STATISTICS

Descriptive (On a one sample)

Inferential (on a large scale)

* DESCRIPTIVE STAT : collection, organization, analysis, results

- It consists of the collection, Organization, summarization and presentation of data.

* Collection : DATA → Information

Qualitative Data

categorical data

E.g: Male / Female (M_1, M_2, F_i)

- Data values are characters or qualities (old, young) etc.

Quantitative Data

Numeric Data

E.g: Age (continuous but numeric)

- Data values are continuous weight etc.

Discrete

$x = 0, 1, 2$

E.g. no of children

continuous

$x = (1 - 10)$

$x_2 = (10 - 20)$

E.g: weight

* Organizing Data :

We can organize our data (Quali/Quant) by using frequency distribution.

Frequency

- Raw Data : Rough collection of data is raw data.

Day/ Date:

- Frequency Distribution:
  It is used to organize raw data into tableted Form. (Table form)

- Columns in freq Dist:
  1) Data Values (n)
  2) Frequency (f) (No of repetations)
  3) Relative Frequency (f/n) 4) Percentage = (f/n x100)
  5) Cummolative Frequency (c.f)

- Organizing Qualitative Data:

Q. Construct a frq. dis of raw data given below:

D	R	O	R	R	R	R	R
D	O	R	D	O	O	R	D
D	R	O	D	R	R	O	R
D	O	D	D	R	R	O	D
O	R	D	R	R	R	R	D

n = 40

Freq Distribution

	Data values	Freq (f)	Relative freq (f/n)	Percentage (f/n x100)	Commulative (c.f)
1)	D	13	.325	32.5	13
2)	R	18	.45	45	13+18 = 31
3)	O	9	.225	22.5	40 = 31+9
	Σ	=40 = n	=1.00	= 100 %	

Day/ Date:

* Organizing Quantitative Data:

Raw data

ungrouped ⟷ Grouped

Ungrouped

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

$n = 32$

Grouped

25	42	82	48	85
70	54	29	44	82
99	29	55	47	85
86	82	47	86	83

$n = 20$

- To check grouped or ungrouped we check smallest or highest value of data.

Ungroup data.

1	1	1	2	6	3	3	4	2	4
3	2	1	5	2	1	3	6	2	2
3	1	1	4	3	2	2	2	2	3
0	3	1	2	1	2	3	1	1	3
3	2	1	2	1	1	3	1	5	1

$n = 50$

Data Values	(f)	(f/n)	(f/n×100)	C.F
0	1	.02		1
1	16	.32		17
2	14	.28		31
3	12	.24		43
4	3	0.06		46
5	2	.04		48
6	2	.04		$\boxed{50}$
	=50	= 1	=100%	=

Dazzle

Day/Date:

Grouped Data : (classes banaty hain frequencies ky liye)

				64	89	87	65
70	64	99	55	60	69	78	39
62	38	67	70	99	68	95	86
75	56	71	51	55	81	80	98
57	53	47	50	85	79	83	70
51	36	63	66				

$n = 40$ Smallest $= 36$ Max $= 99$

- First we decide ky kitni no of classes chahiye, your choice.

class diff $=$ Width $= \frac{H-L}{\text{no of class}} = \frac{99-36}{7} = 10.5 \cong 10$

- 0.5 lower ~~or upper~~ limits mai sy minus kardo or upper limit mai add kardo

(To make data cont)

classes	class boundaries	f	f/n	f/n x100	c.f	Mid Point
1. 30-39	29.5 – 39.5	3	0.075		3	34.5
2. 40-49	39.5 – 49.5	1	0.025		4	44.5
3. 50-59	49.5 – 59.5	8	0.2		12	54.5
4. 60-69	59.5 – 69.5	10	0.25		22	64.5
5. 70-79	69.5 – 79.5	7	0.175		29	74.5
6. 80-89	79.5 – 89.5	7	0.175		36	84.5
7. 90-99	89.5 – 99.5	4	0.1		40	94.5
		= 40	= 1			

Dazzle

Day/ Date:

Q. Group Data

112 100 127 120 134 118 105 110 109 112
110 118 117 116 118 112 114 114 105 109
107 112 114 115 118 117 118 122 106 110
116 108 110 121 113 120 119 111 104 111
120 113 120 117 105 110 118 112 114 114

$n = 50$

$$\text{Width} = \frac{H-L}{\text{no of classes}} = \frac{134-100}{7} = 4.85 \simeq 5$$

	Classes	class-Bound.	(f)	f/n	f/n x100	C.F	M.P
1.	100-104	99.5-104.5	2	0.04		2	102
2.	105-109	104.5-109.5	8	0.16		10	107
3.	110-114	109.5-114.5	18	0.36		28	112
4.	115-119	114.5-119.5	13	0.26		41	117
5.	120-124	119.5-124.5	7	0.14		48	122
6.	125-129	124.5-129.5	1	0.02		49	127
7.	130-134	129.5-134.5	1	0.02		50	132
			= 50	= 1			

Day/ Date:

* Graphs Using freq dis

qualitative   ungroup   group
(x - Data V/CB)

1) Histogram (x,y axis)

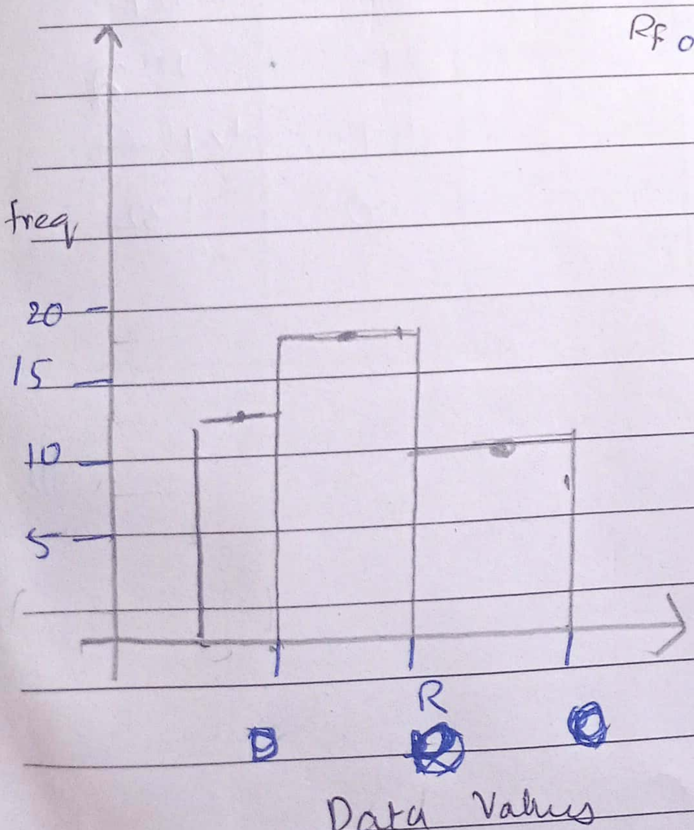
from   R.freq   —x   y py Relative freq'
Histo   Histo py data values ya class boundary

y py freq (F)

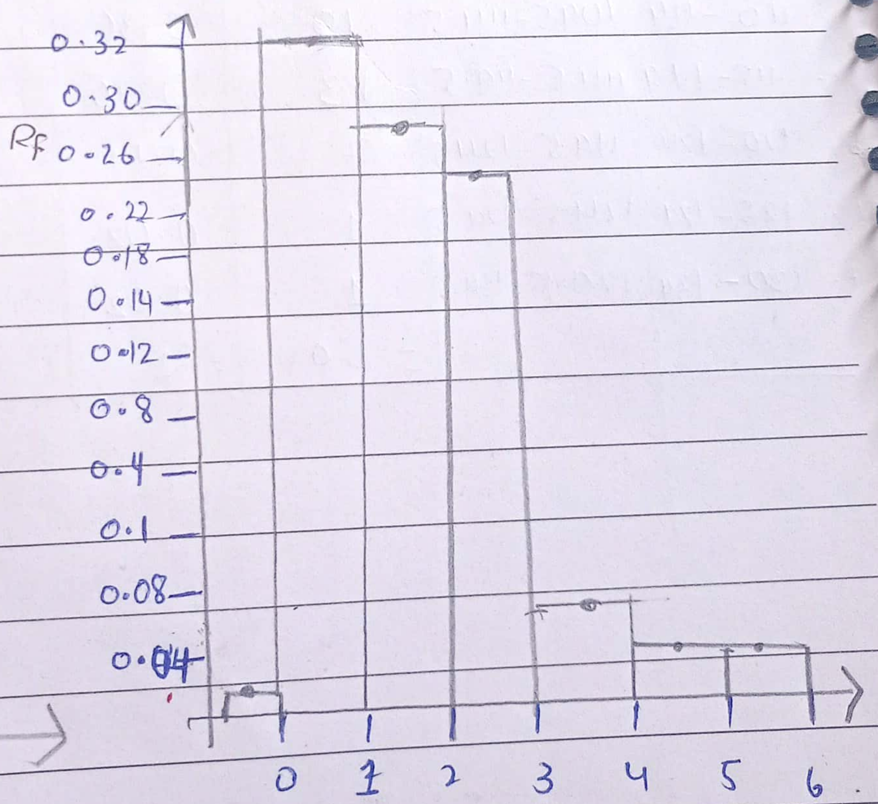
GRAPHS FOR UNGROUP/CATEGORICAL DATA

- Only Histograms are formed.

1) freq- Histo



2) Relative Freq.



Day/ Date:

GRAPHS FOR GROUP DATA

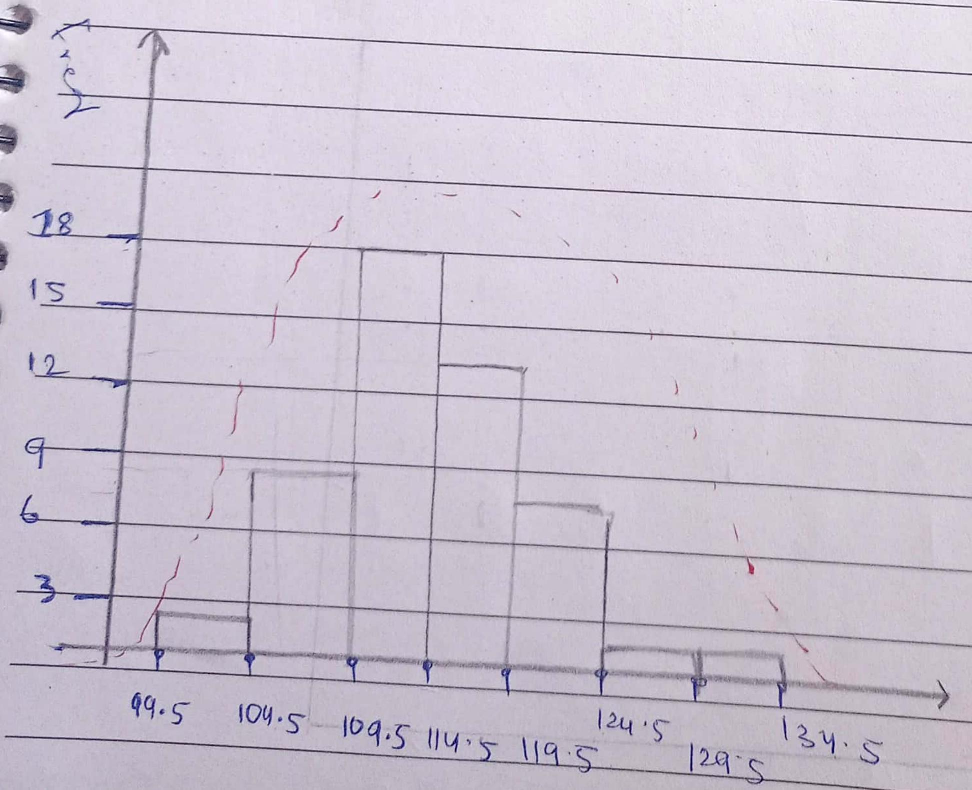
* Histogram, freq. polygons and Ogives.

Class	CB	Freq	f/n	C.f	MP
1) 100-104	99.5-104.5	2	0.04	2	102
2) 105-109	104.5-109.5	8	0.16	10	107
3) 110-114	109.5-114.5	18	0.36	28	112
4) 115-119	114.5-119.5	13	0.26	41	117
5) 120-124	119.5-124.5	7	0.14	48	122
6) 125-129	124.5-129.5	1	0.02	49	127
7) 130-134	129.5-134.5	1	0.02	50	132

– Histogram (Vertical Bars).

- CB are set on x-axis
- frequencies/RF/CF on y-axis

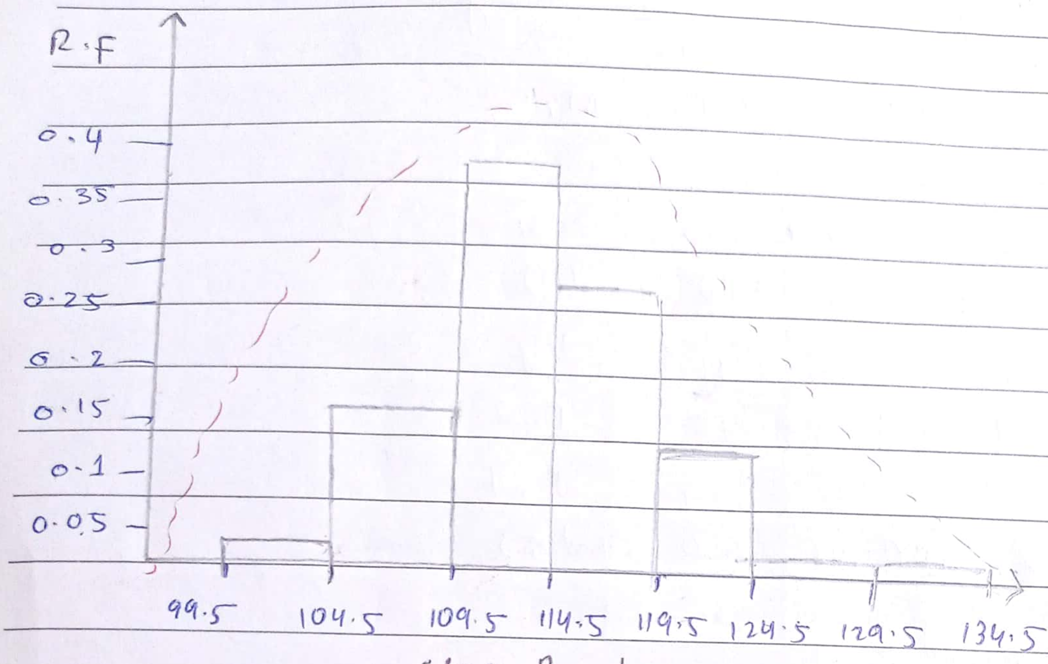
1) Frequency Histogram: Using freq col on y-axis



Dazzle

Day / Date:

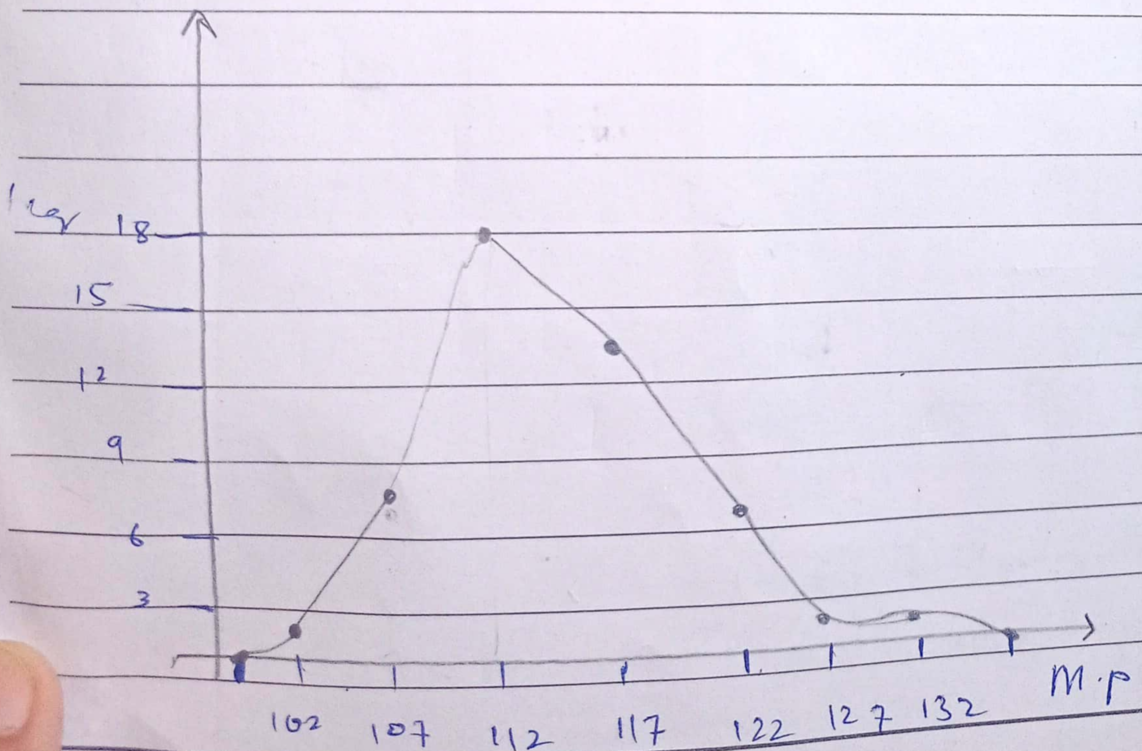
2) Relative - frequency Histogram : (using rel.f on y-axis)



Frequency - Polygon (always connected on x-axis)

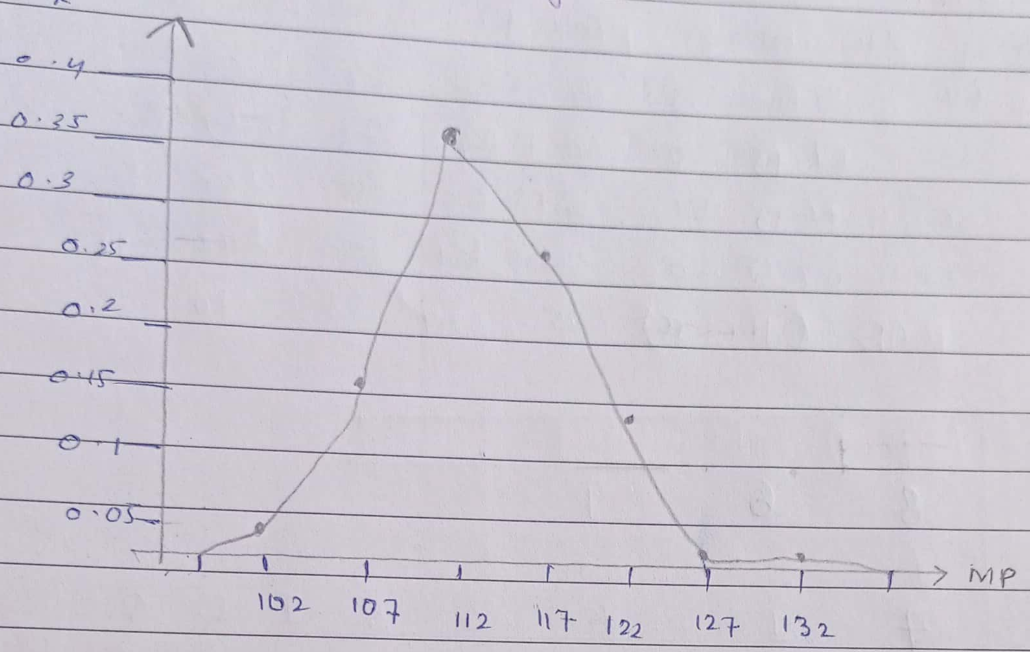
1) Frequency - Polygon

- X-axis py mid points
- y-axis py freq.



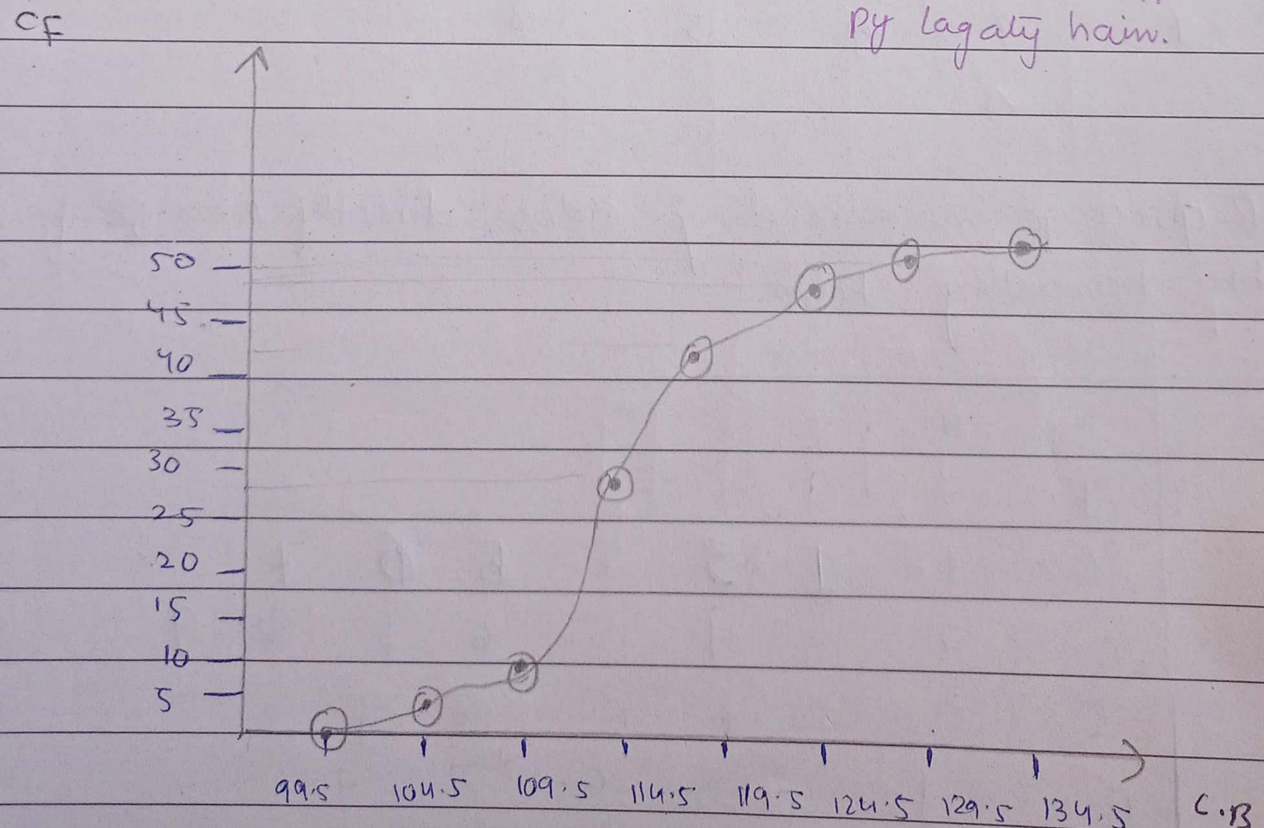
Day/ Date:

iii) Relative freq - polygon:



O-give (C.F, CB)

- x-axis py Cb , y-axis C.F Point hamesha upper limit py lagaty hain.



Day/ Date:

STEM AND LEAF PLOT (Depends on digits)

- No freq distribution plotted.

70	64	99	55	64	89	87	65
62	38	67	70	60	69	78	39
75	56	71	51	99	68	95	86
57	53	47	50	55	81	80	98
51	36	63	66	85	79	83	70

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

: Rearrange in ascending order, directly arrange form bhi banaskty hain.

Stem	Level.
3	6 8 9
4	7
5	0 1 1 3 5 5 6 7
6	0 2 3 4 5 6 7 8 9
7	0 0 0 1 5 8 9
8	0 1 3 5 6 7 9
9	5 8 9 9

Dazzle

Day/ Date:
Q.

210	209	212	208
217	207	210	203
208	210	210	199
215	221	213	218
202	218	200	214.

n = 20

Stem	leaves
19	9
20	0 2 3 7 8 8 9
21	0 0 0 0 2 3 4 5 7 8 8
22	1

- Ek digit hamesha leaf ky liye rakhty hain.
- Can give statement ky konsi shape hai histo ki.

Day/ Date:

Q. ~~6~~ ~~38~~ D R O R R R R R

~~7~~ ~~42~~ D O R D O O R D

~~29~~ D R O D R R O R

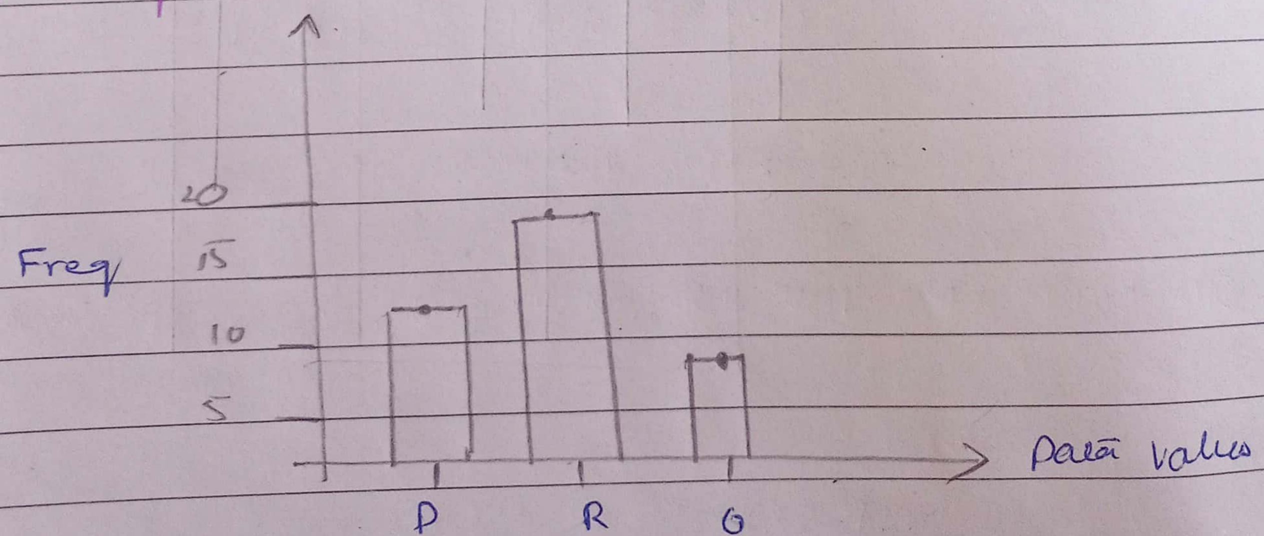
~~35~~ D O D D D R O D

~~38~~ O R D R R R R D

* BAR CHART (only for qualitative)

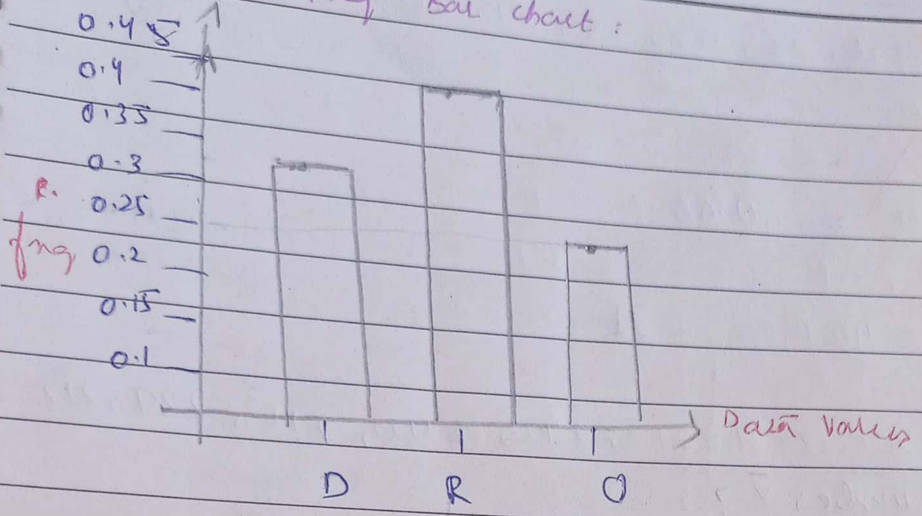
Data Values	freq (f)	RF	C.F
D	13	0.325	13
R	18	0.452	31
O	9	0.225	40
	= 40	= 1	

1) Freq Bar chart :



Day/ Date:

2) Relative. Freq Bar chart :



UNGROUP DATA

1) Mean : (Average value)

i) Population Mean : Poori population ko consider karty hain. $(\mu) = \Sigma X/N$

ii) Sample : sample means 2 ek jagah sy baki do dusri jagah sy : $(\bar{X}) \rightarrow$ Data values ka average.

$$\bar{X} = \frac{\Sigma x}{n}$$

2) Median (middle value) (MD) :

i) Arrange data

ii) middle value = both + both/2.

3) Mode :

- Most repeated Value

4) Mid Range :

$$MR = \frac{\text{Highest val} + \text{lowest Value}}{}$$

Dazzle

Day/ Date:

Q. 20, 26, 40, 36, 23, 42, 35, 24, 30. Find Average and MR.

$\bar{X}$ = Average = $\frac{276}{9}$ = 30.66

Mid-Range = MR = $\frac{42+20}{2}$ = 31

Q. 209, 223, 211, 227, 213, 240, 240, 211, 229, 212.
Find MD, mode, $\bar{X}$?.

For MD :

209, 211, 211, 212, 213, 223, 227, 229, 240, 240

Two middle vales so

MD = $\frac{213+223}{2}$ = 218

For mode :

Mode 1 : 211 Mode 2 : 240

$\bar{X}$ = $\frac{2215}{10}$ = ~~24.615~~ 221.5

Q. 713, 300, 618, 595, 311, 401, 292. Find MR and MD and Mode.

MD : 292, 300, 311, 401, 595, 618, 713

MD = 401

Mode : None.

MR : $\frac{292+713}{2}$ = 502.5

Dazzle

Day/ Date:

Q. 104 104 104 104 104
107 109 109 109 110
109 111 112 111 109
find Mode:
Mode (1): 104 => f = 5
mode (2): 109 => f = 5

Q.

Data	Freq (no of students) (Qualitative)
Buisness	1425
Liberal Arts	878
Computer Sci	632
Education	471
General Studies	95

Find $\bar{X}$, Mode

$\bar{x} = 700.2$

~~Mode = 1425 (highest frequency)~~ mode = none

MD = Median = $\frac{\text{sum}}{2} = \frac{3501}{2} = 1750.5$

GROUPED DATA

Q. find mean, median, mode for grouped data using freq. distribution.

	C.B	f	X_m Mid.Point	f × MP	C.f
1)	5.5 – 10.5	1	8	8	1
2)	10.5 – 15.5	2	13	26	3
3)	15.5 – 20.5	3	18	54	6
4)	20.5 – 25.5	5	23	115	11
5)	25.5 – 30.5	4	28	112	15
6)	30.5 – 35.5	3	33	99	18
7)	35.5 – 40.5	2	38	76	20

$n = 20$ $\Sigma(f.MP) = 490$

Dazzle

Day/ Date:

i) Average $\bar{x} = \frac{\Sigma(f.MP)}{n} = \frac{490}{20} = 24.5$

$\bar{x}$ lies in class (20.5 - 25.5)

ii) $MD = \frac{n}{2} = \frac{20}{2} = 10$ MD lies in class (20.5 - 25.5) (Check C.F)

iii) Modal class is (20.5 - 25.5)

Mode = 5 (Highest freq)

DATA TYPES OF MEAN, MEDIAN, MODE :

1) Raw Data/small data with less frequency:

Data : $\overset{x_1}{209}$, $\overset{x_2}{211}$, $\overset{x_3}{211}$, 212, 213, 223, 227, 229, 240, 240.

We will simply add and divide for average

$\bar{x} = \Sigma x_i/n$; If odd = middle term

Median = 212+213/2 = 218   Mode(1) = 211, 240.

2) Categorical Data with given frequencies:

1)

Department	No of student (f)	(x_i)
Buisness	1425	
Liberal Art	878	
Comp. Sci	632	
Education	471	Median = Middle value
General studies	95	
Mode =	none	

2)

Staff	Salary (x_i)
Owner	50,000
Manager	20,000
Sales Person	12,000 → MD
Technician	9000
2	9000 } Mode

$\bar{X} = \frac{\Sigma x_i}{n}$

Mode = Most repeated value = 9000 = Technician

Dazzle

Day/ Date

3) Ungrouped Data with given Frequencies:

1)

Salary	(f)	f.x	CF
300	6		6
400	2		8
450	2		10
800	1		11
940	1		12
1050	1		13
	n=13		

$\bar{x} = \frac{\Sigma(x.f)}{n}$

$MD = \frac{n}{2}$ (search Cf)   Mode = 300 (highest freq)

(2)

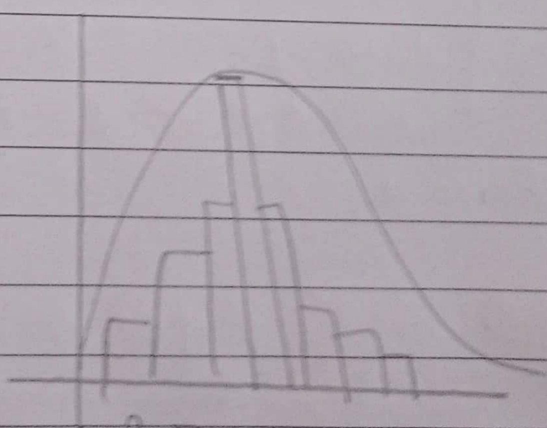
x	(f)	f.(x)	Cf	
0	3	0	3	
1	4	4	7	
2	6	12	13	
3	7	21	20	
4	10	40	30	←
5	6	30	36	Median Model Average class
6	5	30	41	
7	5	35	46	
8	3	24	49	
9	1	9	50	

2) $\bar{x} = \frac{205}{10} = 20.5$ Median $= \frac{50}{2} = 4.1$ Mode = 4.

~~Mode (MD)~~

Mean > MD

2)



Both bell shape or +ve skewed.

Dazzle

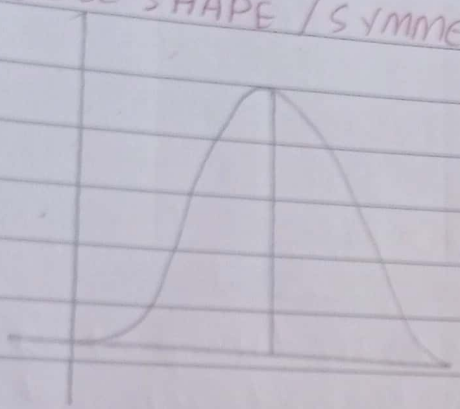
Day / Date:

SKEWNESS & KURTOSIS

Symmetric   +ve skewed

Skewness not checked on categorical data.

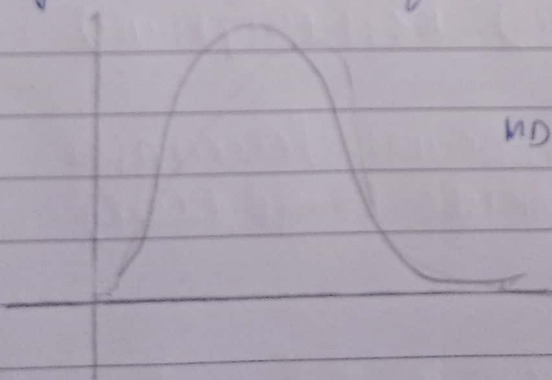
1) BELL SHAPE / SYMMETRIC / ZERO SKEWNESS :



- Can check through Histogram (graphical)
- Can check Mean and Median. If Mean = Median we can comment data set is symmetric

2) POSITIVE SKEWED :

- If mean > Median +ve skewed.
- If through Histogram, first increase then decrease. Tail going to +ve side of x-axis.

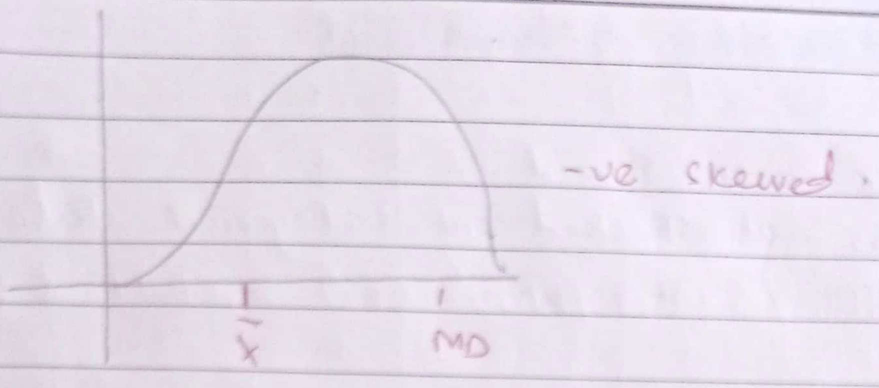


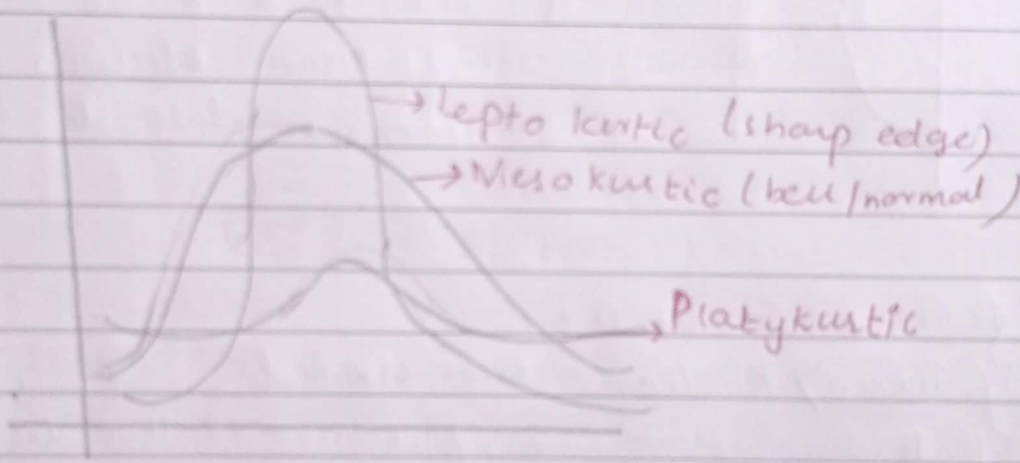
MD < $\bar{x}$

3) NEGATIVE SKEWED :

- If mean < Median, -ve skewed
- Through histogram, tail towards -ve x.

Day/ Date:



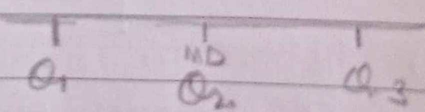


• kurtosis represented by 'k'. If $k>3$ (lepto) $k=3$ (Meso kurtic) $k<3$ (Platy kurtic)

QUARTILES, DECILES, PERCENTILES & INTERQUARTILE RANGE (IQR) (UNGROUP)

* Formulas:

i) Quartiles : (Q_1, Q_2, Q_3)



$$Q_i = \left[\frac{i \times (n+1)}{4}\right] \text{th value of observation.}$$

• Breaking Data Set into 4^{th} portion. (means 3 portion)

Day/Date:

ii) Deciles : $(D_1, D_2, \ldots\ldots, D_9)$

$$D_i = \left[\frac{i \times (n+1)}{10}\right] \text{th value}$$

iii) Percentiles : $(P_1, P_2, \ldots P_{99})$

$$P_i = \left[\frac{i \times (n+1)}{100}\right] \text{th value}$$

iv) IQR :

$IQR = Q_3 - Q_1$ (only Quartiles)

(V) Box-Plot : Only for ungroup

(low) L, Q_1, Q_2, Q_3, h (highest)

Q. Find all Possible Quartiles and then find IQR, also plot Box-Plot.

215, 138, 78, 48, 39, 89, 47, 164, 296, 30 $n = 10$

$Q_1 = \Big[$ 30, 39, $\boxed{47}$, 48, $\boxed{78, 89}$, 138, $\boxed{164}$, 215, 296

Md

For Q_1:

$$Q_1 = \left[\frac{1(11)}{4}\right] \text{th} = 2.75 \text{th value}$$

$= 2$ nd $+ 0.75(\text{ratio})(3^{rd} - 2^{nd})$

$= 39 + 0.75(47 - 39)$

$= 39 + 0.75(8) = 39 + 6 = 45$

$\boxed{Q_1 = 45}$

Day/Date: ____

For Q_2:

$Q_2 = \left(\frac{2 \times 11}{4}\right)$th value $= 5.5$th value

$= 5 + 0.5(6^{th} - 5^{th})$

$= 5 + 0.5(78 - 189)$

$Q_2 = \cancel{89.5}$ $\cancel{24.5 + 78}$

$Q_2 = 83.5 = 84$

For Q_3 =

$Q_3 = \left(\frac{3 \times 11}{4}\right)$th value $= 8.25$

$= 8^{th} + 0.25(7^{th} - 8^{th})$

$= 164 + 0.25(215 - 164)$ $\cancel{164 - 138}$

$= 164 + 0.25(51)$

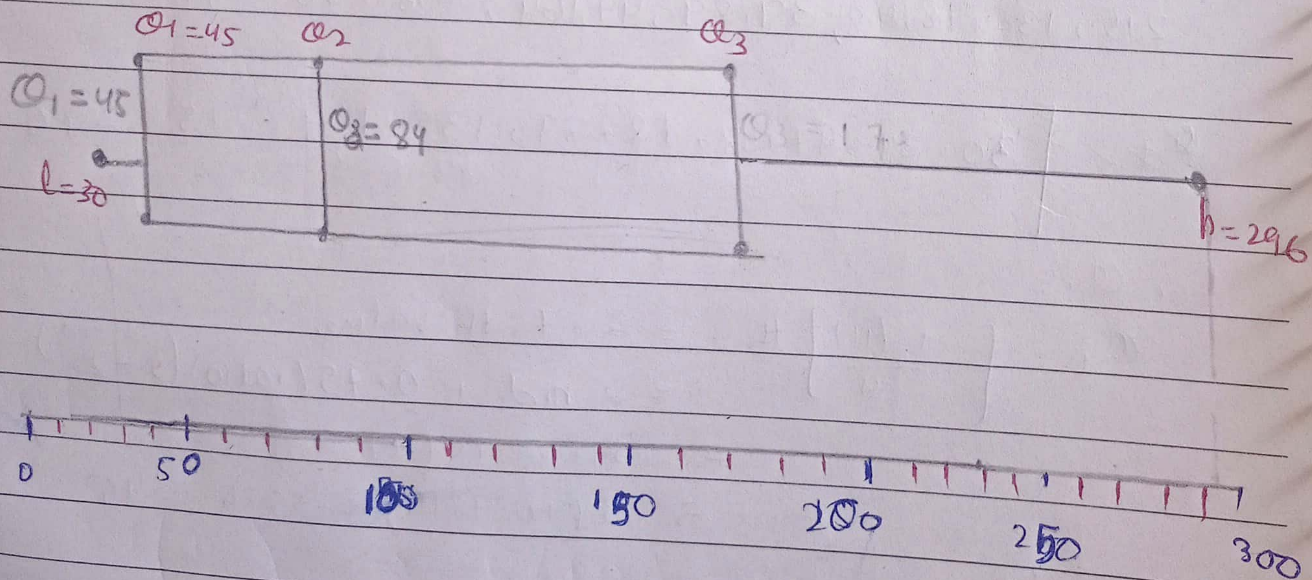
$= 164 + 12.75$

$Q_3 = 176.75 = 177$

Box - PLOT:

l, Q_1, Q_2, Q_3, h

30, 45, 84, 177, 296



Day/ Date:

GROUP DATA (Quartiles, Percentiles, Deciles)

Formulas : (only commulative frequency required)

1) Quartiles :

$l_{i+1} - l_i$

~~breadth of class~~

sum of freq $= \Sigma f$

$$Q_i = l + \frac{h}{f}\left[\left(\frac{i \times n}{4}\right) - C.f_{upp}\right]$$

lower limit of selected class

Freq of selected class

(upper class ki C.F likhy gy)

ii) Percentiles :

$$P_i = l + \frac{h}{f}\left[\frac{(i \times n)}{100} - C.F_{upp}\right]$$

iii) Deciles :

$$D_i = l + \frac{h}{f}\left[\frac{(i \times n)}{10} - C.f_{upp}\right]$$

Scanned with CamScanner

$Q_2 = D_5 = P_{50}$

Day/ Date:

Q. Find P_{20}, P_{50}, Q_2 by using given freq distrn.

width

Groups 2-12, 12-22, 22-32, 32-42, 42-52, 52-62, 62-72, 72-82, 82-92, 92-102.

Freq 2, 5, 8, 12, 15, 20, 16, 14, 10, 8

$n = 110$

Groups	f	C.F
2 - 12	2	2
12 - 22	5	7
22 - 32	8	15
32 - 42	12	27 ← P_{20}
42 - 52	15	42
52 - 62	20	62 ← P_{50} and Q_2
62 - 72	16	78
72 - 82	14	92
82 - 92	10	102
92 - 102	8	110
	$n = 110$	

* For Selection of Classes:

1) $P_{20} = ?$

$\Rightarrow i\left(\frac{n}{100}\right) \Rightarrow 20\left(\frac{110}{100}\right) = 22$ (in C.F)

Day/ Date:

2) P_{50}

$$50\left(\frac{110}{100}\right) = 55$$

3) Q_2

$$2\left(\frac{110}{4}\right) = 55$$

Applying Formula :

i) P_{20} :

$$P_{20} = 32 + \frac{10}{12}\left[\left(\frac{20 \times 110}{100}\right) - 15\right]$$

$$= \frac{197}{6} \cdot \frac{10}{12}\left[22-15\right]$$

$$= \frac{197}{6}(7) = 37.83$$

ii) P_{50} :

$$P_{50} = 52 + \frac{10}{20}\left[\frac{50 \times 110}{100} - 42\right]$$

$$= 58.5$$

iii) $Q_2 = 52 + \frac{10}{20}\left[\frac{2 \times 110}{4} - 42\right] = 58.5$

Day/ Date:

FORMULAS

1) RANGE :

Range = Highest Value - Lowest Value

E.g;

Staff	Salary
Owner	100,000
Manager	40,000
Sales Representative	30,000
Workers	25,000
	15,000
	18,000

Range = 100,000 - 15,000
= 85,000

2) POPULATION VARIANCE & STANDARD DEVIATION : (σ^2, σ)

Formula :

1) Popl = Variance :

$$\sigma^2 = \frac{\Sigma (X-\mu)^2}{N}$$

X = Individual Observation

μ = Population Mean = $\frac{\Sigma X}{N}$

N = Population Size.

Day/ Date:

2) Standard ~~Variance & Standar~~ deviation:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\Sigma (X-\mu)^2}{N}}$$

3) Sample Variance & Standard deviation (s^2, s):

1) Sample Variance:

$$s^2 = \frac{\Sigma (x - \bar{x})^2}{(n-1)}$$

where,

X = Individual Observation

$\bar{x}$ = sample mean = $\frac{\Sigma x_i}{n}$

n = Sample size

2) Standard deviation:

$$s = \sqrt{s^2} = \sqrt{\frac{\Sigma (x-\bar{x})^2}{(n-1)}}$$

4) ~~Standard~~ Short-cut Formulas for Sample Variance and Standard deviation (s^2, s):

1) Sample variance:

$$s^2 = \frac{n(\Sigma x^2) - (\Sigma x)^2}{n(n-1)}$$

Day/ Date:

4) Standard Deviation:

$$S = \sqrt{s^2}$$

5) Sample Variance & Standard deviation for Grouped data (s^2, s):

1) Variance:

$$s^2 = \frac{n(\Sigma f \cdot X_m^2) - (\Sigma f \cdot X_m)^2}{n(n-1)}$$

where,

X_m = Midpoint of each class.

2) Sample Standard dev. for grouped data:

$$s = \sqrt{s^2} = \sqrt{\frac{n(\Sigma f \cdot X_m^2) - (\Sigma f \cdot X_m)^2}{n(n-1)}}$$

6) Coefficient of Variations (for both population & Sample)

(CVar)

1) CVar ~~exad~~ for Samples:

$$CVar = \frac{s}{\bar{X}} \cdot 100\%$$

where, s = Standard Deviation, $\bar{X}$ = mean

2) CVar for population:

$$CVar = \frac{\sigma}{\mu} \cdot 100\%$$

where, σ = standard devi

μ = Mean,

Dazzle

Day/ Date:

Q. A Testing lab wishes to test two experimental brands of outdoor paint to see how long each will last before fading. The testing lab makes 6 Gallons of each paint to test since diff. chemical agents added to each group and only 6 cans are involved. These two groups constitute two small population. The results in months are shown

i) Find Mean of each group ($\mu = \Sigma X/N$)

ii) Find Range of Brand A and B.

iii) Find variance and standard deviation of Brand A

iv) " " " " " B

Brand (A) — 10, 60, 50, 30, 40, 20

Brand (B) — 35, 45, 30, 35, 40, 25

i) μ For Brand A and Brand B:

$\mu_A = \frac{\Sigma x}{N} = \frac{210}{6} = 35$ $\mu_B = \frac{\Sigma x_B}{N} = \frac{210}{6} = 35$

6 → No of gallons.

ii) Range For Both:

Range = Highest - lowest = $60-10 = 50$ (A) $45-25=20$ (B)

iii) Variance and St. dev for A:

$$\sigma^2 = \frac{\Sigma(X_A - \mu_A)^2}{N} = \frac{1750}{60} = 291.666$$

$$\sigma = \sqrt{\frac{\Sigma(X_A - \mu_A)^2}{N}} = 17.078.$$

• Coefficient of Variations:

X	$X_A - \mu_A$	$(X_A - \mu_A)^2$
10	-25	625
20	-15	225
30	-5	25
40	5	25
50	15	225
60	25	625
		$\Sigma 1750$

$$CV_{ar\,A} = \frac{\sigma_A}{\mu_A} \cdot 100$$

$$CV_{ar_A} = 48.79\%$$

Dazzle

Day/ Date:

iv) Variance and st. deviation for Brand B (σ^2, σ):

$$\sigma^2 = \frac{\Sigma(X_B - \mu_B)^2}{N} = \frac{250}{6} = 41.66666$$

$$\sigma = 6.4549$$

X	$X_B - \mu_B$	$\cancel{\Sigma}(X_B - \mu_B)^2$
25	-10	100
30	-5	25
35	0	0
35	0	0
40	5	25
45	10	100
		$\Sigma(X_B - \mu_B)^2 = 250$

Coefficient of Variance

$$CVar\,B = \frac{\sigma_B}{\mu_B} \cdot 100 = 18.44\%$$

Day / Date:

Q. Find the sample variance and standard deviation for the given data.

11.2, 11.9, 12.0, 12.8, 13.4, 14.3 $= \frac{75.6}{6} = 12.6$

$n(n-1) = 6(6-1) = 30$

$$S^2 = \frac{\Sigma(x_i - \bar{X})^2}{(n-1)} = \frac{6.38}{30} = 0.21266$$

$$s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{(n-1)}} = 0.4611$$

X	$X - \bar{X}$	$(X - \bar{X})^2$
11.2	-1.4	1.96
11.9	-0.7	0.49
12.0	-0.6	0.36
12.8	0.2	0.04
13.4	0.8	0.64
14.3	-1.7	2.89
		⇒ 6.38

$$C\,Var = \frac{S}{\bar{X}} 100\,\% = \frac{0.4611}{12.6} \times 100$$

$= \cancel{12.668778} = 3.659968$

Day/ Date:

Q. Data.

classes	5.5-10.5	10.5-15.5	15.5-20.5	20.5-25.5	25.5-30.5	30.5-35.5	35.5-40.5
(f)	1	2	3	5	4	3	2

Find Variance (s^2) and stand. deviation (s).

Solution:

$$s^2 = \frac{n \sum (f \cdot X_m^2) - [\sum (f \cdot X_m)]^2}{n(n-1)}$$

sum of frequency

Classes	X_m	f	$f \cdot X_m$	X_m^2	$f \cdot X_m^2$
5.5-10.5	8	1	8	64	64
10.5-15.5	13	2	26	676	1352
15.5-20.5	18	3	54	2916	8748
20.5-25.5	23	5	115	13225	66125
25.5-30.5	28	4	112	12544	50176
30.5-35.5	33	3	99	9801	29403
35.5-40.5	38	2	76	5476	11552
		$n=20$	$=490$		$\Rightarrow 167420$

$$s^2 = \frac{\cancel{20}(167420 - 240100)}{\cancel{20}(19)}$$

$$s^2 = \frac{-72680}{19}$$

$$\boxed{s^2 = -3825.263158}$$

Day/ Date:

PROBABILITY (walpole)

- Chances of occurences, possible outcome/result.

SAMPLE SPACE (S)

The set of all possible outcomes of a statistical experiment is called sample space and it is denoted by "S".

Q. Find sample space (S) for the given experiment.

1) Toss one coin:

$S = \{T, H\}$

* Probability with same chances called equally directed probability.

2) Roll a dice:

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

3) Answer a True/False:

$S = \{True, false\}$

4) Toss two coins:

$S = \{HH, HT, TH, TT\}$

5) Roll two dice:

$S = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (2,2), (2,3), (2,4), (2,5), (2,6), (3,1), (3,2), (3,3), (3,4), (3,5), (3,6), (4,1), (4,2), (4,3), (4,4), \cancel{(4,5), (4,6),} \ldots (6,6)\}$

Q. Find sample space of die is roll and outcomes are even and odd.

$S = \{even, odd\}$

Dazzle

Scanned with CamScanner

Day / Date:

TREE DIAGRAM

A Tree diagram is a representation or a procedure to find sample space more easily.

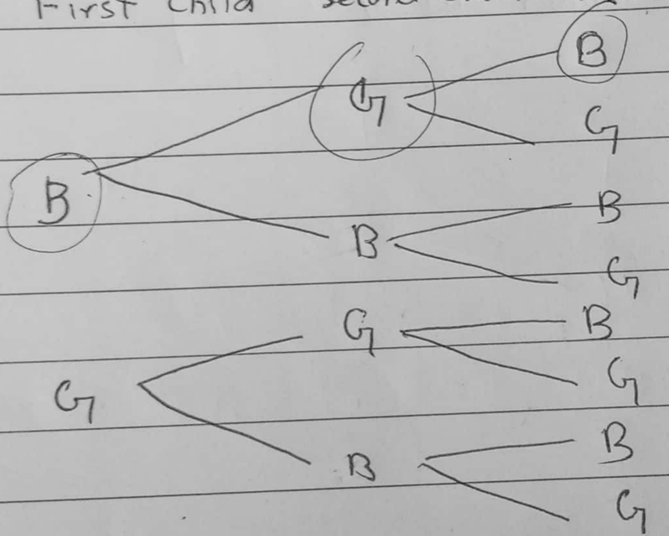
Q. Find a sample space for the gender of three children in a family, and then use tree diagram to solve the same problem.

Possibilities = 2 (Boy or Girl)   Experiment = 3

S = {BBB, GGG, BGG, GBG, GGB, BBG, GBB, BGB}

Tree diagram

First Child   Second Child   Third Child   Sample Points

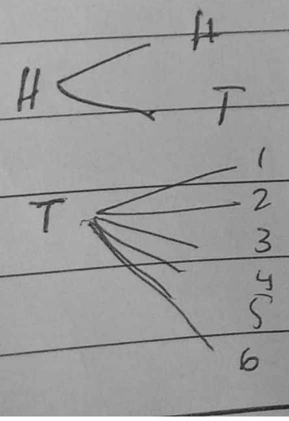


BGB
BGG
BBB
BBG
GGB
GGG
GBB
GBG

S = {BGB, BGG, BBB, BBG, GGB, GGG, GBB, GBG}.

Example 2.2 :

First outcome   Second outcome   Sample Points



S = {HH, HT, T1, T2, T3, T4, T5, T6}

Dazzle

Day/ Date:

EVENTS

Subset of 'S' called Events, represented by E.

$S \subseteq E$

Q. Roll a die and event only evens/odds occurs

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

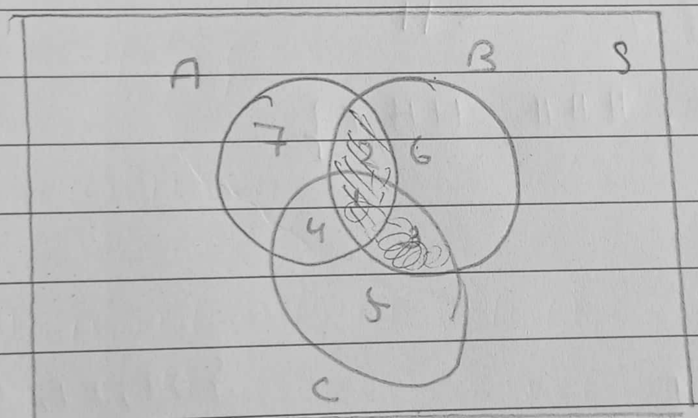
E_1 = Even No $\{2, 4, 6\}$

E_2 = Odd No $\{1, 3, 5\}$

E_3 = No of greater than 3 $\{4, 5, 6\}$.

VENN DIAGRAM

- Sample space by rectangle box
- Events by circle.
- Regions



1) $A \cap B = \{1, 2, 4, 7\} \cap \{1, 2, 3, 6\}$

Region 1 and Region 2

$= \{1, 2\}$

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

$= \{1, 3\}$

Day/Date:

3) $A \cup C = \{1, 2, 4, 7\} \cup \{1, 3, 4, 5\}$

$A \cup C = \{1, 2, 3, 4, 5, 7\}$

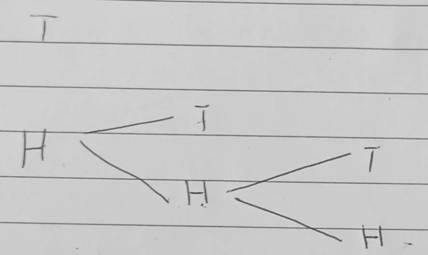
4) $B' \cap A = \{4, 5, 7\} \cap \{1, 2, 4\}$

$= \{4, 7\}$

5) $A \cap B \cap C =$ region 1   pg #42

2.1 (c)

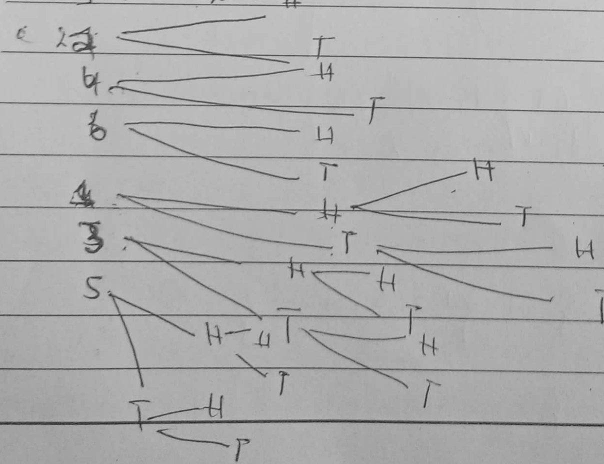
1.0   2.0   3.0



$S = \{T, HT, HHH, HHT\}$

2.4, 2.5

1.0   2.0



$\{2H, 2T, 4H, 4T, 6H,$
$6T, 1HH, 1HT, 1TH, 1TT,$
$3HH, 3HT, 3TH, 3TT,$
$5HH, 5HT, 5TH, 5TT\}$

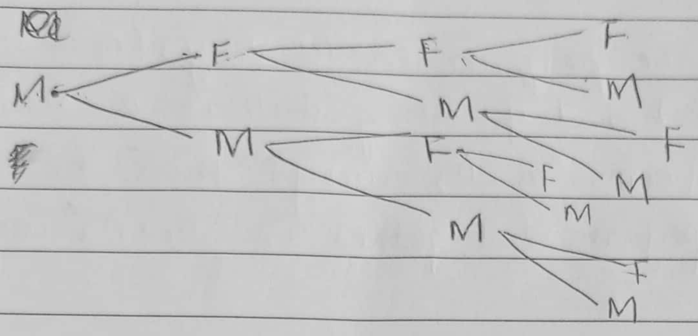
sample points $= 18$

Day/ Date:

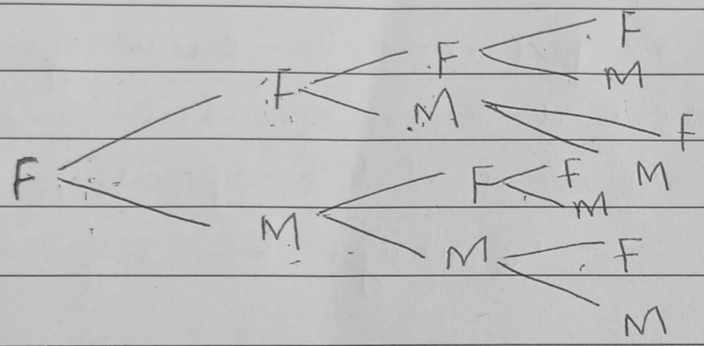
2.6, 2.7

Q2.7

1.0 2.0 3.0 4.0



MFFF, MFFM, MFMF, MFMM, MMFF, MMFM, MMMF, MMMM



FFFF, FFFM, FFMF, FFM,M, FMFF, FMFM, FMMF, FMMM.

2.8

Dazzle

Day/ Date:

sec 2.3

COUNTING SAMPLE POINTS

Rule # 2.1 :

If an operation can be perform in n_1 ways (n_1 = no of sample points or possibilities) and if along with it a second operation can be performs in n_2 ways then the two operations can be performed together in "$n_1 \cdot n_2$" ways.

Rule # 2.2 :

If an operation can be performed in n_1 ways and if for each of these a second operation and third and fourth and so on and k^{th} operation is performed in n_2, n_3 ———— n_k ways then together

$(n_1)(n_2) \cdots (n_k)$ ways

- If we flip '5' coins together, then

$(n_1)(n_2)(n_3)(n_4)(n_5)$

$2 \times 2 \times 2 \times 2 \times 2 = 2^5 = 32$

Q. How many sample points/ways are there in the sample space when pair of dice is rolled one?

$n_1 \times n_2 = 6 \times 6 = 36$

Q. A developer of a new subdivision offers prospective home buyers a choice of tudor, rustic, colonial and traditional exterior styling in ranch, two story, and split level floor plans. In how many diff ways can a buyer order one of these homes?

Day/ Date:

$n_1 = 4$ $n_2 = 3$

Exterior style Floor plans

Tudor Ranch

Rustic two story

colonial split-level

Traditional

$n_1 \times n_2 = 4 \times 3 = 12$ ways.

Q. If a 22 member club needs to elect a chair and a treasurer. In how many diff ways can these be selected.

$n_{CH} = 22$

$n_{Trees} = 21$

$n_{CH} \times n_{treas.} = 462$ different ways.

Q. How many even 4 digit number can be formed from the digits 0,1,2,5,6,9.

First case:

Th	H	T	U
0			0
1			2
2			6
5			
6			
9			

if $u = 0$

$n_u = 1$

$n_{Th} = 5$

$n_H = 4$

$n_T = 3$

$(1)(5)(4)(3) = 60$

There are '60' possible ways to get 4 digit even no when units is at 0.

Dazzle

Day/ Date:

2nd case:

if $U \neq 0$, let $U = 2$

$n_U = 1$

$n_{Th} = 4$

$n_H = 4$

$n_T = 3$

$(n_{TH})(n_H)(n_T)(n_U) = (4)(4)(3)(1) = 48.$

for non zero total $= 96$

Total ways $= 60 + 96 = 156$ ways.

There are '96' possible ways to form 4 digit even no when 2 or 6 is at "U".

1) Multiplication Rule:

Ek saath occur herhy hon to.

And $\Rightarrow (n_1)(n_2)(n_3)$

2) Additive Rule:

Or $\Rightarrow (n_1) + (n_2) + (n_3)$

n^{k} $k \to$ Possible, $n \to$ spaces (when repetition)

2.3: PERMUTATION

→ Order matters

→ All possible combinations.

A permutation is an arrangement of "All" or "Part" of ~~an obj~~ "n-objects"

Q. a, b, c: i) Find all Possible Permutations. ii) No of Permutations.

⇒ ii) $3! = 6$ i) {abc, acb, bac, bca, cab, cba}

no of permutations of n-objects

Dazzle

Day/ Date:

- By Multiplication Rule:

$\underset{n_1}{\underline{3}} \times \underset{n_2}{\underline{2}} \times \underset{n_3}{\underline{1}} = 3! = 6$

• no of permutations of n-objects $= n!$

$\underset{n_1\ n_2\ n_3}{3 \times 3 \times 3} = 27$

1) No of permutations $= n!$

2) No of permutations of n-objects = taken "r" at a time =

$^nP_r = n!/(n-r)!$

Q. a, b, c, d find all possible permutations if taken '2' at a time.

$\Rightarrow$ {ab, ac, ad, ba, bc, bd, ca, cb, cd, da, db, dc}

no. of permutations taken '2' at a time $= 12$

$\Rightarrow \underset{n_1}{\underline{4}} \times \underset{n_2}{\underline{3}} = 12$

$\Rightarrow {}^nP_r = \frac{n!}{(n-r)!} = \frac{4!}{(4-2)!} = \frac{4!}{2!} = 12$

$^nP_n = \frac{n!}{0!} = n!$

$\Rightarrow \frac{n!}{n_1!\, n_2! \cdots n_k!}$ (Permutation of n objects) (repeatition)

Pehle kinds mei jo points ruky hain.

Day / Date:

2.4 ki example 20 : football wali example.

$$\Rightarrow \frac{10!}{1!\,2!\,4!\,3!} \Rightarrow 12600$$

2.5

$$\binom{n}{n_1, n_2, \ldots n_r} = \frac{n!}{n_1!\, n_2! \ldots n_k!}$$

No of ways of partioning set of n objects.

21.

Eg $\binom{7}{3, 2, 2} = \frac{7!}{3! \times 2! \times 2!} = 210$

COMBINATIONS

$$\binom{n}{r} = \frac{n!}{r!\,(n-r)!}$$

Day/ Date:

2.4 PROBABILITY OF AN EVENTS

- Probability always lie b/w 0 and 1.
- " " " taken of events.

$P(\emptyset) = 0$

$P(E) \Rightarrow 0 \leq P(E) \leq 1$.

$P(S) = 1$ (whole sample space ki)

- $0.5 =$ moderate $0.2 =$ less chance $1 = 0.8 = 0.9$ More chance.

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

$A \cup B \neq A + B = (1, 2, 3, 4, 5) \neq S$

If $A_1, A_2, \text{———} A_n$ are mutually exclusive Events

No common event

then;

$P(A_1 \cup A_2 \cup A_3 \cup \text{———} \cup A_n) = P(A_1) + P(A_2) + \text{———} P(A_n)$

~~$P(A \cup B) = P(A) + P(B) - P(M)$~~

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

Not mually exclusive events.

Q. A coin is tossed twice. What is the probability that head occures atleast 1?

$S = \{HH, HT, TH, TT\} \rightarrow N = 4$

$P(A) = P($atleast one head$)$

$A = \{HH, HT, TH\}$ $n = 3$

$P(E) = \frac{n}{N} = \frac{3}{4} = 0.75$

Dazzle

Day/Date:

Q. A die is ~~loaded~~ loaded in such a way that an even no is twice as likely to occur as an odd. If 'E' is an event that a no less than '4' occurs on a single toss of a die. Find probability of E.

$S = \{1, 2, 2, 3, 4, 4, 5, 6, 6\}$ $N = 9$

$E = \{1, 2, 2, 3\} \rightarrow n = 4$

$P(E) = \frac{n}{N} = \frac{4}{9} =$

Q. Referred to previous example let A be the event that even no turns up and let B be the event that no divisible by 3. Find Probability of $A \cup B$ and $A \cap B$?

$S = \{1, 2, 2, 3, 4, 4, 5, 6, 6\}$ $N = 9$

$P(A \cap B) = P(A \& B)$

$A = \{2, 2, 4, 4, 6, 6\} \Rightarrow n = 6$

$B = \{3, 6, 6\} \Rightarrow n = 3$

i) $P(A \cup B) = ?$

$(A \cup B) = \{2, 2, 3, 4, 4, 6, 6\} \rightarrow n = 7$

$P(A \cup B) = \frac{n}{N} = \frac{7}{9} =$

2) $P(A \cap B) = ?$

$A \cap B = \{6, 6\}$, $n = 2$

$P(A \cap B) = \frac{2}{9}$

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

$= \frac{6}{9} + \frac{3}{9} - \frac{2}{9}$

$= \frac{7}{9}$

Day/Date:

2.5 " ADDITIVE RULES "

1) If A and B are two events

$$P(A \cup B) = P(A \text{ or } B) = P(A) + P(B) - P(A \cap B)$$

2) If A and B are mutually exclusive

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

3) If $A_1, A_2, \ldots, A_n$ Mutually exclusive

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + P(A_3) + \ldots + P(A_n)$$

4) If $A_1, A_2, \ldots, A_n$ are partition of 'S'

$$\Rightarrow P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(S) = 1$$

5) If A and A' then

$$P(A) + P(A') = 1$$

$\Rightarrow$

$P(A) = 1 - P(A')$, if A' is found easily

Scanned with CamScanner

Day/Date:

Q. John is going to graduate from an industrial engineering department in a university by the end of semester after being interviewed at two companies he likes, he ~~assess that~~ assesses that his probability of getting an offer from company A is 0.8 and his prob of getting an offer from comp B is 0.6, If he believes the prob that he will get offers from both companies is 0.5. What is the prob that he'll get atleast '1' offer from these two companies?

$\Rightarrow$ $P(A) = 0.8$ $P(B) = 0.6$ $P(A \cap B) = 0.5$

$P(A \cup B) = 0.8 + 0.6 - 0.5 = 0.9$

Q. What is the prob. of getting a total of '7' or '11' when a pair of dice is rolled.

$A = \{(1,6),(6,1),(2,5),(5,2),(4,3),(3,4)\}$ $n = 6$

$B = \{(5,6),(6,5)\} \Rightarrow n = 2$

$P(A \cup B) = ?$

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

$P(A) = \frac{n}{N} = \frac{6}{36}$ $P(B) = \frac{n}{N} = \frac{2}{36}$

$P(A \cup B) = 1/6 + 1/18 = 2/9$

Dazzle

Day/ Date:

Q. If the probabilities are respectively 0.09, 0.15, 0.21, and 0.23 that a person purchasing a new automobile will choose the color green, white, red or blue what is the probability that a given buyer will purchase a new automobile that comes in one of those color.

=)

$$P(A_1 \cup A_2 \cup A_3 \cup A_4) = 0.09 + 0.15 + 0.21 + 0.23$$
$$= 0.68$$

Q. If the probabilities that an automobile mechanic will service 3, 4, 5, 6, 7, or 8 or more cars on any given workday are respectively 0.12, 0.19, 0.28, 0.24, 0.10 and 0.07. What is the probability that he will service '5' cars on next day.

(kam nhi horha or less hai to uses cauplement vala formula).

$$P(A) + P'(A) = 1$$
$$P(A) = 1 - P(A')$$
$$= 1 - 0.31$$
$$P(A) = 0.69$$

Q. Suppose the manufacturer's specifications (Known Values) for the length of the certain type of computer cable are 2000 ± 10 millimeter

$1990 \leq X \leq 2010$. In this industry it is known that small cable is just as likely to be defective (not meeting specifications) as large

Day/ Date:

cable, there is the probability of randomly producing a cable with length exceeding 2010 millimeters is equal to the prob of producing a cable with length smaller than 1990.

The Prob that the production procedure meets specifications is known $P(S) = P(L)$ to be 0.99.

$$P(M.s) = P(1990 \le x \le 2010) = 0.99$$

a) what is the probability that a cable selected randomly is too large?

$$P(L) = 1 - \{P(L') = P(N.M.s)\} = 1 - 0.99 \} =$$

$$= 1 - 0.99 = 5 \times 10^{-3} \qquad P(L) = P(S) = 5 \times 10^{-3}$$

2(2 possibilities this)

b) what is the prob that a randomly selected cable is larger than 1990.

$$P(x > 1990) = (0.99 + 5 \times 10^{-3}) = 0.995$$