

## Section 1.1 - Introduction

**Raw Data** is numerical and/or non-numerical value's of information or observations before it is arranged or analyzed in a useful manner.

**Statistics** involves the procedures associated with the data collection process, the summarizing and interpretation of data, and the drawing of inferences or conclusions based upon the analysis of the data.

**Descriptive Statistics** uses numerical and/or visual techniques to summarize or describe the data in a clear and effective manner.

The **population** is the entire collection of all individuals or objects of interest.

The **sample** is the portion of the population that is selected for study.

**Inferential Statistics** is the process of using **sample information** to draw inferences or conclusions about the **population**.

A **representative sample** is a sample that has the pertinent characteristics of the population in the same proportion, as they are included in that population.

A **statistic** is a number that describes a characteristic of a sample.

A **parameter** is a number that describes a characteristic of a population.

## Section 1.2 - Why Sample?

**Sampling** is the process of selecting a portion, or sample, of the entire population.

A survey that includes every item or individual of the population is called a **census**.

### Why Sample?

Samples are taken when it is either impossible or impractical to examine the entire population. Studying an entire population is impractical when the population is extremely large.

**Sample Size** – The size of a sample, denoted by  $n$ , is the number of data values in the sample.

**Population Size** – The size of a finite population, denoted by  $N$ , is the number of data values in the population.

### Reasons for Sampling

1. A census is impossible
2. Cost
3. Time
4. Accuracy

### Section 1.3 - Sampling Techniques

A **simple random sample** is a sample of data values selected from a population in such a way that every sample of size  $n$  has an equal probability of being selected and every data value of the population has the same chance of being selected for the sample.

Selecting a **non-representative** or **biased** sample from a population can lead to invalid inferences about the population.

### Section 1.4 - Uses of Statistics

Statistics are used in many different fields such as economics, finance, psychology, sociology, education, and the physical sciences just to name a few.

Extensive data collection and distribution activities are performed by the federal and other governmental and private agencies in areas such as education, employment, health, crime, prices, housing, medical care, manufacturing, agriculture, construction, transportation, etc.

The development of the computer has led to a revolution in the area of data collection and analysis. Statistical analysis can now be applied to vast amounts of data accurately and quickly. Such diverse problems as weather forecasting, economic stabilization, and disease control are today being solved using statistical analysis.

### Section 1.5 - Misuses of Statistics

The following techniques illustrate how information can be presented in a misleading fashion:

1. Misleading graphs - can be very misleading based on the scale.
2. Non-representative samples (Biased) - purposely choosing a sample that favors your desired outcome.  
*ex:* Surveying people with a personal interest, such as their livelihood is dependent on a certain product  
*ex:* Surveying many people but only using the ones who say what you are looking for.
3. Inappropriate comparisons  
*ex:* Publishing that destination X was a more popular vacation spot for the month of August than destination Y, when destination Y is in its off-season during the month of August and destination X is not.
4. The omission of variation about an average

**Review the following examples in your textbook:**

*Example 1.1* - pg. 6, *Example 1.2* - pg. 7, *Example 1.3* - pg. 12, *Example - Literary Digest* - pg. 11, *Example 1.13* - pg. 18, *Example 1.15* - pg. 18-19, *Example 1.19* - pg. 20, *Example 1.20* - pg. 21, *Example 1.21* - pg. 21, *Example 1.22* - pg. 21

**Class Worksheet - Chapter 1**

Complete the following questions:

1. A survey of 1,000 undergraduates at a midwestern university of 40,000 undergraduates was taken to determine the percentage of undergraduates living at home. The survey concluded that 20% of the students live at home.
  - a. What is the population?
  - b. What is the sample?
  - c. The percentage value 20% represents a \_\_\_\_\_?
2. Parameters describe the characteristics of a \_\_\_\_\_.
3. The use of sample information to estimate a population characteristic is the primary objective of \_\_\_\_\_ statistics. Thus, the researcher must take care to insure that the sample is \_\_\_\_\_ of the sample.
4. Most psychology texts reports that the average IQ score for an adult male is between 90 and 110. This average is an example of a \_\_\_\_\_.
5. The branch of statistics that uses numerical and/or visual techniques to summarize data is called \_\_\_\_\_.
6. The collection, presentation, analysis and interpretation of data is called \_\_\_\_\_.
7. A study of all working women in San Francisco was conducted to determine their average annual salary. A sample of 250 working women was selected, and their salaries were recorded.
  - a. The population is \_\_\_\_\_.
  - b. The sample is \_\_\_\_\_.
  - c. The population parameter being studied is \_\_\_\_\_.
8. Determine if each of the following sampling procedures will produce a random sample. (*Example 1.3 - pg. 12*)
  - a. The population is all females between the ages of 17 and 35 living in NY State. The sampling procedure is devised to select a sample of 100 female college students with in the NY State.
  - b. The population is defined to be all the people who shop at Green Acres Mall during the week before Christmas. The sampling procedure is devised to select every 25<sup>th</sup> shopper who enters the Mall at only one of its ten entrances.
  - c. The population is all the faculty members teaching at the College of Bill and Mary. The sampling procedure will select a sample of 25 faculty members using the following procedure: The name of each faculty member will be placed in a large drum. The drum is rotated several times before each of the 25 faculty names are selected.
  - d. An advertising agency decides to conduct a marketing poll of people living in a New England community. The sampling procedure will draw a sample of people living within this community by selecting the names of the people from the community telephone directory.