Unit 8, Lesson 4

The Mean

Let’s explore the mean of a data set and what it tells us.

4.1 Which Three Go Together: Division

Which three go together? Why do they go together?

A

$$\frac{5+5+5+5}{4}$$

B

$$\frac{10+6+4}{4}$$

C

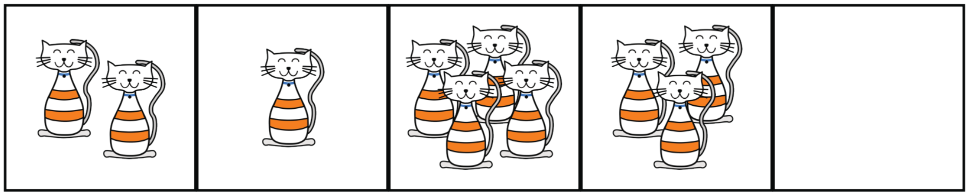
$$\frac{10+8+6+4}{4}$$

D

$$\frac{7+6+4+2+1}{5}$$

4.2 Spread Out and Share

1. The stuffed toy kittens in a preschool room are placed in 5 boxes.



   a. The preschool teacher wants the kittens distributed equally among the boxes. How might that be done? How many kittens will end up in each box?

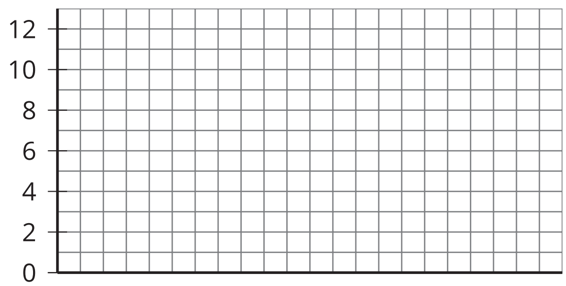
   b. The number of kittens in each box after they are equally distributed is called the **mean** number of kittens per box, or the **average** number of kittens per box. Explain how the expression $10 \div 5$ is related to the average.

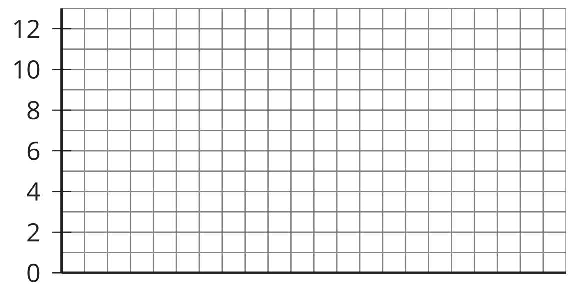
   c. Another preschool room has 6 boxes. No 2 boxes have the same number of kittens, and there is an average of 3 kittens per box. Draw or describe at least 2 different arrangements of kittens that match this description.

2. Five servers are scheduled to work the number of hours shown. They decide to share the workload, so each one would work equal hours.

Server A: 3 Server B: 6 Server C: 11 Server D: 7 Server E: 4

a. On the first grid, draw 5 bars whose heights represent the hours worked by Servers A, B, C, D, and E.





Then, think about how you would rearrange the hours so that each server gets a fair share. On the second grid, draw a new graph to represent the rearranged hours. Be prepared to explain your reasoning.

b. Based on your second drawing, what is the average, or mean, number of hours that the servers will work?

c. Explain why we can also find the mean by finding the value of the expression $31 \div 5$.

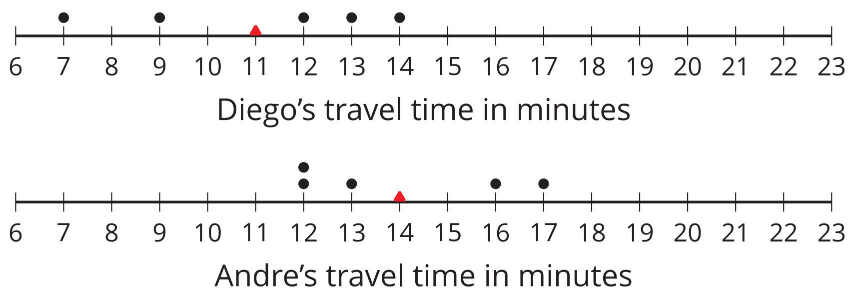
d. Which server will see the biggest change to work hours? Which server will see the least change?

Are you ready for more?

Server F, working 7 hours, offers to join the group of five servers, sharing their workload. If server F joins, will the mean number of hours worked increase or decrease? Explain how you know.

4.3 Travel Times (Part 2)

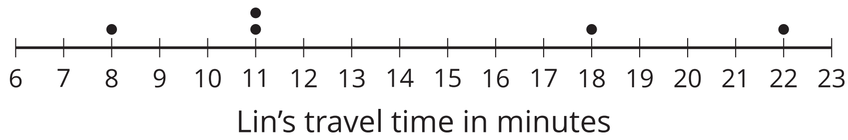
1. Here are dot plots showing how long Diego's trips to school took in minutes and how long Andre's trips to school took in minutes. The dot plots include the means for each data set, and those means are marked by triangles.



   a. Which of the two data sets has a larger mean? In this context, what does a larger mean tell us?

   b. Which of the two data sets has larger sums of distances to the left and right of the mean? What do these sums tell us about the variability in Diego's and Andre's travel times?

2. Here is a dot plot showing lengths of Lin's trips to school.



a. Calculate the mean of Lin's travel times.

b. Complete the table with the distance between each point and the mean as well whether the point is to the left or right of the mean.

time in minutes	**distance from the mean**	**left or right of the mean?**
22		
18		
11		
8		
11		

c. Find the sum of distances to the left of the mean and the sum of distances to the right of the mean.

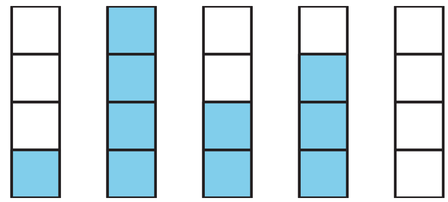
d. Use your work to compare Lin's travel times to Andre's. What can you say about their average travel times? What about the variability in their travel times?

Lesson 4 Summary

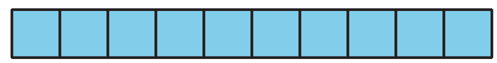
Sometimes a general description of a distribution does not give enough information, and a more precise way to talk about center or spread would be more useful. The **mean**, or **average**, is a number we can use for the center to summarize a distribution.

We can think about the mean in terms of "fair share" or "leveling out." That is, a mean can be thought of as a number that each member of a group would have if all the data values were combined and distributed equally among the members.

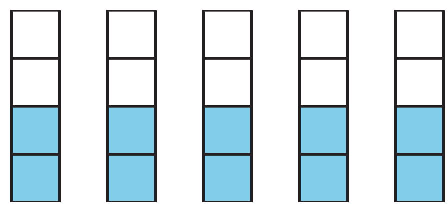
For example, suppose there are 5 containers, each of which has a different amount of water: 1 liter, 4 liters, 2 liters, 3 liters, and 0 liters.



To find the mean, first we add up all of the values. We can think of this as putting all of the water together: $1 + 4 + 2 + 3 + 0 = 10$.



To find the "fair share," we divide the 10 liters equally into the 5 containers: $10 \div 5 = 2$.

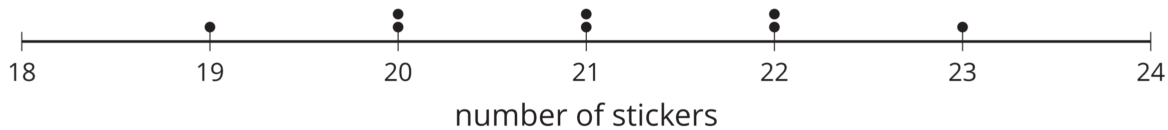


In general, to find the mean of a data set with n values, we add all of the values and divide the sum by n.

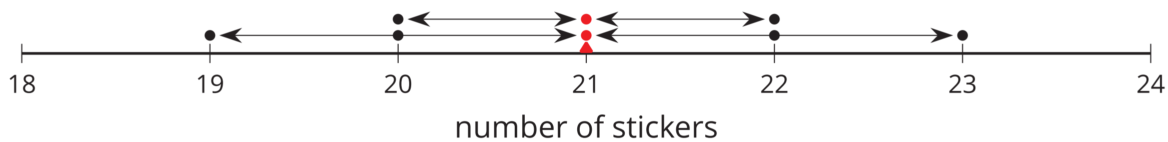
The mean is often used as a **measure of center** of a distribution. One way to see this is that the mean of a distribution can be seen as the "balance point" for the distribution. Why is this a good way to think about the mean? Let's look at a very simple set of data on the number of stickers that are on 8 pages:

19 20 20 21 21 22 22 23

Here is a dot plot showing the data set.



The distribution shown is completely symmetrical. The mean number of stickers is 21, because $(19 + 20 + 20 + 21 + 21 + 22 + 22 + 23) \div 8 = 21$. If we mark the location of the mean on the dot plot, we can see that the data points could balance at 21.



Even when a distribution is not completely symmetrical, the distances of values below the mean, on the whole, balance the distances of values above the mean.

