Chapter 1: Basics
1.1 Sampling

______________________________ is the study of procedures for collecting, describing, and drawing conclusions from information. The information may be collected in various ways.

Definitions:

**Population:**

**Sample:**

**Statistic:**

**Parameter:**

**Example**

a) If we are talking about just our students in this class, our class survey would yield ______________? (Parameters or statistics)

b) If we are using our student survey to give us information about Cabrillo College students in general, our class survey would yield ______________? (Parameters or statistics)

**Sampling**

When using a sample, we would like our sample to represent the population as closely as possible. Unfortunately, there are no methods that can guarantee that a sample will represent the population well. The best we can do is to use a method that makes it very likely that the sample will be similar to the population.
Simple random sample:

Example
Survey polls exactly 750 males and 750 females. Is this a simple random sample?
Solution

Example
A survey decides which 150 Cabrillo students to poll by putting all of the names of all Cabrillo students into a hat and then selecting names from the hat. Is this a simple random sample?
Solution

In some cases, it is difficult or impossible to draw a sample in a truly random way. Therefore, we need other methods.
Sample of convenience:

Stratified sampling:

Cluster sampling:

Systematic sampling:

Voluntary response sampling:
Example
Classify the following as simple random, voluntary, stratified, cluster, systematic, or voluntary response samples.

a) A pollster gets the list of all the addresses in a town, and uses a random number generator to choose 150 of them. The pollster visits each of the 150 households and interviews all of the adults in each house about TV viewing habits.
Solution

b) You attempt to sample people by yelling out your window at people walking down the street.
Solution

c) Police at a sobriety checkpoint pull over every 5th car to determine whether the driver is sober.
Solution

d) A cell phone company draws a random sample of 200 people with Blackberries, 200 people with Samsungs, and 200 people with IPhones.
Solution

e) 500 raffle tickets are placed in to a hat, mixed up, and 10 winners are drawn.
Solution

f) A cell phone company invites cell phone users to register their opinions on an online survey about fee increases.
Solution

Do You Know
- What is statistics?
- Difference between a population and a parameter?
- What is a simple random sample?
- Difference between stratified sampling, cluster sampling, systematic sampling, and voluntary response sampling?
- Difference between a statistic and a parameter?
1.2 Types of Data

Introduction
In the previous section, we described various methods of collecting information by sampling. Once the information has been collected, the collection is called a ________________.

Structure of a Data Set
Information is collected on ________________.

The characteristics of the individuals about which we collect information are called ________________.

The values of the variables that we obtain are the ________________.

Types of Variables
There are two types of variables:
Qualitative variables:

Quantitative variables:

Example
Which of the following variables are qualitative and which are quantitative?
(a) Model of a car
(b) A person’s height
(c) Whether a person is left or right handed
(d) # of students in a class

Qualitative variables can be divided into:
Ordinal:

Nominal:
**Example**
Which of the following variables are ordinal and which are nominal?
(a) City you were born in

(b) left or right handed-ness

(c) Questions on a survey scored 1 through 5, with 1 strongly disagree and 5 strongly agree.

Recall that qualitative variables can be divided into two categories, so can quantitative variables. Quantitative variables can be classified as;

**Discrete:**

**Continuous:**

**Example**
Which of the following variables are discrete and which are continuous?
(a) The number of students in a class

(b) The weight of a person

**Do You Know**
- Difference between qualitative and quantitative variables?
- Difference between the types of qualitative variables?
- Difference between discrete and continuous variables?

**1.3 Design of Experiments**
Two main types of experiments. Entire classes are devoted to this topic, so here is the quick version:

In an __________________________, specific characteristics are measured and observed, but subjects are not manipulated.

In an __________________________, a treatment is applied and effects are observed.
1.4 Bias
A study conducted by a procedure that produces the correct result on the average is said to be _______________________.

Studies conducted with methods that tend to overestimate or underestimate the true value are said to be _______________________.

Sources of Bias
Voluntary response surveys:

Example: Ratemyprofessors.com

Self interest bias:

Example: Toyota motor trend car of the year mentioned on commercials, but not massive recalls

Social acceptability bias:

Example: Have you ever drank and driven?

Loaded question bias:

Example: a)Less than 2% of landfills are disposable diapers. Should disposable diapers be banned from landfills?
b)More than 2.5 million tons of waste in landfills consists of disposable diapers. Should disposable diapers be banned from landfills?

Non-response bias:
**Example:** Almost every night when I get a call from an 800 number and there is a long pause after I say “Hello.”

**Sampling bias:**

**Example:** SPCA calls people who have adopted animals before to poll them.

**Do You Know**

- What is bias in a study?
- The various sources of bias?