AGRICULTURAL TVET COLLEGE

SMALL SCALE IRRIGATION DEVELOPMENT LEVEL-III

MODEL TTLM

Learning Guide- 05

Unit of competence: Analyze and interpret irrigation related data

Module title: Analyzing and interpreting irrigation related data

LG code: AGR SSI3 M05 LO1-LO3

TTLM Code: AGR SSI3 TTLM 1218V1

Nominal Duration: 40 Hours

Instruction sheet	Learning guide 05

This learning guide is developed to provide you the necessary information regarding the following content coverage and topics –

- ✓ Identify and categorize irrigation data
- ✓ Analyze and interpret data with descriptive statistics
- ✓ Present data

Thisguidewillalsoassistyoutoattainthelearningoutcomestatedinthecoverpage.specifically, upon completion of this learning guide, you will be able to: –

- ✓ Develop suitable formats
- ✓ Identify, organize & categorize data for analysis & interpretation
- ✓ Assess data held by the unit
- ✓ Realize methods of data collecting
- ✓ Update, modify, maintain & store data
- ✓ Define & consistency objectives of analysis
- ✓ Realize methods of data analysis
- ✓ Clear, justify & consistency assumptions
- ✓ Support conclusions & contribute to business objectives
- ✓ Use descriptive statistics
- ✓ Clear & conform structure & format of reports
- ✓ Report findings
- ✓ Obtain feedback & comments

Learning activities

1. Read the specific objectives of this learning guide.
2. Read the information written in the “information sheets.
3. Accomplish the “self-check” at the end of each learning outcomes.
4. If you earned as atisfactory evaluation proceed to the next “information sheet” .however, if your rating is unsatisfactory, see your teacher for further instructions or go back to the learning activity.
5. Submit your accomplished self-check. This will form part of your training portfolio

Informationsheet-1	Identify and categorize irrigation data

Introduction

What are data, information, knowledge and wisdom?

Data: - is any information collected as part of a research proposal and expressed as numbers. In practice, most measurements are classified into qualitative or quantitative data.

Data refers to raw material such as facts and figures that could be collected by an information system. Data are raw facts, neutral and context free, internal meaning is irrelevant.

Information: - refers to analyzed data, often presented in a form that is specifically designed for a given decision-making task, and transmitted to/received by decision makers.

Is the pattern people impose on the phenomena they are able to observe? Much of the information needed to solve farmers' problems with their decision making will come from research, although some will come from others farmers and from policy makers. Is data refined for some purpose or that have been given some degree of interpretation, often costly to produce but easy to add value to through processing.

Knowledge: - refers to the subsequent absorption, assimilation, understanding and appreciation of that information is information to which value has been added by interpretation based on a particular context, experience, and purpose; it is often acquired over substantial periods and involvement in many projects?

Wisdom– used in the context of decisions made.

1.1 Developing suitable formats

- **Data types & file formats**

After defining what we mean by data, it is helpful to consider what types of data you create and/or work with, and what format those data take. Your data stewardship practices will be dictated by the types of data that you work with, and what format they are in.

- **Data Formats**

Research data comes in many varied formats: text, numeric, multimedia, models, software languages, discipline specific (e.g. crystallographic information file (CIF) in chemistry), and instrument specific.

Formats more likely to be accessible in the future are:
- Non-proprietary
- Open, documented standards
- In common usage by the research community
- Using standard character encodings (ASCII, UTF-8)
- Uncompressed (desirable, space permitting)

Ideas expressed in paragraphs, letters, memos, or reports should be organized in some logical, efficient order. Collected data has to be organized in a suitable form to understand the information gathered (organization of data). There are two possible methods format: **inductive** and **deductive.**

A. Inductive (indirect). This ways supplies examples, facts, or reasons first and then draws conclusions from them or from specific to general. Inductive organization is useful when readers are uninformed or when resistance or antagonism is expected. For example, a report written to convince management to fund an employee fitness program might begin with the advantages of a fitness program: improved job satisfaction, reduced absenteeism and turnover, improved productivity, and lower health care costs. After describing the benefits, the report writer could draw the conclusion that a company-sponsored fitness program is a wise investment. Starting with the main idea first risks the chance that readers opposed to the idea will read no further. Persuasive memos, letters, and reports often follow the indirect plan.

B. Deductive (direct). This organizational plan presents the main idea or conclusions and recommendations first. Examples, reasons, and clarification follow or from general to specific. Most business writing is deductive because this method presents information clearly and openly. Use this plan for routine messages, such as those that convey favorable or neutral information. For example, to inform students of campus parking regulations, a straightforward announcement should be made. But if students must be persuaded to pay an extra fee for parking in preferred

locations, a letter describing the proposal might be written inductively with the assessing the spatial distribution of crop production using a generalized cross-entropy approach

1.2.Identifying, organizing & categorizing data for analysis & interpretation

Here are four other possible plans plus additional information on the inductive and deductive methods of organizing.

I. **Chronological: -** Information presented in time sequence, either forward or backward, is chronological. This organization pattern works well in explaining events over time such as monthly sales figures for the past year. Chronological ordering is also desirable to describe history or development, such as background leading to a personnel/management disagreement. Chronological sequencing of ideas is necessary to show time relationships. However, it is overused by individuals who sequence their writing to reflect their thinking about a problem. Whatever was thought about first is written about first. Such organization prevents flexibility in the writing process and ignores the techniques of emphasis. An important thought may be submerged in a pool of chronologically listed ideas.

II. **Geographical/spatial: -** If you were planning a vacation trip touring the east show zones, you might work out the trip in a geographical order. Within a large business building, custodial services might be assigned according to location—first floor, second floor, and so forth. Another form of geographical organization is illustrated by the division of a business into sales by state, county, or city. Geographical sequencing of ideas is less usable than other methods because of its rigidity; moreover, relatively few topics lend themselves to such organization.

III. **Value/size: -** The logical order for some topics begins with the most valuable or the largest item first. E.g. A report on entrepreneur, might discuss its largest-volume franchise first, followed in descending order by others. Realtors have found that listing their properties from the most expensive to the least expensive (or vice-versa) is helpful to buyers and sellers. Imagine how frustrating it would be to find properties in the same price range if they were listed chronologically.

IV. **Simple to complex: -** For difficult, technical, or abstract topics, the best plan of organization is often from simple to complex. Good teachers, for example, begin with simple, basic concepts and proceed to more complex topics. More complex ideas can follow the foundation built on simple concepts. Whenever readers or listeners are unfamiliar with a subject or when the topic is theoretical, simple-to-complex organization is effective.

- **Types of data**
  - A. quantitative
  - B. qualitative

A. **Quantitative data: -** when the variables being studied in a numerical form and it can be ranked or ordered. So it said to be quantitative data.

B. **Qualitative data: -**when variables being studied in a non-numeric form it is called qualitative data. Examples -color, religious preferences sex and, etc

1.3 Assessing the quality, accuracy and relevance of data

Accuracy: Data should be sufficiently accurate for the intended use and should be captured only once, although it may have multiple uses. Data should be captured at the point of activity.

Data is always captured at the point of activity. Performance data is directly input into Performance Plus[1] (P+) by the service manager or nominated data entry staff. Access to P+ for the purpose of data entry is restricted through secure password controls and limited access to appropriate data entry pages. Individual passwords can be changed by the user and which under no circumstances should be used by anyone other than that user. Where appropriate, base data, i.e. denominators and numerators, will be input into the system which will then calculate the result. These have been determined in accordance with published guidance or agreed locally. This will eliminate calculation errors at this stage of the process, as well as provide contextual information for the reader. Data used for multiple purposes, such as population and number of households, is input once by the system administrator.

Relevance: Data captured should be relevant to the purposes for which it is to be used. This will require a periodic review of requirements to reflect changing needs. We have a duty to collect and report performance information against a wide range of statutory indicators. These are set

out in the context of the Government's White Paper – Strong and Prosperous Communities. Where appropriate each service will identify reliable local performance indicators to manage performance and drive improvement. These are reviewed on an annual basis to ensure relevance.

Validity: Data should be recorded and used in compliance with relevant requirements, including the correct application of any rules or definitions. This will ensure consistency between periods and with similar organizations, measuring what is intended to be measured. Relevant guidance and definitions are provided for all statutory performance indicators. Service Heads are informed of any revisions and amendments within 24 hours of receipt from the relevant government department. Local performance indicators comply with locally agreed guidance and definitions

Timeliness: Data should be captured as quickly as possible after the event or activity and must be available for the intended use within a reasonable time period. Data must be available quickly and frequently enough to support information needs and to influence service or management decisions.

Performance data is requested to be available within one calendar month from the end of the previous quarter and is subsequently reported to the respective Policy and Scrutiny Panel on a quarterly basis. As a part of the ongoing development of Performance Plus it is intended that performance information will be exported through custom reporting and made available via the Three Rivers DC website. This will improve access to information and eliminate delays in publishing information through traditional methods.

Completeness: - Data requirements should be clearly specified based on the information needs of the organization and data collection processes matched to these requirements. Checks will be made to ensure for completeness of data. An annual assessment of this is undertaken by Internal Audit.

Characteristic of information

1) Relevance: the information that a person uses must be relevant to the current problem at hand. If the information is not relevant it would not help to solve the problem at all.
2) Completeness: good agricultural information is complete. This means that it provides intended users with all the information that is necessary to fulfill their information needs and requirements. Completeness also suggests that all necessary information is included in any report that the organization produces. The assumption is that there would be no error of

omission in the information. If the information is not complete and only partial information is received, it may lead to wrong decisions being made as only half of an entirety of the information is known.

3) Accuracy: - the information must be accurate as inaccurate information would lead to many different major problems. It leads individuals to wrong decisions. Agricultural information should be accurate. This does not suggest that you must always state figures and facts down to the last penny or detail. What it means is that information should be accurate enough for its intended purpose (or user), without being unnecessarily detailed. Inaccurate information cannot provide a valid representation of reality and can limit the effectiveness or worth of decisions based on it.
4) Timely (current):- the information must be current as a fact of yesterday may not be a fact of today. For example, the price of wheat per kilo today may not be the same with tomorrows. As the stock market is ever changing, current information is really a must. Since information has an objective, there are usually periods within which these objectives operate. Good information neither is produced too frequently nor is it compiled after it is needed most. For instance, information that reaches a decision-maker after the decision is of limited use in the context of the decision-making process.
5) Economical: - the information used must be of economical use to farmers, different organizations, companies that use them. If the information used is not economical it would result in users making losses. That is why economical information is important.

1.4 Realizing methods of data collecting

Reliability: Data should reflect stable and consistent data collection processes across collection points and over time. Progress toward performance targets should reflect real changes rather than variations in data collection approaches or methods. Source data is clearly identified and readily available from manual, automated or other systems and records. Protocols exist where data is provided from a third party, such as Hertfordshire Constabulary and Hertfordshire County Council

When you collect data, you will need to keep these key issues in mind:

- Are your measures credible?
- Are your measures valid?

- Are your measures measuring what counts?
- Are your measures reliable?
- Are your measures precise?

Credibility: refers to how reliable or believable your data collection is. In other words, are the data that you are collecting giving you information about the actual situation? As well, it is important to make sure that the data you are collecting are relevant and they measure the most important information. Be sure to avoid the trap of measuring what is easy instead of measuring what you need. For example, teacher opinions may not be the most credible measure for learning the reasons for high dropout rates. The opinions of the dropouts are a more relevant measure.

Validity: is a term used to describe if a measurement actually measures what it is supposed to measure. Are the questions you are asking giving you information about the issues you want to measure? For example, using waiting lists as a measure of the demand for early childhood education may not be valid. Waiting lists are frequently out of date and parents place children on multiple waiting lists. Two kinds of validity are face validity and content validity:

Face validity: addresses the extent to which the contents of the test or procedure *look* like they are measuring what they are supposed to measure. For example, if you were measuring health status or physical fitness, the measure of how fast one runs 100 meters, may indeed look like it could be a measure of health status or at least physical fitness.

Content validity: addresses the extent to which the content of the test or procedure adequately represents all that is required for validity. Again using the example of health status, if a researcher was trying to develop such a measure, then he or she should allow other competent people to examine the content of the proposed test to ensure that all relevant measures are included and that all are weighted appropriately for the proposed test.

Reliability: is a term used to describe the stability of your measurement: that it measures the same thing, in the same way, in repeated tests. For example, the measurement tools for some sporting events need to be reliable. The clock, stopwatch, or tape that measures the distance of a jump, must be measure the time or distance in the same way, each time it is used. If it does, it is considered a reliable measure. If it does not, the results of the study (the competition) would be flawed and results of the event could be questioned. Birth weights of newborn infants are an

example of a reliable measure, assuming the scales are calibrated. Attendance rates are an example of a measure with low reliability unless it is precisely defined. Attendance rates on the first day of school and three quarters of the way through the school year are known to vary substantially.

Precision: is a term used to describe how the language used in the data collection matches the measure. For example, if the question is about countries, then the measures must be at the national level. If the question is about people, then the measures must be on the individual level.

Data collection is the first operational stage in the information system. The objective is to ensure that event data entering the system are valid, complete and free from material errors. In many respects, this is the most important stage in the system. Two rules govern the design of data collection procedure: relevance and efficiency. The information system should capture only relevant data. A fundamental task of the system designer is to determine what is and what is not relevant. He or she does so by analyzing the user's needs. Only data that ultimately contribute to information are relevant. The data collection stage should be designed to filter irrelevant facts from the system.

1.4.1 Methods of collecting irrigation data

The researcher should clearly identify and indicate on what specific key variables data are required, the method(s) of measurement and the unit of measurement. The selection of the variables should depend on what is to be done in the study.

The researcher should also pay due attention to: - The choice of methods of data collection is largely based on **the efficiency** and **accurac**y with which the information will be collected and its **practicability of the methods** or techniques to the respondents.

- **The nature of data** (quantitative and /or qualitative,)

A. **Quantitative data;** when the variables being studied in a numerical form and it can be ranked or ordered. So it said to be quantitative data.

Example; height, body temperature, money, age weights, numbers of students in the class, distance from Addis to Alage Atvet College

B. **Qualitative data:** when variables being studied in a non-numeric form it is called qualitative data. Examples -color, religious preferences sex and, etc.

1.4.2. Data collection methods

Regardless of the kinds of data involved, data collection in a qualitative study takes a great deal of time. The researcher needs to record any potentially useful data thoroughly, accurately, and systematically, using field notes, sketches, audiotapes, photographs and other suitable means. The data collection methods must observe the ethical principles of research.

Different ways of collecting evaluation data are useful for different purposes, and each has advantages and disadvantages. Various factors will influence your choice of a data collection method: the questions you want to investigate, resources available to you, your timeline, and more.

- ✓ **Quantitative data collection methods -**rely on random sampling and structured data collection instruments that fit diverse experiences into predetermined response categories. They produce results that are easy to summarize, compare, and generalize.
- **Typical quantitative data gathering strategies include:**

  Administer surveys with closed-ended questions (e.g., face-to face and telephone interviews, questionnaires etc).

  I. Interviews: In quantitative research (survey research); interviews are more structured than in qualitative research. In a structured interview, the researcher asks a standard set of questions and nothing more and uses different types of interviews. These are the followings:

Types of interview

A)Face -to -face interviews: has a distinct advantage of enabling the researcher to establish rapport with potential participants and therefore gain their cooperation. These interviews yield highest response rates in survey research. They also allow the researcher to clarify ambiguous

answers and when appropriate, seek follow-up information. Disadvantages include impractical when large samples are involved time consuming and expensive.

B) Telephone interviews : are less time consuming and less expensive and the researcher has ready access to anyone on the planet that has a telephone. Disadvantages are that the response rate is not as high as the face-to- face interview as but considerably higher than the mailed questionnaire. The sample may be biased to the extent that people without phones are part of the population about whom the researcher wants to draw inferences.

C) Computer assisted personal interviewing (CAPI): is a form of personal interviewing, but instead of completing a questionnaire, the interviewer brings along a laptop or hand-held computer to enter the information directly into the database. This method saves time involved in processing the data, as well as saving the interviewer from carrying around hundreds of questionnaires. However, this type of data collection method can be expensive to set up and requires that interviewers have computer and typing skills.

II. Questionnaires

A. Paper-pencil-questionnaires: can be sent to a large number of people and saves the researcher time and money. People are more truthful while responding to the questionnaires regarding controversial issues in particular due to the fact that their responses are anonymous. But they also have drawbacks. Majority of the people who receive questionnaires don't return them and those who do might not be representative of the originally selected sample.

B. web based questionnaires: a new and inevitably growing methodology is the use of internet based research. This would mean receiving an e-mail on which you would click on an address that would take you to a secure web-site to fill in a questionnaire. This type of research is often quicker and less detailed. Some disadvantages of this method include the exclusion of people who do not have a computer or are unable to access a computer. Also the validity of such surveys is in question as people might be in a hurry to complete it and so might not give accurate responses. Questionnaires often make use of checklist and rating scales. These devices help simplify and quantify people's behaviors and attitudes. **A** checklist is a list of behaviors, characteristics, or other entities that researcher is looking for. Either the researcher or survey participant simply checks whether each item on the list is observed, present or true or vice versa. A **rating scale** is more useful when a behavior needs to be evaluated on a continuum

- Observing and recording well-defined events (e.g., counting the number of patients waiting in emergency at specified times of the day).
- Obtaining relevant data from management information systems.
- Experiments/clinical trials.

Qualitative data collection methods: play an important role in impact evaluation by providing information useful to understand the processes behind observed results and assess changes in people's perceptions of their well-being. Furthermore qualitative methods can be used to improve the quality of survey-based quantitative evaluations by helping generate evaluation hypothesis; strengthening the design of survey questionnaires and expanding or clarifying quantitative evaluation findings.

These methods are characterized by the following attributes:

- They tend to be **open-ended and** have **less structured protocols** (i.e., researchers may change the data collection strategy by adding, refining, or dropping techniques or informants)
- They rely more heavily on interactive interviews; respondents may be interviewed several times to follow up on a particular issue, clarify concepts or check the reliability of data
- They use **triangulation** to increase the credibility of their findings (i.e., researchers rely on multiple data collection methods to check the authenticity of their results)
- Generally their findings **are not generalization** to any specific population, rather each case study produces a single piece of evidence that can be used to seek general patterns among different studies of the same issue

The qualitative methods most commonly used in evaluation can be classified in three broad categories:

1. **In-depth interview** - is a technique designed to elicit a vivid picture of the participant's perspective on the research topic. During in –depth interviews the person being interviewed is considered the expert and the interviewer is considered as the student. Researchers engage with participants by posing questions in a neutral manner, listening attentively to participants' responses, and asking follow –up questions and probes based on those responses. In- depth interviews are usually conducting face- to face and involve one interviewer and one participant. In- depth interviews are useful for learning about the perspectives of individuals, as opposed to, for example, group norms of a community, for which focus groups are more appropriate. They are an effective qualitative method for getting people to talk about their personal feelings, opinions, and experiences. They are also an opportunity for us to gain insight into how people interpret and order the world

2. **Observation methods** –is a fundamental and highly important method in all qualitative inquiry. Observation entails the systemic nothing and recording of events, behaviors and artifacts

(objects) in the social setting chosen for study. The observation record is frequently referred to as field notes-detailed, nonjudgmental, concrete descriptions of what has been observed

3. Document review -documentation (to document) - is the term used in several different ways. Generally, it refers to the process of providing evidence. Documentation include

- Written information for any read, projection or technical performing,
- Data media of any format and for any reproduction,
- Other content.

1.5. Updating, modifying, maintaining & storing data

Most database maintenance requires running the commands. You can also display information from a command.

For example

- Upgrade the data store
- Move the data in the data store to a different database server
- Migrate the data store from a microsoft access database to a microsoft server database
- Change the name of the file

With the exception of microsoft access connects to the data store, although some affect the data store. Citrix strongly recommends creating a backup copy of the data store. Without a backup, you must manually recreate all of the farm policies, settings, accounts, and other persistent data in the data store. If the data store fails, each farm server can run off the data in its local host cache indefinitely (provided it can contact the license server). However, you cannot make any modifications to the farm or use the access management console or advanced configuration tool. To restore a backup database or migrate to a new server, use the migrate command. Set up to configure the data store. If you use the same name as the previous data store, you do not need to reconfigure the farm servers.

Self-Check 1	Written Test

Name: _____ **Date:** _____

Directions: **Answer all the questions listed below. Illustrations may be necessary to aid some explanations/answers**

1. Define data, information, knowledge and wisdom?(**5pt**)
2. How to organize the data for analysis? (**5pts**)
3. List the types of interview and questionnaires used to collect irrigation data?(**5pts**)
4. What are the Realizing methods used to data collecting **?(5pt)**
5. What is the difference between qualitative and quantitative data?(**5pt**)

Note: **Satisfactory rating –12.5 points and above Unsatisfactory - below 12.5 points**

You can ask your teacher for the copy of the correct answer.

Informationsheet-2	Analyze and interpret data with descriptive statistics

2.1 defining & consistency objectives of analysis

What do we mean by analyzing data?

Analyzing information involves examining it in ways that reveal the relationships, patterns, trends, etc. that can be found within it. That may mean subjecting it to statistical operations that can tell you not only what kinds of relationships seem to exist among variables, but also to what level you can trust the answers you're getting. It may mean comparing your information to that from other groups (a control or comparison group, statewide figures, etc.), to help draw some conclusions from the data. The point, in terms of your evaluation, is to get an accurate assessment in order to better understand your work and its effects on those you're concerned with, or in order to better understand the overall situation. There are two kinds of data analyzing you're to be working with, although not all evaluations will necessarily include both.

Analysis of data: The basic purpose of data analysis is to make it useful for certain conclusions. This analysis may simply be critical observation of data to draw conclusions about it or it may involve highly complex and sophisticated mathematical techniques.Once data are collected, the next step is to analyze the data. A plan for analyzing your data should be developed well before

it is time to conduct analysis. The best time to develop your analysis plan is when you are first identifying your key evaluation questions and determining how you will collect the needed information. It's important to match the analysis strategy to the type of information that you have and the kinds of evaluation questions you are trying to answer.

Interpretation of data; based on analyzed data conclusion have to be drawn, that conclusion is called interpretation.

The objectives of analysis data:

- Is to determine a model that best describes the experimental results.
- Is to take all variables into account in a model.

Obtaining information on the variables for the entire population may not be possible. For instance, what qualities do farmers of a particular region of the country value most in a new insecticide? Is it the long lasting effect of the chemical or the environmental impact of the insecticide, or can it be the price?

2.2 realizing methods of data analysis

Analysis should be approached as a critical, reflective, and iterative process that cycle between data and an overarching research framework that keeps the big picture in mind

Managing Data: -Regardless of data type, managing your data involves Analysis of data inorder to meet the desired purpose of investigation, data has to be analyzed.

The purpose of analyzing data is to high light information which is useful for decision making.

⇒ Effective data analysis involves

- keeping your eye on the main game
- managing your data
- engaging in the actual process of quantitative and / or qualitative analysis
- presenting your data
- drawing meaningful and logical conclusions

Both qualitative and quantitative data analyzed –

- **Qualitative data** analyzed by description, interpretation, tabulation, figure and map
- **Quantitative data** analyzed by simple statistical methods such as percentages , averages, means, mode and etc

Stages in analysis

1. **Transcription** –almost all qualitative research studies involves some degree of transcription the data may be tape recorded interviews, focus groups, video recordings or hand written field notes. It is not appropriate, usually to write up summery notes from a tape recording.
2. **Organizing your data-** after transcription, it is necessary to organize your data into easily retrievable sections. You may wish to give each interview a number or code or to break up field notes into sections identified by date, or by context.
3. **Familiarization-** by this we mean the researcher listening to tapes and watching video material, reading and re-reading the data, making memos and summaries before the formal analysis begins. It is important if the main researcher has not gathered all the data themselves.
4. **Coding-**this process of deciding how to conceptually decided up raw data. These sections contain data which the researcher is interested in the exploring and analyzing further.
5. **Themes (categories) -** a theme is generated when similar issues and ideas expressed by participants with in data are brought together by the researcher into a single categories or cluster.

2.3- clearing, justifying & consistency assumptions

Assumption: is something that is taken for granted or advanced as fact. In the logical frame work, important assumptions are key threats to the project that exist in the external environment. Good data should satisfy the criteria of:-

- ❖ They should serve a definite purpose
- ❖ Be easy to complete

- ❖ Be up to date

So that any action needed can be taken as early as possible

Note: Assumptions used in analyses are clear, justified and consistent

2.4 supporting conclusions & contributing to business objectives.

Once the data are presented, analysis is needed to explain what the data mean. Observations and conclusions based on the data are used to produce a summary account of stream over a given time period. Analyzing data will reveal information about a particular issue. It can also point out differences between issues or between sampling periods. If historical data are available, analysis will reveal how the current situation compares with that of previous years. Long-term comparisons may be crucial in identifying trends, especially those that are gradual.

The first step in analyzing any data is to identify the obvious. In other words, as you look at the data, presented in graphs and tables, what starts to stand out at you? Are any trends evident? Are there notably high and low values for a certain parameter or do values remain fairly constant? Do changes in one parameter coincide with changes in another?

Analyzing and interpreting data means asking questions, looking for and explaining answers, and drawing conclusions. To help you in the process, a list of questions follows. The list is by no means complete, but serves as a starting-off place. After you've analyzed the data and made some interpretations, keep in mind that your interpretations are just that

2.5 Using descriptive statistics

Statistics is a set of scientific principles and techniques that are useful in reaching conclusions about populations and processes when the available information is both limited and variable; that is, statistics is the science of learning from data. Almost everyone including corporate presidents, marketing representatives, social scientists, engineers, medical researchers, and consumers deals with data. These data could be in the form of quarterly sales figures, percent increase in juvenile crime, contamination levels in water samples, survival rates for patients undergoing medical therapy, census figures, or input that helps determine which brand of car to purchase. In this text, we approach the study of statistics by considering the four steps in learning from data:

(1) Designing the data collection process,

(2) Preparing data for analysis (summarization, models),

(3) Analyzing the data, and

(4) Reporting the conclusions obtained during data analysis.

Descriptive statistics: descriptive statistics are used to summarize the basic feature of a data set through:-

- Measures of central tendency (mean, mode, and median)
- Dispersion (range, quartiles, variance, and standard deviation)
- Distribution (skewness and kurtosis)

Data types: different data types demand discrete treatment, so it's important to be able to distinguish variables by cause and effect (dependent or independent) measurement scales (nominal, ordinal, interval, and ratio)

Statistical distribution measures

When studying the data distribution graphically, it is apparent that the distributions of the Variables can look quite different. Instead of looking at countless data distributions, it maybe desired to characterize the data distribution by a number of parameters in a table. What is required? First the central value of the distribution needs to be identified, together with a measure of the spread (variation) of the data. Furthermore, the quartiles and different percentiles of the distribution may be of interest (i.e. above what value fall the uppermost two, five, or ten Per cent of the data).when working with "ideal" data, two further measures are often provided in statistical tabulations: skewness and kurtosis. Skewness is a measure of the symmetry of the data distribution, kurtosis provides an expression of the curvature – the appearance of the density trace can be flat or steep. Such "summary values" are frequently used to compare data from different investigations. For real data there is often the problem that the presence of outliers and/or multimodal distributions will bias these measures. Both can be easily recognized from the graphics described in the previous chapter.

Central value: what is the most appropriate estimator for the central value of a data distribution? What is actually the central value of a distribution? It could be the "center of gravity", it could be the most likely value, it could be the most frequent value, it could be the value that divides the

samples into two equal halves. Accordingly there exist several different statistical measures of the central value (location) of a data distribution.

Using this form of calculation should be avoided as rounding errors may occur due to the Arithmetic precision limitations of computers. A preferable method is to use logarithms, then:

The mode: the mode is the value with the highest probability of occurrence. There is no simple formula to estimate the mode; however, the mode is often estimated from a histogram or density trace the mode being the value where the histogram or density function shows a maximum.

The median: the median divides the data distribution into two equal halves. The data are sorted from the lowest to the highest value, and the central value of the ordered data is the median. In the case that *n* is an even number, there exist two central values and the average of these two values is taken as the median.

This may best be demonstrated by:

- **a simple example**:2.3 2.7 1.9 2.1 1.8 2.4 2.0 5.9.
- The data are then sorted: 1.8 1.9 2.0 2.1 2.3 2.4 2.7 5.9.
- The two central values are:1.8 1.9 2.0 **2.1 2.3** 2.4 2.7 5.9,and
- the median is (2.1 + 2.3)/2 = **2.2**
- the mean of these eight values is 2.64,
- the geometric mean, g, is 2.44,
- Range = max min.

Example

Dataset: 1, 9, 5, 6, 9

Sum of data = 1 + 9 + 5 + 6 + 9 = 30

Total numbers= 5 therefore the mean for this dataset is 30/5==6

Example

Dataset: 1, 9, 5, 6, 9

1) Organize data in numerical order = 1, 5, 6, 9, 9

2) Find the middle value = 6

The median is 6.

Example

Dataset: 1, 9, 5, 6, 9

The mode is 9.

Example

Dataset: 1, 9, 5, 6, 9

1) Organize data in numerical order = 1, 5, 6, 9, 9

2) calculate the range by subtracting the lowest value from the highest value;

Range = 9-1 = 8the range is 8.

Self-Check 2	Written Test

Name: _____ **Date:** _____

Directions: **Answer all the questions listed below. Illustrations may be necessary to aid some explanations/answers**

1. What do you mean by analyze data? (5pts)
2. List the objective of data analysis? (5pts)
3. What are stages in analysis? 5pts
4. What is descriptive statics?5pts

Note: **Satisfactory rating –10 points and above Unsatisfactory - below 10 points**

You can ask your teacher for the copy of the correct answer.

Informationsheet-3	Present data

3.1 Preparing data using suitable business technology

Good scientific research involves a sound methodology and a novel idea that can be tested simply and repeatedly to give valid, trustworthy results. However, even the most clinically significant research is useless if it is not communicated successfully. Scientific ideas are novel, sometimes simple in theory, but most always complex in technique. These attributes of research make it necessary to use all available means of presentation. The most common media for scientists to communicate with the general public is primary journal articles. However, posters and oral presentations are also affective because they allow scientists to be in direct contact with their audience. This provides both parties an opportunity to ask pertinent questions to add clarity to the work being presented.

A poster is an exciting way for scientists to present their research. It, just as a primary research article, includes all aspects of the scientific method. A title that is brief, but specific, an abstract, an introduction, material and methods, results, and a conclusion are some headings that can appear on a poster. Also, references and acknowledgments are sometimes are included. A poster

is different from a written manuscript or an oral presentation because it is mostly graphical. As such, it is important to design a poster that is visually pleasing by focusing on charts, graphs, and pictures and minimizing lengthy introductions and discussions. Highlighting all significant information with the use of bullets is essential because if further explanation is needed the audience will simply ask for it.

Oral presentations are yet another avenue for scientists to share their findings with the world. Although it can be challenging to present years of works within fifteen minutes, oral presentations can be a rewarding experience because you are the only one front of an audience whose attention you know have. Of course this emphasizes the need to speak clearly and concisely with choice words that engross the audience. Again, just as with written manuscript and posters the format of oral presentations can also vary, but essentially it must include logical, easy-to-understand events that are presented in a matter with respect to the scientific method.

How to presenting data?

Data are prepared in an appropriate format, style and structure using different tools of data presentation methods.

Presentation of data- in graphs and diagrams may also use to give the data a vivid meaning and make the presentation attractive. Data are presented in the tabular or graphical or diagrams form of presentation.

Large data will be presented in tables in a very summarized and condensed manner. Discussion helps to interpret the results of the study in proper perspectives and to relate them with other relevant studies, including the hypothesis rejected and accepted.

A .Tabular method of data presentation

A table which shows the list of all values obtained in the data and the frequency with which these values occur in the data is called a frequency distributions table. A table that grouped data that can be easily visualized is called frequency distributions. The row data which have been collected and edited will not usually give you sufficient information unless they are put in to a form that will make them eases to understand and interpret. Consider the following raw data

collected from the scores of 40 samples students in a tea, coffee and spice production examination.

Table3 .1.Score of 40 students of exam results

56	78	62	37	54	39	62	60
30	67	72	39	56	78	64	56
42	55	39	54	78	56	62	78
56	55	37	56	47	42	30	37
47	54	42	55	46	46	55	80

We can say about the performance of the 40 sample students unless the above raw data are put into an ordered array in ascending or descending order so that they can be looked at more objectively.

The above data may be set up and tallied like this ungrouped frequency distribution (see table **Table 3.2 U**ngrouped frequency distribution.

Score(X)	Students no (f)
30	2
37	3
39	3
42	3
46	2
47	2
54	3
55	4
56	6
60	1
62	3

64	1
67	1
72	4
80	1

N=40

X=scores

F= frequency

N= total number of students

Raw data are better understood if they are summarized in same defines order or sequence. After having summarized the scores of 40 students you can say more about the performance of the student, the score of the students range from 30 – 80 they rather dispersed the most frequently occurring score is 56.If you had a large number of scores say thousands or more, the bode method of summarizing data would be very lengthy and quality in efficient in cases where we have a very large data, it is necessary to condense them in to as suitable number of groups or classes and then assigning the combined frequencies of these values to their respective classes. For example, the data about the marks of the 40 students can be condensed into five classes as shown below.

Table 3.3 grouped frequency distribution or grouped data

Range of scores	Frequency
30-40	8
41-50	7
51-60	14
61-70	5
71-80	6

N= 40 (total)

The table used for the arrangement of data into groups of classes, together with the respective class frequencies is called a grouped frequency distribution table (see the above table 3.3)The grouping of data in a grouped frequency distribution destroys mach of the original detail of the data. For instance, in ungrouped frequency distribution table we know that there is only one student who scored 64. But in grouped frequency distribution table tells us that there is only one student who has scored in the rang 61 – 70 we do not know from grouped frequency distribution whether that single student has scored 61,62,63,64,65,66,67,68,69, or 70 .

Some basic terminologies

Class interval(k):- it used to categories (classify) data into similar characteristics. Stages rules used to determine the class interval (k)

$K=1+3.322(\log_{10}^{n})$, where n is the total no of observation. In the above table n=40 then $k=1+3.322(\log 1o^{40})$ **K=6**

Example: - in table 3.3, the symbol 30 – 40 used to denote the first class is a class interval of 30 – 40.

Range: is the difference between maximum value of data and minimum value collect data.

In the above table 3.3 range **= maximum value - minimum value**

R = 80-30=50

Width (w), w=range/k =maximum value -minimum value/class interval=50/6=~8

Class limits: - the first and last elements of a given class interval are called class limits. Or by taking the last value of an observation as the lower boundary of the first class and obtain the other by adding the class width.

Example: - in the interval of 30 – 40, the value 30 and 39 are class limits. 30 is the lower class limit and 40 is the upper class limit of the interval 30 – 40.

Class marks (m):- the middle value of any particular class is referenced to as the class mark or class midpoint. It is the average (mean) of the upper and lower class limits.

Class mark (m) = $\frac{\text{lower class limit + upper class limit}}{2}$

E.g. M = (30 – 40) $= \frac{30 + 40}{2} = 35$

B. Graphical methods of data presentation

Obviously constructing a frequency distribution table is not the end of our work. Rather, it is the beginning. Commonly, the next step is to present the grouped data in charts and graphs so that the reader may easily determine the essential features of a frequency distribution and compare one with another it desired.

Why do we present data in charts and graphs?

As mentioned above, presenting data in charts and graphs is important it have great memorizing value in that you:

1. Has a clear idea about the nature of the data
2. Easily understand patterns &trends of grouped data
3. Facilitate comparisons between two or more sets of data

The most commonly used charts are: **bar charts** and **pie charts**. Similarly graphs that are most commonly used are**: histogram**, **frequency polygon and cumulative frequency curve**. We usually use charts (diagrams) to illustrate data which are discrete while we use graphs to represent data which are of continuous nature. Graphs are simply visual aids for thinking about and discussing statistical problems.

Charts (diagrams)

I. Bar charts

It is applicable for discrete variable, i.e. the data given according to time, places, etc. The period of time and place are putting in x-axis and the corresponding frequencies at the y- axis. Bars are vertical lines, where the lengths of the bars are proportional to their corresponding frequencies or numerical values. Bar charts are called one-dimensional. This means only the length of the bar (rectangle) is important not the width.

There are three types of bar charts

1. Simple (deviation) bar chart
2. Component bar chart
3. Multiple bar charts

1. Simple (deviation) bar chart

When data contained both **positive and negative numbers** such data like profit, net revenue, present change in deviation, then deviation bar chart is appropriate.

Simple bar chart is usually used to illustrate the nature of a single item without considering the components in the item.

Example. Suppose a merchant has the record of net profit (loss) of three commodities in the given period of time below, and then construct deviation bar chart for these records?

Commodity	Net profit(loss)
Soap	80
Sugar	-95
Coffee	125

2. Component bar chart:

This is a bar chart which illustrates the component parts of the whole. The rectangles (bars) are subdivided into components depending up on the type of information to be shown in the chart.

Example. Let us decompose the total of the number of peasant by status to further subdivision n in year.

Status	Male	Female	Total
Single	90	10	100
Maries	30	40	70
Divorced	1	29	30

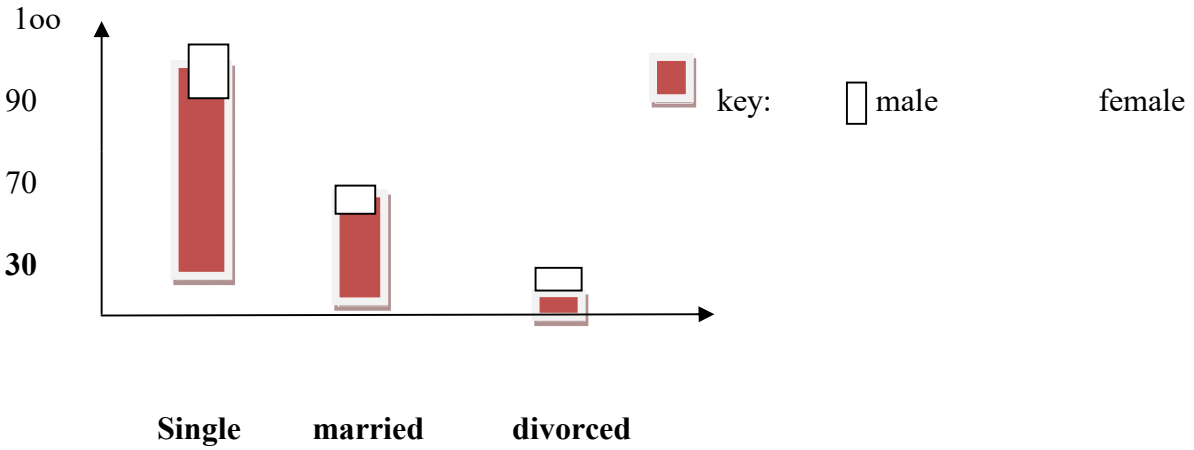


Fig 3.4 component bar chart

As you may have noted in the component bar chart, the two components, shaded red color represents the total numbers of male status and the white un-shaded white is the status of female status subdivision in n year.

3. Multiple bar charts

This is a bar chart showing the various components of an item side by side. Example. The following item of expenditure of two families to different item

Items of expenditure	Family a	Family b
Food	540	240
Clothing	180	100
House	100	53
Education	65	23

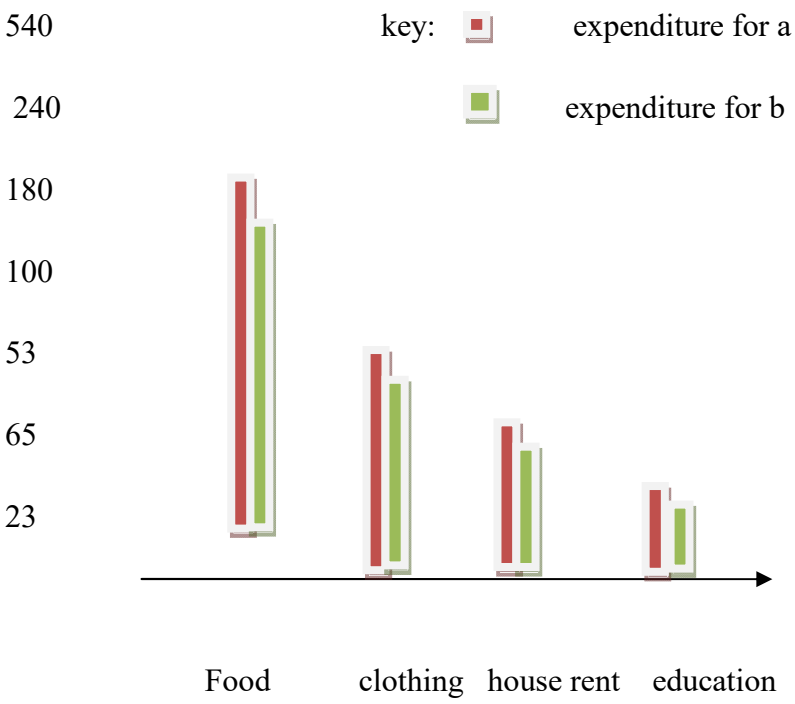


Fig 3.3 multiple bar chart

Ii.Pie chart

A pie chart is a circle divided from its center into several component parts to show how the whole is divided up. It is called "pie" because the entire chart looks like a pie and its components resemble slices cut from it. The size of a slice represents the proportion of the component out of the total. A pie chart is a circle divided in components sectors according to the breakup of components given in percentage.

Example of income generated from different sector is given below and then construct **a pie-chart**.

Sectors	**Income in million(f)**	**Percentile (%)**
Agriculture(agric)	3200	93.348
Industry(indi)	28	0.82
Services(ser)	200	5.84
Total(n)	3428	100%

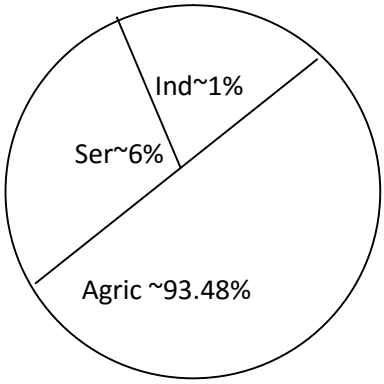
How to calculate the percentages (**%)** of income generated from each sector

Formula of %=fx100/n agric%= 3200x100/3428=93.48

F=frequency indi%=28x100/3428=0.82

N=total number ser%=200x100/3428=5.834~ 6

Fig 3.2 pie **chart**



3.2 clearing & conforming structure & format of reports

A structure generated from a production version or a bill of material in that contains information about the production cycle and the component assignment for the production of a product; the production data structure can be used in as a source of supply for in-house production.

Structure: the production data structure consists of the following list of component with information:

- Consumed total quantity
- Type of consumption
- List of available capacities with reference to resources
- List of activities with data for setup
- List of modes with data for duration and assignment to activities
- Relationships

IntegrationThe production data structure is generated during the transfer of master data from the connected. The production data structure can be generated from the following original data:

- Production version with routing and bill of material.

- Production version based on a master recipient.

3.3 Reporting findings

A report: is an official record of a given period activities that presents a summary of a work plan implementation and performance reporting. Progress reports are essential mechanisms for project implementers to inform partners and donors on the progress, difficulties, and problems encountered and lessons learned during the implementation of project activities.

Reports are designed to:

Enable the assessment of progress in the implementation process and achievement of Results.

- ✓ .focus activities and therefore improve subsequent work plans.
- ✓ It serves as evidence
- ✓ It helps for follow up evaluation
- ✓ It provides input to plan the training for future
- ✓ Even if there is staff turnover the information is not hidden from the new comer ,since it is documented

A report may include the activity name, worker name and date, activity description, progress of activities, major issues, expenditure and any future activities that may need to be planned.

- ❖ Reports on progress of implementation of activities
- Are the activities being implemented as planned?
- ❖ Reports on progress in delivering outputs
- Is delivery of outputs on schedule?
- ❖ Reports on achievement of results
- Evidence of progress towards intermediate effects (based on indicators)..
- Evidence of progress towards outcome indicators.
- Evidence of progress towards impacts based on indicators

Reports are made monthly, quarterly and annually. Some projects prefer to have monthly reports, especially on the individual work plans. The individuals responsible for activities and tasks

prepare monthly monitoring reports that include an assessment of the level of achievement, difficulties experienced and suggestions for the next month

Presenting the findings

- ✓ Only make claims that your data can support
- ✓ The best way to present your findings depends on the audience, the purpose, and the data gathering and analysis undertaken
- ✓ Graphical representations (as discussed above) may be appropriate for presentation
- ✓ Other techniques are:
    - Rigorous notations, e.g. UML
    - Using stories, e.g. to create scenarios
    - Summarizing the findings

Recording production data

In the description of the recording techniques contained in the appendices, some of the details given may seem petty and supper fulvous, they are deliberately in clouded, however, because they are often not appreciated, and one of the most important reasons for farmers not keeping records that would be valuable to them is not they do not know precisely what needs to be done. Sometimes recording is started but later abandoned because it proves to be too difficult or not to be giving the information required often these faults can be rectified by a dears appreciation of recording procedures.

Purpose of recording

- ❖ To check on performance
- ❖ To guide future reference
- ❖ To provide planning data

Planning involves both selecting the right technique & obtaining the necessary data. It is usually the later that provides the greater problem for advisers "standard data are available to some extent but are rarely sufficiently detailed to be applied with full confidence to individual. Even if standard data were available in for greater detail than at present, circumstances on the individual farm would frequently fall between different categories, since it would be optimize to expect data to be provided to cover a continuous range of differences.

Method of data recording

Dates are recorded in different methods: - for instance

A. Field data records: - can be kept in on ordinary bound ledger, or on cards. There is no page or cord, per field per year. The virtue of loose leaf files or cords is that successive years for each field can be put together to make continuous record. Also, data for and single crop can easily be assembled as required.

If only a few items are recorded, one sheet or card may be used for more them one year, thus saving on space requirement over a period of years, these records will from a type of diary for each field for subsequent reference. The items to be recorded include

- Crop grown
- Variety
- Fertilizers and sprays used
- Date sown and
- Date of harvested.

Yield should be recorded as accurately as possible without going to excessive trouble trying to keep lost separate.Often these can be only roughly estimated, as is usually the case with cereals, but an approximate figure is undoubtedly better them nothing, notes should be added, recording observations about the field, e.g. Concerning drainage, fertility or soil, pea laities many of these items may first be written in note books carried by a member of the form staff the head tractor driver, arable fore man or the farm manager and subsequently transferred to the field records.

I. Rotational data record:-

This should be kept an a single sheet of paper, unless the number of fields is so large as to make it impossible the names and/or numbers of the field(and their areas if required)are listed down the left hand side and the remaining space is divided into. Say ten columns, representing, successive years, the crop grow on each field in each year is entered in the appropriate space.

II. Labor data recording

On farm with elaborate grapping systems and at peak times(and on many horticultural units labor may be fully deployed on essential productive work for most on even all of the year) the task may be recorded on time sheets, but these are laborious to analysis where a large staff is

employed and are often in accurate, time sheets are more useful as a means of recording ordinary time and ever time worked for calculating wages, and as a means of transferring useful information from the field to the office, them for recording the labor hours spent on different jobs. On farms with a limited number of major tasks at peak times, the information required about these can be recorded in a way that requires far less subsequent analysis than time sheet, the method involves recording certain details of each task, either in a pocket book kept by the head tractor driver, arable foreman or manager, or specially prepared small cards, one page or cord is needed for each job, unless it covers along period, in which case two or more pages cards, each covering one or weeks, will be required per job, the task such as main crop harvesting is written at the top. Normally there will be one entry per day. Columns are needed for the date, the gang size (i.e. Numbers of workers employed, hours worked and hectors completed.

- **Preparing appropriate format for data recording**

For data recording appropriate format preparation is necessary: - data are recorded in different waist may be undisruptive way or in way or in table or way or in graphical, chart, figurative

E.g.: Format of controlling coffee growth by furrow irrigation

Spacing (m)	Plants /ha	Coffee type	No of sucks	Date of plotting	Date of measuring the plant	Height of plant	frequency of irrigation
2x2 2x1.8	250 1000	Open Compact	3 4	10/1/2004 10/3/2004	Monday September Friday October 16/2005	6cm 9cm	

Nb. Format for recording data is differing according to the user and their necessary, for example format used for recording daily laborers controlling and recording data for plant management is not the same.

3.4 obtaining feedback & comments

Customer feedback: This is particularly important with regards to providing poor service. People are much more likely to tell people that they have received poor service from a store than they are to tell people they received exceptional service. If you provide poor service you are likely to lose current customers, and they will tell others which may prevent them from using your company.

Feedback –is the response a receiver gives to the message received. It is the action –reaction interdependence in communication. It is an error correcting mechanism that can overcome noise/barriers of production factors. Is the process in which knowledge of the surroundings or of the consequences of actions of a system lead to adjustments of future actions, seen in the light of achieving a certain goal

Feedback play an important part because it tells both the source and the receivers, their messages are being interpreted. It provides the source with information concerning his or her success in accomplishing the objective. Feedback describes the situation when output from (or information about the result of) an event or phenomenon in the past will influence an occurrence or occurrences of the same development of the original phenomenon in the present and future. An example feedback, an extension agent who is helping farmers to achieve higher crop yields can improve the farmers' control over their decision making process by teaching them how to measure crop yields. This form of control or feedback should be carried out quickly.

Self-Check 3	**Written Test**

Name: _____ **Date:** _____

Directions: **Answer all the questions listed below. Illustrations may be necessary to aid some explanations/answers**

1. List the types of charts and graphs used to preset data? 5pts
2. How do you present the finding? 5pts
3. What is the purpose of recording? 5pts

4. What are the method of data recording?5pts

Note: **Satisfactory rating –10 points and above Unsatisfactory - below 10 points**

You can ask your teacher for the copy of the correct answer.

Reference

- ✓ Judd, Charles and, McCleland, Gary (1989). Data Analysis. Harcourt Brace Jovanovich. ISBN 0-15-516765-0.
- ✓ John Tukey-The Future of Data Analysis-July 1961
- ✓ Rankin, J. (2013, March 28). How data Systems & reports can either fight or propagate the data analysis error epidemic, and how educator leaders can help. Presentation conducted from Technology Information Center for Administrative Leadership (TICAL) School Leadership Summit.
- ✓ Chambers, John M.; Cleveland, William S.; Kleiner, Beat; Tukey, Paul A. (1983). Graphical Methods for Data Analysis, Wadsworth/Duxbury Press. ISBN 0-534-98052-X
- ✓ Fandango, Armando (2008). Python Data Analysis, 2nd Edition. Packt Publishers.
- ✓ NIST/SEMATECH (2008) Handbook of Statistical Methods,