DATA SCIENCE - EXTRACTING INSIGHTS
Open Journal

Review

Data Analysis in Social Science Research

Chunuram Soren, PhD*

Research Scholar in Law, Sambalpur University, Odisha, India

****Corresponding author***
Chunuram Soren, PhD
Research Scholar in Law, Sambalpur University, Odisha, India

Article information
Received: May 5th, 2023; **Revised:** August 7th, 2023; **Accepted:** August 26th, 2023; **Published:** September 6th, 2023

Cite this article
Soren C. Data analysis in social science research. 2023; 1(1). doi: https://doi.org/10.70705/ppp.dsei.2023.v01.i01.pp38-42

ABSTRACT

Data analysis in social science provides the statistical concepts and programming skills to conduct and evaluate social science research studies. Data analysis provides social science researchers with the tools to unlock insights and understand complex social phenomena. Researchers can interpret the data and uncover relationships and patterns to address human behaviour and social experiences. Data analysis summarizes the collected data. It involves the interpretation of data gathered through the use of analytical and logical reasoning to determine patterns, relationships, or trends. Social science researchers focus on expanding their knowledge of social dynamics. Data analysis in social science research provides empirical evidence to dig deeper and explore human experience, attitudes, interactions, and social structures. The social science data seeks researchers with training or a demonstrable aptitude for social science work and programming to refine and extend their skills through the generation, analysis, and critique of large-scale social data. The tools for such an approach are multifaceted and evolve quickly. Data analysis encompasses both quantitative and qualitative methods. Quantitative methods in social science research allow objective insights with the help of statistical analysis. Qualitative methods in social science research provide exploratory insights with the help of textual analysis. The goal of data analysis in social science is to interpret and summarise findings. Statistical analysis is playing a major role in social science research. In the present era, the top tier journal editors expect the researchers to use a scientific approach to analyse the data collected through primary or secondary sources. This knowledge will help social science researchers and teachers understand the process of analysing the data and reporting the results in their research projects using different software.

Keywords
Social science; Data; Analysis; Interpretation.

INTRODUCTION

We do not generally associate data with mathematics. However, data is the base of all operations in statistics. So let us learn more about data collection, primary data, secondary data, and a few other important terms.

Data can be defined as a systematic record of a particular quantity. It is the different values of that quantity represented together in a set. It is a collection of facts and figures to be used for a specific purpose such as a survey or analysis. When arranged in an organized form can be called information. The source of data (primary data, secondary data) is also an important factor.

Research data comes in many different formats and is gathered using a wide variety of methodologies. In this module, we will provide you with a basic definition and understanding of what research data are. We'll also explore how data fits into the scholarly research process.

Many people think of data-driven research as something that primarily happens in the sciences. It is often thought of as involving a spreadsheet filled with numbers. Both of these beliefs are incorrect. Research data are collected and used in scholarship across all academic disciplines and, while it can consist of numbers in a spreadsheet, it also takes many different formats, including videos, images, artifacts, and diaries. A psychologist collecting survey data to better understand human behavior, an artist using data to generate images and sounds, or an anthropologist using audio files to document observations about different cultures, scholarly research across all academic fields is increasingly data-driven.

In our Data Literacy Modules, we will demonstrate the ways in which research data are gathered and used across various academic disciplines by discussing it in a very broad sense. We define research data as: any information collected, stored, and processed to produce and validate original research results. Data might be used to prove or disprove a theory, bolster claims made in research, or to further the knowledge around a specific topic or problem.

II. Meaning of Data:

One definition of research data is: "the recorded factual material commonly accepted in the scientific community as necessary to validate research findings."

Data, information, knowledge, and wisdom are closely related concepts, but each has its role concerning the other, and each term has its meaning.

According to a common view, data is collected and analyzed; data only becomes information suitable for making decisions once it has been analyzed in some fashion. One can say that the extent to which a set of data is informative to someone depends on the extent to which it is unexpected by that person. The amount of information contained in a data stream may be characterized by its Shannon entropy.

Knowledge is the awareness of its environment that some entity possesses, whereas data merely communicate that knowledge. For example, the entry in a database specifying the height of Mount Everest is a datum that communicates a precisely-measured value. This measurement may be included in a book along with other data on Mount Everest to describe the mountain in a manner useful for those who wish to decide on the best method to climb it. An awareness the characteristics represented by these data is knowledge.

Data is often assumed to be the least abstract concept, information the next least, and knowledge the most abstract. In this view, data becomes information by interpretation; e.g., the height of Mount Everest is generally considered "data", a book on Mount Everest geological characteristics may be considered "information", and a climber's guidebook containing practical information on the best way to reach Mount Everest's peak may be considered "knowledge". "Information" bears a diversity of meanings that ranges from everyday usage to technical use. This view, however, has also been argued to reverse how data emerges from information, and information from knowledge. Generally speaking, the concept of information is closely related to notions of constraint, communication, control, data, form, instruction, knowledge, meaning, mental stimulus, pattern, perception, and representation. Beynon-Davies uses the concept of a sign to differentiate between data and information; data is a series of symbols, while information occurs when the symbols are used to refer to something.

Research data covers a broad range of types of information (see examples below), and digital data can be structured and stored in a variety of file formats. Note that properly managing data (and records) does not necessarily equate to sharing or publishing that data.

Data is a set of values of subjects with respect to qualitative or quantitative variables. Data is raw, unorganized facts that need to be processed. Data can be something simple and seemingly random and useless until it is organized.

When data is processed, organized, structured or presented in a given context so as to make it useful, it is called information. Information, necessary for research activities are achieved in different forms.

III. Concept of data collection:

Data collection is the process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer stated research questions, test hypotheses, and evaluate outcomes. The data collection component of research is common to all fields of study including physical and social sciences, humanities, business, etc. While methods vary by discipline, the emphasis on ensuring accurate and honest collection remains the same.

Data collection is one of the most important stages in conducting a research. You can have the best research design in the world but if you cannot collect the required data you will be not be able to complete your project.

Data collection is a very demanding job which needs thorough planning, hard work, patience, perseverance and more to be able to complete the task successfully. Data collection starts with determining what kind of data required followed by the selection of a sample from a certain population. After that, you need to use a certain instrument to collect the data from the selected sample. It enables researcher to get the answer of Research Question, to test Hypothesis, Analysis & Interpretation of data, specially help investigator to arrive at conclusion.

IV. Types of Data:

Data may be qualitative or quantitative. Once you know the difference between them, you can know how to use them.

Qualitative Data: They represent some characteristics or attributes. They depict descriptions that may be observed but cannot be computed or calculated. For example, data on attributes such as intelligence, honesty, wisdom, cleanliness, and creativity collected using the students of your class a sample would be classified as qualitative. They are more exploratory than conclusive in nature.

Qualitative research methods, on the contrary, do not involve numbers or mathematical calculations. Qualitative research is closely associated with words, sounds, feeling, emotions, colours and other elements that are non-quantifiable.

Qualitative studies aim to ensure greater level of depth of understanding and qualitative data collection methods include interviews, questionnaires with open-ended questions, focus groups, observation, game or role- playing, case studies etc.

Quantitative Data: These can be measured and not simply observed. They can be numerically represented and calculations can be performed on them. For example, data on the number of students playing different sports from your class gives an estimate of how many of the total students play which sport. This information is numerical and can be classified as quantitative.

Quantitative data collection methods are based on mathematical calculations in various formats. Methods of quantitative data collection and analysis include questionnaires with closed-ended questions, methods of correlation and regression, mean, mode and median and others.

Quantitative methods are cheaper to apply and they can be applied within shorter duration of time compared to qualitative methods. Moreover, due to a high level of standardisation of quantitative methods, it is easy to make comparisons of findings.

V. Data Collection:

Depending on the source, it can classify as primary data or secondary data. Let us take a look at them both.

Primary Data

Data that has been collected from first-hand-experience is known as primary data. Primary data has not been published yet and is more reliable, authentic and objective. Primary data has not been changed or altered by human beings; therefore its validity is greater than secondary data. Importance of Primary Data: In statistical surveys it is necessary to get information from primary sources and work on primary data.

For example, the statistical records of female population in a country cannot be based on newspaper, magazine and other printed sources. A research can be conducted without secondary data but a research based on only secondary data is least reliable and may have biases because secondary data has already been manipulated by human beings. One of such sources is old and secondly they contain limited information as well as they can be misleading and biased. Sources of Primary Data: Sources for primary data are limited and at times it becomes difficult to obtain data from primary source because of either scarcity of population or lack of cooperation. Following are some of the sources of primary data.

These are the data that are collected for the first time by an investigator for a specific purpose. Primary data are 'pure' in the sense that no statistical operations have been performed on them and they are original. An example of primary data is the Census of India.

Primary data is the type of data that has not been around before. Primary data is unique findings of your research. Primary data collection and analysis typically requires more time and effort to conduct compared to the secondary data research. Primary data collection methods can be divided into two groups: quantitative and qualitative.

Your choice between quantitative or qualitative methods of data collection depends on the area of your research and the nature of research aims and objectives.

Secondary Data

Secondary data is the data that is collected from the primary sources which can be used in the current research study. Collecting secondary data often takes considerably less time than collecting primary data where you would have to gather every information from scratch. It is thus possible to gather more data this way. Secondary data can be obtained from two different research strands –

Quantitative: Census, housing, social security as well as electoral statistics and other related databases.

Qualitative: Semi-structured and structured interviews focus groups' transcripts, field notes, observation records and other personal, research-related documents.

Secondary data is often readily available. After the expense of electronic media and internet the availability of secondary data has become much easier. Published Printed Sources: There are varieties of published printed sources. Their credibility depends on many factors. For example, on the writer, publishing company and time and date when published. New sources are preferred and old sources should be avoided as new technology and researches bring new facts into light. Books: Books are available today on any topic that you want to research.

The use of books starts before even you have selected the topic. After selection of topics books provide insight on how much work has already been done on the same topic and you can prepare your literature review. Books are secondary source but most authentic one in secondary sources. Journals/periodicals: Journals and periodicals are becoming more important as far as data collection is concerned. The reason is that journals provide up-to-date information which at times books cannot and secondly, journals can give information on the very specific topic on which you are researching rather talking about more general topics. Magazines/Newspapers: Magazines are also effective but not very reliable. Newspapers on the other hand are more reliable and in some cases the information can only be obtained from newspapers as in the case of some political studies.

They are the data that are sourced from someplace that has originally collected it. This means that this kind of data has already been collected by some researchers or investigators in the past and is available either in published or unpublished form. This information is impure as statistical operations may have been performed on them already. An example is information available on the Government of India, the Department of Finance's website or in other repositories, books, journals, etc.

Secondary data is a type of data that has already been published in books, newspapers, magazines, journals, online portals etc. There is an abundance of data available in these sources about your research area in business studies, almost regardless of the nature of the research area. Therefore, application of appropriate set of criteria to select secondary data to be used in the study plays an important role in terms of increasing the levels of research validity and reliability.

These criteria include, but not limited to date of publication, credential of the author, reliability of the source, quality of discussions, depth of analyses, the extent of contribution of the text to the development of the research area etc. Secondary data collection is discussed in greater depth in Literature Review chapter.

Secondary data collection methods offer a range of advantages such as saving time, effort and expenses. However they have a major disadvantage. Specifically, secondary research does not make contribution to the expansion of the literature by producing fresh (new) data.

VI. Discrete and Continuous Data:

Discrete Data: These are data that can take only certain specific values rather than a range of values. For example, data on the blood group of a certain population or on their genders is termed as discrete data. A usual way to represent this is by using bar charts.

Continuous Data: These are data that can take values between a certain range with the highest and lowest values. The difference between the highest and lowest value is called the range of data. For example, the age of persons can take values even in decimals or so is the case of the height and weights of the students of your school. These are classified as continuous data. Continuous data can be tabulated in what is called a frequency distribution. They can be graphically represented using histograms.

VII. Data Interpretation:

Data interpretation refers to the process of using diverse analytical methods to review data and arrive at relevant conclusions. The interpretation of data helps researchers to categorize, manipulate, and summarize the information in order to answer critical questions.

The importance of data interpretation is evident and this is why it needs to be done properly. Data is very likely to arrive from multiple sources and has a tendency to enter the analysis process with haphazard ordering. Data analysis tends to be extremely subjective. That is to say, the nature and goal of interpretation will vary from business to business, likely correlating to the type of data being analyzed. While there are several types of processes that are implemented based on individual data nature, the two broadest and most common categories are "quantitative and qualitative analysis".

When interpreting data, an analyst must try to discern the differences between correlation, causation, and coincidences, as well as many other biases – but he also has to consider all the factors involved that may have led to a result. There are various data interpretation methods one can use to achieve this.

The interpretation of data is designed to help people make sense of numerical data that has been collected, analyzed, and presented. Having a baseline method for interpreting data will provide your analyst teams with a structure and consistent foundation. Indeed, if several departments have different approaches to interpreting the same data while sharing the same goals, some mismatched objectives can result. Disparate methods will lead to duplicated efforts, inconsistent solutions, wasted energy, and inevitably – time and money. In this part, we will look at the two main methods of interpretation of data: qualitative and quantitative analysis.

This involves asking a series of questions about your data that relate to your study question(s). Your answers to these questions are organized as findings and conclusions. Based on these, you may develop recommendations for action or further study.

Interpreting your data is a process that involves answering a series of questions about it. We suggest the following steps:

1) Review and interpret the data "in-house" to develop preliminary findings, conclusions, and recommendations.

2) Review the data and your interpretation of it with an advisory group or technical committee. This group should involve local, regional, and state resource people who are familiar with monitoring and with your river. They can verify, add to, or correct your interpretation of the results.

3) Review the data and your interpretation of it with the people who will use your data -- for example, the public, river users, and government officials. Ultimately, your interpretation of the data relates back to the questions your monitoring program is trying to answer. For example, does the river meet state water quality standards?

VIII. Process of data analysis:

Partington (2003) states that there is a little standardization with no absolutes, where a specific type of qualitative data relates to a specific type of analysis.

Neuman (2011) further opines that no single qualitative data analysis approach is widely accepted. Posit that there are always variations in the number and description of steps for the same process of data analysis by different authors. From the preceding views, it can thus be inferred that each qualitative data analysis to some extent will be a uniquely designed event. With the preceding in mind, the qualitative data analysis of this research (responses from the semi structured interviews) was done according to a qualitative content analysis process that integrated

Qualitative content analysis involved the following procedures:

1) Recording of data was done by audio recording on a digital voice recorder, while audio recording on another tape recorder served as backup of electronic failure and faults; and to ensure that all voices could be heard. Taking notes served as further backup and provided the context to the interviews.

2) Verbatim transcription of the responses from the interview commenced as soon as possible; and was done by an expert to ensure a speedy completion. To ensure that the researcher became acquainted with the data for the purpose of analysis and interpretation, the original interview of the completed verbatim transcription was listened to again. Transcription notation.

IX. Conclusion:

As we reach the end of this insightful post about data interpretation and analysis we hope you have a clear understanding of the topic. We've covered the definition, and given some examples and methods to perform a successful interpretation process.

Data interpretation and analysis is an important aspect of working with data sets in any field or research and statistics. They both go hand in hand, as the process of data interpretation involves the analysis of data. Data interpretation is very important, as it helps to acquire useful information from a pool of irrelevant ones while making informed decisions. It is found useful for individuals, businesses, and researchers.

The importance of data interpretation is undeniable. Dashboards not only bridge the information gap between traditional data interpretation methods and technology, but they can help remedy and prevent the major pitfalls of the process. As a digital age solution, they combine the best of the past and the present to allow for informed decision-making with maximum data interpretation.

The interpretation of data is designed to help people make sense of numerical data that has been collected, analyzed, and presented. Having a baseline method for interpreting data will provide your analyst teams with a structure and consistent foundation. Indeed, if several departments have different approaches to interpreting the same data while sharing the same goals, some mismatched objectives can result. Disparate methods will lead to duplicated efforts, inconsistent solutions, wasted energy, and inevitably – time and money. In this part, we will look at the two main methods of interpretation of data: qualitative and quantitative analysis.

REFERENCES

1. P.M. Bakshi, "Legal Research and Law Reform", S.K. Verma and M. Afzal Wani editions Legal Reseach and Methodology, Indian Law Institute, New Delhi, 2nd. Ed.,2011.

2. Raj Kumari Agarwala, "Indian Legal Research", 1982-JSTOR.

3. B. S. Murty, "Socio-Legal Research", 1982-JSTOR.

4. S.N. Jain, "Doctrinal and Non-Doctrinal research, Indian Law Institute, New Delhi, 1983.

5. S.N. Jain, "Legal Research Methodology", Journal International Law Institute, New Delhi, 1972.

6. R.B. Jain, "Sampling Methods in Legal Research", S.K. Verma and M. Afzal Wani editions Legal Research and Methodology, Indian Law Institute, New Delhi, 2nd. Ed., 2011.

7. B.N. Ghose, "Scaling Techniques", S.K. Verma and M. Afzal Wani editions Legal Research and Methodology, Indian Law Institute, New Delhi, 2nd. Ed., 2011.

8. K.K. Garg, "Research Methodology", Omega Publications, New Delhi, 2006.

9. C.R. Kathari, "Research Methodology: Methods and Techniques", New age International Publication, 2006.