Survey Design

Arevik Avedian
Harvard Law School
aavedian@law.harvard.edu
October 15, 2014

Overview

Basics of survey research

Measurement levels

Key definitions

Types of survey designs

Randomization and probability sampling methods

Collecting data using computerized survey design

Avoiding error and bias

Assessing reliability and validity

Resources

Learning Qualtrics

Purpose of surveys

- A **survey** is a systematic method for gathering information from (a sample of) entities for the purposes of constructing quantitative descriptors of the attributes of the larger population of which the entities are members.
- **Surveys** are conducted to gather information that reflects population's attitudes, behaviors, opinions and beliefs that cannot be observed directly.

- The success of survey research depends on how closely the answers that people give to survey questions match how people think and act in reality.

Surveys by type of study design

Design - Planning/implementing a study

- Sample survey or experiment?
- How to choose people (subjects) for the study, and how many?
- What questions to ask to find answers to our research questions?

- *Descriptive* – Graphical and numerical methods for summarizing (describing) the data.
  - Describe phenomena and summarize them.
    - Graphs, tables and numerical summaries are all examples of descriptive statistics

- *Inferential* – Making predictions based on the data.
  - Inferential Statistics uses methods for making predictions about a population (total set of subjects of interest), based on data from a sample (subset of the population on which study collects data).
    - Measure associations, e.g. income and quality of life.

Survey designs

- *Cross-sectional surveys:*
  - Data collected at one point in time selected to represent a larger population.

- *Longitudinal surveys:*
  - Trend:
    - Surveys of sample population at different time points.
  - Cohort:
    - Study of sample population each time data are collected but samples studied maybe different.
  - Panel:
    - Data collection at various time points with the same sample of respondents.

Questionnaire construction

Questionnaire - a document containing questions and other types of items designed to solicit information appropriate for analysis.

- The format of a questionnaire can influence the quality of data collected.
- A clear format for contingency questions is necessary to ensure that the respondents answer all the questions in the questionnaire.
- The order of items and wording in a questionnaire can influence the responses given.
- Clear instructions are important for getting appropriate responses in a questionnaire.
- Questionnaires should be pretested before being administered to the study sample.

Questions to think about before starting a survey

Before designing a survey a researcher should ask:

- Is this survey necessary?
- Is the purpose of the survey to evaluate people or programs?
- Can the data be obtained by other means?
- What level of detail is required?
- What type of survey is most appropriate and/or viable given funding and time (e.g. self-administered, telephone, face-to-face, or Internet)?
- Is the survey ethically possible?
- Is this a one-time survey or will the researcher repeat the survey in different settings and/or occasions (e.g. follow-up mailing in a self-administered questionnaire, follow-up calls in a telephone survey)?
- Will the researcher have a target completion rate?
- How will the researcher deal with non-respondents?

Guidelines for asking questions

- The form and meaning of questions should be appropriate to the project.
- The questions must be clear and precise.
  - Negative terms should be avoided
  - Double-barreled questions (multiple questions enclosed within one) should be avoided.
- Questions should be relevant to the respondent.
- Respondents must be competent and willing to answer the questions.
- The order and wording of questions should be set in a manner to avoid biased responses.

Types of surveys

- *Interview Surveys*
  - Interviewers must be neutral in appearance and actions; their presence in the data-collection process or personal opinion must have no effect on respondents' choice of answers.
  - Interviewers must be adequately trained to be familiar with the questionnaire, to follow the question wording and question order exactly, and to record responses exactly as they are given.
  - Interviewers can use probes to elicit an elaboration on an incomplete or ambiguous response. Probes should be neutral. Ideally, all interviewers should use the same probes.

- *Telephone Surveys*
  - Telephone surveys can be cheaper and more efficient than face-to-face interviews, and they can permit greater control over data collection.
  - Random-digit dialing (RDD) is a useful technique for eliminating potential bias in selecting numbers.

- *Online Surveys*
  - This method can be even cheaper than telephone and interview surveys, however must be used with caution because respondents may not be representative of the intended population.

Methods of computerized data collection

CAPI – computer-assisted personal interviewing, in which the computer displays the questions on screen, the interviewer reads them to the respondent and then enters the respondent's answer.

ACASI – audio computer-assisted self-interviewing, in which the respondent operates a computer, the computer displays the question on its screen and plays recordings of the questions to the respondent, who then enters his/her answers.

CATI – computer-assisted telephone interviewing, which is the telephone counterpart to CAPI.

IVR – interactive voice response, the telephone counterpart to ACASI, in which the computer plays recordings of the questions to respondents over the telephone who then respond by using the keypad of the telephone or saying their answers aloud.

Web – internet surveys (e.g. Qualtrics), in which a computer administers the questions online.

Privacy and ethics in survey research

Researchers should:

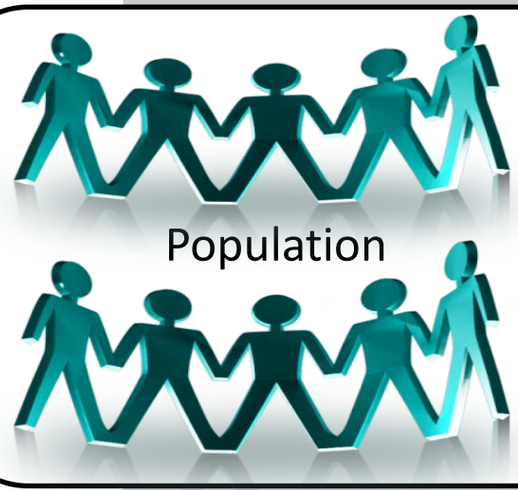
- keep confidential private information about survey participants,
- minimize the possibility of causing psychological discomfort or harm to respondents,
- when possible use paper-based, self-administered questionnaires (SAQ), instead of face-to-face survey to elicit information of a sensitive nature.

Key definitions

- *Population:* Is an entire collection of people, firms, states or things, that we are interested in, which we wish to describe, explain or predict. Population distribution is usually unknown; we make inferences about its characteristics such as the parameter.

- *Sample:* A sample that is representative of the population that we actually observe and is used to infer about the population. Sample value we find from surveys is called statistic.

Sampling and inference



Population

Random Sample

1. Find sample statistic $\bar{y}$
2. Infer population parameter μ

Randomization and probability sampling methods

Randomization – the mechanism for achieving reliable data by reducing potential bias.

Simple random sample – in a sample survey, each possible sample of size *n* has the same probability of being selected.

Systematic Random Sample – (1) selects a subject at random from the first k names in the sampling frame, and (2) selects every k^{th} subject listed after that one. The number k is called the skip number.

- Population size is N, sample size is n, k = N/n.

Stratified Random Sample – divides the population into separate groups, called strata, and then selects a simple random sample from each stratum.

- Can be proportional (proportionate to population parameters) or disproportional.

Cluster random sampling – divides the population into a large number of clusters, such as city blocks. Selects a simple random sample of the clusters. Uses all the subjects in those clusters as the sample.

Multistage Sampling – uses combination of sampling methods.

Sampling error

The *sampling error* of a statistic equals the error that occurs when we use a sample statistic to predict the value of a population parameter.

- Ex. Random survey is estimating percentage of US population favoring Obama before the elections. If the survey estimated 54 and the actual rating was 59%, the sampling error would equal 5%.

Note: In practice the true sampling error is unknown, because the population parameters are unknown.

Randomization protects against bias; direction and extent of bias is unknown for studies that cannot employ randomization.

- Major polling organizations predict outcomes with ±3% accuracy (margin of error) when n is about 1000.

Sampling variability and possible bias

Other factors besides sampling error can cause results to vary from sample to sample:

- ***Sampling bias*** (nonprobability sampling, undercoverage)
  - Volunteer sampling

- ***Response bias*** (e.g., poorly worded questions, order of questions, approval of the interviewer)

During the Cold War a study asked: "Do you think the US should let Russian newspaper reporters come here and send back what ever they want?" and "Do you think Russia should let American newspaper reporters come in and send back whatever they want?" The percentage of yes responses to the first question increased from 36% to 73% when asked second ("*Tainted Truth: The Manipulation of Fact in America,"* Crossen, 1994).

- ***Nonresponse bias*** (missing data, respondents can't be reached or refuse to participate)
  - Results of any sample with over 20% nonresponse rate should be questionable.

Results of surveys may depend greatly on wording

ex. 2006 *New York Times* poll:

"Do you favor a gasoline tax?"
12% yes

"Do you favor a gasoline tax
- to reduce U.S. dependence on foreign oil?" 55% yes
- to reduce global warming?" 59% yes

66. Would you favor or oppose an increased federal tax on gasoline?

	Favor	Oppose	DK/NA
2/22-26/06	12	85	3

67. What if your payroll taxes or income taxes were reduced as a result of the increased gasoline tax, then would you favor or oppose an increased federal tax on gasoline?

	Favor	Oppose	DK/NA
2/22-26/06	28	63	10

68. What if the increased tax on gasoline would reduce the United States' dependence on foreign oil, then would you favor or oppose an increased federal tax on gasoline?

	Favor	Oppose	DK/NA
2/22-26/06	55	37	8

69. What if the increased tax on gasoline would cut down on energy consumption and reduce global warming, then would you favor or oppose an increased federal tax on gasoline?

	Favor	Oppose	DK/NA
2/22-26/06	59	34	7

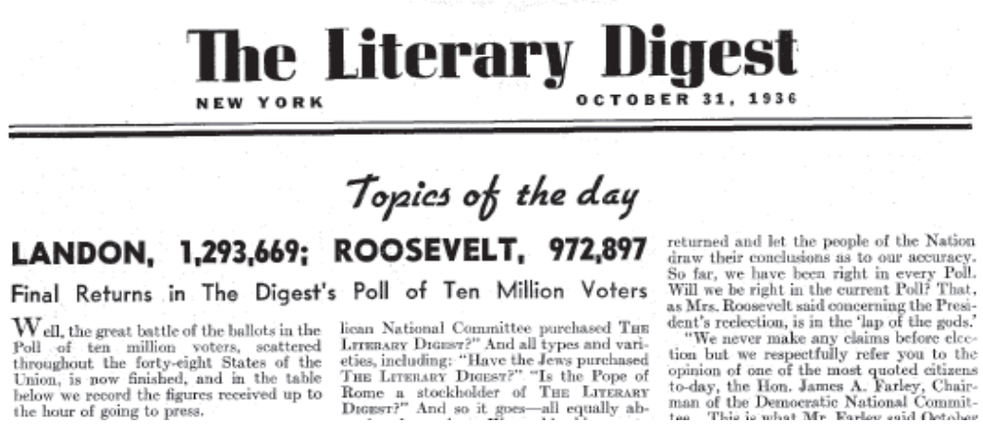
http://www.nytimes.com/packages/pdf/national/20060228_poll_results.pdf

The case of the 1936 Literary Digest Poll

- The presidential election of 1936 pitted Alfred Landon, the Republican governor of Kansas, against the incumbent President, Franklin D. Roosevelt.
  - The year 1936 marked the end of the Great Depression, and economic issues such as unemployment and government spending were the dominant themes of the campaign.

- The Literary Digest magazine mailed a questionnaire to 10 million people (2.3 million replied) just before the presidential election.
  - It was based on every telephone directory in the United States, lists of magazine subscribers, rosters of clubs and associations, etc.

The prediction was that Landon would get 57% of the vote against Roosevelt's 43% (these are the *statistics* that the poll measured).



The Literary Digest

NEW YORK — OCTOBER 31, 1936

Topics of the day

LANDON, 1,293,669; ROOSEVELT, 972,897

Final Returns in The Digest's Poll of Ten Million Voters

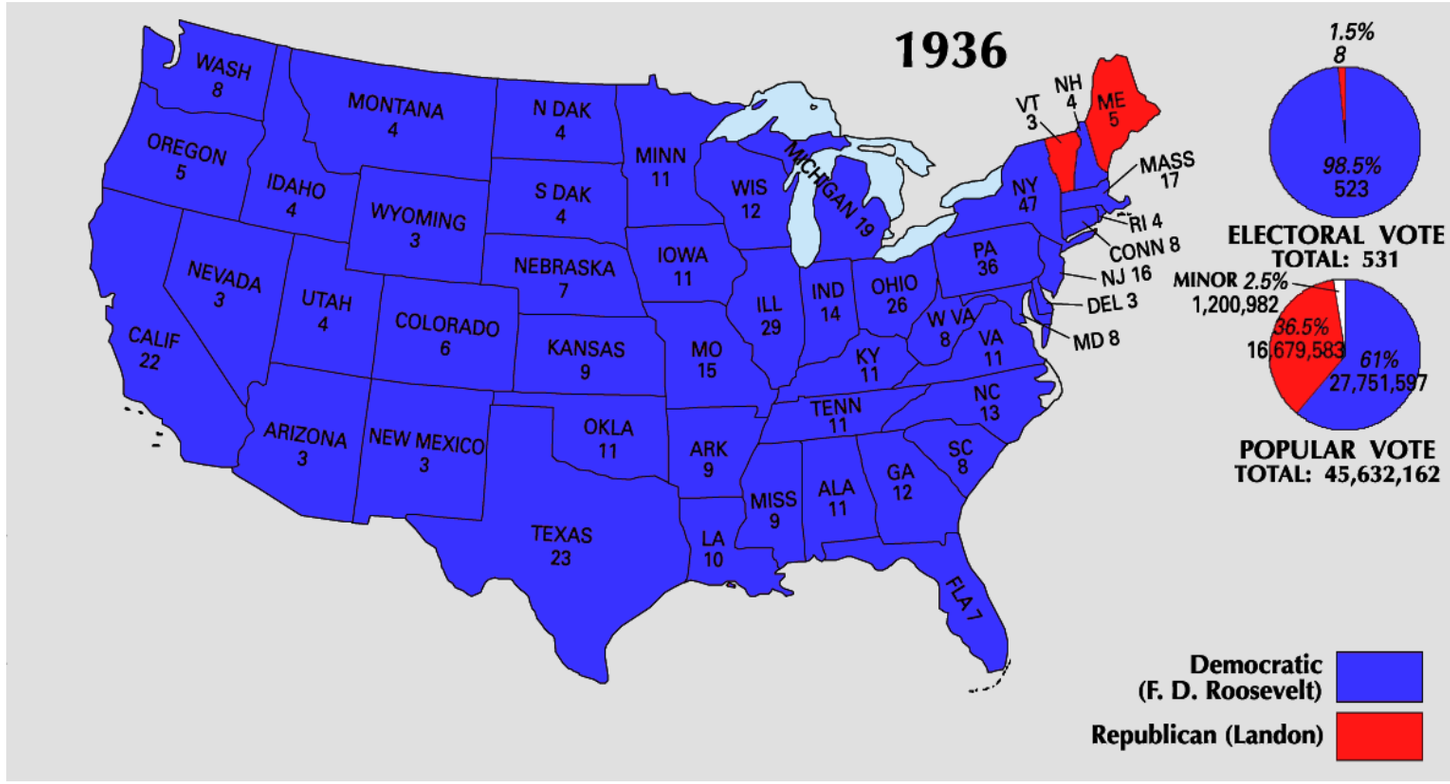
Well, the great battle of the ballots in the Poll of ten million voters, scattered throughout the forty-eight States of the Union, is now finished, and in the table below we record the figures received up to the hour of going to press.

lican National Committee purchased THE LITERARY DIGEST?" And all types and varieties, including: "Have the Jews purchased THE LITERARY DIGEST?" "Is the Pope of Rome a stockholder of THE LITERARY DIGEST?" And so it goes—all equally ab-

returned and let the people of the Nation draw their conclusions as to our accuracy. So far, we have been right in every Poll. Will we be right in the current Poll? That, as Mrs. Roosevelt said concerning the President's reelection, is in the 'lap of the gods.'

"We never make any claims before election but we respectfully refer you to the opinion of one of the most quoted citizens to-day, the Hon. James A. Farley, Chairman of the Democratic National Committee. This is what Mr. Farley said October

Who won?



The actual results of the election were 61% for Roosevelt against 37% for Landon (these were the *parameters* the poll was trying to measure).

- Gallup American Institute of Public Opinion achieved national recognition by correctly predicting the result of the election within about 1%, using a much smaller sample size of 50,000.

What went wrong with the polls?

There were two basic causes of the *Literary Digest's* downfall: *selection bias* and *nonresponse bias*.

The first major problem with the poll was in the nonrandom selection process for the names on the mailing list, which were taken from telephone directories, club membership lists, lists of magazine subscribers, etc.

- Such a list was guaranteed to be slanted toward middle- and upper-class voters in 1936, and by default to exclude lower-income voters.

The second problem was that out of the 10 million people whose names were on the original mailing list, only about 2.4 million responded to the survey (that's a 76% nonresponse rate!).

- People who respond to surveys are different from people who don't, not only in the obvious way (their attitude toward surveys) but also in more subtle and significant ways.

Validity and reliability of survey measures

Reliability: Whether the measure will produce a similar value when the measuring instrument is reapplied.

External validity: Whether (causal) relationships can be generalized to different measures, persons, settings, and times.

Internal Validity: Whether the effects observed in a study are due to the independent variable of interest and not some other "confounding" factor.

Types of survey question formats

- Open-ended question: Questions for which the respondent is asked to provide his or her own answers. In-depth, qualitative interviewing relies almost exclusively on open ended questions.
  - Disadvantage of open-ended questions is more complex data analysis.

- Closed-ended question: Survey questions in which the respondent is asked to select an answer from among a list provided by the researcher.
  - Popular in survey research because they provide a greater uniformity of responses and are more easily processed than open-ended questions.

Open-ended questions

Open-ended question formats provide a blank space or box where respondents type or write in their response using their own words (or numbers).

Pros:

- Some information that researchers are seeking may be impossible to obtain with closed-ended questions and are revealed through open-ended format questions.

Cons:

- Coding the answers for statistical analysis.

Examples of open-ended questions:

Have you ever made an error in judgment that you had to address with your employer? How did you handle it?

When you are under a lot of stress, what is your typical reaction?

Closed-ended questions

Closed-ended question formats or scalar questions provide respondents with a list of answer choices from which they must choose to answer the question.

Types of closed-ended answers:

- *Interval scale* questions provide answers that possess the properties of order and constant units of distance.
  - E.g. age, income
- *Ordinal scale* questions provide answers with ordered categories (difference between the categories is not the same).
  - E.g. intensity of opinion or pain, frequency of events or behaviors
- *Nominal scale* questions provide answers with categories that are unranked and unordered. Nominal scale does not possess order, distance, or origin.
  - E.g. Ethnicity, color of hair, select all that apply

Interval:

Here is a scale of incomes. We would like to know in what group your household is, counting all wages, pensions and other incomes that come in.

- ❑ Up to 20,000
- ❑ 20 001 to 40,000
- ❑ 40,001 to 60,000
- ❑ 60,001 to 80,000
- ❑ 80,001 to 100,000
- ❑ 100,001 or more

Ordinal:

How likely are you to vote for Obama in the November 2012 Presidential election?

- ❑ Not likely at all
- ❑ Somewhat likely
- ❑ Very likely

Nominal:

For which major candidate do you plan to vote in the November 2012 Presidential election?

- ❑ Obama
- ❑ Romney
- ❑ Other
- ❑ I will not vote

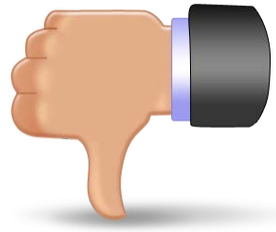
Mixed-method (combines open & closed formats)

EMPLOYEE BENEFITS SURVEY

Statement	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Health Benefits					
I am satisfied with my health plan options.					
I am satisfied with my dental plan options.					
I am satisfied with my vision plan options.					
I am satisfied with my long-term disability insurance.					
I am satisfied with my short-term disability insurance.					
I am satisfied with my options for life insurance.					
Overall, I am satisfied with my health benefits.					
Financial Benefits					
I am satisfied with my retirement plan.					
I am satisfied with my salary.					
I am satisfied with the Employee Stock Purchase Program.					
I am satisfied with my opportunities for promotion, raises, and bonuses.					
Overall, I am satisfied with my financial benefits.					
Paid Time Off					
I am satisfied with the number of vacation, sick, and personal days that I receive.					
Overall, I am satisfied with my paid time off.					
Overall					
I understand my benefit options.					
I know where to find information about my benefits.					
I know whom to call if I have questions about my benefits.					
Overall, I am satisfied with my employee benefits.					
Additional Comments:					

Make questions clear

Questionnaire items should be precise so that the respondent knows exactly what the researcher is asking.

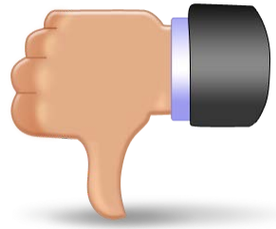


- What was your income last week?
  - Respondent may only consider weekdays, while researcher means full week.
- Are you employed full time?
  - Respondent may not know exactly what is considered full time for the researcher.

- What was your income for the entire period of 10/01/14 to 10/08/14?
- Are you employed at least 35 hours a week?

Avoid double-barrel questions

Avoid asking for a single answer to a question that actually has multiple parts. As a general rule, whenever the word *and* appears in a question check whether you are asking a double-barreled question.

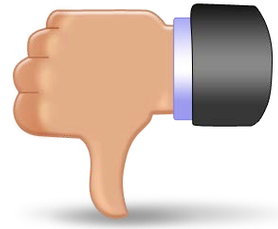


- The United States should withdraw from Afghanistan and spend the money on domestic programs.
  - While many respondents would unequivocally agree with the statement and others would unequivocally disagree, some may be unable to answer, as they agree only with one part of the question.

- In your opinion, should the US withdraw from Afghanistan?
- Would you like the US to spend more money on domestic programs?

Short questions are best

Respondent should be able to read a question quickly, understand its intent and select or provide an answer without difficulty. In general, assume that respondents will read questions quickly and give quick answers.



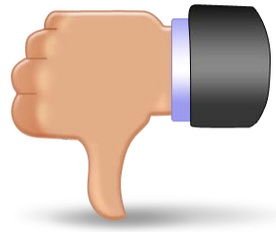
- Example from a survey conducted by Harris Poll in 1986:
  - "If Libya now increases its terrorist acts against the US and we keep inflicting more damage on Libya, then inevitably it will all end in the US going to war and finally invading that country which would be wrong."

Respondents were given the opportunity of answering "Agree," "Disagree," or "Not sure" with the following statements:

- 1. Will Libya increase its terrorist acts against the U.S.?
- 2. Will the U.S. inflict more damage on Libya?
- 3. Will the U.S. inevitably or otherwise go to war against Libya?
- 4. Would the U.S. invade Libya?
- 5. Would that be right or wrong?

Avoid negative and double negative questions

Negation in a question paves the way for easy misinterpretation. Negative words to avoid are “not”, “prohibit”, “impossible,” etc. Double-negative questions often make it unclear for respondents whether to put a “yes” or “no.”



- PhD students should not be required to take qualifying exams to graduate.
- An actual example of double negative:
  - Would you favor or oppose a bill that would prevent any foreign-owned company from owning cargo operations at seaports in the United States? (Gallup, March 10-11, 2006)
    - 58% opposed

- Should PhD students be required to take qualifying exams to graduate?
- After correcting the double negative:
  - Would you favor or oppose a bill that would allow only U.S. companies to own cargo operations at seaports in the United States? (Gallup, March 13-16)
    - 25% opposed

Target the vocabulary of the population to be surveyed

- For studies within a specific organization, use the jargon used in that organization.
- Be careful not to use language that may not be familiar to the respondents.
- Avoid unnecessary abbreviations.
- Use simple words.
- Avoid biased items and terms.

  – Rasinski (1989) analyzed several General Social Survey (GSS) studies and found that the way programs were identified had an impact on the amount of public support they received: Here are some of the comparisons:

More support	Less support
“Assistance to the poor”	“Welfare”
“Dealing with drug addiction”	“Drug rehabilitation”
“Improving conditions of blacks”	“Assistance to blacks”
“Protecting social security”	“Social security”

Describing surveys - Example 1: SOC

Survey Name	**Survey of Consumers (SOC)**
Sponsor	University of Michigan
Collector	Survey Research Center, University of Michigan
Purpose	Main objectives are to: • Measure changes in consumer attitudes and expectations • Understand why such changes occur • Evaluate how they relate to consumer decisions to save, borrow, or make discretionary changes
Year Started	1946
Target Population	Noninstitutionalized adults in the coterminous United States (excludes Hawaii and Alaska)
Sampling Frame	Coterminous US telephone households, through lists of working area codes and exchanges
Sampling Design	List-assisted random-digit dial sample, randomly selected adults
Sample Size	500 adults
Use of Interviewer	Interviewer administered
Mode of Administration	Telephone Interview
Computer Assistance	Computer-assisted telephone interviewing (CATI)
Reporting Unit	Randomly selected adult
Time Dimension	Two-wave panel of persons
Frequency	Conducted monthly
Interviews per Round of Survey	Two: reiniterview conducted six months after initial interview subset of wave 1 respondents
Levels of Observation	Person
Web Link	http://sca.isr.umich.edu

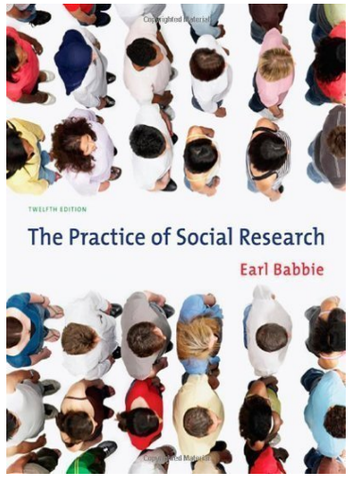
Source: Survey Methodology, Groves et al.

Describing surveys - Example 2: NSDUH

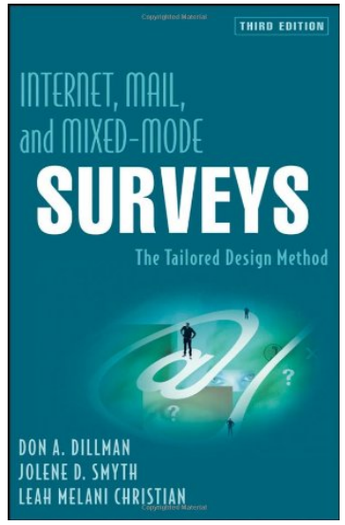
Survey Name	**National Survey of Drug Use and Health (NSDUH)**
Sponsor	Substance Abuse and Mental Health Services Administration (SAMHSA)
Collector	RTI International
Purpose	Main objectives are to: • Provide estimates of rates of use, number of users, and other measures related to illicit drug, alcohol, and tobacco use at the state and national level • Improve the nation's understanding of substance abuse • Measure the nation's progress in reducing substance abuse
Year Started	1971 (formerly names National Household Survey on Drug Abuse)
Target Population	Noninstitutionalized population of the United States aged 12 years or older
Sampling Frame	U.S. households, enumerated through U.S. counties, blocks and list of members of the households
Sampling Design	Multistage, stratified clustered area probability sample within each state.
Sample Size	141,487 housing units; 67870 persons (2007 NSDUH)
Use of Interviewer	Interviewer administered, with some self-administered questionnaire sections for sensitive questions
Mode of Administration	Face-to-face interview in respondent's home, with portions completed by respondent alone
Computer Assistance	Computer-assisted telephone interview (CATI), with audio computer-assisted self-interview by respondent alone
Reporting Unit	Each person age 12 or older in household reports for self. Respondents may allow more knowledgeable family member to complete Health Insurance and Income sections of survey for them.
Time Dimension	Repeated cross-sectional survey
Frequency	Conducted annually
Interviews per Round of Survey	One
Levels of Observation	Person, household
Web Link	http://www.samsha.gov

Source: Survey Methodology, Groves et al.

Recommended readings

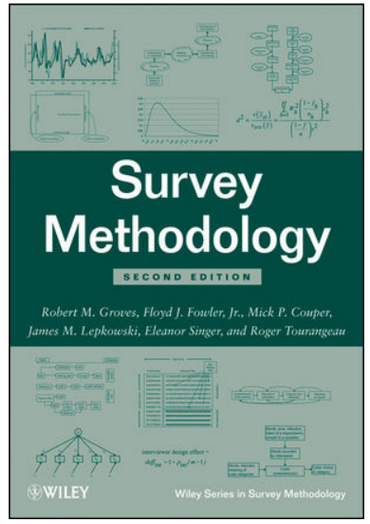


Babbie, E. R. (2009) *The Practice of Social Research* (12th Edition). Belmont, CA: Wadsworth Publishing. ISBN-13: 9780495598411



Dillman, D.A., Smyth, J.D., & Christian, L.M. (2008) *Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method* (3rd Edition). Hoboken, N.J.: Wiley & Sons.

ISBN-13: 978-0-471-69868-5



Groves, R.M., Fowler, F.J. Jr., Couper, M.P., Lepkowski, J.M., Singer, E., Tourangeau, R. (2009) Survey Methodology (2nd Edition). Hoboken, N.J.: Wiley & Sons

ISBN-13: 978-0-470-46546-2

Working with Qualtrics

harvard.qualtrics.com

Support
For resources, including training and documentation, please visit the Qualtrics webpage at
www.qualtrics.com/university/researchsuite/

Learn Qualtrics in 5 Steps at: Five step training program