

## 8.5

# Putting the Pieces Together

## Analyzing and Interpreting Data

### LEARNING GOALS

In this lesson, you will:

- Analyze and interpret data graphically and numerically.
- Determine which measure of central tendency and spread is most appropriate to describe a data set.

### KEY TERMS

- stem-and-leaf plot
- side-by-side stem-and-leaf plot

Taking a trip on an airplane is always exciting. However, the process of flying can sometimes be frustrating. One of the most challenging tasks is boarding the plane before take-off. The most common method used to board passengers is boarding people by zone or row so that passengers in the back of the plane board first. This seems like it should be the most efficient way to board because people in the front won't be blocking the way. However, this is not necessarily the case. An astrophysicist used a computer simulation to try and determine the best method for loading passengers. After many simulations he found that passengers in even-numbered window seats near the back should board first, followed by even-numbered window seats in the middle, and even-numbered window seats in the front. This trend then continues through even-numbered middle seats, and even-numbered aisle seats. The whole process is then repeated with odd numbered seats.

So why does this work? It seems that allowing passengers a row between each other gives them more space to load their luggage and allows them to move if a passenger needs to get past them. This is not the only method that works, but it is the simplest for passengers to understand. Do you think airlines should try to change their methods for loading to this one? How much time do you really think it would save?

## 8

**PROBLEM 1** Go For the Gold

When a participant takes part in the Special Olympics, each person receives a number. The chart shown represents the first twenty people labeled by their participation number and the number of gold medals each participant won.

Participation Number	Gold Medals Won
001	6
002	14
003	1
004	6
005	0
006	0
007	9
008	1
009	1
010	9
011	5
012	10
013	1
014	2
015	2
016	5
017	4
018	3
019	4
020	2

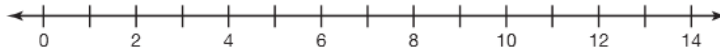
© 2012 Carnegie Learning



1. Analyze the data. Calculate the mean and standard deviation, and then interpret the meaning of each in terms of this problem situation.



2. Construct a box-and-whisker plot of the data and include any outliers.



3. Interpret the IQR.

4. Which measure of central tendency and spread should you use to describe this data? Explain your reasoning.

5. What conclusions can you draw about the number of gold medals participants won?

8

6. Shelly states the median and standard deviation should be used to describe the data because the standard deviation is less than the IQR. Is Shelly correct? Explain why or why not.



**PROBLEM 2** Flying High



Data were collected from two rival airlines measuring the difference in the stated departure times, and the times the flights *actually* departed. The average departure time differences were recorded for each month for one year. The results are shown in the *side-by-side stem-and-leaf plot* given.

Difference in Departure Times (minutes)						
My Air Airlines				Fly High Airlines		
	5	0	0	7	8	
9	5	1	1	4	5	6
6	0	0	2	4	7	9
4	3	3	3	0	2	
		0	4	5	9	

2|4 = 24 minutes

A **stem-and-leaf plot** is a graphical method used to represent ordered numerical data. Once the data is ordered, the stem and leaves are determined. Typically, the stem is all the digits in a number except the right most digit, which is the leaf. A **side-by-side stem-and-leaf plot** allows a comparison of two data sets. The two data sets share the same stem, but have leaves to the left and right of the stem.

Oh I remember stem-and-leaf plots! There should be a key somewhere which represents the value of each data point.



1. Describe the distribution of each data set.



© 2012 Carnegie Learning

2. Based on the shape of the data, calculate an appropriate measure of central tendency and spread for each data set.

8

3. What conclusions can you draw from the measure of central tendency and spread you calculated?



4. You are scheduling a flight for an important meeting and you must be there on time. Which airline would you schedule with? Explain your reasoning.

## 8

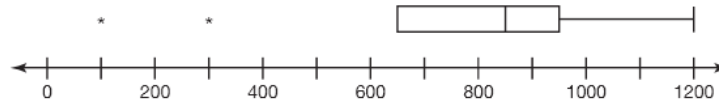
## Talk the Talk

When analyzing data it is important to use both graphs and numbers to describe the data.

- The mean describes the average data point.
- The median describes the middle data point.
- Standard deviation describes the spread of the data from the mean.
- The interquartile range (IQR) describes the spread of the data from the median.
- For data that is symmetric, the mean is the most appropriate measure of central tendency and the standard deviation is the most appropriate measure of spread.
- For data that is skewed, the median is the most appropriate measure of central tendency and the IQR is the most appropriate measure of spread.



1. Analyze the box-and-whisker plot shown.



- a. Amina's teacher wants her students to create a list of data values that could result in the box plot shown. Amina states that she can just use the data values graphed as her list. She lists 100, 300, 700, 850, 950, and 1200 as her list. Is Amina's thinking correct? If yes, will this work for all box-and-whisker plots. If no, explain why not.

- b. Create a list of values that when graphed would result in the given box-and-whisker plot shown.

- c. Describe the data using an appropriate measure of central tendency and spread.

2. A data set ranges from 10 to 20. A value of 50 is added to the data set.
- a. Explain how the mean and median are affected by this new value.



- b. Which measure of central tendency and spread would you use to describe the original data set before the new value is added? Explain your reasoning.
- c. Which measure of central tendency and spread would you use to describe the data set after the new value is added? Explain your reasoning.



Be prepared to share your solutions and methods.