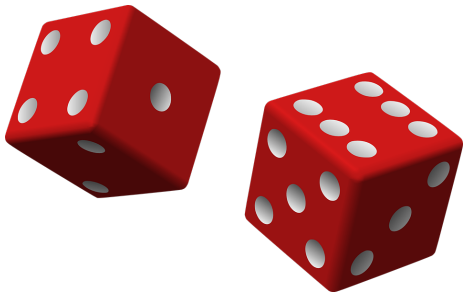
Sampling People, Records, & Networks

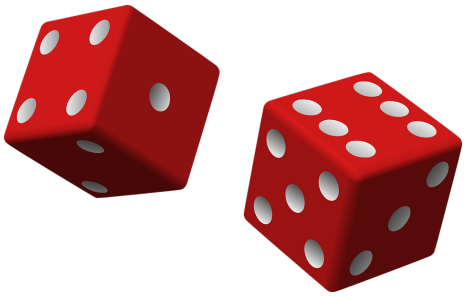
Jim Lepkowski, PhD
Professor & Research Professor *Emeritus*
Institute for Social Research, University of Michigan
Research Professor,
Joint Program in Survey Methodology, University of Maryland



Unit 1

- 1 Research designs ...
- 2 Surveys
- 3 Why sample?
- 4 Why randomize?
- 5 Types of sampling
- 6 Evaluating samples
- 7 Units sampled

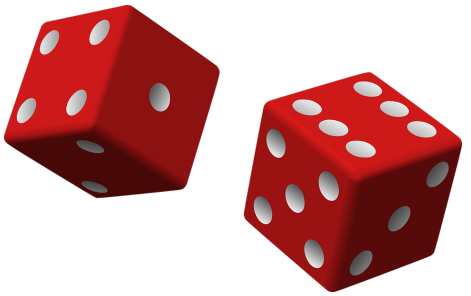
- **Unit 1: Sampling as a research tool**
  - Lecture 1 – Research design & sampling
  - Lecture 2 – Surveys & sampling
  - Lecture 3 -- Why sample at all?
  - Lecture 4 – Why might we randomize, and how do we do it?
  - Lecture 5 – What happens when we randomize?
  - Lecture 6 – How do we evaluate how good the sample is?
  - Lecture 7 – What kinds of things can we sample?
- **Unit 2: Mere randomization**
- **Unit 3: Saving money**
- **Unit 4: Being more efficient**
- **Unit 5: Simplifying sampling**
- **Unit 6: Some extensions & applications**



Unit 1

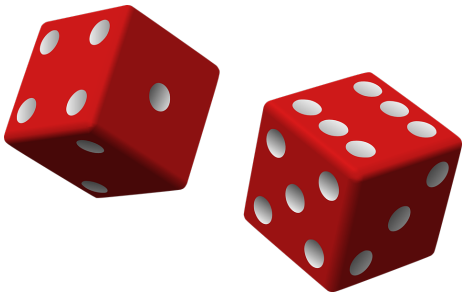
- 1 Research designs ...
- 2 Surveys
- **3 Why sample?**
- 4 Why randomize?
- 5 Types of sampling
- 6 Evaluating samples
- 7 Units sampled

- **Unit 1: Sampling as a research tool**
  - Lecture 1 – Research design & sampling
  - Lecture 2 – Surveys & sampling
  - **Lecture 3 -- Why sample at all?**
  - Lecture 4 – Why might we randomize, and how do we do it?
  - Lecture 5 – What happens when we randomize?
  - Lecture 6 – How do we evaluate how good the sample is?
  - Lecture 7 – What kinds of things can we sample?
- **Unit 2: Mere randomization**
- **Unit 3: Saving money**
- **Unit 4: Being more efficient**
- **Unit 5: Simplifying sampling**
- **Unit 6: Some extensions & applications**



- Census or sample
- Accuracy
- Probabilities
- Frames
- Techniques
- Deficiencies
- Complex design

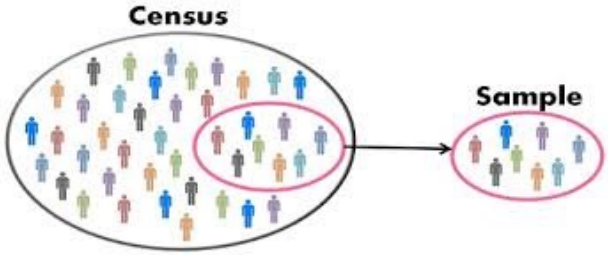
- **Unit 1: Sampling as a research tool**
  - Lecture 1 – Research design & sampling
  - Lecture 2 – Surveys & sampling
  - **Lecture 3 -- Why sample at all?**
  - Lecture 4 – Why might we randomize, and how do we do it?
  - Lecture 5 – What happens when we randomize?
  - Lecture 6 – How do we evaluate how good the sample is?
  - Lecture 7 – What kinds of things can we sample?
- **Unit 2: Mere randomization**
- **Unit 3: Saving money**
- **Unit 4: Being more efficient**
- **Unit 5: Simplifying sampling**
- **Unit 6: Some extensions & applications**



- Census or sample
- Accuracy
- Probabilities
- Frames
- Techniques
- Deficiencies
- Complex design

- During conceptualization, a researcher considers the RELEVANT POPULATION for evaluating the theory/hypothesis

- In designing the data collection, the researcher has two concerns in mind:
  - External validity
  - Cost/benefit calculations for the overall cost of the study

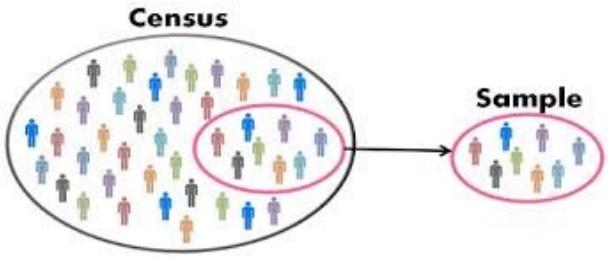


- Census or sample
- Accuracy
- Probabilities
- Frames
- Techniques
- Deficiencies
- Complex design

A census involves an enumeration of a population. When the population is large:

1. It is costly
2. It is time consuming
3. It may not be feasible with complete precision

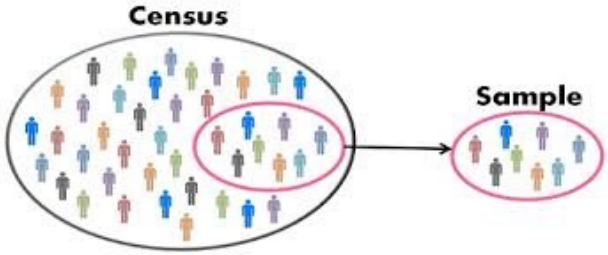
(US Census as an example)



- Census or sample
- Accuracy
- Probabilities
- Frames
- Techniques
- Deficiencies
- Complex design

A sample involves a selection of a representative subset of a population in order to draw inferences to the population

Collecting data from a sample of a large population is FAR LESS costly and FAR LESS time consuming

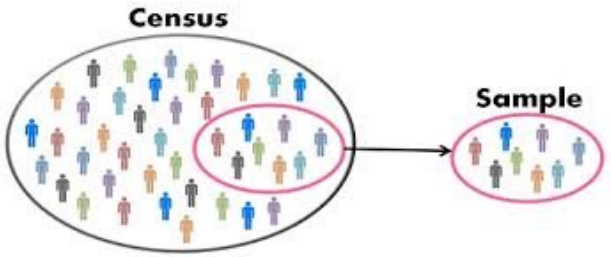


- Census or sample
- Accuracy
- Probabilities
- Frames
- Techniques
- Deficiencies
- Complex design

Recruitment directly – volunteer samples

Lists, selection, & then recruitment

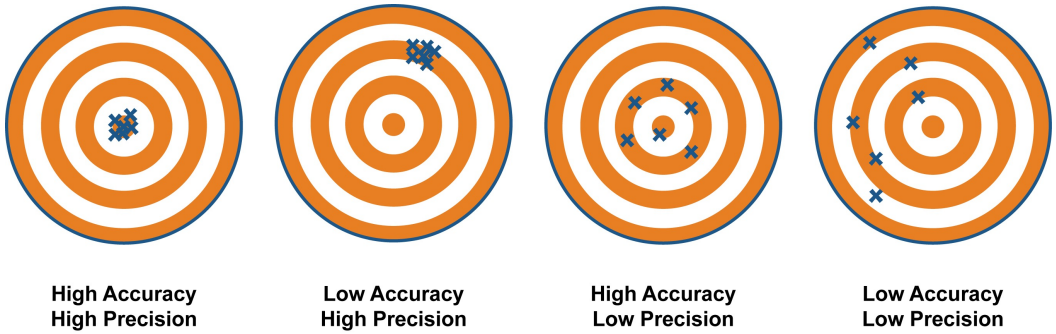
Lists, selection, recruitment, & nonresponse



- Census or sample
- Accuracy
- Probabilities
- Frames
- Techniques
- Deficiencies
- Complex design

- **Because of the cost savings, sampling allows a researcher to devote**
  - **More resources to the collection of more data (variables)**
  - **The reduction of error in measurement (reliability and validity)**
  - **Better coverage of the units of analysis**

- **This fits in with what is called a Total Survey Error perspective**



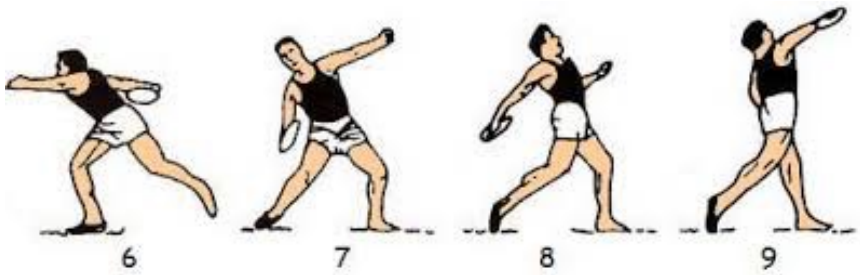
- **Non-probability sampling**
  - **Haphazard, convenience, or accidental sampling**
  - **Purposive sampling or expert choice**
  - **Quota sampling**
  - **Substitution (for nonresponse)**
  - **Online panels**
  - **River sampling**

- **Probability sampling**
  - **Simple random selection**
  - **Stratified selection**
  - **Cluster samples**
  - **Systematic samples**
  - **More complex samples: probabilities proportionate to size**

- List frame
- Area frame
- Problems
  - Missing elements
  - Duplicate listings
  - Clusters
  - Blanks or ineligibles

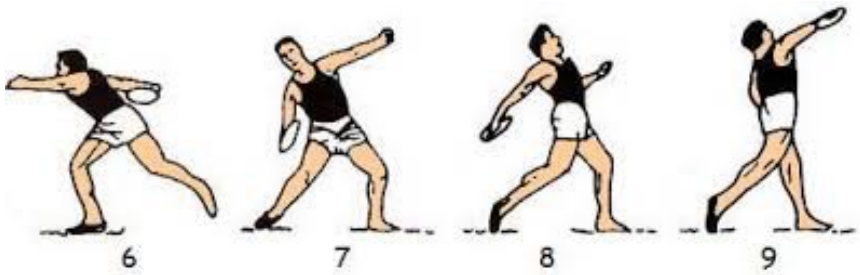
- Census or sample
- Accuracy
- Probabilities
- Frames
- **Techniques**
- Deficiencies
- Complex design

- **Simple random sampling**
- **Systematic sampling**
- **Stratified sampling**
  - **Proportionate allocation**
  - **Disproportionate allocation**



- Census or sample
- Accuracy
- Probabilities
- Frames
- **Techniques**
- Deficiencies
- Complex design

- **Cluster sampling**
- **Two-stage sampling**
- **Probability proportionate to size sampling**
- **Stratified probability proportionate to size sampling**
- **Multistage sampling**
- **Multiple phase sampling**



- **Nonresponse**
  - **Total/unit**
  - **Item**
- **Noncoverage**
- **Compensation: weighting**
  - **Unequal probabilities**
  - **Nonresponse**
  - **Noncoverage (poststratification)**
    - **Make the sample distribution conform to known population distribution**

- **Complex designs typically involve one or more of …**
  - **Stratification**
  - **Clusters**
  - **Weights**
- **Estimation becomes complex**
  - **Even a simple mean or proportion requires non-standard techniques**

- Standard software cannot handle complex sample designs correctly
- Estimating precision becomes more complex as well
- Methods of variance estimation must be considered
  - Taylor series approximation
  - Balanced or Jackknife repeated replication
- Computer software available for these methods
  - Requires stratum, cluster, and weight on each sample record

Unit 1

- 1 Research designs ...
- 2 Surveys
- 3 Why sample?
- **4 Why randomize?**
- 5 Types of sampling
- 6 Evaluating samples
- 7 Units sampled

- **Unit 1: Sampling as a research tool**
  - Lecture 1 – Research design & sampling
  - Lecture 2 – Surveys & sampling
  - Lecture 3 -- Why sample at all?
  - **Lecture 4 – Why might we randomize, and how do we do it?**
  - Lecture 5 – What happens when we randomize?
  - Lecture 6 – How do we evaluate how good the sample is?
  - Lecture 7 – What kinds of things can we sample?
- **Unit 2: Mere randomization**
- **Unit 3: Saving money**
- **Unit 4: Being more efficient**
- **Unit 5: Simplifying sampling**
- **Unit 6: Some extensions & applications**