

Covariance:-

* Covariance is a measure of the joint variability of 2 random variables.

* Basically it describes the relationship between 2 numerical variables.

* Covariance doesnot quantify any value regarding variables, it just tell the relation that is positive, negative, or none.

* There are 3 types of Covariance

i) Positive Covariance

ii) Negative Covariance

iii) No relation.

Formula:-

$$\text{Cov}_{x,y} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

* It is just the same formula as the Variance, the only difference is we can find square in Variance (one variable). Here in Covariance we have 2 variables.

Positive covariance:-

* If the x and y increase or decrease together then x and y ~~are positively~~ has ~~complete~~ Positive covariance.

* For example if the temperature of the day increase then the ice cream sales of that day will increase.

* If the temperature is getting down (winter season) then the sales of ice cream will decrease.

* The above both cases ~~are~~ ^{has} ~~positive~~ Positive covariance.

Sample:-

x	y	$x - \bar{x}$	$y - \bar{y}$	$(x - \bar{x})(y - \bar{y})$
30	5	-3	-1	3
35	8	+2	+2	4
40	8	+7	+2	14
25	4	-8	-2	16
35	5	+2	-1	-2
Sum =	<u>33</u>	<u>6</u>		<u>35</u>

$$\text{COV}(x, y) = \sigma_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

$$= \frac{35}{4} = 8.75$$

* The Value 8.75 doesn't have significance
like ~~the~~ higher the Value higher the covariance,
its not like that

~~* The value obtained from the formula~~
~~just~~

* The Sign of the Value is Very important
which shows it is positive or negative.

Basic intuition:-

x y
(+ve) (+ve)

Both increase

Positive cov

(-ve) (-ve)

Both decrease

Positive cov

Negative covariance:-

* If x increase and y decrease or
 y increase and x decrease then x
and y has negative covariance.

* The Value obtained from the formula
will have negative Value.

Basic intuition

x	y	
(+ve)	(-ve)	$x \uparrow y \downarrow$
(-ve)	(+ve)	$x \downarrow y \uparrow$

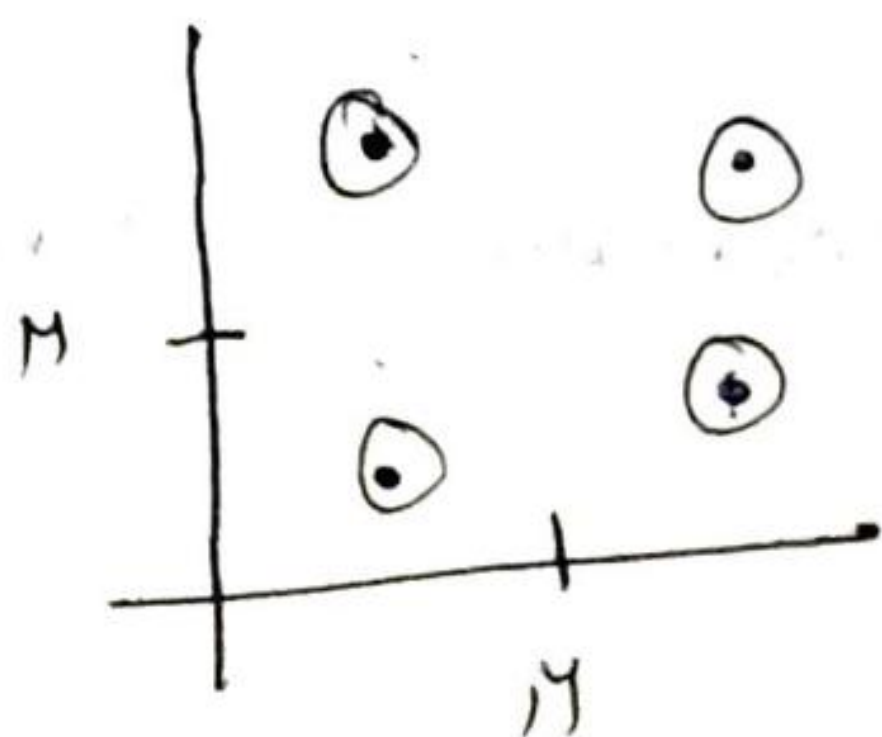
negative cov
negative cov

No covariance or no relation:-

* If $x = y$ means if the 2 features are same then covariance is 0.

* If 2 features are not related and does not follow any linear path, then it has no covariance or no relation.

example.



* Here the data points ~~are~~ does not follow any linear path.

* The data were spread out.

Here the values ~~with~~ above and below mean will cancel out each other so the covariance is zero, means there is no relation.

NOTE:-

If the values given with probability the formula would be

$$\text{COV}(x, y) = \sum [(x - E(x))(y - E(y))] P(x)$$

The covariance value doesn't show whether the line is steeper or flatter.

It just shows and gives whether the slope of the line is positive or negative.