

Basic Concepts

1. The facts or figures, which are numerical or otherwise, collected with a definite purpose are called *data*. Data is the plural form of the Latin word '*datum*'.
2. Statistics is the area of study dealing with the presentation, analysis and interpretation of data.
3. The data which are collected by an investigator personally to fulfill his objective is called *primary data*.
4. The data which are collected by someone else but used by the investigator for his own purpose is called *secondary data*.
5. Putting the data in the form of tables, in condensed form, is known as the *presentation of data*.
6. The number of times an observation occurs is called its *frequency*.
7. The tabular arrangement of data, showing the frequency of each observation, is called a *frequency distribution*.
8. There are two types of frequency distributions:
 - (a) **Exclusive form (or continuous form)**: A frequency distribution in which upper limit of each class is excluded and lower limit is included, is called an exclusive form.
Consider the classes 0 – 10, 10 – 20, etc. In class 10 – 20, we include 10 and exclude 20.
 - (b) **Inclusive form (or discontinuous form)**: A frequency distribution in which each upper limit as well as lower limit is included, is called an inclusive form. Thus, we have classes of the form 0 – 10, 11 – 20, 21 – 30, etc. In 0 – 10, both 0 and 10 are included.
9. Each group into which the raw data is condensed, is called a class interval. Each class is bounded by two figures, which are called *class limits*. The figure on the left side of a class is called its *lower limit* and that on the right is called its *upper limit*.

10. In the case of exclusive classes the upper and lower are respectively known as its *true upper limits* and *true lower limits*.

But in the case of inclusive classes, the true lower and upper limits are obtained by subtracting 0.5 from the lower limit and adding 0.5 to the upper limit.

Thus in the case of classes 5 – 10, 10 – 15, 15 – 20,

True lower limit of the class 10 – 15 = 10

True upper limit of the class 10 – 15 = 15

In the case of class 1 – 10, 11 – 20, 21 – 30,

True lower limit of the class 11 – 20 = $11 - 0.5 = 10.5$

True upper limit of the class 11 – 20 = $20 + 0.5 = 20.5$

True upper limits and true lower limits are also known as the boundaries of the classes.

11. The difference between the true upper limit and the true lower limit of a class is called the *size of the class*.

12. The value which lies midway between lower and upper limits of a class is known as *mid – value* or *class mark*.

i.e., Class mark =
$$\frac{\text{lower limit of class} + \text{upper limit of class}}{2}$$

13. The difference of the highest and the lowest values in the data is called the *range* of the data.

14. Bar diagram or bar graph

A set of bars (thick lines and narrow rectangles) representing variable and frequency constitute a bar graph. The following points should be kept in mind while drawing bar graphs.

- (i) The bars should be at equal distances from each other.
- (ii) All the bars in a diagram should be of the same width.
- (iii) Both axes should be adequately labelled.

15. Histogram

We use a histogram to represent grouped data by representing class boundaries along the horizontal axis and the corresponding frequencies along the vertical axis. Thus, rectangles are constructed with base as the class size and their heights representing the frequencies.

16. Frequency Polygon

A frequency polygon is obtained by joining the mid – points of the respective tops of the rectangles in a histogram. To complete the polygon

join the mid – points of two more classes (called imaginary classes) one at each end.

17. Median is the middle item in the arrayed data.

Median = $\left(\frac{n+1}{2}\right)$ th item, if n is odd.

Median = mean of $\left[\frac{n}{2} \text{th item} + \left(\frac{n}{2} + 1\right) \text{th item}\right]$, if n is even.

18. Mode is the most frequently occurring observation.

Empirical formula for mode is:

Mode = 3 Median – 2 Mean

cbseassistance.com