CA FOUNDATION

MARATHON

JUNE 2024

Quantitative Aptitude

By Anurag Chauhan

Chapters
to be covered

1. Statistical Description of Data & Sampling
2. Central Tendency and Dispersion
3. Correlation & Regression
4. Index Number

2022-J → 9 marks
2022-D 4 marks
2023-J → 5 marks
2023-D ⇒ 7 marks

Statistical Description of Data

Statistics Originated From

Status $\Rightarrow$ Latin

Statista $\Rightarrow$ Italian

Statistik $\Rightarrow$ German

Statistique $\Rightarrow$ French

Parabola
$\Downarrow$
Latus Rectum
$\Downarrow$
Status

Swastik

Queen
$\Downarrow$
models
$\Downarrow$
pants

Definition of Statistics

Singular Sense

Plural Sense

Definition of Statistics (Singular sense)

Scientific Method
- -Collection Of data
- -Classification
- -Presentation
- -Analysis
- -Interpretation

Definition of Statistics (Plural sense)

Systematically collected numerical facts

Limitation of Statistics

-Deals in aggregates only
-Does not deal in qualitative data
-Not always true
-Expert consultation required
-They can be misused

DATA

⇓

Information of some character

Qualitative character

⇓

Attributes

Eg Colour

Eg Nationality

Eg Caste

Quantitative character

⇓

Variable

Descrete

Eg No. of Accidents

Eg Goals in a match

Eg Salary

Continous

Eg weight

Eg Height

Eg Age

COLLECTION OF DATA

Primary

Collected For First Time

Secondary

Public (Existing) Information used by others

COLLECTION OF PRIMARY DATA

- 1) Interview Method

Personal
(Direct)
(Natural Calamity)
{ Earthquake
{ cyclone

Indirect
(Train Accident)

Telephonic
(Fastest)

2) Questionnaire Method

- i) Mailed Questionnaire Method

  Questionnaire is mailed to all target respondents

- ii) Enumerator Method

  Questionnaire filled by enumerator
  On behalf of respondents

Customer Satisfaction Survey Questionnaire

I. Questions

Directions: Please indicate your level of agreement or disagreement with each of these statements regarding QRZ Family Restaurant. Place an "X" mark in the box of your answer.

Q1: How many times per year do you visit QRZ Family Restaurant?

____

Q2: Do you visit QRZ Family Restaurant with family or friends?

□ Yes □ No

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
1. The store is accessibly located.					
2. Store hours are convenient for my dining needs.					
3. Advertised dish was in stock.					

- 3) Observation Method

Investigator's own direct observation without asking respondents

Used for research together data about people, objects, events & behaviors etc.

Accurate Data
Time Consuming

COLLECTION OF SECONDARY DATA

International Sources

- WHO
- IMF
- WORLD BANK

Govt Sources

- Ministry of agriculture

Private and Semi govt

Unpublished Sources

जाँच

Scrutiny Of Data

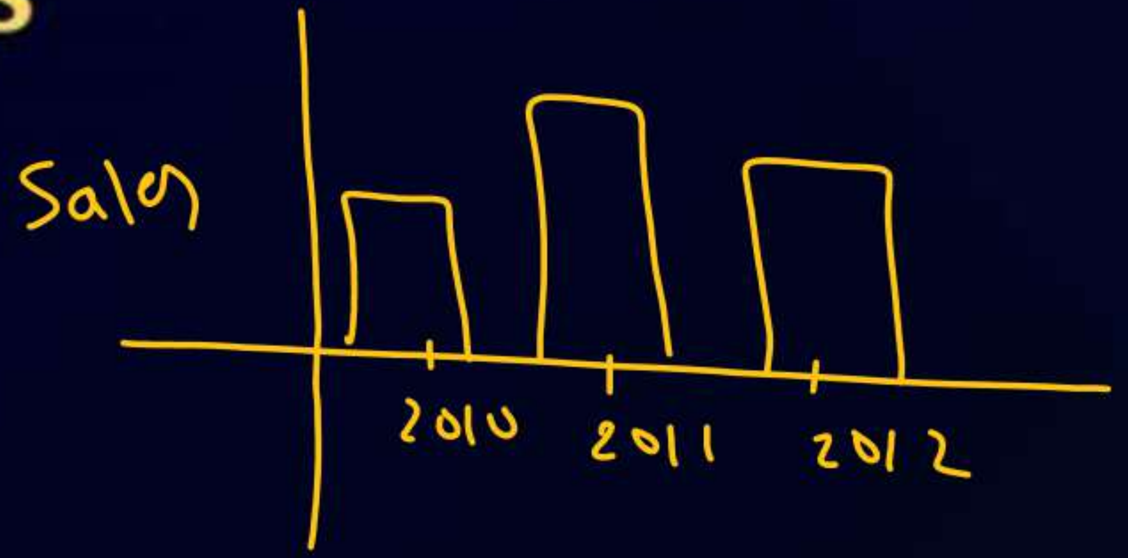
Verification of

- Accuracy
- Consistency Of Data

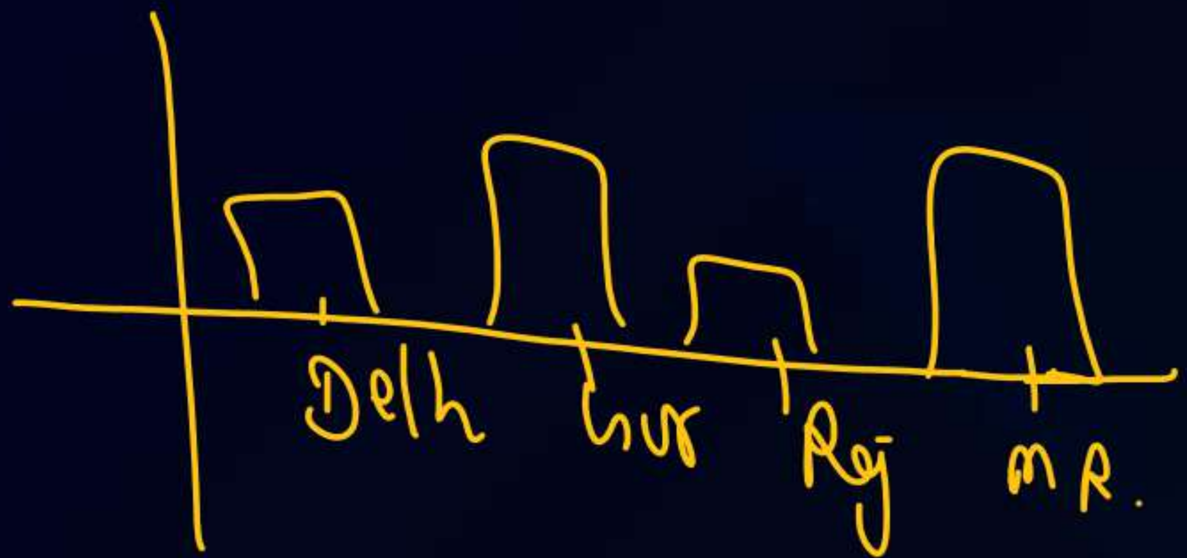
Classification Of Data

Data may be classified as

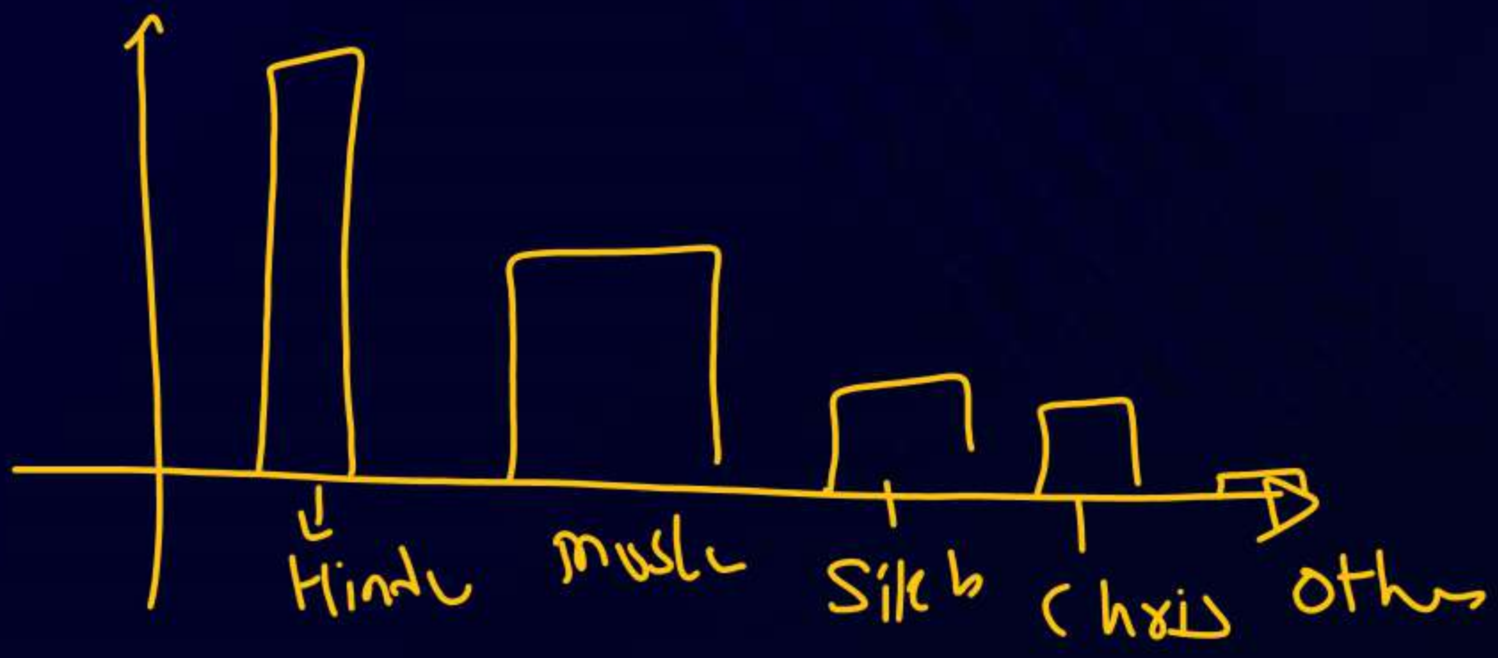
- i)Temporal (Chronological) → on the Basis of time



- ii)Geographical (Spatial) → on the Basis of Place

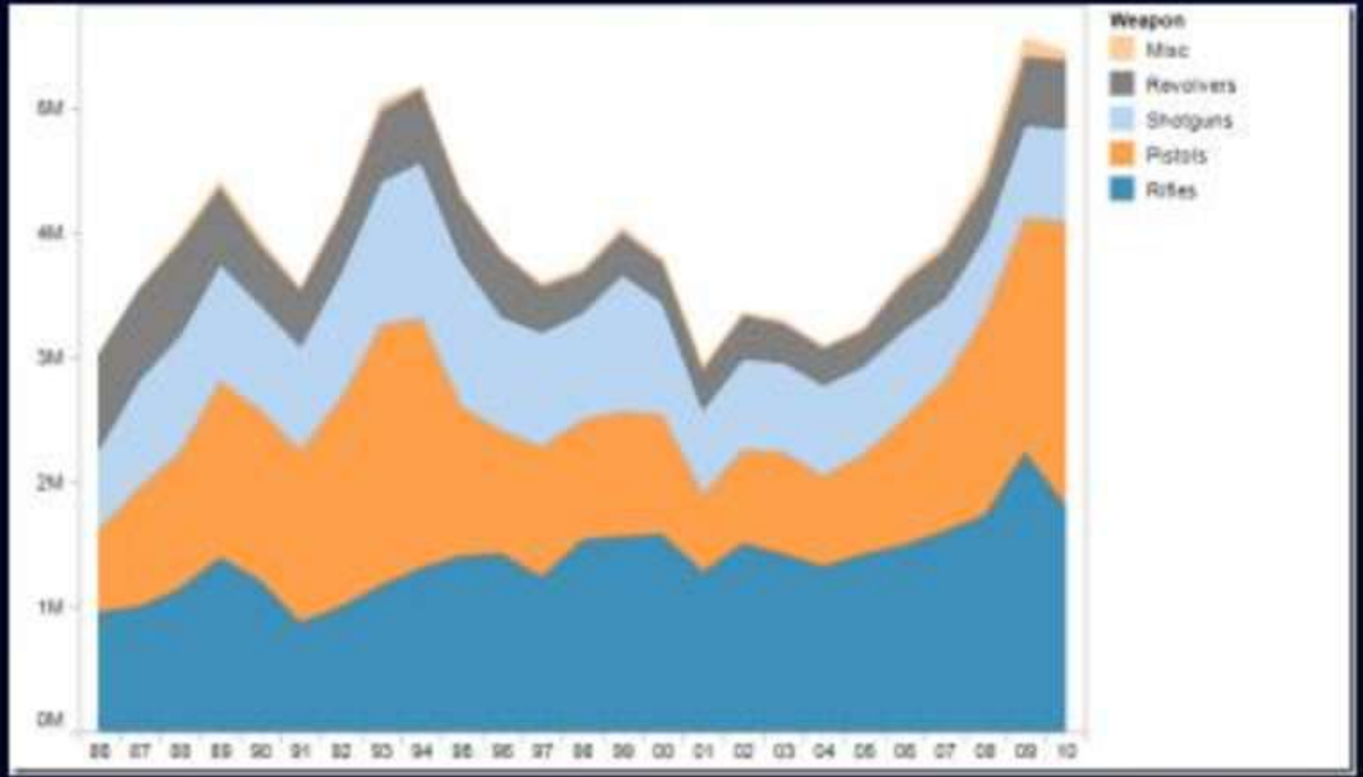


iii) Qualitative
(Ordinal)



iv)Quantitative
(Cardinal)

Presentation of Data



1) Textual Presentation : E.g. Paragraph

2) Tabular Presentation : E.g. Table
(This Is Best Method)

3) Diagrammatic Presentation : Eg Bar diagram , Pie Chart
(This is most attractive) (Hidden Trend Can Be Noticed)

#Textual Presentation (Worst method)

Paragraph

In 2009, out of a total of five thousand workers of Roy Enamel Factory, four thousand and two hundred were members of a Trade Union. The number of female workers was twenty per cent of the total workers out of which thirty per cent were members of the Trade Union. In 2010, the number of workers belonging to the trade union was increased by twenty per cent as compared to 2009 of which four thousand and two hundred were male. The number of workers not belonging to trade union was nine hundred and fifty of which four hundred and fifty were females.

#Table

Best method

1)Table Number

2)Title –Explains content of table

3)Head Note –Given in brackets provides information of units

4)Column Heading-(Caption)- Describes columns and subcolumns

5)Row Heading (Stubs) – Describes Rows

6)Body Of Table - Numeric Information

7)Source

8)Footnote –Specific feature which is not explained

Column Heading (Caption)

Table 1.1 ← Table Number

Literacy in Bihar by Sex and Locations ← Title

(Percent) ← Head Note

Row Heading (Stub) →

Sex	LOCATION		Total
	Rural	Urban	
Males	69.67	82.56	71.20
Females	44.30	61.95	51.50
Total	59.78	76.86	61.80

← Row Total

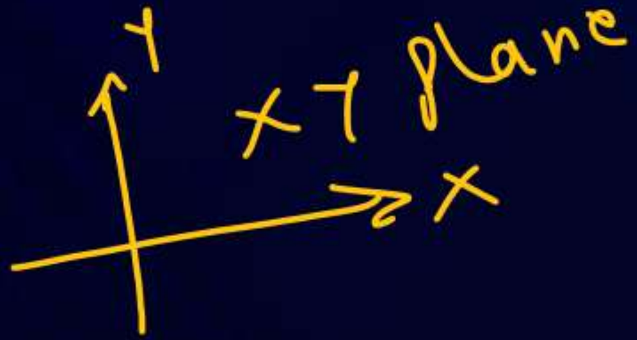
Row Entries (Stub Entries) →

Column Total →

Source Note → ***Source Note:*** *Census of India 2011, Provisional Population* ***Totals***

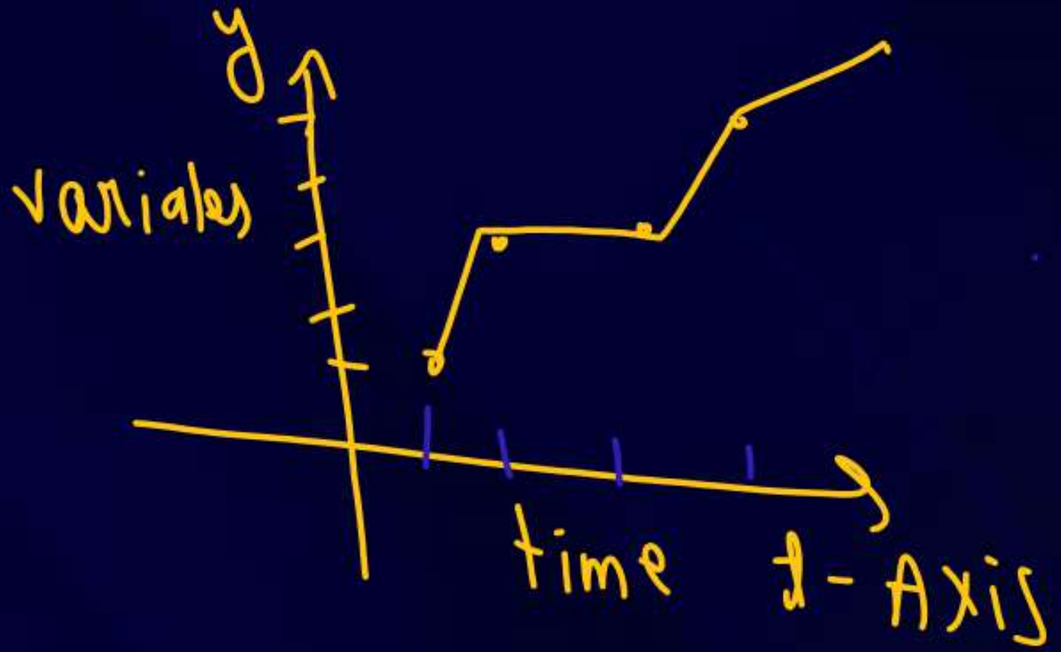
Footnote → ***Note:*** *Figures are rounded to two digits after decimal.*

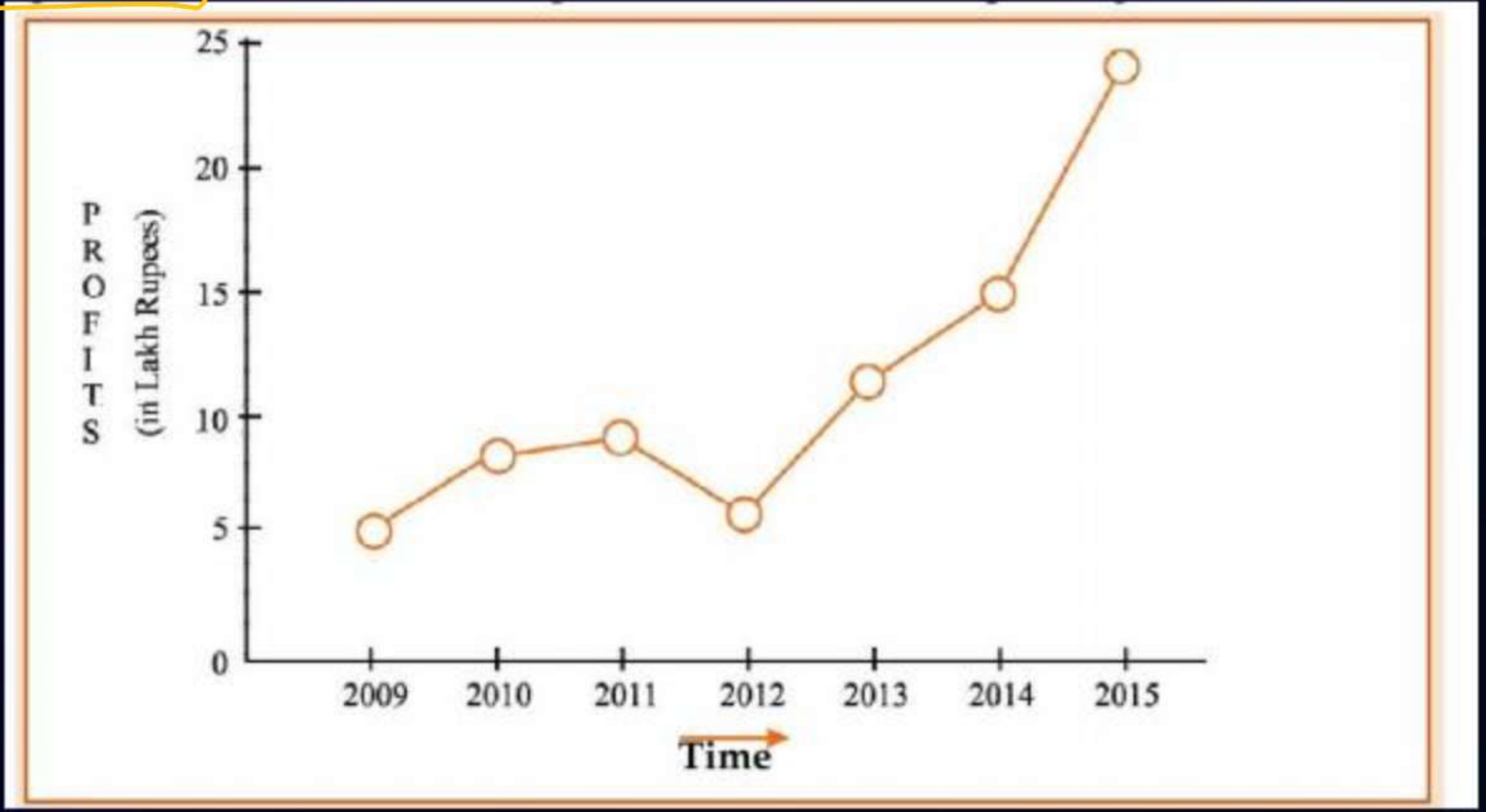
#Diagrammatic Presentation



1) Line Diagram(Historiagram)

- -One Dimensional
- -ty Plane
- -Time Series

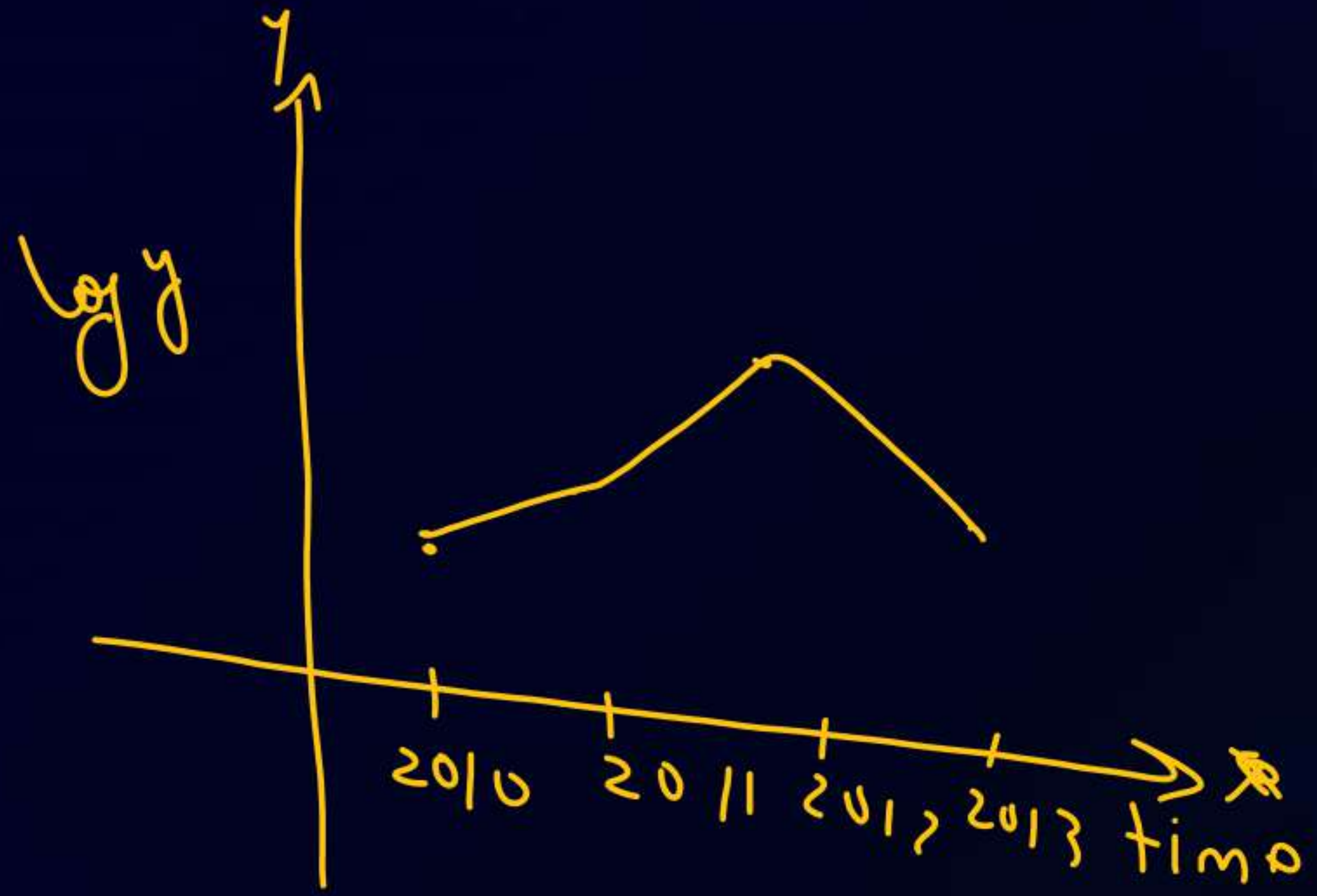




2) Ratio Chart

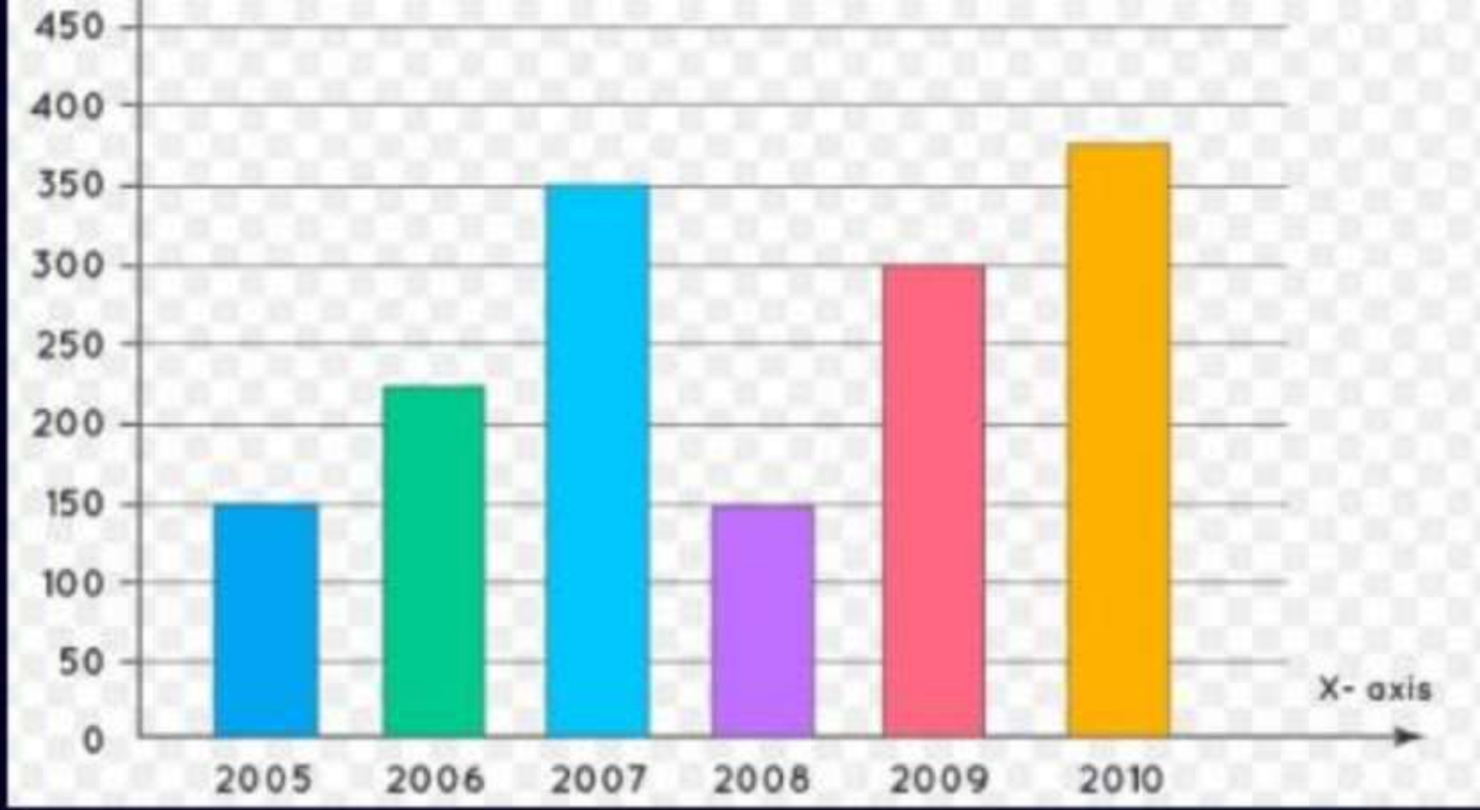
When time series have fluctuations we use ratio chart

(t) time	(y) Sales	log y_i
2010	20	1.3010
11	30	1.4771
12	50	.
13	60	—



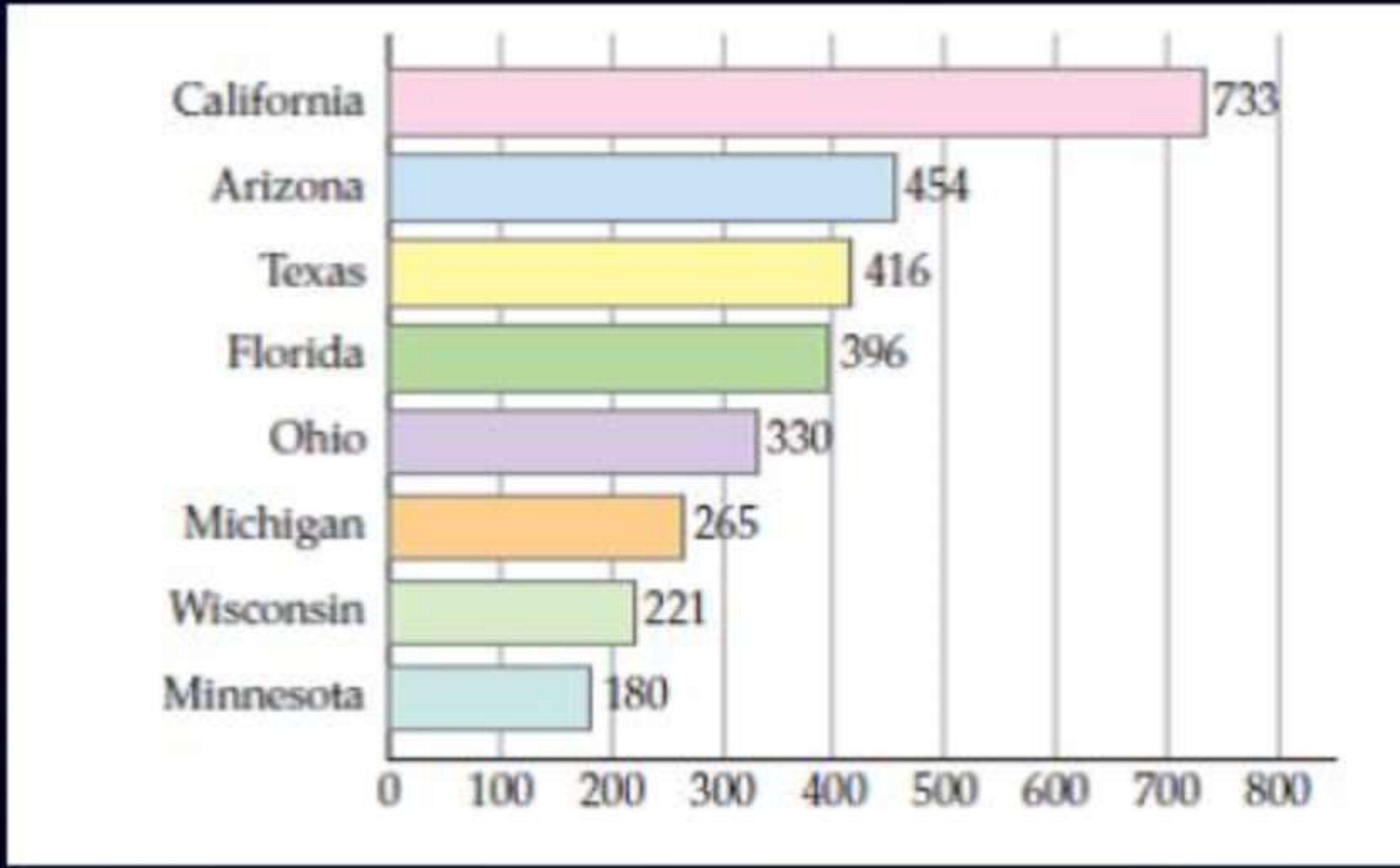
3) Bar Diagram

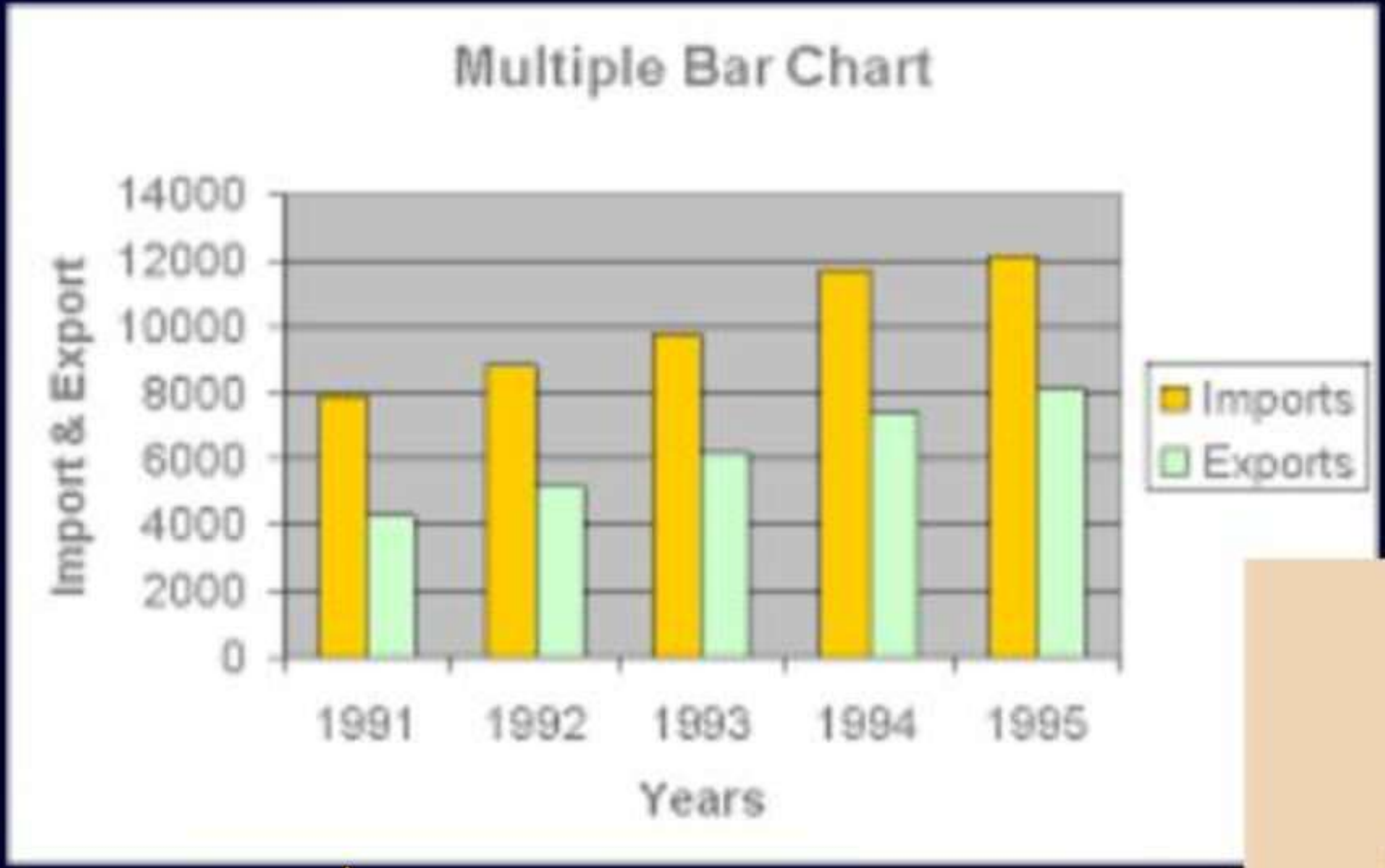
-Rectangular bars(Only Length matters)
- One Dimensional



Vertical Bar Diagram – Quantitative data
(When data vary over time)
(Temporal)

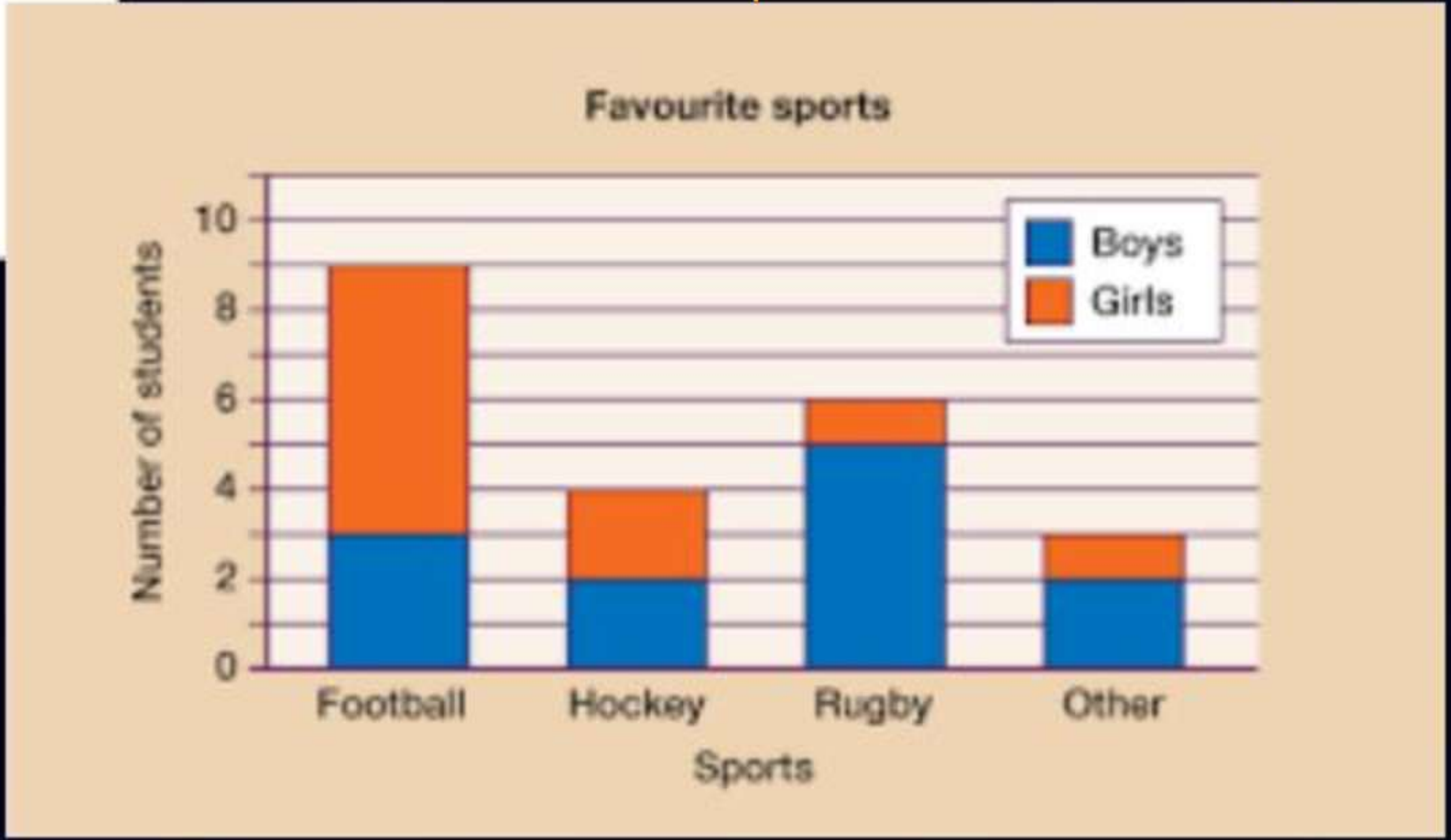
Horizontal Bar Diagram – Qualitative Data
(When data vary over space)
Spatial





Multiple Bar Diagram

Component Bar Diagram

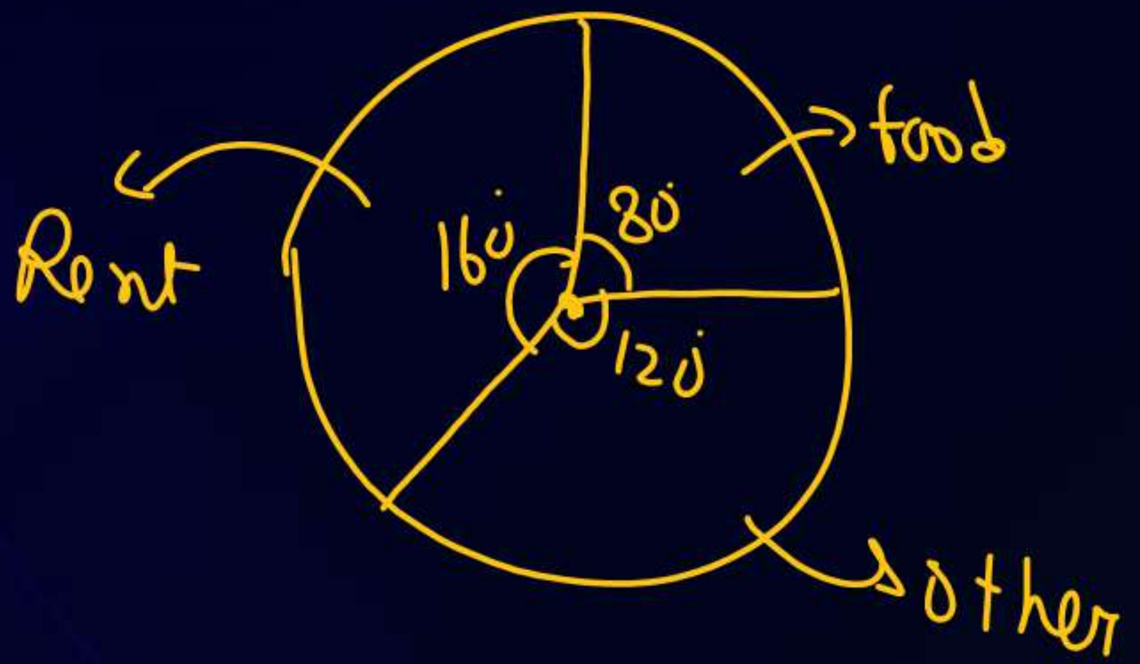


4)Pie Chart

Displays data in a circular-shaped graph

- 2 Dimensional

	Expenses	Central angle
Food	2000	$\frac{2000}{9000} \times 360 = 80$
Rent	4000	$\frac{4000}{9000} \times 360 = 160$
Other	3000	$\frac{3000}{9000} \times 360 = 120$
	9000	360°



Frequency Distribution

Tabular presentation of Statistical Data

Number of accidents in last 15 days

1,2,1,1,1,2,3,3,2

1,1,1,3,2,2

No. of Accidents (x_i)	Tally Bar	frequency
1	~~||||~~ ||	7
2	~~||||~~	5
3	|||	3

Types of Series

Individual series

⇓

Marks: 1, 2, 3, 1, 4,
5, 2, 4, 5, 1, 3
6, 1, 1, 2, 5, 2, 2

Descrete series

⇓

Marks (x_i)	frequency (f_i)
1	2
2	3
3	4
4	2
5	1
	N = 12

Continous series

⇓

Intervals

Continous series

Exclusive ⇓

C I	fi
0 - 2	3
2 - 4	2
4 - 6	5
6 - 8	1

Inclusive

C I	fi
1 - 6	3
7 - 12	6
13 - 18	4
19 - 24	1

Open ended series

C I	fi
les than 10	2
10 - 12	3
12 - 14	1
14 - 16	4
more than 16	5

CI	fi
3 – 8	6
9 – 14	2
15 – 20	5

CI	fi
2.5 – 8.5	6
8.5 – 14.5	2
14.5 – 20.5	5

3 – 8

3 → lower class limit (LCL)

8 → upper class limit (UCL)

2.5 – 8.5

2.5 → L.C.L & lower class Boundary

8.5 → upper class Boundary

frequency Density = $\dfrac{\text{Frequency of class interval}}{\text{Class Length}}$

Relative frequency = $\dfrac{\text{Frequency class interval}}{\text{Total frequency}}$

Eg

CI	f_i	Frequency Density	Relative frequency
0-5	3	$\frac{3}{5} = 0.6$	$\frac{3}{20} = 0.15$
5-10	6	$\frac{6}{5} = 1.2$	$\frac{6}{20} = 0.30$
10-15	2	$\frac{2}{5} = 0.4$	$\frac{2}{20} = 0.10$
15-20	9	$\frac{9}{5} = 1.8$	$\frac{9}{20} = 0.45$
	20		1

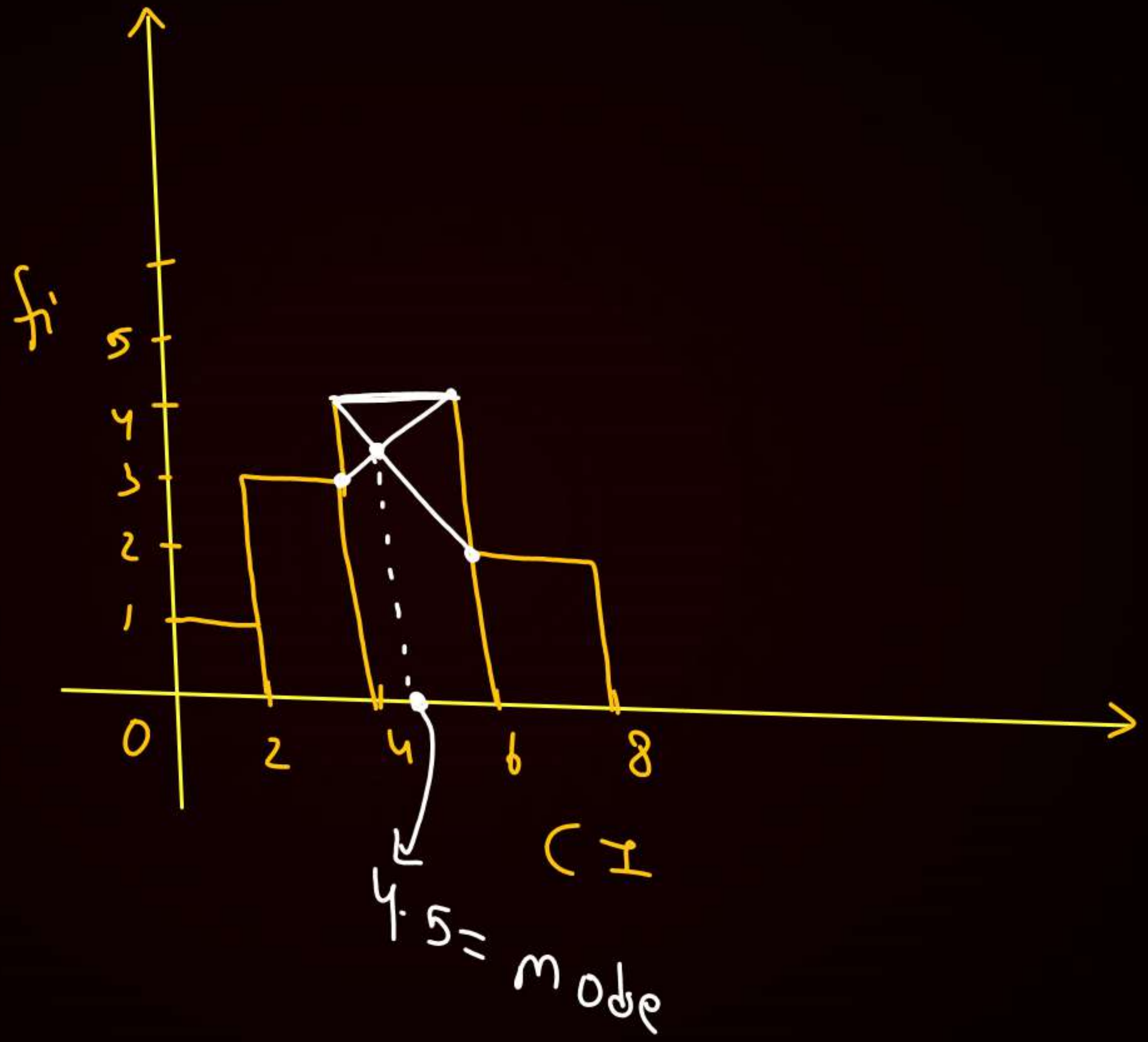
Graphical Representation of a Frequency Distribution

Histogram
(Area Diagram)
⇓
2 Dimension

frequency
Polygon
one-Dimensional

Cumulative frequency curve
(O-give)
one Dimensional

CI	fi
0-2	1
2-4	3
4-6	4
6-8	2

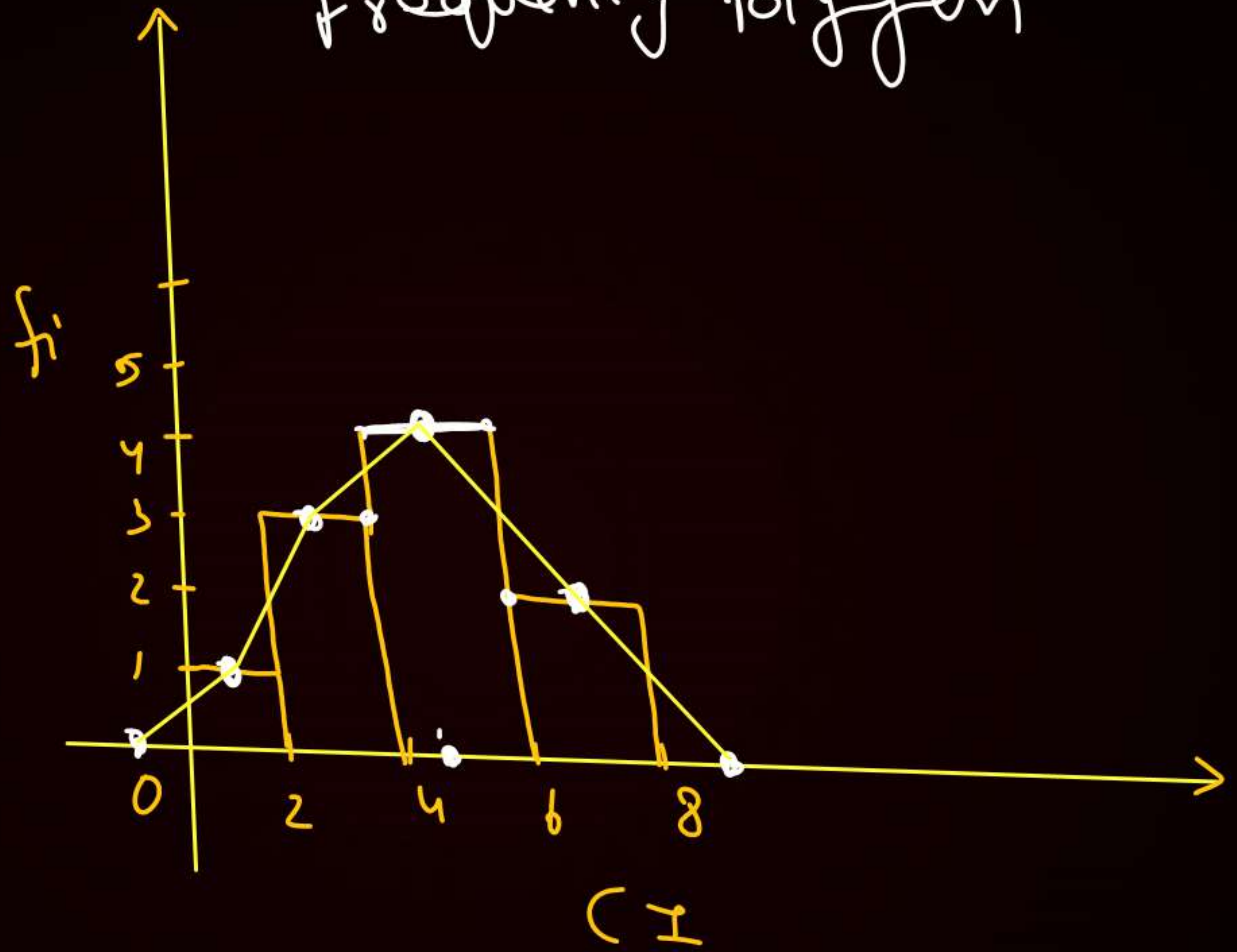


Eg

CI	f_i
0 - 2	3
2 - 4	2
4 - 8	6
8 - 10	1

Frequency Polygon

CI	fi
0-2	1
2-4	3
4-6	4
6-8	2

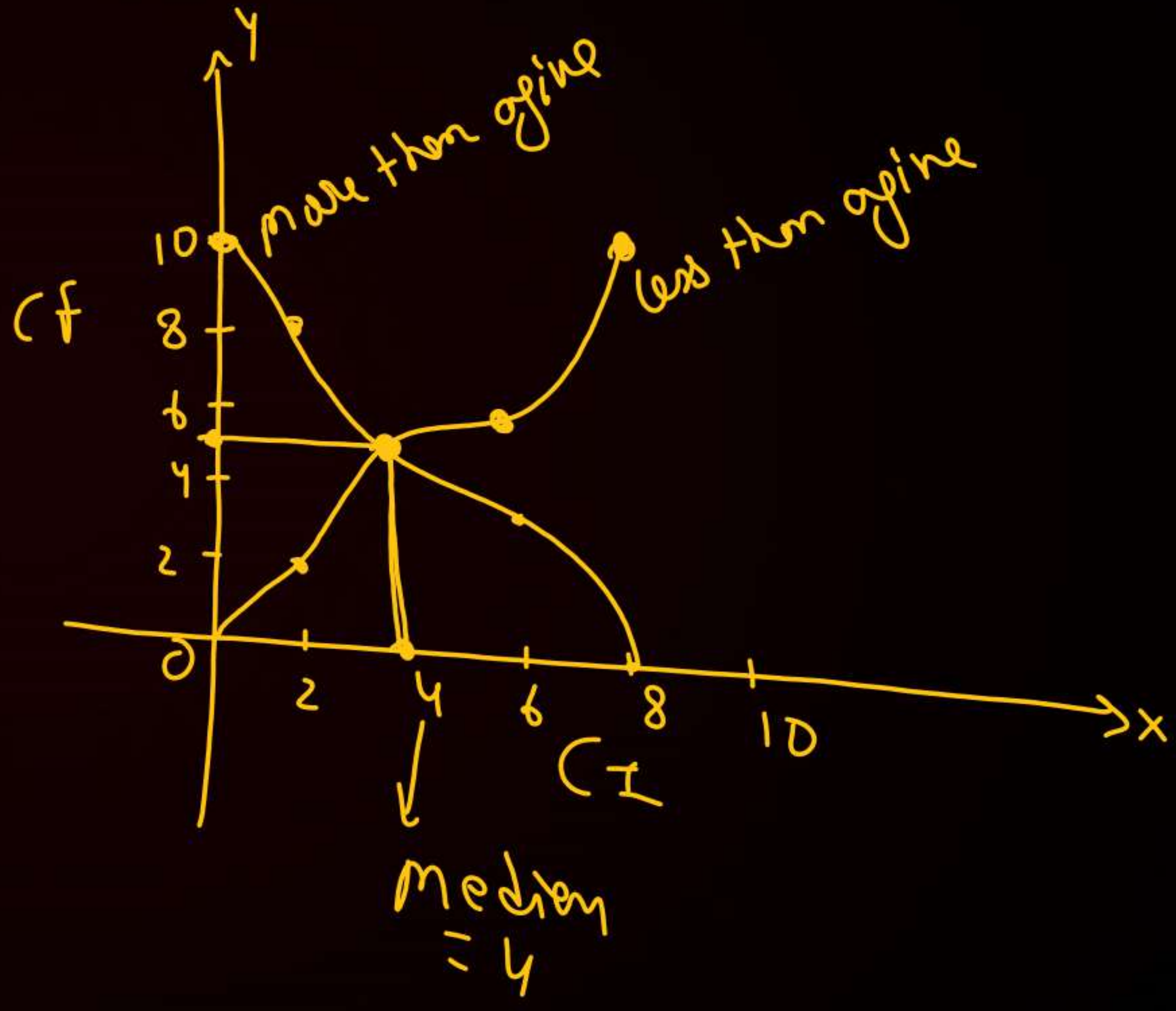


Cumulative frequency curve (ogive)

CI	fi	less than	cf	more than	cf
0-2	2	2	2	0	10
2-4	3	4	5	2	8
4-6	1	6	6	4	5
6-8	4	8	10	6	4

$N = 10$

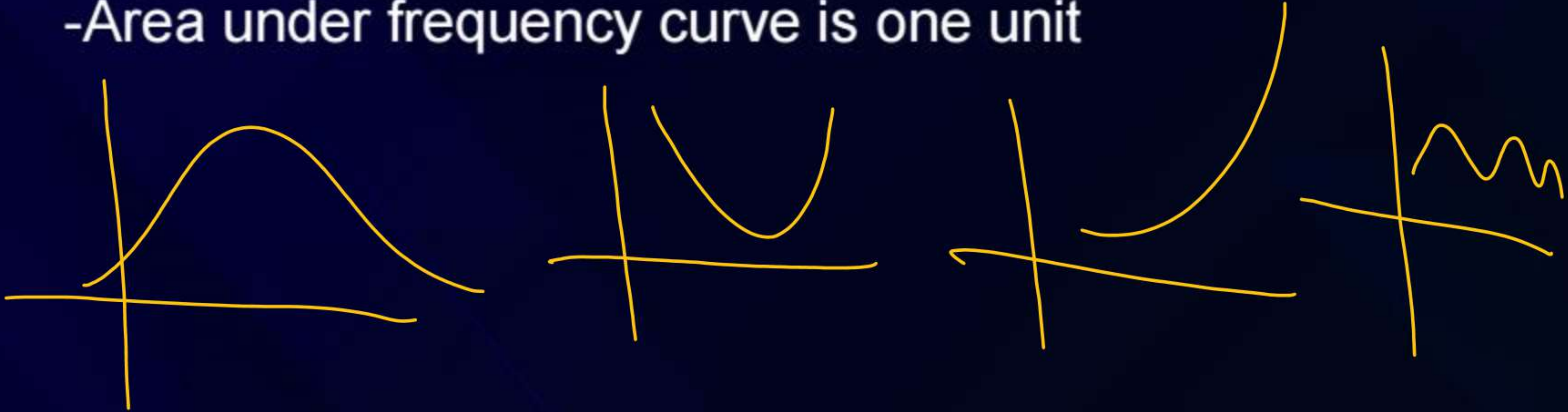
$\frac{N}{2} = 5$



Frequency Curve

-Smooth curve obtained by joining midpoint of upper side of rectangles in histogram

-Area under frequency curve is one unit



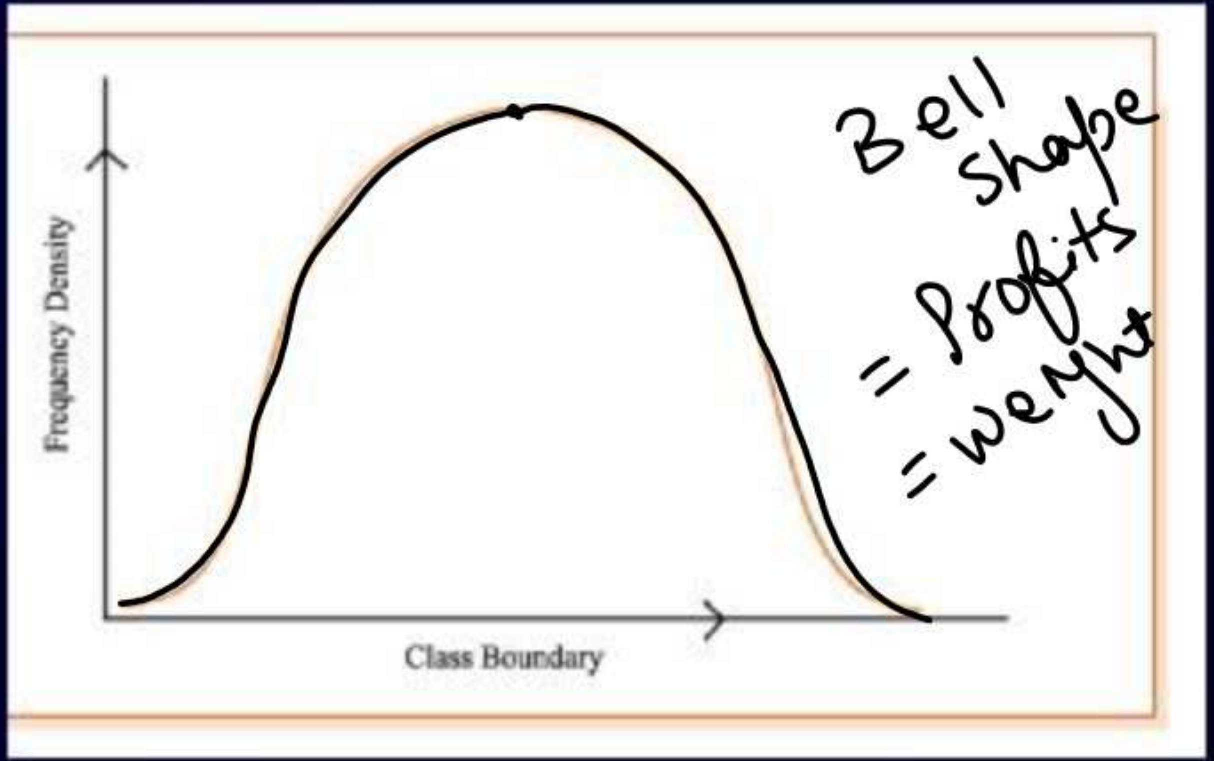
There exist four types of frequency curves namely

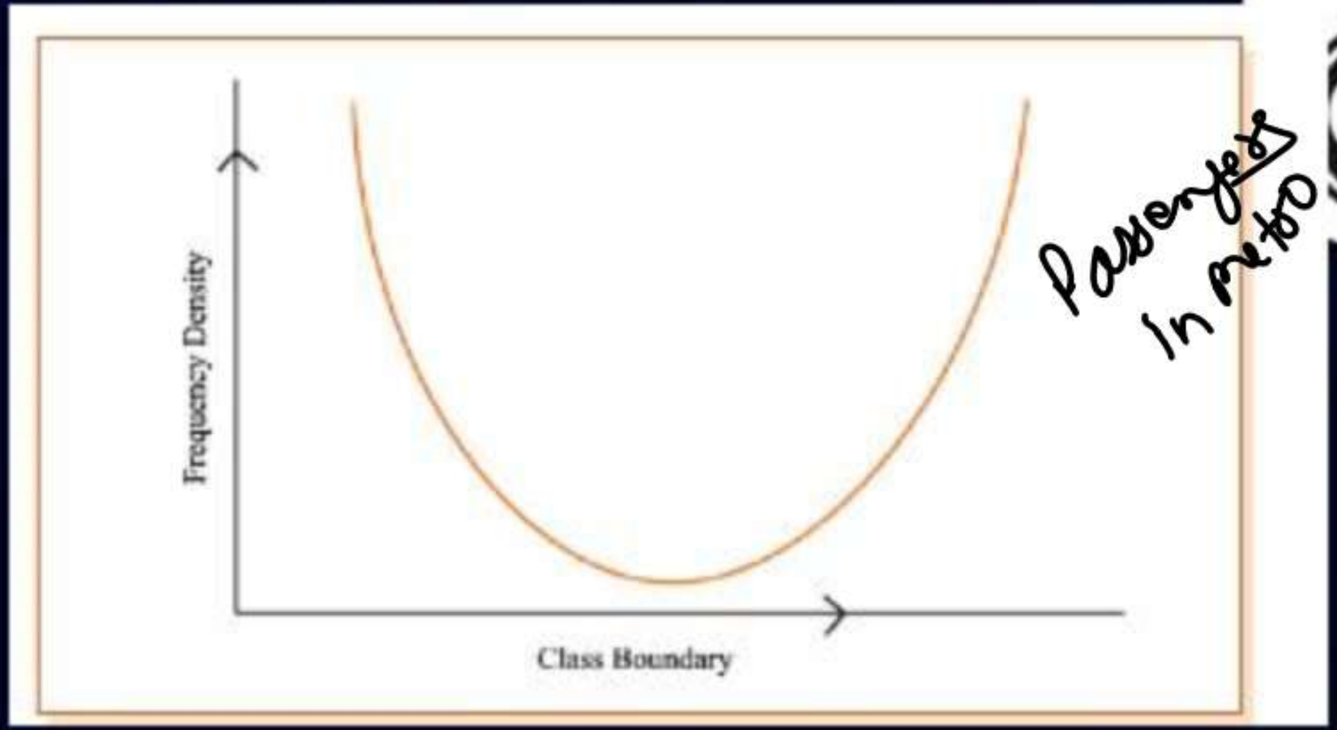
(a) Bell-shaped curve;

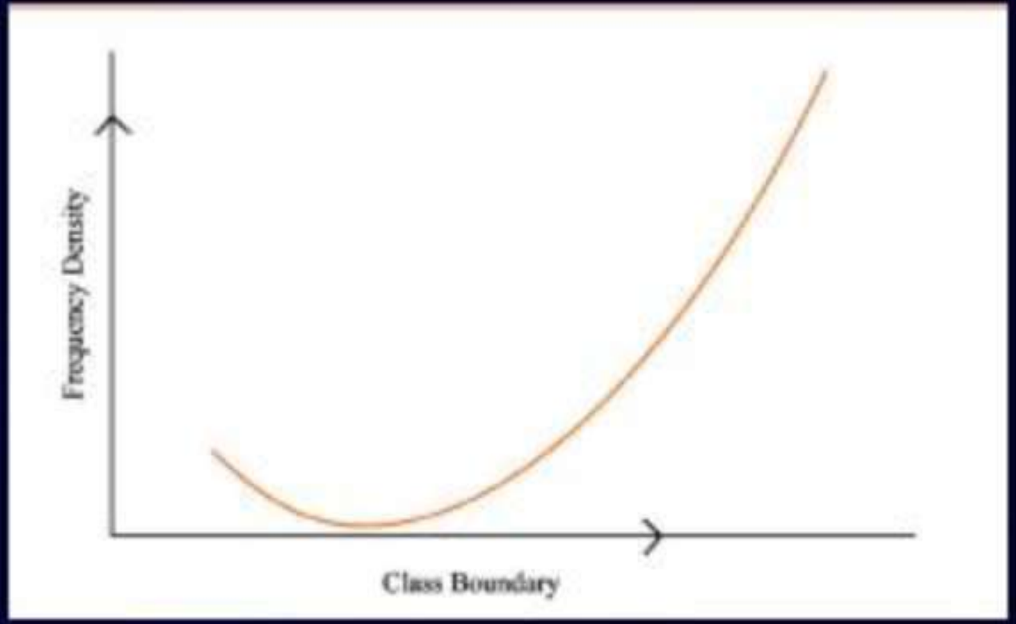
(b) U-shaped curve;

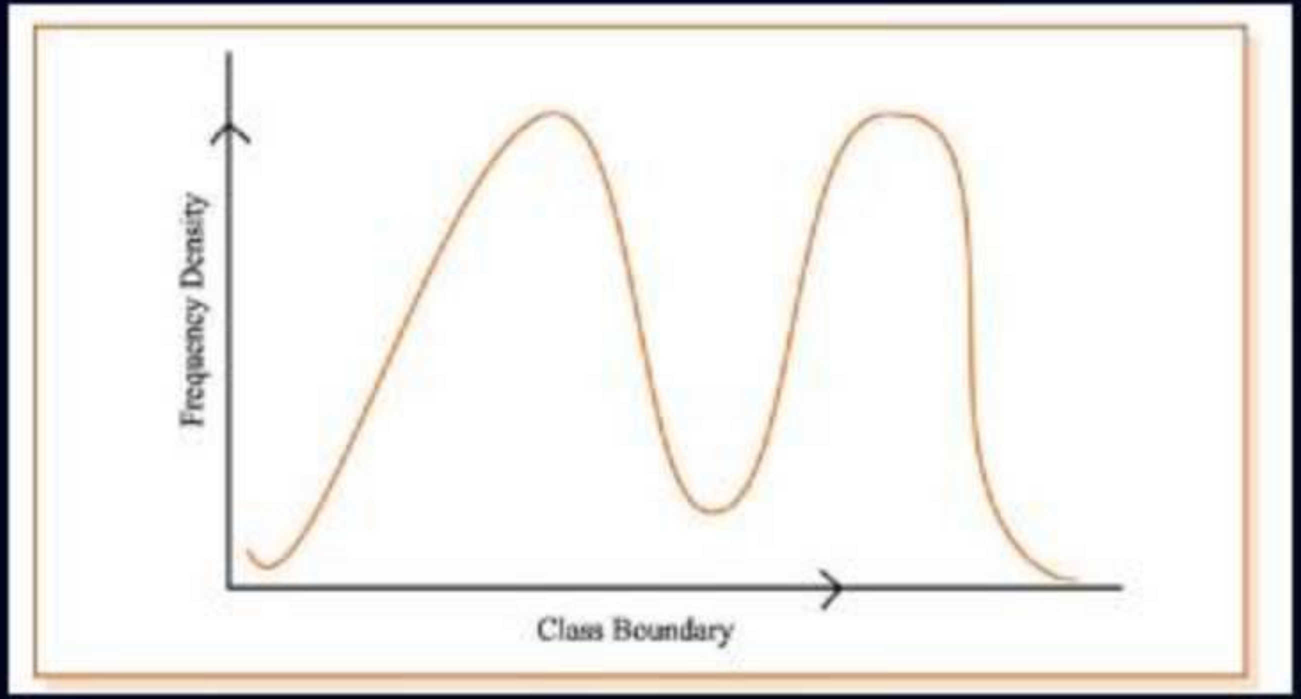
(c) J-shaped curve;

(d) Mixed curve.


Frequency Density
Class Boundary
Bell Shape
= Profits
= weight


Frequency Density
Class Boundary
Passengers
In metro


Frequency Density
Class Boundary


Frequency Density
Class Boundary

QUESTION

The quickest method to collect primary data is

(a) Personal Interview

(b) Indirect Interview

(c) Mailed Questionnaire Method

(d) Telephonic Interview

QUESTION

CA

Which of the following statement is true?

(a) Statistics is derived from the French word ‘Statistik’

(b) Statistics is derived from the Italian word ‘Statista’.

(c) Statistics is derived from the Latin word ‘Statistique’.

(d) None of these

QUESTION

Relative frequency for a particular class lies between :

(a) 0 and 1
(b) 0 and 1, both inclusive
(c) – 1 and 0
(d) – 1 and 1

QUESTION

CA

The graphical representation of a cumulative frequency distribution is called :

(a) Histogram

(b) Ogive

(c) Both

(d) None

QUESTION

Cost of sugar in a month under the heads raw materials, labour, direct production and others were 12, 20, 35 and 23 units respectively. What is the difference between the central angles for the largest and smallest components of the cost of sugar ?

(a) 72^0
(b) 48^0
(c) 56^0
(d) 92^0

12 → $\frac{12}{90} \times 360 = 48$

20

35 → $\frac{35}{90} \times 360 = 140$

23

90

$140 - 48$

$= 92°$

QUESTION

Frequency density corresponding to a class interval is the ratio of :

(a) Class Frequency to the Total Frequency

(b) Class Frequency to the Class Length

(c) Class Length to the Class Frequency

(d) Class Frequency to the Cumulative Frequency

QUESTION

An area diagram is :

(a) Histogram

(b) Ogive

(c) Frequency Polygon

(d) None of these

QUESTION

The distribution of profits of a company follows
(a) J - shaped frequency curve
(b) U - shaped frequency curve
(c) Bell - shaped frequency curve
(d) Any of these

QUESTION

Median of a distribution can be obtained from ;

(a) Histogram (b) Frequency Polygon

(c) Less than type Ogives (d) None of these

QUESTION

The column headings of a table are known as

(a) Body
(b) Stub
(c) Box - head
(d) Caption

QUESTION

From the following data find the number class intervals if class length is given as 5.

73, 72, 65, 41, 54, 80, 50, 46, 49, 53.

(a) 6
(b) 5
(c) 7
(d) 8

(1 mark)

$$\frac{80-41}{5} = \frac{39}{5} = 7.8 \simeq 8$$

41 – 46
46 – 51
51 – 56
56 – 61
61 – 66
66 – 71
71 – 76
76 – 81

QUESTION

The most appropriate diagram to represent the data relating to the monthly expenditure on different items by a family is

(a) Histogram (b) Pie-diagram.

(c) Frequency polygon (d) Line graph. (1 mark)

QUESTION

Which of the following is a statistical data ?

(a) Ram is 50 years old.

(b) Height of Ram is 5'6" and of Shyam and Hari is 5'3" and 5'4" respectively.

(c) Height of Ram is 5'6" and weight is 90kg

(d) Sale of A was more than B and C. (1 mark)

QUESTION

Which of the following is not a two-dimensional figure ?

(a) Line Diagram → 1

(b) Pie Diagram → 2

(c) Square Diagram → 2

(d) Rectangle Diagram → 2

QUESTION

Arrange the dimensions of Bar diagram, Cube diagram, Pie diagram in sequence.

1 3 2

(a) 1, 3, 2 ✓✓

(b) 2, 1, 3

(c) 2, 3, 1

(d) 3, 2, 1

(1 mark)

QUESTION

With the help of histogram one can find.

(a) Mean
(b) Median
(c) Mode
(d) First Quartile

QUESTION

CA

Nationality of a person is :

(a) Discrete variable

(b) An attribute

(c) Continuous variable

(d) None

QUESTION

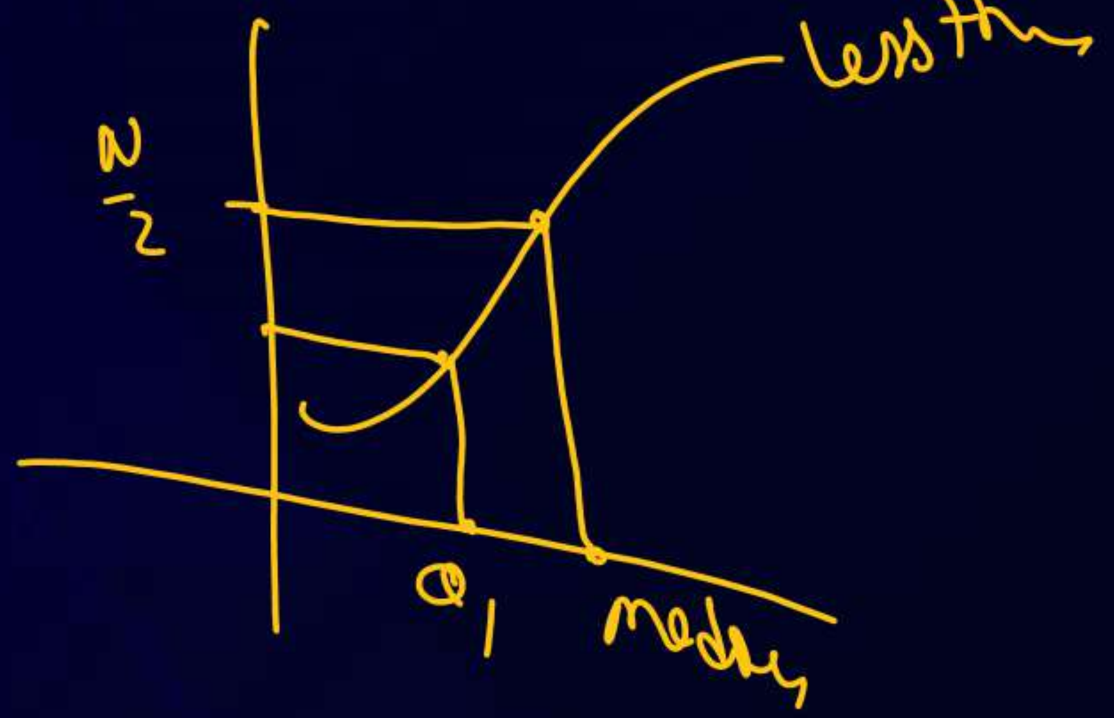
Using Ogive Curve, we can determine

(a) Median (b) Quartile

(c) Both (a) and (b) ✓ (d) None.

$\frac{N}{2}$

$\frac{N}{4}$



QUESTION

Histogram is used for the presentation of the following type of series

(a) Time series

(b) Continuous frequency distribution

(c) Discrete frequency distribution

(d) Individual observation (1 mark)

QUESTION

CA

Classification is of ______ kinds.

(a) Two

(b) Three

(c) One

(d) Four

QUESTION

The chart that uses logarithm of variable is known as:

(a) Ratio chart
(b) Line chart
(c) Multiple line chart
(d) Component line chart

QUESTION

CA

'Stub' of a table is the

(a) Left part of the table describing the columns

(b) Right part of the table describing the columns

(c) Right part of the table describing the rows

(d) Left part of the table describing the rows.

QUESTION

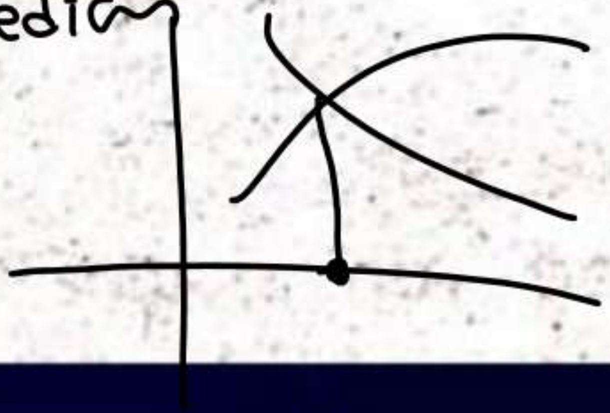
A perpendicular drawn from the point of intersection of two Ogive on the horizontal axis gives the value of:

(a) 2nd Quartile = median

(b) 3rd Quartile

(c) Mode

(d) 1st Quartile



QUESTION

CA

The frequency of visitor in an office is given below:

Time	9 AM-11 AM	11 AM – 1 PM	1 PM- 3 PM	3 PM – 5 PM
Frequency	5	18	7	12

Find the cumulative frequency of visitors for the time 11 AM – 1 PM?

(a) 5
(b) 23
(c) 18
(d) 30

5 23 30 42

Central Tendency & Dispersion

Central Tendency

A central value which represent all the observations
Of a series is called central tendency

→ AM
→ median
→ mode
→ G.M.
→ H.M.

Central Tendency

A central value which represent all the observations
Of a series is called central tendency

Criteria for An ideal Central Tendency

1) Easy calculation
2) Easy to understand
3) Based on all the observations
4) Rigidly defined
5) Have mathematical properties
6) Least affected by extreme values

1 & 2

$\frac{2}{\frac{1}{1}+\frac{1}{2}}$

0 & 2

$\frac{2}{\frac{1}{0}+\frac{1}{2}}$

Arithmetic Mean

Direct method

$$\bar{x} = \frac{\Sigma x_i}{N}$$

$$\bar{x} = \frac{\Sigma f_i x_i}{N}$$

Shortcut method

Assume mean

$$\Rightarrow d_i = x_i - A$$

then

$$\bar{x} = A + \frac{\Sigma f_i d_i}{N}$$

Step Deviation method

$$\Rightarrow u_i = \frac{x_i - A}{h}$$

then

$$\bar{x} = A + \frac{\Sigma f_i u_i}{N} \times h$$

Eg

x_i
1
2
3
4
5

$$\bar{X} = \frac{\Sigma x_i}{N} = \frac{15}{5} = 3$$

Eg

x_i	f_i	f_ix_i
1	3	3
2	2	4
3	4	12
4	1	4
	$\boxed{10}$	23

$$\bar{X} = \frac{\Sigma f_ix_i}{N} = \frac{23}{10} = 2.3$$

Eg

CI	f_i	x_i	f_ix_i
0-2	3	1	3
2-4	8	3	24
4-6	4	5	20
6-8	5	7	35
	$\boxed{N=20}$		82

$$\bar{X} = \frac{\Sigma f_ix_i}{N} = \frac{82}{20} = 4.1$$

Eg

x_i	f_i	$f_i x_i$
6.7	5	33.5
7.7	4	30.8
8.7	3	26.1
9.7	8	77.6
10.7	5	53.5
	$N = 25$	221.5

$$\bar{X} = \frac{\Sigma f_i x_i}{N} = \frac{221.5}{25}$$

$$= 8.86$$

x_i	f_i	$d_i = x_i - 8.7$	$f_i d_i$
6.7	5	−2	−10
7.7	4	−1	−4
$\boxed{8.7}$	3	0	0
9.7	8	1	8
10.7	5	2	10
	$N = 25$		4

$$\bar{X} = A + \frac{\Sigma f_i d_i}{N}$$

$$= 8.7 + \frac{(4)}{25}$$

$$= 8.86$$

Assume mean method

eg

CI	f_i	x_i	f_ix_i
2-5	3	3.5	10.5
5-8	2	6.5	13
8-11	6	9.5	57
11-14	4	12.5	50
$\boxed{N=15}$			130.5

$$\bar{X} = \frac{\Sigma f_ix_i}{N}$$

$$= \frac{130.5}{15}$$

$$= 8.7$$

eg

CI	f_i	x_i	u_i	f_iu_i
2-5	3	3.5	-1	-3
5-8	2	6.5 = A	0	0
8-11	6	9.5	1	6
11-14	4	12.5	2	8
	15			11

$$\bar{X} = A + \frac{\Sigma f_iu_i}{N} \times h$$

$$= 6.5 + \frac{(11)}{15} \times 3$$

$$= 8.7$$

Q

CI	fi
2-5	3
5-8	?
8-11	6
11-14	4

mean = 8.7

find missing frequency.

sol:

x_i	f_i	f_ix_i
3.5	3	10.5
6.5	f	6.5f
9.5	6	57
12.5	4	50
	N = 13+f	117.5+6.5f

$$\bar{x} = \frac{\Sigma f_ix_i}{N}$$

$$8.7 = \frac{117.5+6.5f}{13+f}$$

$$8.7(13+f) = 117.5+6.5f$$

$$113.1+8.7f = 117.5+6.5f$$

$$2.2f = 4.4$$

$$f = \frac{4.4}{2.2}$$

$$\boxed{f=2}$$

Combined mean

Group-A	Group-B
N_1	N_2
$\bar{X}_1$	$\bar{X}_2$

Combined mean $\bar{X}_{12} = \dfrac{N_1\bar{X}_1 + N_2\bar{X}_2}{N_1 + N_2}$

Eg

Girls	Boys
$N_1 = 5$	$N_2 = 15$
$\bar{X}_1 = 25$	$\bar{X}_2 = 30$

find combined mean

Sol: $\bar{X}_{12} = \dfrac{5(25) + 15(30)}{5+15}$

$= \dfrac{575}{20}$

$= 28.75$

Change of origin

$\Downarrow$

If K is added to each item then new mean will be = old mean + K

Change of Scale

$\Downarrow$

If each item of series is multiplied by K

New mean = old mean × K

If $y_i = a + bx_i$

then $\overline{y} = a + b\overline{x}$

	Change of origin	Change of scale
mean	✓	✓
median	✓	✓
mode	✓	✓

Q $\bar{x} = 25$

$y_i = 2x_i + 10$

find $\bar{y} = ?$

Sol:

$$\bar{y} = 2\bar{x} + 10$$
$$= 2(25) + 10$$
$$= 50 + 10$$
$$= 60$$

If each item is 'k' then A.M. is also 'k'.

$\Sigma(x_i - \bar{x}) = 0$

or

sum of deviations from arithmetic mean is always zero

$\Sigma(x_i - A)^2$ will be minimum when $A = \bar{x}$

Sum of Squares of Deviation is minimum when deviations are taken from A.M.

eg

$\Sigma(x_i - m)^2$

$\Sigma(x_i - mode)^2$

$\Sigma(x_i - \bar{x})^2$

minimum

weighted mean

$$\text{weighted mean} = \frac{\sum w_i x_i}{\sum w_i}$$

eg

x_i	w_i	$w_i x_i$
1	3	3
2	5	10
3	2	6
	10	19

$$\text{weighted mean} = \frac{19}{10} = 1.9$$

QUESTION

CA

The mean of 'n' observation is 'x'. If k is added to each observation, then the new mean is.

(a) k

(b) xk

(c) x-k

(d) x+k

(1 mark)

mean $(\bar{x}) = x$

New mean $= x + k$

QUESTION

If there are 3 observations 15, 20, 25 then the sum of deviation of the observations from their AM is

(a) 0

(b) 5

(c) -5

(d) 10

$$\Sigma (x_i - \bar{x}) = 0$$

(1 mark)

QUESTION

CA

Find the mean of the following data

Class interval	10-20	20-30	30-40	40-50	50-60	60-70	70-80
Frequency	9	13	6	4	6	2	3

(a) 23.7
(b) 35.7
(c) 39.7
(d) 43.7

(1 mark)

x_i	f_i	u_i	f_iu_i
15	9	-3	-27
25	13	-2	-26
35	6	-1	-6
A ⟸ 45	4	0	0
55	6	1	6
65	2	2	4
75	3	3	9
	43		-40

$$\bar{x} = A + \frac{\Sigma f_iu_i}{N} \times h$$

$$= 45 + \frac{(-40)}{43} \times 10$$

$$= 35.69$$

QUESTION

CA

If mean of 5 observations x + 1, x + 3, x + 5, x + 7 and x + 9 is given 15, then the value of x will be:

(a) 10

(b) 12

(c) 8

(d) 11

(1 mark)

x_i

$x+1$

$x+3$

$x+5$

$x+7$

$x+9$

$5x+25$

$$\bar{x} = \frac{\Sigma x_i}{N}$$

$$15 = \frac{5x+25}{5}$$

$$75 = 5x+25$$

$$50 = 5x$$

$$x = 10$$

QUESTION

The mean of the first three terms is 17 and mean of next four terms is 21. Calculate the mean of seven terms.

(a) 18.28
(b) 19.78
(c) 19.58
(d) 19.28

(1 mark)

$$\underbrace{1 \quad 2 \quad 3}_{\bar{x}_1 = 17} \quad \underbrace{4 \quad 5 \quad 6 \quad 7}_{\bar{x}_2 = 21}$$

$n_1 = 3 \qquad n_2 = 4$

$$\bar{x}_{12} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1+n_2} = \frac{17(3)+21(4)}{3+4} = 19.28$$

QUESTION

The mean of a set of 20 observations in 18.3. The mean is reduced by 0.6 when a new observation is added to the set. The new observation is:

(a) 17.6

(b) 18.9

(c) 5.7

(d) 24.6

(1 mark)

$N = 20$

$\bar{x} = 18.3$

$N = 21$

New mean $= 18.3 - 0.6$

$= 17.7$

$$\sum^{20} x_i = N\bar{x} = 20 \times 18.3 = 366$$

$$\sum^{21} x_i = 21 \times 17.7 = 371.7$$

$$371.7 - 366 = 5.7$$

QUESTION

A Professor has given assignment to students in a statistics class. A student computer the arithmetic mean and standard deviation for 100 students as 50 and 5 respectively. Later on She paints out the student that he has made mistake in taking one observation as 100 instead of 50. What would be the consent mean if the wrong observation is correct?

(a) 50.5
(b) 49.9
(c) 49.5
(d) 50.1

(1 mark)

$\bar{x} = 50$

$\sigma = 5$

$N = 100$

$\bar{x} = \frac{\Sigma x_i}{N}$

$50 = \frac{\Sigma x_i}{100}$

wrong $\Sigma x_i = 5000$

Correct $\Sigma x_i = 5000 - 100 + 50$

$\Sigma x_i = 4950$

Correct mean

$= \frac{4950}{100}$

$= 49.50$

QUESTION

CA

There are n numbers. When 50 is subtracted from each of these number the sum of the numbers so obtained is - 10. When 46 is subtracted from each of the original n numbers, then the sum of numbers. So obtained is 70. What is the mean of the original n numbers?

(a) 56.8
(b) 25.7
(c) 49.5
(d) 53.8

(1 mark)

$$\Sigma(x_i - 50) = -10 \quad \& \quad \Sigma(x_i - 46) = 70$$

$$\Sigma x_i - 50n = -10 \quad \& \quad \Sigma x_i - 46n = 70$$

$$\Sigma x_i = -10 + 50n \quad \& \quad \Sigma x_i = 46n + 70$$

$$\Sigma(x_i - k)$$

$$\Sigma x_i - kn$$

$$50n - 10 = 46n + 70$$

$$4n = 80$$

$$\boxed{n = 20}$$

NU

$$\Sigma x_i = -10 + 50n$$
$$= -10 + 50(20)$$
$$= -10 + 1000$$
$$= 990$$

$$\bar{x} = \frac{\Sigma x_i}{N}$$
$$= \frac{99\cancel{0}}{2\cancel{0}}$$
$$= 49.5$$

CA WALLAH

Median $\Rightarrow 50\%$

$\frac{1}{2}$ median $\frac{1}{2}$

Individual series or Descrete series

$n = \text{odd}$

$$\text{median} = \left(\frac{n+1}{2}\right)^{th} \text{term}$$

$n = \text{even}$

$$\text{median} = \frac{\left(\frac{n}{2}\right)^{th} + \left(\frac{n}{2}+1\right)^{th}}{2}$$

Continous series

→ find $\frac{N}{2}$

→ Locate $\frac{N}{2}$ in cf

→ select median class

→ $$\text{median} = l + \left\{ \frac{\frac{N}{2} - cf}{f} \right\} \times h$$

QUESTION

For the given data set: 5,10,3,6,4,8,9,3,15,2,9,4,19,11,4, what is the median

(a) 8

(b) 6

(c) 4

(d) 9

(1 mark)

x_i : 2, 3, 3, 4, 4, 4, 5, $\boxed{6}$, 8, 9, 9, 10, 11, 15, 19

$N = 15$ (odd)

$\text{median} = \left(\frac{15+1}{2}\right)^{th} = 8^{th} = 6$

Eg

x_i	f_i	cf
1	3	3 (1-3)
2	2	5 (4-5)
3	5	10 (6-10)
4	4	14 (11-14)
5	2	16 (15-16)
	$N=16$	

find median

Sol: $N = 16$ (even)

1, 1, 1, 2, 2, 3, 3, $\boxed{3, 3}$, 3, 4, 4, 4, 4, 5, 5

$$\text{median} = \frac{\left(\frac{16}{2}\right)^{th} + \left(\frac{16}{2}+1\right)^{th}}{2}$$

$$= \frac{8^{th} + 9^{th}}{2}$$

$$= \frac{3+3}{2} = \frac{6}{2} = 3$$

QUESTION

The Median of the following frequency distribution is:

x	0-10	10-20	20-30	30-40	40-50
f(x)	3	5	20	12	7

(a) 27.75
(b) 9.35
(c) 8.25
(d) 10.01

CI	fi	cf
0-10	3	3
10-20	5	8
20-30	20	28 (8-28)
30-40	12	40
40-50	7	47
	47	

median class

$\frac{N}{2} = \frac{47}{2} = 23.5$

median

$= l + \left\{ \frac{\frac{N}{2} - cf}{f} \right\} \times h$

$= 20 + \left\{ \frac{23.5 - 8}{20} \right\} \times 10$

$= 27.75$

$\boxed{\sum |x_i - A| \text{ will be minimum if } A = \text{median}}$

Sum of absolute deviations taken from median is always minimum

$\sum |x_i - \bar{x}|$ or $\sum |x_i - \text{mode}|$ or $\sum |x_i - \text{median}|$

× ×

$\Downarrow$

Minimum

QUESTION

Mean deviation is minimum when deviations are taken from:

(a) Mean

(b) Median

(c) Mode

(d) Range

QUESTION

If two variable 'x' and 'y' are related as $2x - y = 3$, if the median of 'x' is 10, what is median of 'y'?

(a) 4

(b) 17

(c) 5

(d) 6

$2\bar{x} - \bar{y} = 13$

Med. of $x = 10$

Now $2x - y = 3$

$2x - 3 = y$

$y = 2x - 3$

Med of $y = 2(10) - 3$

$= 17$

Fractiles

Quartiles

$\Downarrow$

Divide the series in 4 parts

Q_1	Q_2	Q_3
25%	50%	75%
$\downarrow$	$\downarrow$	$\downarrow$
$\frac{N}{4}$	$\frac{N}{2}$	$\frac{3N}{4}$
	= median	

Individual Series / Descrete

$$Q_1 = \left(\frac{N+1}{4}\right)^{th} \quad \& \quad Q_3 = \left[3\left(\frac{N+1}{4}\right)\right]^{th}$$

Continous series

→ locate $\frac{N}{4}$ & $\frac{3N}{4}$ in cf

→ Select Q_1 & Q_3 class

→ $$Q_1 = l + \left\{\frac{\frac{N}{4} - cf}{f}\right\} \times h$$

& $$Q_3 = l + \left\{\frac{\frac{3N}{4} - cf}{f}\right\} \times h$$

Eg

Marks

2

8

12

25

30

36

39

50

$N = 8$

Q_1

$= \left(\frac{N+1}{4}\right)^{th}$

$= \left(\frac{8+1}{4}\right)^{th}$

$= (2.25)^{th}$

$= 2^{nd} + 0.25(3^{rd} - 2^{nd})$

$= 8 + 0.25(12 - 8)$

$= 8 + 1$

$= 9$

Q_3

$= \left[3\left(\frac{N+1}{4}\right)\right]^{th}$

$= \left[3\left(\frac{8+1}{4}\right)\right]^{th}$

$= (6.75)^{th}$

$= 6^{th} + 0.75(7^{th} - 6^{th})$

$= 36 + 0.75(39 - 36)$

$= 36 + 2.25$

$= 38.25$

1
1
1
2
2
2
2
2
3
3
3
3
4
4
4
4
4
4

eg

x_i	f_i	cf
1	3	3 (1-3)
2	5	8 (4-8)
3	4	12 (9-12)
4	6	18 (13-18)
5	5	23 (19-23)

$\boxed{N = 23}$

$$Q_1 = \left(\frac{23+1}{4}\right)^{th}$$

$$= 6^{th}$$

$$= 2$$

$$Q_3 = \left[3\left(\frac{23+1}{4}\right)\right]^{th}$$

$$= 18^{th}$$

$$= 4$$

CI	f_i	cf
0-8	2	2
8-16	4	6
16-24	6	12
24-32	2	14
32-40	6	20

Q_1

Q_3

$\boxed{N = 20}$

$\frac{N}{4} = \frac{20}{4} = 5$

$\frac{3N}{4} = \frac{3 \times 20}{4} = 15$

$Q_1 = l + \left\{ \frac{\frac{N}{4} - cf}{f} \right\} \times h$

$= 8 + \left\{ \frac{5-2}{4} \right\} \times 8$

$= 8 + 6$

$= 14$

$Q_3 = l + \left\{ \frac{\frac{3N}{4} - cf}{f} \right\} \times h$

$= 32 + \left\{ \frac{15-14}{6} \right\} \times 8$

$= 33.33$

Deciles

D1 D2 D3 D4 D5 D6 D7 D8 D9

$D_1 \downarrow \frac{N}{10} \downarrow 10\%$

$D_2 \downarrow \frac{2N}{10} \downarrow 20\%$

$D_3 \downarrow \frac{3N}{10} \downarrow 30\%$

$\boxed{D_5} \downarrow \frac{N}{2} \downarrow 50\%$

median

Individual series or Descrete series

$$D_1 = \left(\frac{N+1}{10}\right)^{th} \qquad D_6 = \left[6\left(\frac{N+1}{10}\right)\right]^{th}$$

$$D_3 = 3\left(\frac{N+1}{10}\right) \qquad D_9 = \left[9\left(\frac{N+1}{10}\right)\right]^{th}$$

Continous

for D_1

→ locate $\frac{N}{10}$ in cf

→ Select D_1 class

→ $D_1 = l + \left\{\frac{\frac{N}{10} - cf}{f}\right\} \times h$

for D_7

→ locate $\frac{7N}{10}$ in cf

→ $D_7 = l + \left\{\frac{\frac{7N}{10} - cf}{f}\right\} \times h$

Percentiles

$\Downarrow$

Divides series in 100 parts

$\Downarrow$

50%

$\Downarrow$

median

Individual | Descrete series

$$P_1 = \left(\frac{N+1}{100}\right)^{th} , \quad P_{62} = 62\left(\frac{N+1}{100}\right)^{th}$$

Continous serey

for P_3

$\rightarrow$ locate $\frac{3N}{100}$ in cf

$$\rightarrow P_3 = l + \left\{ \frac{\frac{3N}{100} - cf}{f} \right\} \times h$$

QUESTION

The 3rd decile for the numbers 15, 10, 20, 25, 18, 11, 9, 12 is

(a) 13

(b) 10.70

(c) 11.00

(d) 11.50

9, 10, 11, 12, 15, 18, 20, 25

$N = 8$

$$D_3 = \left[3\left(\frac{N+1}{10}\right)\right]^{th}$$

$$D_3 = 3\left(\frac{8+1}{10}\right)^{th}$$

$$= (2.7)^{th}$$

$$= 2^{nd} + 0.7(3^{rd} - 2^{nd})$$

$$= 10 + 0.7(11 - 10)$$

$$= 10 + 0.7$$

$$= 10.7$$

Mode

Observation with Highest frequency

Eg marks

2, 3, 1, 5, 2, 1, 3, 2, 5, 2, 1

1, 2, 1, 2, 4, 2, 5, 3, 2

mode = 2

Eg

marks x_i	No of students (f_i)
1	4
2	6
3	2
4	5
5	1

mode = 2

CI	f_i
0 – 4	3
4 – 8	4 $\rightarrow f_0$
$\boxed{8 - 12}$	7 $\rightarrow f_1$
12 – 16	5 $\rightarrow f_2$
16 – 20	1

modal class

$$\text{mode} = l + \left\{\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right\} \times h$$

$$= 8 + \left\{\frac{7-4}{2(7)-4-5}\right\} \times 4$$

$$= 8 + \left\{\frac{3}{5}\right\} \times 4$$

$$= 10.4$$

3 median = mode + 2 mean

QUESTION

Which of the following measures of central tendency cannot be calculated by graphical method?

(a) Mean (b) Mode

(c) Median (d) Quartile (1 mark)

Ogive
→ median
→ Quarti
→ Decile
→ Percentile

Histogram
⇓
mode

QUESTION

Relation between mean, median and mode is

(a) mean-mode = 2 (mean-median)

(b) mean-median = 3(mean-mode)

(c) mean-median = 2 (mean-mode)

(d) mean-mode = 3(mean-median)

3 median = mode + 2 mean

QUESTION

One hundred participants expressed their opinion on recommending a new product to their friends using the attributes : most unlikely, not sure, likely, most likely. The appropriate measure of central tendency that can be used here is

(a) Mean

(b) Mode

(c) Geometric mean

(d) Harmonic mean

mU
NS
mL

(1 mark)

QUESTION

If the mean and median of a moderately asymmetrical series are 26.8 and 27.9 respectively, then the most probable mode is:

(a) 35.4
(b) 30.1
(c) 34.3
(d) 70.8

$\bar{X} = 26.8$

$\text{med} = 27.9$

$\text{mode} = ?$

$3 \text{ median} = \text{mode} + 2 \text{ mean}$

$3(27.9) = \text{mode} + 2(26.8)$

$83.7 - 53.6 = \text{mode}$

$\text{mode} = 30.1$

Geometric Mean

$$G.M. = [x_1 \times x_2 \times x_3 \times \dots \times x_n]^{1/n}$$

Eg GM of 2 & 8

$= (2 \times 8)^{1/2}$

$= (16)^{1/2}$

$= 4$

Eg GM of 2, 10 & 25

$= (2 \times 10 \times 25)^{1/3}$

$= (500)^{1/3}$

$= 7.945$

$(x)^{1/n} = ?$

→ $\sqrt{\ }$ 12 times

→ -1

→ $\div n$

→ $+1$

→ $\boxed{x=}$ 12 times

eg

x_i	f_i
2	4
3	2
4	3
	9

GM

$$= (2\times2\times2\times2\times3\times3\times4\times4\times4)^{1/9}$$

eg

CI	f_i	x_i
0-4	2	2
4-8	1	6
8-12	3	10

$$GM = (2\times2\times6\times10\times10\times10)^{1/6}$$

⇛ If all observation are same (let k) then GM is also k

Eg GM of 2, 2, 2, 2, 2 is also 2

$\log(GM) = \dfrac{\sum \log x_i}{N}$

$GM = AL\left(\dfrac{\sum \log x_i}{N}\right)$

$GM(x \cdot y) = GM \text{ of } x \times GM \text{ of } y$

$GM\left(\dfrac{x}{y}\right) = \dfrac{G.M \text{ of } x}{GM \text{ of } y}$

QUESTION

The geometric mean of 3,7,11,15,24,28,30,0 is:

(a) 6

(b) 0

(c) 9

(d) 12

$$\left(3\times 7\times 11\times 15\times 24\times 28\times 30\times 0\right)^{1/8}$$

$$= (0)^{1/8} = 0$$

Harmonic Mean

"Reciprocal of Average of the reciprocal of all observations"

$$HM = \frac{1}{\left(\dfrac{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \cdots + \frac{1}{x_n}}{n}\right)}$$

$$Hm = \frac{N}{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \dots + \frac{1}{x_n}}$$

or

$$Hm = \frac{N}{\Sigma\left(\frac{1}{x_i}\right)}$$

$$\boxed{Hm = \frac{N}{\Sigma\left(\frac{f_i}{x_i}\right)}}$$

Eg Hm of 2 & 8

$$= \frac{2}{\frac{1}{2} + \frac{1}{8}}$$

$$= \frac{2}{\frac{4+1}{8}}$$

$$= \frac{16}{5}$$

$$= 3.2$$

Eg HM of 2, 5 & 10

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

Eg

x_i	f_i	$\frac{f_i}{x_i}$
1	4	4
2	2	1
3	6	2
	12	7..

$$HM = \frac{N}{\Sigma\left(\frac{f_i}{x_i}\right)}$$

$$= \frac{12}{7}$$

$$= 1.7142$$

Weighted Harmonic mean

$$= \frac{\Sigma w}{\Sigma\left(\frac{w_i}{x_i}\right)}$$

Combined HM

$$= \frac{N_1 + N_2}{\frac{N_1}{H_1} + \frac{N_2}{H_2}}$$

$AM \geq GM \geq HM$

If all the items are same then $AM = GM = HM$

If all the items are different then $AM > GM > HM$

For any two numbers

$$HM \times AM = (GM)^2$$

QUESTION

A man travels at a speed of 20 km/hr and then returns at a speed of 30 km/ hr. His average speed of the whole journey is :

(a) 25 km/ hr (b) 24.5 km/hr

(c) 24 km/hr (d) None (1 mark)

$$\xrightarrow{\quad 20\,km/h \quad}$$

$$\xleftarrow[\quad 30\,km/h \quad]{}$$

Avg speed

$$= HM \text{ of } 20 \,\&\, 30$$

$$= \frac{2}{\frac{1}{20}+\frac{1}{30}}$$

$$= \frac{2}{\frac{30+20}{600}} = \frac{120\cancel{0}}{5\cancel{0}} = 24$$

QUESTION

Given the weights for the numbers 1,2,3.....n are respectively $1^2, 2^2, 3^2 \ldots n^2$ then weighted HM is ________.

(a) $\frac{2n+1}{4} = \frac{7}{4}$

(b) $\frac{2n+1}{6} = \frac{7}{6}$

(c) $\frac{2n+1}{3} = \frac{7}{3}$

(d) $\frac{2n+1}{2} = \frac{7}{2}$

(1 mark)

x_i	w_i	$\frac{w_i}{x_i}$
1	1	1
2	4	2
3	9	3
	14	6

Weighted HM

$$= \frac{\Sigma w}{\Sigma\left(\frac{w_i}{x_i}\right)}$$

$$= \frac{14}{6} = \frac{7}{3}$$

QUESTION

The harmonic mean A and B is 1/3 and harmonic mean of C and D is 1/5. The harmonic mean of ABCD is

(a) 8/15

(b) 1/4

(c) 1/15

(d) 5/3

(1 mark)

A & B	C & D	Combined HM
$H_1 = \frac{1}{3}$	$H_2 = \frac{1}{5}$	$= \dfrac{N_1 + N_2}{\frac{N_1}{H_1} + \frac{N_2}{H_2}}$
$N_1 = 2$	$N_2 = 2$	$= \dfrac{2+2}{\left(\frac{2}{1/3}\right) + \left(\frac{2}{1/5}\right)} = \dfrac{4}{6+10} = \dfrac{4}{16} = \dfrac{1}{4}$

QUESTION

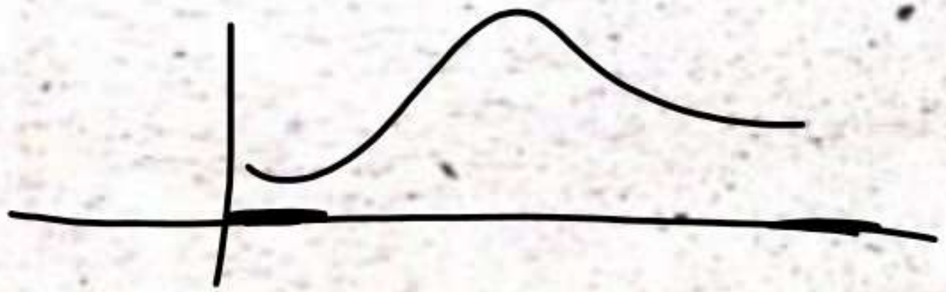
Which one of these is least affected by extreme values?

(a) Mean

(b) Median

(c) Mode

(d) None



(1 mark)

QUESTION

From the record on sizes of shoes sold in a shop, one can compute the following to determine the most preferred shoe size.

(a) Mean
(b) Median
(c) Mode
(d) Range

(1 mark)

QUESTION

If A.M. and G.M. of two positive numbers a and b are 12 and 12, respectively, find the numbers.

(a) 18 and 6

(b) 15 and 9

(c) 16 and 8

(d) 12 and 12

(1 mark)

AM = 12

GM = 12

a = b = 12

QUESTION

Which measure is suitable for open-end classification?

(a) Median

(b) Mean

(c) Mode

(d) GM

CI	fi
less than 10	1
10-12	6
12-14	7
14-16	8
more than 16	2

QUESTION

The AM and HM of two numbers are 5 and 3.2 respectively, then GM will be:

(a) 4.4

(b) 4.2

(c) 4.0

(d) 3.8

(1 mark)

$AM = 5$

$HM = 3.2$

$GM = ?$

$GM^2 = AM \times HM$

$= 5 \times 3.2$

$GM^2 = 16$

$GM = 4$

QUESTION

Which of the following measure does not possess mathematical properties?

(a) Arithmetic mean

(b) Geometric mean

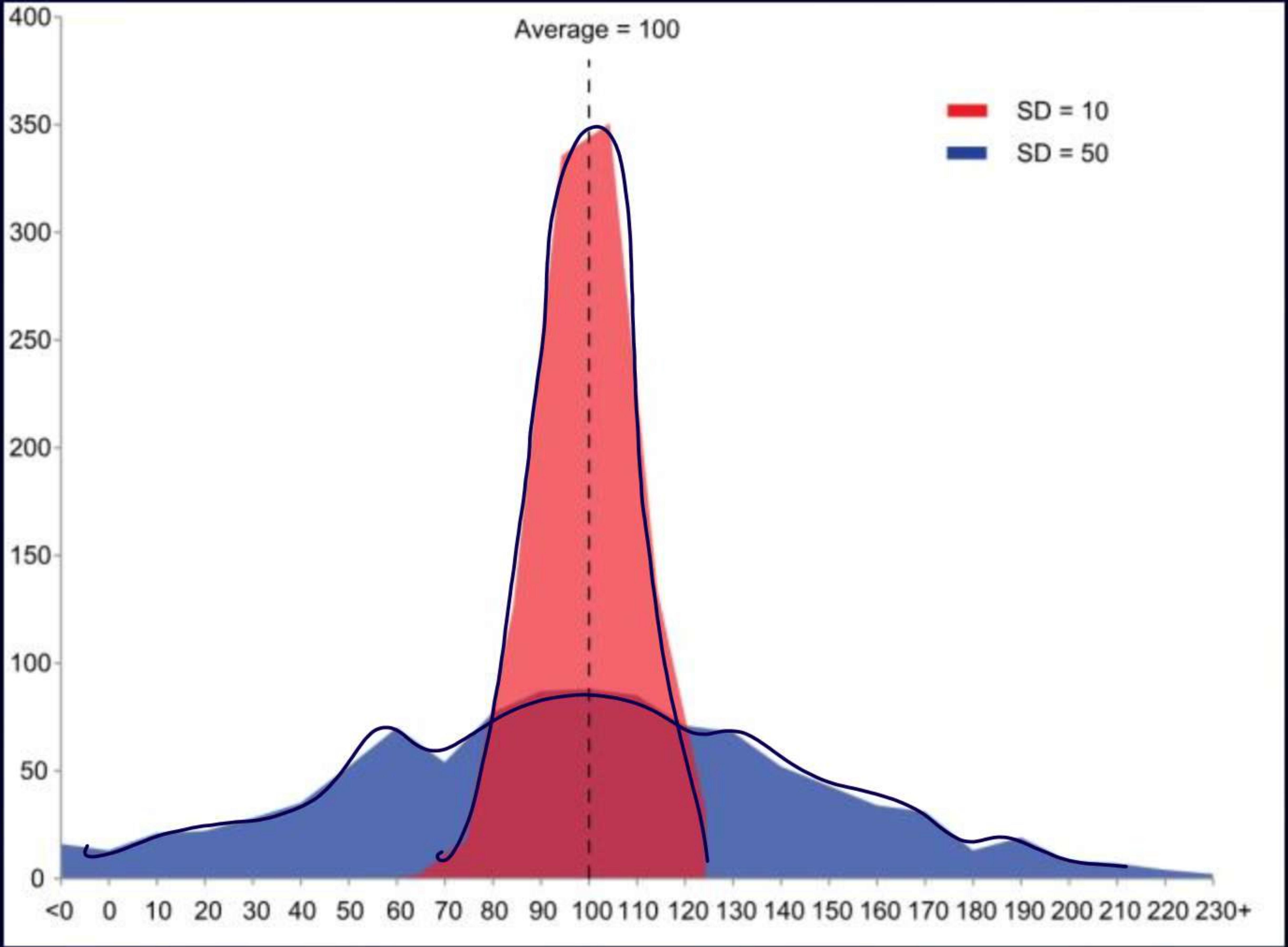
(c) Harmonic mean

(d) Median

(1 mark)

Dispersion

"The Degree Of the **scatterness** or **spread** or **variation** of the variable about a **central value** is called **dispersion**"


Average = 100
SD = 10
SD = 50
400
350
300
250
200
150
100
50
0
<0 0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230+

Measures of Dispersion

Absolute measurement

⇓

→ Range

→ Q.D.

→ M.D.

→ S.D.

Relative measurement

⇓

→ cofficient of Range

→ coff of Q.D

→ coff of MD

→ coeff of variance

RANGE

$$\text{Range} = \text{Largest} - \text{Smallest}$$

$$\text{Range} = L - S$$

$$\text{Coff. of Rang} = \left(\frac{L - S}{L + S}\right) \times 100$$

Eg marks: 2, 5, 7, 8, 9

$$\text{Range} = 9 - 2 = 7$$

$$\text{Coeff of Range} = \frac{9-2}{9+2} \times 100$$

$$= 63.63\%$$

Eg

x_i	f_i
S ← 1	6
2	3
3	5
4	2
L ← 5	8

Range = 5 − 1 = 4

$$\text{Coff of Range} = \frac{5-1}{5+1} \times 100$$

$$= 66.67\%$$

Eg

CI	f_i
2 − 4	3
4 − 6	5
6 − 8	1
8 − 10	3

S L

Range = 10 − 2 = 8

$$\text{Coff of Range} = \frac{10-2}{10+2} \times 100$$

Range, Q.D, MD & SD

→ Change of origin ⟹ No impact

Eg Range of $x_i = 5$

$y_i = x_i + 10$

Range of $y_i = ?$

Sol: Range of y_i = Range of x_i

Eg

x_i	$y_i = x_i + 10$
1	11
2	12
3	13
6	16
Range $= 6 - 1 = 5$	Range $= 16 - 11 = 5$

Range, QD, MD, SD

→ Change of scale

if all the items are multiplied by any number 'k'

then new Range $= |k| \times$ old Range

New SD $= |k| \times$ old SD

New MD $= |k| \times$ old MD

New QD $= |k| \times$ old QD.

Eg Range of $x_i = 5$

$y = 2x_i$

Range of y_i

$= 2 \times$ Range of x_i

$= 2 \times 5$

$= 10$

Eg Range of $x_i = 5$

$y_i = -2x_i$

Range of y_i

$= |-2| \times$ Range of x_i

$= 2 \times 5$

$= 10$

QUESTION

CA

If the range of a data is 20 and its smallest value is 5, then what is the largest value of data is?

(a) 20
(b) 25
(c) 5
(d) 30

(1 mark)

Range = 20

$L - S = 20$

$L - 5 = 20$

$L = 25$

QUESTION

CA

If the relationship between x and y is given by $2x + 3y = 10$ and the range of y is 10, then what is the range of x?

(a) 10

(b) 18

(c) 8

(d) 15

(1 mark)

$$2x + 3y = 10$$

find value of x in terms of y

$$2x = -3y + 10$$

$$x = \frac{-3}{2}y + \frac{10}{2}$$

Range of $y = 10$

Range of $x = \left|\frac{-3}{2}\right| \times$ Range of y

$$= \frac{3}{2} \times 10$$

$$= 15$$

Quartile Deviation

Q_1	Q_2	Q_3
$\frac{N}{4}$	$\frac{2N}{4}$	$\frac{3N}{4}$

Quartile Range $= Q_3 - Q_1$

Quartile Deviation $= \dfrac{Q_3 - Q_1}{2}$

(Semi Quartile Range)

Coeff of Q.D $= \left(\dfrac{Q_3 - Q_1}{Q_3 + Q_1}\right) \times 100$

If Distribution is symmetric

$$\frac{Q_1 + Q_3}{2} = Q_2 = \text{median}$$

QUESTION

If the quartile deviation is 12 and the first quartile is 25, then the value of the third quartile is:

(a) 37

(b) 49

(c) 61

(d) 60

(1 mark)

$QD = 12$ & $Q_1 = 25$, $Q_3 = ?$

$$\frac{Q_3 - Q_1}{2} = 12$$

$$\frac{Q_3 - 25}{2} = 12 \Rightarrow Q_3 = 49$$

QUESTION

CA

If 'x' and 'y' are related as $3x - 4y = 20$ and the quartile deviation of 'x' is 12, then the quartile deviation of 'y' is:

(a) 9
(b) 8
(c) 7
(d) 6

Sol:

$3x-4y=20$

QD of $x=12$

QD of $y=?$

$3x-4y=20$

$-4y=-3x+20$

$y=\frac{-3x}{-4}+\frac{20}{-4}$

$y=\frac{3}{4}x-5$

Now

QD of $y=\frac{3}{4}\times$ QD of x

$=\frac{3}{4}\times 12$

$=9$

Mean Deviation

MD about mean

MD about median

$$MD_{\bar{x}} = \frac{\Sigma f_i |x_i - \bar{x}|}{N}$$

$$MD_m = \frac{\Sigma f_i |x_i - m|}{N}$$

$$\text{coff of } MD = \frac{MD_{\bar{x}}}{\bar{x}} \times 100$$

$$\text{coff of } MD = \frac{MD_m}{m} \times 100$$

QUESTION

Find mean deviation about mean for the data
12,16,21,30,35,39,40

(a) 9.14
(b) 9.63
(c) 8.91
(d) 9.81

x_i	$\lvert x_i - \bar{x} \rvert$
12	15.57
16	11.57
21	6.57
30	2.43
35	7.43
39	11.43
40	12.43
193	67.43

$$\bar{x} = \frac{\Sigma x_i}{N} = \frac{193}{7} = 27.57$$

$$MD = \frac{\Sigma \lvert x_i - \bar{x} \rvert}{N}$$

$$= \frac{67.43}{7}$$

$$= 9.63$$

Q x_i : 2, 6, 8, 12, 15, 17

find MD about median

Sol:

x_i	$\lvert x_i - 10 \rvert$
2	8
6	4
8	2
12	2
15	5
17	7
	28

$N = 6$ (even)

Median

$$= \frac{\left(\frac{6}{2}\right)^{th} + \left(\frac{6}{2}+1\right)^{th}}{2}$$

$$= \frac{3^{rd} + 4^{th}}{2}$$

$$= \frac{8+12}{2} = 10$$

$$MD_m = \frac{\Sigma \lvert x_i - m \rvert}{N}$$

$$= \frac{28}{6}$$

$$= 4.6666$$

Standard Deviation (σ —)

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

$$SD(\sigma) = \sqrt{\frac{\Sigma f_i x_i^2}{N} - \left(\frac{\Sigma f_i x_i}{N}\right)^2}$$

or

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

$$d_i = x_i - A$$

$$SD = \sqrt{\frac{\Sigma f_i d_i^2}{N} - \left(\frac{\Sigma f_i d_i}{N}\right)^2}$$

$$u_i = \frac{x_i - A}{h}$$

$$SD = \sqrt{\frac{\Sigma f_i u_i^2}{N} - \left(\frac{\Sigma f_i u_i}{N}\right)^2} \times h$$

x_i	x_i^2
2	4
4	16
6	36
12	56

$$\text{mean} = \frac{\Sigma x_i}{N} = \frac{12}{3} = 4$$

$$SD = \sqrt{\frac{\Sigma x_i^2}{N} - \left(\frac{\Sigma x_i}{N}\right)^2}$$

$$= \sqrt{\frac{56}{3} - \left(\frac{12}{3}\right)^2}$$

$$= 1.63299$$

$$SD = \sqrt{\frac{n^2-1}{3}}$$

$$= \sqrt{\frac{3^2-1}{3}}$$

$$= 1.63299$$

g

x_i	$d_i = x_i - 0.23$	d_i^2
1.23	1	1
2.23	2	4
3.23	3	9
	6	$\boxed{14}$

$$SD = \sqrt{\frac{\Sigma d_i^2}{N} - \left(\frac{\Sigma d_i}{N}\right)^2}$$

$$= \sqrt{\frac{14}{3} - \left(\frac{6}{3}\right)^2} = 0.8164$$

$$SD = \sqrt{\frac{n^2-1}{12}}$$

$$= \sqrt{\frac{3^2-1}{12}}$$

$$= 0.8164$$

Q

CI	f_i	x_i	$f_i x_i$	$f_i x_i^2$
0-4	3	2	6	12
4-8	2	6	12	72
8-12	1	10	10	100
12-16	4	14	56	784
	10		84	968

$$SD = \sqrt{\frac{\Sigma f_i x_i^2}{N} - \left(\frac{\Sigma f_i x_i}{N}\right)^2}$$

$$= \sqrt{\frac{968}{10} - \left(\frac{84}{10}\right)^2} = 5.1224$$

CI	f_i	x_i	u	$f_i u_i$	$f_i u_i^2$
0-4	3	2	-1	-3	3
4-8	2	6 = A	0	0	0
8-12	1	10	1	1	1
12-16	4	14	2	8	16
	$\boxed{10}$			$\boxed{6}$	20

$$SD = \sqrt{\frac{\Sigma f_i u_i^2}{N} - \left(\frac{\Sigma f_i u_i}{N}\right)^2} \times h$$

$$= \sqrt{\frac{20}{10} - \left(\frac{6}{10}\right)^2} \times 4$$

$$= 1.28 \times 4$$

$$= 5.1224$$

$$\text{Variance} = (S.D.)^2$$

$$S.D. = \sqrt{\text{Variance}}$$

coefficient of variance

$$= \frac{S.D.}{\text{mean}} \times 100$$

more C.V.
more variability
& less consistency

less C.V.
→ less variability
→ more consistency

$$\boxed{QD : MD : SD = 10 : 12 : 15}$$

$$\frac{QD}{MD} = \frac{10}{12} \quad \Bigg| \quad \frac{MD}{SD} = \frac{12}{15} \quad \Bigg| \quad \frac{QD}{SD} = \frac{10}{15}$$

S.D. of two numbers a & b = $\frac{|a-b|}{2}$

Eg SD of 2 & 7 = $\left|\frac{2-7}{2}\right| = 2.5$

Eg

x_i	x_i^2
2	4
7	49
$\boxed{9}$	53

$$SD = \sqrt{\frac{53}{2} - \left(\frac{9}{2}\right)^2} = 2.5$$

S.D. of first 'n' natural no $= \sqrt{\frac{n^2-1}{12}}$

S.D of first 'n' even natural no or odd natural no $= \sqrt{\frac{n^2-1}{3}}$

Q

x_i	x_i^2
1	1
3	9
5	25
9	35

$$SD = \sqrt{\frac{35}{3} - \left(\frac{9}{3}\right)^2}$$

$$= 1.63299$$

if all the observations are same

then Range = 0

MD = 0

QD = 0

SD = 0

Combined S.D.

$$= \sqrt{\frac{N_1(\sigma_1^2 + d_1^2) + N_2(\sigma_2^2 + d_2^2)}{N_1 + N_2}}$$

where

$d_1 = \overline{X}_{12} - \overline{X}_1$

$d_2 = \overline{X}_{12} - \overline{X}_2$

Eg

$N_1 = 5$	$N_2 = 5$
$\sigma_1 = 2$	$\sigma_2 = 1$
$\bar{X}_1 = 25$	$\bar{X}_2 = 35$

combined SD = ?

Sol:

$$\bar{X}_{12} = \frac{N_1\bar{X}_1 + N_2\bar{X}_2}{N_1 + N_2}$$

$$= \frac{5(25) + 5(35)}{5+5}$$

$$\bar{X}_{12} = \frac{30\cancel{0}}{1\cancel{0}} = 30$$

Now

$d_1 = \bar{X}_{12} - \bar{X}_1$

$= 30 - 25$

$= 5$

$d_2 = \bar{X}_{12} - \bar{X}_2$

$= 30 - 35 = -5$

Now

$$\sigma_{12} = \sqrt{\frac{N_1(\sigma_1^2 + d_1^2) + N_2(\sigma_2^2 + d_2^2)}{N_1 + N_2}}$$

$$= \sqrt{\frac{5(4+25) + 5(1+25)}{5+5}}$$

$$= \sqrt{\frac{275}{10}}$$

$$= 5.24$$

QUESTION

If the sum of square of the values equals to 3390, Number of observations are 30 and Standard deviation is 7, what is the mean value of the above observations?

(a) 14

(b) 11

(c) 8

(d) 5

(1 mark)

$\Sigma x_i^2 = 3390$

$N = 30$

$\sigma = 7$

$\bar{x} = ?$

$\sigma = \sqrt{\frac{\Sigma x_i^2}{N} - \left(\frac{\Sigma x_i}{N}\right)^2}$

$7 = \sqrt{\frac{339\cancel{0}}{3\cancel{0}} - (\bar{x})^2}$

$49 = 113 - (\bar{x})^2$

$(\bar{x})^2 = 64$

$\boxed{\bar{x} = 8}$

QUESTION

CA

If the Standard Deviation of data 2,4,5,6,8,17 is 4.47, then Standard Deviation of the data 4,8,10,12,16,34 is

(a) 4.47

(b) 8.94

(c) 13.41

(d) 2.24

(1 mark)

x_i	$2x_i$
2	4
4	8
5	10
6	12
8	16
17	34

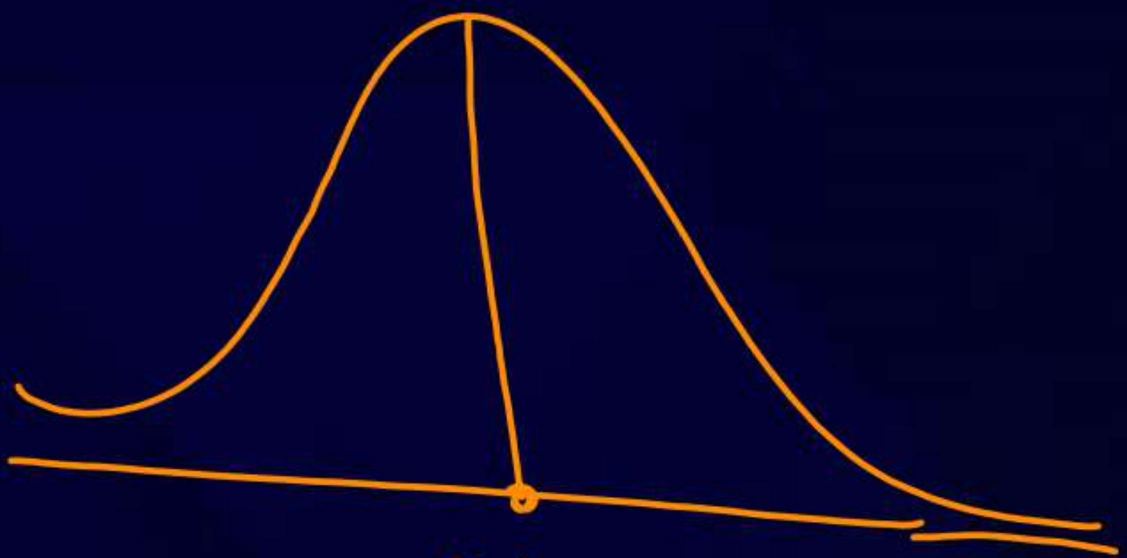
$SD = 4.47$ $\quad$ $SD = 2 \times 4.47 = 8.94$

QUESTION

For the given set of normally distributed data, the following statistical data are known: Mean = 6; Standard Deviation = 2.6; Median = 5 and Q deviation =1.5, then the coefficient of quartile deviation equals to

(a) 30 (b) 32

(c) 25 (d) 39 (1 mark)



Mean = median = mode

Skewness = zero

$\bar{x} = 6$

$\sigma = 2.6$

median $= 5$

$QD = 1.5$

Coeff of $QD = ?$

Coeff of QD

$= \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$

$= \frac{\frac{Q_3 - Q_1}{2}}{\frac{Q_3 + Q_1}{2}} \times 100$

$= \frac{QD}{\text{Median}} \times 100$

$= \frac{1.5}{5} \times 100$

$= 30$

QUESTION

If two samples of sizes 30 and 20 have means as 55 and 60 and variances as 16 and 25 respectively, then what would be the S. D. of the combined sample size 50 ?

(a) 5.33 (b) 5.17

(c) 5.06 (d) 5 (1 mark)

$N_1 = 30$ $N_2 = 20$

$\bar{X}_1 = 55$ $\bar{X}_2 = 60$

$\sigma_1^2 = 16$ $\sigma_2^2 = 25$

$\sigma_1 = 4$ $\sigma_2 = 5$

Combined SD = ?

Coefficient of Variance

$$CV = \frac{SD}{mean} \times 100$$

QUESTION

If the coefficient of variation and standard deviation are 30 and 12 respectively, then the arithmetic mean of the distribution is:

(a) 40 (b) 36

(c) 25 (d) 19 (1 mark)

$CV = 30$

$\sigma = 12$

$\bar{x} = ?$

$CV = \frac{\sigma}{\bar{x}} \times 100$

$30 = \frac{12}{\bar{x}} \times 100$

$\bar{x} = 40$

QUESTION

The probable value of mean deviation when $Q_3 = 40$ and $Q_1 = 15$ is:

(a) 15

(b) 18.75

(c) 17.50

(d) 0

(1 mark)

$MD = ?$

$Q_3 = 40, Q_1 = 15$

$$QD = \frac{Q_3 - Q_1}{2}$$

$$= \frac{40-15}{2}$$

$$= 12.5$$

$QD : MD : SD = 10 : 12 : 15$

$$\frac{QD}{MD} = \frac{10}{12}$$

$$\frac{12.5}{MD} = \frac{10}{12}$$

$$MD = 15$$

QUESTION

Which one of the following is not a method of measures of dispersion?

(a) Standard deviation

(b) Mean deviation

(c) Range

(d) Concurrent deviation method

QUESTION

CA

The standard deviation of 1 to 9 natural number is:

(a) 6.65

(b) 2.58

(c) 6.75

(d) 5.62

$$\sqrt{\frac{n^2-1}{12}}$$

$$= \sqrt{\frac{9^2-1}{12}}$$

$$= 2.58$$

QUESTION

CA

It is given that the mean ($\overline{X}$) is 10 and standard deviation (s.d.) is 3.2. If the observations are increased by 4, then the new mean and standard deviations are:

(a) $\overline{X} = 10$, s.d. = 7.2

(b) $\overline{X} = 10$, s.d. = 3.2

(c) $\overline{X} = 14$, s.d. = 3.2

(d) $\overline{X} = 14$, s.d. = 7.2

(1 mark)

$\overline{X} = 10$

$\sigma = 3.2$

$y_i = x_i + 4$

origin

New mean = 10 + 4 = 14

New SD = 3.2

QUESTION

The best statistical measure used for comparing two series is

(a) Mean absolute deviation ~~X~~

(b) Range ~~X~~

(c) Coefficient of variation

(d) Standard deviation

(1 mark)

QUESTION

CA

The sum of the squares of deviations of a set of observations has the smallest value, when the deviations are taken from their :

(a) A . M.
(b) H. M.
(c) G. M.
(d) None
(1 mark)

$\Sigma (x_i - A)^2$ is small (minimum)

$A = Am$

QUESTION

CA

Which of the following result hold for a set of distinct positive observations?

(a) A . M. $\geq$ G. M. $\geq$ H. M.

(b) G. M. > A. M. > H. M.

(c) G. M. $\geq$ A. M. $\geq$ H. M.

(d) A . M. > G. M. > H. M. (1 mark)

QUESTION

CA

If variance of x is 5, then find the variance of (2 – 3x)

(a) 10

(b) 45

(c) 5

(d) –13

(1 mark)

$\sigma_x^2 = 5$

$\sigma_x = \sqrt{5}$

SD of $(2-3x)$

$= |-3| \times$ SD of x

SD $= 3 \times \sqrt{5}$

Variance $= (3\sqrt{5})^2$

$= 9 \times 5 = 45$


THANK YOU

Type Heading Here

For Normal Text

Font Type : Poppins

Font Color : #FFFFFF

Font Size : 18

Font Type : Poppins Bold

Font Color : #FFFFFF

Font Size : 28

- Those who don't know, Press (Alt) + (=) button to type in Cambria Math for equations like -

Equation

Font Color :#FFF600

Font Size : 18

Font Style : Italic

PS – This ppt is given for design purpose only and for your reference that where to type what or how to go with your final ppt content. So please follow the written instructions given in each placeholder and remove the given 'type here' text to type your related content on the same place. **Do use Microsoft PowerPoint slideshow** for its best view & If you have any issues, please connect with our ppt team.

Sampling

Branches Of Statistics

Sample

Descriptive Stats-Gathering , Classifying & summarizing the data from samples

अनुमान

Inferential Stats - It involves drawing conclusions, generalization or making prediction from the gathered data

Population

- Group of people
- Group of objects
- Group of events & observations

Ex. Height of male students

Ex. Blood pressure of females b/w age 40 & 60 years

N = 50 lakh

50,00,000

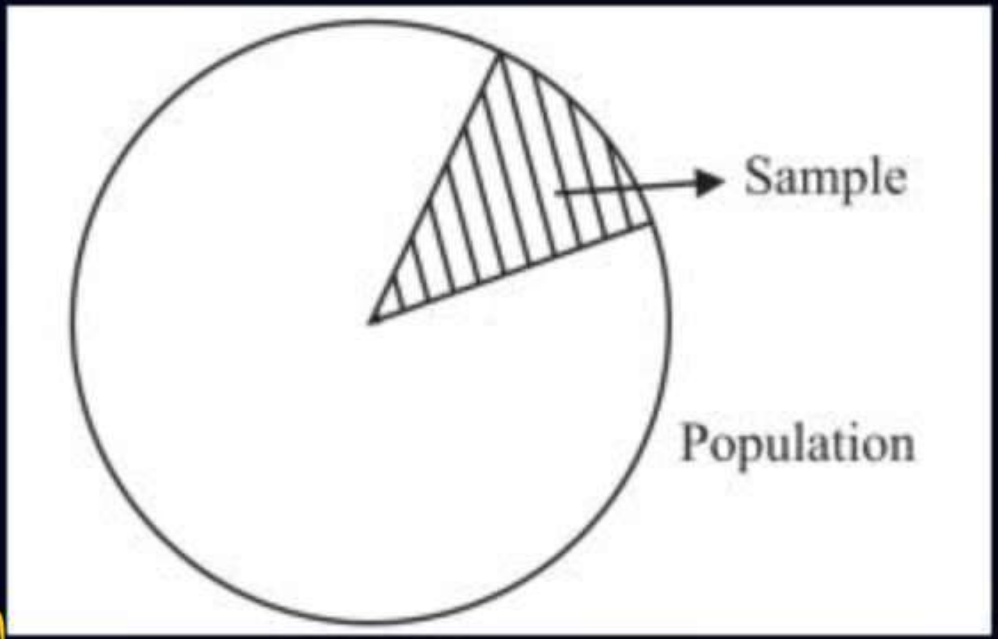
1 lakh

n = 1 lakh

Sample

- It is a subset of Population

"Small group of elements selected from population"



Number of elements in a sample is called the sample size.

Parameters:

Measurable characters of populations are called parameters

Eg Population Mean $= \mu$ $= \mu$

Eg Population Variance $= \sigma^2 = \sigma^2$

S.D. $= \sigma$

Proportion $= P$

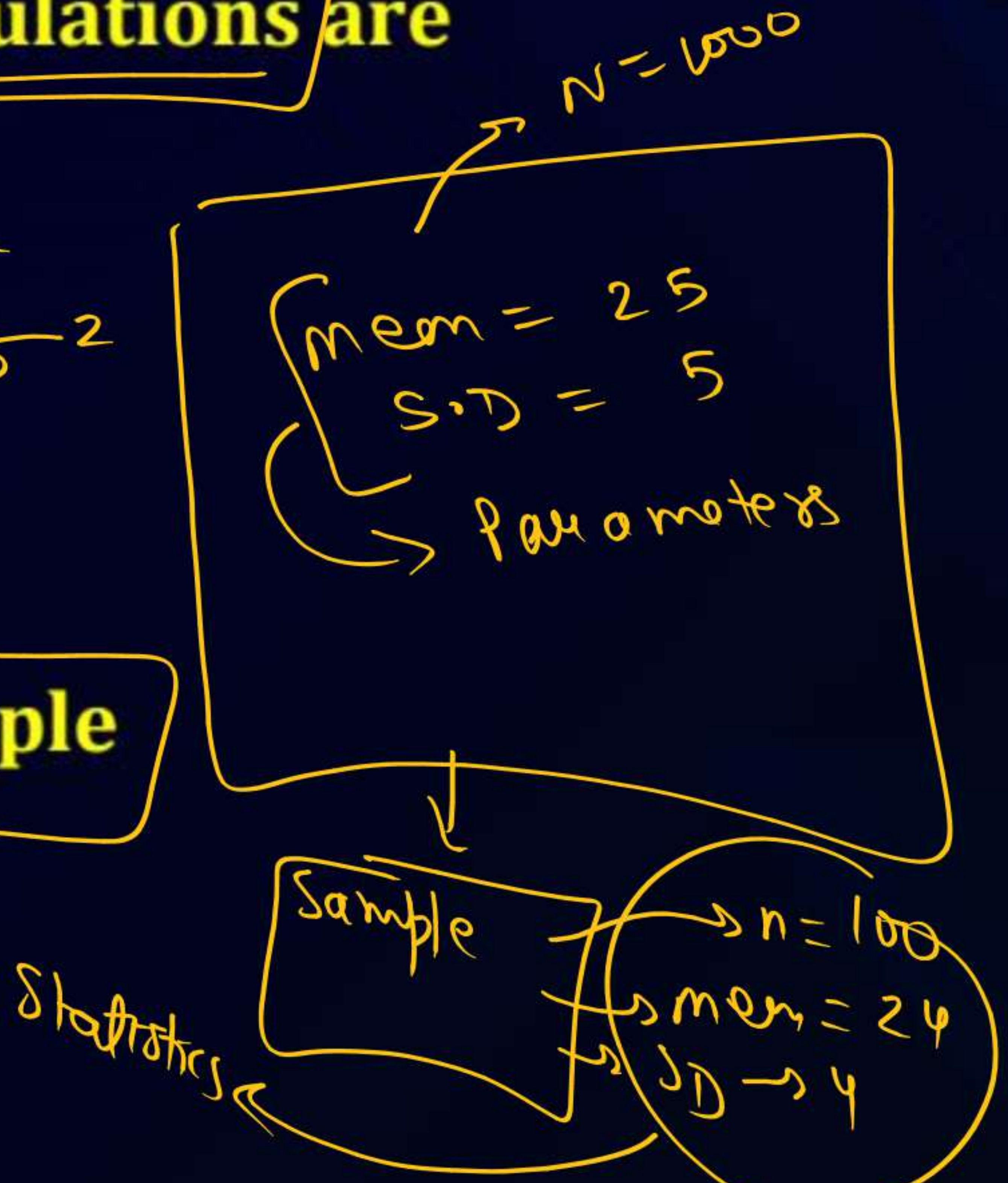
Statistics:

Measurable characters of Sample

Eg Sample Mean $= \bar{x}$

Eg Sample Variance $= s^2$

S.D $=$ S



Parameter is a characteristic of:

(a) Population

(b) Sample

(c) Probability distribution

(d) Both (a) and (b).

Sampling is preferred than population in the following case:

(a) Testing of items is destructive in nature

(b) Testing of items needs equipment of high costs

(c) Population is very large

(d) All of the above.

Principles Of Sampling

Sampling is the procedure of selecting elements for a sample from population so that inferences can be ~~down~~ drawn about population from sample

Some Basis principles of sampling

1. Law of statistical Regularity :

This law suggests that if a large sample is taken randomly from population, it will possess almost same characters of population.

2. Law of Inertia of Large Numbers:

This law is the corollary of "Law of statistical Regularity"
This law says that "Larger the size of sample, more accurate the results one"

3. Principle of optimization:

Maximum efficiency at minimum cost can be achieved only when appropriate "sampling design" is selected

→ method of selection of sample

max profit
minimum cost
Time minimum

4. Principle of validity:

According to this law sampling Design is valid only if it is possible to obtain accurate estimates about population

Probabilistic method

Sampling v/s Census:

1. Speed : Sampling takes less time

2. Cost : Sampling is least expensive

3. Accuracy & Reliability : Census gives more Accurate & Reliable results unless there is no bias or error in collecting information

Types Of Sampling

(Random)

Probability Sampling

when every element of population has equal chances of being selected in sample

Non Probability Sampling

Probability (Random Sampling)

i) Simple Random sampling (SRS):

In this sampling each member of population has an equal chance of being selected in sample.

Ex. There are 1000 students in a college. You assign a number to every student & Then Randomly select a sample of 100 students.

Simple Random sampling without Replacement

SRSWOR

$\Downarrow$

Element selected once can not be selected again.

Eg M, R, S, C

selection of 2 students

$= MR, MS, MC$
RS, RC, SC $= {}^4C_2 = \frac{4!}{2!\,2!} = 6$

Eg Selection of 3 elements from 5 elements

$= {}^5C_3$

$= \frac{5!}{3!\,2!}$

$= 10$

$\boxed{\text{Total ways of selecting 'n' elements from a population of 'N' elements} = {}^NC_n}$

Simple Random sampling with Replacement

SRSWR

element selected once can be selected again.

Eg A, B, C, D

selection of 2 elements

AA	BA	CA	DA
AB	BB	CB	DB
AC	BC	CC	DC
AD	BD	CD	DD

Total ways of selecting 'n' elements from a population of 'N' elements $= (N)^n = 4^2 = 16$

Eg Selection of a sample of 3 elements with Replacement from a population of 5 elements

Sol: $5^3 = 125$

Simple random sampling is

(a) A probabilistic sampling

(b) A non- probabilistic sampling

(c) A mixed sampling

(d) Both (b) and (c).

(ii) Systematic Sampling(Quasi):

In this sampling every member of the population is assigned a number. The first member is selected randomly & then instead of choosing other Randomly we chose them in regular intervals.

Ex. There are 1000 employees in a company, they are assigned numbers from 1 to 1000 we randomly select number '6' from first 10 numbers & after that every 10th person is selected that is 6, 16, 26, 36, ………

$\frac{1000}{100} = 10$

6, 16, 26, 36, 46, - - - - | 7, 17, 27, 37, 47, - - -

50 Peopele

⇓

10 sample

$$\frac{50}{10} = 5$$

1, 2, 3, 4, 5

2, 7, 12, 17, 22, 27, 32, 37, 42, 47

CA WALLAH

[redacted]

T m
S

$\Rightarrow 1000$
$\Downarrow$
100

2 12 22 32 42
T T T T T

Which sampling is affected most if the sampling frame contains an undetected periodicity?

(a) Simple random sampling (b) Stratified sampling

(c) Multistage sampling (d) Systematic sampling

(iii) Stratified Sampling :

In this sampling members are divided into sub group called strata based on gender, age & income etc. After that members are selected using Random or systematic sampling from each sub group.

Ex. He say there are 1000 employees out of which 800 are males & 200 females. A sample of 100 employees reflecting gender balance of the company is made by Dividing the employees in males & females then selecting 80 male & 20 female employer.

100

80 20

$\boxed{\text{1000 Emp}}$

Strata ← $\boxed{\text{200 F}}$ $\boxed{\text{800 m}}$ → Strata

1 : 4

$\boxed{\text{Bowley}}$

$\boxed{\text{Neyman}}$

$\boxed{\text{Sample of 100}}$

f : 20 m : 80

Dehli = 500 UP = 1000

150

$\boxed{\text{50 \& 100}}$

Which sampling provides separate estimates for population means for different segments and also an over all estimate?

(a) Multistage sampling
(b) Stratified sampling
(c) Simple random sampling
(d) Systematic sampling

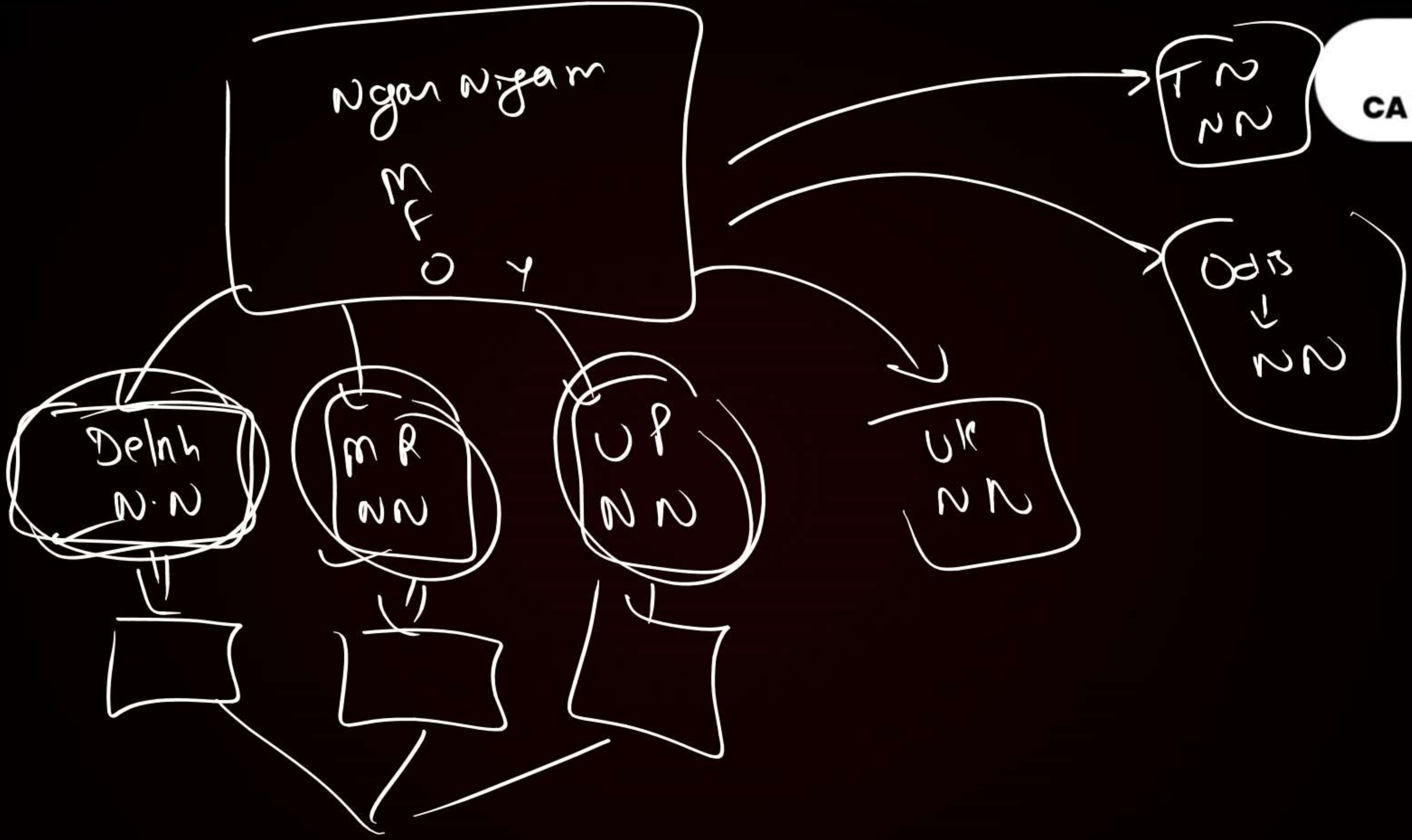
According to Neyman's allocation, in stratified sampling

(a) Sample size is proportional to the population size
(b) Sample size is proportional to the sample SD
(c) Sample size is proportional to the sample variance
(d) Population size is proportional to the sample variance.

(iv) Cluster Sampling (Multi Stage Sampling)

When population size is large, divide the population in sub groups (each sub group has similar characteristic of the whole population) .Then some groups are selected randomly & then members are selected from them for sample.

Ex. You want to study the behavior of Govt. employees of Nagar Nigam. Nagar Nigam Dept. of each state is a cluster (Sample). Then you select any 3 Department of Nagar Nigam & then select 20 members from each Department. Thus sample of 60 item is formed.


Ngar Nigam
M
F
O
Y
Delhi
N.N
MP
NN
UP
NN
UK
NN
TN
NN
Odis
NN

CA WALLAH

Which sampling adds flexibility to the sampling process?

(a) Simple random sampling

(b) Multistage sampling

(c) Stratified sampling

(d) Systematic sampling

Non Probability Sampling

1) Purposive Judgement Sampling: Based on the opinion of expert
Ex. Indian idol

Which sampling is subjected to the discretion of the sampler?

(a) Systematic sampling
(b) Simple random sampling
(c) Purposive sampling
(d) Quota sampling.

2) Convenience Sampling: Those elements are selected which are easily accessible to researcher. Ex. Asking your students to complete survey regarding services provided by universities.

(3) Volunteer response sampling:
People who are themselves ready to conduct the survey collect the sample Data

(4) Snow ball sampling: First select some members, then with the help of them select some more & process continues.

Sampling Errors:

Difference b/w Sample statistics & population parameter because sample was not the true representative of population

- Faulty sampling method
- Faulty Demarcation of sampling units
- Replacing sampling units with unsuitable unit
- Wrong choice of statistic

N = 100000
Population
mean = 25

n = 100
Sample
mean = 22

Non Sampling errors:

These are human errors

census & sampling both can have these errors.

- Lapse of memory
- Preference for certain units
- Wrong measurements
- Untrained interviewer
- Biased opinion

Population

-Aggregate of all units under consideration
-Population size is denoted by "N"

Finite Population

Infinite Population

Population

(Imaginary)
कल्पना

Existent Population

⇓

→ No of students in a schools

→ Male population of Delhi

Hypothetical Population

⇓

Eg A coin is tossed 100 times

Eg A dice is thrown 200 times

- Sample size is denoted by 'n'

- Detailed & Complete list of all potential sampling units is known as Sampling Frame

Sampling frame is a term used for
(a) a list of random numbers
(b) a list of voters
(c) a list of sampling units of population
(d) None of the above.

Population
= 10,000
Audition

100 → Sampling frame

10

Parmeter:

Characteristic of Population

$\Rightarrow$ Population mean (u) = $\dfrac{\sum_{i=1}^{N} Xi}{N}$

$\mu = \dfrac{\Sigma x_i}{N}$

$\Rightarrow$ Population Proportion (P) = $\dfrac{X}{N}$

$P = \dfrac{X}{N}$

$\Rightarrow$ Population S.D (σ) = $\sqrt{\dfrac{\Sigma(xi-\mu)^2}{N}}$

$\underset{(\sigma)}{SD} = \sqrt{\dfrac{\Sigma(x_i-\mu)^2}{N}}$

Statistics:

Measurable Characters of Sample

$\hat{\mu}$ = $\underbrace{\text{Sample mean}}$ = $\bar{x}$ = $\bar{x} = \frac{\Sigma x_i}{n}$

$\bar{x}$ = Sample S.D = S = $\sqrt{\frac{\Sigma (x_i - \bar{x})^2}{n}}$

$\hat{p}$ = Sample Proportion = P = $\frac{x_i}{n}$

Total no of samples with replacement = $(N)^n$ $= (N)^n$

Total no of Sample without Replacement = ${}^{N}C_n$ $= {}^{N}C_n$

Sampling Fluctuation: Value of Sample Statistics may be different in different samples, this variation is called sampling fluctuation

Q 1, 3, 5

2 numbers are selected without Replacement.

Write down all possible samples & their A.M.

$^3C_2 = \frac{3!}{2!\,1!} = 3$

Sol:

		mean
S_1	1 & 3	$\frac{4}{2} = 2$
S_2	1 & 5	$\frac{6}{2} = 3$
S_3	3 & 5	$\frac{8}{2} = 4$

➢ A population comprises of following units : a, b, c & d

Draw all possible samples of

(i) Size two without replacement

(ii) Size two with replacement

Population = a, b, c, d

i) w/o Replacement

a, b
a, c
a, d
b, c
b, d
c, d

$= {}^{N}C_{n}$

$= {}^{4}C_{2}$

$= \frac{4!}{2!\,2!} = 6$

ii) with Replacement

$= (N)^{n} = (4)^{2} = 16$

aa	ba	ca	da
ab	bb	cb	db
ac	bc	cc	dc
ad	bd	cd	dd

Sampling Distribution

- We can make many samples of same size with a given population.
- Sample statistic will be different in each sample.
- If this sample statistic is considered as a Random variable we can make probability distribution. .
- This probability Distribution is known as sampling Distribution

Eg Population: 1, 3 & 5
2 elements selected for sample.
with Replacement

X_i: Random variable (sample mean) = 1, 2, 3, 4, 5

CA WALLAH

Sol:

S.no.	Sample	Mean ($\bar{X}$)
1	(1,1)	1
2	(1,3)	2
3	(1,5)	3
4	(3,1)	2
5	(3,3)	3
6	(3,5)	4
7	(5,1)	3
8	(5,3)	4
9	(5,5)	5

Sampling Distribution

X_i	P_i	P_iX_i	$P_iX_i^2$
1	1/9	1/9	1/9
2	2/9	4/9	8/9
3	3/9	9/9	27/9
4	2/9	8/9	32/9
5	1/9	5/9	25/9
	1	$\frac{27}{9} = 3$	$\frac{93}{9} = 10.33$

Mean
$= E(x)$
$= \Sigma P_iX_i$
$= 3$

S.D (Standard Error)

$= \sqrt{\Sigma P_iX_i^2 - (\Sigma P_iX_i)^2}$

$= \sqrt{\frac{93}{9} - (3)^2}$

$= 1.15$

➢ The mean of Sampling Distribution is known as Expectation denoted by E(x) $= \Sigma p_i x_i$

➢ Standard deviation of this sampling Distribution is known as "Standard Error" $= \sqrt{\Sigma p_i x_i^2 - (\Sigma p_i x_i)^2}$

Standard Error of mean $(\bar{x}) = \dfrac{\sigma}{\sqrt{n}}$ for SRSWR

Standard Error of mean $= \dfrac{\sigma}{\sqrt{n}} \sqrt{\dfrac{N-n}{N-1}}$ for SRSWOR

SRSWR $\rightarrow$ Simple Random Sampling with Replacement

SRSWOR $\rightarrow$ Simple Random Sampling without Replacement

Eg 1, 3, 5

If samples of size 2 are made with replacement find Standard Error of mean.

Sol:

x_i	x_i^2
1	1
3	9
5	25
9	35

$$SD(\sigma) = \sqrt{\frac{\Sigma x_i^2}{N} - \left(\frac{\Sigma x_i}{N}\right)^2}$$

$$\sigma = \sqrt{\frac{35}{3} - \left(\frac{9}{3}\right)^2}$$

$$= \sqrt{2.6666}$$

$$\sigma = 1.63299$$

S.E of mean

$$= \frac{\sigma}{\sqrt{n}}$$

$$= \frac{1.63299}{\sqrt{2}}$$

$$= 1.1546$$

CA WALLAH

Standard Error for Proportion:

$$SE(P) = \sqrt{\frac{Pq}{n}} \qquad \text{SRSWR}$$

$$SE(P) = \sqrt{\frac{Pq}{n}}\sqrt{\frac{N-n}{N-1}} \qquad \text{SRSWOR}$$

A measure of precision obtained by sampling is given by

(a) Standard error
(b) Sampling fluctuation
(c) Sampling distribution
(d) Expectation.

As the sample size increases, standard error

(a) Increases
(b) Decreases
(c) Remains constant
(d) Decreases proportionately.

Standard error can be described as

(a) The error committed in sampling
(b) The error committed in sample survey
(c) The error committed in estimating a parameter
(d) Standard deviation of a statistic.

A Population has 3 elements 1, 5 & 3
Draw all possible sample of size two
(i) With Replacement (ii) Without Replacement

W

N^n

$= 3^2$

$= 9$

W/

^NC_n

$= {}^3C_2$

$= 3$

A population comprises 5 members. The number of all possible samples of size 2 that can be drawn from it with replacement is

(a) 100 (b) 15

(c) 125 (d) 25

$$N^n = 5^2 = 25$$

If from a population with 25 members, a random sample without replacement of 2 members is taken, the number of all such samples is

(a) 300 (b) 625

(c) 50 (d) 600

$$^NC_n = {}^{25}C_2 = \frac{25!}{2!\,23!} = \frac{25 \times 24}{2} = 300$$

Sol: $N = 3$

$SD(\sigma) = \sqrt{\frac{8}{3}}$

$n = 2$

i) WR

$S.E = \frac{\sigma}{\sqrt{n}}$

$= \frac{\sqrt{\frac{8}{3}}}{\sqrt{2}} = \frac{1.632}{1.41} = 1.1539$

ii) WOR

$SE = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$

$= \frac{\sqrt{\frac{8}{3}}}{\sqrt{2}} \sqrt{\frac{3-2}{3-1}}$

$= \frac{\sqrt{\frac{8}{3}}}{\sqrt{2}} \times \frac{1}{\sqrt{2}} = 0.8164$

QUESTION

Sampling fluctuations may be described as:

(a) The variation in the values of a sample

(b) The differences in the values of a parameter

(c) The variation in the values of a statistic

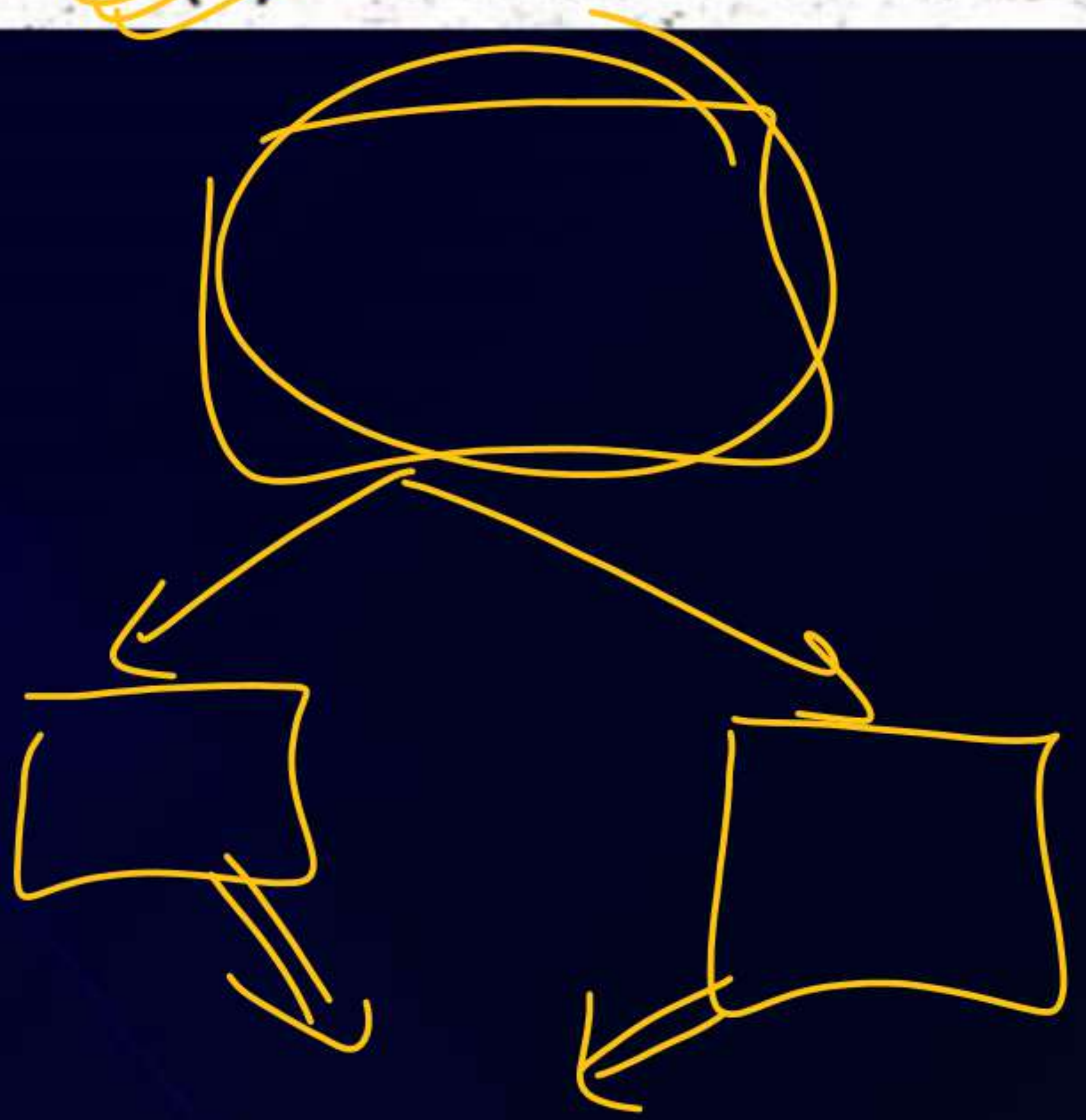
(d) The variation in the values of observations.

QUESTION

Which sampling provides separate estimates for population means for different purposes and also an over all estimate ?

(a) Multistage sampling (b) Simple random sampling

(c) Systematic sampling (d) Stratified sampling



QUESTION

If the population S.D. is known to be 5 for a population containing 80 units, then the standard error of sample mean for a sample of size 25 without replacement is :

(a) 0.83

(b) 0.80

(c) 0.93

(d) 0.74

$\sigma = 5$

$N = 80$

$n = 25$

$$SE = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

$$= \frac{5}{\sqrt{25}}\sqrt{\frac{80-25}{80-1}}$$

$$= \frac{5}{5} \times \sqrt{\frac{55}{79}} = 0.83$$

QUESTION

A Population comprises of 20 members. The number of all possible samples If size 2 that can be drawn from it without replacement.

(a) 210

(b) 380

(c) 190

(d) 400

$^{20}C_2$

QUESTION

Distribution formed of all possible value of statistics is called______.

(a) Sampling Distribution
(b) Classification
(c) Tabulation
(d) None

QUESTION

In sampling, standard error is:

(a) Standard deviation

(b) Quartile deviation

(c) Mean deviation

(d) Coefficient of variation

QUESTION

CA

If every 9^{th} unit is selected from universal set then this type of sampling is known as:

(a) Quota Sampling

(b) Systematic Sampling

(c) Stratified Sampling

(d) None of these

QUESTION

The method of sampling in which each unit of the population has an equal chance of being selected in the sample is

(a) Random sampling
(b) Stratified sampling
(c) Systematic sampling
(d) None of the above.

QUESTION

The standard deviation of.sampling distribution is know

(a) Standard error
(b) Mean
(c) Variance
(d) Mode

QUESTION

If a random sample of size 2 with replacement is taken from the population containing the units 3, 6 and 1, then the samples would be

(a) (3, 6), (3, 1), (6, 1)

(b) (3, 3), (6, 6), (1, 1)

(c) (1, 1), (1, 3), (1, 6,), (6, 1), (6, 2), (6, 3), (6, 6), (6, 1)

(d) (1, 1), (1, 3), (1, 6), (3, 1), (3, 3), (3, 6), (6, 1), (6, 3), (6, 6)

3, 6, 1

$N^n = 3^2 = 9$

QUESTION

Method used to test the human blood is called in Statistical terminology___

(a) Census Investigation
(b) Blood Investigation
(c) Sample Investigation
(d) None of these