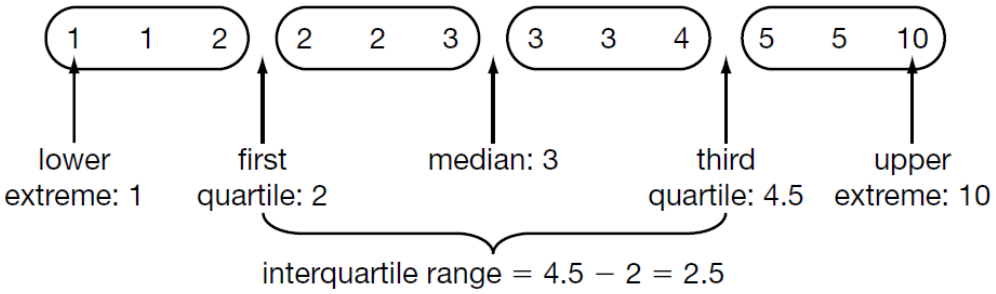
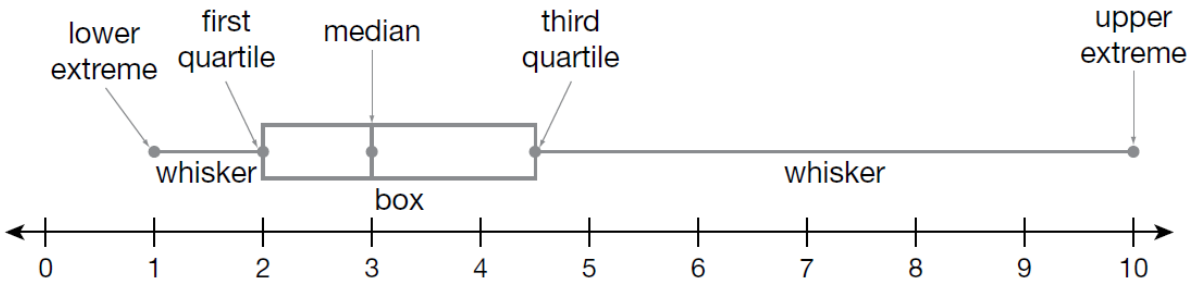
Box (and Whisker) Plots

UNDERSTAND Besides mean absolute deviation, another measure of spread is the **interquartile range (IQR)**. The interquartile range measures the variability of the middle 50% of the data, which is bounded by the **first quartile (Q_1)** and the **third quartile (Q_3)**. Recall that the median, *M*, divides a set of data into two halves. The first quartile is the median of the lower half of the data set. The third quartile is the median of the upper half of the data set.

The diagram below shows the **lower extreme, upper extreme,** quartiles, and median of a data set, as well as the interquartile range. It also helps illustrate how the median and quartiles divide a data set into four discrete sets of data.



A **box plot** (sometimes called a box-and-whisker plot) is an excellent way to display the extremes, quartiles, and median of a data set.



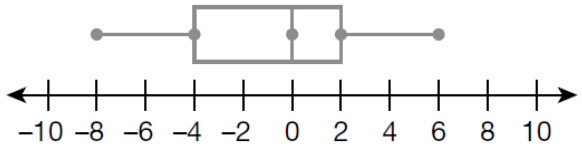
The box contains the middle 50% of the data, bounded by the first and third quartiles. The whiskers are on either end of the box. The left whisker contains the lower 25% of the data, and the right whisker contains the upper 25% of the data.

Outliers have much less effect on median and IQR than they do on mean and mean absolute deviation. Consider the example above and imagine that the upper extreme is 110, instead of 10. The first quartile, median, third quartile, and IQR would remain the same. The mean and mean absolute deviation, however, would be much greater.

In general, mean absolute deviation is a better measure of spread for data with a symmetric distribution and without outliers. For data with a skewed, or nonsymmetric, distribution and with outliers, IQR is usually a better measure of variability.

Practice

Use the box plot for questions 1–5.



1. What is the median? _____

HINT The median is the middle value in the set.

2. What is the lower extreme? _____

3. What is the upper extreme? _____

4. What is the first quartile? _____

5. What is the third quartile? _____

Find the median (M), first quartile (Q_1), and third quartile (Q_3) of the data.

6. 1, 2, 3, 5, 7, 9, 10

M = _____

Q_1 = _____

Q_3 = _____

REMEMBER The median divides the data set into two halves.

7. 10, 12, 12, 15, 17, 19, 21, 25

M = _____

Q_1 = _____

Q_3 = _____

8. −2, −1, 2, 3, 4, 6, 7, 7, 9

M = _____

Q_1 = _____

Q_3 = _____

9. 25, 35, 40, 45, 45, 50, 60, 65, 75, 95

M = _____

Q_1 = _____

Q_3 = _____

10. 15, 12, 18, 25, 36, 48, 28, 15

M = _____

Q_1 = _____

Q_3 = _____

11. 1.5, 2.5, 4.5, 8.5, 3.5, 0.5, 0.75, 2.25, 3.25

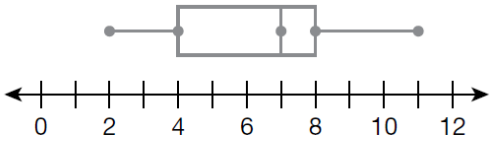
M = _____

Q_1 = _____

Q_3 = _____

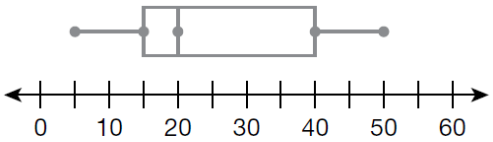
Calculate the interquartile range of the data.

16.



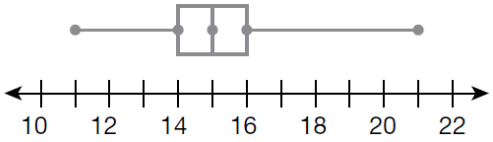
IQR = _____

17.



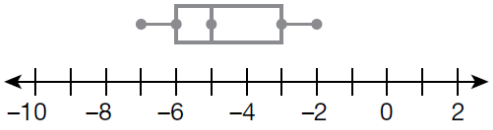
IQR = _____

18.



IQR = _____

19.

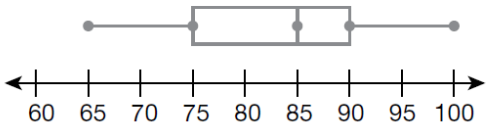


IQR = _____

Choose the best answer.

20. The box plot shows the test scores earned by students in a biology class. Which statement about the test scores is **not** true?

Biology Test Scores



A. The scores ranged from 65 to 100.

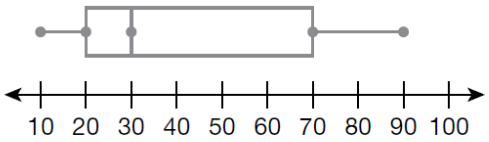
B. The median score earned was an 85.

C. 25% of students scored less than 75 points on the test.

D. 50% of students had scores that ranged from 75 to 85 points.

21. The box plot shows the prices of 20 skirts for sale at a boutique. Which statement about the prices is true?

Skirt Prices (in dollars)



A. The highest-priced skirt costs $100.

B. The median price of a skirt is $70.

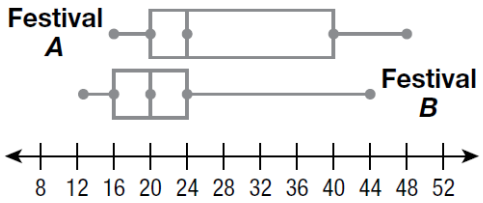
C. Half the skirts have prices that range from $20 to $70.

D. The prices of the skirts are close to the median and not very variable.

Use the box plots and information below for questions 22 and 23.

Music festival *A* and music festival *B* each had 100 volunteers. The box plots show the ages of the volunteers at each festival.

Ages of Music Festival Volunteers



22. Compare the median ages of volunteers at each festival.

_____

23. Which festival has more variability in the ages of its volunteers? Explain your answer.

_____

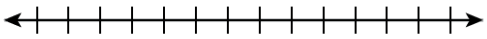
_____

Use the information below for questions 29 and 30.

Mrs. Heath visited her aunt in Nome, Alaska, for the first ten days of January 2012. She recorded the daily low temperature, in degrees Fahrenheit (°F), each day:

$-27, -27, -31, -33, -34, -33, -34, -25, -29, -26$

29. ORGANIZE Organize these data by displaying them in a box plot. Use the number line provided below.



30. JUSTIFY Mrs. Heath said, "The weather was very, very cold and did not vary much during the trip." Is her statement accurate? Use one or more measures of variability to justify your answer.

_____

_____

_____

_____