- NOTES WITH MIND MAPS -

MATHEMATICS

(STATISTICS)

9TH

Statistics

1. Facts or figures collected with a definite purpose are called **data**.

2. **Statistics** deals with collection, presentation, analysis and interpretation of numerical data.

3. Arranging data in an order to study their salient features is called **presentation of data**.

4. Data arranged in ascending or descending order is called arrayed data or an **array**.

5. When an investigator with a definite plan or design in mind collects data first handedly, it is called primary data.

6. Data when collected by someone else, say an agency or an investigator, comes to you, is known as the secondary data.

7. **Range** of the data is the difference between the maximum and the minimum values of the observations.

8. The small groups obtained on dividing all the observations are called classes or **class intervals** and the size is called the **class size** or class width.

   Class size = Upper limit – Lower limit

9. **Class mark** of a class is the mid-value of the two limits of that class.

10. The number of times an observation occurs in the data is called the **frequency** of the observation.

11. A **frequency distribution** in which the upper limit of one class differs from the lower limit of the succeeding class is called an Inclusive or discontinuous frequency distribution.

12. A frequency distribution in which the upper limit of one class coincides with the lower limit of the succeeding class is called an exclusive or **continuous frequency distribution**.

13. In case of continuous frequency distribution, the upper limit of a class is not to be included in that class while in discontinuous both the limits are included.

14. The **cumulative frequency** of a class-interval is the sum of frequencies of that class and the classes which precede (come before) it.

15. A data can be represented **graphically** through:

    (i) Bar graph (ii) Histogram (iii) Frequency polygon.

16. A **bar graph** is a diagram showing a system of connections or interrelations between two or more things by using bars.

17. In a bar graph, rectangular bars of uniform width are drawn with equal spacing between them on one axis, usually the x-axis. The value of the variable is shown on the other axis that is the y-axis.

18. A **histogram** is a graphical representation of a frequency distribution in the form of rectangles with class intervals as bases and heights proportional to the corresponding frequencies such that there is no gap between any two successive rectangles.

19. If classes are not of equal width, then the height of the rectangle is calculated by the ratio of the frequency of that class, to the width of that class

20. **Frequency polygons** are a graphical device for understanding the shapes of distributions.

21. If both a histogram and a frequency polygon are to be drawn on the same graph, then we should first draw the histogram and then join the mid-points of the tops of the adjacent rectangles in the histogram with line-segments to get the frequency polygon.

22. A measure of central tendency tries to estimate the central value which represents the entire data.

23. The three **measures of central tendency** for ungrouped data are mean, mode and median.

24. The **mean** value of a variable is defined as the sum of all the values of the variable divided by the number of values.

25. If $x_1, x_2, x_3 \ldots, x_n$ are n values of a variable X, then the arithmetic mean of these values is given by:

$$\text{Mean } (\bar{x}) = \frac{1}{n} \sum_{i=1}^{n} x_i$$

If a variate X takes values $x_1, x_2, x_3 \ldots, x_n$ with corresponding frequencies $f_1, f_2, f_3, \ldots f_n$ respectively, then arithmetic mean of these values is given by

$$\text{Mean } (\bar{x}) = \frac{\sum f_i x_i}{\sum f_i}$$

26. **Median** is the value of middle most observation(s).

27. The median is calculated only after arranging the data in ascending order or descending order.

If n is odd, then median $= \left(\dfrac{n+1}{2}\right)^{th}$ observation

If n is even, then median $= \left\{\dfrac{\left(\dfrac{n}{2}\right)^{th} \text{observation} + \left(\dfrac{n}{2}+1\right)^{th} \text{observation}}{2}\right\}$

28. **Mode** of a statistical data is the value of that variate which has the maximum frequency.

29. The variate corresponding to the highest frequency is to be taken as the mode and not the frequency.

30. The disadvantage of arithmetic mean is that it is affected by extreme values.

31. The disadvantage of mode is that it is not uniquely defined in many cases.

Introduction to Statistics

A study dealing with the collection, presentation and interpretation and analysis of data is called as statistics.

Data

- Facts /figures numerical or otherwise collected for a definite purpose is called as data.

- data collected first-hand data:- Primary

- Secondary data: Data collected from a source that already had data stored

Frequency

The number of times a particular instance occurs is called frequency in statistics.

Ungrouped data

Ungrouped data is data in its original or raw form. The observations are not classified in groups.

Grouped data

In grouped data, observations are organized in groups.

Class Interval

- The size of the class into which a particular data is divided.

- E.g divisions on a histogram or bar graph.

- Class width = upper class limit – lower class limit

Regular and Irregular class interval

Regular class interval: When the class intervals are equal or of the same sizes.

E.g 0-10, 10-20, 20-30….. 90-100

Irregular class interval: When the class intervals are of varying sizes.

E.g 0-35, 35-45, 45-55, 55- 80, 80-90, 90-95, 95-100

Frequency table

A frequency table or distribution shows the occurrence of a particular variable in a tabular form.

Sorting

- Raw data needs to be sorted in order to carry out operations.

- Sorting ⇒ ascending order or descending order

Ungrouped frequency table

When the frequency of each class interval is not arranged or organised in any manner.

Grouped frequency table

The frequencies of the corresponding class intervals are organised or arranged in a particular manner, either ascending or descending.
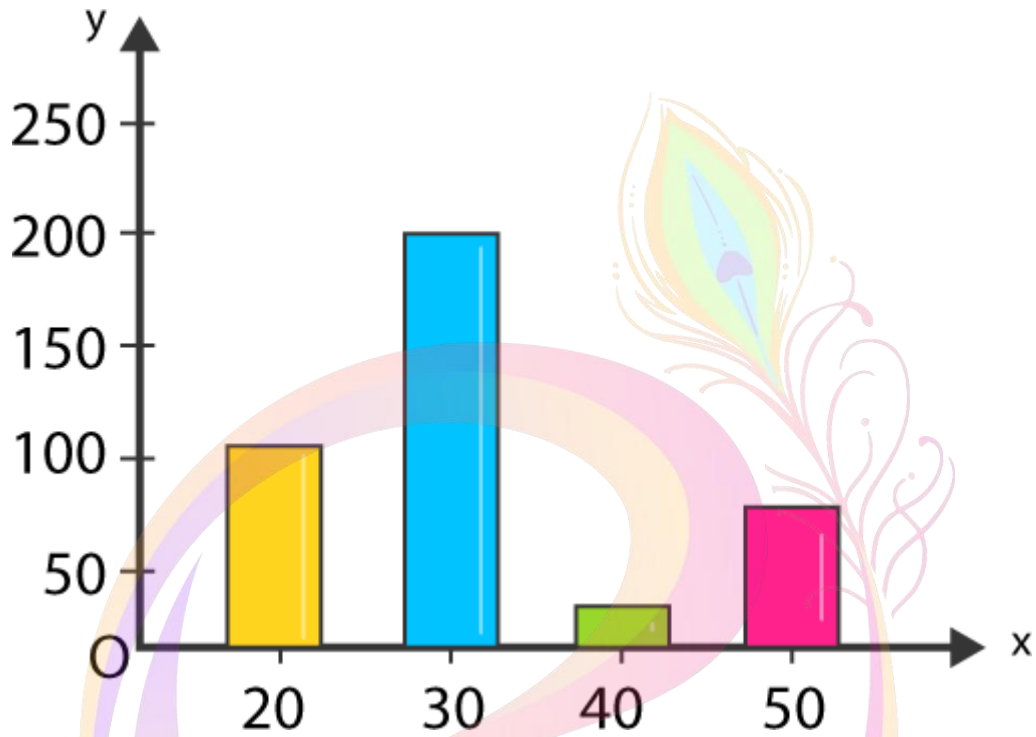
Graphical Representation of Data

Bar graphs

Graphical representation of data using bars of equal width and equal spacing between them (on one axis). The height

Savings (in percentage)	Number of Employees (Frequency)
20	105
30	199
40	29

50	73
Total	400

The data can be represented as:


Variable being a number

- A variable can be a number such as 'no. of students' or 'no. of months'.

- Can be represented by bar graphs or histograms depending on the type of data.

Discrete → bar graphs

Continuous → Histograms

Histograms

- Like bar graphs, but for continuous class intervals.

- Area of each rectangle is a Frequency of a variable and the width is equal to the class interval.
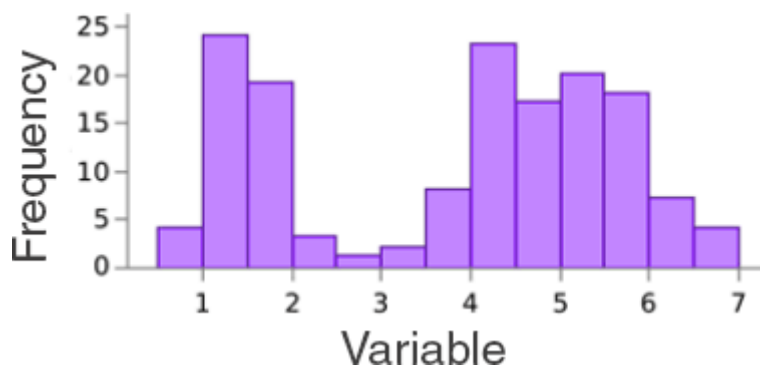
Frequency polygon

If the midpoints of each rectangle in a histogram are joined by line segments, the figure formed will be a frequency polygon.

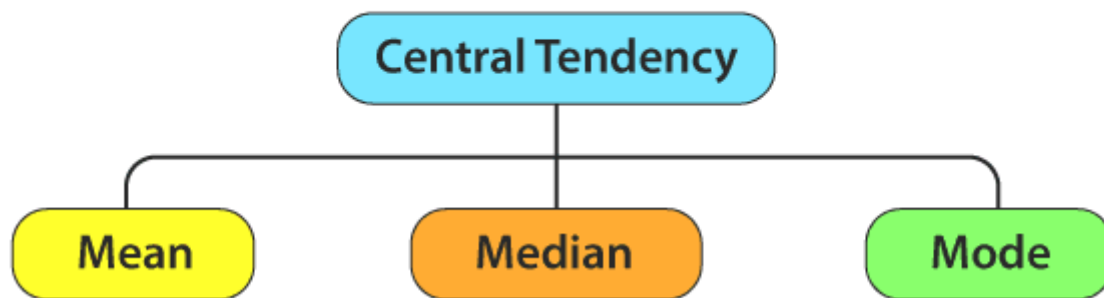Can be drawn without histogram. Need midpoints of class intervals
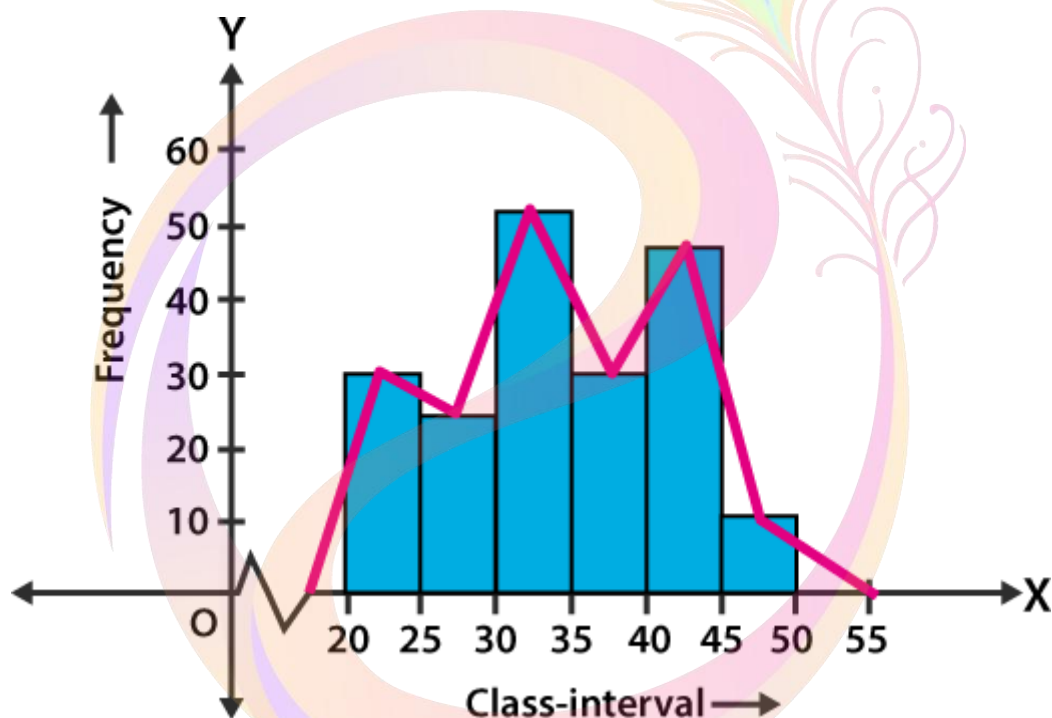
Midpoint of class interval

The midpoint of the class interval is called a class mark

Class mark = (Upper limit + Lower limit)/2

CENTRAL TENDENCY

Central Tendency
- Mean
- Median
- Mode

Equality of areas

Addition of two class intervals with zero frequency preceding the lowest class and succeeding the highest-class intervals enables to equate the area of the frequency polygon to that of the histogram (Using congruent triangles.)

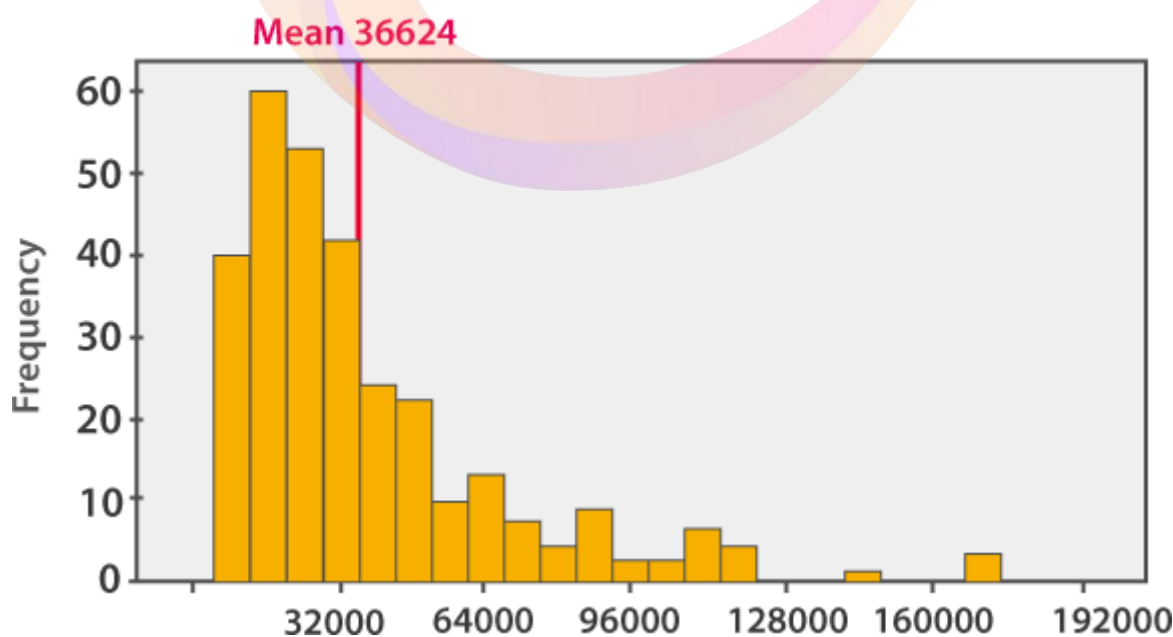**Measures of Central Tendency**

Average

The average of a number of observations is the sum of the values of all the observations divided by the total number of observations.
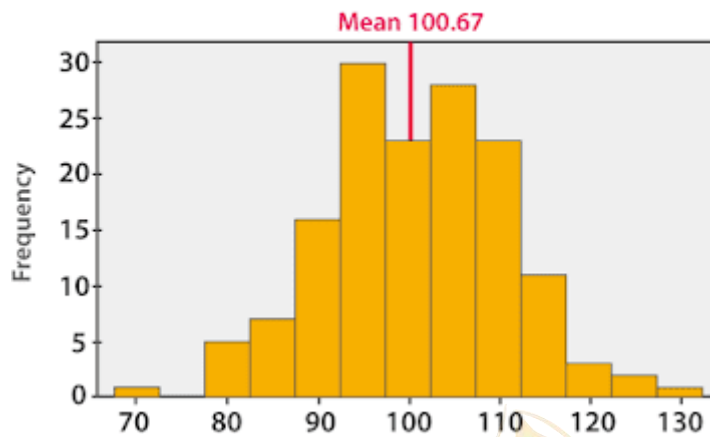
Mean

Mean for ungrouped frequency distribution, $\bar{x} = \frac{\sum x_i f_i}{f_i}$

where f_i is the frequency of ith observation x_i

Histogram of skewed continuous

Mean 36624

Histogram of symmetric continuous

Mode

- The most frequently occurring observation is called the mode.

- The class interval with the highest frequency is the modal class

Median

- Value of the middlemost observation.

- If n(number of observations) is odd, Median $=[(n+1)/2]^{th}$ observation.

- If n is even, the Median is the mean or average of $(n/2)^{th}$ and $[(n+1)/2]^{th}$ observation.

Median odd
23
21
18
16
15
13
12
10
9
7
6
5
2

Median even
40
38
35
33
32
30
29
27
26
24
23
22
19
17

28