[4] Histogram :

- A **histogram** is a graphical display of data using bars of different heights. In a **histogram**, each bar groups numbers into ranges. Taller bars show that more data falls in that range. A **histogram** displays the shape and spread of continuous sample data. In a histogram no gaps between the bars.

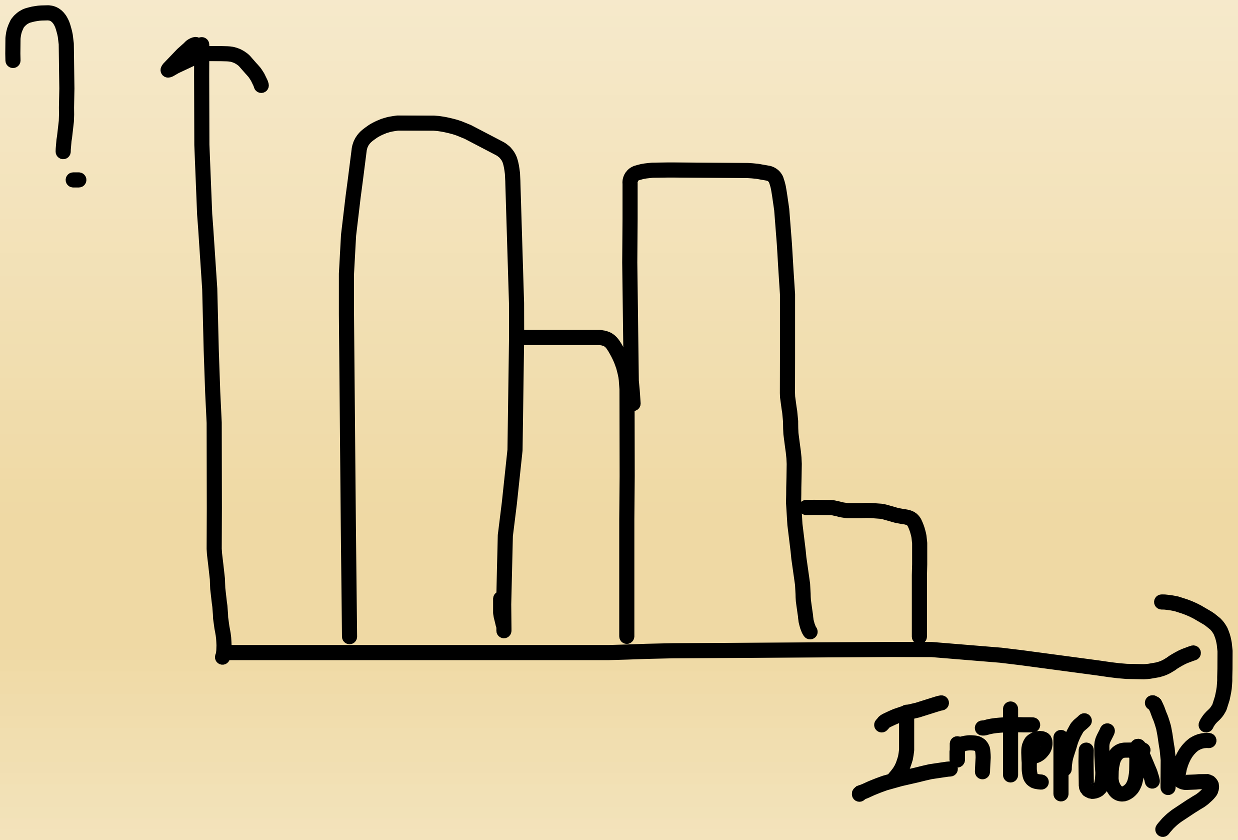
- **We have two cases:**

(1) Histogram for grouped data:

(a) Intervals ‘ classes ’ with equal width:

X – axis ⟶ Intervals.

Y – axis ⟶ Frequencies.



Example (1) : The following table shows a grouped frequency distribution for the statistics grades:

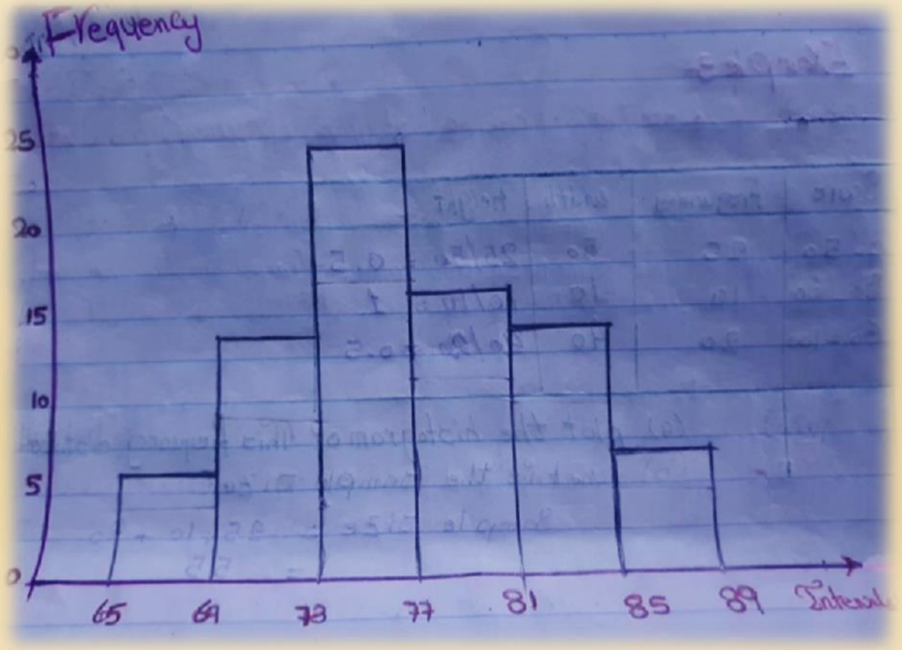
Intervals	Frequency	Width = upper limit – lower limit
65 – 69	6	4
69 – 73	13	4
73 – 77	24	4
77 – 81	16	4
81 – 85	14	4
85 – 89	7	4

a) What is the sample size ?

sample size = the sum of all frequencies

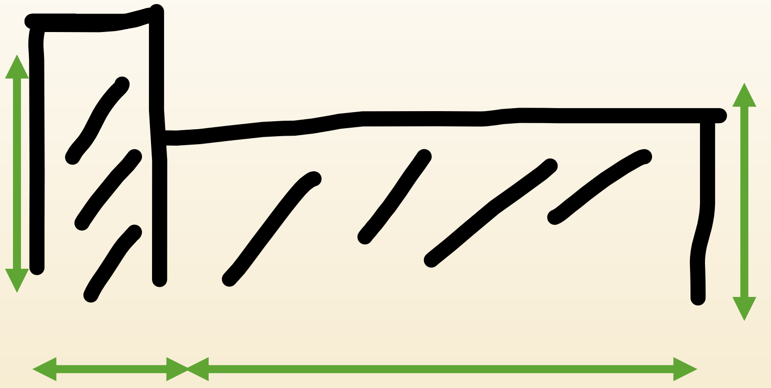
$= 6 + 13 + 24 + 16 + 14 + 7 = 80$

b) Plot the histogram of this frequency distribution table?



(1) Histogram for grouped data:

(b) Intervals ‘ classes ’ with unequal width:



X – axis $\longrightarrow$ Intervals.

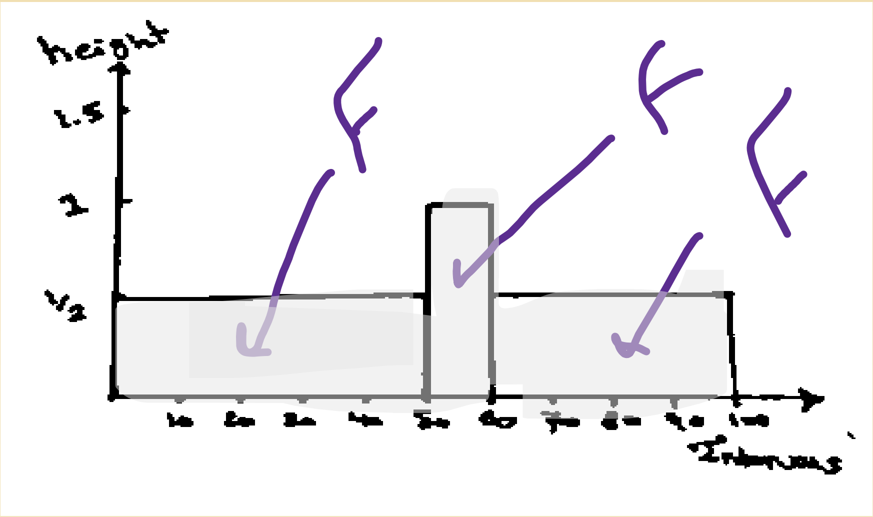
Frequency = Area of rectangle

= Height × Width $\quad \Rightarrow \quad \text{Height} = \dfrac{\textit{Frequaency}}{\textit{Width}}$

Y – axis $\longrightarrow$ Height.

Example (2) : The following table shows a grouped frequency distribution for the statistics grades:

Intervals	Frequency	Width = upper limit – lower limit	Height $= \dfrac{\textit{Frequaency}}{\textit{Width}}$
0 – 50	25	50	$1/2$
50 – 60	10	10	1
60 – 100	20	40	$1/2$



(2) Histogram for ungrouped data:

Step 1 : Find the smallest and largest data point.

Step 2 : Decide how many bins you need using

$\sqrt{n}$, where n: Sample size.

let $k = \sqrt{n}$ = number of intervals.

Step 3 : Divided the range ‘ R ’of the data (R = largest value – smallest value) by the number of intervals ‘ k ‘ .

$$L = \text{The width of each interval} = \frac{largest - smallest}{number\ of\ intervals} = \frac{R}{K}$$

Step 4 : create the bin boundaries by starting with your smallest number ‘ or less ‘, and adding the bin size (L). The last interval should contain the largest number.

Sheet (1)

6. The following data represent the length of life in years, measured to the nearest tenth, of 30 similar fuel pumps:

2.0	3.0	0.3	3.3	1.3	0.4
0.2	6.0	5.5	6.5	0.2	2.3
1.5	4.0	5.9	1.8	4.7	0.7
4.5	0.3	1.5	0.5	2.5	5.0
1.0	6.0	5.6	6.0	1.2	0.2

6

5

Set up a frequency and a relative frequency distribution histogram.

Answer

(1) Largest value $=$ 6.5 , Smallest value $=$ 0.2

(2) n $=$ Sample size = $6 \times 5 = 30.$

k $=$ number of intervals $= \sqrt{n} = \sqrt{30} = 5.4 = 6.$

(3) L $=$ Width of each interval $= \dfrac{largest - smallest}{k} = \dfrac{6.5-0.2}{6} = 1.05 = 1.$

0.2 — 1.2

1.2 — 2.2

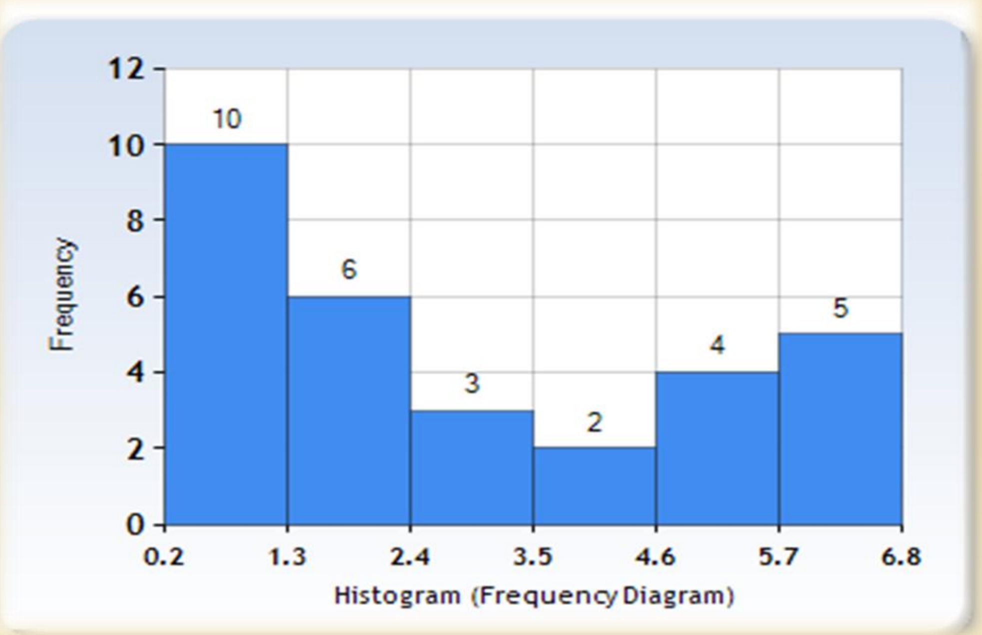
2.2 — 3.2

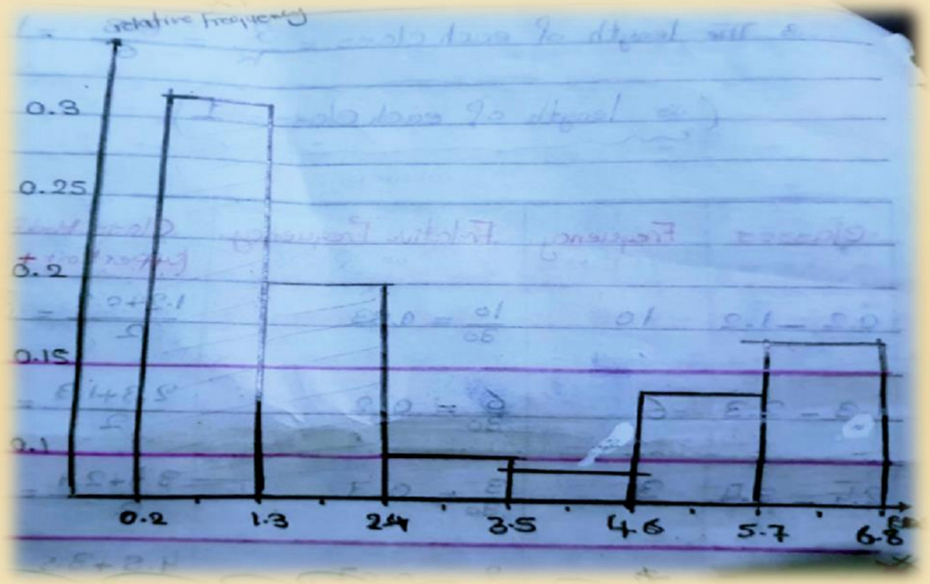
3.2 — 4.2

4.2 — 5.2 $\xrightarrow{x}$ 6.5

5.2 — 6.2

Intervals	Frequency ‘ F ’	Relative Frequency ‘ RF’ = $\frac{F}{n}$
0.2 – 1.2	10	$^{10}/_{30} = 0.33$
1.3 – 2.3	6	$^{6}/_{30} = 0.2$
2.4 – 3.4	3	$^{3}/_{30} = 0.1$
3.5 – 4.5	2	$^{2}/_{30} = 0.07$
4.6 – 5.6	4	$^{4}/_{30} = 0.13$
5.7 – 6.7	5	$^{5}/_{30} = 0.17$





Sheet (1)

7. The following data represent the length of life in seconds, of 50 fruit flies subject to a new spray in a controlled laboratory experiment:

10

5

17	20	10	9	23	13	12	19	18	24
12	14	6	9	13	6	7	10	13	7
16	18	8	13	3	32	9	7	10	11
13	7	18	7	10	4	27	19	16	8
7	10	5	14	15	10	9	6	7	15

Set up a frequency distribution histogram. Draw an estimate of the graph of the distribution?

Answer

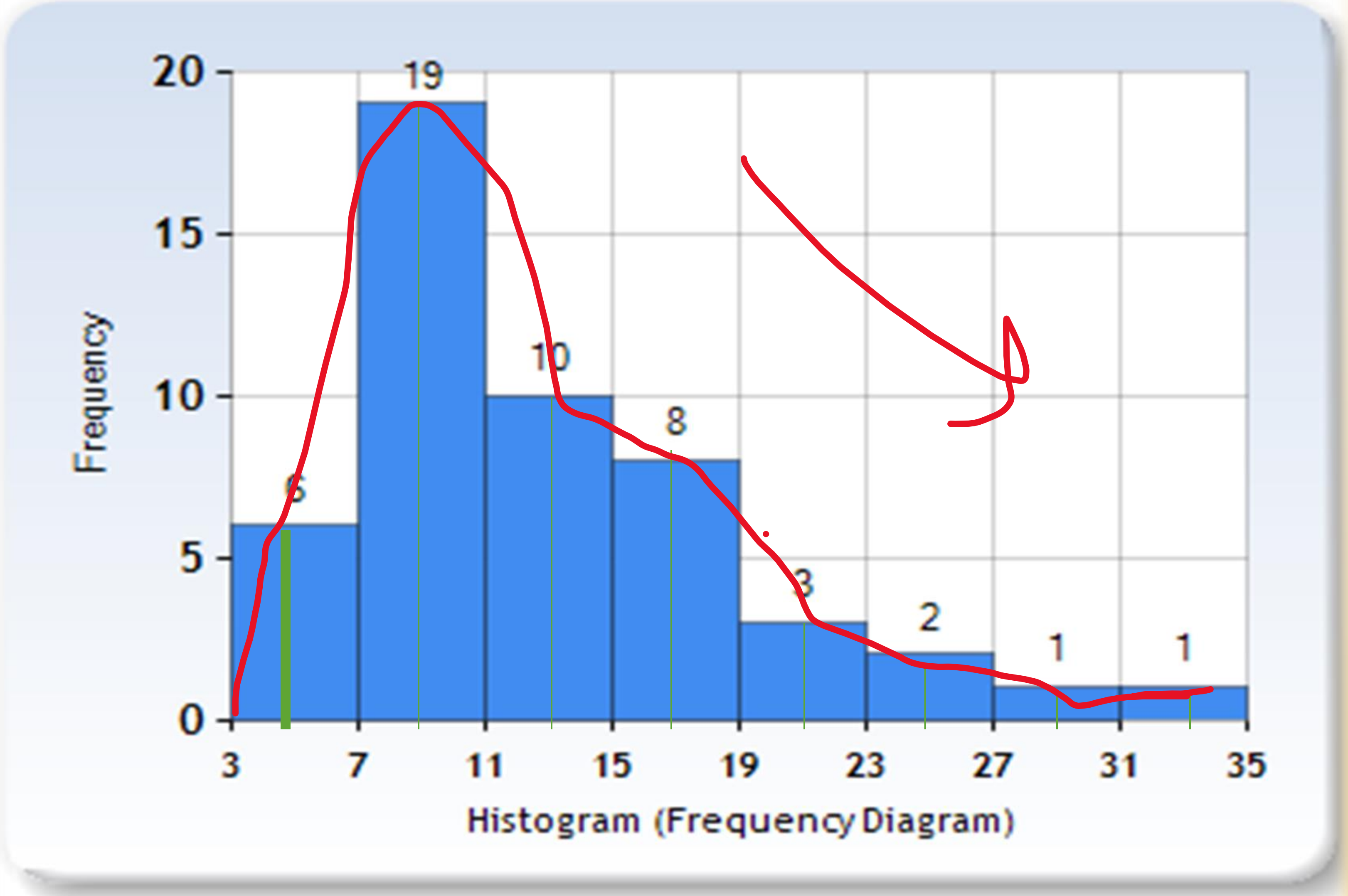
(1) Largest value $= 32$, Smallest value $= 3$

(2) $n =$ Sample size = $10 \times 5 = 50.$

k = number of intervals $= \sqrt{n} = \sqrt{50} = 7.07 = 8.$

(3) L = Width of each interval $= \frac{largest - smallest}{k} = \frac{32-3}{8} = 3.6 = 4.$

Intervals	Frequency ' F '	Class midpoint $= \frac{lower\ limit + upper\ limit}{2}$
3 – 7	6	$3 + 7/_2 = 5$
7 – 11	19	$7 + 11/_2 = 9$
11 – 15	10	13
15 – 19	8	17
19 – 23	3	21
23 – 27	2	25
27 – 31	1	29
31 – 35	1	33


20
15
10
5
0
Frequency
19
10
8
6
3
2
1
1
3
7
11
15
19
23
27
31
35
Histogram (Frequency Diagram)