**PGD001 – Postgraduate Diploma in Monitoring and Evaluation**

**Name: Wycliffe Otwori Nyakundi**

**Date of submission: 15th January 2020**

**Assignment Number: Module 5**

**Module 5 Questions:**

**Q1.** Explain the difference between data collection and data capture (10mrks)

Data collection is the process of gathering and measuring raw information (data) on variables of interest, in an established systematic fashion that enables determination of stated project performance or research questions, test hypotheses, and or evaluate outcomes. It uses specific tools for collection purposes. For efficient data collection and depending on the discipline or field, the nature of the information being sought, and the objective or goal of users, the methods of data collection will vary. However, the data collection tools should have the following characteristics:

* They should be as simple and as clear as possible
* They should be pre-tested to ensure they are user-friendly
* They should have and stick to a minimum set of questions needed to measure indicators

On the other hand, data capture is a process of keying in information into an M&E system, designed specifically for the purpose of data collection and eventual analysis. This can also include a process of converting data from forms into a format that can be interpreted and analyzed.

Thera are many multiple methods which can be used for capturing data from unstructured documents (letters, invoices, email, fax, forms etc.)! to a consumable state and while this is done, it is advisable in the first instance to consider the original documents, to determine if the document or form can be updated to improve the capture/recognition process and method. Data capture tools[[1]](#footnote-1) should:

* provide ways to organise and structure data files
* have data validation components which ensure captured data meet required types and ranges etc.
* enable open and flexible formats where good conversion tools exist. Proprietary formats are OK if they are well documented and an API or conversion tool exists
* allow data to be moved to its destination efficiently and with high quality.

**Q 2:** Explain the benefits of correctly interpreting data in an M&E process. (5 mrks)

* Informed decision-making: A decision is only as good as the knowledge that formed it. Informed data decision are often arrived at, based on good interpretation which must be approached without personal bias or pre-conceived opinions. By interpreting data objectively, the correct conclusion is reached.
* Anticipating needs with trends identification and future predictions: Data insights provide knowledge, and knowledge is power. The insights obtained have the ability to set trends for future program/project implementation in similar circumstances if they are properly analysed and interpreted.
* Cost efficiency: Proper data interpretation improves the M&E processes by efficient cost reductions. This will be possible in situations where, there are informed decisions that results to better approaches which are cost-effective in implementation
* Clear foresight: organizations that interpret data correctly 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 like graphs and pie-charts etc is faster.

**Q3.** Explain the main concerns for a data analyst while undertaking the task of data analysis. (10 mrks)

Data in and of itself will not provide any meaning unless it can be delivered in a proper way. It is therefore important that the data analyst have the following considerations before undertaking his/her tasks;

* + - * 1. Meaningfulness of data: The data analyst should ask if, the data is data meaningful. Data analysis starts with collecting the right data to analyze. The data should pertain to the goals and objectives of the analysis. No amount of statistical analysis, regardless of the level of the sophistication, will correct poorly defined objective outcome measurements. Whether done unintentionally or by design, this practice increases the likelihood of clouding the interpretation of findings, thus potentially misleading readers.
        2. Data collection and recording method: The data analyst should consider the method in which data was collected and recorded. For example, data documented by recording audio and/or video and transcribing later or where participants themselves are requested to take notes, compile and submit them to researchers, there could be issues of objectivity and subjectivity may be raised when data is analyzed.
        3. Reliability and Validity: Researchers performing analysis on either quantitative or qualitative analyses should be aware of challenges to reliability and validity. For example, in the area of content analysis, Gottschalk (1995) identifies three factors that can affect the reliability of analyzed data:
* stability , or the tendency for coders to consistently re-code the same data in the same way over a period of time
* reproducibility , or the tendency for a group of coders to classify categories membership in the same way
* accuracy , or the extent to which the classification of a text corresponds to a standard or norm statistically

The potential for compromising data integrity arises when researchers cannot consistently demonstrate stability, reproducibility, or accuracy of data analysis

* + - * 1. Measurability of data: The analyst should consider if the data measurable. Actually, the first step to success is defining an objective. Data analysis requires objective measurable facts hence, without concrete measurable data the analyst will not be able to see whether success is achievable. Thus there is need to make sure the data can be defined and quantified.

Other questions that the analyst looks at are:

* + - * 1. Is the data transformable? The data analyst needs to be fluent in the important tools of the information age. Proper tools will allow the analyst to sift through data quickly and achieve the desirable results. Proper data transformation will lead to meaningful information for the analyst's audience.
        2. Is the data beneficial? This is probably the most important question to ask in data analysis. In other words, is the data analysis presenting itself in a meaningful way to its intended audience? Remember that data is only data until it becomes information.

**Q4.** Describe key measures that are mandatory for data quality assurance at program level and explain the value of data quality assurance. (15 mrks).

The key measures that are mandatory for data quality assurance at program level are;

* Increasing the HR capacity both at Coordination and field levels, for the M&E functions. This will provide a conducive environment to ensure all processes in regard to data quality are addressed timely and effectively without compromise
* Strengthen national mechanisms on data quality through supportive supervision. The national mechanisms aim to achieve coordination, responsibility and accountability at national level. Further, they allow creation and harmonization of tools that are used in programs
* Establish an electronic/ web based data capturing, reporting and management system that will help to minimize on data errors. When properly used and developed, electronic databases will always minimize data errors e.g duplication, missing data etc thus improving data quality within organizations
* Periodic reviews and revision of data collection and reporting tools at all levels, regular updates to review and enhance tools while re-orientating staff on them. It provides not only more knowledge sharing but also reduces errors and program mistakes due to increased understanding and revision of tools.
* Provide training and mentorship in Monitoring and evaluation including regular updates focusing on data collection, analysis and use of data to field staff to improve their capacity in data quality assurance.
* Provide technical support to assist field staff develop good data storage at their level and at all service delivery points. This happens during periodic assessment and as on-going support according to the needs of field staff.

Equally, the value of data quality assurance can not be overemphasized since it is now becoming clear that very few organizations can be able to keep accurate information about their clients and projects. There are five components[[2]](#footnote-2) that will ensure data quality; completeness, consistency, accuracy, validity, and timeliness. When each of these components are properly executed, it will result in high-quality data. And thus, proper data quality assurance will enhance the component and thereby ensure more efficiency in driving an organization’s success because of the dependence on fact-based decisions, instead of habitual or human intuition. More specifically the value of data quality assurance include;

* Regular reviews of data quality built into a system of checks for an organization/ programme reporting systems as part of a feedback cycle that identifies errors and thus provide ways of developing reliable and better systems and tools;
* Data quality assurance focuses on specific core set of tracer indicators in order to identify gaps and errors in reporting and the plausibility of trends in reported data, thus able to refine and provide more details on indicators for improved programming
* Through data quality assurance, analysed results can be used to provide better and working strategies and approaches for an institution including those that can enhance accountability and reduces operation costs
* Data quality assurance results to knowledge and experience sharing which can provide better project implementation approaches thus improved, timely and efficiency in activity implementation

**Q5:** In about 350 words, describe the main challenges to effective data interpretation and analysis. (10 mrks)

Since data analysis and interpretation depends on many factors, it is highly notable that it is associated with many challenges.

**Some analysis and interpretation always have correlation mistaken for causation. This happens when** 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. When interpreting data, an analyst should therefore try to discern the differences between correlation, causation and coincidences, as well as many other bias – but s/he also has to consider other factors for the results. In addition, data tends to be extremely subjective depending on organizations however much the nature and goal of interpretation is likely to correlate with the type of data being analyzed.

Also and common is the realization that the interpretation of data assigns a meaning to the information analyzed and determines its significance and effects. This is why it data interpretation should be done properly with minimal biases and objectivity of the assignment at hand.

More often, data arrive from multiple sources and has a tendency to enter the analysis process with haphazard ordering. At the same time the different types of processes that are implemented will affect the methodology of analysis i.e whether “quantitative analysis” or “qualitative analysis”. The analysis method hereafter influences the method of interpretation; for example qualitative data is widely open to interpretation and should be “coded” so as to facilitate the grouping and labeling of data into identifiable themes

Again, before any serious data interpretation inquiry can begin, it should be understood that visual presentations of data findings can be irrelevant unless a sound decision is made regarding scales of measurement. The scales of measurement of data i.e nominal, ordinal and ration must be thought through to effectively determine the analysis and interpretation, which has a long-term impact. It is also critical to ensure that we have a baseline method (or methods) and approaches for interpreting data which will provide structure and consistent foundation. Lastly, since large data is no longer centrally stored, based on many digital platforms, it becomes very difficult for analysts to differentiate relevant from irrelevant data for the problems that they are trying to correct.

**References**

1. Gottschalk, L. A. (1995). Content analysis of verbal behavior: New findings and clinical applications. Hillside, NJ: Lawrence Erlbaum Associates, Inc
2. Shamoo, A.E. (1989). Principles of Research Data Audit. Gordon and Breach, New York.
3. Smeeton, N., Goda, D. (2003). Conducting and presenting social work research: some basic statistical considerations. Br J Soc Work, 33: 567-573.

1. <https://www.ands.org.au/working-with-data/data-management/data-capture> [↑](#footnote-ref-1)
2. <https://cerasis.com/data-quality/> [↑](#footnote-ref-2)