Presentation and Description of Statistical Data

Statistic: in the last time the definition of statistic is means the collecting of data and arrange it by a tables or presented it by graphs

Statistic: in the present date is mean the scientific tools can be used in the several kinds of science such as engineering, management, agriculture, …… etc for collecting data and analysis these data by many several mathematical forms to make decision.

Method of presentation data

There are several method to present data there are as follows

1. Table method

It is concluded by putting the data in table form contain the information about data in each column.

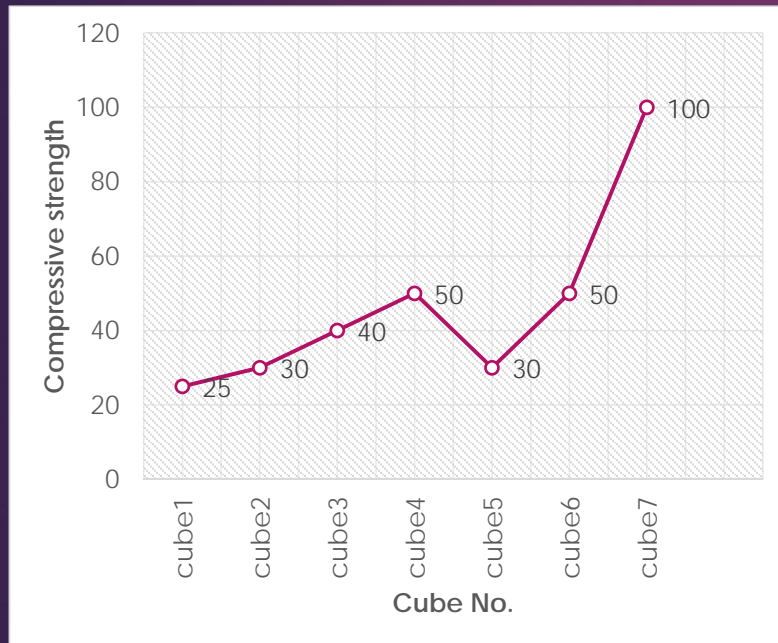| Concrete cube (No.)	Compressive strength (kPa.)	Young modulus (kPa.)
1	25	20000
2	30	25000
3	40	28000
4	50	30000
5	30	25000
6	50	30000
7	100	36000

2. Bar chart
   It is concluded by graph the data in a rectangular form.
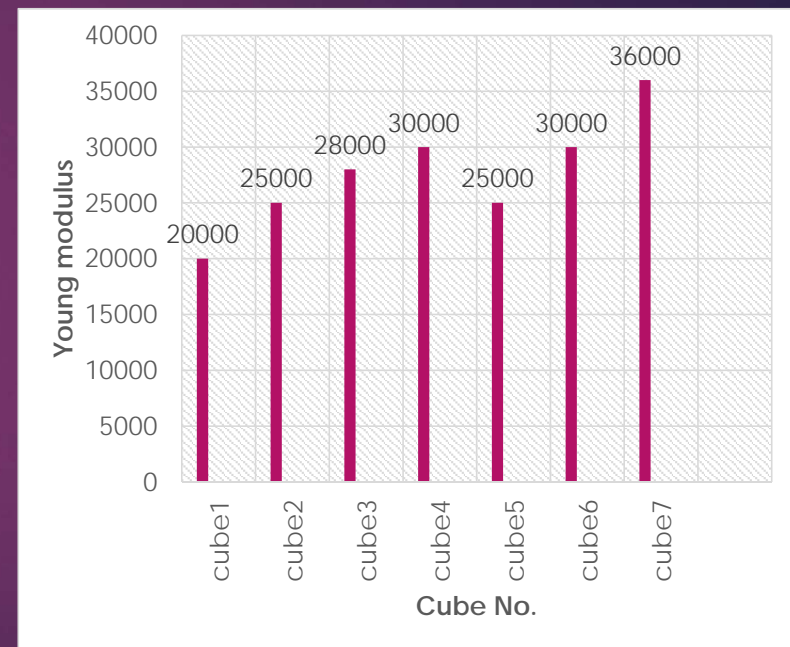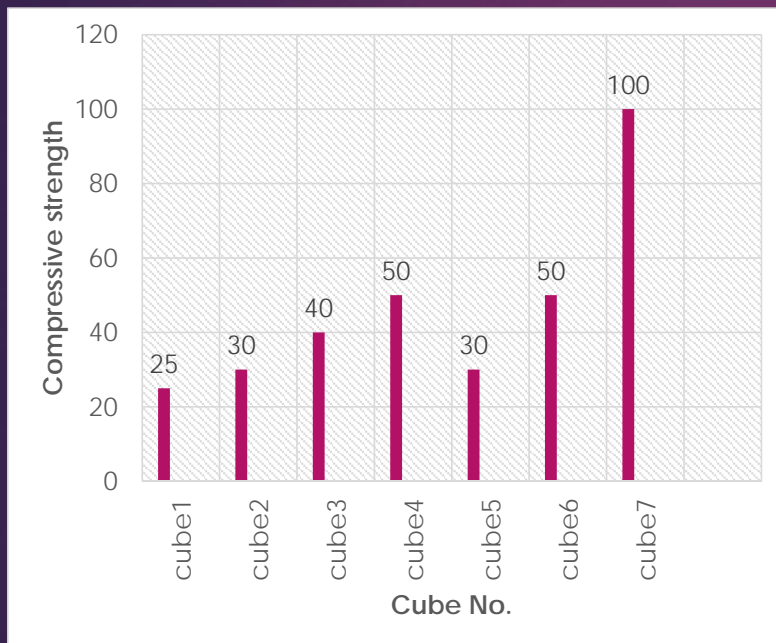
3. Broken line chart
It is concluded by graph the data in a broken line by connect each point to another point by straight line form.
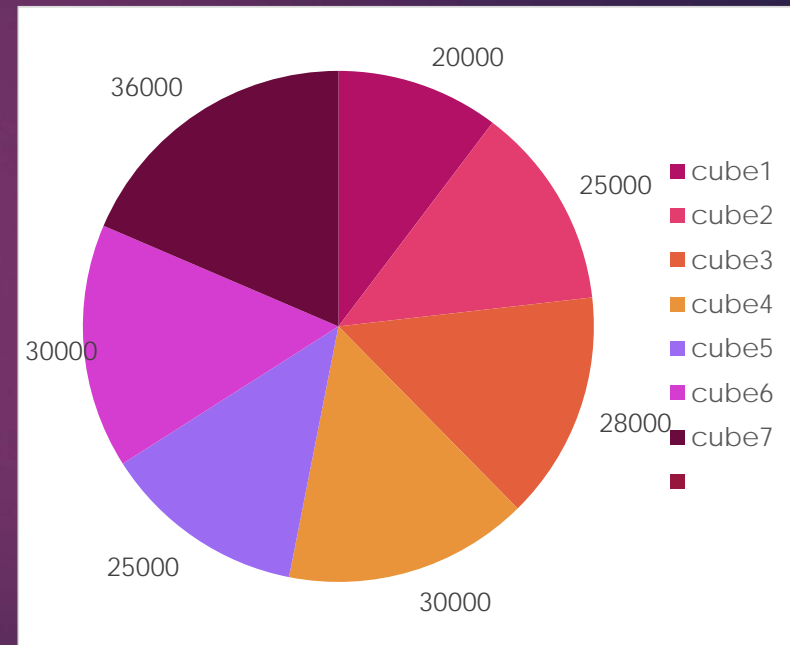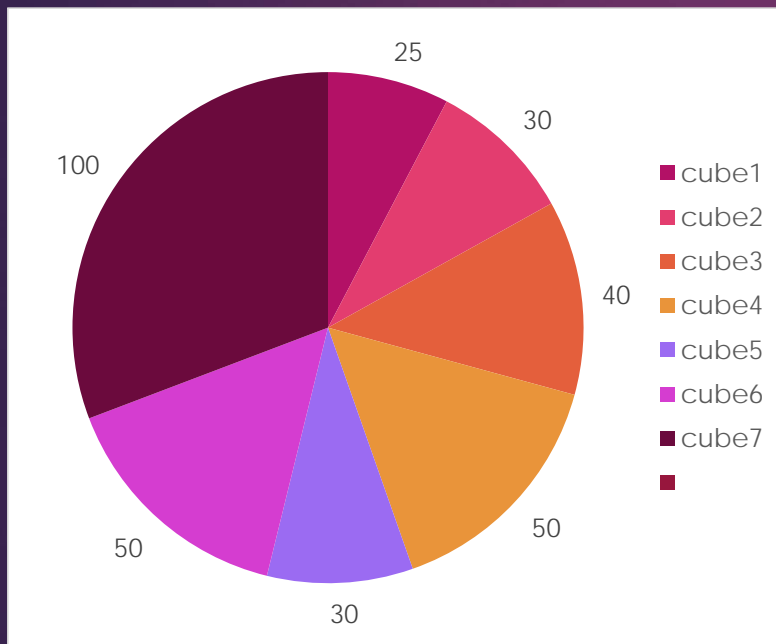
4. Curved line chart
It is concluded by graph the data in a broken line by connect each point to another point by straight line form.

5. Pie chart
It is concluded by graph the data in a circular form.

Frequency distribution:
When found a large number of data of an problem for such field the frequency distribution is the one of important form to contribution these data.

Example:
The following data are considered the values of columns load when tested several samples in structural laboratory construct a frequency table for these data.

100kN, 90kN, 120kN, 60kN, 140kN, 110kN, 80kN, 70kN, 110kN, 130kN, 140kN, 90kN, 70kN, 80kN, 140kN, 50kN, 148kN, 120kN, 130kN, 140kN, 80kN, 70kN, 130kN, 140kN, 60kN, 90kN, 70kN.

Solution:
There are several forms of frequency tables as follows.
1. Simple frequency table

(Column Load (kN), (x	(Frequency (f
50	1
60	2
70	4
80	3
90	3
100	1
110	2
120	2
130	3
140	5
148	1

2. Interval frequency table

in this stage we are must limiting the length of interval, assume interval length   a= 10,

Columns Load (X1 – X2) Interval limit	Frequency (fi)	Real interval limit
50 - 59	1	49.5 – 59.5
60 - 69	2	59.5 – 69.5
70 - 79	4	69.5 – 79.5
80 – 89	3	79.5 – 89.5
90 – 99	3	89.5 – 99.5
100 – 109	1	99.5 – 109.5
110 – 119	2	109.5 – 119.5
120 – 129	2	119.5 – 129.5
130 – 139	3	129.5 – 139.5
140 – 149	6	139.5 – 149.5

b=20

Columns Load (X1 – X2) Interval limit	Frequency (fi)	Real interval limit (X1 – X2)
50 - 69	3	49.5 – 69.5
70 - 89	7	69.5 – 89.5
90 - 109	4	89.5 – 109.5
110 – 129	4	109.5 – 129.5
130 – 149	9	129.5 – 149.5

2. Interval frequency table continue

in this stage we are must limiting the length of interval, assume interval length   a= 10,

Columns Load (X1 – X2)	Frequency (fi)	Interval center (Xi)
50 - 59	1	54.5
60 - 69	2	64.5
70 - 79	4	74.5
80 – 89	3	84.5
90 – 99	3	94.5
100 – 109	1	104.5
110 – 119	2	114.5
120 – 129	2	124.5
130 – 139	3	134.5
140 – 149	6	144.5

b=20

Columns Load (X1 – X2)	Frequency (fi)	Interval center (Xi)
50 - 69	3	59.5
70 - 89	7	79.5
90 - 109	4	99.5
110 – 129	4	119.5
130 – 149	9	139.5

$$Xi = \frac{X1 + X2}{2}$$

3. Relative frequency distribution

Is the same of frequency distribution but we are fixed the frequency as a rate of total frequency. In the last example, were interval length 20 the relative frequency table becomes, as follows.

Columns Load (X1 – X2)	Interval center (Xi)	Frequency (fi)	Relative frequency (fr)	Percentage frequency (fp)
50 - 69	59.5	3	0.11111	11.111
70 - 89	79.5	7	0.25926	25.926
90 - 109	99.5	4	0.14815	14.815
110 – 129	119.5	4	0.14815	14.815
130 – 149	139.5	9	0.33333	33.333
		ft = 27		

$$ft = \sum_{1}^{n} fi$$

$$fr = \frac{fi}{ft}$$

$$fp = \frac{fi}{ft}\%$$

4. Cumulative frequency distribution

Is the same of frequency distribution but we are fixed the frequency as a cumulative arrangement. In the last example, were interval length 20 the cumulative frequency table becomes, as follows.

Columns Load (X1 – X2)	Real interval limit (X1 – X2)	Interval center (Xi)	Frequency (fi)	Cumulative frequency
50 - 69	49.5 – 69.5	59.5	3	3
70 - 89	69.5 – 89.5	79.5	7	10
90 - 109	89.5 – 109.5	99.5	4	14
110 – 129	109.5 – 129.5	119.5	4	18
130 – 149	129.5 – 149.5	139.5	9	27

$$f_{ic} = fi + fic_{\_1}$$

5. Frequency graphs

It is a new form to display the frequency distribution by graphs method, and divided to the following:

a) Frequency histogram.
b) Frequency polygon.
c) Frequency curve.
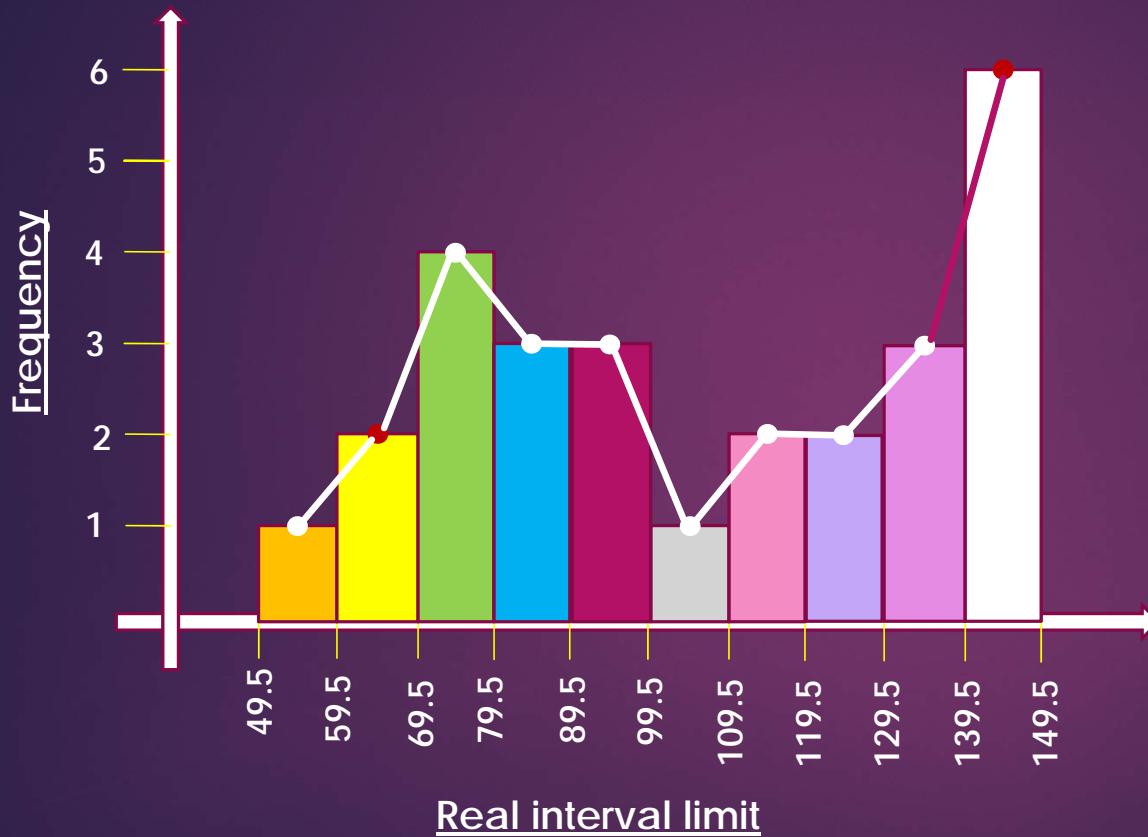d) Cumulative frequency polygon.

Example:

For the following date in the beside table draw the frequency histogram, frequency polygon, frequency curve, and cumulative frequency polygon.

Columns Load (X1 – X2)	Real interval limit	Frequency (fi)
50 - 59	49.5 – 59.5	1
60 - 69	59.5 – 69.5	2
70 - 79	69.5 – 79.5	4
80 – 89	79.5 – 89.5	3
90 – 99	89.5 – 99.5	3
100 – 109	99.5 – 109.5	1
110 – 119	109.5 – 119.5	2
120 – 129	119.5 – 129.5	2
130 – 139	129.5 – 139.5	3
140 – 149	139.5 – 149.5	6
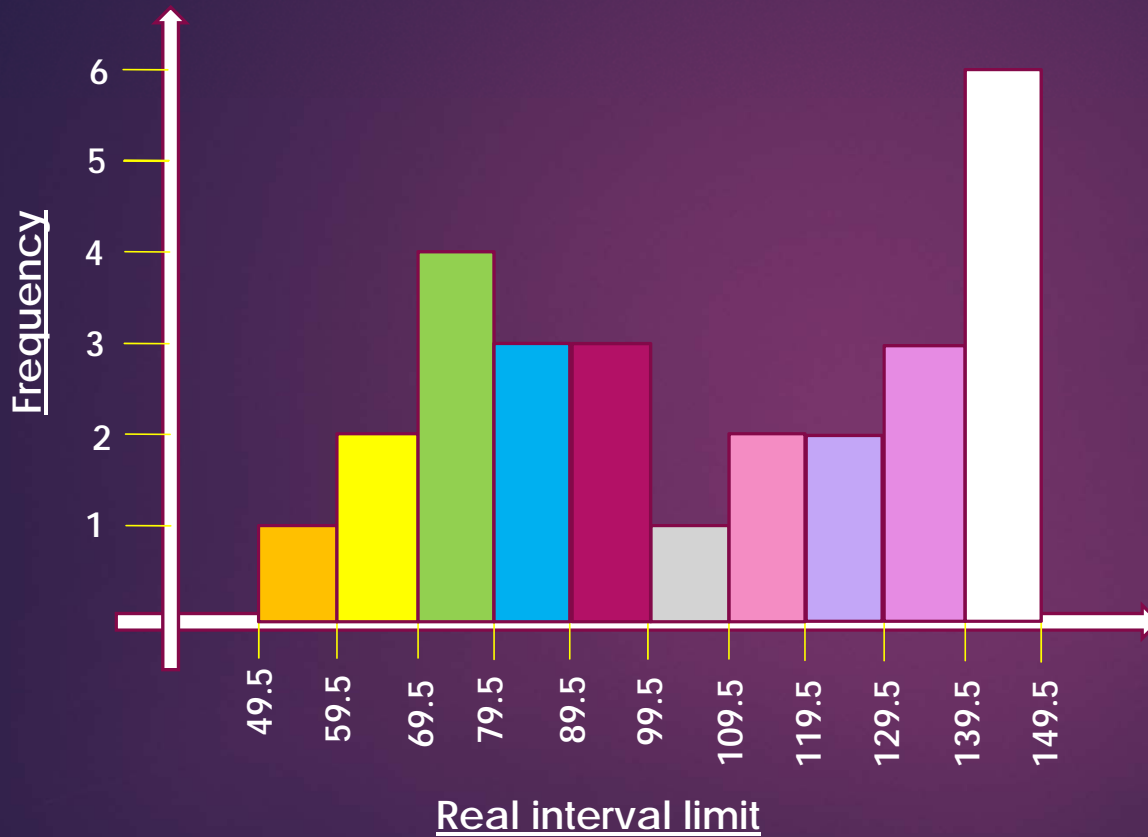
Real interval limit	Frequency (fi)
49.5 – 59.5	1
59.5 – 69.5	2
69.5 – 79.5	4
79.5 – 89.5	3
89.5 – 99.5	3
99.5 – 109.5	1
109.5 – 119.5	2
119.5 – 129.5	2
129.5 – 139.5	3
139.5 – 149.5	6

Frequency histogram

Real interval limit	Frequency (fi)	Interval center (Xi)
49.5 – 59.5	1	54.5
59.5 – 69.5	2	64.5
69.5 – 79.5	4	74.5
79.5 – 89.5	3	84.5
89.5 – 99.5	3	94.5
99.5 – 109.5	1	104.5
109.5 – 119.5	2	114.5
119.5 – 129.5	2	124.5
129.5 – 139.5	3	134.5
139.5 – 149.5	6	144.5

Frequency polygon
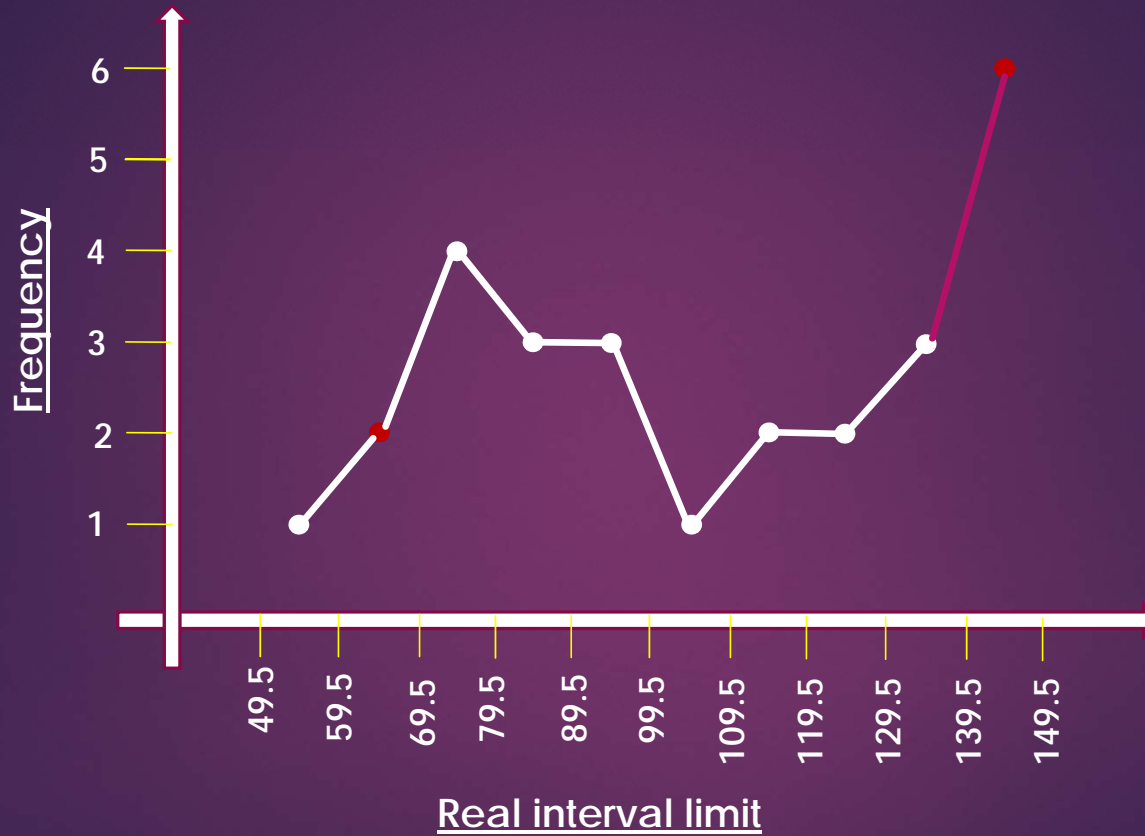
Frequency polygon

| Real interval limit	Frequency (fi)	Interval center (Xi)
49.5 – 59.5	1	54.5
59.5 – 69.5	2	64.5
69.5 – 79.5	4	74.5
79.5 – 89.5	3	84.5
89.5 – 99.5	3	94.5
99.5 – 109.5	1	104.5
109.5 – 119.5	2	114.5
119.5 – 129.5	2	124.5
129.5 – 139.5	3	134.5
139.5 – 149.5	6	144.5

Frequency curve

Frequency curve

Cumulative frequency polygon

Columns Load (X1 – X2) Interval limit	Real interval limit	Frequency (fi)	Interval center (Xi)	Cumulative frequency (fi)
50 - 59	49.5 – 59.5	1	54.5	1
60 - 69	59.5 – 69.5	2	64.5	3
70 - 79	69.5 – 79.5	4	74.5	7
80 – 89	79.5 – 89.5	3	84.5	10
90 – 99	89.5 – 99.5	3	94.5	13
100 – 109	99.5 – 109.5	1	104.5	14
110 – 119	109.5 – 119.5	2	114.5	16
120 – 129	119.5 – 129.5	2	124.5	18
130 – 139	129.5 – 139.5	3	134.5	21
140 – 149	139.5 – 149.5	6	144.5	27
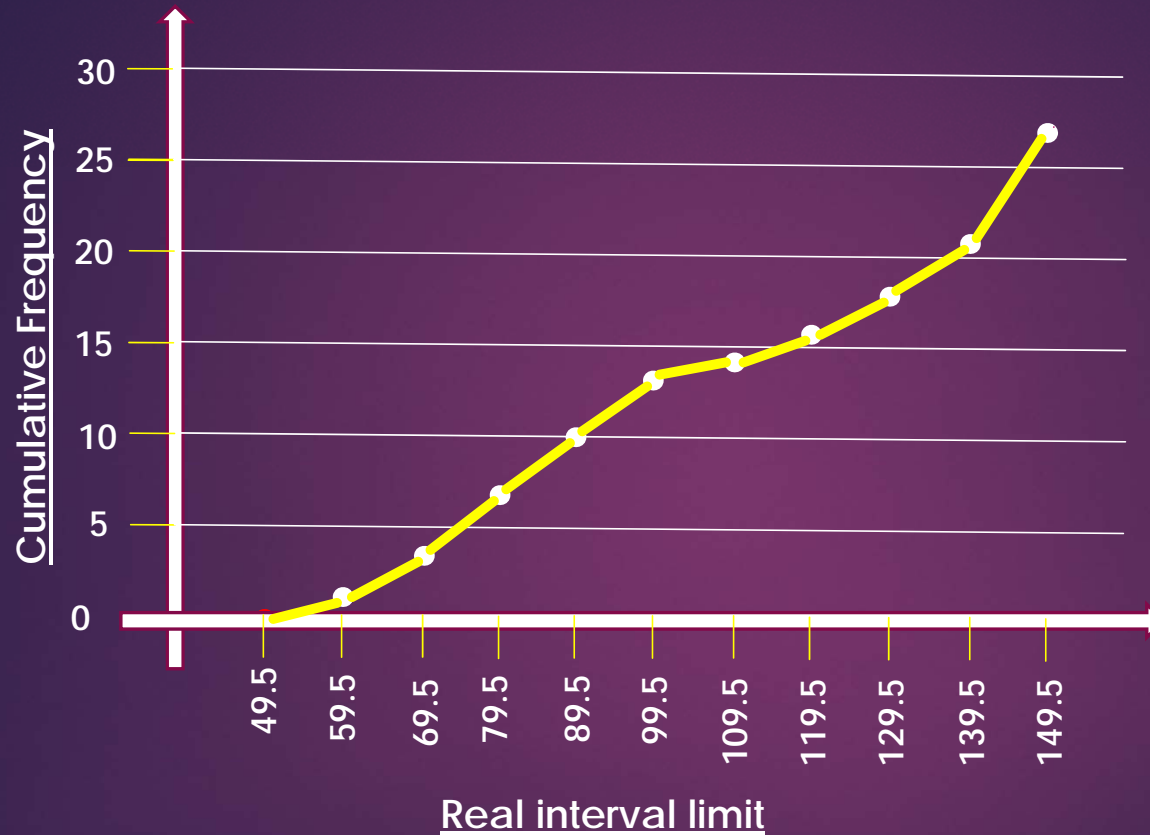
Cumulative frequency polygon

This graph is concluded by plotting the points, between cumulative frequency and the upper limit of each interval, but the first point only is located between the lower limit of interval and zero frequency.

Stem and Leaf Display

If we have the following data and we need to review in the form of stem and leaf method.
20, 42, 22, 24, 26, 38, 39, 40, 41, 52, 53, 54, 55, 56, 48, 49, 26, 27, 28, 29, 30, 31, 32, 36, 37, 44, 45, 36

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

Example:

The ages of 30 concrete cubes admitted to a material laboratory during a particular week were as follows:

48, 31, 54, 37, 18, 64, 61, 43,
40, 71, 51, 12, 52, 65, 53, 42,
39, 62, 74, 48, 29, 67, 30, 49,
68, 35, 57, 26, 27, 58.

Construct a stem-and-leaf display from the data and list the data in an array

Solution:

A scan of the data indicates that the observations range (in age) from 12 to 74. We use the first (or leading) digit as the stem and the second (or trailing) digit as the leaf. The first observation is 48, which has a stem of 4 and a leaf of 8, the second a stem of 3 and a leaf of 1, etc. Placing the leaves in the order in which they APPEAR in the data, we get the stem-and-leaf display as shown below:

Stem (Leading Digit)	Leaf (Trailing Digit)
1	8 2
2	9 6 7
3	1 7 9 0 5
4	8 3 0 2 8 9
5	4 1 2 3 7 8
6	4 1 5 2 7 8
7	1 4

But it is a common practice to ARRANGE the trailing digits in each row from smallest to highest. In this example, in order to obtain an array, we associate the leaves in order of size with the stems as shown below:

in ascending order
12, 18, 26, 27, 29, 30, 31, 35, 37, 39, 40, 42, 43, 48, 48, 49, 51, 52, 53, 54, 57, 58, 61, 62, 64, 65, 67, 68, 71, 74.

STEM AND LEAF DISPLAY

The stem-and-leaf table provides a useful description of the data set and, if we so desire, can easily be converted to a frequency table.

Stem (Leading Digit)	Leaf (Trailing Digit)
1	2 8
2	6 7 9
3	0 1 5 7 9
4	0 2 3 8 8 9
5	1 2 3 4 7 8
6	1 2 4 5 7 8
7	1 4

The frequency table of data in this example becomes as follows when interval length is assumed 10.

Interval limit	Real Interval limit	Frequency
10 – 19	9.5 – 19.5	2
20 – 29	19.5 – 29.5	3
30 – 39	29.5 – 39.5	5
40 – 49	39.5 – 49.5	6
50 – 59	49.5 – 59.5	6
60 – 69	59.5 – 69.5	6
70 - 79	69.5 – 79.5	2

Stem (Leading Digit)	Leaf (Trailing Digit)
1	2 8
2	6 7 9
3	0 1 5 7 9
4	0 2 3 8 8 9
5	1 2 3 4 7 8
6	1 2 4 5 7 8
7	1 4

VARIOUS TYPES OF FREQUENCY CURVES

- the symmetrical frequency curve

- the moderately skewed frequency curve

- the extremely skewed frequency curve

- the U-shaped frequency curve

First of all, the symmetrical frequency curve is of the following shape:

THE SYMMETRIC CURVE
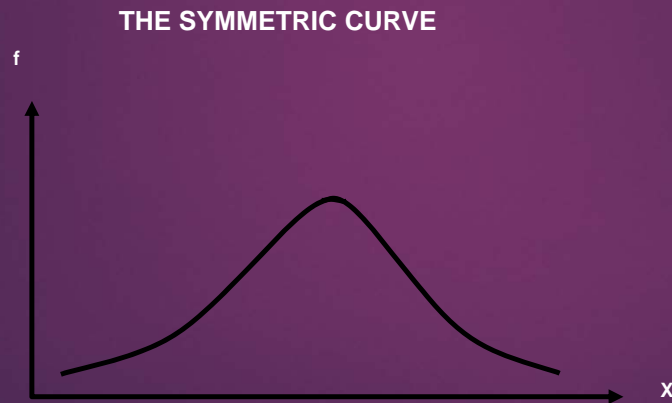
If we place a vertical mirror in the center of this graph, the left hand side will be the mirror image of the right hand side.
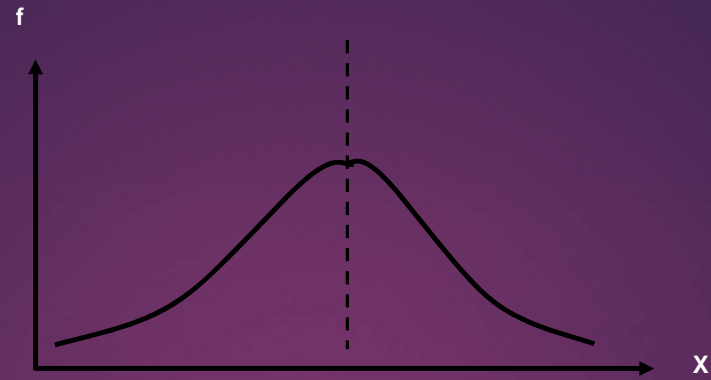
Next, we consider the moderately skewed frequency curve. We have the positively skewed curve and the negatively skewed curve. The positively skewed curve is that one whose right tail is longer than its left tail, as shown below:
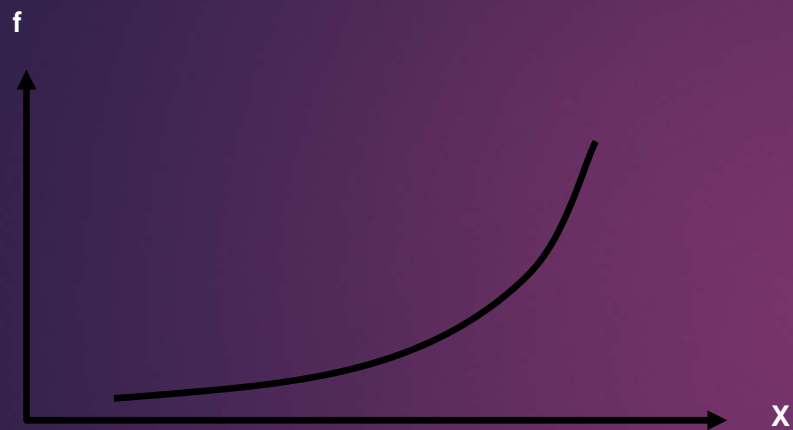
THE POSITIVELY SKEWED CURVE
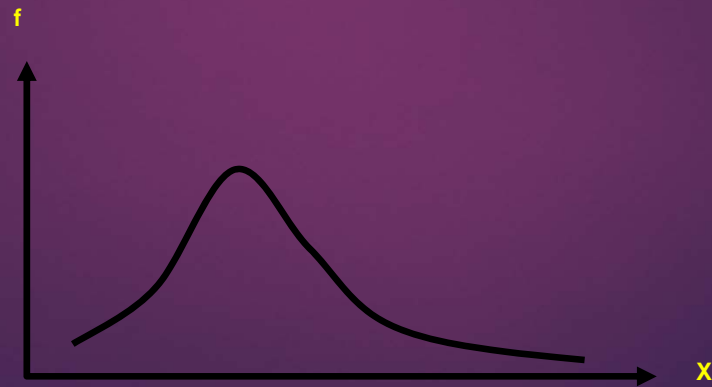
On the other hand, the negatively skewed frequency curve is the one for which the left tail is longer than the right tail.
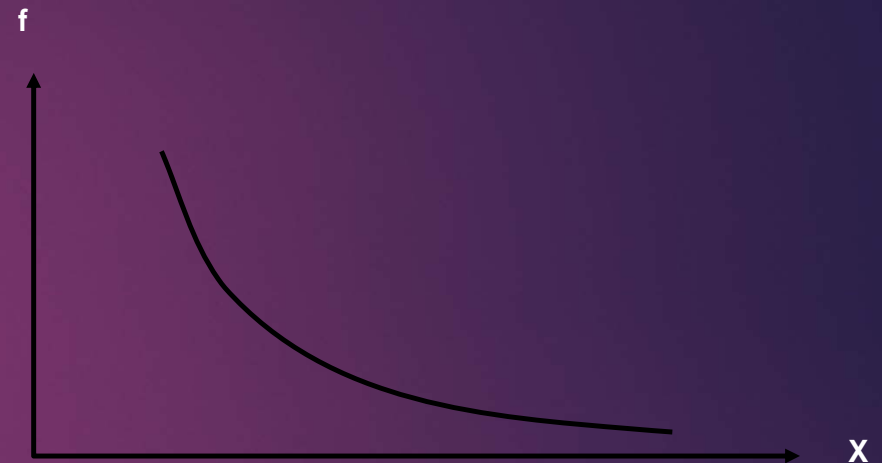
THE NEGATIVELY SKEWED CURVE

Both of these that we have just consider are moderately positively and negatively skewed. Sometimes, we have the extreme case when we obtain the EXTREMELY skewed frequency curve. An extremely negatively skewed curve is of the type shown below:

THE <u>EXTREMELY</u> NEGATIVELY SKEWED (J-SHAPED) CURVE

THE <u>EXTREMELY</u> POSITIVELY SKEWED (REVERSE J-SHAPED) CURVE

f

X

f

X

This is the case when the maximum frequency occurs at the end of the frequency table.
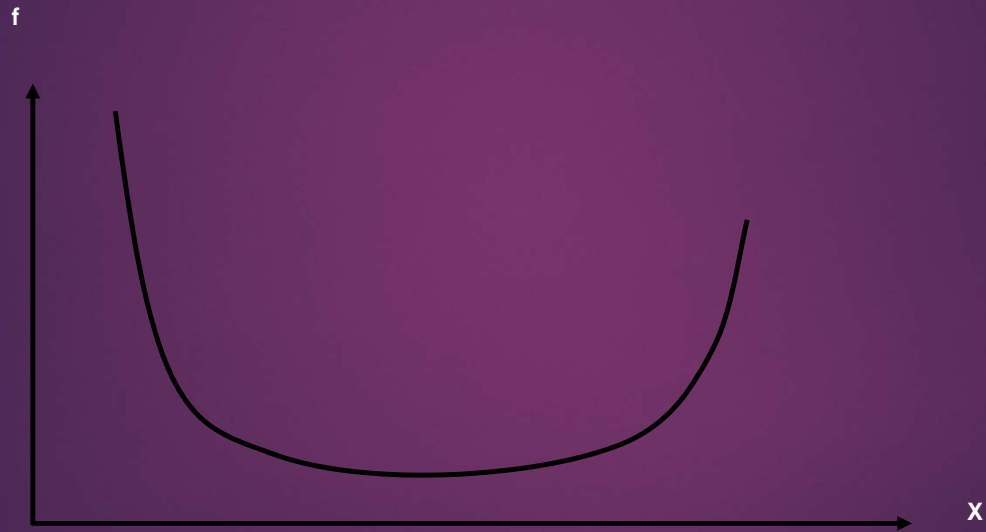
Similarly, the extremely positively skewed distribution is known as the REVERSE J-shaped distribution.

A relatively LESS frequently encountered frequency distribution is the U-shaped distribution
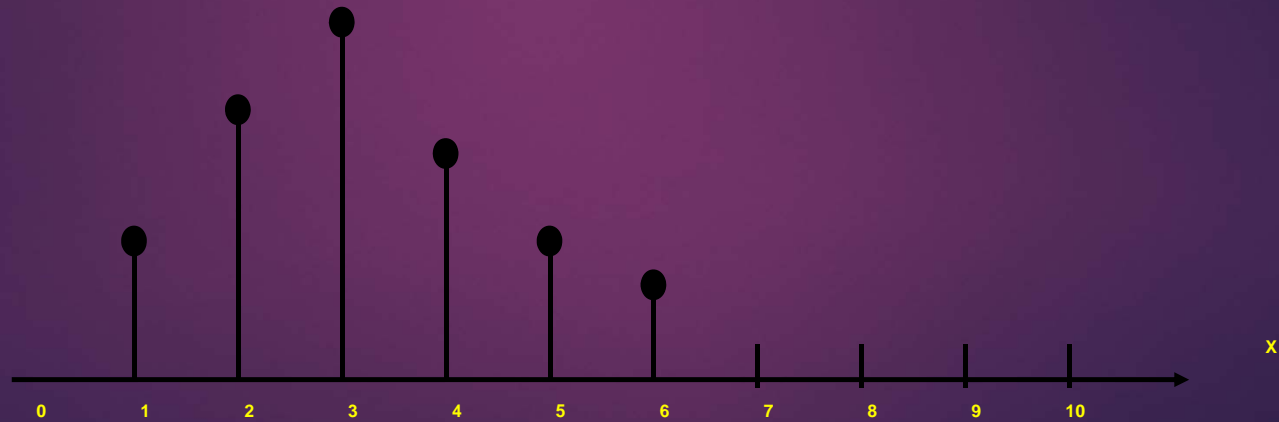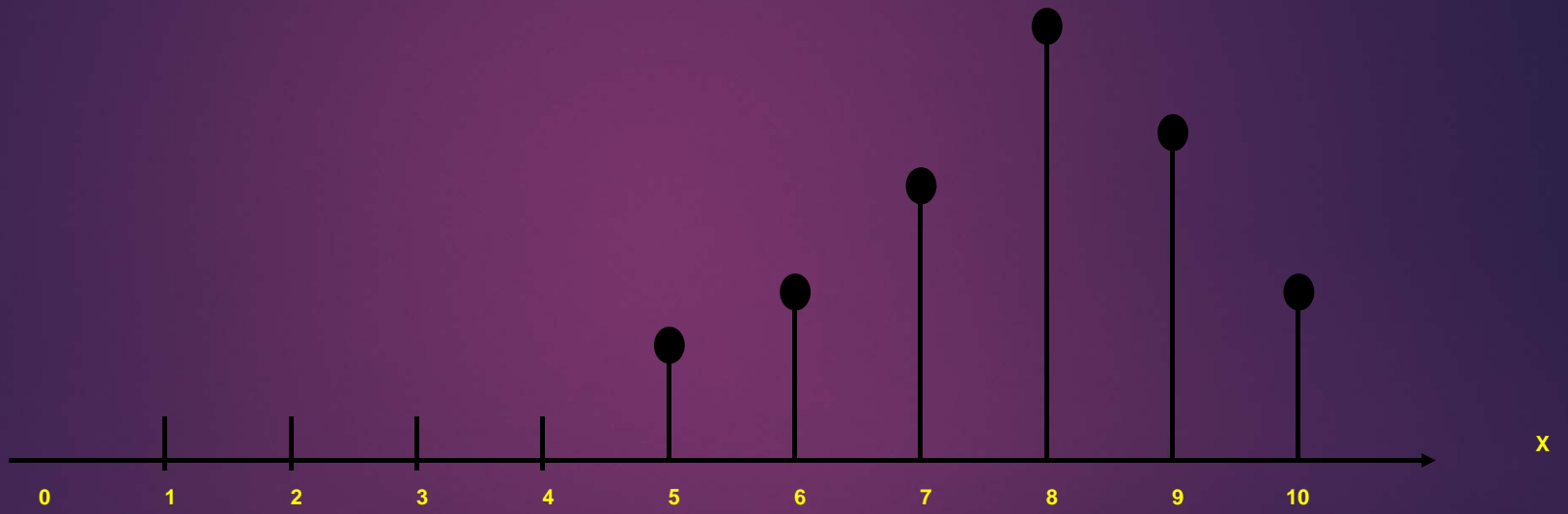
THE U-SHAPED CURVE

Similar shapes are possible for the frequency distribution of a discrete variable.

VARIOUS TYPES OF DISCRETE FREQUENCY DISTRIBUTION

I. **Positively Skewed Distribution**

II. Negatively Skewed Distribution

III. Symmetric Distribution