Date :- 22/12/21

STATASTICS

Measure + Analyse.

* Descriptive (Describe) Data.
* Inferential (Estimate)
  ↓
  inference

(Exploratory Data Analysis)
Analyse Each column
EDA - 60%

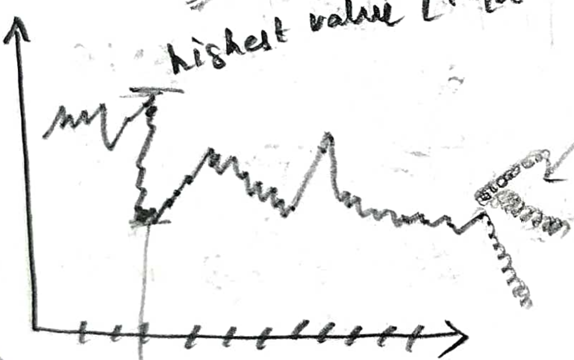
4 Types of Analytics

* Descriptive Analytics
* Diagonistic Analytics
* predictive Analytics
* prescreptive Analytics
  ↓
  Reaching Your estimation

X - health problem
↓
doctor
1. Describe Health Problem
2. Scanning
3. what is Your probl (Estimate)
4. Medicines.

highest value (Max. values

Share Value



Max. Value.
min.
Average.

Infosys 1400/-

Reason of Fall

* Descriptive* Statistics

- collecting, presenting, Describe data

* Inferential* Statistics → Sample Data.

- Drawing Conclusion

* Random Sampling (Probability Sampling)
* Non-Random Sampling (Bias Sampling)

→ Simple Random Sample (Blindly)

→ **** Stratifed (spilit Data into parts) (strata)  Population : 70 students

20 G   50 Bo

* proportionate → 2:5
* Dis proportionate → Sample 1:25 → 4 girls : 10 boy

→ Systematic Random Sample

100 students
1
11
21
31
41
51
61

→ Cluster (or Area) Sampling (parts) (Geographic locations)

Corona +ve Cases

- Green zone
- red zone
- orange zone

100 + red
50-100 (orange
<50 green

* not compulsory
*

N = population size
n = sample size

- N = 20
- n = 4

## Non. Random Sampling

Your

* Convience Sampling (As per ↑ Convimce) Bias

Ex:-

* Judgement Sampling police x — criminal
x — innocent

Ex:-

* quota Sampling * class - 40/only /40 chairs

* Snow ball Sampling Ex: (references/IT) Eaiy qustions

sek → (x) → (y)

*** Imp

Variability → column names / Features → machine Learnin

rows → records

| Date | Name | Email | phone |
|---|---|---|---|
| | | | |

Variable → Continous Variable, Descrete, Time Series Variable

Continous variable

Decimals

Ex: Age
Height
Temperature

Discrete variable

No Decimals

1. Discrete Count Data
(No. decimal)

2. Categorical Data
Text
* Nominal (No Natural sequence) Default
* Ordinal (Natural sequence)

Name / Colours
UK, India, U.S, France, Ger
White, Black, Pink

Small / medium / Large
XXL
XXXL

Letter grade
A A+ B B+

Time Series

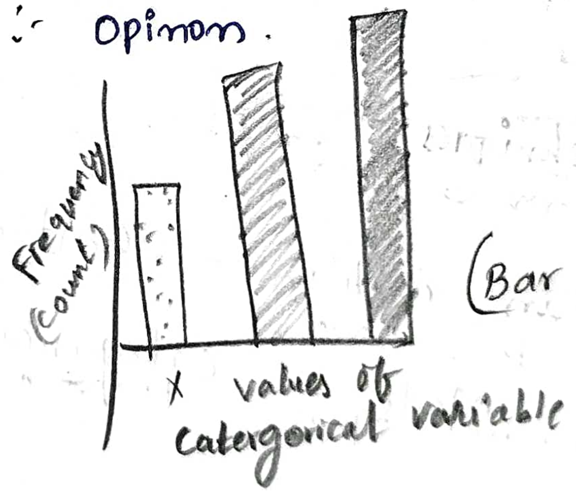
Time / Date

* Data
- Numeric
- Text
- Time Series

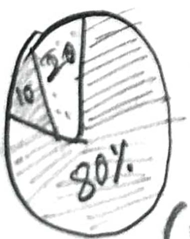
* quantitive variable :- Text (numeric)

* qualitive variable :- Opinon.

Categorical Data
(Discrete variable)

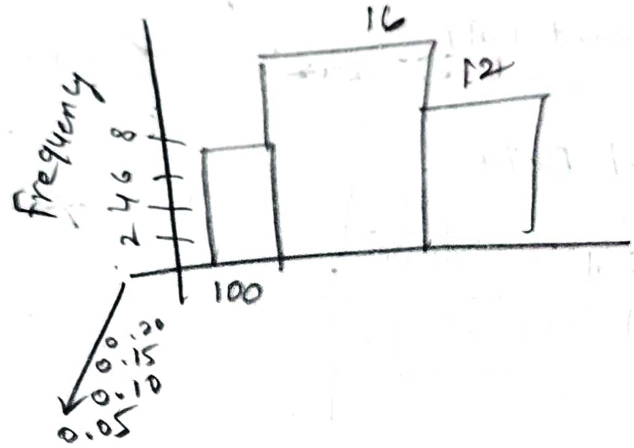



(Bar graph)




(pie chart)

⇒ Frequency Distribution / Histogram




110 – 120

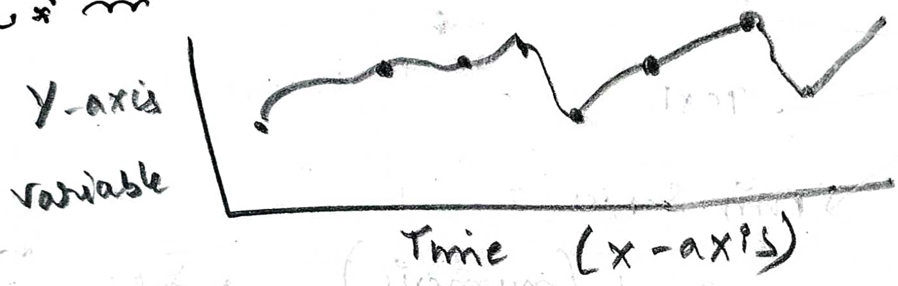
120 – 130

included

Right side is not included

⇒ Relative Frequency : Percentages

⇒ Timeplot




****

⇒ Sampling Techniques

1. Mean, median – Continous
2. Mode – Categeorical Data / Discrete

DaTe:- 23/12/21

# STATISTICS

1. Measures of Central Tendency ($1^{st}$ Bussiness Moment)
(mean, median, Mode)

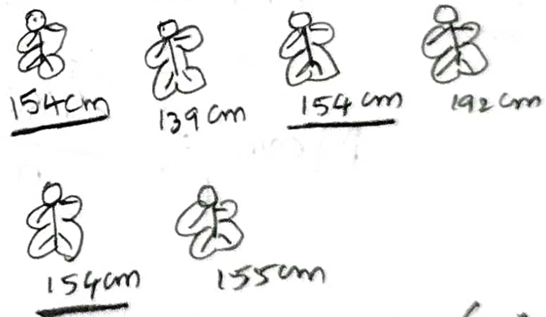
2. Measures of Spread / dispersion ($2^{nd}$ Bussiness Moment)
(range, Variance, STandard deviation)
Overall distribution Data.

3. Measures of Shapes

* Skiwness ($3^{rd}$ Bussiness Moment)
* Kurtosis ($4^{th}$ Bussiness Moment)

* Mode :- refers To The data Value That is most Frequently Observed.
repeated.

Ex :- Measures of Centre - Mode




Mode = 154

* Median :- Refers To The data value That is positioned in middle of an ordered data set.

Mode

Ex:-
3 * 154, 154, 154 → 154 (uni modal)
3 155, 155, 155 → 155 (Bio modal)
3 156, 156, 156 → 156 (Multi modal)

only 1
4
155, 155, 155, 155

Median :- Refers To The Data Value That is positioned in The "middle" of an Ordered Data set.

Focusing on

Ex :-

n = 9

139 (1)
140 (2)
154 (3)
154 (4)
$\boxed{154}$ (5)
155 (6)
180 (7)
192 (8)
192 (9)

if n=10

139
140
154
154
154
155
180
192
192
196

Take average

$\frac{154+155}{2}$

$= 154.5$

Formula :-

no. of data Values

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

Position of The median

if $\left(\frac{n+1}{2}\right) = \frac{10+1}{2} = \frac{11}{2} = 5.5$

(it is not Compulsory, 7th Median value Data Set Vundali ani Ledu)

Kani

Mode 10 Vundali

* Mean :- Arthematic Average

sigma

$$\text{Mean}/\text{Avg} = \boxed{\frac{\Sigma x_i}{n}}$$

$\frac{\text{Summation of all data Values}}{\text{Total no. of Data Values}}$

$\therefore$ Sample Mean $= \bar{x}$

Mue

Population Mean $= \mu = \frac{\Sigma x}{N}$

capital

Ex:- Population Mean

Data Values :- 10, 12, 18, 4, 2

$$\boxed{\mu = \frac{\Sigma x}{N}} = \frac{10+12+18+4+2}{5}$$

$$\mu = \frac{46}{5}$$

$$\boxed{\mu = 9.2}$$ → population mean.

Sample Mean

Data values :- 10, 12, 18, 4, 2

Sample Kabati (12, 18)

$$\boxed{\bar{x} = \frac{\Sigma x}{n}} \qquad \bar{x} = \frac{12+18}{2}$$

$$\bar{x} = \frac{30}{2} = 15$$

$$\bar{x} = 15.$$

Mean/Average :-

| $n=10$ |
|---|
| 139 |
| 140 |
| 154 |
| 154 |
| 154 |
| 155 |
| 180 |
| 192 |
| 192 |
| 196 |

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

$$\bar{x} = \frac{139+140+154+154+154+155+180+192+192+196}{10}$$

$$\boxed{\bar{x} = 165.6}$$

* Weighted Mean

$$= \boxed{\frac{\Sigma W_i X_i}{\Sigma W_i}}$$

Ex:-

Internal → 40%

External → 60%

∴ Internal Exam (40%)

$\frac{100}{80 \text{ (scored)}} \rightarrow 32\%$

External Exam (60%)

$\frac{100}{90 \text{ (scored)}} \rightarrow 54\%$

$\frac{100\%}{86\% \text{ (scored)}}$

(2) 1, 2, 2, 2, 4

$$= \frac{1 + 3(2) + 4}{5}$$

* Range :-

Range = Maximum Value − Minimum Value

Ex:-

$n = 6$

139 → Min

140

154

196 → Max

192

150

$$\boxed{\text{Range} = \text{Max} - \text{Min}}$$

= 196 − 139

= 57

****

# Variance

Ex :-

| | Sachin Runs | Dhoni (X) Runs | Deviation Sachin) $(x-\bar{x})$ | Deviation Dhoni $(x-\bar{x})$ |
|---|---|---|---|---|
| 1st match | 80 | 48 | 80 - 40 $\boxed{=+40}$ | 48-40 $\boxed{=+8}$ |
| 2nd match | 0 | 42 | 0 - 40 $\boxed{=-40}$ | 42-40 $\boxed{=+2}$ |
| 3rd match | 12 | 38 | 12 - 40 $\boxed{=-28}$ | 38-40 $\boxed{=-2}$ |
| 4th match | 68 | 40 | 68 - 40 $\boxed{=+28}$ | 40-40 $\boxed{=0}$ |
| 5th match | 40 | 32 | 40 - 40 $\boxed{=0}$ | 32-40 $\boxed{=-8}$ |
| Sum = | 200 | 200 | ToTal = 0 | ToTal = 0 |

mean/Avg = $\frac{80+0+12+68+40}{5} = 40\bar{x}$

Avg: Dhoni Runs = $\frac{48+42+38+40+32}{5} = 40\bar{x}$

Deviation = $(x-\bar{x})$

Mean Deviation = $\frac{\Sigma(x-\bar{x})}{n}$

If Sachin/dhoni deviation = $\frac{\Sigma(x-\bar{x})}{n}$

$= \frac{0}{5}$

$= 0$

(-) value ni (+ve) value chesthey deviation vasthadi

Machine Learning

* Eppudina Deviations Takkuva Vundali

* Convert (-) = +ve (value

* Modulus | |

|-ve| = +ve

* Squaring ( $^2$ )

$(-ve)^2 = +ve$

* Deviation = $(x-\bar{x})$

*****

* If Your adding (absolute) to Deviation, it is called

||

$$\boxed{\text{Absolute Deviation} = |x-\bar{x}|}$$

****

* If your squarring The Deviation, is called as

$$\boxed{\text{Square Deviation} = (x-\bar{x})^2}$$

****

∴ Deviations Should be "Minimum".

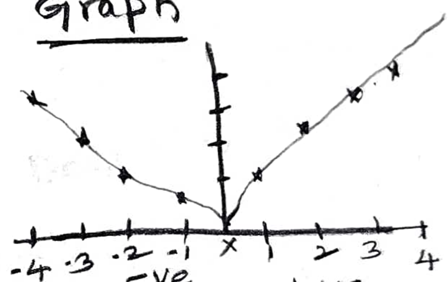
Imp

Max = $\dfrac{d(y)}{dx}$ at $x = 0$

- If you Take Absolute deviation

$= \dfrac{d(|x-\bar{x}|)}{dx}$ at $x = 0$ → $\{\infty\}$ infinite

Graph




∴ $y = |x|$ → Absolute

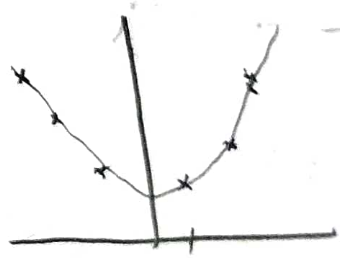
Graph to $\dfrac{dy}{dx}$ at $x = 0 \Rightarrow \infty$

| | |
|---|---|
| $y = +1$ | $y = -1$ |
| $y = +2$ | $y = -2$ |
| $y = +3$ | $y = -3$ |

- If you Take square deviation

$= \dfrac{d((x-\bar{x})^2)}{dx}$ at $x = 0$ → $2x = 0$

Graph



∴ $\dfrac{dy}{dx} = 2x$ at $x = 0$ ← $y = x^2$

| | | |
|---|---|---|
| $+1 = 2$ | $+3 = 9$ | $-1 = 2$ |
| $+2 = 4$ | | $-2 = +4$ |
| | | $-3 = +9$ |

∴ Dhoni Runs Deviation

| Deviation $(x-\bar{x})$ | In Square deviation $(x-\bar{x})^2$ |
|---|---|
| +8 | 64 |
| +2 | 4 |
| -2 | 4 |
| 0 | 0 |
| -8 | 64 |
| 0 | 136 |

∴ Mean Square Deviation

$$Var = \frac{\Sigma(x-\bar{x})^2}{}$$

They Derived Formula n From This :

$$\boxed{\text{Variance} = \frac{\Sigma[x-\bar{x}]^2}{n}}$$

$$\therefore \boxed{\text{Variance} = \frac{\Sigma(x-\bar{x})^2}{n}}$$

→ $(x-\bar{x})$

Runs = 48 - 40

R = 8

In order To Remove Square Add Route

$$\text{Variance} = \frac{\Sigma(8R)^2}{n} = \frac{\Sigma(64R^2)}{n} = \sqrt{\frac{\Sigma 64R^2}{n}}$$

Then, it called as,

$$* \text{ Standard Deviation} = \sqrt{\frac{\Sigma(x-x)^2}{N}}$$

Consistancy

Low Vundali

$= \_\_ R$

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

Anduky squaring chestham

$(x-\bar{x})^2 = ✓$

* Variance in

## → Population Variance

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

sigma — mue

From Center

Standard . Deviation.

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

Entire population Data Vuntey 'work chestham.

## → Sample Variance

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

**** Imp Besel Correction *

Standard deviation

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

view

* Besel correction * :- Why $n-1$ ? / why not $n$ ?

Ex :- $x+y+z = 90$

Eppudu $x+y+z$ (infinite options)

$30+y+z = 90$

Same so $y+z$ (infinite options)

$30+20+z = 90$

Eppudu manaki okatey option Vundhi Adhi $z = 40$

2) T Shirts :- Enni options Vunnaye Choice

* Mon – 7 (No repeat)
* Tue – 6 options
* Wen – 5
* Thu – 4
* Fri – 3
* Sat – 2
* Sun – 1 (Mandatory) ($n-1$)

Variance *

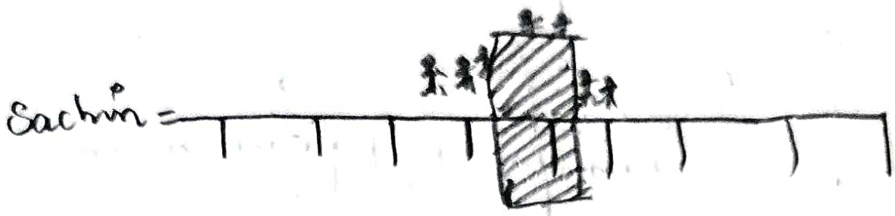
SRK — 8L, 8L 9L 1L

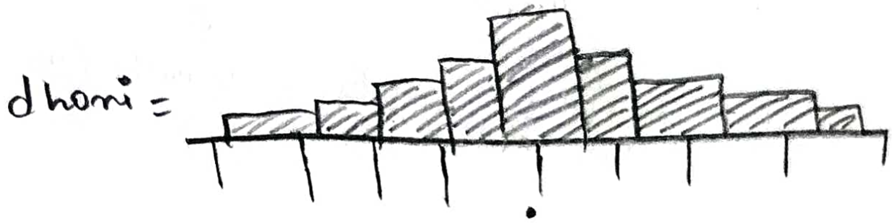
KSL (Trainer) — 8L, 8L, 3L, 4L

Data points Eppudu "Mean" ki dhaggara Vundali → Best

* What Does STanDarD Deviation Tell us ?

A: How close The Value in DaTa Set are to The Mean/Average.




* Percentile :- position value

Describes The PercenTage of DaTa Values That Fall at $\boxed{\text{or}}$ Below another DaTa Value

Ex :- 70 students

65 - Nakantey Takkuva/below vunnaru study

$= \frac{65}{70}$ = 10 % percentaile / Top '10' position

Percentage :
* Score = $(31/50)(100) = 62\%$

* Percentile (Position wise)
- Roughly - 23rd percentile

Ex :- Class size ⇒ $n = 15$

Scores out of 50

* Ascending order

50th percentile

15 20 25 31 34 34 35 35 36 36 36 38 38 39 43

25% percentile ⇓ $Q_1$ = Quartile (1)

Median ↓ Quartile (2)

75% percentile ⇓ → $Q_3$ = Quartile (3)

* quartile Locations

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 65,600 | 73,600 | 500,000 | 70,400 | 54,000 | 99,200 | 94,700 | 54,000 | 78,800 | 91,200 | 29,500 | 80,400 |

- Ascending Order lo Pettali ↵

↑ 50th percentile
Quartile (2)

6.5
$= 6 + .5(7-6)$
$(73,600 + .5(78,800 - 73,600)$
$= 76,200$

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 29,500 | 54,000 | 54,000 | 65,600 | 70,400 | 73,600 | 78,800 | 80,400 | 91,200 | 94,700 | 99,200 | 500,000 |

Median ⇓
$\frac{6+7}{2} = 6.5$

25th Percentile
Quartile(1) 3.25
$= 3 + 0.25(4-3)$
$= 54,000 + 0.25(65,600 - 54,000)$
$= 56,900$

75th Percentile
Quartile(3)
9.75
$= 9 + 0.75(10-9)$
$= 93,825$

## * Location Formula

→ To find Accurate Percentile.

$$\boxed{L_p = \frac{P}{100}(n+1)}$$ → Calculation of Percentile Value.

Ex :- Location of quartiles

$n = 12$

We want $25^{th}$ position percentile.

$\therefore L_{25^{th}} = \frac{25}{100}(12+1) = 3.25$

$\therefore L_{50^{th}} = \frac{50}{100}(12+1) = 6.5$ (or) $\frac{6+7}{2} = 6.5$

$\therefore L_{75^{th}} = \frac{75}{100}(12+1) = 9.75$ Analysis

Ex :-

3 .25 .50 .75 4

We want 3.25 Value

Calculate :- $3 + (4-3) \times 0.25$

* App. One quater (25%) of Workers Have Salaries at / below = \$ 56,900
* App. Half (50%) of workers have Salaries at or below = \$ 76,200
* App. Three forth (75%) of workers Have Salaries at or below = \$ 93,825

## * Interpreting quartiles

25 Percentile, 50 percentile, 75 Percentile

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 29,500 | 54,000 | 54,000 | 65,600 | 70,400 | 73,600 | 78,800 | 80,400 | 91,200 | 94,750 | 99,200 | 500,000 |

56,900 76,200 93,825

* Percentile of Value

Q: What is you would Like To know The percentile of a Specific Value ?

$$\text{Percentile} = \frac{x + 0.5y}{n}$$

Percentile of Data (5) = 70,400 (in quartile)

$$\text{Percentile}_{70,400} = \frac{x + 0.5y}{n}$$

46,500 70,400 73,500

okatey value yundi

1 2 3 4 5

Kantey Enni Takkuva Vunnaye

$$= \frac{4 + 0.5(1)}{12}$$

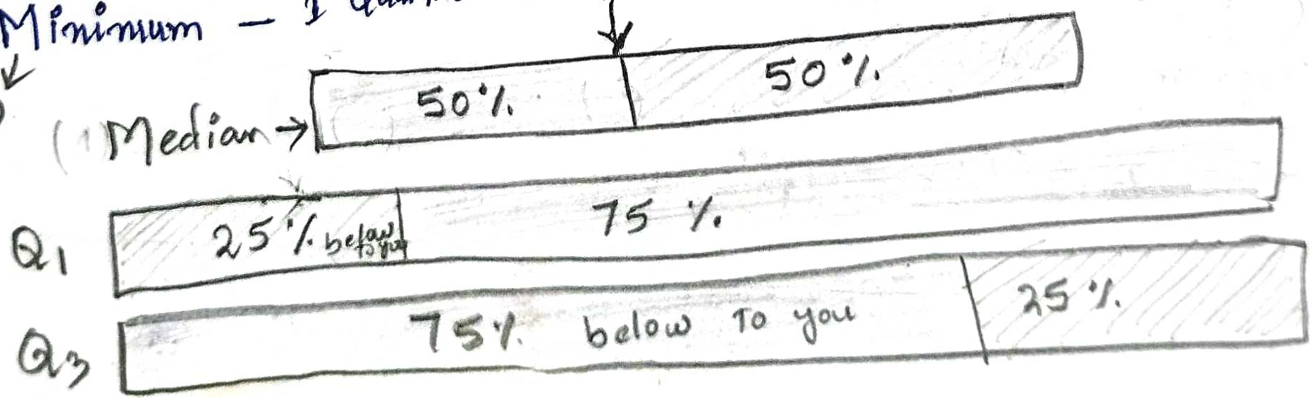
$$= \frac{4.5}{12}$$

$$= 0.375 \Rightarrow 38^{th}.$$

* Five number Summary

Minimum — 1st Quartile — Median — 3rd Quartile — Maximum

below to you 0 ... below to you < 100



* **Five number Summary**

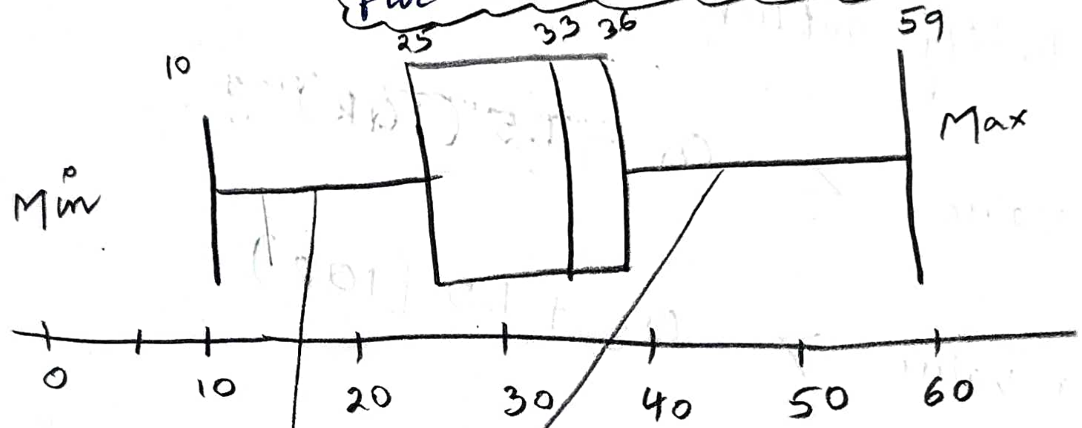
Minimum — 1st Quartile — Median — 3rd Quartile — Maximum

Minimum: 10

Maximum: 59

10 11 12 25 $\boxed{25}$ 27 31 $\boxed{33}$ 34 34 35 $\boxed{36}$ 43 50 59

* **Box plot** :- Gives us a visual Representation of The → Five number Summary.




Whiskers

∴ Box whisker plot

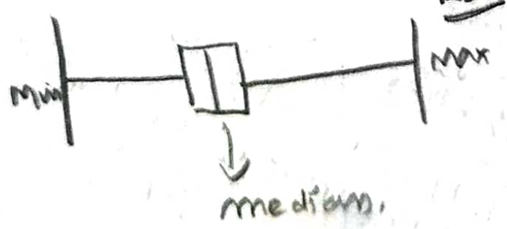
∴ interquatile range *****

$IQR = Q_3 - Q_1$

Ex:-

SRK → Having Max. no. of students


KSL



Date :- 24/12/21 Fri

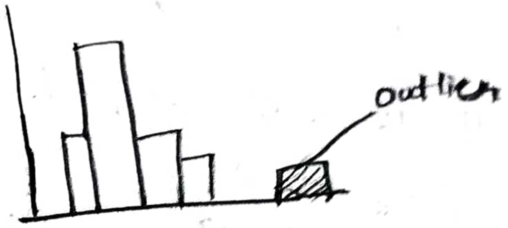
STATISTICS

25 Sat, Su Holiday

→ Outlier

Numerically Distant From a Data Set.

Ex :-




* "Mean" is getting mainly affected by "outlier

* How To indentify outlier

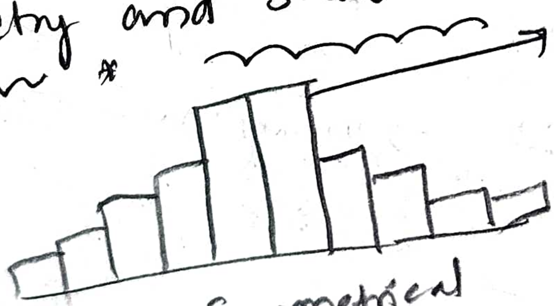
→ Data Value $<$ $Q_1 - 1.5\,(IQR)$

→ Data value $>$ $Q_3 + 1.5\,(IQR)$

* Outlier = From Boxplot can be indentify by "•"

→ Symmetry and Skewness → mean = median.

*



Symmetrical

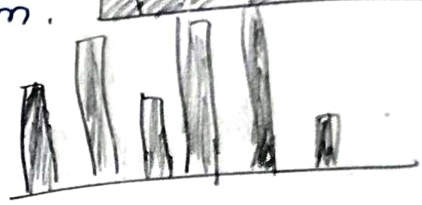
* Asymmetry (non - symmetry)

Due class 24/12/21 (Outlier, Skewness, Symmetric...).

Measured.

* Skewness

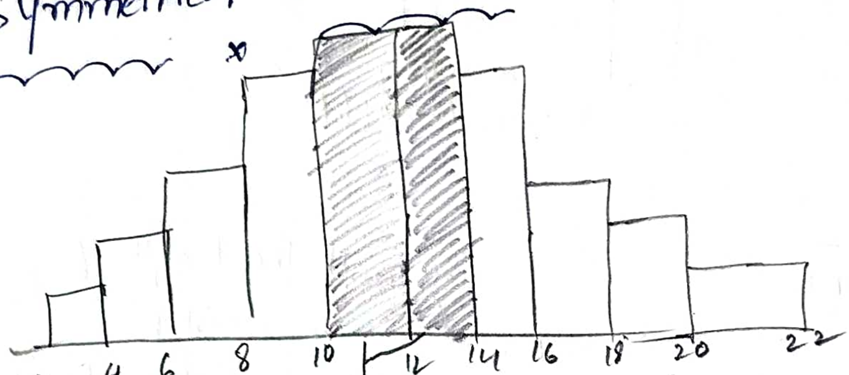
→ Continuous Data ⇒ Histogram.



→ Discrete Data ⇒ Bargraph

Counted

TCS
* Symmetrical Distribution



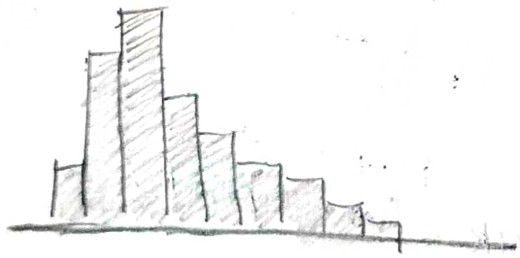
(Mean = Median)

* Equal vantage

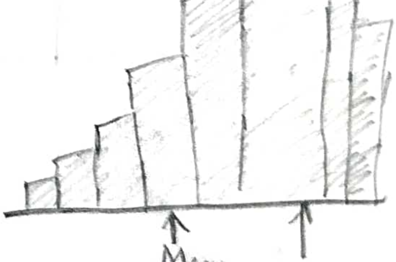
* Mean = Centre of g...
* Median = Centre in points

* Asymetry :- (Not symetrical) ⇒ Refers To "Sweknes"

Skewed To right



Skewed To left.



Mean Median

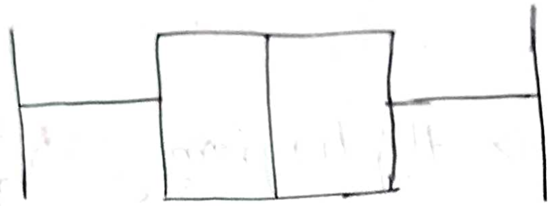
⇒ For Right Skewed :-

Median is Less Than Mean (Median < Mean)

TCS
→ For Left Skewed :-

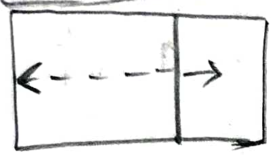
Mean is Less Than Median (Mean < Median)

## * Skewness in Boxplot :-

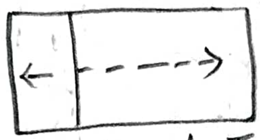


* Equal Boxes with Same whisker length.

* Un Equal Boxes

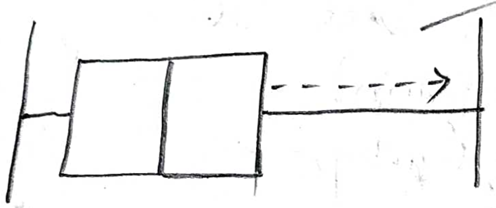


→ Skewed To Left



→ Skewed To right

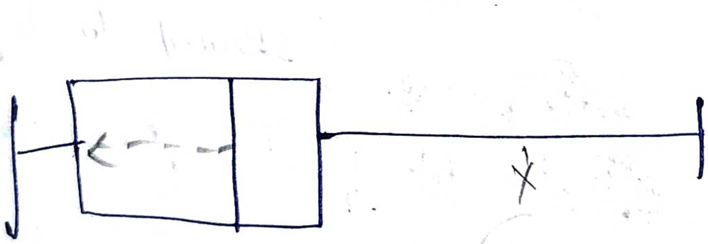
(or)



→ with The length of The whisker can also determined.

Skewed To right

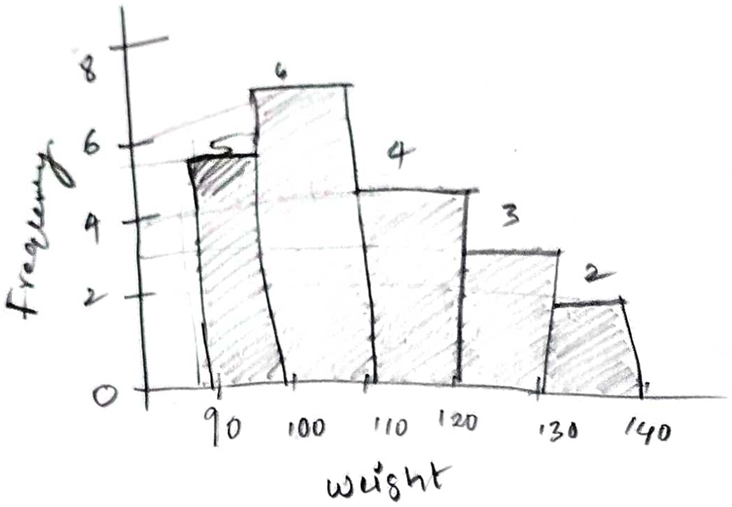
Example

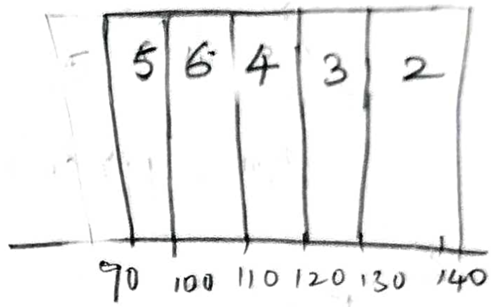


its Left Skewed

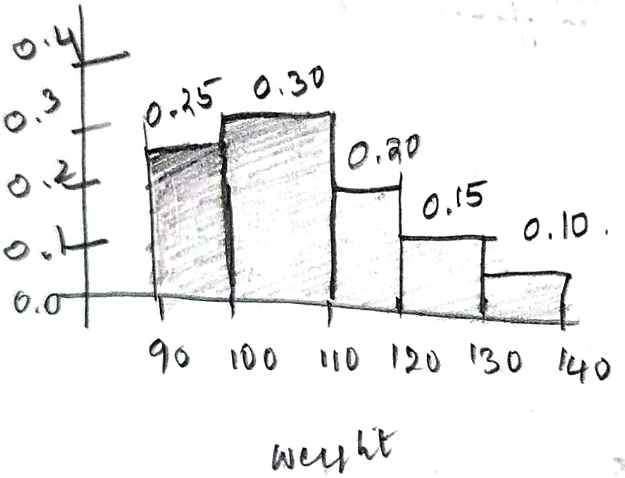
Because we should only check Box First Than whisker

## * Regular Frequency Distribution :-




## * Relative Frequency Distribution




## * Probability Of Given Data

$5+6+4+3+2 = 20$

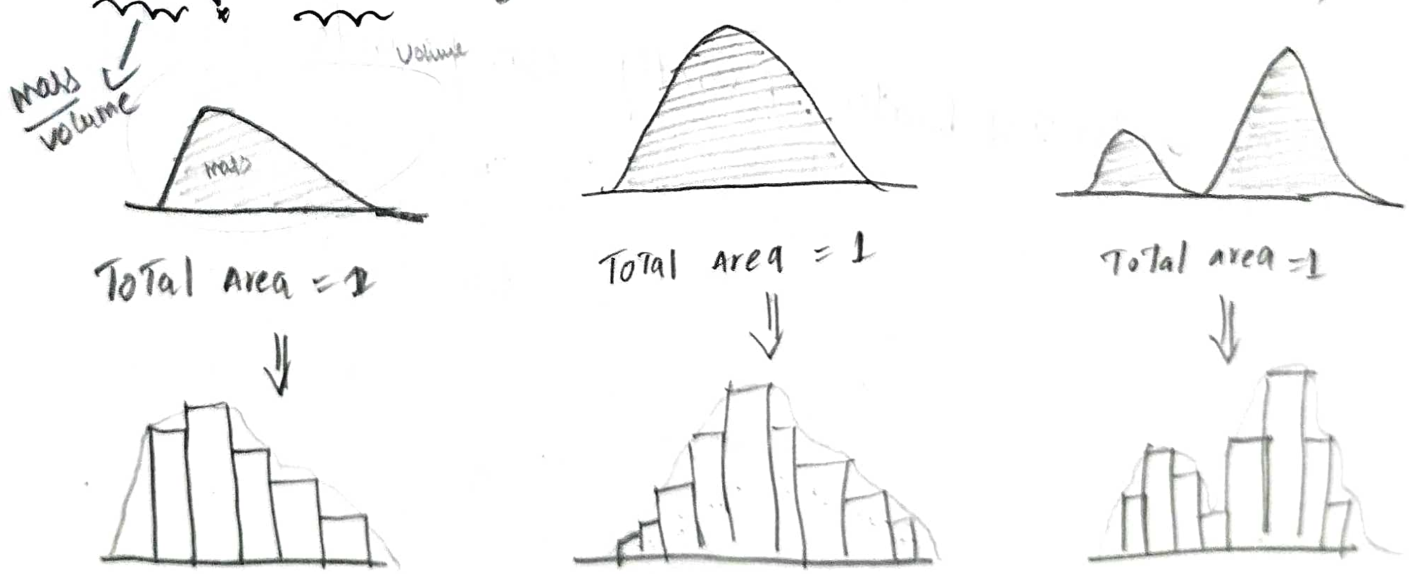
$n = 20$

$= \frac{5}{20} + \frac{6}{20} + \frac{4}{20} + \frac{3}{20} + \frac{2}{20}$

$= 0.25 + 0.30 + 0.20 + 0.15 + 0.10$

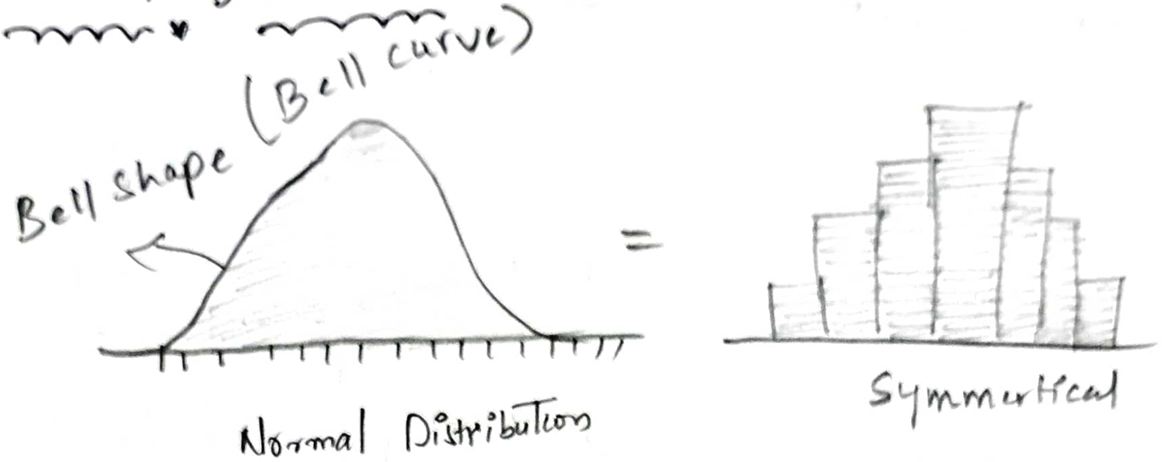
Total Area $= 1$

## * Density Curves :- Below the curve is 1 (Only Horizontal axis)




## * Relative Frequency Distribution

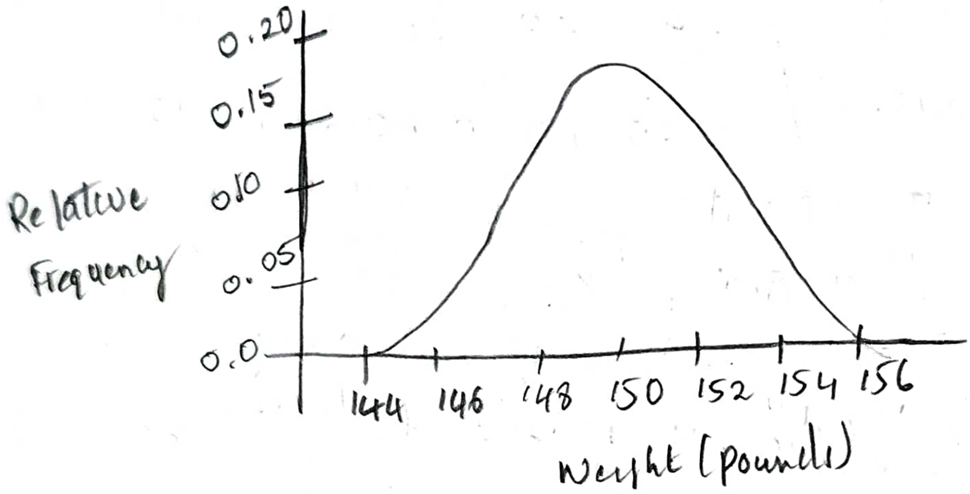
## * Normal Distribution



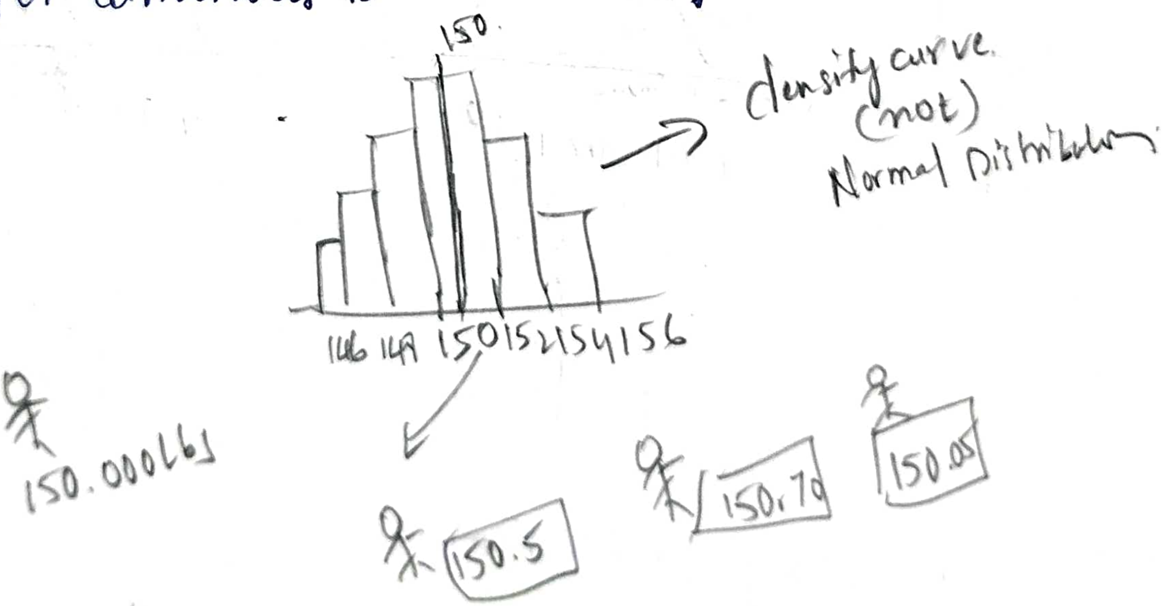

STATISTICS

⇒ interview question

1. For The Density Curve below, approximately what percentage of people weigh exactly 150 pounds ?




A :- For Continous Data Exactly 150 pounds = $\boxed{0}$



* For Continous Data what is the probability of Single Value ?

A : For Continous Data single value probability is "Zero" = 0.

* For Discrete Data, Probability of single value is $(P(H)) = 1/2$

Ex :- For coins

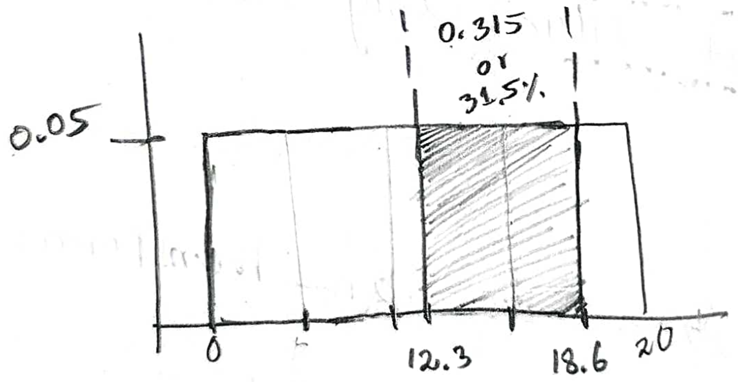
$$\{H, T\} = 1/2$$

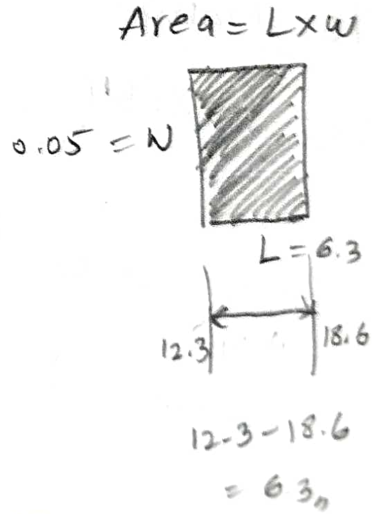
* For Dice

Equally Distributed.

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

* For $\boxed{\text{Uniform Distribution below}}$. what proportion of Values are located between 12.3 and 18.6 ?




Where Area = length x width

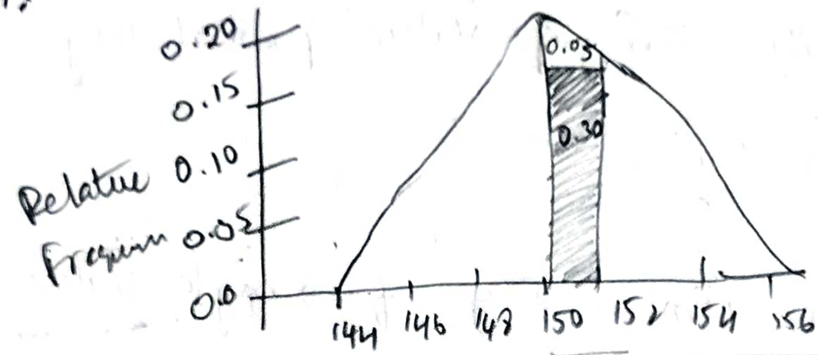
$= 6.3 \times 0.05$

$= 0.315$ (or) 31.5 %

Entran...

* What is The percentage weight between 150 and 152 pounds

A:




2 (distance)

150 - 152

= 2

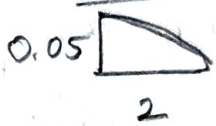
* Rectangle

Area = L×W

= 0.15 × 2

= 0.30

* Triangle




= ½ b h

Area = Length × width

= 0.05 × 2

= 0.1

= ½ b×h

= ½ (0.1)

= $\frac{0.1}{2}$ = 0.05

Total Area = 0.05 + 0.30

= 0.35

curve

Than Answer is

Approximately ✓ = 35 %

Not

Accurately ✗

* To know The Exact Value Only Normal Distribution

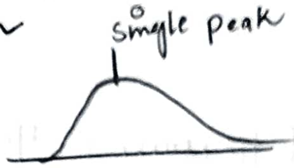
on possible

⇒ From skew distribution we can transform into

Normal Distribution.

* **Normal Distributions** *

→ Unimodal




→ it is $\boxed{\text{Symmertrical}}$ about its $\boxed{\text{mean}}$ (or) Centre.


→ it is Characterize by the parameters of $\boxed{\mu \text{ (Mue)}}$ and $\boxed{\sigma \text{ (sigma)}}$

→ $X \sim N(\mu, \sigma)$

⇓

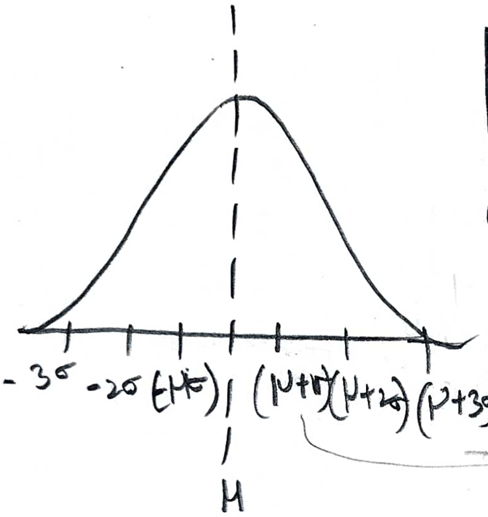
$$\boxed{X \sim N(\mu, \sigma)}$$

variable

Mean

(normal distribution)

STANDARD Deviation




$(\mu + 1\sigma)$ (Mue + 1 (sigma))

* 68 – 95 – 99.7 Rule

Ex: Wento University Measuring Heights of The Students

where (i) Mean $\mu$ = 5.5

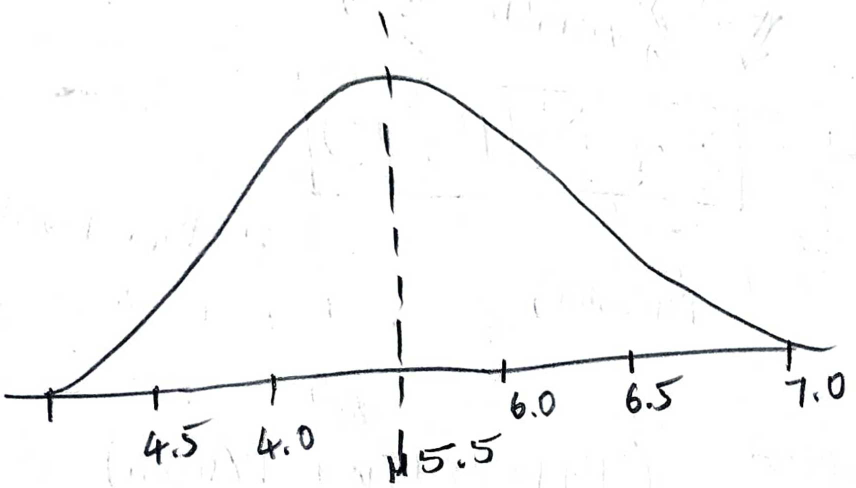
(ii) Standard deviation → Sigma ($\sigma$) = 0.5

(iii) X = Height (ft)

These Students forms "Normal Distribution"

$\mu = 5.5$

$\sigma = 0.5$




* $\mu + 1\sigma$

$= 5.5 + 1(0.5)$

$= 6$

* $\mu - 1\sigma$

$= 5.5 - 1(0.5)$ = ½

$= 5$

* $\mu + 2\sigma$

$= 5.5 + 2(0.5)$

$= 6.5$

* $\mu - 2\sigma$

$= 5.5 - 2(0.5)$ 1

$= 4.5$

* $\mu + 3\sigma$

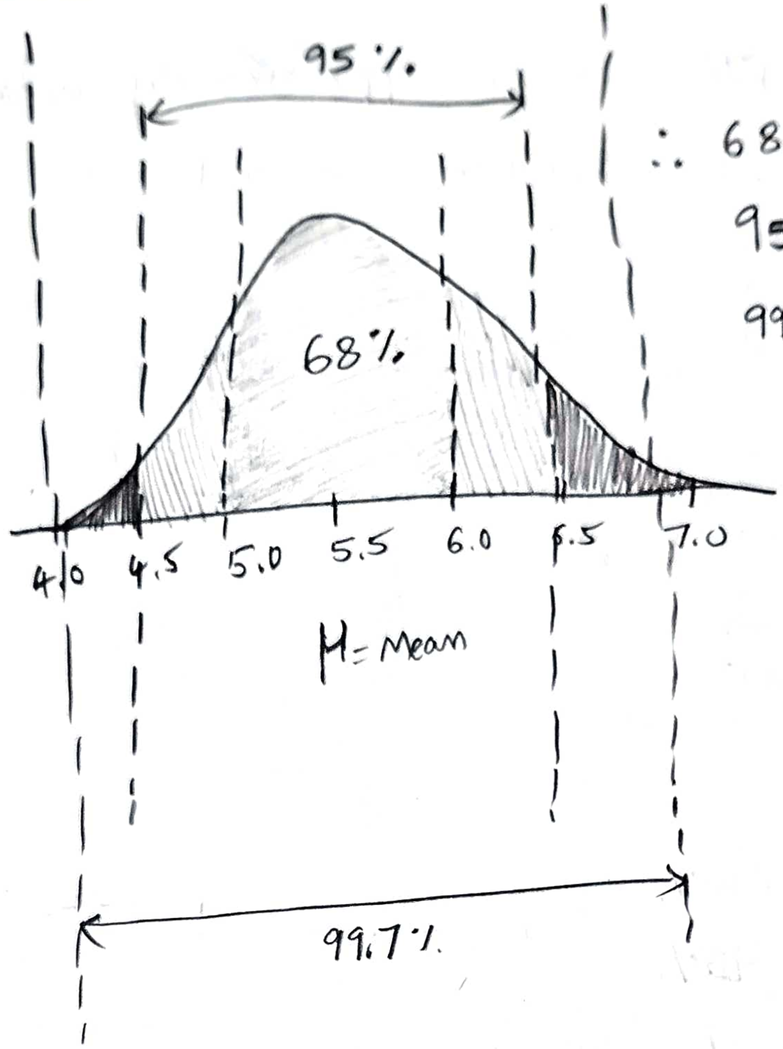
$= 5.5 + 3(0.5)$

$= 7$

* $\mu - 3\sigma$

$= 5.5 - 3(0.5)$ 1.5

$= 4$




$\therefore$ 68 % = $\pm 1\sigma$

95 % = $\pm 2\sigma$ (I.T)

99.7 % = $\pm 3\sigma$

5 % Error rate

$\therefore$ $\pm 4\sigma$ = 99.97 %

$\pm 5\sigma$ = 99.997 %

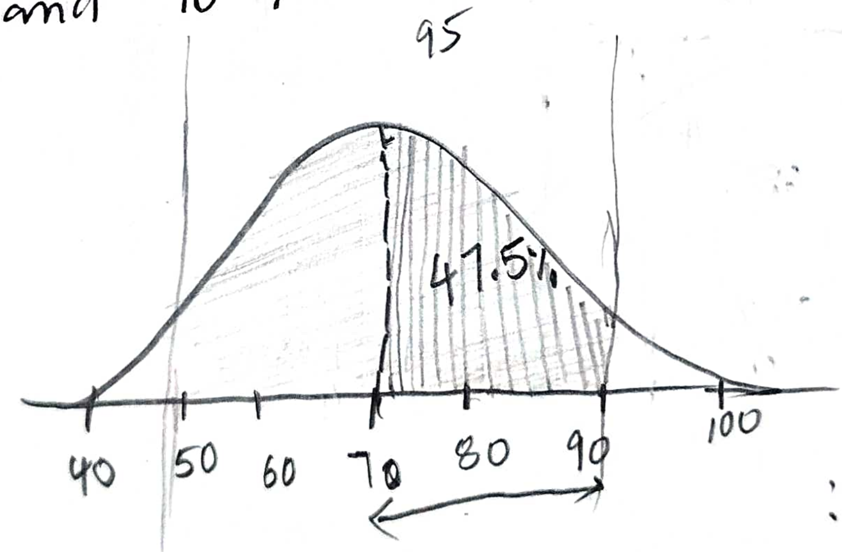
$\pm 6\sigma$ = 99.9999999 %

Dabba wala.

1 crore = 3 Defect

T.C.S

* The normal Distribution below has a standard Deviation of 10. Approximately what area is contained between 70 and 90?




$\mu = 70$

$\sigma = 10$

$\therefore$ 70 − 90

$95/2 = 47.5\%$ = $\pm 2\sigma$

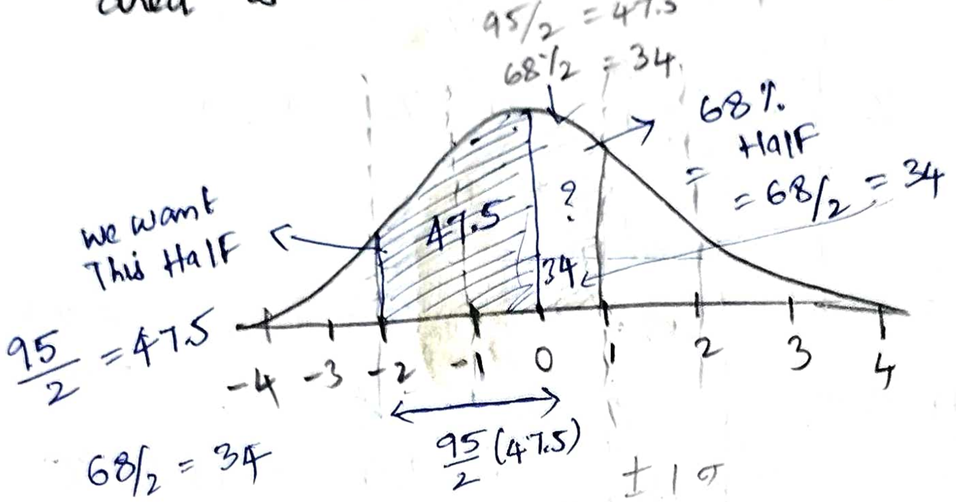
Where $\pm 2\sigma$

70 − 90 $(\pm 2\sigma)$ = −50, −60, $\boxed{70}$, +80, +90

Mean

= Half = 95 % ⇒ $95/2 = 47.5\%$

* For normal Distribution below, approximately what area is contained between -2 and 1 ?




$\mu = 0$

$\sigma = 1$

∴ where,

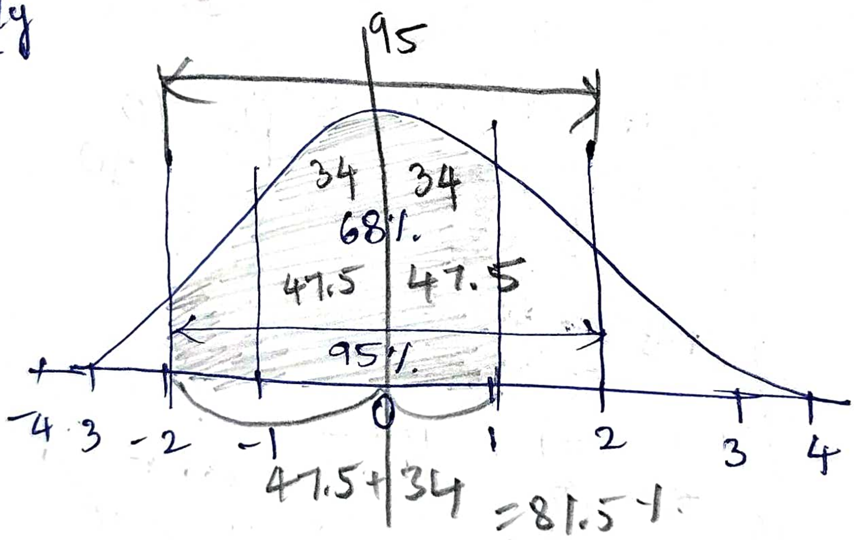
-1 to 1 = 68%

-2 to +2 = 95%

∴ $\frac{95}{2} = 47.5\%$, $\frac{68}{2} = 34\%$

= 81.5 %

* <u>clarity</u>




Manaki kavalisindhi

-2 to 1 so, 47.5 + 34 = 81.5 %

1 Hr.08 min (Z. scores).

Formula :- $\boxed{Z = \dfrac{X - \mu}{\sigma}}$ STanDarization formula.

$X$ = Value

$\mu$ = Mean

$\sigma$ = STanDard deviation.

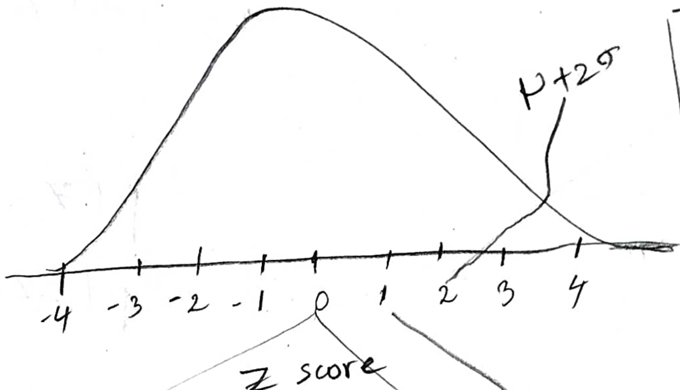
TCS

* STanDard Normal Distribution :- IF Every Data point Converted into 'z' score is called "STandard normal distribution".

Formula implementation

$\Rightarrow \mu = 0$

$\Rightarrow \sigma = 1$




$\boxed{Z = \dfrac{X - \mu}{\sigma}}$

$Z = \dfrac{\mu + 2\sigma - \mu}{\sigma}$

$Z = \dfrac{2\sigma}{\sigma}$

$Z = 2$

So, "0" Ela vachindhi dhanni formula lo pediThey

$X = 0$

$\boxed{Z = \dfrac{X - \mu}{\sigma}}$

$Z = \dfrac{0 - \mu}{\sigma}$

$Z = \dfrac{0 - 0}{\sigma}$

$Z = 0$

$\mu + 1\sigma$

$Z = \dfrac{X - \mu}{\sigma}$

$Z = \dfrac{\mu + 1\sigma - \mu}{\sigma}$

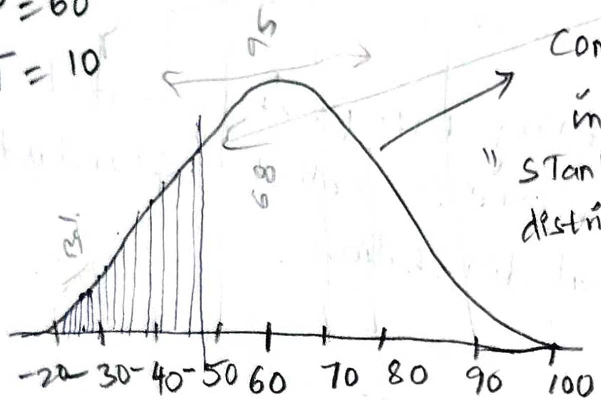
$Z = \dfrac{1\sigma}{\sigma}$

$Z = 1$

Ex:-

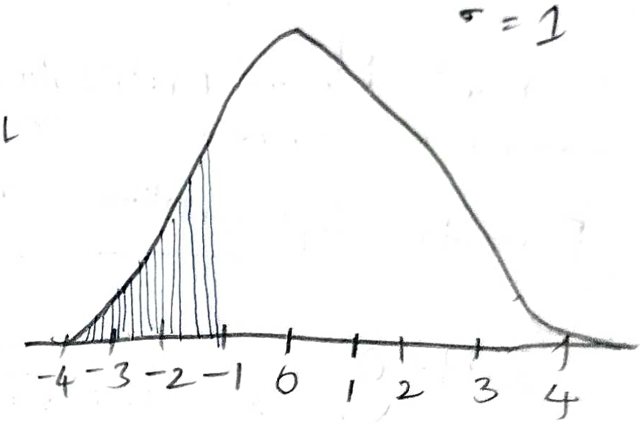
Suppose That we gathered data from last year's Final Chemistry Exam and found that it followed a normal distribution with mean of 60, standard deviation is 10. What proportion of Students Scored less Than 49 in exam

$\mu = 60$
$\sigma = 10$




Convert it into "STandard normal distribution".

$\mu = 0$
$\sigma = 1$




Only 98% value Thelusthadi with 68-95-98 Rule. Ndukuday "Z" score.

$$Z = \frac{X - \mu}{\sigma}$$

$$Z = \frac{X - 60}{10}$$

$Z = \frac{60 - 60}{10}$, $Z = \frac{0}{10}$, $Z = 0$

$Z = \frac{\mu + 80 - 60}{10}$, $Z = \frac{\mu + 20}{10}$, $Z = 2\mu$

$Z = \frac{\mu + 70 - 60}{10}$, $Z = \frac{\mu + 10}{10}$, $Z = 1\mu$

$P(z < -1.1) = ?$

$P(z < -1.1) = 0.13567$

$P(x < 49) = ?$

Convert into "Z" score.

$$Z = \frac{49 - 60}{10}$$

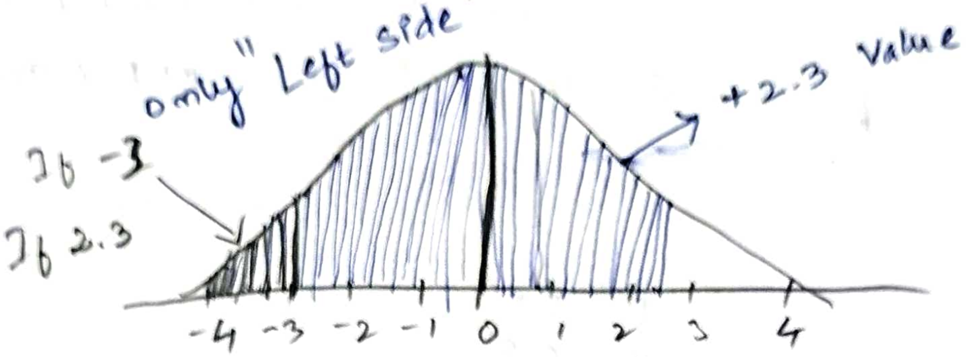
$$Z = \frac{-11}{10} \Rightarrow -1.1$$

Suppose -1.14

In "Z" Table check The value

| z | 0 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 |
|---|---|---|---|---|---|---|
| -1.1 | $\boxed{.13567}$ | .13350 | .13136 | .12924 | .12714 | .1250 |

* Answer always gives the "Left side Only".



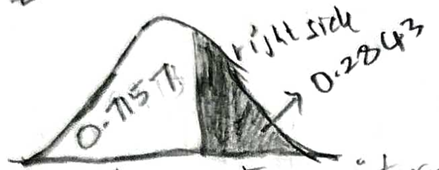

∴ If we want "Right side" (+2.3) Value

$= 1 - (+2.3) \Rightarrow$ From "z" Table (+2.3) = 0.98214

$= 1 - (\text{Area}_{\text{Left}}) \Rightarrow \text{Area}_{(\text{right})}$

$= 1 - 0.982$

Z=0.57

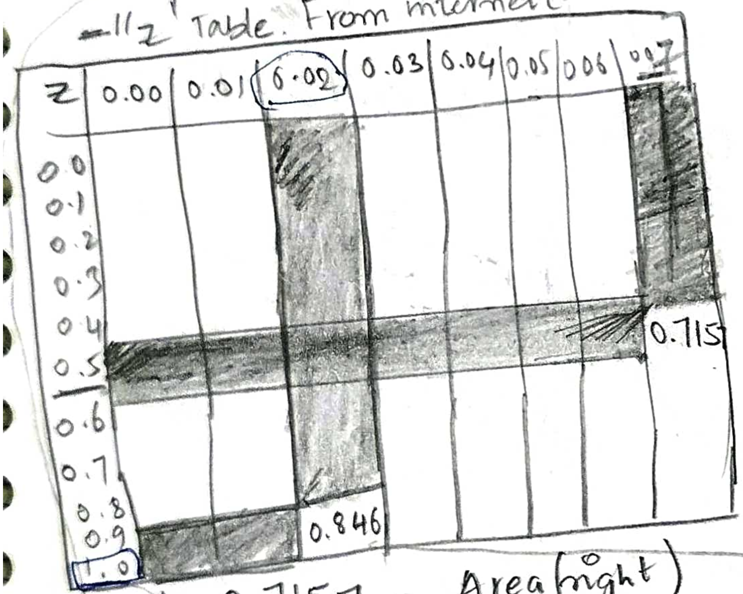


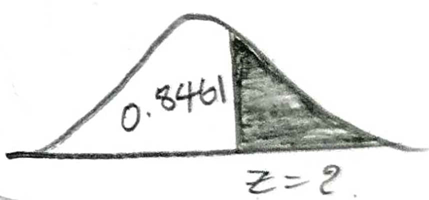

*****
Imp

* From "probability" we can Calcul "z" Score.

* From z score we can calc Probability".

"z" Table From internet (chrome)




= 1 - 0.7157 = Area (right)

= 0.2843

(Density curve = 1) Area

Total area = 1

z = 1.02

* Leftside = 0.7157

* right side = 0.2843

Total area → 1

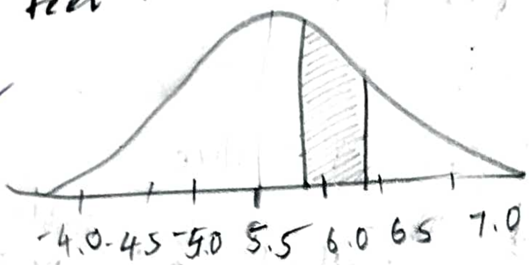
* If we Had "Probability" = 0.8461 In "z" Table we can Find "z" score = 1.02

* "Z" Scores

When measuring the Heights of all Students at a local University, it was found that it was normally Distributed with mean Height of 5.5 Feet, and Standard deviation 0.5 Feet. what proportion of students are between 5.81 Feet and 6.3 Feet Tall?

$P(5.81 < x < 6.3) = ?$

Sol:-




$\mu = 5.5$
$\sigma = 0.5$

"2" Z scores

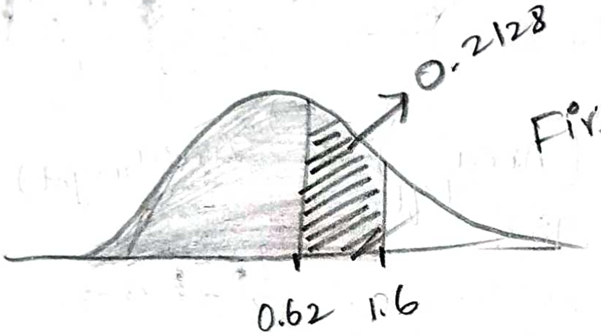
First (Big Area - Small area)

1.6 - 0.62

(z score)

0.9452 - 0.7324

$\Rightarrow$ 0.2128




$Z = (5.81 < x < 6.3)$

$\boxed{Z = \frac{X - \mu}{\sigma}}$ $X = 5.81$

$Z = \frac{5.81 - 5.5}{0.5}$

$\boxed{Z = 0.62}$

Z Table

| | 0.01 | 0.02 |
|---|---|---|
| 0.6 | | 0.7324 |

$\boxed{P(Z < 0.62) = 0.7324}$

$X = 6.3$

$Z = \frac{6.3 - 5.5}{0.5}$

$\boxed{Z = 1.6}$

Z - Table

| | 0.00 | 0.01 | 0.02 |
|---|---|---|---|
| 1.6 | 0.9452 | | |

$= 0.9452 - 0.7324$
$\Rightarrow 0.2128$

$\boxed{P(Z < 1.6) = 0.9452}$

$P(0.62 < Z < 1.6)$
$= 0.2128$