6. class width : The difference between the upper and lower boundaries of a class interval is called the width or size of the class.

7. Frequency density : The frequency density of a class is the frequency per unit width of the class.

$$\text{i.e., frequency density} = \frac{\text{class frequency}}{\text{width of the class interval}}$$

Frequency densities are used for comparing the concentration of frequencies in different classes, particularly when the class are of unequal width

Now, let us consider the construction of frequency distribution of a continuous variable and the relevant guidelines. Suppose we are given n values of a continuous variable. To prepare a frequency distribution with the given values, we proceed as follows:

We first pick up the smallest and greatest of the given values. Their difference gives the range of variation. The range is then divided into a suitable number of classes depending on the total ~~frequency~~ frequency.

In determining the classes, we have to bear in mind the following points:

1. The classes should be exhaustive so that no values escapes classification.

2. The classes should be mutually exclusive (i.e., non-overlapping) so that no values comes under more than one class.

3. The number of classes should not be very large. If the number of classes be large, the primary object of classification, namely summarisation, is defeated. It should also be mentioned that by considering a large number of classes one may introduce an irregular pattern in the frequencies which may be absent in actual population.

4. The number of classes, again, should not be very small. If there are too few classes, the true nature of the distribution may be obscured. Moreover, in such cases, different statistical measures would involve large error due to grouping (arising out of the assumption that all the frequencies in a class are concentrated at the mid-value of the class).

There is no hard and fast rule regarding number of classes. As a working rule, one can take 15 to 20 classes when the total frequency is well above 1000. 10 to 15 classes for total frequency around 1000. With total frequency far smaller than 1000, one may take fewer classes; however 7 or 8 classes will be sufficient in case total frequency is near 200.

5. Equal width should preferably be maintained for different ~~classes~~ classes. This will enable one to compare the class frequencies. Moreover, the computation of different statistical measures will be comparatively easy. However,

this condition of equal class-width is not to be rigidly followed. There are situations (as in the case of income distribution) where classes of varying width are preferred in order to make the classification more significant.

Keeping the above points in mind, we are to divide the range into a suitable number of classes, defined in terms of class limits. For convenience, sometimes we are required to take a range which is slightly bigger than the one obtained from the data. Next, the given values are taken one by one and recored in their respective classes with the help of tally marks. The procedure is continued until all the values are considered. In order to facilitate counting tally marks are kept in groups of five, as in the case of discrete variable. The table will look like:

class limits	Tally marks
$a-(a+c-d)$	𝍸 || ... ...
$(a+c)-(a+2c-d)$	𝍸 𝍸 ... ...
... ...	... ...
... ...	... ...
$(a+\overline{k-1}\,c)-(a+kc-d)$	||| ... ...

Here $a \leq$ the smallest given values;

c is the desired width of the classes;

k is the number of classes;

$a+kc-d \geq$ greatest given values;

$d = 1, 0.1$ or 0.01 etc. according as the values are given

in integers; upto one place after decimal or upto two places after decimal, etc.

After completing the table of tally marks we prepare the final frequency table where classes are given in terms of class boundaries and frequencies of different classes are noted against them.

For any class,

Lower boundary = lower limit - $d/2$

Upper boundary = upper limit + $d/2$

The final table will look like the following:

class boundaries	frequency
$(a-d/2) - (a+c-d/2)$	f_1
$(a+c-d/2) - (a+2c-d/2)$	f_2
... ...	...
... ...	...
$(a+\overline{k-1}c - d/2) - (a+kc-d/2)$	f_k
Total	n

A frequency distribution may also be represented in terms of relative frequencies (i.e., proportions) or cumulative frequencies. The cumulative frequencies are found by successively adding the class frequencies, where the addition begins from top (i.e., lowest class) or bottom (i.e the highest class) depending on less-than type or more-than type cumulative frequencies are obtained. In fact, the former kind of cumulative frequency of a class indicates the number of values less than