

Q 2 What is orthogonal?

correlation coefficient

covariance

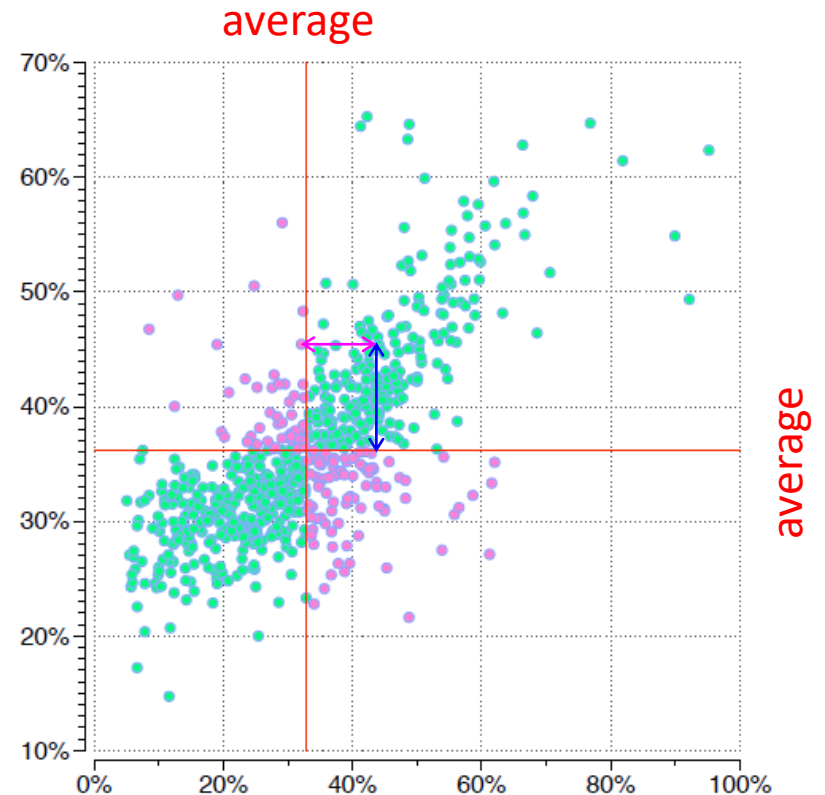
$$(x_i - \bar{x})(y_j - \bar{y})$$

	x	$x_i - \bar{x}$	y	$y_j - \bar{y}$	
a	43.9	11.0	46.1	9.9	108.9
b	31.7	-1.2	36.6	0.4	-0.5
c	29.5	-3.4	42.0	5.8	-19.7
...	40.5	7.6	40.3	4.1	31.2
...	32.4	-0.5	37.9	1.6	-0.8
...	59.0	26.1	48.0	11.8	308.0
...	46.6	13.8	43.7	7.5	103.5
...	28.4	-4.5	31.5	-4.7	21.2
...	28.7	-4.2	32.1	-4.1	17.2
...	29.5	-3.4	32.0	-4.2	14.3
...	35.9	3.0	32.5	-3.7	-11.1
...	32.6	-0.3	33.0	-3.2	1.0
...	38.6	5.7	38.0	1.8	10.3
...	31.0	-1.9	30.8	-5.4	10.3
...	18.5	-14.3	25.7	-10.5	150.2
...	45.3	12.4	40.0	3.8	47.1
...	25.2	-7.7	36.7	0.5	-3.9
...	50.3	17.4	45.0	8.8	153.1
...	54.7	21.9	42.4	6.2	135.8
...	30.0	-2.8	32.7	-3.5	9.8
...	50.2	17.4	49.5	13.3	231.4
...	44.9	12.0	41.6	5.4	64.8
...	27.6	-5.2	29.2	-7.0	36.4
...	50.0	17.1	42.6	6.4	109.4
...	47.3	14.4	40.6	4.4	63.4
...	...	...	...	...	...

Average

84.3

the percentage of unmarried persons



the percentage of rented houses

covariance

$$\sum (x_i - \bar{x})(y_j - \bar{y})$$

correlation coefficient



$$\text{covariance} = \frac{84.3}{14.4 \times 8.1} = 0.723$$

↑
covariance of x

↑
covariance of y

correlation coefficient

$$\frac{\sum (x_i - \bar{x})(y_j - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_j - \bar{y})^2}}$$

correlation coefficient

$$-1 \leq \text{correlation coefficient } r \leq 1$$

correlation coefficient r vector X' Y' inner product

$$= \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}}$$

the lengths of the vectors X' and Y'

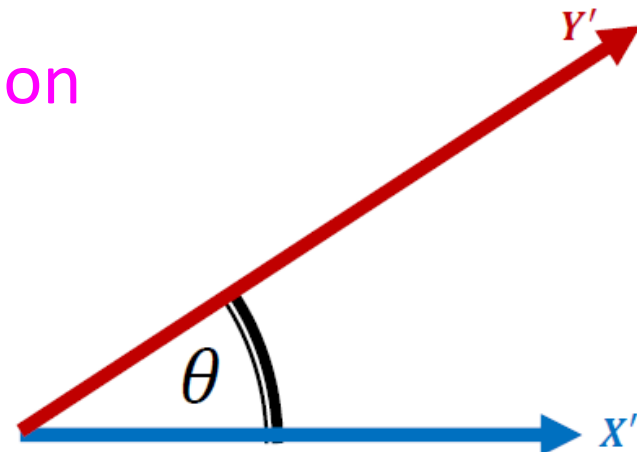
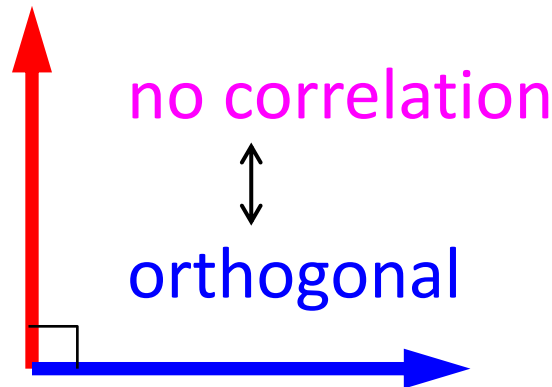
$$= \frac{X' \cdot Y'}{\|X'\| \|Y'\|}$$

$$= \cos \theta$$

$$\theta = 0^\circ \quad \cos \theta = 1 \quad r = 1$$



$$\theta = 90^\circ \quad \cos \theta = 0 \quad r = 0$$



$$\begin{matrix} X' & Y' \\ \begin{pmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ x_3 - \bar{x} \\ x_4 - \bar{x} \\ x_5 - \bar{x} \\ x_6 - \bar{x} \\ x_7 - \bar{x} \\ x_8 - \bar{x} \\ \vdots \\ x_{N-3} - \bar{x} \\ x_{N-3} - \bar{x} \\ x_{N-2} - \bar{x} \\ x_{N-1} - \bar{x} \\ x_N - \bar{x} \end{pmatrix} & \begin{pmatrix} y_1 - \bar{y} \\ y_2 - \bar{y} \\ y_3 - \bar{y} \\ y_4 - \bar{y} \\ y_5 - \bar{y} \\ y_6 - \bar{y} \\ y_7 - \bar{y} \\ y_8 - \bar{y} \\ \vdots \\ y_{N-3} - \bar{y} \\ y_{N-3} - \bar{y} \\ y_{N-2} - \bar{y} \\ y_{N-1} - \bar{y} \\ y_N - \bar{y} \end{pmatrix} \end{matrix}$$

$$\text{correlation coefficient} = \frac{\text{covariance}}{\text{covariance1} \times \text{covariance2}}$$

	A	B	C	D
Ex.1	10	-2	-3	1
Ex.2	10	5	-5	-1
Ex.3	10	-3	8	0
Ex.4	-6	-2	-5	0
Ex.5	-6	5	8	1
Ex.6	-6	-3	-3	-1
Ex.7	-4	-2	8	-1
Ex.8	-4	5	-3	0
Ex.9	-4	-3	-5	1

	control factor				SN比
	A	B	C	D	
Ex.1	1	1	1	1	
Ex.2	1	2	2	2	
Ex.3	1	3	3	3	
Ex.4	2	1	2	3	
Ex.5	2	2	3	1	
Ex.6	2	3	1	2	
Ex.7	3	1	3	2	
Ex.8	3	2	1	3	
Ex.9	3	3	2	1	

	A	B	C	D
level1	10	-2	-3	1
level2	-6	5	-5	-1
level1	-4	-3	8	0

	B×C	C×D	B×D	D×A	A×B	A×C
covariance	0	0	0	0	0	0
covariance1	3.6	5.7	3.6	0.8	7.1	7.1
covariance2	5.7	0.8	0.8	7.1	3.6	5.7
correlation coefficient	0	0	0	0	0	0

When focusing on the effect of control factor A
The correlation coefficients of the columns B and C, C and D,
and B and D are all zero.



The columns of B, C and D are orthogonal to each other