STATISTICS

- **Central Tendency** : A measure of central tendency is a value around which all the observations have a tendency to from a cluster.

  i. Arithmetic Mean (A M)
  ii. Median
  iii. Mode
  iv. Geometric Mean (GM)
  v. Harmonic Mean (HM)

i. Arithmetic Mean ($\bar{x}$) : $\sum_{i=1}^{n} \frac{x_i}{n}$, where n is number of observations.

$$\bar{x} = \frac{\sum_{i=1}^{k} f_i x_i}{\sum_{i=1}^{k} f_i}$$

$$\boxed{\text{Weighted Mean } (\bar{x}_w) = \frac{\sum w_i x_i}{\sum w_i}}$$

* NOTE :- i. Median is better indicator than mean if :

a. % of outliers from total sample is higher than expected.

b. Distance b/w mean and outliers is greater than expected.

c. Distribution of values is still Gaussian (normal distribution) but skewed to one side.

Question : calculate the mean, median, variance and standard deviation

class	0 - 10	10 - 20	20 - 30	30 - 40	40 - 50	50 - 60
freq.	27	10	7	5	4	2

- Measures of Dispersion

Lets consider the given data :

(i) 5, 5, 5, 5, 5 (mean = 5, Median = 5)

(ii) 3, 4, 5, 6, 7 (mean = 5, median = 5)

(iii) 1, 3, 5, 7, 9 (Mean = 5, Median = 5)

1. Median = $\left(\frac{n+1}{2}\right)^{th}$, when n is odd

$$= \left[\left(\frac{n}{2}\right)^{th} \text{observation} + \left(\frac{n}{2}+1\right)^{th} \text{observation}\right] \times \frac{1}{2}$$

when n is even

* Median is not helpful when we consider Relative importance or weight assignment.

2. Geometric Mean - It is defined as n^{th} root of the product of value of n items in a given list.

$$G.M. = \sqrt[n]{\Pi x_i} = \sqrt[n]{x_1 \cdot x_2 \cdots x_n}$$

$$= (x_i)^{1/n}$$

Example 2 : Investment which earns 10% first year, 50% second year, 30% in third year. what is average rate of return.

$$\text{Geometric Mean} = (1.10 \times 1.50 \times 1.30)^{1/3}$$

$$= 1.283$$

~~average rate of return~~

Let the Money be x ,

$$\cancel{x} \times 1.10 \times 1.50 \times 1.30 = \cancel{x} \times f \times f \times f$$

$$1.10 \times 1.50 \times 1.30 = f^3$$

$$f = (1.10 \times 1.50 \times 1.30)^{1/3}$$

* We use Geometric Mean when we have to calculate AVERAGE PERCENT OF CHANGE.

3. Mean Deviation : It is an average of ~~frequencies~~ diffrences of the values of items from some average of series.

$$(\delta_x) \text{ Mean Deviation} = \frac{\Sigma |x_i - \bar{x}|}{N}$$

$$\text{Coefficient of mean deviation} = \frac{\delta_x}{\bar{x}}$$

* It is less sensitive to extreme Scopes, so it may not reliable in outline the source.

4. Standard Deviation :

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

$$\text{coeff. of standard deviation} = \frac{\sigma}{\bar{x}}$$

Standard Deviation for sampled data $= \sqrt{\frac{1}{(n-1)} \sum_{i=1}^{n} (x_i - \bar{x})}$

$\bar{x}$ is sample mean

x is No. of items in sample

5. Variance :

$$\sigma^2 = \frac{\sum_{i=1}^{N} (x_i - \bar{x})^2}{N}$$ for population (complete dataset).

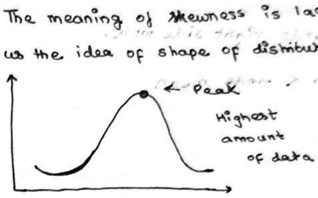
$$= \frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i}$$

for sampled data, (Incomplete dataset)

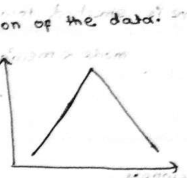
$$\text{variance } (s^2) = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

- Measures of Skweness

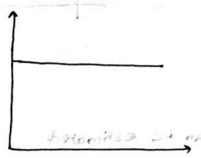
- The meaning of skewness is lack of symmetry, skewness gives us the idea of shape of distribution of the data.



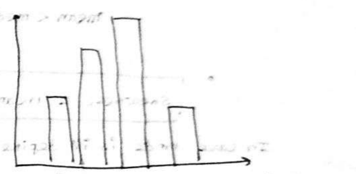
(i)



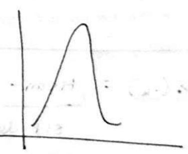
(ii)



(iii)



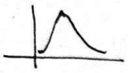
- Skewed distribution : a dataset has an skewed distribution when mean, median, mode (if ~~exits~~ exists) are not the same. In such case the plot of distribution is stressed to one side than to the other side.



(i) +ve skewness

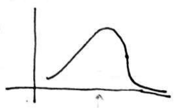
curve is streched towards right side more.

mode < median < ~~mode~~ mean



(ii) -ve skewness

curve is streched towards left side more.



mean < median < mode

- $\boxed{\text{Skewness} = \text{Mean} - \text{Mode}}$

In case mode is ill defined then it can be estimated from mean and median.

$\boxed{\text{Mode} = 3\,\text{Median} - 2\,\text{Mean}}$

- COMPARISION OF SKEWNESS OF TWO DATASETS

we can compare the datasets using coefficient of skewness.

$$\boxed{\text{coefficient of skewness } (S_k) = \frac{\text{Mean} - \text{Mode}}{\text{std deviation}}}$$

- Univariate, Bivariate, Multivariate population

In Bivariate population two variables are represented by X and Y, and paired observations are (x_1, y_1), (x_2, y_2) (x_3, y_3), ... (x_m, y_m)

covariance :-

$$\boxed{\text{cov}(x, y) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}$$

- Population (Universe) : collection of all items about which the information is desired is called as population. The Population or universe can be finite or infinite.

- Parameter : Any characteristics of measure of population units is known as parameter. Population mean, population standard deviation are commonly studied parameters.

μ - mean

σ - std deviation

P, π - population proportion

- **STATISTICS** :- Any characteristic of any sample of item is known as statistics.

  s : sample standard deviation

  x : sample mean

  p : sample population

- **SAMPLING AND NON SAMPLING ERRORS**

* **Sampling error** arises on account of sampling and it generally happens to be random variations in the sample estimates around the true population values.

  Sampling error is inversly related to size of sample.

* **Non sampling error** may creep in during the process of collecting the actual information.

- **Stem and Leaf Plot**

A stem and leaf plot also known as stem and leaf diagram is a way to arrange and represent data so that it is simple to see how frequenty various data value occur.

- Digits of data value are divided into stem (first few digits) and a ~~last~~ leaf (usually last digit).

- Symbol "|" is used to split and illustrate steam and leaf values.

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

Key : 2 | 0 = 20

- **CONSTRUCTION**

Step 1 : classify data values in terms of number of digits.

Step 2 : fix the key for stem and leaf plot.

Example = 2 | 5 = 25, 3 | 2 = 3.2 etc...

Step 3 : Consider 1^{st} digits as stems and last digit as leafs.

Step 4 : Sort the data from highest to lowest.

Step 5 : place stem on left side and leaf on right side.

Step 6 : List leaf values in col. against the stem from lowest to highest horizontally.

Example 1 : A table shows the duration of calls that ROSY makes each day. Represent using stem and leaf plot.

Data	9	9	9	10	10	10	10	11	11	11	11	11
Minute	56	3	6	14	19	5	23	36	30	23	10	2

Step 1 : sort the data (minutes)

2, 3, 5, 6, 10, 14, 19, 23, 23, 30, 36, 56

Step 2 : Let assume,

Tens digits for $\boxed{\text{stem}}$,

ones digit for $\boxed{\text{leaf}}$,

Step 3 : key Defination,

3 | 6 = 36 min, 3 | 6 = 31 min

Step 4 : phone call lengths

Stem	Leaf
0	2, 3, 5, 6
1	0, 4, 9
2	3, 3,
3	0, 6
5	6

Example 2 : The stem and leaf plot below shows the quiz score of students, (i) find no. of students who scores less than 9 points.

(ii) find no. of students who scored a minimum of 9.

Quiz Score

Stem	Leaf
6	6
7	0 5 7 8
8	1 1 3 4 4 6 8 8 9
9	0 2 9
10	0

$\boxed{\text{key } 9 \mid 2 = 9.2 \text{ points.}}$

Solⁿ : (i) No. of students who scored less than 9 = 14 students

6.6, 7.0, 7.5, 7.7, 7.8, 8.1, 8.1, 8.3, 8.4, 8.4, 8.6, 8.8, 8.8, 8.9

(ii) No. of students with a minimum 9 score = 4 students.

9.0, 9.2, 9.9, 10.0

• HISTOGRAM

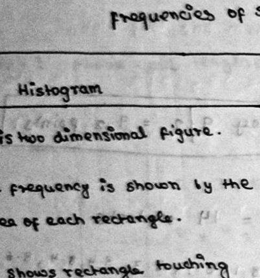
A Histogram is a graphical represention of a grouped frequency distribution with continuose classes.

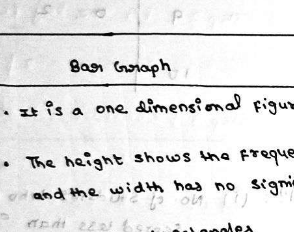
It is an area diagram and can be defined as a set of rectangles with bases along with the intervals class boundaries and with area propotional to frequencies in the corresponding classes.

(i) All the rectangles are adjacent.

(ii) Heights of rectangles are propotional to corresponding frequencies of similar class.

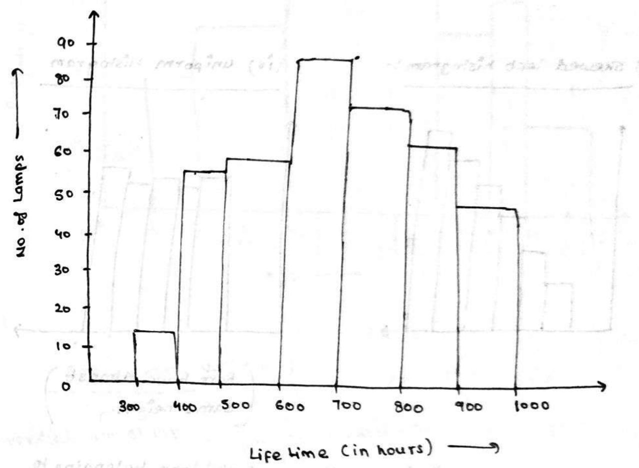
Histogram	Bar Graph
• It is two dimensional figure.	• It is a one dimensional figure.
• The frequency is shown by the area of each rectangle.	• The height shows the frequency and the width has no significance.
• It shows rectangle touching each other. • NO GAPS b/w rectangles	• It consists of rectangles seperated from each other with equal spaces.





Ques: Following table gives life time of 400 lamps. Draw histogram of below data.

Life time	300 - 400	400 - 500	500 - 600	600 - 700	700 - 800	800 - 900	900 - 1000
No. of Lamps	14	56	60	86	74	62	48

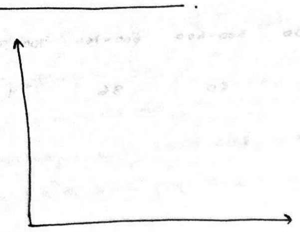


• Histogram Shapes :-

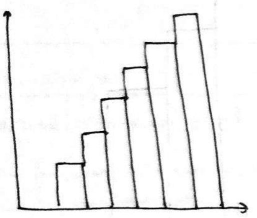
1. Bell - Shaped Histogram

• Left and Right part has the same symmetry

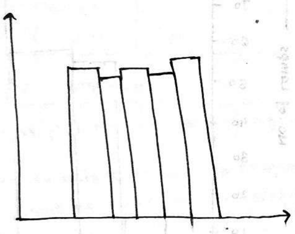
ii. **Bimodal Histogram**



(iii) **Skewed left histogram** :-



(iv) **Uniform Histogram**



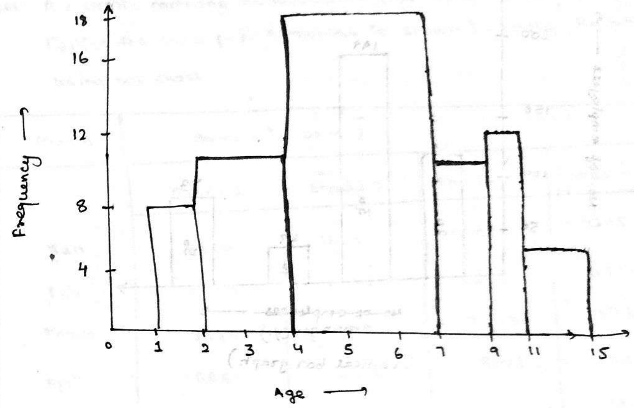
(Bars have almost same height.)

Ques: A random survey is done on the no. of children belonging to diff. age group who play in govt. park and the information is tabulated below:

Age (years)	1 - 2	2 - 4	4 - 7	7 - 9	9 - 11	11 - 15
Frequency	8	10	18	10	12	6

i. Draw a histogram graph of representative data.

ii. Identify No. of children belonging to age group 2, 3, 4, 5, 6 and 7 who play in Govt. parks. (28)



• **Bar Diagram variations**

(i) vertical bar chart
(ii) Horizontal bar chart

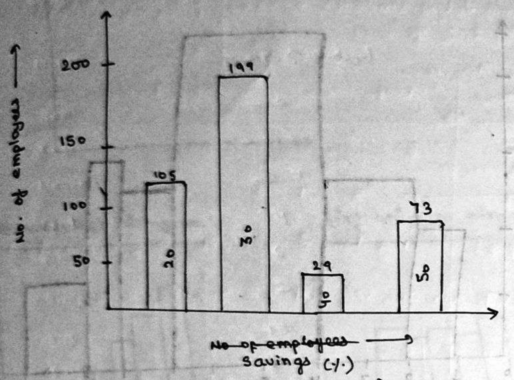
} → main categories

(iii) Grouped bar graph

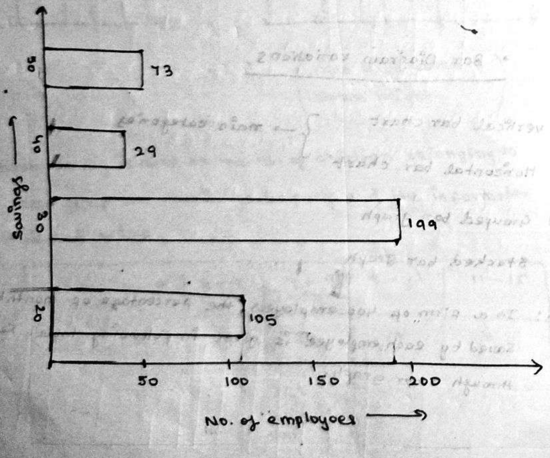
(iv) Stacked bar graph

Ques: In a firm of 400 employees, the percentage of monthly salary saved by each employee is given in following table. Represent through a bar graph.

Savings (in %)	10	20	30	40	50	Total
No. of employees	0	105	199	29	73	400



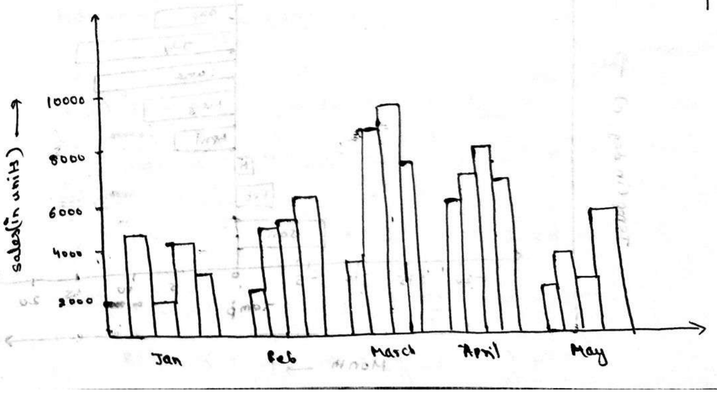
(vertical bar graph)



- **GROUPED BAR GRAPH** : It is also called clustered bar graph. It is used to represent the discrete value for more than one object that shares the same category.

Ques: A cosmetic company manufactures four different shades of lipstick the sale for 6 months is shown in table. Represent using bar chart.

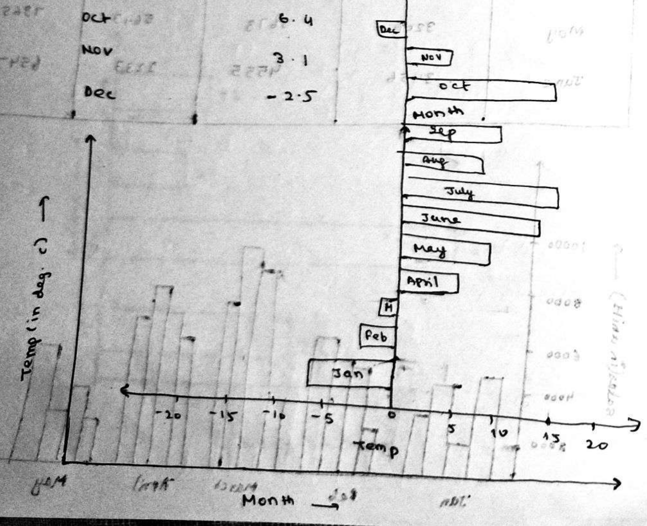
Month	Sales (in units)			
	Shade 1	Shade 2	Shade 3	Shade 4
Jan	4500	1600	4400	3245
Feb	2870	5645	5675	6754
March	3985	8900	9768	7786
April	6865	8976	9008	8965
May	3200	5678	5643	7865
June	3456	4555	2233	6547



- **STACKED BAR GRAPH** : It divides the aggregate into diff. Parts.

Ques: The variation of temperature in a region during a year is given below. Draw the bar graph :

Month	Temp (in deg c)
Jan	− 6
Feb	− 3.5
March	−2.7
April	4
May	6
June	12
Jul	15
Aug	8
Sep	7.9
Oct	6.4
Nov	3.1
Dec	− 2.5



- **Pie chart** : It is a type of graph that represents the data in circle graph. The slices of pie show the relative size of data.

Ques: A teacher surveys his class on the basis of scores of students. Construct a pie chart of the data given below.

Football	Hockey	Cricket	Basket Ball	Badminton
100	5	5	10	10

Total Students = 100 + 5 + 5 + 10 + 10 = 130

Football = $\frac{10\cancel{0}}{13\cancel{0}} \times 360 = \frac{3600}{13} = 276.92^\circ$

Hockey = $\frac{5}{13\cancel{0}} \times 36\cancel{0} = \frac{180}{13} = 13.84^\circ$

Basketball = $\frac{10}{13\cancel{0}} \times 36\cancel{0} = \frac{360}{13} = 27.69^\circ$

Badminton = $\frac{1\cancel{0}}{13\cancel{0}} \times 360^\circ = \frac{360}{13} = 27.69^\circ$

To get percentage value for each Sports :-

Football = $\frac{100}{13\cancel{0}} \times 10\cancel{0} = \frac{1000}{13}\% = 76.92\%$

Hockey = $\frac{5}{13\cancel{0}} \times 10\cancel{0} = \frac{50}{13}\% = 3.84\%$

Basketball = $\frac{10}{13\cancel{0}} \times 10\cancel{0} = \frac{100}{13}\% = 7.69\%$

Badminton = $\frac{10}{130} \times 100 = \frac{100}{13}\% = 7.69\%$

• **Box Plot** : when we display data distribution in a standardized way using five summary (minimum, maximum, Q_1 (first quartile), Q_2 (Median), Q_3 (Third quartile) and maximum) it is called box plot.

" A box plot is a special type of diagram that shows quartiles in a box and the line extending from lowest to highest value."

Ques: Let there is a series 11, 8, 10, 14, 13, 17, 29, 27, 33, 49, 22, 20, 29.

Soln : 8, 10, 11, 13, 14, 17, 20, 22, 24, 27, 29, 35, 49

Min = 8 , Max = 49 , Median = 20

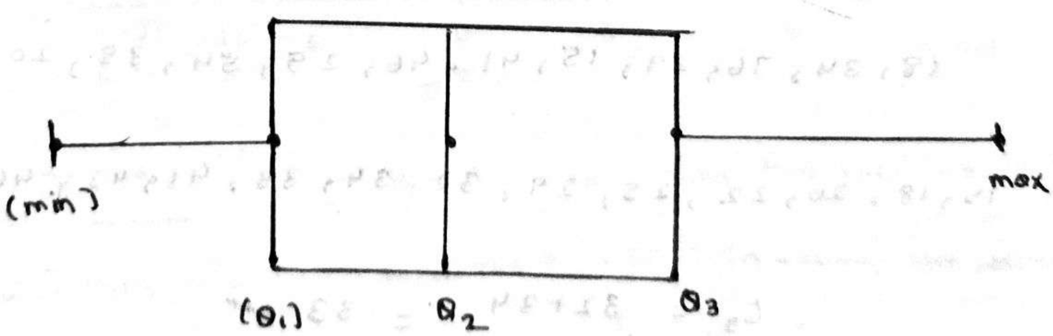
First Quartile (Q_1) = $\left(\frac{11+13}{2}\right) = \frac{24}{2} = 12$

Median of 1st of the data

Second Quartile (Q_2) = 20 (Median of data)

Third Quartile (Q_3) = $\left(\frac{27+29}{2}\right) = \frac{56}{2} = 28$

Median of 2nd half of data.



check min (3) and max (49) are outlier, calculate

$\boxed{\text{Outlier Range} = Q_1 - 1.5 IQR \text{ to } Q_3 + 1.5 IQR}$

$\boxed{IQR = Q_3 - Q_1}$ Inter Quartile Range

$IQR = Q_3 - Q_1$

$= 28 - 12 = 16$

upper extreme = $1.5 \times IQR$ | Lower extreme = $1.5 \times IQR$

upper bound = $Q_1 - 1.5 IQR$ | lower bound = $Q_3 + 1.5 IQR$

Outlier Range

$Q_1 - 1.5 IQR = 12 - (1.5)(16) = -12$ (min)

$Q_3 + 1.5 IQR = 28 + (1.5)(16) = 32$ (max)

Range = -12 to 32

Anything outside this range will be considered as outlier.

Ques2 : Given data plot an box plot

18, 34, 76, 29, 15, 41, 46, 25, 54, 38, 20, 32, 43, 22

Soln: 15, 18, 20, 22, 25, 29, 32, 34, 38, 41, 43, 46, 54, 76

$Q_2 = \frac{32+34}{2} = 33$

$Q_1 = 22$, $Q_3 = 43$

$IQR = Q_3 - Q_1 = 43 - 22 = 21$

Outlier Range

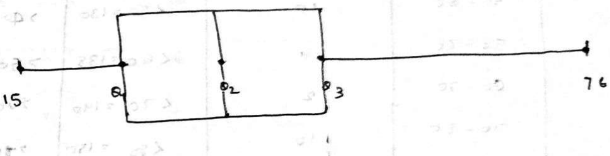
$Q_1 - 1.5 IQR = 22 - (1.5) \times (21) = -9.5$

$Q_3 + 1.5 IQR = 43 + (1.5)(21) = 74.5$

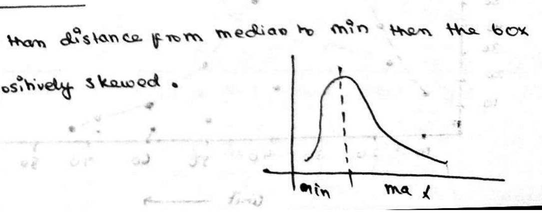
Min = 15, max = 76, $Q_1 = 22$, $Q_2 = 33$, $Q_3 = 43$

$IQR = 21$,

Outlier range = -9.5 to 74.5



• Box plot on a normal distribution :



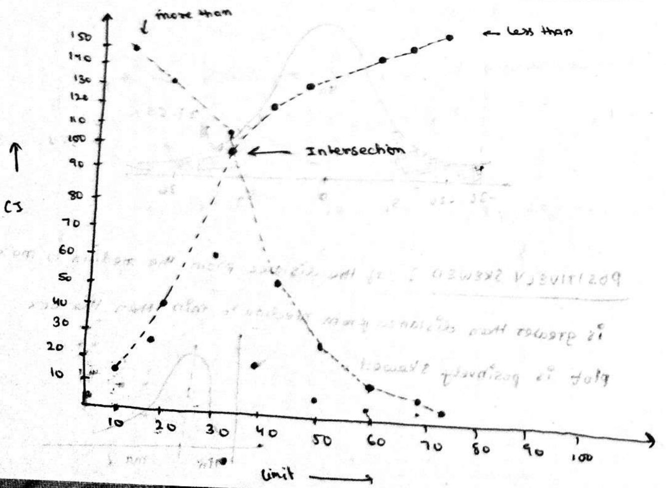
POSITIVELY SKEWED : If the distance from the median to max is greater than distance from median to min then the box plot is positively skewed.

• **O give** : It is an cummlative frequency graph (Total frequency till some point). O give is a graph that shows cummlative frequency of data usually in form of line graph.

Ques :

Marks	Frequency (No. of Students)	CMF Less than	More than
0 - 10	15	< 10 = 15	> 0 = 150
10 - 20	25	< 20 = 40	> 10 = 135
20 - 30	60	< 30 = 100	> 20 110
30 - 40	20	< 40 = 120	> 30 50
40 - 50	10	< 50 = 130	> 40 30
50 - 60	8	< 60 = 138	> 50 20
60 - 70	2	< 70 = 140	> 60 12
70 - 80	10	< 80 = 150	> 70 10

Total = 150

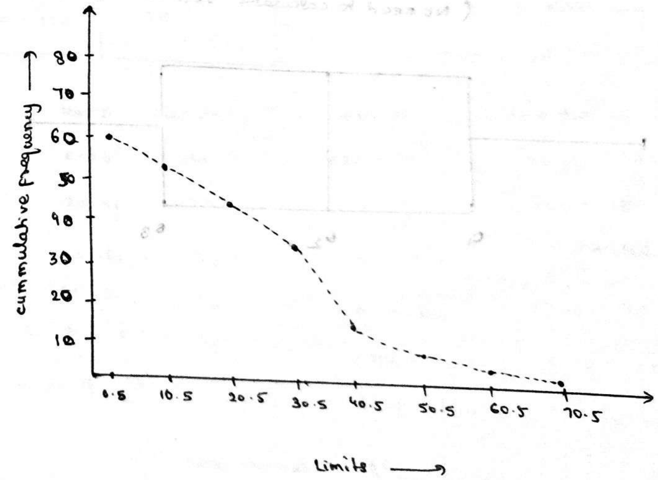


Ques :

Marks	1 - 10	11 - 20	21 - 30	31 - 40	41 - 50	51 - 60	61 - 70	71 - 80
Freq	3	8	12	14	10	6	5	2

Soln :

Marks	Frequency	more than CF
0.5 - 10.5	3	> 0.5 = 60
10.5 - 20.5	8	> 10.5 = 57
20.5 - 30.5	12	> 20.5 = 49
30.5 - 40.5	14	> 30.5 = 37
40.5 - 50.5	10	> 40.5 = 23
50.5 - 60.5	6	> 50.5 = 13
60.5 - 70.5	5	> 60.5 = 7
70.5 - 80.5	2	> 70.5 = 2



To draw box plot we need to find Q_1, Q_2, Q_3,

$Q_1 = 25\%$, $Q_2 = 50\%$, $Q_3 = 75\%$.

$Q_1 = 25\%$ of total $= \frac{25}{100} \times 60 = \cancel{30}\ 15$

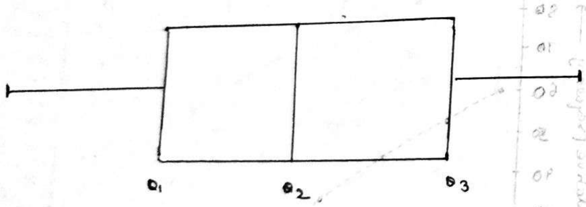
$Q_2 = 50\%$ of $60 = \frac{50}{100} \times 60 = 30$ (Median)

$Q_3 = 75\%$ of $60 = \frac{75}{100} \times 60 = 45$

$IQR = Q_3 - Q_1 = 45 - 15 = 30$

upper extreme $= Q_1 - 1.5\ IQR$ | Lower extreme $= Q_3 + 1.5\ IQR$
$= 15 - 1.5 \times 30$ | $= 45 + 1.5 \times 30$
$= -30$ | $= 90$

(No need to calculate IQR in Ogive)



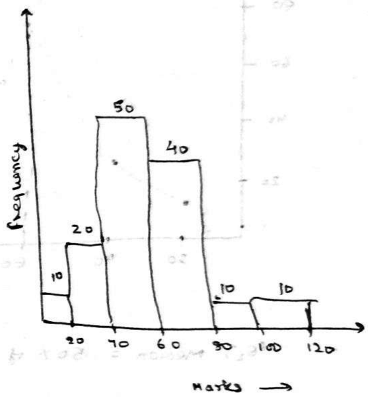
Ques: Draw a cf graph ogive from the histogram showing the distribution of marks in a math test for 140 students.

Maximum marks is 120.

(i) find median, mode, lower quartile, upper quartile.

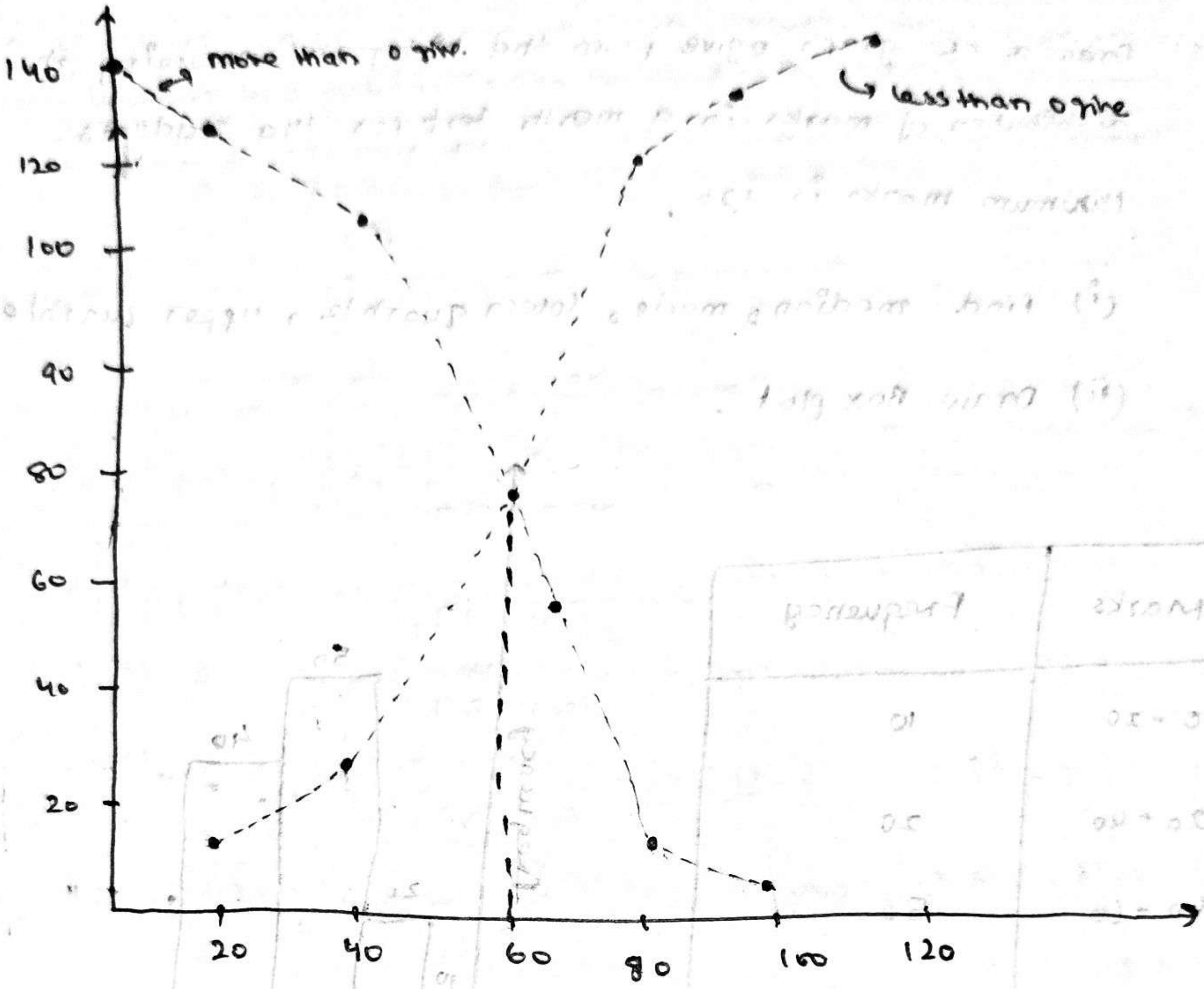
(ii) Draw Box plot.

Marks	Frequency
0 - 20	10
20 - 40	20
40 - 60	50
60 - 80	40
80 - 100	10
100 - 120	10



Marks	frequency	less than cf	more than cf
0-20	10	<20 = 10	>0 = 140
20-40	20	<40 = 30	>20 = 130
40-60	50	<60 = 80	>40 = 110
60-80	40	<80 = 120	>60 = 60
80-100	10	<100 = 130	>80 = 20
100-120	10	<120 = 140	>100 = 10

Marks interval $= \frac{120}{6} = 20$



(Q_2) Median $= 50\%$ of $140 = \frac{5\cancel{0}}{1\cancel{0}\cancel{0}} \times 14\cancel{0} = 70$

cf → marks $= 60$

$Q_1 = 25\%$ of $140 = \frac{\cancel{25}^{5}}{\cancel{100}_{2}} \times \cancel{140}^{7} = 35$

$Q_3 = 75\%$ of $140 = \frac{\cancel{75}^{15}}{\cancel{100}_{2}} \times \cancel{140}^{7} = \frac{105}{\cancel{2}} = 105$

PROBABILITY

- Probability theory is a Mathematical modeling of phenomena of chance or randomness.
- A probabilistic mathmatical model of random phenomena is defined by assigning probabilities to all the possible outcomes of experiments.
- Let a coin is tossed in a random manner :
  - * There is 50% chance of getting head and 50% chance of getting tail.
  - * Let S be the number of times heads appear when coin is tossed n times.

$$\boxed{\text{Ratio } f = s/n}$$

* SAMPLE SPACE AND EVENTS

- The set S of all possible outcomes of a given experiment is called the sample space.
- A particular outcome, an element in S is called sample Point.
- An even A is a set of outcomes or in other words a subset of the sample space S.

NOTE : (i) Let $A \cup B$ is the event that occurs iff A occurs or B occurs (or both).

(ii) $A \cap B$ is the event that occurs iff A occurs and B occurs.

(iii) A^c, compliment of A is the event that occurs iff A does not occur.

(iv) if $A \cap B = \phi$

Ques1: Toss a die and observe the number that appears on top

$S = \{1, 2, 3, 4, 5, 6\}$

(i) Let A be the event an even No. occurs

(ii) B be the event an odd No. occurs,

(iii) C be the event a prime No. occurs.

$A = \{2, 4, 6\}$ $B = \{1, 3, 5\}$ $C = \{2, 3, 5\}$

$A \cup C = \{2, 3, 4, 5, 6\}$

$B \cap C = \{3, 5\}$

$\bar{C} = \{1, 4, 6\}$

Ques2: Toss a coin three times and observe the sequence of head (H) and tails (T).

$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$

(i) A = Two or more heads appears consecutivly

$A = \{HHH, HHT, THH\}$

(ii) Let event B be a event such that all tosses are same.

$B = \{HHH, TTT\}$

- let S be a finite sample space, a finite probability space or probability model is obtained by assigning to each point a_i in S a real number p_i. called probability of a_i

$$S = \{a_1, a_2, a_3, \ldots, a_n\}$$

statisfying the following :

i. Each p_i is non - negative, $p_i \geq 0$

ii. sum of p_i is 1. i.e $p_i + p_2 + \ldots + p_n = 1$

Then probability of event A written as P(A) is defined to be the sum of probabilities of the point in A.

$P(a_i)$ is probability of a_i.

Ques 3: Let 3 coins are tossed and No. of heads are observed.

$$S = \{0, 1, 2, 3\}$$

$P(0) = 1/8$ $\quad$ $P(2) = \frac{3}{8}$

$P(1) = 3/8$ $\quad$ $P(3) = 1/8$

Probability of 3 heads.

let A be the event at least 1 head appears

$$P(A) = 7/8$$

let B be the event that all heads or all tail's appear

$$P(B) = P(0) + P(3) = \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$$

Ques 4: Let a card be selected from an ordinary deck of 52 playing cards. Let

A = {the card is spade}

B = {card is a face card}

$P(A) = \frac{13}{52} = \frac{1}{4}$, $P(B) = \frac{12}{52} = \frac{3}{13}$.

$P(A \cap B) = \frac{3}{52}$

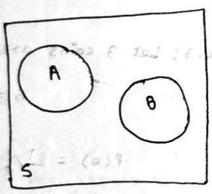
Theorem :

i. For every event A, $0 \leq P(A) \leq 1$

ii. $P(S) = 1$

iii. If A and B are mutually exclusive then,

$P(A \cup B) = P(A) + P(B)$



iv. $P(\phi) = 0$

v. $\cancel{P(A/B)} = P(A) - P(A \cap B)$

$P(A \setminus B)$ or $P(A - B)$

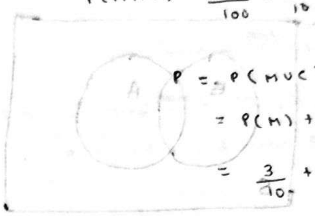
vi. If $A \subseteq B$ then $P(A) \leq P(B)$

vii. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Ques 5: Suppose a student is selected at random from 100 students where 30 are taking maths, 20 are taking chemistry and 10 are opting for both subjects. Find probability p such that student is taking mathematics and chemistry.

$P(M) = \frac{30}{100}$, $P(C) = \frac{20}{100}$,

$P(M \cap C) = \frac{10}{100} = \frac{1}{10}$,



$P = P(M \cup C)$

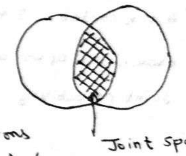
$= P(M) + P(C) - P(M \cap C)$

$= \frac{3}{10} + \frac{1}{5} - \frac{1}{10} = \frac{2}{5}$

JOINT PROBABILITY

- It refers to statistical measure that calculates the likelyhood of two events occuring together and at the same time.
- Both event must be independent of other, Means they are not conditional or dnot re on eachother

Ques 6: In a hospital two posts are advertised one for eye specialist and one for ENT. Total 5 applications are recieved for eye specialist and 6 for ENT.



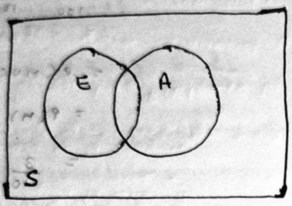
Joint space

Ques 7: Two posts of MBBS doctors are advertised. Total 11 appln are recieved.

what is the probability that both husband (eye specialist) and wife (ENT) would be selected in above Ques 6 & 7.

CONDITIONAL PROBABILITY: Let E is an event in a sample space S with $P(E) > 0$. Probability that an event A occurs once E has occured.

$$P(A/E) = \frac{P(A \cap E)}{P(E)}$$



Let $n(A)$ denotes No. of elements in even A then,

$$P(A \cap E) = \frac{n(A \cap E)}{n(S)}, \quad P(E) = \frac{n(E)}{n(S)}$$

So, $\boxed{P(A/E) = \frac{n(A \cap E)}{n(E)}}$

Ques: A box contains 3 red and 2 black balls two balls are drawn one after another without replacement. what is Joint probability that first is red and second is blue?

[Ans. 3/10]

Ques: In a class of 30 students, 18 boys and 12 girls 10 students are selected at random what is the probability that 6 are boys and 4 are girls.

Ans = $\frac{{}^{18}C_6 \times {}^{12}C_4}{{}^{30}C_{10}}$

Ques: A card is drawn from deck 52 cards. what is the probability that the card is king consording that it is a face card.

Ans = $\frac{4}{12}$

Ques: In a survey 40% like tea, 30% like coffee and 10% like both, what is probability that the person like coffee given that he likes tea.

Ans: $(1/4)$

Ques: Two dice are rolled what is probability that first shows even and second shows greater than 4.

[Ans: 1/6]

Ques: A jar has 5 red balls and 3 green balls. one ball is drawn at random. If it known that ball is not green what is the probability that ball is red.

[Ans: 1]

Ques: A couple had two children with probability 1/4 for each. Find probability that both children are boys know that older one boys.

{ GG, GB, BG, BB }

$\frac{1}{2}$

Ques 10: A pair of dice is tossed. The sample space S consists 36 ordered pairs (a, b) where a and b can be any integer from 1 to 6. Find probability that one dice is 2 if sum is 6.

Sol^n: $S = \{(1,5), (2,4), (3,3), (4,2), (5,1)\}$

$P = 2/5$

E = { sum is 6 }, A = { 2 Appears on at least one dice }

(2, 4) and (4, 2) belong to A,

$A \cap E = \{ (2,4), (4,2) \}$

$$P(A/E) = \frac{P(A \cap E)}{P(E)} = \frac{2}{5}$$

Ques11: A ~~couple has two~~ lot contains 12 items of which four are defective. Three are drawn at random one after another. Find the probability that all are none defective.

First item is Non-defective = $8/12$

Second item is Non-defective = $7/11$

Third item is Non-defective = $6/10$

$$P(A) = \frac{\cancel{8}^{\cancel{4}\,2}}{\cancel{12}_{2}} \times \frac{7}{11} \times \frac{\cancel{6}}{\cancel{10}_{5}}$$

$$= \frac{14}{55} = 0.25$$

RANDOM VARIABLE

- It may be convinient to assign 1 to h and 0 to t for tossing of a point.
- In tossing of a pair of dice, we may want to assign the sum of two Integers
- such an assignment of numerical value is called random variable.

"A random variable "X" is a rule that assigns a numerical value to each outcome in a sample space S."

Let R_X denotes the set of numbers assigned ~~by~~ a random variable is refered as a range space.

Ques12: A pair of dice is tossed. The sample space S consists of the 36 ordered pairs (a, b) where a and b can be any integers between 1 to 6.

$S = \{ (1,1), (1,2), \ldots, (6,6) \}$

Let X assign to each point in S the sum of the numbers then X is a random variable with range space R_X

$R_X = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$

Let Y assign to each point the maximum of the two numbers then Y is a random variable 1 to 6

$R_Y = \{1, 2, 3, 4, 5, 6\}$

Ques13: A box contains 12 items of which 3 are defective a sample of three items is selected from the box. The sample space S consist of the ${}^{12}C_3$ different samples of size 3. Let X denote the number of defective items in the sample then X is a random variable with range space

$R_X = \{0, 1, 2, 3\}$

PROBABILITY DISTRIBUTION OF A RANDOM VARIABLE

Let $R_X = \{x_1, x_2, x_3, \ldots, x_n\}$ be the range space of random variable X defined on a finite sample space S. Then X ~~includes~~ induces an assignment of probabilities on the range space R_X as

$$p_i = P(x_i) = P(X = x_i) = \text{sum of probabilities}$$

of point S whose image is x_i.

$$p_i = P(x_i) = \frac{\text{No. of points in S whose image is } x_i}{\text{No. of points in S}}$$

x_i	x_2	....	x_t
p_1	p_2	.....	p_x

Ques 14: Refer Question 12,

$P(2) = \frac{1}{36}$, There are two outcomes (1, 2) and (2, 1) whose sum is 3. So

$P(3) = \frac{2}{36} = \frac{1}{18}$,

There are three outcomes (1, 3), (2, 2), (3, 1) whose sum is 4, $P(4) = \frac{3}{36} = \frac{1}{12}$

There are 4 outcomes whose sum is 5,

$P(5) = \frac{4}{36} = \frac{1}{9}$

$P(6) = \frac{5}{36}$, . . . , $P(12) = \frac{1}{36}$

The distribution of X,

x_i	2	3	4	5	6	7	8	9	10	11	12
p_i	$1/36$	$2/36$	$3/36$	$4/36$	$5/36$	$6/36$	$5/36$	$4/36$	$3/36$	$2/36$	$1/36$

Ques 15: Refer Ques 13, Total sample = 220, ${}^{12}C_3$

Let X be the random variable,

${}^{9}C_3 = 84$ samples of size 3 with no defect.

$P(0) = \frac{84}{220}$, $P(1) = \frac{3 \times {}^{9}C_2}{220} = \frac{108}{220}$ samples of size 3 containing one defective item.

Two defectives:

$P(2) = \frac{9 \times {}^{3}C_2}{220} = \frac{27}{220}$ Samples of size 3 with two defectives.

Three defective:

$P(3) = \frac{1}{220}$,

x_i	0	1	2	3
p_i	$\frac{84}{220}$	$\frac{108}{220}$	$\frac{27}{220}$	$\frac{1}{220}$

Expectation of random Variable

Let X be a random variable.

Let V or μ_X is mean of X,

let σ or σ_X is standard deviation of X.

Mean is also called the expectation of X, written as $E(X)$.

Let $S = \{a_1, a_2, \ldots, a_n\}$ X is a random variable on the sample space S.

The mean or expectation of X is defined as —

$$m = E(X) = X(a_1)\cdot P(a_1) + X(a_2)\cdot P(a_2) + \ldots + X(a_n)$$

$$P(a_m) = \sum X(a_i)\cdot P(a_i)$$

$$\mu = E(X) = x_1 P_1 + x_2 P_2 + \ldots + x_m P_m$$

$$= \sum x_i \cdot P_i$$

x_i	x_1	x_2	$\ldots$	x_m
P_i	P_1	P_2	$\ldots$	P_m