## **CHAPTER 12**

### Statistics



ALWAYS LEARNING

Copyright © 2015, 2011, 2007 Pearson Education, Inc.

## 12.1

### Sampling, Frequency Distributions, and Graphs

ALWAYS LEARNING

Copyright © 2015, 2011, 2007 Pearson Education, Inc.

### Objectives

- 1. Describe the population whose properties are to be analyzed.
- 2. Select an appropriate sampling technique.
- 3. Organize and present data.
- 4. Identify deceptions in visual displays of data.

## **Statistics**

**Statistics**: Method of collecting, organizing, analyzing, and interpreting data, as well as drawing conclusions based on the data. Methodology is divided into two main areas.

**Descriptive Statistics**: Collecting, organizing, summarizing, and presenting data.

**Inferential Statistics**: Making generalizations about and drawing conclusions from the data collected.

### **Populations and Samples**

**Population**: Set containing all the people or objects whose properties are to be described and analyzed by the data collector.

**Sample:** Subset or subgroup of the population.

**Representative Sample**: A sample that exhibits characteristics typical of those possessed by the target population.

ALWAYS LEARNING Copyright © 2015, 2011, 2007 Pearson Education, Inc. PEARSON Section 12.1, Slide 5

## **Example: Populations and Samples**

A group of hotel owners in a large city decide to conduct a survey among citizens of the city to discover their opinions about casino gambling.

a. Describe the population.

### Solution

The population is the set of all the citizens of the city.

**b**. It is suggested to survey all the people at six of the largest nightclubs in the city on a Saturday night about their opinion on casino gambling. Is this a good idea?

### Solution

This is a bad idea. The nightclub set is more likely to have a positive attitude toward casino gambling than the general population.

## **Random Sampling**

A **random sample** is a sample obtained in such a way that every element in the population has an equal chance of being selected.

Methodology:

Identify each element in the population.

Assign numbers to each element in the population.

Randomly select numbers.

Assign the elements in the population who have those numbers to the sample set.

### Example: Selecting an Appropriate Sampling Technique

Which of the following is the best way to select a random sample to find out how the city's citizens feel about casino gambling?

a. Randomly survey people who live in the oceanfront condominiums in the city.

Many hotels lie along the oceanfront and these people might object to the increased traffic and noise which may result. It also does not give each citizen of the city an equal chance.

Which of the following is the best way to select a random sample to find out how the city's citizens feel about casino gambling?

b. Survey the first 200 people whose names appear in the city's telephone directory.

Does not give each citizen in the city an equal chance.

c. Randomly select neighborhoods of the city and then randomly survey people within the selected neighborhoods.

Yes. Gives each citizen in the city an equal chance.

# Example: Constructing a Frequency Distribution

Construct a frequency distribution for the data of the age of maximum yearly growth for 35 boys: 12, 14, 13, 14, 16, 14, 14, 17, 13, 10, 13, 18, 12, 15, 14, 15, 15, 14, 14, 13, 15,

16, 15, 12, 13, 16, 11, 15, 12, 13, 12, 11, 13, 14, 14.

Frequency Distribution for a Boy's Age of Maximum Yearly Growth

Age of Maximum Growth	Number of Boys (Frequency)	
10	1	
11	2	
12	5	
13	7	
14	9	
15	6	
16	3	
17	1	
18	1	
<b>Total:</b> $n = 35$		

ALWAYS LEARNING

Copyright © 2015, 2011, 2007 Pearson Education, Inc.

What are some of the conclusions we can draw from this example?

- Maximum growth for most subjects occurred between ages 12 and 15.
- The number of boys who attain their maximum yearly growth at a given age increases until age 14 and decreases after that.

### Example: Constructing a Grouped Frequency Distribution

Here are the statistics test scores for a class of 40 students:

82 47 75 64 57 82 63 93

76 68 84 54 88 77 79 80

94 92 94 80 94 66 81 67

75 73 66 87 76 45 43 56

57 74 50 78 71 84 59 76

Group the frequencies into classes that are meaningful for the data.

Since letter grades are given based on 10-point ranges, use the classes 40–49, 50–59, 60–69, 70–79, 80–89, 90–99.

Class	Frequency	
40-49	3	
50-59	6	
60-69	6	
70-79	11	
80-89	9	
90-99	5	
Total: $n = 40$ 40, the sum of the frequencies, is the number of data liems.		

The class 40-49 has 40 as the lower class limit and 49 as the upper class limit. The class width is 10. It is sometimes helpful to vary the width of the first or last class to allow for items that fall above or below most data.

ALWAYS LEARNING

Copyright © 2015, 2011, 2007 Pearson Education, Inc.

## **Histograms and Frequency Polygons**

**Histogram**: A bar graph with bars that touch can be used to visually display the data. **Frequency Polygon**: A line graph formed by connecting dots in the midpoint of each bar of the histogram.



ALWAYS LEARNING

Copyright © 2015, 2011, 2007 Pearson Education, Inc.

### **Stem-and-Leaf Plots**

This plot is constructed by separating each data item into two parts:

Stem consists of the ten's digit

Leaf consists of the units' digit.

ALWAYS LEARNING

Copyright © 2015, 2011, 2007 Pearson Education, Inc.

#### **Example: Constructing a Stem-and-Leaf Plot**

Use the data below for a stem-and-leaf plot:

82477564578263937668845488777980949294809466816775736687764543565774507871845976

ALWAYS LEARNING

Copyright © 2015, 2011, 2007 Pearson Education, Inc. PEARS

## Enter the data items in the first row:

Stem	Leaves	5
4	7	
5	7	
6	4 3	
7	5	
8	22	
9	3	

Continue with all the remaining rows. The final plot:



ALWAYS LEARNING

Copyright © 2015, 2011, 2007 Pearson Education, Inc.

## **Deceptions in Visual Displays of Data**

Percentage of People in the United States Living below the Poverty Level, 2000-2004



Graphs can be used to distort the underlying data

The graph on the left stretches the scale on the vertical axis to create an impression of a rapidly increasing poverty rate.

The graph on the right compresses the scale on the vertical axis to create an impression of a slowly increasing poverty rate.

### Things to Watch for in Visual Displays of Data

- 1. Is there a title that explains what is being displayed?
- 2. Are numbers lined up with tick marks on the vertical axis that clearly indicate the scale? Has the scale been varied to create a more or less dramatic impression than shown by the actual data?
- 3. Do too many design and cosmetic effects draw attention from or distort the data?
- 4. Has the wrong impression been created about how the data are changing because equally-spaced time intervals are not used on the horizontal axis?

### Things to Watch for in Visual Displays of Data continued

- 5. Are bar sizes scaled proportionately in terms of the data they represent?
- 6. Is there a source that indicates where the data in the display came from? Does the data come from an entire population or a sample? Was a random sample used and, if so, are there possible differences between what is displayed in the graph and what is occurring in the entire population? Who is presenting the visual display and do they have a special case to make for or against the trend shown by the graph?

ALWAYS LEARNING

Copyright © 2015, 2011, 2007 Pearson Education, Inc.

## **Examples of Misleading Visual Displays**

Cosmetic effects of home with equal heights, but different frontal additions and shadow lengths, make it impossible to tell if they proportionately depict the given areas. Time intervals on the horizontal axis are not uniform in size, making it appear that dwelling swelling has been linear from 1980 through 2010.

The data indicate that this is not the case. There was a greater increase in area from 1980 through 1990, averaging 34 square feet per year, than from 1990 through 2010, averaging approximately 15.6 square feet per year.



ALWAYS LEARNING