

## Box and Whisker Plots

A data set can be described using mean, median, mode, and range. It can also be divided into quartiles and displayed using a box and whisker plot.

Quartiles divide the data set into four equal parts. First the data set is divided in half by finding the median (Q2).

For example, let's consider the data set:

$$\{9, 11, 12, 15, 16, 16, 17, 18, 20, 21, 22\}$$

which has already been ordered from least to greatest.

What is the median? \_\_\_\_\_ The median represents quartile 2 (Q2). Circle and label it above.

Next, find the median of the first half of the data. The first half of the data does not include the Q2 value. This represents quartile 1 (Q1)

Then, find the median of the second half of the data. The second half of the data does not include the Q2 value. This represents quartile 3 (Q3)

Using this data to create a box-and-whisker plot we would first draw a number line large enough to include the minimum and maximum values in the data set. Then, we draw a box from Q1 to Q3 and a vertical line at the median value. Finally, we draw whiskers extending from Q1 to the minimum value and from Q3 to the maximum value.

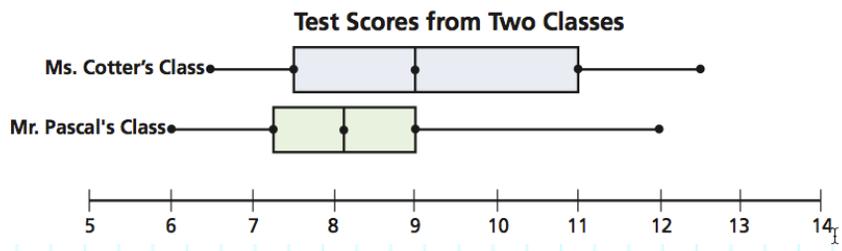
The first whisker represents the lowest  $\frac{1}{4}$  of the data. The box represents the middle half of the data. The last whisker represents the highest  $\frac{1}{4}$  of the data.

Box and whisker plots not only show the minimum, Q1, median (Q2), Q3, and the maximum, but also the range of the data and the interquartile range.

The **interquartile range** is found by subtracting Q1 from Q3. The width of the box above is the interquartile range for this set of data.

The interquartile range for this set of data is 8 (20-12). The advantage of interquartile range as a measure of variability is that, unlike range, it is not affected by outliers.

Consider the double box and whisker plot below to answer the following questions.



Which class had the higher median score?

Which class had its scores grouped more closely around its median?

What word could you use to describe a data set with data grouped closely around the median?

What word could you use to describe a data set with data grouped far from its median?

Which class had the better test scores?

Construct a box and whisker plot for this data set.

**Ages of children already signed up  
for swimming classes:  
10, 12, 9, 7, 10, 12, 14, 14, 10, 16**

