

ELEVENTH EDITION

EXPLORING
Marketing Research



Barry Babin | William Zikmund

Chapter 12

Sampling and Statistical Theory

LEARNING OUTCOMES

After studying this chapter, you should

1. Explain reasons for taking a sample rather than a complete census
2. Describe the process of identifying a target population and selecting a sampling frame
3. Compare random sampling and systematic (nonsampling) errors with an emphasis on how the Internet is intertwined with this issue
4. Identify the types of nonprobability sampling, including their advantages and disadvantages
5. Summarize various types of probability samples
6. Discuss how to choose an appropriate sample design

Introduction

- A sample is a subset or some part of a larger population
- Purpose of sampling is to estimate an unknown characteristic of a population
- Sampling is defined in terms of the population being studied

Introduction (cont'd.)

- A population (universe) is any complete group sharing some common set of characteristics
- The term population element refers to an individual member of the population
- A census is an investigation of all the individual elements making up the population—a total enumeration rather than a sample

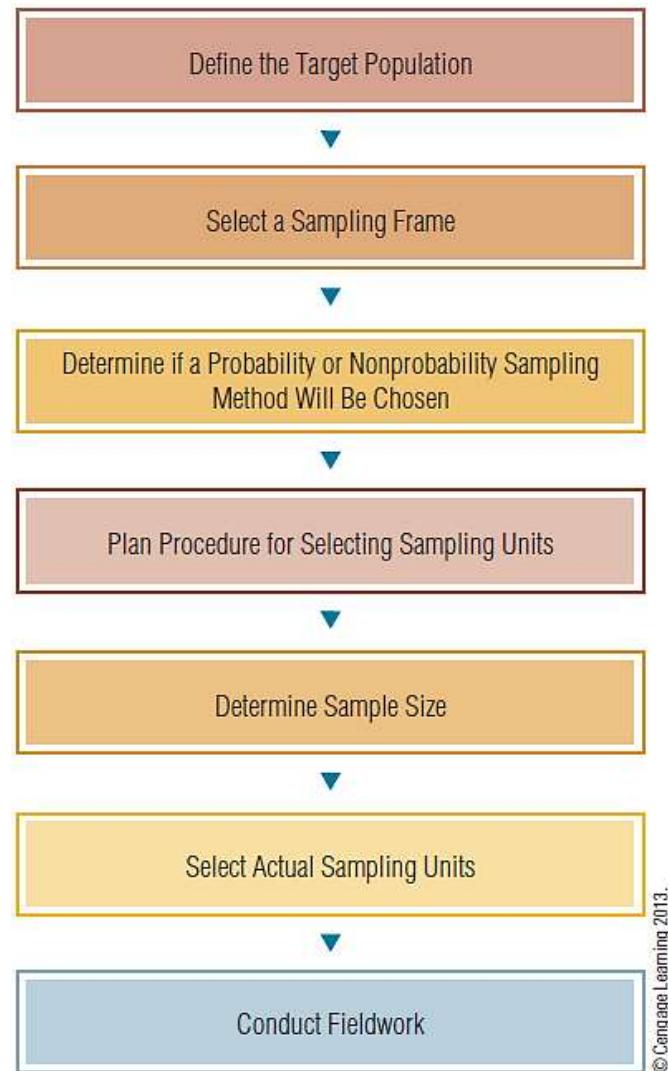
Why Sample?

- Pragmatic reasons
 - Sampling cuts costs, reduces labor requirements, and gathers vital information quickly
- Accurate and reliable results
 - A sample on occasion is more accurate than a census
 - ❖ Increased volume of work in a census may lead to interviewer mistakes, tabulation errors, and other nonsampling errors
- Destruction of test units
 - Occurs in the process of the research project
 - Provides the case against using a census

Identifying a Relevant Population and Sampling Frame

- Defining the target population
 - Once the decision to sample has been made, the first question concerns identifying the target population
 - For consumer research, the appropriate population element frequently is the household
 - The sample is implemented using the tangible, identifiable characteristics that also define the population
 - If the population members cannot be reached by an appropriate communication method, they cannot be part of a sample

EXHIBIT 12.2 Stages in the Selection of a Sample



© Cengage Learning 2013.

The Sampling Frame

- A list of elements from which the sample may be drawn
 - Also called the working population
- Sampling frame errors
 - Occur when certain sample elements are excluded or when the entire population is not accurately represented in the sampling frame
- Almost every list excludes some members of the population

Sampling Services

- Firms specializing in providing lists or databases of specific populations
 - Also called list brokers
- Equifax City Directory provides complete, comprehensive, and accurate business and residential information
- A reverse directory provides listings by city and street address or by phone number
 - Useful when a researcher wishes to survey only a certain geographical area

Online Panels

- Lists of respondents who have agreed to participate in marketing research
 - Generally contain millions of potential respondents
 - The more specific the profile requested, the more expensive the panel

Sampling Frames for International Marketing Research

- The availability of sampling frames around the world varies dramatically
- Not every country's government conducts a census of the population

Sampling Units

- The sampling unit is a single element or group of elements subject to selection in the sample
- If the target population has first been divided into units, additional terminology must be used
 - Primary sampling unit (PSU) designates units selected in the first stage of sampling
 - Secondary sampling units describes units in successive stages of sampling
- When there is no list of population elements
 - The sampling unit is generally something other than the population element

Random Sampling and Nonsampling Errors

- Statistical error—a difference exists between the value of a sample statistic and the value of the corresponding population parameter
- Two basic causes of differences
 - Random sampling errors
 - Systematic (nonsampling) errors
- Random sampling error—the difference between the sample result and the result of a census conducted using identical procedures

Random Sampling Error

- A technical term that refers only to statistical fluctuations that occur because of chance variations in the elements selected for the sample
- A function of sample size
 - As sample size increases, random sampling error decreases
 - Margin of error is determined by the sample size

Systematic Sampling Error

- Systematic (nonsampling) errors result from nonsampling factors, primarily the nature of a study's design and the correctness of execution
 - These errors are not due to chance fluctuations
- Sample biases account for a large portion of errors in marketing research
- Errors due to sample selection problems are nonsampling errors

Systematic But Not Obvious Sampling Error

- Internet surveys allow researchers to reach a large sample rapidly
 - Both an advantage and a disadvantage
- Due to the flood of online questionnaires, frequent Internet users may be more selective about which surveys they bother answering

Website Visitors

- These unrestricted samples are clearly not random samples
 - May not be representative because of the haphazard manner by which many respondents arrived at a particular website or because of self-selection bias
- A better technique for sampling website visitors is to randomly select sampling units

Website Visitors (cont'd.)

- Survey software can be used to trigger a pop-up survey for every N th visitor or on information gathered on the respondent's Web behavior
- Randomly selecting Website visitors can cause a potential problem
 - May over-represent the more frequent visitors to the site
 - Programming techniques and technologies that can help accomplish more representative sampling based on site traffic ► "cookies," registration data, or pre-screening

Panel Samples

- Consumer panels provide a practical sampling frame in many situations
 - There is some concern regarding the representativeness of these samples
 - Researchers must take more steps to ensure that the sampling units do indeed represent the population

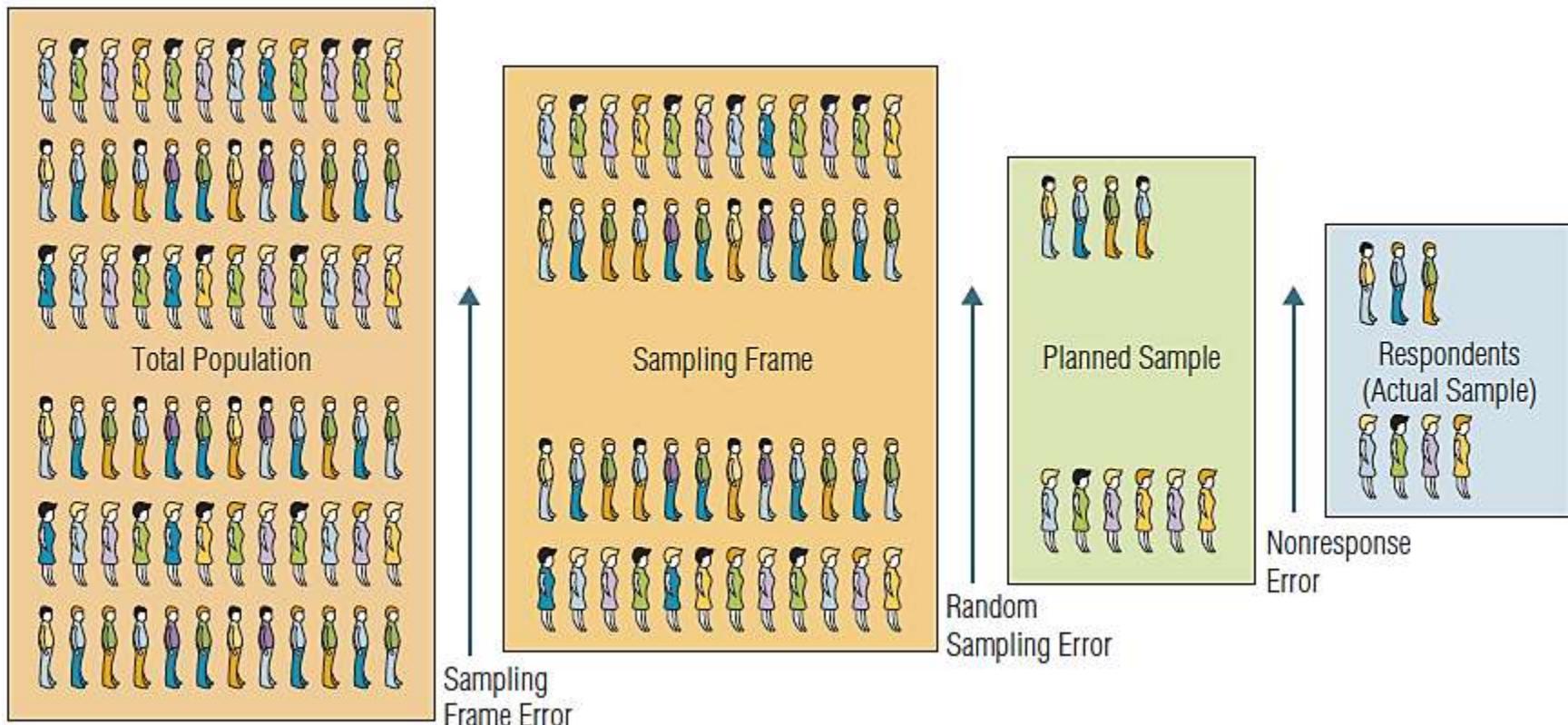
Opting In

- Opt in—to give permission to receive selected e-mail from a company with an Internet presence
 - Spamming is not tolerated by experienced Internet users and can backfire
- Sites like Amazon's Mechanical Turk provide another opportunity for respondents to opt in to surveys
 - These respondents participate as an unscreened, paid respondent; therefore, not random

Less Than Perfectly Representative Samples

- Random sampling errors and systematic errors associated with the sampling process may combine to yield a sample that is less than perfectly representative of the population
- Additional errors will occur if individuals refuse to be interviewed or cannot be contacted
 - Such nonresponse error may also cause the sample to be less than perfectly representative

EXHIBIT 12.3 Errors Associated with Sampling



Probability Versus Nonprobability Sampling

- Several alternative ways to take a sample are available
- The main alternative sampling plans may be grouped into two categories
 - Probability techniques
 - Nonprobability techniques

Probability Versus Nonprobability Sampling (cont'd.)

- Probability sampling—every population element has a known, nonzero probability of selection
 - Simple random sample is the best-known probability sample
- Nonprobability sampling—probability of any member of the population being chosen is unknown
 - The selection of sampling units is quite arbitrary
- Nonprobability samples are pragmatic and are used in market research

Convenience Sampling

- Sampling by obtaining people or units that are conveniently available
- Used to obtain a large number of completed questionnaires quickly and economically, or when obtaining a sample through other means is impractical
- Employed when research is looking for cross-cultural differences in organizational or consumer behavior

Judgment Sampling

- A nonprobability technique in which an experienced individual selects the sample based on his or her judgment
 - Also called purposive sampling
- The consumer price index (CPI) is based on a judgment sample
- Test-market cities often are selected because they are viewed as typical cities whose demographic profiles closely match the national profile

Quota Sampling

- Used to ensure that the various subgroups in a population are represented on pertinent sample characteristics to the exact extent that the investigators desire
- In quota sampling, the interviewer has a quota to achieve
- Aggregating the various interview quotas yields a sample representing the desired proportion of the subgroups

Quota Sampling: Possible Sources of Bias

- Respondents are selected according to a convenience sampling procedure rather than on a probability basis (as in stratified sampling)
 - The haphazard selection of subjects may introduce bias
- Quota samples tend to include people who are easily found, willing to be interviewed, and middle class

Advantages of Quota Sampling

- The major advantages
 - Speed of data collection
 - Lower costs
 - Convenience
- Careful supervision of the data collection may provide a representative sample for analyzing the various subgroups within a population
 - May be appropriate when the researcher knows that a certain demographic group is more likely to refuse to cooperate with a survey

Snowball Sampling

- A variety of procedures in which initial respondents are selected by probability methods
- Additional respondents are then obtained from information provided by the initial respondents
 - This technique is used to locate members of rare populations by referrals
 - Reduced costs and sample sizes are clear-cut advantages of snowball sampling
 - Possible bias due to the referred member being similar to the first person who made the referral

Probability Sampling

- Based on chance selection procedures
 - Eliminate the bias inherent in nonprobability sampling procedures because the probability sampling process is random
- Randomness characterizes a procedure whose outcome cannot be predicted because it depends on chance
- Randomness is the basis of all probability sampling techniques

Simple Random Sampling

- A sampling procedure that assures that each element in the population will have an equal chance of being included in the sample
- Sample selection when populations consist of large numbers of elements
 - Utilizes tables of random numbers or computer-generated random numbers

Systematic Sampling

- An initial starting point is selected by a random process; then every n th number on the list is selected
- While this procedure is not actually a random selection procedure, it does yield random results if the arrangement of the items in the list is random in character
- The problem of periodicity occurs if a list has a systematic pattern, that is, if it is not random in character

EXHIBIT 12.4 Systematically Sampling from a List

	Observation	Month	Year	Average Temperature (C°)	
Random Starting Point	1	January	2010	4	
	2	February	2010	6	
	3	March	2010	10	
	4	April	2010	15	
	5	May	2010	16	
	6	June	2010	22	
Select	7	July	2010	26	
	8	August	2010	22	
	9	September	2010	27	
	10	October	2010	19	
	11	November	2010	10	
Select	12	December	2010	6	
	13	January	2011	22	
	14	February	2011	9	
	15	March	2011	9	
	16	April	2011	12	
	Select		17	May	2011
			18	June	2011
			19	July	2011
			20	August	2011
			21	September	2011
	Select		22	October	2011
			23	November	2011
			24	December	2011
			25	January	2012
			26	February	2012
	Select		27	March	2012
			28	April	2012
			29	May	2012
			30	June	2012
			31	July	2012
	Select		32	August	2012

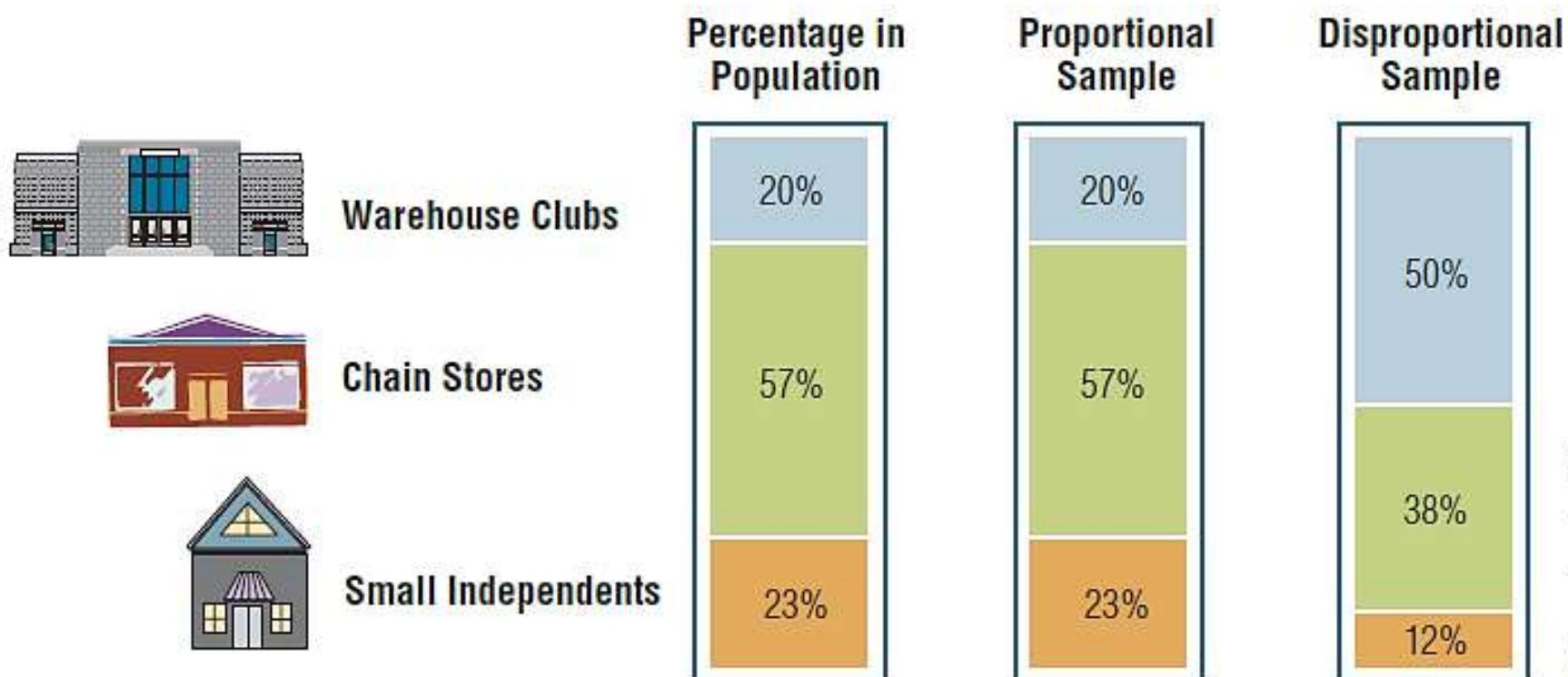
Stratified Sampling

- A subsample is drawn using simple random sampling within each stratum
 - Provides a more efficient sample than would be possible with simple random sampling
- Random sampling error will be reduced
 - Each group is internally homogeneous but there are comparative differences between groups

Proportional Versus Disproportional Sampling

- The number of sampling units drawn from each stratum is in proportion to the relative population size of the stratum
- In a disproportional stratified sample, the sample size for each stratum is not allocated in proportion to the population size but is dictated by analytical considerations
 - Ensures an adequate number of sampling units in every stratum

EXHIBIT 12.5 Disproportional Sampling: Hypothetical Example



Cluster Sampling

- The primary sampling unit is no longer the individual element in the population (e.g., grocery stores) but a larger cluster of elements located in proximity to one another (e.g., cities)
 - The area sample is the most popular type of cluster sample
- Cluster sampling is classified as a probability sampling technique because of either the random selection of clusters or the random selection of elements within each cluster

EXHIBIT 12.6 Examples of Clusters

Population Element	Possible Clusters in the United States
U.S. adult population	States Counties Metropolitan Statistical Areas Census Tracts Blocks Households
College seniors	Colleges
Manufacturing firms	Counties Metropolitan Statistical Areas Localities Plants
Airline travelers	Airports Planes
Sports fans	Football Stadiums Basketball Arenas Baseball Parks

Multistage Area Sampling

- Involves two or more steps that combine some of the probability techniques already described
 - Typically, geographic areas are randomly selected in progressively smaller (lower-population) units
- The Bureau of the Census provides maps, population information, demographic characteristics for population statistics, etc., by several small geographical areas
 - May be useful in sampling

What Is the Appropriate Sample Design?

- Degree of accuracy
 - Cost savings is a trade off for a reduction in accuracy
- Resources
 - If the researcher's financial and human resources are restricted, certain options will have to be eliminated
- Time
 - Time constraints restrict sampling techniques to simpler methods

What Is the Appropriate Sample Design? (cont'd.)

- Advance knowledge of the population
 - A lack of adequate lists may automatically rule out systematic sampling, stratified sampling, or other sampling designs
- National versus local project
 - Geographic proximity of population elements will influence sample design
 - When population elements are unequally distributed geographically, a cluster sample may become much more attractive