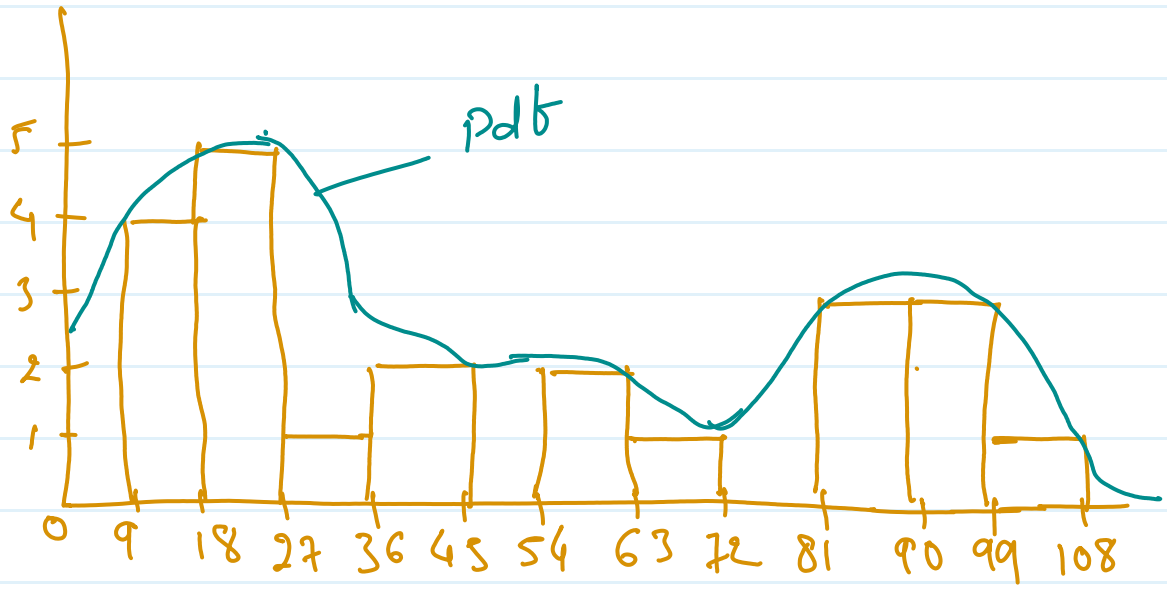
Histogram

Data = {10, 12, 13, 14, 20, 22, 24, 25, 26, 35, 38, 42, 55, 56, 68, 82, 84, 86, 92, 93, 94, 100}

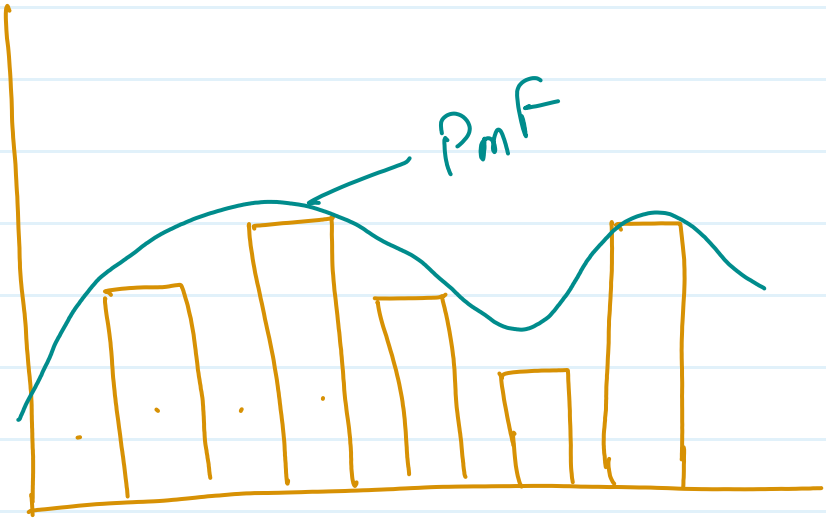
$$\text{Bin Size} = \frac{\max - \min}{\text{no. of bin}}$$

$$= \frac{100 - 10}{10} \Rightarrow 9$$

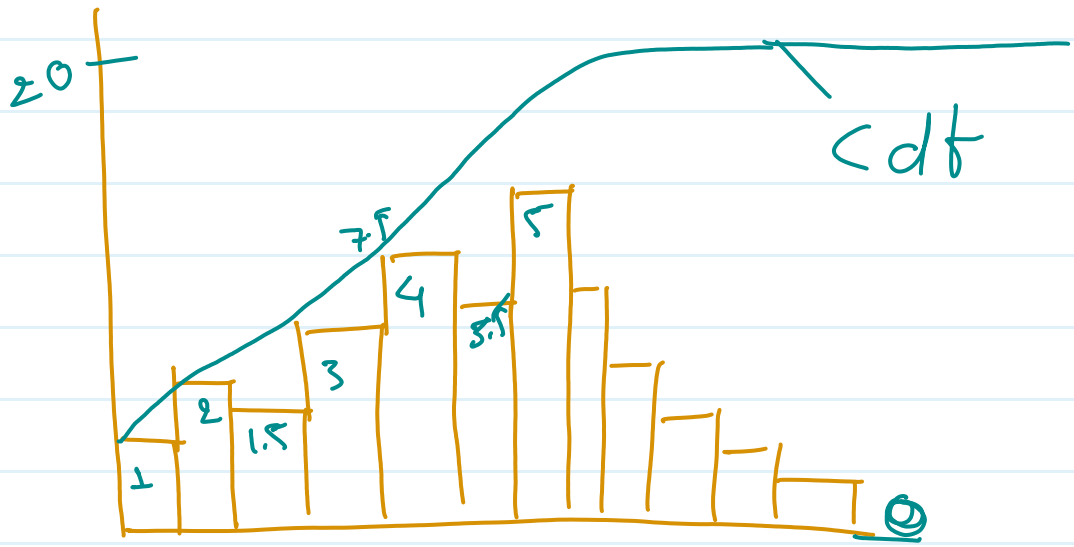
Bin Size = 9



Pdf = Probability distribution function.



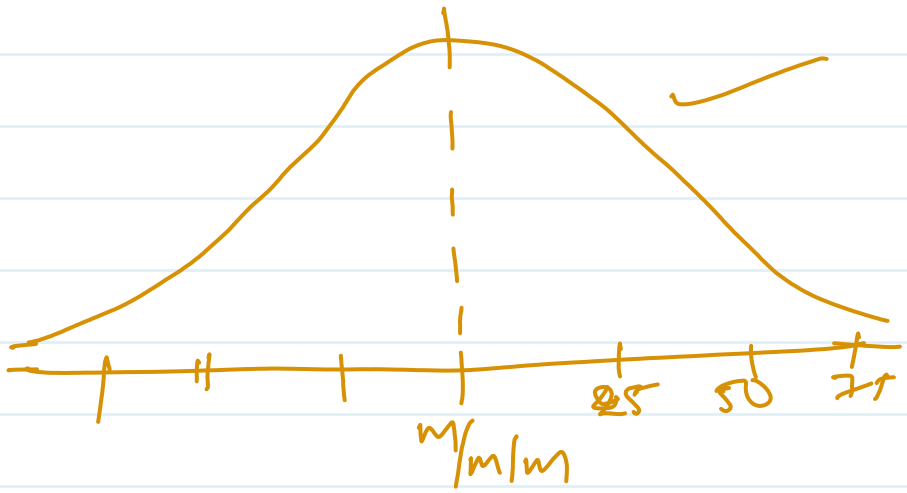
pmf = probability mass function



cdf = Cummulative dist. function

Cummulative sum =>

Standard normal Dist.



$\mu = 40/50/30/60$

SND = $\mu = 0$

$SD = 1$

$$Z_{score} = \frac{x_i - \mu}{\sigma}$$

Eg:- Data = $\{1, 2, 3, 5, 6, 7\}$

$\mu = \frac{28}{7} = 4$ $\quad SD = 1$

$1 = \frac{1-4}{1} \Rightarrow -3$

$2 = \frac{2-4}{1} \Rightarrow -2$

$3 = \frac{3-4}{1} \Rightarrow -1$

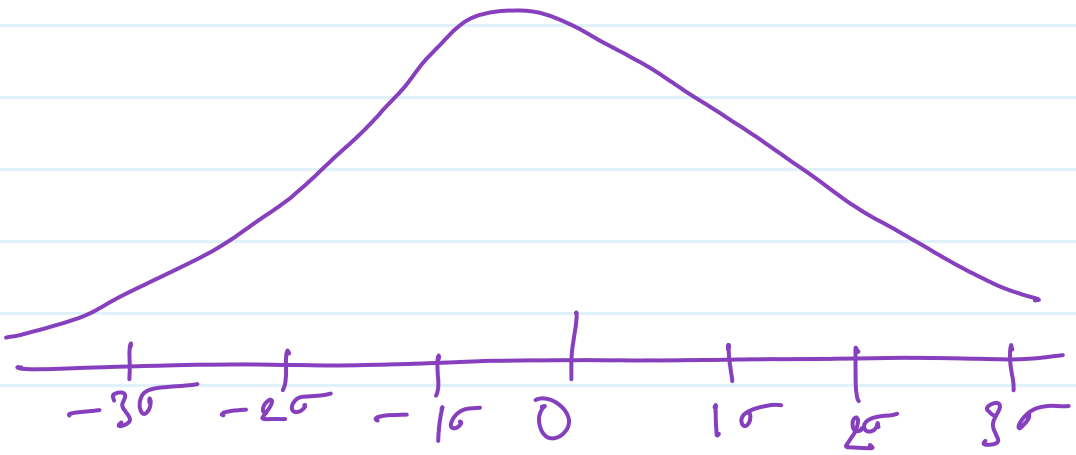
$4 = \frac{4-4}{1} \Rightarrow 0$

$5 = \frac{5-4}{1} \Rightarrow 1$

$6 = \quad \Rightarrow 2$

$7 = \quad \Rightarrow 3$

$\{-3, -2, -1, 0, 1, 2, 3\}$



✱ Covariance

X	Y
Age	weight
14	40
15	45
18	61
20	68
25	74

Covariance

$X\uparrow \quad Y\uparrow$

$X\downarrow \quad Y\downarrow$

} +ve variance, +ve correlation

$X\uparrow \quad Y\downarrow$

$X\downarrow \quad Y\uparrow$

} -ve var., -ve correlation

$X\uparrow \quad Y=$
$X= \quad Y\uparrow$
} no covari / no correlation

① Covariance :- Quality the relationship b/w x & y numeric value.

$$\text{Population } Cov(X,Y) = \frac{\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})}{N}$$

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

Eg :- Eco. growth | Utility 50%

2.1	8
2.5	12
3.6	10
4.0	14

Cov.

X	Y	$\bar{x}$	$\bar{y}$	$(X-\bar{x})$	$(Y-\bar{y})$
2.1	8	3.05	11	-1	-3
2.5	12			-0.6	1
3.6	10			0.5	-1
4.0	14			0.9	3

$$Cov(x,y) = \sum_{i=1}^{n} \frac{(x_i - \bar{x})(y - \bar{y})}{n-1}$$

$$= \frac{(-1)(-3) + (-0.6)(1) + (0.5)(-1) + (0.9)(3)}{4-1}$$

$$\Rightarrow \frac{4.6}{3} \Rightarrow 1.53$$

$\Rightarrow$ +ve 1.53 X and Y

Covariance $\Rightarrow$ − or +

correlation $\Rightarrow$ −1 or +1

→ 0.8 +0.8 −

−0.1 +0.1 −

−0.02 +0.002

X ≈ Y