



## Lesson Overview

In this TI-Nspire lesson, students use their previous work on describing a single data distribution to compare several distributions with respect to the same attribute when the data were collected from different populations through a census.



Graphs and summary statistics can be used to make comparisons among distributions of the measurements of the same characteristic for different populations.

## Learning Goals

1. Associate a graphical representation of a set of data with measures of center and spread;
2. use statistical calculations of center and spread to compare different distributions of the same measurement collected from different populations;
3. identify the advantages and disadvantages of using different graphical representations to compare distributions of the same measurement collected from different populations;
4. recognize bar graphs are limited as tools for analyzing data but can be used to compare individuals or categories with respect to a given measurement (including frequencies).

## Prerequisite Knowledge

*Comparing Distributions* is the eleventh lesson in a series of lessons that investigates the statistical process. In this lesson, students compare distributions collected from different populations. Prior to working on this lesson students should have completed *Introduction to Data*. Students should understand:

- how to interpret data represented on bar graphs and histograms;
- how to compare different measures of center and spread for a given distribution.



### Vocabulary

- **symmetric:** when one side is the exact image or reflection of the other
- **skewed:** data that clusters towards one end of a graphical display
- **mound shaped:** data that clusters towards the middle of a graphical display
- **bimodal:** data distribution having two equal, most common values
- **median:** the value that separates the upper half of the distribution of a set of data values from the lower half
- **mean:** the sum of all the data values in a set of data divided by the number of data values
- **outlier:** a value that lies outside most of the other values in a set of data
- **interquartile range:** the difference between the upper quartile and the lower quartile
- **mean absolute deviation:** the mean of the absolute values of all deviations from the mean of a set of data

### Lesson Pacing

This lesson should take 50–90 minutes to complete with students, though you may choose to extend, as needed.

### Lesson Materials

- Compatible TI Technologies:



TI-Nspire CX Handhelds,



TI-Nspire Apps for iPad®,



TI-Nspire Software

- Comparing Distributions\_Student.pdf
- Comparing Distributions\_Student.doc
- Comparing Distributions.tns
- Comparing Distributions \_Teacher Notes
- To download the TI-Nspire activity (TNS file) and Student Activity sheet, go to <http://education.ti.com/go/buildingconcepts>.



## Class Instruction Key

The following question types are included throughout the lesson to assist you in guiding students in their exploration of the concept:



**Class Discussion:** Use these questions to help students communicate their understanding of the lesson. Encourage students to refer to the TNS activity as they explain their reasoning. Have students listen to your instructions. Look for student answers to reflect an understanding of the concept. Listen for opportunities to address understanding or misconceptions in student answers.



**Student Activity:** Have students break into small groups and work together to find answers to the student activity questions. Observe students as they work and guide them in addressing the learning goals of each lesson. Have students record their answers on their student activity sheet. Once students have finished, have groups discuss and/or present their findings. The student activity sheet can also be completed as a larger group activity, depending on the technology available in the classroom.



**Deeper Dive:** These questions are provided for additional student practice and to facilitate a deeper understanding and exploration of the content. Encourage students to explain what they are doing and to share their reasoning.



## Mathematical Background

Data can consist of measurements collected from an entire population, or they can be measurements obtained from a sample drawn from the population. Measurements collected from the entire population are typically called a *census*, a complete count of everyone or everything in the population. A *sample* is a subset of the population, typically selected to represent the population. A census is conducted when it is possible and important to count everyone, for example in considering the number of people in a country or region. The first censuses were conducted in 3800 BC by the Babylonians to record the number of people and their property. By an act of the first Congress, the US conducts a census of everyone in the country every 10 years.

If all measurements in a population are known through a census, it is still important to consider variability, so any descriptions of a distribution of data should involve calculated measures of center and spread. Thus, the results of a survey of every student in your school about how much time they spent reading for fun last month would typically be reported in terms of the mean and a deviation from that mean or the median and interquartile range.

The data on life expectancy, per capita income and literacy rate come from the sources below, typically collected from 2008 to 2013, and may vary from other sources, as collecting and reporting these kinds of data occur in different ways and at different times. The general patterns should be the same, while some individual values might be different. Note that these data are not summary data for an entire region but for countries within the region. Note also that life expectancy is the average number of years a person born can expect to live if the conditions were to remain the same. The literacy rate is the number of people over the age of 15 who can read and write. Comparisons of national income are typically calculated based on a mean and using a method involving purchasing power parity (PPP) to adjust for differences in the purchasing power for an equivalent value of the currency in different countries.

### Data Sources:

Countries by GDP per capita

CIA World Factbook (2013):

<https://www.cia.gov/library/publications/the-world-factbook/rankorder/2004rank.html>

United Nations Dept. of Economic and Social Affairs:

<http://esa.un.org/unpd/wpp/Download/Standard/Mortality/>

<http://www.nationmaster.com/country-info/stats/Education/Average-years-of-schooling-of-adults>

<http://www.indexmundi.com/factbook/countries>

UNESCO:

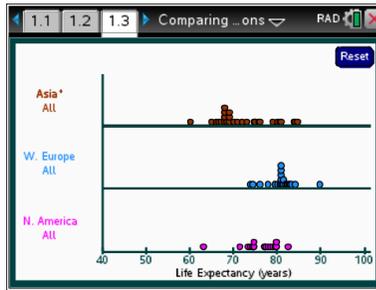
[http://data.uis.unesco.org/Index.aspx?DataSetCode=EDULIT\\_DS&popupcustomise=true&lang=en](http://data.uis.unesco.org/Index.aspx?DataSetCode=EDULIT_DS&popupcustomise=true&lang=en)



### Part 1, Page 1.3

Focus: Dot plots, box plots, and histograms can be used to compare the shape, center and spread of different populations with respect to a characteristic such as life expectancy.

Page 1.3 shows dot plots displaying a single attribute of up to 3 regions of the world.



**Region** chooses a region of the world and the gender of interest.

**Attribute** chooses the characteristic of interest.

The **up/down arrows** move from line to line.

Selecting a point or tabbing through a dot plot reveals the country associated with that point.

**Graph type** chooses among dot plot, boxplot, and histogram.

**Bin width** allows the bin width to be changed when a histogram is displayed.

**Clear** produces a blank screen.

**Reset** returns to the original screen.

### TI-Nspire Technology Tips

**menu** accesses page options.

**+** Adds a new region or changes the active graph.

**tab** cycles through points on a dot plot.

**enter** selects a point on the graph.

**esc** lets go of all selected points and closes the dropdown menus.

**ctrl del** resets the page.



### Class Discussion

*Life expectancy is the average number of years a person born can expect to live if the conditions were to remain the same. The plots show the life expectancy for countries in three different regions of the world.*

- **What strikes you as interesting?**

Answers will vary. Life expectancy for countries in Western Europe seems to be the longest and those countries also seem to have the least variability among them. The countries in North America seem to have the next longest life expectancies, but the distribution of life expectancies for countries in Asia has the largest range with some countries having a very long life expectancy and others below 70 years.

- **Identify the country you think has the smallest life expectancy, and the greatest.**

Answer: Afghanistan has the smallest, and Monaco has the largest.



## Class Discussion (continued)

- ***Are the questions in the question above statistical questions? Why or why not?***  
Answer: They are not statistical questions because there is no variability in the responses.
- ***Create a statistical question you might be able to answer from looking at the distributions. Share your question with someone to see if they agree with you.***  
Answers will vary. Students might ask, “How do the typical life expectancies vary in the regions or which region has the smallest variability in life expectancies?”
- ***How do the life expectancies for the countries in the three regions of the world compare? Use what you know about distributions from your earlier work to help you answer the question.***  
Answers will vary. Students should notice that the distribution of life expectancies for countries in Asia has the largest range in life expectancies, about 24 years, from about 60 to 84 years. The range for life expectancies in countries in Western Europe is about 18 years, from about 72 to 90 years. The typical life expectancy seems to be the largest for countries in Western Europe, then North America and then Asia. The distribution of life expectancies for countries in Western Europe is mound shaped and relatively symmetric, but the distribution for the countries in Asia is skewed right, with most Asian countries having life expectations somewhere around 68 to 70 years. Western Europe and North America have what looks like an “outlier”: the life expectancy in Haiti of about 63 years is 8 or 9 years less than other countries in North America; and people in Monaco typically live to 90 years, about 5 years longer than the other Western Europe countries.

**Teacher Tip** Note you may want students to use **menu> Graph Type> Boxplot** and see for sure the outliers in the regions; four countries have extreme life spans compared to the rest of the countries in Western Europe, three below (Serbia, Latvia, and Estonia) and one above (Monaco).



## Class Discussion (continued)

Choose menu> Summary Measures> Med, IQR.

- ***Write a sentence for each of the three distributions describing what the segments on the graph indicate about the life expectancy for countries in the region.***
  
- ***How do the countries in the three regions compare with respect to life expectancy?***

Answers will vary: For countries in North America, the median life expectancy is about 77 years with half of the North American countries having about a four or five year difference in life expectancy. LQ is approximately 74 years and UQ is around 78 years. In countries in Western Europe, the median life expectancy is about 82 years, and the variation among the middle half of the countries' life expectancies is only about 3 years. The median life expectancy for countries in Asia is around 72 years, and the distribution of life expectancies for countries in Asia has the largest interquartile range, about 10 years. (Note that the exact values will vary because they are estimated from the plots.)

Answers will vary. The countries in Western Europe have a much higher life expectancy than the countries in the other two regions and the smallest IQR, which means the life expectancies of the countries in the middle half are really close together. The lowest life expectancy in a country in Western Europe is above the median of all the countries in Asia. 75% of the countries in Western Europe have a life expectancy greater than 75% of the countries in Asia. 50% of the countries in Western Europe have a life expectancy greater than 75% of the countries in North America. Countries in North America have the next largest median and next smallest IQR. The countries in Asia have a large difference in life expectancies, with people from Singapore and Hong Kong having life expectancies greater than most countries in Western Europe and North America, and those in Afghanistan the least of all the countries in these three regions. The median life expectancy for countries in Asia is just over 70 years.



## Class Discussion (continued)

- How do you think the Mean $\pm$  MAD will differ for each distribution? Use menu> Summary Measures and Mean $\pm$  MAD to check your answer.**

Answers will vary. Some might suggest that the mean will be smaller than the median for North America, larger for Asia and about the same for Western Europe. The interval for the Mean $\pm$  MAD will be larger for Asia and North America and about the same as the IQR for Western Europe.
- Select menu> Graph Type> Box plot and then Histogram (try several bin widths). Which of the graphs seemed to be easier to use to compare the distributions? Explain your reasoning.**

Answers will vary. Some might choose the box plots because you can see the median and IQR right away on the plot and get an overall picture of how the three compare. You can also clearly see the countries where the life expectancies are outliers for the region. Others might select the histogram because if the bin width is 3 or 5 you can see the gaps and where the data bunch up, which you cannot really see on the box plot. The dot plot seems small and the data scrunched.
- How do you think life expectancy for countries in Africa will compare to the life expectancies for countries in Asia and in Western Europe**

Answers will vary. Students may think it will be lower.
- Select North America and replace it with Africa. Was your conjecture in the previous question correct?**

Answers will vary. Yes, because the mean life expectancy for countries in Africa was around 60 years, way below those of the other two regions.
- What other features of the distribution of life expectancies for countries in Africa do you notice?**

Answers will vary. Half of the countries in Africa have a life expectancy that is lower than nearly every country in the other two regions. Even the countries with the highest life expectancy in Africa have life expectancies below most of the countries in Western Europe. Another observation is that Africa has many more countries than the other two regions.
- Replace Asia with Eastern Europe. How does the life expectancy for countries in Eastern Europe compare to the life expectancy for countries in Western Europe?**

Answer: The mean and median life expectancy are smaller for Eastern Europe; the MADs are about the same, maybe 1.5 years, but the IQR for Eastern Europe is a bit larger than the IQR for Western Europe (the distribution of years is not very symmetric.)



## Class Discussion (continued)

**Reset.** Use the TNS activity to compare the life expectancy of women and men for countries in the designated region.

- **North America**

Answers will vary. Women typically live longer than men. The median life expectancy for men in countries in North America is around 74 compared to 79 for the women. Except for Haiti, females in all of the countries in North America have a longer life expectancy than half of the males. The IQR is about 5 years for both, but the MAD for both is a bit larger than the IQR because Haiti has such a low life expectancy for both men and women.

- **South America**

Answers will vary. The life expectancies for women in countries in South America are considerably higher than men; half of the men in a South American country have life expectancies of around 72 years or more, while half of the women have life expectancies of about 79 years or more. The mean life expectancies for both men and women are lower than the median because French Guiana has such low numbers.

In  $\frac{3}{4}$  of the countries, women have a life expectancy greater than 75 years while in all the countries except Chile all of the men have a life expectancy less than 75 years.



## Student Activity Questions—Activity 1

1. Clear the data on page 1.3. Add the *Region* > *North America* > *All*, and then choose *Attribute* > *Education*. The plot shows the literacy rate for the countries in North America. The literacy rate is the percent of the population over the age of 15 that can read and write. Add the region *South America* > *All*.

a. How do the literacy rate for the countries in the two regions compare? (Use different graphs and summary statistics to support your thinking.)

Answers will vary. Basically, it looks like the countries in South America all have literacy rates between about 90% to 100%. The literacy rates for countries in North America are much more spread out than for the countries in South America, with a range of about 48%, from about 52% to 100%. The median of the countries in both regions are close, around 95 or 96%, but the IQR is about 3% in South America and about 18% in North America. The mean literacy rate for countries in North America is considerably lower than the median, about 89%, which reflects the fact that the distribution of literacy rates is skewed left.

b. Add *Eastern Europe* > *All* to the plot. How do the literacy rates for countries in Eastern Europe compare to those for countries in North and South America.

Answer: The countries in Eastern Europe have very literate people, way more so than the countries in North and South America. Most of the Eastern Europe countries have a literacy rate of 100%, and the mean is around 99%.

c. Make a conjecture about the literacy rate for countries in Western Europe. Then replace *South America* with *Western Europe* and check your conjecture.

Answers will vary. When checking, students should note that the means and medians of the countries in the two regions are about the same, with only two countries in Western Europe having lower literacy rates than countries in Eastern Europe. There are also more countries in Western Europe.

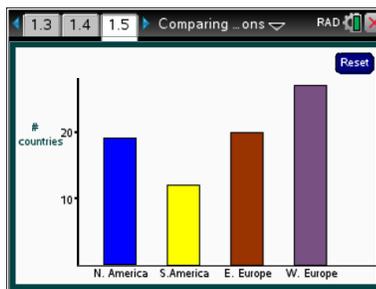
### Part 2, Page 1.5

Focus: Bar graphs convey some information that can be used to compare distributions but are limited as a data analysis tool.

Page 1.5 shows a bar graph displaying the number of countries in each region.

**Region** allows you to add another region.

**Order** has options to rearrange the bars.



### TI-Nspire Technology Tips

**menu** accesses page options.

**tab** cycles among the bars.

**ctrl del** resets the page.



### Class Discussion

#### Have students...

- **Describe the information you can read from the graph on page 1.5.**
- **Select menu> Order> Small to Large. Which region of the world displayed in the graph had the smallest numbers of countries? The largest?**
- **This is called a bar graph, similar to ones you have seen in earlier work. How is a bar graph different from a histogram? How does this graph differ from a histogram?**

#### Look for/Listen for...

Answer: You can tell the number of countries in each of the regions of the world in the data set.

Answer: South America has the smallest and Africa has the largest number of countries.

Answer: The horizontal axis only has regions of the world and is not a number line. You can move the bars around and order them, but you cannot move the bars in a histogram. Spaces between the bars in a histogram can only happen if no data values occur in that interval on the number line, while spaces between the bars in a bar graph don't have any meaning and can be as big or as little as you want.

### Part 2, Page 1.7

Focus: Bar graphs convey some information that can be used to compare distributions but are limited as a data analysis tool.

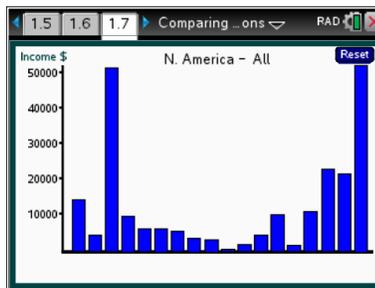
Page 1.7 shows the countries from a single region in a bar graph.

**Region** chooses a different region.

**Attribute** changes attribute among income, life expectancy, and literacy.

**Order** chooses the order of display

**Clear** displays a blank graph.



### TI-Nspire Technology Tips

**menu** accesses page options.

**tab** cycles among the bars.

**ctrl del** resets the page.



### Class Discussion

The data in the TNS activity represent this adjusted per person income, sometimes called per capita income. Select menu> Region> North America> All

- **What strikes you about the graph?**

Answers will vary. Two bars are much taller than the others.



### Class Discussion (continued)

- *Move through the bars until you find one labeled United States. Explain what the bar represents.*
- *Does your answer indicate that everyone in the United States has that income?*

Answer: The graph indicates that the per person income for people in the United States is a bit over \$50,000.

Answers will vary. Students should recognize that this is based on a mean, and therefore a few people with a lot of money will raise the mean. You might give an example of 10 people, 9 with \$1 and 1 with \$100. The mean amount per person will be \$10.90, which is more than anyone but one person had.

- *Explain why the menu does not have an option to display the mean, MAD, median, and IQR.*

Answer: These numbers make sense only on a number line and you don't have one on the horizontal axis in a bar graph. The vertical axis is a number line, but no points are plotted on it like in a dot plot that turns into a histogram.

Select menu> Region> Middle East.

- *Move through the bars from left to right. Explain what the graph of these bars represents.*
- *Select the highest bar. What does this bar represent?*
- *The tallest bar belongs to Qatar. The tallest bar in the bar graph for North America belonged to the United States. Is the per capita income the same for Qatar as the United States?*
- *Select menu> Order> Large to Small. Identify three countries that have very low income per person.*
- *Find a region of the world where the per person income for at least five countries is greater than that of the United States. Identify the countries.*

Answer: Each bar represents a country in the Middle East and the height of the bar represents the adjusted income per person in a country in the Middle East.

Answer: This bar represents the adjusted income per person in Qatar, which is a little over \$100,000.

Answer: No, because the scales are different. The per capita income in the United States is a bit more than half as much as the per capita income in Qatar.

Answers will vary. The countries might be Iran, Palestine, Lebanon, Syria, Jordan, Yemen, Iraq.

Answer: Monaco, Liechtenstein, Luxemburg, Norway and Switzerland in Western Europe all have a larger per person income.



## Class Discussion (continued)

Select menu> Region> Western Europe> Income and order the countries from those with the largest per capita income to those with the smallest.

- **Which country in this region has the largest per capita income? How does it compare to the per capita income for the United States?**  
Answer: Slovenia, with a per capita income of about \$22,000, is less than half of the per capita income in the US.
- **Estimate the difference in the per capita income for the Ukraine and for Russia.**  
Answer: The difference is about \$10,000 because Ukraine is about \$4,000 and Russia is about \$14,000.
- **Think of a statistical question related to the income in the countries in different regions of the world. Use the data in the files to answer your question. Make a poster to share with the class describing your question, your answer and the graphs and statistics that helped you figure out your answer.**  
Answers will vary.



## Student Activity Questions—Activity 2

1. **Decide whether each statement is sometimes true, always true or never true. Give a reason for your answer.**
  - a. **Bar graphs are an efficient way to compare a characteristic across countries in a region.**  
Answer: Sometimes. When there are lots of categories, it can be really difficult to get them all on one sheet and still be able to interpret each bar.
  - b. **The horizontal axis in a bar graph can be used to illustrate the interquartile range.**  
Answer: Never. The IQR is a number that represents an interval on a number line. Bar graphs do not have number lines on the horizontal axis.
  - c. **A bar graph shows how many categories are being investigated.**  
Answer: Always, if it is a bar graph of all the data.
  - d. **Bar graphs are good ways to display the most frequent occurrence of a particular category.**  
Answer: True, the tallest bar represents the category with the greatest frequency (or percent, value, etc.).



## Student Activity Questions—Activity 2 (continued)

2. What are advantages and disadvantages of dot plots, box plots, histograms and bar graphs in comparing the countries within one region to the countries within another region with respect to a certain characteristic? Explain why you made your choice.

Answers will vary.

- The box plot shows a five number statistical summary and some sense of how the data are distributed across the range. The box plot does not show exact data values, gaps or clusters in the distribution, and the number of countries.
- The dot plot shows all of the data and how they are distributed including gaps and clusters, but the dots get jumbled up when the data are close together.
- A histogram is a nice way to get an idea of the shape and spread of the distribution, and you can see gaps and clusters if the bin widths are not too large. You lose the exact data values but you can tell how many countries by looking at the frequencies for each of the bars. A histogram is not too useful if you only have a few data values; they are more useful when you have a very large data set.
- A bar graph keeps the name of the country along with the measurement, and you can arrange the bars so you can see the most to the least. You cannot easily see the shape of the distribution of all the measurements, clusters or gaps, nor estimate any measures of spread or center.



## Deeper Dive — Page 1.3

- ***Compare the literacy rate for females in different regions of the world. Which region has the highest literacy rate for females? The lowest? Explain which graphs helped you find your answer.***
- ***Which region seems to have the most countries with around the same per capita income? Justify your thinking.***

Answers may vary. The countries in Africa have the lowest typical literacy rate with a mean of 64% and an MAD of about 17%. Eastern and Western Europe both have a mean literacy rate of about 99% and a MAD of about 1%. One way is to look at the individual countries using the bar graphs (and checking out Africa on page 1.3).

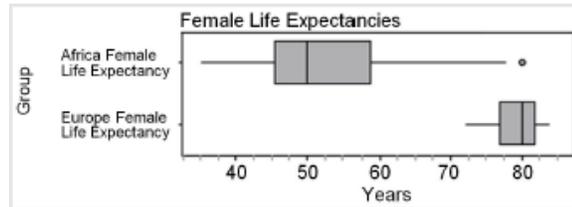
Answers may vary. South America has one country, French Guiana that has about a \$25,000 per capita income, but the other countries in that region are between \$4,000 (Bolivia) and \$15,000 (Chile). The other regions have several countries with high per capita income, and others with very low incomes.



## Sample Assessment Items

After completing the lesson, students should be able to answer the following types of questions. If students understand the concepts involved in the lesson, they should be able to answer the following questions without using the TNS activity.

Life expectancies in countries in Africa and Europe



**Figure A**

(Adapted from Progressions for statistics and probability, 2011)

1. Which of the statements are true given the graph above?
  - a. The life expectancy of females in European countries is greater than  $\frac{3}{4}$  of the life expectancies of females in African countries.

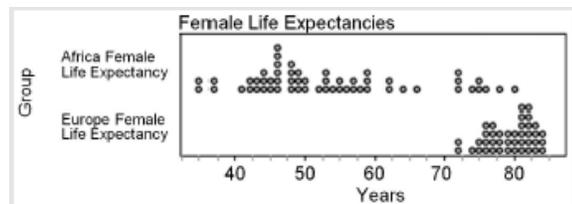
**Answer: True**

- b. There is less variability in life expectancy among females in European countries than among females in African countries.

**Answer: True**

- c. The mean life expectancy among females in African countries is about 50 years.

**Answer: False**



**Figure B**

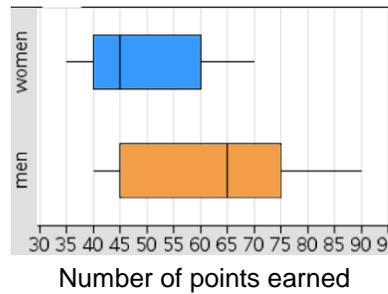
(Adapted from Progressions for statistics and probability, 2011)

2. Name one thing you can learn about the distribution of life expectancies from the graph in Figure A that you cannot easily see in Figure B and one thing you know about the distribution of life expectancies from Figure B that you cannot tell from Figure A.

**Answer: You can see the median and the interquartile range in figure A and also see the outlier in female life expectancy in Africa. In Figure B, you can see the shape of the distribution, like the gap between 68 and 72 years for the female life expectancies in countries in Africa. You have a better sense of the number of countries in Figure B.**



3. The plot shows the number of points earned in a contest by men and women.



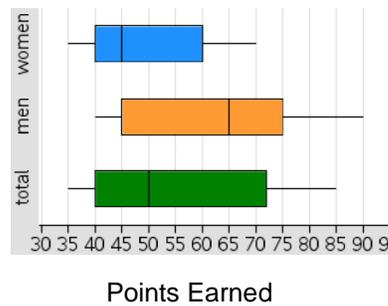
a. What is the least number of points earned in the contest by either men or women?

**Answer: 35 points**

b. What is the difference between the two medians?

**Answer: 20 points**

c. Could the box plot labeled “total” in the graph “Points Earned” represent the number of points earned by both the men and the women? Explain why or why not.



**Answers may vary. No because at least one man earned 90 points, and the maximum value for the box plot for both is only 85.**



## Student Activity Solutions

In these activities you will compare data collected from different populations and analyze the advantages and disadvantages of using different graphical representations. After completing the activities, discuss and/or present your findings to the rest of the class.



### Activity 1 [Page 1.3]

1. Clear the data on page 1.3. Add the **Region> North America> All**, and then choose **Attribute> Education**. The plot shows the literacy rate for the countries in North America. The literacy rate is the percent of the population over the age of 15 that can read and write. Add the region **South America> All**.

- a. How do the literacy rate for the countries in the two regions compare? (Use different graphs and summary statistics to support your thinking.)

*Answers will vary. Basically, it looks like the countries in South America all have literacy rates between about 90% to 100%. The literacy rates for countries in North America are much more spread out than for the countries in South America, with a range of about 48%, from about 52% to 100%. The median of the countries in both regions are close, around 95 or 96%, but the IQR is about 3% in South America and about 18% in North America. The mean literacy rate for countries in North America is considerably lower than the median, about 89%, which reflects the fact that the distribution of literacy rates is skewed left.*

- b. Add **Eastern Europe> All** to the plot. How do the literacy rates for countries in Eastern Europe compare to those for countries in North and South America.

*Answer. The countries in Eastern Europe have very literate people, way more so than the countries in North and South America. Most of the Eastern Europe countries have a literacy rate of 100%, and the mean is around 99%.*

- c. Make a conjecture about the literacy rate for countries in Western Europe. Then replace South America with Western Europe and check your conjecture.

*Answers will vary. When checking, students should note that the means and medians of the countries in the two regions are about the same, with only two countries in Western Europe having lower literacy rates than countries in Eastern Europe. There are also more countries in Western Europe.*



### Activity 2 [Page 1.7]

1. Decide whether each statement is sometimes true, always true or never true. Give a reason for your answer.

- a. Bar graphs are an efficient way to compare a characteristic across countries in a region.

*Answer: Sometimes. When there are lots of categories, it can be really difficult to get them all on one sheet and still be able to interpret each bar.*



- b. The horizontal axis in a bar graph can be used to illustrate the interquartile range.

*Answer: Never. The IQR is a number that represents an interval on a number line. Bar graphs do not have number lines on the horizontal axis.*

- c. A bar graph shows how many categories are being investigated.

*Answer: Always, if it is a bar graph of all the data.*

- d. Bar graphs are good ways to display the most frequent occurrence of a particular category.

*Answer: True, the tallest bar represents the category with the greatest frequency (or percent, value, etc.).*

2. What are advantages and disadvantages of dot plots, box plots, histograms and bar graphs in comparing the countries within one region to the countries within another region with respect to a certain characteristic? Explain why you made your choice.

*Answers will vary.*

- The box plot shows a five number statistical summary and some sense of how the data are distributed across the range. The box plot does not show exact data values, gaps or clusters in the distribution, and the number of countries.*
- The dot plot shows all of the data and how they are distributed including gaps and clusters, but the dots get jumbled up when the data are close together.*
- A histogram is a nice way to get an idea of the shape and spread of the distribution, and you can see gaps and clusters if the bin widths are not too large. You lose the exact data values but you can tell how many countries by looking at the frequencies for each of the bars. A histogram is not too useful if you only have a few data values; they are more useful when you have a very large data set.*
- A bar graph keeps the name of the country along with the measurement, and you can arrange the bars so you can see the most to the least. You cannot easily see the shape of the distribution of all the measurements, clusters or gaps, nor estimate any measures of spread or center.*