**Introduction.**

 **STATISTICS:**

 The word statistics is derived from a word “Status”, which means you have complete information about a related problem/area. It is Latin word. It can be define as:

* The collection of methods and facts.

OR

* The presentation of data in such a way that a common man can easily understand.

OR

* The collection of data, analysis of data, summarization of data, interpretation of data, tabulation of data to get a precise result is called statistics.

OR

* A numerical data relating to an aggregate of facts is called statistics.

In statistics we draw our conclusion by using statistical influence with the help of data or aggregate of data/group of data.

**Uses of Statistics & Importance of Statistics:**

 The Statistical information are used for a variety of reason. Some of them are:

1. To inform general public.
2. To explain things those have happened.
3. To justify a claim.
4. To provide general comparisons.
5. To predict the decision regarding future outcomes.
6. To estimate the unknown quantities.
7. To establish association/relationship between factors.
8. To find out regression correction between variables.
9. The statistical techniques are useful to draw inference about unknown parameters.
10. To predict the unknown values on the basis of available information.

**Importance of Statistics:**

 Statistics is perhaps a subject that is used by everyone. The following functions and uses of statistics in most diverse filed serve to indicate its importance:

1. Statistics assists in summarizing the large sets of data in a form that is easily understandable.
2. Statistics assists in the efficient design of laboratory and field experiments as well surveys.
3. Statistics assists in a sound and effective planning in any field of inquiry.
4. Statistics assists in drawing general conclusions and in making predictions of how much of a thing will happen under given conditions.
5. Statistical techniques being powerful tool for analysis numerical data are used in almost every branch of learning.
6. A businessman, an industrialist and a research worked all employee statistical methods in their work. Banks, Insurance companies and government all have their statistics departments.
7. A modern administrator whether in public to provide basis for decision.
8. A politician uses statistics advantageously to lend support and credence to his arguments while elucidating the problem he handles.
9. A social scientist uses statistical methods in various areas of socio-economic life of a nation. It is sometimes said that:

“A social scientist without an adequate understanding of statistics is often like the blind man groping in dark room for a black cat that is not there”.

1. Simply in every field of life we need statistics.

**SAMPLE:**

* The representative part of the population is called sample.

OR

* The part which contains some of the characteristics of the population is called sample.

OR

* The sub part of the population is called sample.

It is denoted by “n”.

**POPULATION:**

* The aggregate of samples is called population.

OR

The collection of all the samples into one place is called population.

It is denoted by “N”.

**PARAMETER:**

* The numerical/theoretical results calculated from population data are called parameters.

OR

* That characteristics/results calculated from population data is called parameters.

It is constant in nature. It is denoted by “N”.

e.g: $µ, Ω , x^{2}$

**Types/Phases of Statistics:**

 There are two types/phases of statistics:

1. Descriptive Statistics
2. Inferential Statistics

**1. Descriptive Statistics:**

 Descriptive statistics can be defined as:

* The collection of data, analysis of data, summarization of data, interpretation of data, tabulation of data at last we get a precise result in numerical form is called descriptive statistics

OR

* Descriptive statistics is concerned with the summarization and describing a body of data

OR

* Descriptive statistics is that branch of statistics which deals with concepts and methods concerned with summarization and description of important aspect of numerical data.

**2. Inferential Statistics:**

* Inferential Statistics is a branch of statistics through which we collect the data, analysis the data, summarize the data, interpretate the data and tabulate the data to get precise result in non-numerical form.

OR

* The process of reaching generalizations about the whole by examining a portion is called inferential statistics.

OR

* By using inferential statistics we draw inference about the characteristics of related problem and our inference gives non-numerical results.

**Variable:**

* A characteristic that varies with an individual or an object is called variable.

For example, age is a variable as it varies from person to person.

**Types of variable:**

Variable may be classified into qualitative and quantitative according to the form of the characteristics of interest.

**Qualitative Variable:**

If the characteristic is non-numerical such as education, gender, eye- color, intelligence etc., the variable is referred to as qualitative variable.

**Quantitative Variable:**

A variable is called quantitative variable when a characteristic can be express numerically.

Such as age, weight, income etc.

A quantitative variable may be classified as Discrete OR continuous.

**Discrete Variable:**

A discrete variable is one that can any a discrete set of integers or whole number that is the values are taken by jump or breaks, e.g number of persons in a family, numbers of rooms in a house etc.

**Continuous Variable:**

 A variable is called continuous variable if it can take on any value fractional or integer with in a given internal.

 For example age of a person, the temperature at a place etc.

**Data & Types of Data:**

**Data:**

 The collection of raw facts and figures is called data.

1. Primary data ii. Secondary data

**i. Primary Data:**

The data collected for the first time and original in character is called primary data.

OR

The Data that have been originally collected and have not under-gone any sort of statistical treatment are called primary data.

OR

The data is original raw form is called primary data.

**ii. Secondary Data:**

The data that have been collected, classified, tabulated or presented in some form for a certain purpose are called secondary data.

OR

The Data which is not primary called secondary data.

**Sources of Primary Data:**

1. Direct personal investigation.
2. Indirect investigation
3. Interview method
4. Collection through Enumerators.
5. Questioner method
6. Collection through local sources
7. Computer interview method

**Sources of Secondary Data:**

1. Official method.
2. Semi-official method.
3. Data collected from Research organization.
4. Data collected from Newspapers, books, magazine.
5. Data collected from media.
6. Data collected from publication, Article.
7. Data from NADRA.

**Mistake:**

A wrong calculation or adoption of wrong procedure in the collection, completion, analysist and interpretation of data e.g. 6+2-1 = 4.

**Error:**

Error can be define as “The difference between the true value and approximate value”

Or

 The difference between what is acceptable as a true figure and what is taken for an estimate or approximation is called error.

**Absolute Error:**

 It is define as “the actual difference between an estimate value and the true value is called absolute error”

**Relative Error:**

 It can be calculated by:

$$Relative error= \frac{Absolute Error}{Estimate value } x 100$$

**Biased:**

 In series of items the error are all in one direction, the result is called biased.

**Measurement Scales:**

 By Measurement, we usually mean the assigning of number to observations and sealing is a process of measuring.

**Types of scales:**

1. **Nominal Scale:**

It can be define as “the classification of the observation into mutually exclusive qualitative classes is said to be nominal scale”

E.g:

1. Students are classified as male and female. We may use number 1 and 2.
2. Rainfall may be classified as heavy, moderate and light.

We may use number 1,2, and 3

 The numbers when they are used, only identify the categories. In this scale no particular order is used.

1. **Ordinal Scale:**

It is very similar to nominal scale and addition property of ordering of measurement.

Or

In ordinal Scale we classify the data in form of order is called ordinal scale.

E.g:

1. The performance of students or players is rated as excellent, good, fair or poor etc, numbers 1,2,3,4,5, etc are used to indicates ranks.
2. Education, Primary, middle, high, inter, Bachelor, Master etc giving numbers, 1,2,3,4,5,6 etc.
3. **Interval Scale:**

“Interval scale indicate the distance between the objects”

Or

 A measurement scale has a constant distance but not a true zero point is called interval scale.

e.g Representation of temperature.

1. **Ratio Scale:**

The ratio scale satisfies all four of the properties of scales of measurement.

It is a special kind of an interval scale where the scale of measurement has a true zero point as its origin.

 It is used to measure weight, volume, length, distance, money etc. zero point make a difference between interval and ratio Scales.

**PRACTICAL EXAMPLES.**

1. In the following examples, the facts and figures usually called statistics presented in the media almost every day are given:
	1. Children who brush their teeth with brand XYZ toothpaste have 60% fewer countries.
	2. The Bureau of census projects the population of Pakistan to be 170.1 million in the year 2010.
	3. Eight out of ten Pakistanis do not have skills.
	4. The prevalence of diabetes is nearly 3 times as high in overweight people as compared to normal people.
	5. In 1980 it was estimated that 0.1% of people have tired any sort of drug, where as in 2008, it was estimated that 10%had done so.
2. Suppose we want to determine the best teacher at Govt. College University, Lahore. How should we decide this?. This could be done by asking Govt. College University students who the best teacher is to do so, we collect the data, analyze the results and make the decision. Now various questions are.
	1. Should we survey every student?
	2. How will the survey be conducted?
	3. How will the data be analyzed?
	4. How will the best teacher be determined? Etc.

In order to answer these and other questions, statistical techniques are used.

1. A TV station claims that on advertisement of a product on their channel attracts more customers compared to all other TV channels, Now if this claim is based on data, there it can be used to market the TV Channel. Suppose we have some doubts about the claim, in order to remove the doubts, we might gather relevant information analyze the result using appropriate statistical technique and make a decision regarding the claim.
2. Suppose university of the Punjab is planning expansion program of its physical facilities. To draw up on effective course of action, the university authorities decide that it needs to answer this question, how many college students will we need to accommodate over the next ten years? The question can be further broken down into many smaller questions. How many college students will then be in the Punjab? How many will want to attend the university of the Panjab? Etc. once again statistical method can assist in evaluation and planning of expansion program.
3. State whether each of the following is a population or sample.
	1. Total number of absentees by all students in a college during the last month.
	2. Number of color TV sets owned by all families in Lahore.
	3. Monthly salaries of all employees of the company.
	4. Wheat yield per acre for 5 pieces of land.
	5. Number of computer sold during the last month at all the computers stores in Lahore.

**Solution:**

* + 1. Population.
		2. Population.
		3. Population.
		4. Sample
		5. Population.
1. Identify each of the following as example of Attribute, discrete or Continuous Variables.
	1. The hair color of children.
	2. Length of time required for a wound to heal.
	3. The numbers of telephone calls arriving at a switch board per 1 hour period.
	4. The breaking strength of a given type of a string.
	5. The number of questions answered correctly on a test.
	6. The number of stop sign in the city of Lahore.
	7. The color of your eye.
	8. The number of children in your family.

**Solution:**

* + 1. Attribute.
		2. Continuous.
		3. Discrete
		4. Continuous
		5. Discrete
		6. Discrete
		7. Attribute
		8. Discrete.