



UNIVERSITAS AIRLANGGA  
Excellence with Morality

# Korelasi Antara 2 Variabel

Tim Dosen Data dan Pustaka







# Pokok Bahasan

1. Kekuatan dan arah korelasi
2. Mengestimasi korelasi data kategori
3. Mengestimasi korelasi data kontinum
4. Perbedaan antara korelasi dengan hubungan sebab-akibat (*causal inference*)





# Definisi Korelasi

Hubungan antara 2 variabel X dan Y

X  $\longleftrightarrow$  Y  
hubungan

X  $\longrightarrow$  Y  
pengaruh

Dalam simbol notasi standar:

X disebut sebagai variabel independent

Y disebut sebagai variabel dependent

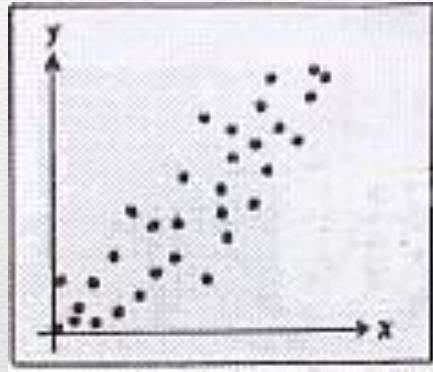
Ukuran yang digunakan untuk mengetahui derajat hubungan disebut **koefisien korelasi**



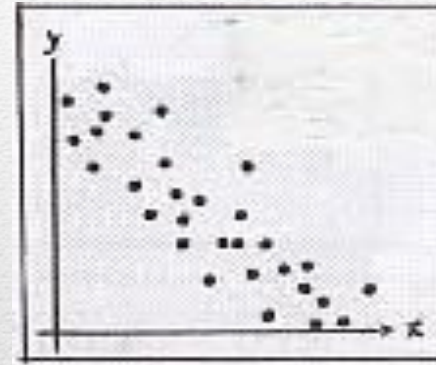


UNIVERSITAS AIRLANGGA  
Excellence with Morality

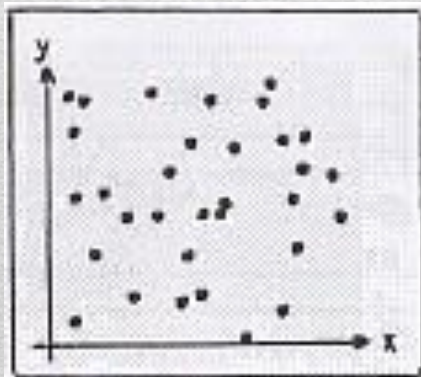
# Pola Hubungan



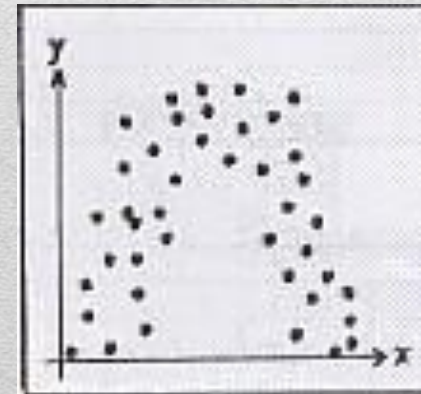
Korelasi positif



Korelasi negatif



Tidak berkorelasi



Korelasi tidak linier





# Ukuran Korelasi/ Koefisien Korelasi





# Interpretasi nilai koefisien korelasi

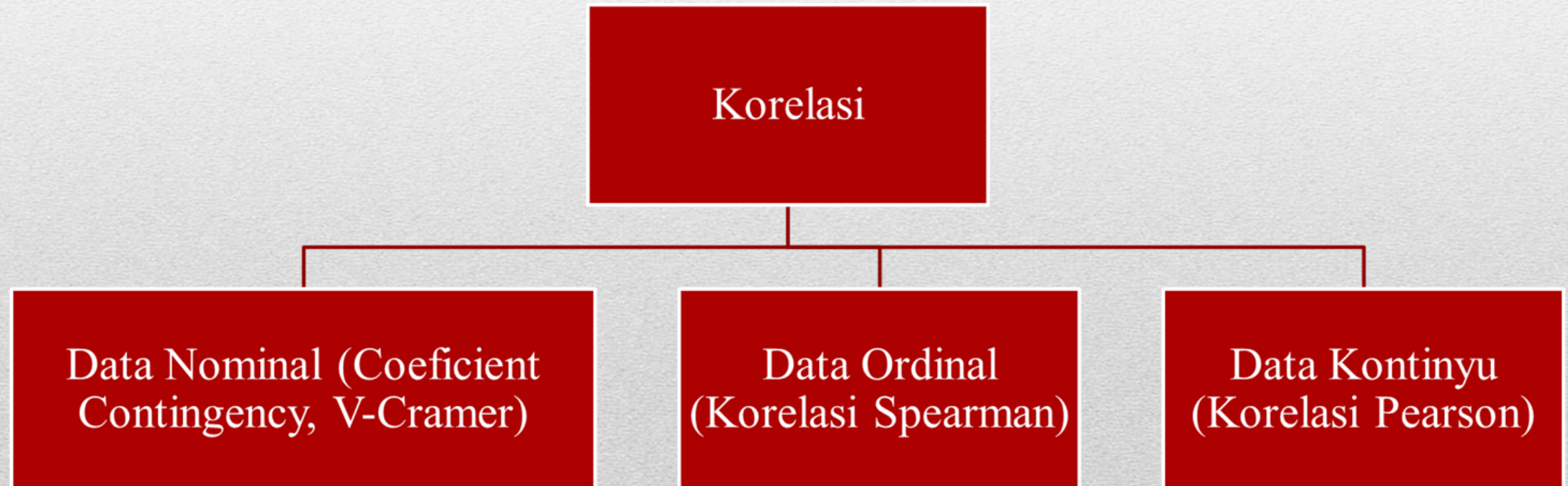
Interval Koefisien	Tingkat Hubungan
0,00 - 0,199	Sangat rendah
0,20 - 0,399	Rendah
0,40 - 0,599	Sedang
0,60 - 0,799	Kuat
0,80 - 1,000	Sangat kuat

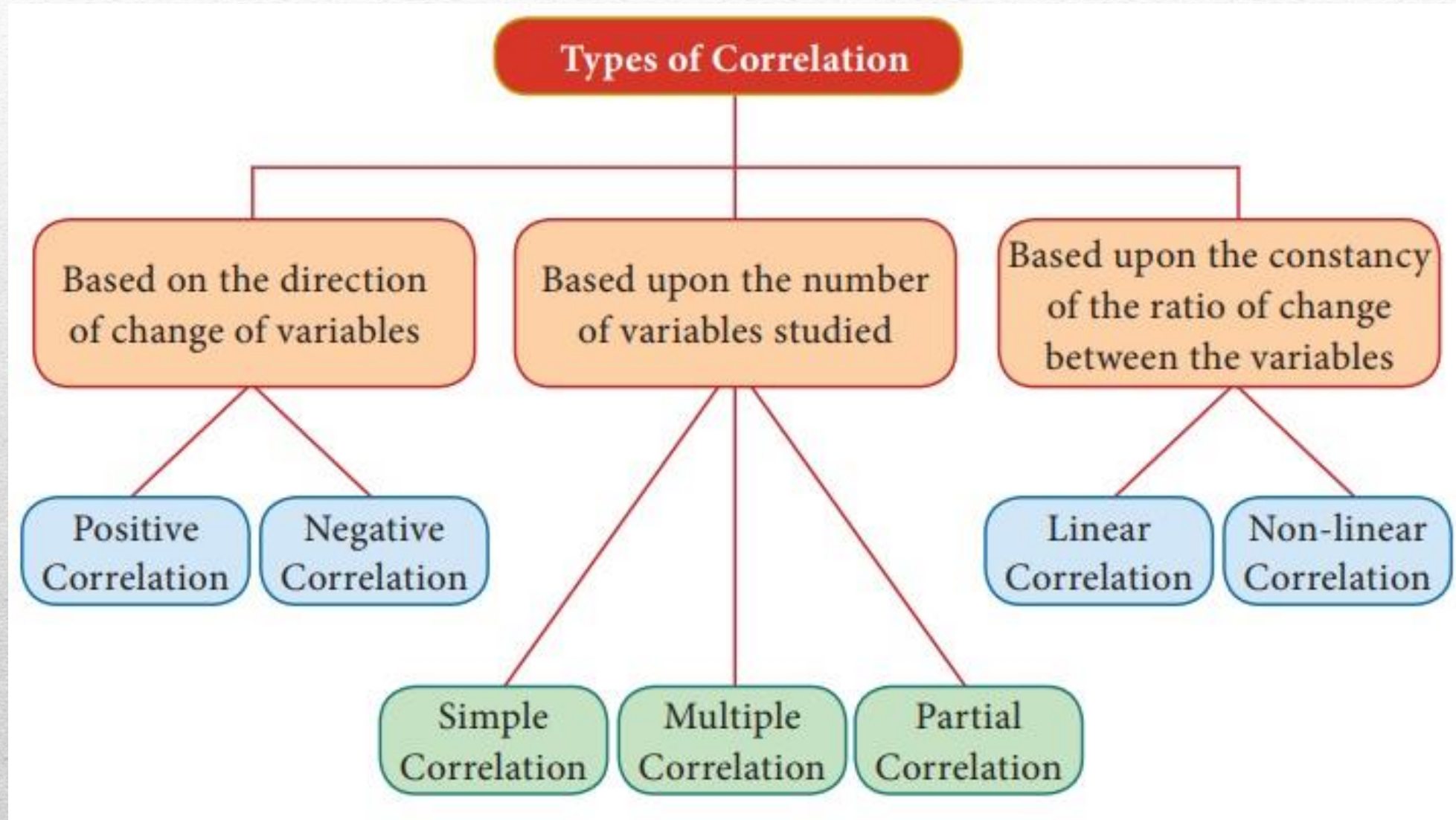
Berlaku juga untuk nilai minus





# Uji Korelasi









# ***Koefisien Korelasi Pearson Product Moment***

***Syarat :***

1. Data berskala minimal interval (data kontinyu)
2. Data berdistribusi normal

***Rumus :***

$$r = \frac{n \sum_{i=1}^n x_i y_i - \left( \sum_{i=1}^n x_i \right) \left( \sum_{i=1}^n y_i \right)}{\sqrt{\left[ n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2 \right] \left[ n \sum_{i=1}^n y_i^2 - \left( \sum_{i=1}^n y_i \right)^2 \right]}}$$





# ***Makna Koefisien Korelasi Pearson***

Tanda (-) dan (+) hanya menunjukkan arah hubungan

- (+) Jika nilai variabel X naik maka nilai pada variabel Y juga akan naik, ***Atau*** Jika nilai variabel X turun maka nilai pada variabel Y juga akan turun
- (-) Jika nilai variabel X naik maka nilai pada variabel Y akan turun, ***Atau*** Jika nilai variabel X turun maka nilai pada variabel Y akan naik





## Korelasi PEARSON PRODUCT MOMENT

X (10.000)	Y	X <sup>2</sup>	Y <sup>2</sup>	XY
18	17	324	289	306
23	20	529	400	460
28	23	784	529	644
32	27	1074	729	864
41	32	1681	1024	1312
59	46	3481	2116	2714
86	63	7396	3969	5418
99	74	9801	5476	7326
$\Sigma X$	$\Sigma Y$	$\Sigma X^2$	$\Sigma Y^2$	$\Sigma XY$
386	302	25020	14532	19044

Hitung berapa koefisien korelasi Pearson pada data tabel di samping ini

Interpretasikan makna hasil koefisien korelasi tersebut





# Korelasi Parsial

Ukuran hubungan linier antara variabel secara parsial (dengan menganggap variabel lain tetap) misalnya :

Korelasi antara Y dan  $X_2$ , dengan  $X_1$  dibuat tetap dilambangkan dengan  $r_{yx2.x1}$

Jika korelasi sederhana melibatkan satu variabel dependent dan satu variabel independent, maka korelasi parsial melibatkan lebih dari satu variabel independent dan satu variabel dependent.





# Koefisien Korelasi Parsial

$$r_{yx2.x1} = \frac{r_{yx2} - r_{yx1}r_{x1x2}}{\sqrt{(1 - r_{yx1}^2)(1 - r_{x1x2}^2)}}$$

dimana :

$r_{yx2}$  = korelasi antara y dan  $x_2$

$r_{yx1}$  = korelasi antara y dan  $x_1$

$r_{x1x2}$  = korelasi antara  $x_1$  dan  $x_2$

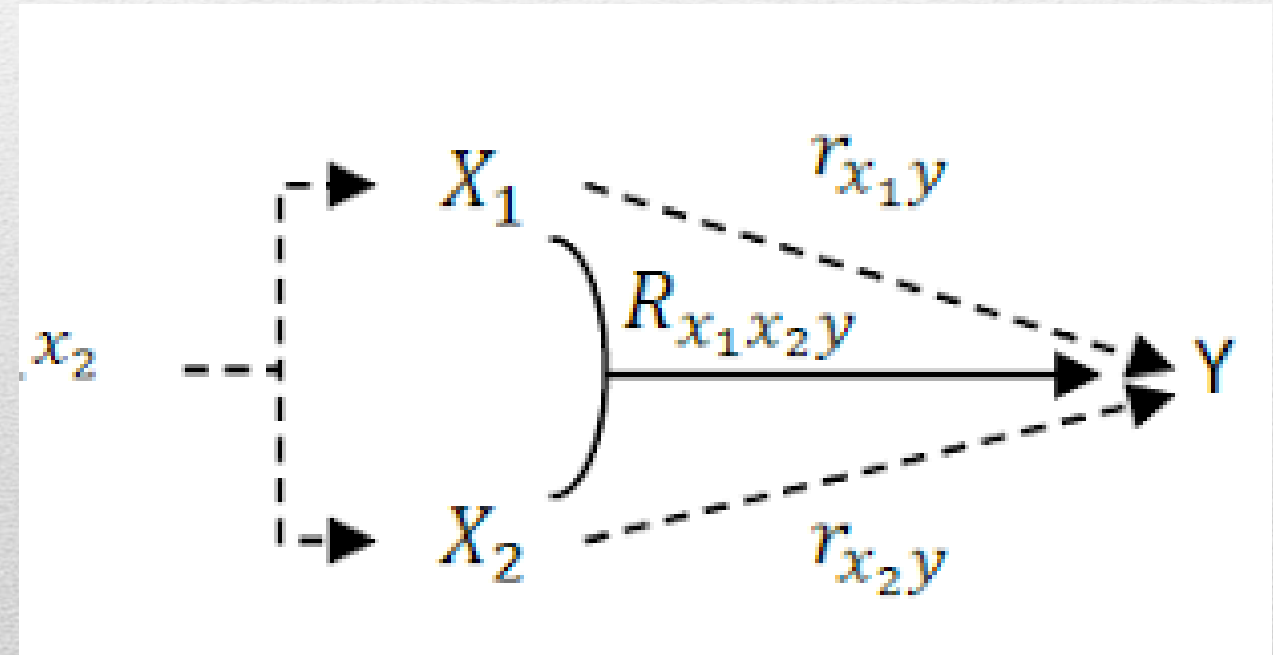




# Korelasi Berganda

Hubungan yang terjadi antara variabel dependent dengan lebih dari 1 variabel independent.

Misalnya: korelasi antara  $Y$  dengan  $X_1$  dan  $X_2$







$$R_{yx1x2} = \sqrt{\frac{r^2_{yx1} + r^2_{yx2} - 2r_{yx1}r_{yx2}r_{x1x2}}{1 - r^2_{x1x2}}}$$

dimana :

$r_{yx2}$  = korelasi antara  $y$  dan  $x_2$

$r_{yx1}$  = korelasi antara  $y$  dan  $x_1$

$r_{x1x2}$  = korelasi antara  $x_1$  dan  $x_2$





# Bahan Diskusi

Apa perbedaan korelasi parsial dengan korelasi ganda?





# Korelasi Spearman

Korelasi spearman digunakan untuk **menganalisis hubungan pada variabel dengan skala data minimal ordinal.**

Bisa juga merupakan uji alternatif dari uji korelasi Pearson. Data dengan skala kontinyu juga dapat dihitung menggunakan formula Korelasi Spearman.





# Koefisien Korelasi Spearman

$$r = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

dimana :  $\sum d_i^2 = \sum_{i=1}^n [R(X_i) - R(Y_i)]^2$





## Contoh soal :

Terdapat 2 orang juri yang diminta untuk menilai dalam lomba memasak makanan. Jumlah makanan yang dilombakan 10. Hasil penilaian juri dalam bentuk score tercantum pada tabel.

Hitunglah koefisien korelasi Spearman pada kasus tersebut!

Makanan	Juri I	Juri II
1	9	6
2	6	7
3	5	6
4	7	8
5	4	5
6	3	4
7	2	2
8	8	9
9	7	8
10	6	6





Makanan	Juri I ( $X_i$ )	Juri II ( $Y_i$ )	Rangking ( $X_i$ )	Rangking ( $Y_i$ )	$D_i$	$d_i^2$
1	9	8	10	8	-2	4
2	6	7	5.5	6	0.5	0.25
3	5	6	4	4.5	0.5	0.25
4	7	8	7.5	8	0.5	0.25
5	4	5	3	3	0	0
6	3	4	2	2	0	0
7	2	2	1	1	0	0
8	8	9	9	10	1	1
9	7	8	7.5	8	0.5	0.25
10	6	6	5.5	4.5	-1	1
Jumlah					<b>0</b>	<b>7</b>

$$r = 1 - \frac{(6)(7)}{10(10^2 - 1)} = 1 - 0,04 = 0,96$$





# Coefficient Contingency C

Coefficient contingency untuk mengukur keeratan hubungan antara 2 variabel dengan skala data nominal. Data nominal diubah dalam bentuk tabel kontingensi.

Coefficient contingency bernilai antara 0 - 1. Dimana semakin koefisien menuju nilai 1 semakin kuat hubungan antara kedua variabel.





# Rumus Coeficient Contingency C

$$C = \sqrt{\frac{\chi^2}{N + \chi^2}}$$

dimana :  $\chi^2$  = Chi-square yang dihitung dari tabel kontingensi  
N = jumlah data



# Cara Menghitung Nilai $\chi^2$

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^k \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Contoh penghitungan nilai chi-square ada di link berikut





# Korelasi vs Hubungan Causalitas (sebab akibat)

Apakah sama antara korelasi dengan hubungan sebab akibat.

Korelasi antara musim hujan dengan jumlah kecelakaan lalu lintas.

Artinya setiap hujan deras maka kecelakaan lalu lintas juga akan terjadi. Pada waktu yang sama terdapat 2 variabel yang muncul secara bersamaan. Namun apakah kemudian hujan deras merupakan mempunyai hubungan kausalitas dengan kecelakaan lalu lintas.

Silakan didiskusikan.

correlation does not imply causation





# Requirement for causal relationship

## Necessary Condition:

1. Empirical association (korelasi)
2. Appropriate time order: sebab (faktor independen) ada sebelum dampak ada (faktor dependen)
3. Nonspuriousness: bukan terjadi karena variabel lain

## Important Condition:

1. Mechanism
2. Context

Contoh:

- Media violence would increase aggression only among individuals who were already predisposed to aggression
- children reacted more aggressively after observing men committing violent acts than after observing women committing these same acts.



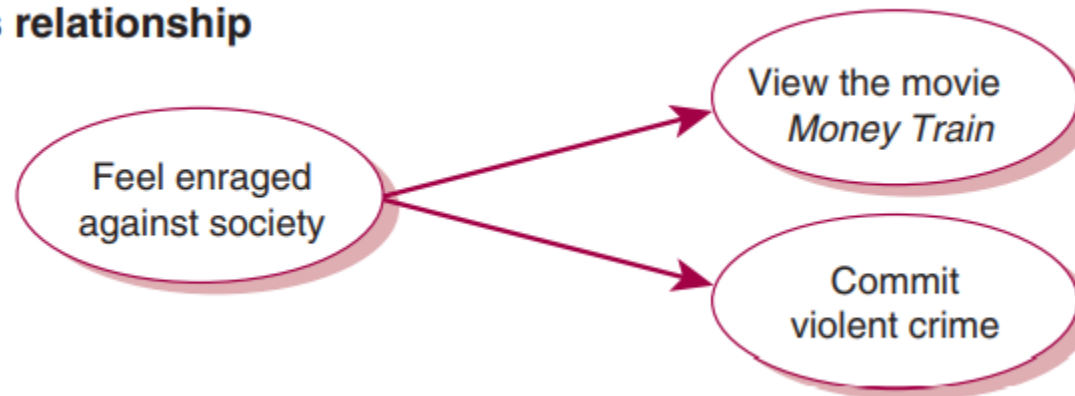


# Spurious relationship

## Spurious relationship



## The extraneous variable creates the spurious relationship

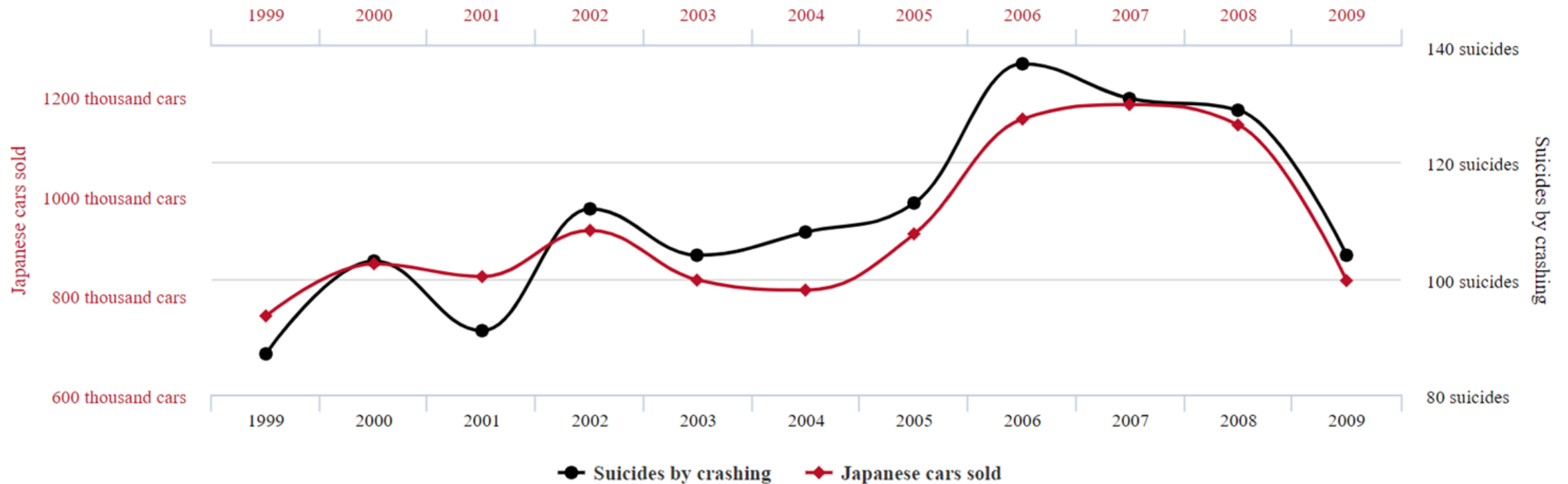






# Japanese passenger cars sold in the US correlates with Suicides by crashing of motor vehicle

Correlation: 93.57% ( $r=0.935701$ )



Data sources: U.S. Bureau of Transportation Statistics and Centers for Disease Control & Prevention

tylervigen.com

<https://tylervigen.com/spurious-correlations>





UNIVERSITAS AIRLANGGA  
Excellence with Morality

Thank You!