

Boxplot // Box-and-Whiskers

Example

13 9 8 8 10 7 10 9 8 9 8 9

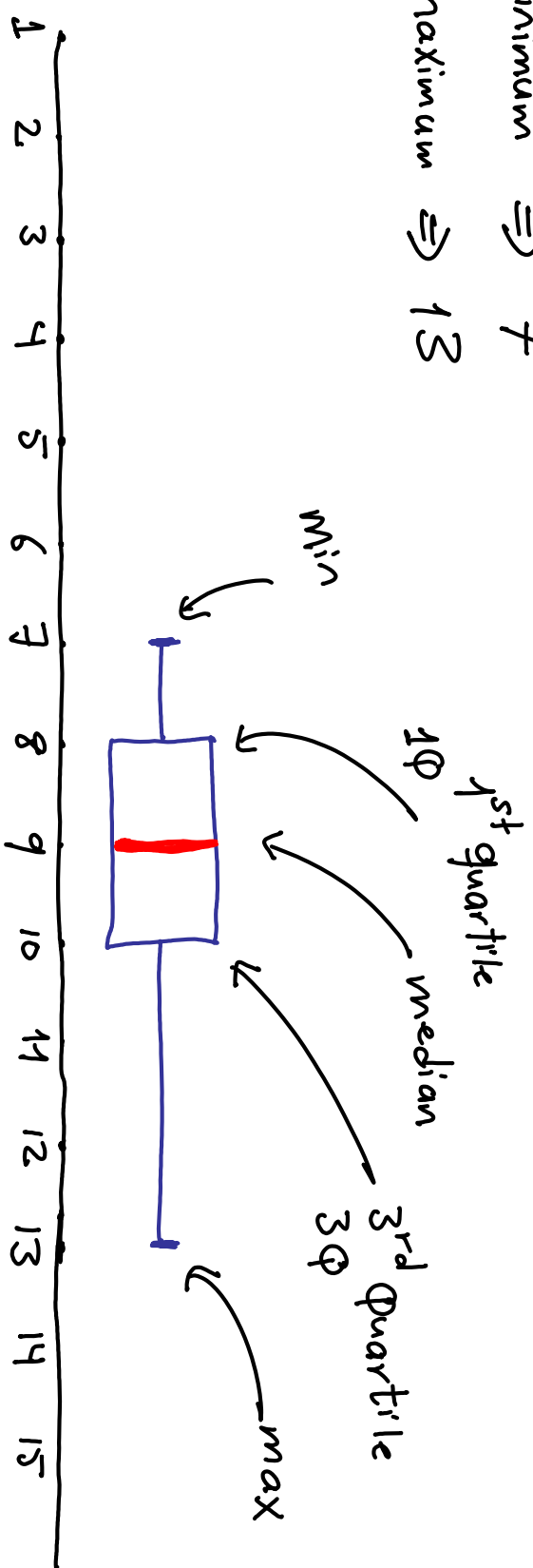
In-order $\Rightarrow$ 7 8 8 8 9 9 9 9 10 10 10 13

Median $\Rightarrow$ 9

1st quartile $\Rightarrow Q_1 \Rightarrow 25\% \Rightarrow$ 8

3rd quartile $\Rightarrow Q_3 \Rightarrow 75\% \Rightarrow$ 10

Minimum $\Rightarrow 7$
Maximum $\Rightarrow 13$



(Ex)

The owner of a restaurant wants to find out more about where his patrons are coming from. One day he decided to gather data about the distance (in miles) that people commuted to get to his restaurant. people reported the following distances

14 6 3 2 4 15 11 8 1 7 2 1 3 4 10 22 20

He wants to create a graph that helps him understand the spread of distances (and the median distance) that people travel

Note

Whisker $\Rightarrow$ spread of data

Range $\Rightarrow$ highest - lowest

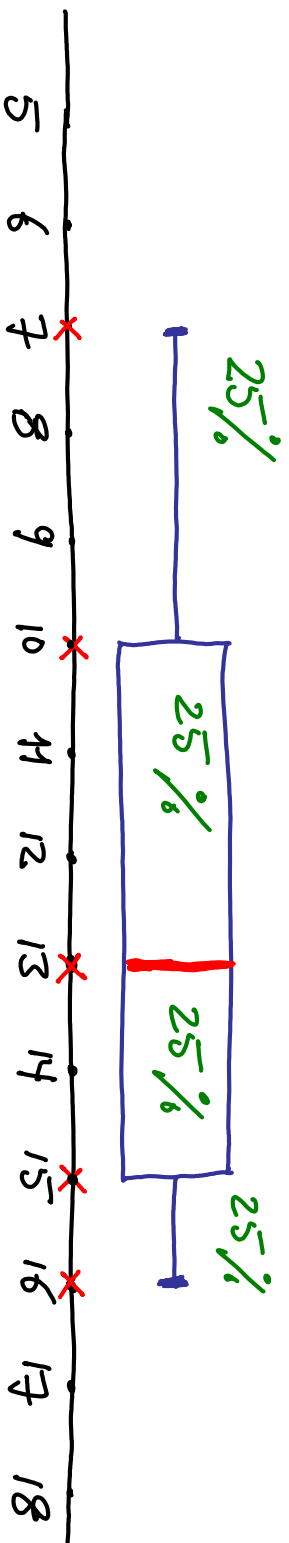
half of ages are less than median

$Q_1 \Rightarrow$ 1st quartile

IQR $\Rightarrow$ interquartile Range

$$= Q_3 - Q_1$$

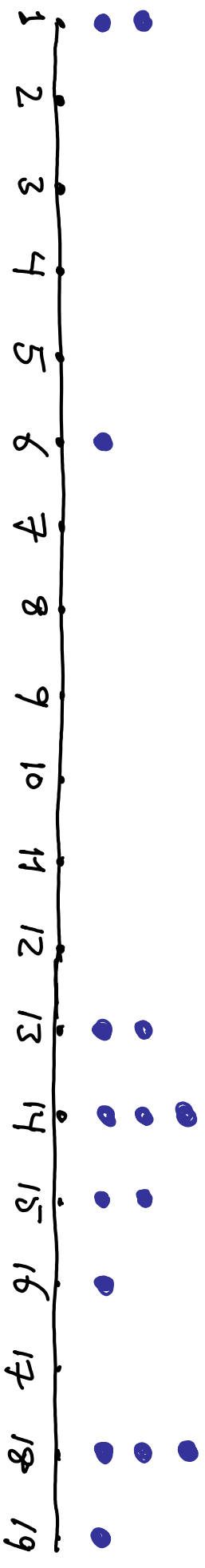
Interpreting boxplot



- ⊗ All of students are less than 17 years-old
- ⊗ At least 75% of the student are 10 years old or older
- ⊗ Exactly half the students are older than 13

Judging outlier in a dataset

1 1 6 13 13 14 14 14 15 15 16 18 18 18 18 19



* Visually $\Rightarrow$ may be 1, 6 are outlier

by Rule $\Rightarrow Q_1 - 1.5 * IQR > \text{outlier} > Q_3 + 1.5 * IQR$

* Outlier is a data point that lie outside the overall pattern in a distribution

* Outlier can be represented as dots that are separate from the rest of the plot

* Whisker extend to the farthest point in the dataset that was not an outlier

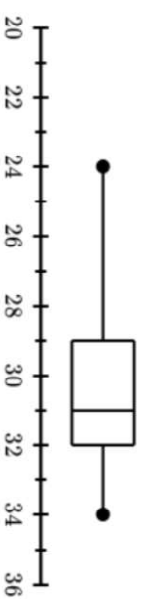
The data below represents the number of customers at each Slurpee Sam's Spaghetti Shop.

24, 25, 29, 30, 31, 31, 32, 34, 34

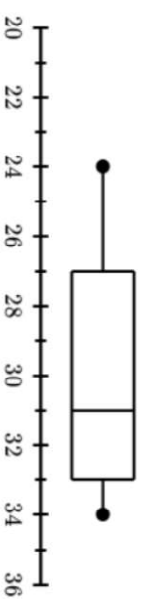
Which box plot correctly summarizes the data?

Choose 1 answer:

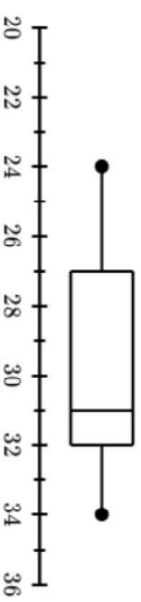
(A)



(B)

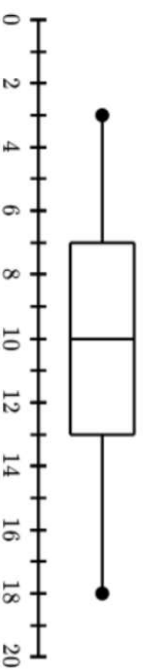


(C)



Choose 1 answer:

Which data set could be represented by the box plot shown below?



Choose 1 answer:

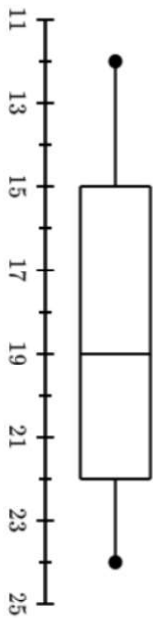
- ☐ A 3, 4, 8, 9, 9, 10, 12, 13, 13, 16, 18
- ☐ B 3, 4, 7, 9, 9, 10, 12, 13, 13, 16, 18
- ☐ C 2, 4, 7, 9, 9, 10, 12, 13, 13, 16, 18
- ☐ D 3, 4, 8, 9, 9, 12, 12, 13, 13, 16, 18

The data below represents the number of runners on each cross country team in the Northern Conference.

12, 15, 15, 17, 19, 19, 20, 22, 24

Which box plot correctly summarizes the data?

(A)



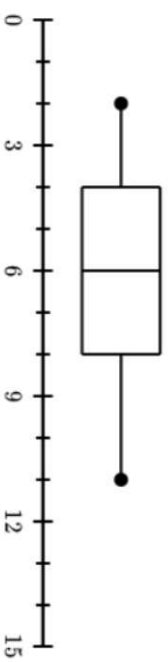
(B)



(C)



Which data set could be represented by the box plot shown below?



Choose 1 answer:

☐ A 2, 3, 5, 5, 6, 7, 8, 10, 11

☐ B 3, 3, 5, 5, 6, 7, 8, 9, 11

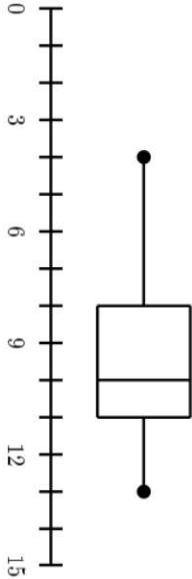
☐ C 2, 3, 5, 5, 7, 7, 8, 10, 11

☐ D 2, 3, 5, 5, 6, 7, 8, 8, 11

Find the interquartile range (IQR) of the data in the box plot below.

rebounds

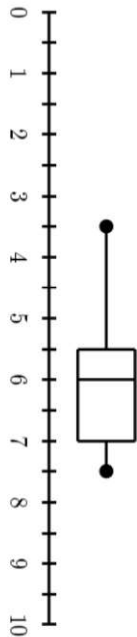
Juan's number of rebounds each game



Find the range of the data in the box plot below.

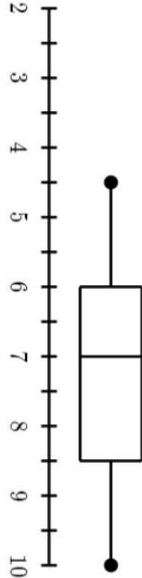
cm

Amount of rain in Vietnam cities
(centimeters)

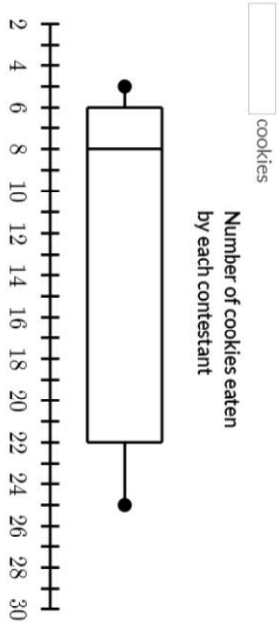


Find the median of the data in the box plot below.

kg Masses of Farmer Faye's pumpkins
(kilograms)



Find the interquartile range (IQR) of the data in the box plot below.



The five-number summary for the number of gym members at each gym in Shanesville is shown in the following table.

Min	Q_1	Median	Q_3	Max
50	56	64	70	76

About what percent of gyms in Shanesville have fewer than 56 gym members?

Choose 1 answer:

☐ A 0%

☐ B 25%

☐ C 50%

☐ D 75%

The five-number summary for the number of teams in each of Brad's fantasy football leagues is shown in the following table.

Min	Q_1	Median	Q_3	Max
4	7	10	14	18

The five-number summary suggests that about 50% of Brad's fantasy football leagues have fewer than how many teams?

Choose 1 answer:

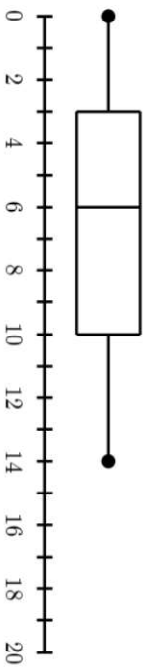
☐ A 4

☐ B 7

☐ C 10

☐ D 14

Number of aces Olga served
during each tennis match



The box plot suggests that Olga served fewer than what number of aces during about 75% of tennis matches?

Choose 1 answer:

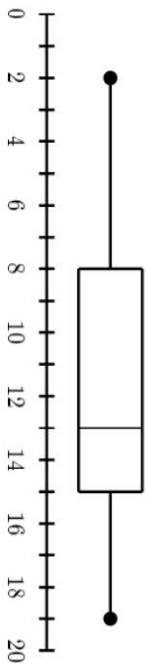
☐ A 0

☐ B 3

☐ C 6

☐ D 10

Number of surfers at each beach in Miami



The box plot suggests that 75% of beaches in Miami have more than what number of surfers?

Choose 1 answer:

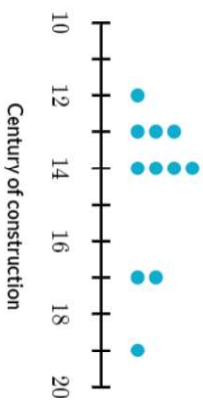
☐ A 2

☐ B 8

☐ C 13

☐ D 15

The following dotplot shows the centuries during which the 11 castles whose ruins remain in Somerset, England were constructed. Each dot represents a different castle.



Here is the five-number summary for these data:

Five-number summary				
min	Q_1	median	Q_3	max
12	13	14	17	19

According to the $1.5 \cdot IQR$ rule for outliers, how many high outliers are there in the data set?

Choose 1 answer:

☐ A 0

☐ B 1

☐ C 2

☐ D 3

