

Descriptive Statistics

Descriptive statistics is a branch of statistics that is concerned with describing the characteristics of the known data. Descriptive statistics provides summaries about either the population data or the sample data. Apart from descriptive statistics, inferential statistics is another crucial branch of statistics that is used to make inferences about the population data.

Descriptive statistics can be broadly classified into two categories - measures of central tendency and measures of dispersion. In this article, we will learn more about descriptive statistics, its various types, formulas, and see associated examples.

Types of Descriptive Statistics

Measures of central tendency and measures of dispersion are two types of descriptive statistics that are used to quantitatively summarize the characteristics of grouped and ungrouped data. When an experiment is conducted, the raw data obtained is known as ungrouped data. When this data is organized logically it is known as grouped data. To visually represent data, descriptive statistics use graphs, charts, and tables. Some important types of descriptive statistics are given below.

Descriptive Statistics	
Measures of Central Tendency	Measures of Dispersion
Mean	Range
Median	Standard Deviation
Mode	Quartile Deviation
	Variance
	Absolute Deviation

Measures of Central Tendency

In descriptive statistics, the measures of central tendency are used to describe data by determining a single representative central value. The important measures of central tendency are given below:

Mean: The mean can be defined as the sum of all observations divided by the total number of observations.

Median: The median can be defined as the center-most observation that is obtained by arranging the data in ascending order. The formulas for the median are given as follows:

Ungrouped data Median (n is odd): $[(n + 1) / 2]^{\text{th}}$ term

Ungrouped data Median (n is even): $[(n / 2)^{\text{th}}$ term + $((n / 2) + 1)^{\text{th}}$ term] / 2

Grouped data Median: $l + [((n / 2) - c) / f] \times h$

l is the lower limit of the median class given by $n / 2$, c is the cumulative frequency, f is the frequency of the median class and h is the class height.

Mode: The mode is the most frequently occurring observation in the data set.

Measures of Dispersion

In descriptive statistics, the measures of dispersion are used to determine how spread out a distribution is with respect to the central value. The important measures of dispersion are given below:

Range: The range can be defined as the difference between the highest value and the lowest value. The formula is given as follows:

$$\text{Range} = \text{HV} - \text{LV}$$

HV is the highest value and LV is the lowest value in a data set.

Variance: The variance gives the variability of the distribution with respect to the mean.

Standard Deviation: The square root of the variance will result in the standard deviation. It helps to analyze the variability in a data set in a more effective manner as compared to the variance. The formula is given as follows:

Standard Deviation: $S.D. = \sqrt{\text{Variance}} = \sigma$

Mean Deviation: The mean deviation will give the average of the absolute value of the data about the mean, median, or mode. It is also known as absolute deviation.

Quartile Deviation: Half of the difference between the third and first quartile gives the quartile deviation.

