below.



- a** Approximately what percentage of scores lie between each of these values?

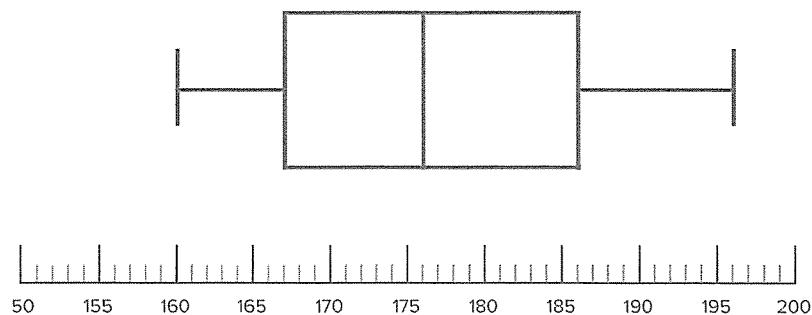
Q_1 and Q_3 %

Minimum and Q_1 %

Median and Maximum %

Q_1 and Maximum %

Minimum and Q_3 %



a What is the range of the cyclist's heart rates and what does it represent?

The cyclist's heart rates varied by 10 beats per minute (bpm).

A

The cyclist's heart rates varied by 20 beats per minute (bpm).

B

The cyclist's heart rates varied by 29 beats per minute (bpm).

C

The cyclist's heart rates varied by 36 beats per minute (bpm).

D

b What is the interquartile range of the cyclist's heart rates and what does it represent?

The middle half of the cyclist's heart rates varied by 19 beats per minute.

A

The middle half of the cyclist's heart rates varied by 10 beats per minute.

B

The middle half of the cyclist's heart rates varied by 26 beats per minute.

C

The middle half of the cyclist's heart rates varied by 10 beats per minute.

D

15. Consider the summary statistics below.

	Mean	Median	IQR	Range
Data set A	10	14	11	15

	Mean	Median	IQR	Range
Data set B	15	15	13	16

a What is the shape of the data set A?

Symmetrical

A

Skewed left

B

Skewed right

C

b What is the shape of the data set B?

Symmetrical

A

Skewed right

B

Skewed left

C

c Which data set has the larger measure of centre?

A

A

B

B

d Which data set has the larger measure of spread?

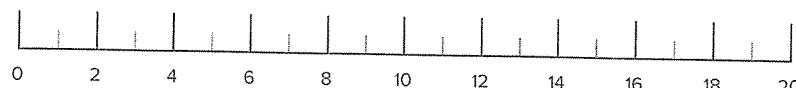
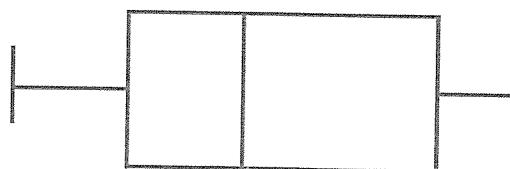
A

A

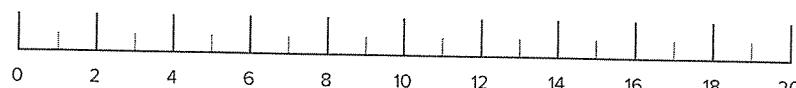
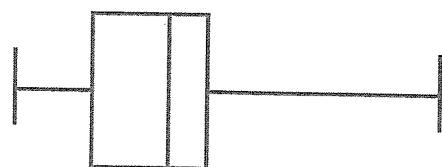
B

B

16. A mathematics test is given to two classes. The marks out of 20 received by students in each class are represented in the box plots below.



Class 9P



Class 9Q

- a Complete the following table using the two box plots:

Class 9P Class 9Q

Median

Lower Quartile

Upper Quartile

Range

Interquartile Range

- b Which class tended to score better marks?

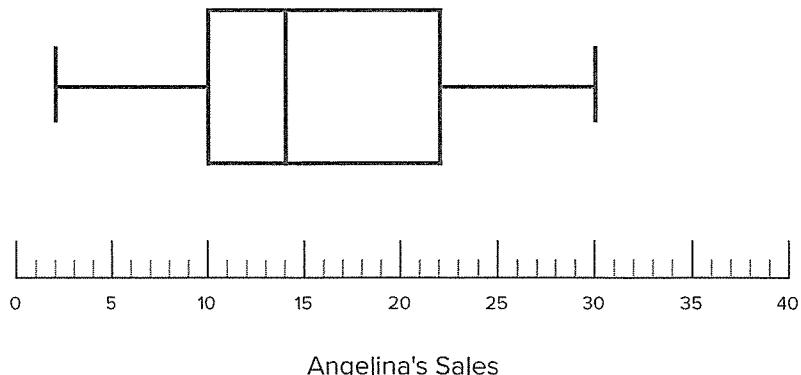
Class 9P

A

Class 9Q

B

17. The box plots below represent the daily sales made by Carl and Angelina over the course of one month.



- a What is the range in Angelina's sales?
- b What is the range in Carl's sales?
- c By how much did Carl's median sales exceed Angelina's?

- d Considering the middle 50% of sales for both sales people, whose sales were more consistent?

Angelina

A

Carl

B

- e Which salesperson had a more successful sales month?

Angelina

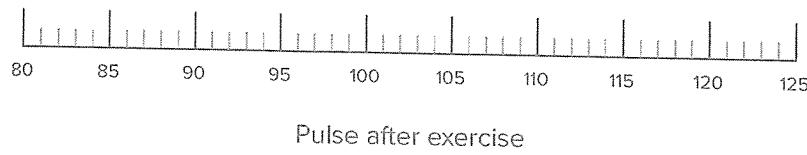
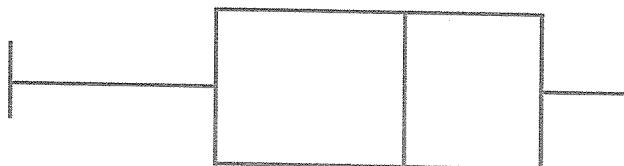
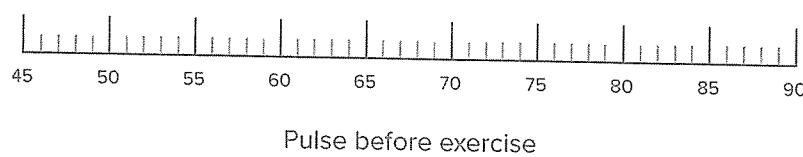
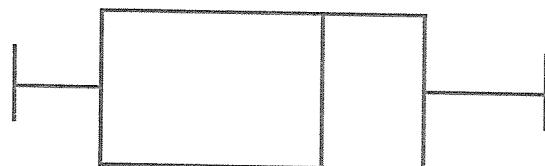
A

Carl

B

18. Ten participants had their pulse measured before and after exercise with results shown in different boxplots.

Which of the following best compares the participants' pulse rates before and after exercise?



B

The typical pulse rates increases from about 54 to about 115 and are spread out over a larger range after exercise.

A

The typical pulse rates increases from about 73 to about 96 and are spread out over a larger range after exercise.

D

The typical pulse rates increases from about 49 to about 120 and are spread out over a larger range after exercise.

C

The typical pulse rates increases from about 67 to about 107 and are spread out over a larger range after exercise.

19. The test scores of 11 students in Maths and Music are listed below.

Maths: 63, 59, 53, 82, 77, 81, 75, 67, 93, 73, 85

Music: 88, 69, 87, 82, 88, 73, 81, 79, 67, 97, 63

- a Construct a boxplot for the Maths scores.

- b Construct a boxplot for the Music scores.

- c In which subject did students perform better overall?

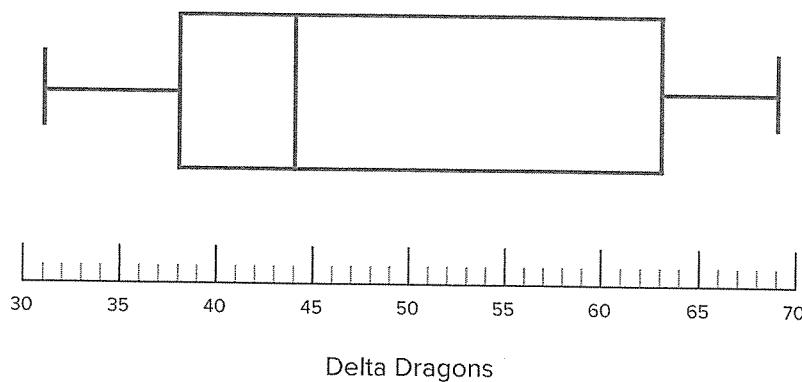
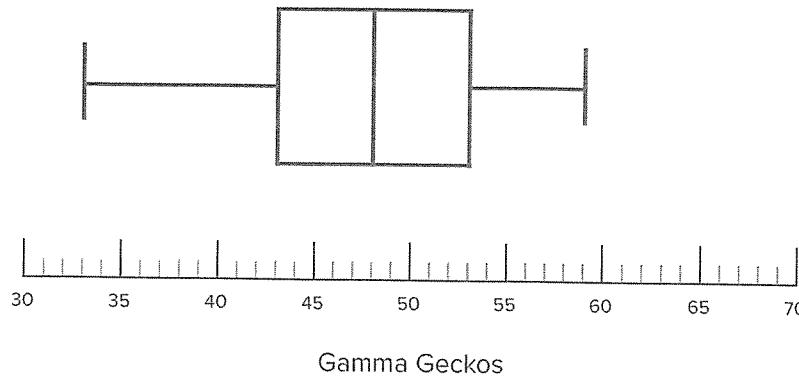
Music

A

Maths

B

20. These boxplots show the number of points scored by two basketball teams in each of their matches.



- a Compare the median score of the Gamma Geckos and the Delta Dragons.

The median score of the Gamma Geckos is 8 points lower than the median of the Delta Dragons.

A

The median score of the Gamma Geckos is 4 points lower than the median of the Delta Dragons.

B

The median score of the Gamma Geckos is 8 points higher than the median of the Delta Dragons.

C

The median score of the Gamma Geckos is 4 points higher than the median of the Delta Dragons.

D

- b Compare the range of scores for the Gamma Geckos and the Delta Dragons.

The range of the Delta Dragons is 24 points wider than the Gamma Geckos.

A

The range of the Gamma Geckos is 12 points wider than the Delta Dragons.

B

The range of the Delta Dragons is 12 points wider than the Gamma Geckos.

The range of the Gamma Geckos is 24 points wider than the Delta Dragons.

- c Compare the interquartile range for the Gamma Geckos and the Delta Dragons.

The interquartile range of the Gamma Geckos is 15 points wider than the interquartile range of the Delta Dragons.

A

The interquartile range of the Gamma Geckos is 30 points wider than the interquartile range of the Delta Dragons.

B

The interquartile range of the Delta Dragons is 15 points wider than the interquartile range of the Gamma Geckos.

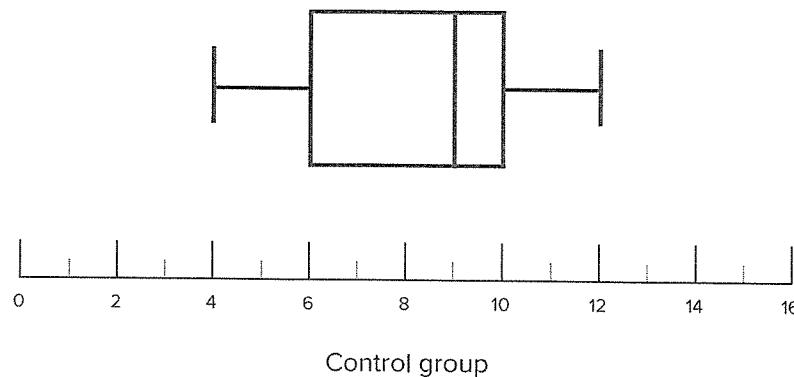
C

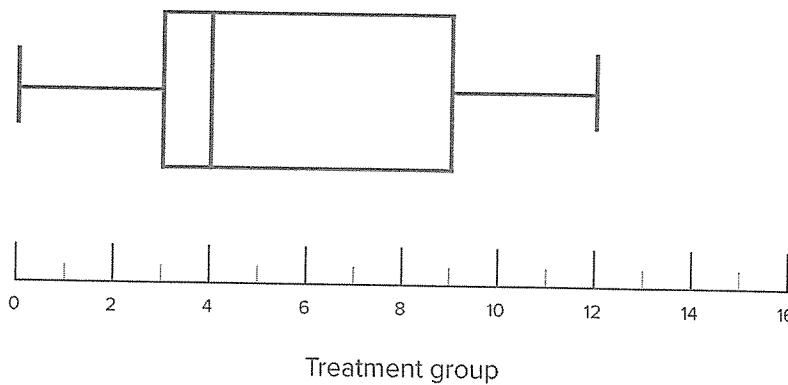
The interquartile range of the Delta Dragons is 30 points wider than the interquartile range of the Gamma Geckos.

D

21. The boxplots summarise results from a medical study. The treatment group received an experimental drug to relieve cold symptoms, and the control group received a placebo.

The boxplots show the number of days each group continued to report symptoms.





a Which of the following best describes the shape of the data from the control group?

Symmetrical

A

Positively skewed

B

Negatively skewed

C

b Which of the following best describes the shape of the data from the treatment group?

Negatively skewed

A

Positively skewed

B

Symmetrical

C

c Does the drug have a positive effect on patient recovery?

No, the treatment group had a higher median and most people in the treatment group had symptoms for more than 9 days.

A

No, both groups had the same median and the distribution of symptom duration was similar across the treatment and control groups.

B

Yes, both groups had the same median but the number of people with symptoms lasting

C

Yes, the treatment group had a lower median and most people had symptoms for less than 9

D

less than 9 days was about the same in both groups.

days, whereas the control group had half the people having symptoms for at least 9 days.

22. The marks in an end-of-year exam for a class of students are shown:
54, 58, 62, 64, 66, 78, 80, 82, 82, 84, 85, 88, 90, 92, 96, 97

- a Construct a boxplot for the data.
 - b Calculate the interquartile range.
 - c What percentage of marks lie in the range 89 to 97?

d Which values do the lowest 75% of scores lie between?

and

23. Construct a boxplot for the data set:

Range	15
Lowest score	2
<i>IQR</i>	8
25% of scores	Below 4

50% of scores Above 8

24. For each of the conditions, use five of the digits from 1 to 9 with no repetition to complete a five-number summary.

- a The smallest possible interquartile range with the largest possible range.

Minimum

Lower quartile

Median 3

Upper quartile

Maximum

- b The range is three times the length of the interquartile range.

Minimum 2

Lower quartile

Median

Upper quartile 5

Maximum

C.T 12.05C The shape of data (17Q'S)

1. A set of data is strongly positively skewed. If the median is 60, the mean is :

Less than 60

A

Equal to 60

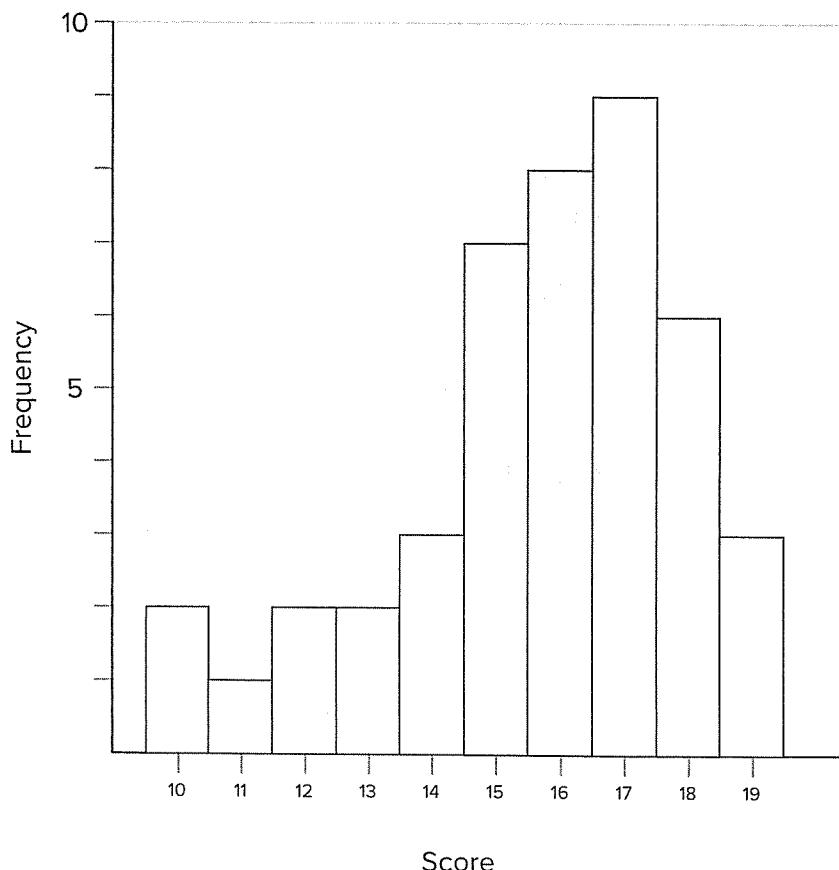
B

Greater than 60

C

2. State whether the scores in each histogram are positively skewed, negatively skewed or symmetrical (approximately).

a



Negatively skewed

A

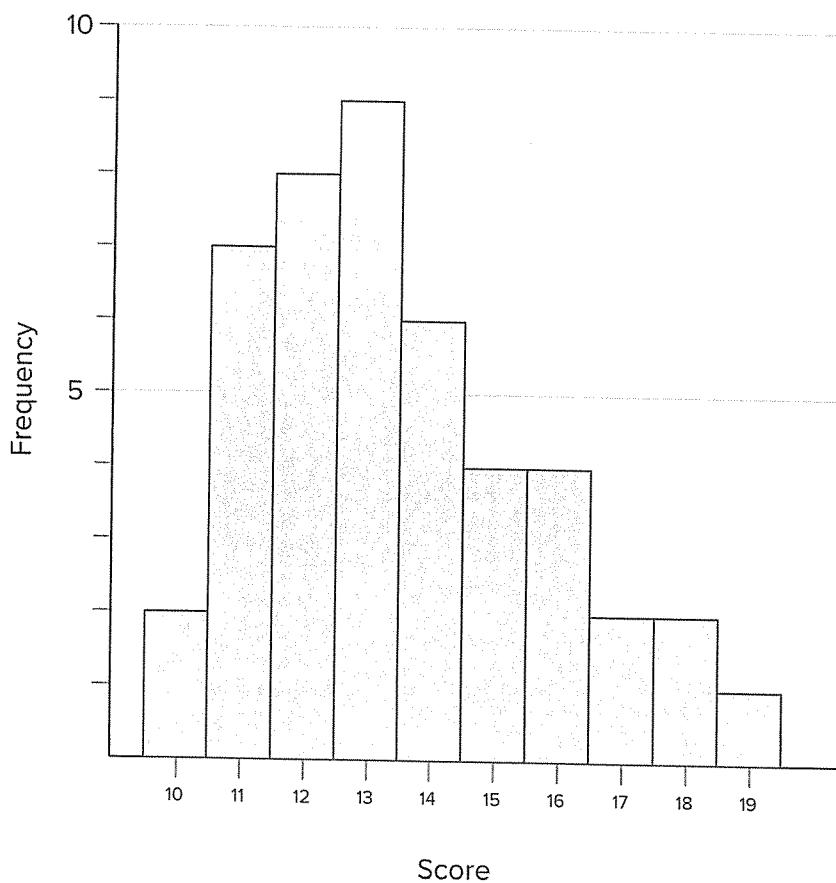
Positively skewed

B

Symmetrical

C

b



Positively skewed

A

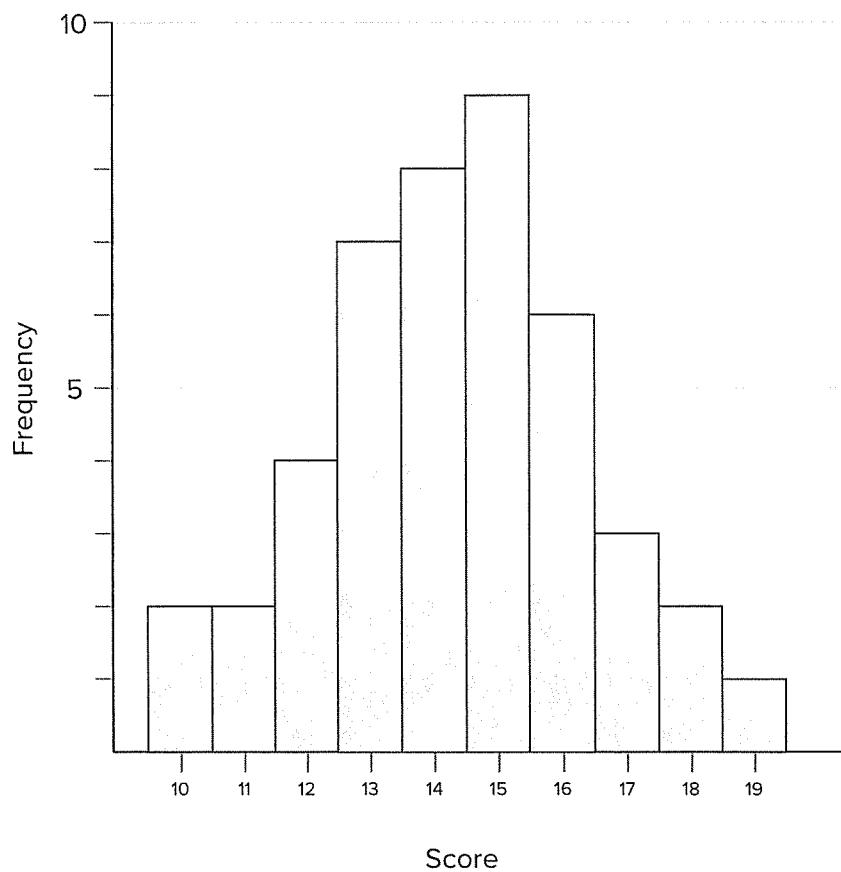
Symmetrical

B

Negatively skewed

C

C



Negatively skewed

A

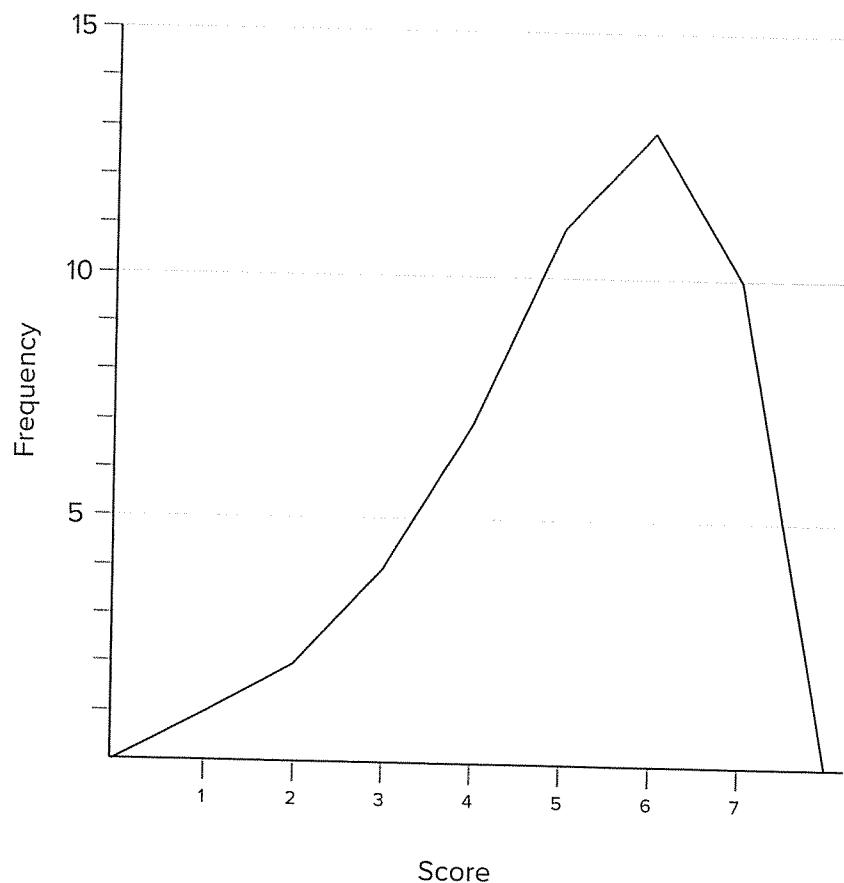
Symmetrical

B

Positively skewed

C

3. What type of skew does the graph show?



Positively skewed

A

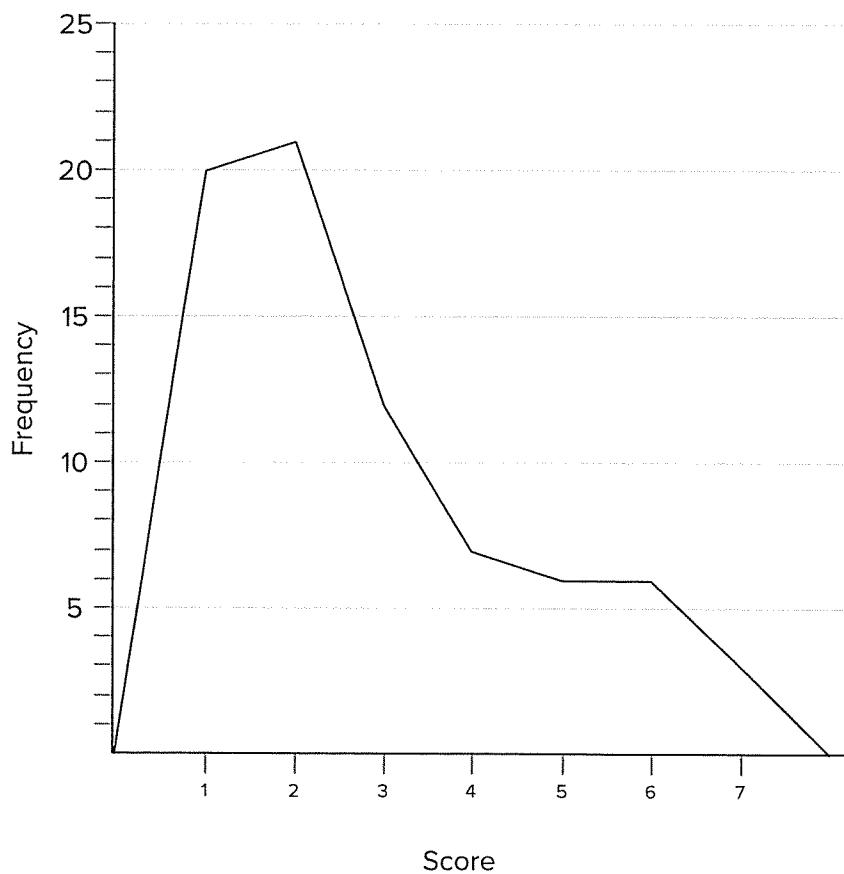
Negatively skewed

B

Symmetrical

C

4. What type of skew is shown in the graph below?



Positively skewed

A

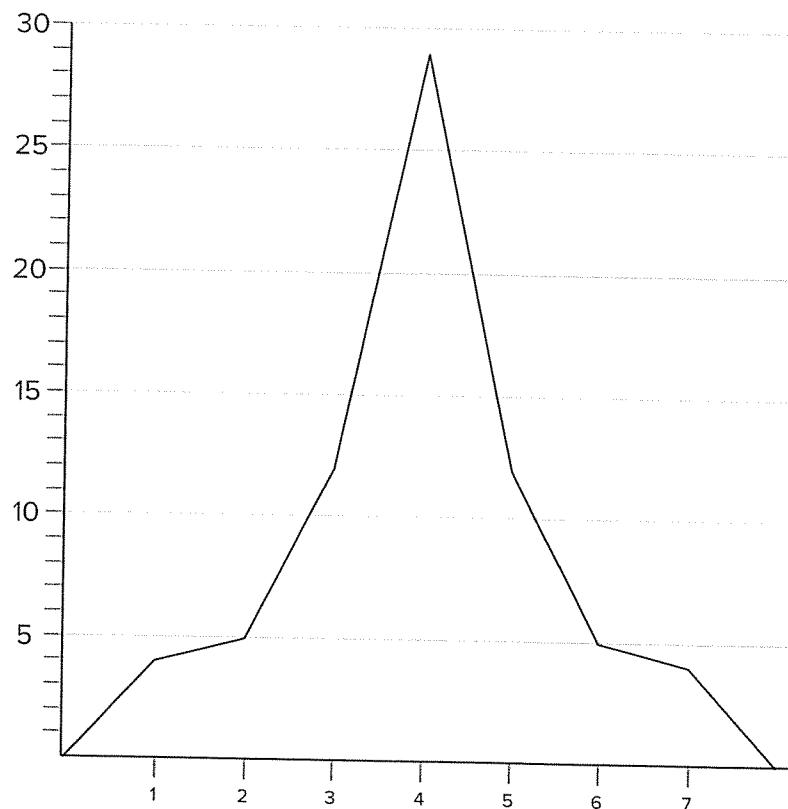
Negatively skewed

B

Symmetrical

C

5. Which one of the following could be used to describe the graph?



Negatively skewed

A

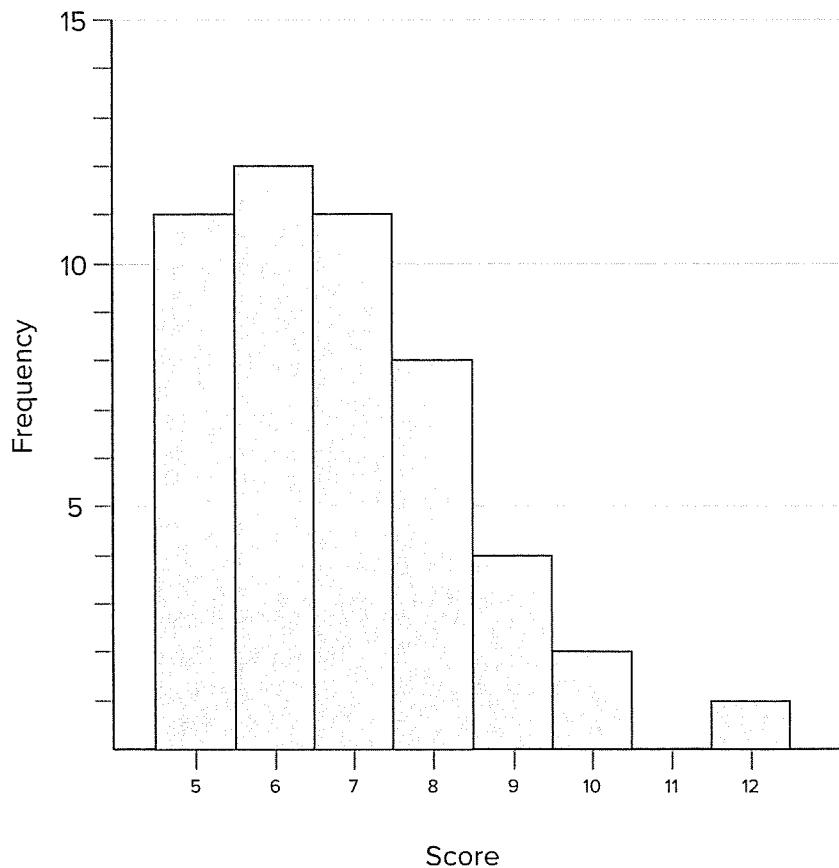
Symmetrical

B

Positively skewed

C

6. What type of skew is shown in the graph below?

Histogram

Negatively skewed

A

Positively skewed

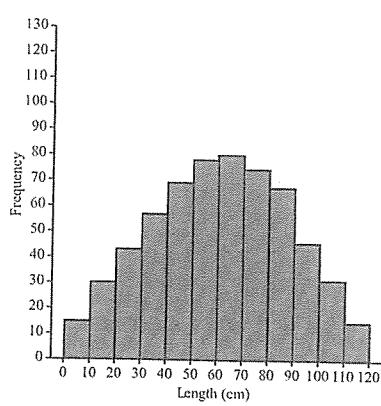
B

Symmetrical

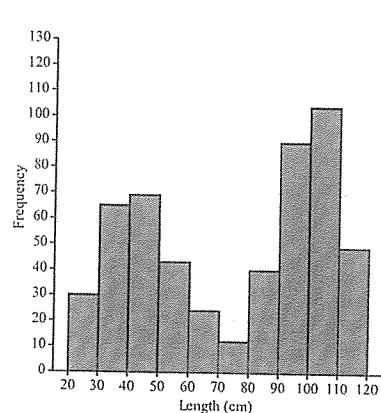
C

7. Which of these graphs shows a bimodal distribution?

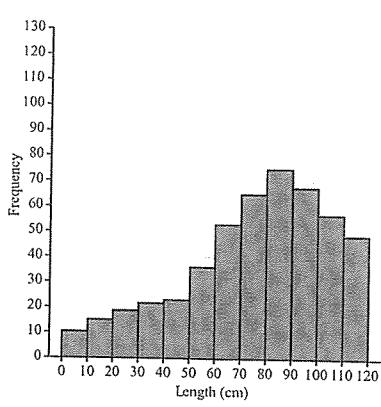
A



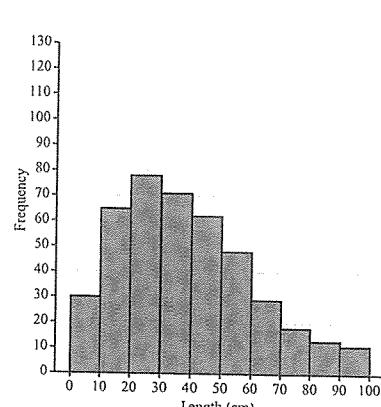
B



C



D



8. The stem and leaf plot below shows the age of people to enter through the gates of a concert in the first 5 seconds.

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

Key: 1 | 2 = 12

years old

- a What was the median age?

- b What was the difference between the lowest age and the median?

- c What is the difference between the highest age and the median?

- d What was the mean age? Round your answer to two decimal places if needed.

- e Is the data positively or negatively skewed?
Positively skewed A Negatively skewed B

9. For the stem-and-leaf plot attached:

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

Key: 1 | 2 = 12

a Is there any clustering of data?

Yes

A

No

B

b Where does the clustering occur?

0s – 10s

A

20s – 30s

B

30s – 40s

C

c What is the mode of the stem-and-leaf plot?

d What is the shape of the distribution?

Positively skewed

A

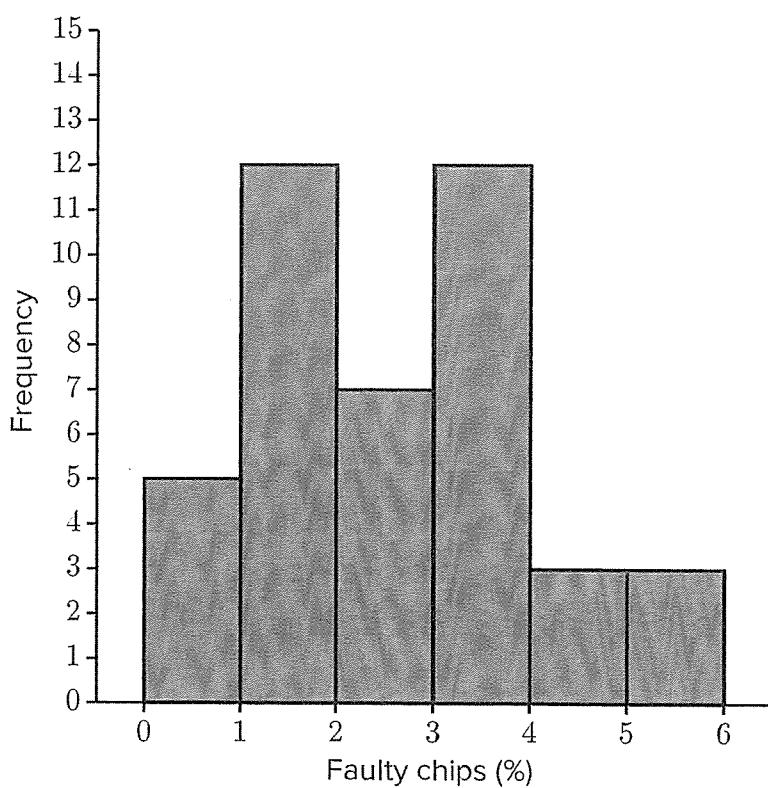
Symmetrical

B

Negatively skewed

C

10. The percentage of faulty computer chips in 42 batches were recorded in the histogram below.



- a Which of the following makes this statement true?

The distribution is:

Uni-modal

A

Bi-modal

B

Multi-modal, but not bi-modal C

- b Which of the following are the modal classes? Select all that apply.

0 – 1

A

1 – 2

B

2 – 3

C

3 – 4

D

4 – 5

E

5 – 6

F

11. The stem-and-leaf plot below displays the number of books borrowed by students from the school library over the past term.

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

Key: $2 | 5 = 25$

Determine the two regions where clustering of data occurs.

Around 70 – 79

A

Around 20 – 29

B

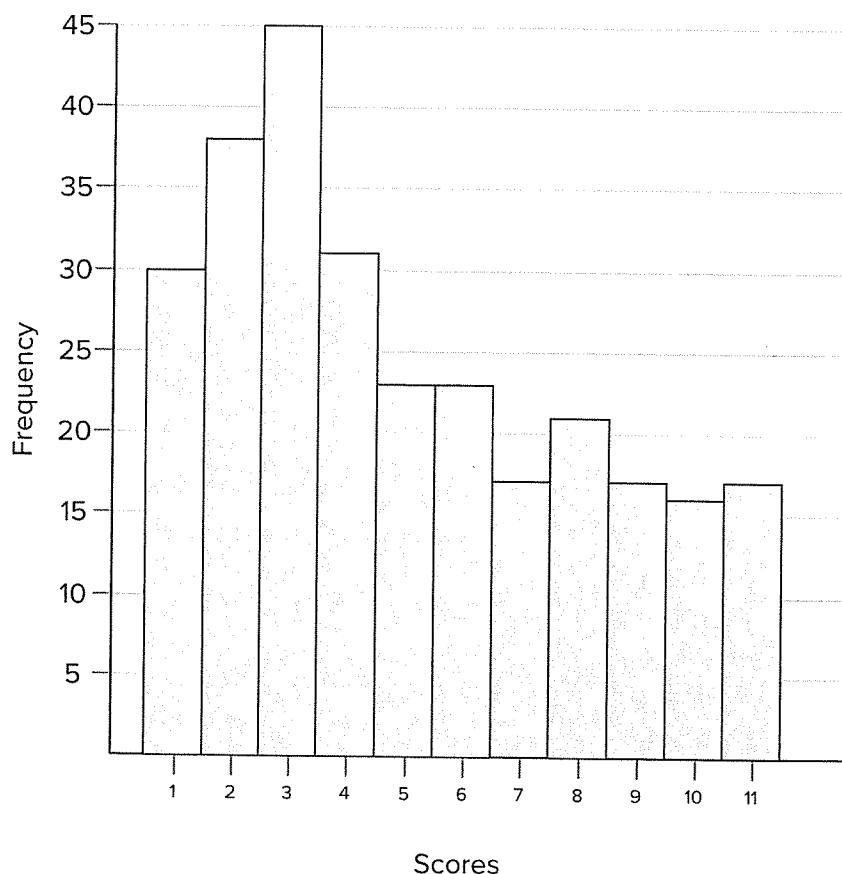
Around 50 – 59

C

Around 10 – 19

D

12. Consider the data shown in the column graph below.



a Are there any outliers?

Yes

A

No

B

b Is there any clustering of data?

Yes

A

No

B

c Where does the clustering occur?

8 to 11

A

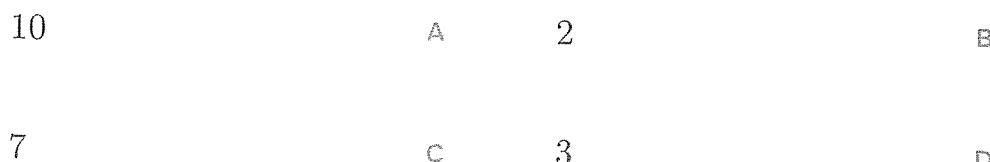
1 to 4

B

5 to 7

C

d What is the mode?



e The distribution of the data is:

Symmetrical.

A

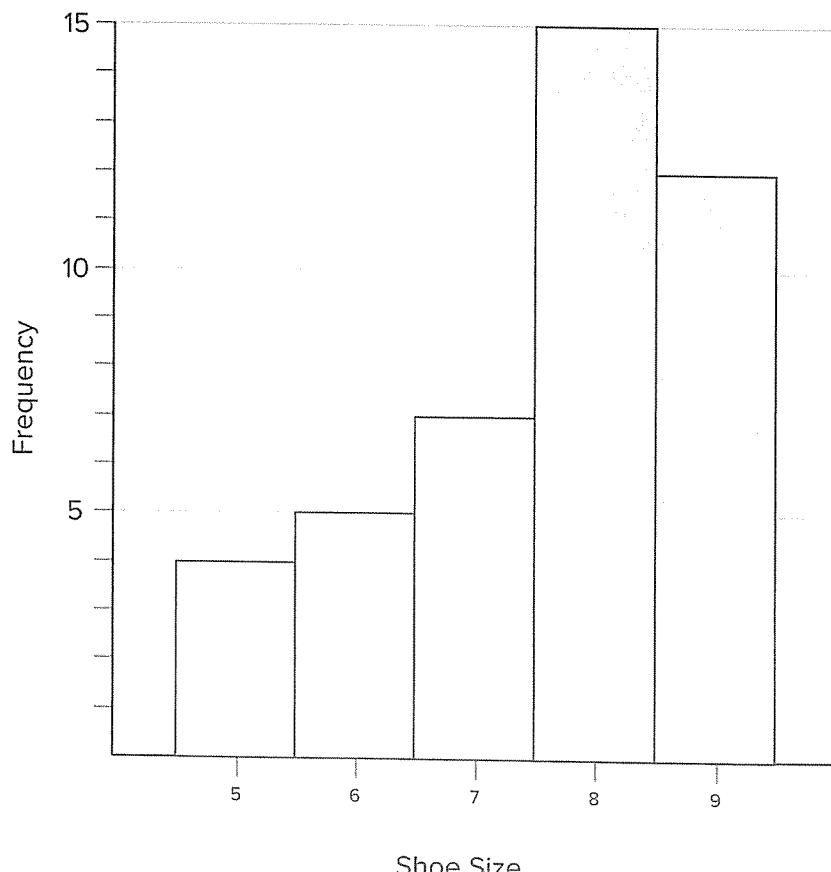
Negatively skewed.

B

Positively skewed.

C

13. The shoe sizes of all the students in a class were measured and the data was presented in a bar graph.



a Are there any outliers?

Yes

A

No

B

b Is there any clustering of data?

Yes

A

No

B

c Where does the clustering occur?

5 - 6

A

8 - 9

B

6 - 7

C

d What is the modal shoe size?

7

A

5

B

8

C

9

D

e The distribution of the data is:

Negatively skewed.

A

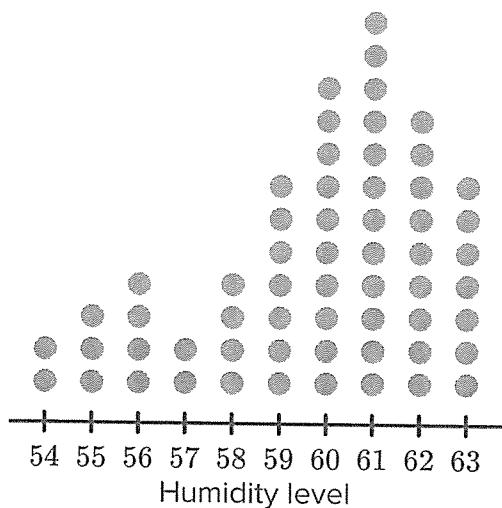
Positively skewed.

B

Symmetrical.

C

14. The dot plot shows the humidity levels recorded in a greenhouse over a period of time.



a Are there any outliers?

Yes

A

No

B

b Is there any clustering of data?

Yes

A

No

B

c Where does the clustering occur?

59 – 63

A

56 – 60

B

54 – 58

C

d What is the modal humidity level?

e The distribution of the data is:

Positively skewed

A

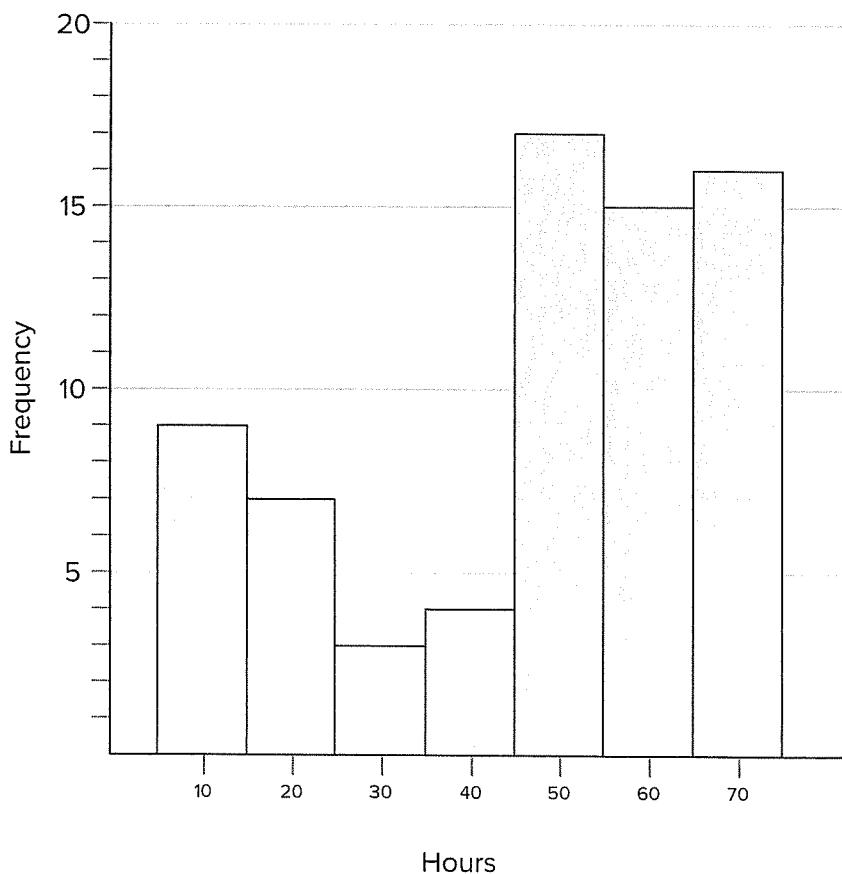
Symmetrical

B

Negatively skewed

C

15. The histogram represents the number of hours students have dedicated to community service in a given month:



- a Does the histogram most likely represent grouped data or individual scores?

Individual scores

A

Grouped data

B

b Estimate the value of the mean, rounded to one decimal place.

c What best describes the shape of the distribution?

Symmetrical

A

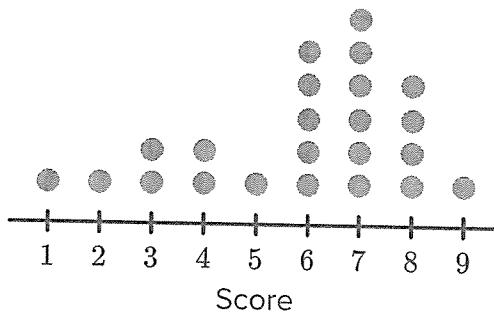
Negatively skewed

B

Positively skewed

C

16. What is the shape of the distribution in the dot plot below?



Symmetrical

A

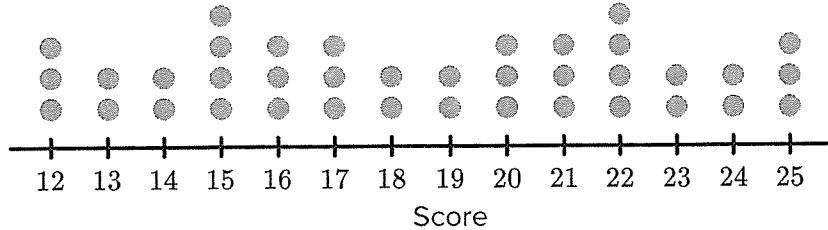
Negatively skewed

B

Positively skewed

C

17. For the dot plot:



a Is there any clustering of data?

Yes

A

No

B

b Find the modal scores, writing each value on the same line separated by commas.

c Describe the shape of the data.

Negatively skewed

A

Positively skewed

B

Symmetrical

C

Worksheet 15 08 The Spread of Data

Spread of data sets

1 Calculate the range of the following sets of data:

- a 22, 35, 20, 21, 40, 32
- c 30.1, 4.7, 14.6, 45.2, 11.3

- b 20, 17, 19, 18, 24, 21, 25, 22
- d 9.75, 45.5, 30.25, 17.5, 42.75, 37.5, 16.25

2 Calculate the interquartile range of the following sets of data:

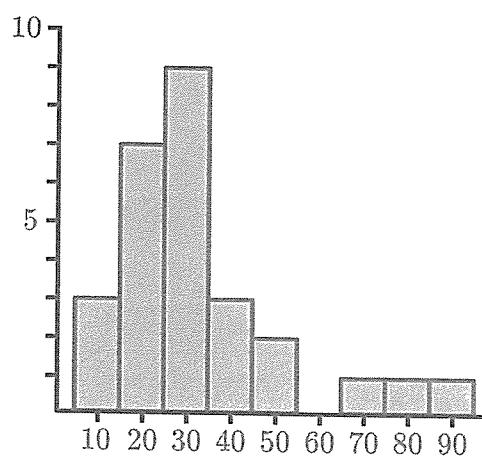
- a 21, 17, 16, 18, 22, 25, 15, 20
- c 31, 30, 25, 28, 37, 39, 20, 38, 35, 24

- b 21, 39, 23, 32, 22, 31, 33
- d 23.5, 38, 25.5, 29.5, 24, 35.5, 26.5, 20

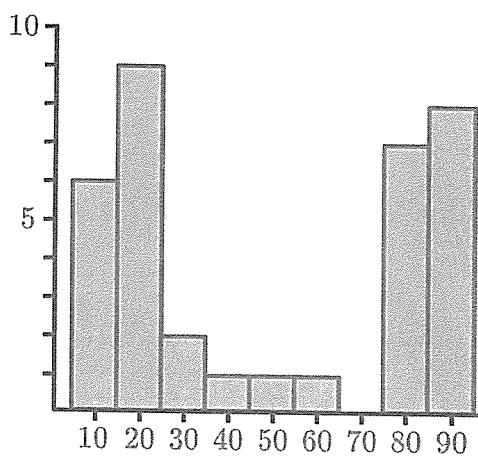
Match box plots to graphs

3 Construct a box plot for each of the following histograms:

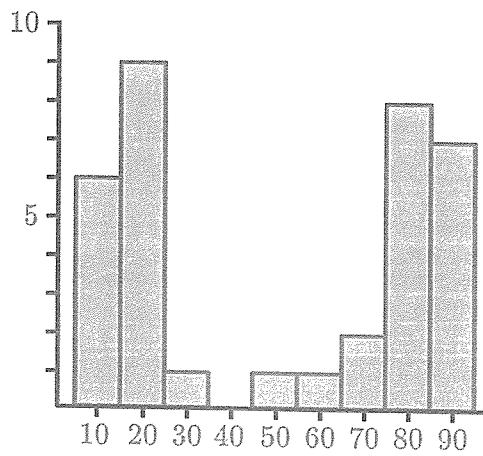
a



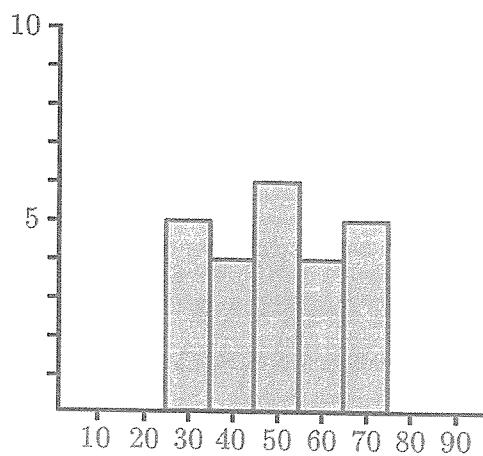
b

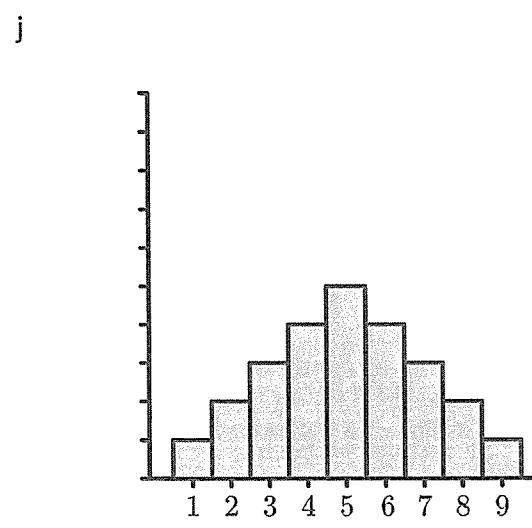
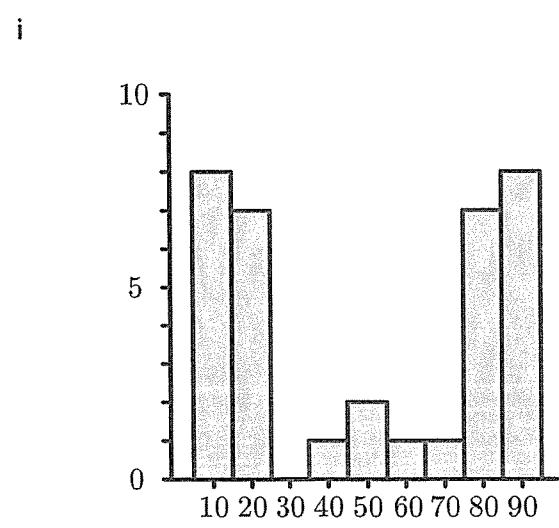
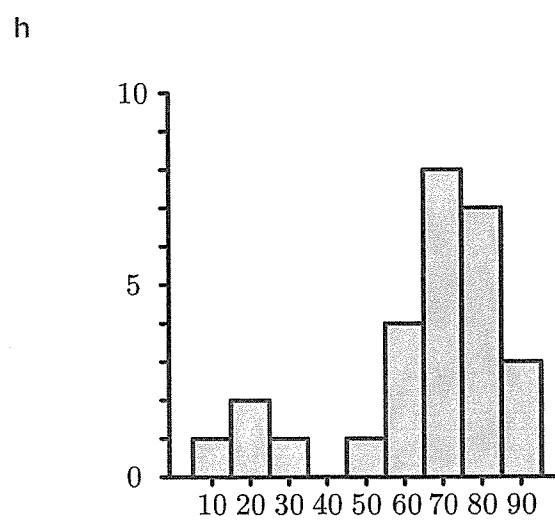
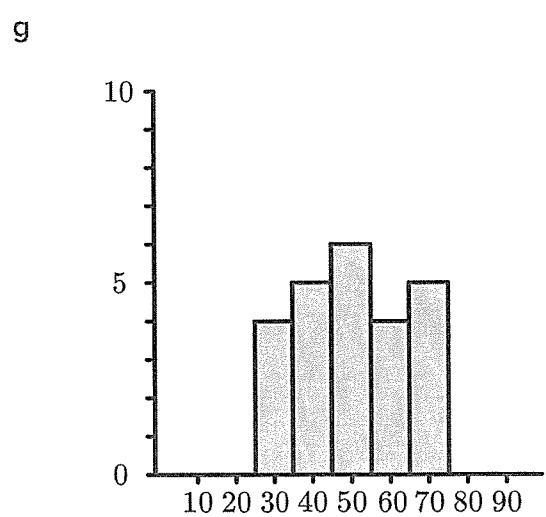
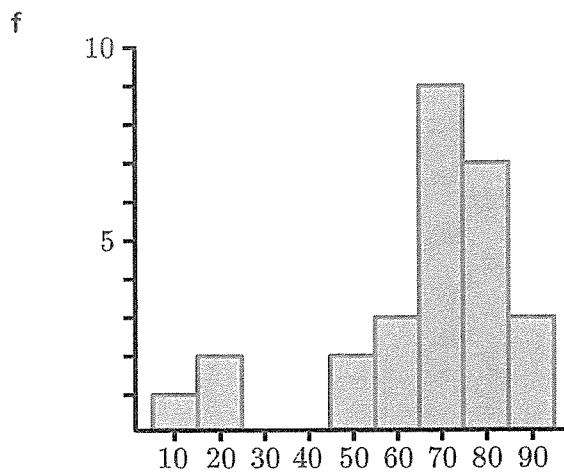
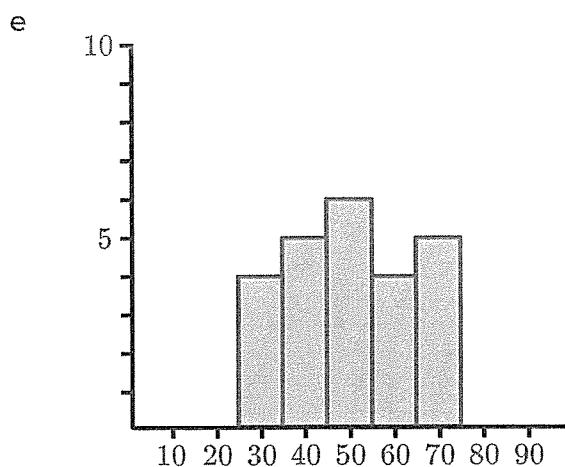


c



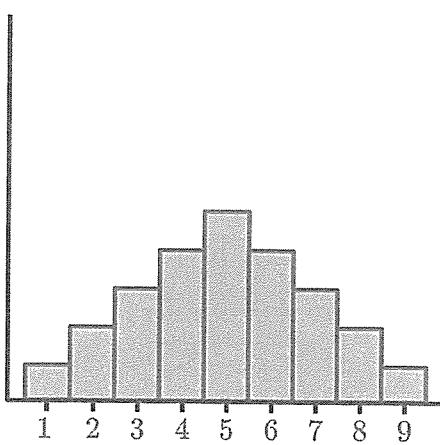
d



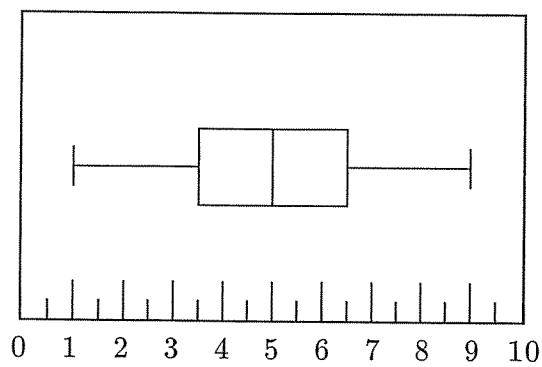


- 4 Match the histograms on the left to the corresponding box plots on the right:

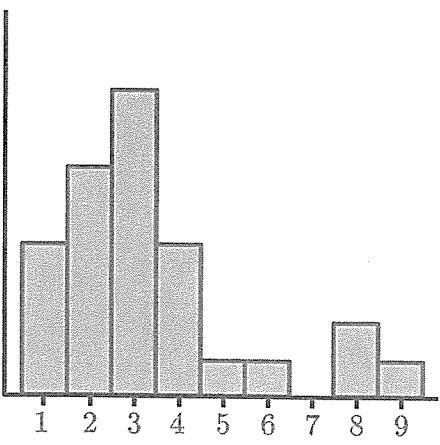
Histogram A



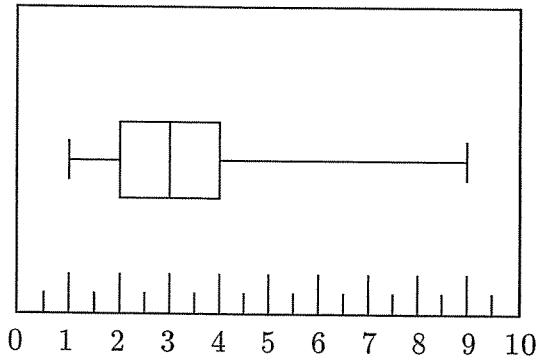
Box Plot 1



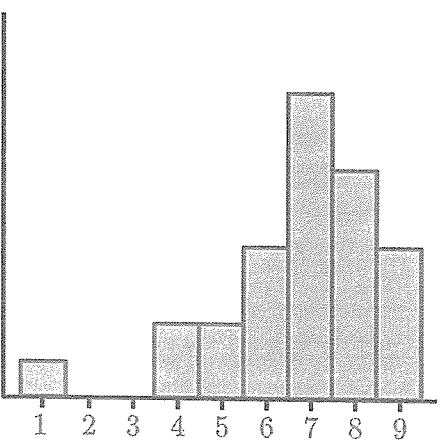
Histogram B



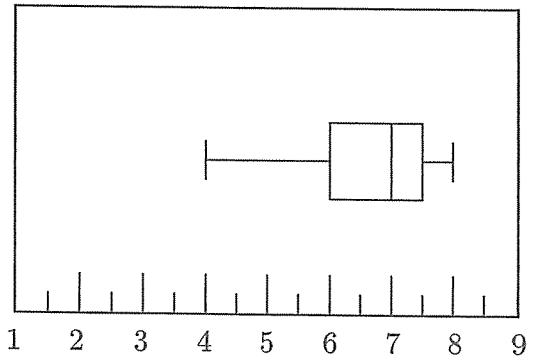
Box Plot 2

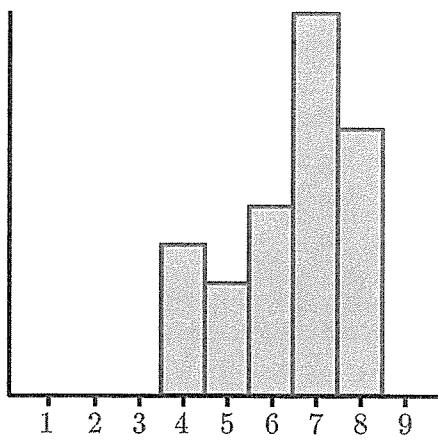
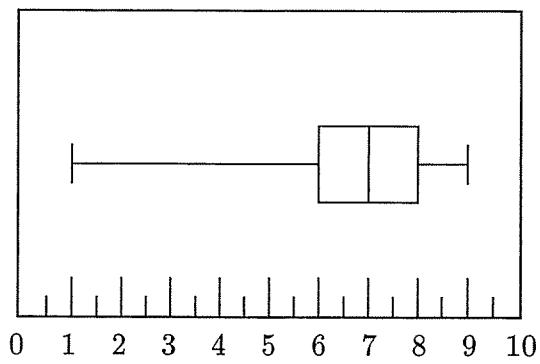


Histogram C



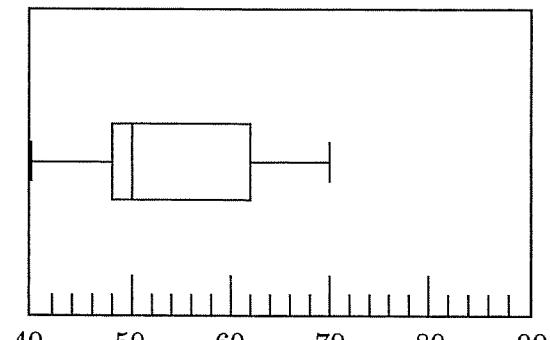
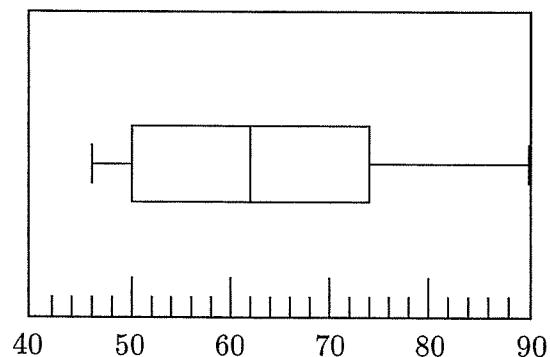
Box Plot 3



Histogram D**Box Plot 4**

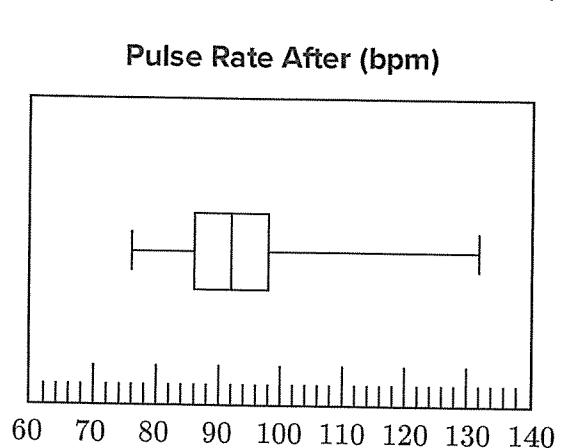
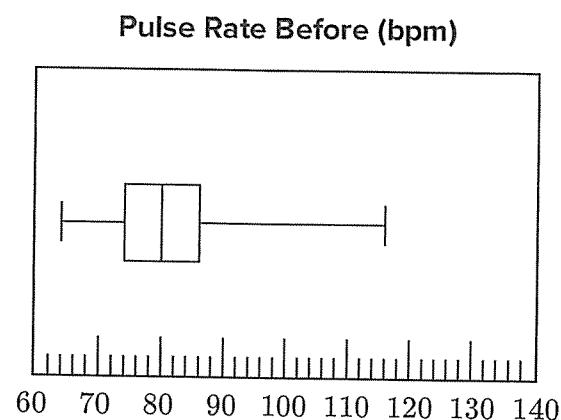
Parallel box plots

- 5 Two groups of people, athletes and non-athletes, had their resting heart rate measured. The results are displayed in the following pair of box plots:
- Find the median heart rate of athletes.
 - Find the median heart rate of the non-athletes.
 - Using this measure, which group has the lower heart rates?
 - Find the interquartile range of the athletes' heart rates.
 - Find the interquartile range of the non-athletes' heart rates.
 - Using this measure, which group has more consistent heart rate measures?

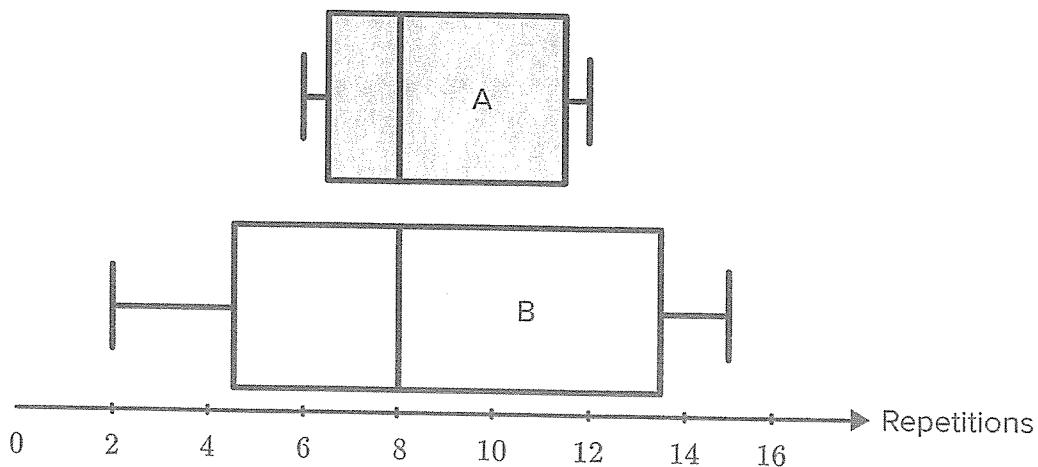
Athletes**Non-athletes**

- 6 At every training session of the season, a cyclist measured her pulse rate before a sprint and after a sprint. The before and after rates, measured in beats per minute (bpm), recorded throughout the season are presented in the following box plots:

- a How much greater was her median pulse rate after the sprint than before the sprint?
- b Find the interquartile range of her pulse rate before the sprint.
- c Find the interquartile range of her pulse rate after the sprint.
- d Find the range of her pulse rate before the sprint.
- e Find the range of her pulse rate after the sprint.
- f Are her pulse rate readings more consistent before or after the sprint?
- g In the last session of the season, the cyclist recorded her highest pulse rate of the season both before and after the sprint. By how much did her pulse rate increase during this particular training session?

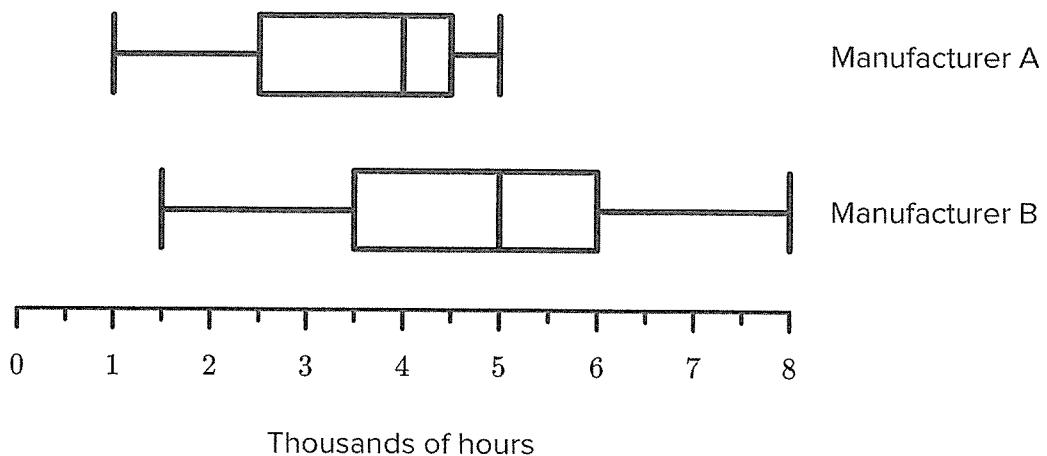


- 7 Two weightlifters both record their number of repetitions with a 70 kg bar over 30 days. The results are displayed in the following box plots:



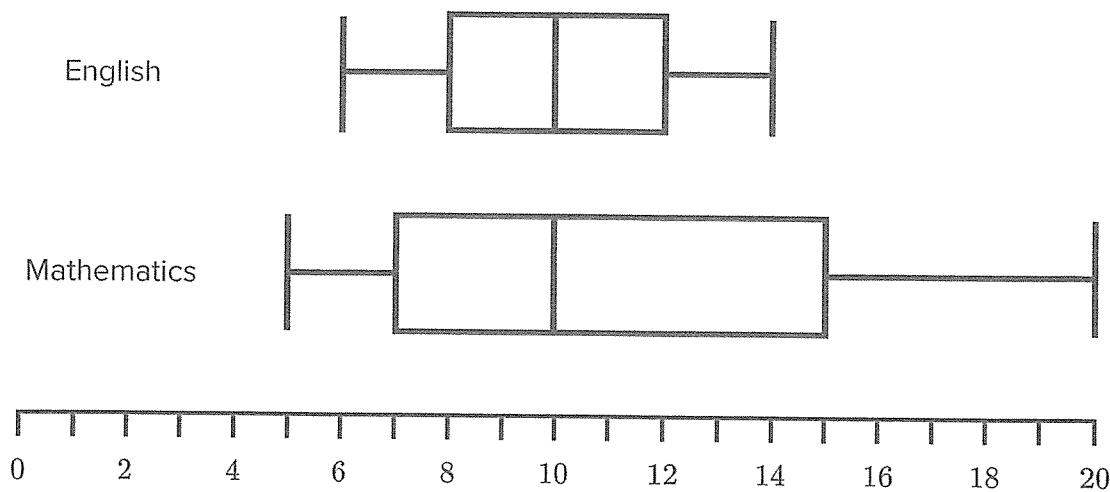
- a Which weightlifter has the more consistent results?
- b Determine which of the following statistics supports your answer:
 - A The mean.
 - B The range.
 - C The mode.
 - D The graph is positively skewed.

- c Which statistic is the same for each weightlifter?
- d Which weightlifter can do the most repetitions of 70 kg?
- 8 The parallel box plots below shows the data collected by the manufacturers on the life-span of light bulbs, measured in thousands of hours:



- a Complete the following table using the two box plots:
- | | Manufacturer A | Manufacturer B |
|---------------------|----------------|----------------|
| Median | | |
| Lower quartile | | |
| Upper quartile | | |
| Range | | |
| Interquartile range | | |
- b Hence, which manufacturer produces light bulbs with the best lifespan? Explain your answer.

- 9 A class took an English test and a Mathematics test. Both tests had a maximum possible mark of 20. The following parallel box plots shows the results of the tests:



- a Complete the following table using the two box plots:

	English	Mathematics
Median	10	10
Lower quartile	8	7
Upper quartile	12	15
Range	7	14
Interquartile range	4	8

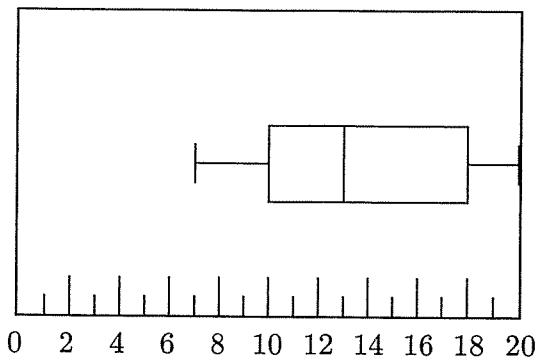
- b In which test did the class tend to score better marks? Explain your answer.

- 10 A mathematics test is given to two classes. The marks out of 20 received by students in each class are represented in the following box plots:

a Complete the following table:

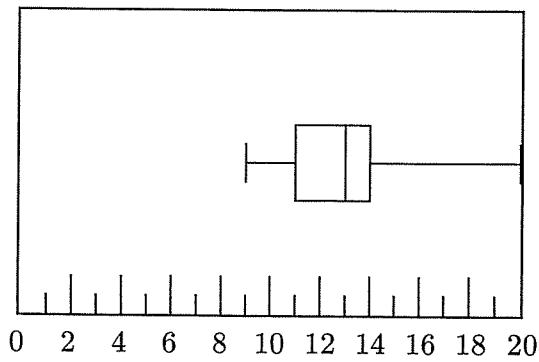
	Class 10P	Class 10Q
Median		
Lower quartile		
Upper quartile		
Range		
Interquartile range		

Class 10P



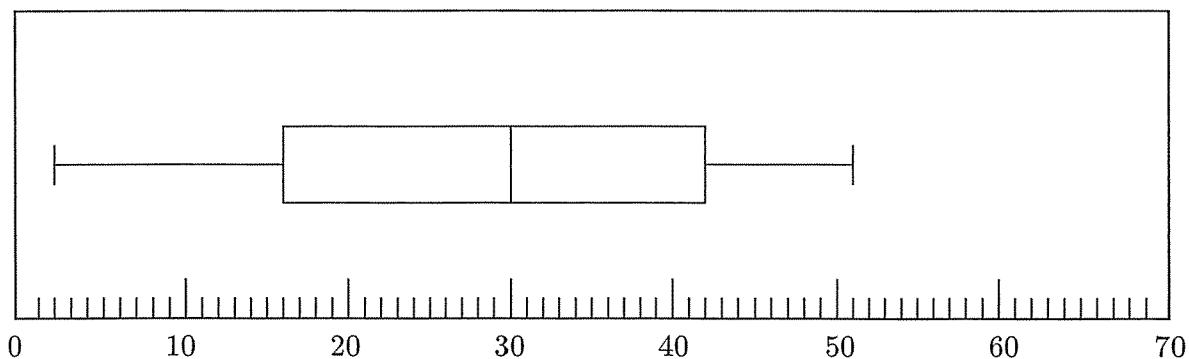
- b Which class tended to score better marks? Explain your answer.

Class 10Q

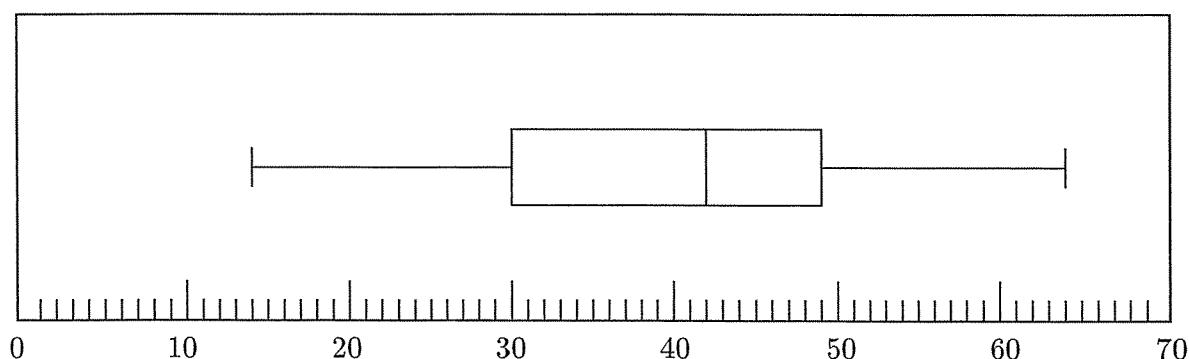


- 11 The following box plots represent the daily sales made by Carl and Angelina over the course of one month:

Angelina's sales

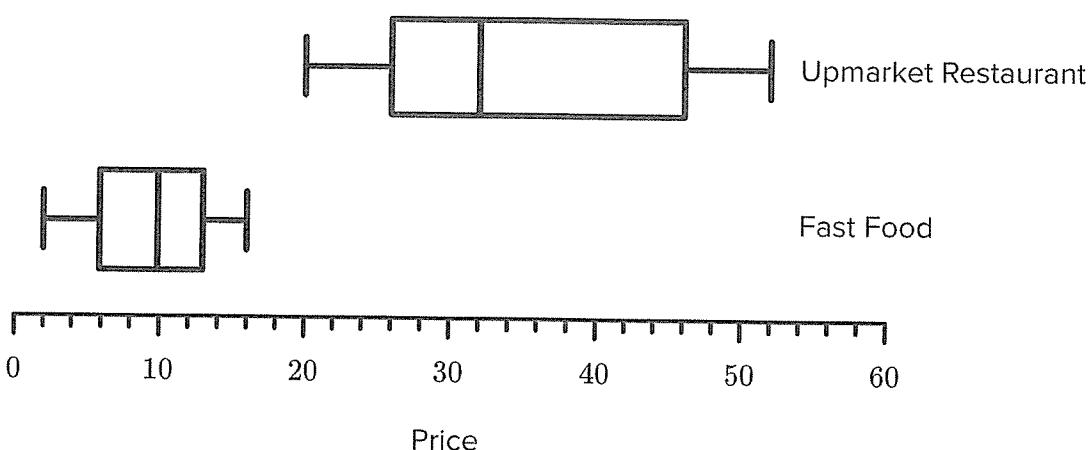


Carl's sales

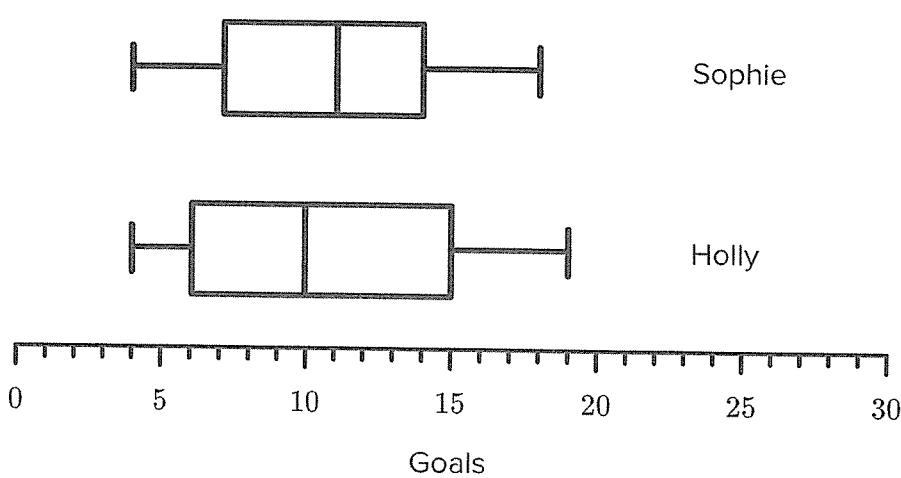


- Calculate the range for Angelina's sales.
- Calculate the range for Carl's sales.
- By how much did Carl's median sales exceed Angelina's?
- Considering the middle 50% of sales, which salesperson had the more consistent number of sales?
- Which salesperson had a more successful sales month?

- 12 The parallel box plots shows the prices, in dollars, of the items on the menu of an upmarket restaurant and the menu of a fast food restaurant:

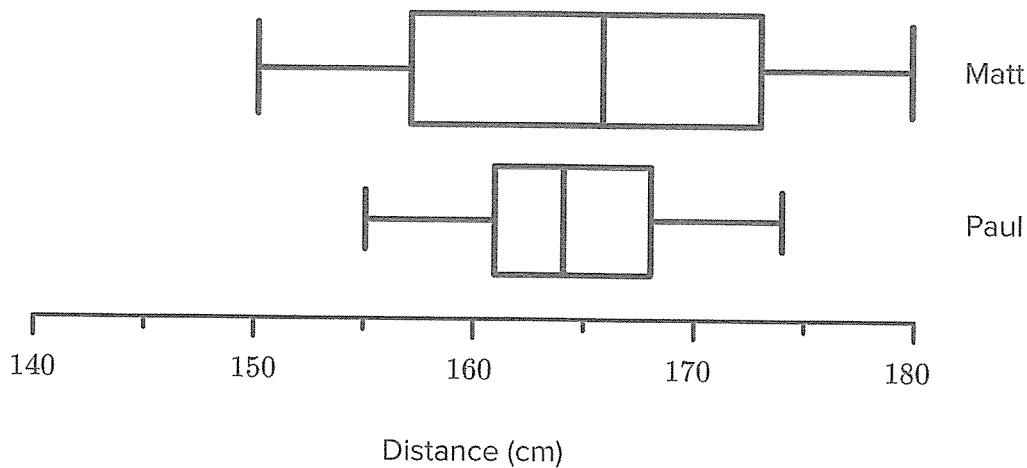


- a Which restaurant has the higher median price for the items they sell?
 - b What is the difference between the median prices of the items sold by each restaurant?
 - c Which restaurant has a greater price range for the items on their menu?
 - d What is the price difference between the most expensive items sold by each restaurant?
 - e What amount of the cheapest item at the fast food restaurant could be bought for the same price of the most expensive item at the upmarket restaurant?
- 13 The parallel box plots show the number of goals scored by two football players in each season:

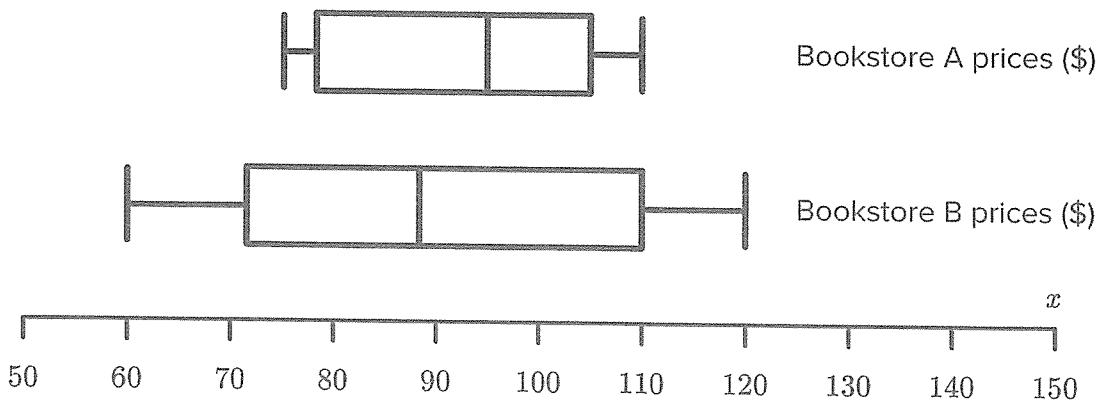


- a Who scored the most goals in a season?
- b How many more goals did Holly score in her best season compared to Sophie in her best season?
- c What is the difference between the median number of goals scored in a season by each player?

- d What is the difference between the interquartile range for both players?
- 14 The parallel box plots shows the distances, in centimetres, jumped by two high jumpers:



- a Who had a higher median jump?
- b Who made the highest jump?
- c Who made the lowest jump?
- 15 Two bookstores recorded the selling price of all their books. The results are presented in the parallel box plots:



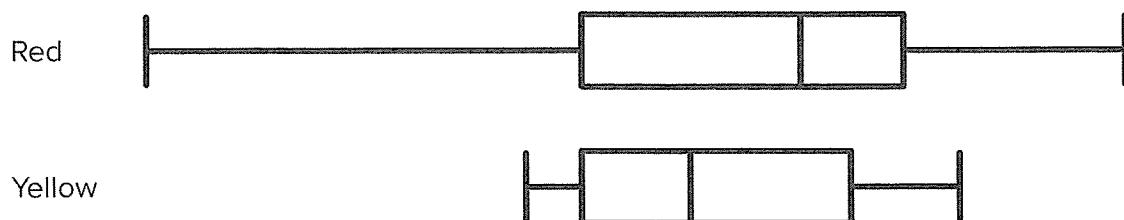
- a Which bookstore had the more consistent prices? Explain your answer.
- b Comparing the most expensive books in each store, how much more expensive is the one in store B?

c State whether the following statements are true or false:

- i 25% of the books in Bookstore B are at least as expensive than the most expensive book in Bookstore A.
- ii 25% of the books in Bookstore B are cheaper than the cheapest book in Bookstore A.

16 A builder can choose between two different types of brick that are coloured red or yellow.

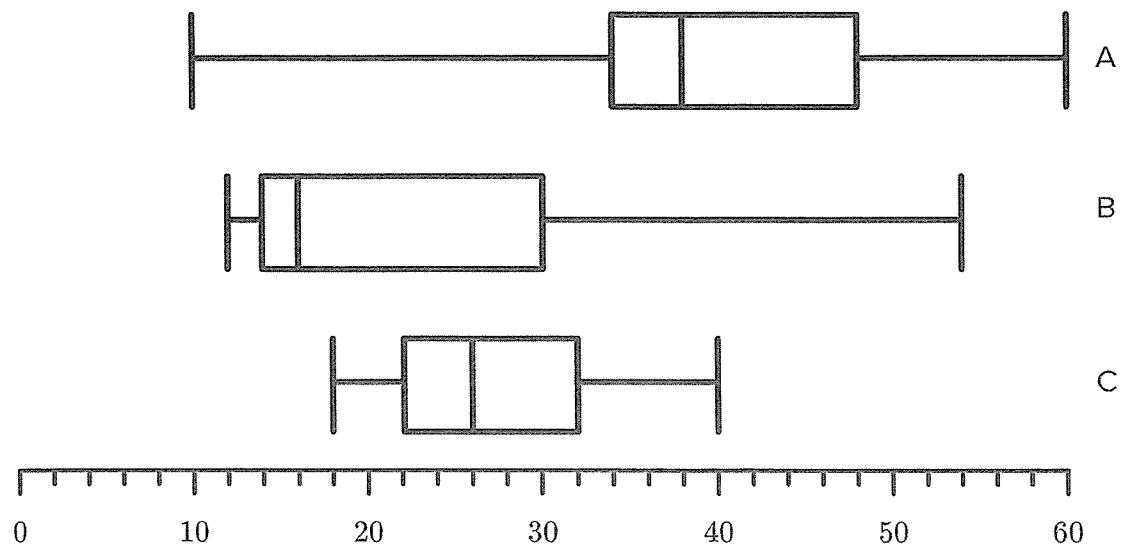
The following parallel box plots shows the results of tests on the strength of the bricks:



a Using the box plot, explain why a builder might prefer to use the red bricks.

b Using the box plot, explain why a builder might prefer to use yellow bricks.

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



- a Which film do you think has an adults only rating, restricting it to viewers 18 years of age and older? Explain your answer.
- b Which film would you recommend for a group of 15 year olds to watch? Explain your answer.
- c Which film would you recommend to a family of two parents in their 40's and two teenagers? Explain your answer.

C.T 12.09P Samples and populations (23Q'S)

1. For a statistical survey the population is deemed to be all the students that attend the local high school.

Which three of the following are samples of that population?

All the seniors.

A

The first 30 students to arrive at school that day.

B

All the students that were in the library at lunch time.

C

All the teachers.

D

Some of the teachers.

E

2. For a statistical survey the population is deemed to be all people aged under 18 , who live in Sydney.

Which two of the following are samples of that population?

100 students chosen from a university in Sydney.

A

Every person who lives in a particular suburb in Sydney.

B

All students of a primary school in Melbourne.

C

All students of a primary school in Sydney.

D

100 children chosen from the
Sydney under 15s baseball
tournament.

E

3. A sample of 50 people is drawn from a population. In this sample the youngest is 18 years old, and the oldest is 64.
Which two of the following might be populations that this sample was drawn from?

Students from a primary school.

A

People stuck in traffic due to roadworks.

B

Residents of a retirement home.

C

Employees at a bank.

D

4. For a statistical survey the population is deemed to be all people in a city who play in any organised sporting competition.

Which three of the following are samples of that population?

- 500 spectators chosen from a weekend sports match. A
- The members of 3 teams chosen from the local hockey tournament. B
- 100 people chosen at random from a local park. C
- All students from a local school who compete in a school sports competition. D
- All the active members of a local football club. E
5. Lachlan asks 120 Year 12 students at his school how much time they spend on homework per night. 78 Year 12 students say they do more than 3 hours. At a meeting of the student council Lachlan reports "65% of students at this school do too much homework."
- Which one of the following explains why this is misleading?
- The survey does not represent the population of the school. A
- The question should have been multiple choice. B
- The question was biased. C
- The sample size was too small. D
6. True or false: A census of the whole population is always possible.
- True A
- False B

7. True or false: A sample should have the same diversity as the population.

True

A

False

B

8. True or false: A convenience sample is a good sampling method to use.

True

A

False

B

9. If the population has 73 000 members, would a sample size of 7 be large enough?

Yes

A

No

B

10. If the population has 62 000 members, would a sample size of 6 000 be large enough?

Yes

A

No

B

11. A study is being done on physical activity in a small town.

Determine if each of the following characteristics would be important to consider when selecting the sample:

a Name

Yes

A

No

B

b Age

Yes

A

No

B

12. Antony wants to know what teams the students at his school support. He finds that 30 students support the Sydney Swans, and 55 students support the Melbourne Demons.

What is the population?

The students who support the
Sydney Swans and Melbourne
Demons

A

The 30 students who support
the Sydney Swans

B

All the students at his school

C

The 55 students who support
the Melbourne Demons

D

13. For a statistical survey, the population is deemed to be all the students that attend the local secondary school.

Determine if each of the following are a sample of that population:

a The first 30 students to arrive at school that day

Yes

A

No

B

b Some of the teachers

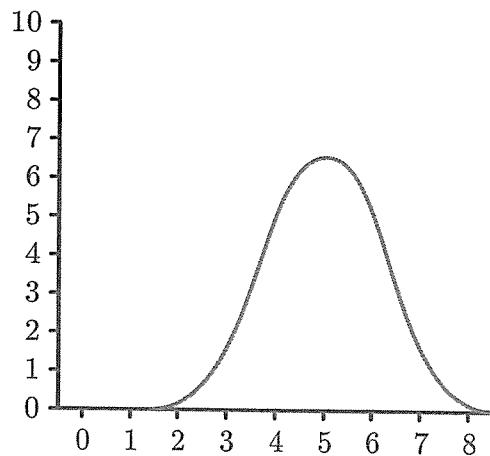
Yes

A

No

B

14. Match each sample histogram to the population it was taken from.



10

9

8

7

6

5

4

3

2

1

0

0

1

2

3

4

5

6

7

8

A

B

10

9

8

7

6

5

4

3

2

1

0

0

1

2

3

4

5

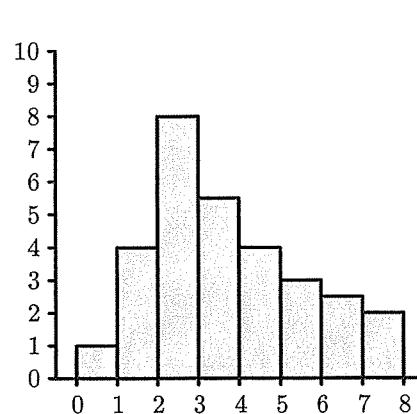
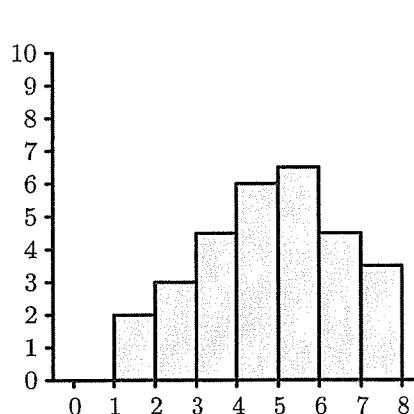
6

7

8

A

B



15. The population is all people in a city who play in any organized sporting competition. Which three of the following are examples of a biased sample of this population?

A Surveying athletes from various sports and skill levels

B Surveying only members of the most popular soccer league in the city

C Surveying only players from one specific age group

D Surveying only participants from a single sporting event

16. Sharon wants to know what year 10 students think of their Science class, so she polls 70 random year 10 students. Is the sample chosen biased or fair?

Biased

A

Fair

B

17. At a certain chocolate factory, 20% of products contain nuts. 700 chocolates are tested to check if they meet the required quality for sale. Of those tested, 78% contained nuts.

a What is the population?

The 700 chocolates tested

A

The chocolates that meet the required quality for sale

B

The chocolates that contain nuts

C

All chocolates produced in the factory

D

b What proportion of the population contains nuts?

c What proportion of the sample contains nuts?

d Do the 700 chocolates tested represent a random sample?

Yes

A

No

B

18. An ecologist collects 300 of a particular bee species in a meadow, marks them, and then releases them. In a follow-up collection, the ecologist gathers 250 bees, and 97 of them are found to be marked.

Estimate the total bee population in the meadow.

Give your estimate as a whole number.

19. A television station wants to estimate the number of viewers it had for a new show.

After randomly selecting 5 000 people from a population of 12 540 000, it was determined that 400 of those people had watched the show.

Estimate the number of people in the entire population who watched the show.

20. Uther is studying the performance of students in two different classes over a term. He found two secondary sources of data:

- Source A: This data set provides the individual grades of every student from both classes.
- Source B: This data set provides the average final grade of students from both classes.

- a If Uther wants to know how many students from one class scored above 80%, which source should he use?

Source A

A

Source B

B

- b Which source should Uther use if he wants to calculate the median grade for only one class?

Source A

A

Source B

B

- c Source B indicates that the average grade for one class is 85%. Can Uther determine how many students scored exactly 85% using Source B?

Yes

A

No

B

21. You are writing an essay on Australia's role in World War II. While researching, you come across two sources of information:

- Source A: A school textbook published in 1990 detailing Australia's military engagements, key figures, and notable battles during the war.
- Source B: A documentary released in 2015 that focuses on the personal stories of Australian soldiers, interviews with war historians, and unseen archival footage.

- a If you find conflicting information between the two sources, how might you determine which one is more accurate? Select the two correct options.

Trust the more recent source because it likely has access to newer information.

A

Disregard the conflicting information since both sources could be biased.

B

Use the source that provides the most detailed and vivid description.

C

Consider the authorship and expertise of the sources to evaluate credibility.

D

- b Why is it important to use multiple secondary sources when trying to get a comprehensive understanding of a historical event? Select the three correct options.

It prevents over-reliance on a single source, which could be biased or incomplete.

A

It allows for a more balanced perspective by comparing different viewpoints.

B

Different sources may provide unique details that others might omit.

C

It ensures that you can cite a variety of sources in your essay to meet academic standards.

D

22. For a statistical survey, the population is considered to be all of the athletes at an Australian Football League (AFL) game.

- a Select two different samples that could be taken at the game to represent this population.

A random sample of athletes who are substitutes during the game

A

A random sample of athletes from both teams playing in the game

C

A sample of athletes from only one team that played in the previous game

B

A sample of the starting players from both teams

D

- b Select two samples that could be taken at this game that do not belong to this population.

A sample of spectators in the audience, because they are not part of the population of athletes being surveyed

A

A sample of athletes from a different sport playing in the same venue, because they are not part of the AFL athletes population

B

A sample of players who were not on the field at the time of the survey, because focusing only on non-active players does not represent the athletes playing in the game.

C

A sample of athletes from only one of the AFL teams, because excluding athletes from the other team makes this sample incomplete.

D

23. Farma, an Australian pharmaceutical company, conducted a trial of a new medication with 4 000 participants in Sydney. 15% of participants reported mild side effects like headaches and fatigue. Farma plans to introduce the medication to Sydney's 8 million residents.

- a Based on the trial data, estimate the number of Sydney residents who might experience side effects if the medication were given to the entire target population.

b Which two ethical concerns might arise from releasing this prediction to the public?

It may lead to higher sales of the medication due to increased awareness.

A

It ensures complete transparency, which is always beneficial.

B

The information might be misinterpreted, leading to mistrust in the medication.

C

It could cause unnecessary fear among the public.

D

c Is the trial's sample size sufficient to represent Sydney's population? Why or why not?

Select the two correct options.

No, because a smaller sample size can lead to biased results.

A

Yes, because 4 000 participants is a large enough sample.

B

Yes, because the percentage of side effects would be similar in any group size.

C

No, because 4 000 participants may not accurately represent the diversity of 8 million people.

D

C.T REVIEW Univariate data (10Q'S)

1. Consider the following set of data:

27, 50, 24, 37, 47, 41, 27, 126, 44, 27

- a Fill in this table of summary statistics.

Mean

Median

Mode

Range

- b Which data value is an outlier?

- c Fill in this table of summary statistics after removing the outlier.

Mean

Median

Mode

Range

d

Let A be the original data set and B be the data set without the outlier.

Fill in this table using the symbols $>$, $<$ and $=$ to compare the statistics before and after removing the outlier.

With outlier Without outlier

Mean: A B

Median: A B

Mode: A B

Range: A B

2. Use technology to determine the standard deviation for the data represented by the frequency table.

Score	Frequency
15	13
16	9
17	23
18	19
19	8
20	13

Round your answer to two decimal places.

3. The life of two brands of batteries are tested using a sample of 10 batteries from each brand. Their battery lives (in hours) are shown below.

- Brand X : 23.3, 19.7, 20.7, 25.3, 22.5, 19.1, 20.0, 20.7, 20.7, 20.9
- Brand Y : 23.2, 27.5, 25.0, 24.5, 22.7, 29.8, 22.9, 26.0, 26.4, 22.6

- a Calculate the mean battery life of each brand.

Brand X : hours

Brand Y : hours

- b Calculate the standard deviation for the battery life of Brand X .

Round your answer to one decimal place.

- c Calculate the standard deviation for the battery life of Brand Y .

Round your answer to one decimal place.

- d Which brand produces batteries that generally last longer?

Brand X

A

Brand Y

B

- e Which brand produces batteries that are more consistent?

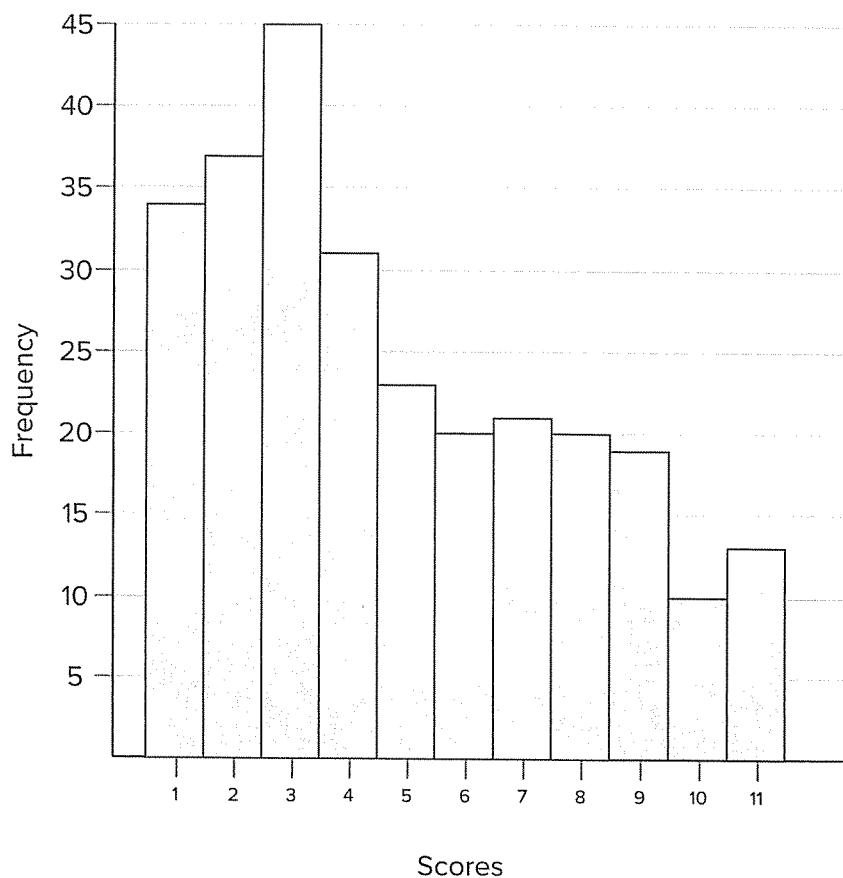
Brand X

A

Brand Y

B

4. Consider the data shown in the column graph below.



a Are there any outliers?

Yes

A

No

B

b Is there any clustering of data?

Yes

A

No

B

c Where does the clustering occur?

5 to 7

A

8 to 11

B

1 to 4

C

d What is the mode?

7

A

2

B

3

C

10

D

- e The distribution of the data is:

Negatively skewed.

A

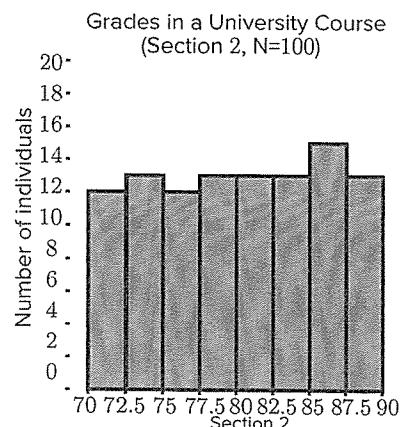
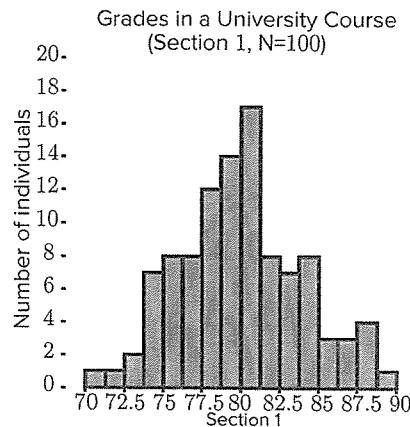
Symmetrical.

B

Positively skewed.

C

5. Consider the two histograms below which show the grade distributions in two university courses.



- a Which distribution has a clear mode?

Section 1

A

Section 2

B

- b Which distribution would be described as “uniform”?

Section 1

A

Section 2

B

c Which distribution has the highest range?

Section 1

A

Section 2

B

They have the same range. C

d How do you expect the mean and median of Section 1 to compare?

Median > mean

A

Mean > median

B

Mean = median

C

e Which distribution would have the highest standard deviation?

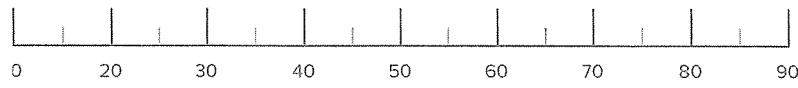
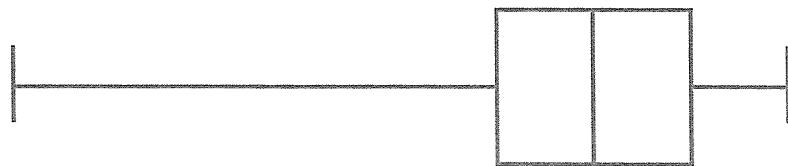
Section 1

A

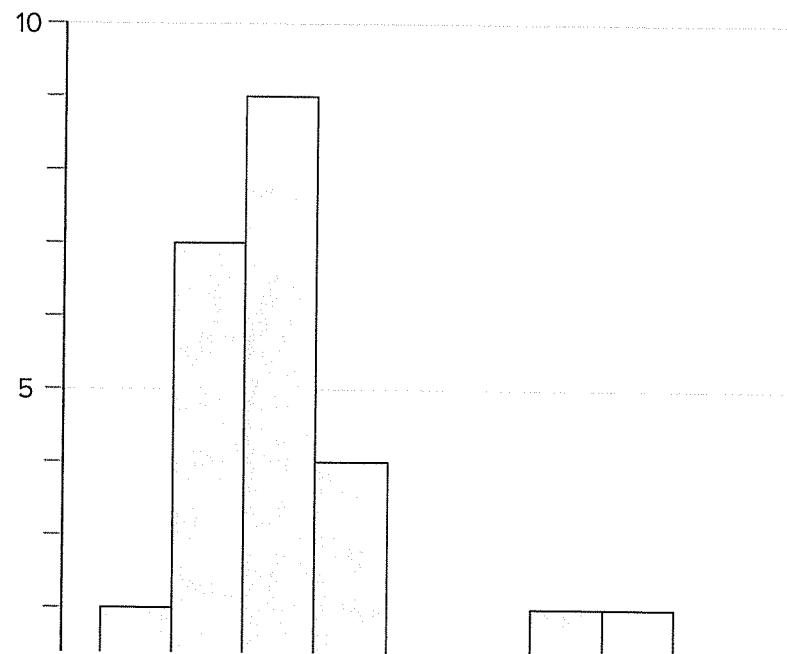
Section 2

B

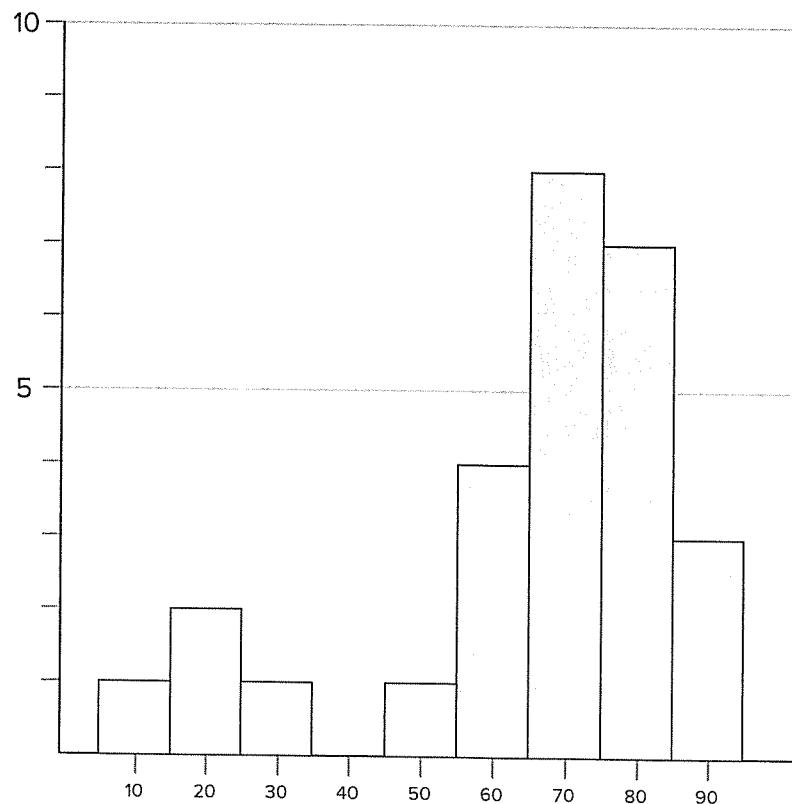
6. Match the box plot shown to the correct histogram.



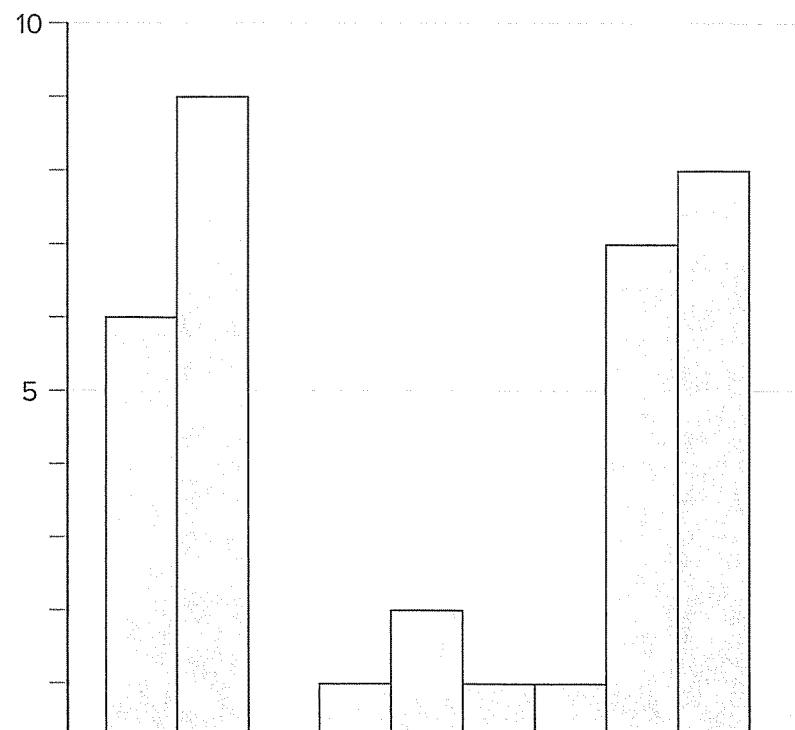
A



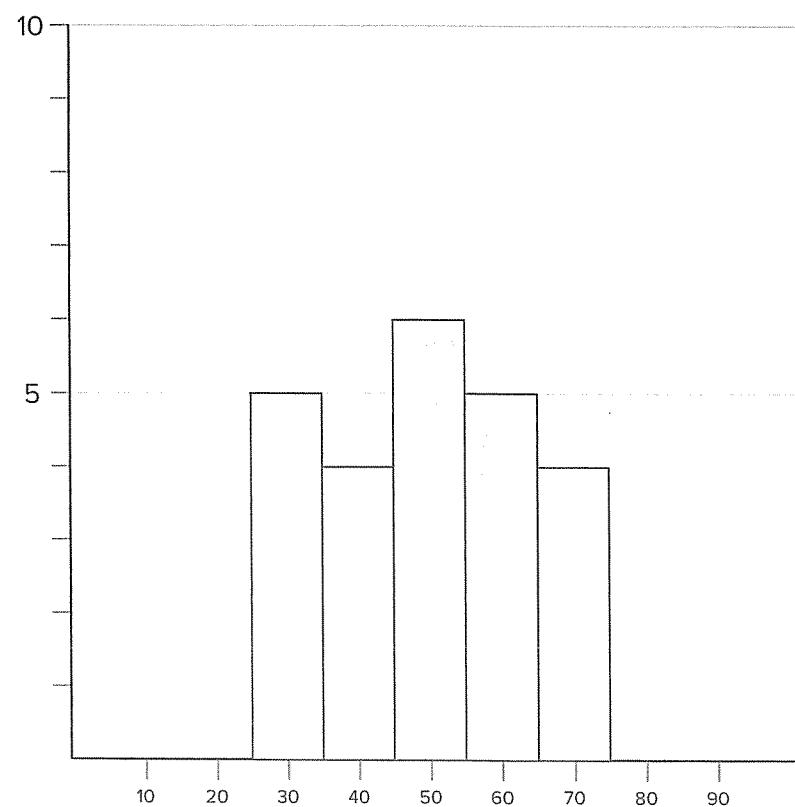
B



C



D



7. For each of the following, select whether they are a census or a sample.

- a Lucy has asked everyone in her office what snacks should be provided in the office.

Census

A

Sample

B

- b James asks a few of his friends how they did in the test to see if he is above average in his class.

Census

A

Sample

B

- c Joanne finds the height of the entire class to try to find the average height of 15 year old students in Australia.

Census

A

Sample

B

8. A factory produces 1 820 TVs every day. How many TVs are tested daily if the factory tests a systematic sample of:

- a every 13th TV

- b every 70th TV

9. Four lucky people from a group of 215 each stand to win an iPad. Every contestant is given a different number between 1 and 215, and the winners are selected by generating a random number uniformly between 0 and 1.

To ensure there is an equal chance of each contestant winning, the number is multiplied by 215 and then rounded up. In this case the numbers generated

were:

0.152 0.534 0.352 0.795

- a What is the number of the winning contestant selected using the first random number generated?
 - b What is the number of the winning contestant selected using the second number in the list?
 - c What is the number of the winning contestant selected using the third number in the list?
 - d What is the number of the winning contestant selected using the fourth number in the list?
10. The owner of a movie cinema wants to use stratified sampling in their survey of people that come to their cinema.
- Which two of the following methods would be considered as stratified sampling?

Interview 10% of the people who used the candy bar and 10% of people who didn't.

A

Interview every person that sees a romantic movie.

B

Interview 10% of the people from each movie.

C

Interview every 10th person that purchases a ticket.

D

