

Rappresentare graficamente la correlazione e la predizione

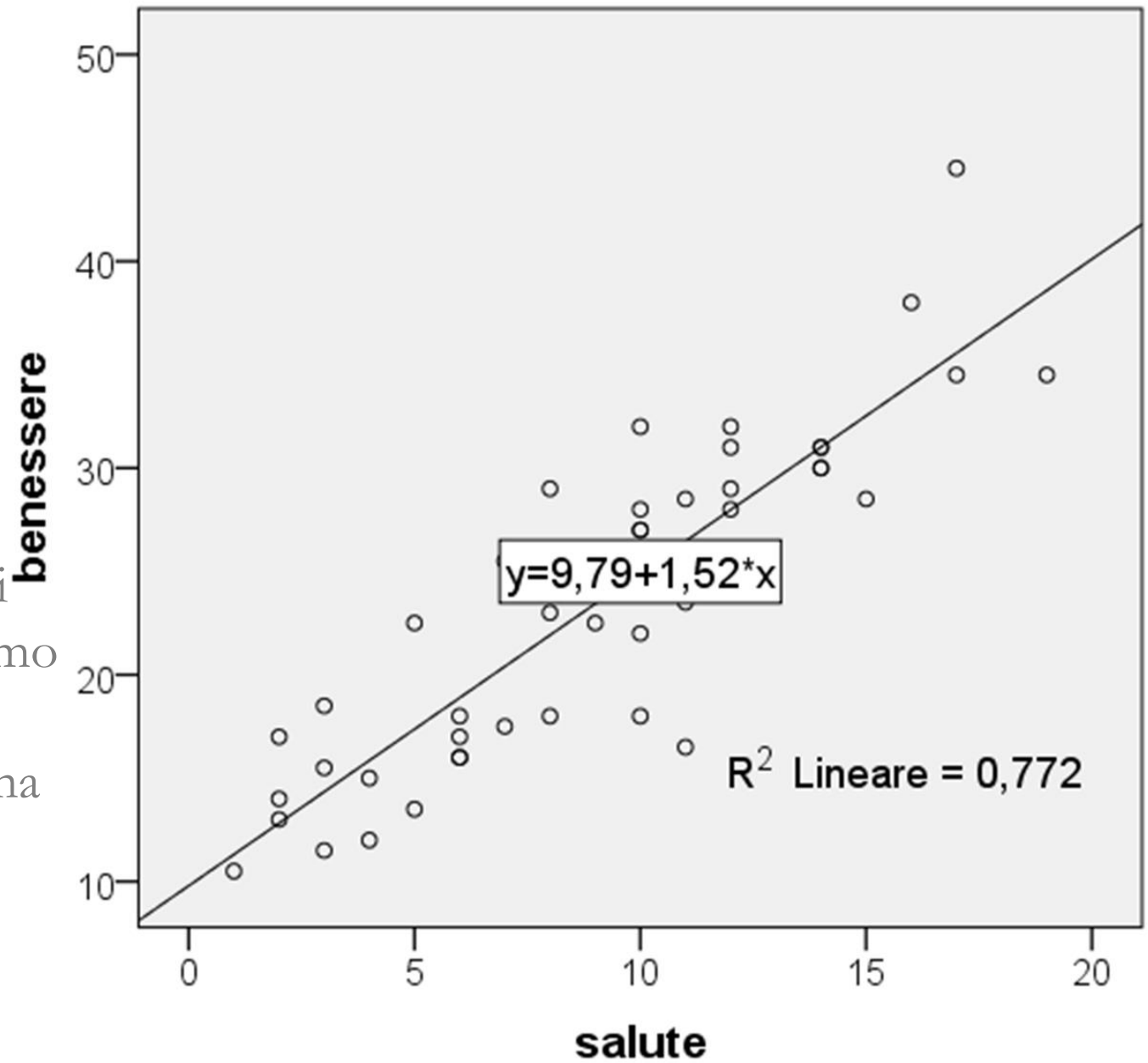
Giovanni Battista Flebus

Lezioni di Psicometria

Consideriamo due misurazioni, per esempio
benessere e salute.

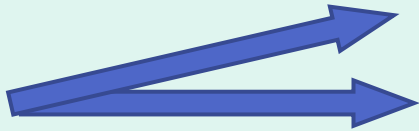
Le due variabili sono simili. Chi ha punteggi sopra
la media in una variabile, tende ad avere punteggi
sopra la media anche nell'altra variabile, e viceversa

- Il grafico con coordinate cartesiane rappresenta le coppie di osservazioni, e solamente dall'insieme dei dati noi possiamo dedurre l'esistenza di una relazione

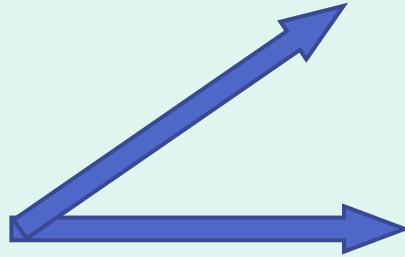


Il **vettore grafico** può essere usato per rappresentare graficamente questa relazione: un vettore è un **segmento orientato** che ha una **forza** e una **direzione**

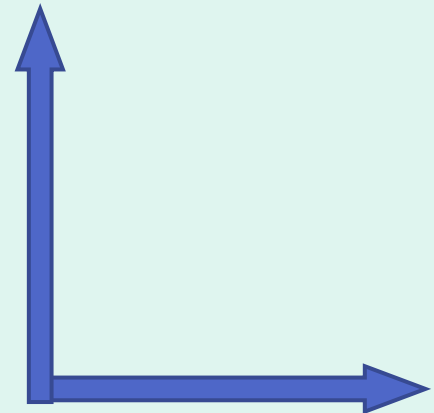
Tuttavia, per rappresentare una correlazione come vettore, faremo ricorso solo alla **direzione**



A



B

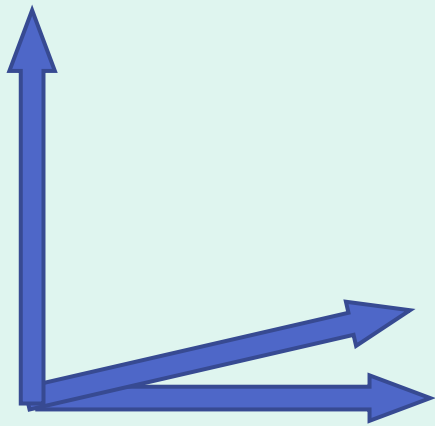


C

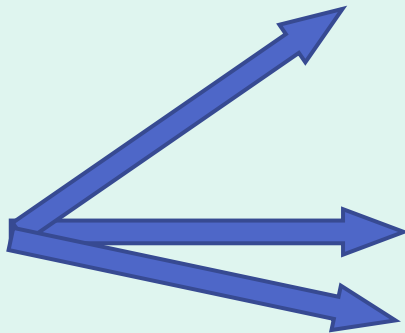
Il caso A rappresenta due variabili altamente correlate, il caso C due variabili indipendenti e il caso B una situazione intermedia

Consideriamo tre variabili, due molto correlate e
una indipendente.

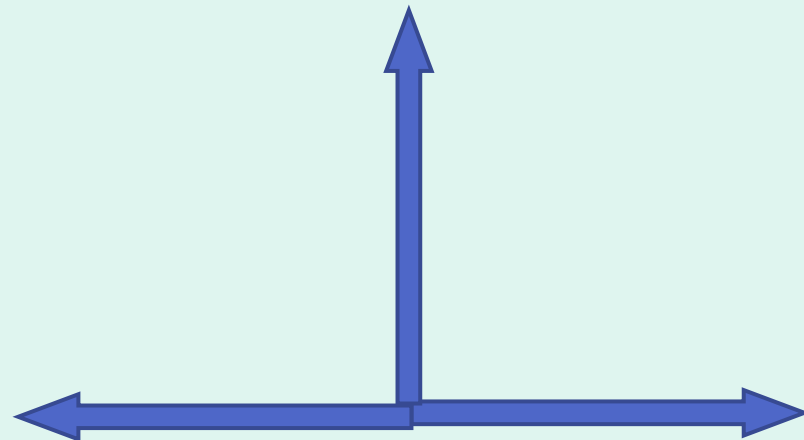
Come si disporrebbero i vettori?



A



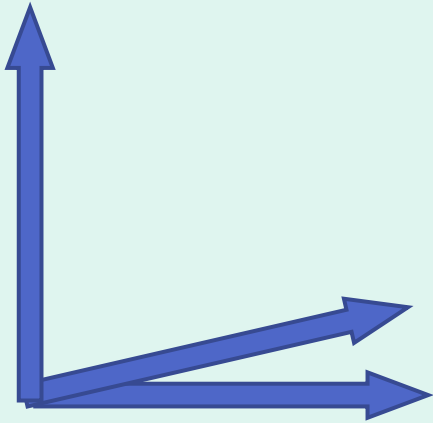
B



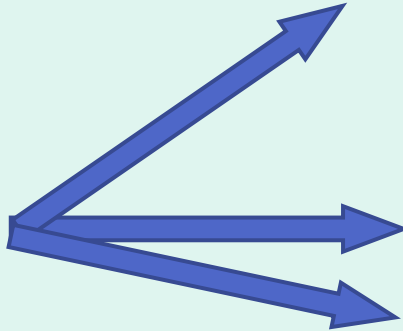
C

Come si rappresenterebbe una scala di stress scolastico, uno di stress generale e una di livello socio-economico?

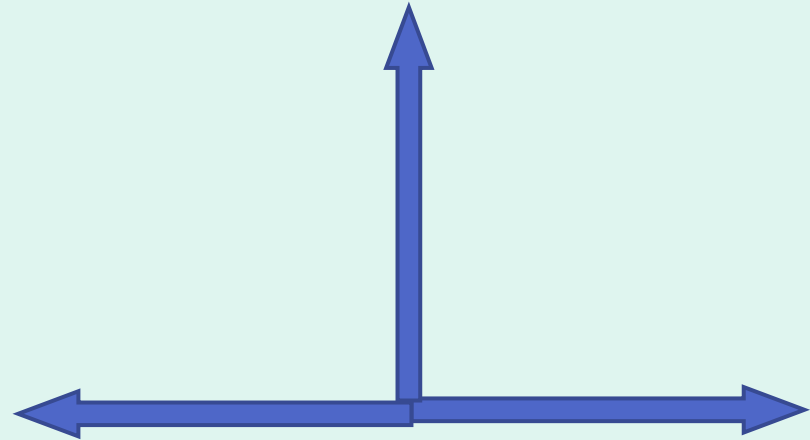
A, B oppure C?
Perché?



A

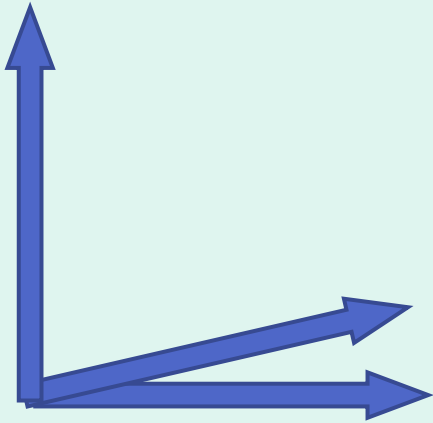


B

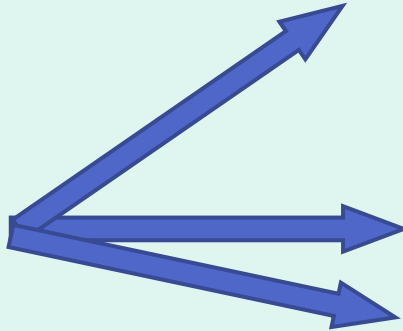


C

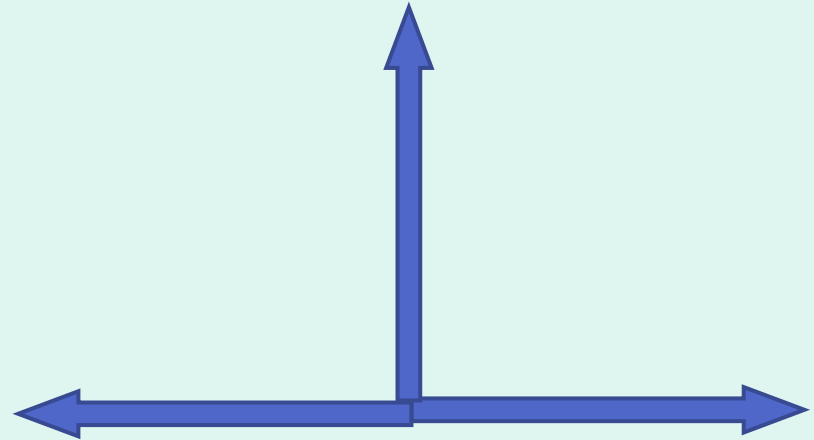
Due variabili sono simili fra di loro e molto diverse dalla terza



A

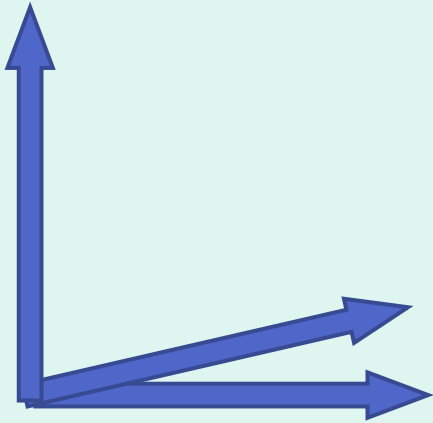


B

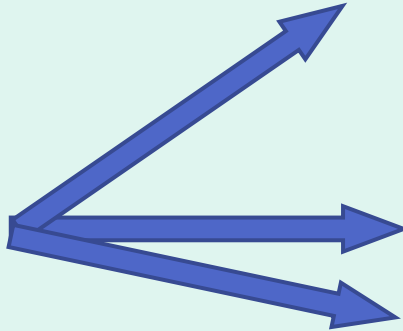


C

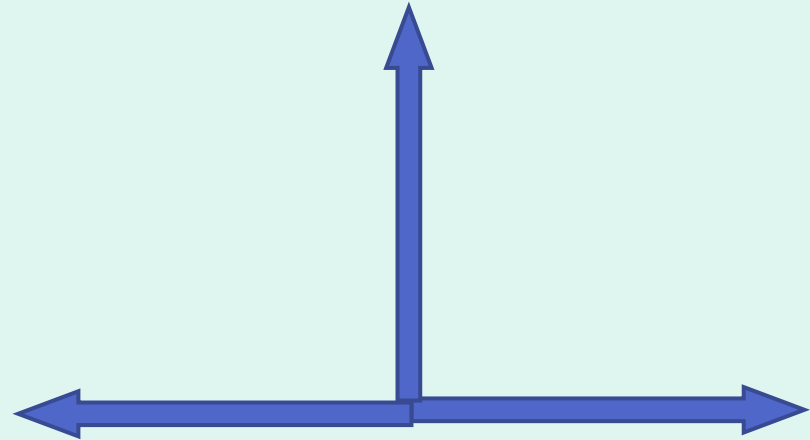
Due variabili sono simili fra di loro e non molto diverse dalla terza.



A



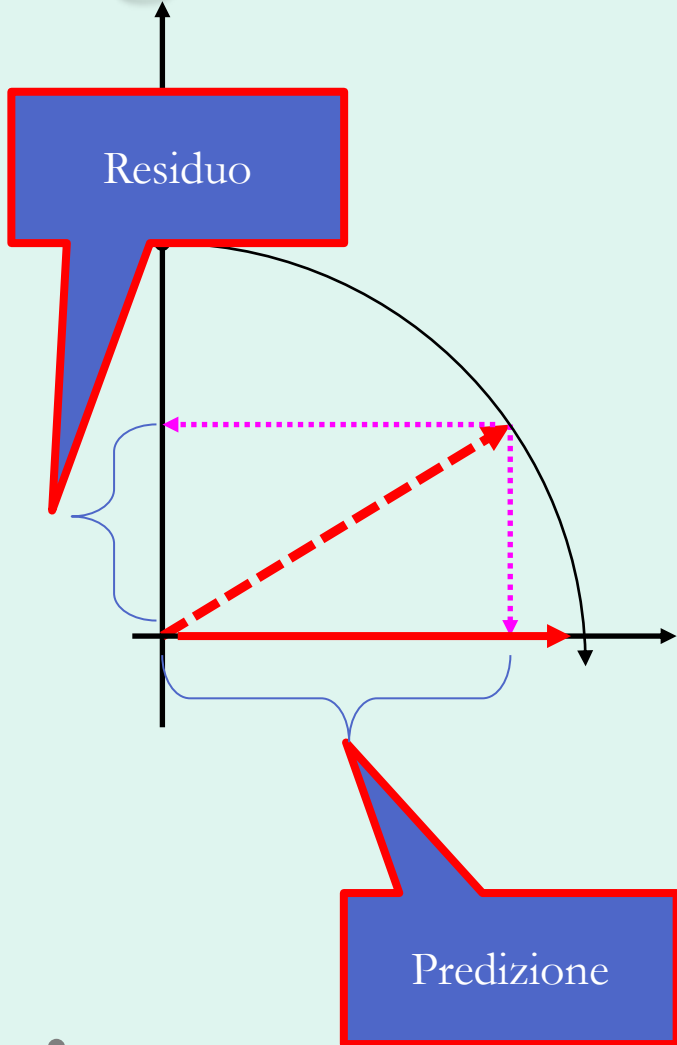
B



C

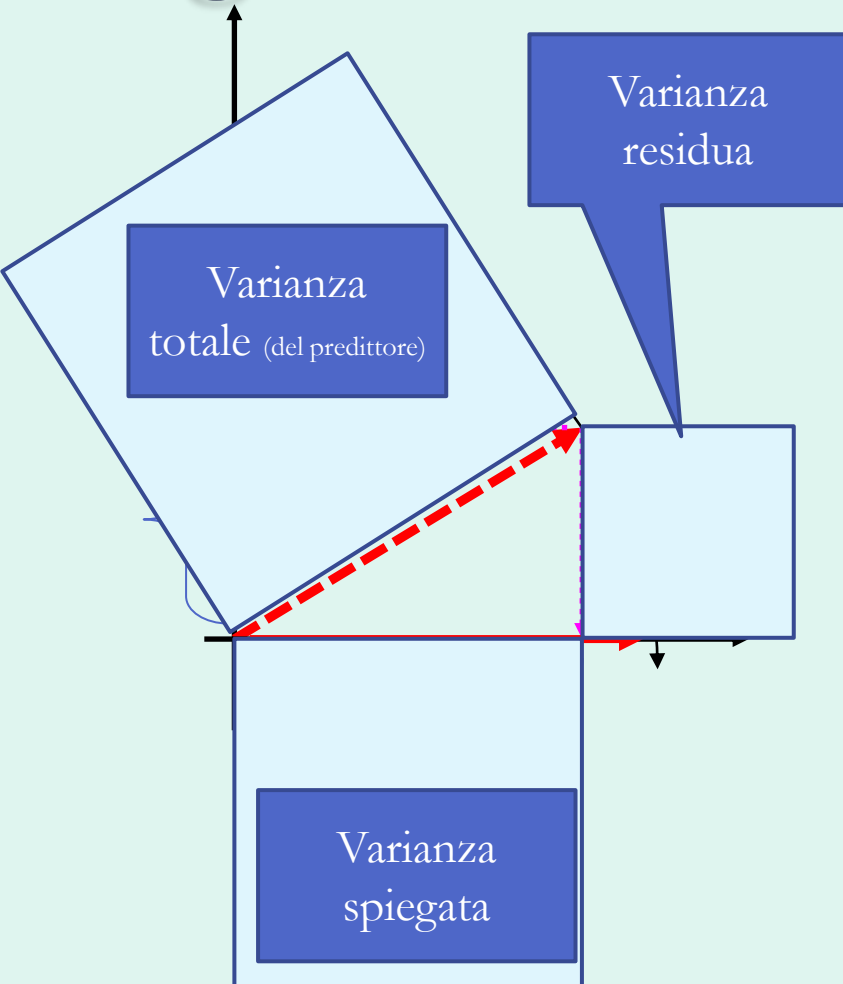
Due variabili sono uguali ma opposte e indipendenti con una terza

Rappresentazione grafica della **predizione**



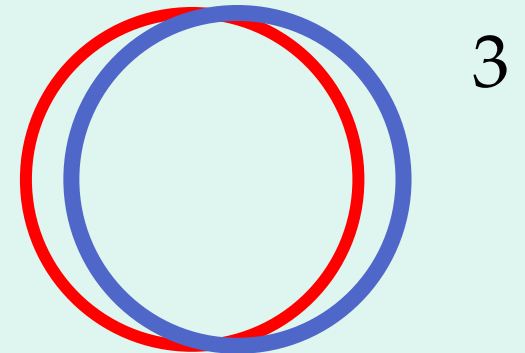
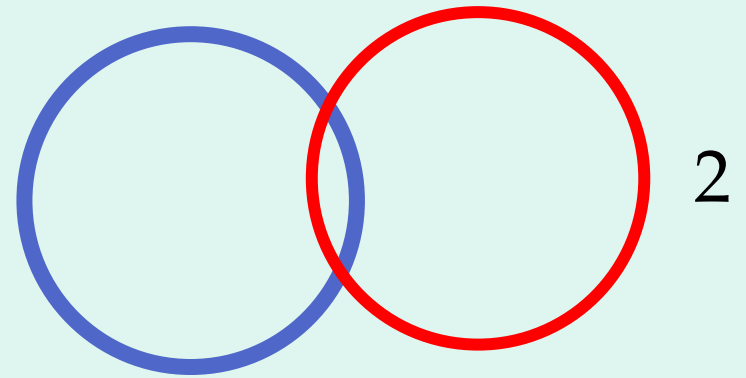
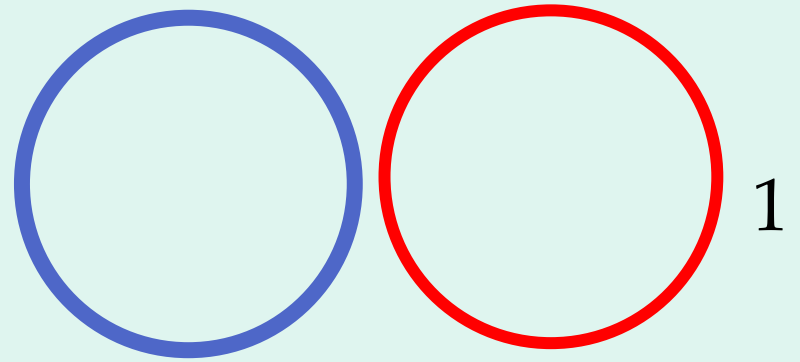
- Il predittore viene trasformato con l'equazione di regressione in due variabili: la prima è perfettamente correlata con il predittore, ha correlazione uguale a uno ma una deviazione standard inferiore a quella del predetto
- La seconda variabile è costituita dai residui
- È indipendente dalla variabile rossa (predittore)

Rappresentazione grafica della **predizione**



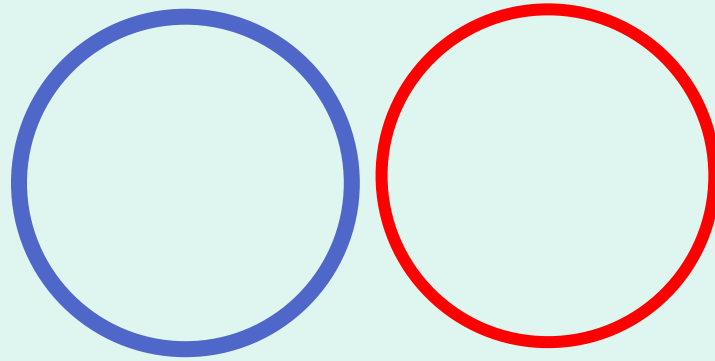
- La varianza totale è sempre scomponibile in due: varianza spiegata dalla regressione e varianza residua o dell'errore

Le correlazioni come diagrammi di Venn

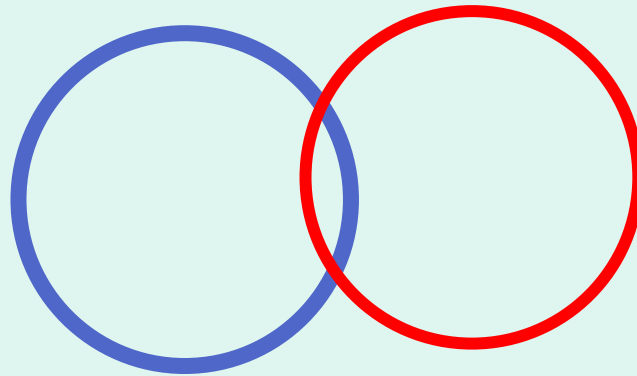
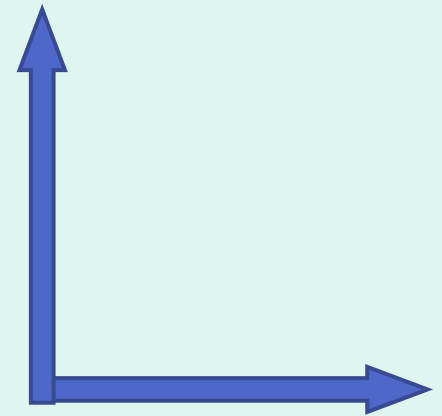


- A) Indipendenti?
- B) Molto correlate?
- C) poco correlate?

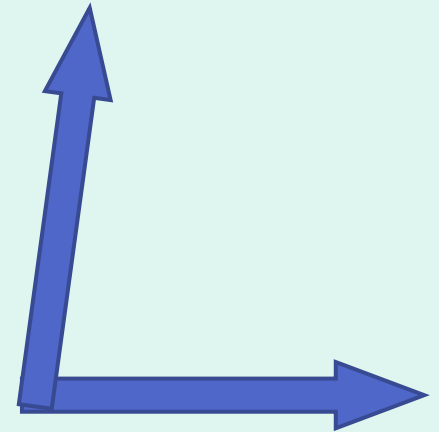
Vettori e
diagramm
o di Venn



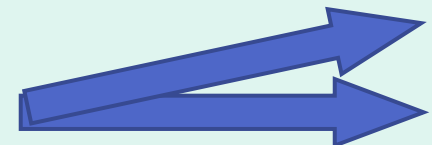
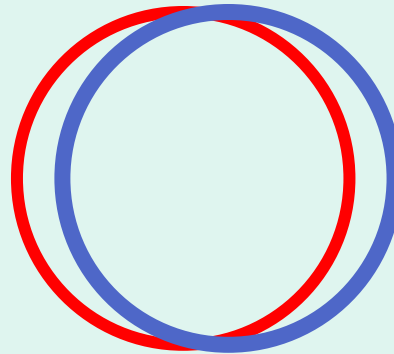
1



2



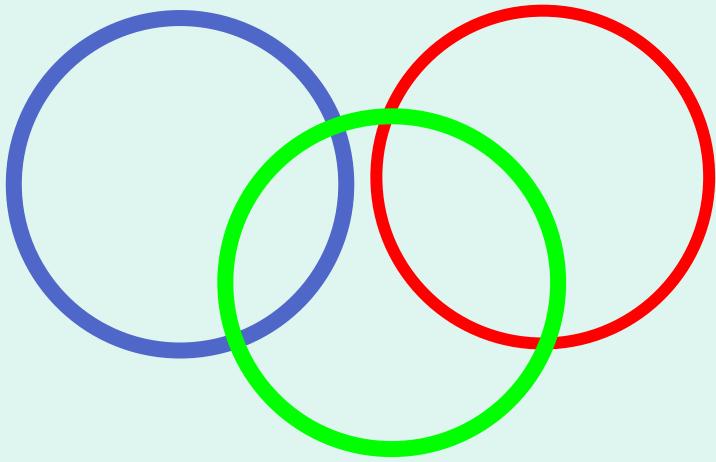
3



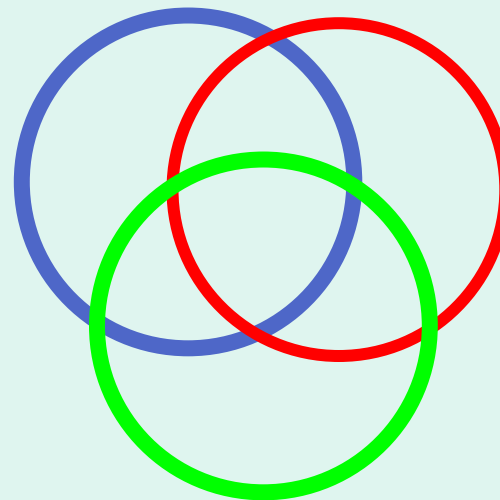
Che succede aggiungendo una terza variabile?

La terza variabile **VERDE** può essere **poco, molto o per niente** correlata con le altre due

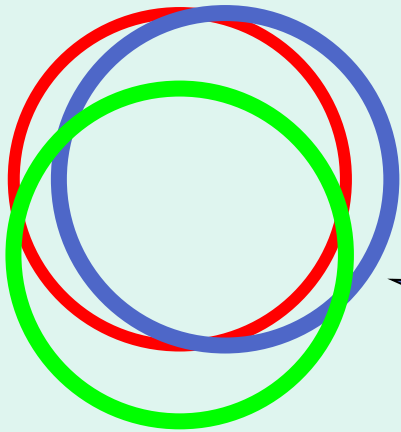
Il suo apporto nella regressione multipla dipende dalla varianza che può spiegare in modo autonomo



La nuova variabile –verde- spiega nuova
varianza, mentre la rossa non spiega nulla



La variabile verde spiega una quota di varianza,
ma una parte è già spiegata dalla variabile rossa



La nuova variabile verde ha pochissima
varianza da spiegare, in aggiunta a quella
spiegata dalla rossa.

