RESEARCH METHODOLOGY

UNIT- 3

Syllabus - UNIT-III : DATA PROCESSING AND ANALYSIS

Processing of data: editing, coding and Tabulation - Problems - use of computer in social research. Analysis of data: Statistical analysis; diagrammatic and graphic representation. Interpretation of results.

Expected Leaning Outcome:

1. To learn the skills of analyzing qualitative and quantitative information.
2. To understand the Process of Data
3. To know application Computer in Social Research
4. To know the art of diagrammatic and graphic representation of Results

DATA PROCESSING

After collecting data, the method of converting raw data into meaningful statement; includes data processing, data analysis, and data interpretation and presentation. Data reduction or processing mainly involves various manipulations necessary for preparing the data for analysis. The process (of manipulation) could be manual or electronic. It involves editing, categorizing the open-ended questions, coding, computerization and preparation of tables and diagrams. Data processing is concerned with editing, coding, classifying, tabulating and charting and diagramming research data. The essence of data processing in research is data reduction.

Data reduction involves winnowing out the irrelevant from the relevant data and establishing order from chaos and giving shape to a mass of data. Data processing in research consists of five important steps

1. **Editing of data**
2. **Coding of data**
3. **Classification of data**
4. **Tabulation of data**
5. **Data diagrams**

Data Collection, Processing and Analysis

Acquiring data: Acquisition involves collecting or adding to the data holdings.

There are several methods of acquiring data:

1. **collecting new data**
2. **using your own previously collected data**
3. **reusing someone others data**
4. **purchasing data**
5. **acquired from Internet (texts, social media, photos**)

Data processing: A series of actions or steps performed on data to verify, organize, transform, integrate, and extract data in an appropriate output form for subsequent use. Methods of processing must be rigorously documented to ensure the utility and integrity of the data.

Data Analysis involves actions and methods performed on data that help describe facts, detect patterns, develop explanations and test hypotheses. This includes data quality assurance, statistical data analysis, modeling, and interpretation of results.

Results: The results of above mentioned actions are published as a research paper. In case the research data is made accessible, one has to prepare the data set for opening up.

DATA PROCESSING

Data processing occurs when data is collected and translated into usable information. Usually performed by a data scientist or team of data scientists, it is important for data processing to be done correctly as not to negatively affect the end product, or data output. Data processing starts with data in its raw form and converts it into a more readable format (graphs, documents, etc.), giving it the form and context necessary to be interpreted by computers and utilized by employees throughout an organization.

Six stages of data processing

1. Data collection

Collecting data is the first step in data processing. Data is pulled from available sources, including data lakes and data warehouses. It is important that the data sources available are trustworthy and well-built so the data collected (and later used as information) is of the highest possible quality.

2. Data preparation

Once the data is collected, it then enters the data preparation stage. Data preparation, often referred to as “pre-processing” is the stage at which raw data is cleaned up and organized for the following stage of data processing. During preparation, raw data is diligently checked for any errors. The purpose of this step is to eliminate bad data (redundant, incomplete, or incorrect data) and begin to create high-quality data for the best business intelligence.

3. Data input

The clean data is then entered into its destination and translated into a language that it can understand. Data input is the first stage in which raw data begins to take the form of usable information.

4. Processing

During this stage, the data inputted to the computer in the previous stage is actually processed for interpretation. Processing is done using machine learning algorithms, though the process itself may vary slightly depending on the source of data being processed (data lakes, social networks, connected devices etc.) and its intended use (examining advertising patterns, medical diagnosis from connected devices, determining customer needs, etc.).

5. Data output/interpretation

The output/interpretation stage is the stage at which data is finally usable to non-data scientists. It is translated, readable, and often in the form of graphs, videos, images, plain text, etc.). Members of the company or institution can now begin to self-serve the data for their own data analytics projects.

6. Data storage and Report Writing

The final stage of data processing is storage. After all of the data is processed, it is then stored for future use. While some information may be put to use immediately, much of it will serve a purpose later on. Plus, properly stored data is a necessity for compliance with data protection legislation like GDPR. When data is properly stored, it can be quickly and easily accessed by members of the organization when needed.

DATA ANALYSIS TOOLS

Data analysis tools make it easier for users to process and manipulate data, analyze the relationships and correlations between data sets, and it also helps to identify patterns and trends for interpretation. Here is a complete list of tools.

Types of Data Analysis: Techniques and Methods

There are several types of Data Analysis techniques that exist based on business and technology. **However, the major types of data analysis are:**

- ✓ Text Analysis
- ✓ Statistical Analysis
- ✓ Diagnostic Analysis
- ✓ Predictive Analysis
- ✓ Prescriptive Analysis
- ✓ Text Analysis

Text Analysis is also referred to as Data Mining. It is a method to discover a pattern in large data sets using databases or data mining tools. It used to transform raw data into business information. Business Intelligence tools are present in the market which is used to take strategic business decisions. Overall it offers a way to extract and examine data and deriving patterns and finally interpretation of the data.

Statistical Analysis -Statistical Analysis shows "What happen?" by using past data in the form of dashboards. Statistical Analysis includes collection, Analysis, interpretation, presentation, and modeling of data. It analyses a set of data or a sample of data. There are two categories of this type of Analysis - Descriptive Analysis and Inferential Analysis.

Descriptive Analysis - analyses complete data or a sample of summarized numerical data. It shows mean and deviation for continuous data whereas percentage and frequency for categorical data.

Inferential Analysis =analyses sample from complete data. In this type of Analysis, you can find different conclusions from the same data by selecting different samples.

Diagnostic Analysis - Diagnostic Analysis shows "Why did it happen?" by finding the cause from the insight found in Statistical Analysis. This Analysis is useful to identify behavior patterns of data. If a new problem arrives in your business process, then you can look into this Analysis to find similar patterns of that problem. And it may have chances to use similar prescriptions for the new problems.

Predictive Analysis - Predictive Analysis shows "what is likely to happen" by using previous data. The simplest example is like if last year I bought two dresses based on my savings and if this year my salary is increasing double then I can buy four dresses. But of course it's not easy like this because you have to think about other circumstances like chances of prices of clothes is increased this year or maybe instead of dresses you want to buy a new bike, or you need to buy a house.So here, this Analysis makes predictions about future outcomes based on current or past data. Forecasting is just an estimate. Its accuracy is based on how much detailed information you have and how much you dig in it.

Prescriptive Analysis - Prescriptive Analysis combines the insight from all previous Analysis to determine which action to take in a current problem or decision. Most data-driven companies are utilizing Prescriptive Analysis because predictive and descriptive Analysis are not enough to improve data performance. Based on current situations and problems, they analyze the data and make decisions.

Quantitative Data Analysis:

As the name suggests, the quantitative analysis is used for the quantification of data which allows the generalization of the results obtained from a sample to a population of interest.Simply put, statistical methods of data analysis are used to collect raw data and transform it into numerical data. Some of the methods that fall under that Quantitative Analysis are:

Mean: Also known as the average, Mean is the most basic method of analyzing data where the sum of a numbers' list is divided by the number of items on that list. It is useful in determining the overall trend of something.

Hypothesis Testing: Majorly used in business research and is done to assess if a certain theory or hypothesis for a population or data set is true.

Sample Size Determination: When doing research on a large population like workforce for your company, small sample size is taken and then analyzed, and the results are considered almost same for every member of the population.

DATA ANALYSIS PROCESS

The Data Analysis Process is gathering information by using a proper application or tool which allows you to explore the data and find a pattern in it. Based on that information and data, you can make decisions, or you can get ultimate conclusions.

Data Analysis consists of the following phases:

1. **Data Requirement Gathering**
2. **Data Collection**
3. **Data Cleaning**
4. **Data Analysis**
5. **Data Interpretation**
6. **Data Visualization**

Data Requirement Gathering - First of all, you have to think about why do you want to do this data analysis? All you need to find out the purpose or aim of doing the Analysis. You have to decide which type of data analysis you wanted to do! In this phase, you have to decide what to analyze and how to measure it, you have to understand why you are investigating and what measures you have to use to do this Analysis.

Data Collection - After requirement gathering, you will get a clear idea about what things you have to measure and what should be your findings. Now it's time to collect your data based on requirements. Once you collect your data, remember that the collected data must be processed or organized for Analysis. As you collected data from various sources, you must have to keep a log with a collection date and source of the data.

Data Cleaning - Now whatever data is collected may not be useful or irrelevant to your aim of Analysis, hence it should be cleaned. The data which is collected may contain duplicate records, white spaces or errors. The data should be cleaned and error free. This phase must be done before Analysis because based on data cleaning, your output of Analysis will be closer to your expected outcome.

Data Analysis - Once the data is collected, cleaned, and processed, it is ready for Analysis. As you manipulate data, you may find you have the exact information you need, or you might need to collect more data. During this phase, you can use data analysis tools and software which will help you to understand, interpret, and derive conclusions based on the requirements.

Data Interpretation - After analyzing your data, it's finally time to interpret your results. You can choose the way to express or communicate your data analysis either you can use simply in words or maybe a table or chart. Then use the results of your data analysis process to decide your best course of action.

Data Visualization

Data visualization is very common in your day to day life; they often appear in the form of charts and graphs. In other words, data shown graphically so that it will be easier for the human brain to understand and process it. Data visualization often used to discover unknown facts and trends. By observing relationships and comparing datasets, you can find a way to find out meaningful information.

METHODS OF DATA PROCESSING IN RESEARCH

Data processing is that procedure in which research frame collected data through editing, coding, classifying, tabulating, charting, and diagramming. The purpose of data processing in research is data reduction or minimization. This processing transforms irrelevant data to relevant. Basically it works with 5 steps that is given below.

Validation - Covers five areas:

1. Fraud
2. Screening
3. Procedure
4. Completeness
5. Courtesy

EDITING OF DATA - Editing is the first step of data processing. Editing is the process of examine the data collected through questionnaire or any other method. It starts after all data collection to check it or reform into useful data.

1. Raw data is checked for mistakes made by either the interviewer or the respondent
2. By reviewing completed interviews from primary research, the researcher can check several areas of concern:
3. Asking the proper questions
4. Accurate recording of answers
5. Correct screening of respondents
6. Complete and accurate recording of open-ended questions

Mildred B. Parten in his book points out that the editor is responsible for seeing that the data are;

1. Accurate as possible,
2. Consistent with other facts secured,
3. Uniformly entered,
4. As complete as possible,
5. Acceptable for tabulation and arranged to facilitate coding tabulation.

There are different **types of editing**. They are:

1. **Editing for quality** asks the following questions: are the data forms complete, are the data free of bias, are the recordings free of errors, are the inconsistencies in responses within limits, are there evidences to show dishonesty of enumerators or interviewers and are there any wanton manipulation of data.
2. **Editing for tabulation** does certain accepted modification to data or even rejecting certain pieces of data in order to facilitate tabulation. or instance, extremely high or low value data item may be ignored or bracketed with suitable class interval.

3. **Field Editing** is done by the enumerator. The schedule filled up by the enumerator or the respondent might have some abbreviated writings, illegible writings and the like. These are rectified by the enumerator. This should be done soon after the enumeration or interview before the loss of memory. The field editing should not extend to giving some guess data to fill up omissions.
4. **Central Editing** is done by the researcher after getting all schedules or questionnaires or forms from the enumerators or respondents. Obvious errors can be corrected. For missed data or information, the editor may substitute data or information by reviewing information provided by likely placed other respondents. A definite inappropriate answer is removed and "no answer" is entered when reasonable attempts to get the appropriate answer fail to produce results.

Editors must keep in view the following points while performing their work:

1. They should be familiar with instructions given to the interviewers and coders as well as with the editing instructions supplied to them for the purpose,
2. While crossing out an original entry for one reason or another, they should just draw a single line on it so that the same may remain legible,
3. They must make entries (if any) on the form in some distinctive color and that too in a standardized form,
4. They should initial all answers which they change or supply,
5. Editor's initials and the data of editing should be placed on each completed form or schedule.

CODING OF DATA - Coding is the process of categories data according to research subject or topic and the design of research. In coding process researcher set a code for a particular things like male - M, Female- F that indicate the gender in questionnaire without writing full spelling same as researcher can be use colors to highlight something or numbers like 1+, 1-. this type of coding makes easy to calculate or evaluate result in tabulation.

1. Grouping and assigning values to various responses from the survey instrument
2. Codes are numerical
3. Can be tedious if certain issues are not addressed prior to collecting the data

Four-step process to develop codes for responses:

1. Generate a list of as many potential responses as possible
2. Consolidate responses
3. Assign a numerical value as a code
4. Assign a coded value to each response

CLASSIFICATION OF DATA - Classification or categorization is the process of grouping the statistical data under various understandable homogeneous groups for the purpose of convenient interpretation. A uniformity of attributes is the basic criterion for classification; and the grouping of data is made according to similarity. Classification becomes necessary when there is a diversity in the data collected for meaningless for meaningful presentation and analysis. However, it is meaningless in respect of homogeneous data. A good classification should have the characteristics of clarity, homogeneity, equality of scale, purposefulness and accuracy.

Objectives of Classification are below:

1. The complex scattered and haphazard data is organized into concise, logical and intelligible form.
2. It is possible to make the characteristics of similarities and dis – similarities clear.
3. Comparative studies is possible.
4. Understanding of the significance is made easier and thereby good deal of human energy is saved.
5. Underlying unity amongst different items is made clear and expressed.
6. Data is so arranged that analysis and generalization becomes possible.

TABULATION OF DATA - Tabulation is the process of summarizing raw data and displaying it in compact form for further analysis. Therefore, preparing tables is a very important step. Researcher can be tabulation by hand or in digital mode. The choice is made largely on the basis of the size and type of study, alternative costs, time pressures, and the availability of computers, and computer programmes. If the number of questionnaire is small, and their length short, hand tabulation is quite satisfactory. The counting the number of observations (cases) that are classified into certain categories

1. One-way tabulation: Categorization of single variables existing in a study
2. Cross-tabulation: Simultaneously treating two or more variables in the study
3. Categorizing the number of respondents who have answered two or more questions consecutively

Table may be divided into: (i) Frequency tables, (ii) Response tables, (iii) Contingency tables, (iv) Uni-variate tables, (v) Bi-variate tables, (vi) Statistical table and (vii) Time series tables.

Generally a research table has the following parts:

(a) table number,
(b) title of the table,
(c) caption
(d) stub (row heading),
(e) body,
(f) head note,
(g) foot note.

As a general rule the following steps are necessary in the preparation of table:

Title of table: The table should be first given a brief, simple and clear title which may express the basis of classification.

Columns and rows: Each table should be prepared in just adequate number of columns and rows.

Captions and stubs: The columns and rows should be given simple and clear captions and stubs.

Ruling: Columns and rows should be divided by means of thin or thick rulings.

Arrangement of items; Comparable figures should be arranged side by side.

Deviations: These should be arranged in the column near the original data so that their presence may easily be noted.

Size of columns: This should be according to the requirement.

Arrangements of items: This should be according to the problem.

Special emphasis: This can be done by writing important data in bold or special letters.

Unit of measurement: The unit should be noted below the lines.

Approximation: This should also be noted below the title.

Foot – notes: These may be given below the table.

Total: Totals of each column and grand total should be in one line.

Source : Source of data must be given. For primary data, write primary data.

DATA DIAGRAMS - Diagrams are charts and graphs used to present data. These facilitate getting the attention of the reader more. These help presenting data more effectively. Creative presentation of data is possible. The data diagrams classified into:

1. **Charts:** A chart is a diagrammatic form of data presentation. Bar charts, rectangles, squares and circles can be used to present data. Bar charts are uni-dimensional, while rectangular, squares and circles are two-dimensional.
2. **Graphs:** The method of presenting numerical data in visual form is called graph, A graph gives relationship between two variables by means of either a curve or a straight line. Graphs may be divided into two categories. (1) Graphs of Time Series and (2) Graphs of Frequency Distribution. In graphs of time series one of the factors is time and other or others is / are the study factors. Graphs on frequency show the distribution of by income, age, etc. of executives and so on.

Problems in Processing of data:

The problem concerning "Don't know" (or DK) responses: While processing the data, the researcher often comes across some responses that are difficult to handle. One category of such responses may be 'Don't Know Response' or simply DK response. When the DK response group is small, it is of little significance. But when it is relatively big, it becomes a matter of major concern in which case the question arises: Is the question which elicited DK response useless? The answer depends on two points viz., the respondent actually may not know the answer or the researcher may fail in obtaining the appropriate information. In the first case the concerned question is said to be alright and DK response is taken as legitimate DK response. But in the second case, DK response is more likely to be a failure of the questioning process.

How DK responses are to be dealt with by researchers? The best way is to design better type of questions. Good rapport of interviewers with respondents will result in minimising DK responses. But

what about the DK responses that have already taken place? One way to tackle this issue is to estimate the allocation of DK answers from other data in the questionnaire. The other way is to keep DK responses as a separate category in tabulation where we can consider it as a separate reply category if DK responses happen to be legitimate, otherwise we should let the reader make his own decision. Yet another way is to assume that DK responses occur more or less randomly and as such we may distribute them among the other answers in the ratio in which the latter have occurred. Similar results will be achieved if all DK replies are excluded from tabulation and that too without inflating the actual number of other responses.

USES OF COMPUTER IN SOCIAL RESEARCH.

Computer have always assisted to solve the problems faced by the mankind since the time of invention, the size of the computers have drastically reduced from that of a room to that can be accommodated in a human palm. The word computer means something which computes or a machine for performing calculations automatically, but, today computer means not merely a calculator. It does vast variety of jobs with tremendous speed and efficiency. Today people use computers in almost every walk of life. Electronic computers have now become an indispensable part of every profession: so do research. Computers have a very important role to play in all research activities.

The practice of computing in sociology has evolved rapidly. Computers have been applied to practically every research task, including such unlikely ones as field note-taking, interviewing, and hundreds of other tasks (Brent and Anderson 1990). The many diverse uses of computing technology in social research are difficult to categorize because applications overlap and evolve unpredictably. Nonetheless, it is necessary to discuss different categories of applications in order to describe the state of the art of computing in sociology.

The importance of computers in scientific research is exceptionally high and the use of a computer can help scientific and Social Science research immensely, and is an almost invaluable tool. There are many reasons why computers are so important in scientific research and here are some of the main reasons:

SPEED: computer can process numbers and information in a very short time. So researcher can process and analyze data quickly. By saving time researcher can conduct further research. Calculation that may take a person several hours to process will take computer mere minutes, if not seconds.

ACCURACY: Computer is incredibly accurate. Accuracy is very much important in scientific research. Wrong calculation could result an entire research or project being filled with incorrect information.

ORGANIZATION: We can store millions of pages of information by using simple folders, word processors & computer programs. Computer is more productive & safer than using a paper filing system in which anything can be easily misplaced.

CONSISTENCY: computer cannot make mistakes through "tiredness" or lack of concentration like human being. This characteristic makes it exceptionally important in scientific research.

. **Writing and Publishing.** Once equated with the secretarial pool, word processing now is an activity of nearly every graduate student and professional in sociology. It consists not only of writing but preparing tables, "typesetting" mathematical equations, and resizing objects, such as three-dimensional graphs embedded within text. Social researchers are using such capabilities and moving rapidly toward workstation environments that obscure the transition between data analysis and manuscript preparation (Steiger and Fouladi 1990). Not only do researchers use their computers for writing papers, but word processing software plays a central role in the refinement of data collection instruments, especially questionnaires and codebooks, which allows for rapid production of alternative forms and multiple drafts.

Importance of Computer in research:

Internet: Before you start research, you often want to quickly learn about possible issues or topics of study by searching available sources of information. Nearly all academic journals are available online, and many are organized into online databases. Government agencies often have demographic or economic information online you can use in your research.

Information Storage: Computers store vast amounts of information. You can quickly and efficiently organize and search information, making for easier retrieval than paper storage. You can store your raw data in multiple formats. Some researchers conduct their research online, often through the use of surveys.

Computational Tools: Computers began as powerful calculators, and that service is important to research today. Regardless of the amount of data you have, you can do more with it with a computer's

help. Statistical programs, modeling programs and spatial mapping tools are all possible because of computers. Researchers can use information in new ways, such as layering different types of maps on one another to discover new patterns in how people use their environment.

Communication; Building knowledge through research requires communication between experts to identify new areas requiring research and debating results. Before computers, this was accomplished through papers and workshops. Now, the world's experts can communicate via email or webchats. Information can be spread by virtual conferences. Knowledge from marginalized groups, such as African scholars, is now more visible.

Mobility Researchers can take computers anywhere, making it easier to conduct field research and collect data. New areas of research in remote areas or at a community level are opened up by the mobility of computers. Social media sites have become a new medium for interaction and information.

COMPUTER APPLICATIONS IN THE RESEARCH PROCESS

Research process consists of series of actions or steps necessary to effectively carry out research and the desired sequencing of these steps. The following order concerning various steps provides a useful procedural guideline regarding the research process:

(1) Formulating the research problem;
(2) Extensive literature survey;
(3) Developing the hypothesis;
(4) Preparing the research design;
(5) Determining sample design;
(6) Collecting the data;
(7) Execution of the project;
(8) Analysis of data;
(9) Hypothesis testing;
(10) Generalisations and interpretation, and
(11) Preparation of the report or presentation of the results, i.e., formal write-up of conclusions reached.

There are five major phases of the research process. They are:

1. Conceptual phase

2. Design and planning phase
3. Data collection phase
4. Data Analysis phase and
5. Research Publication phase

1. Role of Computer in Conceptual Phase: The conceptual phase consists of formulation of research problem, extensive literature survey, theoretical frame work and developing the hypothesis. Use of computers in extensive literature review: computers help for searching the literatures (for review of literature) and bibliographic reference stored in the electronic database. This has the advantage over searching the literatures in the form of books, journals and other newsletters at the libraries which consume considerable amount of time and effort.

2. Role of Computers in Design and Planning Phase: This phase consists of research design preparation and determining sample design. Design and planning phase also consists of population, research variables, sampling plan, reviewing research plan and pilot study. Role of Computers for Sample Size Calculation: Several software's are available to calculate the sample size required for a proposed study. The standard deviation of the data from the pilot study is required for the sample size calculation.

3. Role of Computers in Data collection phase: This Empirical phase consists of collecting and preparing the data for analysis: In research studies, the preparation and inputting data is the most labor-intensive and time consuming aspect of the work. Typically the data will be initially recorded on a questionnaire or record for suitable for its acceptance by the computer. To do this the researcher in conjunction with the statistician and the programmer, will convert the data into Microsoft word file or excel spreadsheet or any statistical software data file. These data can be directly opened with statistical software's for analysis. Data collection and Storage: The data obtained from the subjects are stored in computes are word files or excel spread sheets or any statistical software data file. This has the advantage of making necessary corrections or editing the whole layout of the tables if needed, which is impossible or time consuming incase of writing in papers. Thus, computers help in data entry, data editing, data management including follow up actions etc. computers also allow for greater flexibility in recording the data while they are collected as well as greater ease during the analysis of these data. Examples of editors are WordPad, SPSS data editor, word processors, others like ultraedit etc. Data exposition: Most researchers are anxious about seeing the data: what they look like; how

they are distributed etc. you can also examine different dimension of variables or plot them in various charts using a statistical application.

4. Role of Computers in Data Analysis: This phase consist of the analysis of data, hypothesis testing and generalisations and interpretation. Data analysis phase mainly consist of statistical analysis of the data and interpretation of results. Data analysis: many software's are now available to perform the mathematical part of the research process i.e. the calculations using various statistical methods. Softwares like SPSS and spreadsheets are the widely used. They can be like calculating the sample size for a proposed study, hypothesis testing and calculating the power of the study. Familiarity with any one package will suffice to carry out the most intricate statistical analysis. Computers are useful not only for statistical analysis, but also to monitor the accuracy and completeness of the data as they are collected. These software's also display the results in graphical char or graph form.

5. Role of Computer in Research Publication: This phase consists of preparation of the report or presentation of the results, i.e., formal write-up of conclusions reached. This is the research publication phase. The research article, research paper, research thesis or research dissertation is typed in word processing software and converted to portable data format (PDF) and stored and/or published in the world wide web. Online sites are available through we can convert our word file into any format like html, pdf etc. Various online applications are also available for this purpose. Even we can prepare our document using online word processing software and can store/edit/access it from anywhere using internet.

Issues and Challenges in practice of computing in social research

The practice of computing in social research has evolved rapidly. Computers have been applied to practically every research task, including such unlikely ones as interviewing, and hundreds of other tasks (Brent and Anderson 1990). This variety of computer applications will continue to evolve with the newer Internet-based technologies.

The application of computing to sociology is not without problems. Errors in data and software abound yet rarely do social scientists check their results by running more than one program on the same data. Data and software tend to be very costly, but there are many impediments to the sharing of these critical resources. Better software is needed but graduate students often are discouraged from

programming new software. Nonetheless, new breakthroughs in computer technology will continue, and major new opportunities will emerge. Many of the advances in sociological computing during the next few years undoubtedly will follow the lines of progress already described: hypertext networks; integrated, high performance, graphic data analysis stations; software for computer-supported cooperative work; and neural networks for complex models of social systems.

Perhaps the most exciting challenge for the future involves a concert of these innovations directed at the problem of modeling and analyzing vast amounts of social data. One solution would incorporate three-dimensional, multicolored, dynamic graphical representations of complex social data structures. But new techniques for analyzing these data will require new models of dynamic social structures as well as parallel social processes. Computer representations of these models tend to require extremely fast processing. Access to such models on the Web, supplemented with audio and video displays, may evolve into an important part of the sociologist's tool kit of the future.

ANALYSIS OF DATA

Definition - Data Analysis:

According to *LeCompte* and *Schensul*, **research data analysis is a process used by researchers for reducing data to a story and interpreting it to derive insights**. The data analysis process helps in reducing a large chunk of data into smaller fragments, which makes sense.

The systematic application of statistical and logical techniques to describe the data scope, modularize the data structure, condense the data representation, illustrate via images, tables, and graphs, and evaluate statistical inclinations, probability data, to derive meaningful conclusions, is known as Data Analysis.

Three essential things take place during the data analysis process-the first data organization. Summarization and categorization together contribute to becoming the second known method used for data reduction. It helps in finding patterns and themes in the data for easy identification and linking. Third and the last way is data analysis – researchers do it in both top-down or bottom-up fashion.

Marshall and Rossman, on the other hand, describe data analysis as a messy, ambiguous, and time-consuming, but a creative and fascinating process through which a mass of collected data is being brought to order, structure and meaning.

We can say that "the data analysis and interpretation is a process representing the application of deductive and inductive logic to the research and data analysis.

Why analyze data in research?

Researchers rely heavily on data as they have a story to tell or problems to solve. It starts with a question, and data is nothing but an answer to that question. But, what if there is no question to ask? Well! It is possible to explore data even without a problem – we call it *'Data Mining'* which often reveal some interesting patterns within the data that are worth exploring.

Irrelevant to the type of data, researchers explore, their mission, and audiences' vision guide them to find the patterns to shape the story they want to tell. One of the essential things expected from researchers while analyzing data is to stay open and remain unbiased towards unexpected patterns, expressions, and results. Remember, sometimes, data analysis tells the most unforeseen yet exciting stories that were not expected at the time of initiating data analysis. Therefore, rely on the data you have at hand and enjoy the journey of exploratory research.

Types of data in research

Every kind of data has a rare quality of describing things after assigning a specific value to it. For analysis, you need to organize these values, processed and presented in a given context, to make it useful. Data can be in different forms; here are the primary data types.

- **Qualitative data:** When the data presented has words and descriptions, then we call it qualitative data. Although you can observe this data, it is subjective and harder to analyze data in research, especially for comparison. **Example:** Quality data represents everything describing taste, experience, texture, or an opinion that is considered quality data. This type of data is usually collected through focus groups, personal interviews, or using open-ended questions in surveys.
- **Quantitative data:** Any data expressed in numbers of numerical figures are called quantitative data. This type of data can be distinguished into categories, grouped, measured, calculated, or ranked. **Example:** questions such as age, rank, cost, length, weight, scores, etc. everything

comes under this type of data. You can present such data in graphical format, charts, or apply statistical analysis methods to this data. The (Outcomes Measurement Systems) OMS questionnaires in surveys are a significant source of collecting numeric data.

- **Categorical data:** It is data presented in groups. However, an item included in the categorical data cannot belong to more than one group. **Example:** A person responding to a survey by telling his living style, marital status, smoking habit, or drinking habit comes under the categorical data. A chi-square test is a standard method used to analyze this data.

Data analysis in qualitative research

Data analysis and qualitative data research work a little differently from the numerical data as the quality data is made up of words, descriptions, images, objects, and sometimes symbols. Getting insight from such complicated information is a complicated process. Hence it is typically used for exploratory research and data analysis.

Finding patterns in the qualitative data

Although there are several ways to find patterns in the textual information, a word-based method is the most relied and widely used global technique for research and data analysis. Notably, the data analysis process in qualitative research is manual. Here the researchers usually read the available data and find repetitive or commonly used words.

For example, while studying data collected from African countries to understand the most pressing issues people face, researchers might find *"food"* and *"hunger"* are the most commonly used words and will highlight them for further analysis.

The keyword context is another widely used word-based technique. In this method, the researcher tries to understand the concept by analyzing the context in which the participants use a particular keyword.

For example, researchers conducting research and data analysis for studying the concept of *'diabetes'* amongst respondents might analyze the context of when and how the respondent has used or referred to the word 'diabetes.'

The scrutiny-based technique is also one of the highly recommended text analysis methods used to identify a quality data pattern. Compare and contrast is the widely used method under this technique to differentiate how a specific text is similar or different from each other.

For example: To find out the "importance of resident doctor in a company," the collected data is divided into people who think it is necessary to hire a resident doctor and those who think it is

unnecessary. Compare and contrast is the best method that can be used to analyze the polls having single answer questions types.

Metaphors can be used to reduce the data pile and find patterns in it so that it becomes easier to connect data with theory.

Variable Partitioning is another technique used to split variables so that researchers can find more coherent descriptions and explanations from the enormous data.

Data Analysis Techniques

There are different techniques for Data Analysis depending upon the question at hand, the type of data, and the amount of data gathered. Each focuses on strategies of taking onto the new data, mining insights, and drilling down into the information to transform facts and figures into decision making parameters. Accordingly, the different techniques of data analysis can be categorized as follows:

1. Techniques based on Mathematics and Statistics

Descriptive Analysis: Descriptive Analysis takes into account the historical data, Key Performance Indicators, and describes the performance based on a chosen benchmark. It takes into account past trends and how they might influence future performance.

Dispersion Analysis: Dispersion in the area onto which a data set is spread. This technique allows data analysts to determine the variability of the factors under study.

Regression Analysis: This technique works by modeling the relationship between a dependent variable and one or more independent variables. A regression model can be linear, multiple, logistic, ridge, non-linear, life data, and more.

Factor Analysis: This technique helps to determine if there exists any relationship between a set of variables. In this process, it reveals other factors or variables that describe the patterns in the relationship among the original variables. Factor Analysis leaps forward into useful clustering and classification procedures.

Discriminant Analysis: It is a classification technique in data mining. It identifies the different points on different groups based on variable measurements. In simple terms, it identifies what makes two groups different from one another; this helps to identify new items.

Time Series Analysis: In this kind of analysis, measurements are spanned across time, which gives us a collection of organized data known as time-series.

2. Techniques based on Artificial Intelligence and Machine Learning

Artificial Neural Networks: a Neural network is a biologically-inspired programming paradigm that presents a brain metaphor for processing information. An Artificial Neural Network is a system that changes its structure based on information that flows through the network. ANN can accept noisy data and are highly accurate. They can be considered highly dependable in business classification and forecasting applications.

Decision Trees: As the name stands, it is a tree-shaped model that represents a classification or regression models. It divides a data set in smaller subsets simultaneously developing into a related decision tree.

Evolutionary Programming: This technique combines the different types of data analysis using evolutionary algorithms. It is a domain-independent technique, which can explore ample search space and manages attribute interaction very efficiently.

Fuzzy Logic: It is a data analysis technique based on probability which helps in handling the uncertainties in data mining techniques.

3. Techniques based on Visualization and Graphs

Column Chart, Bar Chart: Both these charts are used to present numerical differences between categories. The column chart takes to the height of the columns to reflect the differences. Axes interchange in the case of the bar chart.

Line Chart: This chart is used to represent the change of data over a continuous interval of time.

Area Chart: This concept is based on the line chart. It additionally fills the area between the polyline and the axis with color, thus representing better trend information.

Pie Chart: It is used to represent the proportion of different classifications. It is only suitable for only one series of data. However, it can be made multi-layered to represent the proportion of data in different categories.

Funnel Chart: This chart represents the proportion of each stage and reflects the size of each module. It helps in comparing rankings.

Word Cloud Chart: It is a visual representation of text data. It requires a large amount of data, and the degree of discrimination needs to be high for users to perceive the most prominent one. It is not a very accurate analytical technique.

Gantt Chart: It shows the actual timing and the progress of activity in comparison to the requirements.

Radar Chart: It is used to compare multiple quantized charts. It represents which variables in the data have higher values and which have lower values. A radar chart is used for comparing classification and series along with proportional representation.

Scatter Plot: It shows the distribution of variables in the form of points over a rectangular coordinate system. The distribution in the data points can reveal the correlation between the variables.

Bubble Chart: It is a variation of the scatter plot. Here, in addition to the x and y coordinates, the area of the bubble represents the 3rd value.

Gauge: It is a kind of materialized chart. Here the scale represents the metric, and the pointer represents the dimension. It is a suitable technique to represent interval comparisons.

Frame Diagram: It is a visual representation of a hierarchy in the form of an inverted tree structure.

Rectangular Tree Diagram: This technique is used to represent hierarchical relationships but at the same level. It makes efficient use of space and represents the proportion represented by each rectangular area.

Map

Regional Map: It uses color to represent value distribution over a map partition.

Point Map: It represents the geographical distribution of data in the form of points on a geographical background. When the points are the same in size, it becomes meaningless for single data, but if the points are as a bubble, then it additionally represents the size of the data in each region.

Flow Map: It represents the relationship between an inflow area and an outflow area. It represents a line connecting the geometric centers of gravity of the spatial elements. The use of dynamic flow lines helps reduce visual clutter.

Heat Map: This represents the weight of each point in a geographic area. The color here represents the density.

DATA ANALYSIS TOOLS IN COMPUTER

There are several data analysis tools available in the market, each with its own set of functions. The selection of tools should always be based on the type of analysis performed, and the type of data worked. Here is a list of a few compelling tools for Data Analysis.

1. Excel

It has a variety of compelling features, and with additional plugins installed, it can handle a massive amount of data. So, if you have data that does not come near the significant data margin, then Excel can be a very versatile tool for data analysis.

2. Tableau

It falls under the BI Tool category, made for the sole purpose of data analysis. The essence of Tableau is the Pivot Table and Pivot Chart and works towards representing data in the most user-friendly way. It additionally has a data cleaning feature along with brilliant analytical functions.

3. Power BI

It initially started as a plugin for Excel, but later on, detached from it to develop in one of the most data analytics tools. It comes in three versions: Free, Pro, and Premium. Its PowerPivot and DAX language can implement sophisticated advanced analytics similar to writing Excel formulas.

4. Fine Report

Fine Report comes with a straightforward drag and drops operation, which helps to design various styles of reports and build a data decision analysis system. It can directly connect to all kinds of databases, and its format is similar to that of Excel. Additionally, it also provides a variety of dashboard templates and several self-developed visual plug-in libraries.

5. R & Python

These are programming languages which are very powerful and flexible. R is best at statistical analysis, such as normal distribution, cluster classification algorithms, and regression analysis. It also performs individual predictive analysis like customer behavior, his spend, items preferred by him based on his browsing history, and more. It also involves concepts of machine learning and artificial intelligence.

6. SAS

It is a programming language for data analytics and data manipulation, which can easily access data from any source. SAS has introduced a broad set of customer profiling products for web, social media, and marketing analytics. It can predict their behaviors, manage, and optimize communications.

7. SPSS

"Statistical Package for the Social Sciences - SPSS is a widely used program for statistical analysis in social science. It is also used by market researchers, health researchers, survey companies, government, education researchers, marketing organizations, data miners, and others.

INTERPRETATION OF RESULTS

Data interpretation refers to the implementation of processes through which data is reviewed for the purpose of arriving at an informed conclusion. The interpretation of data assigns a meaning to the information analyzed and determines its signification and implications.

MEANING OF INTERPRETATION

Interpretation refers to the task of drawing inferences from the collected facts after an analytical and /or experimental study. In fact, it is a search for broader meaning of research findings.

The task of interpretation has two major aspects viz.,

1. the effort to establish continuity in research through linking the results of a given study with those of another, and
2. The establishment of some explanatory concepts. "In one sense, interpretation is concerned with relationships within the collected data, partially overlapping analysis.
3. Interpretation also extends beyond the data of the study to include the results of other research, theory and hypotheses.

Thus, interpretation is the device through which the factors that seem to explain what has been observed by researcher in the course of the study can be better understood and it also provides a theoretical conception which can serve as a guide for further researches.

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 different types of processes that are implemented based on individual data nature, the two broadest and most common categories are “quantitative analysis” and “qualitative analysis”.

Yet, before any serious data interpretation inquiry can begin, it should be understood that visual presentations of data findings are irrelevant unless a sound decision is made regarding scales of measurement. Before any serious data analysis can begin, the scale of measurement must be decided for the data as this will have a long-term impact on data interpretation ROI. The varying scales include:

Nominal Scale: non-numeric categories that cannot be ranked or compared quantitatively. Variables are exclusive and exhaustive.

Ordinal Scale: exclusive categories that are exclusive and exhaustive but with a logical order. Quality ratings and agreement ratings are examples of ordinal scales (i.e., good, very good, fair, etc., OR agree, strongly agree, disagree, etc.).

Interval: a measurement scale where data is grouped into categories with orderly and equal distances between the categories. There is always an arbitrary zero point.

Ratio: contains features of all three.

For a more in-depth review of scales of measurement, read our article on data analysis questions. Once scales of measurement have been selected, it is time to select which of the two broad interpretation processes will best suit your data needs. Let’s take a closer look at those specific data interpretation methods and possible data interpretation problems.

How to Interpret Data?

Illustration of data interpretation on blackboard

When interpreting data, an analyst must try to discern the differences between correlation, causation and coincidences, as well as many other bias – 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.

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 (or methods) for interpreting data will provide your analyst teams a structure and consistent foundation. Indeed, if several departments have different approaches to interpret 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: with a qualitative and a quantitative analysis.

Qualitative Data Interpretation

Qualitative data analysis can be summed up in one word – categorical. With qualitative analysis, data is not described through numerical values or patterns, but through the use of descriptive context (i.e., text). Typically, narrative data is gathered by employing a wide variety of person-to-person techniques. These techniques include:

Observations: detailing behavioral patterns that occur within an observation group. These patterns could be the amount of time spent in an activity, the type of activity and the method of communication employed.

Documents: much like how patterns of behavior can be observed, different types of documentation resources can be coded and divided based on the type of material they contain.

Interviews: one of the best collection methods for narrative data. Enquiry responses can be grouped by theme, topic or category. The interview approach allows for highly-focused data segmentation.

A key difference between qualitative and quantitative analysis is clearly noticeable in the interpretation stage. Qualitative data, as it is widely open to interpretation, must be "coded" so as to facilitate the grouping and labeling of data into identifiable themes. As person-to-person data collection techniques can often result in disputes pertaining to proper analysis, qualitative data analysis is often summarized through three basic principles: notice things, collect things, think about things.

Quantitative Data Interpretation

If quantitative data interpretation could be summed up in one word (and it really can't) that word would be "numerical." There are few certainties when it comes to data analysis, but you can be sure that if the research you are engaging in has no numbers involved, it is not quantitative research.

Quantitative analysis refers to a set of processes by which numerical data is analyzed. More often than not, it involves the use of statistical modeling such as standard deviation, mean and median. Let's quickly review the most common statistical terms:

Mean: a mean represents a numerical average for a set of responses. When dealing with a data set (or multiple data sets), a mean will represent a central value of a specific set of numbers. It is the sum of the values divided by the number of values within the data set. Other terms that can be used to describe the concept are arithmetic mean, average and mathematical expectation.

Standard deviation: this is another statistical term commonly appearing in quantitative analysis. Standard deviation reveals the distribution of the responses around the mean. It describes the degree of consistency within the responses; together with the mean, it provides insight into data sets.

Frequency distribution: this is a measurement gauging the rate of a response appearance within a data set. When using a survey, for example, frequency distribution has the capability of determining the number of times a specific ordinal scale response appears (i.e., agree, strongly agree, disagree, etc.). Frequency distribution is extremely keen in determining the degree of consensus among data points.

Typically, quantitative data is measured by visually presenting correlation tests between two or more variables of significance. Different processes can be used together or separately, and comparisons can be made to ultimately arrive at a conclusion. Other signature interpretation processes of quantitative data include:

Regression analysis

Cohort analysis

Predictive and prescriptive analysis

Now that we have seen how to interpret data, let's move on and ask ourselves some questions: what are some data interpretation benefits? Why do all industries engage in data research and analysis? These are basic questions, but that often don't receive adequate attention.

Why Data Interpretation Is Important

Illustrating quantitative data interpretation with charts & graphs

The purpose of collection and interpretation is to acquire useful and usable information and to make the most informed decisions possible. From businesses, to newlyweds researching their first home, data collection and interpretation provides limitless benefits for a wide range of institutions and individuals.

Data analysis and interpretation, regardless of method and qualitative/quantitative status, may include the following characteristics:

1. Data identification and explanation
2. Comparing and contrasting of data
3. Identification of data outliers
4. Future predictions

Data analysis and interpretation, in the end, helps improve processes and identify problems. It is difficult to grow and make dependable improvements without, at the very least, minimal data collection and interpretation. What is the key word? Dependable. Vague ideas regarding performance enhancement exist within all institutions and industries. Yet, without proper research and analysis, an idea is likely to remain in a stagnant state forever (i.e., minimal growth). So… what are a few of the business benefits of digital age data analysis and interpretation? Let's take a look!

1) Informed decision-making: A decision is only as good as the knowledge that formed it. Informed data decision making has the potential to set industry leaders apart from the rest of the market pack. Studies have shown that companies in the top third of their industries are, on average, 5% more productive and 6% more profitable when implementing informed data decision-making processes. Most decisive actions will arise only after a problem has been identified or a goal defined. Data analysis should include identification, thesis development and data collection followed by data communication.

If institutions only follow that simple order, one that we should all be familiar with from grade school science fairs, then they will be able to solve issues as they emerge in real time. Informed decision making has a tendency to be cyclical. This means there is really no end, and eventually, new

questions and conditions arise within the process that need to be studied further. The monitoring of data results will inevitably return the process to the start with new data and sights.

2) Anticipating needs with trends identification: data insights provide knowledge, and knowledge is power. The insights obtained from market and consumer data analyses have the ability to set trends for peers within similar market segments. A perfect example of how data analysis can impact trend prediction can be evidenced in the music identification application, Shazam. The application allows users to upload an audio clip of a song they like, but can't seem to identify. Users make 15 million song identifications a day. With this data, Shazam has been instrumental in predicting future popular artists.

When industry trends are identified, they can then serve a greater industry purpose. For example, the insights from Shazam's monitoring benefits not only Shazam in understanding how to meet consumer needs, but it grants music executives and record label companies an insight into the pop-culture scene of the day. Data gathering and interpretation processes can allow for industry-wide climate prediction and result in greater revenue streams across the market. For this reason, all institutions should follow the basic data cycle of collection, interpretation, decision making and monitoring.

3) Cost efficiency: Proper implementation of data analysis processes can provide businesses with profound cost advantages within their industries. A recent data study performed by Deloitte vividly demonstrates this in finding that data analysis ROI is driven by efficient cost reductions. Often, this benefit is overlooked because making money is typically viewed as "sexier" than saving money. Yet, sound data analyses have the ability to alert management to cost-reduction opportunities without any significant exertion of effort on the part of human capital.

4) Clear foresight: companies that collect and analyze their data gain better knowledge about themselves, their processes and performance. They can identify performance challenges when they arise and take action to overcome them. Data interpretation through visual representations lets them process their findings faster and make better-informed decisions on the future of the company.

Data Interpretation Problems

The oft-repeated mantra of those who fear data advancements in the digital age is "big data equals big trouble." While that statement is not accurate, it is safe to say that certain data interpretation problems or "pitfalls" exist and can occur when analyzing data, especially at the speed of thought. Let's identify three of the most common data misinterpretation risks and shed some light on how they can be avoided:

1) Correlation mistaken for causation: our first misinterpretation of data refers to the tendency of data analysts to mix the cause of a phenomenon with correlation. It is the assumption that because two actions occurred together, one caused the other. This is not accurate as actions can occur together absent a cause and effect relationship.

Digital age example: assuming that increased revenue is the result of increased social media followers… there might a definitive correlation between the two, especially with today's multi-channel purchasing experiences. But, that does not mean an increase in followers is the direct cause of increased revenue. There could be either a common cause or an indirect causality.

Remedy: attempt to eliminate the variable you believe to be causing the phenomenon.

2) Confirmation bias: our second data interpretation problem occurs when you have a theory or hypothesis in mind, but are intent on only discovering data patterns that provide support, while rejecting those that do not.

Digital age example: your boss asks you to analyze the success of a recent multi-platform social media marketing campaign. While analyzing the potential data variables from the campaign (one that you ran and believe performed well), you see that the share rate for Facebook posts were great, while the share rate for Twitter Tweets were not. Using only the Facebook posts to prove your hypothesis that the campaign was successful would be a perfect manifestation of confirmation bias.

Remedy: as this pitfall is often based on subjective desires, one remedy would be to analyze data with a team of objective individuals. If this is not possible, another solution is to resist the urge to make a conclusion before data exploration has been completed. Remember to always try to disprove a hypothesis, not prove it.

3) Irrelevant data: the third and final data misinterpretation pitfall is especially important in the digital age. As large data is no longer centrally stored, and as it continues to be analyzed at the speed

of thought, it is inevitable that analysts will focus on data that is irrelevant to the problem they are trying to correct.

Digital age example: in attempting to gauge the success of an email lead generation campaign, you notice that the number of homepage views directly resulting from the campaign increased, but the number of monthly newsletter subscribers did not. Based on the number of homepage views, you decide the campaign was a success when really it generated zero leads.

Remedy: proactively and clearly frame any data analysis variables and KPIs prior to engaging in a data review. If the metric you are using to measure the success of a lead generation campaign is newsletter subscribers, there is no need to review the number of homepage visits. Be sure to focus on the data variable that answers your question or solves your problem and not on irrelevant data.

Interpretation of Data: The Use of Dashboards Bridging the Gap

As we have seen, quantitative and qualitative methods are distinct types of data analyses. Both offer a varying degree of return on investment (ROI) regarding data investigation, testing and decision-making. Because of their differences, it is important to understand how dashboards can be implemented to bridge the quantitative and qualitative information gap. How are digital data dashboard solutions playing a key role in merging the data disconnect? Here are a few of the ways:

1) Connecting and blending data. With today's pace of innovation, it is no longer feasible (nor desirable) to have bulk data centrally located. As businesses continue to globalize and borders continue to dissolve, it will become increasingly important for businesses to possess the capability to run diverse data analyses absent the limitations of location. Data dashboards decentralize data without compromising on the necessary speed of thought while blending both quantitative and qualitative data. Whether you want to measure customer trends or organizational performance, you now have the capability to do both without the need for a singular selection.

2) Mobile Data. Related to the notion of "connected and blended data" is that of mobile data. In today's digital world, employees are spending less time at their desks and simultaneously increasing production. This is made possible by the fact that mobile solutions for analytical tools are no longer standalone. Today, mobile analysis applications seamlessly integrate with everyday business tools. In

turn, both quantitative and qualitative data are now available on demand where they're needed, when they're needed and how they're needed.

3) Visualization. Data dashboards are merging the data gap between qualitative and quantitative methods of interpretation of data, through the science of visualization. Dashboard solutions come "out of the box" well-equipped to create easy-to-understand data demonstrations. Modern online data visualization tools provide a variety of color and filter patterns, encourage user interaction and are engineered to help enhance future trend predictability. All of these visual characteristics make for an easy transition among data methods – you only need to find the right types of data visualization to tell your data story the best way possible.

To give you an idea of how a market research dashboard fulfils the need of bridging quantitative and qualitative analysis, and helps in understanding how to interpret data in research thanks to visualization, have a look at the following one. It brings together both qualitative and quantitative data knowledgeably analyzed and visualizes it in a meaningful way that everyone can understand, thus empowering any viewer to interpret it:

To see more data analysis and interpretation examples, visit our library of business dashboards. Now that we have an understanding of how business intelligence dashboards can help bridge the gap between traditional quantitative and qualitative data methods, let's have a little summary of the data interpretation methods mentioned all along this article.

Data Interpretation Methods Summary List & Tips

Data analysis and interpretation are critical to develop sound conclusions and make better informed decisions. As we have seen all along this article, there is an art and science to the interpretation of data. Hereafter is a list-summary of how to interpret data and some tips:

1. Collect your data and make it as clean as possible.
2. Choose the type of analysis to perform: qualitative or quantitative, and apply the methods respectively to each.
3. Qualitative analysis: observe, document and interview notice, collect and think about things.

4. Quantitative analysis: you lead a research with a lot of numerical data to be analyzed through various statistical methods – mean, standard deviation or frequency distribution for instance.
5. Take a step back: and think about your data from various perspectives, and what it means for various participants or actors of the project.
6. Reflect on your own thinking and reasoning: and be aware of the many pitfalls data analysis and interpretation carries. Correlation versus causation, subjective bias, false information and inaccurate data, etc.

Key Terms to Remember

Coding - The process of assigning values, typically numeric values, to the different levels of a variable.

Data Analysis -The process by which data are organized to better understand patterns of behavior within the target population.

Data Collection - The observation, measurement, and recording of information in a research study.

Data Imputation - A method used to fill in missing values (due to nonresponsive) in surveys.

Data Reduction - The basic concept is the reduction of large amounts of data down to the meaningful parts.

Deduction - The process of reasoning from the more general to the more specific.

Important Questions

1. Describe, in brief, the importance of editing, coding, classification, tabulation and presentation of data in the context of research study.
2. Discuss the different aspects of classification of data.
3. What are the likely problems encountered in the classification and how they can be handled?
4. Why tabulation is considered essential in a research study?
5. Give the characteristics of a good table.
6. Write briefly about the different forms of data presentation devices.
7. Discuss the fundamental rules of code construction.
8. Discuss with the help of suitable examples various steps involved in data processing.

MCQ – Questions

1. What is a hypothesis?

a) A statement that the researcher wants to test through the data

b) collected in a study.

c) A research question the results will answer.

d) A theory that underpins the study.

e) A statistical method for calculating the extent to which the resultscould have happened by chance.

2. Qualitative data analysis is still a relatively new and rapidly developing branch of research methodology.

a) True

b) False

3.. The process of marking segments of data with symbols,descriptive words, or category names is known as _____.

a) Concurring

b) Coding

c) Colouring

d) Segmenting

4. What is the cyclical process of collecting and analysing data during a single research study called?

   a) Interim analysis

   b) Inter analysis

   c) Inter-item analysis

   d) Constant analysis

5. The process of quantifying data is referred to as _____.

   a) Typology

   b) Diagramming

   c) Enumeration

   d) Coding

6. An advantage of using computer programs for qualitative data is that they _____.

   a) Can reduce time required to analyse data (i.e., after the data are transcribed)

   b) Help in storing and organising data

   c) Make many procedures available that are rarely done by hand due to time constraints

   d) All of the above

7. Boolean operators are words that are used to create logical combinations.

   a) True

   b) False

8. _____ are the basic building blocks of qualitative data.

   a) Categories

   b) Units

   c) Individuals

   d) None of the above

9. This is the process of transforming qualitative research data from written interviews or field notes into typed text.

   a) Segmenting

   b) Coding

   c) Transcription

   d) Mnemoning

10. A challenge of qualitative data analysis is that it often includes data that are unwieldy and complex; it is a major challenge to make sense of the large pool of data.
    a) True
    b) False
11. Hypothesis testing and estimation are both types of descriptive statistics.
    a) True
    b) False
12. A set of data organised in a participants(rows)-byvariables(columns) format is known as a "data set."
    a) True
    b) False
13. A graph that uses vertical bars to represent data is called a ___
    a) Line graph
    b) Bar graph
    c) Scatterplot
    d) Vertical graph
14. _____ are used when you want to visually examine the relationship between two quantitative variables.
    a) Bar graphs
    b) Pie graphs
    c) Line graphs
    d) Scatterplots
15. The denominator (bottom) of the z-score formula is
    a) The standard deviation
    b) The difference between a score and the mean
    c) The range
    d) The mean

16. Which of these distributions is used for a testing hypothesis?

a) Normal Distribution

b) Chi-Squared Distribution

c) Gamma Distribution

d) Poisson Distribution

17. A statement made about a population for testing purpose is called?

a) Statistic

b) Hypothesis

c) Level of Significance

d) Test-Statistic

18. If the assumed hypothesis is tested for rejection considering it to be true is called?

a) Null Hypothesis

b) Statistical Hypothesis

c) Simple Hypothesis

d) Composite Hypothesis

19. If the null hypothesis is false then which of the following is accepted?

a. Null Hypothesis

b. Positive Hypothesis

c. Negative Hypothesis

d. Alternative Hypothesis.

20. Alternative Hypothesis is also called as?

a) Composite hypothesis

b) Research Hypothesis

c) Simple Hypothesis

d) Null Hypothesis