ENCYCLOPEDIA OF

MEASUREMENT

AND

STATISTICS

VOLUME **3**

EDITED BY
NEIL J. SALKIND

ENCYCLOPEDIA OF
MEASUREMENT
AND
STATISTICS

VOLUME 3

EDITOR
NEIL J. SALKIND
UNIVERSITY OF KANSAS

MANAGING EDITOR
KRISTIN RASMUSSEN
UNIVERSITY OF KANSAS

Contributors

Francisco J. Abad
Universidad Autonoma de Madrid

Inmaculada Aban
University of Alabama at Birmingham

Hervé Abdi
University of Texas at Dallas

Phillip L. Ackerman
Georgia Institute of Technology

Demetrios S. Alexopoulos
University of Patras, Greece

Audrey Amrein-Beardsley
Arizona State University

Lauren E. Auld
DePauw University

Carrie R. Ball
University of Wisconsin–Madison

Kimberly A. Barchard
University of Nevada, Las Vegas

Jonathan Barzilai
Dalhousie University

Edward J. Bedrick
University of New Mexico

Mark L. Berenson
Montclair State University

Dongsheng Bi
University of Nebraska Lincoln

Damian P. Birney
University of Sydney

David M. Boynton
Saint Michael's College

Bruce A. Bracken
College of William & Mary

Jennifer Bragger
Montclair State University

Gary G. Brannigan
State University of New York–Plattsburgh

Ernest W. Brewer
University of Tennessee

Carolyn Brodbeck
Chapman University

Sarah Brookhart
American Psychological Society

Duane Brown
University of North Carolina, Chapel Hill

Jennifer Ann Brown
University of Canterbury

Shawn T. Bubany
University of Minnesota

A scree plot is a two-dimensional plot with the number of factors on the horizontal axis and eigenvalues on the vertical axis. Its purpose is to provide a useful visual aid for determining an appropriate number of components to retain. The plot shows the fraction of total variance in the data as explained, or represented, by each factor or component. The plot is called a scree plot because it looks like a sloping mass of loose rocks (scree) at the base of a cliff. The scree plot allows one to pick the number of components or factors on the basis of the point at which "elbow," or separation, is observed or where a plateau begins. The plot in Figure 1 shows eigenvalues versus number of factors with annotation of the first five values of the eigenvalues and their corresponding cumulative percent of variance. The eigenvalues are based on the results from Proc Factor in SAS, and the plot is created by Minitab software.

In this plot, an elbow occurs at the third point and is followed by a plateau. That is, the eigenvalues after the third value are all relatively small. Thus, we see that three principal components should be retained for further use. There is a clear separation between the first two components and the remaining components.

—*Kyoungah See*

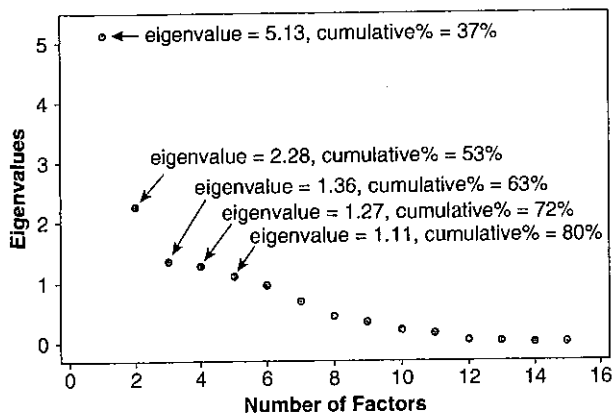*See also* Eigenvalues; Factor Analysis; Factor Scores

5 - o ← eigenvalue = 5.13, cumulative% = 37%

eigenvalue = 2.28, cumulative% = 53%
eigenvalue = 1.36, cumulative% = 63%
eigenvalue = 1.27, cumulative% = 72%
eigenvalue = 1.11, cumulative% = 80%

Eigenvalues (y-axis)
Number of Factors (x-axis)

Figure 1   Scree for Nursing Home Quality Data

Further Reading

Johnson, R. A., & Wichern, D. (2002). *Applied multivariate statistical analysis* (5th ed.). Englewood Cliffs, NJ: Prentice Hall.

Mardia, K. V., Kent, J. T., & Bibby, J. M. (1979). *Multivariate analysis.* New York: Academic Press.

Exploratory factor analysis primer: http://www.upa.pdx.edu/IOA/newsom/semclass/ho_efa.doc

Scree plots: http://www.rsd-associates.com/mtxscree.htm

SECONDARY DATA ANALYSIS

If locating a sample on which to conduct research poses problems, it might be worthwhile to consider conducting a secondary data analysis. During the past 25 years, an increasing number of researchers in the United States have turned to secondary data rather than collecting original data for their research. More recently, some researchers have begun to explore uses of secondary data for qualitative research.

At its most basic level, secondary data analysis involves using data collected in previous research to address a different research question. Special considerations in any secondary analysis include differences in operational definitions, unavailability of raw data, and various ethical issues. This entry explores these issues as well as the uses, advantages, and disadvantages of secondary data analysis. To demonstrate how secondary data analysis has been used to address various topics, summaries of some secondary analysis studies have been included at the end of this entry.

Similarities and Distinctions Between Secondary Data Analysis and Meta-Analysis

Graves considered meta-analysis to be a special case of secondary analysis. However, because of differences in the two methods, some researchers do not agree with Graves. A comparison of these two methods reveals both similarities and distinctions. Regarding similarities, both methodologies use data collected by others. Both also offer the temptation to

rely on published findings, which means that each method can be subject to publication bias.

However, at least two distinctions exist between secondary analysis and meta-analysis. First, secondary data analysis uses previously collected data to address problems other than those for which the data were collected. In contrast, meta-analysis combines results from multiple studies with similar hypotheses to clarify findings on the same problem. Second, whereas meta-analysis entails using data from multiple studies, secondary data analysis might use data from a single (usually large) study. Because of these differences between the two methods, some researchers have viewed secondary analysis as distinct from all other types of research.

Practical Basis for Secondary Data Analysis

Governments, businesses, schools, and other organizations have gathered much quantitative information. Therefore, when available information gathered for primary research can serve other research purposes, it has made sense to use that information, rather than repeatedly attempting massive data collection efforts. Researchers have highlighted at least four practical reasons for using secondary data.

Efficiency

Secondary data has been less costly and frequently more easily available than primary data. This has been confirmed by a variety of researchers. Other secondary data analysis scholars concurred that, with research funding shortages, secondary data banks have provided a cost-efficient source of data.

Strengthened Confidence

Both Brannigan and Cowton noted that similar trends found by different sources strengthen confidence in findings. This phenomenon, known as triangulation, often can be achieved efficiently through multiple secondary sources. For triangulation purposes, K. J. Kiecolt and L. E. Nathan also recommended combining original research with analysis of secondary data.

Special Need for Some Studies

The textual data found in newspaper accounts, letters, and other qualitative sources has been essential for historical research. Katz (in "How satisfied are the self-employed: A secondary analysis approach"; see the last research example in this entry) made the point that secondary data might be less biased than some types of primary data, such as self-report interviews. The reason proposed for reduced bias was that self-report is more subjective than third-party reporting.

Another example of using qualitative data to meet special research needs is found in Lindenmann's award-winning public health study. Since public health information often is presented to general audiences in qualitative form, Lindenmann chose newspaper articles as the venue for his study. To learn how newspapers reported on matters affecting public health, he studied two national and five regional newspapers' articles that related to public health issues.

Protection of Research Pool

Finally, secondary data analysis might have particular application for graduate students. Fitzpatrick claimed that by using the analysis from secondary information while learning to do research, potential research participants are protected.

Research Uses of Secondary Data Analysis

Quantitative researchers have found broad applications for secondary analysis because of the availability of summative data in national data banks. These data can help in addressing a host of educational, organizational, medical, environmental, and social problems. Also, as noted elsewhere in this article, graphic and textual secondary data have become widely available to address qualitative research questions. Qualitative data banks include those found in periodicals, newspapers, and other print media.

Quantitative Uses

Trochim and Van Dierendonck, Schaufeli, and Buunk noted that multiple secondary sources were used for some research questions. For example, Trochim noted that a researcher could study crime patterns in different parts of the United States using both census and crime data; likewise, a study reported by Van Dierendonck and colleagues noted use of five sources of secondary data for their examination of causal relationships among the three dimensions of burnout. Similarly, census data could be combined with discipline-specific databases to examine relevant issues in other fields. Some recent examples of uses of quantitative secondary analysis are studies of job satisfaction, work environment, gender, and worker decision making.

Qualitative Uses

Historical studies, obviously, have depended on secondary sources. Common sources for secondary data have included ethnographies, legislation, treaties, annual reports of various agencies, and letters and diaries. These sources have been used to study such topics as trends in social behavior; effects of substance abuse; health care; business; and work, school, and family issues. Also, mixed-methods studies have used both qualitative and quantitative as well as primary and secondary data.

Secondary Data Analysis Process

The process for secondary data analysis, as noted in Figure 1, is similar to that for other types of research except for the methods of gathering and working with the data. For original research, researchers (a) formulate research questions, (b) conduct literature reviews, (c) develop research designs, (d) collect data, (e) analyze data, and (f) report findings. For secondary analyses, they (a) formulate research questions, (b) conduct literature reviews, (c) establish criteria for inclusion, (d) locate summarized data collected by others *or* use their own primary data for a different research purpose, (e) analyze data, and (f) report findings.
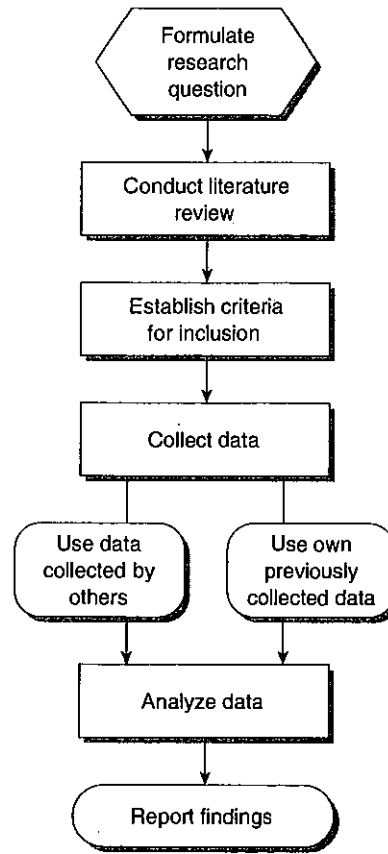

Figure 1   Steps in Quantitative Secondary Data Analysis

In the data gathering and recording stages of secondary data analysis, researchers establish size of study, research design, quality controls, and so forth, for the data to be included. Next, they search data archives for data meeting their criteria. These data, frequently subsets of larger data sets, are transported into SPSS or another statistical program. Variables are established, often by combining other variables or by constructing new variables. Once the secondary data have been analyzed, secondary researchers compare their findings with findings of primary researchers and report their findings.

Variations of the Basic Secondary Data Analysis Process

There have been some variations, or modifications, of the basic secondary data analysis process. These process changes have arisen from the unique needs of

research projects. A few of these modifications are discussed in this section, including (a) use of secondary and primary data in the same study, (b) secondary data analysis using one's own data, and (c) studies that include both qualitative and quantitative secondary data.

Using secondary data analysis to complement primary data studies. Heaton ("Children and Families Team") used a combination of primary and secondary data. This approach can strengthen confidence in the findings of secondary data analysis while adding depth to answers to research questions.

Secondary data analysis of researcher's own data. Researchers can also reuse their own data for a new research purpose. Using this variation, researchers could avoid some of the challenges ordinarily involved in secondary data analysis because they would understand the primary research fully and would have access to the complete data set.

Studies using both qualitative and quantitative secondary data. Finally, Janet Heaton, in "Secondary analysis of qualitative data," discussed the option of including both qualitative and quantitative secondary data in studies. For example, a researcher might use secondary data analysis to identify aspects of job satisfaction in middle-management positions in the accounting industry. The researcher might then conduct interviews with several middle managers in selected accounting firms to add another dimension to the study. This variation can reinforce and clarify the findings of primary studies.

Major National Archives

Several large databases have been developed and made available to researchers. These may be classified according to research disciplines or by agency; some examples are included in this discussion. However, Mertens advised that some national databases have limitations relating to ways they can be manipulated. For example, a researcher might find it impossible to disaggregate data according to demographic variables such as gender or race. Also, some demographics that interested a researcher—such as type of disability, educational level, and so forth—might not even be included in the study.

Nongovernment Databases

ISI Web of Knowledge

Developed by ISI (http://www.isinet.com/), this site was introduced in 2001. ISI, which is owned by the Thomson Corporation, has provided resources to researchers since the 1950s. It has become a worldwide organization with offices in the United States, Ireland, Japan, and China. The site contains qualitative and quantitative data banks on a variety of scientific, business, and social topics, and information is available in several languages.

Shuji Kaneko, an employee of Thomson, believes that a major advantage of data on the Web is its direct connection to related information through the link function. ISI, he said, can create links from secondary databases to primary sources. Kaneko also believes that ISI used adequate critical selection criteria to provide reliable and useful information. The criteria he listed were authority, accuracy, currency, navigation and design, applicability and content, scope, audience level, and quality of writing.

General Social Survey

The General Social Survey was one of three databases for social science research recommended by researchers. The General Social Survey is updated every year or so through major national surveys conducted through the University of Chicago. The purpose of the surveys has been to make fresh data available to researchers. Data has included international information as well as that collected in the United States.

Poll Service Databases

The Roper Center for Public Opinion Research, with facilities located at the University of Connecticut, has conducted public opinion polls on a variety of socially relevant issues, including information from the United States as well as 70 other nations. Other well-known opinion poll services include the Gallup

Poll and national news service polls. While sharing agreements for these data vary, interested researchers could find helpful information through these and other poll services.

Commercial Databases

In "Secondary analysis in entrepreneurship: An introduction to databases and data management," Katz recommended marketing databases for research in appropriate areas. Some of those included Dun and Bradstreet, TriNet, and BRS. Katz also recommended that researchers consider the Interuniversity Consortium for Political and Social Research (http://www.icpsr.umich.edu/) Guide to Resources and Services for data-sharing opportunities.

Hard Copy Data Sources

According to Lawrence Neuman, excellent hard copy sources for social and political research include the *Almanac of American Politics, America Votes: A Handbook of Contemporary American Election Statistics,* and *Vital Statistics on American Politics.* These sources are available to researchers through their national congressional representatives. The Almanac of American Politics publishes congressional voting records, America Votes has county voting information for most statewide and national offices, and Vital Statistics provides information on campaign spending and a variety of related issues. Information gathered from these databases could be used to amend current practices or to develop policy for new initiatives.

Government Databases

Unlike some of the databases mentioned earlier, government databases can be accessed without charge. Another advantage of government data has been that it may be less biased than other data sources. A description of some government databases follows.

The Data Archive

This is a United Kingdom site. It is a major storage for quantitative computer files and their documentation in the United Kingdom, according to the University of Southampton. The Data Archive also contains qualitative material—primarily research reports in the broad areas of humanities and social sciences (i.e., education, labor, history, politics, household finance, child development, crime, and public health).

FedWorld

This site (http://www.fedworld.gov/) promises to make it easy to locate government information. It has links to countless resources for secondary data analysis. These resources include the census and education databases, discussed next.

Census Databases

The U.S. Bureau of the Census database could be useful to researchers in a variety of disciplines, including business, social science, politics, health, and education. The national census conducted at the beginning of each decade gathers information on population numbers, ages, educational and income levels, housing, employment, and other areas. Inferential statistics can be derived from the vast store of current and past information from the censuses made available through the U.S. Departments of Commerce, Education, and Labor as well as directly through the U.S. Bureau of the Census.

Survey of Income and Program Participation

The Survey of Income and Program Participation database is available through a collaborative effort between the U.S. Census Bureau and the U.S. Social Services Administration. This database contains frequently updated information on income, employment, and participation in various government financial assistance programs.

Education Databases

The National Center for Education Statistics has developed a special database that has been of great assistance to persons preparing grant proposals for educational programs. This database can be found through the U.S. Department of Education Web site

(at http://www.ed.gov/index.jhtml). It provides information on demographics according to state, age, level of education, and so forth.

Another valuable resource for education research is the database for the Higher Education Research Institute (http://www.gseis.ucla.edu/heri/heri.html). It has information available at a minimal charge.

Advantages of Secondary Data Analysis

Various authors, such as Katz (in "Secondary analysis in entrepreneurship: An introduction to databases and data management") and Trochim, have discussed advantages of secondary data analysis. This section contains a discussion of some of the most notable reasons for those advantages. Those reasons are efficiency, conservation of participant pools, and the advantage of using unobtrusive measures to obtain the data.

Efficiency

Some authors, including Jean Barclay, A. Dale and colleagues, E. J. Graves, R. Katz (in "Secondary analysis in entrepreneurship: An introduction to databases and data management"), and Kiecolt and Nathan, have stated that conservation of time, effort, and money is the strongest reason for secondary analysis. They pointed out that, especially for studies involving very large populations, individual researchers did not have the necessary money or access to obtain data firsthand. For example, Katz recommended secondary analysis for business studies involving entrepreneurial issues.

Conservation of Participant Pools

J. Fitzpatrick noted that large numbers of graduate students could "wear down" significant pools of potential research participants. Wimmer and Dominick noted that although novice researchers, such as students, can gain benefits from developing and conducting research, this particular type of analysis doesn't usually produce results that are externally valid.

Unobtrusive Measures

Several researchers noted that, since secondary data analysis involves data collection in an indirect manner, they could avoid some contamination of the data. This observation might apply especially to quantitative data. However, it also could invite some of the challenges of secondary analysis because the burden of selecting high-quality studies would become increasingly important.

Challenges of Secondary Data Analysis

In addition to being a very useful tool, secondary analysis presents challenges to the researcher.

Getting a good fit. Accessing data collected for one research purpose and using it for a different purpose require carefully thought-out criteria for selection to assure the validity of the secondary analysis study. Also, there are limitations on the kinds of research questions and hypotheses that can be used for secondary data analysis purposes. A researcher must adapt the study to questions that can be answered using available data, and the primary study or studies must meet criteria for responsible research.

Loss of control of data collection. Obviously, when a researcher has not been involved in developing the research design or in the collection of data, a loss of control over the process of collecting the data necessarily occurs. In addition, the secondary analyst usually has access only to summarized data.

Ethical issues. Heaton (in "Secondary analysis of qualitative data") noted that ethical issues could become important in secondary analysis. For example, should participants be contacted for permission to use their information for a different research purpose? Would the new study involve any compromise of either the identity of participants or the integrity of the data?

Establishing new variables. To meet the objectives of a new research purpose, a researcher often must

manipulate the secondary data. This could involve taking subsets of data from databases. It also could require a researcher to collapse information into new variables.

Research Applications of Secondary Data Analysis

Internationally, researchers have used secondary data to study education, business, and social issues. This section contains a review of human resource studies, including decision making at the middle management level, job satisfaction among volunteers, and job satisfaction among entrepreneurs.

Managerial Decision Making

An applied ethics study by Harris examined relationships between courage and middle management decision making. Harris used secondary qualitative data from 610 newspaper items he found in four leading international newspapers. Specifically, the data were examples of courageous behavior and situations in which the virtue of courage was needed for making management decisions. He stated he used reporter accounts because he believed that the data sources were less biased than self-reporting data would have been.

Harris conducted a content analysis of the gathered data. The process involved him in (a) selecting newspaper reports that provided the necessary data, (b) using special software to code the data, and (c) analyzing and evaluating the data. Harris also conducted a check for face validity using the phrases he had coded, noting that the percentage, reliability coefficient, and agreement coefficient were calculated. Reliability coefficients ranged from 77% to 97%; agreement coefficients ranged from 70% to 90%. According to Harris, his validity and reliability results were consistent with or superior to those of other researchers on the variable of courage.

Findings from this study indicated that accounts of courage could be recognized in the descriptions of courage and courageous activity in major newspapers throughout the world. However, Harris found

insufficient data to provide useful information relative to differences in professions or particular ethical values or choices.

Job Satisfaction of Volunteers

Silverberg and his colleagues studied job satisfaction among volunteers in the public parks and recreation industry. The purpose of their study was to check the reliability and the validity of the questionnaire used in a primary study of almost 6,000 volunteers who responded to a "modified employee job satisfaction scale." Silverberg and colleagues noted that 50% of American adults are volunteers in nonprofit organizations and that volunteer effectiveness depended on competence of volunteers and low turnover.

They determined that the instrument was a reliable, valid instrument for making inferences about volunteer job satisfaction in the parks and recreation industry. They also found that volunteer satisfaction was related to job setting and to the psychological needs met through the volunteer job function. Specifically, volunteers in this study wanted to increase their understanding of governmental and organizational operations. Working to benefit others was not seen as a prime motivator.

Recommendations from the study by Silverberg and colleagues included (a) instituting regular meetings between managers and volunteers to gain the feedback necessary to ensure high job satisfaction among volunteers and (b) trying to ensure a good match between volunteer assignments and volunteer motives.

Job Satisfaction Among Entrepreneurs

In "How satisfied are the self-employed: A secondary analysis approach," Katz used secondary analysis of three previous studies involving more than 1,200 participants to study job satisfaction among the self-employed. He concluded that self-employed workers had higher job satisfaction rates than did salaried workers. Katz also noted that having access to secondary data was insufficient for secondary analysis. Also necessary was detailed information on

(a) how the primary researcher handled missing data, (b) the procedures followed in sample and subsample selection, and (c) coding variations that would affect analysis. Katz recommended that journal editors allow more space so that researchers could report the details needed for secondary analysis. He also suggested that there be changes in research practices in the area of entrepreneurship to enable other researchers to use their data to best effect.

—*Ernest W. Brewer*

Further Reading

Cowton, C. J. (1998). The use of secondary data in business ethics research. *Journal of Business Ethics, 17*, 423–434.

Harris, H. (2001). Content analysis of secondary data: A study of courage in managerial decision-making. *Journal of Business Ethics, 34*(3/4), 191–208.

Katz, J. A. (1992). Secondary analysis in entrepreneurship: An introduction to databases and data management. *Journal of Small Business Management, 30*(2), 74–75.

Silverberg, K. E., Marshall, E. K., & Ellis, G. D. (2001). Measuring job satisfaction of volunteers in public parks and recreation. *Journal of Park and Recreation Administration, 19*(1), 79–92.

Wimmer, R. D., & Dominick, J. R. (2000). *Mass media research: An introduction.* Boston: Wadsworth.

SECTION 504 OF THE REHABILITATION ACT OF 1973

Section 504 of the Rehabilitation Act of 1973, which traces its origins to the wake of World War I, at which time the U.S. government sought to provide vocational rehabilitation to injured soldiers, is the oldest federal law addressing the needs of the disabled. According to Section 504, in language that is similar to that in Title VI of the Civil Rights Act of 1964 and Title IX of the Educational Amendments of 1972: "[n]o otherwise qualified individual with a disability in the United States . . . shall, solely by reason of her or his disability, be excluded from the participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving [f]ederal financial assistance . . . " (29 U.S.C.A. § 794(a)).

Section 504 differs significantly from the Individuals with Disabilities Education Act (IDEA). For example, while Section 504 applies to school systems that receive federal financial assistance in the form of money, books, or free lunches, school systems do not receive additional funds under its provisions, as do boards that serve children who qualify for IDEA services. Further, Section 504 protects individuals under the broader notion of impairment rather than the IDEA's reliance on the statutorily concept of disability, and Section 504 has no age limitation. Moreover, Section 504 covers students, employees, and others, including parents, while the IDEA focuses on the rights of children.

Section 504 defines an individual with a disability as one "who (i) has a physical or mental impairment which substantially limits one or more of such person's major life activities, (ii) has a record of such an impairment, or (iii) is regarded as having such an impairment" (29 U.S.C.A. § 706(7)(B)). In order to have "a record of impairment," one must have a history of, or have been identified as having, a mental or physical impairment that substantially limits one or more major life activities, including learning and working (45 C.F.R. § 84.3(j)(2)(i)). Once an individual is identified as having an impairment, educators must consider whether the child is "otherwise qualified." Unlike the IDEA, Section 504 neither requires that individuals submit to psychological or other examinations nor affords extensive due process protections.

In order to be "otherwise qualified," persons must be "(i) of an age during which nonhandicapped persons are provided such services, (ii) of any age during which it is mandatory under state law to provide such services to handicapped persons, or (iii) [one] to whom a state is required to provide a free appropriate public education [under the IDEA]" (45 C.F.R. § 84.3(k)(2)). Individuals who are "otherwise qualified" (meaning that, impairment aside, they are eligible to participate in programs or activities) must be permitted to partake as long as they can be provided with "reasonable accommodations."

Even if one appears to be "otherwise qualified," educators can rely on three defenses to avoid being charged