Chapter = 02

Presentation of Data.

* Frequency distribution :- A mass of statiscal data in its original form → Raw Data / ungrouped data

* F·D is a classification according to the no. of processing the same values of the variables. It is simply a table in which data are grped into classes & the no. of cases which fall in each class is recorded. The no. → Frequencies.
They are 2 type →

1) Discrete Freq D :- If we have a large no. of items in the data, it is better to prepare a freq array. & contents the data further. Freq array is prepared by listing once & consiquently all the values occuring in the series & not in the no. of times each such value occures.

eg:- 1, 4, 3, 2, 1, 2, 0, 2, 1, 2, 3, 2
1, 0, 2, 3, 0, 3, 2, 1, 2, 2, 1, 4, 2

Data	Freq.
0	3
1	6
2	10
3	4
4	2

2) Continuous Freq. D :- Here the data are classified according to class intervals.

- Following are the rules for forming a freq table

1) Note the difference b/w largest & smallest value in the given set of observations.
2) Determine the no. cls into which difference can be divided.
3) classes should be exclusive.
4) 3 ~~the~~ columns → classes, Tally mrk & freq.
5) write class in 1st column.
6) put Tally mrks
7) write sum of tally mrks of each cls.
8) Sum of freq of all clses = No. of observations.

→ Concepts of freq table :-

1) Class limits = observations which constitute a cls → c.Limits.
   left hand side observation → lower limit
   right " " → upper limit

2) working cls = clses of form 0-9, 10-19, 20-29...
   → w. cls. / nominal cls.

3) Actual cls = If we r leaving either the upper limit / lower limit frm each cls → exclusive method of classification.

- we can convert w. cls to Actual cls →

a) note the difference one ulimit & the next l.limit.
b) ÷ the difference by 2.
c) Subtract that value from L.limit & + same to U.limits.

4) Class boundaries = cls limits of actual cls → c.boundaries / actual cls. limits
5) class mark = Avg of U.limit & l.limit of that cls → c.mrk / mid value.
6) Class interval = Difference b/w u.limit & l.limit of actual cls.
7) Tally Sheet & entry form = In freq table, freq are noted by a storke.

→ Cumulative Frequency Distribution :-

How many observations are lying below / above a particular value / in b/s 2 specified values → C.Freq.

we get less c.f (<cf) & greater c.f (>cf)

<cf give no of observations falling below the upper limit of a cls & >cf → no of observations lying above l.limit of cls.

<cf → + freq of all previous clses.

cls	F	UL	>CF		LL	<CF	
0-10	2	10	2	2	0	3+7+10+8+5+2	35
10-20	5	20	2+5	7	10	7+10+8+5+3	33
20-30	8	30	2+5+8	15	20	10+8+7+3	28
30-40	10	40	2+5+8+10	25	30	10+7+3	20
40-50	7	50	2+5+8+10+7	32	40	7+3	10
50-60	3	60	2+5+8+10+7+3	35	50	3	3

→ Graphical presentation of data :-

2 types of graphs →
1) Freq distribution (2) Time series.

1) Graph of freq distributions :-

a) Histogram :- It has adjoint bars | □ erected on x axis. Depending on width of cls interval - equal or unequal.

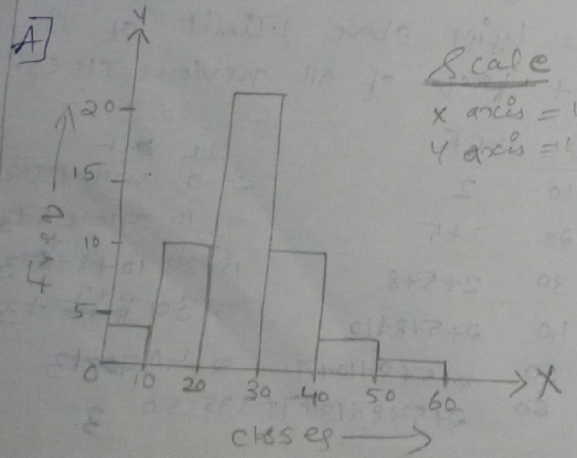
(i) Equal cls-intervals :- For this, mark the actual cls on x axis & on y axis take the freq. Now over each cls erect □ whose lengths are proportional to corresponding freq.

It enables the location of mode for equal cls intervals.

eg= Construct a histogram —

Cls →	0-10	10-20	20-30	30-40	40-50	50-60
freq →	4	10	21	9	4	2.

A



Scale
x axis = 1 cm = 10 units
y axis = 1 cm = 5 units

(ii) Unequal cls-intervals :- cls interval determine the base of □. & the height of bars are determined by cls densities.

* cls density / freq density is the ratio of cls freq to its cls interval → f/c

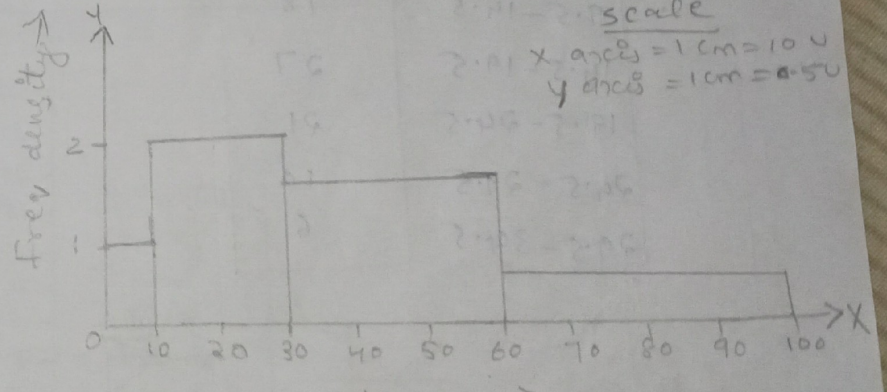
eg= construct an histogram.

Cls -	0-10	10-30	30-60	60-100
freq -	10	40	45	20

A

classes	freq	f/c.
0-10	10	1 (10/10)
10-30	40	2 (40/20)
30-60	45	1.5
60-100	20	0.5

c = width of cls



scale
x axis = 1 cm = 10 u
y axis = 1 cm = 0.5 u

(b) Freq polygon :- It can be constructed by drawing the histogram. If we r using histogram, join the mid points of the top of □ of histogram using straight line. Freq polygon can be drawn by

joining consecutive points of x & y axis. 2 / more freq polygon can be drawn on the same graph, bt 2 histograms cannot.

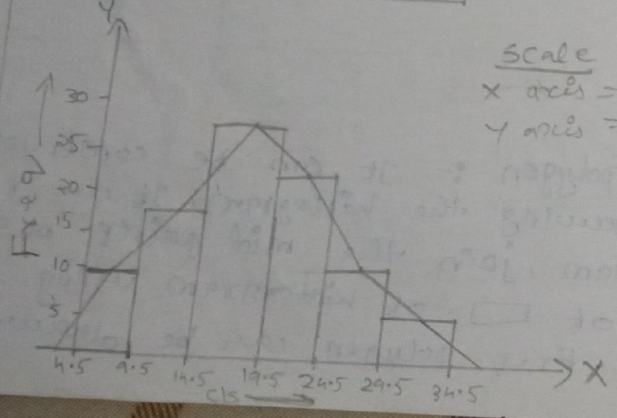
eg:-

cls	Freq
5–9	8
10–14	18
15–19	27
20–24	21
25–29	10
30–34	6

not equal L→−0.5 U→+0.5

A]

cls	Freq
4.5–9.5	8
9.5–14.5	18
14.5–19.5	27
19.5–24.5	21
24.5–29.5	10
29.5–34.5	6



Scale
x axis = 1cm = 5 units
y axis = 1cm = 5 units

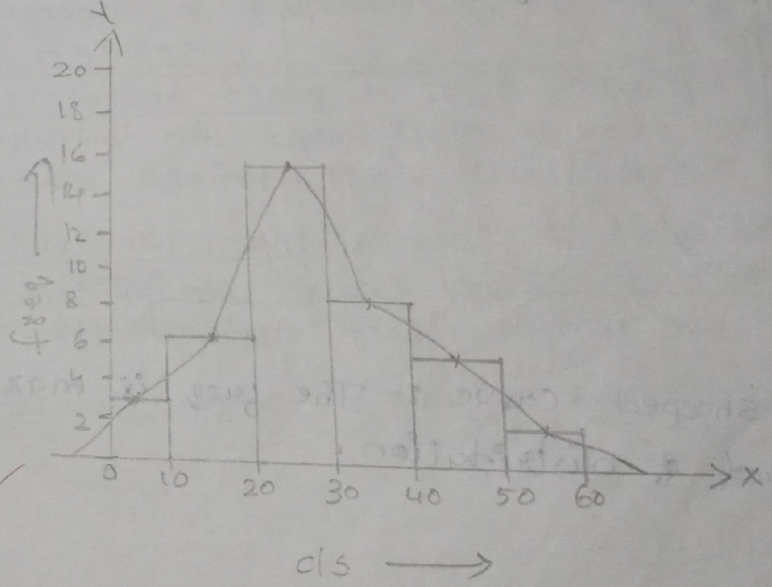
1) ~~construct a freq polygon~~

→ Frequency curve :- The method of drawing freq curve is similar to the construction of freq polygon.

1st construct the histogram. Then join the mid points of the top of ▭'s of the histogram using a smooth curve by free hand method.

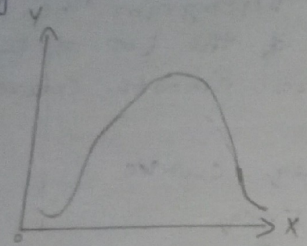
1) construct a freq curve

cls	Freq
0–10	3
10–20	6
20–30	15
30–40	9
~~40~~–50	5
50–60	2



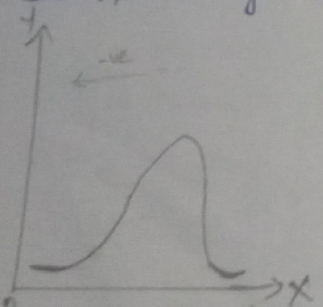
→ Different shapes of freq. curve :-

a) Symmetrical Distribution :- freq vises equally on either side of the central value hving the max freq.

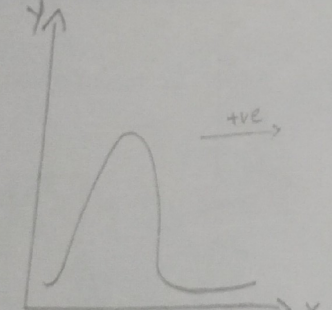


Symmetric

b) Moderately skewed / Slightly Assymtrical Distri :-
The rate of use of frequencies on 1 side of the max value is huge than the corresponding rate on the other side.

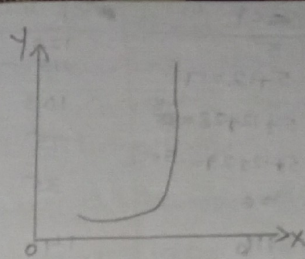


-ve skewed

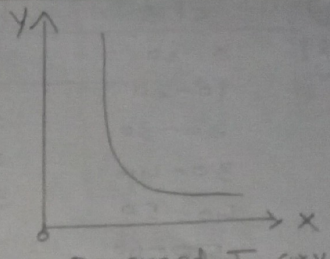


+ve skewed

c) J. shaped curve :- The freq is max at 1 end of distribution.

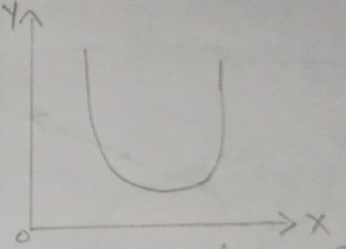


J. shaped curve



Reversed J curve.

d) U. shaped curve :- The freq is max at both ends. & min at centere.



U. shaped curve.

→ Graphs of cumilative freq. Distribution / ogives :-

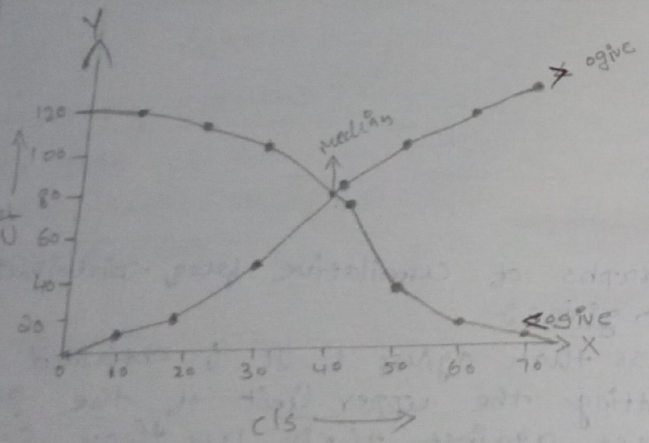
a) Less than ogives :- It is obtained by plotting the upper limit of the actual clses against their less than c.Freq (<cf)

b) More than ogives :- It is obtained by plotting the lower limit of the actual clses against their >CF.

1) Draw < ogive & > ogive for the data.

A)

cls	freq	CF (L.S)	CF (G.S)
0-10	5	5	120
10-20	12	5+12=17	115 (wrong)
20-30	28	5+12+28=45	103
30-40	40	5+12+28+40=85	75
40-50	21	106	35
50-60	10	116	14
60-70	4	120	4



Note :-
ogives for more than & less than type distribution intersect at median

→ Graph of Time Series :- It is a set of values of a variable collected & recorded in chronological order (yr base) of time intervals. (i.e) the data depending on time. It is also called historical series.

plot the time values against the corresponding values & join these points by a smooth line → Time series graph or historigram.

1) Draw a time series graph.

yr :	1991	1992	1993	1994	1995	1996	1997
freq :	80	90	92	83	94	99	92

A)