Lesson 9: The Mean Absolute Deviation (MAD)

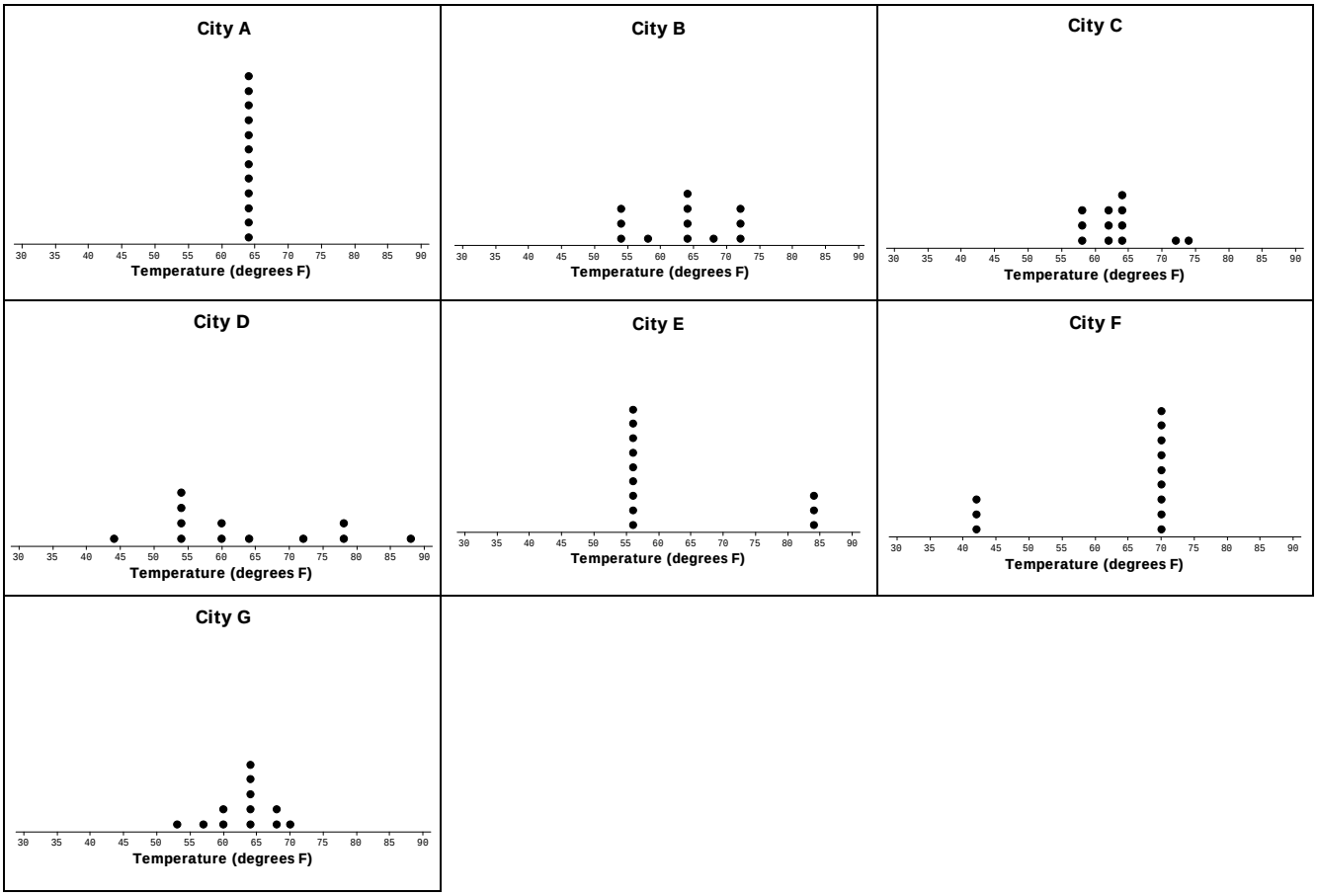
Classwork

Example 1: Variability

In Lesson 8, Robert tried to decide to which of two cities he would rather move, based on comparing their mean annual temperatures. Since the mean yearly temperature for New York City and San Francisco turned out to be about the same, he decided instead to compare the cities based on the variability in their monthly temperatures from the overall mean. He looked at the two distributions and decided that the New York City temperatures were more spread out from their mean than were the San Francisco temperatures from their mean.

Exercises 1–3

The following temperature distributions for seven other cities all have a mean temperature of approximately 63 degrees. They do not have the same variability. Consider the following dot plots of the mean yearly temperatures of the seven cities in degrees Fahrenheit.



1. Which distribution has the smallest variability of the temperatures from its mean of $\mathbf{63}$ degrees? Explain your answer.

2. Which distribution(s) seems to have the most variability of the temperatures from the mean of $\mathbf{63}$ degrees? Explain your answer.

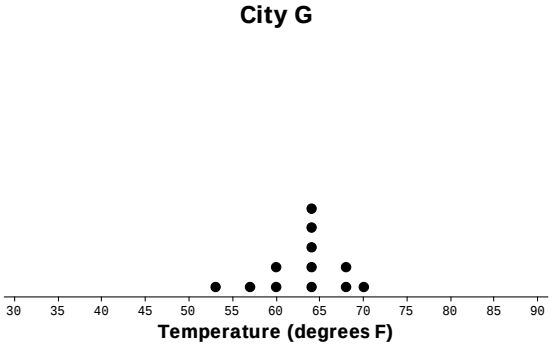
3. Order the seven distributions from least variability to most variability. Explain why you listed the distributions in the order that you chose.

Example 2: Measuring Variability

Based on just looking at the distributions, there are different orderings of variability that seem to make some sense. Sabina is interested in developing a formula that will give a number that measures the variability in a data distribution. She would then use the formula for each data set and order the distributions from lowest to highest. She remembers from a previous lesson that a deviation is found by subtracting the mean from a data point. The formula was summarized as: deviation = data point – mean. Using deviations to develop a formula measuring variability is a good idea to consider.

Exercises 4–6

The dot plot for the temperatures in City G is shown below. Use the dot plot and the mean temperature of 63 degrees to answer the following questions.

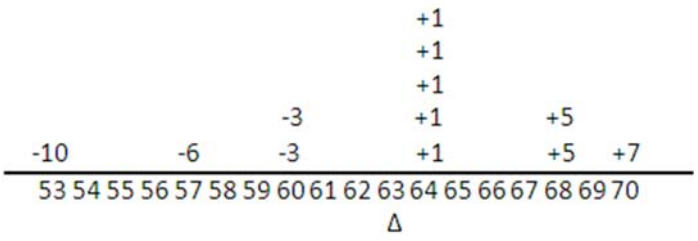


4. Fill in the following table for City G temperature deviations.

Temp	Deviation	Result
53	$53-63$	-10
57	$57-63$	-6
60	$60-63$	-3
60	$60-63$	-3
64	$64-63$	$+1$
64	$64-63$	$+1$
64		
64		
64		
68		
68		
70		
Sum		

5. Why should the sum of your deviations column be equal to zero? (Hint: Recall the balance interpretation of the mean of a data set.)

6. Another way to graph the deviations is to write them on a number line as follows. What is the sum of the positive deviations (the deviations to the right of the mean)? What is the sum of the negative deviations (the deviations to the left of the mean)? What is the total sum of the deviations?



Example 3: Finding the Mean Absolute Deviation (MAD)

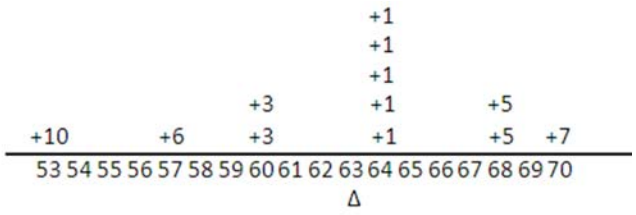
By the balance interpretation of the mean, the sum of the deviations for any data set will always be zero. Sabina is disappointed that her idea of developing a measure of variability using deviations isn't working. She still likes the concept of using deviations to measure variability, but the problem is that the sum of the positive deviations is cancelling out the sum of the negative deviations. What would you suggest she do to keep the deviations as the basis for a formula but to avoid the deviations cancelling out each other?

Exercises 7–8

7. One suggestion to possibly help Sabina is to take the absolute value of the deviations.
   a. Fill in the following table.

Temp	Deviation	Result	Abs
53	$53-63$	-10	$+10$
57	$57-63$	-6	$+6$
60	$60-63$	-3	$+3$
60	$60-63$	-3	$+3$
64	$64-63$	$+1$	$+1$
64	$64-63$	$+1$	$+1$
64			
64			
64			
68			
68			
70			

   b. From the following graph, what is the sum of the absolute deviations?



c. Sabina suggests that the mean of the absolute deviations could be a measure of the variability in a data set. Its value is the average distance that all the data values are from the mean temperature. It is called the Mean Absolute Deviation and is denoted by the letters, MAD. Find the MAD for this data set of City G temperatures. Round to the nearest tenth.

d. Find the MAD for each of the temperature distributions in all seven cities, and use the values to order the distributions from least variability to most variability. Recall that the mean for each data set is 63 degrees. Does the list that you made in Exercise 2 by just looking at the distributions match this list made by ordering MAD values?

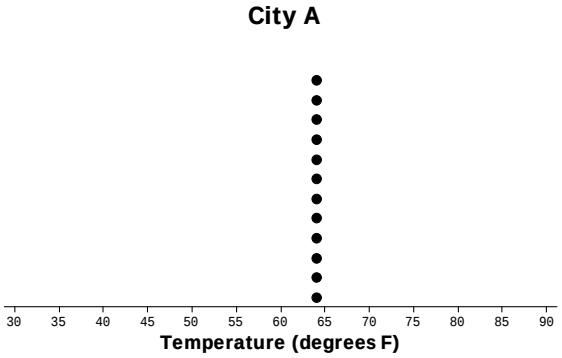
e. Which of the following is a correct interpretation of the MAD?

i. The monthly temperatures in City G are spread 3.7 degrees from the approximate mean of 63 degrees.

ii. The monthly temperatures in City G are, on average, 3.7 degrees from the approximate mean temperature of 63 degrees.

iii. The monthly temperatures in City G differ from the approximate mean temperature of 63 degrees by 3.7 degrees.

8. The dot plot for City A temperatures follows.



a. How much variability is there in City A's temperatures? Why?

b. Does the MAD agree with your answer in part (a)?

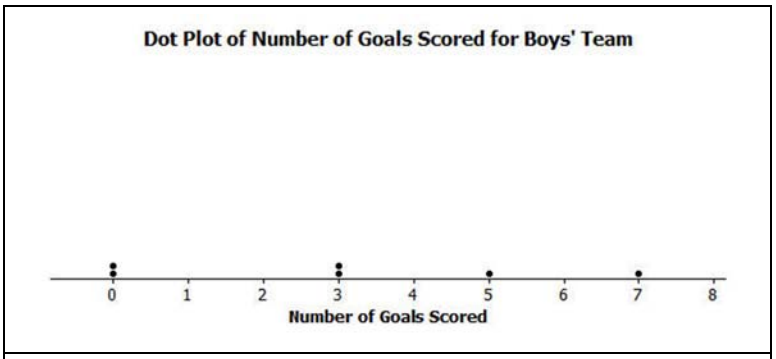
Lesson Summary

In this lesson, a formula was developed that measures the amount of variability in a data distribution.

- The absolute deviation of a data point is how far away that data point is from the mean.
- The Mean Absolute Deviation (MAD) is computed by finding the mean of the absolute deviations in the distribution.
- The value of MAD is the average distance that all the data values are from the mean.
- A small MAD indicates that the distribution has very little variability.
- A large MAD indicates that the data points are spread far away from the mean.

Problem Set

1. Suppose the dot plot on the left shows the number of goals a boys' soccer team has scored in six games so far this season, and the dot plot on the right shows the number of goals a girls' soccer team has scored in six games so far this season. The mean for both of these teams is 3.



   a. Before doing any calculations, which dot plot has the larger MAD? Explain how you know.

b. Use the following tables to find the MAD number of goals for each distribution. Round your calculations to the nearest hundredth.

Boys' Team		
#Goals	Deviations	Absolute Deviations
0	-3	
0	-3	
3	$3-3=0$	
3		
5		
7		
Sum		

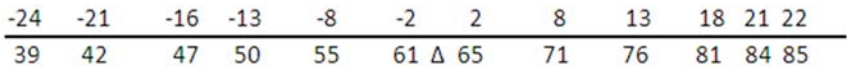
Girls' Team		
#Goals	Deviations	Absolute Deviations
2		
2		
3		
3		
3		
5		
Sum		

c. Based on the computed MAD values, for which distribution is the mean a better indication of a typical value? Explain your answer.

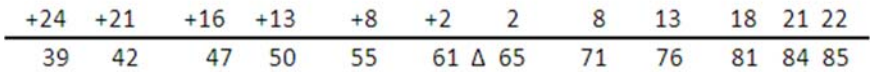
2. Recall Robert's problem of deciding whether to move to New York City or to San Francisco. The table of temperatures (in degrees Fahrenheit) and deviations for the New York City distribution is as follows:

NYC	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Temp	39	42	50	61	71	81	85	84	76	65	55	47
Deviation	-24	-21	-13	-2	8	18	22	21	13	2	-8	-16

a. The dot plot below is written with the deviations above each of the monthly temperatures. What is the sum of all of the deviations? Are you surprised? Explain.



b. The absolute deviations for the monthly temperatures are shown below. Use this information to calculate the MAD. Explain the MAD in words for this problem.



c. Complete the following table and then use the values to calculate the MAD for the San Francisco data distribution.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Temp	57	60	62	63	64	67	67	68	70	69	63	58
Deviations				-1	0	$+3$						
Absolute Deviations												

d. Comparing the MAD values for New York City and San Francisco, which city would Robert choose to move to if he is interested in having a lot of variability in monthly temperatures? Explain using the MAD.

3. Consider the following data of the number of green jellybeans in seven bags sampled from five different candy manufacturers (Awesome, Delight, Finest, Sweeties, YumYum). Note that the mean in each distribution is 42 green jellybeans.

	Bag 1	Bag 2	Bag 3	Bag 4	Bag 5	Bag 6	Bag 7
Awesome	40	40	41	42	42	43	46
Delight	22	31	36	42	48	53	62
Finest	26	36	40	43	47	50	52
Sweeties	36	39	42	42	42	44	49
YumYum	33	36	42	42	45	48	48

a. Complete the following table of the deviations of the number of green jellybeans from the mean number of green jellybeans in the seven bags.

	Bag 1	Bag 2	Bag 3	Bag 4	Bag 5	Bag 6	Bag 7
Awesome	-2	-2	-1	0	0	$+1$	$+4$
Delight	-20	-11	-6				
Finest	-16						
Sweeties							
YumYum							

b. Based on what you learned about MAD, which manufacturer do you think will have the lowest MAD? Calculate the MAD for the manufacturer you selected.