

Statistics :- is a study that estimates the status of a population by using information gained from samples.

Population :- is a collection of all possible individual objects, or measurements of interest.

Sample :- is a portion of the population of interest.

Data Matrix // Overview of the cases and variables.

Data Matrix examples :-

Student name	Age	Mail	weight	Grade
Mohamed	15	@	55	100
Ali	20	@	60	86
Daaa	23	@	70	84
Jane	25	@	45	93

Cases + variables

= Data Matrix

→ Cases

→ Variables

Types of variables "Data"

(1) Categorical "Qualitative"

Sub: _____

Date: _____

Data cannot be measured, describe characteristics.

examples of the categorical &

Types of a car, eye color, gender.

(2) Numerical "Quantitative"

Data can be measured, deals with numbers.

examples of the numerical data &

number of children in a family, minutes remaining in a class.

Type of Numerical Data &

(1) Discrete variable. (1, 2, 3, ..., etc)

(2) Continuous variable. (Interval)

Graphic Presentation &

Describing Data using graph,
to understand the data.

Ungrouped data & ungrouped data have not been summarized in any way.

Grouped data & grouped data have been organized into a frequency distribution.

Frequency distribution & is a grouping of data into mutually exclusive categories showing the number of observations in each class.

Class midpoint: Class mid-point is a point that divides a class into two equal parts

$$\text{Class midpoint} = \frac{\text{Lower limit} + \text{Upper limit}}{2}$$

Graphical presentation of ungrouped data

[1] Dot plot / Dot chart: Describe numerical data

a dot plot is a graphical display of data using dots
X-axis → what is being measured?

Example 8.8.1

How long does it take to eat breakfast

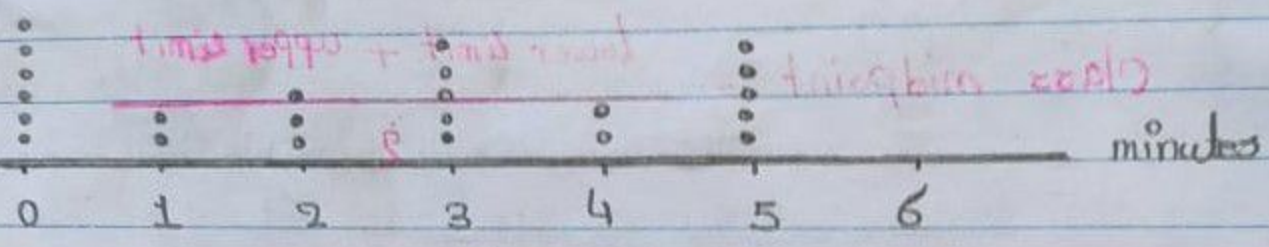
minutes	0	1	2	3	4	5
people	6	2	3	5	2	5

(a) what is the sample size?

(b) Construct a dot plot graph?

(a) Sample size = $6 + 2 + 3 + 5 + 2 + 5 = 23$

(b) dot plot graphs



Sheet 1 :-

1. The following measurements were recorded for the drying time, in hours, of a certain brand of latex paint.

3.4	2.5	4.8	2.9	3.6
2.8	3.3	5.6	3.7	2.8
4.4	4.0	5.2	3.0	4.8

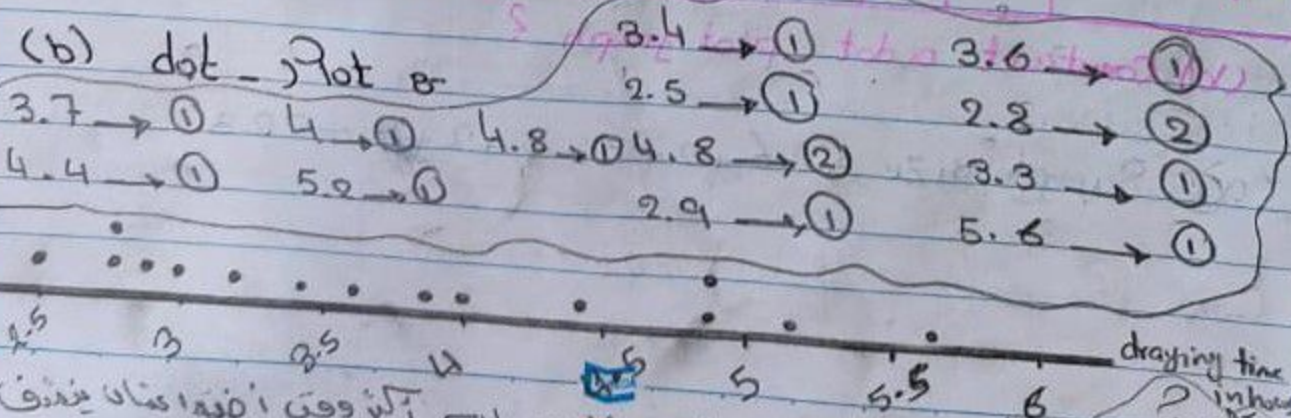
Assume that the measurements are a simple random sample.

(a) What is the sample size for the above example?

(b) Plot the data by way of a dot plot?

(a) Sample size = 15

(b) dot plot



لو جيت داه و تالت ايت اكر وقت اذنوا زمان فيدوق ايت

كله تيرورة
واحد 2.8, 4.8

5

[2] Bar graph & "Bar chart" :- Describe categorical data

A bar graph is a graphical display of data using bars "rectangle" of different heights. A bar graph can be used to show something changes over time or to compare items.

How to construct a bar graph :-

- X-axis :- what is being measured. \rightarrow cat
- Y-axis :- has a number for the amount of stuff being measured. \rightarrow variable
- There exist gaps between the bars.

Example :-

Q. Which kind of movie they liked best

movie	Comedy	action	romance	Drama
people <small>variable</small>	4	5	6	1

- (a) what is the sample size for the above example?
- (b) pbt the data by way of a bar-graph?

(a) Sample size = $4 + 5 + 6 + 1 = 16$

(b) Bar-graph

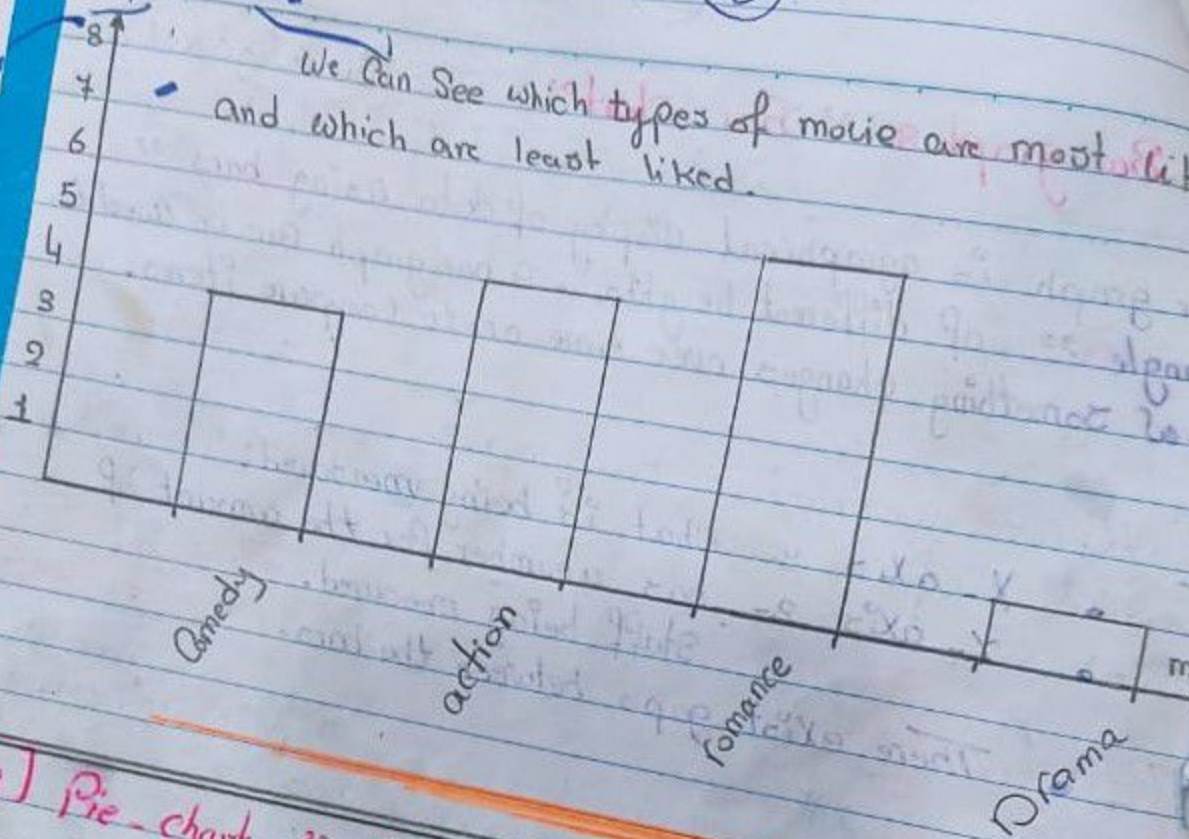
Sub:

Date:

People = height of the bar

(6)

We can see which types of movie are most liked and which are least liked.

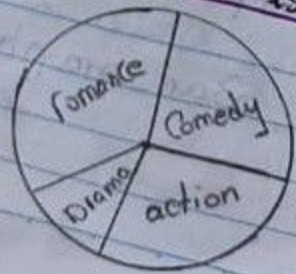


[3] Pie-chart "circle chart" 8. Describe categorical data

special chart that uses "pie slices" to show relative size.

movie	Comedy	action	romance	Drama	Total
people	4	5	6	1	16
relative size	$\frac{4}{16} \times 100 = 25\%$	$\frac{5}{16} \times 100 = 31.25\%$	$\frac{6}{16} \times 100 = 37.5\%$	$\frac{1}{16} \times 100 = 6.25\%$	

Construct a pie-chart graph?



انگریزی میں لکھو اور اردو میں لکھو

7

Sheet 1

4. according to the pie graph, which of the following statements is false?



(a) More than half the animals on the farm are cows? True

(b) one quarter of the animals on the farm are chicken? False True

(c) There are more pigs than cats on the farm? True

(d) Fewer than one quarter of the animals on the farm are pigs True

(e) no cats on the farm have given birth to cows? True

Graphic presentation of a grouped data

[1] Histogram

is a graph in which

- (1) The classes are marked on the horizontal axis (X-axis)
- (2) The class frequencies on the vertical axis (Y-axis)
- (3) No gaps between the bars.
- (4) display of data using bars (rectangle)

We have two cases:

- (1) classes (Interval with equal class interval (widths))

Sub: _____

Date: _____

(8)

Frequency

Class interval (width, length) :-
upper limit - lower limit

Histogram in the case of the class intervals are equal :-

- 1] X-axis → Intervals or class mid point.
- 2] Y-axis → The class Frequencies.

Examples-

The following table shows a grouped frequency distribution for the statistics grades :-

Classes	65-69	69-73	73-77	77-81	81-85	85-89
Frequency	6	13	24	16	14	7

a) plot The histogram of this frequency distribution ?

Classes	Frequency	width
65-69	6	4
69-73	13	4
73-77	24	4
77-81	16	4
81-85	14	4
85-89	7	4

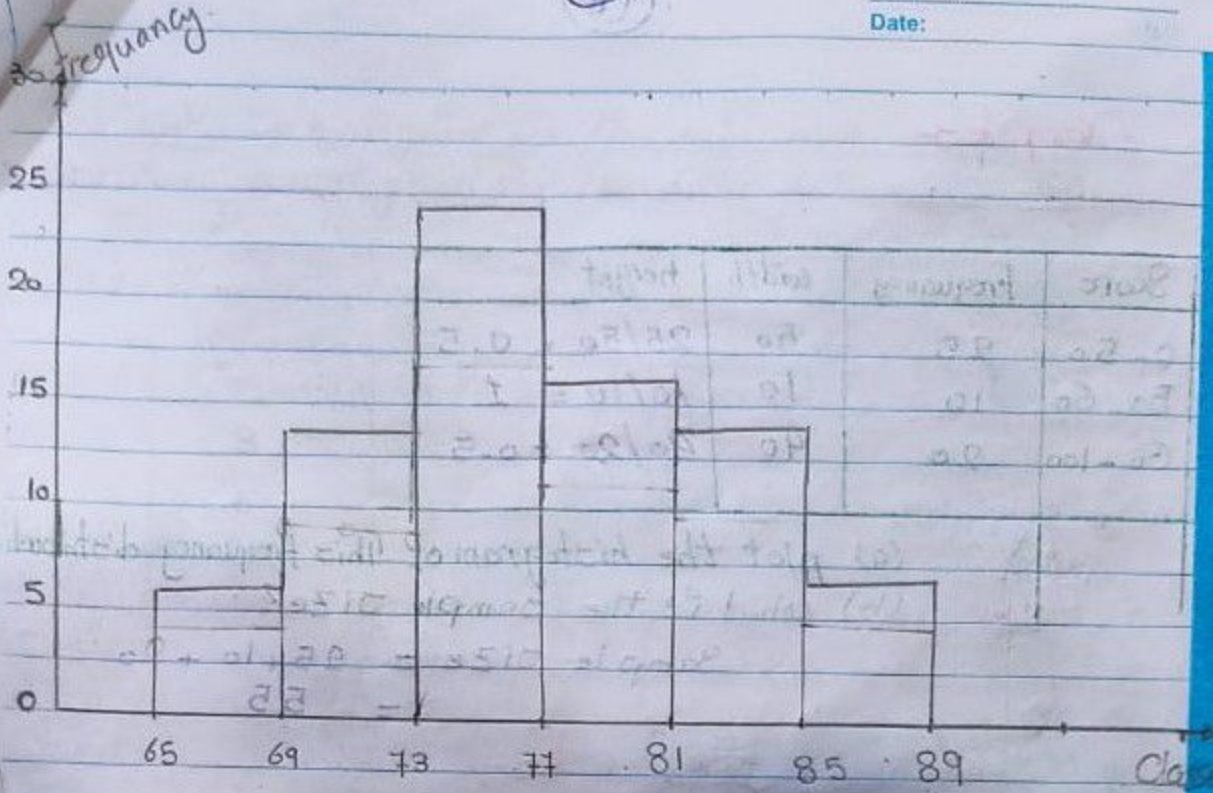
(b) what is The sample size ?

$$\begin{aligned} \text{Sample size} &= \text{Sum of all frequencies} \\ &= 6 + 13 + 24 + 16 + 14 + 7 \\ &= 80 \end{aligned}$$

بما ان العرض اربعة يبقى بقى اقول 6
 ال France هو الدول المنتجة
 تقارن بالقول ان العرض

تقارن بالقول ان العرض

اثنو رونا



(2) classes (Interval) with unequal class Interval (width), length

- 1] Find The length (width) of each classe (Interval).
- 2] area of each bar = frequency.

area of rectangle = height \times width.

frequency = height \times width

height = $\frac{\text{frequency}}{\text{width}}$

3] X-axis \rightarrow Intervals or class mid point

4] Y-axis \rightarrow height = $\frac{\text{Frequency}}{\text{width}}$

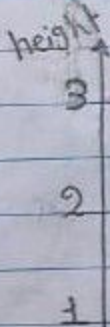
Sub: _____

Date: _____

(10)

Example 3-

Score	Frequency	width	height
0-50	25	50	$25/50 = 0.5$
50-60	10	10	$10/10 = 1$
60-100	20	40	$20/40 = 0.5$



(a) plot the histogram of this frequency distribution
 (b) what is the sample size?

$$\text{Sample size} = 25 + 10 + 20 = 55$$

كبيره (a) رسم المخطط التكراري لهذا التوزيع
 (b) ما هو حجم العينة؟

Sheet 1 3-

3. Different cats weighed, we had a tough time getting them all to sit on the scale without scratching us half to death. anyway, we finally found them to have the following weights in pounds:

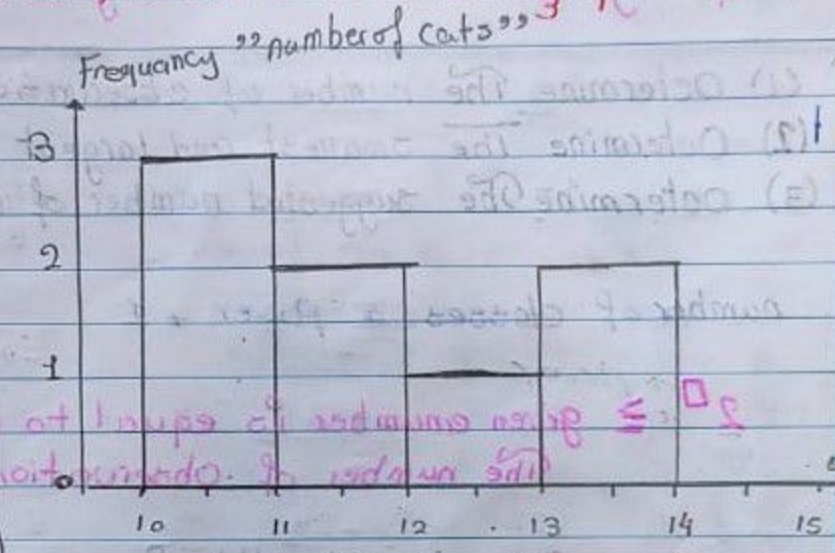
- 10, $10\frac{1}{4}$, $10\frac{1}{2}$, $11\frac{1}{4}$, $11\frac{2}{3}$, $12\frac{3}{4}$, 13, $13\frac{1}{2}$



(a) Draw a histogram for the data with 1 pound.
 (b) Draw a histogram for the data with $\frac{1}{3}$ pounds.

(a)

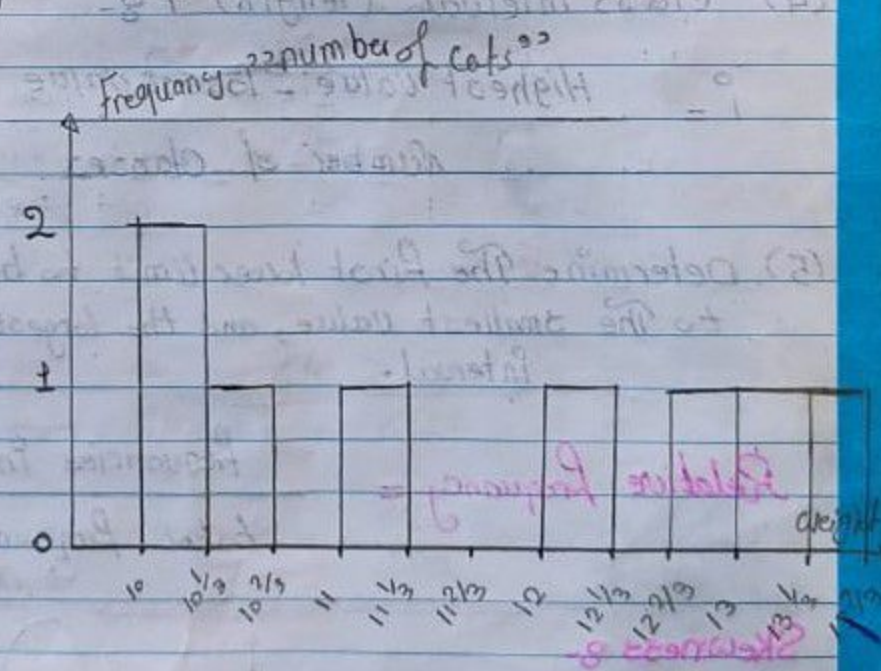
Weight	Frequency
10 - 11	3
11 - 12	2
12 - 13	1
13 - 14	2



10-11 } 2 cats
 12-13 } 1 cat
 13-14 } 2 cats

(b)

Weight	Frequency
10 - 10 $\frac{1}{3}$	2
10 $\frac{1}{3}$ - 10 $\frac{2}{3}$	1
10 $\frac{2}{3}$ - 11	0
11 - 11 $\frac{1}{3}$	1
11 $\frac{1}{3}$ - 11 $\frac{2}{3}$	0
11 $\frac{2}{3}$ - 12	1
12 - 12 $\frac{1}{3}$	0
12 $\frac{1}{3}$ - 12 $\frac{2}{3}$	0
12 $\frac{2}{3}$ - 13	1
13 - 13 $\frac{1}{3}$	1
13 $\frac{1}{3}$ - 13 $\frac{2}{3}$	1



Frequency distribution and a histogram of ungrouped data

- (1) Determine the number of observation.
- (2) Determine the smallest and largest number in the data
- (3) Determine the suggested number of classes

$$\text{number of classes} = \sqrt[n]{N} + 1$$

$\sqrt[n]{N}$ \geq given number is equal to or greater than the number of observations.

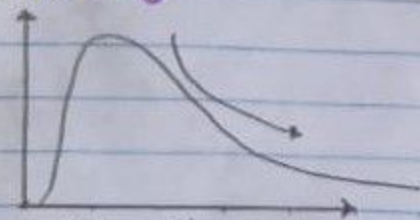
- (4) class interval (length) i is -

$$i = \frac{\text{Highest value} - \text{lowest value}}{\text{number of classes}}$$

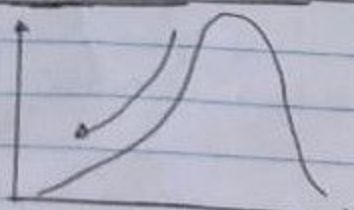
- (5) Determine the first lower limit to be less than or equal to the smallest value, and the largest number is in the last interval.

$$\text{Relative Frequency} = \frac{\text{Frequencies in each interval}}{\text{total frequencies} \rightarrow \text{number of observations}}$$

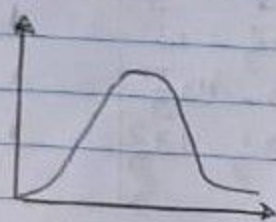
Skewness :-



+ve. Skews to the Right



-ve. Skews to the left



Symmetric

(13)

Sheet 18

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 :-

17	20	10	9	23	13	12	19	18	21
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	24	19	16	8
7	10	5	14	15	10	9	6	7	15

Set up relative frequency distribution and construct a relative frequency histogram? and discuss the skewness of the distribution?

- number of observation = 50
- largest value = 32
Smallest value = 3
- $2^6 \geq 50 \rightarrow 2^6 = 64 > 50$
 \therefore number of classes (intervals) = $6 + 1 = 7$
- length of each interval = $\frac{32 - 3}{7} = \frac{29}{7} = 4.1 \approx 4$
- The first lower limit can be any number from 0 to 3

Relative Frequency of marks

class	frequency	mid-point	relative frequency
0-4	2	$\frac{0+4}{2} = 2$	$\frac{2}{50} = 0.04$
5-9	17	$\frac{5+9}{2} = 7$	$\frac{17}{50} = 0.34$
10-14	16	$\frac{10+14}{2} = 12$	$\frac{16}{50} = 0.32$
15-19	10	$\frac{15+19}{2} = 17$	$\frac{10}{50} = 0.20$
20-24	3	$\frac{20+24}{2} = 22$	$\frac{3}{50} = 0.06$
25-29	1	$\frac{25+29}{2} = 27$	$\frac{1}{50} = 0.02$
30-34	1	$\frac{30+34}{2} = 32$	$\frac{1}{50} = 0.02$
	50		

to
3-7
8-12
13-20
20-27
28-35

7-11
11-15
15-19
19-23
23-27
27-31
31-35

3-7 0-4
8-12 4-8
13-20 8-12
20-27 12-16
28-35 16-20
20-24
24-28
28-

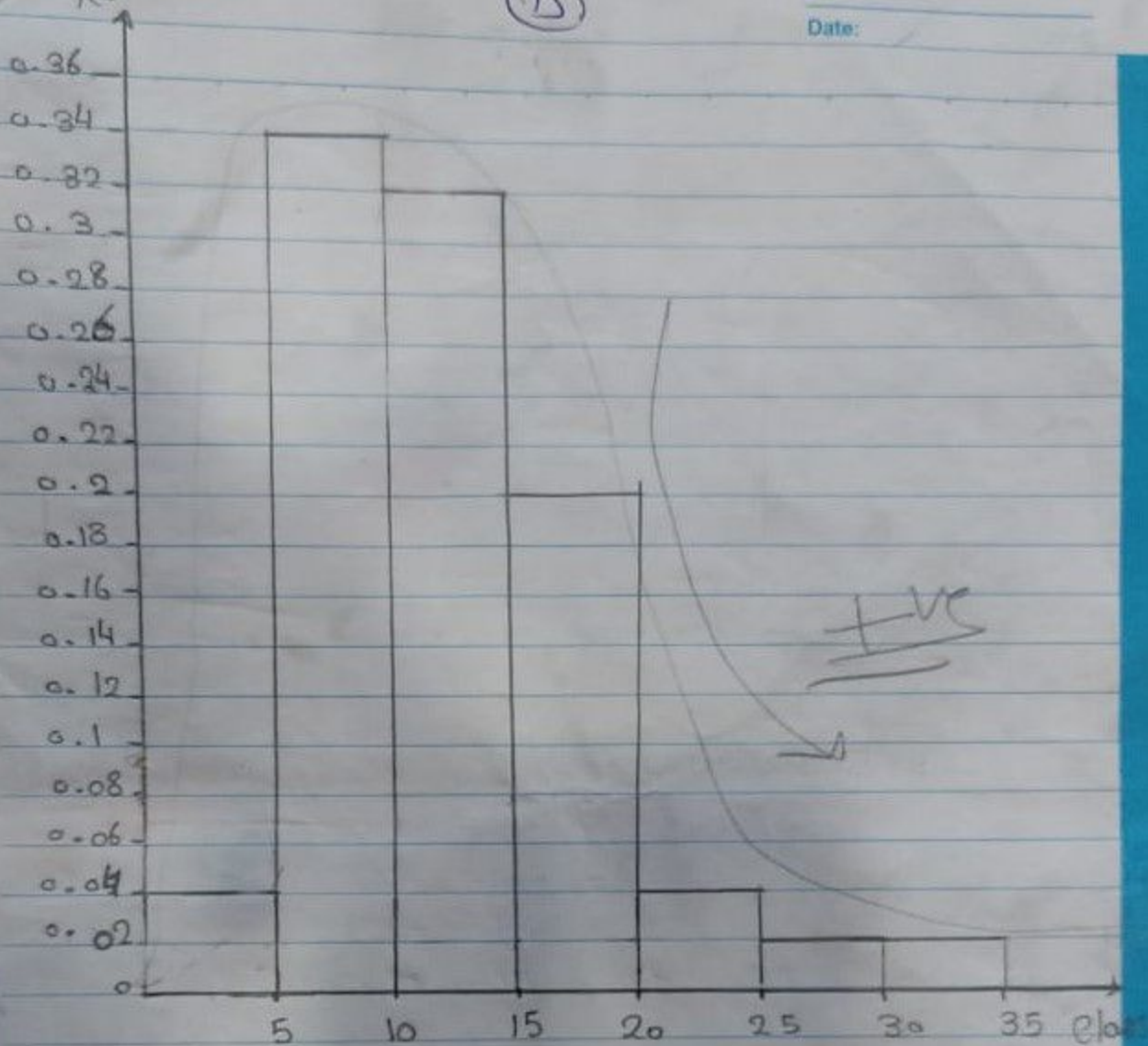


Relative Frequency

15

Sub:

Date:



The distribution skews to The Right.