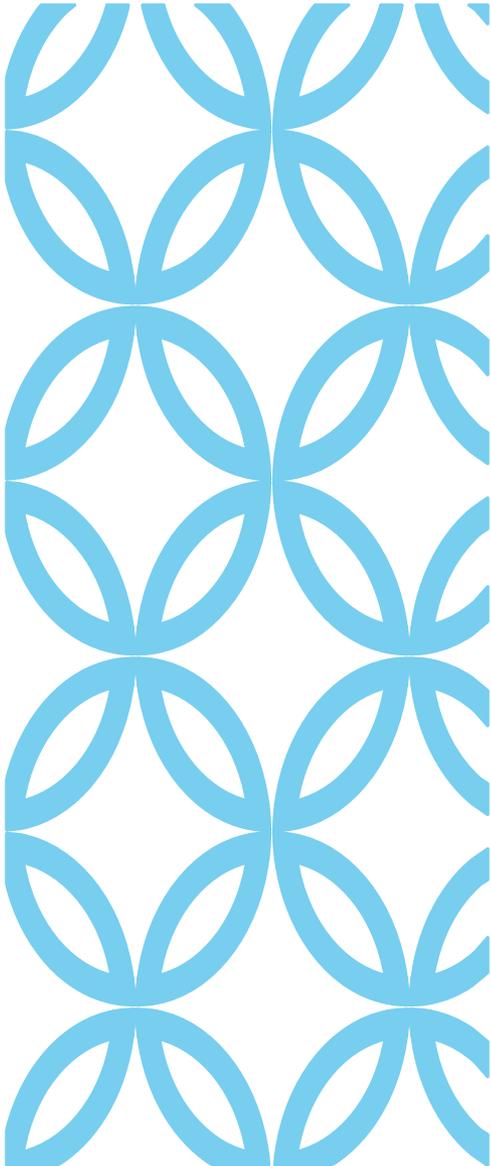


DISTRIBUZIONI

È interessante analizzare il legame di dipendenza esistente tra due (o più) variabili.



In base alla natura delle variabili, possono esistere modi differenti per studiare la dipendenza:

X e Y entrambe qualitative

X e Y entrambe quantitative

X e Y una qualitativa una quantitativa

DISTRIBUZIONI

DISTRIBUZIONI DOPPIE

**X e Y entrambe
quantitative**
variabili

unità	X	Y
1	x_1	y_1
2	x_2	y_2
i	x_i	y_i
n	x_n	y_n

DISTRIBUZIONI DOPPIE

Esempio (1)

X: QI madre

Y: QI figlia

**Esiste un legame tra le
due variabili?**

QI madre	QI figlia
135	121
124	112
115	99
127	131
120	115
112	118
104	106
96	89
94	92
85	90

Rappresentazione grafica

DIAGRAMMA A DISPERSIONE: I due caratteri X e Y corrispondono agli assi ortogonali di un piano cartesiano. Le coppie di modalità osservate (x_i, y_i) sono le coordinate dei punti rappresentati sul piano cartesiano.

Si ottiene la cosiddetta **nuvola** di punti

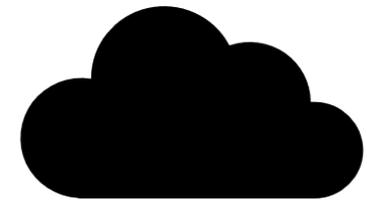
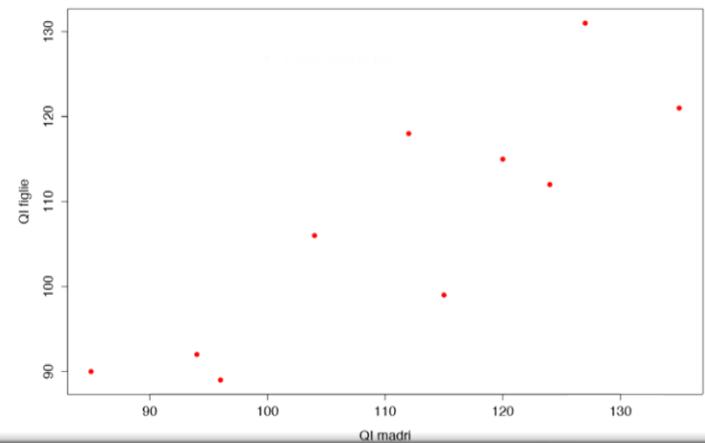


DIAGRAMMA A DISPERSIONE

DIAGRAMMA A DISPERSIONE

Diagramma a dispersione

Esempio (1)



ASSOCIAZIONI POSITIVE E NEGATIVE



Due variabili sono **associate positivamente** (concordi) quando i valori sopra la media di una tendono ad associarsi con i valori sopra la media dell'altra e allo stesso modo si comportano i valori sotto la media.



Due variabili sono **associate negativamente** (discordi) quando i valori sopra la media di una tendono ad associarsi con i valori sotto la media dell'altra e viceversa.

MISURARE L'ASSOCIAZIONE LINEARE

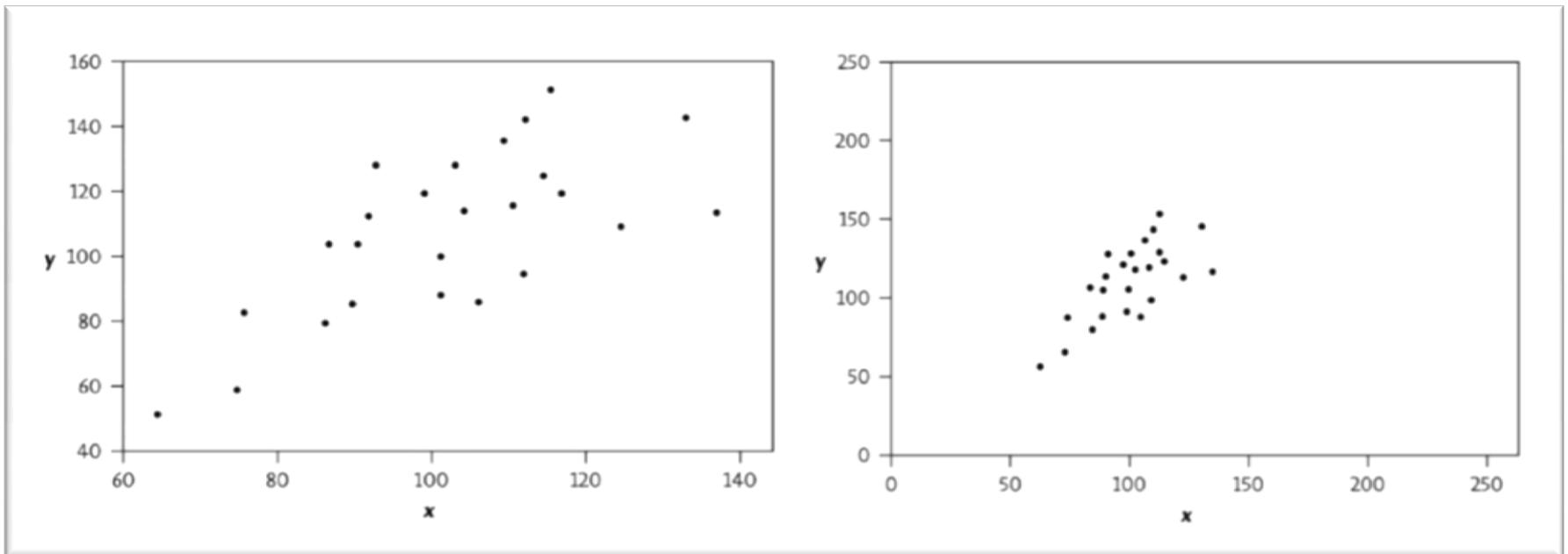
Osservando la nuvola di punti possiamo dedurre informazioni sulla forma, sulla direzione e sulla forza della relazione fra due variabili quantitative.

Le relazioni lineari, ovvero le relazioni la cui forma è rappresentabile tramite una linea retta, sono abbastanza comuni.

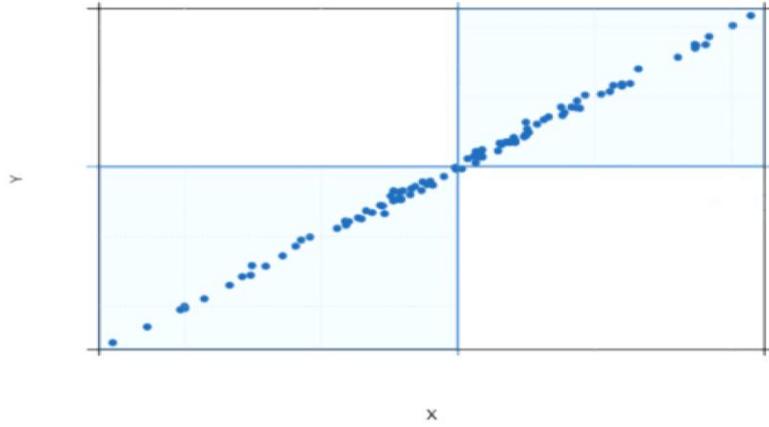
Una relazione lineare è **forte** se i punti si trovano “vicini” ad una linea retta. La relazione è invece **debole** se i punti si trovano molto “lontani” da una retta.

ATTENZIONE: A
OCCHIO È DIFFICILE
GIUDICARE LA
«FORZA»...

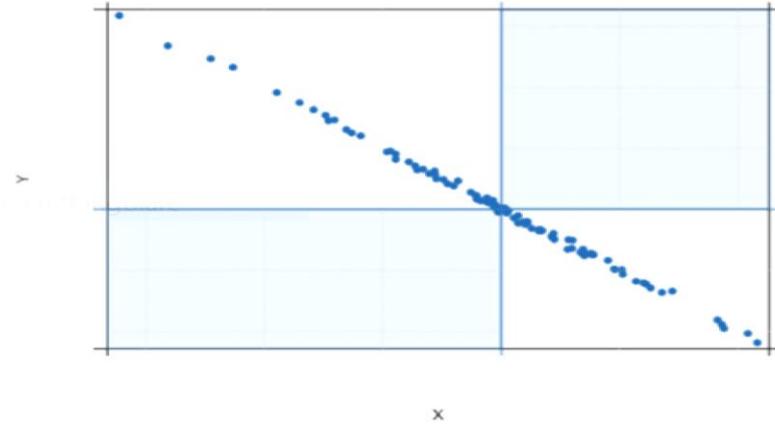
I due grafici a dispersione rappresentano gli stessi dati, ma nella figura di destra la relazione appare più forte a causa dello spazio bianco circostante.



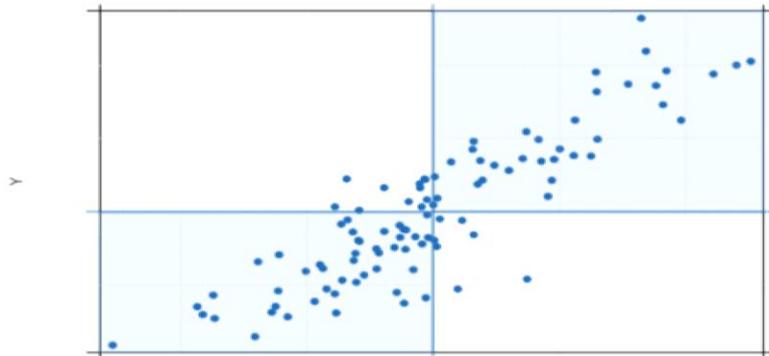
forte relazione lineare positiva :: $r = 0.999$



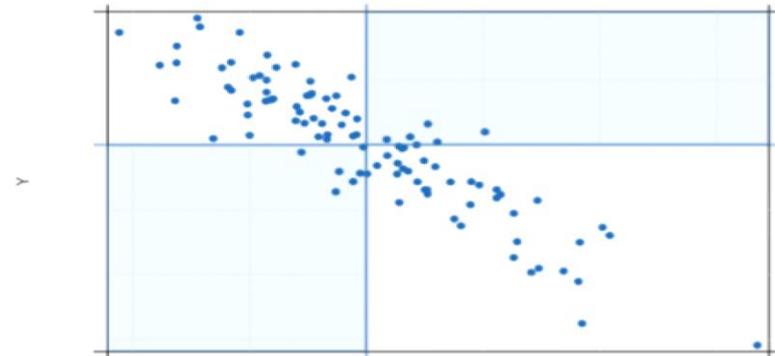
forte relazione lineare negativa :: $r = -0.999$



relazione lineare positiva :: $r = 0.882$

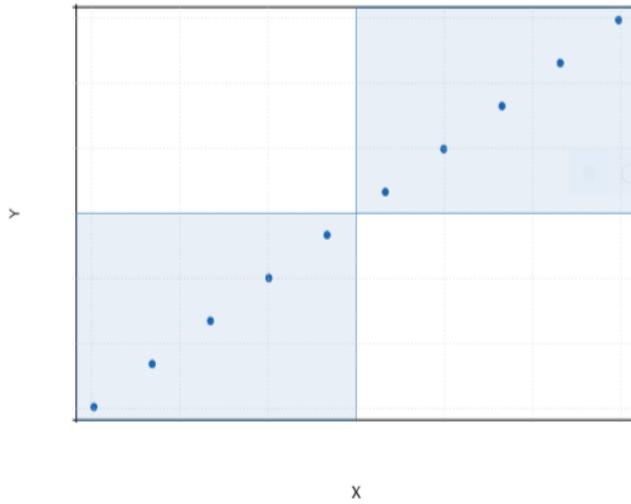


relazione lineare negativa :: $r = -0.901$

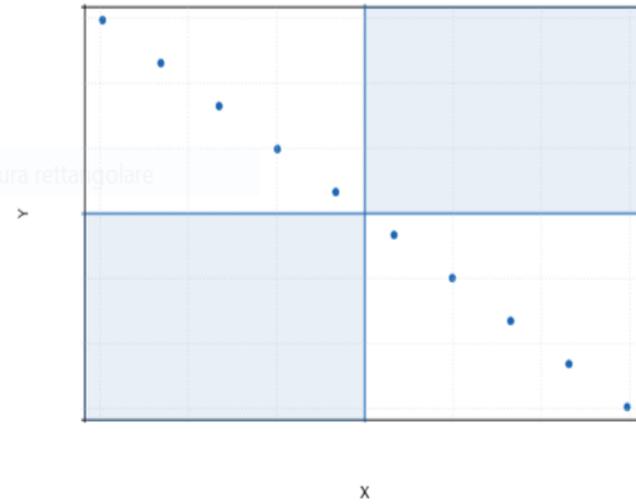


MISURARE L'ASSOCIAZIONE LINEARE

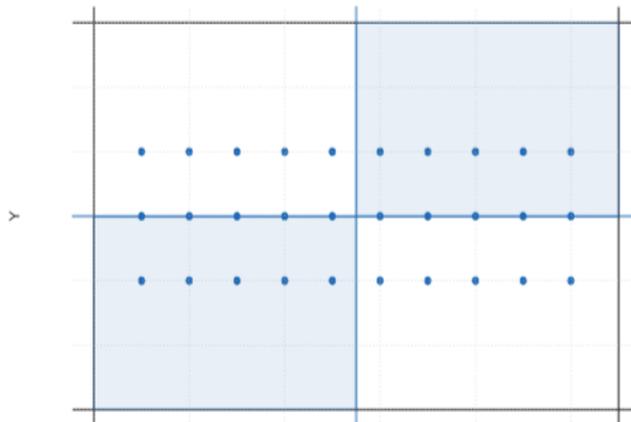
perfetta relazione lineare positiva :: $r = 1$



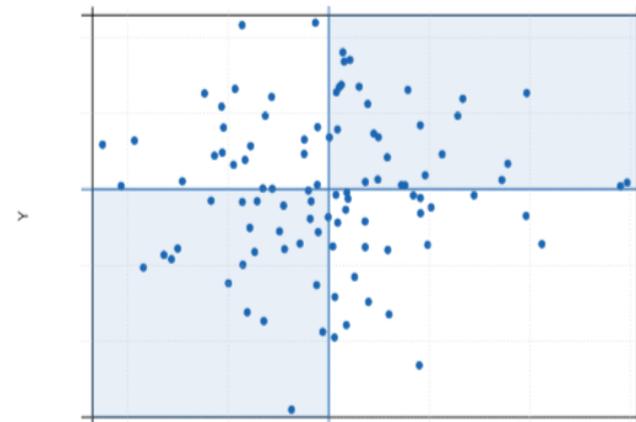
perfetta relazione lineare negativa :: $r = -1$



Indipendenza \Rightarrow Incorrelazione :: $r = 0$



sostanziale incorrelazione :: $r = 0.038$



È OPPORTUNO
UTILIZZARE UN
INDICE
NUMERICO PER
SUPPORTARE
L'INFORMAZIONE
E OTTENUTA
DALL'ESPLORAZIONE
GRAFICA.

Le misure che introduciamo sono la
COVARIANZA s_{XY} e

**IL COEFFICIENTE DI CORRELAZIONE LINEARE
DI BRAVAIS PEARSON** r_{XY} che quantificano
la direzione e la forza della relazione
lineare fra le due variabili quantitative

UN INDICE SIMMETRICO CHE MISURA LA CONCORDANZA O LA DISCORDANZA TRA DUE CARATTERI QUANTITATIVI È LA **COVARIANZA**

$$r_{XY} = \frac{s_{XY}}{s_X s_Y}$$

$$s_{XY} = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y}_n)(x_i - \bar{x}_n)$$

ALTRO INDICE È IL COEFFICIENTE DI CORRELAZIONE LINEARE DI BRAVAIS-PEARSON

**TORNIAMO
ALL'ESEMPIO**

Facciamo un po' di conti