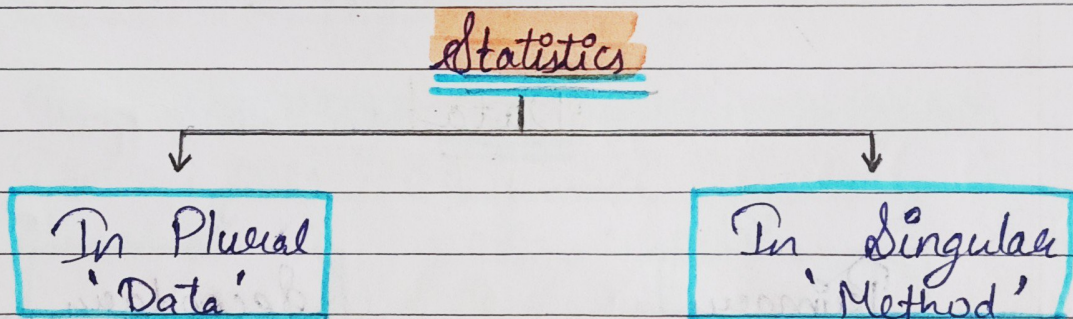


Chapter-13 Unit -1

STATISTICAL DESCRIPTION OF DATA.

* PART - I Introduction

Q. What is Statistics ?



* Uses :-

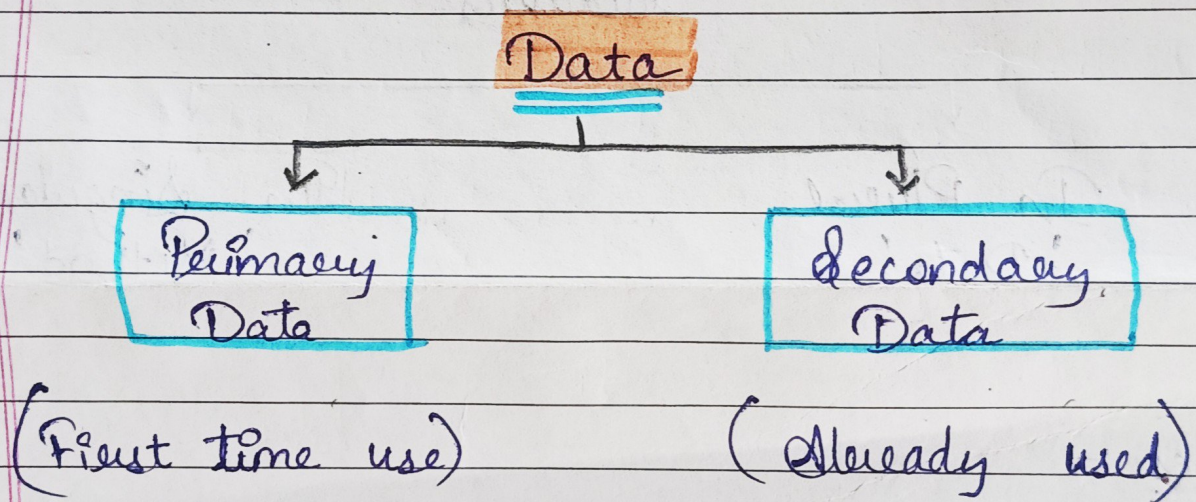
- In Commerce
- In Economics
- In Business Management
- In Industries

* Limitations :-

- It can be used on Group not on individuals.

- It is used on Qualitative and Quantitative both types of Data but more effective on Quantitative Data.

* PART - II COLLECTION OF DATA



* Primary Data Collection Methods:-

① Interview Method:-

(i) Direct Interview (Personal Interview):-

eg:- Natural calamities & Epidemic

(ii) Indirect Interview:-

eg:- Rail Accident

(iii) Telephone Interview :-

- Coverage area is higher than other two.
- Reliability is lower than the other two.
- Quickest among all methods.

② Mailed Questionnaire Method :-

- Its coverage area is highest.
- Its reliability is lowest.
- Amount of non-response is highest.

③ Observation Method :-

- Most accurate Method (Best Method)
(cross check)

④ Questionnaire by Enumerator :-

* Sources of Secondary Data :-

- (1) Government Data
 - (2) International Data
 - (3) Private Data
 - (4) Unpublished Data
- Main

* PART - III

PRESENTATION

OF

DATA

① Textual Presentation :-

- Paragraph is used.
- Even the numerical data is presented in descriptive way.

② Tabular Presentation :-

- Most accurate (Best Method)
- Details can be shown using this method.
- Table has 4 parts.

* Title

Unit			
Stub			

* Foot note

Diagram labels and arrows:

- ① Box head (points to the top row of the table)
- ② Caption (points to the title 'Unit')
- ③ Stub (points to the first column of the table)
- ④ Body (points to the main data area of the table)

(i) Box head :-

- Entire upper part of table.
- Unit is given in this part.

(ii) Caption :-

- Upper part of table.
- Column heading are given in this part.

(iii) Stub :-

- Left part of table.
- Row heading are given in this part.

(iv) Body :-

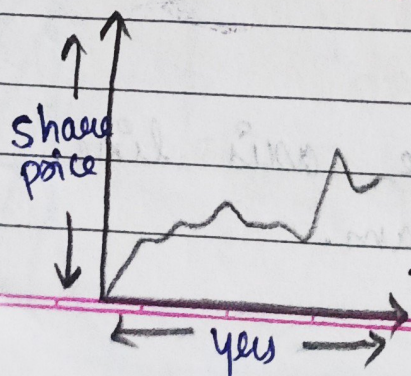
- Main part of table.
- All numerical comparison is done in this part.

* Foot note :- Source of Data is given in this part.

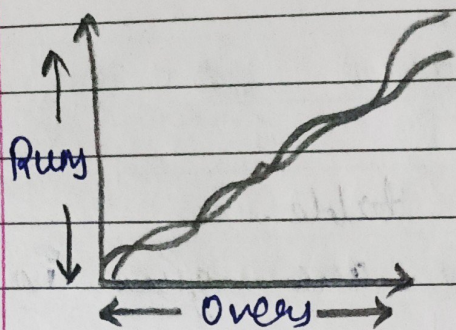
③ Diagram Presentation :-

(Most attractive)

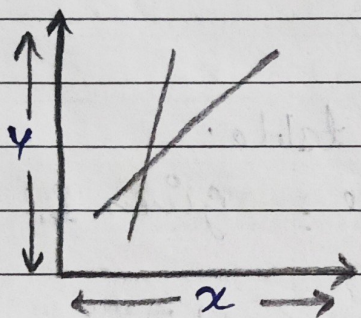
(i) Line Diagram :-



- Mainly used for time series Data.



- For comparison of two or more series it is used.

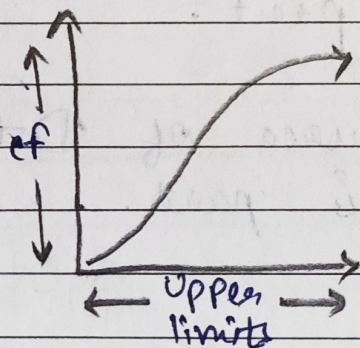


- To know the relation-ship between variables.

eg:-

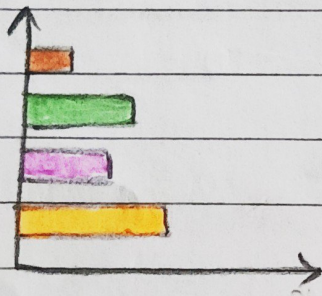
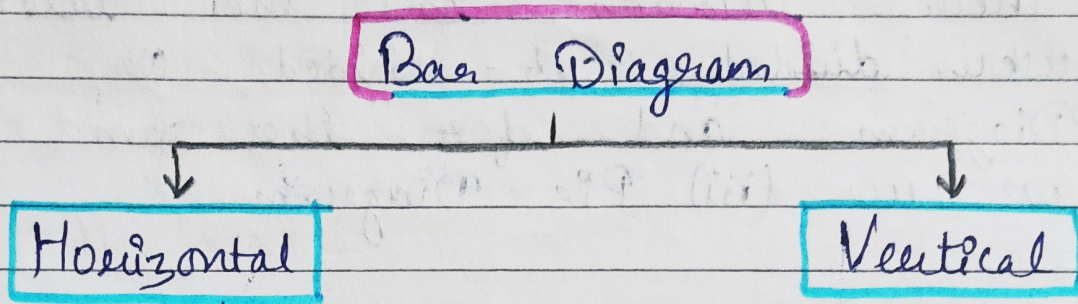
- Graph

- Ogive (cumulative frequency curve)

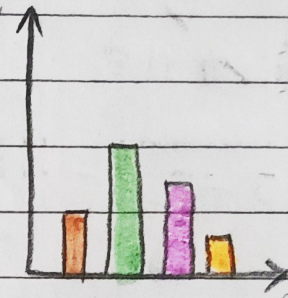


- For single series $\rightarrow$ Simple line Diagram
- For multiple series $\rightarrow$ Multiple line Diagram (with same units)
- For multiple series $\rightarrow$ Multiple axis line Diagram (with different units)

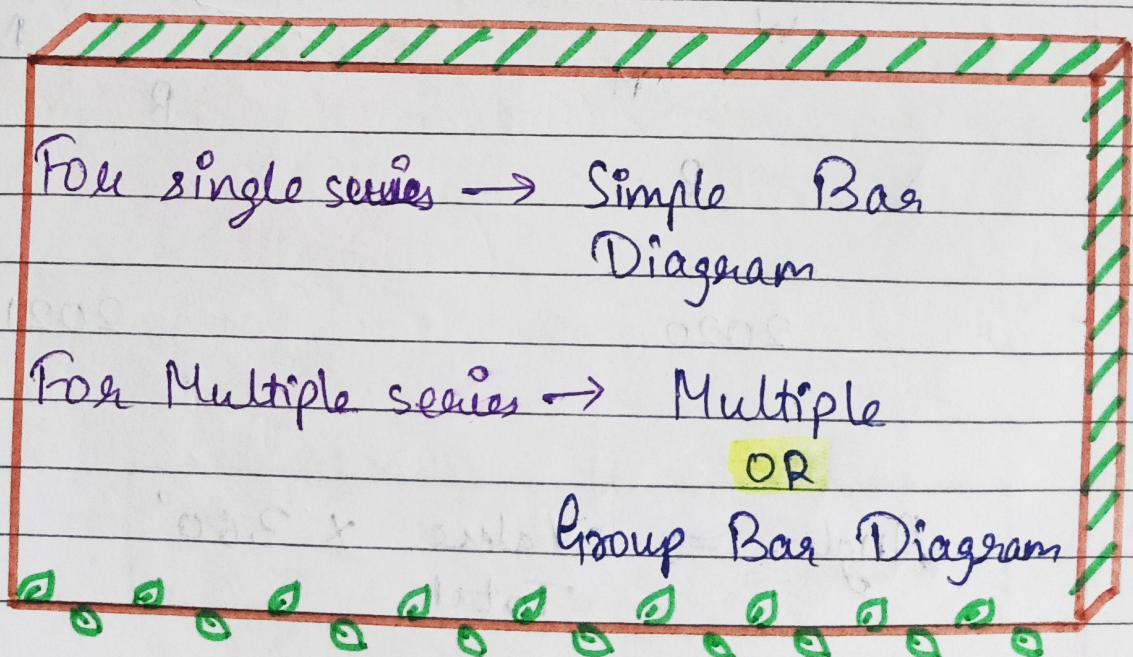
(ii) Bar Diagram :-



- Qualitative Data
(eg:- rating of any app)
- Data vary over space.
(Geographical)

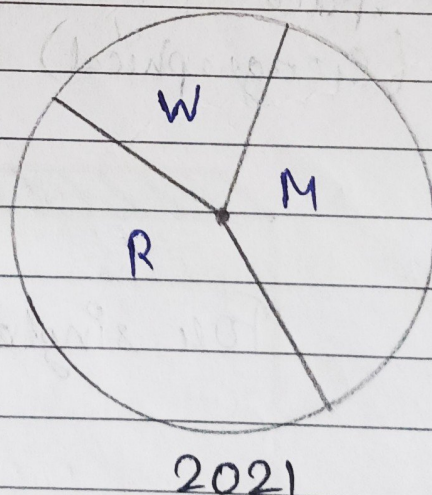
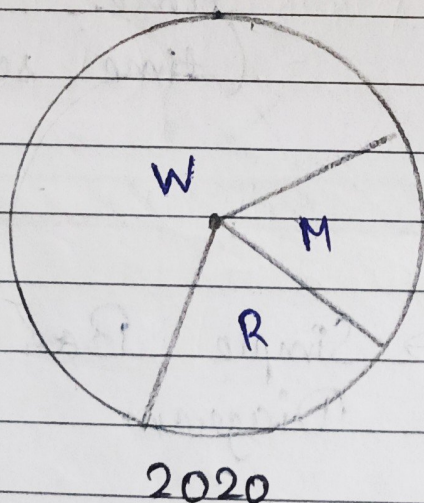
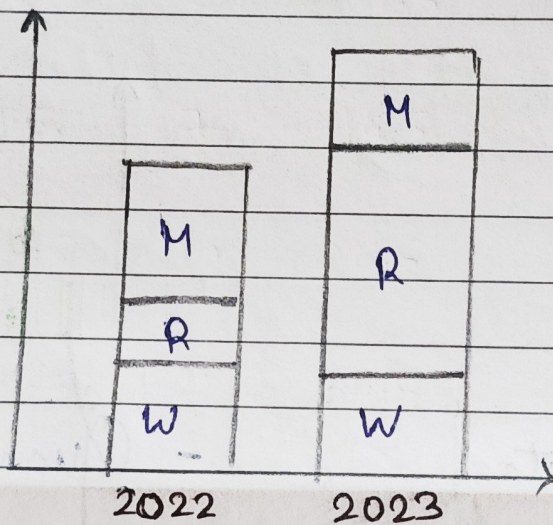


- Quantitative Data
- Data vary over time.
(time series)



Value → change

* If we want to compare different components and also want to know their relation to the total, we use divided, sub-divided or step Bar Diagram and for the same purpose we use (iii) Pie Diagram



$$\text{Angle} = \frac{\text{Value} \times 360^\circ}{\text{Total}}$$

* PART - VI FREQUENCY DISTRIBUTION

Frequency Distribution

Discrete

Continuous

Simple
Frequency
Distribution

Inclusive
Frequency
Distribution

Exclusive
Frequency
Distribution

x	f
1	5
2	7
3	8
4	3

Class	f
0-9	3
10-19	7
20-29	10
30-39	5

Class	f
0-10	5
10-20	10
20-30	15
30-40	5

Both limits
are included

Upper limit is
excluded and
lower limit is
included.

Inclusive → Exclusive
Limits → Boundaries

$$UCB = UCL + \frac{D}{2}$$

$$LCB = LCL - \frac{D}{2}$$

C = Control B = Boundaries D = Difference D = UL - LL

Inclusive $\rightarrow$ Exclusive

Class Boundaries	F	Mid-value (class-Mark) $= \frac{UCB + LCB}{2}$	Width of class (length of class) (size of class) $= UCB - LCB$	Relative Frequency $= \frac{F}{\sum F}$
-0.5 - 9.5 LB UB	3	4.5	$9.5 - (-0.5) = 10$	$\frac{3}{25}$
9.5 - 19.5	7	14.5	10	$\frac{7}{25}$
19.5 - 29.5	10	24.5	10	$\frac{10}{25}$
29.5 - 39.5	5	34.5	10	$\frac{5}{25}$
	$\sum F = 25$			

$\rightarrow$ Relative Frequency always lies between 0 & 1.

$\rightarrow$ For Exclusive distribution

Limits = Boundaries

$\rightarrow$ For Inclusive distribution

Limits $\neq$ Boundaries.

→ Upper Boundary of any class coincides with lower boundary of next class.

→ Upper limit of any class is different from lower limit of next class.

Class limits	Class Boundaries
0-19	- 0.5 - 19.5
20-39	19.5 - 39.5
40-59	39.5 - 59.5
60-79	59.5 - 79.5

→ Excepting first & last all the other class boundaries lie at the mid way between upper limit of class & lower limit of the next class.

→ For the construction of group or Ungroup distribution we use class limits.

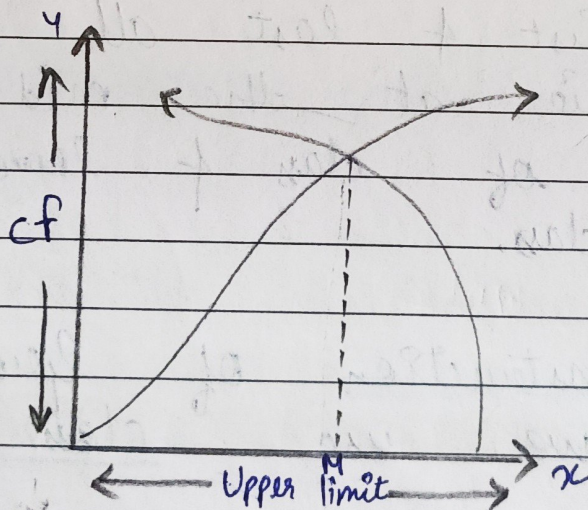
→ For all statistical calculations (eg:- Mode, Calculation / Median calculation) and diagrams we use class boundaries.

coincided = same

Exclusive Distribution

Cumulative Frequency Distribution (statistical table)

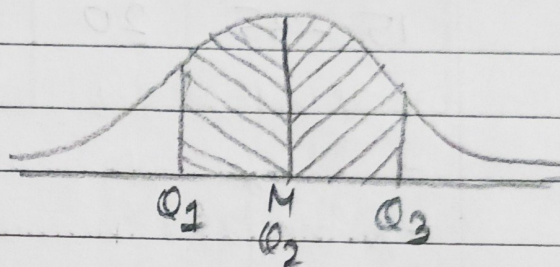
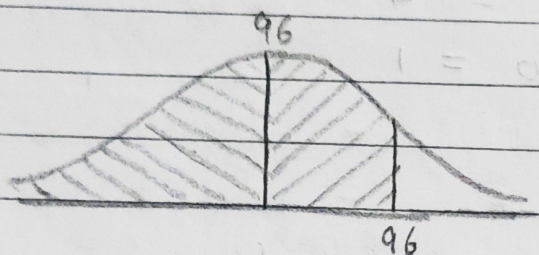
less than	cf	More than	cf
0	0	0	35
10	5	10	30
20	15	20	20
30	30	30	5
40	35	40	0



- Median can be find out by ogive (by intersecting point or if it is single curve then (50.1))
- We can find every positional value through ogive.

→ Cumulative frequency curve.

Median, Quartile, Deciles, Percentiles.



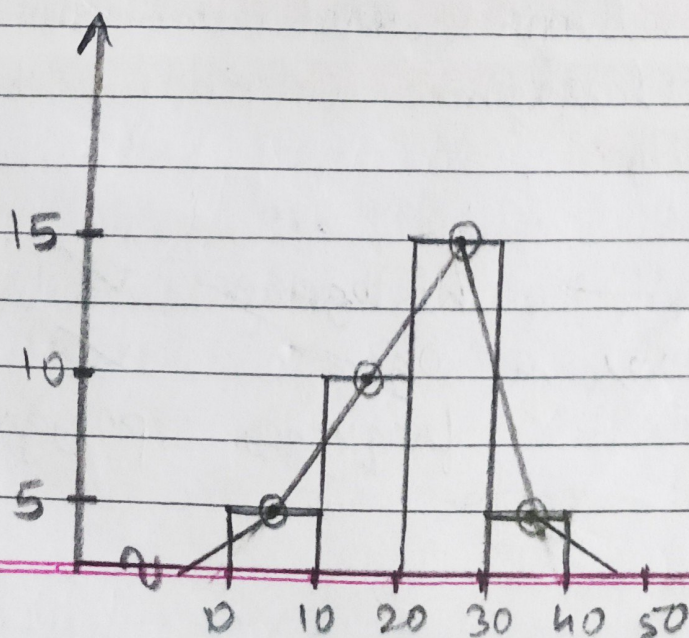
- Exclusive Distribution:-

- ① Histogram
 - ② Frequency Polygon
 - ③ Frequency curve
- (Group or Exclusive)

} smooth limiting from each other

When class length are same then, height of histogram depends upon Frequency.

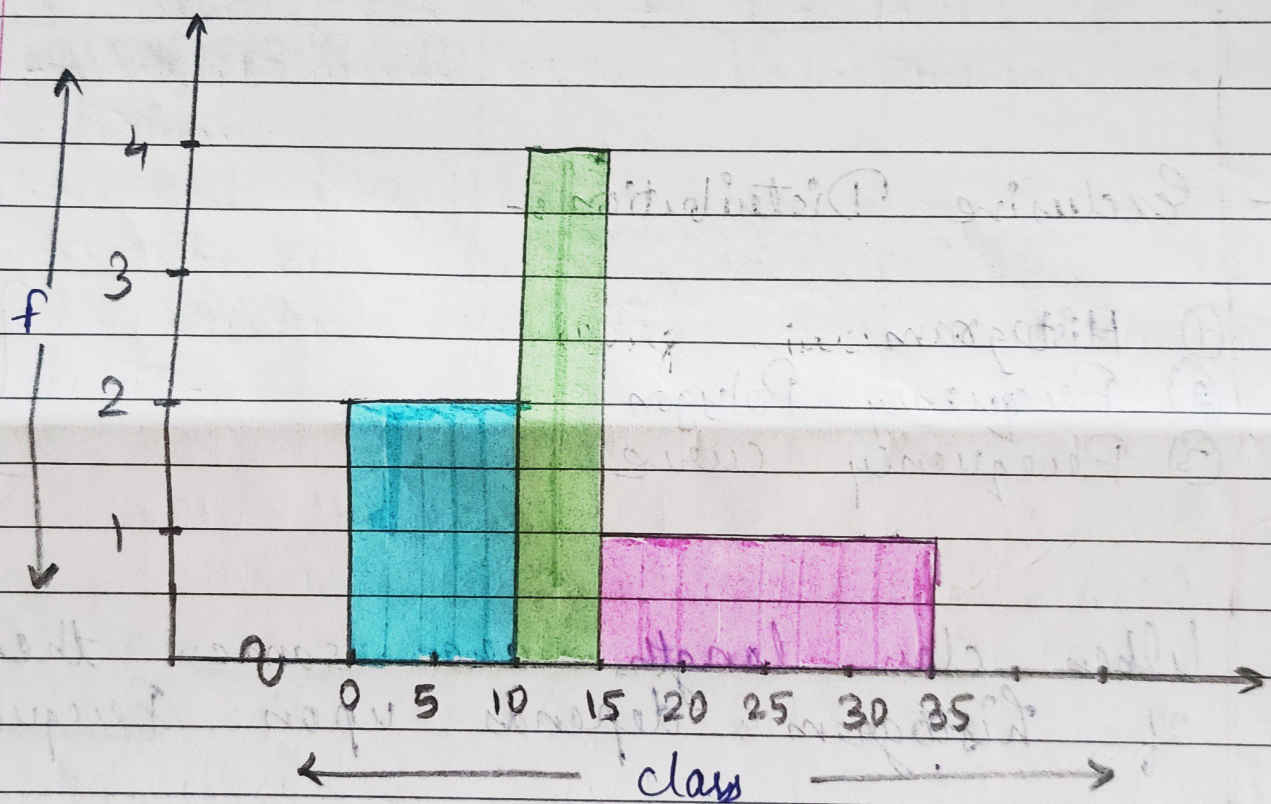
Width is dependent upon Class length.



free hand → curve
line → polygon

→ bcoz class of length are unequal

Class	f	Frequency Density = $\frac{\text{Frequency of class}}{\text{Width of class}}$
0-10	20	$\frac{20}{10} = 2$
10-15	20	$\frac{20}{5} = 4$
15-35	20	$\frac{20}{20} = 1$



- when class of length are unequal then, height of histogram depends upon frequency density.
- Unequal width → histogram ✓
ogive ✓
frequency polygon X

→

Height of histogram depends on

Class length are same

- frequency

Class length are diff.

- frequency density

POINTS

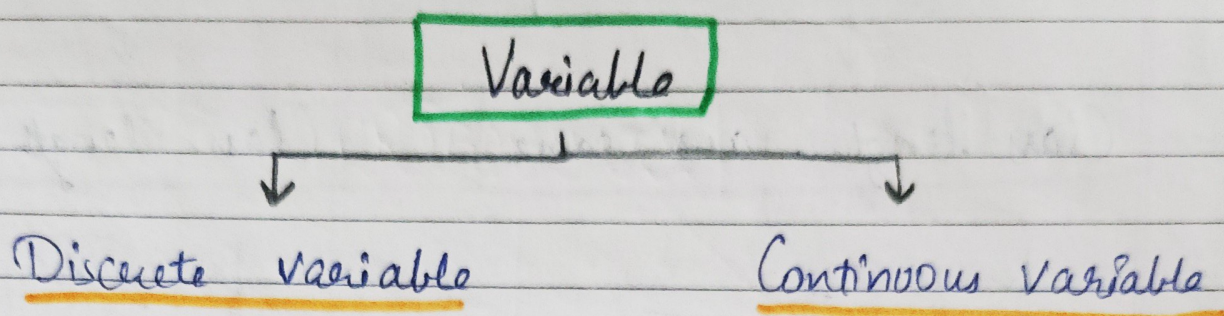
TO

REMEMBER

- Histogram is an Area Diagram.
(exclusive freqⁿ)
- For group frequency distribution mainly histogram is used.
- Histogram, frequency polygon & frequency curve are limiting form of each other and having same area.
- Frequency polygon and frequency curve gives approximate idea about each other.
- Attribute means Quality.
 - Nationality of a student.
 - Habit of Person.
 - Colour of a flower.

limiting → swaroop.

- Variable $\rightarrow$ Quantitative



- To check consistency & accuracy of the data following two methods are used:-

- (i) scrutiny (cross check)
- (ii) By taking number of related series from the data.

- Logarithm is used in ratio chart.

- Hidden trend in the data can be noticed by Diagrammatic Presentation (Trend line).

- Median $\rightarrow$ Ogive

- Mode $\rightarrow$ Histogram

Over-lapping = Continuous

non-overlapping = discrete

* Classification is of 4 kinds.

- (1) Qualitative
- (2) Quantitative
- (3) Geographical
- (4) Time series.

x	f
1	5
2	7
3	10
4	3

