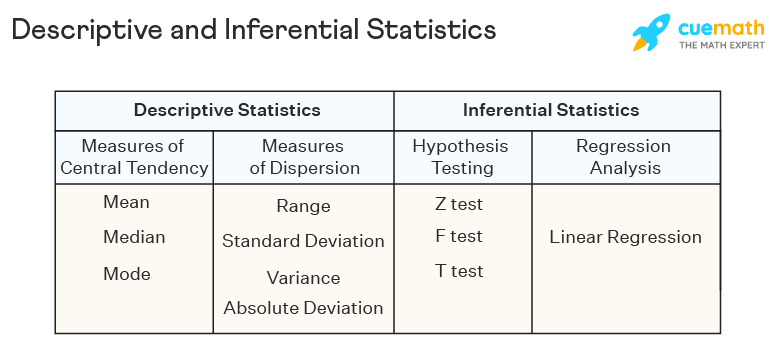
**Descriptive and Inferential Statistics**

Descriptive and inferential statistics are two fields of statistics. Descriptive statistics is used to describe data and inferential statistics is used to make predictions. Descriptive and inferential statistics have different tools that can be used to draw conclusions about the data.

In descriptive and inferential statistics, the former uses tools such as central tendency, and dispersion while the latter makes use of hypothesis testing, regression analysis, and confidence intervals. In this article, we will learn more about descriptive and inferential statistics, its differences, associated formulas and examples.

**What is Descriptive and Inferential Statistics?**

The purpose of descriptive and inferential statistics is to analyze different types of data using different tools. Descriptive statistics helps to describe and organize known data using charts, bar graphs, etc., while inferential statistics aims at making inferences and generalizations about the population data.



**Descriptive Statistics**

Descriptive statistics are a part of statistics that can be used to describe data. It is used to summarize the attributes of a sample in such a way that a pattern can be drawn from the group. It enables researchers to present data in a more meaningful way such that easy interpretations can be made. Descriptive statistics uses two tools to organize and describe data. These are given as follows:

* Measures of Central Tendency - These help to describe the central position of the data by using measures such as [mean](https://www.cuemath.com/data/mean/), [median](https://www.cuemath.com/data/median/), and [mode](https://www.cuemath.com/data/mode/).
* Measures of Dispersion - These measures help to see how spread out the data is in a distribution with respect to a central point. [Range](https://www.cuemath.com/data/range-in-statistics/), standard deviation, [variance](https://www.cuemath.com/data/variance/), quartiles, and [absolute deviation](https://www.cuemath.com/mean-deviation-formula/) are the measures of dispersion.

**Inferential Statistics**

Inferential statistics is a branch of statistics that is used to make inferences about the population by analyzing a sample. When the population data is very large it becomes difficult to use it. In such cases, certain samples are taken that are representative of the entire population. Inferential statistics draws conclusions regarding the population using these samples. Sampling strategies such as simple random sampling, cluster sampling, stratified sampling, and systematic sampling, need to be used in order to choose correct samples from the population. Some methodologies used in inferential statistics are as follows:

* [Hypothesis Testing](https://www.cuemath.com/data/hypothesis-testing/) - This technique involves the use of hypothesis tests such as the [z test](https://www.cuemath.com/data/z-test/), [f test](https://www.cuemath.com/data/f-test/), t test, etc. to make inferences about the population [data](https://www.cuemath.com/data/). It requires setting up the null hypothesis, alternative hypothesis, and testing the decision criteria.
* Regression Analysis - Such a technique is used to check the relationship between dependent and independent variables. The most commonly used type of regression is linear regression.

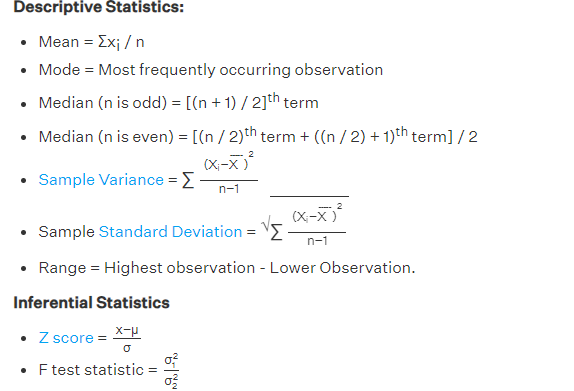
**Difference Between Descriptive and Inferential Statistics**

Both descriptive and inferential statistics are equally important to analyze data. Descriptive statistics are used to order data and describe the sample using the mean, standard deviation, charts, etc. Inferential statistics uses this sample data to predict the trend of the population data. The differences between descriptive and inferential statistics have been outlined in the table given below:

| **Basis** | **Descriptive Statistics** | **Inferential Statistics** |
| --- | --- | --- |
| Definition | Descriptive statistics is used to describe the characteristics of the population using a sample. | Inferential statistics uses various analytical tools to draw inferences about the population using samples. |
| Tools | [Measures of central tendency](https://www.cuemath.com/data/measures-of-central-tendency/) and measures of dispersion. | Hypothesis testing and regression analysis. |
| Use | Organizes, describes and presents data in a meaningful way with the help of charts and graphs. | Tests, predicts, and compares data obtained from various samples. |
| Relevance | It is used to summarize known data in a way that can be used for further predictions and analysis. | It tries to use the summarized samples to draw conclusions about the population. |

**Descriptive and Inferential Statistics Formulas**

There are many [statistical formulas](https://www.cuemath.com/statistics-formulas/) that fall under descriptive and inferential statistics. These are given as follows:

****

**Examples of Descriptive and Inferential Statistics**

Descriptive and inferential statistics need to be used hand in hand so as to analyze the data in the best possible way. Some examples of descriptive and inferential statistics are given below:

* Suppose the scores of 100 students belonging to a specific country are available. The performance of these students needs to be examined. This data by itself will not yield any valuable results. However, by using descriptive statistics, the spread of the marks can be obtained thus, giving a clear idea regarding the performance of each student.
* Now suppose the scores of the students of an entire country need to be examined. Using a sample of, say 100 students, inferential statistics is used to make generalizations about the population.

**Important Notes on Descriptive and Inferential Statistics**

* Descriptive and inferential statistics are used to analyze data, obtain samples and make inferences about the population.
* The tools used in descriptive statistics are measures of central tendency and dispersion.
* The tools used in inferential statistics are hypothesis testing and regression analysis.

How To Find Mean value(average)

**Mean**=Total element/number of element

import numpy as np

x=[2,4,6,7,20,10,22]

y=np.array(x)

print("Mean is : ",round(y.mean(),2))

**Median**

Median(n is odd)= [(n+1)/2]th

n is Total number of element

For Example

x=[2,4,6,7,20,10,22]

n=7

[(7+1)/2]th

[4]th means what 4th index of an array

It will return value from the 4th index

**OUTPUT is 7**

Median(n is Even)=[(n/2)th term + ((n/2)+1)th term]/2

For Example

x=[2,4,6,7,20,10]

n=6

[ (6/2)th term + ((6/2)+1)th term ] / 2

[(3)th term + ((3)+1)th term]/2

[6 + 7]/2

**Output is 6.5**

**Mode:**

Most Repeated number

from statistics import mode

print("Mode is:",mode([1, 1, 2, 3, 3, 3, 3, 4]))

output is 3

**Measure of Dispersion**

Dispersion measure the extent to which the item vary from some central value(Mean value)

Dispersion is know as Scatter , spread or variation

Item is known as Observation

Series 1 [100,100,100,100,100]

Series 2 [98,99,100,101,102]

Series 3 [1,2,3,4,490]

Series 1 total 500 and avg 100 , measure of spread 0,

Series 2 total 500 and avg 100, measure of spread less, stable series

Series 3 total 500 and avg 100, measure of spread very high, unstable series

**Types of Measure of Dispersion**

|  |  |
| --- | --- |
| **Absolute Meaure(Unit same)** | **Relative Measure(Unit Free)** |
| Range | Coefficient of Range |
| Quartile Deviation | Coefficient of Quartile |
| Mean Deviation | Coefficient of Mean Deviation |
| Standard Deviation | Coefficient of Standard Deviation |
| Variance | Coefficient of variance |

**Measure of Dispersion :** Userful for Comparision of Two series

**Absolute Measure :** it is useful for comparision of two series with same unit like KG,

**Relative Measure** : It is useful for Comparision of two series with different unit,

* 1. Range:

Range = Largest – Smallest

Co of range = L-S/L+S \* 100

Individual series

Ex. 13,20,7,15,29,35

L=35, S=7

Range= 35-7 = 28

Co of range= (35-7)\*100/35+7

* 28\*100/42
* 2800/42
* 66.67

-

Uses:

1. Quality control (Accuracy 90-95)
2. Fluctuate in Share Price (day high and day low)
3. Weather Forecasting (min temp, max temp)

**Quartile Deviation:**

**QD=**Q3-Q1/2

Co of QD= (Q3-Q1)\*100/Q3+Q1

Q1=First Quartile(lower quartile)

Q2=Second Quartile

Q3=Third Quartile(Upper quartile)

How to find Q1 individual series

Q1=[(n+1)/4]th term

Q3=3\*[(n+1)/4]th term

Ex.

Individual Series = 5,9,20,35,40,3,10

Arrange: 3, 5, 9, 10, 20, 35, 40

N=7

Q1=(7+1)/4

* (8/4)th term
* 2 th term
* 5