

## Chapter Ten: Sample Surveys

Recall way back from chapters one and two the difference between a population and a sample. A \_\_\_\_\_ is the entire body of possible experimental units or observations of interest. So for example, the U.S Census is a study based on an entire population. A \_\_\_\_\_ is a smaller subset of our population of interest. So if I were to randomly select twenty STAT 101 students, that would be a sample. It's no secret that conducting studies on populations is incredibly difficult. So to combat this issue in statistics we often perform \_\_\_\_\_ on \_\_\_\_\_ instead. However, it is difficult to collect samples or surveys and there is a lot we need to consider:

- 
- 
- 
- 
- 

### Three Big Ideas About Sampling

- 1.
- 2.
- 3.

#### Big Idea One: Take a Sample

##### Population

- 
-

- 

### Sample

- 
- 
- 
- 

**Example** *To estimate the proportion of American college students who support changing the drinking age from 21 to 18, a random sample of 100 college students are asked the question “Would you support a measure to lower the drinking age from 21 to 18?”*

**Population:**

**Sample:**

**Example** *A strawberry farmer wants to estimate the average number of strawberries per plant in her field. She randomly chooses 25 plants and counts the number of strawberries on each plant.* **Population:**

**Sample:**

**Example** *A fishing boat captain examines one days catch of fish in a new location to see if the average weight of fish in that area is large enough to make fishing there profitable.*

**Population:**

**Sample:**

### Parameters and Statistics

Like we covered before in our mathematical symbol review, we use different symbols for mathematical calculations if they are from our sample versus the population. We call summary statistics collected from populations our \_\_\_\_\_ and those collected from samples, \_\_\_\_\_.

### Parameters

- 
-

- 

- 

## Statistics

- 

- 

- 

## Big Idea Two: Select the Sample Randomly

Selecting a \_\_\_\_\_ sample leads to beneficial properties:

- 

-

- 

There are four different types of random sampling that we will cover in this course:

- 
- 
- 
- 

### Simple Random Sample

A \_\_\_\_\_ or SRS is a selection scheme where every combination of possible sample members has an equal chance of being selected for the sample.

Procedure:

- 1.
- 2.
- 3.

Suppose we are interested in learning the different breeds of dogs at the dog park. There are way too many dogs for us to count them all. So let's take an SRS.



To take an SRS, we would label each dog that we encounter at the dog park, then use a random number generator to generate  $n$  random numbers. In this case  $n$  is the \_\_\_\_\_.

### Stratified Random Sample

Often large populations will be made up of smaller homogenous groups. So it is in our best interest to reduce bias by making sure each group is represented in our sample. For a population which can be divided into strata, a stratified random sample is a sample which is obtained by taking a collection of random samples from each strata.

- 
- 
- 

Back to our dog example, if we want to make sure each breed is represented in our sample, we would take a stratified random sample.



In this case we number each member of each strata, then randomly generate  $n$  numbers in each strata to select individuals.

### Cluster Sampling

- 1.
- 2.
- 3.

Cluster sampling is another form of sampling that is easier for population sampling since we only need to sample a few clusters. I often like to think of this sort of sampling as a spatial sampling. An example would be if we want a student survey of freshmen at ISU. We would consider each dorm to be a cluster and we could randomly select  $n$  clusters for our survey.

**Multi-Stage Sampling** This sort of sampling is a complicated combination of SRS, Stratified, and Cluster sampling. Returning to our freshman survey, suppose we select three dorms (clustering) for our sample. Then in each of the three dorms we randomly select to floors (stratified) to deliver the survey. Then on each of these selected floors we randomly select students living on these floors to take the survey (SRS).

**Systematic Sampling** A systematic sampling schema is where a sample is drawn from a population in a systematic fashion.

A systematic sampling schema can be representative if:

- 
- 

So if we were to do a systematic sampling schema for our college freshman survey. We might get a list of all of the freshmen at ISU, then randomly start at a point on the list, then select every student 100 spots after the randomly selected starting point and so on.