Africa Center
For Project Management

School of online and distance learning

COURSE: PGD IN MONITORING AND EVALUATION

MODULE FIVE ASSESSMENT TEST

PRESENTED BY:

MONOJA ANTHONY EZIBON

REG.

NO: ACPM/PGD/259/2019

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

Indeed, there is a very wide difference between the two terminologies as explained in the following paragraphs: -

Data collection is any process whose purpose is to acquire or assist in the acquisition of data. Collection is achieved by requesting and obtaining pertinent data from individuals or organizations via an appropriate vehicle. The data is either provided directly by the respondent (self-enumeration) or via an interviewer. Collection also includes the extraction of information from administrative sources which may require asking the respondent permission to link to administrative records whereas,

Data capture refers to any process that converts the information provided by a respondent into electronic format. This conversion is either automated or involves staff keying the collected data. Data coding is any process that assigns a numerical value to a response. Coding is often automated; however, more complex decisions usually require human intervention (coders).

Data collection is not only the source of information, it is also the main contact a survey-taking agency has with the public who needs to be convinced to participate. Data capture and coding produce the formatted data used as input by all the subsequent survey processes. Data collection, data capture and coding operations often use a large portion of the survey budget and require considerable human and physical resources as well as time. (*Bethlehem, J., F. Cobben, B. Schouten. 2008.)*

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

In the field of Monitoring and Evaluation, there is need to have a clear plan for data analysis and it should be prepared in advance before the actual data collection process. It should be guided by the needs of the different users of the M&E findings and should account for the time frame and characteristics of the data at hand, the methods or data structure required by the statistical tools, necessary statistical tools or templates to achieve stated objectives, manpower responsible for carrying out the analysis and interpretation of the results of study and purpose of the data analysis. (*Kothari C. (2004*))

Once the data has been gathered, it remains raw data and thus still remain meaningless unless certain statistical treatment is given to them so that the same data can be read easily and can be used for further analysis. Data analysis serves the following benefits: -

- To make the raw data meaningful

- To test null hypothesis
- To obtain the significant results
- To draw some inferences or make generalizations and
- To estimate parameters

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

There are several concerns that a data analyst must consider while undertaking data analysis. These concerns once addressed may improve decision-making, increase accountability, benefit financial health and help project the project performance. *Rebecca Webb, (2018)* in one of her article published on Risk Management Blog argued the following: -

The amount of data being collected: With today's data-driven organizations and the introduction of big data, M&E practitioners are often overwhelmed with the amount of data that is collected. An organization may receive information on every incident and interaction that takes place on a daily basis, leaving analysts with thousands of interlocking data sets.

This therefore asserts the need for a data system that automatically collects and organizes information. Manually performing this process is far too time-consuming and unnecessary in today's environment.

Collecting meaningful and real-time data: With so much data available, it's difficult to dig down and access the insights that are needed most. When employees are overwhelmed, they may not fully analyze data or only focus on the measures that are easiest to collect instead of those that truly add value. In addition, if an employee has to manually sift through data, it can be impossible to gain real-time insights on what is currently happening. Outdated data can have significant negative impacts on decision-making.

Visual representation of data: To be understood and impactful, data often needs to be visually presented in graphs or charts. While these tools are incredibly useful, it's difficult to build them manually. Taking the time to pull information from multiple areas and put it into a reporting tool is frustrating and time-consuming.

Data from multiple sources: The next issue is trying to analyze data across multiple, disjointed sources. Different pieces of data are often housed in different systems. Employees may not always realize this, leading to incomplete or inaccurate analysis. Manually combining data is time-consuming and can limit insights to what is easily viewed.

Inaccessible data: Moving data into one centralized system has little impact if it is not easily accessible to the people that need it. Decision-makers and risk managers need access to all of an organization's data for insights on what is happening at any given moment, even if they are working off-site. Accessing information should be the easiest part of data analytics.

Poor quality data: Nothing is more harmful to data analytics than inaccurate data. Without good input, output will be unreliable. A key cause of inaccurate data is manual errors made during data entry. This can lead to significant negative consequences if the analysis is used to influence decisions. Another issue is asymmetrical data: when information in one system does not reflect the changes made in another system, leaving it outdated.

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

Data quality refers is the reliability and effectiveness of data. It is the quality assurance angle of data (DQA) and is the process of verifying the reliability and effectiveness of data. Maintaining data quality requires going through the data periodically and cleaning it in accordance with the principles of accuracy, consistency, completeness and timeliness. At the program level, there are a number of key measures that guarantee data quality: -

- Increasing the Human Resource capacity both at coordination and field levels for the Monitoring and Evaluation functions
- Strengthen national mechanisms on data quality through supportive supervision
- Establish an electronic/ web-based data capturing, reporting and management system that will help to minimize on data errors.
- 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.
- 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.

Results based development programming requires managers to design and implement programs based on evidence. Since data play a central role in establishing effective performance management systems, it is essential to ensure good data quality. Without this, decision makers do not know whether to have confidence in the data, or worse, could make decisions based on misleading data. Data Quality Assurance (DQA) is therefore important in the following ways:-

- Ensuring that limited development resources are used as effectively as possible
- Ensuring that Agency program and budget decisions in Washington and the field are as well informed as practically possible
- Meeting the requirements of the Government Performance and Results Act (GPRA)
- Reporting the impact of USAID programs to external stakeholders, including senior management, OMB, the Congress, and the public with confidence (*USAID 2009, Number 12, 2nd Edition)*

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

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. According to *Mona Lebied in Data Analysis, 2018*, the following challenges were argued out: -

1. **Correlation mistaken for causation:** Here, 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. With the digital age as an example, if 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 both a common cause or an indirect causality.
2. **Confirmation bias:** This 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. For 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.

3. **Irrelevant data:** 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. For 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.

References:

1. *Bethlehem, J., F. Cobben, B. Schouten. 2008. "Indicators for the Representativeness of Survey Response." Proceedings from the 2008 International Symposium on Methodological Issues, Statistics Canada.*
2. *Kothari C (2004)., "Research Methodology – Methods and Techniques 2nd Edition, New Age International Publishers Limited, New Delhi".*
3. *USAID 2009, Number 12, 2nd Edition: "Performance Monitoring & Evaluation Tips, Data Quality Standards")*
4. *Mona Lebied in Data Analysis., 2018. "A Guide to The Methods, Benefits & Problems of The Interpretation of Data"*