



Q1] Compute the coefficient of correlation b/w x & y and interpret your result.

x	23.3	17.5	17.8	20.7	18.1	20.9	22.9	20.8
y	4.2	3.8	4.6	3.2	5.2	4.7	4.4	5.6

Solution:	x	y	x^2	y^2	xy
	23.3	4.2	542.89	17.64	97.86
	17.5	3.8	306.25	14.44	66.5
	17.8	4.6	316.84	21.16	81.88
	20.7	3.2	428.49	10.24	66.24
	18.1	5.2	327.61	27.04	94.12
	20.9	4.7	436.81	22.09	98.23
	22.9	4.4	524.41	19.36	100.76
	20.8	5.6	432.64	31.36	116.48
	$\Sigma x = 162$	$\Sigma y = 35.7$	$\Sigma x^2 = 3315.9$	$\Sigma y^2 = 163.3$	$\Sigma xy = 722.07$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{162}{8} = 20.25, \quad \bar{y} = \frac{\Sigma y}{n} = \frac{35.7}{8} = 4.46$$

$$\text{Cor}(x, y) = \frac{1}{n} \Sigma xy - \bar{x} \cdot \bar{y}$$

$$= \frac{1}{8} (722.07) - (20.25 \times 4.46)$$

$$= -0.1068$$

0.52

$$\sigma_x = \sqrt{\frac{1}{n} \Sigma x^2 - \bar{x}^2} = 2.10475$$

$$\sigma_y = \sqrt{\frac{1}{n} \Sigma y^2 - \bar{y}^2} = 0.708$$



$$r = \frac{\text{cov}(x, y)}{\sigma_x \cdot \sigma_y} = \frac{-0.1068}{2.104 \times 0.708} = -0.072$$

Interpretation: No co-relation.

Q2] calculate Karl Pearson's co-efficient of correlation between x and y .

x : 8 8 7 5 6 2

y : 3 4 10 13 22 8

x	y	x^2	y^2	xy
8	3	64	9	24
8	4	64	16	32
7	10	49	100	70
5	13	25	169	65
6	22	36	484	132
2	8	4	64	16

$$\sum x = 36 \quad \sum y = 60 \quad \sum x^2 = 242 \quad \sum y^2 = 842 \quad \sum xy = 339$$

$$\bar{x} = \frac{\sum x}{n} = \frac{36}{6} = 6, \quad \bar{y} = \frac{\sum y}{n} = \frac{60}{6} = 10$$

$$\text{cov}(x, y) = \frac{1}{n} \sum xy - \bar{x} \cdot \bar{y} = \frac{1}{6} \times 339 - 6 \times 10 = -3.5$$

$$\sigma_x = \sqrt{\frac{1}{n} \sum x^2 - \bar{x}^2} = \sqrt{\frac{1}{6} \times 242 - 36} = 2.081$$

$$\sigma_y = \sqrt{\frac{1}{n} \sum y^2 - \bar{y}^2} = \sqrt{\frac{1}{6} \times 842 - 100} = 6.3508$$

$$r = \frac{\text{cov}(x, y)}{\sigma_x \cdot \sigma_y} = \frac{-3.5}{2.081 \times 6.3508} = 0.2646$$

Q3] Spearman's rank correlation;

X :	52	63	45	36	72	65	45	25
Y :	62	53	51	25	79	43	60	33

Solution:

X	R ₁	Y	R ₂	D = R ₁ - R ₂	D ²
52	4	62	2	2	4
63	3	53	4	-1	1
45	5.5	51	5	0.5	0.25
36	7	25	8	-1	1
72	1	79	1	0	0
65	2	43	6	-4	16
45	5.5	60	3	2.5	6.25
25	8	33	7	1	1

45 repeated 2 times

$$m_1 = 2$$

$$\sum D^2 = 29.5$$

$$R = 1 - 6 \left[\frac{\sum D^2 + \frac{1}{12}(m_1^3 - m_1)}{N^3 - N} \right]$$

$$R = 1 - 6 \left[\frac{29.5 + \frac{1}{12}(8 - 2)}{504} \right]$$

$$R = 1 - 0.357 = 0.643$$

Q4] Compute the Rank correlation coefficient for the data.

X :	98	101	104	107	113	120	125	128
Y :	65	65	67	68	68	69	68	68

X	Y	R ₁	R ₂	D = R ₁ - R ₂	D ²
98	65	1	1.5	-0.5	0.25
101	65	2	1.5	0.5	0.25
104	67	3	3	0	0
107	68	4	5.5	-1.5	2.25
113	68	5	5.5	-0.5	0.25
120	69	6	8	-2	4
125	68	7	5.5	1.5	2.25
128	68	8	5.5	2.5	6.25

$$\sum d^2 = 15.5$$

65 is repeated 2 times $\therefore m_1 = 2$

68 is repeated 4 times $\therefore m_2 = 4$

$$R = 1 - \frac{\sum D^2 + \frac{1}{12}(m_1^3 - m_1) + \frac{1}{12}(m_2^3 - m_2)}{N^3 - N}$$

$$R = 1 - \frac{15.5 + 0.5 + 5}{504}$$

$$R = 0.75$$

- Q5] It is given that the means of x & y are 5 & 10. If the Line regression of y on x is parallel to the Line $20y = 9x + 40$. Estimate value of y for $x = 30$.

Line of regression of y on x is $y - \bar{y} = b_{yx}(x - \bar{x})$

It's slope is b_{yx} . But $20y = 9x + 40$

$$y = \frac{9}{20}x + 2 \quad \text{whose slope is } \frac{9}{20}$$

from data: $\bar{x} = 5$ & $\bar{y} = 10$ Hence, the Eqn of the Line of regression of y on x is $y - 10 = \frac{9}{20}(x - 5)$

$$y = \frac{9}{20}(x) + 155/20$$

when, $x = 30$ | $y = 270/20 + 155/20 = 21.2$

- Q6] find the Lines of regression for following data.

$x :$	65	66	67	67	68	69	70	72
$y :$	67	68	65	66	72	72	69	71

soln):

x	y	xy	x ²	y ²
65	67	4355	4225	4489
66	68	4488	4356	4624
67	65	4355	4489	4225
67	66	4422	4489	4356
68	72	4896	4624	5184
69	72	4968	4761	5184
70	69	4830	4900	4761
72	71	5112	5184	5041

$$\sum x = 544 \quad \sum y = 550 \quad \sum xy = 37426 \quad \sum x^2 = 37028 \quad \sum y^2 = 37864$$

Regression of x on y .

$$x = a + by$$

$$\sum x = na + \sum by$$

$$544 = 8a + 550b \rightarrow (1)$$

$$\sum xy = a \sum y + b \sum y^2$$

$$37426 = 550a + 37864b \rightarrow (2)$$

Solve (1) & (2)

$$x = 33.291 + 0.5048y$$

Regression line of y on x

$$y = a + bx$$

$$\sum y = na + b \sum x$$

$$550 = 8a + 544b \rightarrow (3)$$

$$\sum xy = a \sum x + b \sum x^2$$

$$37426 = 544a + 37028b \rightarrow (4)$$

Solving (3) & (4)

$$a = 19.638 \quad b = 0.72$$

$$y = 19.638 + 0.72x$$

co-efficient of correlation is given by

$$r = \sqrt{b_{xy} \cdot b_{yx}}$$

$$b_{xy} = \frac{\sum xy}{\sum y^2} = \frac{37426}{37864} = 0.9884$$

$$b_{yx} = \frac{\sum xy}{\sum x^2} = \frac{37426}{37028} = 1.0107$$

$$r = \sqrt{0.9884 \times 1.0107} = 0.99$$

- x - x
x