Unit 1, Lesson 4

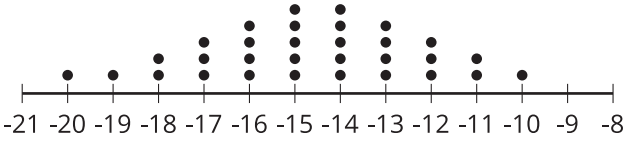
The Shape of Distributions

Let's explore data and describe distributions.

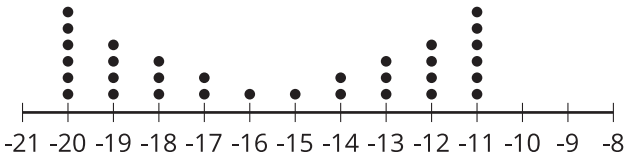
4.1 Which Three Go Together: Distribution Shape

Which three go together? Why do they go together?

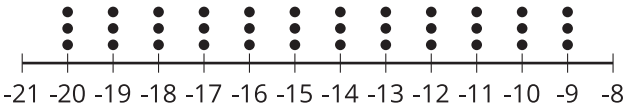
A.



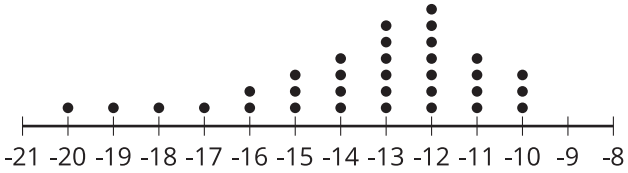
B.



C.



D.



4.2 Card Sort: Matching Distributions

Take turns with your partner matching 2 different data displays that represent the distribution of the same set of data.

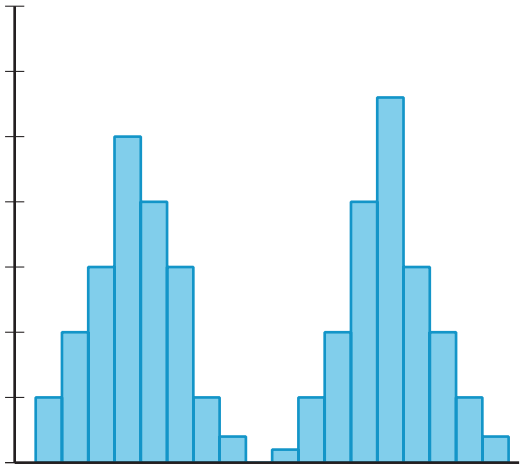
1. For each set that you find, explain to your partner how you know it's a match.
2. For each set that your partner finds, listen carefully to their explanation. If you disagree, discuss your thinking and work to reach an agreement.
3. When finished with all ten matches, describe the shape of each distribution.

4.3 Where Did the Distribution Come From?

Your teacher will assign you some of the matched distributions. Using the information provided in the data displays, make an educated guess about the question that produced this data. Be prepared to share your reasoning.

Are you ready for more?

This distribution shows the length in inches of fish caught and released from a nearby lake.



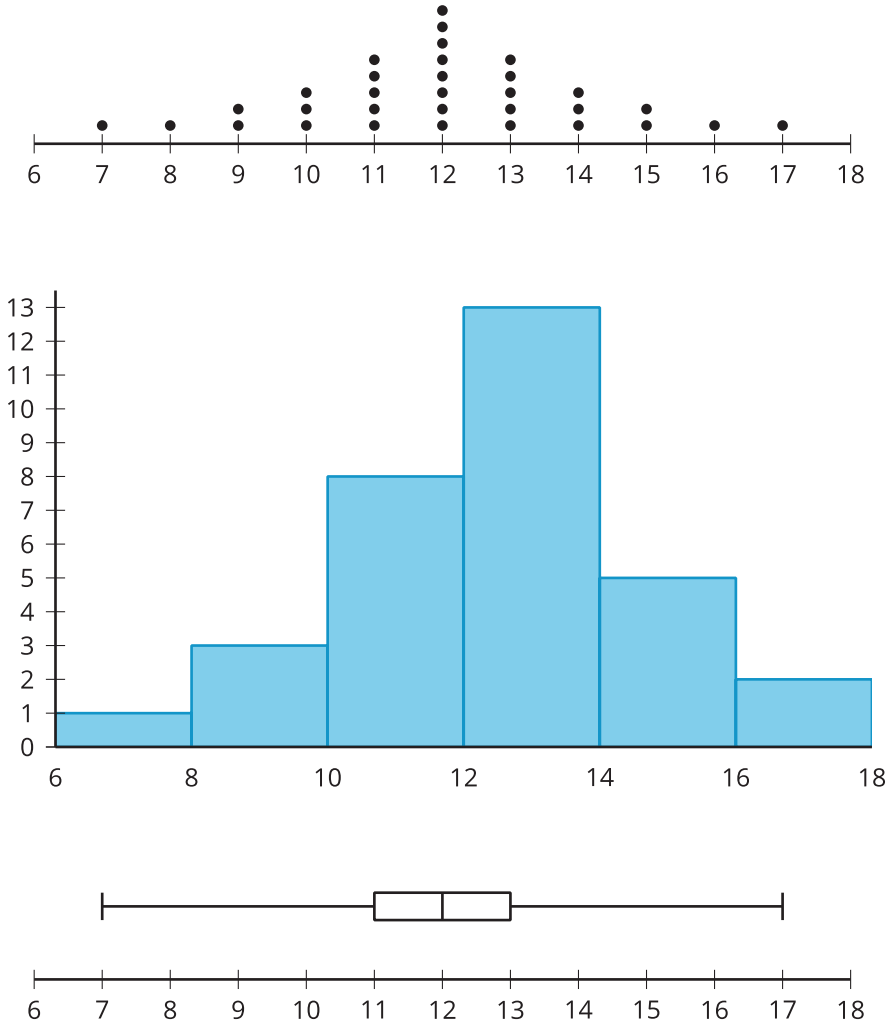
length of fish (inches)

1. Describe the shape of the distribution.

2. Make an educated guess about what could cause the distribution to have this shape.

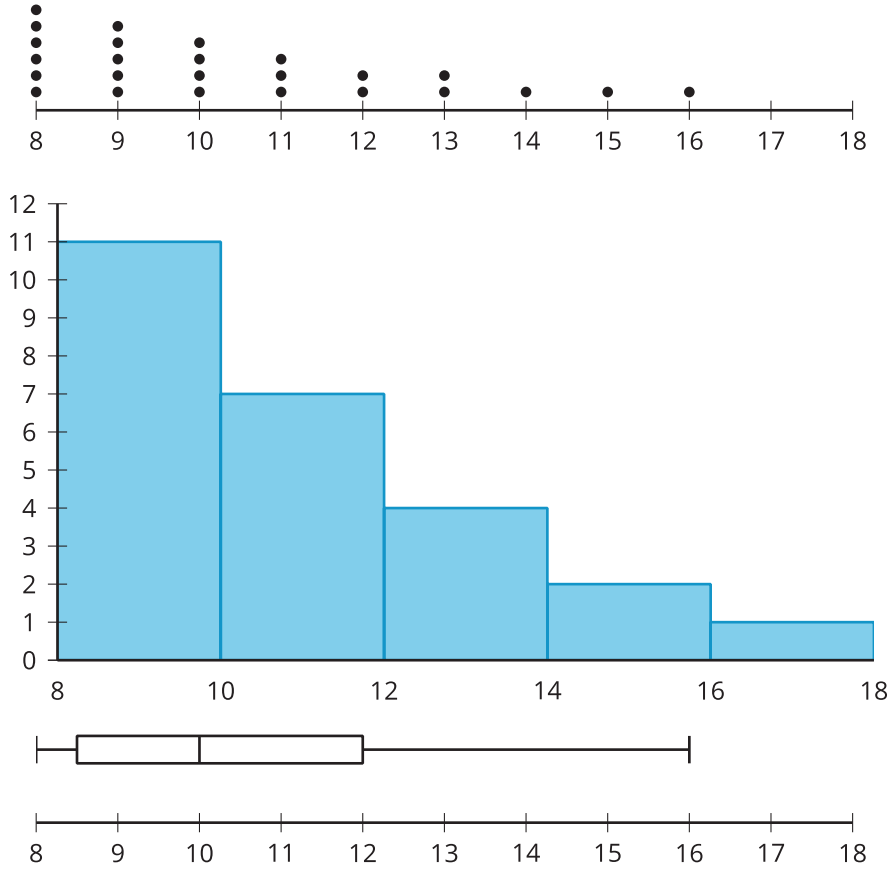
Lesson 4 Summary

We can describe the shape of distributions as *symmetric*, *skewed*, *bell-shaped*, *bimodal*, or *uniform*. The dot plot, histogram, and box plot shown here represent the distribution of the same data set. This data set has a symmetric distribution.



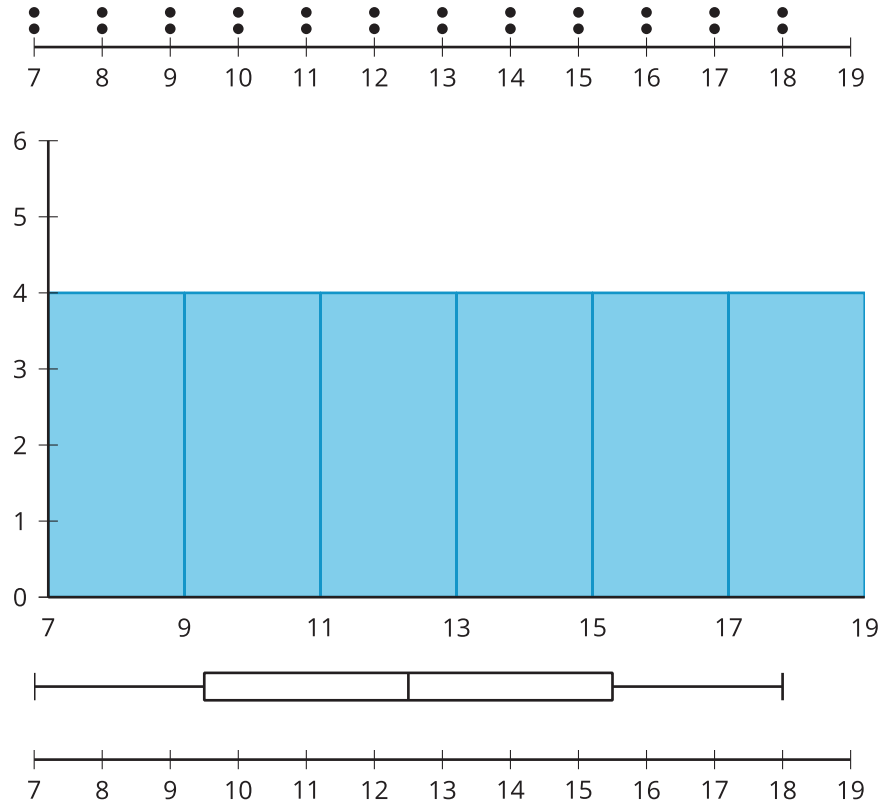
In a **symmetric distribution**, the mean is equal to the median and there is a vertical line of symmetry in the center of the data display. The histogram and the box plot both group data together. Since histograms and box plots do not display each data value individually, they do not provide information about the shape of the distribution to the same level of detail that a dot plot does. This distribution, in particular, can also be called bell-shaped. A **bell-shaped distribution** has a dot plot that takes the form of a bell with most of the data clustered near the center and fewer points farther from the center. This makes the measure of center a very good description of the data as a whole. Bell-shaped distributions are always symmetric or close to it.

The dot plot, histogram, and box plot shown here represent a skewed distribution.



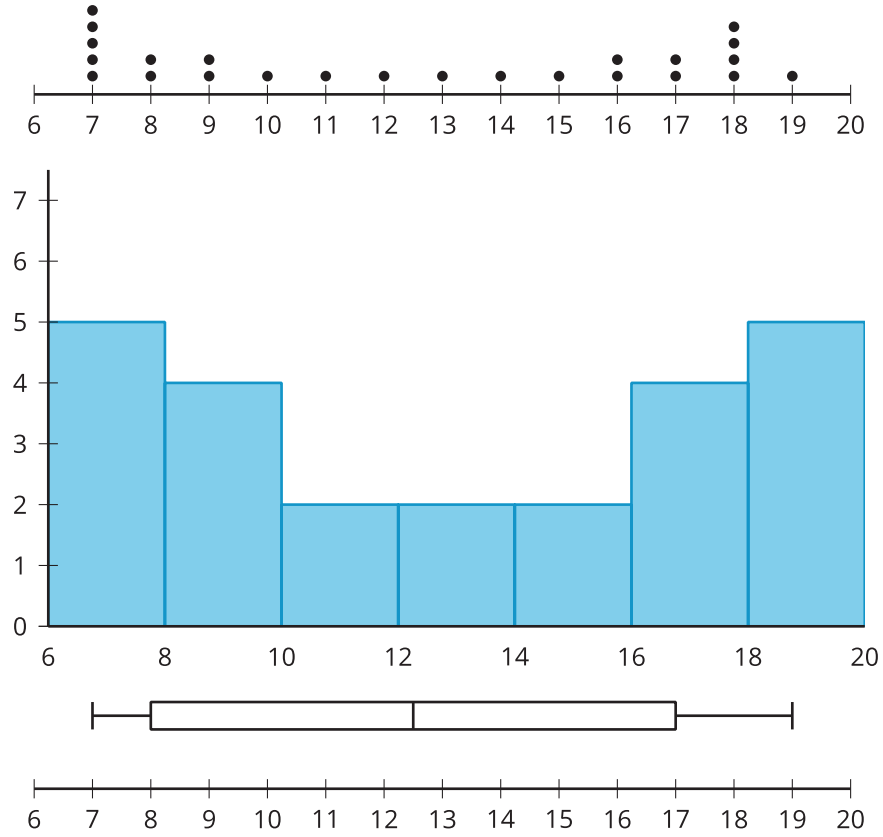
In a **skewed distribution**, one side of the distribution has more values farther from the bulk of the data than the other side has. This results in the mean and median not being equal. In this skewed distribution, the data are skewed to the right because most of the data are near the 8 to 10 interval, but there are many points to the right. The mean is greater than the median. The large data values to the right cause the mean to shift in that direction while the median remains with the bulk of the data. So, the mean is greater than the median for distributions that are skewed to the right. In a data set that is skewed to the left, a similar effect happens but to the other side. Again, the dot plot provides a greater level of detail about the shape of the distribution than do either the histogram or the box plot.

A **uniform distribution** has the data values evenly distributed throughout the range of the data. This causes the distribution to look like a rectangle.



In a uniform distribution, the mean is equal to the median because a uniform distribution is also a symmetric distribution. The box plot does not provide enough information to describe the shape of the distribution as uniform, though the even length of each quarter does suggest that the distribution may be approximately symmetric.

A **bimodal distribution** has two very common data values seen in a dot plot or histogram as distinct peaks.



Sometimes, a bimodal distribution has most of the data clustered in the middle of the distribution. In these cases, the center of the distribution does not describe the data very well. Bimodal distributions are not always symmetric. For example, the peaks may not be equally spaced from the middle of the distribution, or other data values may disrupt the symmetry.