

Monday

PAS

Jan 27, 25

## Statistics

(where there is uncertainty  
there is statistics)



descriptive

Inferential

(Decision making)

### Characteristics:

1. It deals with large groups.
2. It deals with variability, such that no two objects in this universe are exactly alike.
3. It deals with uncertainty or chance variations.
4. It deals with those things which can describe numerically.
5. It deals with those aggregates which are subject to random causes.

→ **Spurious Effect**: third variable effect.

Probability sampling: Each person is included in the survey.

Non-Probability: No Equal opportunity.

→ Snowball sampling:

Sample

$n$  - size

Statistics (estimate)

(A - Z) Latin

Population

$N$  - size

Parameter

( $\alpha, \beta, \gamma, \dots$ ) Greek

**Primary Sources**: First time collection

**Secondary**



## Primary Data Sources:

- 1- Direct personal investigation, (directly relate)
- 2- Indirect investigation, (indirect ways)
- 3- Collection through questionnaire, (google forms...)
- 4- Collection through enumerators, (acc. to area-term)
- 5- Collection through local sources. (local people)

## Study Designs:

1. Cross-sectional (at a particular time)
2. Time-Series (overally diff. time series)

## Errors in Data Collection: (human error)

1. Systematic errors: tools imperfections. By repeating the experiment can remove.
2. Random errors: natural fluctuations (eliminate)
3. Samplig errors: Problem of sample & population  

overestimation  
-ve

underestimation  
+ve

Unbiased  
equal

## Nature of the variable:

Qualitative

Quantitative

variable of interest: jis ko study krna (variable)  
Discrete      Continuous

Finite Population: accessible population.

Infinite " : "

(unis all over world)



**Hypothetical** : All possible values, (understood) outcomes

## Scale of measure:

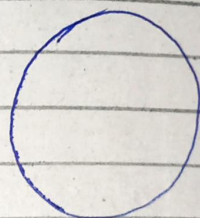
- Nominal (Qualitative nature)
  - Ordinal (Quan / Qual)
  - Interval } Quant.
  - Ratio }
- 0° temperature  
Zero (presence)  
Zero (True) Absence

Relative freq → percent  
 $\frac{\text{Freq}}{\text{total}}$        $\text{RF} \times 360$

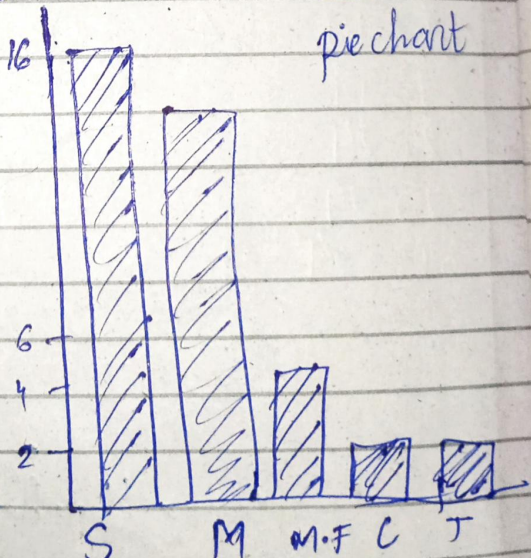
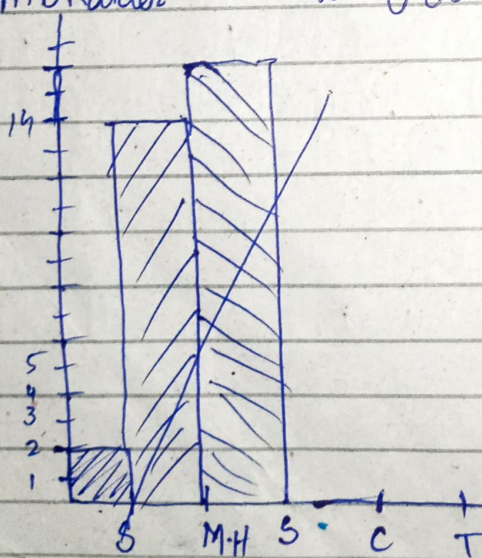
As the crow flies

Series      Freq   R-F   P-F   Angle

As the crow flies	2	0.05	5	1.8
Money Heist	14	0.35	35	126
Suits	16	0.41	41	147.6 ✓
Manifest	5	0.12	12	43.2
Tomb Raider	2	0.05	5	1.8



pie chart



Bar chart

4 significant figures.

(Anderson) → (interpretations)



Feb 3, 25

## Group Data:

- 1) Determine no. of non-overlapping classes.
- 2) Determine class width
- 3) Determine class intervals / Bins.

$$\text{no. of classes} = 1 + 3.3 \log(\text{no. of obs})$$

$$\text{class width} = \text{Range} / \text{no. of classes}.$$

12, 14, 19, 18, 15, 15, 18, 17, 20, 27  
22, 23, 22, 21, 33, 28, 14, 18, 16, 13

$$\text{no. of classes} = 1 + 3.3 \log(20)$$

$$= 5.2 \Rightarrow 5 - 6$$

$$\text{C.W} = \frac{33 - 12}{5.2} = 4.03$$

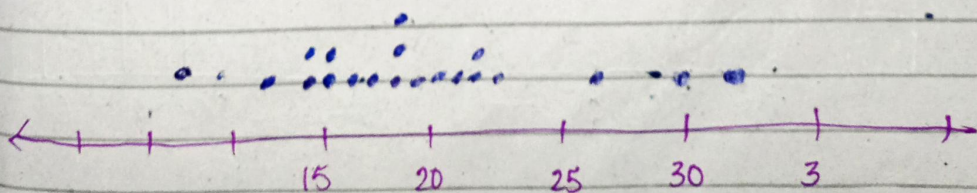
(class width)

$$= \frac{\text{max} - \text{min}}{\text{no. of classes.}}$$

commulative

subtract 0.5  
from lower bound  
and add 0.5  
to upper bound.

classes	f	R-F	P-F	C-F	CRF	CPF	CB
10-14	4	0.2	20	4	4/20	20	9.5-14.5
15-19	8	0.4	40	4+8=12	0.6	60	14.5-19.5
20-24	5	0.25	25	17	0.85	85	19.5-24.5
25-29	2	0.1	10	19	0.95	95	
30-34	1	0.05	5	20	1	100	29.5-34.5
sum	20	1.	100	*	.		



Dot Plot

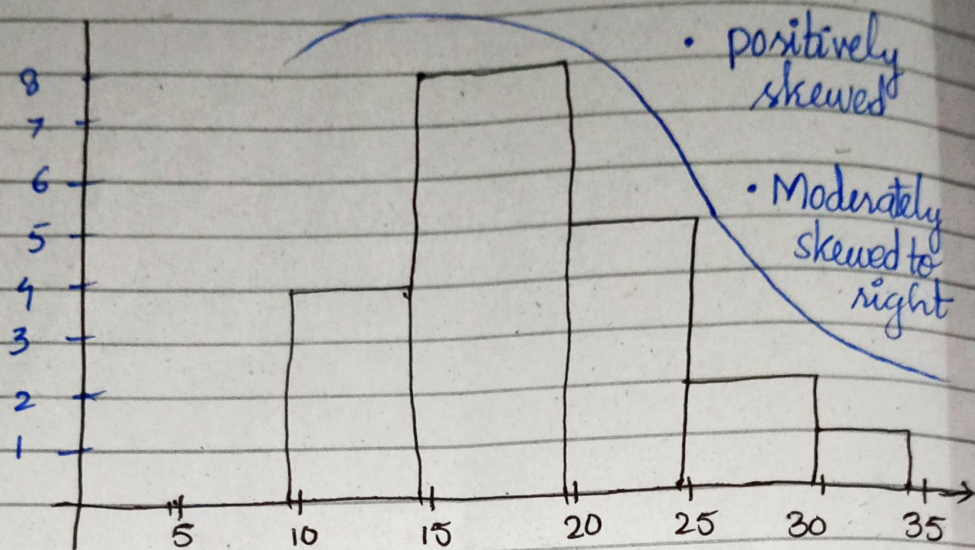


$$\frac{10-14}{15-29}$$

$$\frac{15-14}{2} = 0.5$$

$$\frac{8}{10} \times 100 = 80\%$$

## Histogram. (continuity)



class boundaries.

\* sample mean :  $\bar{X} = \frac{\sum f x}{n}$   $n = \sum f$

↳ (mid pts. x frequencies) / n.

→ Sample variation :  $s^2 = \frac{\sum f (X - \bar{X})^2}{n-1}$

S.d =  $s = \sqrt{\frac{\sum f (X - \bar{X})^2}{n-1}} = \sqrt{\frac{\sum (\text{ind. data pts} - \text{mean})^2}{\text{sample size} - 1}}$

variance.

- C.V =  $\left( \frac{s.d \times 100}{\text{Mean}} \right) \%$

coefficient of variable.

lesser the C.V better the data.

→ \* Average of standard deviation from mean.

higher variation → Less consistency.



$x(\text{midpoint})$	$fx$	$f(x-19)^2$
12	$4 \times 12$	$4(12-19)^2$
17	$8 \times 17$	$8(17-19)^2$
22	$5 \times 22$	$5(22-19)^2$
27	$2 \times 27$	$2(27-19)^2$
32	$1 \times 32$	$1(32-19)^2$
	<u>          </u>	<u>          </u>
		570

$$s^2 = \frac{570}{20-1} = 30$$

$$s = 5.47$$

Cross Tabulation:

$$(3x+6)$$