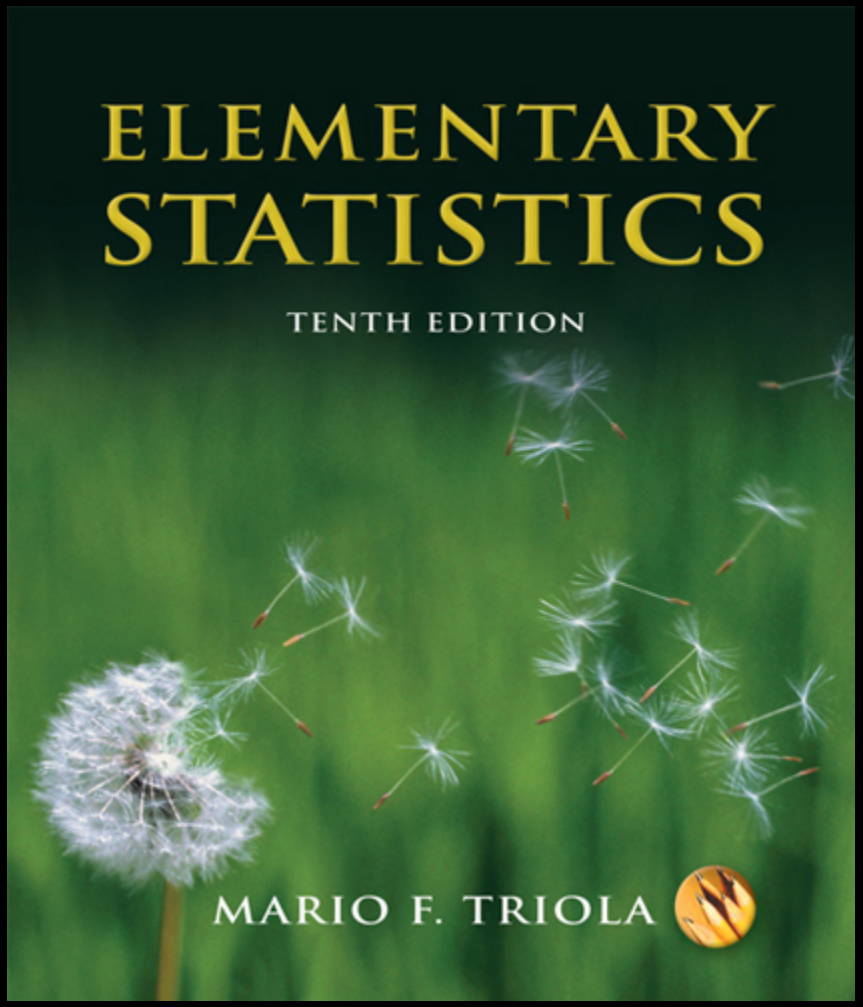
Lecture Slides



Elementary Statistics
Tenth Edition

and the Triola Statistics Series

by Mario F. Triola

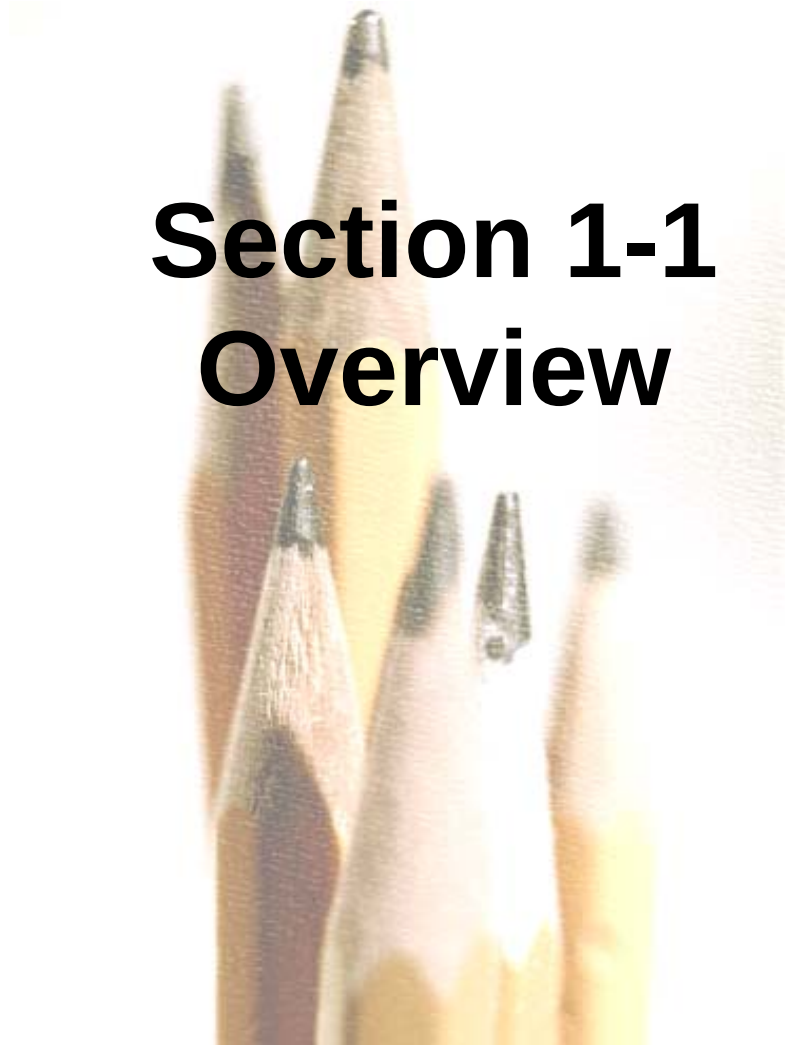
Chapter 1
Introduction to Statistics

1-1 Overview

1-2 Types of Data

1-3 Critical Thinking

1-4 Design of Experiments



Section 1-1 Overview

Created by Tom Wegleitner, Centreville, Virginia

PEARSON
Addison Wesley

Overview

A common goal of studies and surveys and other data collecting tools is to collect data from a small part of a larger group so we can learn something about the larger group.

In this section we will look at some of the ways to describe data.

Definition

- **Data**

  observations (such as measurements, genders, survey responses) that have been collected

Definition

- **Statistics**

  a collection of methods for planning studies and experiments, obtaining data, and then organizing, summarizing, presenting, analyzing, interpreting, and drawing conclusions based on the data

Definition

- **Population**

  the complete collection of all elements (scores, people, measurements, and so on) to be studied; the collection is complete in the sense that it includes all subjects to be studied

Definitions

- **Census**

  Collection of data from every member of a population

- **Sample**

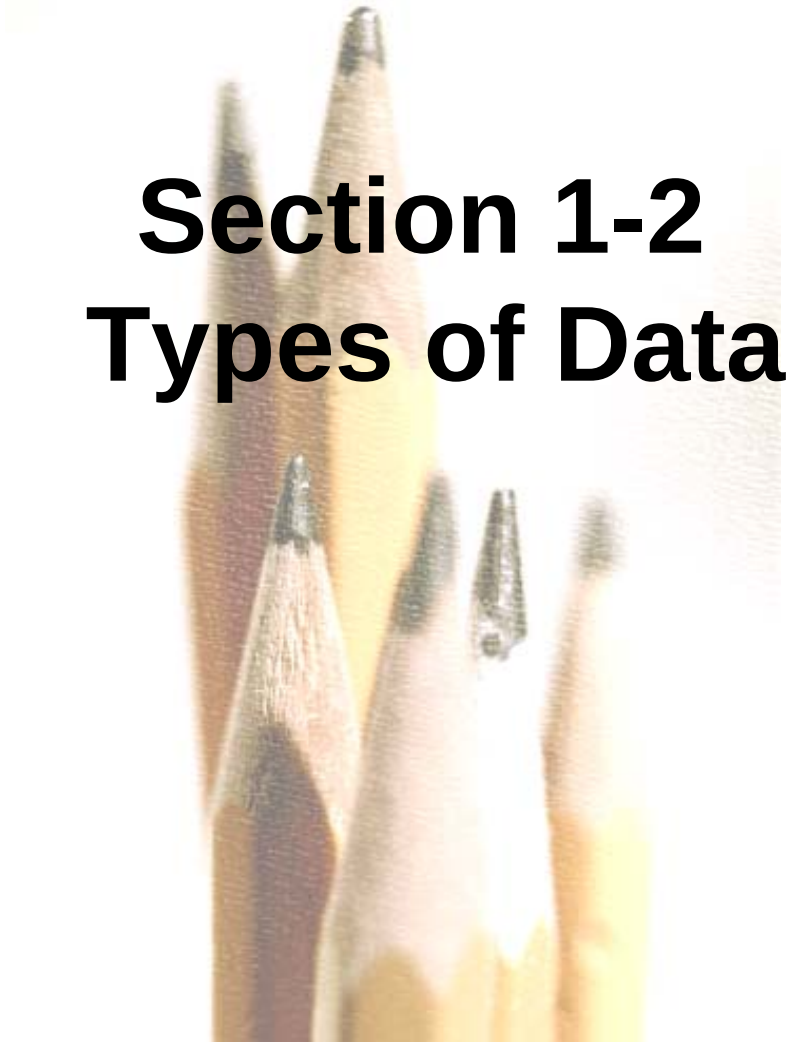
  Subcollection of members selected from a population

Chapter Key Concepts

- **Sample data must be collected in an appropriate way, such as through a process of random selection.**

- **If sample data are not collected in an appropriate way, the data may be so completely useless that no amount of statistical torturing can salvage them.**

Section 1-2
Types of Data



Created by Tom Wegleitner, Centreville, Virginia

PEARSON
Addison Wesley

Key Concept

The subject of statistics is largely about using sample data to make inferences (or generalizations) about an entire population. It is essential to know and understand the definitions that follow.

Definition

- **Parameter**

  a numerical measurement describing some characteristic of a population.

population

parameter

Definition

- **Statistic**

  a numerical measurement describing some characteristic of a sample.

sample

statistic

Definition

- **Quantitative data**

  numbers representing counts or measurements.

  Example: The weights of supermodels

Definition

- **Qualitative (or categorical or attribute) data**

  can be separated into different categories that are distinguished by some nonnumeric characteristic

  Example: The genders (male/female) of professional athletes

Working with Quantitative Data

Quantitative data can further be described by distinguishing between discrete and continuous types.

Definition

Discrete data

result when the number of possible values is either a finite number or a 'countable' number

(i.e. the number of possible values is

0, 1, 2, 3, . . .)

Example: The number of eggs that a hen lays

Definition

- **Continuous (numerical) data**

  result from infinitely many possible values that correspond to some continuous scale that covers a range of values without gaps, interruptions, or jumps

  Example: The amount of milk that a cow produces; e.g. 2.343115 gallons per day

Levels of Measurement

Another way to classify data is to use levels of measurement. Four of these levels are discussed in the following slides.

Definition

- **Nominal level of measurement**

  characterized by data that consist of names, labels, or categories only, and the data cannot be arranged in an ordering scheme (such as low to high)

  Example: Survey responses yes, no, undecided

Definition

❖ Ordinal level of measurement

involves data that can be arranged in some order, but differences between data values either cannot be determined or are meaningless

Example: Course grades A, B, C, D, or F

Definition

- **Interval level of measurement**

  like the ordinal level, with the additional property that the difference between any two data values is meaningful, however, there is no natural zero starting point (where none of the quantity is present)

  Example: Years 1000, 2000, 1776, and 1492

Definition

❖ Ratio level of measurement

the interval level with the additional property that there is also a natural zero starting point (where zero indicates that none of the quantity is present); for values at this level, differences and ratios are meaningful

Example: Prices of college textbooks ($0 represents no cost)

Summary - Levels of Measurement

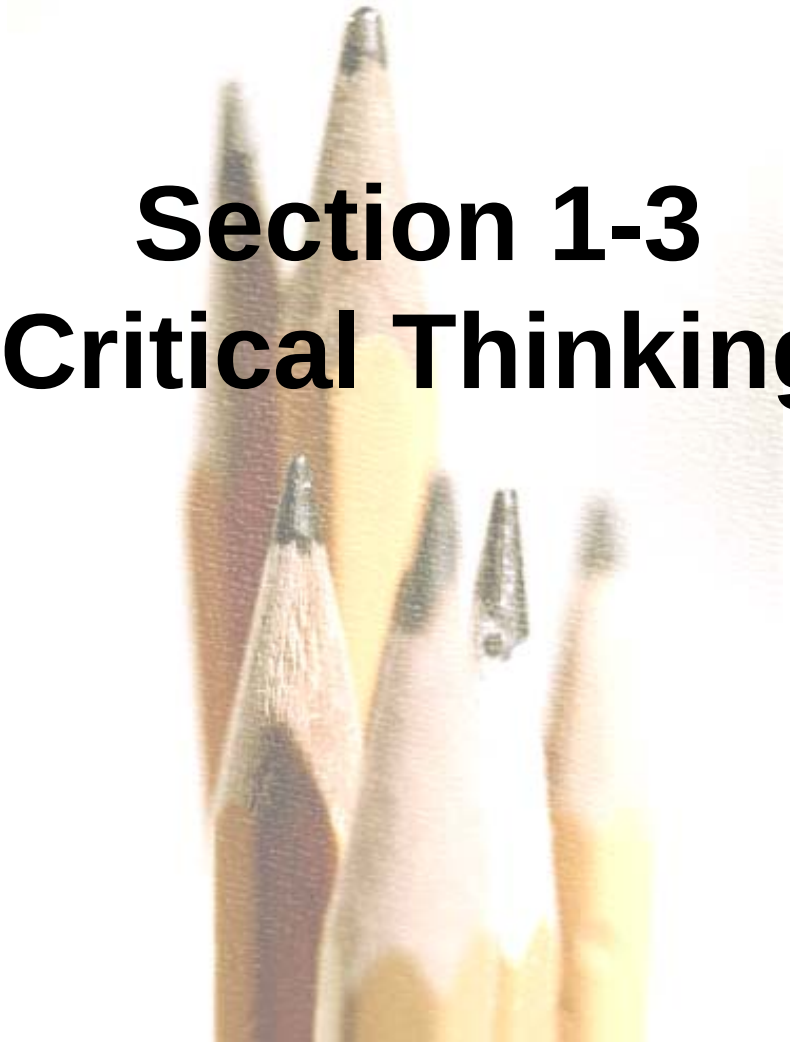
- **Nominal** - categories only
- **Ordinal** - categories with some order
- **Interval** - differences but no natural starting point
- **Ratio** - differences and a natural starting point

Recap

In this section we have looked at:

- **Basic definitions and terms describing data**
- **Parameters versus statistics**
- **Types of data (quantitative and qualitative)**
- **Levels of measurement**

Section 1-3
Critical Thinking



Created by Tom Wegleitner, Centreville, Virginia

PEARSON
Addison Wesley

Key Concepts

- **Success in the introductory statistics course typically requires more common sense than mathematical expertise.**

- **This section is designed to illustrate how common sense is used when we think critically about data and statistics.**

Misuses of Statistics

Misuse # 1- Bad Samples

- **Voluntary response sample (or self-selected sample)**

one in which the respondents themselves decide whether to be included

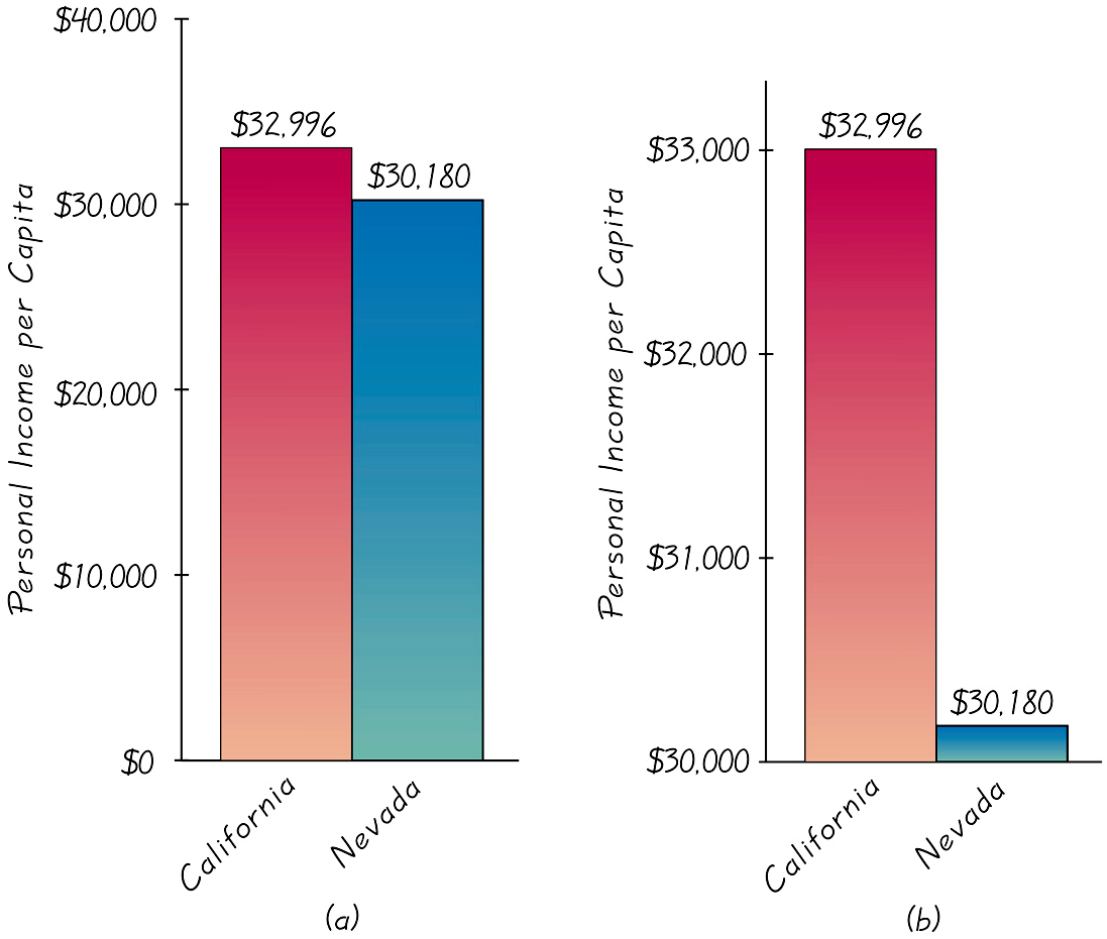
In this case, valid conclusions can be made only about the specific group of people who agree to participate.

Misuse # 2- Small Samples

Conclusions should not be based on samples that are far too small.

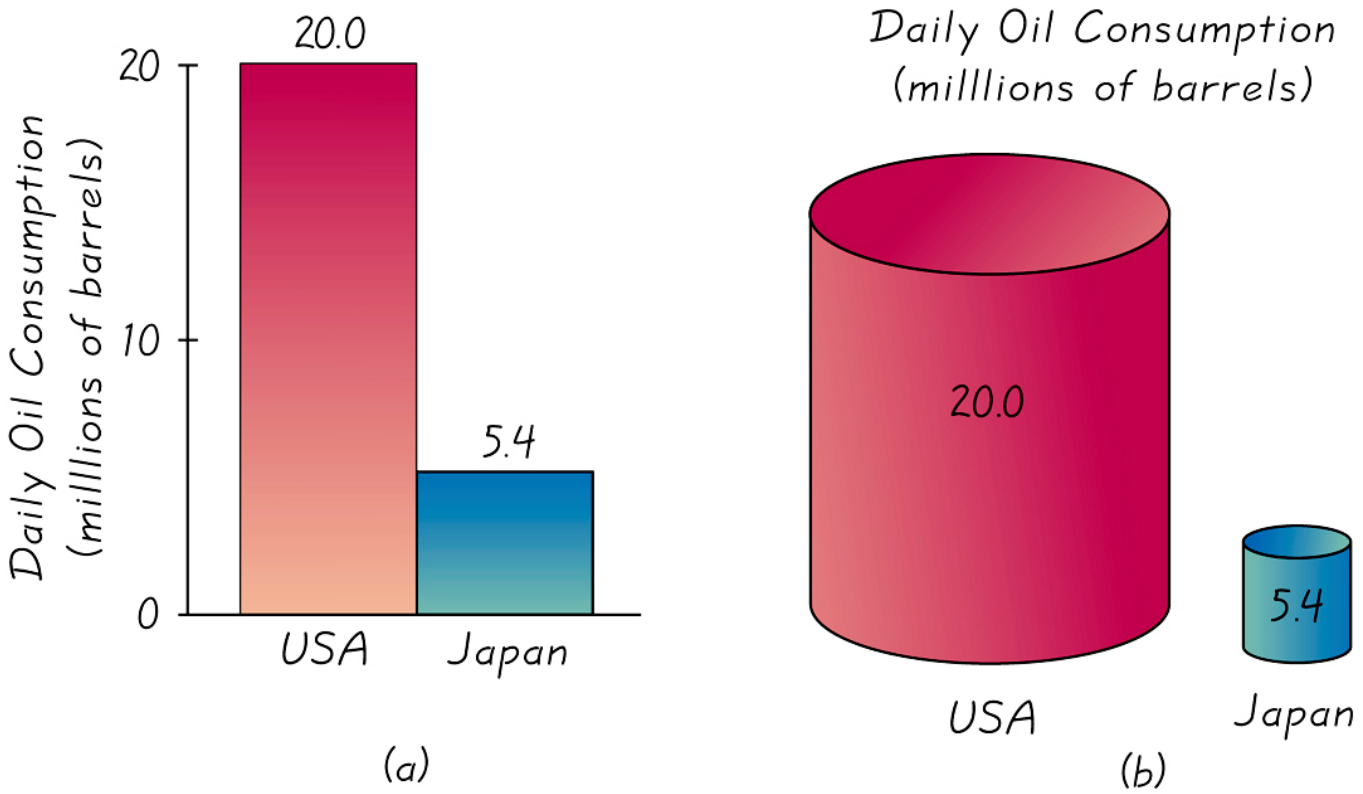
Example: Basing a school suspension rate on a sample of only three students

Misuse # 3- Graphs



To correctly interpret a graph, you must analyze the numerical information given in the graph, so as not to be misled by the graph's shape.

Misuse # 4- Pictographs



Part (b) is designed to exaggerate the difference by increasing each dimension in proportion to the actual amounts of oil consumption.

Misuse # 5- Percentages

Misleading or unclear percentages are sometimes used. For example, if you take 100% of a quantity, you take it all. 110% of an effort does not make sense.

Other Misuses of Statistics

- Loaded Questions
- Order of Questions
- Refusals
- Correlation & Causality
- Self Interest Study
- Precise Numbers
- Partial Pictures
- Deliberate Distortions

Recap

In this section we have:

- **Reviewed 13 misuses of statistics**
- **Illustrated how common sense can play a big role in interpreting data and statistics**

Section 1-4
Design of Experiments

Created by Tom Wegleitner, Centreville, Virginia

PEARSON
Addison Wesley

Key Concept

- **If sample data are not collected in an appropriate way, the data may be so completely useless that no amount of statistical tutoring can salvage them.**

Definition

- **Observational study**

  observing and measuring specific characteristics without attempting to modify the subjects being studied

Definition

Experiment

apply some treatment and then observe its effects on the subjects; (subjects in experiments are called experimental units)

Definitions

- **Cross sectional study**

  data are observed, measured, and collected at one point in time

- **Retrospective (or case control) study**

  data are collected from the past by going back in time

- **Prospective (or longitudinal or cohort) study**

  data are collected in the future from groups (called cohorts) sharing common factors

Definition

Confounding

occurs in an experiment when the experimenter is not able to distinguish between the effects of different factors

Controlling Effects of Variables

- **Blinding**

  subject does not know he or she is receiving a treatment or placebo

- **Blocks**

  groups of subjects with similar characteristics

- **Completely Randomized Experimental Design**

  subjects are put into blocks through a process of random selection

- **Rigorously Controlled Design**

  subjects are very carefully chosen

Replication and Sample Size

Replication

repetition of an experiment when there are enough subjects to recognize the differences from different treatments

Sample Size

use a sample size that is large enough to see the true nature of any effects and obtain that sample using an appropriate method, such as one based on randomness

Definitions

❖ Random Sample

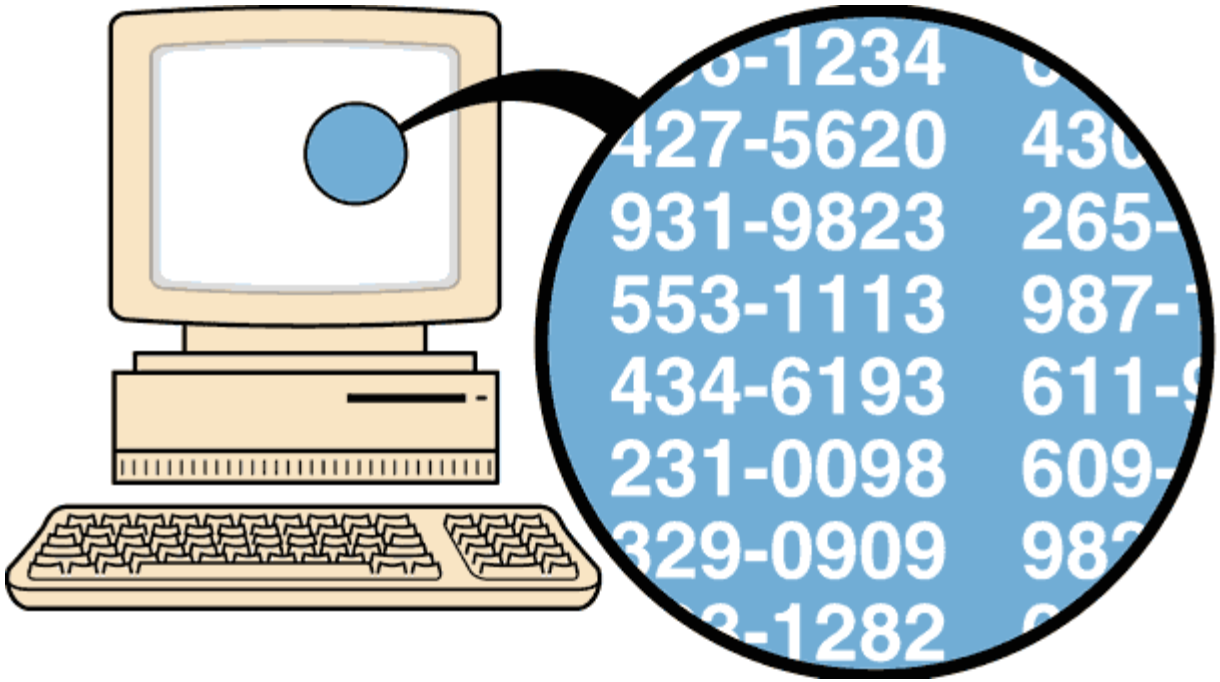
members of the population are selected in such a way that each individual member has an equal chance of being selected

❖ Simple Random Sample (of size n)

subjects selected in such a way that every possible sample of the same size n has the same chance of being chosen

Random Sampling

selection so that each individual member has an equal chance of being selected

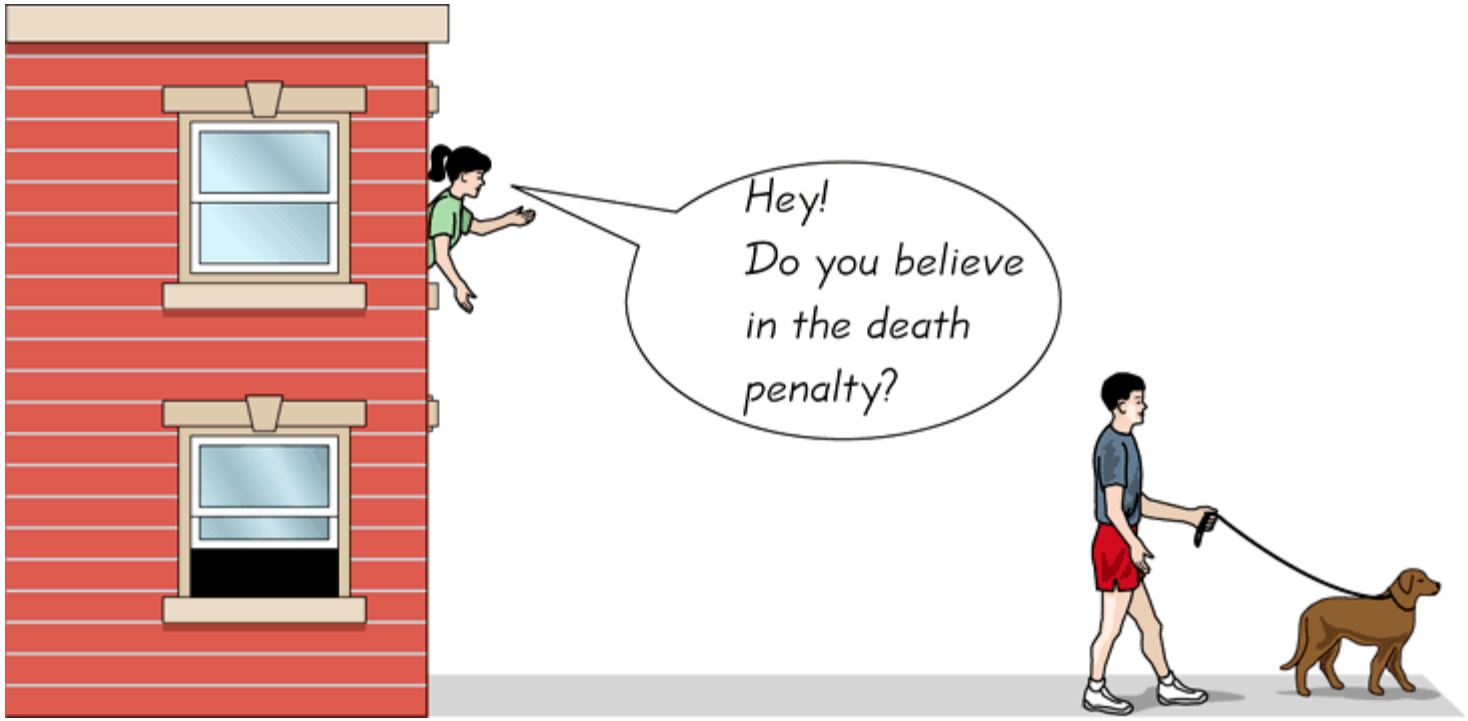


Systematic Sampling

Select some starting point and then select every *k* th element in the population

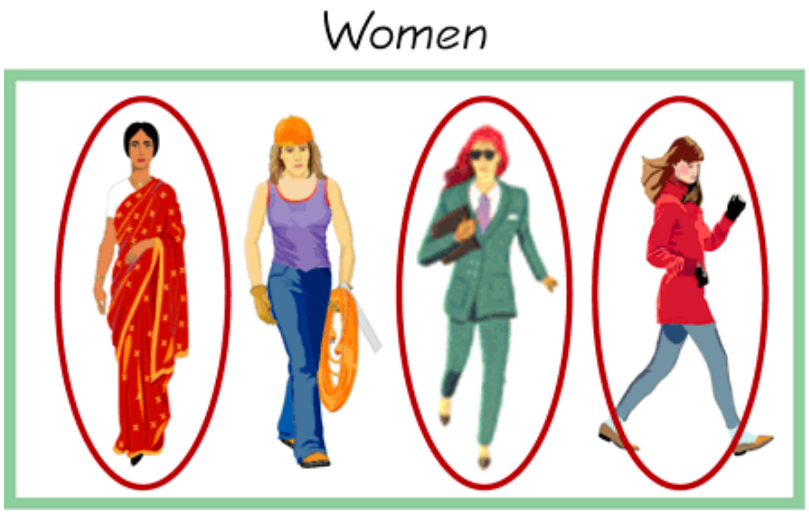
Convenience Sampling

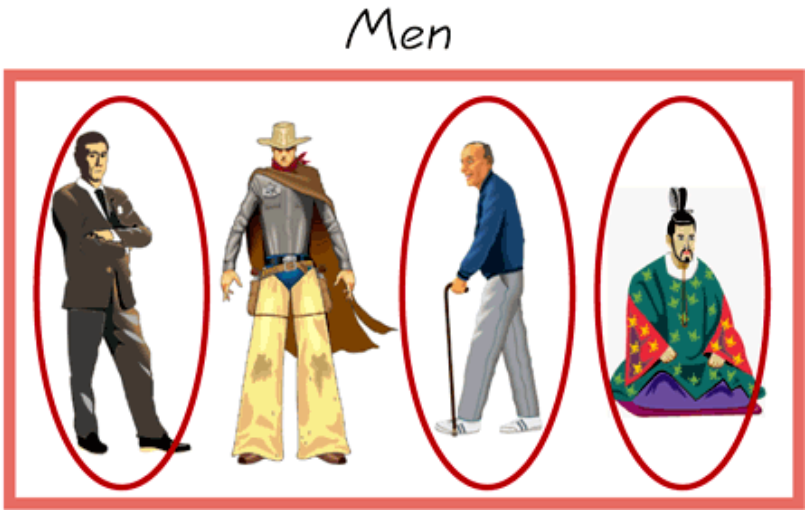
use results that are easy to get



Stratified Sampling

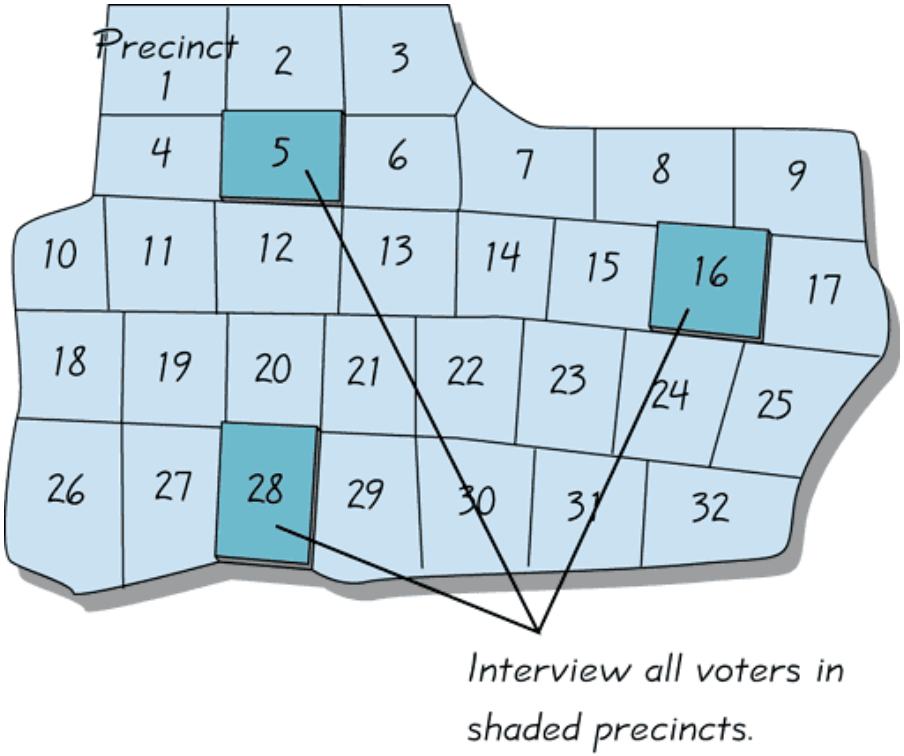
subdivide the population into at least two different subgroups that share the same characteristics, then draw a sample from each subgroup (or stratum)





Cluster Sampling

divide the population into sections (or clusters); randomly select some of those clusters; choose all members from selected clusters



Methods of Sampling - Summary

- Random
- Systematic
- Convenience
- Stratified
- Cluster

Definitions

- **Sampling error**

  the difference between a sample result and the true population result; such an error results from chance sample fluctuations

- **Nonsampling error**

  sample data incorrectly collected, recorded, or analyzed (such as by selecting a biased sample, using a defective instrument, or copying the data incorrectly)

Recap

In this section we have looked at:

- **Types of studies and experiments**
- **Controlling the effects of variables**
- **Randomization**
- **Types of sampling**
- **Sampling errors**