

# Sampling Concepts

## Z Population

- > The total group of people from whom information is needed

## Z Census

- > Data obtained from every member of the population

## Z Sample

- > A subset of the population of interest

**Question:** TO SAMPLE OR NOT?

**Answer:** GENERALLY YES

**TABLE 13.1** Sample vs. Census

**CONDITIONS FAVORING THE USE OF**

	<b>Sample</b>	<b>Census</b>
1. Budget	Small	Large
2. Time available	Short	Long
3. Population size	Large	Small
4. Variance in the characteristic	Small	Large
5. Cost of sampling errors	Low	High
6. Cost of nonsampling errors	High	Low
7. Nature of measurement	Destructive	Nondestructive
8. Attention to individual cases	Yes	No

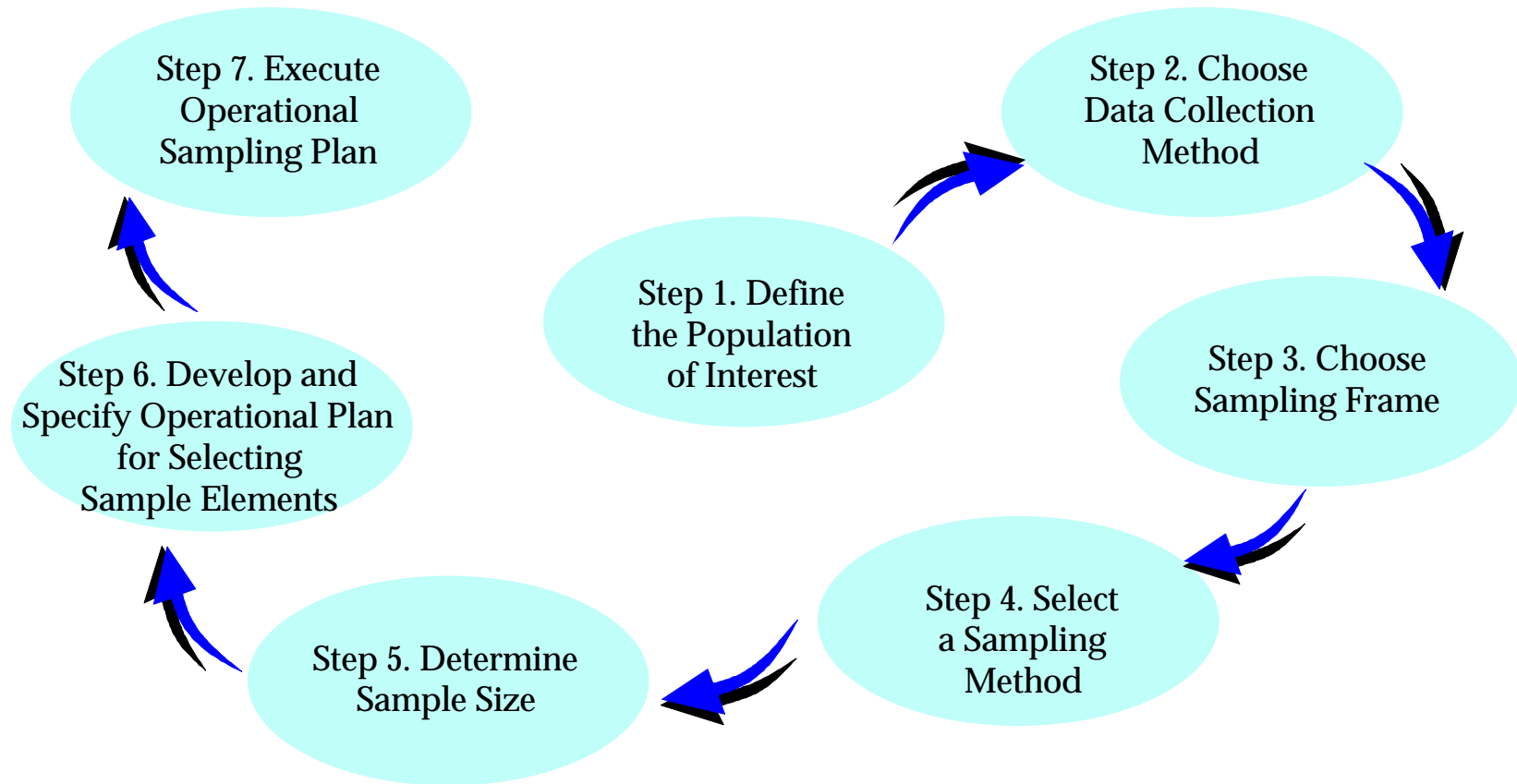
Trade-off

CONFUSING!

**IMPORTANT NOTE:** A sample will usually be more accurate than a census, because of non-sampling error!

$$\text{Total Error} = \text{Sampling error} + \text{N.S. error}$$

# Steps in Developing a Sampling Plan



## Defining the Population

<b>ELEMENT</b>	—————→	<b>Unit (entity) about which information is sought</b>
<b>TARGET POPULATION</b>	————→	<b>Aggregation of all above units (defined prior to SS)</b>
<b>SAMPLING UNIT</b>	—————→	<b>Unit (entity) about which information is actually obtained</b>
<b>STUDY POPULATION</b>	————→	<b>Aggregation of all above units</b>
<b>EXTENT</b>	—————→	
<b>TIME</b>	—————→	

**TABLE 13.1****Some Bases for Defining the Population of Interest**

<b>BASIS</b>	<b>DISCUSSION</b>
<b>Geography</b>	What geographic area is to be sampled? Usually a question of a client's scope of operation. Could be a city, county, metropolitan area, state, group of states, the entire United States, or a number of countries.
<b>Demographics</b>	Given the objectives of the research and the target market for the product, whose opinions, reactions, and so on are relevant? Are we interested in getting information from women over 18; women 18-34; women 18-34 with household incomes over \$35,000 per year, who work, and have preschool children?
<b>Use</b>	In addition to the preceding, the population of interest frequently is defined in terms of some product or service use requirement. This is usually stated in terms of some use versus nonuse or use of some quantity of the product or service or some period of time. The following examples of use screening questions illustrate the point: <ul data-bbox="474 886 1877 1101" style="list-style-type: none"><li>• Do you drink five or more cans, bottles, or glasses of diet soft drinks in a typical week?</li><li>• Have you traveled to Europe for vacation or business purposes in the past two years?</li><li>• Have you or has anyone in your immediate family been in a hospital for an overnight or extended stay in the past two years?</li></ul>
<b>Awareness</b>	We may be interested in surveying those individuals who are aware of the company's advertising to explore what the ad communicated to them about the characteristics of the product or service.

**TABLE 13.2****Example of Screening Question Sequence to Determine Population Membership**

Hello. I'm \_\_\_\_\_ with \_\_\_\_\_ Research. We're conducting a survey about products used in the home. May I ask you a few questions?

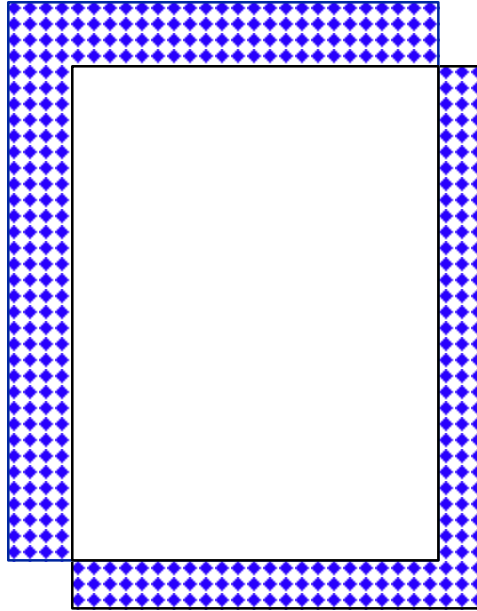
1. Have you been interviewed about any products or advertising in the past ~~3~~ months?
  - Yes TERMINATE AND TALLY
  - No → CONTINUE
2. Which of the following hair care products, if any, have you used in the past month? (HAND PRODUCT CARD TO RESPONDENT. CIRCLE ALL MENTIONS.)
  - 1 Regular Shampoo
  - 2 Dandruff Shampoo
  - 3 Cream Rinse/Instant Conditioner
  - 4 "Intensive" Conditioner

*INSTRUCTIONS:* IF "4" IS CIRCLED--SKIP TO Q.4 AND CONTINUE FOR "INTENSIVE" QUOTA  
 IF "3" IS CIRCLED BUT NOT 4--ASK Q.3 AND CONTINUE FOR "INSTANT" QUOTA

3. You said that you have used a cream rinse/instant conditioner in the past month. Have you used either a cream rinse or an instant conditioner in the past week, or not?
  - Yes (used in the past week) → CONTINUE FOR "INSTANT" QUOTA
  - No (not used in the past week) → TERMINATE AND TALLY
4. Into which of the following groups does your age fall? (READ LIST, CIRCLE AGE.)
  - X Under 18 → CHECK AGE QUOTAS
  - 1 18-24
  - 2 25-34
  - 3 35-44
  - X 45 or over
5. Previous surveys have shown that people who work in certain jobs may have different reactions to certain products. Now, do you or does any member of your immediate family work for an advertising agency, a market research firm, a public relations firm, or a company that manufactures or sells personal care products?
  - Yes TERMINATE AND TALLY
  - No → CONTINUE

**(IF RESPONDENT QUALIFIES, INVITE HER TO PARTICIPATE AND COMPLETE NAME GRID BELOW.)**

## Determine the Sampling Frame



**TARGET POPULATION OF INTEREST (ELEMENTS)**



**STUDY POPULATION (SAMPLING UNITS)**

,

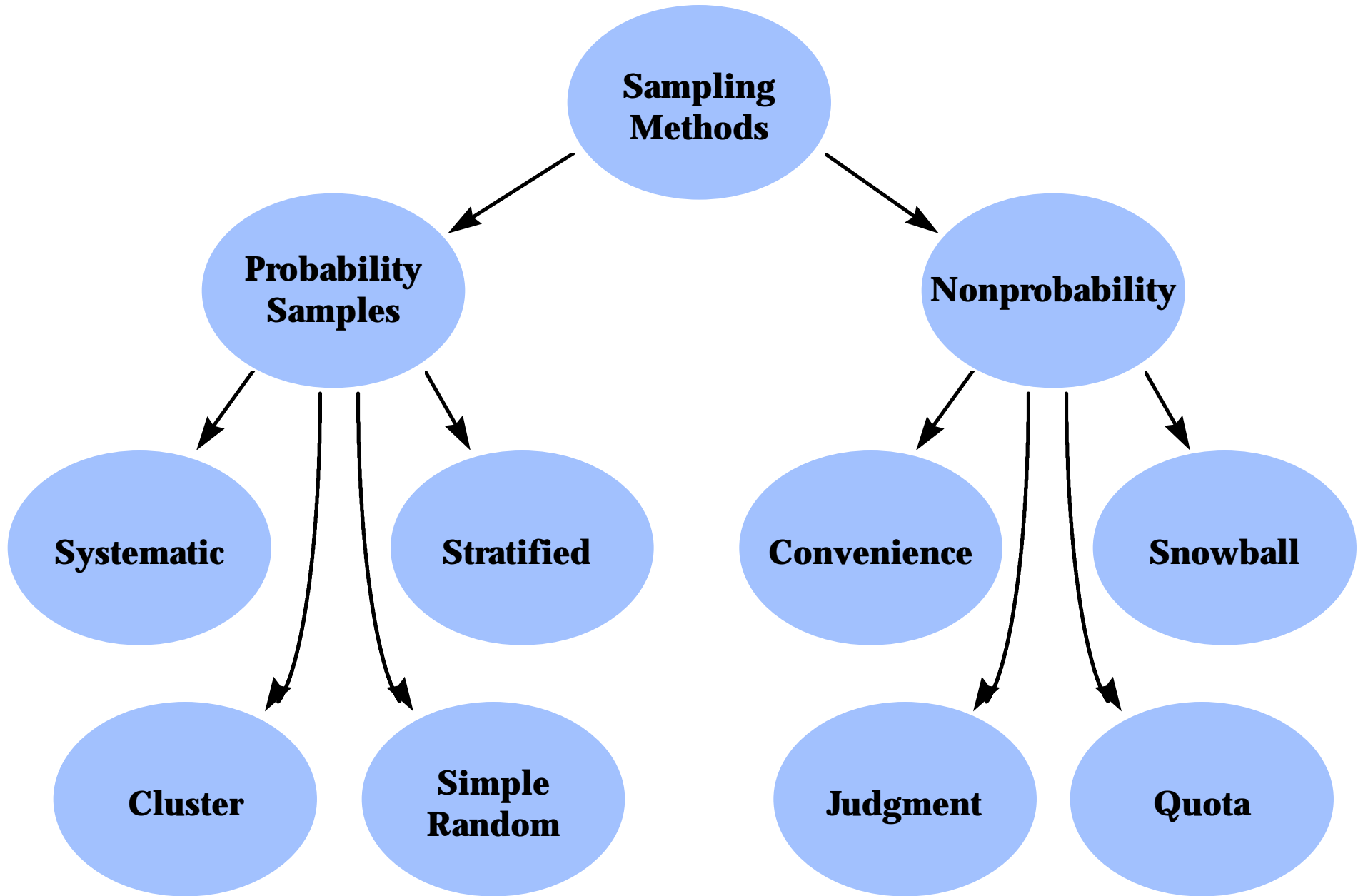
**SAMPLING FRAME ERROR**

**TABLE 13.6****Survey Sampling SSI-LITE Category**

<b>CATEGORY</b>	<b>DESCRIPTION</b>	<b>COUNT</b>
Art/Antiques	interest in clothing	call for count
Astrology/Occult	interest in reading about	call for count
Beer	by usage and by brand	call for count
Bible/Devotional Reading		5,562,000
CD Player	own	7,373,000
Cellular Phone	own	564,000
Cordial/Liqueur Drinkers	by frequency of use	call for count
Crossword Puzzles		2,004,000
Dog Owners	by size of dog	6,804,000
Expecting Child	by date of birth	call for count
Fine Arts/Antiques	interest in	2,945,000
Hemorrhoid Remedy	by frequency of use	call for count
IRA Investment	want	call for count
Mayonnaise	by brand use	call for count
Mouthwash	by brand use	call for count
Plaque Rinse	by frequency of use	call for count
Recreational Vehicles		2,513,000
Students, College	by class year	call for count
Suffers From	back pain	460,170

SOURCE: Courtesy of Survey Sampling, Inc. (c) Copyright Survey Sampling, Inc., 1994

# Classification of Sampling Methods



## **Probability and Nonprobability Sampling**

### ***Probability sampling***



- Objective procedure
- Probability of selection of an element is known and non-zero
- Can calculate sampling error
- Strict procedures to follow
- Yields a representative sample

### ***Probability sampling***



- Subjective procedure
- Probability of selection of an element is unknown
- Cannot calculate sampling error
- Sample not always

**TABLE 13.3****Choosing Nonprobability vs. Probability Sampling**

Factors	<b>CONDITIONS FAVORING THE USE OF</b>	
	Nonprobability Sampling	Probability Sampling
Nature of research	Exploratory	Conclusive
Relative magnitude of sampling and nonsampling errors	Nonsampling errors are larger	Sampling errors are larger
Variability in the population	Homogeneous (low)	Heterogeneous (high)
Statistical considerations	Unfavorable	Favorable
Operational considerations	Favorable	Unfavorable

**IMPORTANT NOTE: Establish trade-off between use of above procedures, before selecting one**

# Sampling Methods

Z Probability Samples → Can Calculate Sampling Error

> Every element of the population has a known nonzero probability of selection

Z Nonprobability Samples → Cannot Calculate Sampling Error

> Selection of specific elements from the population in a nonrandom manner

KNOWN & EQUAL → RANDOM SAMPLE

## **Non-Probability Sampling Methods**

Z Convenience samples

Z Judgement samples

Z Quota samples

Z Snowball samples

## **Probability Sampling Methods**

- Z Simple random sampling
- Z Systematic sampling
- Z Stratified samples
- Z Cluster samples

# Systematic Sampling

## CENSUS OF AGE OF STUDENTS

	(1) Student number	(2) Age ( $X_i$ )		(1) Student number	(2) Age ( $X_i$ )
Graduate students:	1	25	Undergraduate students:	21	21
	<u>2</u>	27		<u>22</u>	19
	3	29		23	24
	4	31		24	22
	5	25		25	20
	6	29		26	22
	7	27		27	19
	8	24		28	20
	9	27		29	19
	10	28		30	24
	11	33		31	25
	<u>12</u>	29		<u>32</u>	22
	13	26		33	20
	14	28		34	21
	15	28		35	21
	16	26		36	23
	17	26		37	21
	18	36		38	23
	19	28		39	18
	20	26		40	21
		41	19		
		<u>42</u>	23		
		43	22		
		44	19		
		45	20		
		46	20		
		47	21		
		48	20		
		49	19		
		50	18		

Pick every 10th (Kth) element after a random start

Adv → No sampling frame is needed

Disadv → Prob. of Periodicity

# Cluster Sampling

---

<b>Group</b>	<b>Population element number</b>
✓ Group 1	<u>1, 2, 3, 4, 5</u>
✓ Group 2	<u>6, 7, 8, 9, 10</u>
Group 3	11, 12, 13, 14, 15
Group 4	16, 17, 18, 19, 20

---

$$N = 20$$

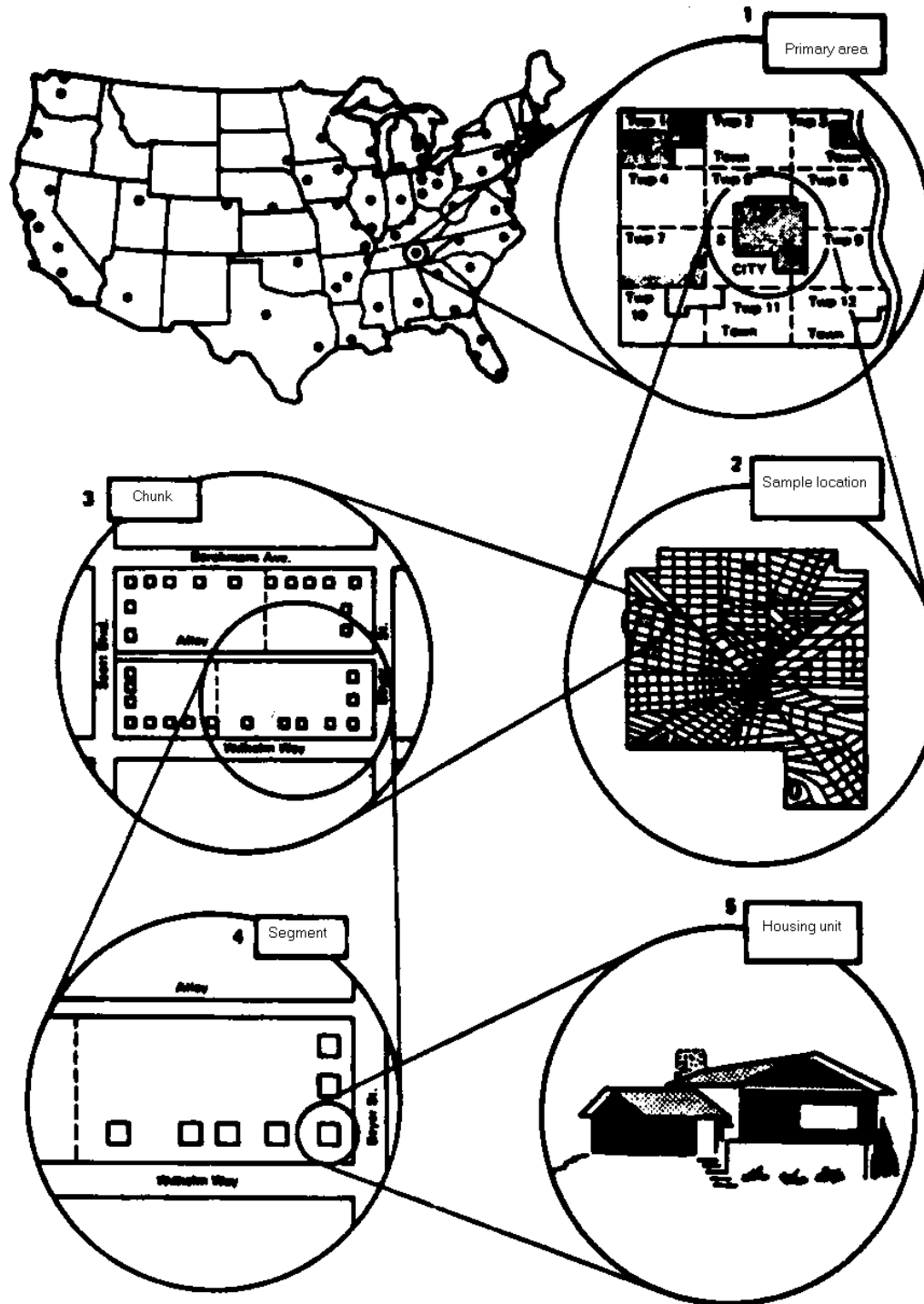
$$n = 10$$

TARGET POPULATION

=

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20
--

# Cluster Sampling



An illustration of area sampling.  
(Source: Adapted from  
*Interviewer's Manual*, rev.ed., Ann  
Arbor: Survey Research Center,  
Institute for Social Research,  
University of Michigan, 1976, p.  
36.)

**TABLE 13.5****Example of Operational Sampling Plan**

---

In the instructions that follow, reference is made to follow your route around a “block.” In cities this will be a city block. In rural areas, a “block” is a segment of land surrounded by roads.

1. If you come to a dead end along your route, proceed down the opposite side of the street, road, or alley, traveling in the other direction. Continue making right turns, where possible, calling at every third occupied dwelling.
2. If you go all the way around a block and return to the starting address without completing four interviews in listed telephone homes, attempt an interview at the starting address. (This should seldom be necessary.)
3. If you work an entire block and do not complete the required interviews, proceed to the dwelling on the opposite side of the street (or rural route) that is *nearest* the starting address. Treat it as the next address on your Area Location Sheet and interview that house only if the address appears next to an “X” on your sheet. If it does not, continue your interviewing to the left of that address. Always follow the right turn rule.
4. If there are no dwellings on the street or road opposite the starting address for an area, circle the block opposite the starting address, following the right turn rule. (This means that you will circle the block following a clockwise direction.) Attempt interviews at every third dwelling along this route.
5. If, after circling the adjacent block opposite the starting address, you do not complete the necessary interviews, take the next block found, *following a clockwise direction*.
6. If the third block does not yield the dwellings necessary to complete your assignment, proceed to as many blocks as necessary to find the required dwellings; these blocks follow a clockwise path around the primary block.

---

SOURCE: Reprinted from interviewer guide by permission of Belden Associates, Dallas, Texas. The complete guide was over 30 pages long and contained maps and other aids for the interviewer.

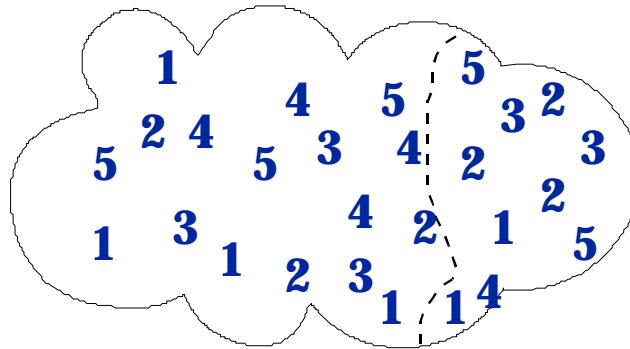
# Simple Random Sampling

## CENSUS OF AGE OF STUDENTS

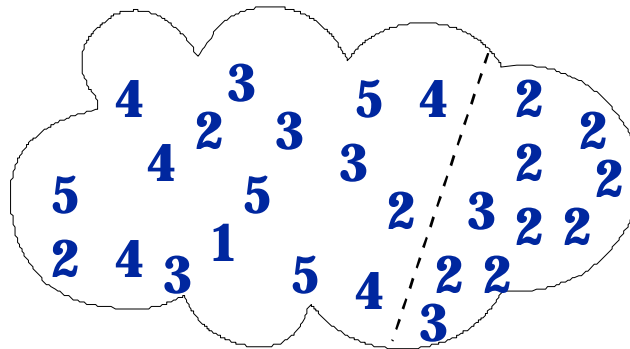
---

	(1) Student number	(2) Age ( $X_i$ )		(1) Student number	(2) Age ( $X_i$ )
Graduate students:	1	25	Undergraduate students:	21	21
	2	27		22	19
	3	29		23	24
	4	31		24	22
	→ 5	25		25	20
	6	29		26	22
	7	27		27	19
	8	24		28	20
	9	27		29	19
	10	28		30	24
	11	33		31	25
	12	29		→ 32	22
	13	26		33	20
	14	28		34	21
	15	28		35	21
	16	26		36	23
	→ 17	26		→ 37	21
	18	36		38	23
	19	28		39	18
	20	26		→ 40	21
		→ 41	19		
		42	23		
		43	22		
		44	19		
		45	20		
		46	20		
		47	21		
		48	20		
		49	19		
		50	18		

# Stratified Sampling



**n = 10**

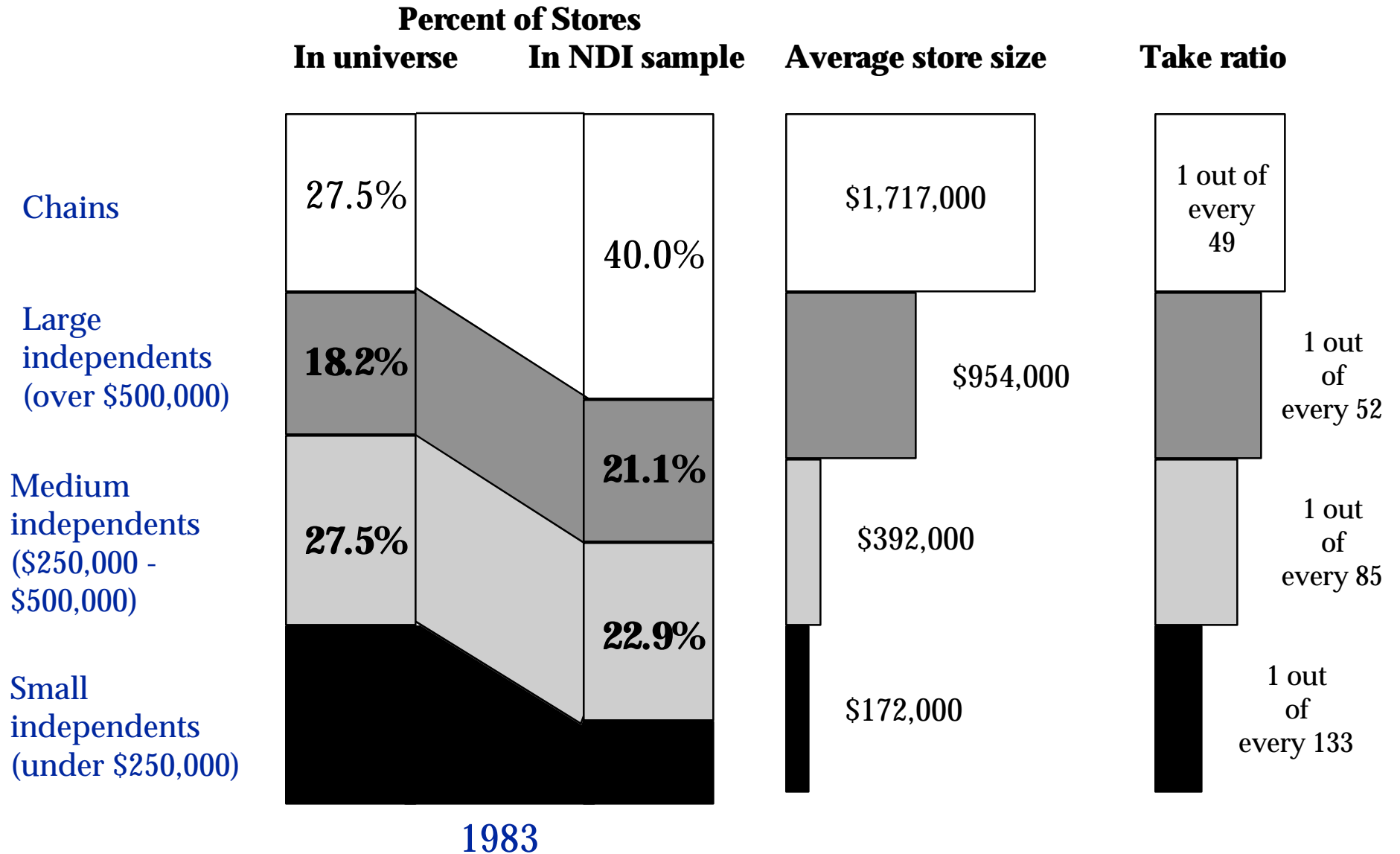


**n = 7**

**PURPOSE** → Reduce sample size for a given level of std. error ... or decrease std. error for a given sample size

Disproportionate stratified sampling in retail store audits. ( *Source: Nielsen Retail Index Services, A.C. Nielsen Company, Northbrook, IL. Reprinted with permission.*)

## DEMONSTRATION OF DISPROPORTIONATE SAMPLING CONCEPT DRUG STORES



**TABLE 13.2****Strengths and Weaknesses of Basic Sampling Techniques**

<b>Technique</b>	<b>Strengths</b>	<b>Weaknesses</b>
<b><i>Nonprobability Sampling</i></b>		
Convenience sampling	Least expensive; least time consuming; most convenient	Selection bias; sample not representative; not recommended for descriptive or causal research
Judgmental sampling	Low cost; convenient; not time consuming	Does not allow generalization; subjective
Quota sampling	Sample can be controlled for certain characteristics	Selection bias; no assurance of representativeness
Snowball sampling	Can estimate rare characteristics	Time consuming
<b><i>Probability Sampling</i></b>		
Simple random sampling (SRS)	Easily understood; results projectable	Difficult to construct sampling frame; expensive; lower precision; no assurance of representativeness
Systematic sampling	Can increase representativeness; easier to implement than SRS; sampling frame not necessary	Can decrease representativeness
Stratified sampling	Includes all important subpopulations; precision	Difficult to select relevant stratification variables; not feasible to stratify on many variables; expensive
Cluster sampling	Easy to implement; cost effective	Imprecise; difficult to compute and interpret results

**Figure 13.3 Improving Response Rates**

**KEY TO SAMPLE  
"REPRESENTATIVENESS"**

