



# Analisi di dati ecologici

**Analisi con molte  
variabili  
(multivariata)**

**DATI**

Tipo

Distribuzione

Forma

**MATRICE DEI  
DATI BRUTI**

degli Stati

delle Storie

Trasformazione

**MATRICE DEI  
DATI  
TRASFORMATI**

Coefficienti di somiglianza

**MATRICE DI  
SOMIGLIANZA**

moda Q

moda R

**ORDINAMENTO**

**CLASSIFICAZIONE**

**Binari** (presenza-assenza)

**Non ordinati** (gilde)

**Semi-quantitativi**  
(ranghi)

**Ordinati**

**Quantitativi**

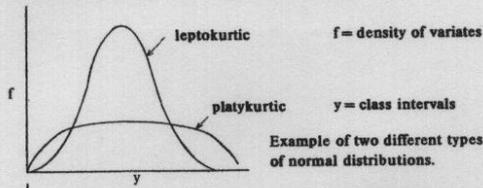
**Metrici** (continui)  
(misure)

**Meristici** (discontinui)  
(conte)

**Multipli**

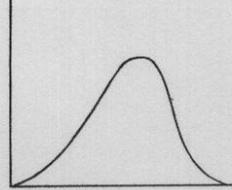
**TIPO  
DI  
DATI**

# DISTRIBUZIONE DI FREQUENZA DEI DATI

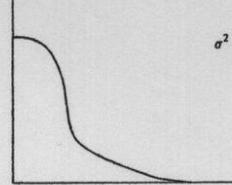


Example of two different types of normal distributions.

approaches the normal distribution as the number of degrees of freedom or df (i.e., sample size) increase. A df of 30 gives a distribution that is approximately normal.



$\sigma^2 = \text{mean}$ , often skewed to the right



skewed to the right

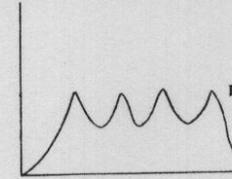
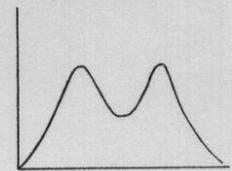
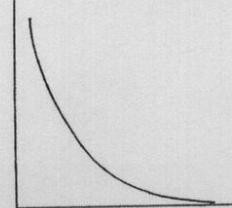


Figure 2-2. Examples of statistical distributions.

Normale

T - distr.

Poisson

Neg-Binomiale

Bimodale

Plurimodale

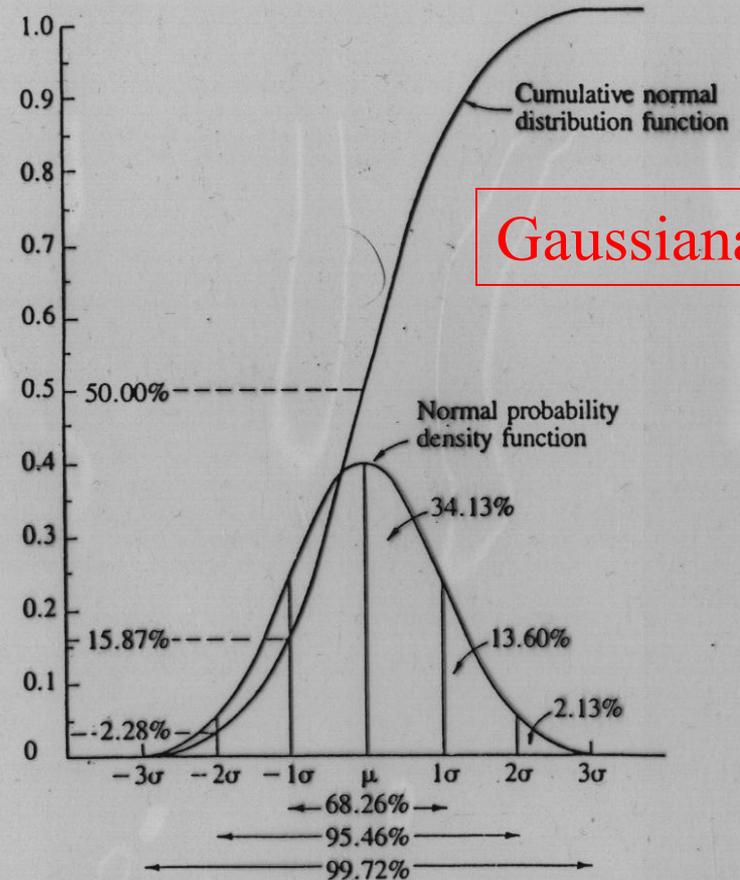


Figure 2-1. Normal distribution (bell-shaped curve) and cumulative normal distribution.

$\mu = \text{mean}$

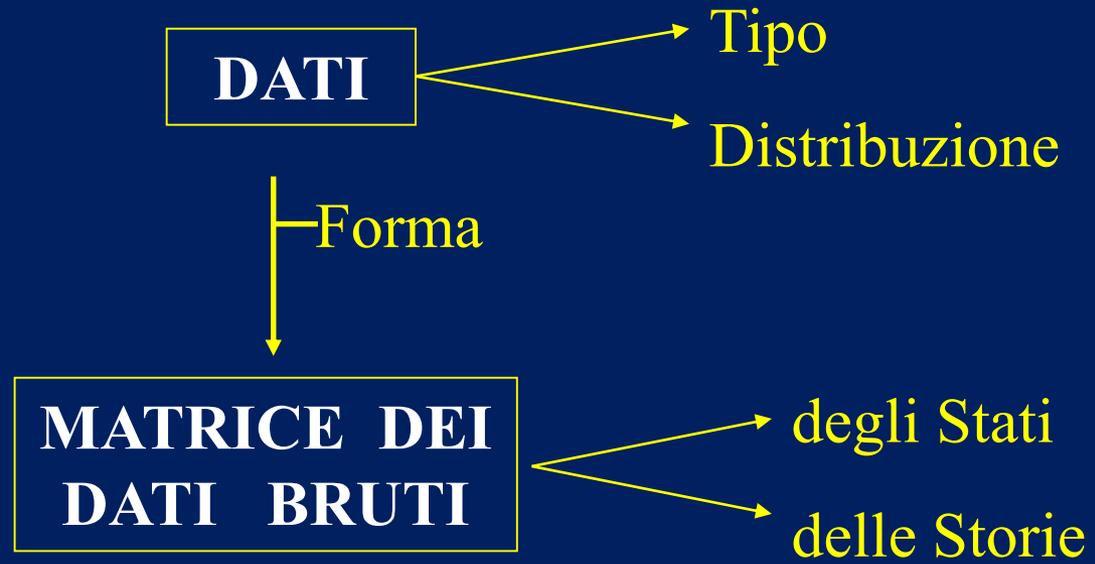
$\sigma = \text{standard deviation}$

Source: modified from Sokal and Rohlf (1981)

Statistica parametrica



# FORMA DEI DATI



# FORMA DEI DATI

## Matrice dei dati bruti

Oggetti (osservazioni)

	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<b>a</b>	$x_{1a}$	$x_{2a}$	$x_{3a}$	$x_{4a}$	$x_{5a}$
<b>b</b>	$x_{1b}$	$x_{2b}$	$x_{3b}$		
<b>c</b>	$x_{1c}$	$x_{2c}$			
<b>d</b>	$x_{1d}$				

Descrittori  
(variabili)

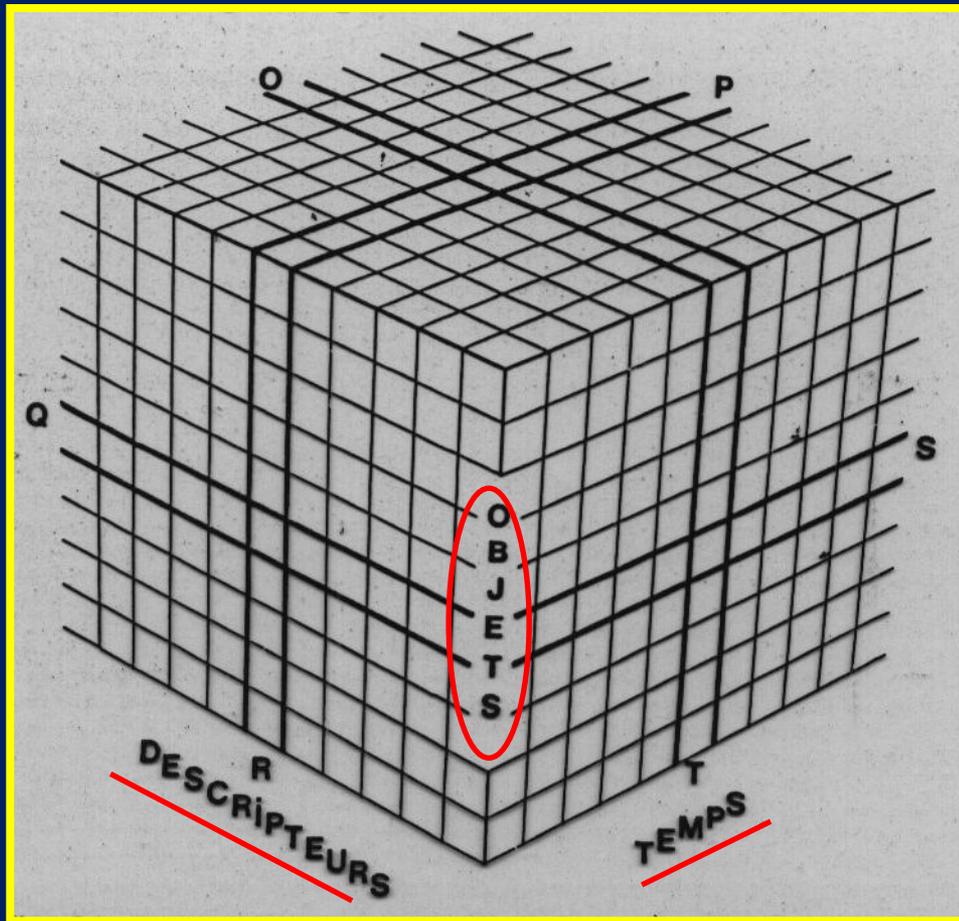
# Matrice dei dati bruti

## Stazioni (osservazioni)

MACROFAUNA GOLFO DI POLICASTRO *****	RI-1	RI-2	RI-3	MA-7	MA-6	MA-5	MA-4	MA-3	MA-2
Gibbula adansoni (PAYRAUDEAU, 1826)	1	0	0	0	0	0	0	0	0
Gibbula ardens (VON SALIS, 1793)	0	0	0	1	0	0	0	0	0
Gibbula philberti (RECLUZ, 1843)	4	0	7	0	0	0	0	0	0
Gibbula rarilineata (MICHAUD, 1829)	0	0	1	0	0	0	0	0	0
Tricolia pullus (L., 1758)	0	0	0	1	0	0	0	0	0
Tricolia speciosa (VON MUHLFELDT, 1824)	8	0	0	0	0	0	0	0	0
Ceratia proxima (ALDER, 1847)	0	0	0	0	0	0	0	0	0
Hyala vitrea (MONTAGU, 1803)	0	0	0	0	0	0	0	0	0
Apicularia guerinii (RECLUZ, 1843)	5	0	0	1	0	0	0	0	0
Rissoa hyalina FREMINVILLE, 1814	32	0	8	0	0	0	0	0	0
Rissoa labiosa (MONTAGU, 1803)	4	0	0	0	0	0	0	0	0
Turboella lineolata (MICHAUD, 1832)	19	0	0	0	0	0	0	0	0
Turritella communis RISSO, 1826	0	0	0	0	0	0	0	4	0
Bittium latreilii (PAYRAUDEAU, 1826)	352	0	0	8	0	0	0	0	0
Aclis minor (BROWN, 1827)	0	0	0	0	0	0	0	0	0
Melanella intermedia (CANTRAINE, 1835)	0	0	0	0	0	0	0	0	1
Strombiformis glaber (DA COSTA, 1778)	0	0	0	0	0	0	0	0	0
Aporrhais serresianus (MICHAUD, 1828)	0	0	0	0	0	0	0	0	0
Lunatia macilenta (PHILIPPI, 1844)	0	0	0	0	0	0	0	4	0
Neverita josephina RISSO, 1826	0	0	0	0	0	0	1	0	0
Naticarius hebraeus (MARTYN, 1784)	1	0	0	0	0	0	0	0	0
Naticarius punctatus (CHEMNITZ, 1789)	0	0	0	0	0	0	0	0	0
Bolinus brandaris (L., 1758)	0	0	0	0	0	0	0	0	0
Phyllonotus trunculus (L., 1758)	42	0	0	0	0	0	0	0	0
Typhinellus sowerbyi (BRODERIP, 1833)	1	0	0	0	0	0	0	0	0
Columbellopsis minor (SCACCHI, 1836)	0	0	0	0	0	0	0	0	0
Columbella rustica (L., 1758)	0	0	0	3	0	0	0	0	0
Cyclope donovani RISSO, 1826	0	0	0	0	0	0	0	0	0
Hinia incrassata (STROM, 1768)	0	0	0	1	0	0	0	0	0
Hinia limata (DESHAYES, 1844)	0	0	0	0	0	0	0	1	0
Hinia pygmaea (LAMARCK, 1822)	0	0	0	0	0	0	2	0	0
Nassarius mutabilis (L., 1758)	104	6	2	0	6	0	0	0	0
Nassarius cuvierii (PAYRAUDEAU, 1826)	37	0	0	0	0	0	0	0	0
Bela brachystoma (PHILIPPI, 1844)	0	0	0	0	0	0	0	0	0
Bela nebula (MONTAGU, 1803)	3	0	2	0	2	0	2	0	0
Bela turgida (FORBES, 1844)	0	0	0	0	0	0	0	0	0

Specie  
(variabili)

# Inserimento del tempo (repliche diacroniche)



Matrice “degli stati”

Matrice “delle storie”

## Matrici dei dati bruti

Come visto, i dati ottenuti durante un campionamento sono organizzati in una matrice dei dati “bruti”, ponendo sulle righe i **descrittori/variabili** (es. specie) e sulle colonne gli **oggetti/osservazioni** (es. stazioni).

Quando, oltre alla dimensione *spaziale* (stazioni), deve essere inserita anche la dimensione *temporale* (repliche delle stesse stazioni nel tempo), le matrici vengono dette **“degli stati”**, se il tempo viene inserito all’interno delle colonne, oppure **“delle storie”**, se il tempo viene inserito all’interno delle righe.

Pertanto, replicando il campionamento nel tempo (ad es. nei mesi di marzo, aprile e maggio) si possono organizzare due tipi di matrici a seconda che le repliche siano inserite nelle colonne (per ottenere la “dinamica delle stazioni”) o nelle righe (per ottenere la “dinamica delle specie”).

## A) Matrice degli stati, inserendo il tempo nelle colonne

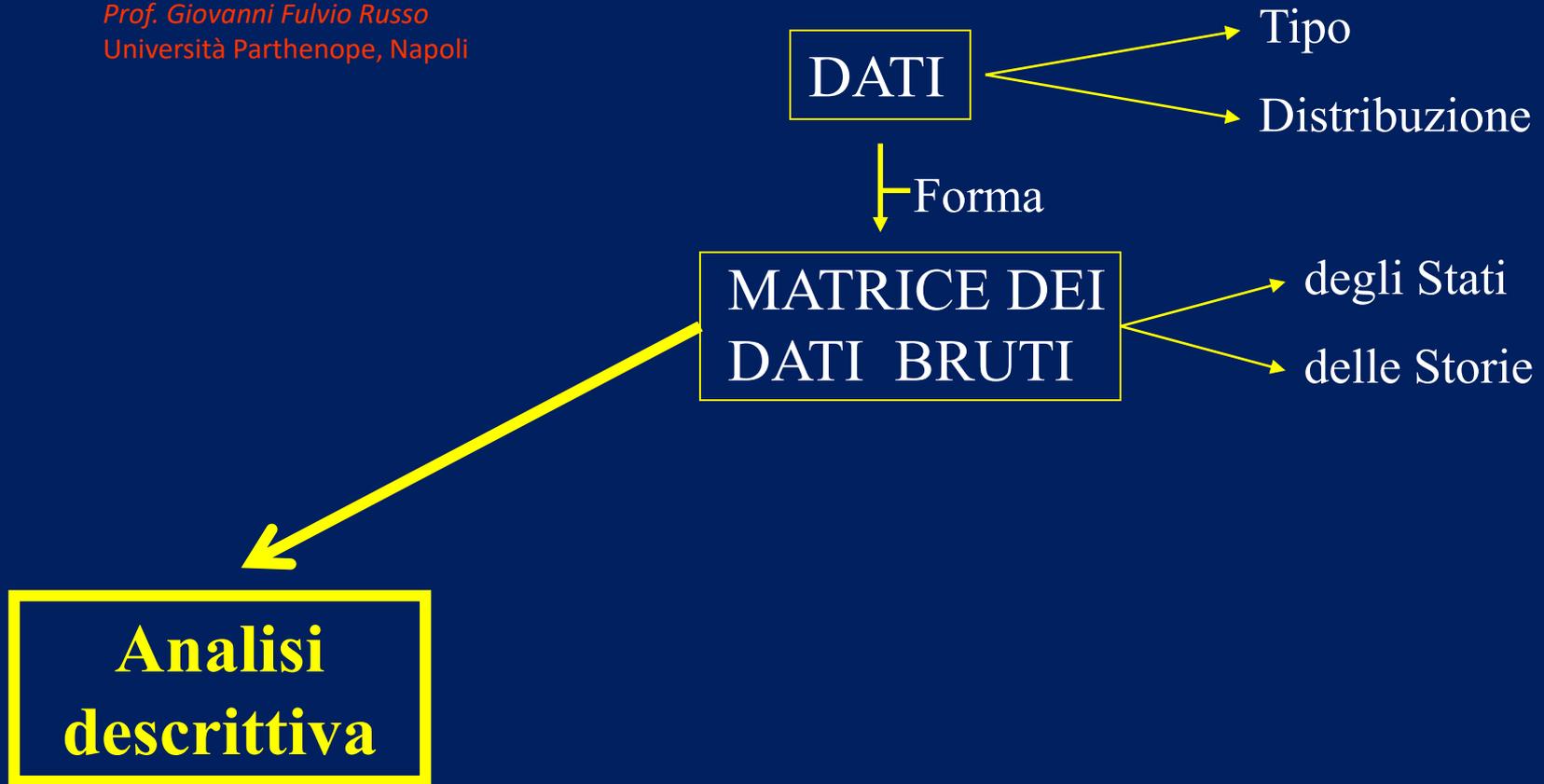
Specie  $x$  (Stazioni  $x$  Tempo)

	St1			St2		
Specie	Mar	Apr	Mag	Mar	Apr	Mag
A	7	5	0	2	1	1
B	3	0	1	5	7	1

## B) Matrice delle storie, inserendo il tempo nelle righe

Stazioni  $x$  (Specie  $x$  Tempo)

Specie	Tempo	St1	St2
A	Mar	7	2
	Apr	5	1
	Mag	0	1
B	Mar	3	5
	Apr	0	7
	Mag	1	1



# Analisi descrittiva della comunità

- Parametri sinecologici
- Abbondanza (A)
- Ricchezza specifica (RS)
- Dominanza quantitativa (DI)
- Dominanza qualitativa (DQ)
- Diversità specifica di Shannon (H')
- Equitabilità di Pielou (J)

$$P_i = \frac{x_i}{N}$$

$$H' = -\sum (P_i \log(P_i))$$

$$J = \frac{H'}{H_{MAX}}$$

$$H_{MAX} = \text{Log}_2 S$$

# Matrice con

## analisi

## descrittiva

## dei dati

- Calcoli di  
colonna (A,  
RS, H', J).

- Calcoli di  
riga (A, F).

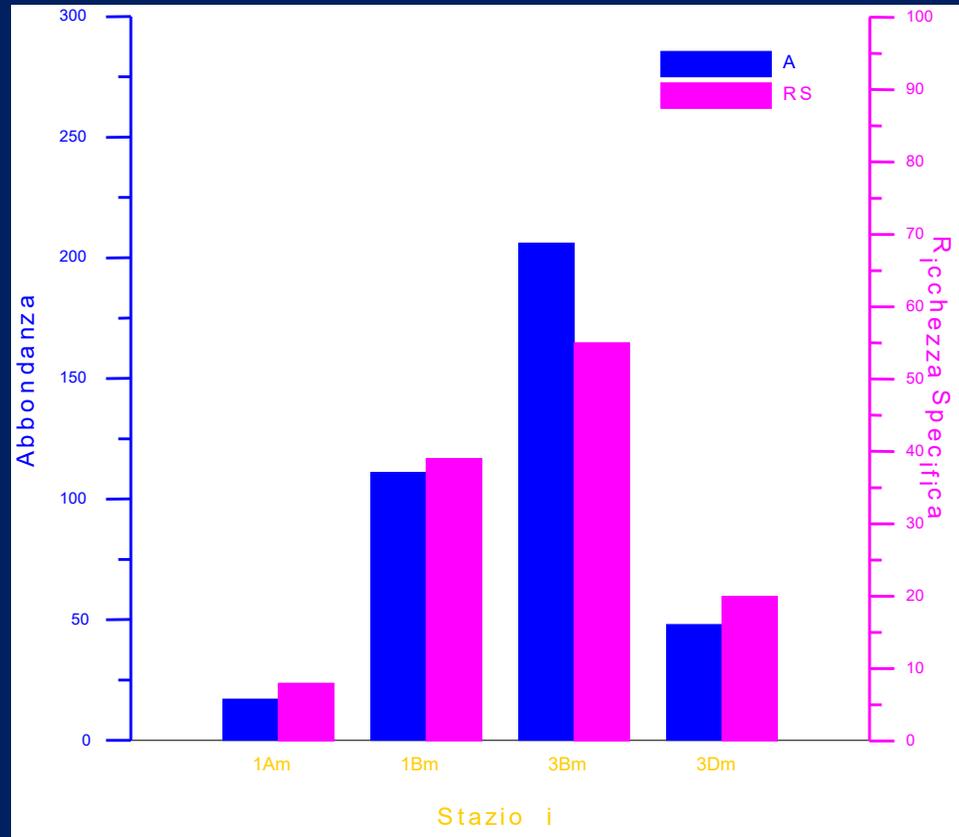
Golfo di Napoli			stazioni	1Bm	1Am	3Dm	3Bm		
Taxa	trof.	Specie	profondità (m)	15 m	4 m	23 m	11,5 m	A	F
Molluschi	E	Acanthochitona fascicularis		0	0	0	1	1	1
	S	Anomia ephippium		0	0	0	0	0	0
	D	Tellina tenuis		0	2	0	0	2	1
	S	Venus verrucosa		1	0	0	1	2	2
	S	Vermetus sp.		0	0	0	0	0	0
Policheti	L	Protoarica oerstedii		0	0	0	0	0	0
	L	Aricidea cerruti		0	0	0	1	1	1
	S	Spiraserpula massiliensis		0	0	0	0	0	0
	S	Vermiliopsis striaticeps		0	0	0	0	0	0
Anfipodi	D	Ampelisca rubella		0	0	0	0	0	0
	D	Amphilocheus neapolitanus		0	0	0	0	0	0
	O	Pariambus typicus		0	0	0	0	0	0
	O	Phtisica marina		0	0	0	0	0	0
Decapodi	DC	Palaemon serratus		0	0	0	0	0	0
	DC	Hippolyte leptocerus		0	0	0	0	0	0
	D	Pagurus anachoretus		0	0	2	0	2	1
	DC	Pagurus cuanensis		0	0	0	1	1	1
	DE	Acanthonyx lunulatus		0	0	0	0	0	0
Isopodi	DC	Paranthura nigropunctata		0	0	0	2	2	1
	DC	Cyathura carinata		0	0	0	0	0	0
	DE	Cleantis prismatica		1	0	0	0	1	1
Anisopodi	DE	Apseudes intermedius		0	0	0	2	2	1
	DE	Apseudes latreillei		0	0	0	0	0	0
	DE	Leptocheila dubia		0	0	0	0	0	0
	DE	Tanais dulongii		0	0	0	0	0	0
Echinodermi	DC	Amphipholis squamata		0	0	0	2	2	1
	S	Ophiothrix fragilis		0	0	0	0	0	0
	DE	Paracentrotus lividus		0	0	0	0	0	0
<b>Abbondanza (A)</b>				<b>2</b>	<b>2</b>	<b>2</b>	<b>10</b>	<b>16</b>	
<b>Ricchezza Specifica (RS)</b>				<b>2</b>	<b>1</b>	<b>1</b>	<b>7</b>	<b>10</b>	
<b>Diversità Specifica (H')</b>				<b>4,80</b>	<b>2,82</b>	<b>3,91</b>	<b>4,78</b>		
<b>Equitabilità (J)</b>				<b>0,91</b>	<b>0,94</b>	<b>0,90</b>	<b>0,83</b>		



# Analisi descrittiva della comunità: forme grafiche

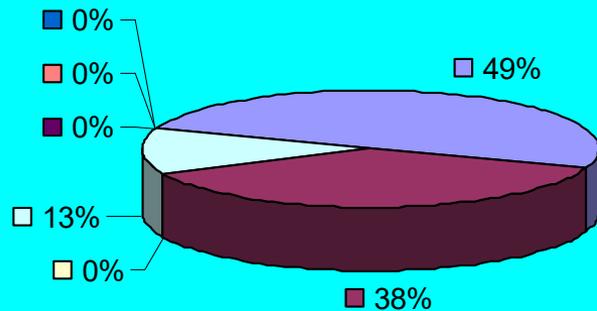
## ISTOGRAMMI

Abbondanza (A)  
Ricchezza specifica (RS)

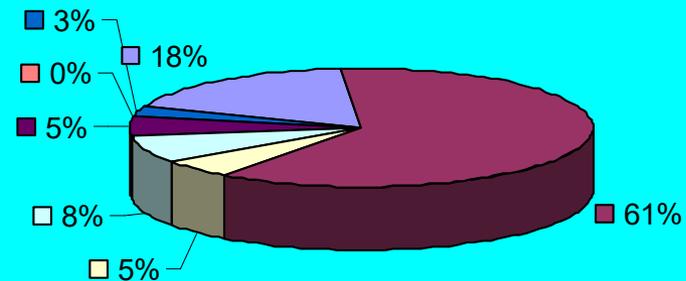


# Dominanza quantitativa (**DI**) o qualitativa (**DQ**)

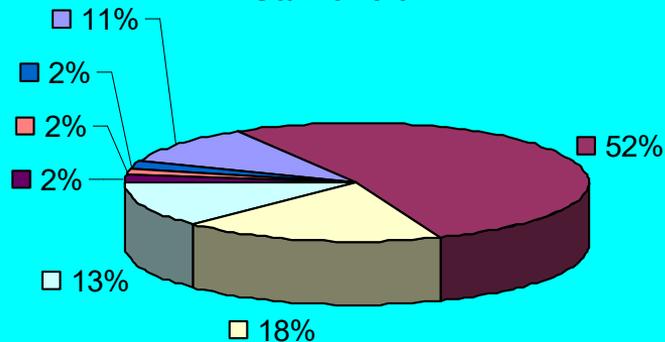
Stazione 1Am



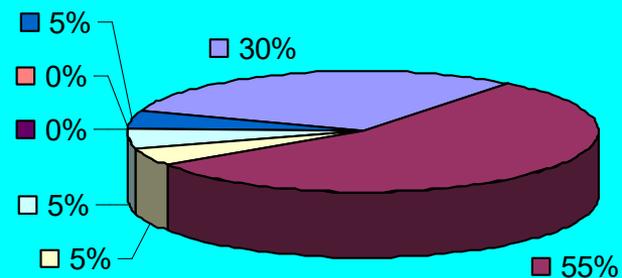
Stazione 1Bm



Stazione 3Bm



Stazione 3Dm



- Molluschi
- Policheti
- Anfipodi
- Decapodi
- Isopodi
- Anisopodi



# TRASFORMAZIONE DEI DATI

## 1) NORMALIZZAZIONE

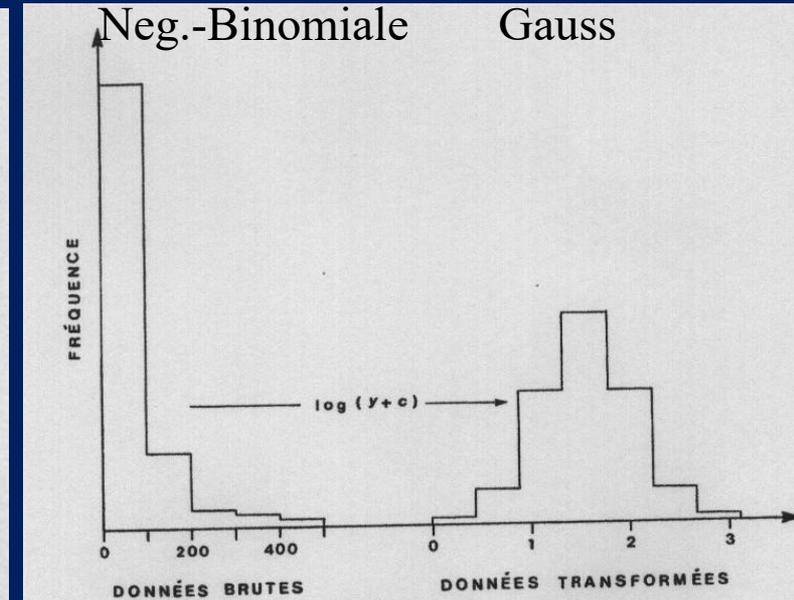
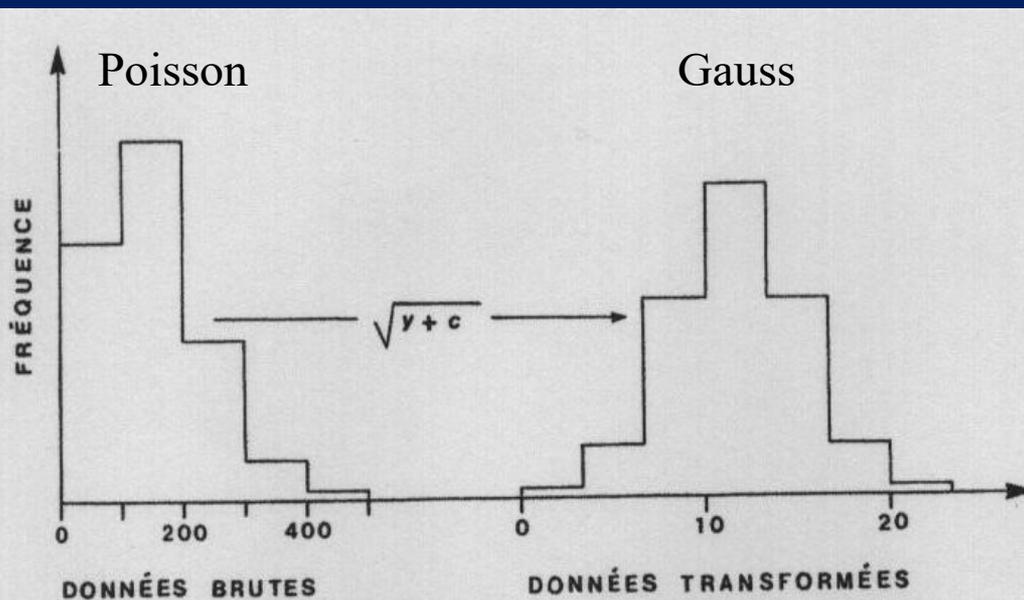
Serve a omogeneizzare la varianza dei dati (*omoscedasticità*), cioè a rendere aleatoria la sua distribuzione, conformemente alla "distribuzione normale" di Gauss.

$$y = \sqrt{x}$$

$$y = \log(x+1)$$

$x$  -> dato bruto

$y$  -> dato trasformato

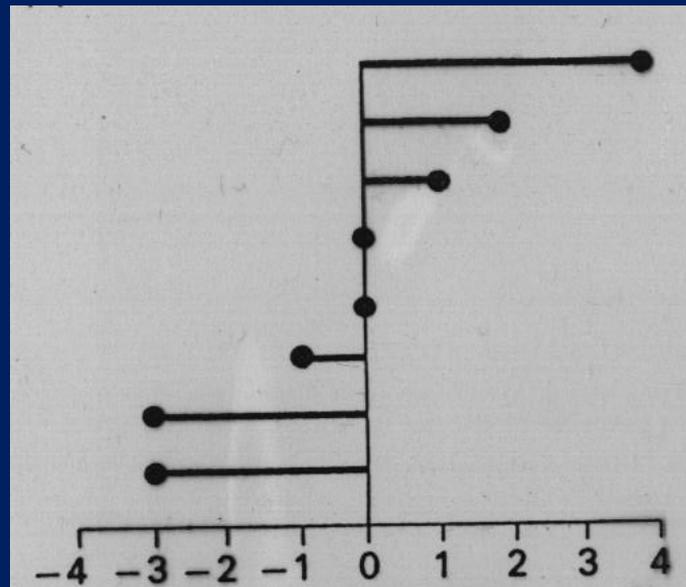


# TRASFORMAZIONE DEI DATI

## 2) STANDARDIZZAZIONE

$$y = (x - \mu) / \sigma$$

Serve per centrare e ridurre i dati, affinché sia migliorata la loro comparabilità, nel caso in cui vi siano alcuni dati aberranti (*outlayers*) rispetto alla restante parte.



Deviazione dalla  
media: “dati centrati”

$$x - \mu$$



## Matrice dei dati trasformati

Oggetti (osservazioni)

	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<b>a</b>	$x_{1a}$	$x_{2a}$	$x_{3a}$	$x_{4a}$	$x_{5a}$
<b>b</b>	$x_{1b}$	$x_{2b}$	$x_{3b}$		
<b>c</b>	$x_{1c}$	$x_{2c}$			
<b>d</b>	$x_{1d}$				

Descrittori  
(variabili)

# Indici di somiglianza

## Indici di somiglianza

### Similitudine

- **Dati binari** (presenza/assenza)
- Perdita d'informazione (no abbondanza)

### Distanza

- Dati quantitativi continui (**misure**)
- Normalizzazione: *non* necessaria
- Standardizzazione: necessaria

### Correlazione

- Dati quantitativi discontinui (**conte**)
- Normalizzazione: necessaria
- Standardizzazione: *non* necessaria

# Indici di Similitudine

- **Indice di Jaccard:** è il più utilizzato e il più semplice

$$J = \frac{a}{a + b + c}$$

**a = numero di specie condiviso da due stazioni**

**b = numero di specie presenti solo in una stazione**

**c = numero di specie esclusive dell'altra stazione**

- **Indice di Sorensen:** rispetto all'indice di Jaccard, considera il doppio delle specie condivise per stabilizzare il dato.

$$S = \frac{2a}{a + b + c}$$

- **Indice del confronto semplice:** si tiene conto anche del valore d

$$C = \frac{a + d}{a + b + c + d}$$

**d = specie assenti nelle due stazioni che stiamo comparando ma che comunque sono presenti nella matrice**

# Indici di Distanza

- Distanza Euclidea

$$ED_{JK} = \sqrt{\sum_{i=1}^S (x_{ij} - x_{ik})^2}$$

$x_j$  = abbondanza della specie  $i$  nella stazione  $j$

$x_k$  = abbondanza della specie  $i$  nella stazione  $K$

- Distanza Quadratica

$$SED_{jk} = \sum_{i=1}^S (x_{ij} - x_{ik})^2$$

- Euclidea Media

$$MED_{jk} = \sqrt{\frac{\sum_{i=1}^S (x_{ij} - x_{ik})^2}{S}}$$

$S$  = numero di specie complessivo

# Indici di Correlazione

Indice di Pearson: mette insieme varianza e covarianza

$$r = \frac{S_{xy}}{S_x^2 S_y^2}$$

$$\text{Varianza} \begin{cases} S_x^2 = \frac{\sum (x - \bar{x})^2}{n - 1} \\ S_y^2 = \frac{\sum (y - \bar{y})^2}{n - 1} \end{cases}$$

$$\text{Covarianza} \begin{cases} \frac{\sum (x - \bar{x})^2 (y - \bar{y})^2}{n - 1} \end{cases}$$

Per dati semi-quantitativi (ranghi):

- Indice di Kendal
- Indice di Spearman



# Matrice di somiglianza

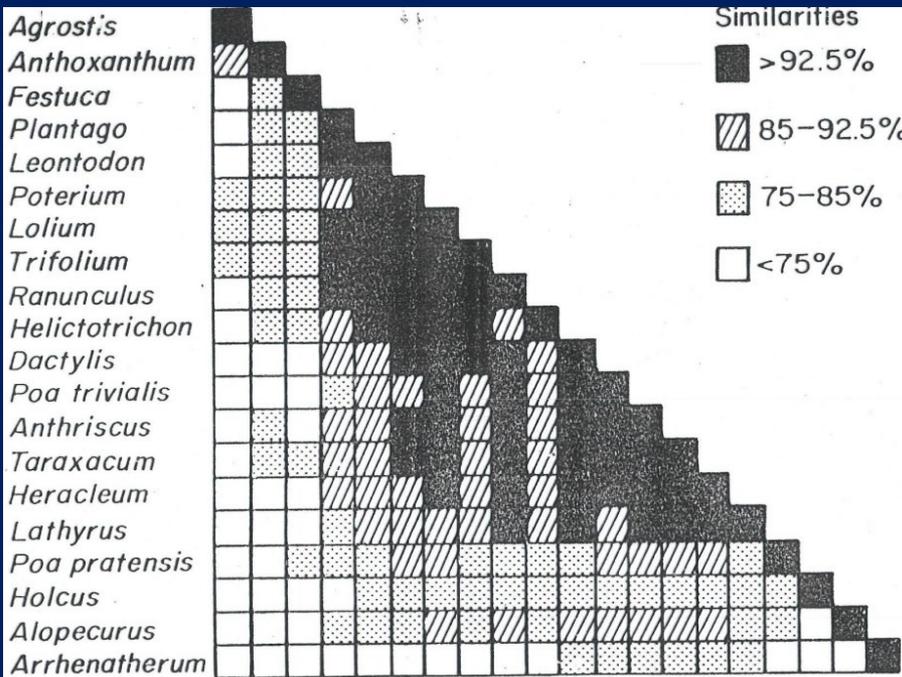
La matrice di somiglianza si ottiene riportando gli indici di somiglianza dei confronti a coppia tra stazioni (matrice di “moda Q”) o tra specie (matrice di “moda R”).

La matrice di somiglianza è di forma triangolare, dato che è simmetrica. Infatti, nella metà al di sopra e in quella al di sotto della sua diagonale (che, poiché riporta il confronto di ciascuna stazione o specie con sé stessa, contiene solo valori 1, cioè 100% di somiglianza), i valori sono gli stessi e quindi si rappresenta solo una metà della matrice.

	VetPrim1	IscaPrim1	TIscaPrim1	VetAut2	IscaAut2	SEAut2	VetPrim3	IscaPrim3	SEPrim3	VetAut4	IscaAut4	SEAut4
VetPrim1												
IscaPrim1	73,62711											
TIscaPrim1	69,30888	69,27039										
VetAut2	71,84796	76,25928	64,75351									
IscaAut2	65,85359	76,54492	66,65855	74,86128								
SEAut2	59,51942	64,48207	68,99114	67,57256	65,51302							
VetPrim3	62,47224	59,37507	54,0829	66,48424	55,79582	48,60488						
IscaPrim3	61,95262	66,50803	61,51754	66,94793	61,41883	64,40557	69,62962					
SEPrim3	62,33871	61,04604	67,82852	64,33988	55,9859	65,40607	63,82166	72,73986				
VetAut4	73,44641	80,11995	75,49565	80,48673	75,39728	71,11471	68,22447	71,68551	71,32032			
IscaAut4	68,88192	76,96725	76,63666	77,67493	77,893	79,79325	58,06785	70,57069	70,35122	84,57342		
SEAut4	62,97244	69,39327	71,38795	70,0525	71,74532	77,08133	54,24477	66,99313	66,04095	76,73499	78,26546	

# Matrice di somiglianza

	VetPrim1	IscaPrim1	TIscaPrim1	VetAut2	IscaAut2	SEAut2	VetPrim3	IscaPrim3	SEPrim3	VetAut4	IscaAut4	SEAut4
VetPrim1												
IscaPrim1	73,62711											
TIscaPrim1	69,30888	69,27039										
VetAut2	71,84796	76,25928	64,75351									
IscaAut2	65,85359	76,54492	66,65855	74,86128								
SEAut2	59,51942	64,48207	68,99114	67,57256	65,51302							
VetPrim3	62,47224	59,37507	54,0829	66,48424	55,79582	48,60488						
IscaPrim3	61,95262	66,50803	61,51754	66,94793	61,41883	64,40557	69,62962					
SEPrim3	62,33871	61,04604	67,82852	64,33988	55,9859	65,40607	63,82166	72,73986				
VetAut4	73,44641	80,11995	75,49565	80,48673	75,39728	71,11471	68,22447	71,68551	71,32032			
IscaAut4	68,88192	76,96725	76,63666	77,67493	77,893	79,79325	58,06785	70,57069	70,35122	84,57342		
SEAut4	62,97244	69,39327	71,38795	70,0525	71,74532	77,08133	54,24477	66,99313	66,04095	76,73499	78,26546	



Per ottenere un'immagine leggibile della somiglianza tra i confronti a coppie, si utilizzano ranghi di tonalità di grigio, che vanno dal bianco (- simile) al nero (+ simile); con questo sistema si ottiene un quadro "analogico" di rapida interpretazione.



# Analisi strutturale della comunità

## Analisi statistiche multivariate (risultati grafici):

- **Classificazione** → dendrogrammi, diagrammi “ad albero” rovesciato.
- **Ordinamento** → scatterplot, nubi di dispersione di punti, in un sistema di assi cartesiani.

**Non gerarchica** (classi di pari livello)  
(*minimum spanning tree*)

**CLASSIFI-  
CAZIONE**

Gerarchica

Divisiva  
(*top-down*)

Sistematica

Agglomerativa  
(*down-top*)

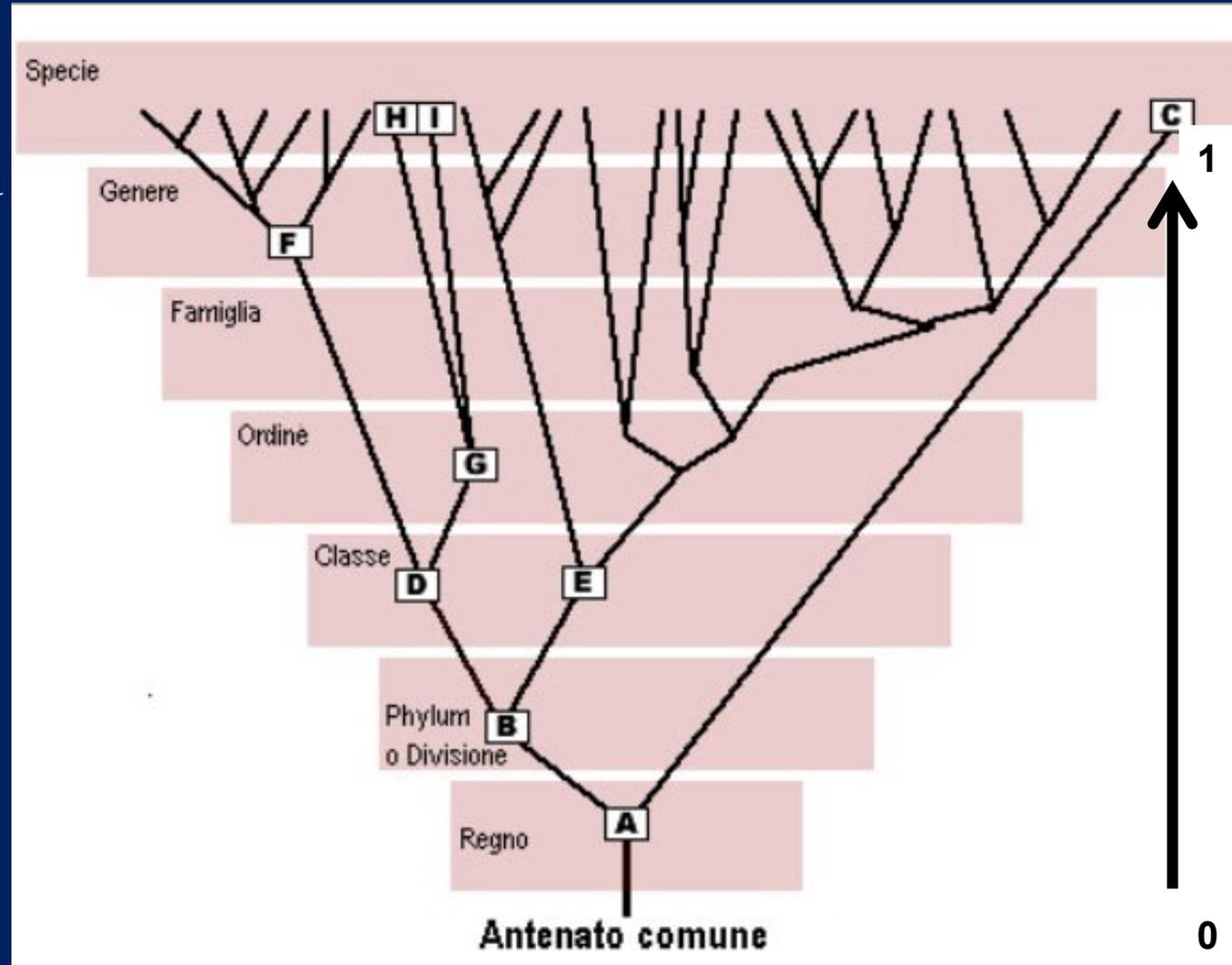
Legame semplice

Legame completo

Legame medio

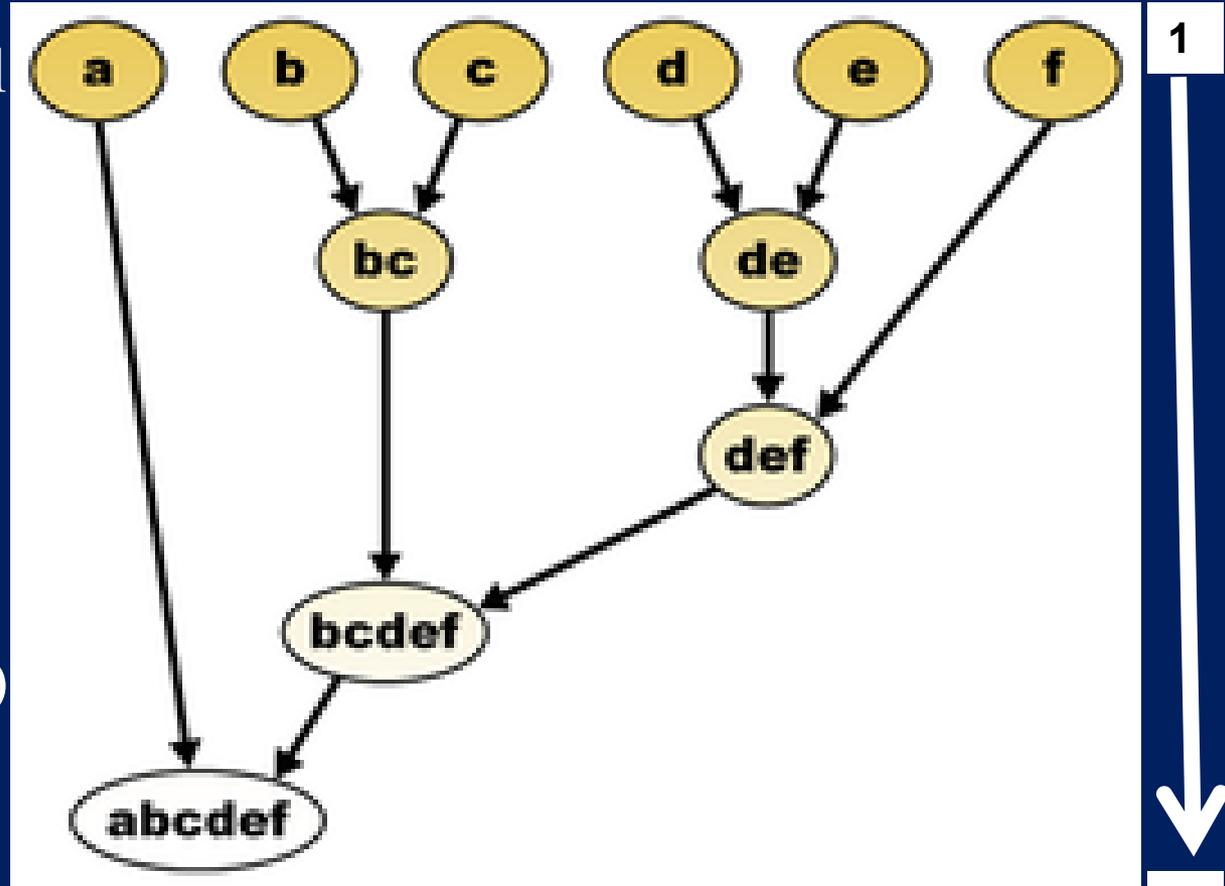
# Classificazione divisiva

Si parte da un insieme più grande e poi lo si suddivide in insiemi sempre più piccoli, contenenti oggetti sempre più somiglianti. Ad es. nella sistematica zoologica si parte da un *phylum* e poi lo si suddivide classi; ogni classe in ordini; ogni ordine in famiglie ecc.



# Classificazione agglomerativa

Si parte da singoli oggetti per poi formare gruppi sempre più grandi col diminuire del grado di somiglianza. In ecologia, si parte da stazioni (o specie) singole per poi raggrupparle per gradi di somiglianza sempre minori. I primi gruppi (*clusters*) che si formano contengono stazioni molto simili fra loro, il gruppo finale le contiene tutte.



somiglianza

# Criteri di Aggregazione

Per stabilire come aggregare i gruppi fra loro, affinché se ne formi uno più grande, bisogna scegliere in ciascuno i valori più simili.  
(a quale gruppo si aggredherà il punto *D* dell'esempio in figura ?)

## Legame Semplice (*Single-Linkage*)

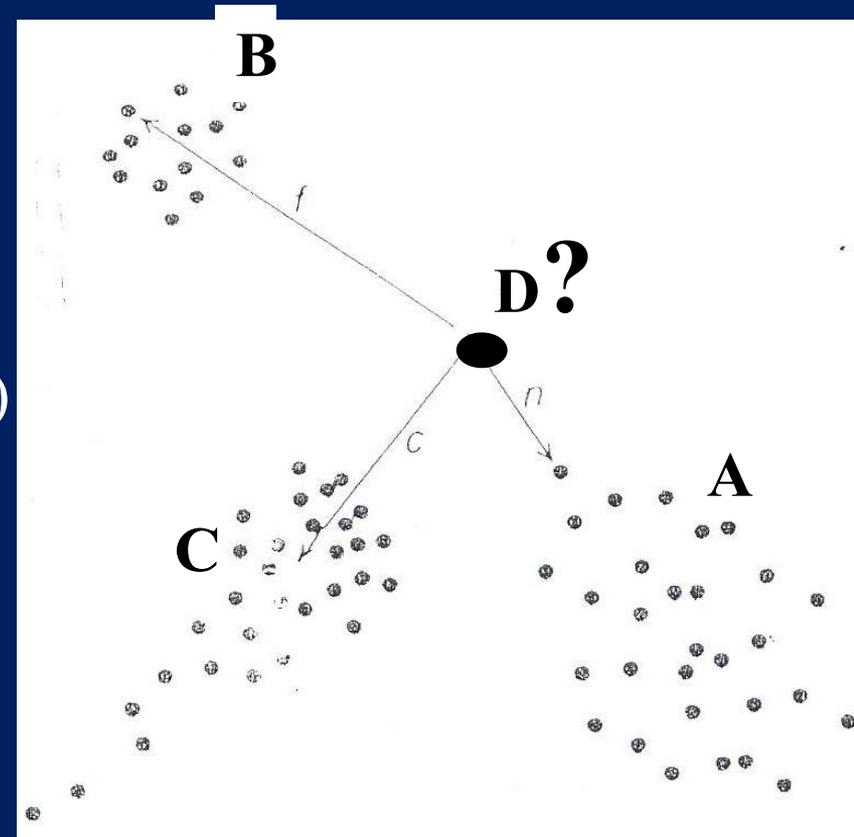
Scelta del valore più simile tra i più simili di ciascun gruppo  
(*D* si aggredherà al gruppo A)

## Legame Completo (*Complete-Linkage*)

Scelta del valore più simile tra i più dissimili di ciascun gruppo  
(*D* si aggredherà al gruppo B)

## Legame Medio (*Average-Linkage*)

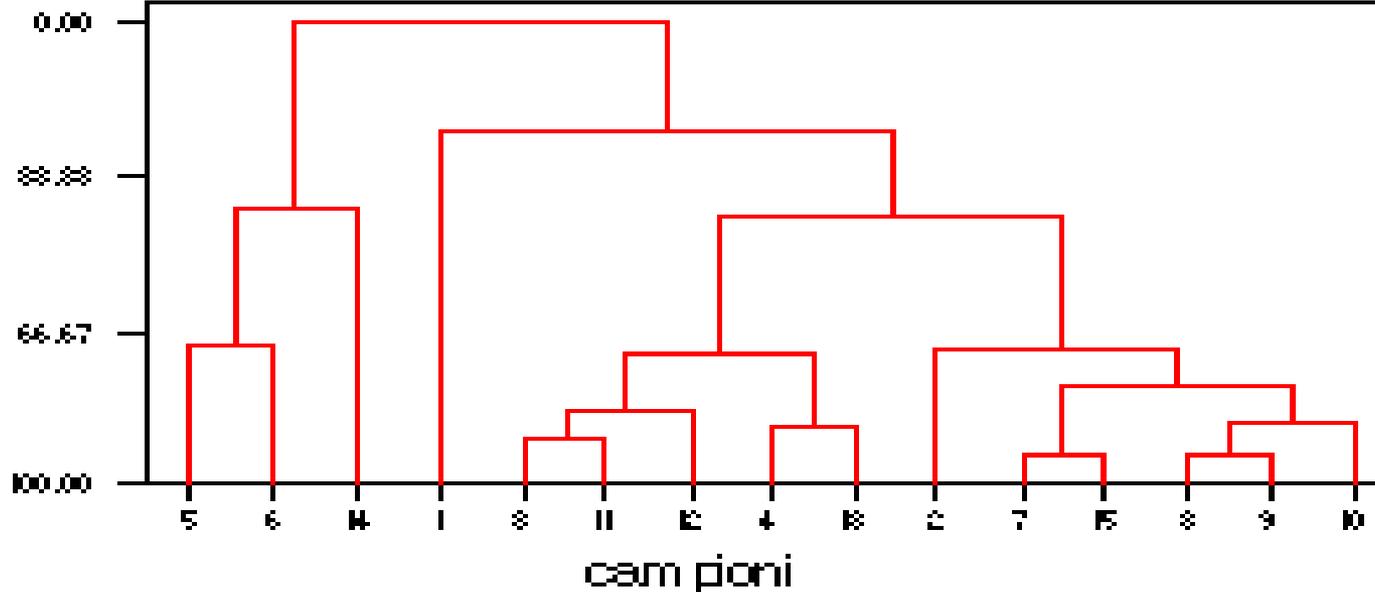
Scelta del valore più simile tra le medie dei valori di ciascun gruppo  
(*D* si aggredherà al gruppo C).

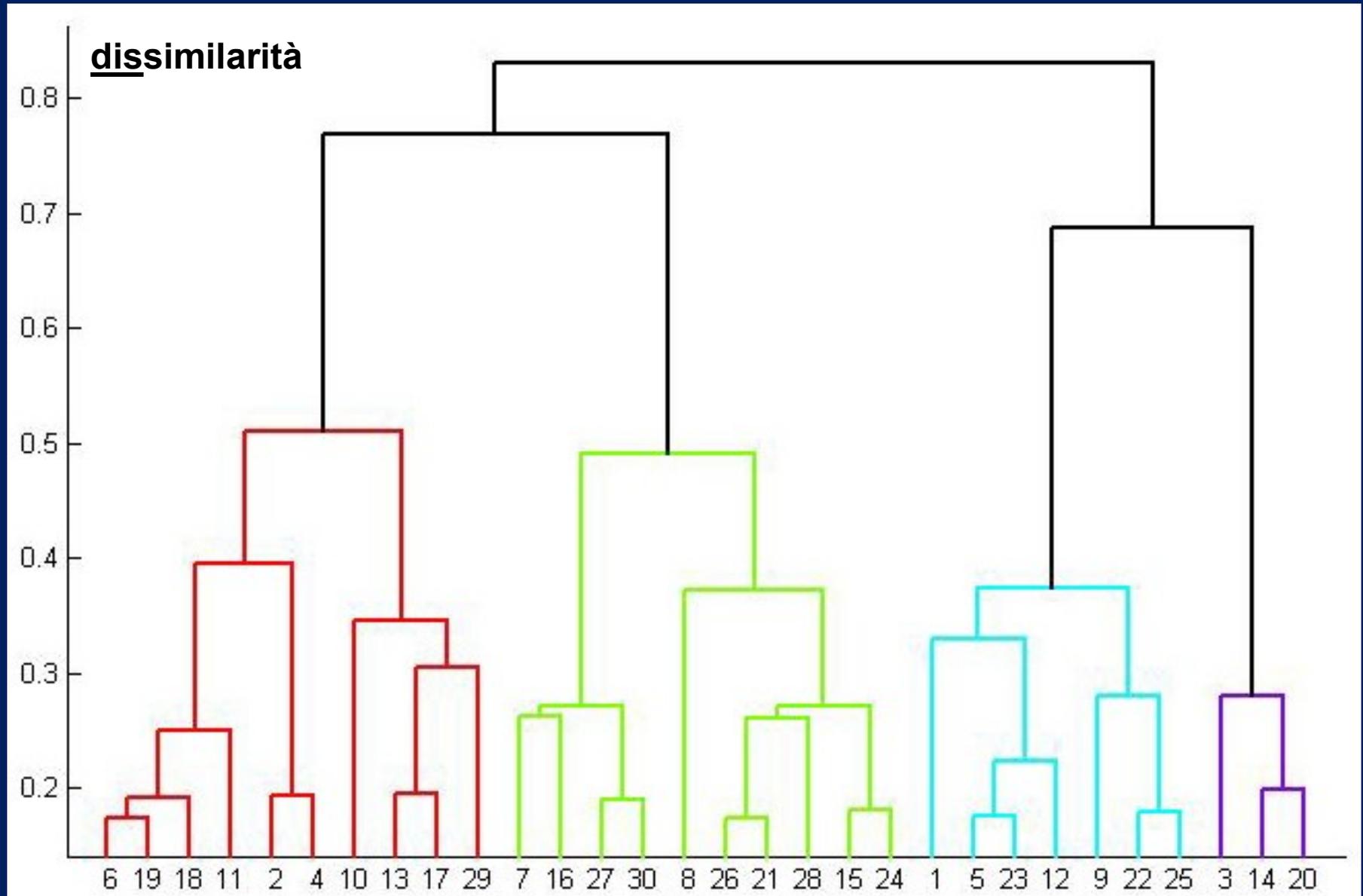


N.B. Nell'esempio con i punti, “più simile” vuol dire “più vicino”.

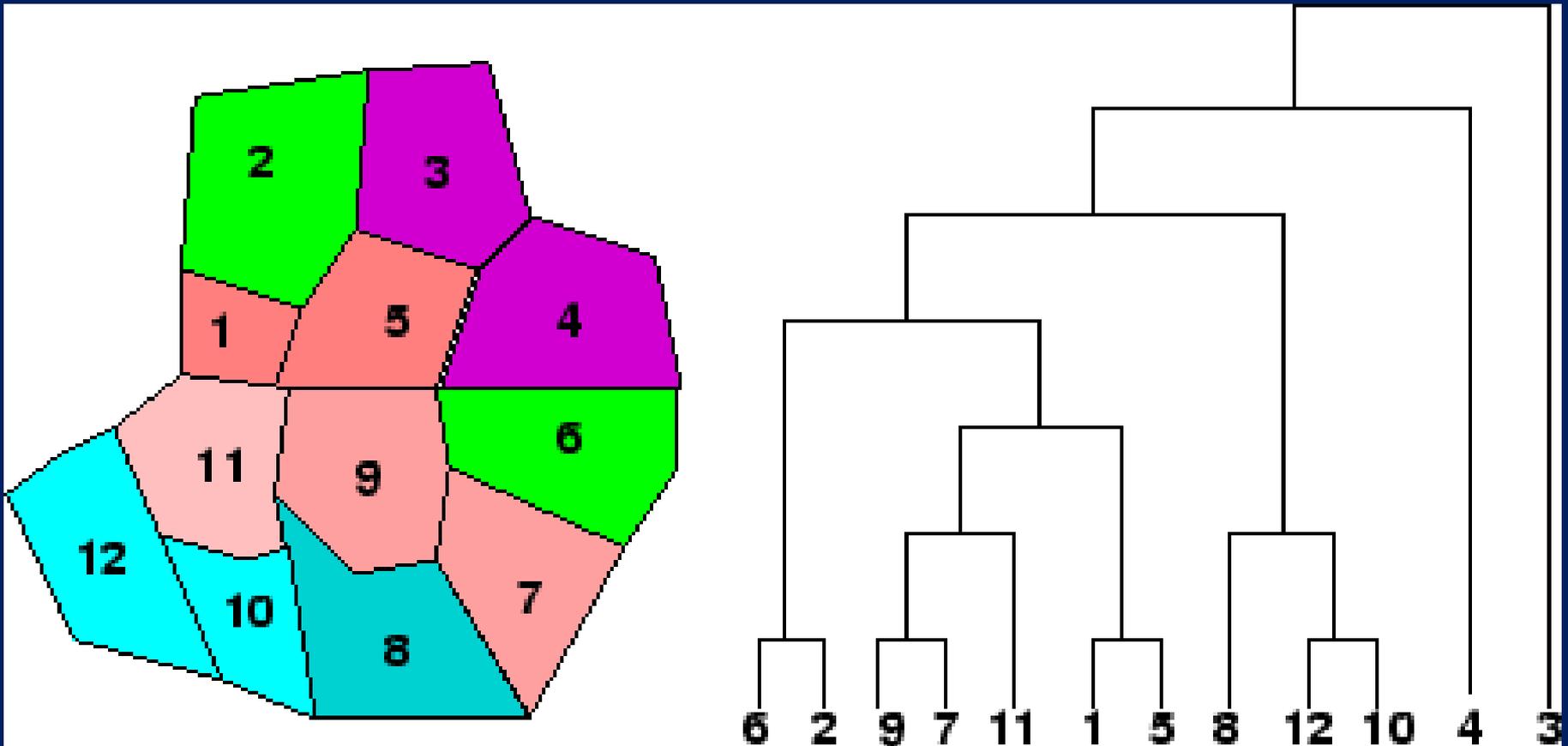
Il dendrogramma, cioè il diagramma ad albero rovesciato, ottenuto da una matrice di somiglianza adottando uno dei tipi di legame, riporta sull'asse orizzontale i singoli oggetti da aggregare (*campioni*) e sull'asse verticale il livello di somiglianza (o *similarità*) a cui avviene l'aggregazione tra due o più oggetti o gruppi.

similarità

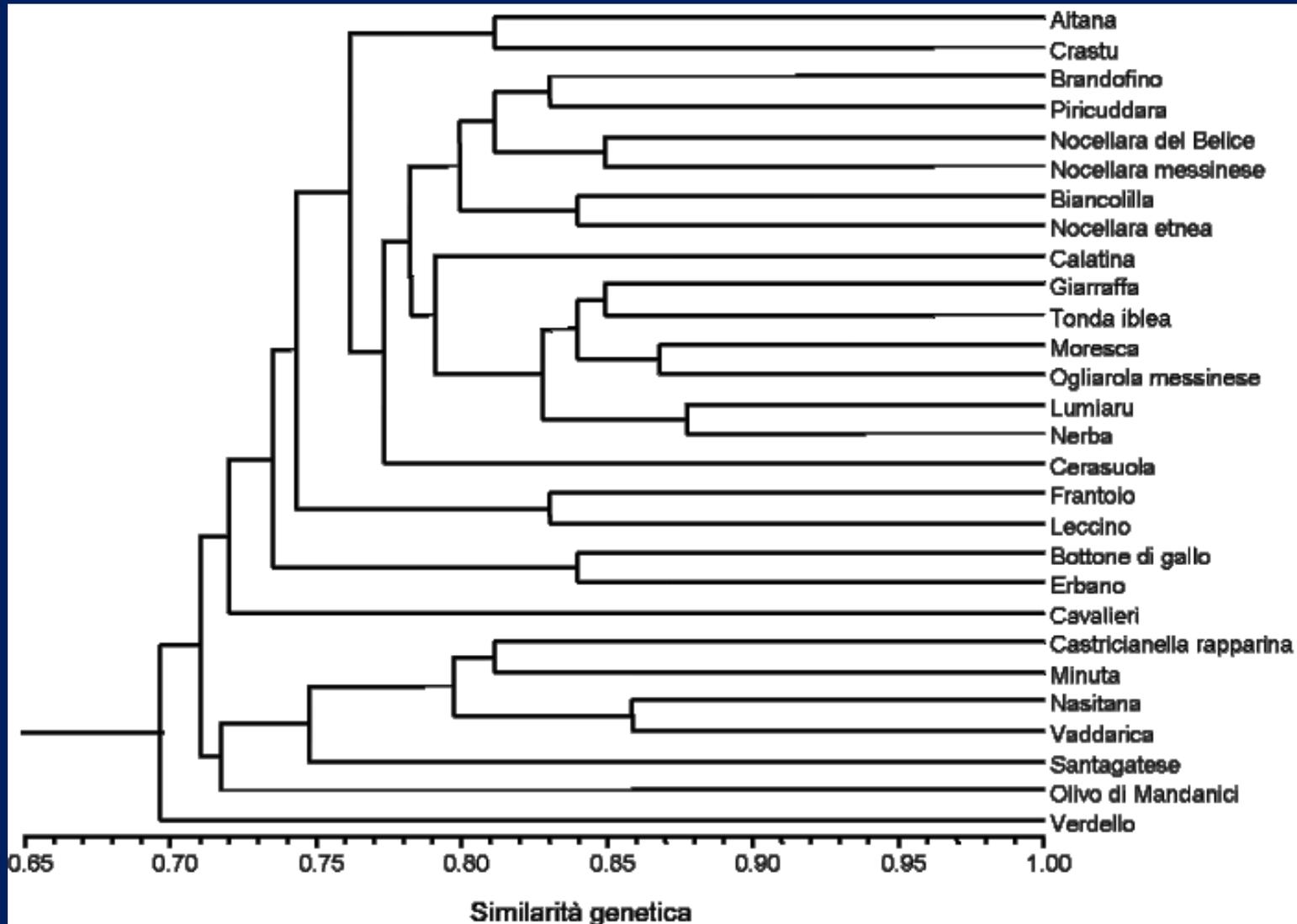




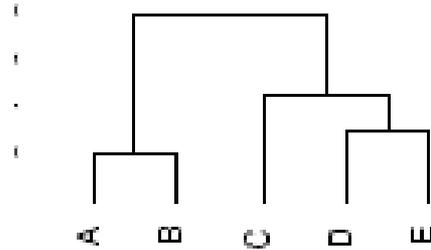
# Dendrogramma delle somiglianze di parcelle di territorio di tonalità diverse di colore.



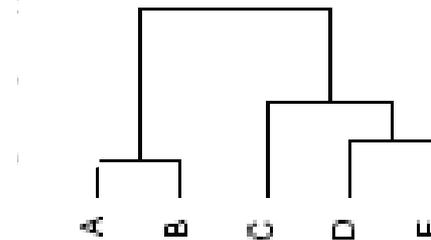
# Dendrogramma delle somiglianze genetiche delle varietà di olivo (in questa rappresentazione sono ruotati gli assi).



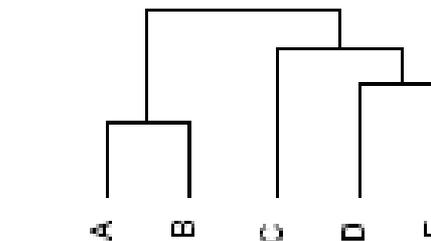
Legame medio



Legame completo



Legame singolo



I livelli di somiglianza a cui avviene l'aggregazione, e talvolta anche gli oggetti o i gruppi che si aggregano, possono cambiare a seconda del tipo di legame che viene adottato (semplice, completo o medio).

In genere, il legame *semplice* produce un dendrogramma “compresso”, il *completo* un dendrogramma “dilatato”, il *medio* un dendrogramma “conservativo”.

L'ordinamento è un diagramma di dispersione di punti, rappresentanti gli oggetti (es. stazioni) e/o i descrittori (es. specie), la cui distanza fra loro è direttamente proporzionale al grado di somiglianza.

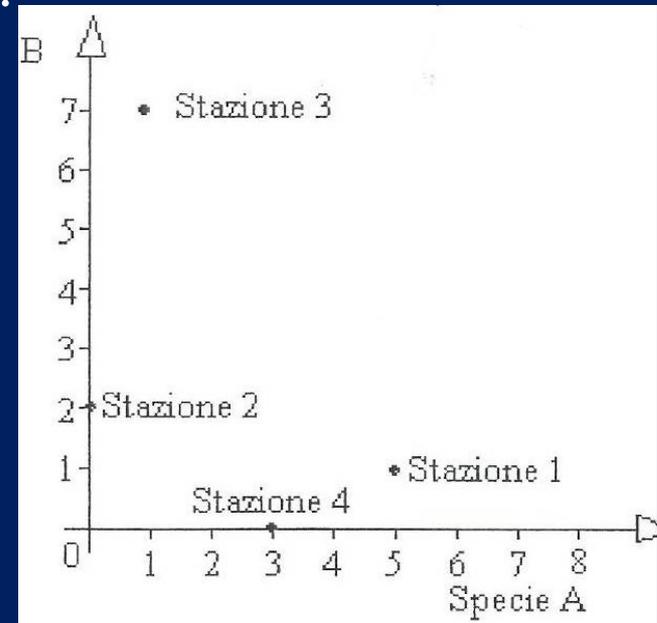


# Ordinamento diretto

L'ordinamento diretto consiste nell'utilizzare i valori di riga della matrice dei dati bruti come coordinate delle stazioni o, viceversa, i dati di colonna come coordinate delle specie.

Per l'ordinamento su di un piano (due assi) i vettori riga o colonna vanno presi a coppie (un vettore per asse).

	1	2	3	4
A	5	0	1	3
B	1	2	7	0
C	0	0	1	0
D	5	0	2	5



	1	2	3	4
A	5	0	1	3
B	1	2	7	0
C	0	0	1	0
D	5	0	2	5



## Ordinamento per estrazione di autovalori

Estraendo gli autovalori e gli autovettori dalla matrice di somiglianza dei dati, si concentra la varianza su pochi assi, rispetto ai quali vengono calcolate le coordinate dei punti-oggetto (stazioni) o dei punti-descrittore (specie).

Nell'ordinamento, ciascun asse esprime il gradiente di un "Fattore ambientale", che determina la distribuzione dei punti ordinati nello spazio cartesiano.

Tuttavia, il fattore ambientale collegato all'asse non è esplicito e deve essere interpretato.

La preminenza di ciascun fattore dipende dalla percentuale di varianza espressa dall'asse cui si associa.

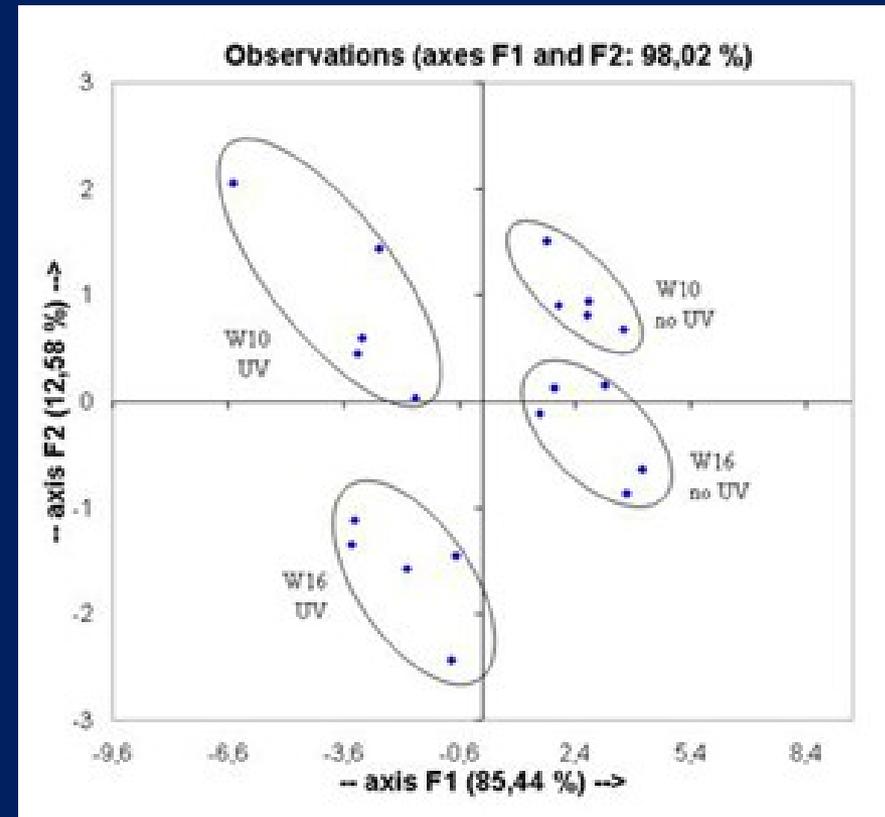
La disposizione degli assi è gerarchica, cioè il 1° (asse  $x$ ) concentra la varianza maggiore, seguito dal 2° (asse  $y$ ) con varianza minore, dal 3° (asse  $z$ ), con varianza ancora minore, ecc.

## Analisi delle componenti principali (PCA)

Le componenti principali (PC) costituiscono gli assi del sistema cartesiano, organizzati in base agli autovalori; nell'esempio, PC1 (Fattore 1) concentra l' 85,44% della varianza, PC2 il 12,58%.

Con questa analisi non è possibile sovrapporre l'ordinamento dei punti-variable con quello dei punti-stazione. Per poter rendere isospaziali i due ordinamenti occorre applicare la procedura del **BIPLOT**.

Questa tecnica è idonea per i dati metrici, quindi si applica a matrici di somiglianza ottenute con gli indici di distanza.

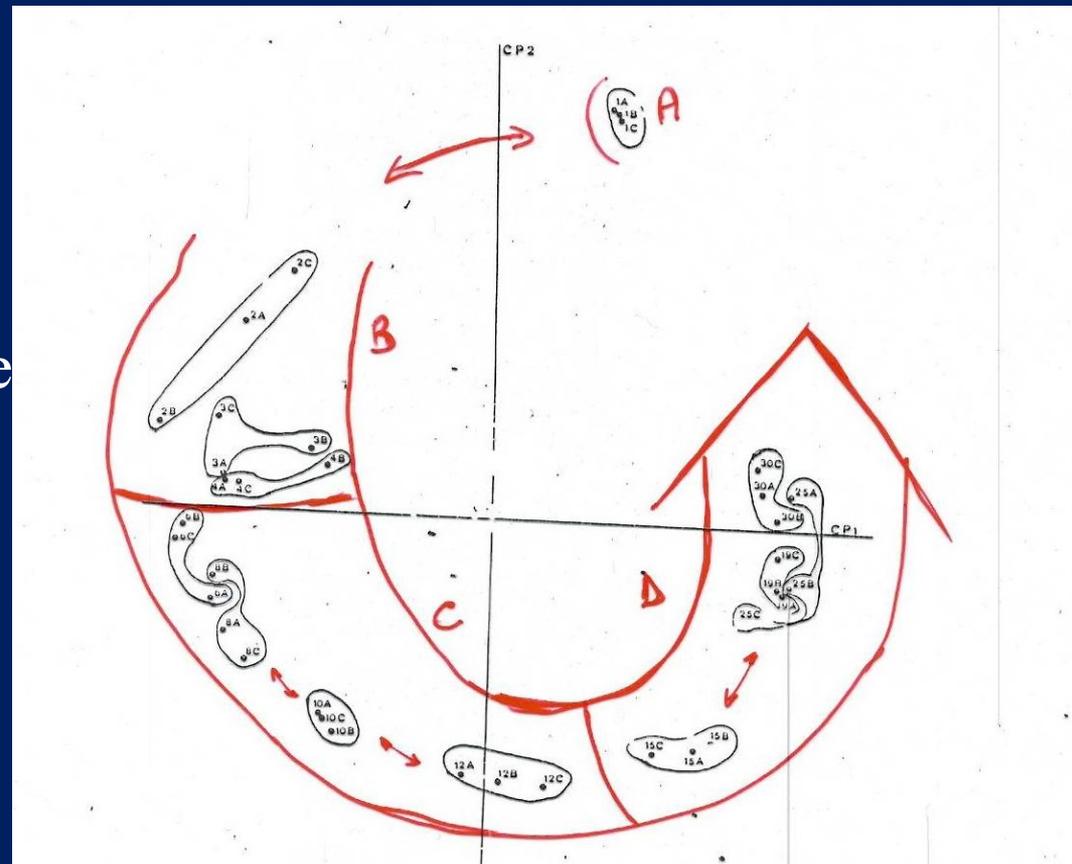


## Effetto Guttman

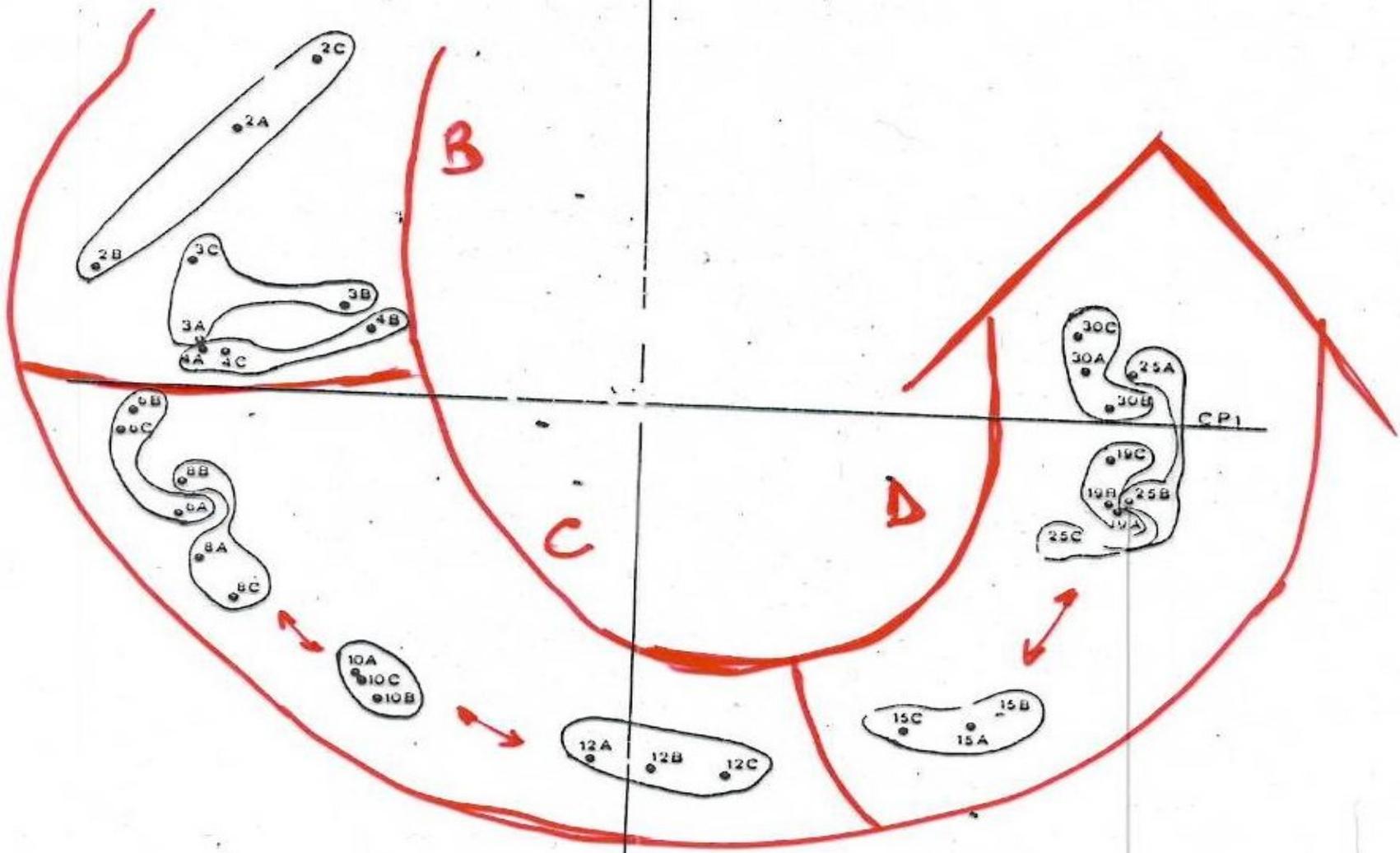
Quando si usa la tecnica della PCA, raramente i punti si dispongono ordinatamente lungo gli assi, ma il più delle volte subiscono delle distorsioni.

La più comune è il cosiddetto “*effetto Guttman*”, che fa assumere alla disposizione dei punti una configurazione ellittica (a “ferro di cavallo”).

Tale effetto dipende dalla differente varianza dei dati. Ad es., nel caso di un ordinamento di punti-stazione (come quello in figura), si ordineranno all’apice della parabola le stazioni in cui sono presenti comunità a più elevata diversità specifica (migliore struttura).



CP2



B

C

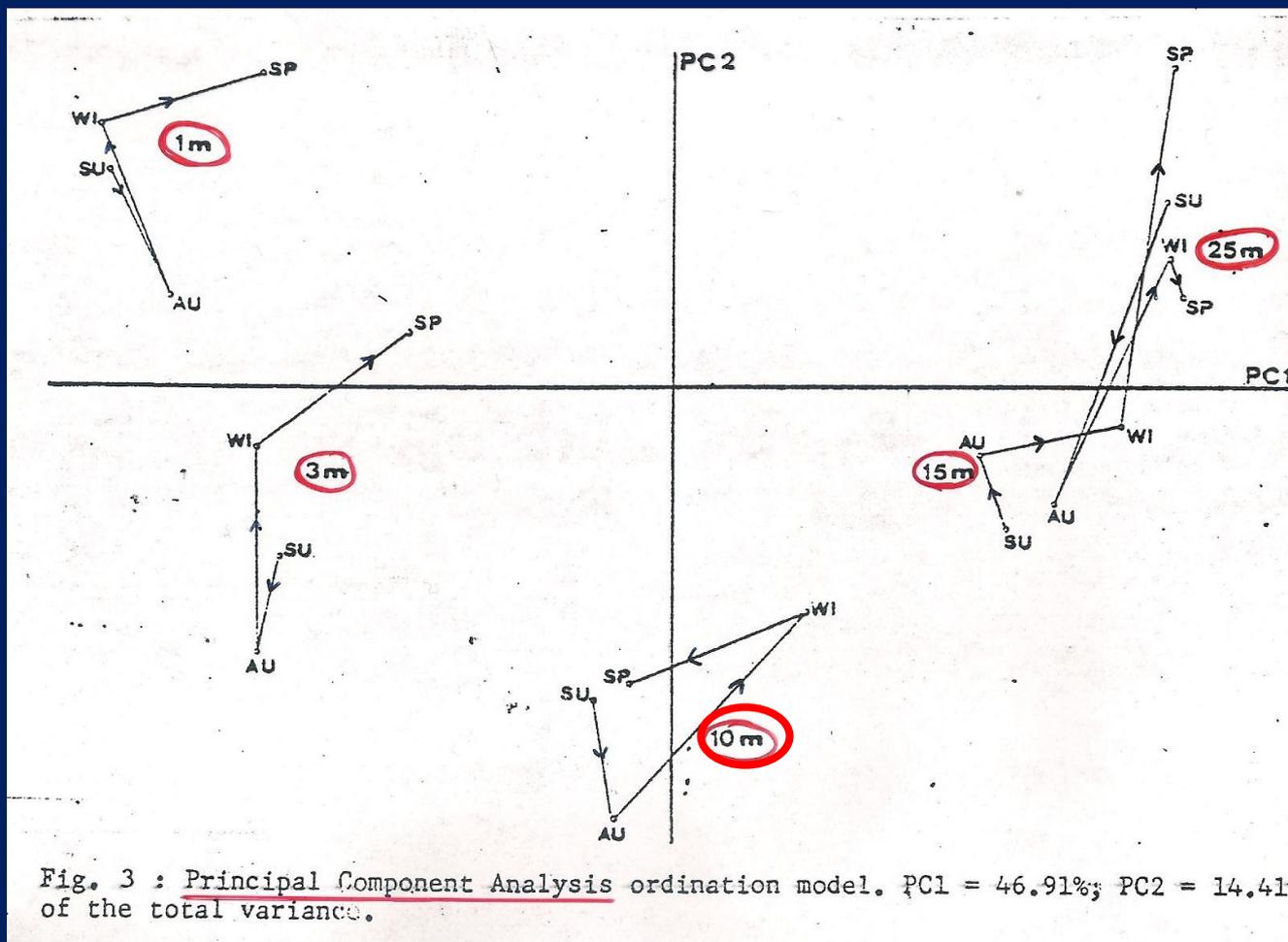
D

CP1

# Ordinamento da matrici Q degli stati

Traiettorie dei *punti-stazione* nel tempo (Su= estate, Au= autunno, Wi= inverno, Sp= primavera).

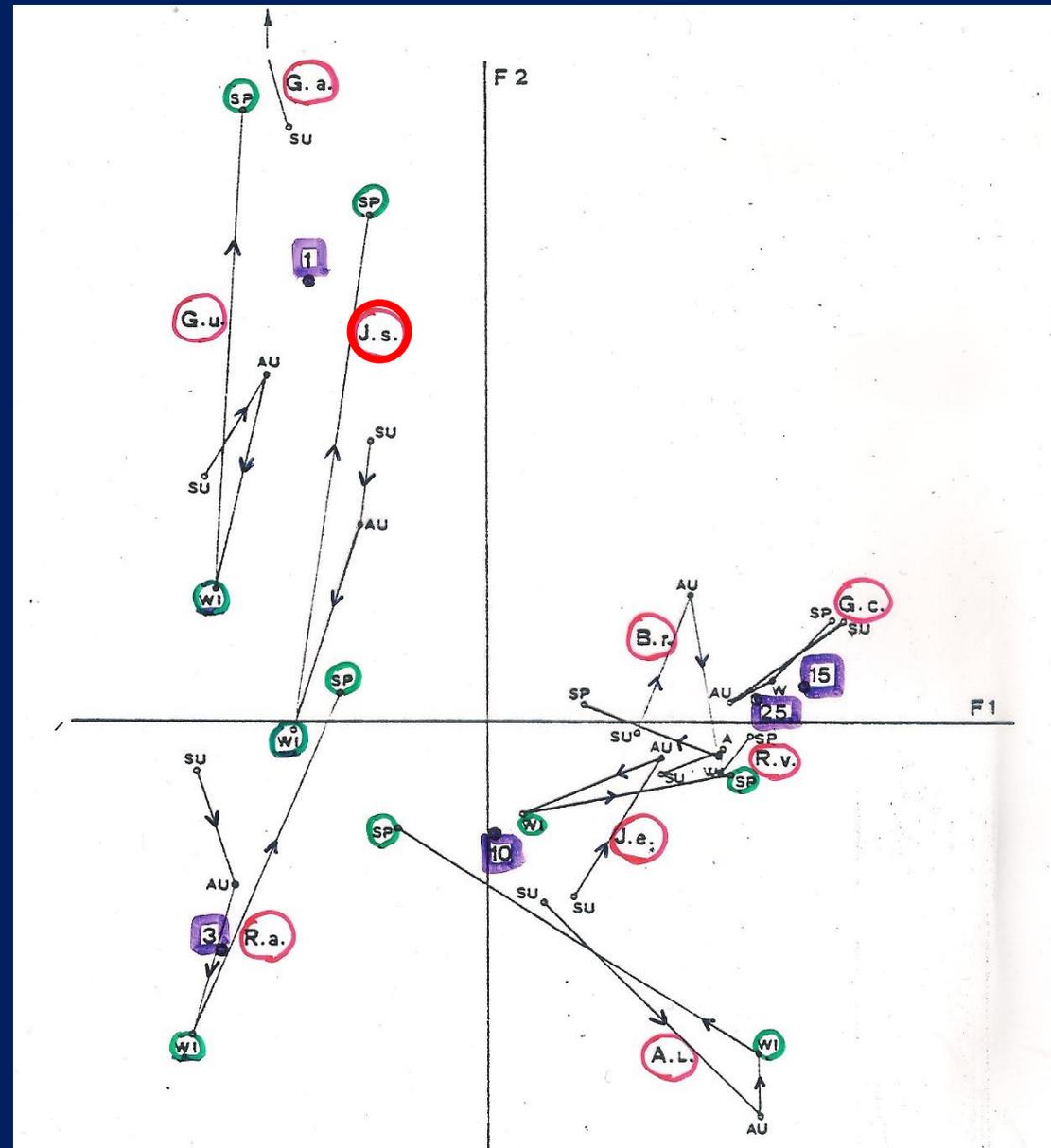
La stazione a *10 m* mostra una dinamica ciclica durante l'anno .



# Ordinamento da matrici R delle “storie”

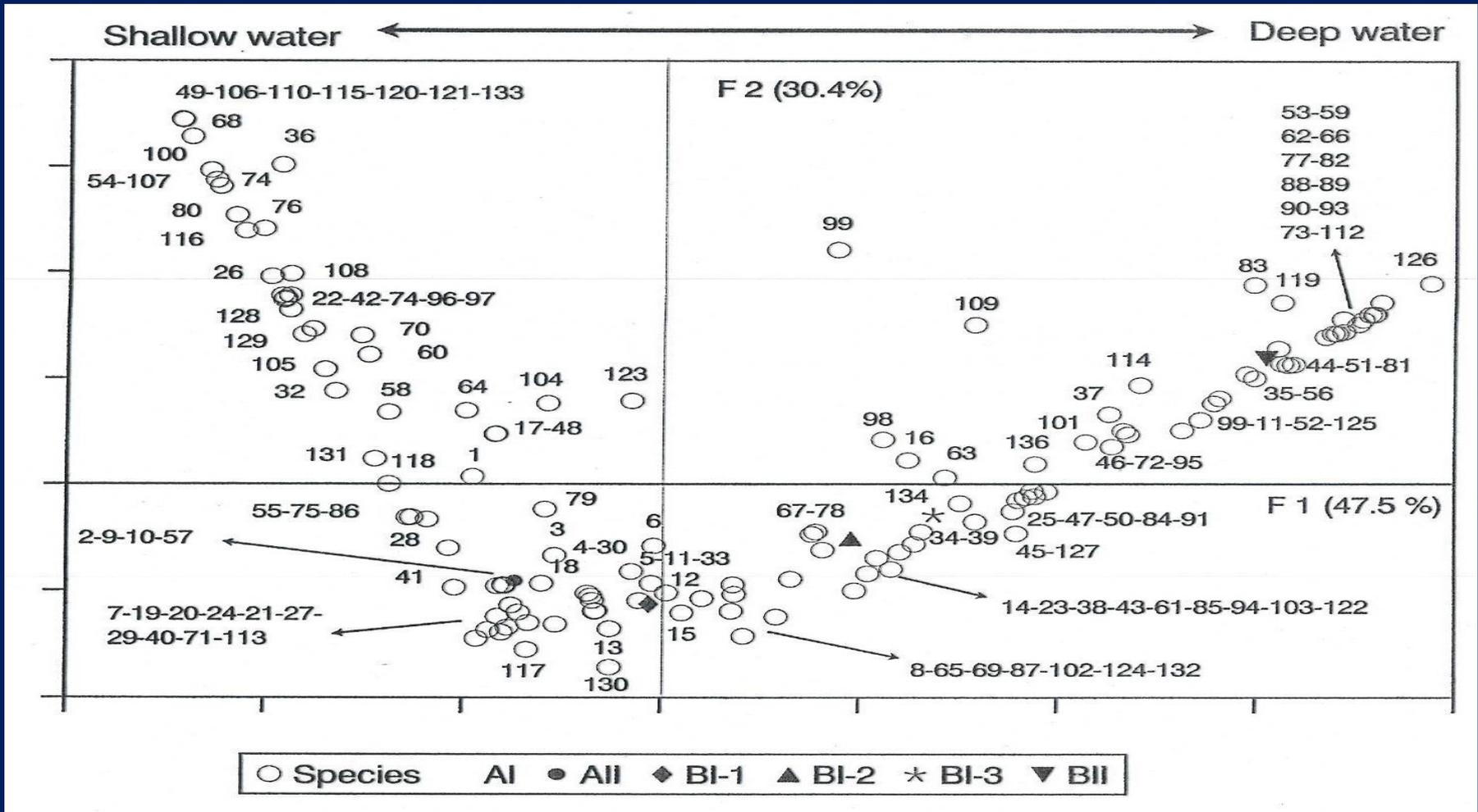
Traiettorie dei *punti-specie* nel tempo (Su= estate, Au= autunno, Wi= inverno, Sp= primavera).

La specie *J. s.* durante la primavera-estate popola soprattutto la stazione a 1 m di profondità, durante l'inverno scende a 3 m di profondità.



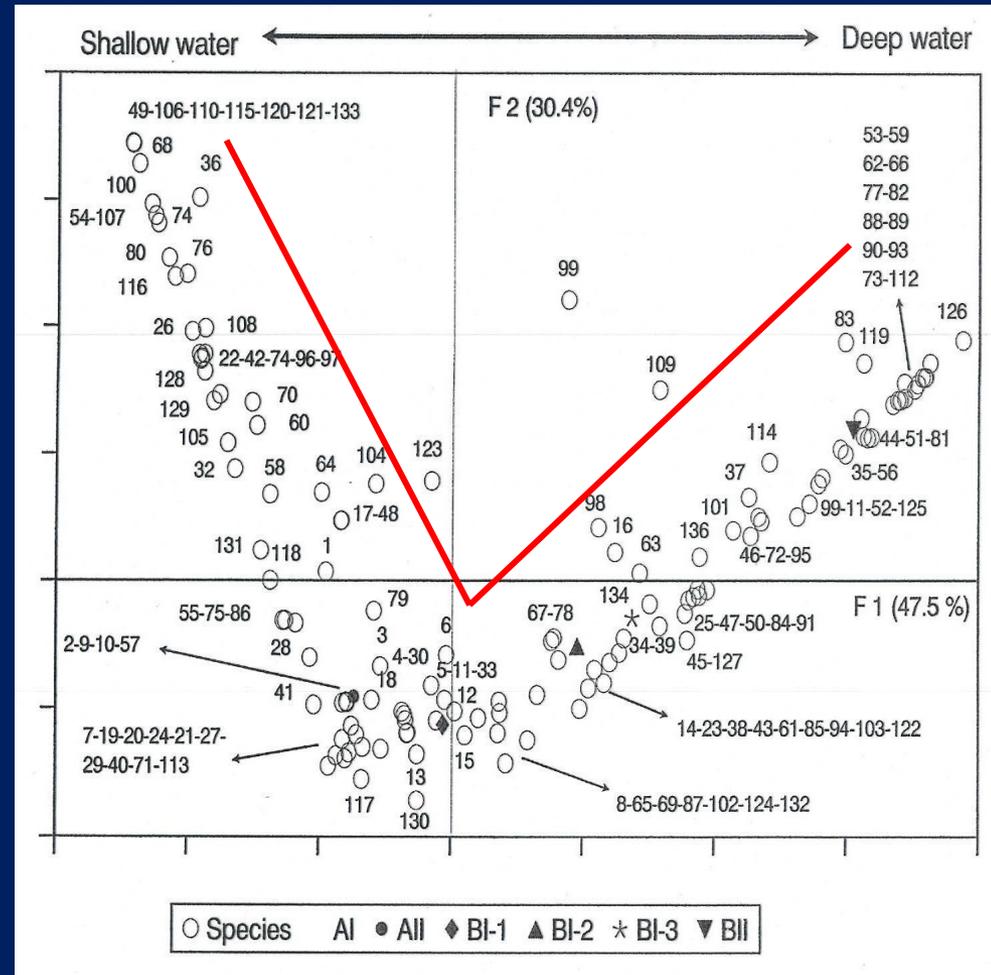
# Analisi fattoriale delle corrispondenze (CA)

E' una tecnica di ordinamento di grande interesse in ecologia. A differenza della PCA, consente di rappresentare simultaneamente sia punti-osservazione (stazioni) sia punti-descrittore (specie).



La CA è preferibile quando i dati sono meristici e, quindi, si applica a matrici di somiglianza ottenute con indici di correlazione.

Anche in questo caso si può avere una distorsione dell'ordinamento per “*effetto Guttman*”, che si manifesta con una distribuzione “a V” dei punti.



Per ovviare all'effetto Guttman, sono state sviluppate le “*detrended techniques*”.

# ORDINAMENTO

- PCA** - Principal Component Analysis
- RQA** - R-Q Analysis (Biplot) (ORLOCI)
- CA** - Correspondence Analysis
  - Reciprocal Averaging (HILL)
  - Analisi Fattoriale delle Corrispondenze (BENECRI)
- CVA** - Canonical Variate Analysis (Discriminant Analysis)  $\Rightarrow$  Tassonomie
- CCA** - Canonical Correlation Analysis  $\Rightarrow$  subsets
  - deti ambientali
  - deti faunistici
- PCO** - Principal Coordinate Analysis (Multidimensional Scaling)

Tecniche per risolvere l'effetto Guttman

## "EFFETTO GUTTMAN"

- DA** - Detrended Analysis
    - DPCA - Detrended Principal Component Analysis
    - DPCO - Detrended Principal Coordinate Analysis
  - NMO** - Non-Metric Ordinations
    - OS - Ordinal Scaling
    - NMDS - Non-Metric Multidimensional Scaling
- $\sigma(R_0, R_c) = 1 - \rho^2(R_0, R_c)$

# non -metric Multi-Dimensional Scaling (nMDS)

E' una tecnica esplorativa dei dati che permette di ottenere una rappresentazione di ' $n$ ' oggetti in uno spazio a ' $k$ ' dimensioni, derivata dalla similarità tra ciascuna coppia di oggetti.

La bontà dell'adattamento, cioè il grado di corrispondenza tra le distanze risultanti dalla mappa  $n$ MDS e quelle della matrice somiglianza di input, è misurata dalla *funzione di Stress*.

Minore è il valore di stress (es.  $< 0.10$ ), maggiore è la bontà dell'adattamento.

