The term “descriptive statistics” refers to the **analysis, summary, and presentation of findings** related to a data set derived from a sample or entire population. Descriptive statistics comprises three main categories –

Frequency Distribution

Measures of Central Tendency

and Measures of Variability.

Descriptive statistics helps facilitate data visualization. It allows for data to be presented in a meaningful and understandable way, which, in turn, allows for a simplified interpretation of the data set in question.

Frequency Distribution

Used for both quantitative and qualitative data, frequency distribution depicts the frequency or count of the different outcomes in a data set or sample. The frequency distribution is normally presented in a table or a graph. Each entry in the table or graph is accompanied by the count or frequency of the values’ occurrences in an interval, range, or specific group.

Frequency distribution is basically a presentation or summary of grouped data categorized based on mutually exclusive classes and the number of occurrences in each respective class. It allows for a more structured and organized way to present raw data.

Common charts and graphs used in frequency distribution presentation and visualization include bar charts, histograms, pie charts, and line charts.

A frequency distribution is a representation, either in a graphical or tabular format, that displays the number of observations within a given interval. The interval size depends on the data being analyzed and the goals of the analyst. The intervals must be mutually exclusive and exhaustive. Frequency distributions are typically used within a statistical context. Generally, frequency distributions can be associated with the charting of a normal distribution.

KEY TAKEAWAYS

A frequency distribution in statistics is a representation that displays the number of observations within a given interval.

The representation of a frequency distribution can be graphical or tabular so that it is easier to understand.

Frequency distributions are particularly useful for normal distributions, which show the observations of probabilities divided among standard deviations.

In finance, traders use frequency distributions to take note of price action and identify trends.

Understanding a Frequency Distribution

As a statistical tool, a frequency distribution provides a visual representation of the distribution of observations within a particular test. Analysts often use a frequency distribution to visualize or illustrate the data collected in a sample. For example, the height of children can be split into several different categories or ranges.

In measuring the height of 50 children, some are tall and some are short, but there is a high probability of a higher frequency or concentration in the middle range. The most important factors for gathering data are that the intervals used must not overlap and must contain all of the possible observations.

Visual Representation of a Frequency Distribution

Both histograms and bar charts provide a visual display using columns, with the y-axis representing the frequency count, and the x-axis representing the variable to be measured. In the height of children, for example, the y-axis is the number of children, and the x-axis is the height. The columns represent the number of children observed with heights measured in each interval.

In general, a histogram chart will typically show a normal distribution, which means that the majority of occurrences will fall in the middle columns. Frequency distributions can be a key aspect of charting normal distributions which show observation probabilities divided among standard deviations.

Frequency distributions can be presented as a frequency table, a histogram, or a bar chart. Below is an example of a frequency distribution as a table.

Height of Children in a School

Interval (Height) 4' 4'5" 5' 5'2"

Frequency 2 5 6 3

Frequency Distribution in Trading

Frequency distributions are not commonly used in the world of investments; however, traders who follow Richard D. Wyckoff, a pioneering early 20th-century trader, use an approach to trading that involves frequency distribution.

Investment houses still use the approach, which requires considerable practice, to teach traders. The frequency chart is referred to as a point-and-figure chart and was created out of a need for floor traders to take note of price action and to identify trends.1

The y-axis is the variable measured, and the x-axis is the frequency count. Each change in price action is denoted in Xs and Os. Traders interpret it as an uptrend when three X's emerge; in this case, demand has overcome supply. In the reverse situation, when the chart shows three O's, it indicates that supply has overcome demand.1

What Are the Types of Frequency Distribution?

The types of frequency distribution are grouped frequency distribution, ungrouped frequency distribution, cumulative frequency distribution, relative frequency distribution, and relative cumulative frequency distribution.

What Is the Importance of a Frequency Distribution?

A frequency distribution is a means to organize a large amount of data. It takes data from a population based on certain characteristics and organizes the data in a way that is comprehensible to an individual that wants to make assumptions about a given population.

How Can I Construct a Frequency Distribution?

To construct a frequency distribution, first, note the specific classes determined by intervals in one column then sum the numbers in each isolated category based on how many times it shows up. The frequency can then be noted in the second column.

The Bottom Line

A frequency distribution is used to display the number of observations within a particular interval. This method, while not always commonly used in investing, is still used by some traders. In this case, the frequency chart is called a point-and-figure chart and is used to identify trends through the observation of price action.

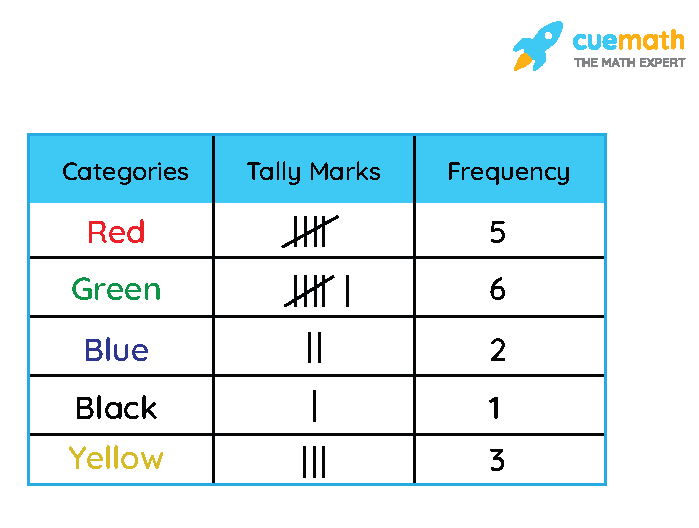
**Types of Frequency Distribution**

There are four types of frequency distribution under statistics which are explained below:

* **Ungrouped frequency distribution:**It shows the frequency of an item in each separate data value rather than groups of data values.
* **Grouped frequency distribution:** In this type, the data is arranged and separated into groups called class intervals. The frequency of data belonging to each class interval is noted in a frequency distribution table. The grouped frequency table shows the distribution of frequencies in class intervals.
* [Relative frequency](https://www.cuemath.com/data/relative-frequency/) **distribution:**It tells the proportion of the total number of observations associated with each category.
* [Cumulative frequency](https://www.cuemath.com/data/cumulative-frequency/) **distribution:**It is the sum of the first frequency and all frequencies below it in a frequency distribution. You have to add a value with the next value then add the sum with the next value again and so on till the last. The last cumulative frequency will be the total sum of all frequencies.

**Frequency Distribution Table**

A frequency distribution table is a chart that shows the frequency of each of the items in a data set. Let's consider an example to understand how to make a frequency distribution table using tally marks. A jar containing beads of different colors- red, green, blue, black, red, green, blue, yellow, red, red, green, green, green, yellow, red, green, yellow. To know the exact number of beads of each particular color, we need to classify the beads into categories. An easy way to find the number of beads of each color is to use [tally marks](https://www.cuemath.com/data/tally-marks/). Pick the beads one by one and enter the tally marks in the respective row and column. Then, indicate the frequency for each item in the table.



Thus, the table so obtained is called a [frequency distribution table](https://www.cuemath.com/data/frequency-distribution-table/).

**Types of Frequency Distribution Table**

There are two types of frequency distribution tables: Grouped and ungrouped frequency distribution tables.

**Grouped Frequency Distribution Table:**To arrange a large number of observations or data, we use grouped frequency distribution table. In this, we form class intervals to tally the frequency for the data that belongs to that particular class interval.

For example, Marks obtained by 20 students in the test are as follows. 5, 10, 20, 15, 5, 20, 20, 15, 15, 15, 10, 10, 10, 20, 15, 5, 18, 18, 18, 18. To arrange the data in grouped table we have to make class intervals. Thus, we will make class intervals of marks like 0 – 5, 6 – 10, and so on. Given below table shows two columns one is of class intervals (marks obtained in test) and the second is of frequency (no. of students). In this, we have not used tally marks as we counted the marks directly.

| **Marks obtained in Test (class intervals)** | **No. of Students (Frequency)** |
| --- | --- |
| 0 – 5 | 3 |
| 6 – 10 | 4 |
| 11 – 15 | 5 |
| 16 – 20 | 8 |
| Total | 20 |

**Ungrouped Frequency Distribution Table:**In the ungrouped frequency distribution table, we don't make class intervals, we write the accurate frequency of individual data. Considering the above example, the ungrouped table will be like this. Given below table shows two columns: one is of marks obtained in the test and the second is of frequency (no. of students).

| **Marks obtained in Test** | **No. of Students** |
| --- | --- |
| 5 | 3 |
| 10 | 4 |
| 15 | 5 |
| 18 | 4 |
| 20 | 4 |
| Total | 20 |

**Important Notes:**

Following are the important points related to frequency distribution.

* Figures or numbers collected for some definite purpose is called data.
* Frequency is the value in numbers that shows how often a particular item occurs in the given data set.
* There are two types of frequency table - Grouped Frequency Distribution and Ungrouped Frequency Distribution.
* Data can be shown using graphs like histograms, bar graphs, frequency polygons, and so on.

A variable is a measurable characteristic which changes from one member of a sample or a population to another, for example, age of a person, GDP of a country.

A continuous variable is a measurable characteristic which potentially can take any value in a continuous range, without any breaks or jumps.

A discrete variable is a measurable characteristic which is restricted to a specific set of values. Discrete data are often represented by bar charts, showing Continuous data are usually represented as histograms, showing frequency density per unit of the variable.

Descriptive statistics has great relevance in maths because simply presenting raw data cannot be understood by people. Descriptive statistics enables people to present the data in a more expressive way, which allows simpler interpretation of the data.

Descriptive statistics makes use of graphical techniques and numerical descriptive measures to summarize and present the data.

**Tabular Methods**

Tabular method of data presentation is wide spread in all spheres of human life. These methods are used to summarize data from a sample or population into table format. Data is grouped into categories and the number (or frequency) of observations in each category is obtained.  
Frequency distribution is a type of tabular method. A frequency distribution is a tabular summary of data showing the frequency of items in each of several non-overlapping classes. The objective is to provide insights about the data that cannot be quickly obtained by looking only at the original data.

**Graphical Methods**

These methods are applied to visually describe data from a sample or population. The shape of sample data can indicate the shape of the population from which it is taken. Graphs provide visual summaries of data which is more quickly and completely describe essential information than tables of numbers.

Graphs are essential as these provide insight for the analyst into the data under scrutiny, and illustrate important concepts when presenting the results to others. A graphical method is developed which signifies the accuracy of test results. The graphs can be constructed from Producer's scores and Consumer's scores on each of the scales of test score, antigen dose and probability of protection against disease.

There are many types of graphical representation:

1. The Bar Chart: To Construct a Bar Chart, place categories on the horizontal axis, then place frequency (or relative frequency) on the vertical axis. After that construct vertical bars of equal width, one for each category. Its height is proportional to the frequency (or relative frequency) of the category.
2. The Pie Chart: For drawing pie chart, make complete circle that represents the total number of measurements. Partition into slices - one for each category. Then, the size of a slice is proportional to the relative frequency of that category. Determine the angle of each slice by multiplying the relative frequency by 360 degree.

Graphical Methods for Quantitative Data include Stem-and-leaf plot and Histogram.

1. Stem-and-leaf plot: Steps for Constructing Stem and Leaf Display are as follows:
   1. Break up each data into two pieces: Stem and leaf. To do this, select one or more leading digits of the data for the stem. The trailing digit or digits become leaves.
   2. List possible stem values in a (vertical) column, with the smallest stem on top.
   3. Record the leaf corresponding to each stem beside it in a row.
   4. Indicate the units for stems and leaves somewhere in the display.
   5. In general, we want the number of stems to be between 5 and 20, if possible.
2. Histogram: A histogram is a graphical representation of a frequency (or relative frequency distribution). A histogram displays the shape of the data. It is useful when it is logical to group data into numerical categories.
3. Quantile plots: These visually portray the quantiles, or percentiles (which equals to the quantiles times 100) of the distribution of sample data. Quantiles of importance such as the median are easily discerned (quantile, or cumulative frequency = 0.5). Main benefits of Quantile plots are as follows:  
   1. Arbitrary categories are not required, as with histograms or S-L's.
   2. All of the data are displayed, unlike a boxplot.
   3. Every point has a distinct position, without overlap.
4. Boxplots: Boxplot is a very useful and brief graphical display for summarizing the distribution of a data set. Boxplots provide visual summaries of the centre of the data (the median-the centre line of the box), the variation or spread (interquartile range-the box height), the skewness (quartile skew-the relative size of box halves) and presence or absence of unusual values ("outside" and "far outside" values). Boxplots are even more useful in comparing these attributes among several data sets.

Graphical representation of reports has numerous benefits.

1. Acceptability: Graphical report is acceptable to people who have busy schedule because it easily highlights about the theme of the report. This helps to avoid wastage of time.
2. Comparative Analysis: Information can be compared in terms of graphical representation. Such comparative analysis helps for quick understanding and attention.
3. Less cost: Information, if descriptive, involves huge time to present properly. It involves more money to print the information but graphical presentation can be made in short but catchy view to make the report understandable. It obviously involves less cost.
4. Decision Making: Business executives can view the graphs at a glance and can make decision very quickly which is hardly possible through descriptive report.
5. Logical Ideas: If tables, design and graphs are used to represent information then a logical sequence is created to clear the idea of the audience.
6. Helpful for less educated Audience: Less literate or illiterate people can understand graphical representation easily because it does not involve going through line by line of any descriptive report.
7. Less Effort and Time: To present any table, design, image or graphs require less effort and time. Furthermore, such presentation makes quick understanding of the information.
8. Less Error and Mistakes: Qualitative or informative or descriptive reports involve errors or mistakes. As graphical representations are exhibited through numerical figures, tables or graphs, it usually involves less error and mistake.
9. A complete Idea: Such representation creates clear and complete idea in the mind of audience. Reading hundred pages may not give any scope to make decision. But an instant view or looking at a glance obviously makes an impression in the mind of audience regarding the topic or subject.
10. Use in the Notice Board: Such representation can be hanged in the notice board to quickly raise the attention of employees in any organization.

**Graphical representation of reports has some drawbacks also:**

1. Expensive: Graphical representations of reports are costly because it involves images, colours and paints. Combination of material with human efforts makes the graphical presentation expensive.
2. More time: Graphical representation involves more time as it requires graphs and figures which are dependent to more time.
3. Errors and Mistakes: Since graphical representations are complex, there is chance of errors and mistake. This causes problems for better understanding to general people.
4. Lack of Privacy: Graphical representation makes full presentation of information which may hamper the objective to keep something secret.
5. Problems to select the appropriate method: Information can be presented through various graphical methods and ways. Which should be the suitable method is very hard to select.
6. Problem of Understanding: All people cannot understand the meaning of graphical representation because it involves various technical matters which are complex to general people.

**Numerical Methods**

These procedures are used to arithmetically describe data from a sample or population. The numerical measures of a sample can be used to estimate the corresponding numerical measures of the population. The numerical methods can be effectively demonstrated in cases dealing with complex problems for which analytical solutions cannot be obtained or hand calculations cannot be made.

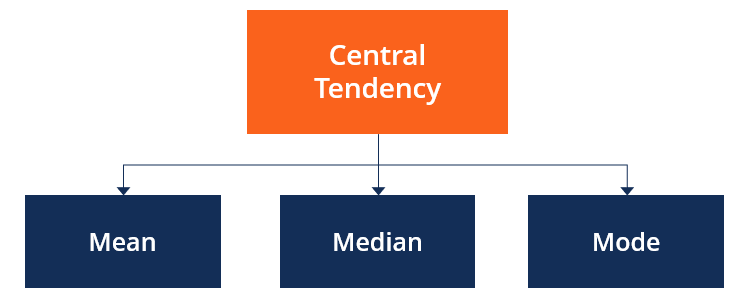
Characteristically, there are two general types of statistic that are used to describe data:  
Measures of central tendency: These are ways of describing the central position of a frequency distribution for a group of data. Measures of central tendency are numbers that tend to cluster around the "middle" of a set of values. These include the mode, median, and mean.

Mean is the average value, calculated by adding all the observations and dividing by the number of observations. A drawback of the mean is that it is heavily influenced by extreme observations. The median is explained as the middle value when observations are arranged in an ascending or descending order. The median is easy to understand, and it is not greatly affected by extreme observations. It is often used in preference to the mean when extreme observations are present. The mode is described as the most common value of individually recorded observations and as the value of the variable for which the frequency density is greatest for grouped data.

Measures of spread: In this type of statistic, group of data is summarized by describing how spread out the scores are. To describe this spread, a number of statistics are available such as the range, quartiles, absolute deviation, variance and standard deviation. When descriptive statistics is used, it is useful to summarize group of data using a combination of tabulated description, graphical description such as graphs and charts and statistical commentary such as discussion of the results.

To summarize, Descriptive statistics consists of statistical procedures that are used to describe the population that are studying. The data could be collected from either a sample or a population, but the results help statistician organize and describe data. Descriptive statistics can only be used to describe the group that is being studying. That is, the results cannot be generalized to any larger group. There are three methods of descriptive statistics that include tabular, graphical and numerical methods.

**Central Tendency**



Central tendency refers to a dataset’s descriptive summary using a single value reflecting the center of the data distribution. Measures of central tendency are also known as measures of central location. The mean, [median](https://corporatefinanceinstitute.com/resources/knowledge/other/median/), and mode are the measures of central tendency.

The mean, considered the most popular measure of central tendency, is the average or most common value in a data set. The median refers to the middle score for a data set in ascending order. The mode refers to the score or value that is most frequent in a data set.

**Measures of Central Tendency**

In descriptive statistics, the [measures of central tendency](https://www.cuemath.com/data/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](https://www.cuemath.com/data/mean/) can be defined as the sum of all observations divided by the total number of observations. The formulas for the mean are given as follows:

Ungrouped data Mean: x̄ = Σxixi / n

Grouped data Mean: x̄ = ∑Mifi∑fi∑Mifi∑fi

Here, xixi is the ith observation, MiMi is the midpoint of the ith interval, fifi is the corresponding frequency and n is the sample size.

**Median:**The [median](https://www.cuemath.com/data/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]th term

Ungrouped data Median (n is even): [(n / 2)th term + ((n / 2) + 1)th term] / 2

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

l is the lower limit of the median class given by n / 2, c is the [cumulative frequency](https://www.cuemath.com/data/cumulative-frequency/), f is the frequency of the median class and h is the class height.

**Mode:**The [mode](https://www.cuemath.com/data/mode/) is the most frequently occurring observation in the data set. The formulas for the mode are given as follows:

Ungrouped data Mode: Most recurrent observation

Grouped data Mode: L + h (fm−f1)(fm−f1)+(fm−f2)(fm−f1)(fm−f1)+(fm−f2)

L is the lower limit of the modal class, h is the class height, fmm is the frequency of the modal class, f11 is the frequency of the class preceding the modal class and f22 is the frequency of the class succeeding the modal class.