

Appendices

Appendix A: Glossary

Term	Definition
Census Division	Census divisions are groupings of states and the District of Columbia that are subdivisions of the four census regions. There are nine census divisions, which the U.S. Census Bureau established in 1910 for the presentation of census data.
Census Region	Census regions are grouping of states and the District of Columbia. The four regions (Northeast, Midwest, South and West) represent areas that were relatively homogeneous when they were established in 1910 and revised in 1950.
Computer Assisted Personal Interviewing (CAPI)	A method of data collection where the interviewer and respondent conduct an interview face-to-face. The survey questions are displayed on the screen of a computer. The interviewer reads the questions to the respondent and then enters the respondent's answers into a database on the computer.
Computer Assisted Telephone Interviewing (CATI)	A method of data collection where the interviewer and respondent conduct an interview over the telephone. The survey questions are displayed on the screen of a computer. The interviewer reads the questions to the respondent and then enters the respondent's answers into a database on the computer.
Consolidated Metropolitan Statistical Area (CMSA)	A metropolitan area is identified as a CMSA when it has a population of one million or more and also has separate component areas [Primary Metropolitan Statistical Areas (PMSAs)] that meet statistical criteria and are supported by local opinion.
Core Based Statistical Area (CBSA)	The term core based statistical area (CBSA) became effective in 2000 and refers collectively to metropolitan and micropolitan statistical areas. The 2000 standards provide that each CBSA must contain at least one urban area of with a population of 10,000 or more.
Cluster sampling	When the basic sampling unit in the population is to be found in groups or clusters (e.g., individual respondents in households), the sampling is sometimes carried out by first selecting a sample of clusters (e.g., households) and then collecting data from all or a sample of the members of each selected cluster.
Equal Probability Sample	An equal probability sample is derived from a random sampling method in which all units on the sampling frame have an equal probability of being selected.
Design Effect	The design effect is the ratio of the actual variance of a sample to the variance of a simple random sample with the same number of elements. The size of the design effect helps researchers understand the extent to which the data collection approach affects the results.

Term	Definition
Disproportionate Stratified Sample	A disproportionate stratified sample is a sampling method where particular units within a stratum have a higher probability of selection than other units. This method is often used when a sample design calls for an oversample of a particular group (i.e., AI/AN population).
Federal Information Procession Standards (FIPS) code	FIPS codes are a standardized set of numeric or alphabetic codes issued by the National Institute of Standards and Technology (NIST) to ensure uniform identification of geographic entities through all federal government agencies. The entities covered include states and statistically equivalent entities, counties and statistically equivalent entities, named populated and related location entities (such as, places and county subdivisions), and American Indian and Alaska Native areas.
Field Period	The field period denotes the start and end date of a data collection effort.
Metropolitan Statistical Area (MSA)	MSAs are metropolitan areas (MAs) that are not closely associated with other MAs. These areas typically are surrounded by nonmetropolitan counties (or county subdivisions in New England).
Mitofsky-Waksberg sample design	The Mitofsky-Waksberg sample design is a multi-stage sampling design method developed to maximize the efficiency of random digit dialing. Telephone numbers that are assigned to residential households tend to be grouped in banks of consecutive telephone numbers. The Mitofsky-Waksberg method groups telephone numbers into banks of 100 consecutive numbers using the area code, the 3-digit prefix, and the first 2 digits of the suffix. In the first stage of sampling, telephone numbers within the 100-banks are selected at random with replacement to determine if the number comes from a bank of residential numbers. If the telephone number is residential, the 100-bank is retained for the second stage of sampling.
Multi-stage sample	A multi-stage sample is derived from a sampling method that uses more than one stage of selection to draw the sample. The first stage of selection involves selecting the largest sampling unit [also called the primary sampling unit (PSU)], while the final stage of sampling involves selecting the smallest sample unit (for example, an individual or household).
Panel survey/study	A panel survey or study is a type of longitudinal study in which a group of subjects are surveyed on more than one occasion.
Primary Metropolitan Statistical Area (PMSA)	Each Primary Metropolitan Statistical Area (PMSA) consists of a large urbanized county or cluster of counties (or cities and towns in New England) that demonstrate very strong internal economic and social links, in addition to close ties to other portions of the larger area. If an area that qualifies as a metropolitan area (MA) has 1 million people or more, two or more PMSAs may be defined within it.
Primary Sampling Unit (PSU)	A primary sampling unit (PSU) is the first sample entity drawn in a multi-stage sample (in a single stage sample, the primary sampling unit is the first and only sampling unit drawn).

Term	Definition
Random Digit Dialing (RDD)	Random Digit Dialing (RDD) is an automated method of dialing telephone numbers at random in a targeted area. Unlisted telephone numbers are included in the sample frame.
Simple Random Sampling (SRS)	Simple Random Sampling (SRS) is a method of sampling that gives all units in the sample frame an equal chance of being selected for inclusion in the sample, and an equal chance for selection for each of all possible samples of the same size.
Standard error	The standard error is a measure of the variation that might be expected to occur merely by chance in the characteristics of samples drawn equally randomly from the same population. It is the standard deviation of the sampling distribution of a statistic. The smaller the standard error, the better the sample statistic is as an estimate of the population parameter.
Stratified Sample	In stratified sampling, an independent sample is selected from each strata of the population. Strata are divisions of the populations into relevant subsets (e.g., race, location).
Two-stage sampling	A two-stage sample is a simple case of a multi-stage sample. The population to be sampled is first classified into primary sampling units (PSUs) (e.g., census tracts) each of which consists of a number of the basic sampling units or the secondary sampling units (SSUs) (e.g., households within the census tract). A sample of these PSUs is taken, constituting the first stage, and then a sample of the SSUs is taken, constituting the second stage.
Wave	Wave is a term used to describe a single field period for a study that involves collecting data more than once (such as in a panel study).

Sources

Kish, L. *Survey Sampling*. 1965. John Wiley & Sons, Inc., New York, NY.

Pocket Dictionary of Statistics. 2002. The McGraw-Hill Companies, Inc., New York, NY.

U.S. Census Bureau online glossary: <http://www.census.gov/main/www/glossary.html>

OECD Glossary of Statistical Terms: <http://stats.oecd.org/glossary/index.htm>