C4C:Coding for Conservation
Community Ecology and Biodiversity Analysis Lecture

Week 4: July 25, 2022

Goals for this lecture

- To define a community
- To define the characteristics of communities
    - Abundance and relative abundance
    - Richness
    - Diversity
- To understand common methods of analysis for community diversity
    - Rarefaction
    - Permutational ANOVA
    - PCA
    - NMDS

Course materials are available at coding4conservation.org/syllabus

I. Basic concepts of community ecology

- **Ecology** is the study of the relationships between living organisms and their physical environment

I. Basic concepts of community ecology

- **Ecology** is the study of the relationships between living organisms and their physical environment

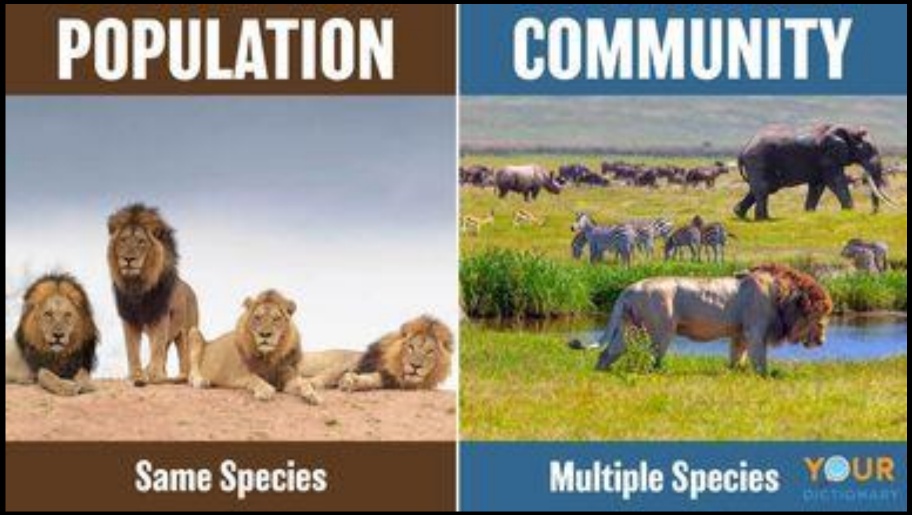
- **Community ecology** is the study of patterns in the diversity, abundance, and composition of species in communities, and the processes underlying these patterns

Definitions from: Vellend, M. 2010. The Quarterly Review of Biology

I. Basic concepts of community ecology

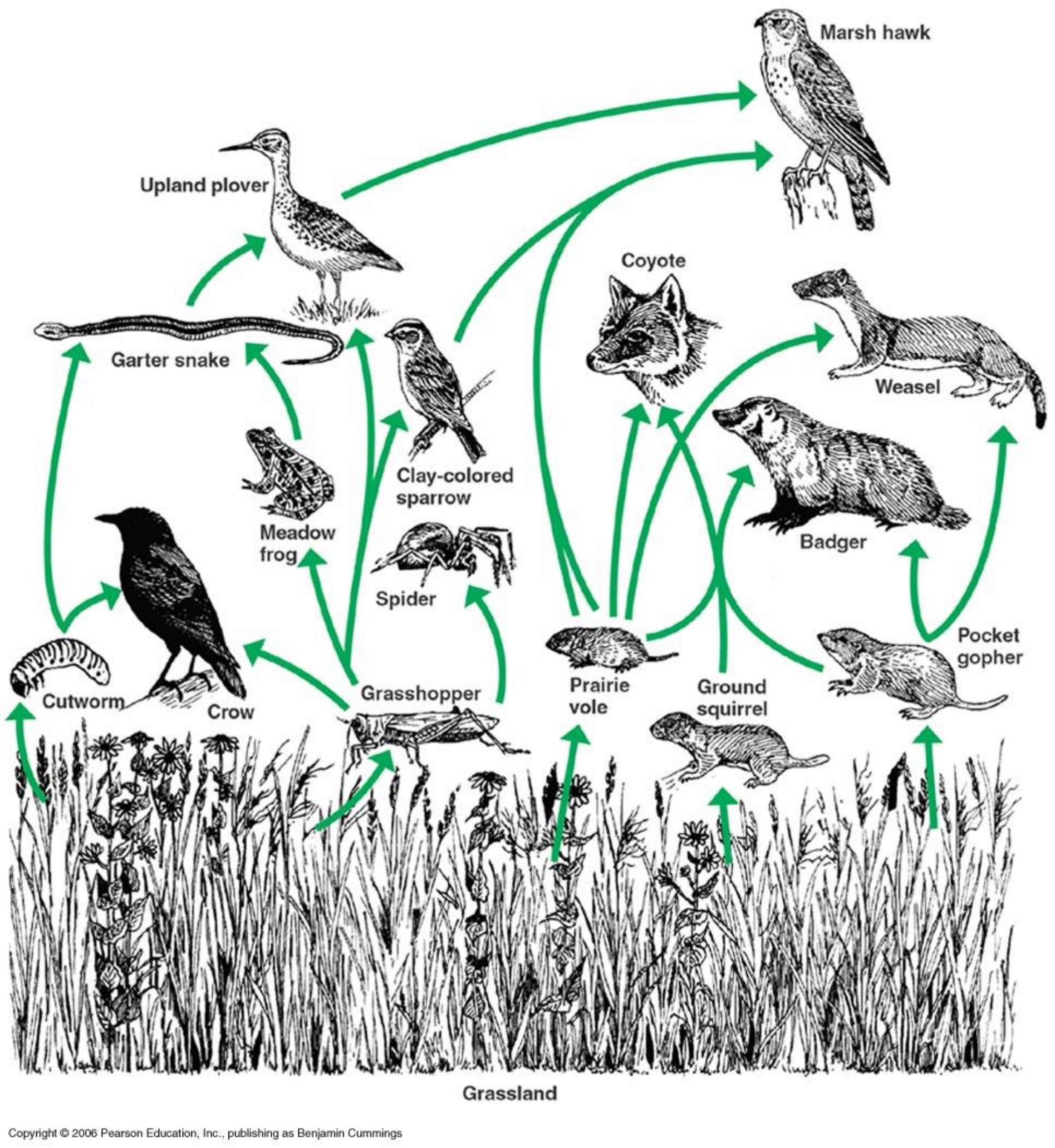
- **Ecology** is the study of the relationships between living organisms and their physical environment

- **Community ecology** is the study of the organization and functioning of communities
  - **Communities** = assemblages of interacting populations of species living within a particular area or habitat



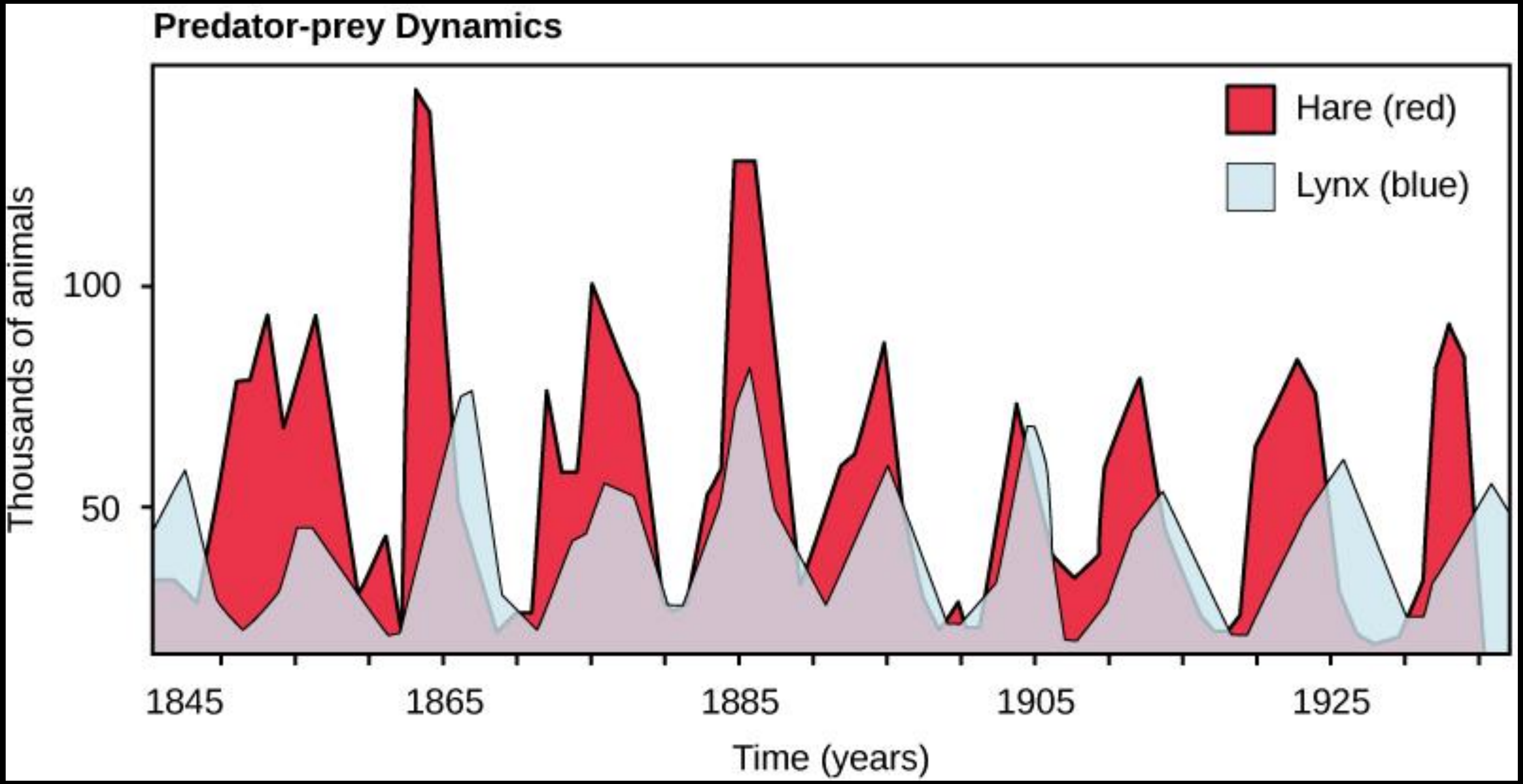
II. Interactions in Community Ecology

- Trophic webs – a succession of organisms in a community that are linked to each other through the transfer of energy and nutrients



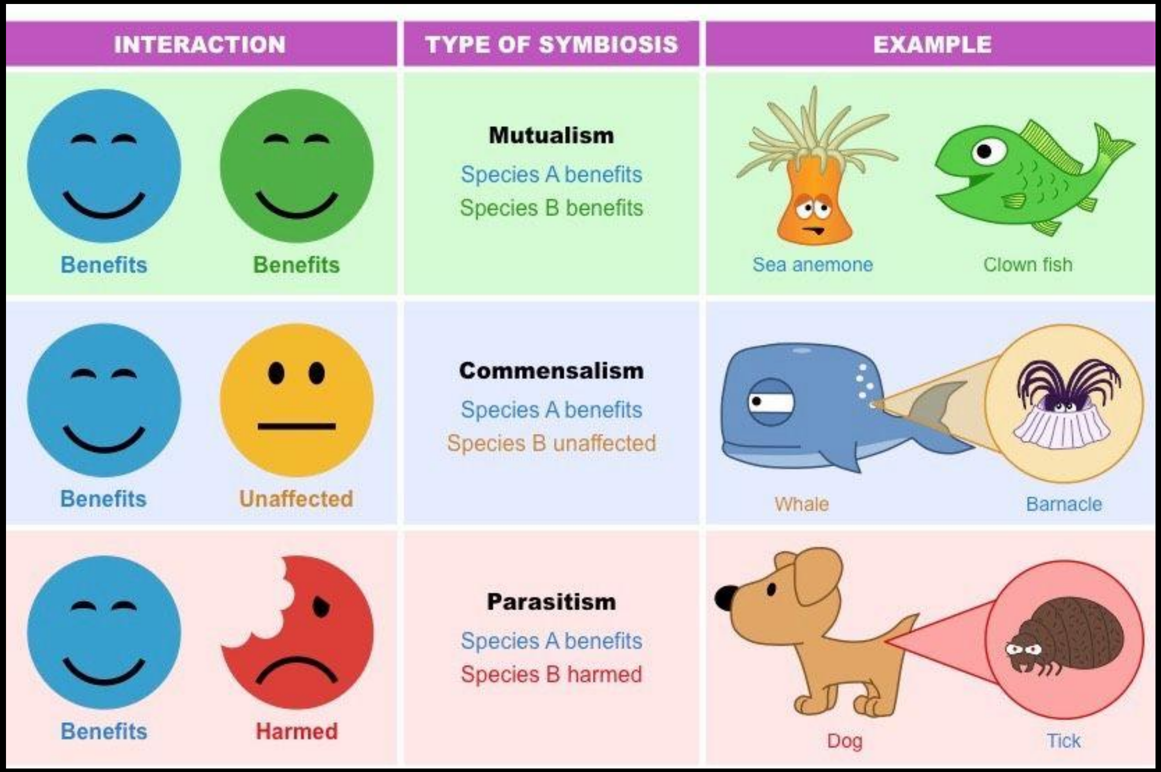
II. Interactions in Community Ecology

- Trophic webs
  - Predator-prey relationships



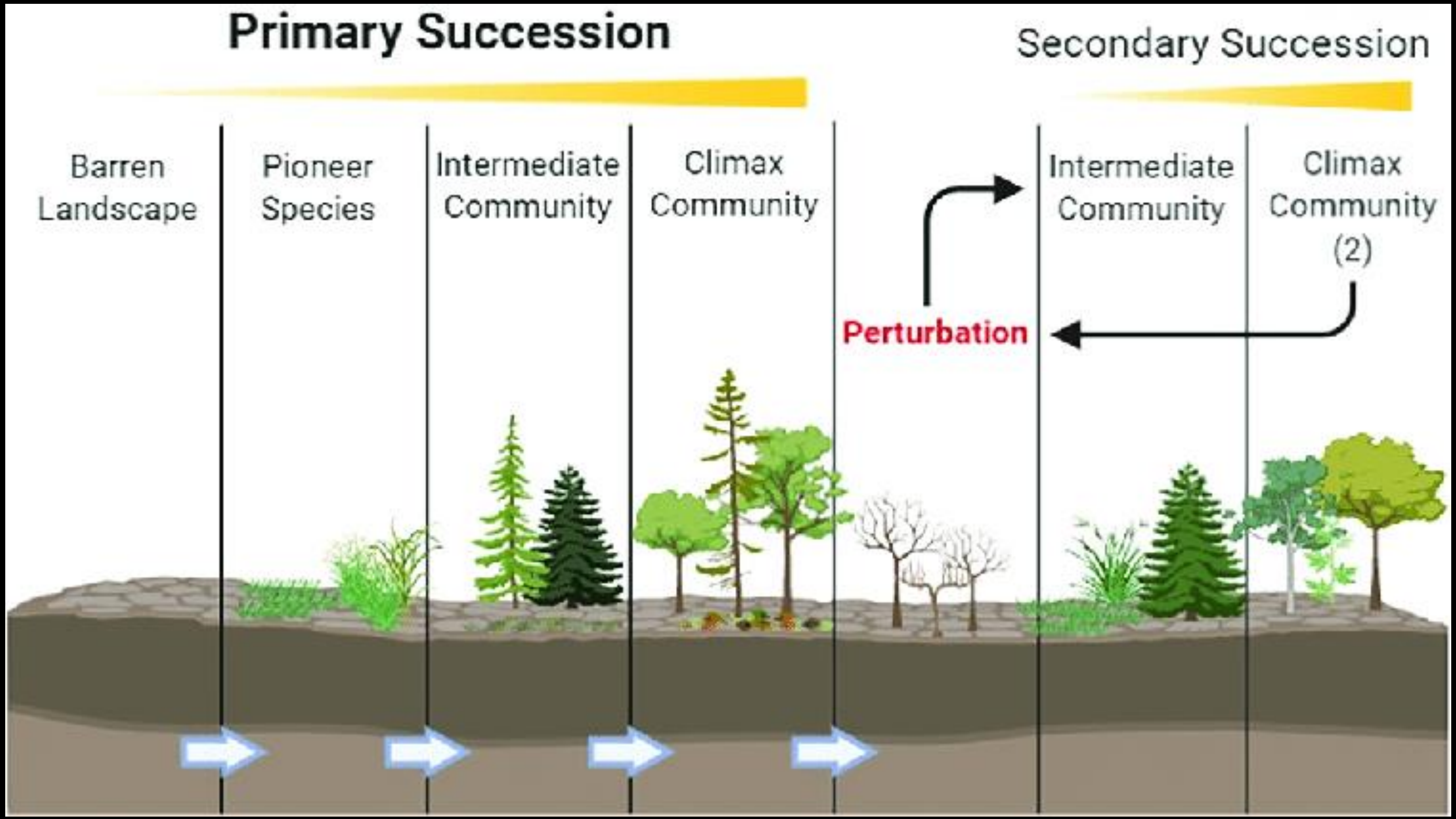
II. Interactions in Community Ecology

- **Symbiosis** – any type of close and long-term interaction between two different biological organisms
  - Mutualism
  - Commensalism
  - Parasitism



II. Interactions in Community Ecology

- **Ecological succession** – A series of progressive changes in the composition of an ecological community over time
  - Foundation species
  - Climax community
  - Keystone species
  - Invasive species
  - Pioneer species

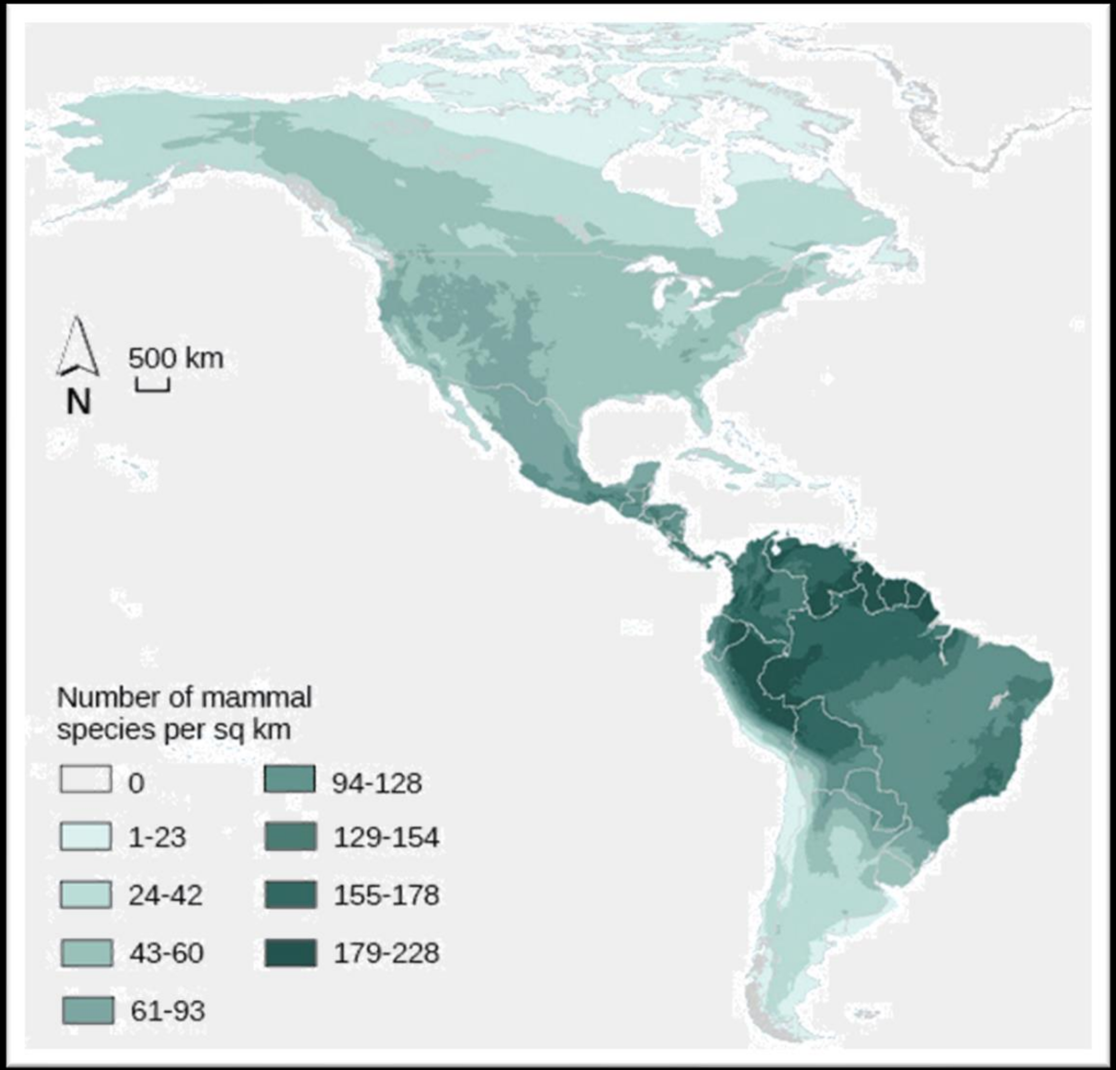


II. Interactions in Community Ecology

- **Biodiversity is the variety of life that can be characterized at various scales and influences community structure.**

- Highest species diversity tends to be found near the equator
    - Lots of solar energy (supporting high primary productivity), warm temperatures, large amounts of rainfall, and little seasonal change
- Lower species diversity tends to lie near the poles
    - Less solar energy and are colder, drier, and less amenable to life.

Global species richness for mammal species

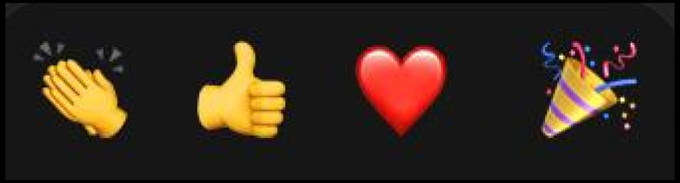


Summary and Check in

Community ecology is the study of patterns in the diversity, abundance, and composition of species in communities, and the processes underlying these patterns.

We can measure diversity, abundance, and composition to answer questions about species interactions

Going well?



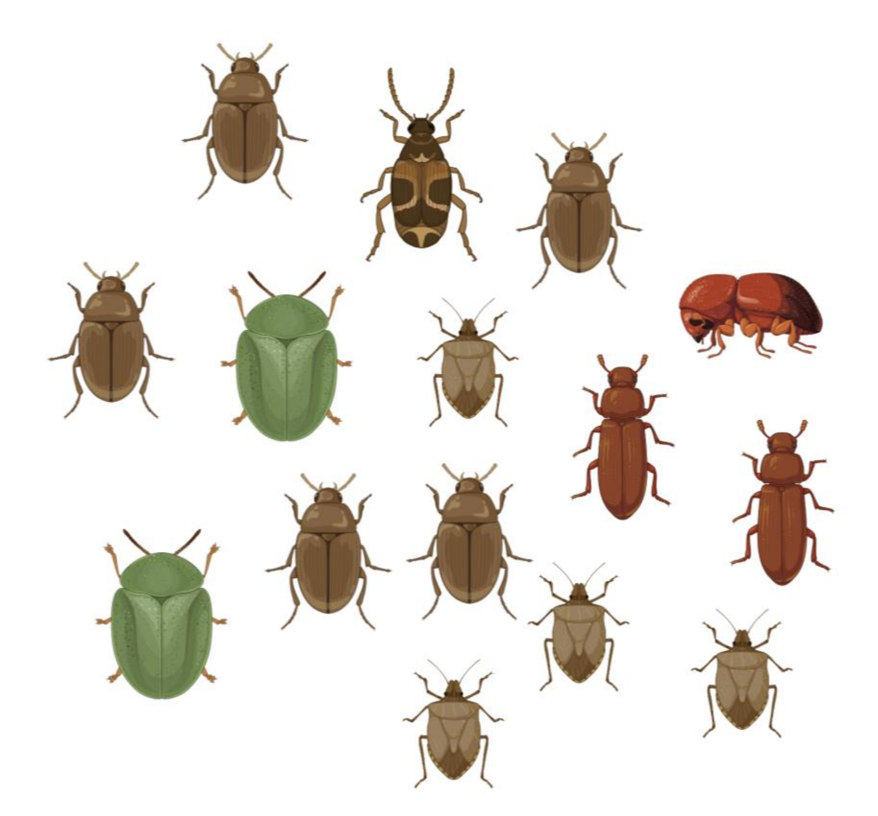
Not going so well?

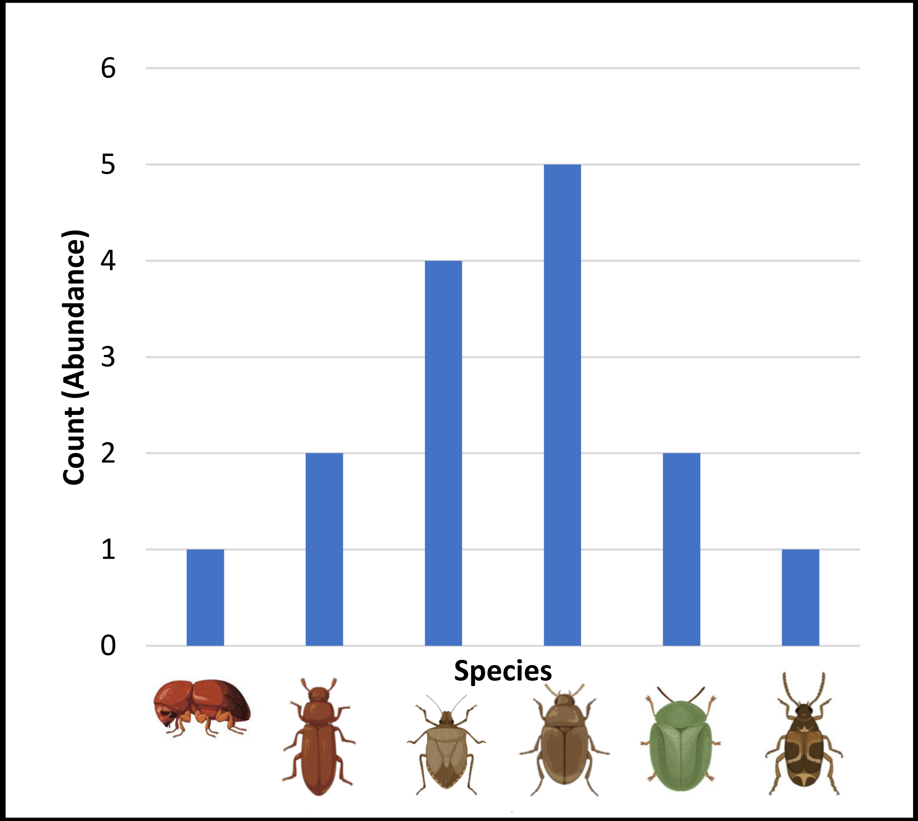
III. Abundance and Relative Abundance

- **Abundance** is total number of individuals per species living in a community

III. Abundance and Relative Abundance

- **Abundance** is total number of individuals per species living in a community
  - It is typically represented as a count of each species in a community



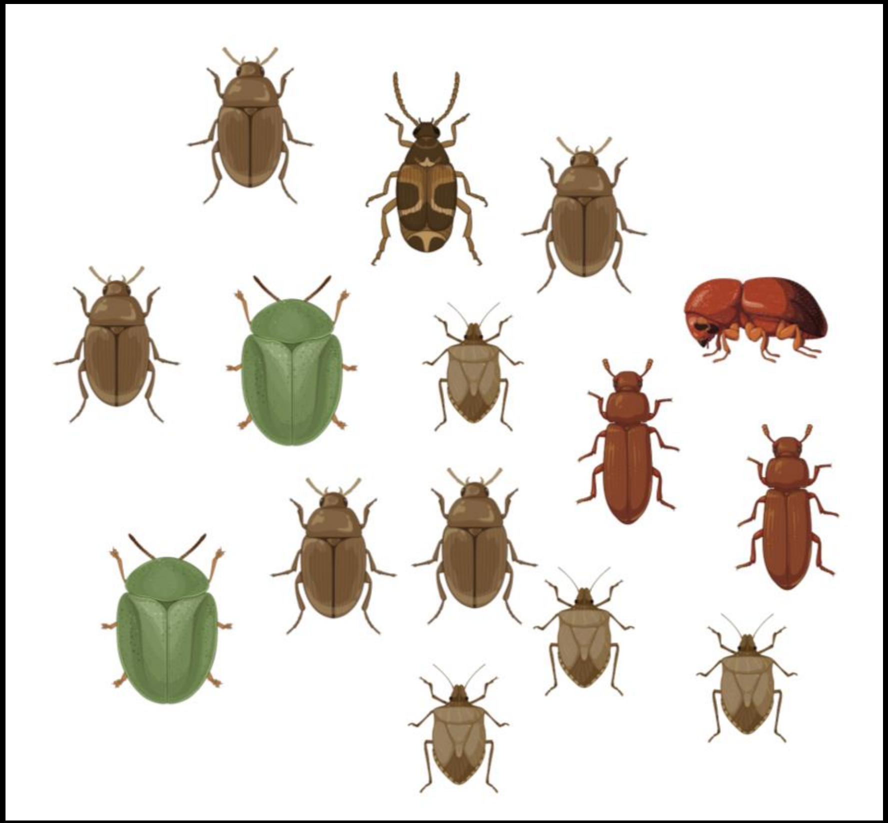


III. Abundance and Relative Abundance

- **Relative abundance** refers to how common or rare a species is relative to other species in a community

III. Abundance and Relative Abundance

- **Relative abundance** refers to how common or rare a species is relative to other species in a community
  - Typically, we see that most species are rare and relatively few species are abundant in a community



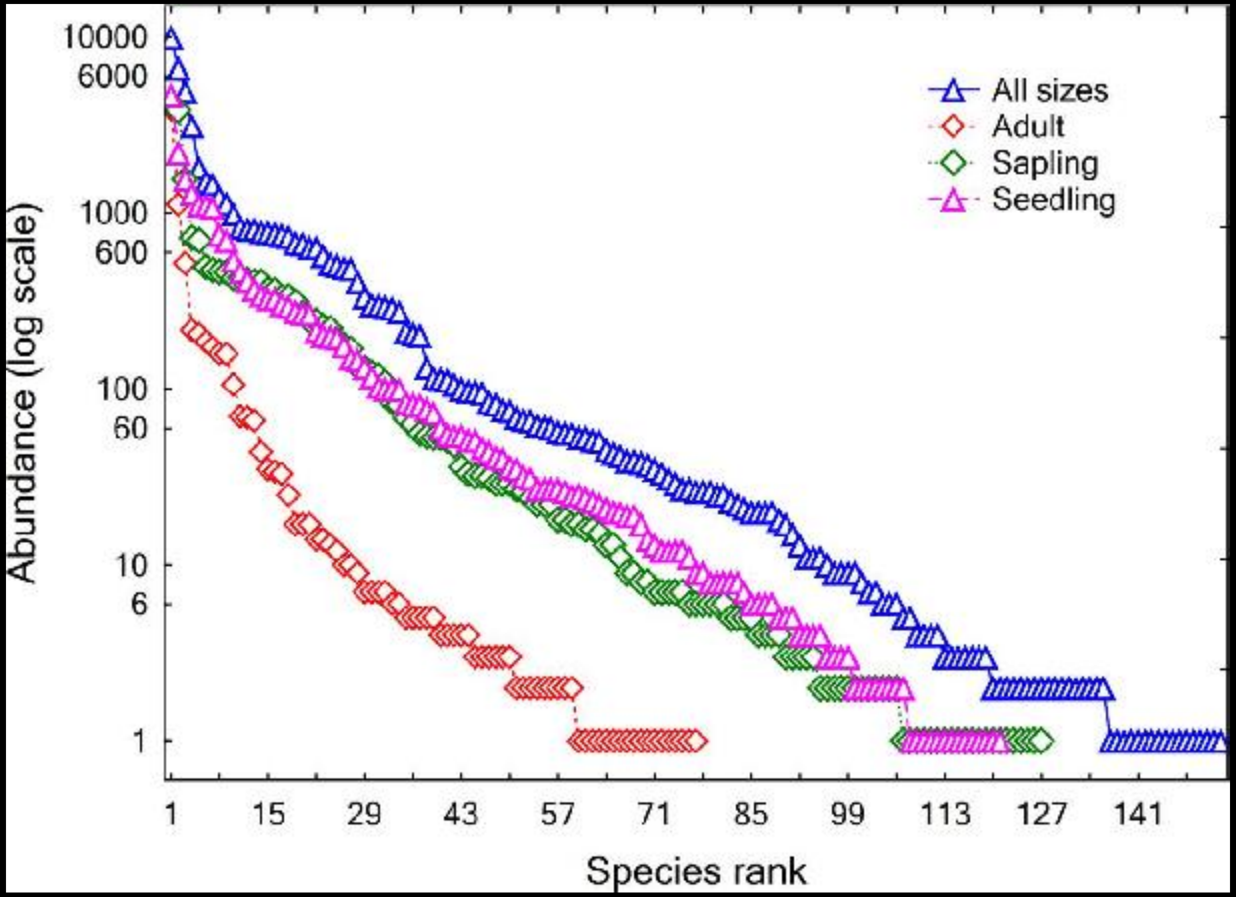
III. Abundance and Relative Abundance

- **Relative abundance** refers to how common or rare a species is relative to other species in a community
  - Typically, we see that most species are rare and relatively few species are abundant in a community

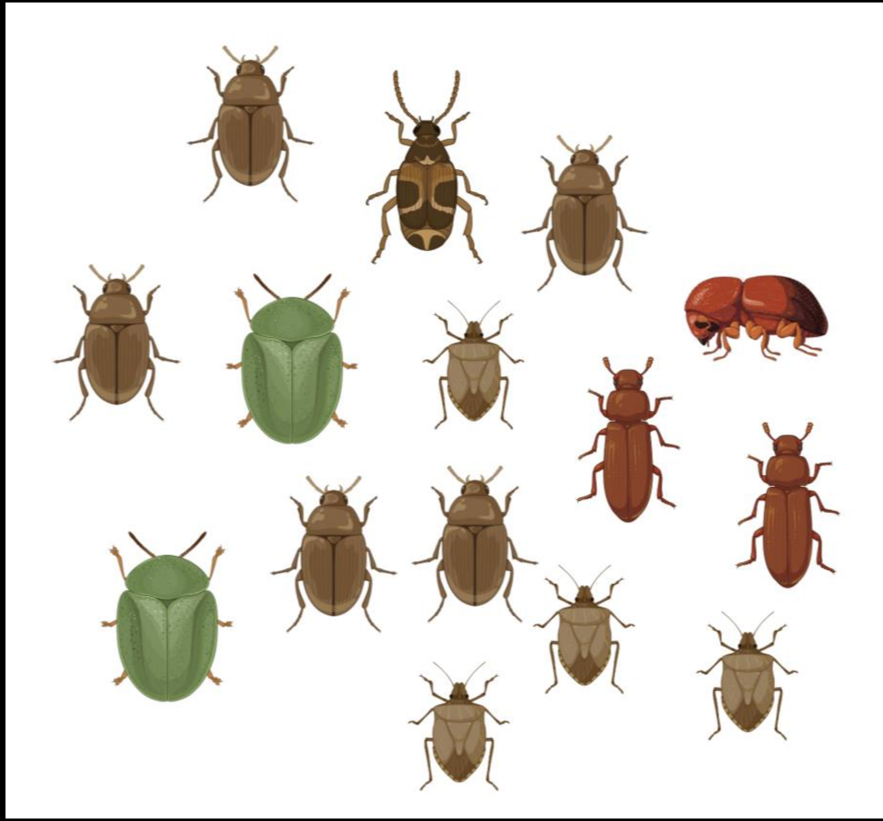
- Relative abundance is most commonly represented as
  - frequency histograms = number of species at a given abundance / log abundance
  - rank-abundance diagrams = log relative abundance / ranking of species from common to rare

III. Abundance and Relative Abundance

- A Whittaker plot or a rank abundance curve shows relative species abundance
  - Steep gradient = high ranking species have higher abundance; therefore, low **evenness.**
  - Shallow gradient = abundance is more similar across ranks; therefore, high **evenness.**

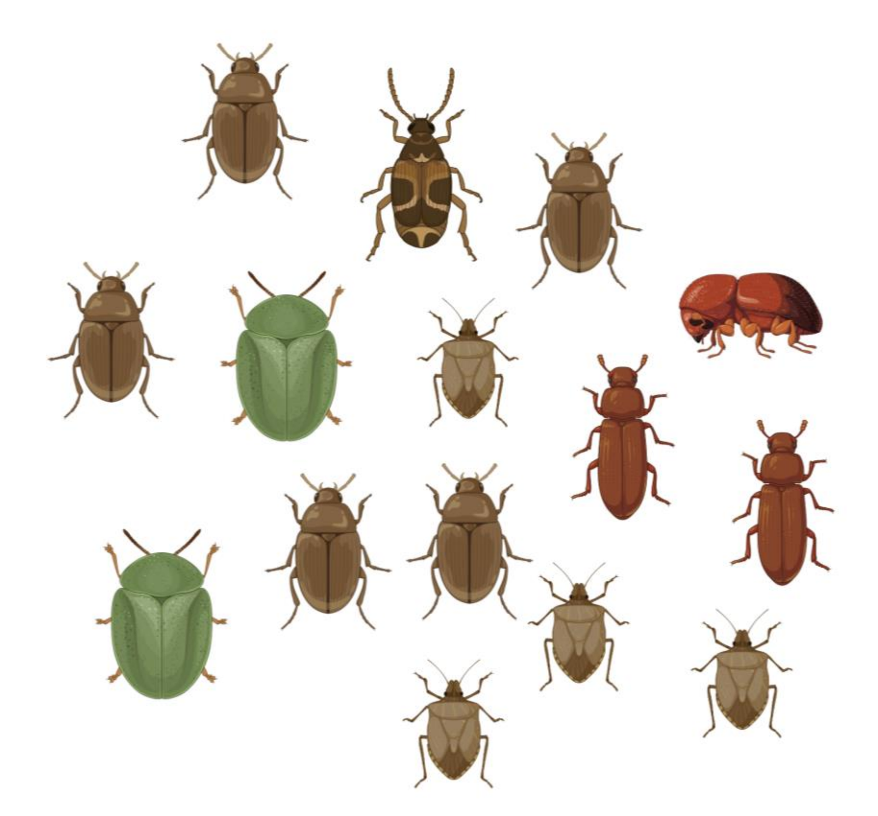


IV. Species Richness (S)



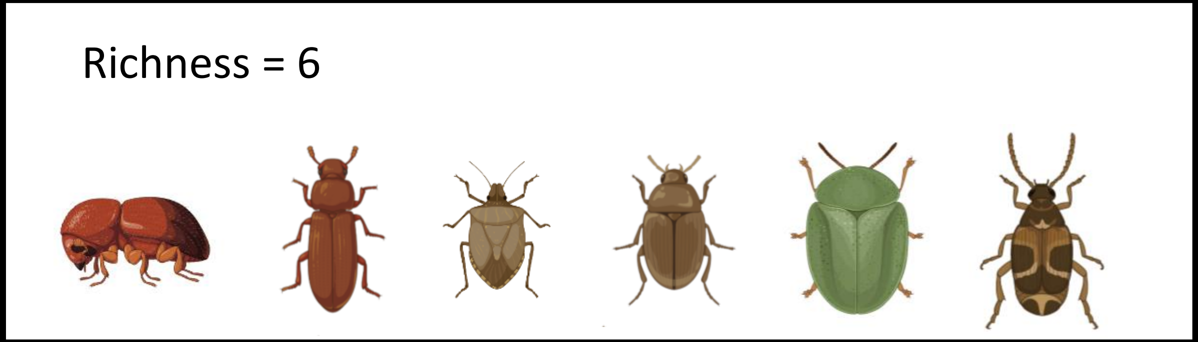
IV. Species Richness (S)

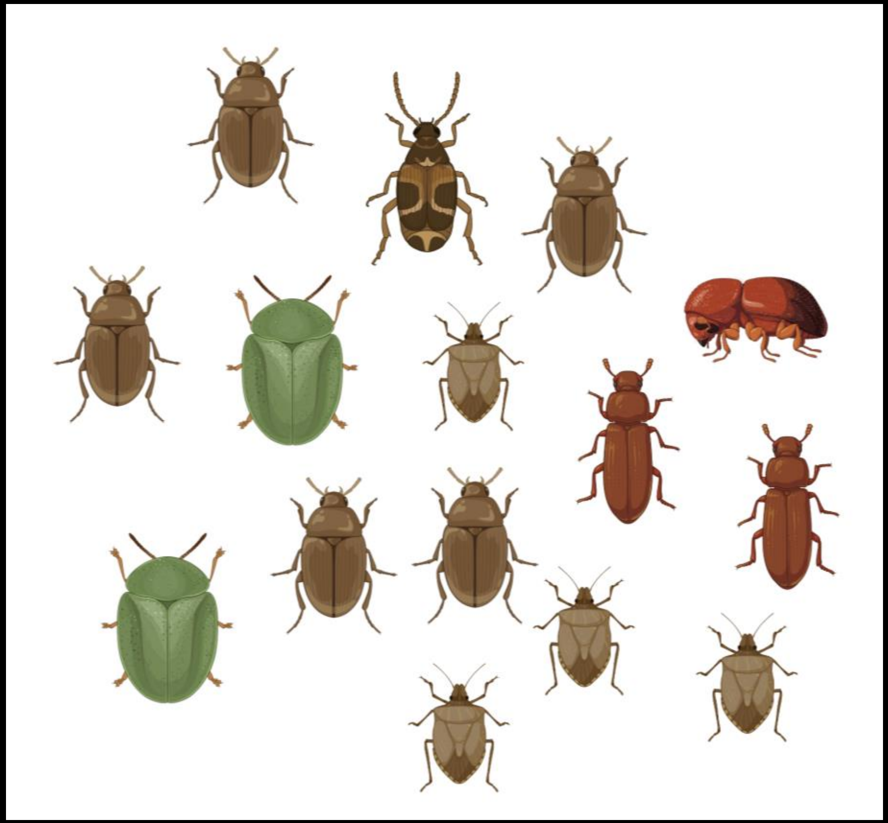
- **Species richness** is the number of species within a community



IV. Species Richness (S)

- **Species richness** is the number of species within a community





IV. Species Richness (S)

How can we know if we properly sampled all species present within a community?

IV. Species Richness (S)

How can we know if we properly sampled all species present within a community?

- Short answer is we can’t, but we can estimate it!

IV. Species Richness (S)

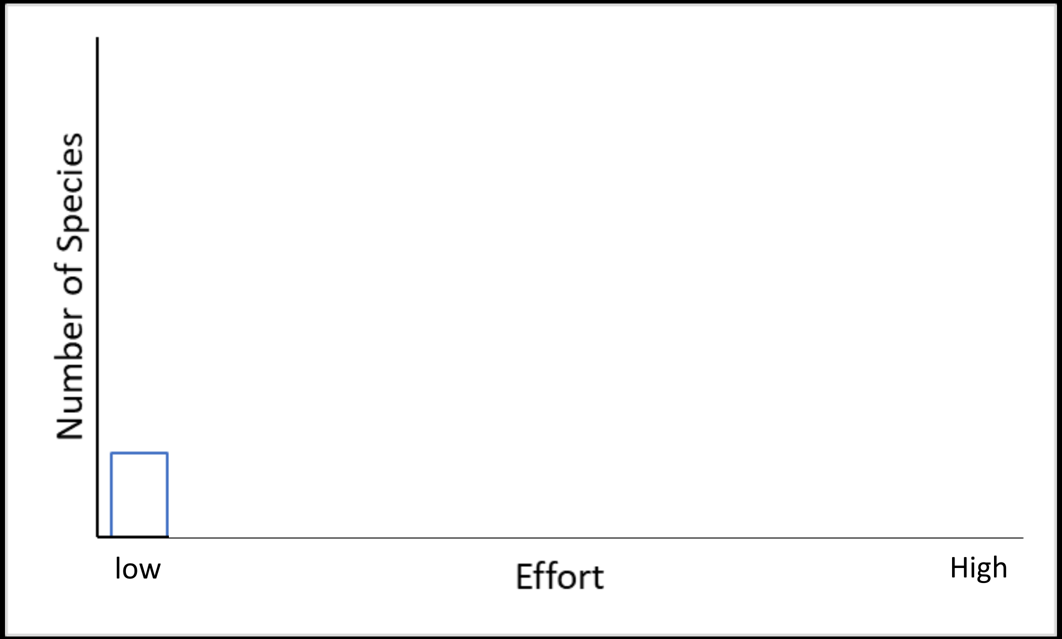
How can we know if we properly sampled all species present within a community?

- Short answer is we can't, but we can estimate it!

Species accumulation curves represent the cumulative number of species recorded in a community as a function of the cumulative effort spent searching for them.

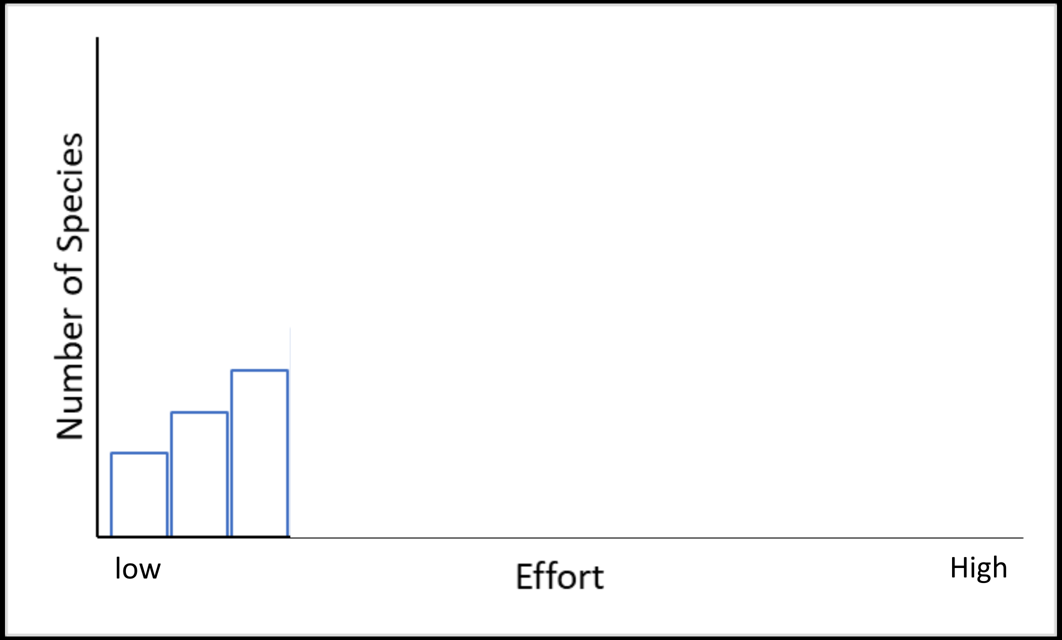
IV. Species Richness (S)

Species accumulation curves represent the cumulative number of species recorded in a community as a function of the cumulative effort spent searching for them.



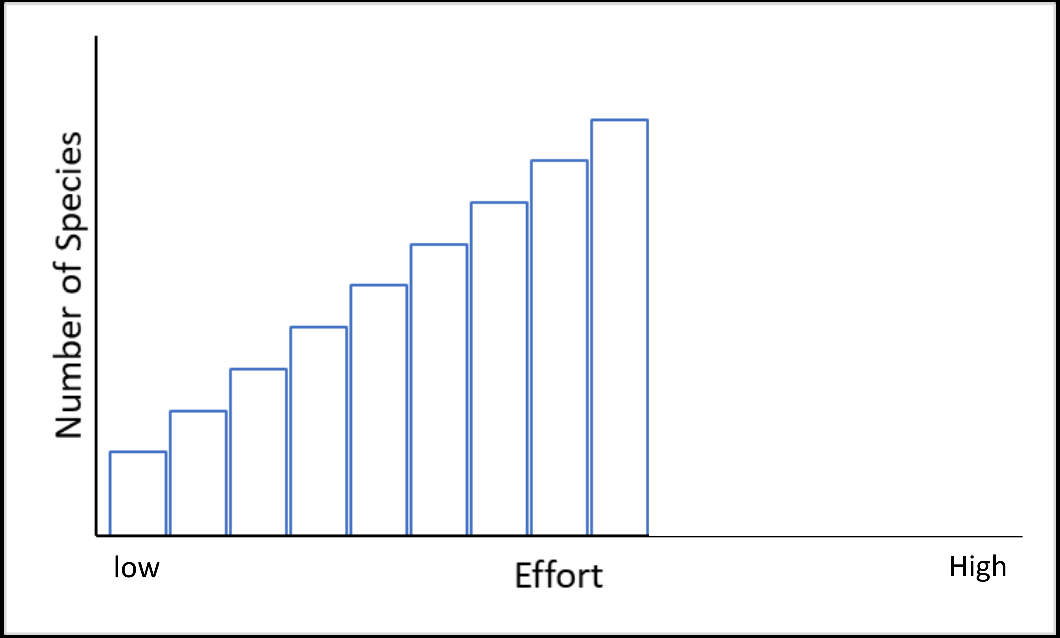
IV. Species Richness (S)

Species accumulation curves represent the cumulative number of species recorded in a community as a function of the cumulative effort spent searching for them.



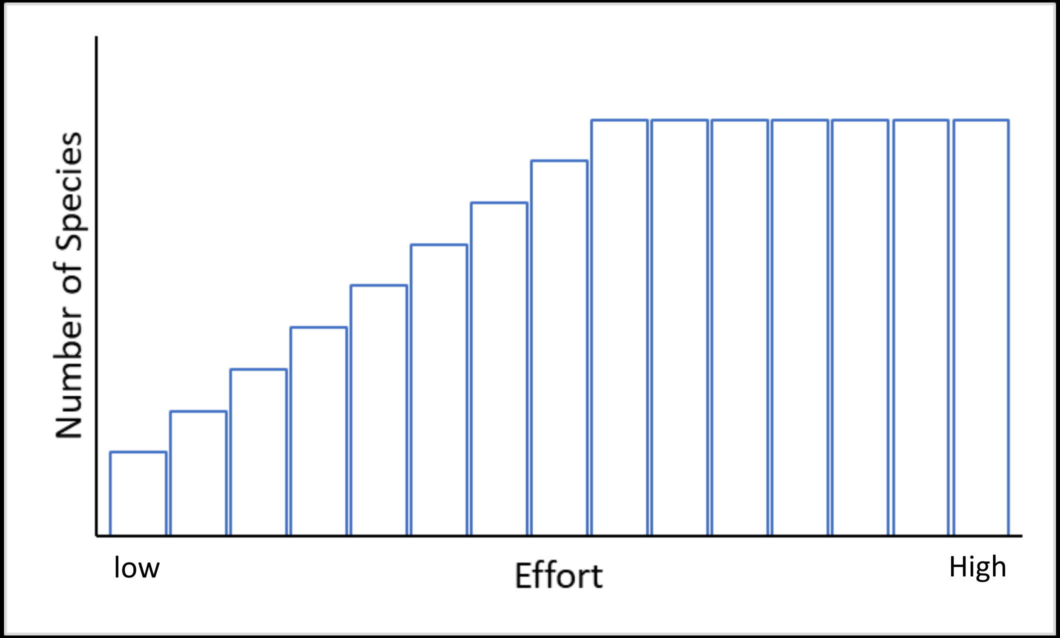
IV. Species Richness (S)

Species accumulation curves represent the cumulative number of species recorded in a community as a function of the cumulative effort spent searching for them.



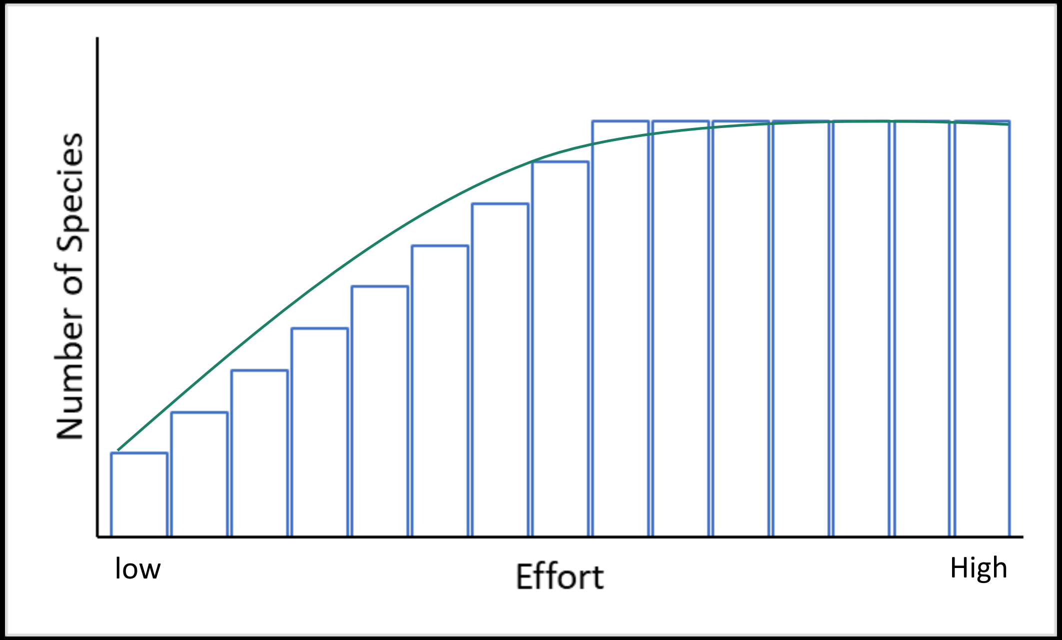
IV. Species Richness (S)

Species accumulation curves represent the cumulative number of species recorded in a community as a function of the cumulative effort spent searching for them.



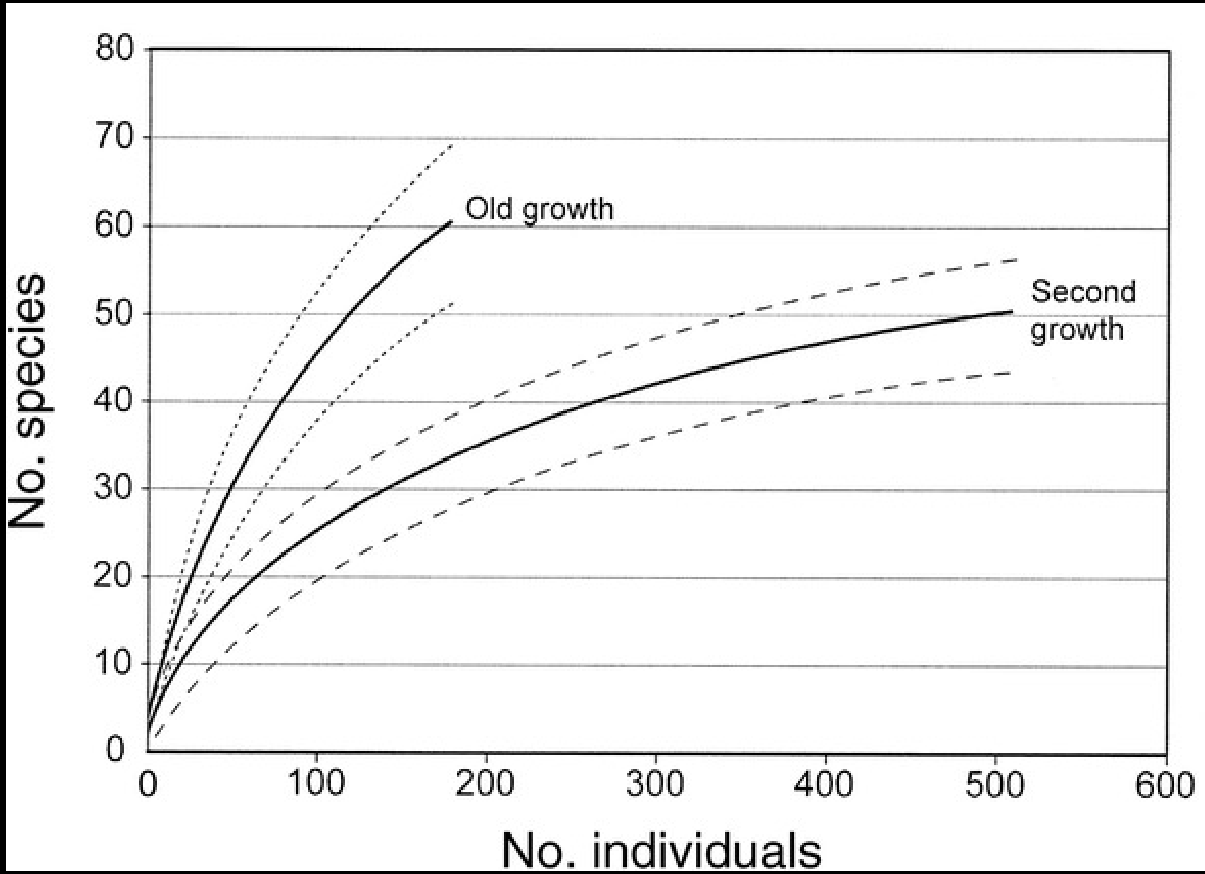
IV. Species Richness (S)

Species accumulation curves represent the cumulative number of species recorded in a community as a function of the cumulative effort spent searching for them.

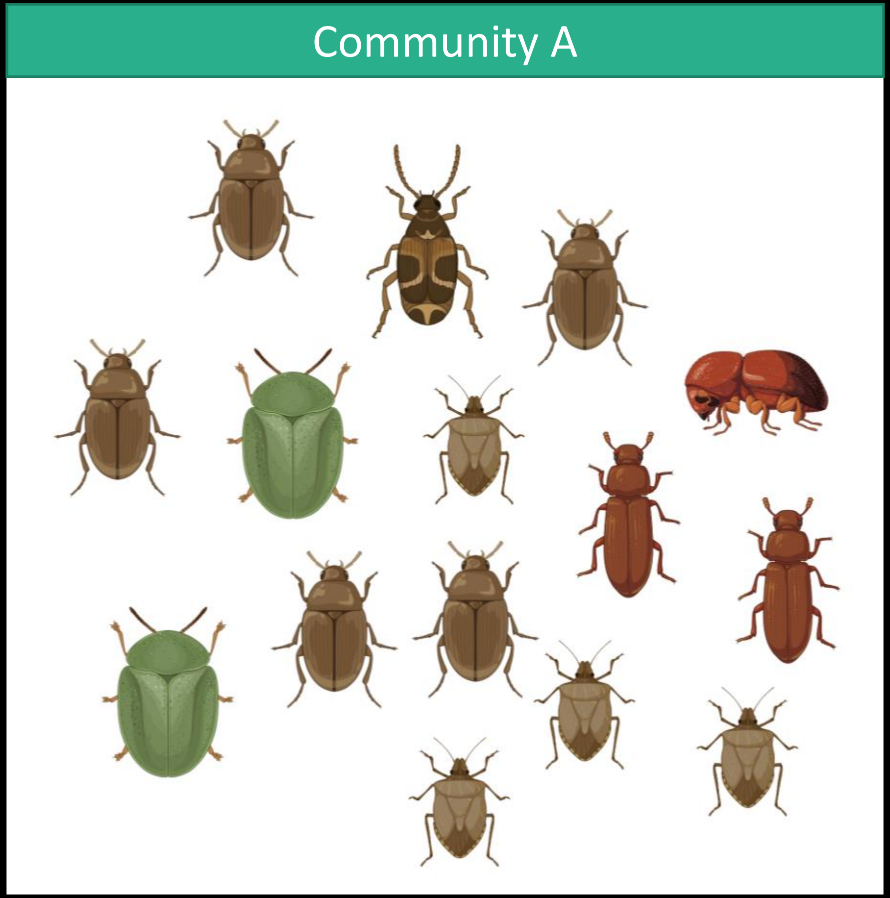


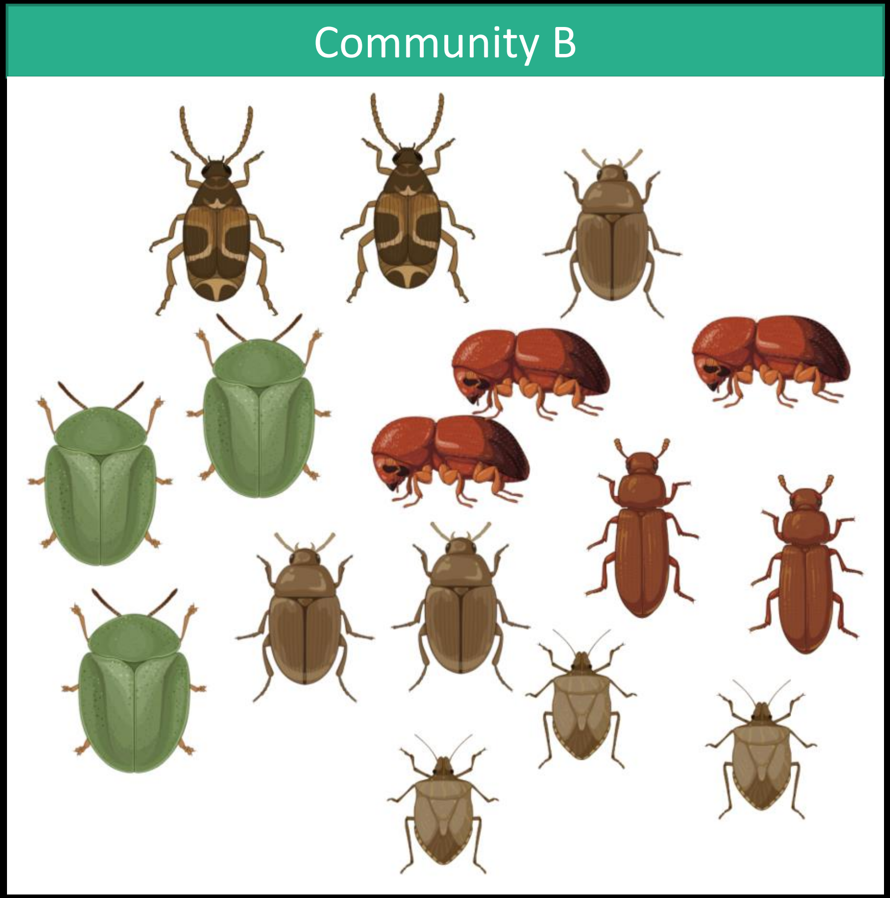
IV. Species Richness (S)

Species accumulation curves from two different forest types



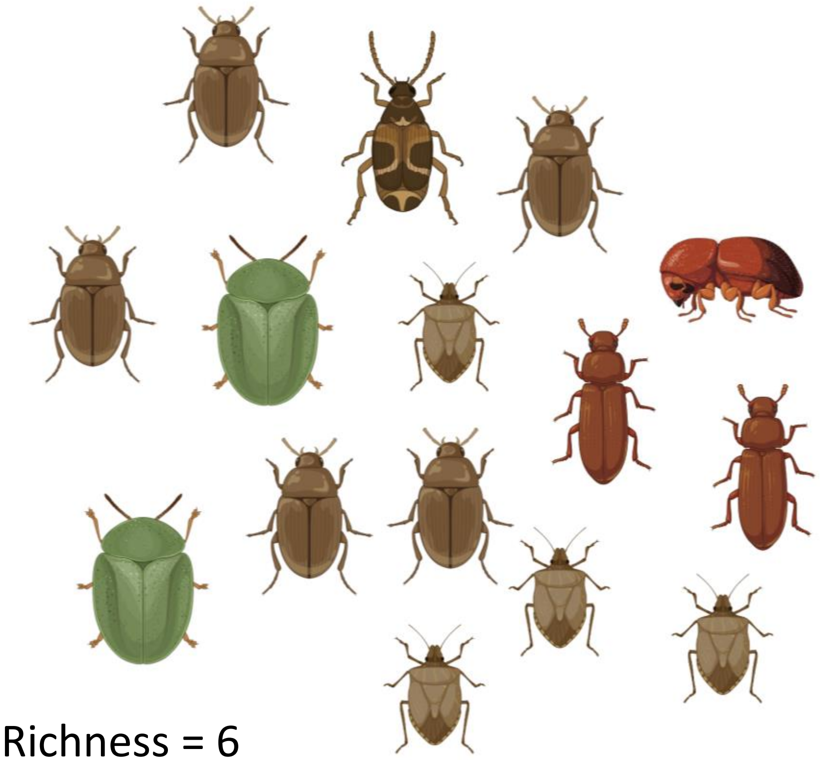
IV. Species Richness (S)





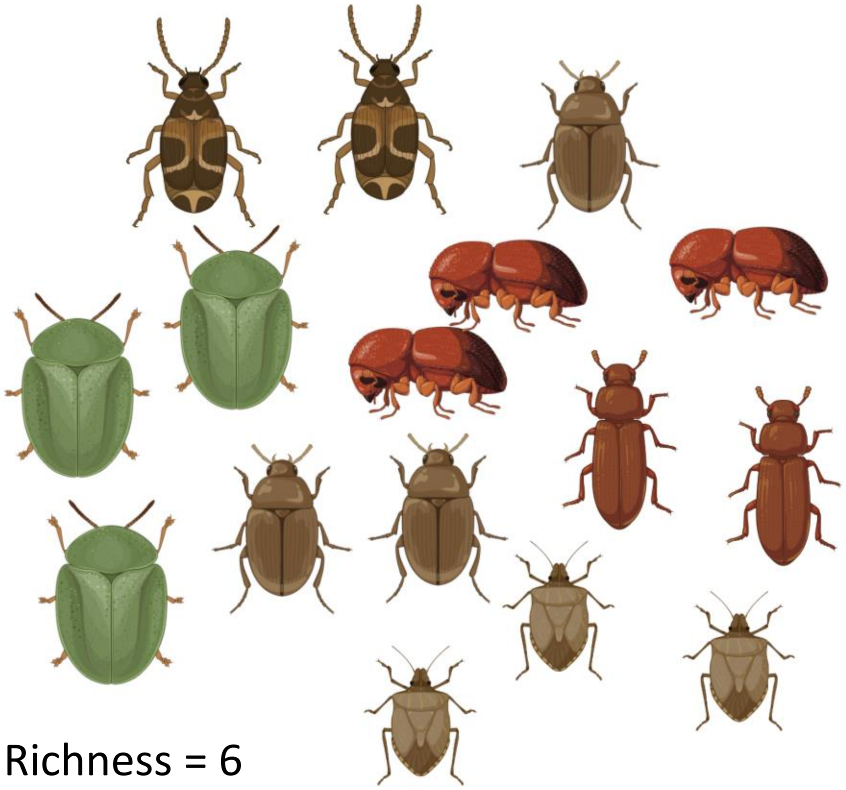
IV. Species Richness (S)

Community A



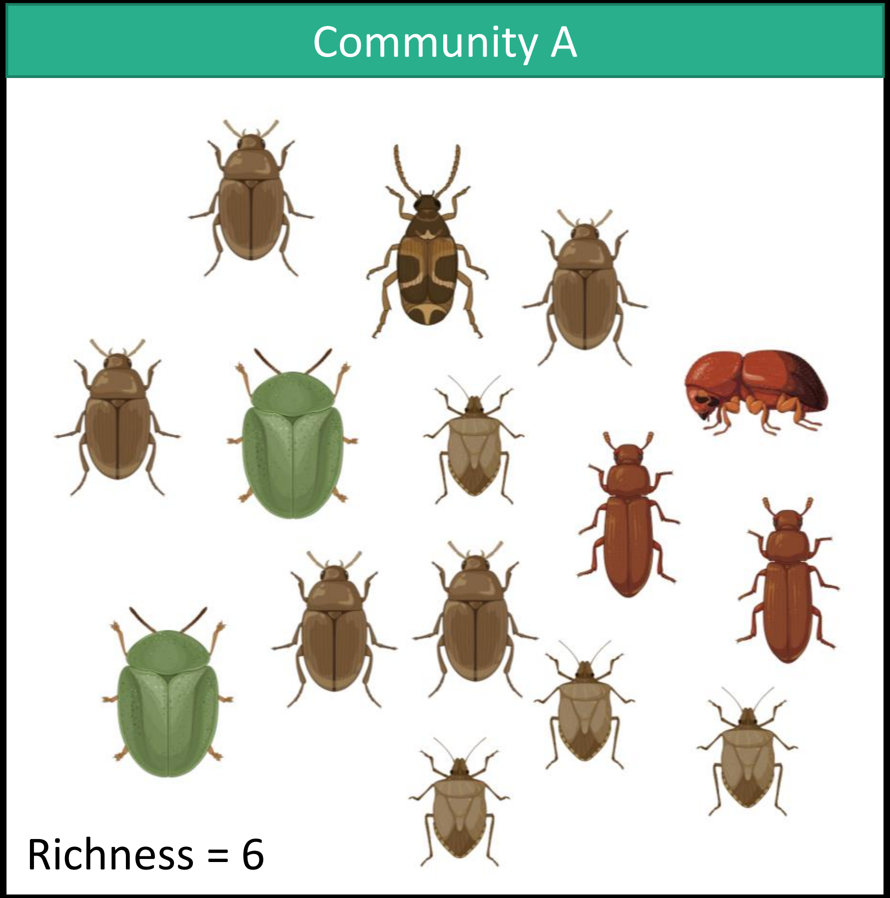
Richness = 6

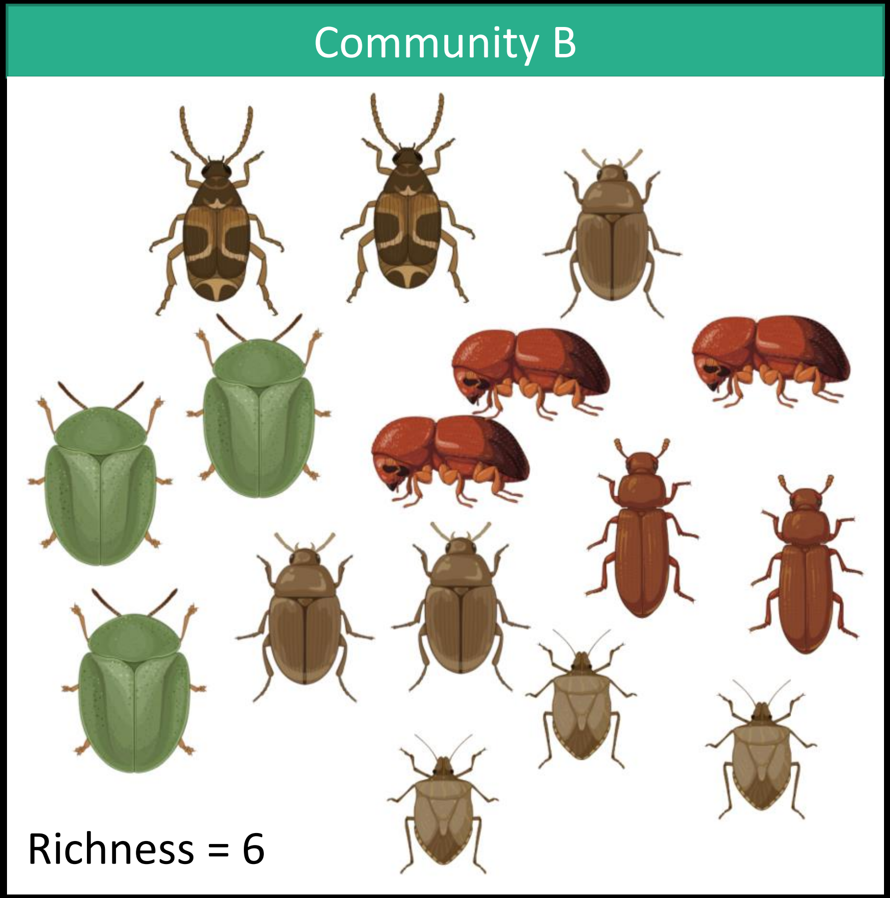
Community B



Richness = 6

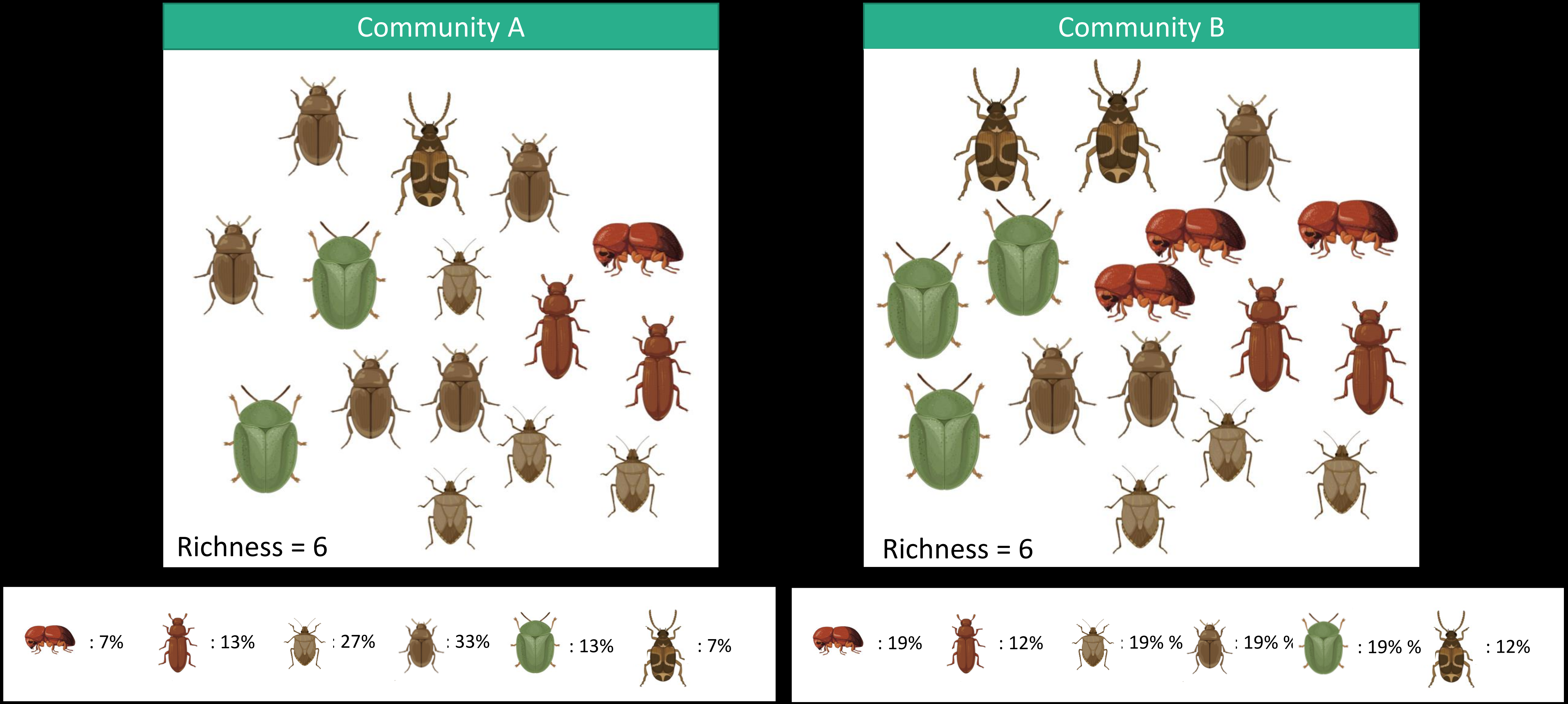
IV. Species Richness (S)



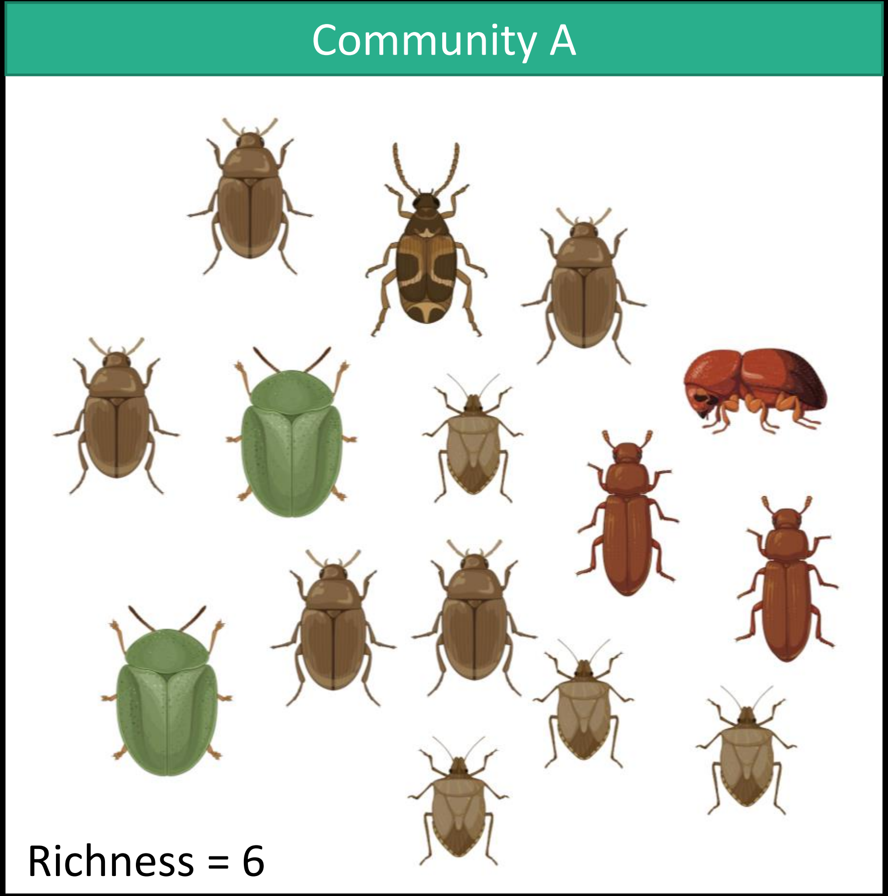


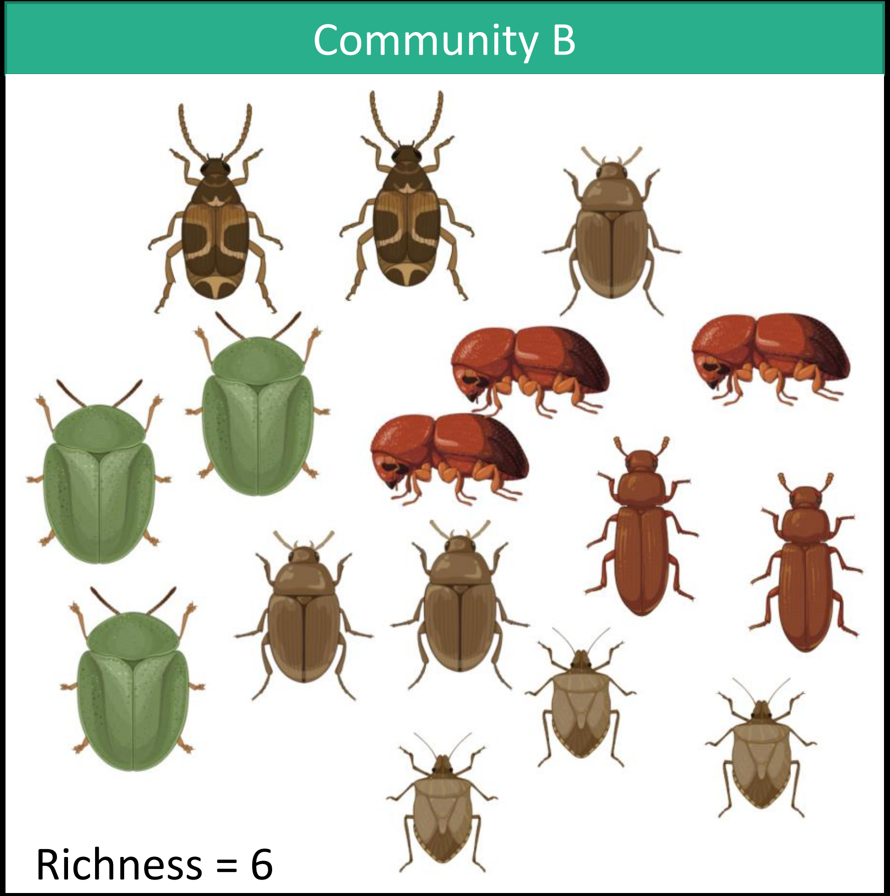
Is species diversity equal between these communities?

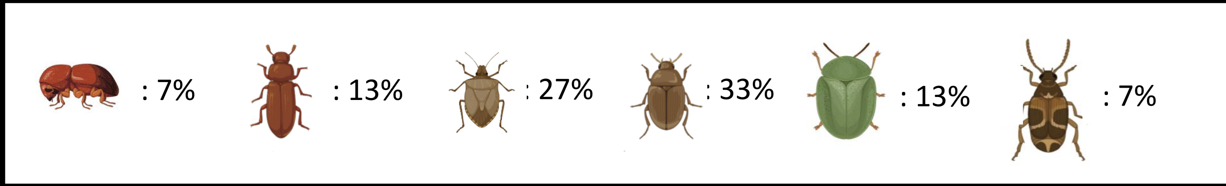
V. Evenness (J)

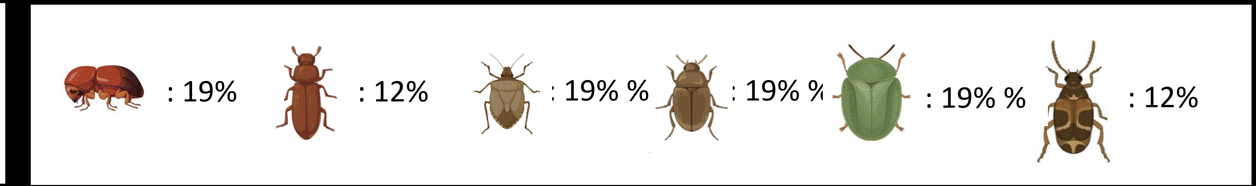

Community A
Richness = 6
Community B
Richness = 6
: 7%
: 13%
: 27%
: 33%
: 13%
: 7%
: 19%
: 12%
: 19% %
: 19% %
: 19% %
: 12%

V. Evenness (J)









- **Evenness** is a measure of the relative abundance in a community

V. Evenness (J)

- Evenness is a **diversity index, a measurement of biodiversity,** that lets one know about how equal two communities are numerically.

V. Evenness (J)

- Evenness is a **diversity index, a measurement of biodiversity,** that lets one know about how equal two communities are numerically.

  - Unlike Richness, evenness takes into account both the number of species present and their abundance.

$$Evenness = \frac{Number\ of\ species}{Total\ abundance}$$

V. Evenness (J)

- Evenness is a **diversity index, a measurement of biodiversity,** that lets one know about how equal two communities are numerically.
  - Unlike Richness, evenness takes into account both the number of species present and their abundance.

$$Evenness = \frac{Number\ of\ species}{Total\ abundance}$$

V. Evenness (J)

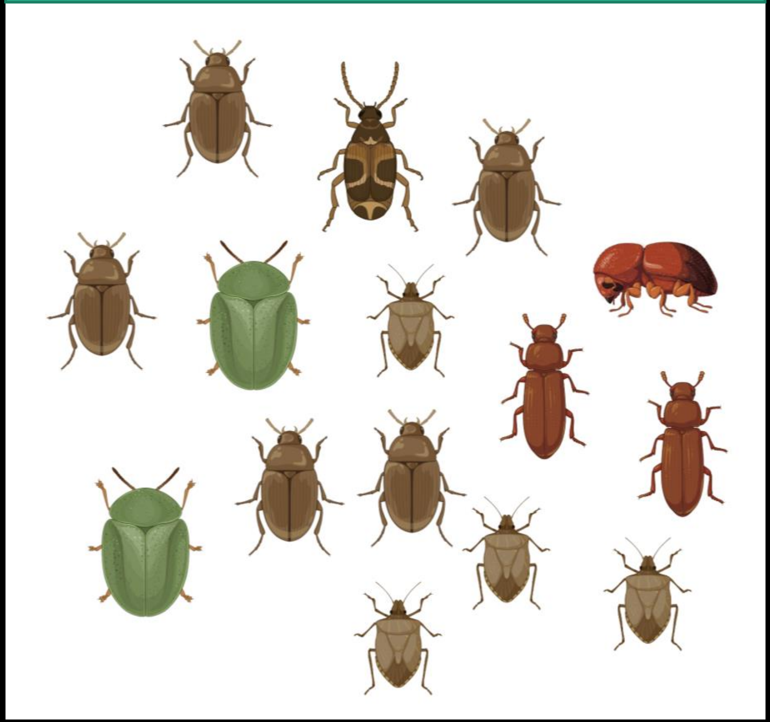
- Evenness is a **diversity index, a measurement of biodiversity,** that lets one know about how equal two communities are numerically.

$$J = \frac{H}{\ln(s)\ or\ Hmax}$$

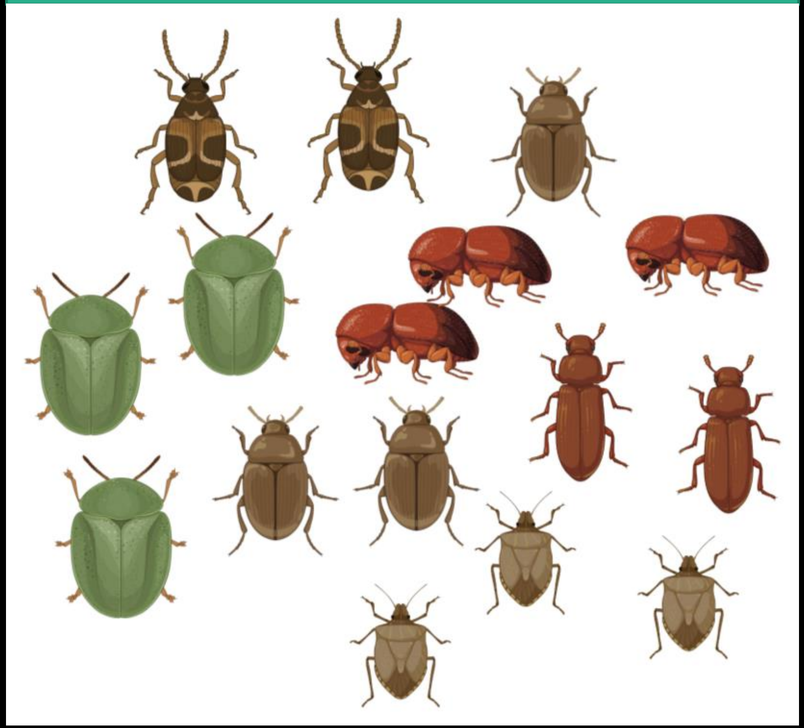
- J = Pielou's evenness
- H = Shannon Wiener diversity index
- Hmax = Maximum Shannon Wiener index
- s = number of species in the sample

Summary and Check in

Community A

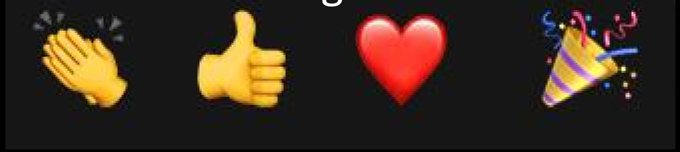


Community B



Abundance: Total count of each species in a community
Richness: Total number of species in a community
Evenness: How equal are the number of species between two communities?

Going well?



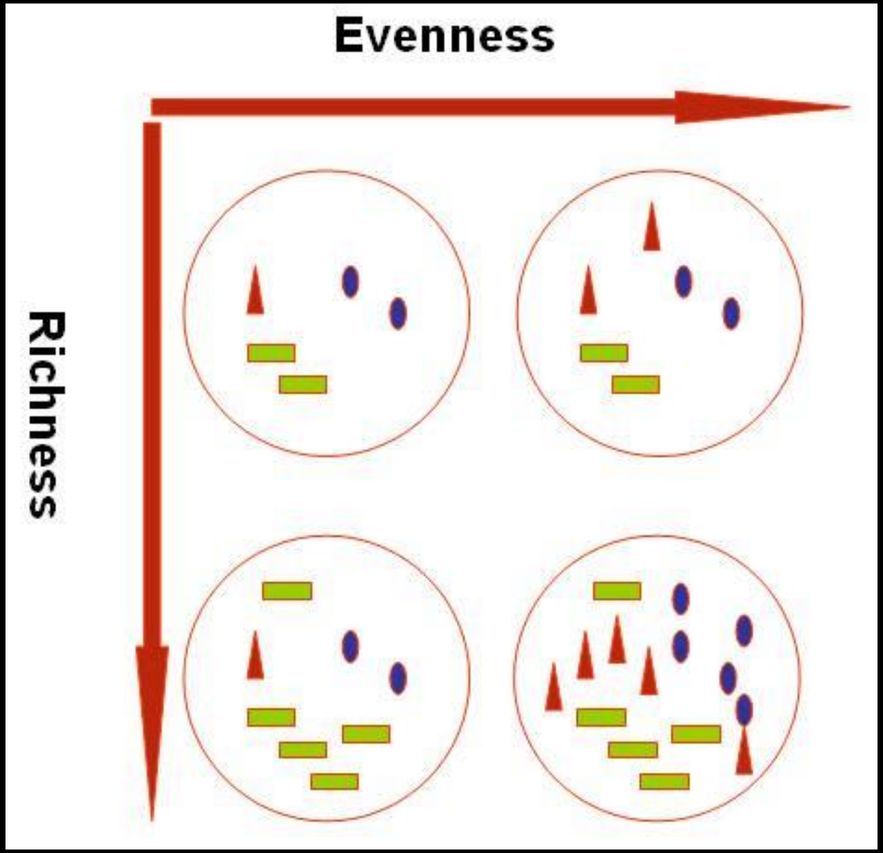
Not going so well?

VI. Diversity

- **Diversity** is measure of community complexity

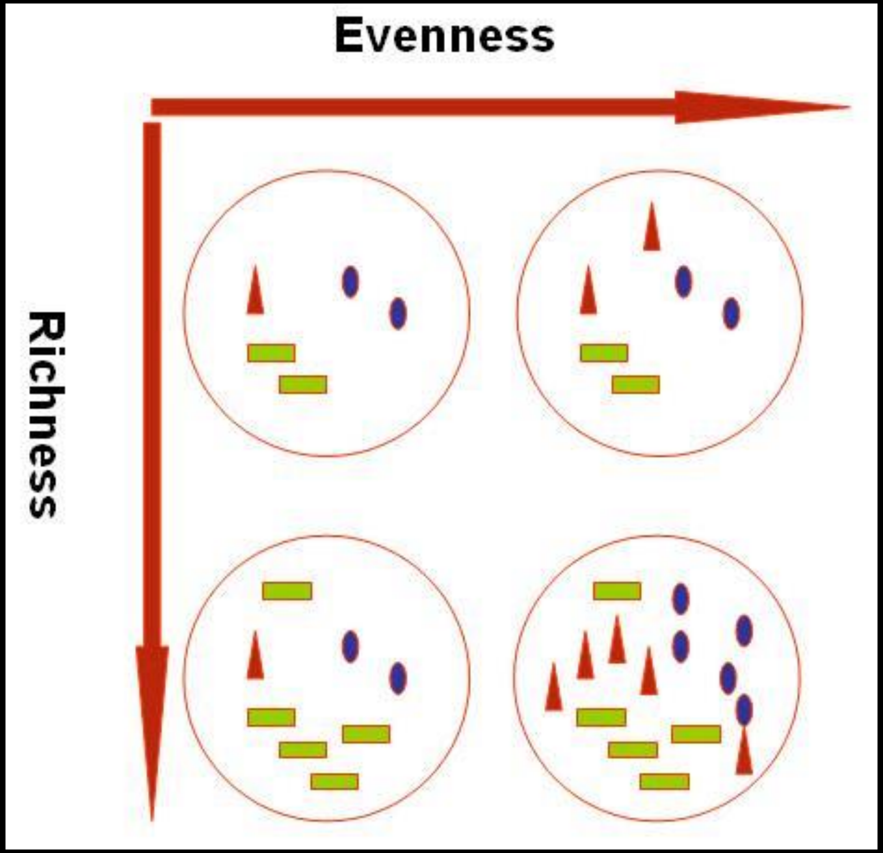
VI. Diversity

- **Diversity** is measure of community complexity
  - It is a function of both species richness (number of species in a community) and their relative abundance (species evenness)



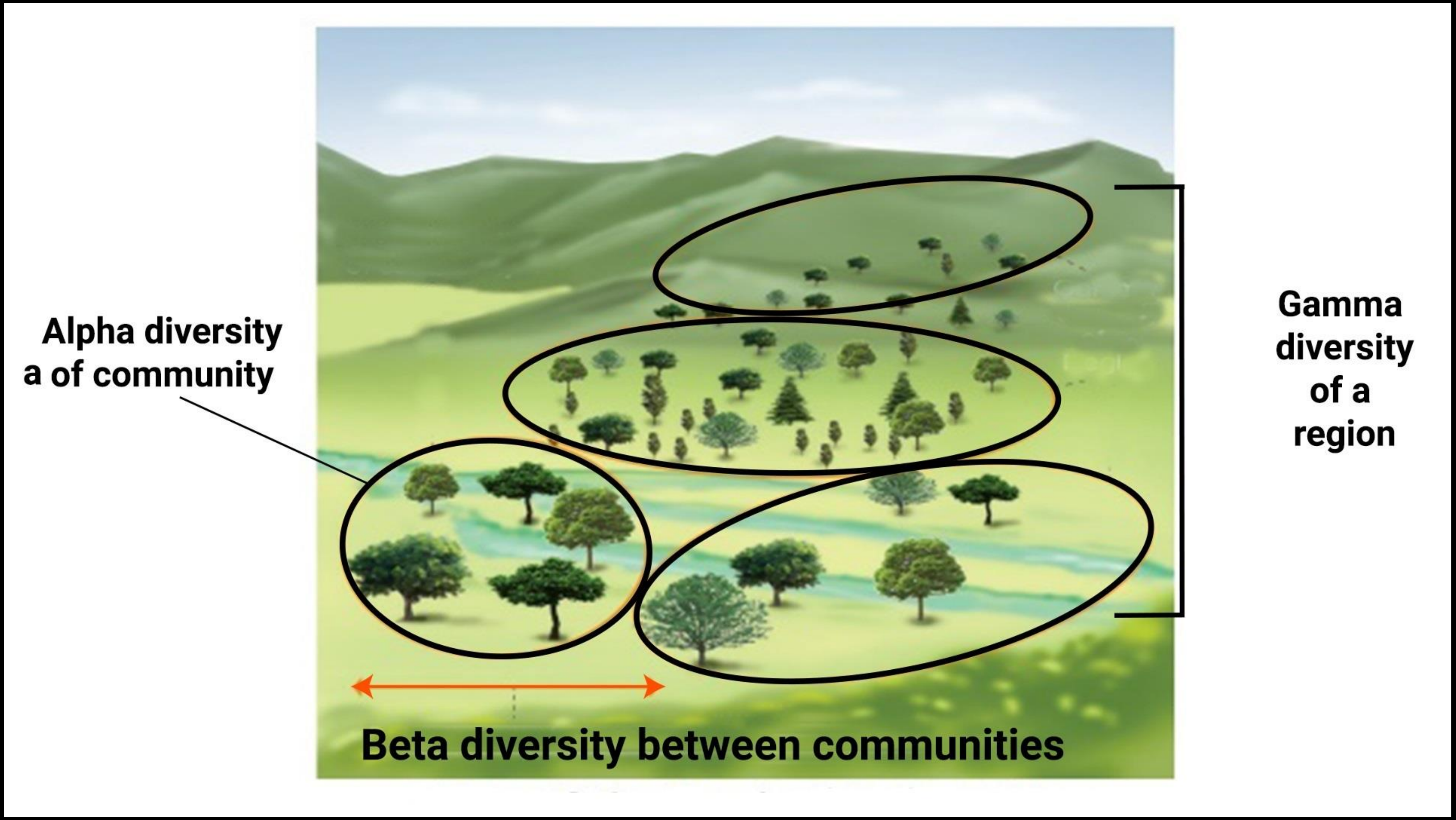
VI. Diversity

- **Diversity** is measure of community complexity
  - It is a function of both species richness (number of species in a community) and their relative abundance (species evenness)



In general, more diverse ecological communities are more stable (better able to recover after disturbance)

V. Diversity is measured at different scales

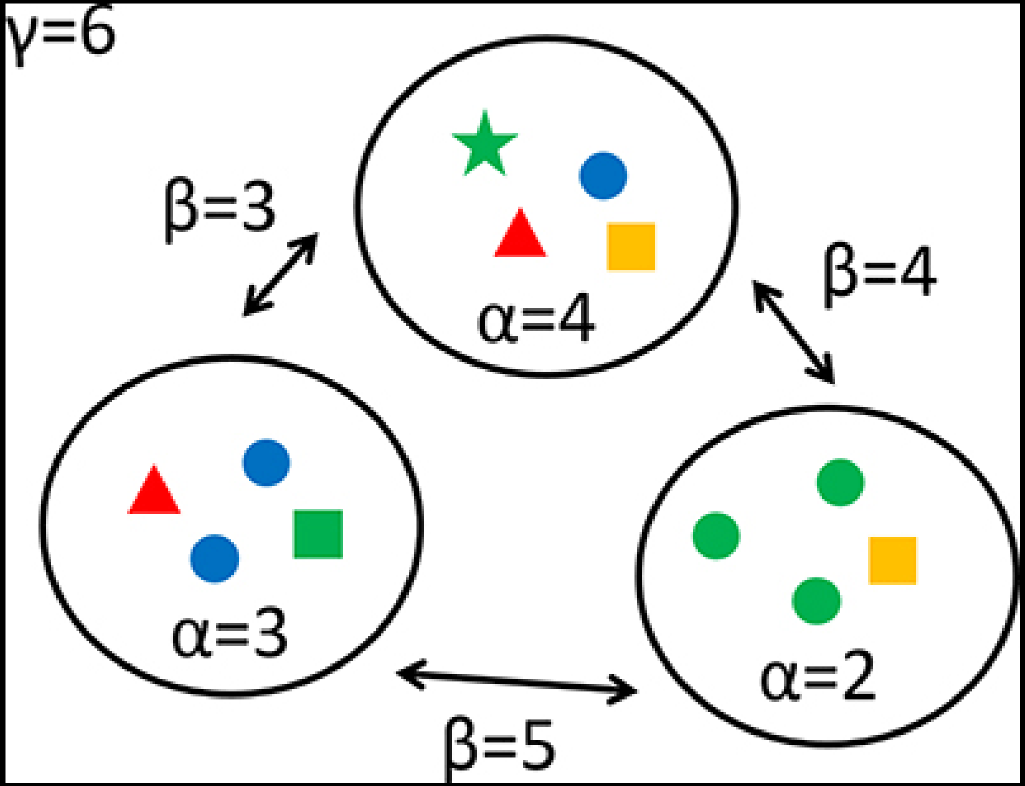


VII. Gamma Diversity (γ-diversity)

- **Gamma Diversity** is the species diversity in a landscape
  - It is determined by alpha diversity and beta diversity.

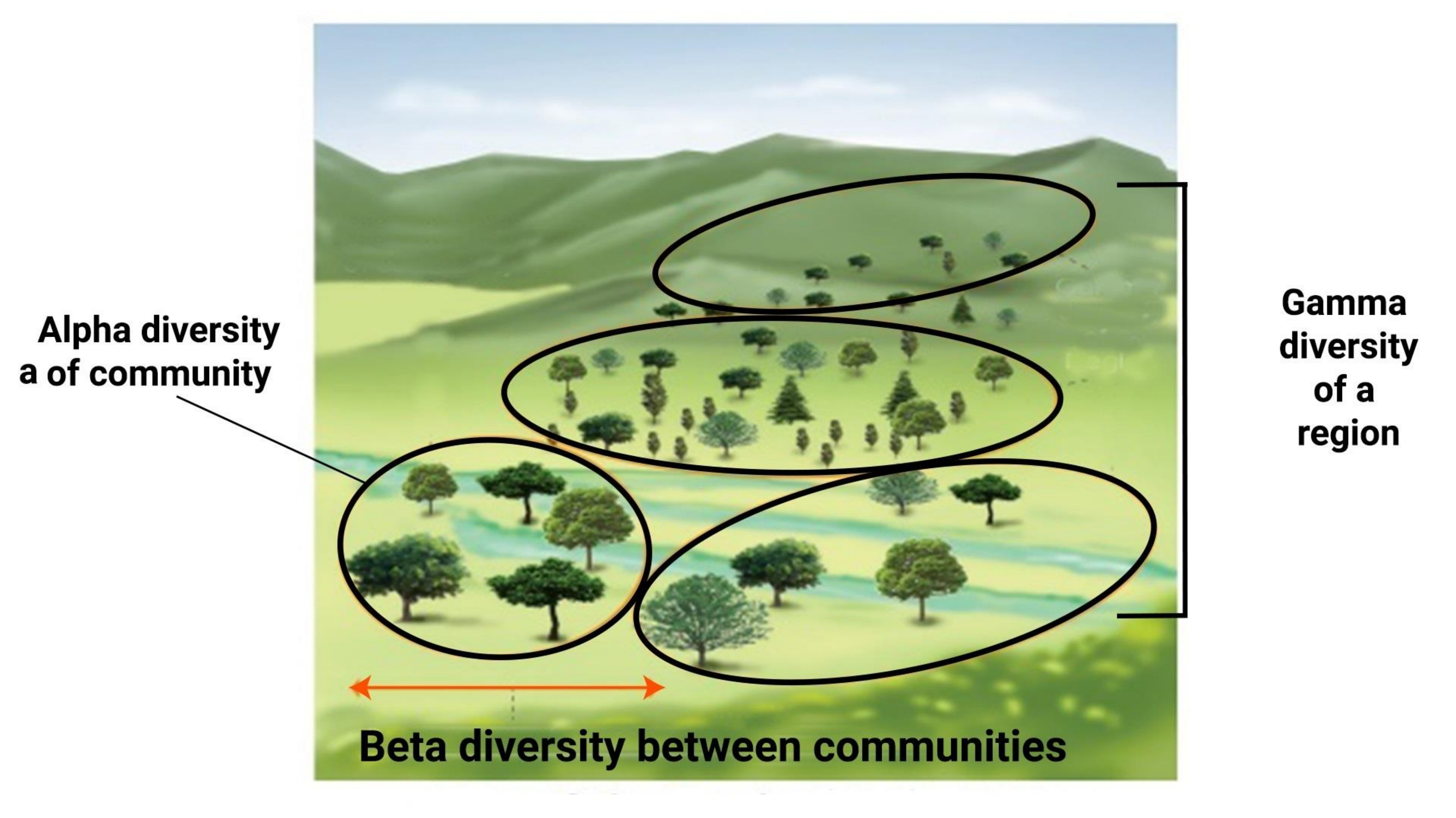
VII. Gamma Diversity (γ-diversity)

- **Gamma Diversity** is the species diversity in a landscape
  - It is determined by alpha diversity and beta diversity.



VIII. Alpha Diversity (α-diversity)

- **Alpha Diversity** is the species diversity present in a site at the local scale



VIII. Alpha Diversity (α-diversity)

- **Alpha Diversity** is measured through **Richness** and **Evenness** within a unit of space (habitat)
  - Diversity indices that consider species proportional abundances can be used to accomplish this.

VIII. Alpha Diversity (α-diversity)

- **A diversity index** is a quantitative measurement of how many species are in a community

VIII. Alpha Diversity (α-diversity)

- **A diversity index** is a quantitative measurement of how many species are in a community

Simpson's Index – The less diversity, the greater the probability that two randomly selected individuals will be the same

$$D = \sum_{i=1}^{S} \left(\frac{n_i}{N}\right)^2$$

n_i = number individuals in species *i*
N = total number of individuals of all species
S = species richness

VIII. Alpha Diversity (α-diversity)

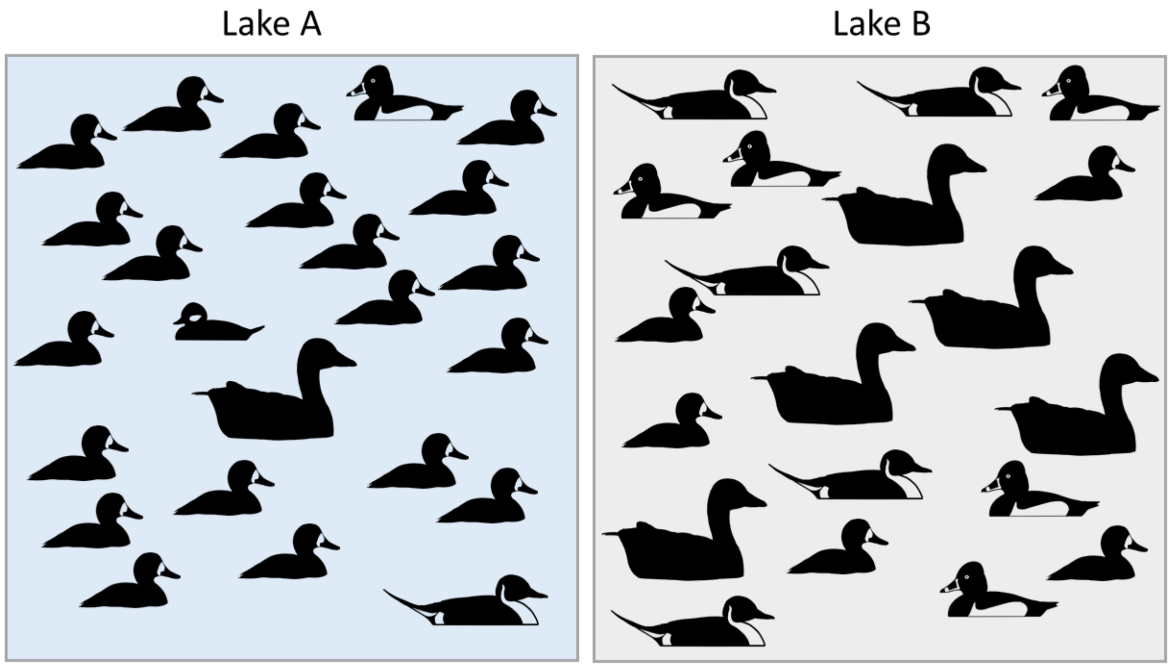
Shannon-Wiener Index – If a community is highly diverse and we choose a species at random, there will be a higher level of uncertainty as to what the species will be.

- Shannon's entropy

$$H = \sum_{i=1}^{S} p_i \times \ln p_i$$

p_i = proportion of individuals of species *i*

S = species richness

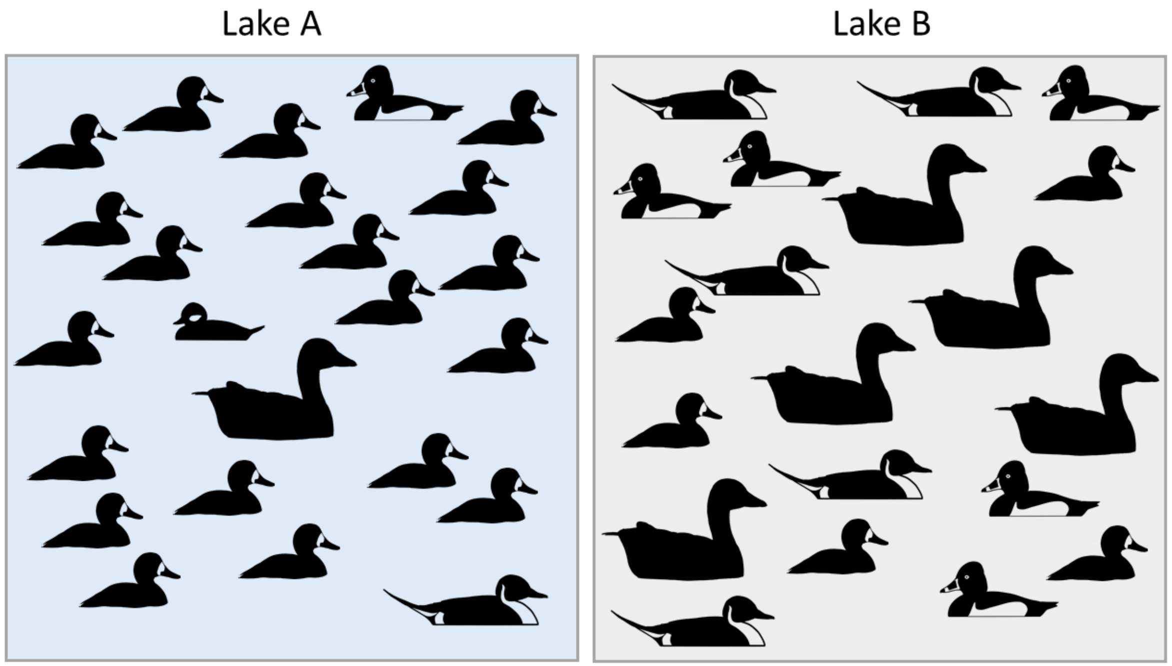


Simpson's Index

	Species (i)	Lake A			Lake B		
		count (n_i)	n_i/N	$(n_i/N)^2$	count (n_i)	n_i/N	$(n_i/N)^2$
	1	1	1/25 = 0.04	0.0016	5	5/25 = 0.2	0.04
	2	1	1/25 = 0.04	0.0016	5	5/25 = 0.2	0.04
	3	1	1/25 = 0.04	0.0016	5	5/25 = 0.2	0.04
	4	1	1/25 = 0.04	0.0016	5	5/25 = 0.2	0.04
	5	21	21/25 = 0.84	0.7056	5	5/25 = 0.2	0.04
	S = 5	N = 25	$D = \sum_{i=1}^{S} \left(\frac{n_i}{N}\right)^2$	0.71	N = 25	$D = \sum_{i=1}^{S} \left(\frac{n_i}{N}\right)^2$	0.2

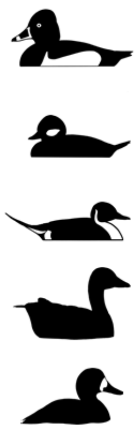
Simpsons index is often expressed as 1-*D*

Therefore, lake B is more diverse because it is less **Dominated** by one species

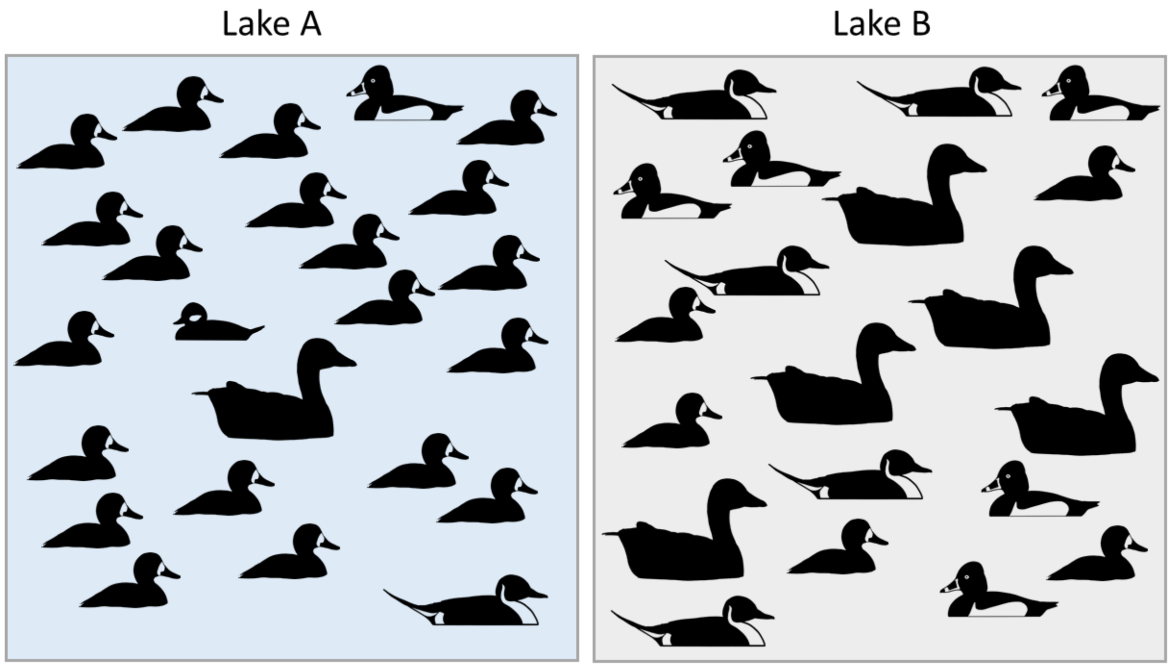


Shannon-Wiener index also suggests lake B is more diverse

Shannon-Weiner Index



Species (i)	Lake A				Lake B			
	count (n_i)	p_i	$\ln(p_i)$	$p_i * \ln(p_i)$	count (n_i)	p_i	$\ln(p_i)$	$p_i * \ln(p_i)$
1	1	0.04	-3.219	-0.129	5	0.2	-1.609	-0.322
2	1	0.04	-3.219	-0.129	5	0.2	-1.609	-0.322
3	1	0.04	-3.219	-0.129	5	0.2	-1.609	-0.322
4	1	0.04	-3.219	-0.129	5	0.2	-1.609	-0.322
5	21	0.84	-0.174	-0.146	5	0.2	-1.609	-0.322
S = 5	N = 25	$H = -\sum_{i=1}^{S} p_i * \ln p_i$		0.661	N = 25	$H = -\sum_{i=1}^{S} p_i * \ln p_i$		1.61



Shannon-Weiner Index

	Species (i)	Lake A count (n_i)	Lake A p_i	Lake A $\ln(p_i)$	Lake A $p_i * \ln(p_i)$	Lake B count (n_i)	Lake B p_i	Lake B $\ln(p_i)$	Lake B $p_i * \ln(p_i)$
	1	1	0.04	-3.219	-0.129	5	0.2	-1.609	-0.322
	2	1	0.04	-3.219	-0.129	5	0.2	-1.609	-0.322
	3	1	0.04	-3.219	-0.129	5	0.2	-1.609	-0.322
	4	1	0.04	-3.219	-0.129	5	0.2	-1.609	-0.322
	5	21	0.84	-0.174	-0.146	5	0.2	-1.609	-0.322
	S = 5	N = 25	$H = -\sum_{i=1}^{S} p_i * \ln p_i$		0.661	N = 25	$H = -\sum_{i=1}^{S} p_i * \ln p_i$		1.61

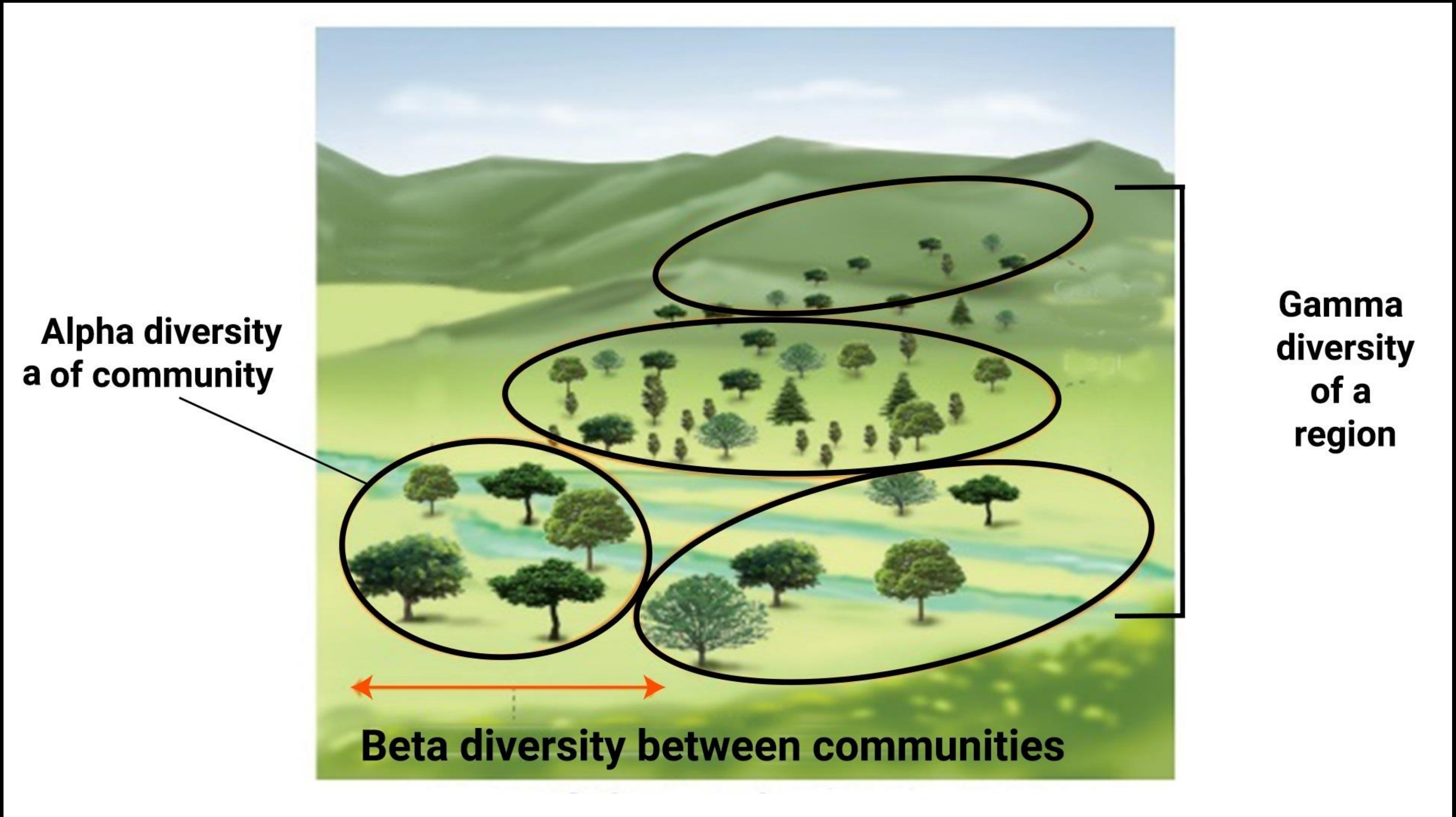
Evenness Index

$$J = \frac{H}{\mathrm{H}(max)}$$

$$J = \frac{0.661}{1.61} = 0.41$$

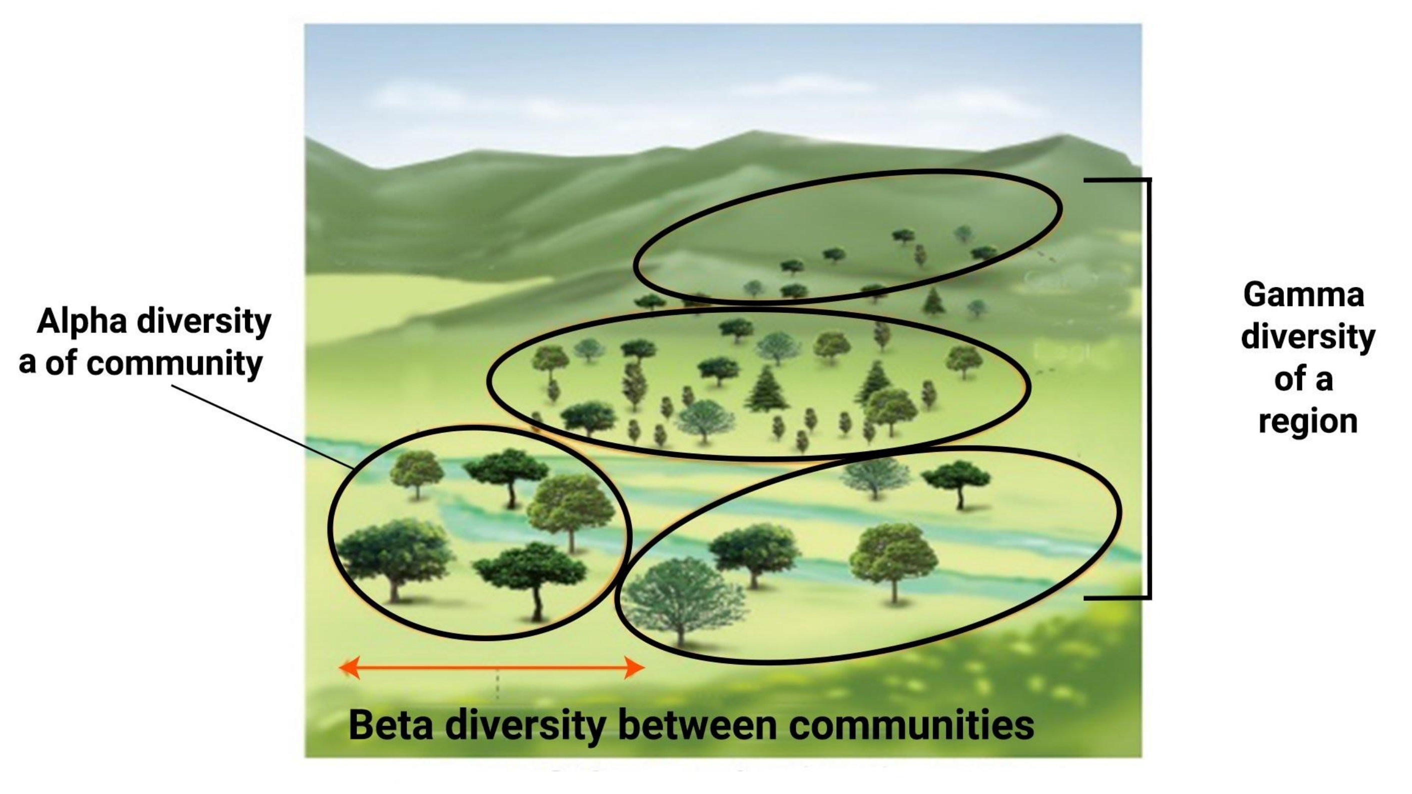
$$J = \frac{1.61}{1.61} = 1$$

IX. Testing for differences in Alpha Diversity



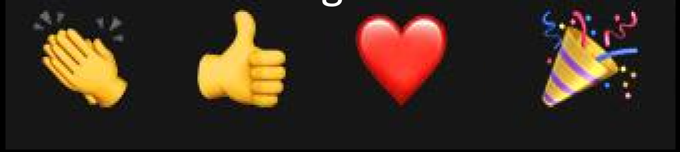
Use ANOVA to test for differences between more than 2 communities

Summary and Check in



Gamma Diversity: Overall species diversity in the landscape
Alpha Diversity: Diversity at a site or habitat
Diversity Indices: Statistical representations of diversity

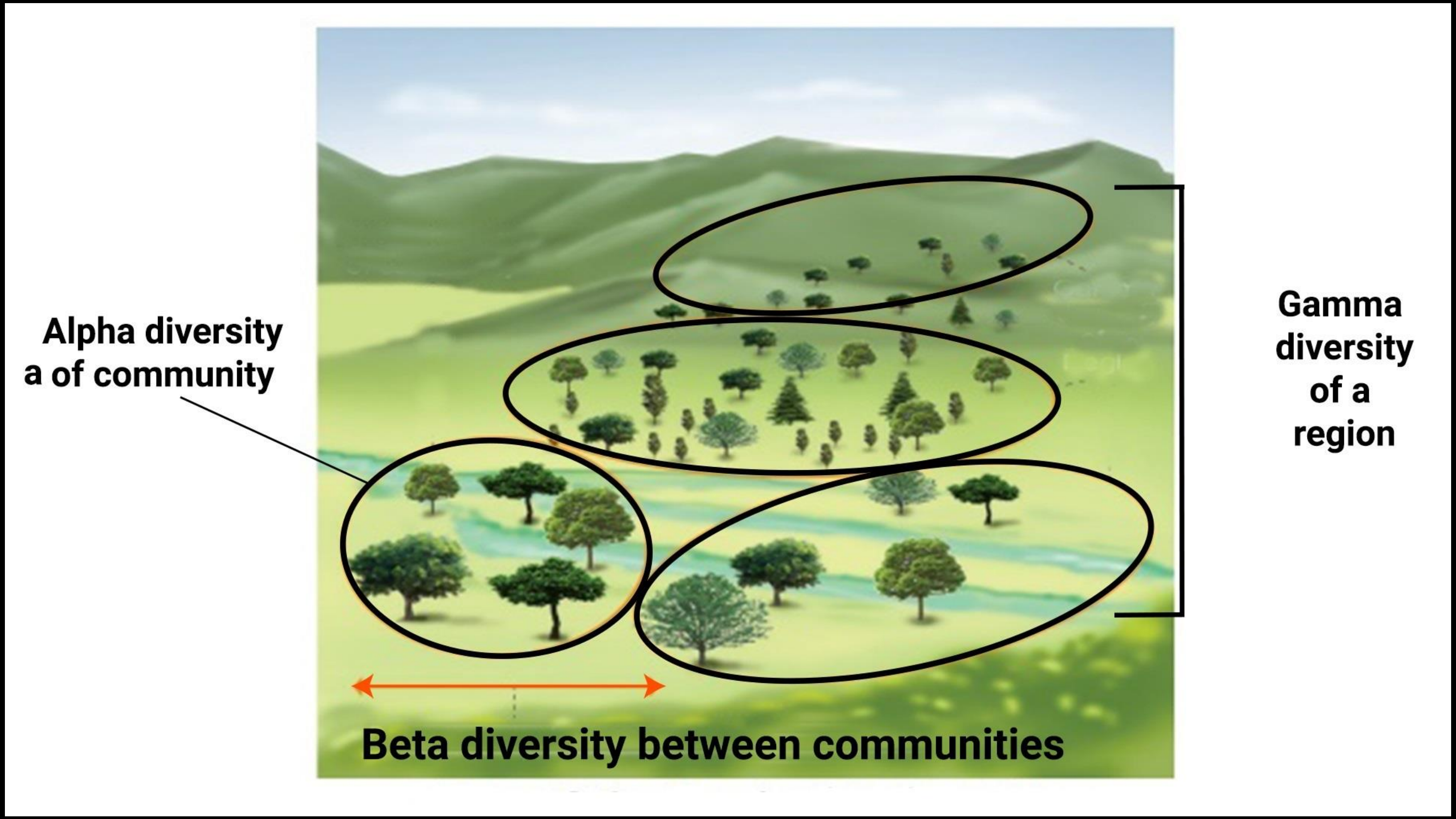
Going well?



Not going so well?

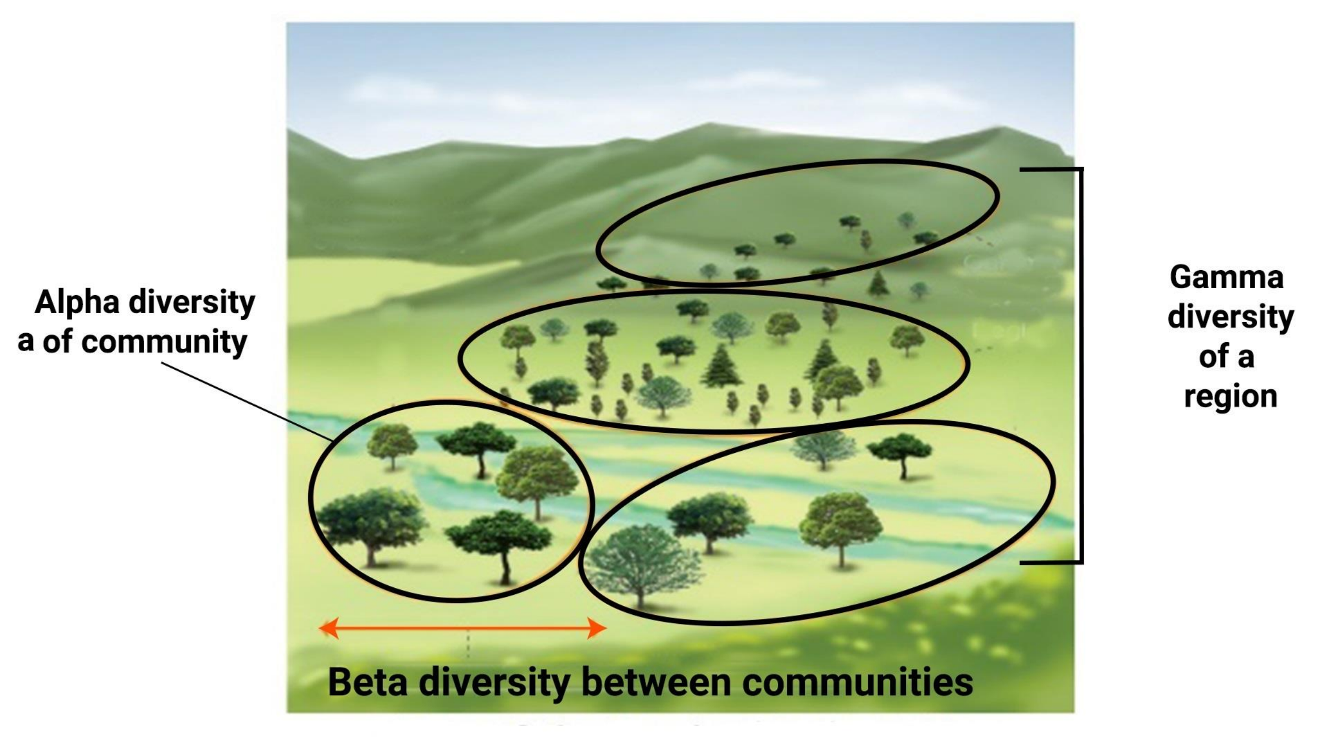
X. Beta Diversity (β-diversity)

- **Beta Diversity** is the rate at which species composition changes across a region and is the link between alpha diversity and gamma diversity



X. Beta Diversity (β-diversity)

- Beta diversity represents the diversity **between** communities
  - If every habitat in a region was inhabited by a **similar** suite of species, then the region will have low beta diversity
  - If some habitats in a region have distinct species composition and do not share much **similarity** with other habitats in the region, then that region has high beta diversity.



X. Beta Diversity (β-diversity)

Bray-Curtis dissimilarity is used to quantify the dissimilarity in community composition between two sites based on species abundances

- Values will be between 0 and 1
  - O = two sites share all the same species
  - 1 = two sites do not share any species.

$$BC_{ij} = 1 - \frac{2C_{ij}}{S_i + S_j}$$

C_{ij} = sum of the lesser counts for species found in both sites
S_i = Total number of species from site i
S_j = Total number of species from site j

X. Beta Diversity (β-diversity)

	A	B	C	D	E
Site 1	4	0	2	7	8
Site 2	3	6	0	4	11

$$BC_{ij} = 1 - \frac{2C_{ij}}{S_i + S_j}$$

C_{ij} = 3+0+0+4+8 = 15
S_i = 4+0+2+7+8 = 21
S_j = 3+6+0+4+11 = 24

Bray-Curtis dissimilarity = 0.33

X. Beta Diversity (β-diversity)

There are many other indices that can be used to calculate Beta diversity

Typically, these indices are either based on abundance or presence/absence of species

Sorenson-index = similarity between two communities

- Usually presence/absence

Jaccard index = similarity between two communities

- Presence/absence

Euclidean distance = physical distance between communities

XI. Testing for differences in Beta Diversity

First, calculate the species similarity/dissimilarity between sites/habitats

You end up with a matrix of "distances" or how different the sites are regarding species composition

Output

	Cleron_07_1	Cleron_07_2	Cleron_07_3	Cleron_08_1	Cleron_08_2
Cleron_07_1	0.0000000	0.2127742	0.3145806	0.5833548	0.4350968
Cleron_07_2	0.2127742	0.0000000	0.2658065	0.5541935	0.4016774
Cleron_07_3	0.3145806	0.2658065	0.0000000	0.4581935	0.3468387
Cleron_08_1	0.5833548	0.5541935	0.4581935	0.0000000	0.3304516
Cleron_08_2	0.4350968	0.4016774	0.3468387	0.3304516	0.0000000

XI. Testing for differences in Beta Diversity

Statistical Test:
PERMANOVA = extension of ANOVA that tests whether the dispersion of groups, defined by a measure of distance, are equivalent.

Unlike ANOVA, PERMANOVA, randomly permutates, or rearranges the order, the distances between groups many times to calculate an F-test statistic

R

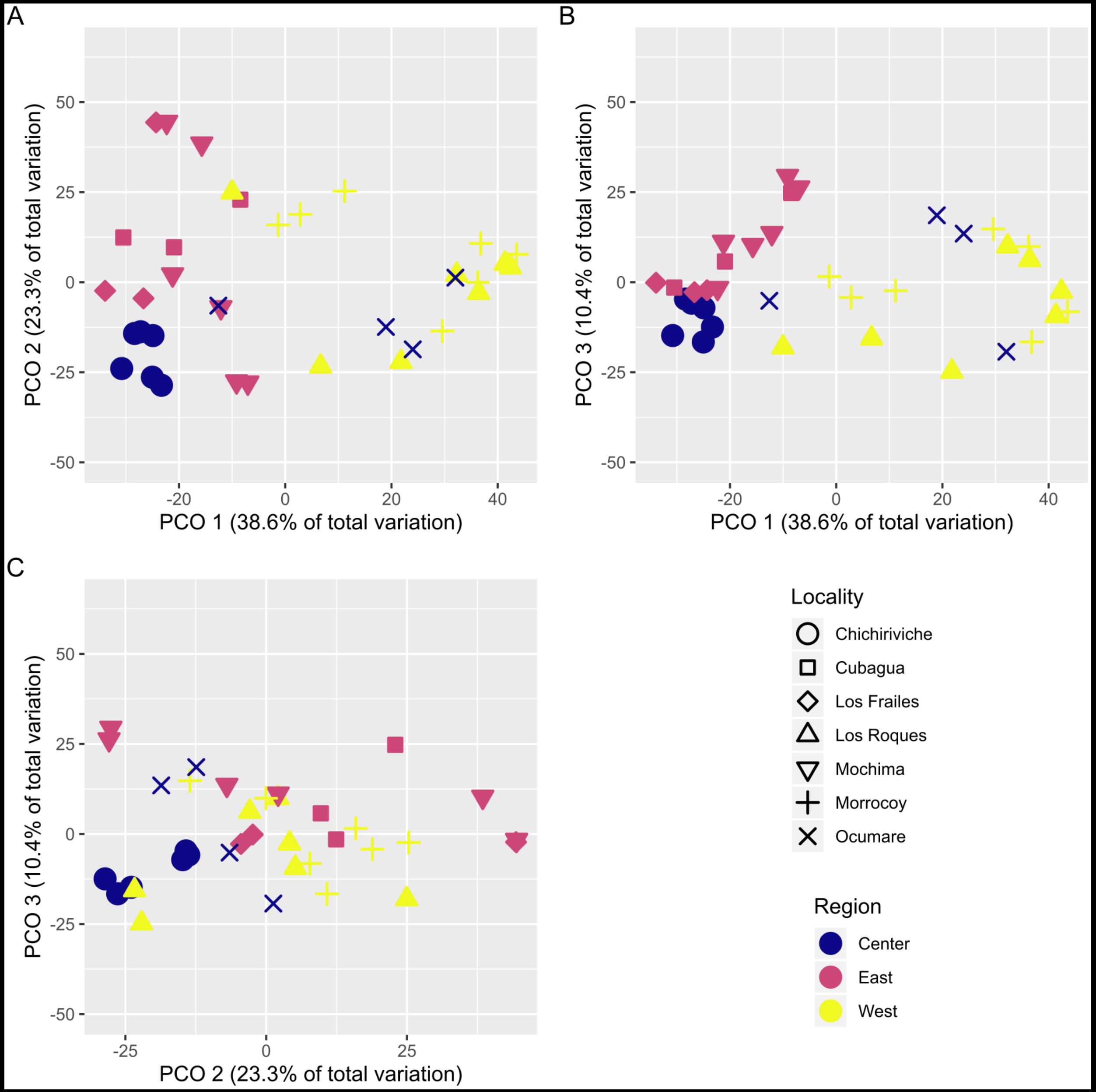
```
# Permanova test using the vegan package
adonis(data_otu_filt_rar~site,data=data_grp, permutations=9999, method="bray")
```

Output

```
Call:
adonis(formula = distance(data_phylo_filt_rar, method = "bray") ~      site, data = metadata)

Permutation: free
Number of permutations: 999

Terms added sequentially (first to last)

          Df SumsOfSqs MeanSqs F.Model      R2 Pr(>F)
site       1   0.63649 0.63649  5.6298 0.26028  0.001 ***
Residuals 16   1.80892 0.11306         0.73972
Total     17   2.44542                 1.00000
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

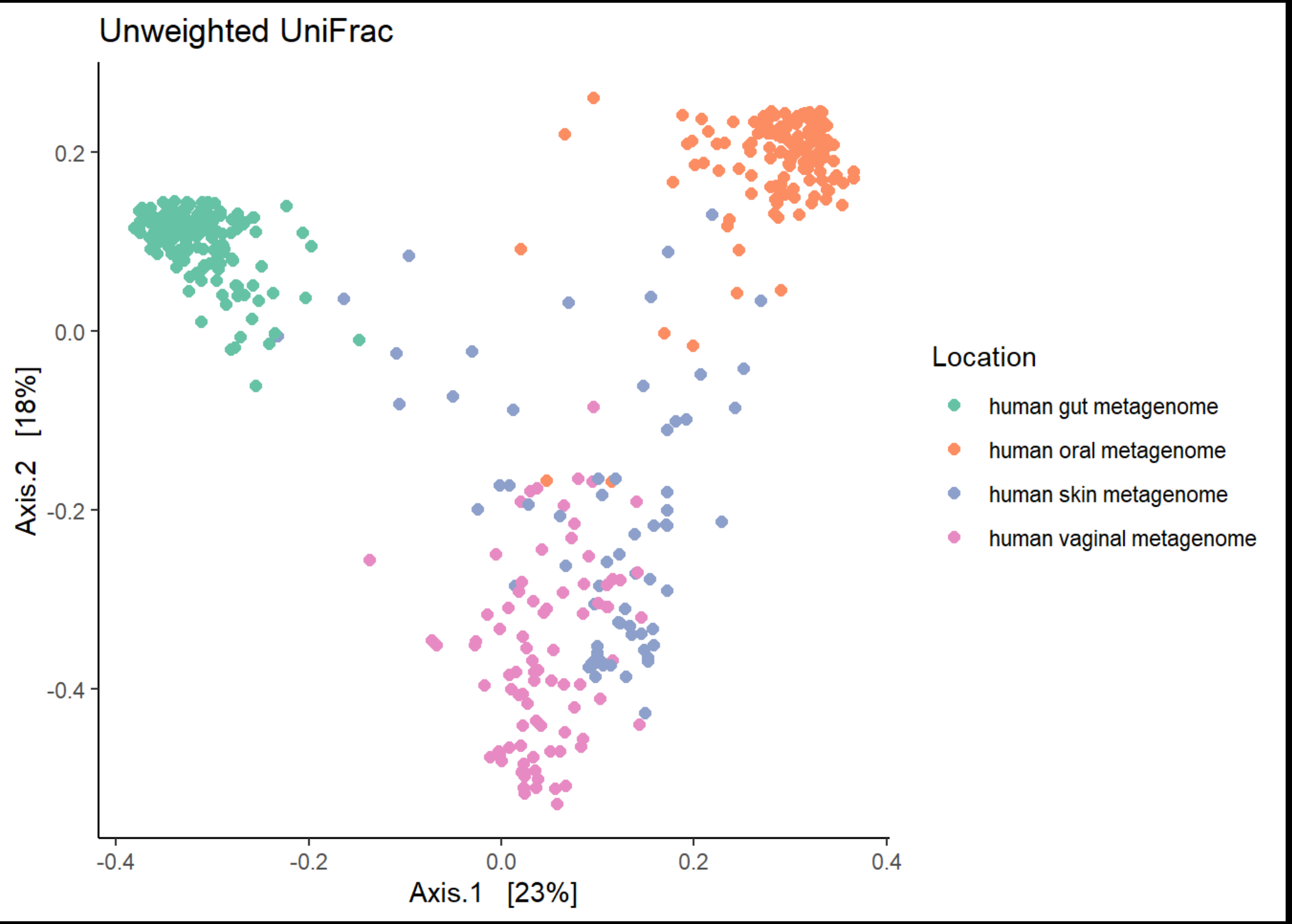
XI. Testing for differences in Beta Diversity

Ordination analysis:
A multivariate technique, use to summarize or reduce a multi-dimensional dataset into a low-dimensional space so intrinsic patterns in the data can be visualized.



XI. Testing for differences in Beta Diversity

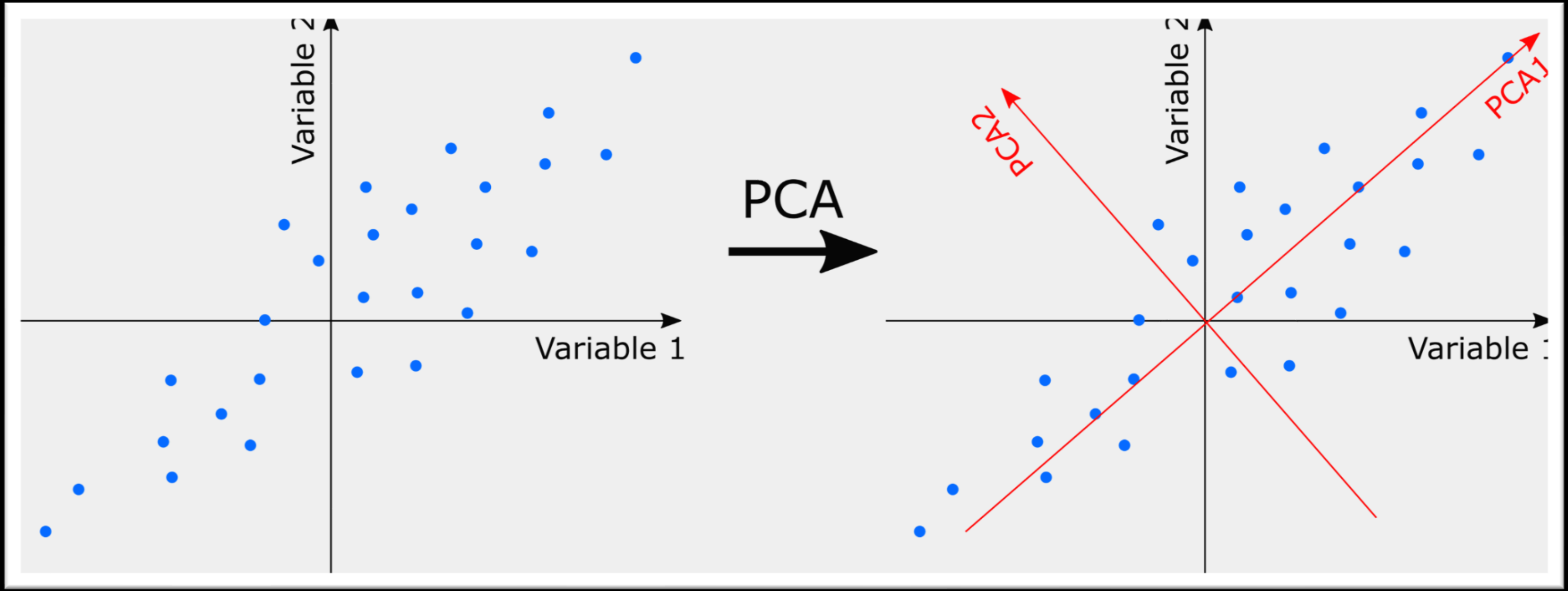
Ordination analysis:
A multivariate technique, use to summarize or reduce a multi-dimensional dataset into a low-dimensional space so intrinsic patterns in the data can be visualized.



XII Relationships between community diversity and the environment

We can use ordination analysis to reduce large datasets into groups of variables with the strongest correlations

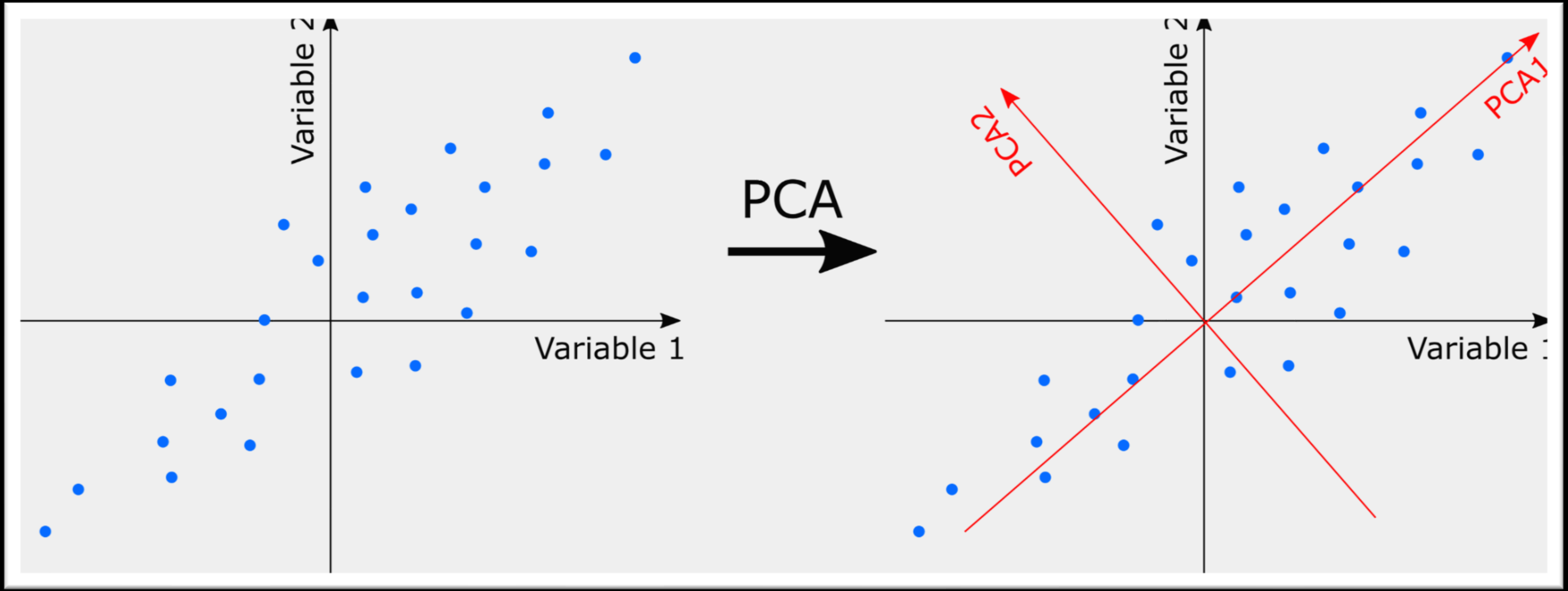
These analyses rotate the original axes of the data to derive new axes, which maximize the variance in the data set



XII. Relationships between community diversity and the environment

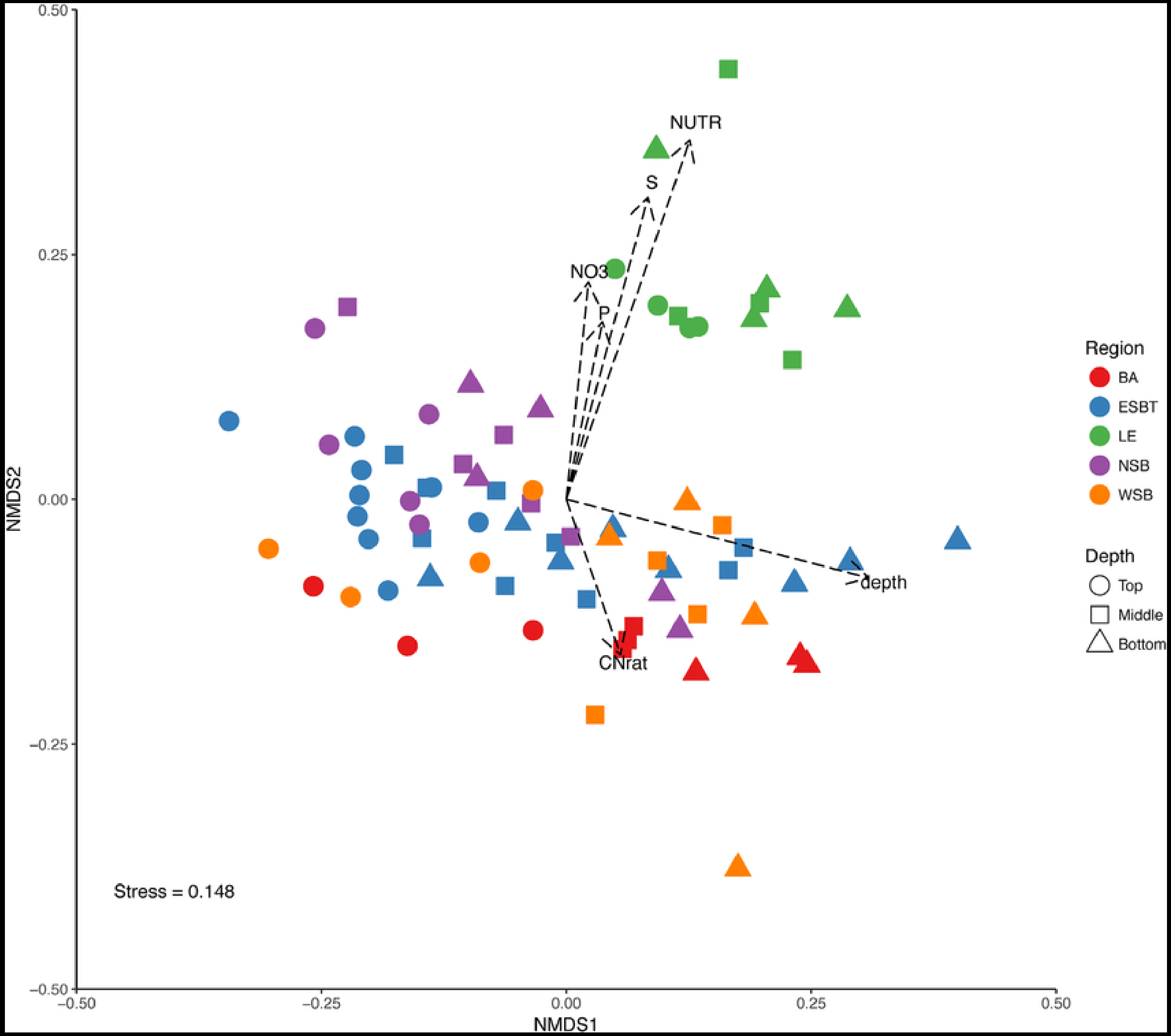
An exploratory analysis that can be used to reduce large datasets into groups of variables with the strongest correlations

It rotates the original axes of the data to derive new axes, which maximize the variance in the data set



XII. Relationships between community diversity and the environment

Using ordination, we can take two matrix of data, species dissimilarity and environmental dissimilarity, and evaluate any correlation between them.



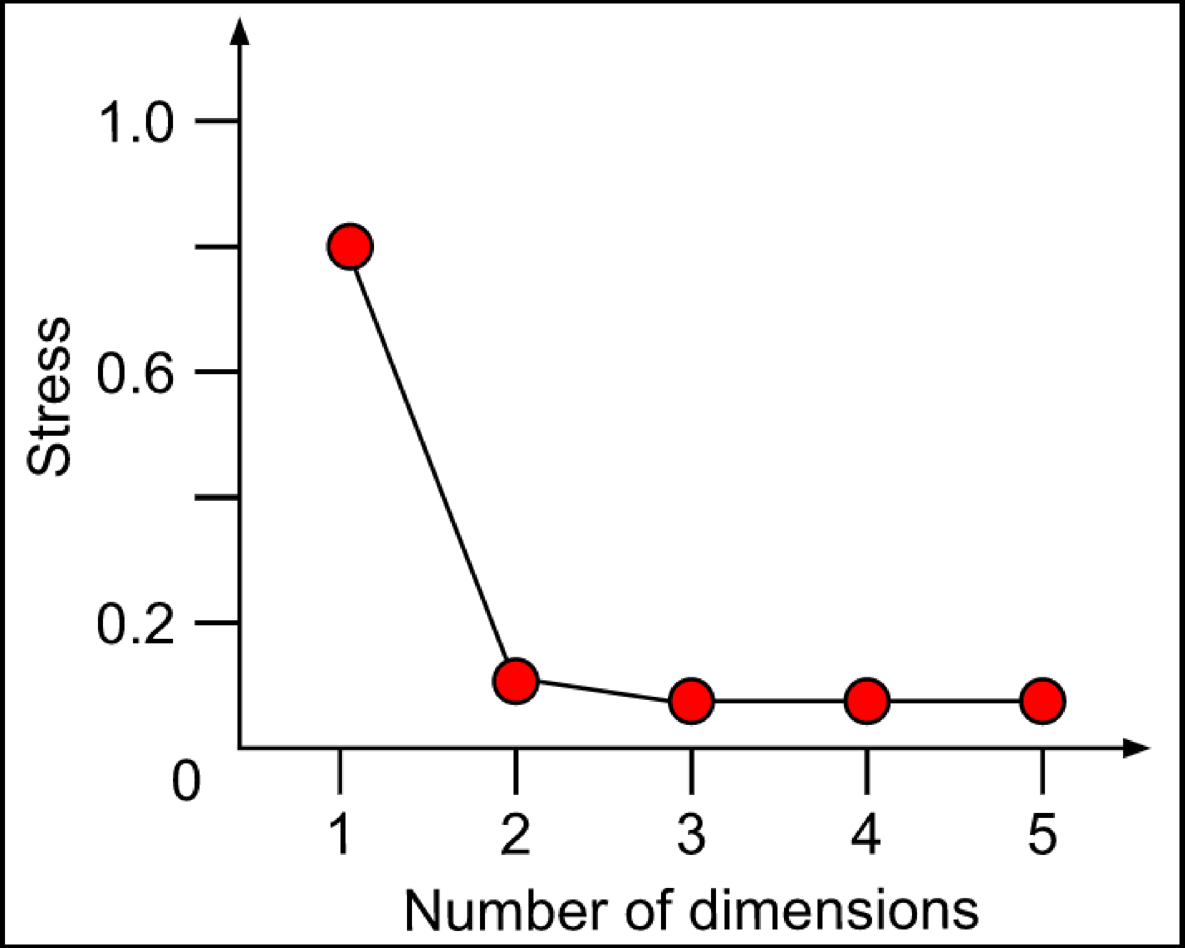
XII. Relationships between community diversity and the environment

A stress test can be used to evaluate the ordination

Measure goodness of fit vs # of dimensions

Stress values > 0.2 = poor fit and potentially uninterpretable

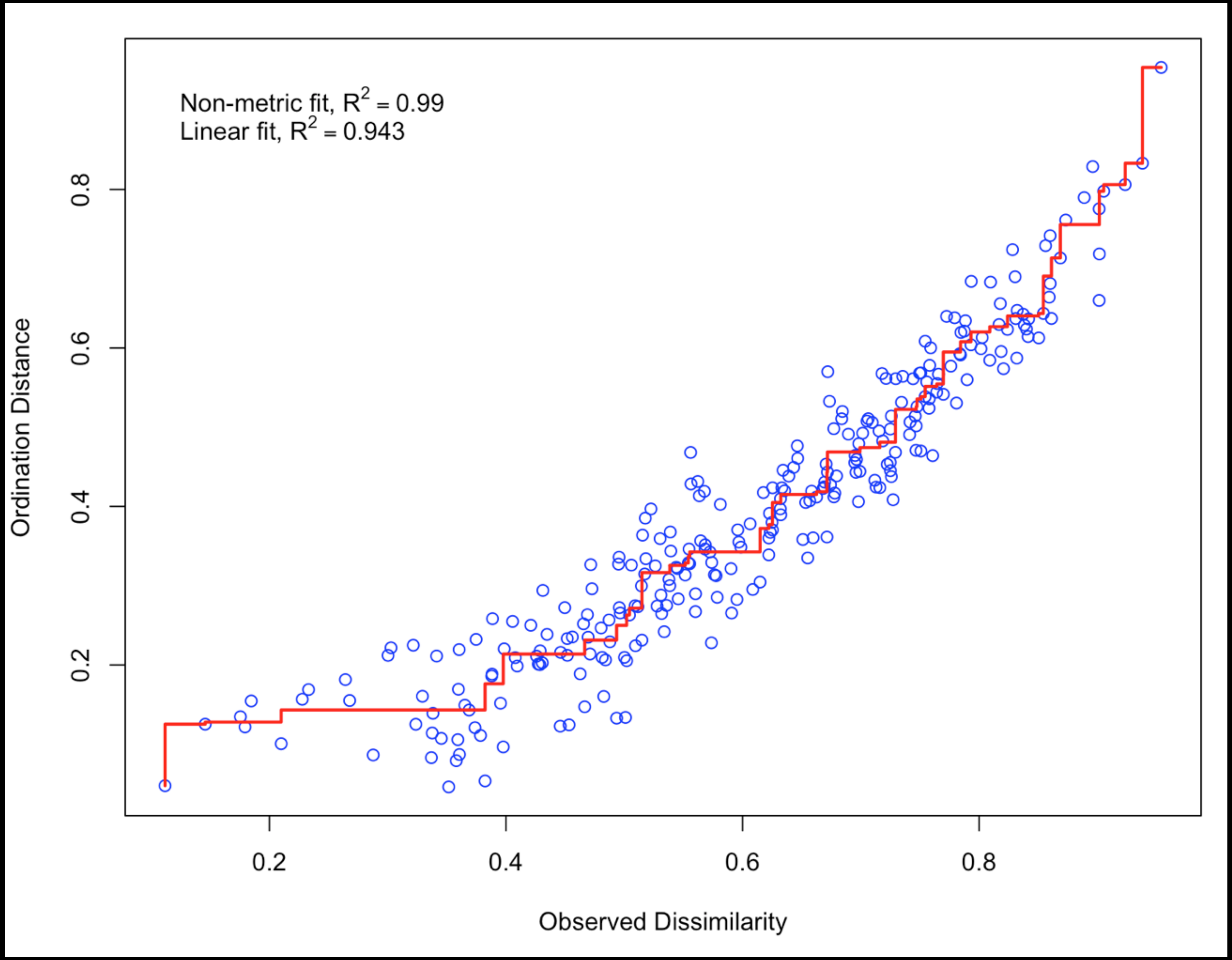
Stress values <0.1 = good and have little danger of misinterpretation



XII. Relationships between community diversity and the environment

A stress test can be used to evaluate the ordination

Measure goodness of fit vs the observed dissimilarities in our communities



XII.I Types of ordination analysis to use in community ecology

Principal component analysis (PCA)

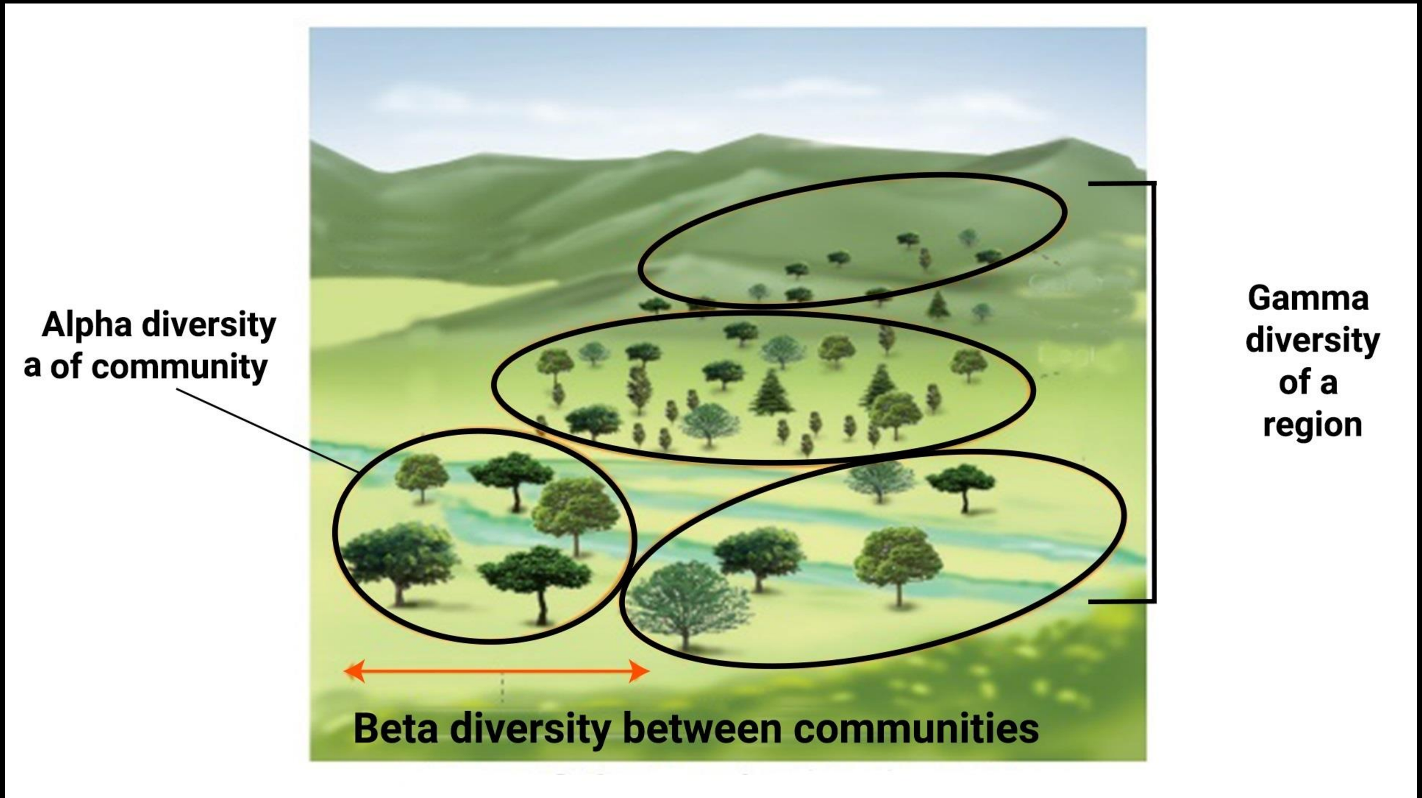
- An exploratory analysis that can be used to reduce large datasets into groups of variables with the strongest correlations
- Eigenanalysis = the matrix is symmetric, axes are orthogonal and independent

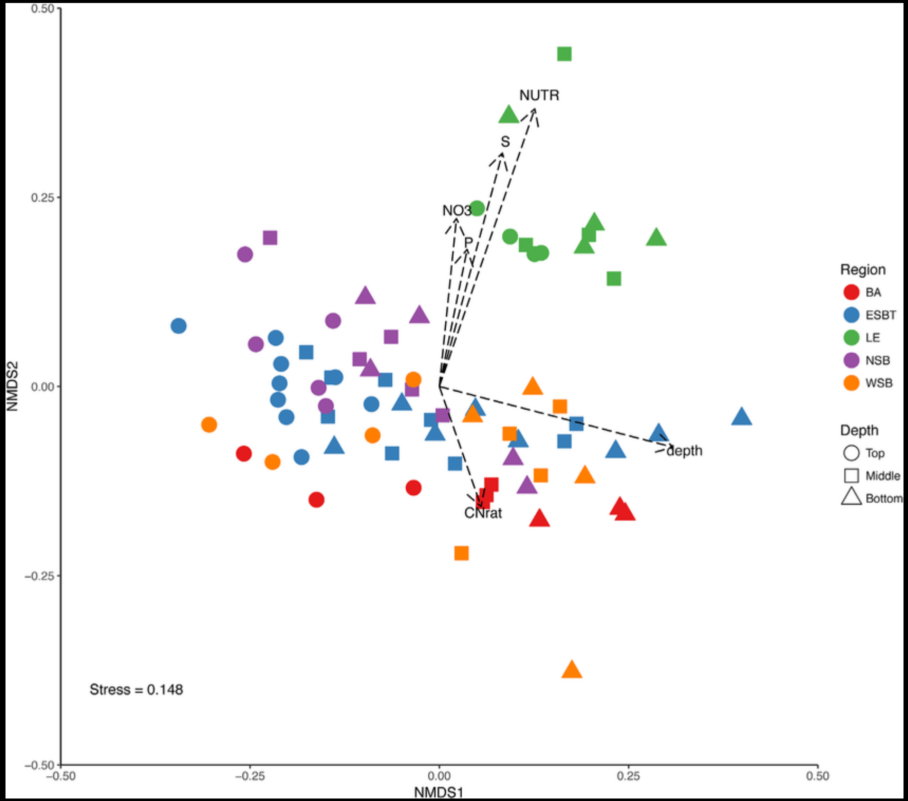
Non-metric Multidimensional Scaling

- Attempts to represent pairwise dissimilarity between objects in low-dimensional space. Best used with a dissimilarity matrix.
- Not an eigenanalysis

Resources for Ordination
https://ourcodingclub.github.io/tutorials/ordination/#section6

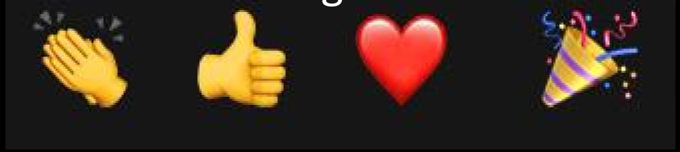
Summary and Check in





Beta Diversity: Difference in species composition between sites
Ordination: Can be used to visualize differences and correlate to environment

Going well?



Not going so well?