Unit 8, Lesson 5

Variability and MAD

Let's use mean and MAD to describe and compare distributions.

5.1 Shooting Hoops (Part 1)

Elena, Jada, and Lin enjoy playing basketball during recess. Lately, they have been practicing free throws. They record the number of baskets they make out of 10 attempts. Here are their data sets for 12 school days.

Elena

2	2	2	2	4	5	5	6	8	9	9	9

Jada

2	4	5	4	6	6	4	7	3	4	8	7

Lin

3	6	6	4	5	5	3	5	4	6	6	7

1. Calculate the mean number of baskets each player made, and compare the means. What do you notice?

2. What do the means tell us in this context?

5.2 Shooting Hoops (Part 3)

The tables show the number of baskets made by Jada and Lin in several games. Recall that the mean of Jada and Lin's data is 5.

1. Record the distance between the number of baskets Jada made in each game and the mean.

Jada	2	4	5	4	6	6	4	7	3	4	8	7
distance from 5												

Now find the average of the distances in the table. Show your reasoning, and round your answer to the nearest tenth.

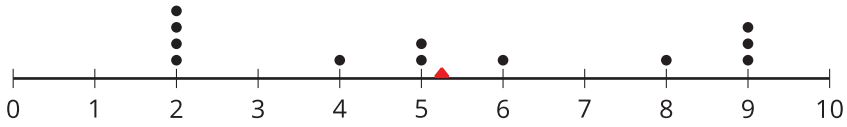
This value is the mean absolute deviation (MAD) of Jada's data. Jada's MAD: _____

2. Find the mean absolute deviation of Lin's data. Round it to the nearest tenth.

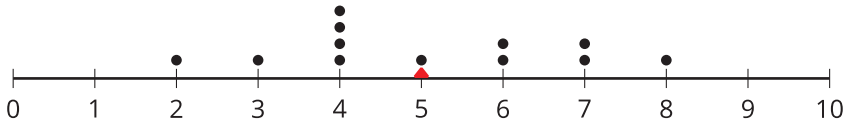
Lin	3	6	6	4	5	5	3	5	4	6	6	7
distance from 5												

Lin's MAD: _____

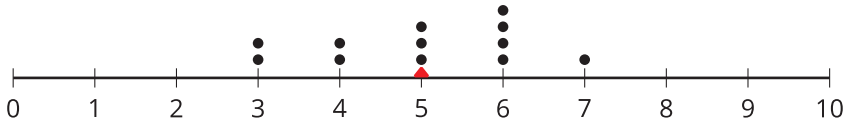
3. Elena's distribution has a MAD of about 2.5. Compare the MADs and dot plots of the three students' data. Do you see a relationship between each student's MAD and the distribution on her dot plot? Explain your reasoning.



number of baskets Elena made



number of baskets Jada made



number of baskets Lin made

Are you ready for more?

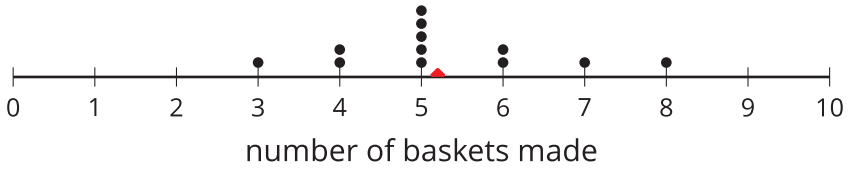
Invent another data set that also has a mean of 5 but has a MAD greater than 2. Remember, the values in the data set must be whole numbers from 0 to 10.

5.3 Which Player Would You Choose?

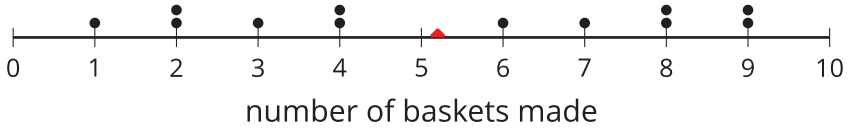
1. Andre and Noah joined Elena, Jada, and Lin in recording their basketball scores. They all record their scores in the same way: the number of baskets made out of 10 attempts. Each person collects 12 data points.
   - Andre's mean number of baskets is 5.25, and his MAD is 2.6.
   - Noah's mean number of baskets is also 5.25, but his MAD is 1.

   Here are two dot plots that represent the two data sets. The triangle indicates the location of the mean.

   Data set A



   Data set B



   a. Without calculating, decide which dot plot represents Andre's data and which represents Noah's. Explain how you know.

   b. If you are the captain of a basketball team and can use 1 more player on your team, do you choose Andre or Noah? Explain your reasoning.

2. An eighth-grade student decides to join Andre and Noah and keeps track of his scores. His data set is shown here. The mean number of baskets he makes is 6.

eighth-grade student	6	5	4	7	6	5	7	8	5	6	5	8
distance from 6												

a. Calculate the MAD. Show your reasoning.

b. Draw a dot plot to represent his data and mark the location of the mean with a triangle.

c. Compare the eighth-grade student's mean and MAD to Noah's mean and MAD. What do you notice?

d. Compare their dot plots. What do you notice about the distributions?

e. What can you say about the two players' shooting accuracy and consistency?

Are you ready for more?

Invent a data set with a mean of 7 and a MAD of 1.

5.4 Swimmers over the Years

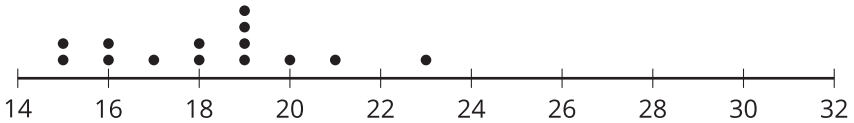
The mean age of swimmers on a 1984 national swim team is 18.2 years and the MAD is 2.2 years. The mean age of the swimmers on the 2016 team is 22.8 years, and the MAD is 3 years.

1. How has the average age of the swimmers on the national team changed from 1984 to 2016? Explain your reasoning.

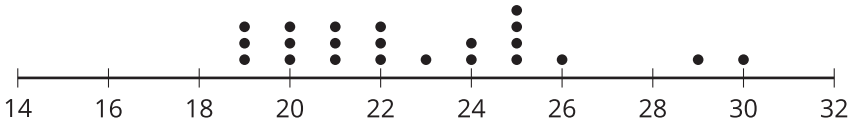
2. Are the swimmers on the 1984 team closer in age to one another than the swimmers on the 2016 team are to one another? Explain your reasoning.

3. Here are dot plots showing the ages of the swimmers on the national swim teams in 1984 and in 2016. Use them to make two other comments about how the team has changed over the years.

1984



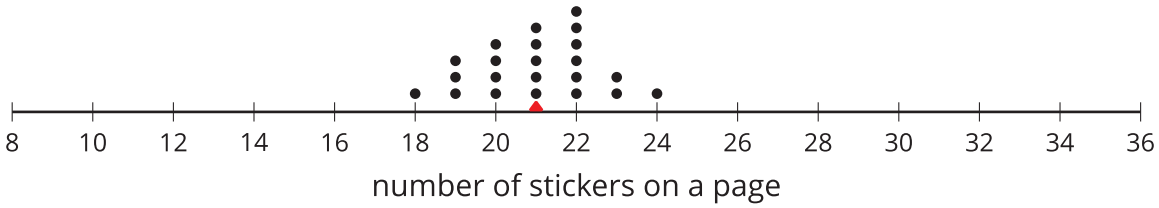
2016



Lesson 5 Summary

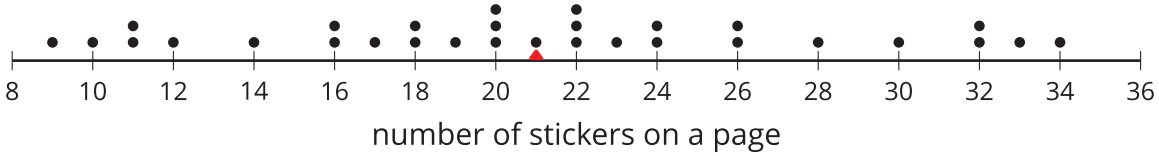
We use the mean of a data set as a "measure of center" of its distribution, but two data sets with the same mean could have very different distributions.

This dot plot shows the number of stickers on each page of a 22-page sticker book.



The mean number of stickers is 21. All the pages have within 3 stickers of the mean, and most of them are even closer. These pages are all fairly close in the number of stickers on them.

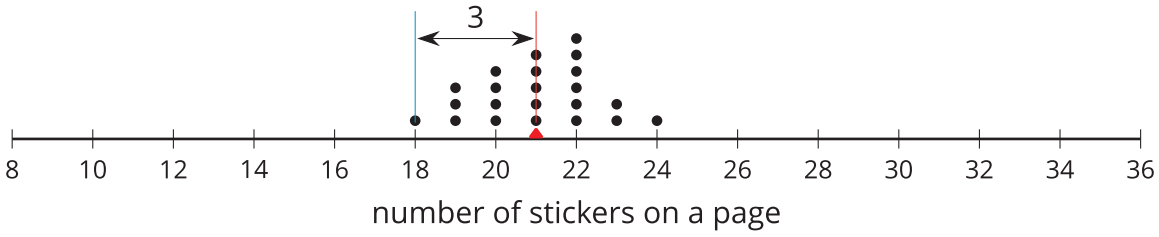
This dot plot shows the number of stickers on each page of another sticker book that has 30 pages.



In this sticker book, the mean number of stickers on each page is also 21, but some pages have less than half that number of stickers and others have more than one-and-a-half times as many. There is a lot more variability in the number of stickers.

There is a number that we can use to describe how far away, or how spread out, data points generally are from the mean. This *measure of spread* is called the **mean absolute deviation (MAD)**.

To find the MAD, we find the distance between each data value and the mean, and then calculate the mean of those distances. For instance, the point that represents 18 stickers is 3 units away from the mean of 21 stickers.



We can find the distance between each point and the mean of 21 stickers and then organize the distances into a table, as shown.

number of stickers	18	19	19	19	20	20	20	20	21	21	21	21	21	22	22	22	22	22	22	23	23	24
distance from mean	3	2	2	2	1	1	1	1	0	0	0	0	0	1	1	1	1	1	1	2	2	3

The values in the first row of the table are the number of stickers on each page in the first book. Their mean, 21, is the mean number of stickers on a page.

The values in the second row of the table are the distances, or absolute deviation, between the values in the first row and 21. The mean of these distances is the MAD of the number of stickers on a page, about 1.2 stickers.

What can we learn from the averages of these distances once they are calculated?

- In the first book, the distances are all between 0 and 3. The MAD is 1.2 stickers, which tells us that the number of stickers are typically within 1.2 of the mean number, 21. We could say that a typical page has between 19.8 and 22.2 stickers.
- In the second book, the distances are all between 0 and 13. The MAD is 5.6 stickers, which tells us that the number of stickers are typically within 5.6 of the mean number, 21. We could say that a typical page has between 15.4 and 26.6 stickers.

The MAD is also called a *measure of the variability* of the distribution. In these examples, it is easy to see that a higher MAD suggests a distribution that is more spread out, showing more variability.

In summary, a measure of center, such as the mean, gives a sense of what is typical for a set of data. A measure of variability, such as the MAD, gives a sense of how consistent the data are.