

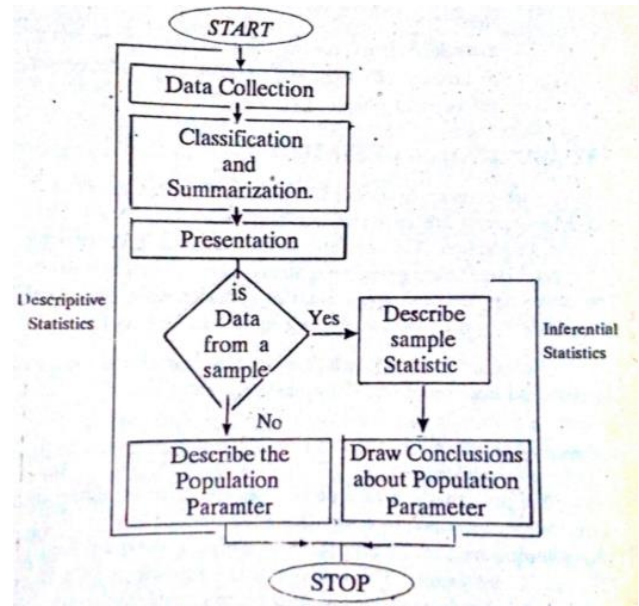
Introduction to Statistics and its types

1 Statistics

Statistics is a branch of math focused on collecting, organizing, and understanding numerical data. It involves analyzing and interpreting data to solve real-life problems.

1.1 Statistics Terminologies

- **Population:** It is actually a collection of a set of individual objects or events whose properties are to be analyzed.
- **Sample:** It is the part of that population which is to be studied.
- **Variable:** It is a characteristic of the individual or of a sample which varies from individual to individual. E.g. heights of the students, marks obtained by the student, price of the shirts in market, titles of book in the library.
- **Data:** Observations are recorded values of a characteristic of every individual of a population or a sample these observations are called Data



1.2 Types of data:

Discrete data: This can only have specific values (number of students in class room, number of defective items etc)

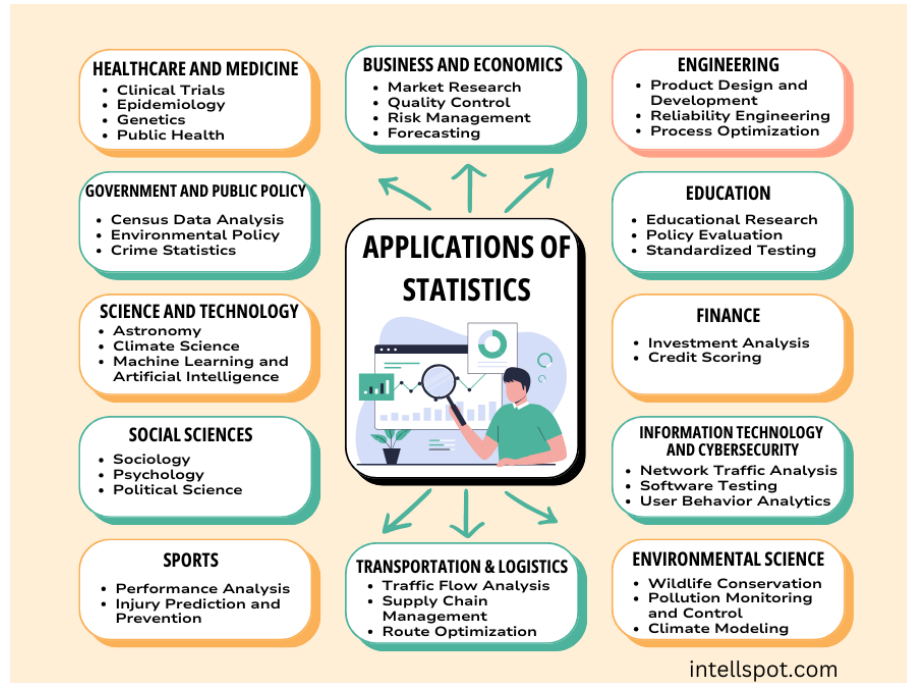
Continuous data: This can have any value. This is often called measurement data. e.g. height, weight, temperature, length .

These two data types can be further subdivided as follows:

	Data type	Definitions	Examples
Discrete	Category data - without order (also called Nominal)	Data has a name only	Street, Road, Male, Female etc
	Category data - ordered (also called Ordinal)	Data has order, but does not have a numerical scale	Very happy, Happy, Unhappy, Very unhappy etc
	Whole number data	Data can have any whole number value	Day 1, Day 2, Day 3, ... (simple time series data) Number of books (0, 1, 2, 3, ...)
Continuous	Measurement data	Data can take any numerical value	Length of a pencil. It can be 8 cm, 9.1 cm, 9.48m. Time in seconds.

1.3 Application of Statistics

Statistics finds applications in diverse fields like weather forecasting, stock market analysis, insurance, betting, and data science.



1.4 Types of Statistics

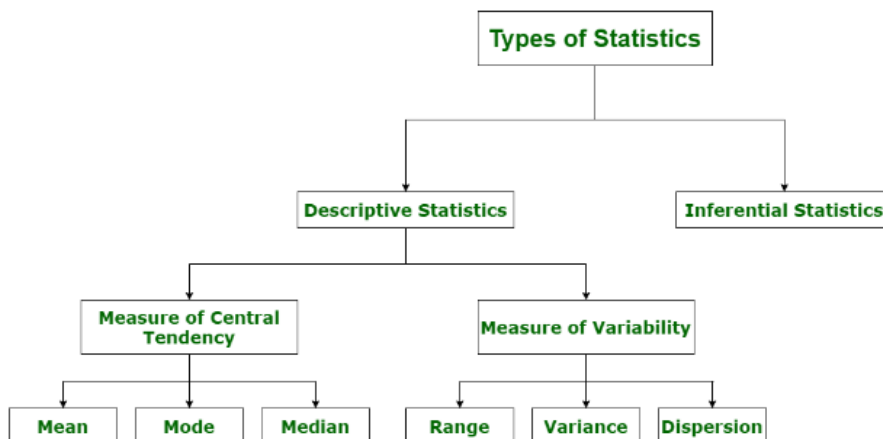
There are 2 types of statistics:

Descriptive Statistics

show and summarize data in a meaningful way. It is a simple way to describe our data. Descriptive statistics is very important to present our raw data ineffective/meaningful way using numerical calculations or graphs or tables.

Inferential Statistics

Used to describe and make inferences about the population. It basically allows you to make predictions by taking a small sample instead of working on the whole population.



1.4.1 Difference between descriptive and inferential Statistics

S.No.	Descriptive Statistics	Inferential Statistics
1.	It gives information about raw data which describes the data in some manner.	It makes inferences about the population using data drawn from the population.
2.	It helps in organizing, analyzing, and to present data in a meaningful manner.	It allows us to compare data, and make hypotheses and predictions.
3.	It is used to describe a situation.	It is used to explain the chance of occurrence of an event.
4.	It explains already known data and is limited to a sample or population having a small size.	It attempts to reach the conclusion about the population.
5.	It can be achieved with the help of charts, graphs, tables, etc.	It can be achieved by probability.

1.4.2 Real life example of descriptive statistics

Imagine an example where a company sells hot sauce. The company gathers data such as the count of sales, average quantity purchased per transaction, and average sale per day of the week. All of this information is descriptive, as it tells a story of what actually happened in the past.

1.4.3. Real life examples of inferential statistics

Here are some examples of how inferential statistics are applied:

- **Healthcare:** Inferential statistics are widely used in healthcare to analyze the effectiveness of new drugs, medical treatments, and interventions. For example, clinical trials use inferential statistics to determine whether a new drug is effective in treating a particular medical condition.
- **Business:** Inferential statistics are used in business to make predictions and forecast trends. Companies use inferential statistics to estimate demand for products and services, forecast sales and revenue, and identify opportunities for growth.
- **Social Sciences:** Inferential statistics are used in the social sciences to test hypotheses, estimate population parameters, and make predictions about social phenomena. For example, inferential statistics are used in psychology to determine whether a particular therapy is effective in treating a mental illness.

1.5 Classification of data

The observation when recorded generally unorganized and without arrangement. The first step towards our studies is to classify these raw data into compact form for further analysis.

1.5.1 Frequency distribution

A grouped frequency distribution is a tabular summary of raw data (ungrouped data) showing the frequencies of values in each of several non-overlapping intervals (or classes). A frequency distribution is referred as grouped data.

(example of frequency distribution is shared in a separate document)

1.5.2 Graphical presentation

It includes histogram, frequency polygons etc

Document Summary

DATA		Facts or figures, which are numerical or otherwise, collected with a definite purpose are called data.	
Types Of Data			
Quantitative Data	Qualitative Data	Primary Data	Secondary Data
These represent numerical value.	These represent some characteristics or attributes.	Data collected for first time.	Data that is sourced by someone other than the user.
These can be numerically computed.	These depict descriptions that may be observed but cannot be computed.	Discrete Data	Continuous Data
		These are the data that can take only specific value.	These are the data that can take values from a given range.
Frequency Distribution Table	A list, table or graph that displays the frequency of various outcomes in a sample of data.		

Frequency Distribution Table

Ungrouped

It is used for small data set.
For eg.

Marks Obtained	Frequency
16	3
17	4
18	8
19	10
20	12
21	6
22	3

Grouped

It is used for large data set.
For eg.

Class Interval	Frequency
0-5	3
5-10	11
10-16	14
15-20	2

