

STATISTICA

Regressione-1

Associazione

Voto per Macron e tasso di disoccupazione

Appartenenza etnica e preferenze politiche

Esposizione ad una data sostanza e insorgenza di malattie

Livello sociale della famiglia e riuscita scolastica dei figli

Metodo colturale e qualità della pro...

Dimensione del cervello e QI

Genere e percorso di studi

ENTRAMBE LE
VARIABILI
QUALITATIVE:

TABELLE DI
CONTINGENZA

Interpretare l'associazione

The screenshot shows a web browser window displaying the AIRC website. The article title is "È vero che la pillola anticoncezionale aumenta il rischio di cancro?". A green callout box on the right contains the text: "L'associazione da sola non implica una diretta relazione di causa-effetto." The article text below the title reads: "Dipende dal tipo di cancro. In alcuni casi può aumentare il rischio, in altri ridurlo. In ogni caso la scelta di farne uso deve comprendere una valutazione generale dei rischi e dei benefici che vanno oltre la sola valutazione del rischio oncologico." Below the text is an image of a blister pack of white pills.

L'associazione da sola non implica una diretta relazione di causa-effetto.

L'associazione può essere causata da una terza variabile non osservata.

Interpretare l'associazione

L'associazione da sola **non** implica una diretta relazione di causa-effetto.

L'associazione può essere causata da una terza variabile non osservata.

Parla con noi 

Pillola e cancro della cervice uterina

Un uso prolungato di contraccettivi orali è stato associato a un aumento del rischio di cancro della cervice uterina che tende a diminuire alla cessazione. Nel 2002, però, uno studio dell'Organizzazione mondiale della sanità ha dimostrato che **l'aumento di rischio è legato alla presenza di un'infezione da virus del Papilloma umano (HPV)**.

Di conseguenza è probabile che fra la pillola e il tumore vi sia un legame indiretto: le donne che usano questo metodo contraccettivo in genere non usano il preservativo e sono quindi a maggior rischio di contrarre l'HPV, il vero responsabile del tumore della cervice uterina. Questo problema dovrebbe ridursi man mano che aumenta la percentuale di persone (femmine e maschi) vaccinate contro l'HPV.

	Tumore	No tumore
Uso pillola		
Non uso pillola		

MIND

MENTE & CERVELLO

Alla ricerca del tempo perduto

Vivere in armonia con l'orologio biologico che regola i nostri ritmi vitali è la chiave per un buon equilibrio del corpo e della mente



44 **Psicologia**
Madri pentite

56 **Fotografia**
Dietro ci si affannava

Il tempo giusto per tutti

L'editoriale del n.160 di Mind
di Marco Cattaneo

Il tempo giusto per tutti



di MARCO CATTANEO

Gufo, decisamente gufo. Secondo i parametri che assegnano un profilo ai nostri ritmi biologici, sono un gufo. Non nel senso di menagramo, ma di inguaribile tiratardi che non andrebbe mai a dormire la sera e non si alzerebbe mai dal letto al mattino.

Nel test che trovate a pagina 42, d'altra parte, ho totalizzato 22 punti. Considerato che si è classificati come gufi sotto i 43 punti, il mio «cronotipo» non è nemmeno in discussione. Ci vegliamo almeno 64 punti invece per essere un'allodola, vale a dire uno che ama alzarsi alle prime luci dell'alba e che non ha perso l'abitudine, e il gusto, di codicarsi dopo Carosello. (Che, a beneficio dei nati dopo il 1977, era un programma televisivo di lunghe clip pubblicitarie con siparietti comici e cartoni animati molto amato dai bambini).

Certo, quel test è poco più che un gioco, ma è evidente che ognuno di noi ha i suoi tempi. A cominciare proprio dall'orologio interno che regola il ciclo sonno/veglia, il più conosciuto dei ritmi circadiani. È talmente ben codificato nel nostro «libretto delle istruzioni» che si conserva anche in condizioni di isolamento in assenza di luce naturale e di orologi, il nostro corpo segue automaticamente un ciclo di circa 25 ore.

Ma non è l'unico, anzi. Diverso da persona a persona è l'orologio che scandisce i tempi della fame, per esempio. E

diversi sono orologi molto meno conosciuti e che, soprattutto, non percepiamo, perché non sono sintomatici. Così, mentre la melatonina regola il ritmo giorno/notte, anche altri ormoni seguono un andamento periodico, che governa i nostri comportamenti. «È probabile - racconta Henrik Oster a pagina 26 - che ciascuna cellula del nostro organismo contenga addirittura un proprio orologio che, in base alla funzione dell'organo in cui si trova, attiva o disattiva durante il giorno e in modo coordinato i programmi biologici adeguati».

Ma c'è di più. Da una parte i nostri ritmi biologici non rimangono inalterati per tutta la vita. In genere siamo allodole da bambini e diventiamo gufi da adolescenti, per poi ritornare gradualmente allodole. E con l'età si riducono le differenze tra il giorno e la notte nei numerosi processi fisiologici guidati dai ritmi circadiani. Dall'altra, ci sono fattori esterni che influenzano sensibilmente i nostri orologi biologici. La luce e l'assunzione di cibo sono i due principali.

A questo proposito, gli studi su chi lavora di frequente facendo turni di notte

hanno evidenziato che lo sfasamento dei ritmi circadiani può portare a una maggiore frequenza di malattie anche gravi, come diabete di tipo 2 e disturbi cardiovascolari. Se poi si tratti soltanto di correlazioni oppure di fattori causali, è ancora da stabilire.

È invece assodato che il lungo inverno artico porti a una maggiore frequenza di depressione nelle popolazioni che vivono ad alte latitudini, a causa di quello che è chiamato «disturbo affettivo stagionale». Questo genere di disturbi, cui contribuisce il malfunzionamento del nostro orologio interno, può essere trattato con la terapia della luce, come racconta Irene Campagna a pagina 36 in un'intervista a Cristina Anna Colombo, direttore del Centro disturbi dell'umore dell'IRCCS Ospedale San Raffaele Turro, a Milano. Diffusa soprattutto nei paesi scandinavi, è arrivata però anche da noi, e viene usata per la depressione grave ricorrente.

Anche se non soffriamo di disturbi dell'umore, però, ascoltare il ritmo dei nostri orologi biologici è importante per il benessere di tutti. Sia che siate gufi, sia che siate allodole.

A questo proposito, gli studi su chi lavora di frequente facendo turni di notte hanno evidenziato che lo sfasamento dei ritmi circadiani può portare a una maggiore frequenza di malattie anche gravi, come diabete di tipo 2 e disturbi cardiocircolatori. **Se poi si tratti soltanto di correlazioni oppure di fattori causali è ancora da stabilire**

Il tempo giusto per tutti

Editoriale del n.160 di Mind
di Marco Cattaneo

Il tempo giusto per tutti



di MARCO CATTANEO

mente gufo. Secondo i parametri che assegnano ai nostri ritmi biologici, sono un gufo. Non nel senso di un animale notturno, ma di inguoribile tiratardi che non andrebbe mai a letto prima di notte e non si alzerebbe mai dal letto al mattino.

rate a pagina 42, d'altronde 22 punti. Considerati come gufi sotto

diversi sono orologi molto meno conosciuti e che, soprattutto, sono meno precisi.

Così, il ritmo circadiano del nostro organismo è governato da un orologio interno, che governa i nostri ritmi biologici. «È probabile che il ritmo circadiano sia governato da una cellula del nostro organismo che, in base alla funzione che svolge, si attiva o disattiva durante il giorno e in modo coordinato con i programmi biologici adeguati».

Ma c'è di più. Da una parte i nostri ritmi biologici non rimangono inalterati per tutta la vita. In genere siamo allodole da bambini e diventiamo gufi da adolescenti, per poi ritornare gradualmente allodole. E con l'età si riducono le differenze tra il giorno e la notte nei numerosi processi fisiologici guidati dai ritmi circadiani. Dall'altra, ci sono fattori esterni che influenzano sensibilmente i nostri orologi biologici. La luce e l'assunzione di cibo sono i due principali.

A questo proposito, gli studi su chi lavora di frequente facendo turni di notte

hanno evidenziato che lo sfasamento dei ritmi circadiani può portare a una maggiore frequenza di malattie anche gravi, come diabete di tipo 2 e disturbi cardiocircolatori. Se poi si tratti soltanto di correlazioni oppure di fattori causali, è ancora da stabilire.

È invece assodato che il lungo inverno artico porti a una maggiore frequenza di depressione nelle popolazioni che vivono ad alte latitudini, a causa di quello che è chiamato «disturbo affettivo stagionale». Questo genere di disturbi, cui contribuisce il malfunzionamento del nostro orologio interno, può essere trattato con la terapia della luce, come racconta Irene Campagna a pagina 36 in un'intervista a Cristina Anna Colombo, direttore del Centro disturbi dell'Umcore dell'IRCCS Ospedale San Raffaele Turro, a Milano. Diffusa soprattutto nei paesi scandinavi, è arrivata però anche da noi, e viene usata per la depressione grave ricorrente.

Anche se non soffriamo di disturbi dell'umore, però, ascoltare il ritmo dei nostri orologi biologici è importante per il benessere di tutti. Sia che siate gufi, sia che siate allodole.

Correlazione

Voto per Macron e tasso di disoccupazione

Appartenenza etnica e preferenze politiche

Esposizione ad una data sostanza e insorgenza di malattie

Livello sociale della famiglia e riuscita scolastica dei figli

Metodo colturale e produzione

Dimensione del cervello e QI

Genere e percorso di studi



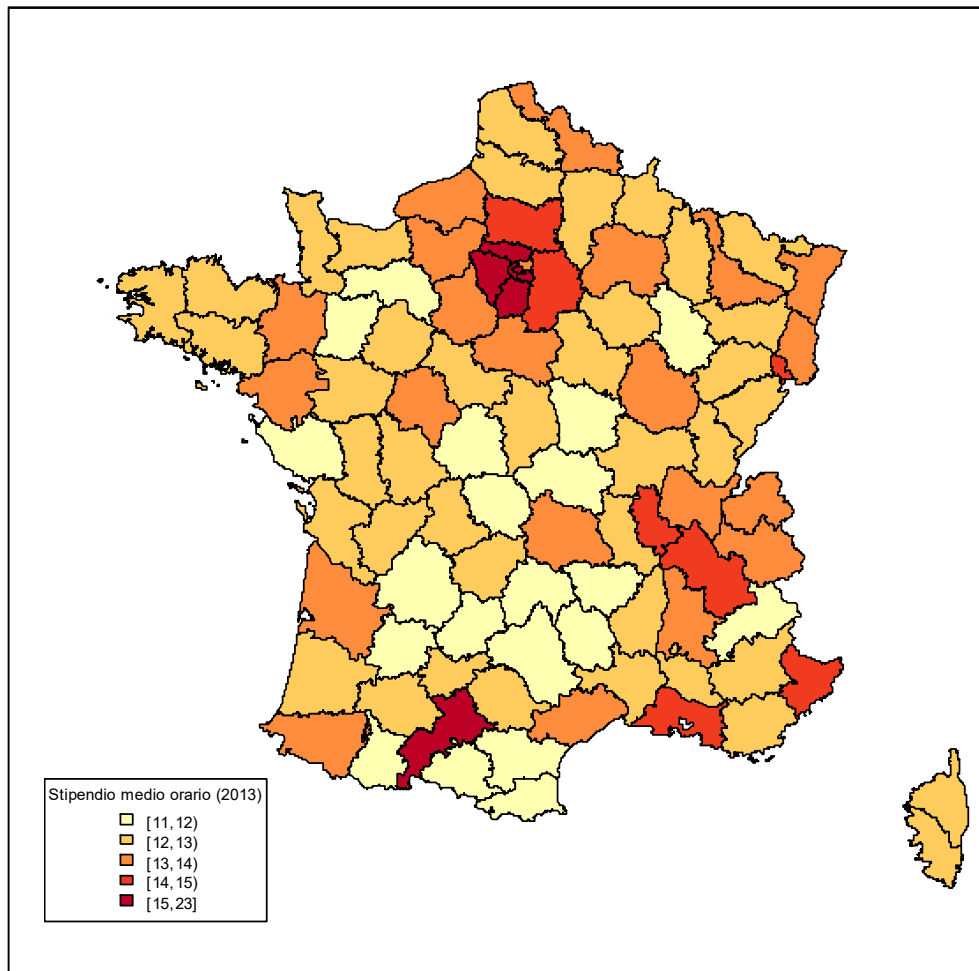
ENTRAMBE LE
VARIABILI
QUANTITATIVE:

ANALISI DI
REGRESSIONE

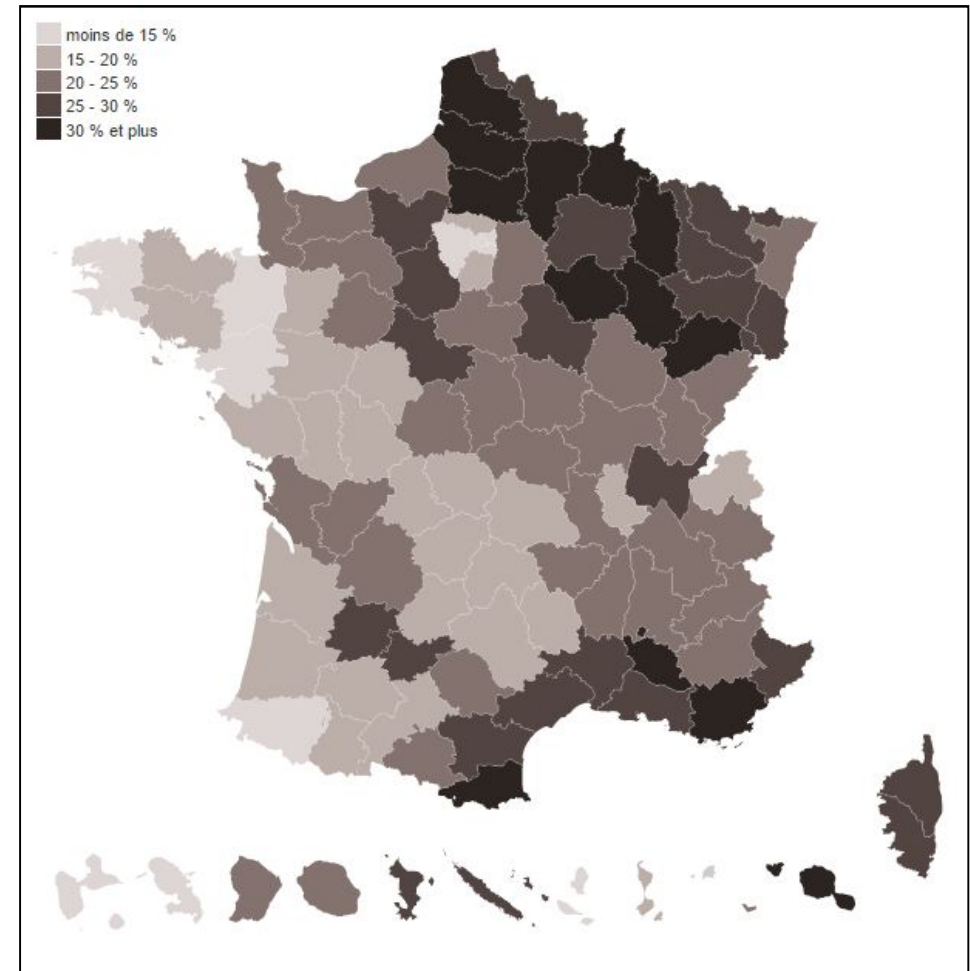
Il voto francese

Elezioni francesi
2017

Stipendio medio orario
2013

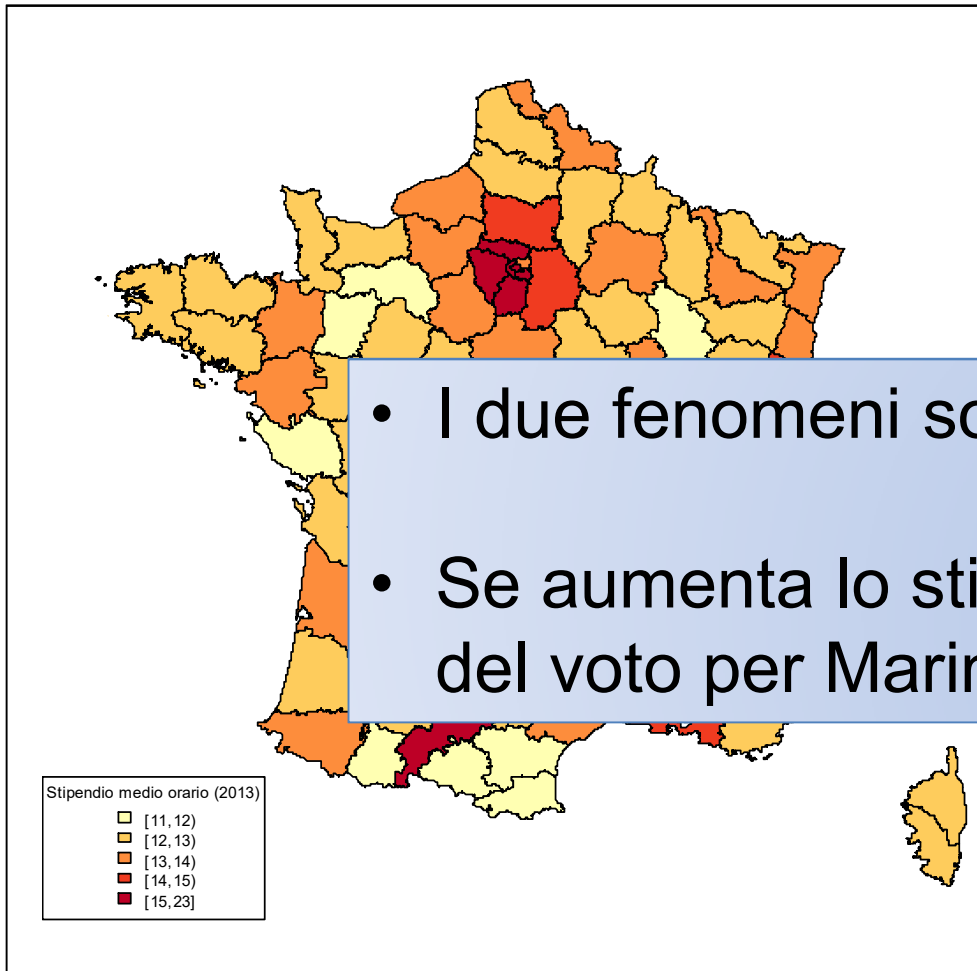


Voto per Le Pen



Il voto francese

Stipendio medio orario
2013

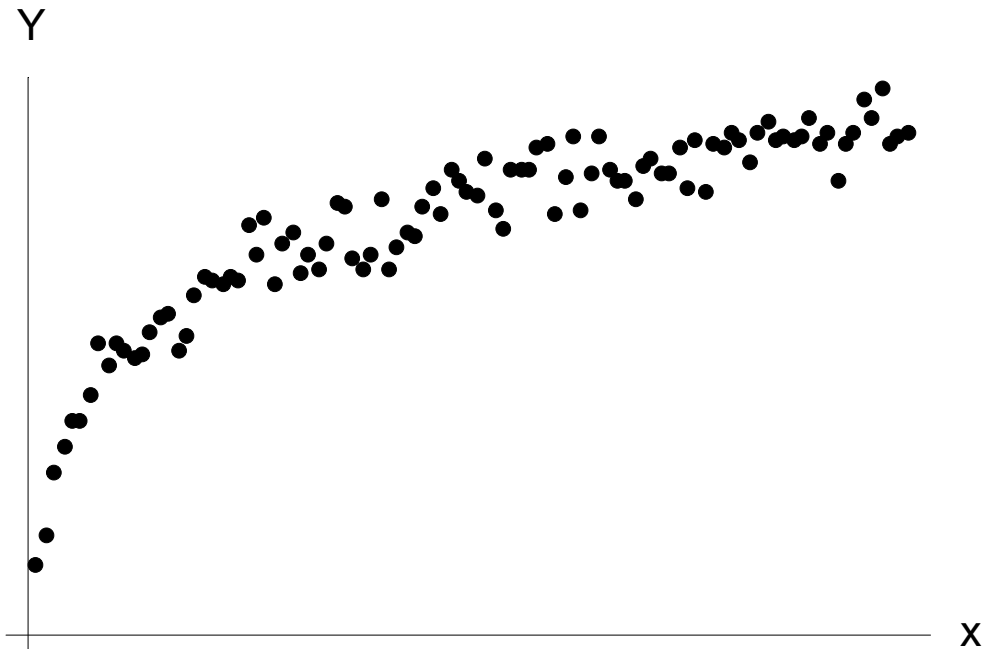


Voto per Le Pen

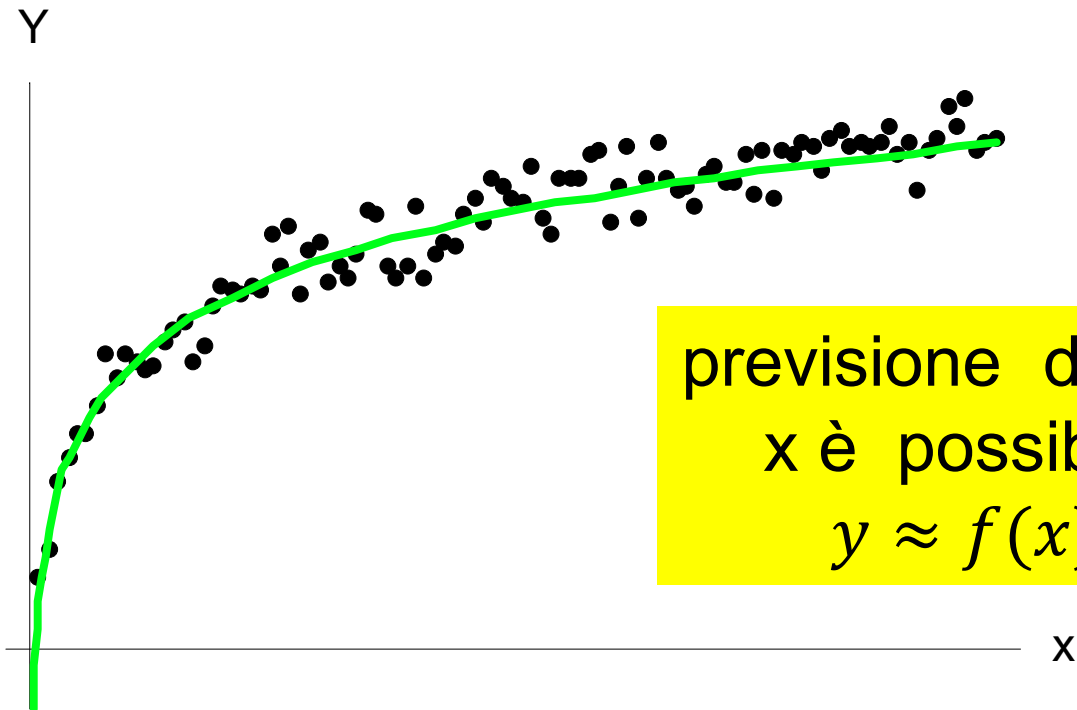


- I due fenomeni sono **correlati**?
- Se aumenta lo stipendio, che ne è del voto per Marine Le Pen?

Obiettivo generale



Obiettivo generale

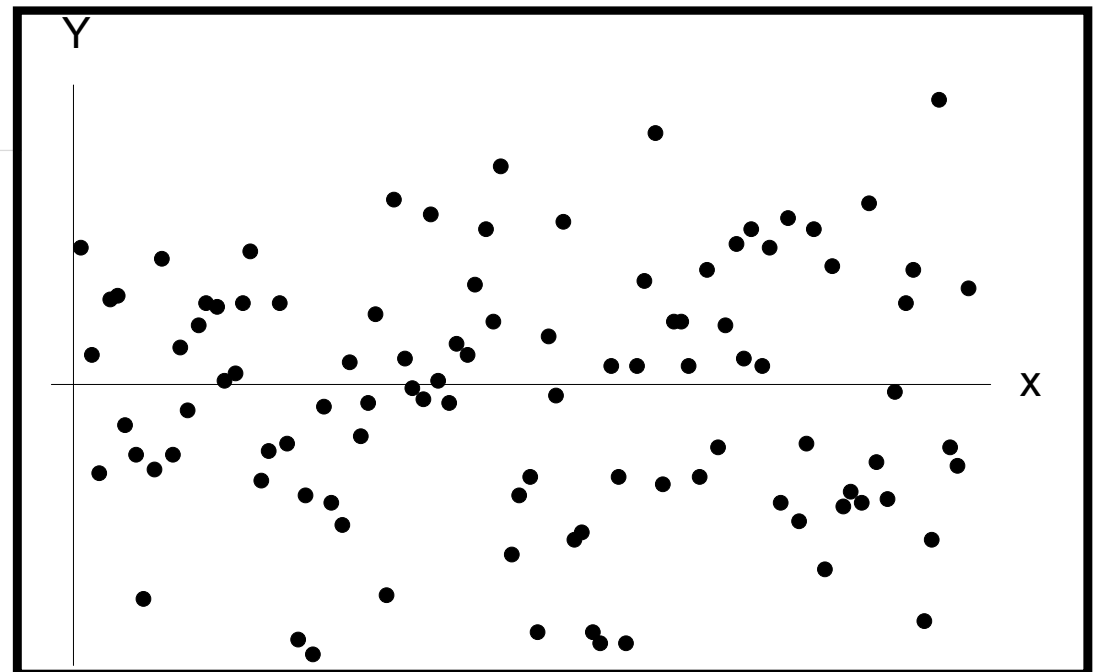


previsione di y da
 x è possibile
 $y \approx f(x)$

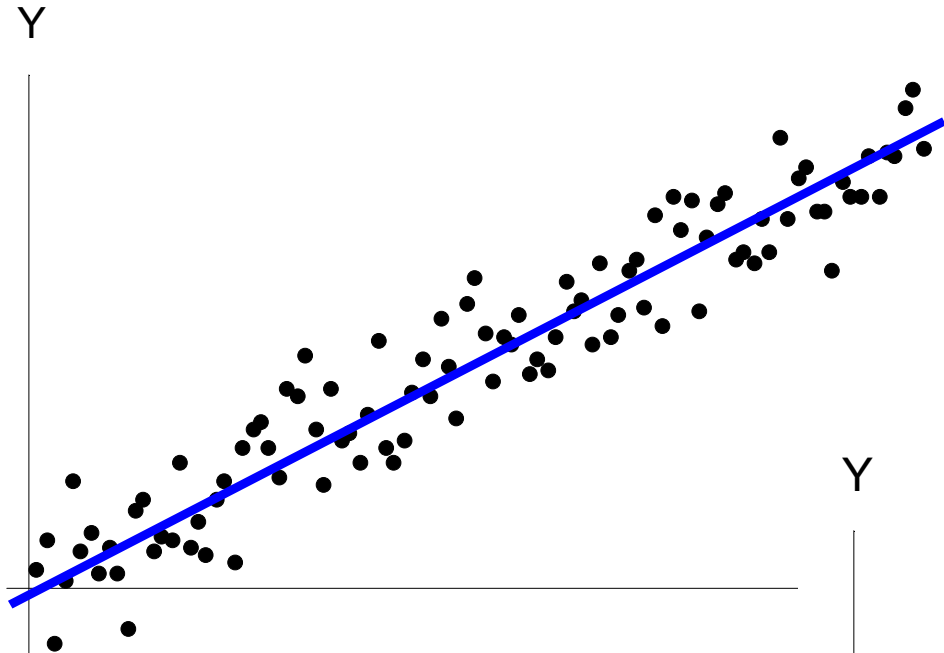
Obiettivo generale



previsione di y
da x *non* è
possibile

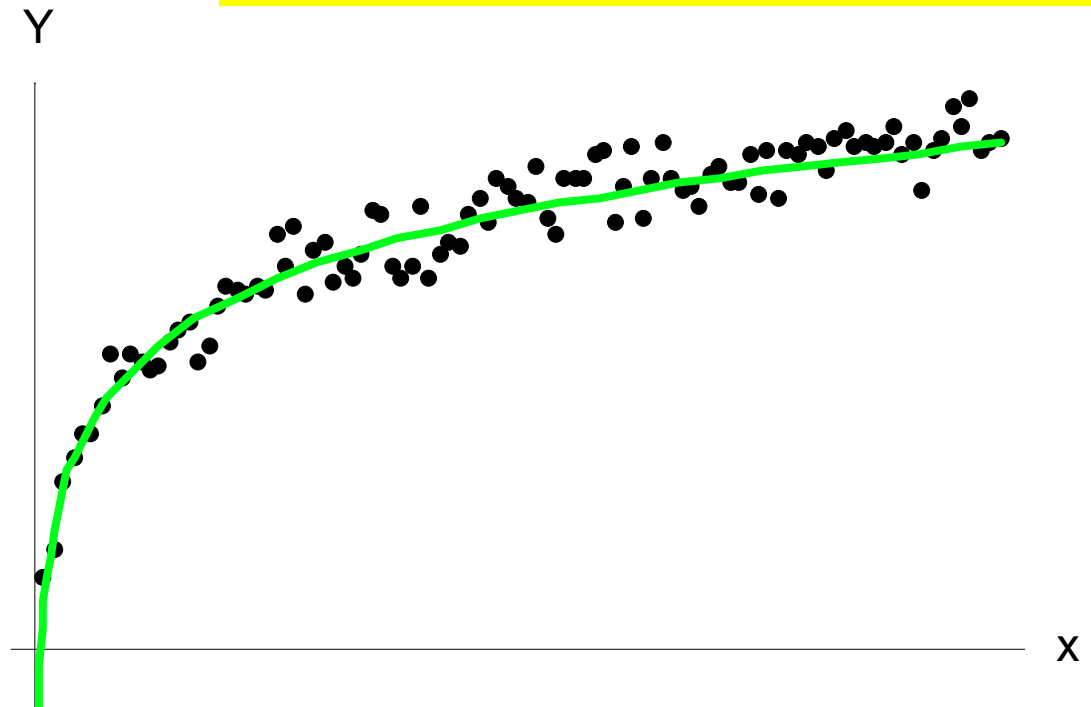


Obiettivo generale

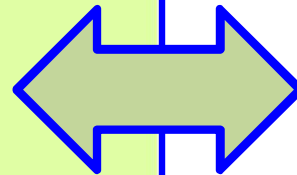
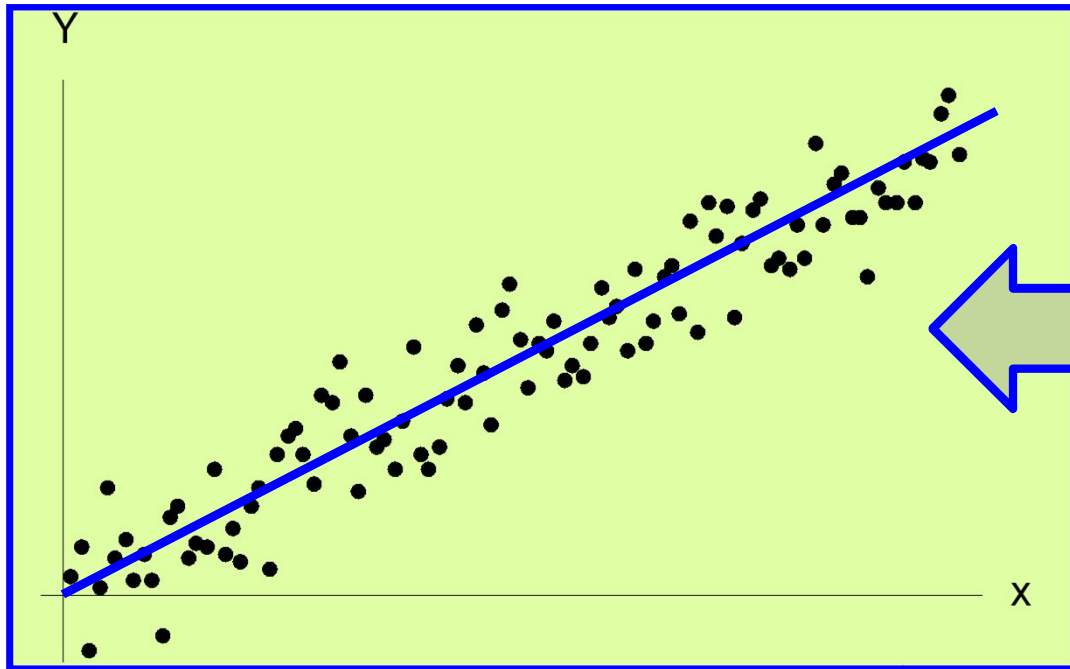


Un **modello** per la previsione di una variabile Y in funzione di un'altra variabile, X :

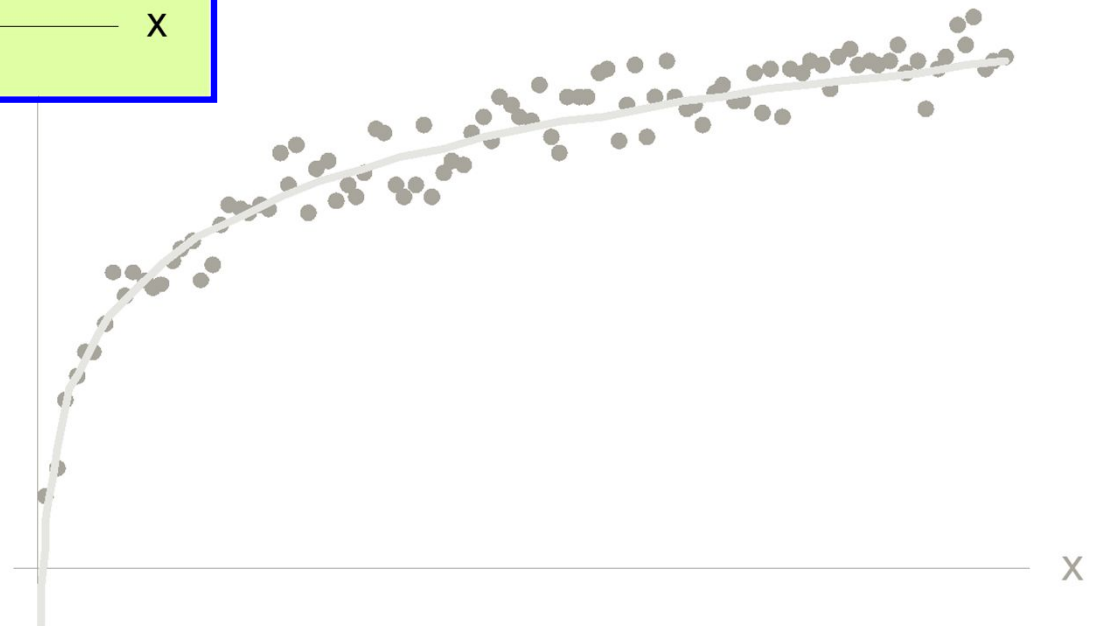
$$y \approx f(x)$$



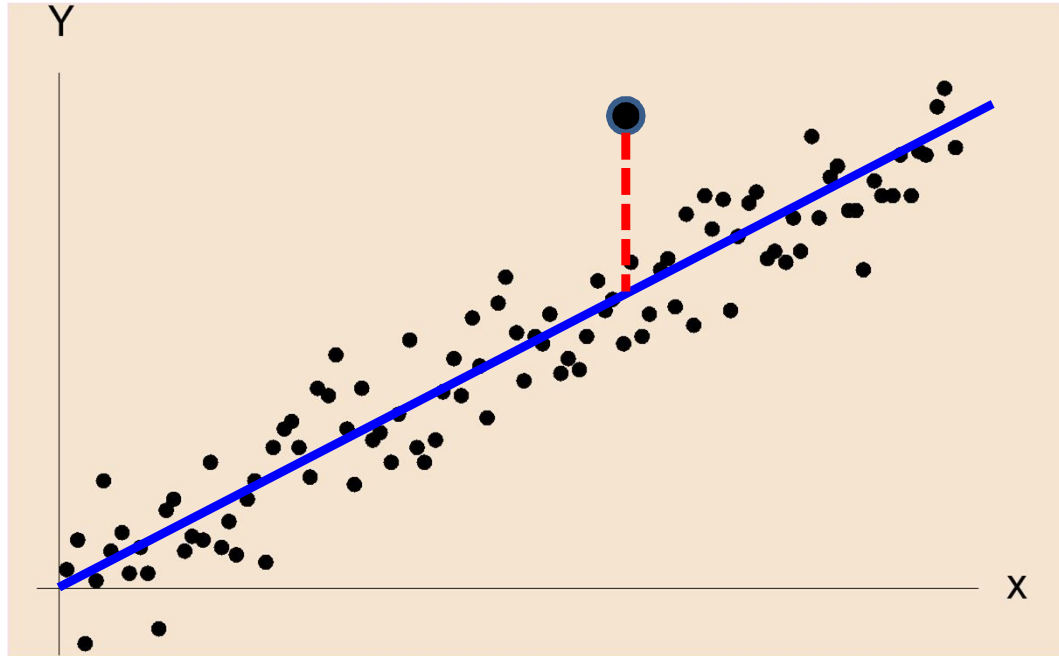
Obiettivo specifico



Il modello della
**regressione lineare
semplice:**



Obiettivo specifico



Il modello della
**regressione lineare
semplice:**

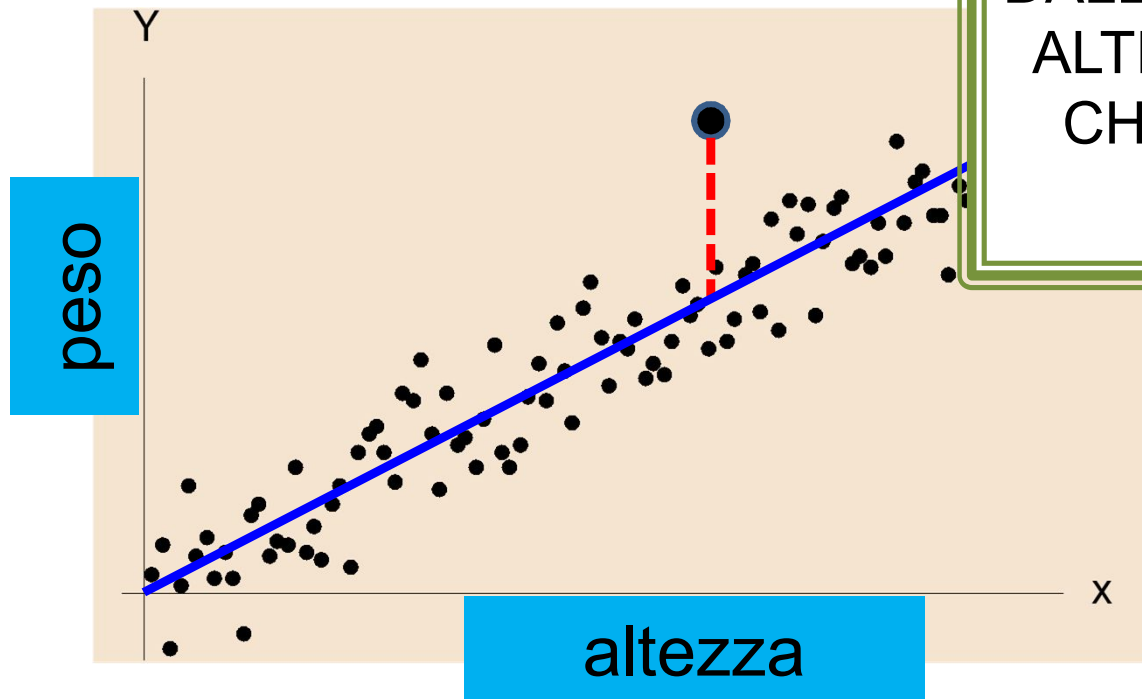
$$f(x_i) = a + bx_i$$
$$\approx \Leftrightarrow \varepsilon_i \sim N(0, \sigma^2)$$

ε_i indipendenti

$$Y_i = a + bx_i + \varepsilon_i$$

Supponiamo l'esistenza di una associazione, **nella popolazione** di riferimento, tra la variabile Y e la variabile X , che si possa approssimare con una relazione lineare a meno di un errore casuale. Dal campione vogliamo stimare l'equazione della retta e la varianza dell'errore.

Obiettivo specifico



IL PESO DIPENDE
DALL'ALTEZZA, MA ANCHE DA
ALTRI FATTORI INDIVIDUALI,
CHE CI SCOSTANO DALLA
RETTA

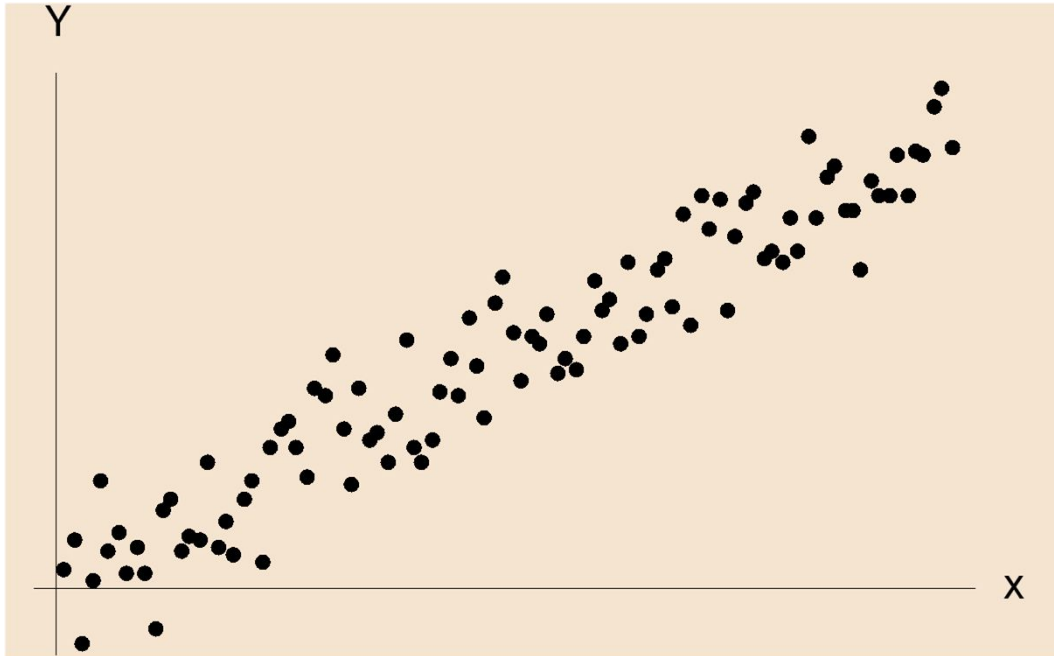
$$Y = a + bx + \varepsilon_i$$

$\approx \Leftrightarrow \varepsilon_i \sim N(0, \sigma^2)$
 ε_i indipendenti

$$Y_i = a + bx_i + \varepsilon_i$$

Supponiamo l'esistenza di una associazione, **nella popolazione** di riferimento, tra la variabile Y e la variabile X , **che si possa approssimare con una relazione lineare a meno di un errore casuale**. Dal campione vogliamo stimare l'equazione della retta e la varianza dell'errore.

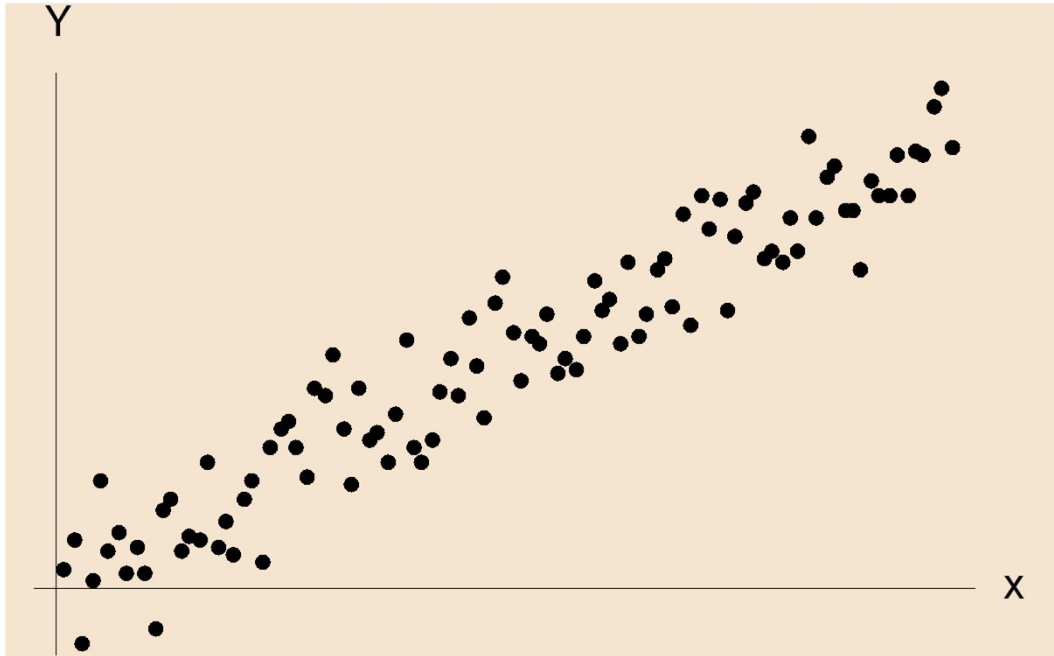
Regressione lineare



A. Valutazione preliminare se una retta possa essere una buona approssimazione

1. GRAFICO di dispersione

Regressione lineare

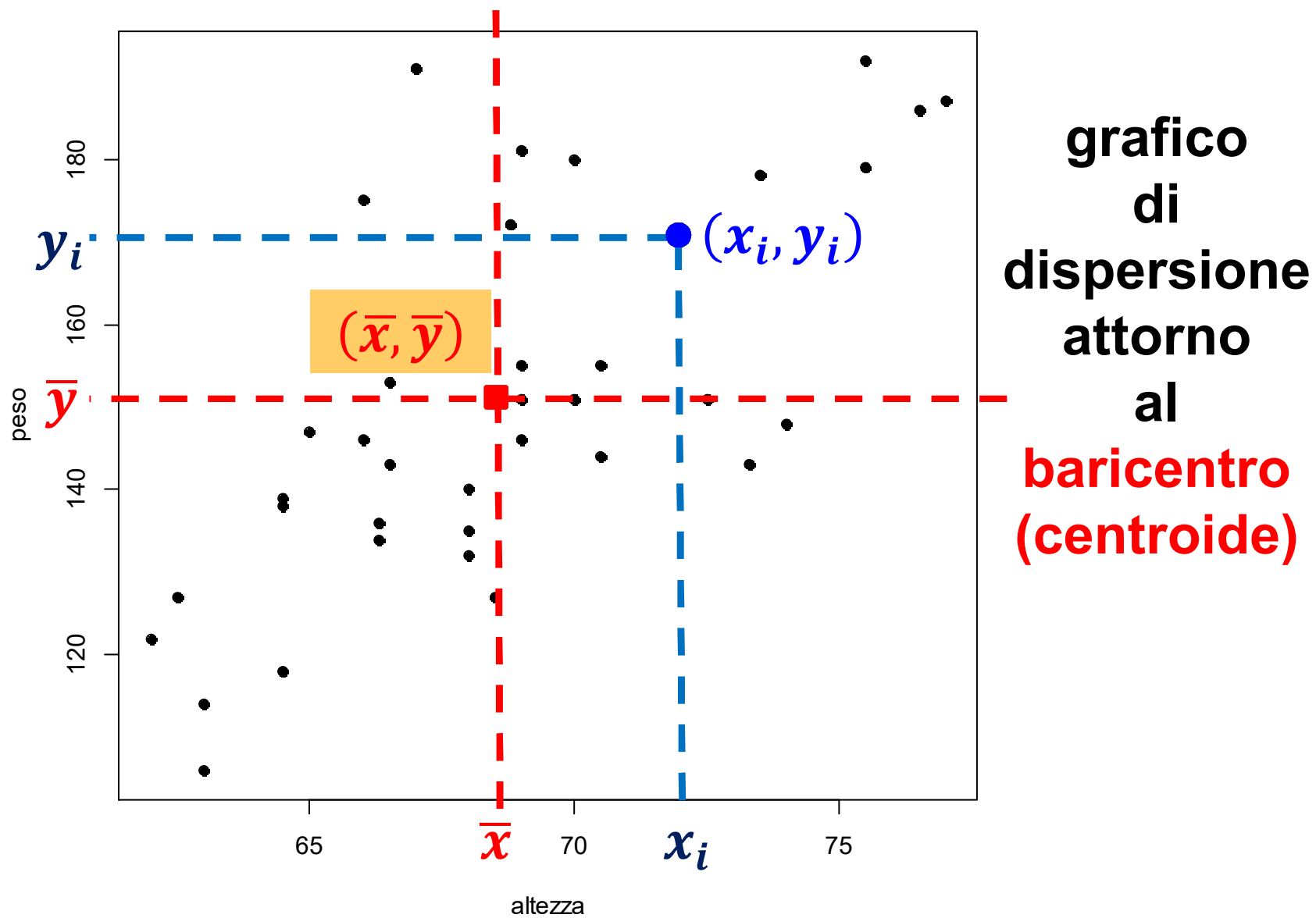


A. Valutazione preliminare se una retta possa essere una buona approssimazione

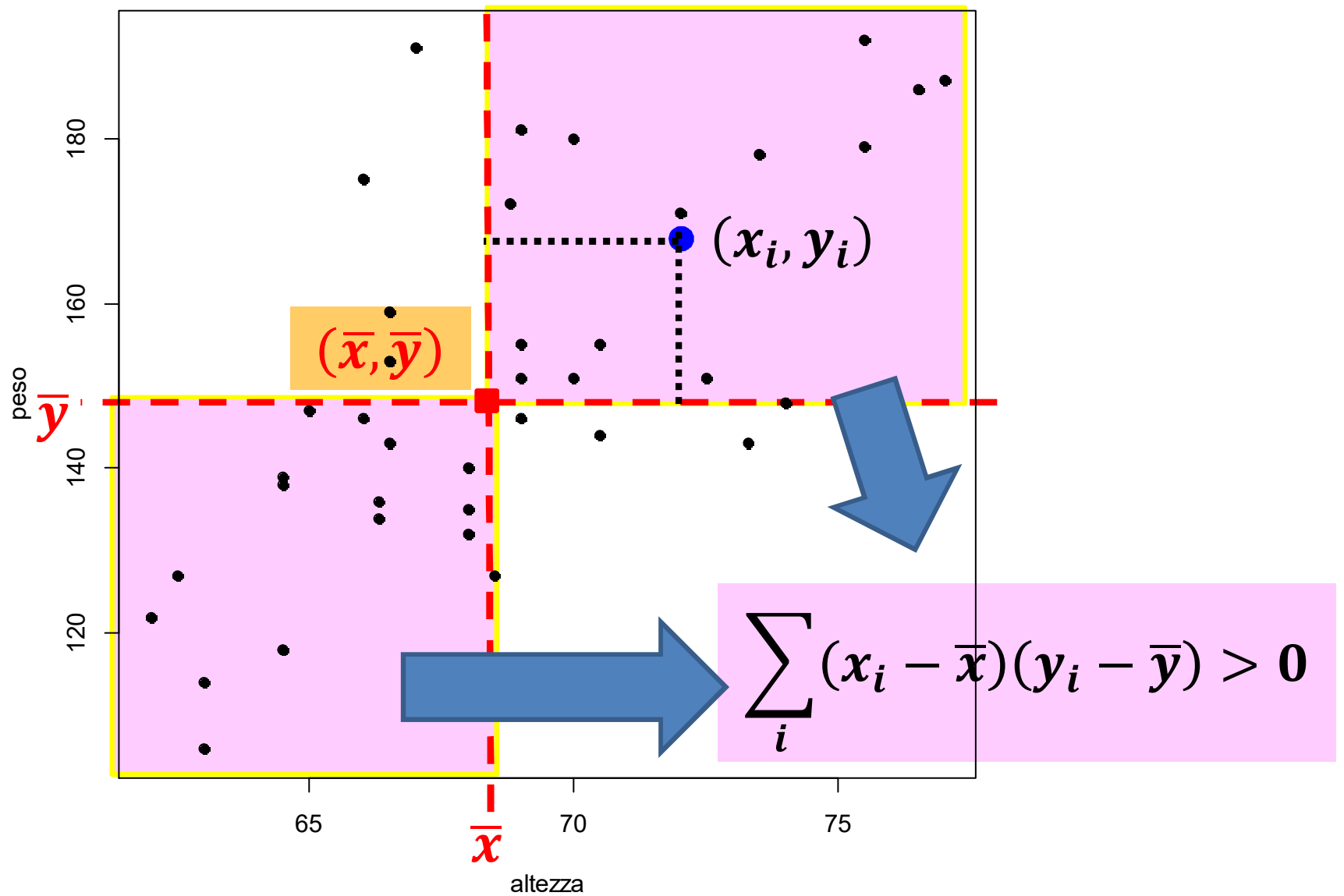
1. GRAFICO di dispersione

2. Indici opportuni

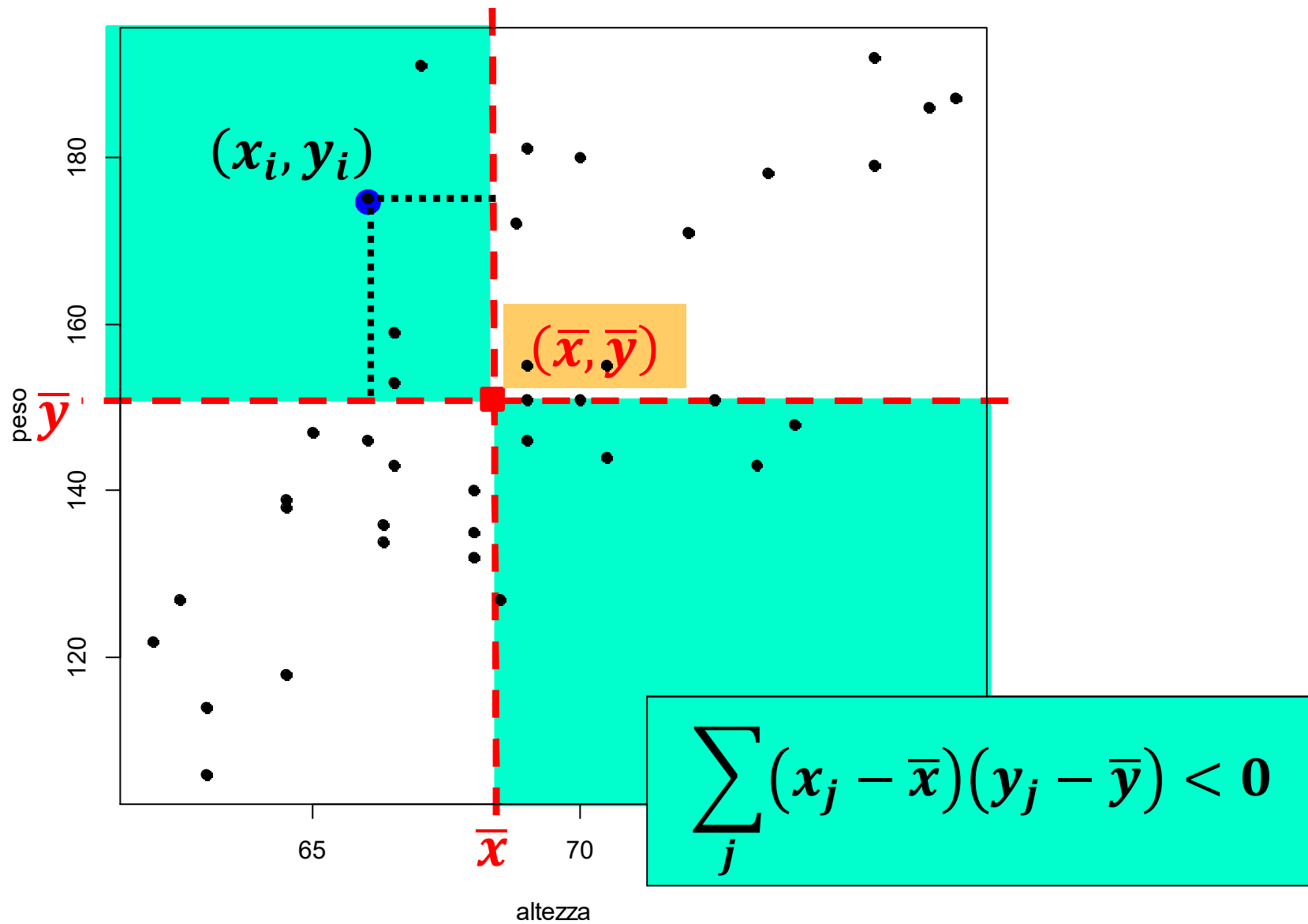
La covarianza



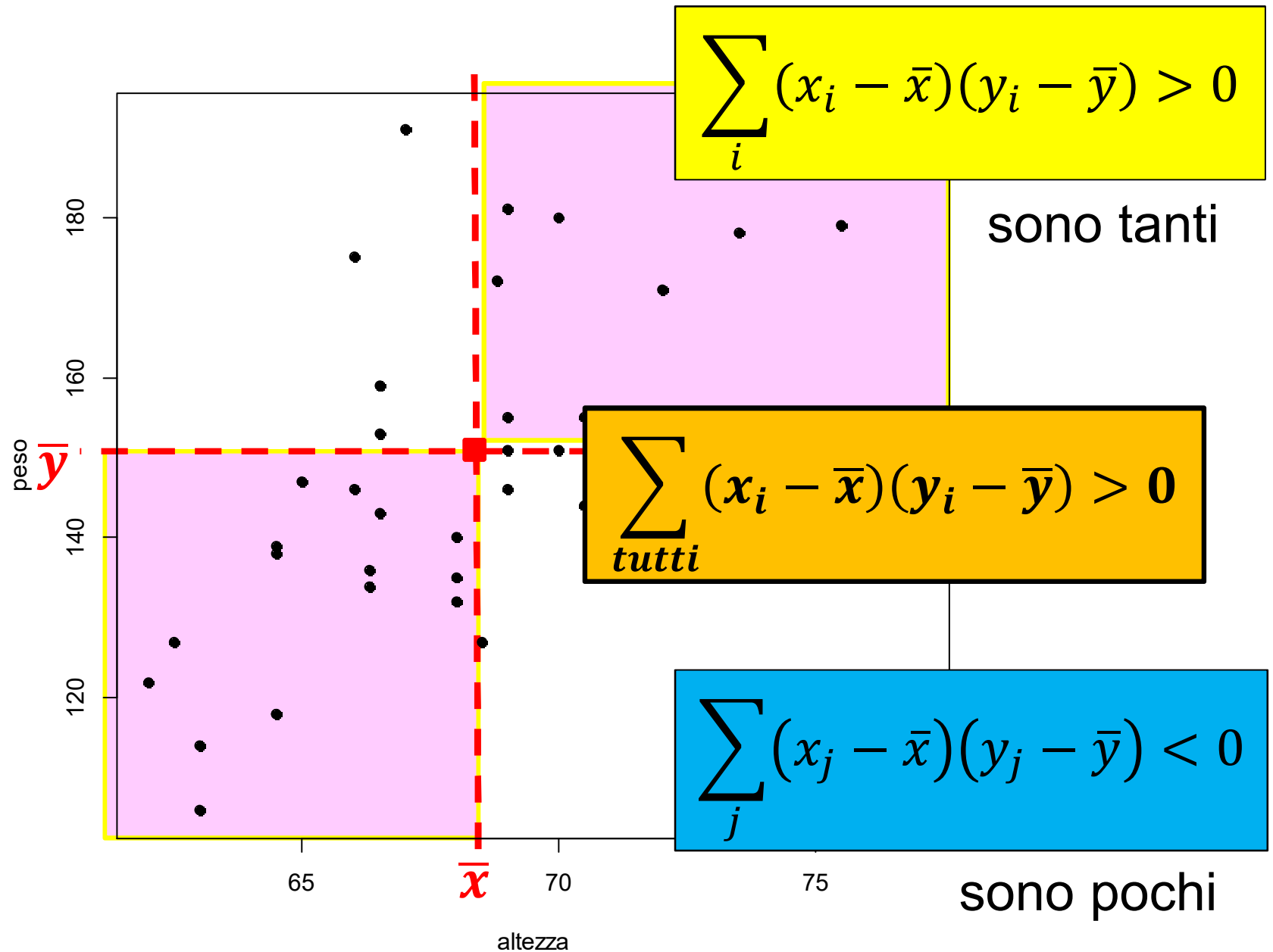
La covarianza



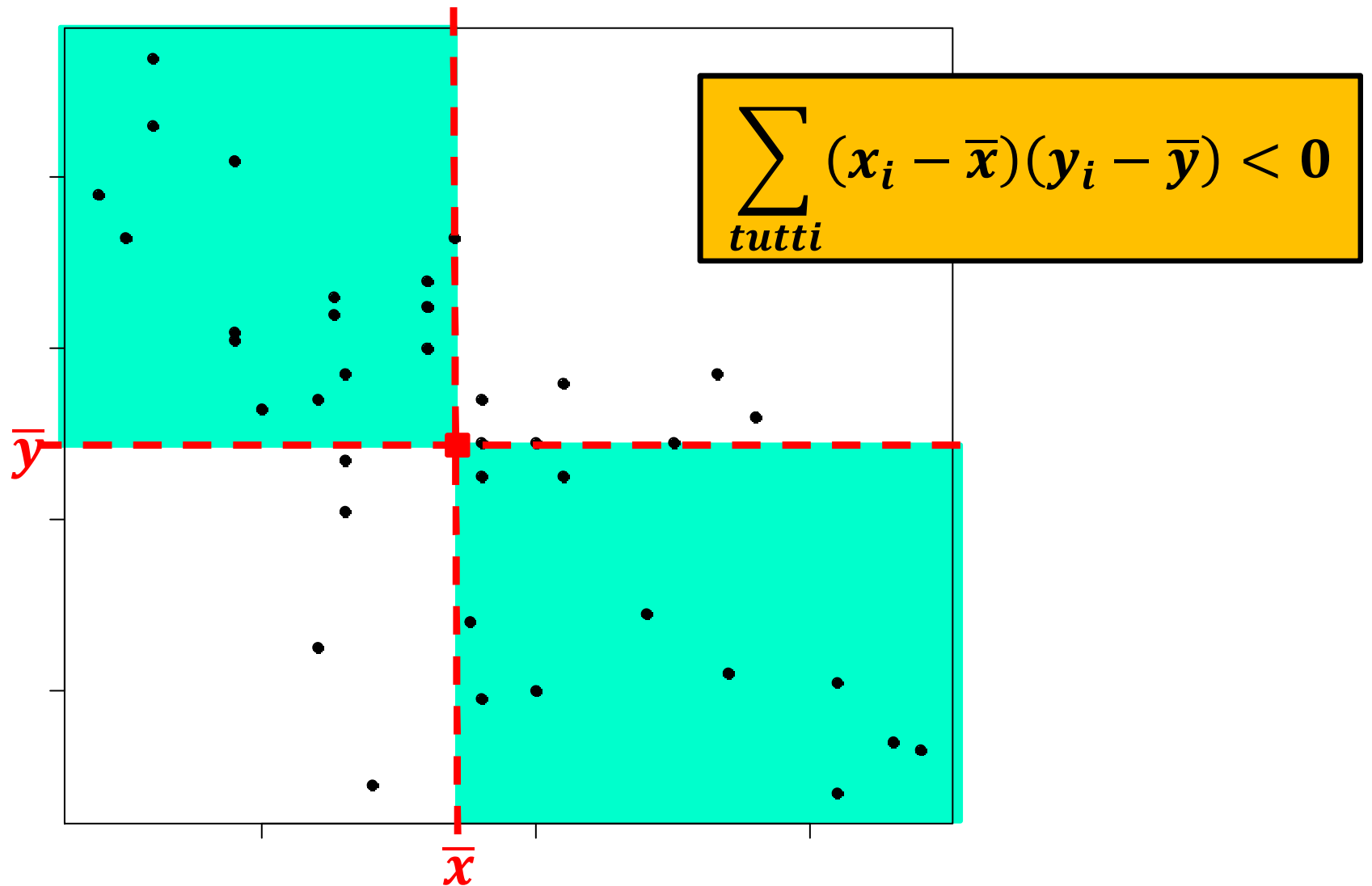
La covarianza



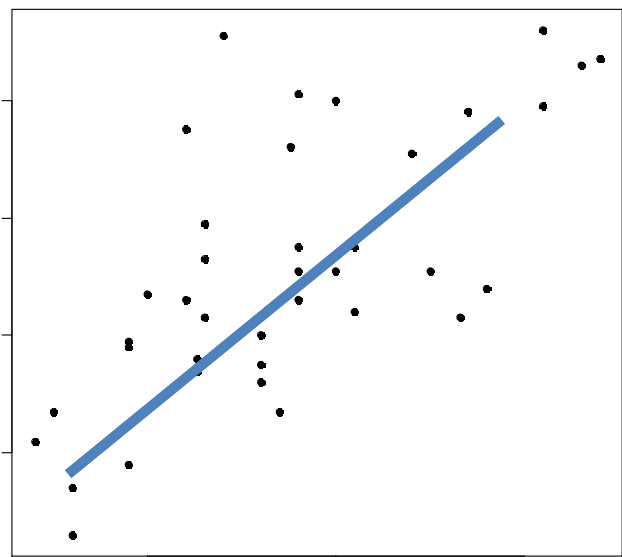
La covarianza



La covarianza



La covarianza

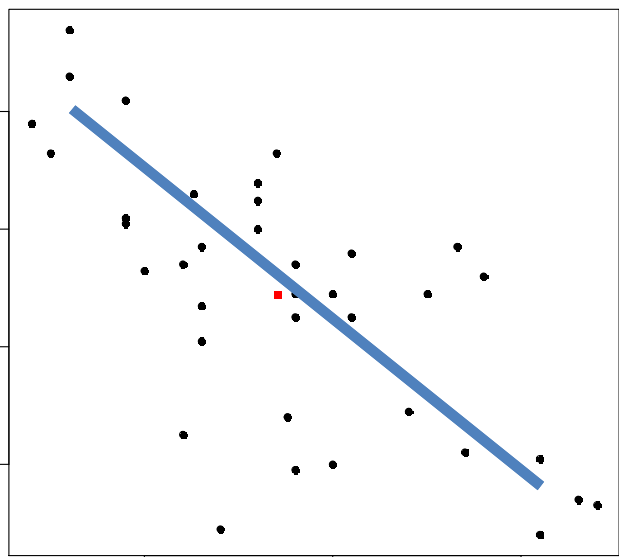


$$\sigma_{xy} = \text{cov}(x, y) = \frac{1}{n} \sum_i (x_i - \bar{x})(y_i - \bar{y})$$

$$\text{cov}(x, y) > 0$$



quando x tende a crescere y fa lo stesso



$$\text{cov}(x, y) < 0$$

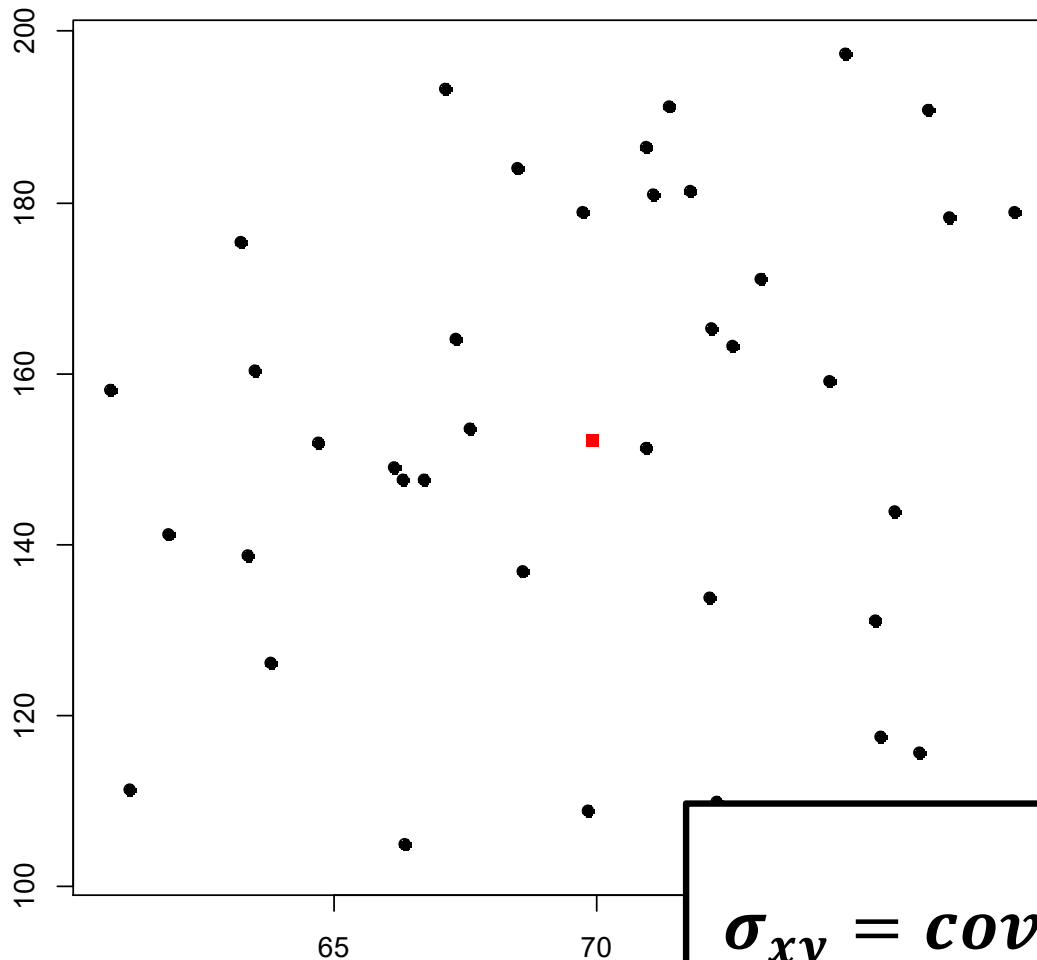


quando x tende a crescere y tende a decrescere

$$\sigma_{xy} = \text{cov}(x, y) = \left[\frac{1}{n} \sum_i x_i y_i \right] - \bar{x} \times \bar{y}$$

La covarianza

$$\sigma_{xy} = \text{cov}(x, y) = \frac{1}{n} \sum_i (x_i - \bar{x})(y_i - \bar{y})$$



$$\text{cov}(x, y) = 0$$

$$\left[\frac{1}{n} \sum_i x_i y_i \right] = \bar{x} \times \bar{y}$$

$$\sigma_{xy} = \text{cov}(x, y) = \left[\frac{1}{n} \sum_i x_i y_i \right] - \bar{x} \times \bar{y}$$

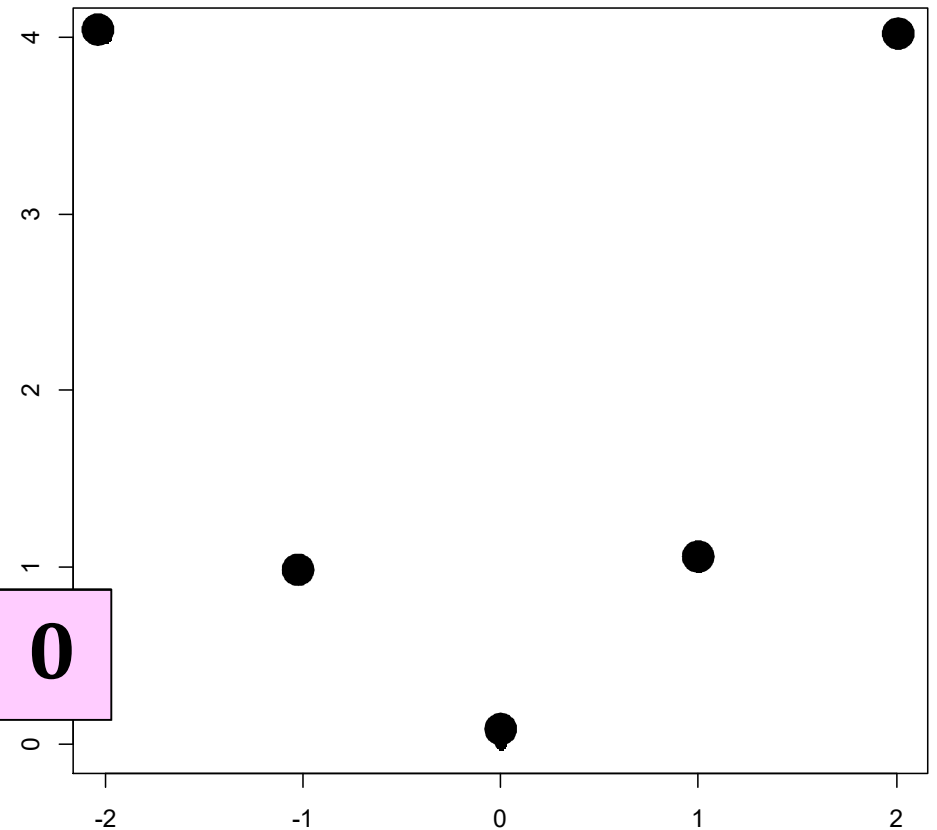
(Contro)Esempio

X	-2	-1	0	1	2
Y	4	1	0	1	4
XY	-8	-1	0	1	8

$$n = 5$$

$$\begin{array}{l} \bar{x} = 0 \\ \bar{y} = 2 \end{array} \Rightarrow \bar{x} \times \bar{y} = 0$$

$$\frac{1}{5} \sum_i x_i y_i = 0! \Rightarrow \sigma_{xy} = 0$$



(Contro)Esempio

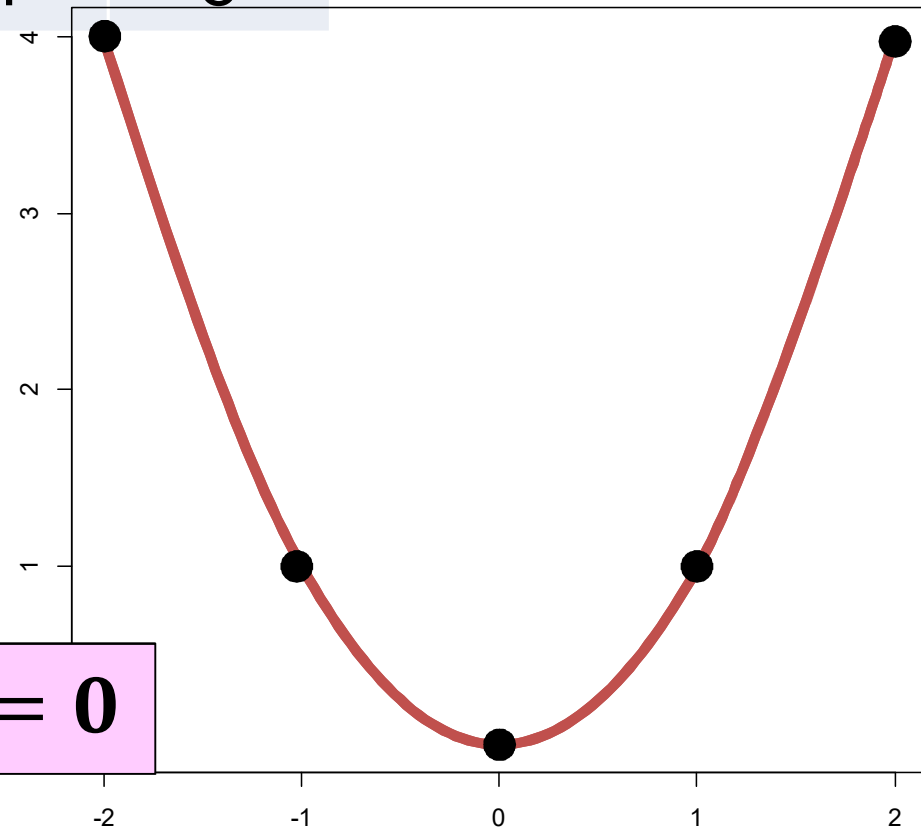
X	-2	-1	0	1	2
Y	4	1	0	1	4
XY	-8	-1	0	1	8

$n =$
 $\bar{x} =$
 $\bar{y} =$
 $\frac{1}{5} \sum_{i=1}^n$

assenza di correlazione lineare...
lineare...
ma chiara correlazione quadratica!

$\bar{y} = 0$

$\sigma_{xy} = 0$



La correlazione lineare

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

$$\sigma_{xy} = \text{cov}(x, y) = \frac{1}{n} \sum_i (x_i - \bar{x})(y_i - \bar{y})$$

$$-1 \leq \rho_{xy} \leq 1$$

$$\rho_{xy} > 0 \quad \text{cov}(x, y) > 0$$

$$\rho_{xy} < 0 \quad \text{cov}(x, y) < 0$$

$$\rho_{xy} = 0 \quad \text{cov}(x, y) = 0$$

coeff. di
correlazione
lineare

$$\sigma_{xy} = \text{cov}(x, y) = \left[\frac{1}{n} \sum_i x_i y_i \right] - \bar{x} \times \bar{y}$$

$$-\sigma_x \sigma_y \leq \sigma_{xy} \leq \sigma_x \sigma_y$$

La correlazione lineare

A spanne...		
Valore di $ \rho $	Interpretazione fine	Meno fine
0.90 - 1.00	Molto alta	forte (talvolta 0.5->0.6)
0.70 - 0.90	Alta	
0.50 - 0.70	Moderata	
0.30 - 0.50	Bassa	moderata
0.00 - 0.30	Poca o non correlazione	> 0.10 debole

Hinkle, Wiersma, & Jurs (2003). Applied Statistics for the Behavioral Sciences (5th ed.).

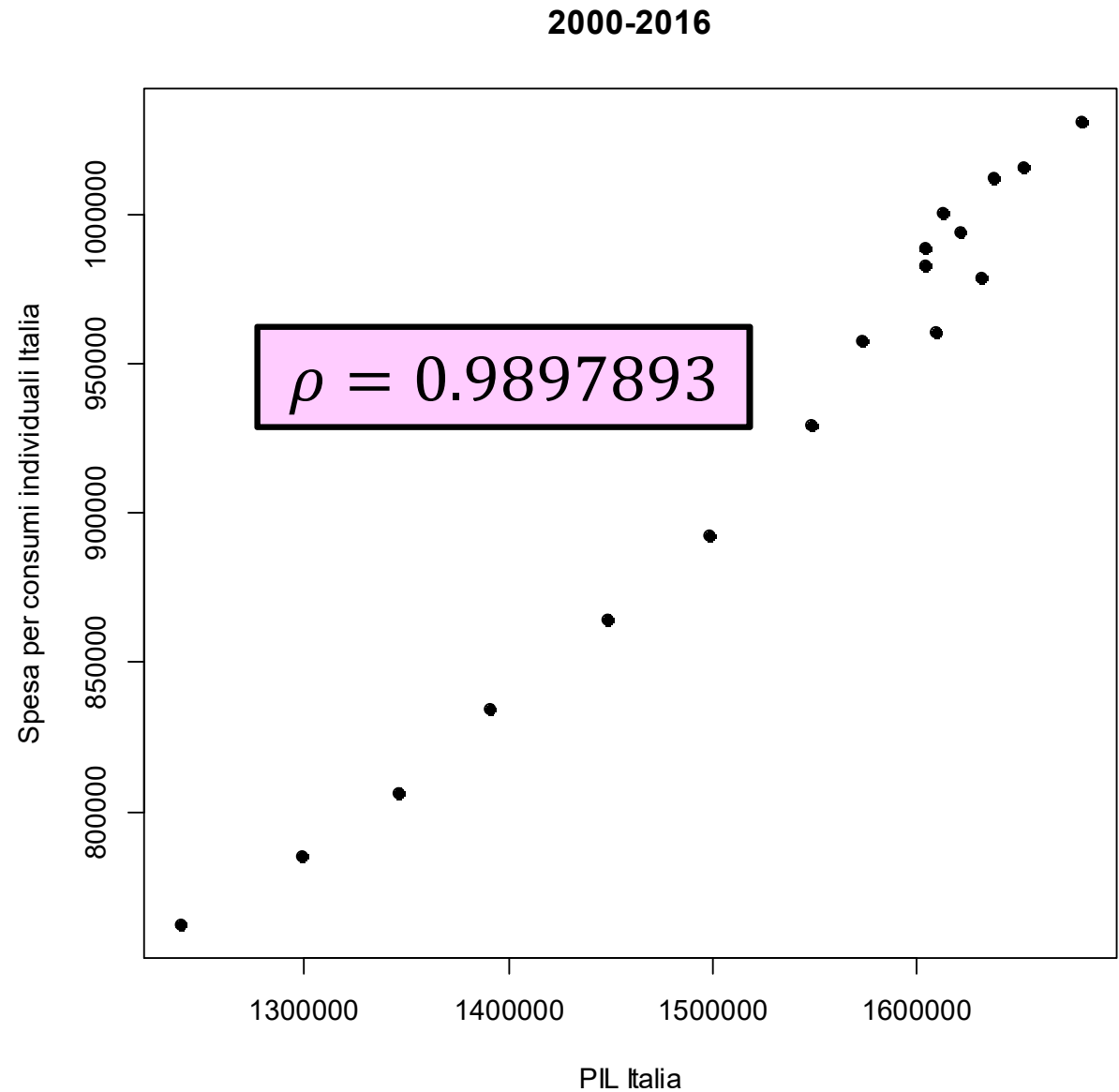
Esempio 1

Istat (2017)

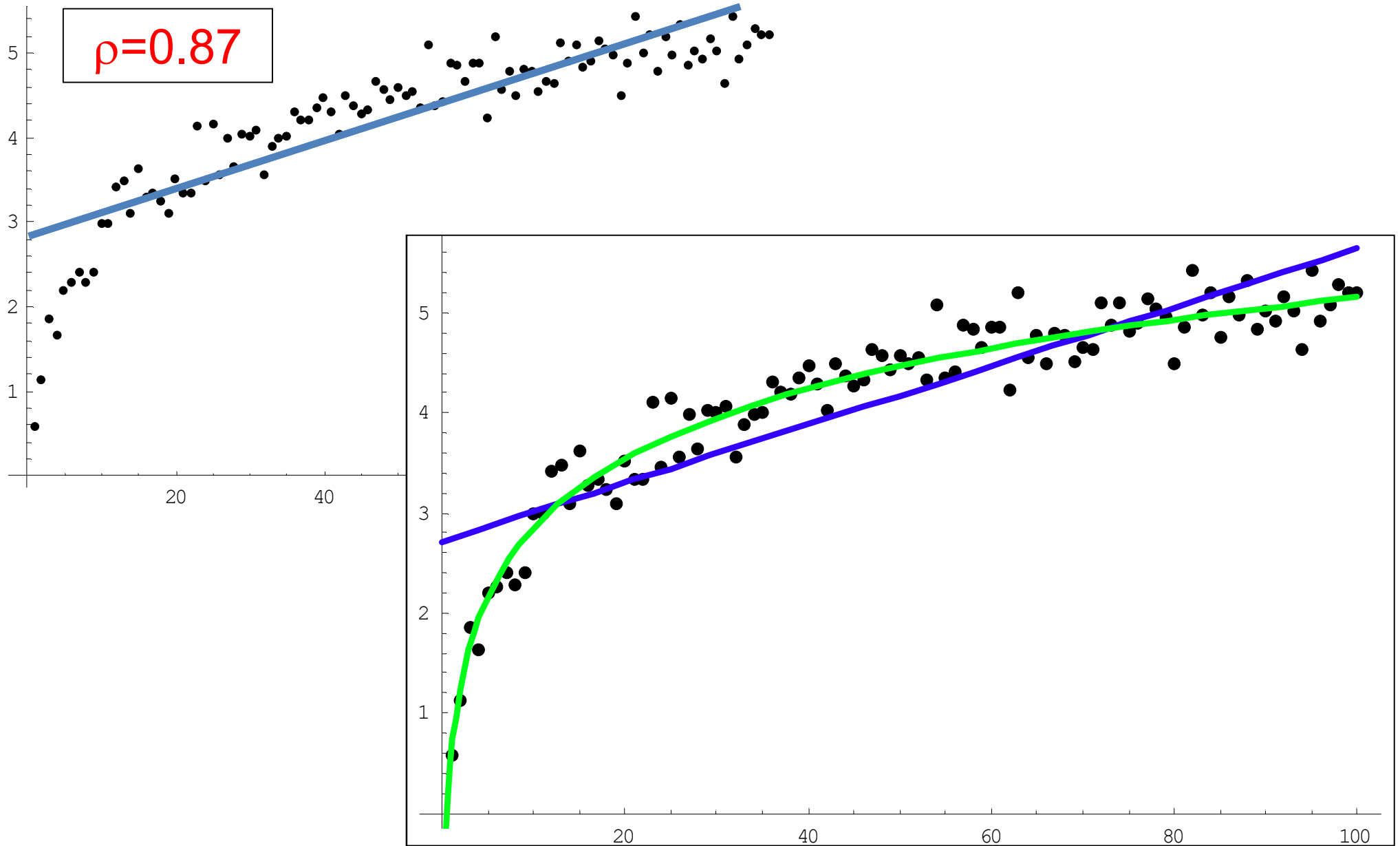
Conti e aggregati
economici
nazionali annuali

Spese per consumi
finali delle famiglie

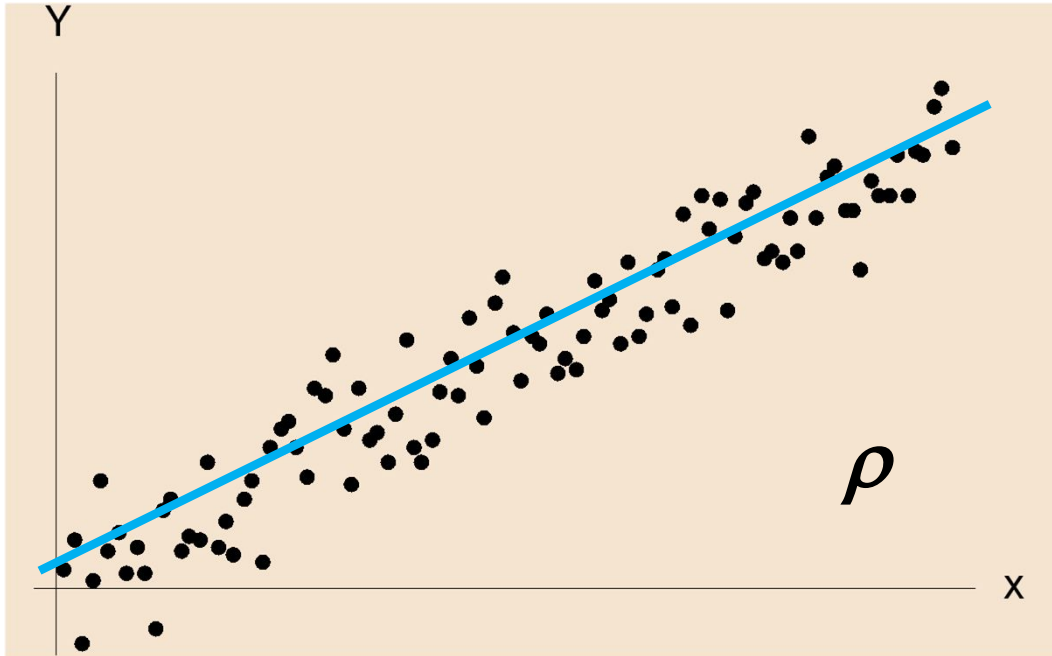
<http://dati.istat.it/>



L'indice non basta



Regressione lineare



A. Valutazione preliminare se una retta possa essere una buona approssimazione

B. Stima dei parametri della retta

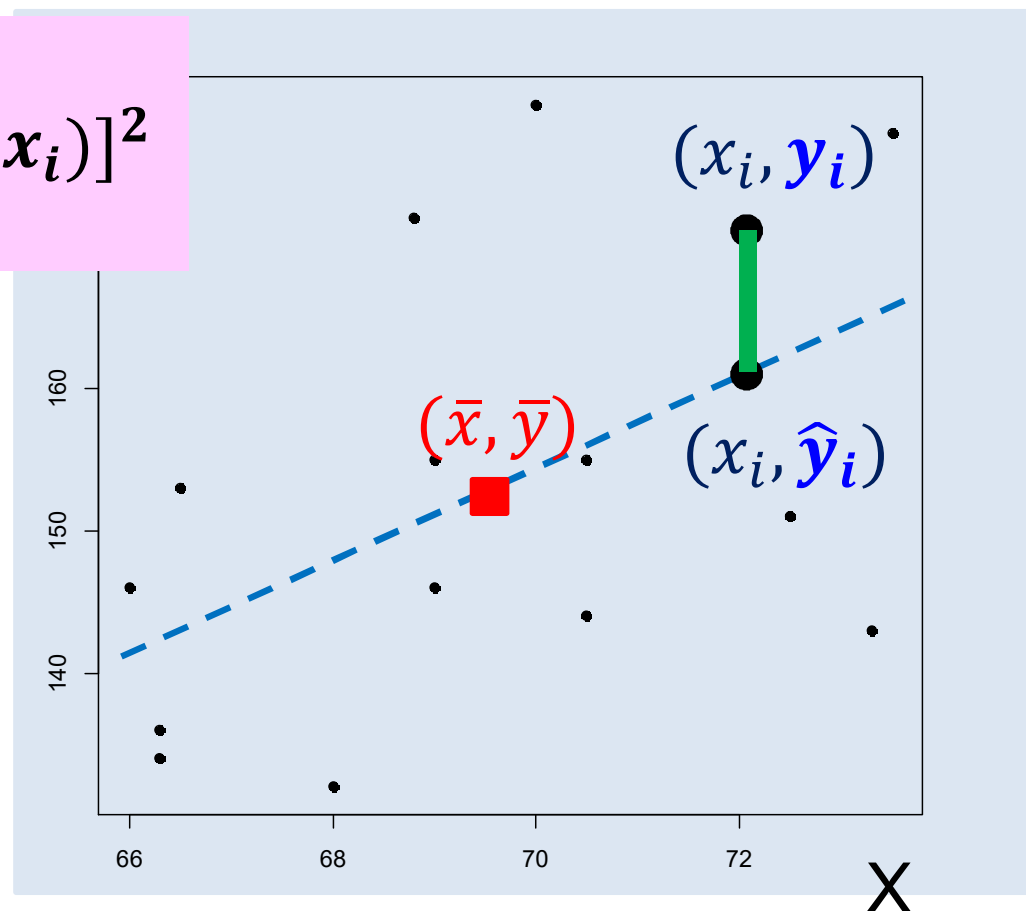
La retta di regressione

Di tutte le rette $y = a + bx$ quella che passa più vicino a tutti i punti, nel senso dei *minimi quadrati*, è quella con coeff. \hat{a} e \hat{b} che rendono minima la quantità:

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n [y_i - (a + bx_i)]^2$$

$$\hat{b} = \frac{\sigma_{xy}}{\sigma_x^2} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

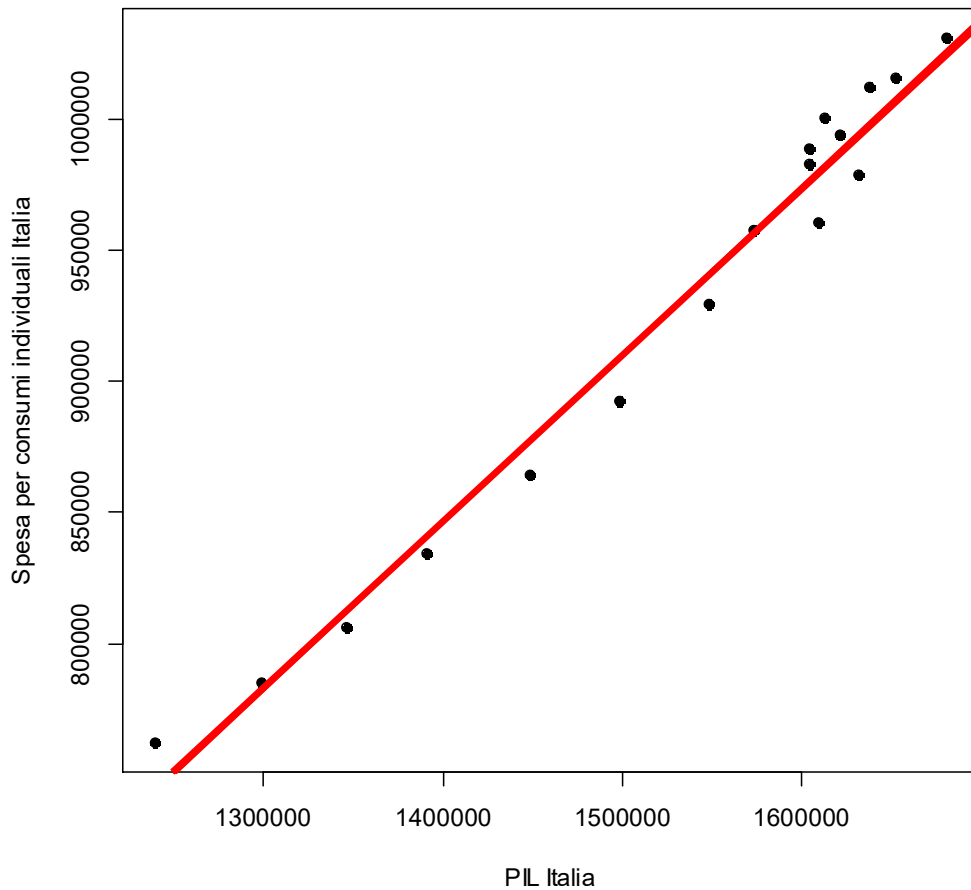
$$\hat{a} = \bar{y} - \hat{b}\bar{x}$$



La retta di regressione

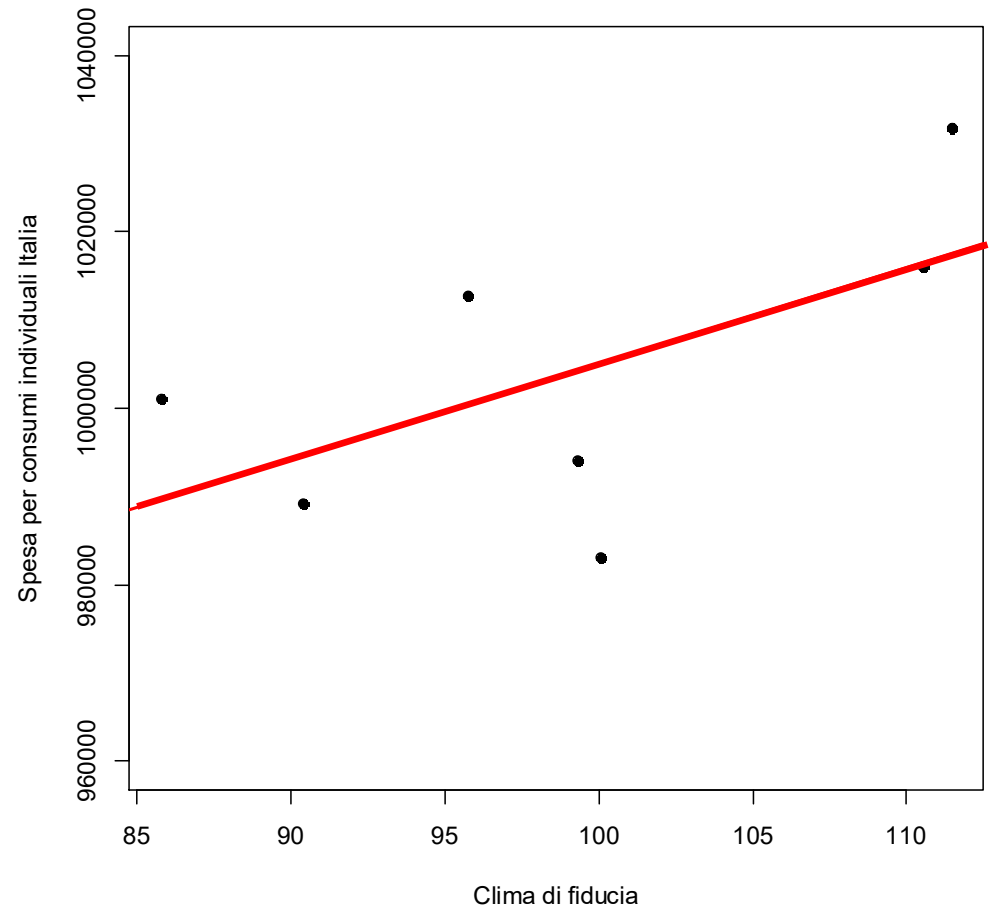
$$\hat{a} = -52723.36$$
$$\hat{b} = 0.642176$$

2000-2016



$$\hat{a} = 896307.76$$
$$\hat{b} = 1086.972$$

2010-2016



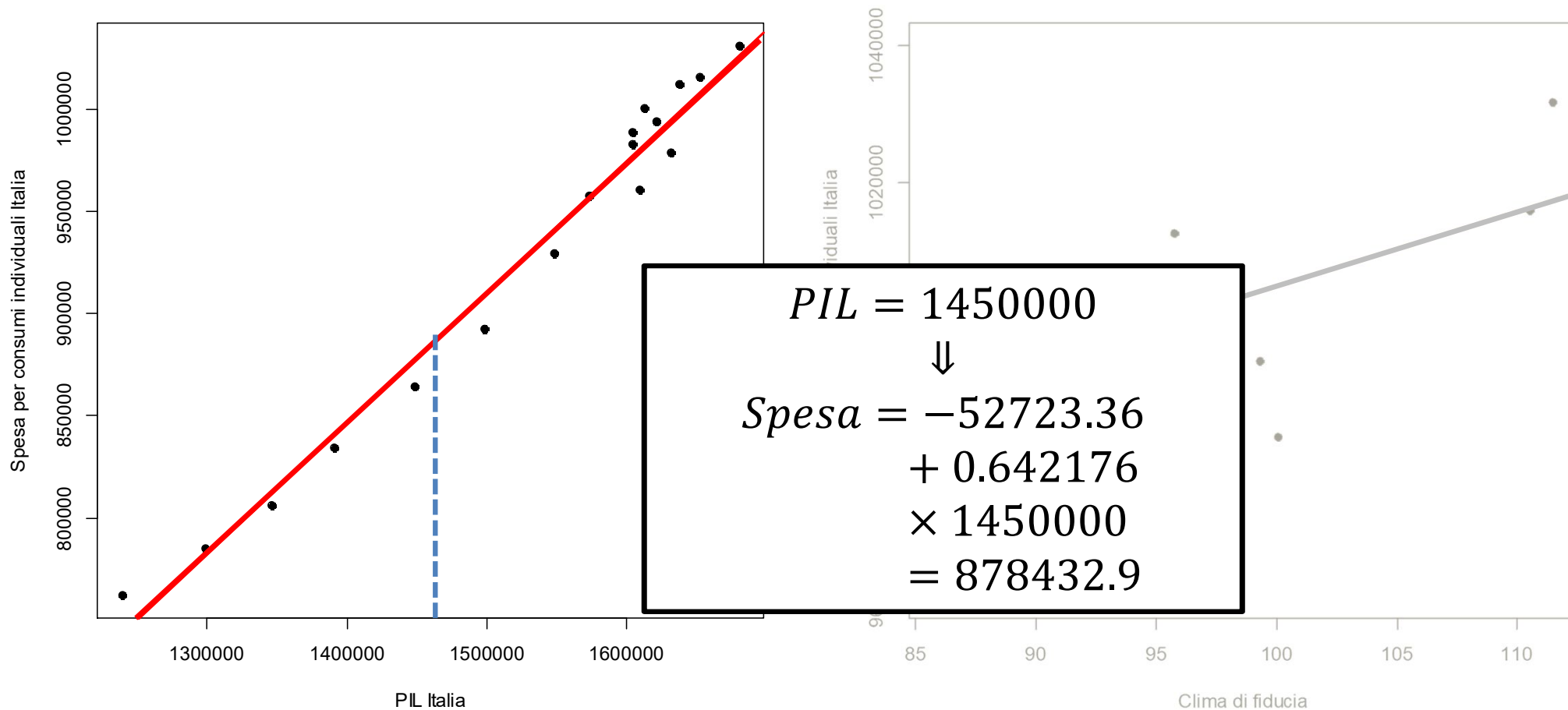
La retta di regressione

$$\hat{a} = -52723.36$$
$$\hat{b} = 0.642176$$

2000-2016

$$\hat{a} = 896307.76$$
$$\hat{b} = 1086.972$$

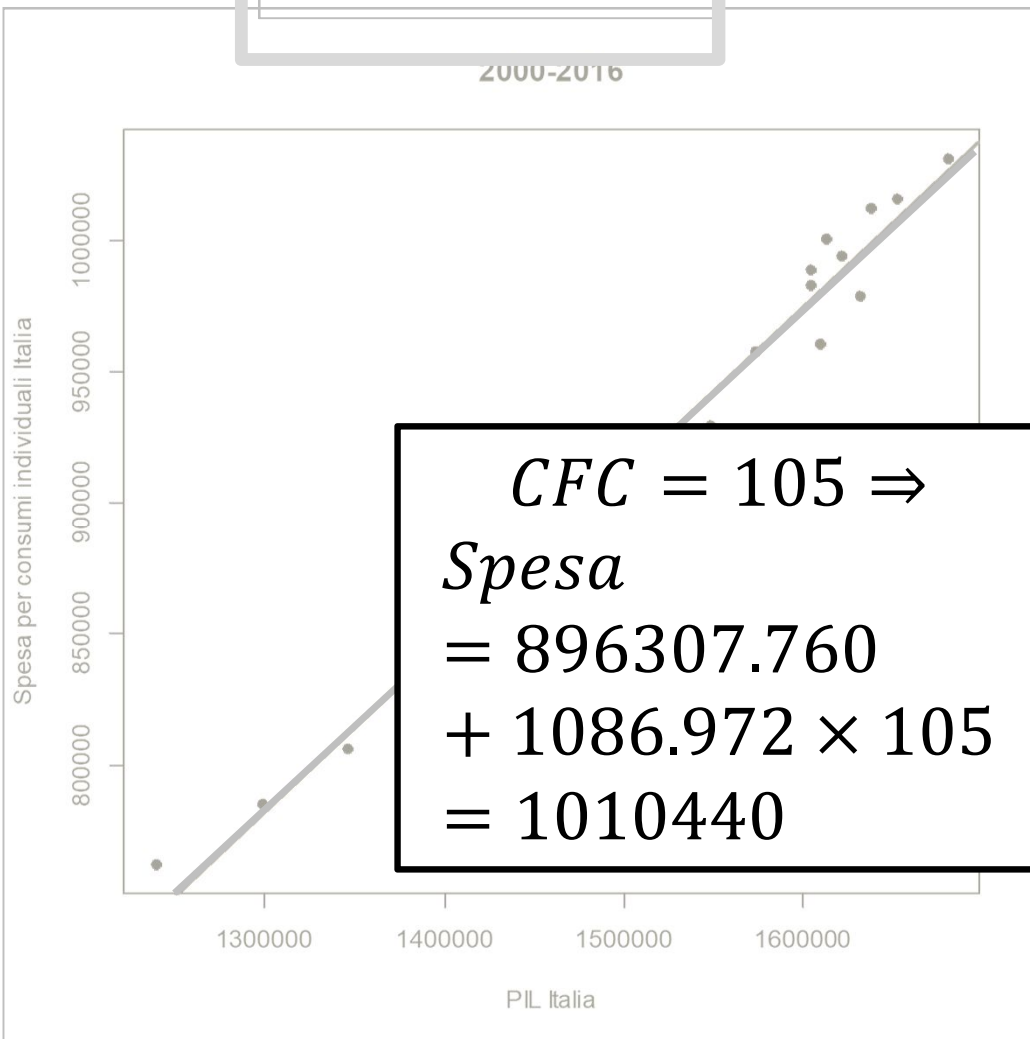
2010-2016



La retta di regressione

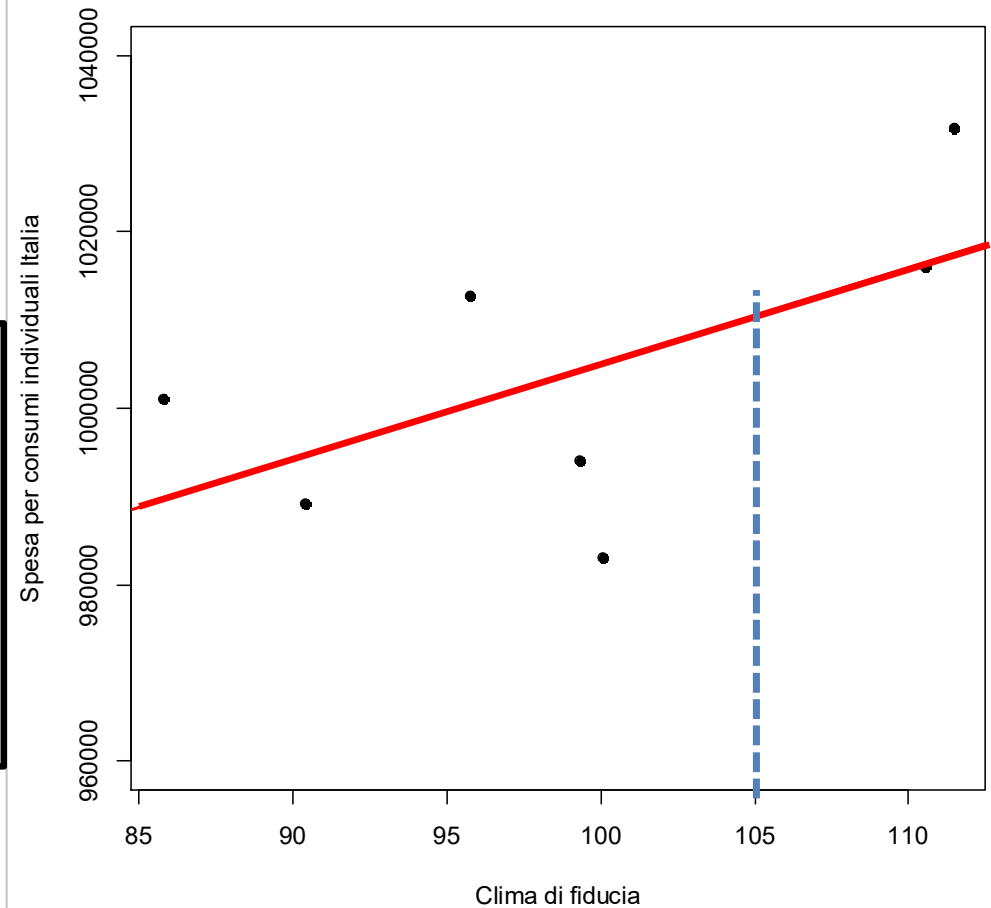
$$\hat{a} = -52723.36$$
$$\hat{b} = 0.642176$$

2000-2016



$$\hat{a} = 896307.76$$
$$\hat{b} = 1086.972$$

2010-2016



La retta di regressione

