

MEASUREMENTS AND EVALUATION

Meaning: The word measurement is used by interchangeably with the other words such as test, assessment and evaluation but actually distinct meaning.

Measurement: is the term used for quantitative data, hence it is used when an observation needs to be quantify. Measurement is that term when empirical observations to be quantify. Then facts and data are gathered.

Some attributes that quantity measure includes but not limited to the following

- height
- Interest
- Intelligence
- Achievement etc.

Measurement requires the qualification of attributes according to specified rules.

For clarification, an attribute about table is measure e.g. the length of the table, the weight of the table, the texture of the table you are not measuring table. Measuring students attributes such as what the students write, watching their performance, listening to what the students say by using senses of sight, hearing, touch, smell and taste to gather information relevant to stated goals

The process of measurement involves three steps

1. Identifying and defining the quality or attributes that is to be measured
2. Determine a set of operation by which the attributes may be made manifest and perceived
3. Establishing a set of procedure or definition for translating observations onto quantitative statements of Degree or amount.

Testing: This is the systematic procedures or techniques of measurement sample of behavior or trait some measurement are obtained without subjecting individuals to task but in test individuals need to go for a specific task. Teacher can keep record of student's behaviour as it occurs in the classroom or on the playground, there is no task to measure the behaviour, this is not test.

Test may be classified as follows:

- How they are administered: individual or group

- How they are scored: objectively or subjectively
- The sort of response they emphasize e.g. power of speed
- What they attempt to measure e.g. sample or sign
- type of response subject must make (e.g. performance or pencil on paper)
- Nature of the group being compared e.g. Teacher made or standardized.

Principles of measurement include:

- ✓ There should be multiple observation of the attributes or trait to be measured
- ✓ Observation must be controlled
- ✓ Multiple observation must be converted to a number according to some set of rules

Level of Measurements

These can be classified into four different levels of scales

- I. **Nominal Scale:** most primitive scale of measurement by placing object or individuals into categories that are qualitatively rather than quantitatively different. These types of measurement only require recognizing distinguishable two or more mutually exclusive categories and knowing the category of placing object or individuals into a category. The number may be used to differentiate but such numbers cannot be used for any calculation such as car plate number, street number, telephone number etc.
- II. **Ordinal Scale:** This rates objects or individual under consideration a high below or above, how much of or how less of attribute under consideration they possess. This usually carries in rank order such as 1st, 2nd, 3rd etc. it may also have names such as high, low, above, below, school, medium etc. Intelligent Quotient (IQ) scores is an ordinal scale in which all scores are only meaningful for comparison.
- III. **Interval Scale:** This is a more precise method of measurement which orders objects or individuals according to the number of attributes they represent and give equal number of intervals between units of measure equal differences in the number represent equal differences in the attribute being measured. Example is the measurement of temperature with Celsius.
- IV. **Ratio Scale:** This possesses the highest level of measurement scale and has properties of an interval scale together with a fixed origin of zero point. Example includes weight, height and time. With ratio scales we can compare both differences and relative magnitude in scores.

ASSESSMENT

Assessment is the collection of data to describe or better understanding an issue. It can include quantitative data (Numbers) and qualitative (picture, word etc). It is a process by which information is obtained relative to some known objects or goals.

Test is special cases assessment but not all assessment information may concern the following

- **Student:** to determine level of understanding
- **Programme:** to ascertain if the programme is achieving its policy on the overall achieving its objectives
- **Educational Policy:** To verify the impact of any educational policy on the overall achievement of educational objectives.

EVALUATION:

This is rather complex concept because, evaluation involves engaging in some process that is design to provide information that will help us make a judgment for both qualitative (non measurement) and quantitative (measurement). Sometimes valuation is usually done a feedback to known the level of compliance or judge what obtained with what we actually intended “feedback” is useful if it aid decision making

Types of Evaluation

- **Formative Evaluation:** it is carried out during the program of instruction. It does not count towards the final grade but as a feedback for the learners to know how they are doing and help them to learn more effectively it is used to permit entry to an examination. The evaluation of this kind include
 - ✓ Test
 - ✓ Assignment
 - ✓ Projects during the course of programme.
- **Summative Evaluation:** This is the kind of evaluation that normally comes up at the end of the programme of instruction or section in order to establish or measure what the learner has achieved. It is usually count towards the final grade, mark or equal or used to determine whether the learner is allow making progress through the course. This is done to evaluate students into category of replacement to satisfy each learner’s merit for the purpose of certification.

Note: The formative evaluation will take note of all important element in unit such as knowledge, comprehension, application analysis, synthesis and evaluation level it must be sequential to certain student satisfaction of one level before the other

Demographic Data Collection and Analysis

Demographic Data Collection: is all the process involves or employ to gather social economic information for the purpose of demographic studies.

The information/Data in this wise included age, sex employment level employment status etc.

Data analysis: This involves the transformation of data into usable information for informed decision making (i.e. to effect decision-making). Data analyses usually employ statistical tools to conceptualize information into usable mathematical process.

TEST AND TESTING

The word test means the presentation of a standard set of questions to be answer.

Testing in another word refers to as systematic procedure or technique of measuring sample of behaviour or traits.

Thus, testing is an objectives method of observation

Common Dimension in Classifying Test

However, Test differs in nature and serves different purposes. Test can be classified into the following classifications:

- Classification based upon the type format used
 - a. Essay
 - b. Objectives Test
- Classification based upon the stimulus material used to present, the task to the testee (student)
 - a. Verbal
 - b. Non-verbal
- Classification based on purpose of the test
 - a. Achievement (b.) Aptitude test
 - a. Diagnosis (b.) Prognosis test
 - a. Criterion referred (b.) Norm referenced
- Classification based upon difficulty level and time schedule for it test
 - a. Standardized (b.) power test
- Classification based upon the number of coverage of the test
 - a. Individual (b.) grouped test

Essay & Objective

Essay Test: require an individual to organize and express in his or her words. It allows free response to task given.

Objective Test: require individual to supply a brief answer or choose the correct answer from the alternative provide. Objective contains more questions but make less time to answer

ADVANTAGES OF ESSAY & ITS FEATURES

Advantages

- Students provides answer with minimum constrain
- Student answers a small number of questions
- Student produces answers having all degree of completeness and accuracy

Features of Essay

- Relatively easy to prepare but more difficult to grade accurately
- Issues is not restricted to a list responses accurately

- Fewer question but lengthy answers.

ADVANTAGES OF TEST & ITS FEATURES

Advantages

- Economic way of obtaining information from students.
- It takes less time to answer than an essay question
- Many questions can be asked with more adequate content resulting to higher reliability with better content validity
- It can be scored more easily and more accurately
- It create incentive for pupils to build up broad base of knowledge skills and ability
- There is high level of objectivity in scoring

Features

- Students operate within an almost completely structured task
- Students selects one of a limited number of alternatives
- Students respond to each of a large sample of items
- Students receive a score for each answer according to predetermined scoring

Types of Objective

- ❖ Select answer
- ❖ Supply answer

Disadvantages of objective test

- It does not measure the higher mental process but encourages rote memory
- It encourage guessing
- It neglects the measurement of writing ability

Disadvantages of Essay

- It has poor content sampling, not cover much of the content , hence content validity is low
- Unreliable scoring, i.e. subjective scoring which may be influenced of many factors
- Encouragement of bluffing where student can present enables packages of irrelevancy

ACHIEVEMENT TEST

This is the test that measures the extent to which a person has learnt or what skills the student has mastered as a result of planned instruction or training. It is course oriented activity.

The main purpose is to measure desirable changes in student both qualitatively and quantitatively.

Importance of Achievement Test

- Judging the pupils mastery of certain essential skills and knowledge
- Measuring growth overtime
- Ranking student in term of their achievement of particular instrumental objects

- Diagnosing pupils difficulties
- Evaluating the teacher's instructional method
- Ascertain the effectiveness of encouraging good study habits
- Motivating students

Aptitude Test

This is the test measure the potential is acquire skills that is a natural tendency to do something well especially one that can be further developed. It is a measurement of quickness and easy in learning.

Importance of Aptitude Test

- It predict future scholastic achievement
- It provide ways shield/performance with other in the same situation
- It provide a profile of assessing strength and weakness
- It assess differences and similarities among individuals
- It uncover hidden talents in individual thereby improving educational opportunities
- It is a valuable tools for working with physical challenged children.

Diagnostic and Prognostic Test

These two are refer to as clinical test

Diagnostic test are done to reveal difficulties in school subject. It is to analyze students' strength and weakness in learning activities to when he or she exposed to. It will serve to identified needs of individuals.

Prognostic test: it is used to discover the degree of success that is likely to attend a pupil's effort in study a particular course or partaking in a programme. It usually serves as an aptitude test. It predict future achievement.

Speed and Power Test

Speed test is done to examine individual differences based on speed of performance very easy to be done but with a very limited time that will be very difficult for student to finish.

Power test has a time limit that permits student to attempt all the items

Uses of Test in Education Administration

- ✓ For placement decision i.e. to decide who can be admitted or rejected selection on job etc.
- ✓ Use by educationist to place student in enrich classes or remedial classes
- ✓ For diagnostic or remedial decision
- ✓ For vocational individual strength and weakness such as teacher, parent, government and other stakeholders
- ✓ To asses effectiveness of instruction, adequacy of curriculum and effectiveness of school program to ascertain modification and improvement.

❖ Feedback Uses

- ✓ Teacher use it to measure his/her performance
- ✓ To guide student learning and improve their performance

- ✓ Use for guidance and counseling purpose to discover special abilities. Or interest in learners.
- ❖ **Research Uses**
- ✓ It is used for theory development e.g. psychology theory
- ✓ Use as a basis for data collection in research
- ❖ **Prediction/Future Purposes**
- ✓ Prediction of future success e.g. entrance examination, JAMB/UTME, SAT etc.
- ❖ **Certification**
- ✓ After necessary of outcome have been taken with meeting required condition by individuals the certificate will be issued

PROJECT TECHNIQUES

Project Techniques: it originate within a clinical setting

- ✓ It provide and permit wide aptitude in its solution going through a relative structured task
- ✓ Individuals will project his characteristics mode of response into a given task
- ✓ The test materials will serve as surf of screen on which respondent project the characteristics thought processes, needs anxieties and conflict.
- ✓ It is used as personality test revealing convert, latent or unconsciousness, aspect of personality.

Types of Project Techniques

- ✓ Ink-blot Techniques
- ✓ Pictorial Techniques
- ✓ Verbal Techniques
- ✓ Performance Techniques

Usefulness

- ✓ Interesting and entertaining, reduces embarrassment and defensiveness, no treat to respondent's prestige any response is right
- ✓ Some of this are specially useful to young children, illiterate, people with speech defects or language difficulties
- ✓ It does not allow faking. The test use to be disquieted individuals may not know how the responses will be scored and interpreted.

Disadvantages of Projective Techniques

- ✓ May be in inadequately standardized with respect to administrating scoring
- ✓ Lack of objectivity in scoring and interpretation as the experiences of the examiner determine the final stage.
- ✓ It lack normative data grossly inadequate or based on vaguely described population
- ✓ The interpretation is usually subjective hence may lead to family interpretation
- ✓ Has a very low red ability test, internal consistency cannot be established.

Characteristics of Projective Test

- ✓ The table provided to the individual are usually both novel and quite unstructured conventional and stereotyped pattern of response
- ✓ The nature of the approval being made is usually well distinguished
- ✓ Most of the tests make little or no demands on literacy on academic skills

Standardized Test and Teacher Made Test

Standardized test are test constructed administered and marked by experts. It is usually have consistency of format content, and administration procedure. Its reliability is verified by statistical evidence

The reliability, validity and fairness must be account for such that how its result will be reported and used must be specified.

e.g. Norm-referenced testing (NRT) criterion referenced testing (CRT) and standard-based testing.

Principle of Standardized Achievement Test

- ✓ They are not constructed in a way to be tricky or confused the test-taker
- ✓ It employ some questions that may seem to require a degree of cunning to answer
- ✓ When writing multiple choice questions for standardized achievement test, there come up with good distractors that can be a challenge.
- ✓ The set of alternatives should not be a challenge
- ✓ The set of alternative should be placed In a particular order or be of such lengths as to provide clue to the test-taker.
- ✓ The correct answer should be randomly placed

Function of Standardized Test

- ✓ Objectives are general to the needs of students in most classroom
- ✓ Items are fixed and are not modifiable only most common areas of the curriculum/syllabus are surveyed.
- ✓ The test developed by its publisher must be followed as with the interest manual.
- ✓ Norms are provided by the publisher to all teachers to compare class performance to different age and grade groups.
- ✓ Data on the quality of examination are provided by the publisher.

The foregoing differentiate standardized test and the teacher made test

The Teacher Made Test

It can be used to measure the academic achievement of students in a particular subject. Teacher-made test constructed, administer and marked by the class teacher. It cannot be used for percentile, Age and grade norm

It can only be used for the following

- ✓ Providing information for career
- ✓ Providing information for reporting purpose
- ✓ Providing the teacher to ascertain individual strength and weakness while the students are studying a particular subject matter.
- ✓ Providing information for job
- ✓ Providing information for replacement into ability levels or educational level.

Feature of Teacher-Made Test

- ✓ Objective are specific to the needs of the students in a given classroom where teaching and learning takes place
- ✓ The content may come from any area of the curriculum. Item may be added, eliminated or modified by the teacher.
- ✓ The tests are determined by the teacher. They should be uniform with the class
- ✓ No norm are provided by may be developed by the teacher from his own class
- ✓ The evaluation procedures are provided by the teacher.

Principle and Technique of Test Construction

In all measurement and evaluation it is required to determine what we are attempting to evaluate or measure. It now depends on the type of test to use.

Majorie and Neavart (2009) test construction developed as follows

1. Planning

- Outline subject matter to be considered
- Identify learning outcome to be measured
- Prepare table of specification
- Choose appropriate test type

2. Preparing the Test

- Write test according to the rules of construction
- Selects the items to be included in the test
- Review and edit item according to guidelines
- Arrange items: decide on
 - a. Grouping of items
 - b. Sequence of items
 - c. Sequence of groupings
- Prepare direction for the test
- Decide on method of scoring

3. Analyzing and revising the Test

- Perform test analysis to determine difficulty, discrimination, and reliability
- Retain, edit or discard

- Revise the test if necessary

Test Preparation

For final draft, reviewing, selecting and editing items

1. Each test item should measure an important learning outcome included in the table of specification
2. Each item type should be appropriate for the particular learning outcome to be measured
3. Each item should present a clearly formulated task
4. The item stated should be in simple and clear language.
5. The Item should be free of extraneous clues
6. The difficulty of the item should be appropriate for the student to be tested
7. Each items should be independent and the items as group should be free from overlapping
8. The items to be included in the test should provided adequate coverage of the table of specification
 - Arranging items
 - Writing Directions
 - Other considerations
 - Reproducing test

Administering and Scoring Test

This is the decision made regarding of time limits, observation of time limits, physical set up proportional distribution of test, scoring sheet, scoring method (hand/machine).

- Assigning grades
- Reporting grades to students
- Ground rules regarding students “Complaint”
- Decision need to be made regarding procedure for appeal by students of well as method of dealing with appeals

Analysis of Testing

The systematic evaluation of the effectiveness of each of a test the analysis will reveal from the following

- Difficulty of the item
- Discriminating power of the item
- The effectiveness of each alternatives

Simplified Item-Analysis

Item analysis require for informal achievement test used in teaching is the simplest ones though very effective.

For illustration: consider 32 test papers to illustrate the steps

- Arrange the test papers in order from highest to the lowest.
- Select approximately one-third of the papers with the highest scores and tags it the “upper group” (10papers)
- Select the same number of papers with the lowest scores and called it “lower group” (10papers)
- Set the middle group (12papers)
- ✓ Poor each item, count the number of students in to upper groups who selected each alternatives make the same count for the lower group

Record the count from the above step in a sheet of paper o n a copy of the test in column to the left of alternatives to which each can refers as follows

Item 1 Alternatives	A	B	C	D	E
Upper	10	0	6	3	1
Lower	10	3	2	2	3

Estimate the item difficulty

- ✓ Get the percentage of students who got the item right, this can be done by using estimate of the student who in chided in the item analysis groups.
- ✓ Sum total of lower and upper group (10 + 10 = 20)
- ✓ The number select the correct answer (6 + 2 = 8)

$$\text{Difficult index } P = \frac{\text{Number of right response (R)}}{\text{Total number exermine (T)}} \times \frac{100}{1}$$

$$\text{i.e. } P = \frac{P}{T} \times \frac{100}{1}$$

$$P = \frac{8}{20} \times \frac{100}{1}$$

$$P = 40\%$$

Index of difficulty

We base estimate on the extreme group of upper and lower, thus this represent an approximate for this entire group.

We can say that the difficult index for this group if 40%, the smaller the percentage figure the more difficult the item.

Establish item: discriminating power- By comparing the number of students in the upper and lower groups who got it right In the above example 6 students from the upper and 2 students from the lower group got the item right) select the correct answer. This is positive discrimination: since the item differentiates between the students in the same way

that the total test score does. This interpret that the students in the higher group test in the item more frequently than the students in the lower **cadre**.

Analysis of inspection may be more appropriate, but an index of discrimination can be easily computed as follows

$$\begin{aligned}\text{Discriminating Power} &= \frac{\text{Those that got it right}}{\text{Number in the upper group} - \text{Number in the lower group}} \\ &= \frac{6-2}{10} = \frac{4}{10} \\ &= 0.4\end{aligned}$$

This is obtained in fraction (decimal)

Hence, maximum discrimination power is given selecting the correct obtain this if all students in the upper group selecting the correct answer and non students from the lower group select correct answer

$$\begin{aligned}\text{i.e. discrimination index} &= \frac{10-0}{10} \\ &= \frac{10}{10} \\ &= 1.00\end{aligned}$$

This item is at 5% level of difficulty i.e. 10 from upper got it right and 10 from the lower got it wrong. Hence, 50% item of difficulty is a good index.

Zero discriminating power: when an equal number of students from both group get the item right.

Negative discriminating power: when more student from the lower and got item right.

Both types of item such with zero discriminating power and those with negative power should be removed from general achievement list an be discard or improved.

Effectiveness of the distracter: By selecting the number of students in the upper and lower groups who select each incorrect alternatives. A good distracter will have more students from lower and who select incorrect/wrong options for an item. In the above example item A & D are good distracter while C is more students from the upper group a poor distracter and alternative E is ineffectiveness since it attracted no one

This analysis test item combine with an inspection of the item itself provides useful information improving the item.

The above procedure steps for analysis item can fill the modified to fill some situation. Sometime is better to inspect the data them computing the difficulty and discriminating power. To select the upper and lower group we can use the top and bottom such as 25% of the group in **large** group or upper and lower halves of the group is small. The important things are to use the large enough fraction of the data/group to provide useful information. Selecting top and bottom 27% of the group and applying other statistical refinement for better analysis.

Objective case of administration

- ✓ Validity
- ✓ Reliability
- ✓ Difficult level
- ✓ Discrimination
- ✓ Usability

Characteristics of a Good Test

The following are good attributes of a good test

1. **Validity:** A valid test will fulfill all the purposes that is to be measure and to the extent desired.

Types Of Validity Test

- Content validity
- Construct validity
- Face validity
- Criterion-related validity (on current & predictive)
- ❖ Concurrent validity
- ❖ Predictive validity
- **CONTENT VALIDITY:** This is the degree to which the sample of items, task or questions on test is representative of some define universe or domain of content.
- **FACE VALIDITY:** This is the physical appearance of the test items which include the quality of the paper used the lay out of the test items, the arrangement of the items of letter used etc. in technical sense.
- **CONSTRUCT VALIDITY:** This is the extent to which test scores measure a psychological construct. That is, to which the test extent the test score reflect the theory behind the psychological construct being measured.
- **CONCURRENT VALIDITY:** This is the relationship between scores on the measure and criteria sources obtained sit the same time.
- **CRITERION-RELATED VALIDITY:** This shows that test scores are systematically related to one or more outcome criteria the two types of criterion include.

- Concurrent and
- Predictive Validity
- **Predictive Validity:** This is the relationship between scores on the measure and the criterion scores available at a future time. Hence, it is the degree to which success or failure in particular test can be used to predict success or failure in a future test or activity.
- **Reliability:** This is the consistency of test ability to measure accurately what it supposes to measure. It is the extent to which a particular measurement is consistent and reproducible. It has as well described as measure of consistency, stability, dependability, predictability, precision and accuracy.

Method of Estimating Reliability Coefficient

These include:

- i. Test-Retest method
 - ii. Equipment/Parallel form method
 - iii. Split-half method *and*
 - iv. Kuder Richardson (20-21) methods
- **Test-retest:** This is a method employed to estimate the stability of a test item. It is the simplest way for internal consistency procedure. It is computed by administering a test item on a group of respondents twice with a given time lag.
The correlation coefficient obtained by this process is called a test-retest reliability coefficient. High correlation coefficient guarantees that the score obtained by individuals at a time will be the same if given at another time.
 - **Equivalent or Parallel Form:** This method makes use of two sets of test where one will serve as equivalent/to parallel similar to the other and measuring something if two forms of test are administered essentially at the same time (in immediate succession) the resulting reliability coefficient is called the coefficient of equivalence.
Although designing alternative forms of a test that are truly equivalent may be problematic and this is the shortest of this technique of estimating reliability.
 - **Split- Half Form:** it is an internal consistency procedure with only a single administration of one form of a test. The requirement to split a test into two equal halves by using even and odd steps to split the already even number of questions to avoid problems of halves using **pearson** moment correlation to correlate the two halves/
The correlation coefficient computed between the two halves-systematically under estimates into reliability of the entire test. To transform the split-half correlation into an appropriate reliability estimate for the whole test, we use the Spearman-Brown test.
 - **Kuder-Richardson (20-21) Reliability:** This is based on the agreement of constituency of all items in the test. The test known form of constituency of all items the test. The best known form of homogeneity is the Richardson-20 formula which is based

on the proportion of correct and incorrect responses to each of the items on a test as the variance of the total scores.

The procedure for computation is of two forms, namely K-R20 and K-R21, with the same result/answer after computation.

$$K-R_{20}: r_{xx} = \frac{n}{n-1} \left(1 - \frac{\sum pq}{SD^2} \right)$$

FACTORS AFFECTING OF A TEST

Length of a Test: Reliability is proportional to the length i.e. the coverage area of the content such that the more sample taken from given area of knowledge or behavior.

Speed of Test' The speed of a test is considered when everyone known virtually what to write on an item but the time is not allowed to finish the answer; this type is considered as pure speed test . Thus the pure speed considered the differences in the score which depends on the number of items attempt.

Few tests are either pure speed or pure power tests. However, to the extent that a test is speed. It is appropriate to estimate reliability through the method of internal consistency and measures of stability or equivalence should be used.

- Examination Malpractices
- Impressionistic marking
- Difficulty of the items: There must be high variability in true scores abstained so that wide, and the score for everyone must not be too poor or extremely high for everyone.

Objectivity: The test author/ examiner must make the test to be fair to the testee where bias test does not portray objectivity and not reliable.

Discrimination: A good test must be able to make distinction between poor and good learner by considering the following criterion

Comprehensiveness: test item that covers much of the content of courses is said to be comprehensive.

Ease of Administration: simple and clear direction, few subjects and moderate time should not be difficult to administer.

Practically and scoring: Gives quantitative score to answer must be practically simple e.g what, why, how etc.

Usability: This depends on the test practical features which include its availability, ease of scoring ease of interpretation and application and availability of equivalent or comparable forms. A good test should be usable, ambiguous and clear stated with one meaning only.

Basic Statistics Measures

Measures of Central tendencies

Measures of spread

Measures of Dispersion

Measures of Variability

Census vital registration in JSS.

Statistics is the scientific methods for collection organizing summarizing, presenting and analyzing data to draw valid conclusion and make reasonable decision.

Statistics deals with set of numerical facts, in the form of numbered which convey or connotes some notion and ideas about something or about some events.

Statistics is also the method of proceeding numerical information (data) to get convey informative notion or the numbers that convey formative notions about some events or concepts. Before we can get information from data, it beds to be organized Summarized and described.

Descriptive Statistics will shows the basis for mauling a particular situation (characteristics, summarizing, generation)

Inferential Statistic: This will give an inference on a particular situation in generalizing.

Some Importance of statistics

- It provides knowledge of the techniques of data collection
- It enable individual to understand statistical procedure or method of analyzing data.
- It will develop the teacher's skills of reducing large quantities of data manageable and understandable form.
- It aids the studying of population and sample
- It improves the efficiency of teacher in determine the amount of relationship among of data
- It aids the teacher in making reliable inferences from observational data.
- It is used in making conclusions from data gathered from research

Note:

The **Description Statistics**: covers the description of the characteristics of a group, summaries of the statistical data and also include graphical or diagrammatic representation of facts such as pie charts, bar charts, histogram, Graphs and Pictogram. It deals with the determination of mean or average characteristics of a group e.g. age weight, height etc. which include measure central tendency (mean, median & mode), measures of variability percentage, stamina etc.

The inferential statistics includes it statistics that provide the basis for making generalization and inferences. It gives deductions or inferences based on the values from a sample. It concerns with primarily with.

- Estimation of Hypothesis
- Testing of Hypothesis

Inferential statistics is majorly the techniques that are used on the data obtained to in far on populations. Randomness and uncertainty in the observation is used to draw inferences about the process or population under study.

The duo of descriptive and inferential statistics.

In using inferential statistics, the amount of error committed can as well be estimated. So that some statement like I am 25% or it is above 95% assurance level will have a meaning to the listening and if a student score above 85% is a pretty good result.

Some Symbol in Statistics

X or Y = Variables

Such as $X = \{x_1, x_2, x_3 \dots\}$

And $Y = \{y_1, y_2, y_3, \dots\}$

Valuables with x and y taking different values.

f_x – Frequency of variables

f_x – Frequency x variables

$$\sum_{i=1}^n xi - (\text{sigma}) - \text{sunnatiion} \frac{xi}{i} \text{ startfrom i t on}$$

\bar{x} – Mean sample

μ - Population mean

Σ For N- Total num of variable

δxi – Standard deviation of sample

δx – Standard deviation of population

γ - Correlation coefficient

β - Beta value

α - Alpha value

Definition

Population: large group of people that contains the entire element that have the common characteristic. Population could be finite (definite number) or infinite (not definite)

Sample: representative of the population taken out to be **studied**

Variables: characteristics which may take several values

Parameter: this is the specific we defined characteristics of a population. Thus a student may required to know some parameters about the population such as

- The most popular student on campus
- The highest temperature measure in the school Obaagun
- The average barometric pressure level at the center of hurricane measured in recorded history
- The department that used to has higher number of enrolment in the school
- The most popular assigned frequency for fin broadcast in Nigeria.

Statistic: A specific characteristics of a sample such simple in the above population include

- ❖ The best student in Edu 223

- ❖ The highest temperature reading by 1pm heat time may 27 in RACOED
- ❖ The barometric measure in pacific hurricane right during the fear 2010 (None female)
- ❖ The department that recorded highest number of enroll during the year 2010(Non female)

Organizing Data

Raw Data: these are the data collected that have not been organized numerically

Example: The set of scores obtained from a test concluded by a teacher for 15 students give the following 5,4,2,4,3,6,5,4,1,0,5,5,8,4. This is a raw data.

Array: it is an arrangement of raw numerical data in ascending or descending order of magnitude.

Range: the difference between the largest and the lowest data entry

Frequency: this is the number of times each score value (or interval of score value) occurs in a distribution. The number (or interval of score value) occurs in a distribution. The number of times that the result or outcomes occurs within a specific sample of population frequency is also know as often

The frequency could be absolute or relative.

Example: if a fair die tossed 6000 times approximately each dots could up like 1000 times each, hence the absolute frequency is 1000 from each face of the die. The relative frequency for each of the six 6 faces is approximately 1 out of 6 i.e. $\frac{1}{6} \times 1000$ percentage i.e 1.67% approximately.

Destitution

A distribution is the description of the set of possible values that a random variable can take. This can be done by noting the absolute relative frequency. Distribution can be illustrated by table or graph.

Frequency distribution presents the scores values and their frequency of occurrence. The score value usually listed in rank order with the lowest score usually at the bottom of the table. The main purpose of frequency distribution is to make the data to be more meaningful and to facilitate ease of understanding and interpreting the scores.

Information from the frequency distributions can now be transform into graphical or pictoral representative with loss of originality but more understanding with the graph picture.

Example: Using the raw data of students scores in Edu 232 examination to present a frequency distribution

4, 2, 14, 8, 5, 5, 3, 7, 5, 6, 5, 4, 0, 4.

Score x	Frequency
8	1
7	1
6	4
5	4
4	1
3	1
2	1
1	1
0	1

Table 1

Measures of central tendency include

- Mean
- Median
- Mode


Estimation of the mean, media & mode

Example I

(1) Use the following array of numbers representing the scores of work for student in a class test 5, 4, 2, 4, 3, 6, 5, 4, 4, 0, 5, 5, 7, 8, 4

Solution

 Arranging the data in ascending order

 0, 1, 2, 3, 4, 4, 4, 4, 5, 5, 5, 5, 6, 7, 8.

$$\begin{aligned}
 \text{Mean } (\bar{x}) &= \frac{\sum Xi}{N} \\
 &= \frac{0+1+2+3+4+4+4+4+5+5+5+5+6+7+8}{15} \\
 &= \frac{63}{15}
 \end{aligned}$$

$$= 4.2$$

Median = the middle number = 4

Mode: (The highest most occurring data entry); 4 appears four times also this data has two mode, hence it is called **bimodal** the mode in this case is the mean of the two mode i.e the mean of 4 and 5

$$\frac{4+5}{2} = 4.5$$

Example II calculate, median and mode of 4, 3, 3, 2, 5, 1, 4 3, 3, 2.

Range of data is the difference between highest and lowest entry

For the entry the distribution in example 1

The range is 8-0

$$= 8$$

But the range in example II is

$$5-1$$

$$=4$$

Solution Table II

Score x	F	fx
8	1	1
7	1	7
6	1	6
5	4	20
4	4	16
3	1	3
2	1	2
1	1	1
0	1	0

$$\sum f = 15 \quad \sum fx = 63$$

$$\text{Mean} = \frac{\sum fx}{\sum f} = \frac{63}{15}$$

$$= 4.2$$

Median = 4

Mode = 4 & 5

$$= \frac{4+5}{2} = 4.5$$

Hence, with frequency distribution of data and taking the data and taking the data as raw, we obtained the same result.

A frequency distribution is otherwise known as list showing the number of time each value or interval score value occurs.

A distribution can be ungroup or group depending on the number of the population.

To present a raw data in a grouped frequency distributing the following steps could be helpful.

Step 1: Decide whether the data could be arranged in ascending or descending order, **stepii:** count the number of times each figure appear in arrays use tally.

StepIII: Determine the range of scores the largest and the smallest numbers in the saw data (if the data contain large masses of raw data). These are often result into grouped frequency distribution with arrangement of data by classes together with the corresponding class frequency.

Example III: Give the sores of 40 students in chemistry test as follows:

25, 6, 9, 10, 12, 15, 19, 9, 21, 15, 21, 19, 14, 14,

12, 10, 12, 14, 19, 14, 14, 12, 14, 15, 19, 14, 15,

10, 14, 15 12, 15, 14, 12, 15, 12, 14, 19, 12, 10.

This data can be represented both as group and ungrouped let the data handled as ungrouped.


Step I: The data is presented in Array by descending order.

Scores (x)	Frequency (F)
25	1

21	2
19	5
15	7
14	10
12	8
10	4
9	2

Table III

How for obtain width for grouped frequency distribution we take following steps

 Find the interval of the data not too wide and too close but usually interval of 10 and 20.

- Range of the score
- Selecting the interval size
- Computing

Before we can consider the data for grouped distribution, we must not the data for a grouped distribution, we must not the following

- That the number of the data large and range is wide this is considered to be large masses of data
- The data could be arranged into class or category
- The frequency of each class can be determined using the data of Table III

$$\text{Range} = 25 - 6 = 19$$

$$\text{Class width} = 4$$

$$\text{Class size} = \frac{\text{Range}}{\text{Class width}} = \frac{19}{4} = 4.9$$

$$\cong 5$$

$$\text{Class interval} \Rightarrow 6 + 4 = 10, 11 + 4 = 15, 16 + 4 = 20, 21 + 4 = 25 \text{ etc.}$$

$$6 - 10, 11 - 15, 16 - 20, 21 - 25 \text{ etc.}$$

Class boundary take by adding ± 0.5 i.e. -0.5 to the lower class boundary and $+0.5$ to the upper class boundary.

i.e. $0 - 0.5 + 6, +0.5 + 10$ etc.

$5.5 - 10.5, 10.5 - 15.5, 15.5 - 20.5$ etc.

From table III, we have the ungrouped frequency distribution and the grouped frequency distribution as follows.

X	f	fx
25	1	25
21	2	42
19	5	95
15	7	105
14	10	140
12	8	96
10	4	40
9	2	18
6	1	6
	<u>40</u>	<u>567</u>

$$\begin{aligned}
 \text{Mean} &= \frac{\sum fx}{\sum f} \\
 &= \frac{567}{40} \\
 &= 14.175 \cong 14.2
 \end{aligned}$$

$$\text{Median} = 14$$

$$\text{i.e. item at } \frac{20}{2} = 20$$

Where the frequency 20 falls

Mode is also 40

The data with the largest frequency

Grouped Data

Class interval	class boundary	class mod point x	frequency(f)	fx
6 – 10	5.5 – 10.5	8	7	56
11 – 15	10.5 – 15.5	13	25	325
16 – 20	15.5 – 20.5	18	5	90
21 – 25	20.5 – 25.5	23	3	69
			<u>$\sum f = 40$</u>	<u>$\sum fx = 540$</u>

$$\begin{aligned}\text{Mean} &= \frac{\sum fx}{\sum f} \\ &= \frac{540}{40} \\ &= 13.6 \cong 14\end{aligned}$$

$$\text{Median } L_1 + \left[\frac{\frac{N}{2} - \sum fi}{f_{\text{median}}} \right] C$$

Where L_1 is the lower class boundary of the median class

N is the size of population i.e. total frequency

$\sum fi$ is the frequency of the median class

From the table V above

$$L_1 = 10.5$$

$$N = 40$$

$$\sum fi = 7$$

$$C = 5$$

$$f_{\text{median}} = 25$$

$$\begin{aligned}\text{Median} &= 10.5 + \left[\frac{\frac{40}{2} - 7}{25} \right] 5 \\ &= 10.5 + \left[\frac{20 - 7}{25} \right] 5 \\ &= 10.5 + \left[\frac{13}{25} \right] 5 \\ &= 10.5 + \frac{13}{5} \\ &= 10.5 + 2.6 \\ &= 13.1 \\ &\cong 13\end{aligned}$$

$$\text{Mode} = 13$$

Standard Deviation

The method to evaluate depression includes range standard deviation variance mean deviation etc.

Range = maximum entry data – minimum entry data

= Highest value – lowest value

Variance written as is δ^2 given as

$$\delta^2 = \frac{\sum fi(xi - \bar{x})^2}{n-1}$$

Standard deviation is positive square root of variance

$$\delta = \sqrt{\delta^2}$$

$$= \sqrt{\frac{\sum fi(xi - \bar{x})^2}{n-1}}$$

Example: use the weight of 24 babies out of 10,000 recorded in an hospital as follows

3000, 3005, 2995, 2990, 2980, 2981

3002, 3003, 2984, 2986, 3000, 2979

3006, 3004, 2981, 2990, 2990, 2901

2904, 2989, 2900, 2989, 3001, 300

By classifying the table into frequency table we have

Weight interval	2900-2919	2920-2939	2940-2959	2960-2979	2980-2999	3000-3019
Frequency	4	0	0	1	10	9

With the class boundary for weight interval

Class width is 20

Range of the above data = 3004 – 2900

= 104

The mean is given as

$$\bar{x} = \frac{x}{n}$$

$$= \frac{(300+3005+\cdots+3001+3009)}{24}$$

$$= \frac{71479}{24}$$

$$= 2978.29$$

$$= 2978 \text{ to the nearest whole number}$$

Using frequency distribution method

$$\bar{x} = \frac{\sum fi}{\sum f}$$

x_i = Average of the class boundaries with the following

Middle weight	2909.5	2929.5	2949.5	2969.5	2989.5	3009.5
Frequency	4	0	0	1	10	9

$$\bar{x} = \frac{\sum fxi}{\sum f}$$

$$= \frac{4 \times (2909.5) + (0) (2929.5) + (0)(2949.5) + (1)(2969.5) + (10)(2989.5) + (9)(3009.5)}{24}$$

$$\frac{71588}{24} = 2982.8$$

$$\bar{x} = 2983 \text{ to the nearest whole number}$$

Middle weight (xi)	frequency (fi)	$x - \bar{x}$	$(x - \bar{x})^2$	$fi(x - \bar{x})^2$
2909.5	4	-73.5	5402.25	21609
2929.5	0	-53.5	2862.25	0
2949.5	0	-33.5	1122.25	0
2969.5	1	-13.5	182.25	182.25
2989.5	10	6.5	42.25	422.5
3009.5	9	26.5	702.25	6320.25
Total				28534.00

$$n = \sum fi = 24$$

$$n-1 = 23$$

$$\delta^2 = \frac{28534}{23}$$

$$\delta^2 = \sqrt{\frac{28534}{23}}$$

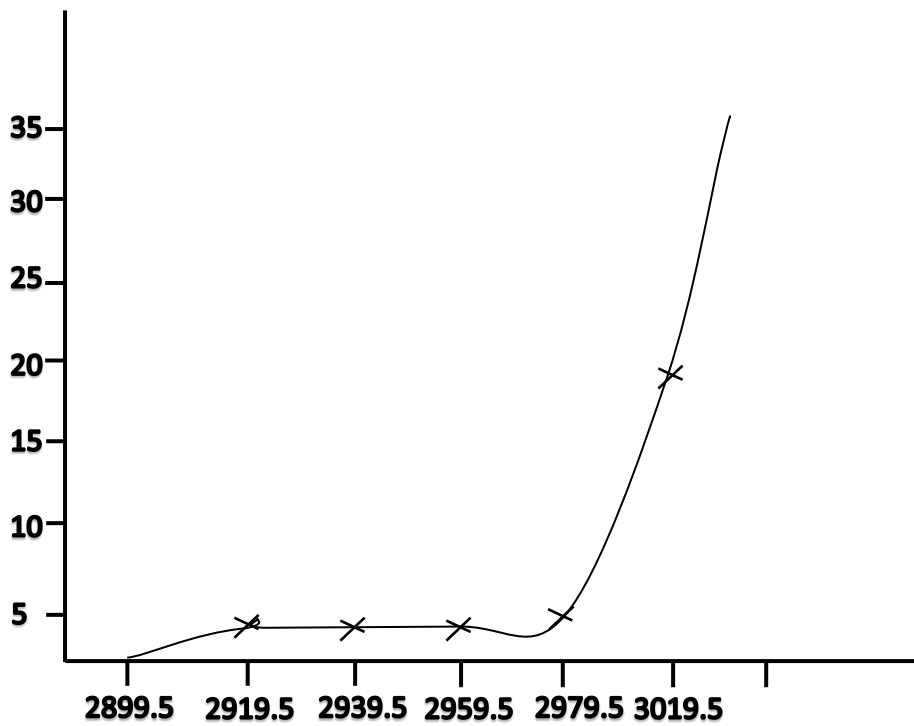
$$= 35.22$$

= 35 to the nearest whole number

35.22 to two decimal places

The cumulative frequency curve (or Ogive)

Wt less than or equal to	
2899.5	0
2919.5	4
2939.5	4
2959.5	4
2979.5	5
3019.5	24



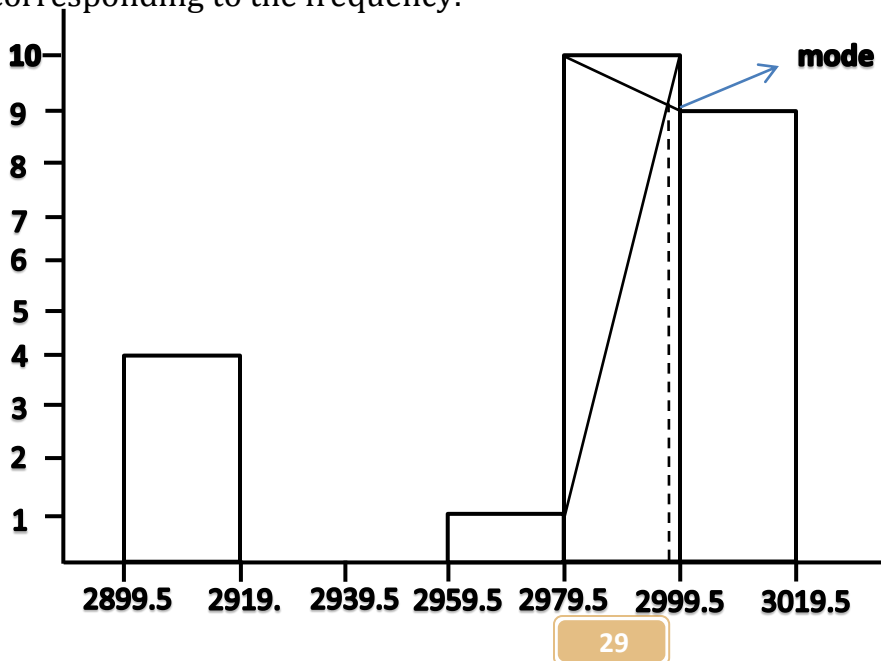
The median of this variation fall between 12th and 15th entry.

$$\begin{aligned} \text{The median} &= \frac{2989+2900}{23} \\ &= 2989.5 \end{aligned}$$

We can as well use Ogive to estimate the median. Since $\sum fi = 24$, half of this is 12 i.e $\frac{1}{2} \sum fi = 12$, we draw a parallel line a at the twelfth frequency parallel to x axis line at the vertical line it meet the weight axis is the median. The median is 2993.5

The Histogram

These are vertical bars joined together with column width and the height corresponding to the frequency.



Only the modal class is 2979.5 – 2999.5. The mode as estimate in the graph is 2998.5

From the formular

$$\text{The mode} = L_1 + \frac{D_1 C}{D_1 + D_2}$$

L_1 – Lower class boundary of the modal class

D_1 - Frequency to the left of the modal class

D_2 - Frequency to the right of the modal class

c - Class width

$$L_1 = 2979.5$$

$$D_1 = 10 - 1 = 9$$

$$D_2 = 10 - 9 = 1$$

$$\text{Mode} = 2979.5 + \frac{D_1 C}{D_1 + D_2}$$

$$= 2979.5 + \frac{(9)20}{9+1}$$

$$= 2979 + \frac{9 \times 20}{10}$$

$$= 2979.5 + 18$$

$$= 2997.5$$

Exercise: 1 A sample survey of the performance of 30 students out of a population of 100 students in chemistry Examination gives the following information.

91, 92, 91, 91, 87, 82, 74, 12, 13, 87

71, 72, 73, 74, 91, 67, 17, 53, 59, 70.

Do the following

1. Classify the information into a frequency distribution table 0-40, 41-50, 51-60, 61-70, 71-100.
- ii. Draw the histogram: the frequency polygon

Mark	0-40	41-50	51 - 60	61-70	71-100
Frequency	2	1	2	2	13

Exercise II: A random survey of height of 30 people out of 10, 000 people revealed the following (in inches)

50.7, 55.9 56.9, 56.6, 57.7, 57.9, 58.1, 69.3, 69.7,

68.7, 68.3, 67.9, 68.3, 67.7, 67.0, 66.6, 65.0, 64.4,

70.5, 70.4 66.4, 69.0, 66.8, 66.9, 68.0, 66.1, 70.1

- Classify the data into 55-53, 54-57, 58-61...
- draw a histogram and frequency polygon for the distribution
- Estimate the mode
- Construct ogive and estimate median height
- Find mean and standard deviation.