DATA PROCESSING

Data processing refers to certain operations such as editing, coding, computing of the scores, preparation of master charts, etc. A researcher has to make his plan for each and every stage of the research process. As such, a good researcher makes a perfect plan of processing and analysis of data. To some researchers data processing and analysis is not a very serious activity. They feel many times that data processing is a job of computer assistants. As a 1 assistants which may not help them to achieve their objectives. To avoid such I situations, it is essential that data processing must be planned in advance and 1 instructed to assistants accordingly.

DATA PROCESSING AND ANALYSIS

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 diagrams1 2 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.1 3 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.2 4 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.2 5 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.3 6 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.3 7 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. 8 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.4 9 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.