

**1. Data presentation:** Data presentation refers to the process of organizing, summarizing, and visually representing data in a format that is easy to understand and interpret.

**2. Ways of data presentation:** We can summarize the raw data in the following ways,

a. Frequency distribution

Classification : Male -Female + Hindu-Muslim [group]

Tabulation : Class + freq + . . . . various table

b. Graphical representation

**Frequency distribution:** A frequency distribution of a data set is a listing of classes or categories and their frequencies. It is a statistical tabulated representation of the number of occurrences of each class/category.

A frequency distribution assumes mainly two different forms depending on the character/types of data.

- i. Frequency distribution of qualitative data ✓
- ii. Frequency distribution of quantitative data ✓

**Relative frequency:** The relative frequency for a particular category/class is the ratio of the frequency of that category to the total number of observations.

Blood type	Frequency	Relative frequency
O	14	$14/40=0.35$
A	10	$10/40=0.25$
B	10	0.25
AB	6	0.15
Total	40	1.00

**Interpretation:** Based on the table, it is evident that O type owners constitute 35% of the total students, followed by A type owners at 25%, B type owners at 25%, and AB type owners at only 15%. The most common blood type among students is O, while the least common type is AB.

**Percentage frequency:** Relative frequency  $\times 100$

**Steps of constructing frequency distribution table:** The process of constructing a frequency distribution table involves the following steps:

a) **Choose the Number of Classes:**

b) **Calculate Class Width/Interval:**

c) **Set individual class limits.**

d) **Tally and Count Data:**

e) *Calculate Frequency:*

**9.** A manager of certain computer selling company wants to study the computer selling behavior in his company. The following data give the number of computers sold by the company for a sample of 40 days

39	37	43	27	59	30	34	41	35	49
78	58	48	64	34	54	37	70	65	36
69	54	53	21	24	57	49	33	42	57
62	41	75	71	37	40	60	47	29	39

Summarize these data into frequency and relative distributions chosen a suitable class interval and write the interpretation.

**10.** Mr. XYZ, Chairman of ABC department, is interested in studying the typical weight of student of his department. The chairman randomly selected 60 students from his department and recorded the weights (kg) which are given below,

53.2	61.7	51.7	61.8	46.7	57.4	51.6	51.8	49.9	66.4
67.4	51.6	45.9	52.8	61.9	71.4	38.8	56.8	54.2	57.7
46.1	51.2	44.7	67.8	56.4	58.8	41.2	48.7	49.2	58.0
53.8	53.4	48.2	51.0	36.2	53.4	49.8	53.6	55.6	51.5
46.6	45.4	63.2	46.8	54.2	52.2	73.4	46.7	59.2	42.5
74.6	61.7	58.5	43.9	68.4	40.2	57.5	55.3	63.7	57.5

Summarize these data into frequency and relative distributions chosen a suitable class interval and write the interpretation.

\*\*\* What do you mean by data representation? Why do you need represent statistical data?

\*\*\* What are the methods/way of representing raw data?

\*\*\* What is meant by classification?

\*\*\* Illustrate the following terms in connection with a frequency table: Class, Class interval, Frequency, Relative frequency, Cumulative frequency.

\*\*\* What do you mean by a frequency distribution? Explain various steps of constructing a frequency distribution.

## **11. Graphical presentations:**

trends relationships within the data.

data visualization  
understand patterns,  
easy-to-interpret

### ***Importance:***

- i. Graphs are more attractive than tables of data and make a clear impression in mind.
  - ii. bird's eye view
  - iii. comparison data sets.
  - iv. pattern of distribution explained.

### ***Limitations:***

- i. detail reduced.
  - ii. cannot compute precise estimates

**12. Graphs of Qualitative data:** Two common methods for graphically displaying qualitative data are “Bar chart” and “Pie chart”.

a) *Bar chart:*

b) *Pie chart:*

data as slices of a circle

ideal for showing proportions and percentages.

### **13. How to construct bar chart:**

- a) 

b)

c)

label the categories → Spacing Saving



b)

c)

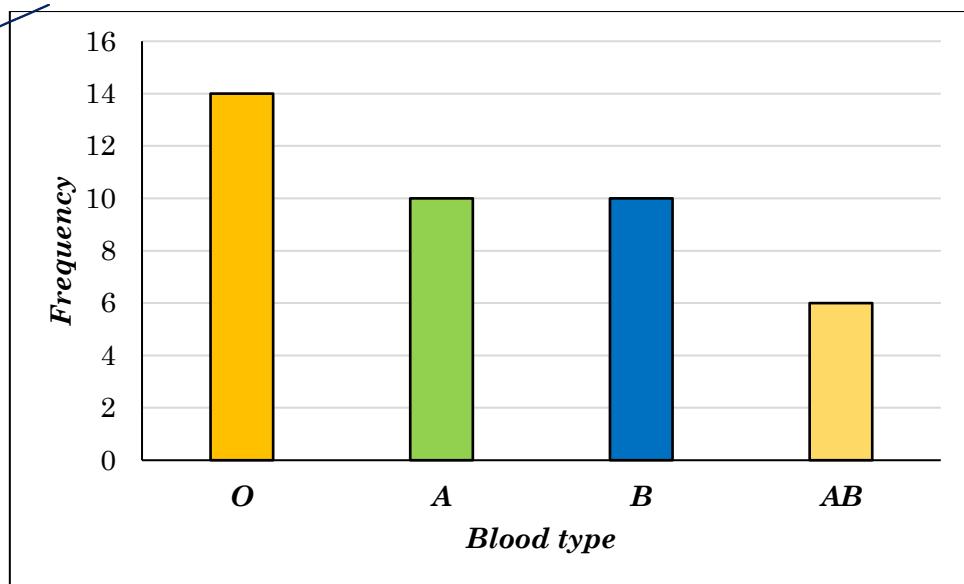
**bars**

**same width**

**do not touch each other.**

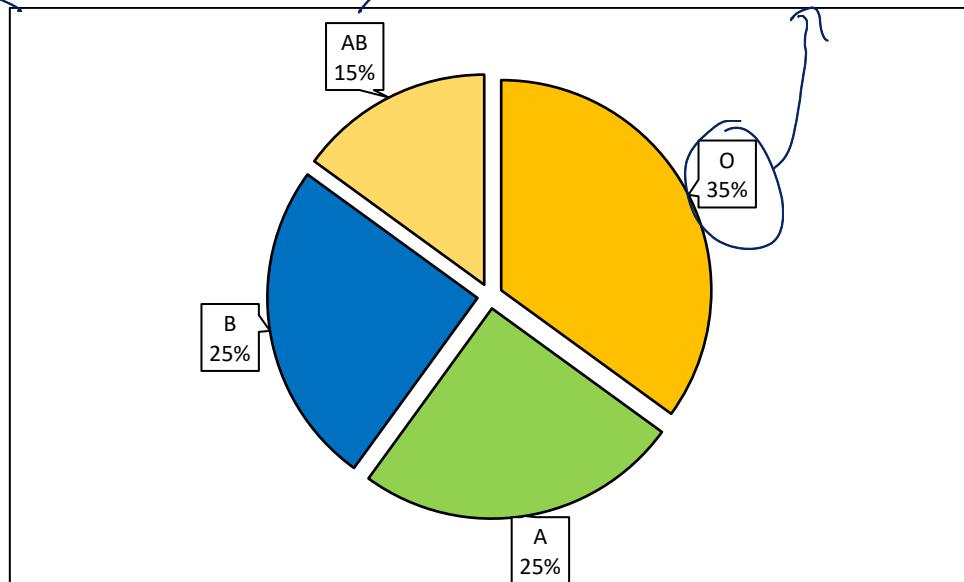


~~Ex:~~ Bar chart of quality data



#### 14. How to construct pie chart:

- Draw a circle
- calculate the "slice" size. *of each quality.* **Slice size =**  
 **$360 \times RF$**
- circle into slices
- Name each slice + (percentage)*



## **Histogram:**

distribution of numerical data.

### **Merits**

Histograms are commonly used to visualize the distribution of continuous data and to identify patterns, trends, and central tendencies within the data. They are particularly useful for exploring the shape of the distribution, including whether it's symmetric, skewed, or has multiple peaks.

a) ***Easy Interpretation:***

b) ***Data Distribution:*** Histograms reveal the shape of the data distribution, symmetric, skewed, bimodal, or multimodal, helps in identifying patterns.

c) ***Central Tendency and Spread:***

central tendency (mean, median) and spread (range, variability) of data.

d) ***Outlier Detection:*** data points significantly different from others, easily identified in histogram as they fall outside typical range.

e) ***Comparison:*** visual comparison multiple data sets informed decisions.

### **Demerits (Disadvantages) of Histograms:**

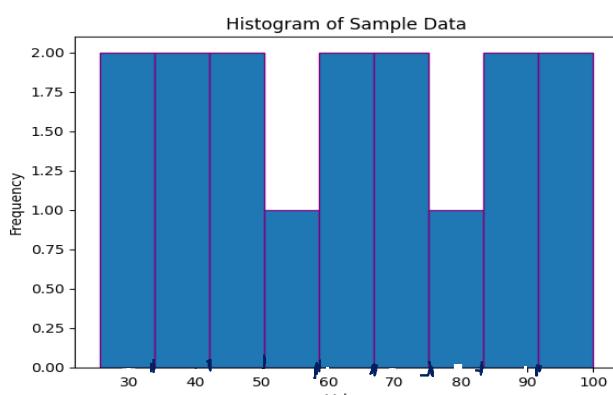
a) ***Data Loss:*** loss of precision.

b) ***Continuous Data Limitation:*** better suited for discrete and continuous data, not for categorical

c) ***Not Ideal for Small Data Sets***

## **16. How to construct histogram:**

- i. horizontal axis class boundaries.
- ii. vertical axis frequencies.
- iii. height equals the frequency
- iv. value or class is centered under its rectangle.



**17. The provided data represents the ages of 40 individuals who have physical disabilities.**  
47, 54, 58, 63, 41, 68, 71, 37, 56, 66, 49, 60, 45, 67, 50, 38, 72, 55, 42, 59, 39, 65, 70, 43, 52, 73, 61, 44, 69, 57, 74, 62, 48, 53, 36, 64, 75, 40, 51, 46.

- a) Construct frequency table.
- b) Draw histogram.

**Histogram vs Bar diagram:** Histograms and bar charts are both graphical representations of data, but they have distinct differences in terms of their purposes, data types, and visual characteristics. Here are four key differences between histograms and bar charts:

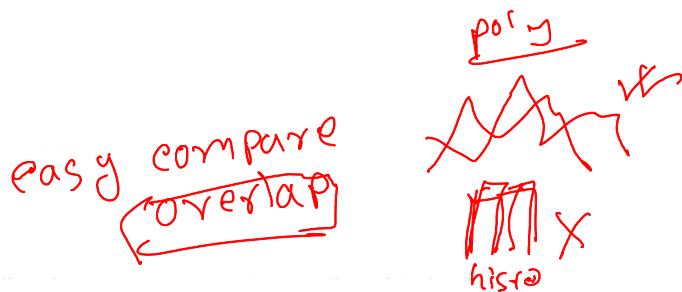
The differences between histogram and bar graph can be drawn clearly on the following grounds:

- a) Histogram
  - frequency of numerical data.
  - bar graph is representation of data that compare categories of data.
- b) histogram
  - frequency distribution of continuous variables
  - bar graph is diagrammatic comparison of discrete variables.
- c) Histogram
  - numerical
  - bar graph categorical
- d) The histogram
  - proper spacing
  - no gap
  - bar graph that indicates discontinuity.

**Frequency polygon:** class frequency plotted against class mid-point. distribution of continuous or grouped data representation of the shape and pattern of data.

### **Merits (Advantages) of Frequency Polygon:**

- **Smooth Visualization:** identify trends and patterns.
- **Easy Comparison:** Overlapping freq polygons can reveal similarities or differences data sets.
- **Identifying Central Tendency:** central peak or mode of the data distribution is clear in a frequency polygon, easier to identify most common values



### **Demerits (Disadvantages) of Frequency Polygon:**

- **Less Intuitive**: familiar with the concept of connecting midpoints with line segments, as compared to the simplicity of histograms or bar charts.
- **Data Loss**: Like histograms, frequency polygons summarize data within intervals, leading to a loss of precision.
- **Subjectivity in Line Drawing**: The visual appearance of a frequency polygon can vary based on the way the lines are drawn. Different interpretations can lead to different impressions of the data distribution.
- **Difficulty in Handling Discrete Data**: Frequency polygons are more suitable for continuous or grouped data. Handling discrete data in a frequency polygon may require additional considerations.

### **22. How to construct:**

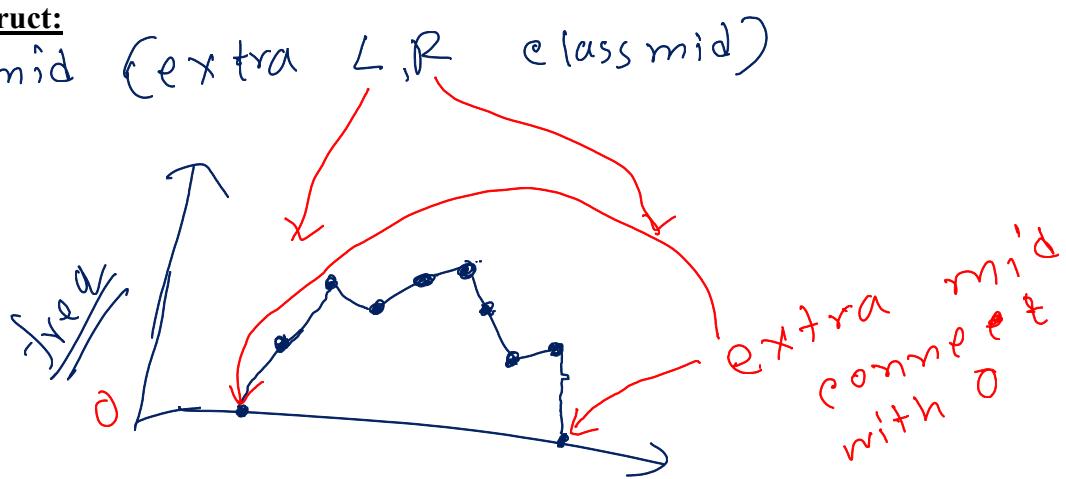
i. Class mid (extra L, R class mid)

ii.

iii.

iv.

v.



frequency polygon  
→ compare  
histogram → single

23. The provided data represents the ages of 40 individuals who have physical disabilities.  
 47, 54, 58, 63, 41, 68, 71, 37, 56, 66, 49, 60, 45, 67, 50, 38, 72, 55, 42, 59, 39, 65, 70, 43, 52, 73, 61, 44, 69, 57, 74, 62, 48, 53, 36, 64, 75, 40, 51, 46.

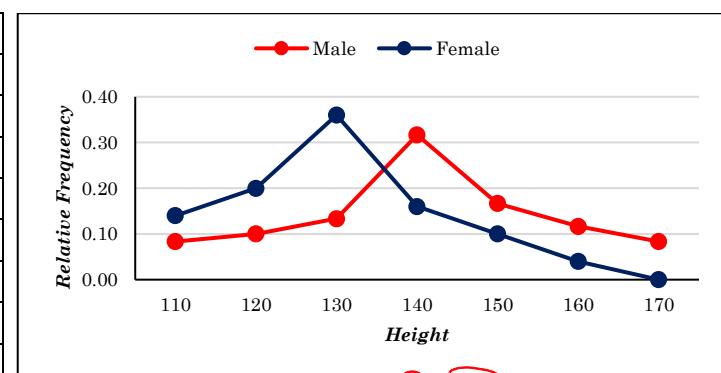
- Construct frequency table.
- Calculate relative frequency.
- Draw line graph using frequency.
- Draw line graph using relative frequency.

**\*\*\* The histogram and the frequency polygon are equally good techniques for presenting frequency distributions.**

**\*\*\* The histogram is more often used when single distribution is presented, while frequency polygon is largely used for comparison of two or more distributions.**

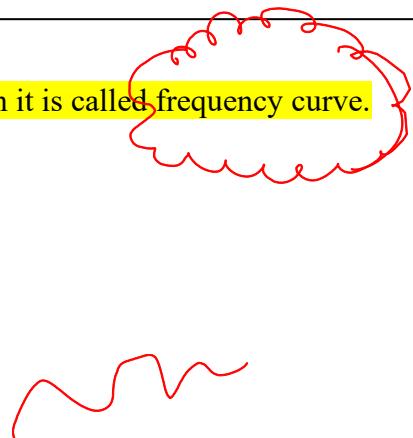
VV  
Compare M-F

Height (in cm)	Male	Female	RF Male	RF Female
105-115	5	7	0.08	0.14
115-125	6	10	0.10	0.2
125-135	8	18	0.13	0.36
135-145	19	8	0.32	0.16
145-155	10	5	0.17	0.1
155-165	7	2	0.12	0.04
165-175	5	0	0.08	0
Total	60	50	1	1



**\*\*\* If we connect the consecutive points with smooth curve, then it is called frequency curve.**

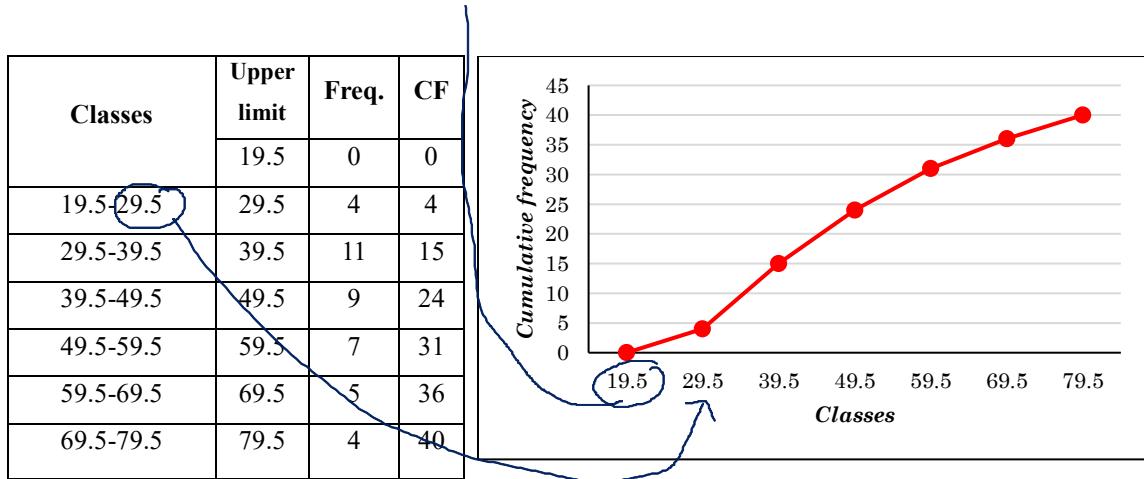
freq poly      freq curve



## **24. Cumulative frequency curve or Ogive:**

## **25. How to construct:**

- i. mark upper boundaries on X axis.
- ii. cumulative frequencies. Y axis.
- iii. point upper boundary height equal to the cumulative frequency
- iv. additional point above the lower boundary for first class at height of zero.



\*\*\* **Relative frequency:** The relative frequency for a particular category is the ratio of the frequency of that category to the total number of observations.

$$RF = \frac{\text{Frequency}}{\text{Total number of observations}}$$

\*\*\* **Percentage frequency:**  $\text{Relative frequency} \times 100$

**Overall solution based on question no 10:**

Here, number of observations = 60 ✓

Largest observation = 74.6, and smallest observation = 36.2

Approximate classes =  $\sqrt{60} = 7.75 \approx 8$

Approximate class width =  $\frac{(74.6 - 36.2)}{8} = 4.8 \approx 5$

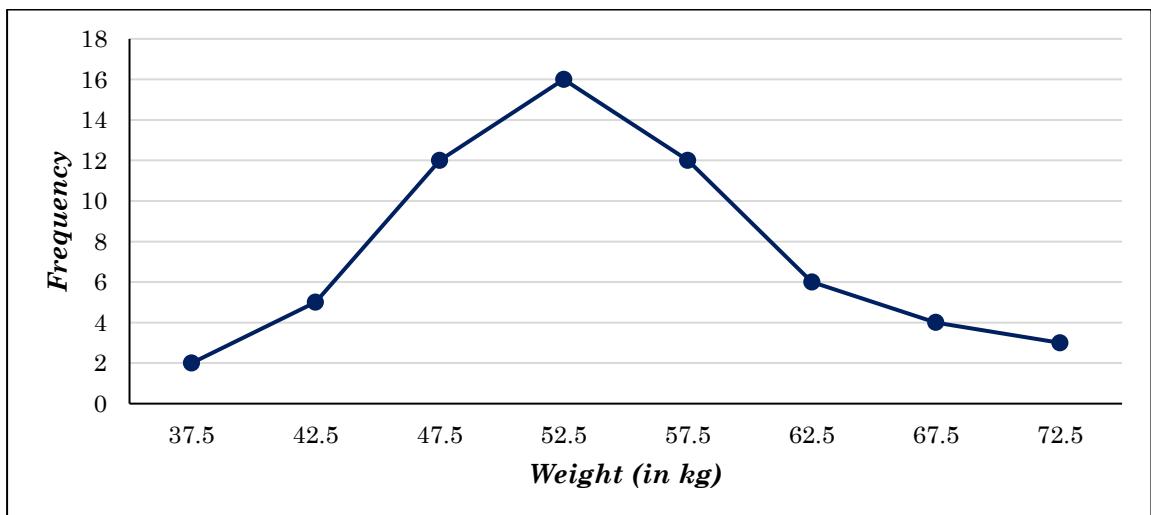
Then our classes will be,

$35 - 40, 40 - 45, 45 - 50, 50 - 55, 55 - 60, 60 - 65, 65 - 70, \text{ and } 70 - 75$

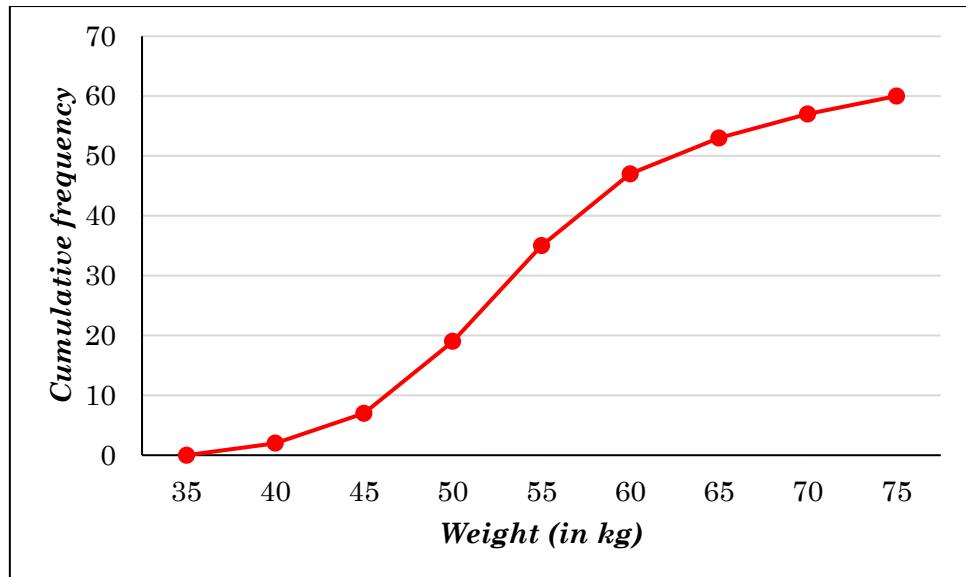
Class ✓ (weight in kg)	Mid value ✓	Tally ✓	Frequency ✓	Relative Frequency ✓	Cumulative frequency ✓
35-40	37.5		2	0.033	2
40-45	42.5		5	0.083	7
45-50	47.5		12	0.200	19
50-55	52.5		16	0.267	35
55-60	57.5		12	0.200	47
60-65	62.5		6	0.100	53
65-70	67.5		4	0.067	57
70-75	72.5		3	0.050	60
Total			n=60	1.00	

a) Histogram:

b) Frequency Polygon:



c) Ogive curve:



**Stem and Leaf Plot:**

**Example 1:**

39	37	43	27	59	30	34	41	35	49
78	58	48	64	34	54	37	70	65	36
69	54	53	21	24	57	49	33	42	57
62	41	75	71	37	40	60	47	29	39

Construct a stem and leaf plot.

Solution: The data ranged from 22 to 79. Because these observations are two-digit numbers, we use the first digit (tens) as the stem and the second digit (ones) as the leaf. We obtain the following stem and leaf plot:

Stem	Leaf
2	7 1 4 9
3	9 7 0 4 5 4 7 6 3 7 9
4	3 1 9 8 9 2 1 0 7
5	9 8 4 4 3 7 7
6	4 5 9 2 0
7	8 0 5 1

After arranging the leaves in order from smallest to largest, the stem and leaf plot finally given in below figure

Stem	Leaf
2	1 4 7 9
3	0 3 4 4 5 6 7 7 7 9 9
4	0 1 1 2 3 7 8 9 9
5	3 4 4 7 7 8 9
6	0 2 4 5 9
7	0 1 5 8

Key: 2|1 represents 21

Figure: Stem and leaf plot for ....

**Example 2:**

The following are the exam scores of 50 students

30.8	30.9	32.0	32.3	32.6	31.7	30.4	31.4	32.7	31.4
30.1	32.5	30.8	31.2	31.8	31.6	30.3	32.8	30.6	31.9
32.1	31.3	32.0	31.7	32.8	33.3	32.1	31.5	31.4	31.5
31.3	32.5	32.4	32.2	31.6	31.0	31.8	31.0	31.5	30.6
32.0	30.4	29.8	31.7	32.2	32.4	30.5	31.1	30.6	31.5

Construct a stem and leaf plot.

Solution: The scores in the data set range from 29.8 to 33.3. Because of these scores are three-digit numbers, we use the first two digit of each number as the stem and third digit as the leaf.

We obtain the following stem and leaf plot,

Stem	Leaf
29	8
30	8 9 4 1 8 3 6 6 4 5 6
31	7 4 4 2 8 6 9 3 7 5 4 3 6 0 8 0 5 7 1 5
32	0 3 6 7 5 8 1 0 8 1 5 4 2 0 2 4
33	3

After arranging the leaves in order from smallest to largest, the stem and leaf plot is displayed in below figure,

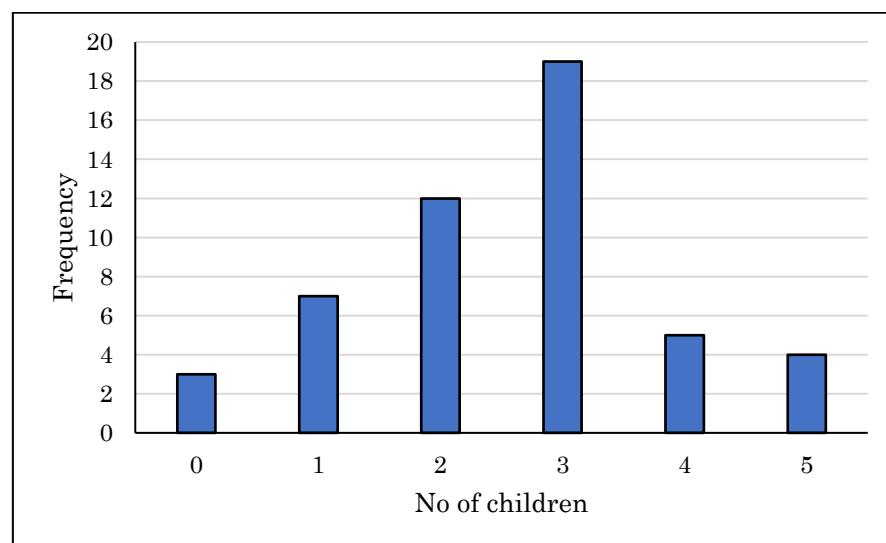
Stem	Leaf
29	8
30	1 3 4 4 5 6 6 6 8 8 9 ✓
31	0 0 1 2 3 3 4 4 5 5 5 6 6 7 7 7 8 8 9 ✓
32	0 0 0 1 1 2 2 3 4 4 5 5 6 7 8 8 ✓
33	3

Key: 29|8 represents 29.8

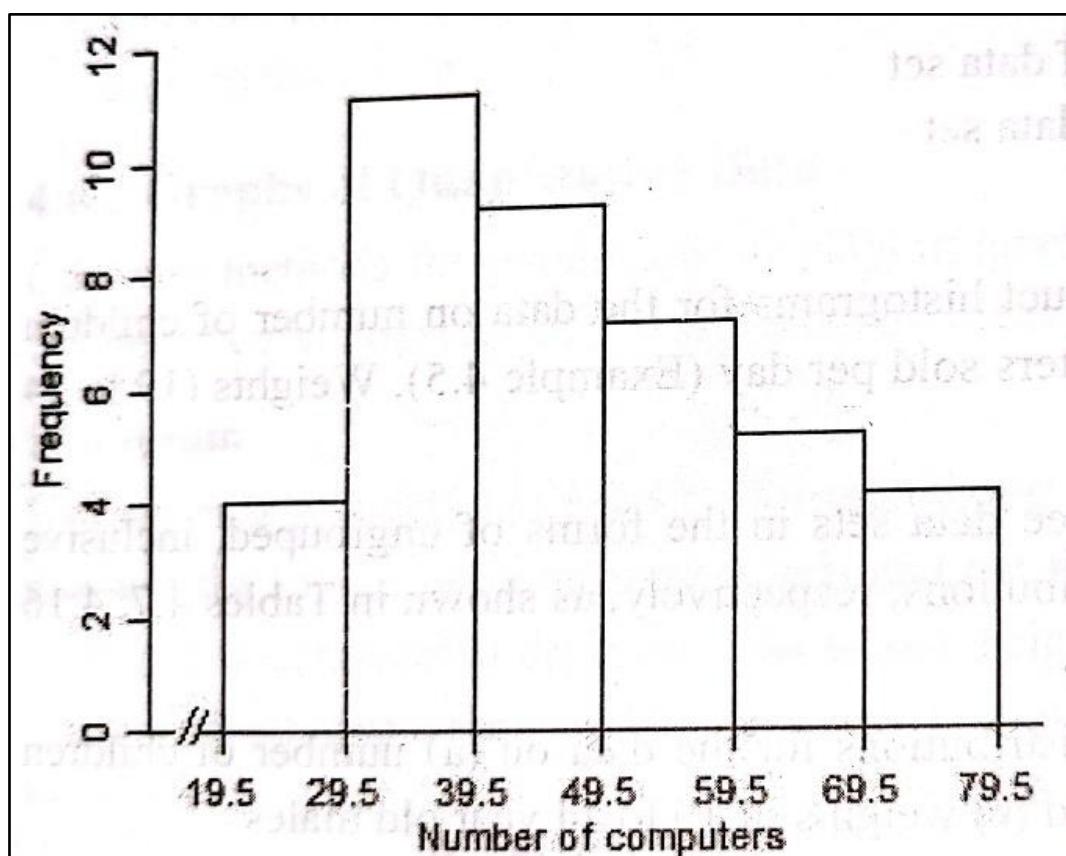
Weight (in KG)	Midpoint	Frequency	Relative frequency	Cumulative frequency	Relative Cumulative freq.
35-40	37.5	2	0.033		
40-45		5	0.083		
45-50	47.5	12			
50-55				35	
55-60		12			
60-65		6			
65-70		4			
70-75		3	0.050		

---

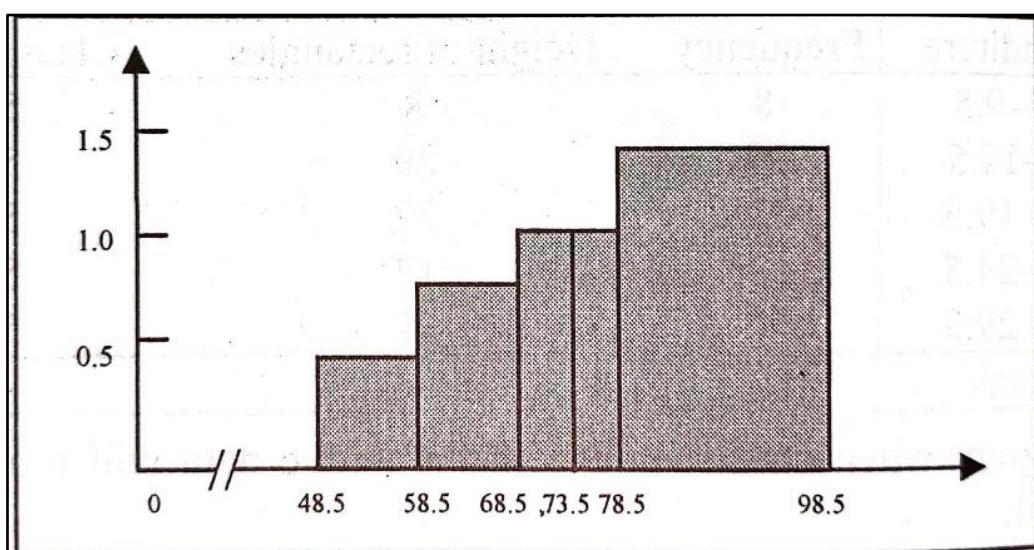
No of children	$f_i$
0	3
1	7
2	12
3	19
4	5
5	4



Computer sold	Class Boundary	$f_i$
20-29	19.5-29.5	4
30-39	29.5-39.5	11
40-49	39.5-49.5	9
50-59	49.5-59.5	7
60-69	59.5-69.5	5
70-79	69.5-79.5	4



Expenditure	$f_i$	Class width	Height of rectangles/ Frequency density
48.5-58.5	4	10	$\frac{4}{10} = 0.4$
58.5-68.5	8	10	0.8
68.5-73.5	5	5	1
73.5-78.5	5	5	1
78.5-98.5	28	20	1.4



Eye Condition

Gender	Near Sighted	Far Sighted	Need Bifocals
Male	12	40	12
Female	12	48	36

**Practice:**

- a) The quiz scores of 13 students from the CSE department of BracU are as follows:

17, 8, 12, 19, 14, 6, 10, 15, 7, 18, 11, 16, 9

- i) Construct the frequency distribution table using appropriate class interval.
- ii) Draw Histogram, Frequency polygon, and Ogive curve
- iii) Determined the Number of Students with Scores greater than or equal 13? (Hints: Ogive curve)

- b) Construct the frequency distribution for below data:

27, 14, 9, 22, 18, 6, 25, 12, 30, 8, 19, 5, 29, 11, 20

And, draw the histogram, frequency polygon and ogive curve.

- c) Utilizing the dataset from (b), create a frequency polygon representing relative frequencies.

**d) Why do we need to group data in the form of a frequency table? Explain.**

Ans: In summary, grouping data into a frequency table enhances data clarity, simplifies complex datasets, facilitates visualization, enables efficient analysis, and promotes effective communication among researchers and analysts.

- e) The following data give the results of a sample survey. The letters A, B, and C represent the three categories.

*A, B, B, A, C, B, C, C, C, A, C, B, C, A, C, C, B, C, C, A*

- i. Prepare a frequency distribution table.
- ii. Calculate the relative frequencies and percentages for all categories.
- iii. What percentage of the elements in this sample belong to category B?
- iv. What percentage of the elements in this sample belong to category A or C?
- v. Draw a bar graph for the frequency distribution.

**f) Difference between frequency polygon and histogram.**

Ans:

- 1) Several frequency distributions can be plotted on the same axis as frequency polygon. However, in the case of a histogram, we must have a separate graph for each distribution.
- 2) Frequency polygon is a continuous curve and it is easy to determine the scope and rate of change estimates. This is not possible in the case of a histogram.