CHAPTER 1
Exploring Data

1.2
Displaying Quantitative Data with Graphs



The Practice of Statistics, 5th Edition
Starnes, Tabor, Yates, Moore

Bedford Freeman Worth Publishers

Displaying Quantitative Data with Graphs

Learning Objectives

After this section, you should be able to:

- ✓ MAKE and INTERPRET dotplots and stemplots of quantitative data
- ✓ DESCRIBE the overall pattern of a distribution and IDENTIFY any outliers
- ✓ IDENTIFY the shape of a distribution
- ✓ MAKE and INTERPRET histograms of quantitative data
- ✓ COMPARE distributions of quantitative data

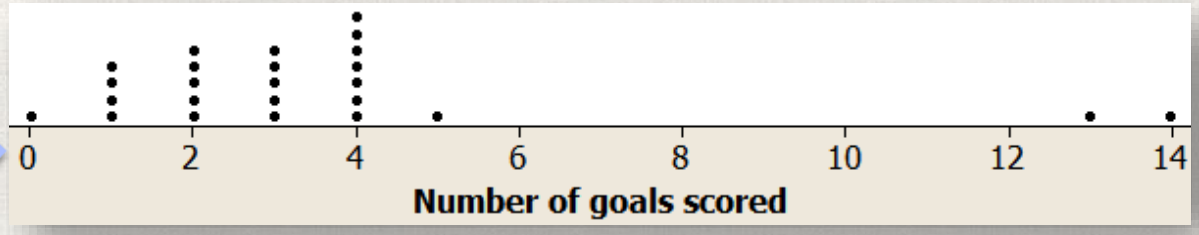
Dotplots

One of the simplest graphs to construct and interpret is a **dotplot**. Each data value is shown as a dot above its location on a number line.

How to make a dotplot:

1) Draw a horizontal axis (a number line) and label it with the variable name.
2) Scale the axis from the minimum to the maximum value.
3) Mark a dot above the location on the horizontal axis corresponding to each data value.

Number of Goals Scored Per Game by the 2012 US Women's Soccer Team												
2	1	5	2	0	3	1	4	1	2	4	13	1
3	4	3	4	14	4	3	3	4	2	2	4	



Examining the Distribution of a Quantitative Variable

The purpose of a graph is to help us understand the data. After you make a graph, always ask, "What do I see?"

How to Examine the Distribution of a Quantitative Variable

1) In any graph, look for the **overall pattern** and for striking **departures** from that pattern.

2) Describe the overall pattern of a distribution by its:
   - **Shape**
   - **Center**
   - **Spread**

Don't forget your SOCS!

3) Note individual values that fall outside the overall pattern. These departures are called **outliers**.

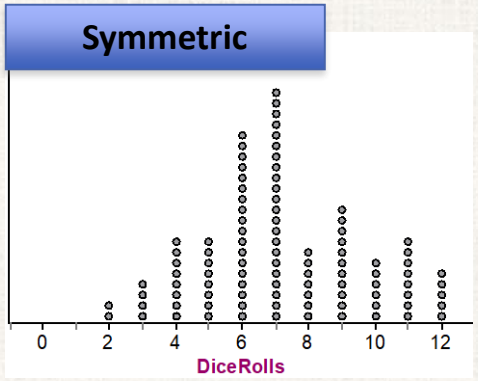
Describing Shape

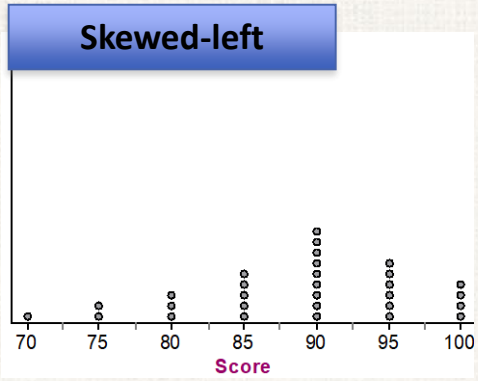
When you describe a distribution's shape, concentrate on the main features. Look for rough **symmetry** or clear **skewness**.

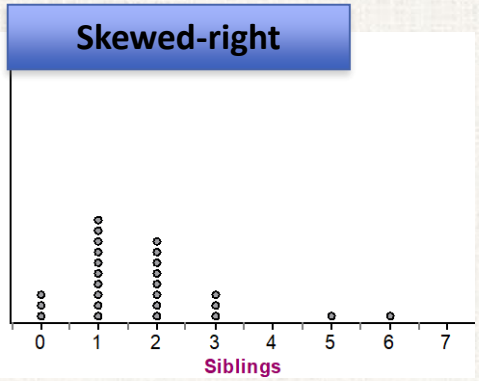
A distribution is roughly **symmetric** if the right and left sides of the graph are approximately mirror images of each other.

A distribution is **skewed to the right** (right-skewed) if the right side of the graph (containing the half of the observations with larger values) is much longer than the left side.

It is **skewed to the left** (left-skewed) if the left side of the graph is much longer than the right side.



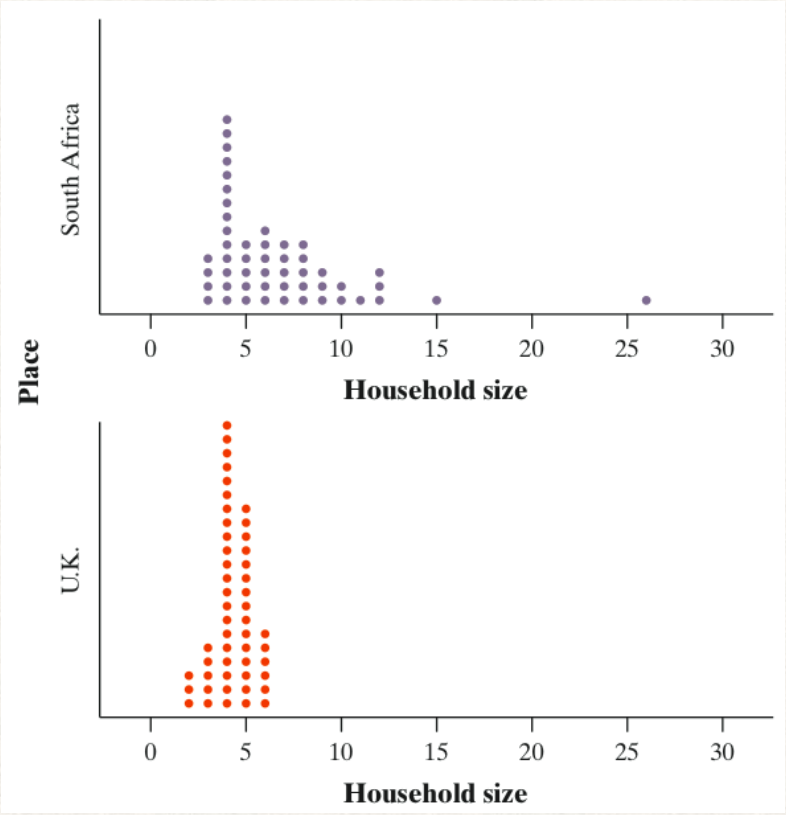




Comparing Distributions

Some of the most interesting statistics questions involve comparing two or more groups.

Always discuss shape, center, spread, and possible outliers whenever you compare distributions of a quantitative variable.



Compare the distributions of household size for these two countries.

Don’t forget your SOCS!

Stemplots

Another simple graphical display for small data sets is a **stemplot**. (Also called a **stem-and-leaf plot.**)
Stemplots give us a quick picture of the distribution while including the actual numerical values.

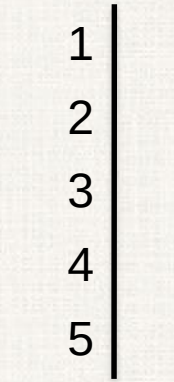
How to make a stemplot:

1) Separate each observation into a **stem** (all but the final digit) and a **leaf** (the final digit).
2) Write all possible stems from the smallest to the largest in a vertical column and draw a vertical line to the right of the column.
3) Write each leaf in the row to the right of its stem.
4) Arrange the leaves in increasing order out from the stem.
5) Provide a key that explains in context what the stems and leaves represent.

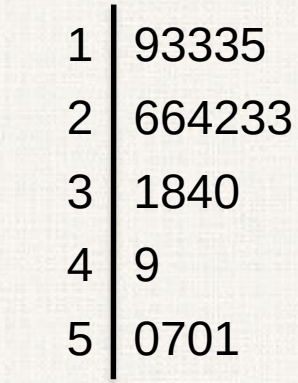
Stemplots

These data represent the responses of 20 female AP Statistics students to the question, “How many pairs of shoes do you have?” Construct a stemplot.

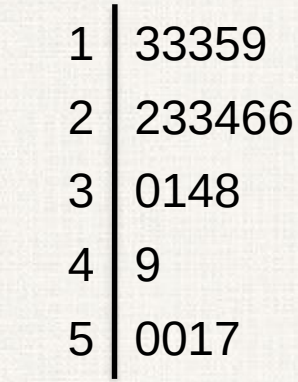
50	26	26	31	57	19	24	22	23	38
13	50	13	34	23	30	49	13	15	51



Stems



Add leaves



Order leaves

Key: 4|9 represents a female student who reported having 49 pairs of shoes.

Add a key

Stemplots

When data values are “bunched up”, we can get a better picture of the distribution by **splitting stems**.

Two distributions of the same quantitative variable can be compared using a **back-to-back stemplot** with common stems.

Females

50	26	26	31	57	19	24	22	23	38
13	50	13	34	23	30	49	13	15	51

Males

14	7	6	5	12	38	8	7	10	10
10	11	4	5	22	7	5	10	35	7

	0		
	0		
	1		
	1		
	2		
	2		
	3		
	3		
	4		
	4		
	5		
	5		

“split stems”

Females		Males
	0	4
	0	555677778
333	1	0000124
95	1	
4332	2	2
66	2	
410	3	
8	3	58
	4	
9	4	
100	5	
7	5	

Key: 4|9 represents a student who reported having 49 pairs of shoes.

Histograms

Quantitative variables often take many values. A graph of the distribution may be clearer if nearby values are grouped together.

The most common graph of the distribution of one quantitative variable is a **histogram**.

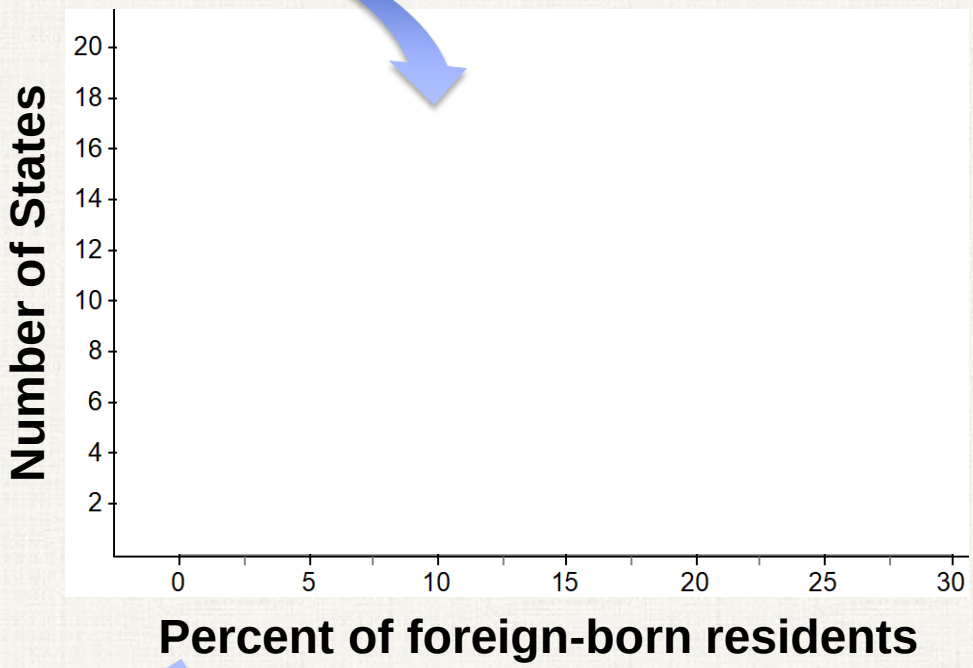
How to make a histogram:

1) Divide the range of data into classes of equal width.
2) Find the count (frequency) or percent (relative frequency) of individuals in each class.
3) Label and scale your axes and draw the histogram. The height of the bar equals its frequency. Adjacent bars should touch, unless a class contains no individuals.

Histograms

This table presents data on the percent of residents from each state who were born outside of the U.S.

Frequency Table	
Class	**Count**
0 to <5	20
5 to <10	13
10 to <15	9
15 to <20	5
20 to <25	2
25 to <30	1
Total	50



Using Histograms Wisely

Here are several cautions based on common mistakes students make when using histograms.

Cautions!

1)Don’t confuse histograms and bar graphs.

2)Don’t use counts (in a frequency table) or percents (in a relative frequency table) as data.

3)Use percents instead of counts on the vertical axis when comparing distributions with different numbers of observations.

4)Just because a graph looks nice, it’s not necessarily a meaningful display of data.

Data Analysis: Making Sense of Data

Section Summary

In this section, we learned how to…

- ✓ MAKE and INTERPRET dotplots and stemplots of quantitative data
- ✓ DESCRIBE the overall pattern of a distribution
- ✓ IDENTIFY the shape of a distribution
- ✓ MAKE and INTERPRET histograms of quantitative data
- ✓ COMPARE distributions of quantitative data