General Education Statistics
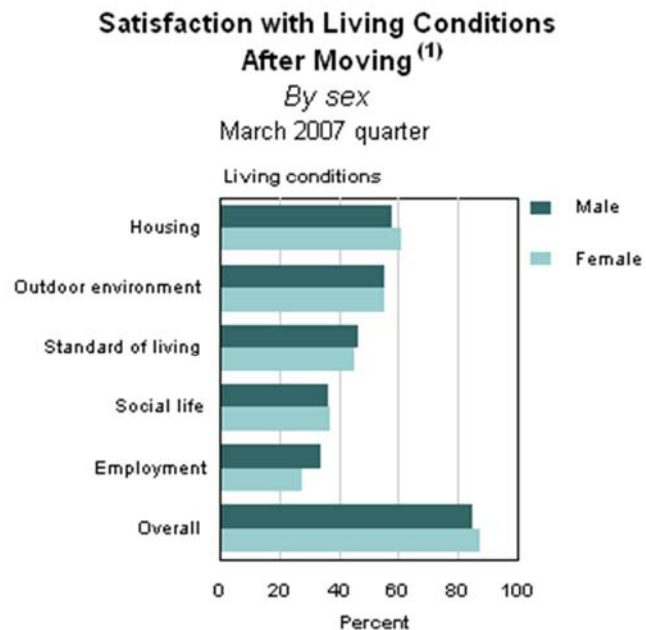Class Notes
Other Sampling Methods (Section 1.4)

We will explore a few other (some good, some bad) methods of sampling.

Simple random sampling is not the only way to sample properly. There are other methods that are less costly in money, time, or other resources. In fact, these other methods may capture the true picture of the population far better than simple random sampling.

Consider this graph that shows the discrepancies between how men and women feel about a recent move.

(Source:
http://archive.stats.govt.nz/browse_for_stats/
population/Migration/internal-migration/benefits-
of-moving-men-women.aspx)

If we were to perform a simple random sample of Americans, we would get an aggregated view of how Americans think. Asking women and men separately shows the differences between the sexes. This is an example of a **stratified random sample**.

Satisfaction with Living Conditions After Moving [1]
By sex
March 2007 quarter

Living conditions

Housing
Outdoor environment
Standard of living
Social life
Employment
Overall

■ Male
■ Female

0   20   40   60   80   100
Percent

(1) Movers who stated that their current living conditions were better or much better than before they moved.

Definition: Stratified Random Sample: A **stratified random sample** is obtained by separating the population into non-overlapping groups called **strata** and then obtaining a simple random sample from each stratum. The individuals within each stratum should be homogeneous (or similar) in some way.

Stratum is singular, strata is plural.

These are very helpful if you have small subgroups within a population, like transgendered people or Native Americans, who may get overlooked otherwise.

1

expl 1: For each proposed survey, determine possible strata that could be used.

a.) survey Americans about their eating habits

Each stratum should be made up of similar individuals.
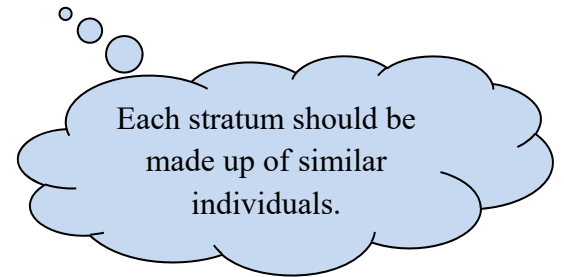
b.) survey the world population about poverty

c.) survey Illinois residents about schools

One big advantage of this technique is that fewer individuals need to be surveyed to get the same information as a simple random sample. You can also report on the differences among groups.

Another technique eliminates the need for a frame for the whole population. Do you remember what a **frame** is?

Definition: Systematic sample: A **systematic sample** is obtained by selecting every k^{th} individual from the population. The first individual selected is a random number between 1 and k.

An example of this would be to survey every fifth house as you walk down a block.

Sometimes this can go awry. I remember hearing of a study where they decided beforehand to survey every third household. The survey team found themselves in a neighborhood of three-story walk-up apartment buildings, each with one apartment per floor. Do you see why this might cause trouble?

Systematic surveys are useful for (voting) exit polls and surveys of customers, among other uses. Advantages over an "srs" include less cost and ease of use. Let's see how they might work.

EXAMPLE Obtaining a Systematic Sample:
A quality control engineer wants to obtain a systematic sample of 10 bottles coming off a filling machine to verify the machine is working properly. Design a sampling technique that can be used to obtain a sample of 10 bottles.

expl 2: Let's assume this machine fills 20,056 bottles a day. We want to sample during one day, obtaining 10 random bottles throughout the day. Follow the steps outlined to determine the sample.

a.) If possible, approximate the population size, N.

b.) Determine the sample size desired, n.

Do you remember k from the definition?

c.) Divide N/n and round down to the nearest integer. This value is k.

d.) Randomly select a number between 1 and k. Call this number p.

e.) The sample will consist of the pth individual and every k^{th} one after that [or individuals p, $p + k$, $p + 2k$, $p + 3k$, …, $p + (n-1)k$]. We will sample these bottles coming off the line.
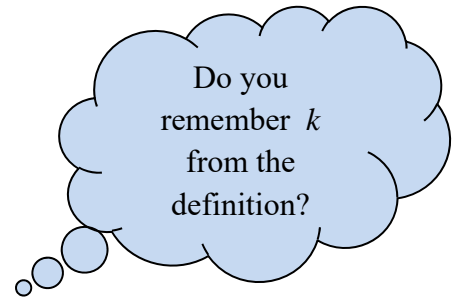
Yet another good technique is the cluster sample.

Definition: Cluster sample: A **cluster sample** is obtained by selecting *all* individuals within a randomly selected collection or group of individuals (clusters).

Here, the technique involves determining clusters (groups of individuals) that you will randomly choose from. Once you have determined the clusters you want included, you survey *every* individual in those clusters.

EXAMPLE Obtaining a Cluster Sample:
A school administrator wants to obtain a sample of students in order to conduct a survey. She randomly selects 10 classes and administers the survey to *all* the students in the selected classes.

Here is a useful illustration of the differences among the four techniques we have studied.
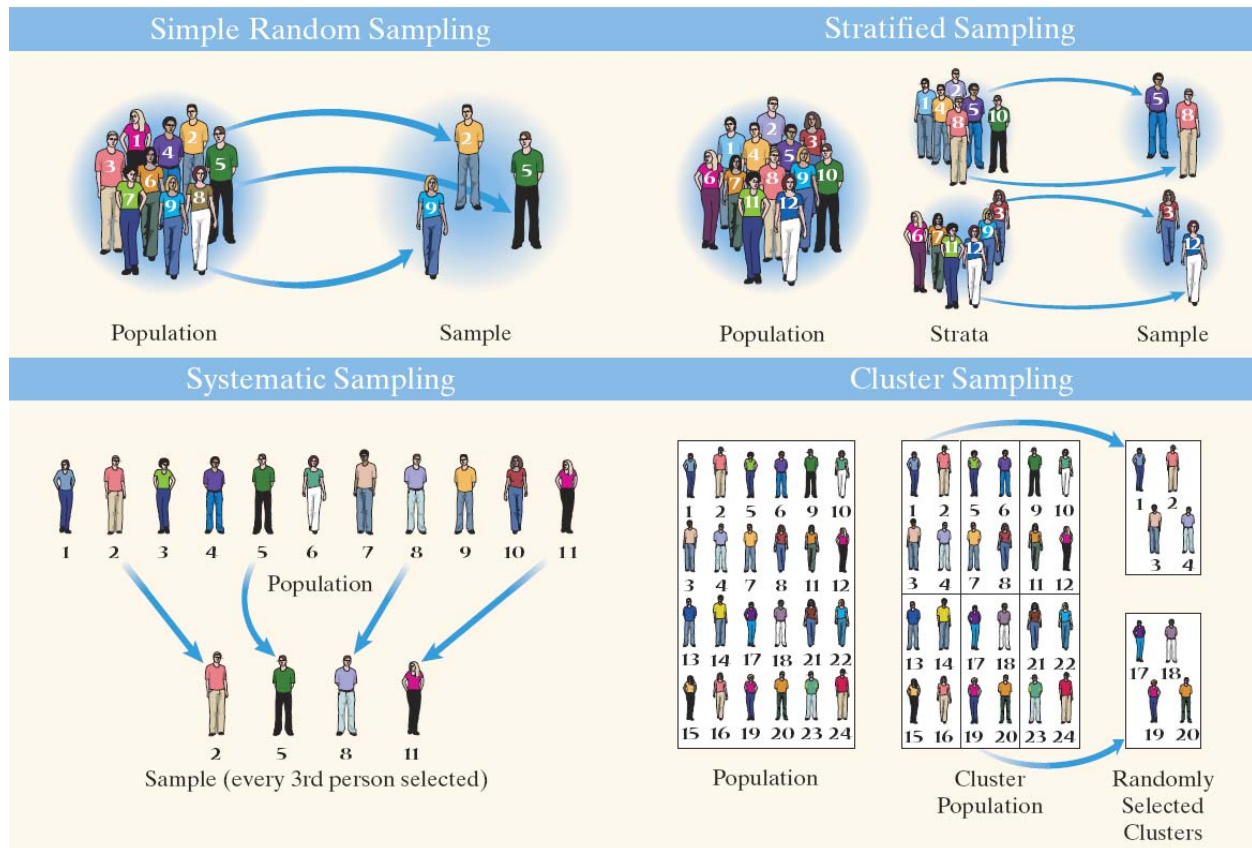
Multistage sampling occurs when a sample uses more than one of these techniques.

Bad Sampling Methods:

We are bombarded by polls that claim to show the views of fellow Americans. Here is one I found online.

(source: https://knss.radio.com/poll-results?page=1)

It seems that every radio and TV show has running polls. The respondents are **self-selected**. What do you think that means?

These are examples of **Voluntary Samples** and are *not* valid. The respondents are *not* chosen randomly and so the results *cannot* be used to infer about the larger population. That does not stop many radio hosts from making that claim.

Another type of bad sample is called the **Convenience Sample**. This is what it sounds like. The individuals in the sample were chosen simply because they were convenient. You have probably been asked to fill out such a survey. They are rampant in colleges as professors give them out for their own research.