

**Examrace: Downloaded from [examrace.com](http://examrace.com)**

For solved question bank visit [doorsteptutor.com](http://doorsteptutor.com) and for free video lectures visit  
[Examrace YouTube Channel](#)

## Competitive Exams: Statistics Basics

Get unlimited access to the best preparation resource for IAS : Get [detailed illustrated notes covering entire syllabus](#): point-by-point for high retention.

### What is Statistics

The word 'Statistics' is derived from the Latin word 'Statis' which means a "political state." Clearly, statistics is closely linked with the administrative affairs of a state such as facts and figures regarding defense force, population, housing, food, financial resources etc. What is true about a government is also true about industrial administration units, and even one's personal life.

The word statistics has several meanings. In the first place, it is a plural noun which describes a collection of numerical data such as employment statistics, accident statistics, population statistics, birth and death, income and expenditure, of exports and imports etc. It is in this sense that the word 'statistics' is used by a layman or a newspaper.

Secondly the word statistics as a singular noun, is used to describe a branch of applied mathematics, whose purpose is to provide methods of dealing with a collections of data and extracting information from them in compact form by tabulating, summarizing and analyzing the numerical data or a set of observations.

The various methods used are termed as statistical methods and the person using them is known as a statistician. A statistician is concerned with the analysis and interpretation of the data and drawing valid worthwhile conclusions from the same.

It is in the second sense that we are writing this guide on statistics.

Lastly the word statistics is used in a specialized sense. It describes various numerical items which are produced by using statistics (in the second sense) to statistics (in the first sense) . Averages, standard deviation etc. Are all statistics in this specialized third sense.

The word 'statistics' in the first sense is defined by Professor Secrit as follows:

"By statistics we mean aggregate of facts affected to a marked extent by multiplicity of causes, numerically expressed, enumerated or estimated according to reasonable standard of accuracy, collected in a systematic manner for a predetermined purpose and placed in relation to each other."

This definition gives all the characteristics of statistics which are (1) Aggregate of facts (2) Affected by multiplicity of causes (3) Numerically expressed (4) Estimated according to reasonable standards of accuracy (5) Collected in a systematic manner (6) Collected for a predetermined purpose (7) Placed in relation to each other.

The word 'statistics' in the second sense is defined by Croxton and Cowden as follows:

“The collection, presentation, analysis and interpretation of the numerical data.”

This definition clearly points out four stages in a statistical investigation, namely:

1. Collection of data
2. Presentation of data
3. Analysis of data
4. Interpretation of data

## Uses

- To present the data in a concise and definite form: Statistics helps in classifying and tabulating raw data for processing and further tabulation for end users.
- To make it easy to understand complex and large data: This is done by presenting the data in the form of tables, graphs, diagrams etc. or by condensing the data with the help of means, dispersion etc.
- For comparison: Tables, measures of means and dispersion can help in comparing different sets of data.
- In forming policies: It helps in forming policies like a production schedule, based on the relevant sales figures. It is used in forecasting future demands.
- Enlarging individual experiences: Complex problems can be well understood by statistics, as the conclusions drawn by an individual are more definite and precise than mere statements on facts.
- In measuring the magnitude of a phenomenon: Statistics has made it possible to count the population of a country, the industrial growth, the agricultural growth, the educational level (of course in numbers) .

## Limitations

- Statistics does not deal with individual measurements. Since statistics deals with aggregates of facts, it can not be used to study the changes that have taken place in individual cases. For example, the wages earned by a single industry worker at any time, taken by itself is not a statistical datum. But the wages of workers of that industry can be used statistically. Similarly the marks obtained by John of your class or the height of Beena (also of your class) are not the subject matter of statistical

study. But the average marks or the average height of your class has statistical relevance.

- Statistics cannot be used to study qualitative phenomenon like morality, intelligence, beauty etc. As these can not be quantified. However, it may be possible to analyze such problems statistically by expressing them numerically. For example we may study the intelligence of boys on the basis of the marks obtained by them in an examination.
- Statistical results are true only on an average: The conclusions obtained statistically are not universal truths. They are true only under certain conditions. This is because statistics as a science is less exact as compared to the natural science.
- Statistical data, being approximations, are mathematically incorrect. Therefore, they can be used only if mathematical accuracy is not needed.
- Statistics, being dependent on figures, can be manipulated and therefore can be used only when the authenticity of the figures has been proved beyond doubt.

## **Distrust of Statistics**

- It is often said by people that, “statistics can prove anything.” There are three types of lies—lies, demand lies and statistics-wicked in the order of their naming. A Paris banker said, “Statistics is like a miniskirt, it covers up essentials but gives you the ideas.”
- Thus by “distrust of statistics” we mean lack of confidence in statistical statements and methods. The following reasons account for such views about statistics.
- Figures are convincing and, therefore people easily believe them.
- They can be manipulated in such a manner as to establish foregone conclusions.
- The wrong representation of even correct figures can mislead a reader. For example, John earned \$ 4000 in 1990 – 1991 and Jem earned \$ 5000. Reading this one would form the opinion that Ram is decidedly a better worker than Sohan. However if we carefully examine the statement, we might reach a different conclusion as Jem’s earning period is unknown to us. Thus while working with statistics one should not only avoid outright falsehoods but be alert to detect possible distortion of the truth.

## **Statistics Can be Misused**

In one factory which I know, workers were accusing the management for not providing them with proper working conditions. In support they quoted the number of accidents. When I considered the matter more seriously, I found that most of the staff was inexperienced and thus responsible for those accidents. Moreover many of the accidents were either minor or fake. I compared the working conditions of this factory to other factories and I found the conditions far better in this factory. Thus by merely noting the number of accidents and complaints of the workers, I would not dare to say that the

working conditions were worse. On the other hand due to the proper statistical knowledge and careful observations I came to conclusion that the management was right.

Thus the usefulness of the statistics depends to a great extent upon its user. If used properly, by an efficient and unbiased statistician, it will prove to be an efficient tool.

Collection of facts and figures and deriving meaningful information from them is an important program.

Often it is not possible or practical to record observations of all the individuals of the groups from different areas, which comprise the population. In such a case observations are recorded of only some of the individuals of the population, selected at random. This selection of some individuals which will be a subset of the individuals in the original group, is called a Sample; i.e.. . Instead of an entire population survey which would be time-consuming, the company will manage with a 'Sample survey' which can be completed in a shorter time.

Note that if a sample is representative of the whole population, any conclusion drawn from a statistical treatment of the sample would hold reasonably good for the population. This will of course, depend on the proper selection of the sample. One of the aims of statistics is to draw inferences about the population by a statistical treatment of samples.

## **Types of Statistics**

As mentioned earlier, for a layman or people in general, statistics means numbers-numerical facts, figures or information. The branch of statistics wherein we record and analyze observations for all the individuals of a group or population and draw inferences about the same is called "Descriptive statistics" or "Deductive statistics" On the other hand, if we choose a sample and by statistical treatment of this, draw inferences about the population, then this branch of statistics is known as Statical Inference or Inductive Statistics.

In our discussion, we are mainly concerned with two ways of representing descriptive statistics: Numerical and Pictorial.

Numerical statistics are numbers. But some numbers are more meaningful such as mean, standard deviation etc.

When the numerical data is presented in the form of pictures (diagrams) and graphs, it is called the Pictorial statistics. This statistics makes confusing and complex data or information, easy, simple and straightforward, so that even the layman can understand it without much difficulty.

## **Common Mistakes Committed in Interpretation of Statistics**

- **Bias:** Bias means prejudice or preference of the investigator, which creeps in consciously and unconsciously in proving a particular point.
- **Generalization:** Some times on the basis of little data available one could jump to a conclusion, which leads to erroneous results.
- **Wrong conclusion:** The characteristics of a group if attached to an individual member of that group, may lead us to draw absurd conclusions.
- **Incomplete classification:** If we fail to give a complete classification, the influence of various factors may not be properly understood.

## Charateristics of Incorrect Classification

- There may be a wrong use of percentages.
- Technical mistakes may also occur.
- An inconsistency in definition can even exist.
- Wrong causal inferences may sometimes be drawn.
- There may also be a misuse of correlation.

## Glossary of Terms

- **Statistics:** Statistics is the use of data to help the decision maker to reach better decisions.
- **Data:** It is any group of measurements that interests us. These measurements provide information for the decision maker (I) . The data that reflects non-numerical features or qualities of the experimental units, is known as qualitative data (ii) . The data that possesses numerical properties is known as quantitative data.
- **Population:** Any well defined set of objects about which a statistical enquiry is being made is called a population or universe. The total number of objects (individuals) in a population is known as the size of the population. This may be finite or infinite.
- **Individual:** Each object belonging to a population is called as an individual of the population.
- **Sample:** A finite set of objects drawn from the population with a particular aim, is called a sample. The total number of individuals in a sample is called the sample size.
- **Characteristic:** The information required from an individual, from a population or from a sample, during the statistical enquiry (survey) is known as the characteristic of the individual. It is either numerical or non-numerical. For e. g. The size of shoes is a numerical characteristic which refers to a quantity, whereas the mother tongue of a person is a non-numerical characteristic which refers to a quality. Thus we have quantitative and qualitative types of characteristics.

- **Variate:** A quantitative characteristic of an individual which can be expressed numerically is called a variate or a variable. It may take different values at different times, places or situations.
- **Attribute:** A qualitative characteristic of an individual which can be expressed numerically is called an attribute. For e. g. The mother-tongue of a person, the color of eyes or the color of hair of a person etc.
- **Discrete variate:** A variable that is not capable of assuming all the values in a given range is a discrete variate.
- **Continuous Variate:** A variate that is capable of assuming all the numerical values in a given range, is called a continuous variate. Consider two examples carefully, viz. The number of students of a class and their heights. Both variates differ slightly, in the sense that, the number of students present in a class is a number say between 0 and 50; always a whole number. It can never be 1.5, 4.33 etc. This type of variate can take only isolated values and is called a discrete variate. On the other hand heights ranging from 140 cm to 190 cm can take values like 140.7, 135.8, 185.1 etc. Such a variate is a continuous variate.

Developed by: [Mindsprite Solutions](#)