CORRELAZIONE

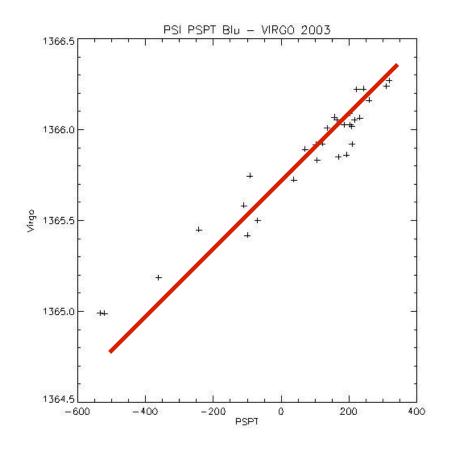
LA CORRELAZIONE LINEARE

La correlazione indica la tendenza che hanno due variabili (X e Y) a variare insieme, ovvero, a covariare. Ad esempio, si può supporre che vi sia una relazione tra il nostro salario ed i soldi spesi in viaggi, nel senso che all'aumentare dell'uno aumenta anche l'altro.

Quando si parla di correlazione bisogna prendere in considerazione due aspetti: *il tipo di relazione esistente* tra due variabili e *la forma della relazione*.

Per quanto riguarda il **tipo di relazione**, essa può essere *lineare* o *non lineare*

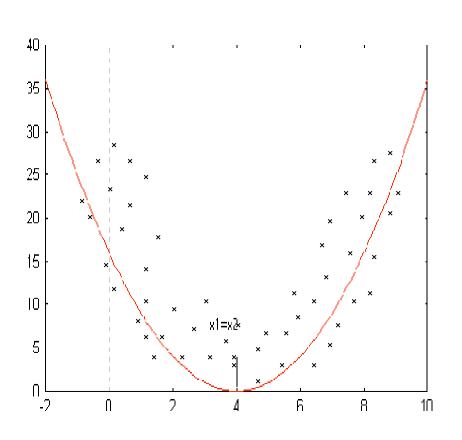
 La relazione è di tipo lineare se, rappresentata su assi cartesiane, si avvicina alla forma di una retta.



In questo all'aumentare (o al diminuire) di X aumenta (o diminuisce) Y

Ad esempio, all'aumentare dell'altezza di una persona aumenta anche il suo peso.

• La relazione è di tipo *non lineare*, se rappresentata su assi cartesiane, ha un andamento curvilineo (parabola o iperbole).

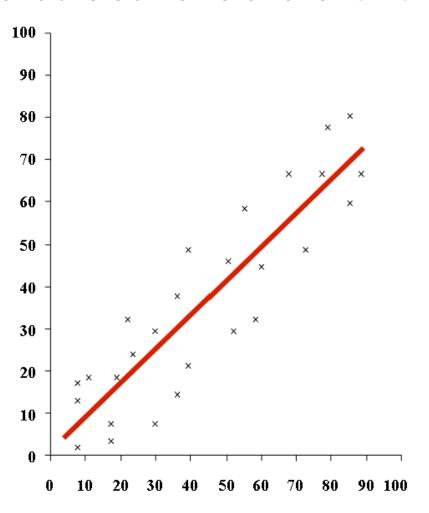


In questo caso a livelli bassi e alti di X corrispondono livelli bassi di Y; mentre a livelli intermedi di X corrispondono livelli alti di Y.

Ad esempio, il tempo impiegato per risolvere un problema è alto quando l'ansia è bassa o alta, è minimo quando l'ansia ha livelli medi

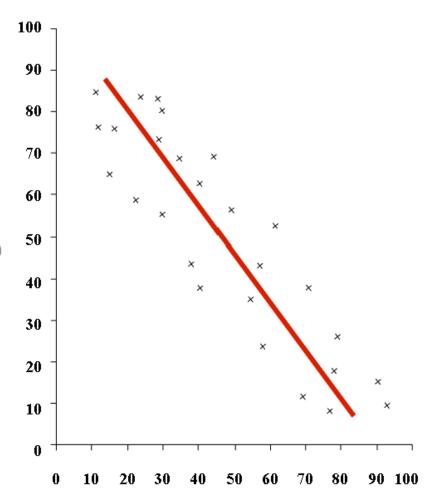
Per quanto riguarda la forma della relazione, si distinguono l'entità e la direzione.

La **direzione** può essere: *positiva*, se all'aumentare di una variabile aumenta anche l'altra.



Ad esempio, all'aumentare dei matrimoni aumentano il numero di nascite

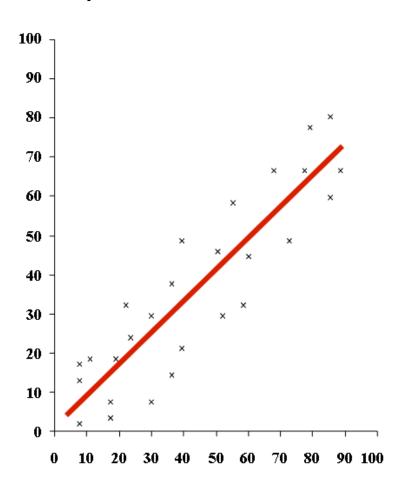
La direzione è negativa se all'aumentare di una variabile diminuisce l'atra.



Ad esempio all'aumentare dei divorzi diminuiscono il numero di nascite

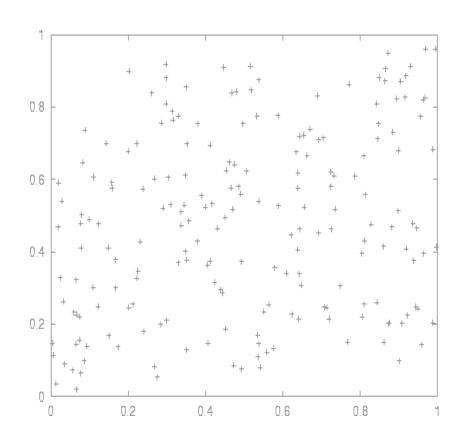
L'entità si riferisce alla forza della relazione esistente tra due variabili.

Quanto più i punteggi sono raggruppati attorno ad una retta, tanto *più forte* è la relazione tra due variabili.



Ad esempio, quanto più elevata è la temperatura, tanto più si suda.

Se i punteggi sono dispersi in maniera uniforme, invece, tra le due variabili non esiste alcuna relazione.



Ad esempio, non esiste alcuna relazione tra la temperatura e la nostra età

Per esprimere la relazione esistente tra due variabili, si utilizza il coefficiente di correlazione

Tale coefficiente è standardizzato e può assumere valori che vanno da **–1.00** (correlazione perfetta negativa) e **+1.00** (correlazione perfetta positiva). Una correlazione uguale a **0** indica che tra le due variabili non vi è alcuna relazione.

Nota. La correlazione non include il concetto di causa-effetto, ma solo quello di rapporto tra variabili. La correlazione ci permette di affermare che tra due variabili c'è una relazione sistematica, ma non che una causa l'altra.

Esistono vari tipi di coefficienti di correlazione a seconda del tipo di scala della variabile.

- Per le scale a intervalli o rapporti equivalenti si usa il coefficiente r di Pearson.
- Per le scale **ordinali** si usano il coefficiente r_s di **Spearman** o il coefficiente *tau* di **Kendall**.

 Per le scale a intervalli o rapporti equivalenti si usa il coefficiente r di Pearson.

Tale coefficiente è calcolato come rapporto tra la covarianza delle due variabili e il prodotto delle loro deviazioni standard

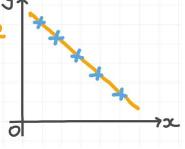
$$-1 \leq
ho_{xy} = rac{\sigma_{xy}}{\sigma_x \sigma_y} = rac{\sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y)}{\sqrt{\sum_{i=1}^n (x_i - \mu_x)^2} \sqrt{\sum_{i=1}^n (y_i - \mu_y)^2}} \leq +1$$

Può assumere valori che vanno da **–1.00** (tra le due variabili vi è una correlazione perfetta negativa) e **+1.00** (tra le due variabili vi è una correlazione perfetta positiva). Una correlazione uguale a 0 indica che tra le due variabili non vi è alcuna relazione.

Between which two values does the product-moment correlation coefficient lie? and 1 P.M.C.C. r20.6? (r) r20? F2-0.8? × X × 7ch 7ch strong negative no correlation moderate positive conelation correlation bertect bositive

conclation

r=-1



correlation r=1

2cf

	CRIM	7	ZN	INDUS	CI	HAS	NOX		RM	AG	E	DIS	R	AD	TAX	PTRATI	о в	LSTAT	Price
CRIM	1.000000	-0.20046	69 O.	406583	-0.0558	892	0.420972	-0.219	247	0.35273	4 -0.3	79670	0.6255	05 (0.582764	0.28994	6 -0.385064	0.455621	-0.38830
ZN	-0.200469	1.00000	00 -0.	533828	-0.0426	697	-0.516604	0.311	991	-0.56953	7 0.6	64408	-0.3119	48 -0	0.314563	-0.39167	9 0.175520	-0.412995	0.36044
INDUS	0.406583	-0.53382	28 1.0	000000	0.0629	938	0.763651	-0.391	676	0.64477	9 -0.7	08027	0.5951	29 (0.720760	0.38324	8 -0.356977	0.603800	-0.48372
CHAS	-0.055892	-0.04269	97 0.0	062938	1.0000	000	0.091203	0.091	251	0.08651	3 -0.0	99176	-0.0073	68 -(0.035587	-0.12151	5 0.048788	-0.053929	0.17526
NOX	0.420972	-0.51660	04 0.	763651	0.0912	203	1.000000	-0.302	188	0.73147	0.7	69230	0.6114	41 (0.668023	0.18893	3 -0.380051	0.590879	-0.42732
RM	-0.219247	0.31199	91 -0.:	391676	0.0912	251	-0.302188	1.000	000	-0.24026	5 0.2	05246	-0.2098	47 -0	0.292048	-0.35550	1 0.128069	-0.613808	0.69536
AGE	0.352734	-0.56953	37 0.0	644779	0.086	518	0.731470	-0.240	265	1.000000	0.7	47881	0.4560	22 (0.506456	0.26151	5 -0.273534	0.602339	-0.37695
DIS	-0.379670	0.66440	08 -0.	708027	-0.099	176	-0.769230	0.205	246	-0.74788	1 1.0	00000	-0.4945	88 -0	0.534432	-0.23247	1 0.291512	-0.496996	0.24992
RAD	0.625505	-0.31194	18 0.:	595129	-0.0073	368	0.611441	-0.209	847	0.45602	2 -0.4	94588	1.0000	00 (0.910228	0.46474	1 -0.444413	0.488676	-0.38162
TAX	0.582764	-0.31456	3 0.	720760	-0.035	587	0.668023	-0.292	048	0.50645	6 -0.5	34432	0.9102	28	1.000000	0.46085	3 -0.441808	0.543993	-0.46853
PTRATIO	0.289946	-0.39167	79 0.:	383248	-0.121	515	0.188933	-0.355	501	0.26151	5 -0.2	32471	0.4647	41 (0.460853	1.00000	0 -0.177383	0.374044	-0.50778
В	-0.385064	0.17552	20 -0.:	356977	0.0487	788	-0.380051	0.128	069	-0.27353	4 0.2	91512	-0.4444	13 -0	0.441808	-0.17738	3 1.000000	-0.366087	0.33346
LSTAT	0.455621	-0.41299	95 0.0	603800	-0.0539	929	0.590879	-0.613	808	0.60233	9 -0.4	96996	0.4886	76 (0.543993	0.37404	4 -0.366087	1.000000	-0.73766
Price	-0.388305	0.36044	45 -0.	483725	0.1752	260	-0.427321	0.695	360	-0.37695	5 0.2	49929	-0.3816	26 -0	0.468536	-0.50778	7 0.333461	-0.737663	1.00000
	CRIM	1 _					-0.22	0.55		0.63									
	OI tilvi																- 0.9		
	ZN	-0.2	1	-0.53	-0.04	-0.52	0.31	-0.57	0.66	-0.31	-0.31	-0.39	0.18	-0.41	0.36		- 0.9		
	INDUS	0.41	-0.53	1	0.06	0.76	-0.39	0.64	-0.71	0.6	0.72	0.38	-0.36	0.6	-0.48				
	CHAS	0.06	0.04	0.06	4	0.00	0.09	0.00	0.1	0.01	0.04	0.12	0.05	0.05	0.19		-06		

- 0.3

- 0.0

- -0.3

- -0.6



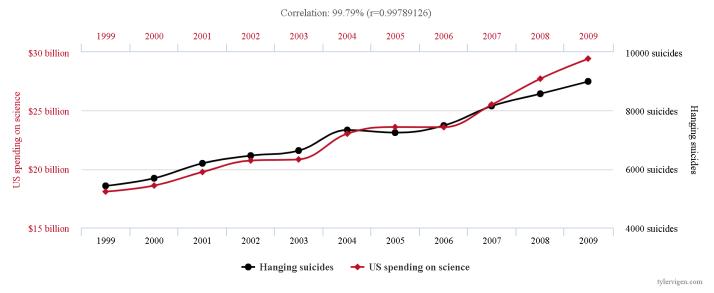


US spending on science, space, and technology

 \equiv

correlates with

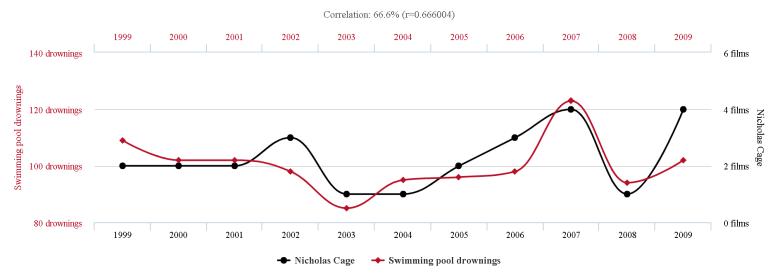
Suicides by hanging, strangulation and suffocation



Number of people who drowned by falling into a pool

correlates with

Films Nicolas Cage appeared in



Divorce rate in Maine

correlates with

Per capita consumption of margarine



People who drowned after falling out of a fishing boat

correlates with

Marriage rate in Kentucky

