

Karl Pearson's Correlation Coefficient

Data Science and A.I. Lecture Series

Bindeshwar Singh Kushwaha

PostNetwork Academy

Problem Statement

Find the Karl Pearson's coefficient of correlation between X and Y for the given data:

X : 6 2 4 9 1 3 5 8

Problem Statement

Find the Karl Pearson's coefficient of correlation between X and Y for the given data:

$X :$	6	2	4	9	1	3	5	8
$Y :$	13	8	12	15	9	10	11	16

Problem Statement

Find the Karl Pearson's coefficient of correlation between X and Y for the given data:

$$\begin{array}{l} X : 6 \quad 2 \quad 4 \quad 9 \quad 1 \quad 3 \quad 5 \quad 8 \\ Y : 13 \quad 8 \quad 12 \quad 15 \quad 9 \quad 10 \quad 11 \quad 16 \end{array}$$

Using the assumed means:

$$u_i = X_i - 5, \quad v_i = Y_i - 12$$

Tabular Data

X_i	Y_i	$u_i = X_i - 5$	$v_i = Y_i - 12$	$u_i v_i$	u_i^2	v_i^2
6	13	1	1	1	1	1
2	8	-3	-4	12	9	16
4	12	-1	0	0	1	0
9	15	-2	3	12	16	9
1	9	4	-3	12	16	9
3	10	-4	-2	4	4	4
5	11	0	-1	0	0	1
8	16	3	4	12	9	16
Sum	-	$\sum u_i = -2$	$\sum v_i = -2$	$\sum u_i v_i = 53$	$\sum u_i^2 = 56$	$\sum v_i^2 = 56$

Table: Calculation Table

The formula for Karl Pearson's coefficient of correlation is:

$$r(X, Y) = \frac{n\sum u_i v_i - (\sum u_i)(\sum v_i)}{\sqrt{[n\sum u_i^2 - (\sum u_i)^2][n\sum v_i^2 - (\sum v_i)^2]}}$$

The formula for Karl Pearson's coefficient of correlation is:

$$r(X, Y) = \frac{n\sum u_i v_i - (\sum u_i)(\sum v_i)}{\sqrt{[n\sum u_i^2 - (\sum u_i)^2][n\sum v_i^2 - (\sum v_i)^2]}}$$

Substituting the values:

$$n = 8, \quad \sum u_i v_i = 53, \quad \sum u_i^2 = 56, \quad \sum v_i^2 = 56,$$

The formula for Karl Pearson's coefficient of correlation is:

$$r(X, Y) = \frac{n\sum u_i v_i - (\sum u_i)(\sum v_i)}{\sqrt{[n\sum u_i^2 - (\sum u_i)^2][n\sum v_i^2 - (\sum v_i)^2]}}$$

Substituting the values:

$$n = 8, \quad \sum u_i v_i = 53, \quad \sum u_i^2 = 56, \quad \sum v_i^2 = 56, \\ \sum u_i = -2, \quad \sum v_i = -2$$

The formula for Karl Pearson's coefficient of correlation is:

$$r(X, Y) = \frac{n\sum u_i v_i - (\sum u_i)(\sum v_i)}{\sqrt{[n\sum u_i^2 - (\sum u_i)^2][n\sum v_i^2 - (\sum v_i)^2]}}$$

Substituting the values:

$$n = 8, \quad \sum u_i v_i = 53, \quad \sum u_i^2 = 56, \quad \sum v_i^2 = 56, \\ \sum u_i = -2, \quad \sum v_i = -2$$

Final Calculation

$$r(X, Y) = \frac{8(33) - (-2)(-2)}{\sqrt{[8(56) - (-2)^2][8(56) - (-2)^2]}}$$

Final Calculation

$$r(X, Y) = \frac{8(33) - (-2)(-2)}{\sqrt{[8(56) - (-2)^2][8(56) - (-2)^2]}}$$

Simplifying:

Final Calculation

$$r(X, Y) = \frac{8(33) - (-2)(-2)}{\sqrt{[8(56) - (-2)^2][8(56) - (-2)^2]}}$$

Simplifying:

$$r(X, Y) = \frac{420}{\sqrt{444 \times 444}} \approx 0.946$$

Final Calculation

$$r(X, Y) = \frac{8(33) - (-2)(-2)}{\sqrt{[8(56) - (-2)^2][8(56) - (-2)^2]}}$$

Simplifying:

$$r(X, Y) = \frac{420}{\sqrt{444 \times 444}} \approx 0.946$$

Thus, the Karl Pearson's coefficient of correlation is approximately 0.946.

Reach PostNetwork Academy

Website

PostNetwork Academy | www.postnetwork.co

YouTube Channel

www.youtube.com/@postnetworkacademy

Facebook Page

www.facebook.com/postnetworkacademy

LinkedIn Page

www.linkedin.com/company/postnetworkacademy

Thank You!