Data Collection

- Sampling studies (surveys)
  - Single-stage
  - Multiple-stage
- Experiments

Sample Surveys

- Idea 1: Examine a part of the whole.

Properties of a Sample

- Would like the sample to be representative of the population.
- This may not be possible, but at least we would like a sample that is not biased.
Sample Surveys

- Idea 2: Random selection
  - Selecting items from the population should be done at random so as to reduce the chance of getting a biased sample.

Sample Surveys

- Idea 3: It’s the sample size!
  - What fraction of the population is sampled is not important.
  - The size of the sample is the important thing.

What about a census?

- Would a census (complete enumeration) of the population be a better way to go?
  - Difficult to do.
  - Populations are often dynamic.
  - Can be more complex.
Example

• Population: All students at ISU.
• Question: Have you posted a video on YouTube?
• Population parameter: Proportion of all ISU students who would answer yes.

Example

• Sample: 400 ISU students.
• Sample statistic: the proportion of the 400 students in the sample who say yes.

How should we select the 400?

• Put an ad in the ISU Daily with the question and ask students to drop off their answers.
• Go to computer labs across campus and ask the first 400 students you meet.
Simple Random Sample (SRS)

- We want a representative sample but will settle for one that is not biased.
- SRS – Each combination of 400 ISU students has the same chance of being the sample selected.

Simple Random Sample

- Sampling Frame
  - A list of all students at ISU (the Registrar has such a list)
  - Use random numbers to select 400 students at random from this list.

Simple Random Sample

- If one were to do this more than once
  - Different random numbers will give different samples of 400 students.
  - We have introduced variability by sampling!
Other Sampling Plans

• Systematic
  – Select in a systematic way from the sampling frame.
  – Select every 60th student on the list from the Registrar.
  – Caution the list should be in random order and the starting point should be selected at random.

Other Sampling Plans

• Stratified Random
  – Divide population into strata (subpopulations that contain similar items) and select a SRS from each strata.
  – Divide ISU students into colleges and select a SRS from each college.

Other Sampling Plans

• Cluster
  – Divide population into clusters, each cluster having a mix of items representative of the population, and select clusters at random as your sample.