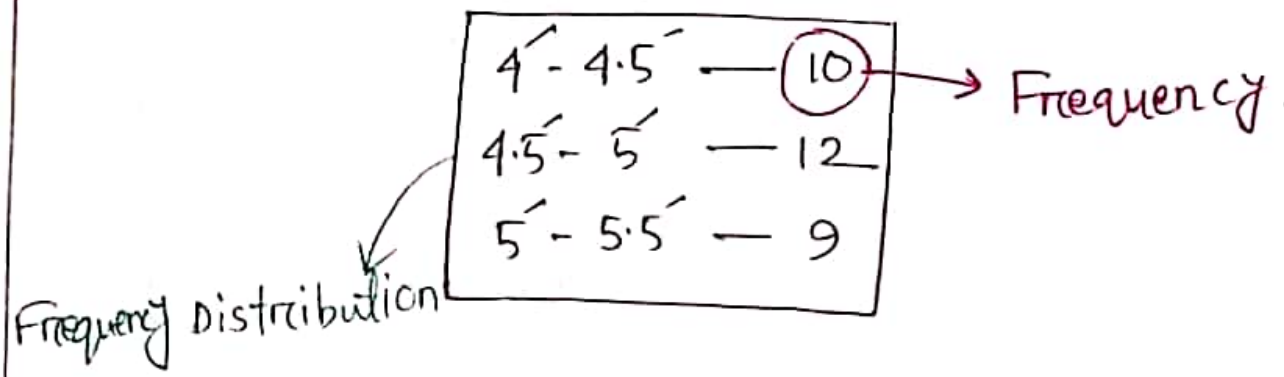
Class No-15 Date → 08.05.2019.

Statistics

Definition:

The science of statistics is essentially a branch of Applied Mathematics dealing with the collection, presentation and analysis of numerical data. Statistics means quantitative data, which are affected to a market extent by multiplicity of causes which is designed to summarize or describe important features of numerical Data by mainly by Tables or charts.

Frequency Distribution:



A set of classes together with the frequencies of occurance of values in each class in a given set of Data, represented in a tabular form, is referred to as a frequency distribution.

Construct of Tables:

Health & Smoking states of 50 workers.

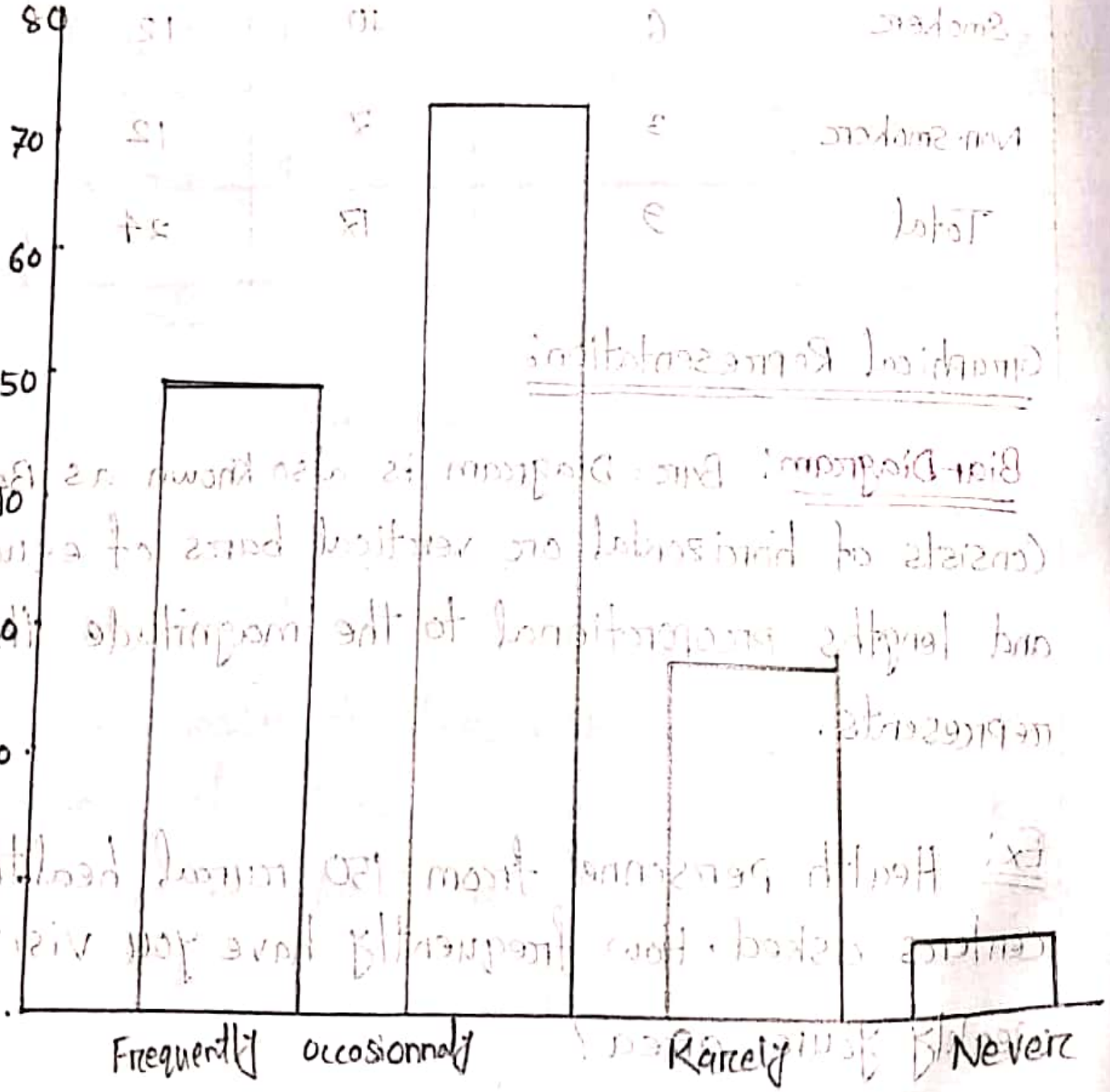
	Health status			
Smoking status	Good	Average	Poor	Total
Smokers	6	10	12	28
Non-Smokers	3	7	12	22
Total	9	17	24	50

Graphical Representation:

Bar Diagram: Bar-Diagram is also known as Bar chart. Consists of horizontal or vertical bars of equal widths and lengths proportional to the magnitude they represents.

Ex: Health personnel from 150 rural health centers asked. How frequently have you visited weakly your area?

Response	Frequency
Frequently	49
Occasionally	71
Rarely	24
Never	6
Total	150



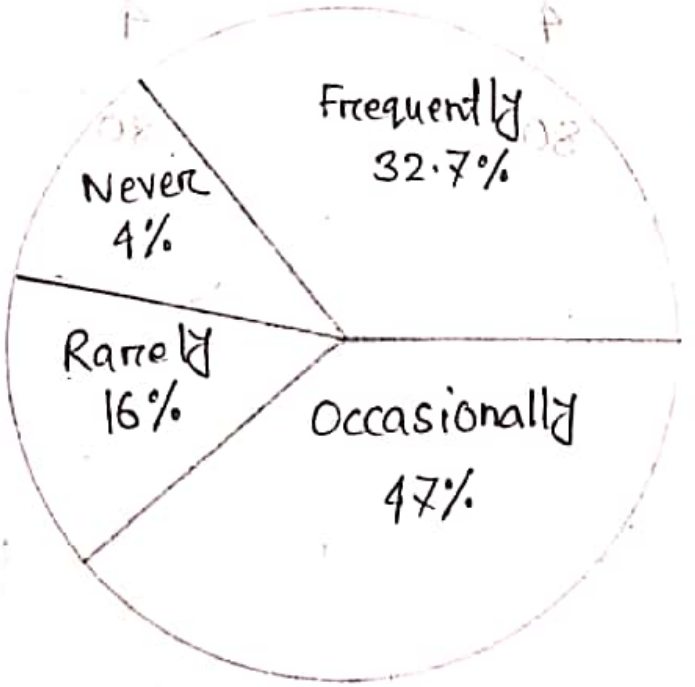
② **Pie chart**: Pie chart also known as Pie Diagram, is an effective way of representing percentage parts when the whole quantity is taken as 100.

$$\text{Frequently} = \frac{49}{150} \times 100\% = 32.7\%$$

$$\text{Occasionally} = \frac{71}{150} \times 100\% = 47.3\%$$

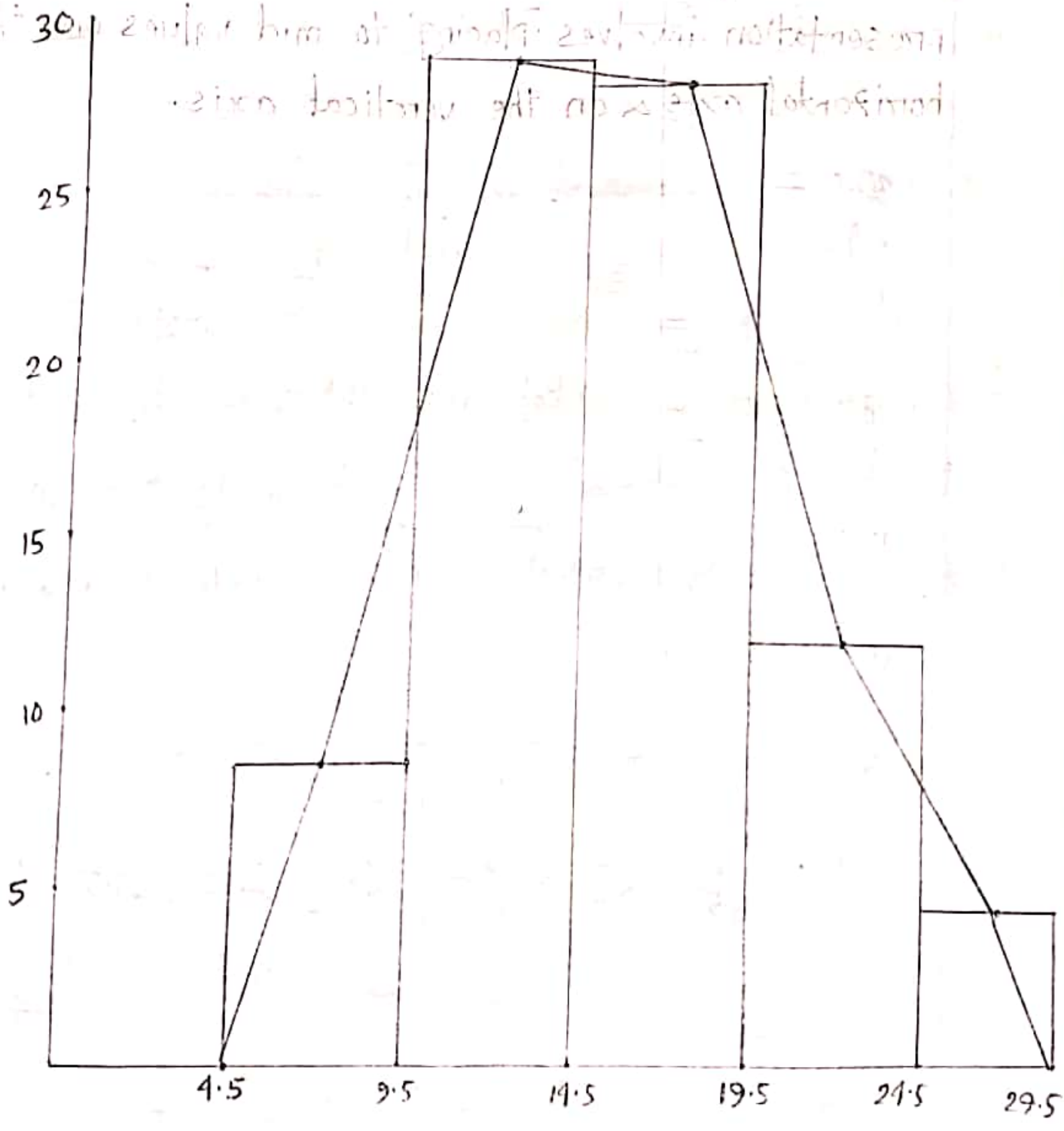
$$\text{Rarely} = \frac{24}{150} \times 100\% = 16\%$$

$$\text{Never} = \frac{6}{150} \times 100\% = 4\%$$

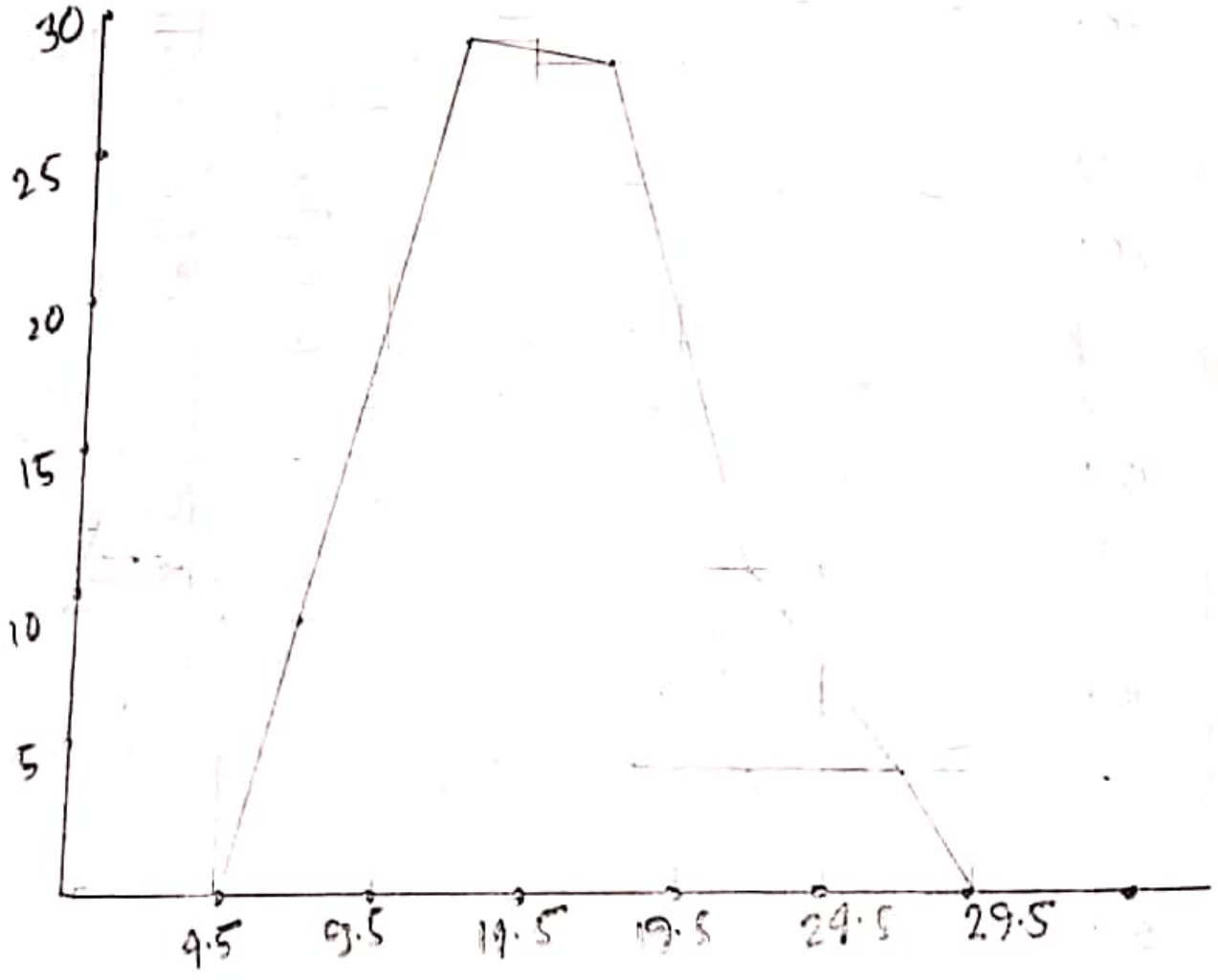


3. Histogram: The most common form of a graphical presentation of a frequency distribution in Histogram.

Expenditure	Class Frequency	Height of Rectangles	class width
4.5 - 9.5	8	8	5
9.5 - 14.5	28	28	5
14.5 - 19.5	27	27	5
19.5 - 24.5	12	12	5
24.5 - 29.5	4	4	5
Total	80	80	



4. Frequency Polygon: Frequency Polygon provides an alternative to a histogram as way of graphical presenting a frequency distribution of a continuous way. The presentation involves placing to mid values on the horizontal axis & on the vertical axis.



5. Cumulative Frequency Polygon or Ogive. (ক্রমযোজিত).

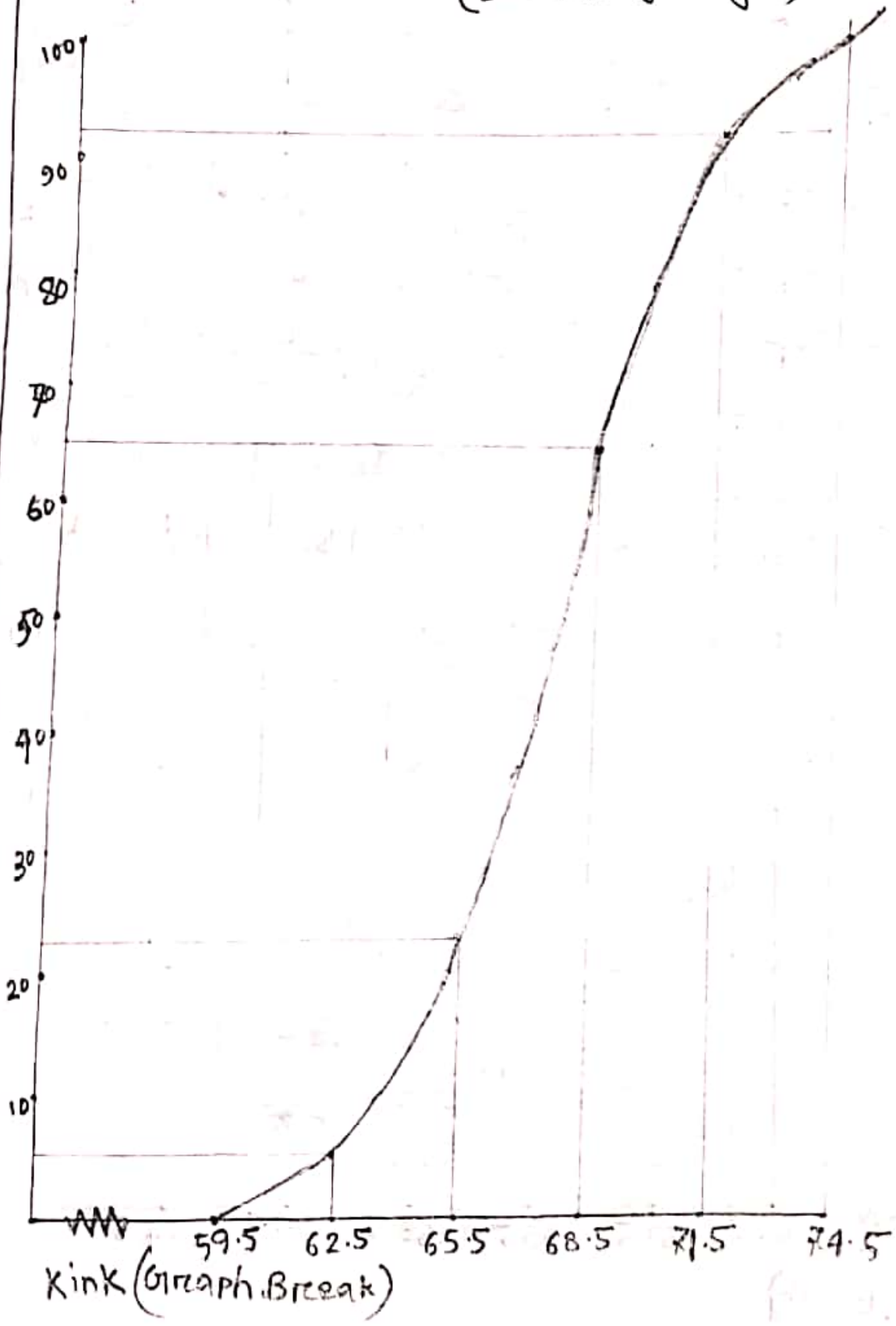
An ogive is based on a cumulative frequency distribution. The frequencies are to be cumulated just by summing the class frequencies. Two types of cumulative distribution are used "less than type" & "more than type".

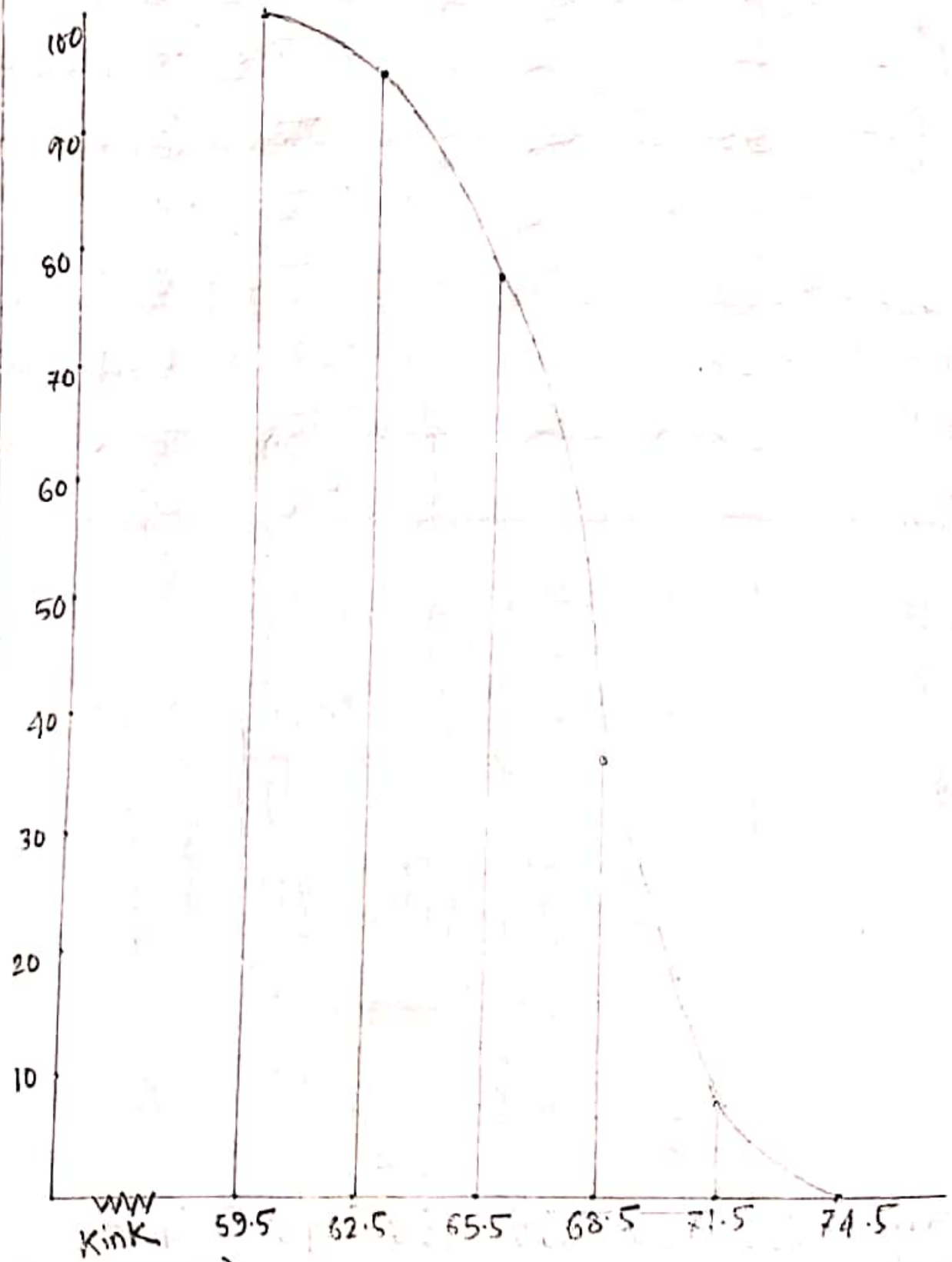
In a less than type cumulative frequency distribution, for each less, the cumulative frequency shot the total number of data items with value. A more than type is also constructed simply by cumulating the frequencies from lower limit of the first class boundary. The first cumulating shows the total frequency in the distribution.

Q table	frequency
59.5-62.5	5
62.5-65.5	18
65.5-68.5	42
68.5-71.5	27
71.5-74.5	8
Total	100

Height (inch)	Frequencies	Cumulative Frequencies	Cumulative frequency Polygon.			
			Less than type		More than type.	
			Height	Frequency	Height	Frequency
59.5-62.5	5	5	less than 59.5	0	More than 59.5	100
62.5-65.5	18	23	" " 62.5	5	" " 62.5	95
65.5-68.5	42	65	" " 65.5	23	" " 65.5	77
68.5-71.5	27	92	" " 68.5	65	" " 68.5	35
71.5-74.5	8	100	" " 71.5	92	" " 71.5	8
Total	100		" " 74.5	100	" " 74.5	0

(Less than type ogive).





(Graph Break)

Average or Measures of Central tendency:

1) The Arithmatic Mean.
2) The Median.
3) The Mode.
4) The Harmonic Mean.
5) The Geometric Mean.

1) $$\bar{x} = \frac{\Sigma x_i}{N}$$

if $x = 1, 2, 3, 9, 7, 10$.

$$\bar{x} = \frac{\Sigma x_i}{N} = \frac{1+2+3+9+7+10}{6} = 5.33$$

For Grouped Data

$$\bar{x} = \frac{\Sigma x_i f_i}{\Sigma f_i}$$

↳ (Arithmatic Mean).

Weekly Wages in Tk	Frequency (f_i)	Mid Value (x_i)	$f_i x_i$
48.5 - 53.5	2	51	102
53.5 - 58.5	2	56	112
58.5 - 63.5	3	61	183
63.5 - 68.5	5	66	330
68.5 - 73.5	5	71	355
73.5 - 78.5	5	76	380
78.5 - 83.5	5	81	405
83.5 - 88.5	7	86	602
88.5 - 93.5	10	91	910
93.5 - 98.5	6	96	576
Total	50		3955

$$\therefore \bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i}$$

$$= \frac{3955}{50}$$

$$= 79.1$$

For grouped data,

Median = মধ্যক

$$\text{Median (Me)} = L_0 + \frac{n/2 - F}{f_0} \times h.$$

L_0 = lower limit of the onedian class.

n = Total number of classes.

F = Cumulative frequency prior the median class.

f_0 = frequency of the median class.

h = class width of the median class.

Calculating the median class.

1) Compute less than type cumulative frequency.
2) Determine $n/2$.
3) Locate median class for which the cumulative frequency is more than $n/2$.
4) Determine the lower limit (L_0) of the median class.
5) Determine the sum of frequencies of all classes prior the median class (F).
6) Determine the frequency of the median class(f_0).
7) Determine class width of the median class(h).

Weakly Wages in Tk	Frequency	Cumulative Frequency
48.5 - 53.5	2	2
53.5 - 58.5	2	4
58.5 - 63.5	3	7
63.5 - 68.5	5	12
68.5 - 73.5	5	17
73.5 - 78.5	5	22
78.5 - 83.5	5	27
83.5 - 88.5	7	34
88.5 - 93.5	10	44
93.5 - 98.5	6	50
Total	50	

$n/2 = \frac{50}{2} = 25$

Median class = 78.5 - 83.5

$L_0 = 78.5$ $F = 22$, $f_0 = 5$, $h = 5$

$(Me) = 78.5 + \frac{25-22}{5} \times 5$

$= 81.5$

Cumulative = ক্রমযোজিত গণসংখ্যা।

1, 2, 3, 1, 5, 6, 1, 5, 1

For Grouped Data

$$\text{Mode } (M_0) = L_0 + \frac{A_1}{A_1 + A_2} \times h.$$

L_0 = Lower limit of the modal class.

A_1 = Absolute difference of frequency between modal & Pre-modal class.

A_2 = Absolute difference of frequency between modal & modal-class.

h = class width of the modal class.

class boundary	Frequency
1.45 - 1.95	2
1.95 - 2.45	1
2.45 - 2.95	4
2.95 - 3.45	15
3.45 - 3.95	10
3.95 - 4.45	5
4.45 - 4.95	3
Total	

$L_0 = 2.95$, $\Delta_1 = 15-4 = 11$

$\Delta_2 = 15-10 = 5$

$h = 0.5$

$$\text{Mode} = L_0 + \frac{\Delta_1}{\Delta_1+\Delta_2} \times h$$

(প্রচুরক)

$$= 2.95 + \frac{11}{11+5} \times 0.5$$

$$= 2.95 + \frac{11}{16} \times 0.5$$

$$= 2.95 + 0.34375$$

$$= 3.29375$$

Case:

প্রচুরক যখন একাধিক থাকে একই

2.45 - 2.95	4
2.95 - 3.45	10
3.45 - 3.95	10
3.95 - 4.45	5

তখন প্রচুরক শ্রেনির গনসংখ্যা (একাধিক) যোগ করে ঐ যোগফল এর সাথে পূর্ববর্তী শ্রেনির বিয়োগ বাদ দিয়ে Δ_1, Δ_2 বের করতে হবে।

& পরবর্তী

Modal class = 2.95 - 3.95

$\Delta_1 = 20-4 = 16$

$\Delta_2 = 20-5 = 15$

$h = 1.00$

Ex: Calculate AM, Median, Mode from the following data.

Age in Years	Number of births
14.5 - 19.5	677
19.5 - 24.5	1908
24.5 - 29.5	1737
29.5 - 34.5	1040
34.5 - 39.5	294
39.5 - 44.5	91
44.5 - 49.5	16

Soln:

Age in Years	Frequency (f_i)	Mid Value (x_i)	$f_i x_i$	Cumulative	
14.5 - 19.5	677	17	11509	677	
19.5 - 24.5	1908	22	41976	2585	
24.5 - 29.5	1737	27	46899	4322	
29.5 - 34.5	1040	32	33280	~~4354~~ 5362	
34.5 - 39.5	294	37	10878	5656	
39.5 - 44.5	91	42	3822	5747	
44.5 - 49.5	16	47	752	5763	
Total	5763		149116		

Arithmatic Mean (AM),

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i} = \frac{149116}{5763} = 25.87$$

$$n/2 = \frac{5763}{2} = 2881.5$$

Median class = 24.5 - 29.5

$L_0 = 24.5$, $f = 2585$

$f_0 = 1737$, $h = 5$

$$(Me) = 24.5 + \frac{2881.5 - 2585}{1737} \times 5$$

$$= 24.5 + \frac{296.5}{1737} \times 5$$

$$= 24.5 + 0.85$$

$$= 25.35$$

Mode: $L_0 = 19.5$, $\Delta_1 = 1908 - 677 = 1231$

$\Delta_2 = 1908 - 1737 = 171$

$h = 5$

$$\text{Mode} = L_0 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h$$

$$= 19.5 + \frac{1231}{1231 + 171} \times 5$$

$$= 19.5 + 4.390$$

$$= 23.89$$

② Calculate AM, Median & Mode from the following data.

Age	No. of Woman
9.5 - 14.5	27
14.5 - 19.5	34
19.5 - 24.5	41
24.5 - 29.5	45
29.5 - 34.5	45
34.5 - 39.5	43
39.5 - 44.5	35
44.5 - 49.5	30

Soln:

Age	Mid Value (x_i)	Frequency (f_i)	$f_i x_i$	Cumulative
9.5 - 14.5	12	27	324	27
14.5 - 19.5	17	34	578	61
19.5 - 24.5	22	41	902	102
24.5 - 29.5	27	45	1215	147
29.5 - 34.5	32	45	1440	192
34.5 - 39.5	37	43	1591	235
39.5 - 44.5	42	35	1470	270
44.5 - 49.5	47	30	1410	300
Total		300	8930	

Arithmatic Mean (AM);

$$\bar{x} = \frac{\sum f_i x_i}{f_i} = \frac{8930}{300} = 29.77$$

$$n/2 = \frac{300}{2} = 150$$

Median class = 29.5 - 34.5

$L_0 = 29.5$, $F = 147$, $f_0 = 45$, $h = 5$

$$(Me) = 29.5 + \frac{150 - 147}{45} \times 5$$

$$= 29.5 + \frac{3}{45} \times 5$$

$$= 29.83$$

Mode;

$L_0 =$

$\Delta_1 = 90 - 41 = 49$

$\Delta_2 = 90 - 43 = 47$

$h = 5$

$$\text{Mode} = L_0 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h$$

Ex: Construct Histogram, frequency Polygon, less than type & more than ogive.

Soln: Next Note Book.

* The standard deviation or other measures of dispersion.
  1) The range.
  2) The quartile deviation.
  3) The Mean Deviation.
  4) The Variance.
  5) The standard deviation.

* The Mean Deviation:

For Grouped Data,

$$MD(\bar{x}) = \frac{\sum f_i |x_i - \bar{x}|}{\sum f_i}$$

where,

x_i = Mid values of class boundary.

$\bar{x}$ = AM (Arithmatic Mean).

f_i = Total frequency.

Compute Mean Deviation (MD).

Class Interval	Frequency
48.5 - 53.5	2
53.5 - 58.5	12
58.5 - 63.5	3
63.5 - 68.5	5
68.5 - 73.5	5
73.5 - 78.5	5
78.5 - 83.5	5
83.5 - 88.5	7
88.5 - 93.5	10
93.5 - 98.5	6
Total	50

Class Interval	Mid Value (x_i)	Frequency (f_i)	$f_i x_i$	$\bar{x}$	$\lvert x_i - \bar{x} \rvert$	$f_i \lvert x_i - \bar{x} \rvert$	$(x_i - \bar{x})^2$	$f_i (x_i - \bar{x})^2$
48.5 - 53.5	51	2	102		28.1	56.2	789.61	1579.22
53.5 - 58.5	56	2	112	$\bar{x} = \frac{\sum f_i x_i}{\sum f_i}$	23.1	46.2	533.61	1067.22
58.5 - 63.5	61	3	183		18.1	54.3	327.61	982.83
63.5 - 68.5	66	5	330		13.1	65.5	171.61	858.05
68.5 - 73.5	71	5	355	$= \frac{3955}{50}$	8.1	40.5	65.61	328.05
73.5 - 78.5	76	5	380	$= 79.1$	3.1	15.5	9.61	48.05
78.5 - 83.5	81	5	405		1.9	9.5	3.61	18.05
83.5 - 88.5	86	7	602		6.9	48.3	47.61	333.27
88.5 - 93.5	91	10	910		11.9	119	141.61	1416.1
93.5 - 98.5	96	6	576		16.9	101.4	285.61	1713.66
Total		50	3955			556.4		8344.5

$$MD(\bar{x}) = \frac{\sum f_i |x_i - \bar{x}|}{\sum f_i}$$

$$= \frac{556.9}{50}$$

$$= 11.128$$

Variance:

$$\sigma^2 = \frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i}$$

$$= \frac{8314.5}{50}$$

$$= 166.89$$

Standard Deviation:

$$\sigma = \sqrt{\frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i}}$$

$$= \sqrt{166.89}$$

$$= 12.91852125$$