Key Concepts

Box-and-Whisker Plots

Objective Teach students to construct box-and-whisker plots from numerical data.

Note to the Teacher *Box-and-whisker plots are a method for displaying data graphically, explicitly showing the median of the data, as well as the upper and lower quartile values. Displaying data sets using a box-and-whisker plot is a convenient way to summarize large amounts of data, since even with very large data sets the procedure does not become more complicated.*

Median, Quartiles, and Extremes

A box-and-whisker plot uses five special values to give a graphic picture of a set of data. These five values are the *median*, the *upper quartile*, the *lower quartile*, the *upper extreme*, and the *lower extreme*. We have already learned how to determine the median of a set of data. How do we find the upper and lower quartiles and the upper and lower extremes?

First let's talk about the quartiles. Remind students that the **median** of a data set is the number that "divides" the set in half. That is, half of the data values are less than the median and the other half of the data values are greater than the median. This is why the median is usually referred to as the "middle" value of the data set. In much the same way, the **lower quartile** is the number such that approximately one-fourth of the data values are less than this number and the other data values are all greater than this number. The lower quartile value can be obtained by finding the median of the set of all data values that are less than or equal to the median; in other words, the *lower quartile is the median of the lower half of the data set.* In the same way, the **upper quartile** is the number such that approximately one-fourth of the data values are greater than this number and the other data values are all less than this number. So, the upper quartile value can be obtained by finding the *median of the upper half of the data set.*

The lower and upper extremes are easier to identify. The **lower extreme** is the least value in the data set and the **upper extreme** is the greatest value. These two values can therefore be found directly from an ordered set of data.

Example 1 A researcher traveled to a variety of stores in a metropolitan area checking on the price of a gallon of milk. The prices, in dollars, are given below.

{2.53, 2.45, 2.42, 2.40, 2.50, 2.40, 2.39, 2.46, 2.48, 2.42, 2.42, 2.44, 2.40, 2.49, 2.45, 2.42}

Find the lower extreme, lower quartile, median, upper quartile, and upper extreme values for this data set.

Solution **Step 1** Write the prices in order from least to greatest.

2.39, 2.40, 2.40, 2.40, 2.42, 2.42, 2.42, 2.42, 2.44, 2.45, 2.45, 2.46, 2.48, 2.49, 2.50, 2.53

Step 2 Read the lower and upper extremes from the list.

Lower extreme: 2.39   Upper extreme: 2.53

Step 3 Find the median of the data values.

There are 16 data values. When a data set has an even number of values, the median is the average of the two middle numbers. In this case, these two numbers are 2.42 and 2.44.

$$\frac{2.42 + 2.44}{2} = \frac{4.86}{2} \text{ or } 2.43$$

Step 4 Find the upper and lower quartile values.

The median separates the data set into two halves. The lower quartile is the median of the lower half of the data. The eight values in the lower half are:

2.39, 2.40, 2.40, 2.40, 2.42, 2.42, 2.42, 2.42.

Since this list has an even number of values, we take the average of the two middle numbers, 2.40 and 2.42, which is 2.41. So, the lower quartile value is 2.41.

The upper quartile is the median of the upper half of the data. The eight values in the upper half are:

2.44, 2.45, 2.45, 2.46, 2.48, 2.49, 2.50, 2.53.

The two middle numbers of this list are 2.46 and 2.48, and their average is 2.47. So, the upper quartile value is 2.47.

So, the five special values for the milk data are:
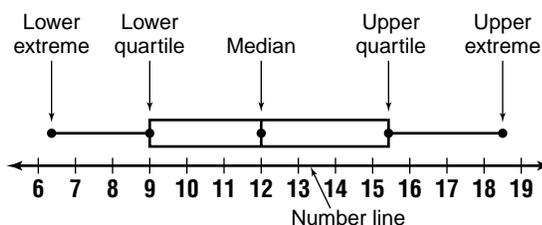
lower extreme: $2.39
lower quartile: $2.41
median: $2.43
upper quartile: $2.47
upper extreme: $2.53

Box-and-Whisker Plots

The box-and-whisker plot is a way to graphically display the median, upper and lower quartiles, and upper and lower extremes of a data set. The figure below shows the general construction of a box-and-whisker plot using the five key values for a data set.


To draw a box-and-whisker plot for a set of data, first draw a number line. Then above the number line, plot each of the five special values identified above. Next draw a rectangle (the "box") whose left side lies at the lower quartile value and whose right side lies at the upper quartile value. Draw a vertical line inside the box at the location of the median. Finally, draw two horizontal line segments (the "whiskers"), one from the lower extreme value to the lower quartile value and the other from the upper quartile value to the upper extreme value.

Work through the following example with students.

Example 2 **For a given data set, the lower extreme is 1.5, the lower quartile is 2.5, the median is 4, the upper quartile is 6, and the upper extreme is 7. Use these values to draw a box-and-whisker plot for the data set.**
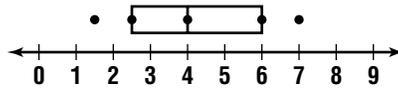
Solution First we draw a number line and plot the five values: 1.5, 2.5, 4, 6, and 7.
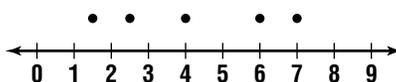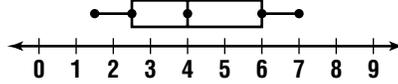

Next we draw the box from the lower quartile to the upper quartile.
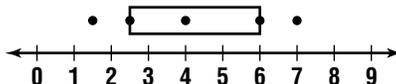
Now we draw a vertical line in the box at the location of the median.

Finally, we draw the whiskers from lower extreme to lower quartile and from upper quartile to upper extreme. This completes the box-and-whisker plot.

Exercise

Have students draw a box-and-whisker plot for the milk price data set given in Example 1.

End of
Lesson