6. Worksheet: Among Site (Beta) Diversity – Part 1

Trang Nguyen; Z620: Quantitative Biodiversity, Indiana University

05 février, 2025

OVERVIEW

In this worksheet, we move beyond the investigation of within-site α-diversity. We will explore β-diversity, which is defined as the diversity that occurs among sites. This requires that we examine the compositional similarity of assemblages that vary in space or time.

After completing this exercise you will know how to:

1. formally quantify β-diversity
2. visualize β-diversity with heatmaps, cluster analysis, and ordination
3. test hypotheses about β-diversity using multivariate statistics

Directions:

1. In the Markdown version of this document in your cloned repo, change "Student Name" on line 3 (above) with your name.
2. Complete as much of the worksheet as possible during class.
3. Use the handout as a guide; it contains a more complete description of data sets along with examples of proper scripting needed to carry out the exercises.
4. Answer questions in the worksheet. Space for your answers is provided in this document and is indicated by the ">" character. If you need a second paragraph be sure to start the first line with ">". You should notice that the answer is highlighted in green by RStudio (color may vary if you changed the editor theme).
5. Before you leave the classroom, **push** this file to your GitHub repo.
6. For the assignment portion of the worksheet, follow the directions at the bottom of this file.
7. When you are done, **Knit** the text and code into a PDF file.
8. After Knitting, submit the completed exercise by creating a **pull request** via GitHub. Your pull request should include this file (**6.BetaDiversity_1_Worksheet.Rmd**) with all code blocks filled out and questions answered) and the PDF output of `Knitr` (**6.BetaDiversity_1_Worksheet.pdf**).

The completed exercise is due on **Wednesday, February 5th, 2025 before 12:00 PM (noon)**.

1) R SETUP

Typically, the first thing you will do in either an R script or an RMarkdown file is setup your environment. This includes things such as setting the working directory and loading any packages that you will need.

In the R code chunk below, please provide the code to:

1) Clear your R environment,
2) Print your current working directory,
3) Set your working directory to your `Week3-Beta/` folder folder, and
4) Load the `vegan` R package (be sure to install first if you have not already).

```
# Clear the environment
rm(list = ls())
getwd()
```

```
## [1] "C:/Users/ttran/OneDrive - Indiana University/SP25 - Quantitative Biodiversity/QB2025_Nguyen/Week
```

```
# setwd("Week3-Beta/")
# library(vegan))

package.list = c("ade4", "vegan", "viridis", "gplots", "BiodiversityR", "indicspecies")
for (package in package.list) {
  if (!require(package, character.only = TRUE, quietly = TRUE)) {
    install.packages(package, dependencies = TRUE)
    library(package, character.only = TRUE)
  }
}
```

```
## Warning: le package 'ade4' a été compilé avec la version R 4.4.2
```

```
## Warning: le package 'vegan' a été compilé avec la version R 4.4.2
```

```
## Warning: le package 'permute' a été compilé avec la version R 4.4.2
```

```
## This is vegan 2.6-8
```

```
## Warning: le package 'viridis' a été compilé avec la version R 4.4.2
```

```
## Warning: le package 'gplots' a été compilé avec la version R 4.4.2
```

```
##
## Attachement du package : 'gplots'
```

```
## L'objet suivant est masqué depuis 'package:stats':
##
##     lowess
```

```
## Warning: le package 'BiodiversityR' a été compilé avec la version R 4.4.2
```

```
## BiodiversityR 2.17-1.1: Use command BiodiversityRGUI() to launch the Graphical User Interface;
## to see changes use BiodiversityRGUI(changeLog=TRUE, backward.compatibility.messages=TRUE)
```

```
## Warning: le package 'indicspecies' a été compilé avec la version R 4.4.2
```

2) LOADING DATA

Load dataset

In the R code chunk below, do the following:

1. load the `doubs` dataset from the `ade4` package, and
2. explore the structure of the dataset.

```
# note, please do not print the dataset when submitting
library(ade4)
data(doubs)
summary(doubs)
```

```
##         Length Class      Mode
## env     11     data.frame list
## fish    27     data.frame list
## xy       2     data.frame list
## species  4     data.frame list
```

```
# dim(doubs$fish)
dim(doubs$xy)
```

```
## [1] 30  2
```

Question 1: Describe some of the attributes of the `doubs` dataset.

a. How many objects are in `doubs`?
b. How many fish species are there in the `doubs` dataset?
c. How many sites are in the `doubs` dataset?

> ***Answer 1a***: There are 4 objects in the `doubs` dataset. They are `env`, `fish`, `xy`, `species` ***Answer 1b***: There are 27 fish species in the `doubs` dataset. ***Answer 1c***: There are 30 sites in the `doubs` dataset. (30 rows in `doubs$xy`)

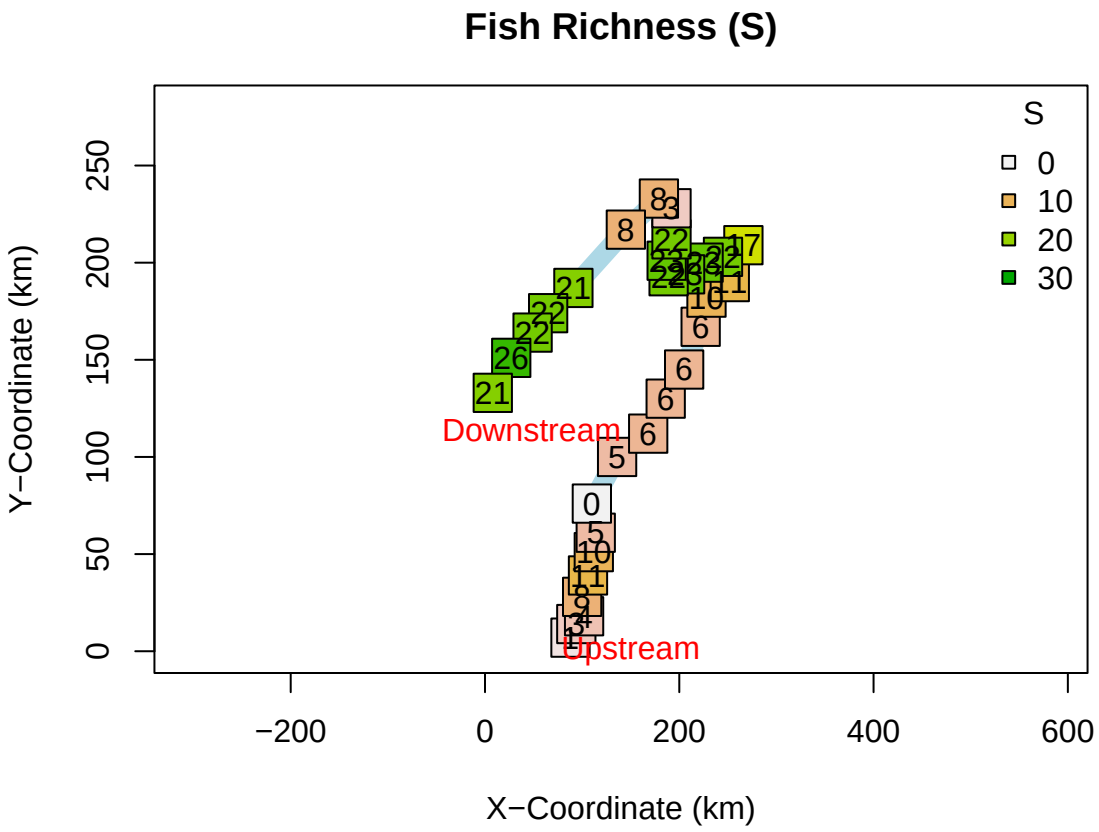
Visualizing the Doubs River Dataset

Question 2: Answer the following questions based on the spatial patterns of richness (i.e., α-diversity) and Brown Trout (*Salmo trutta*) abundance in the Doubs River.

a. How does fish richness vary along the sampled reach of the Doubs River?
b. How does Brown Trout (*Salmo trutta*) abundance vary along the sampled reach of the Doubs River?
c. What do these patterns say about the limitations of using richness when examining patterns of biodiversity?

```
# Observed richness as in previous notebook
spa.S = specnumber(doubs$fish)
spa.S.color = rev(terrain.colors(31))    # Define Richness Color Pallette
spa.N.color = rev(terrain.colors(7))     # Define Abundance Color Pallette
```

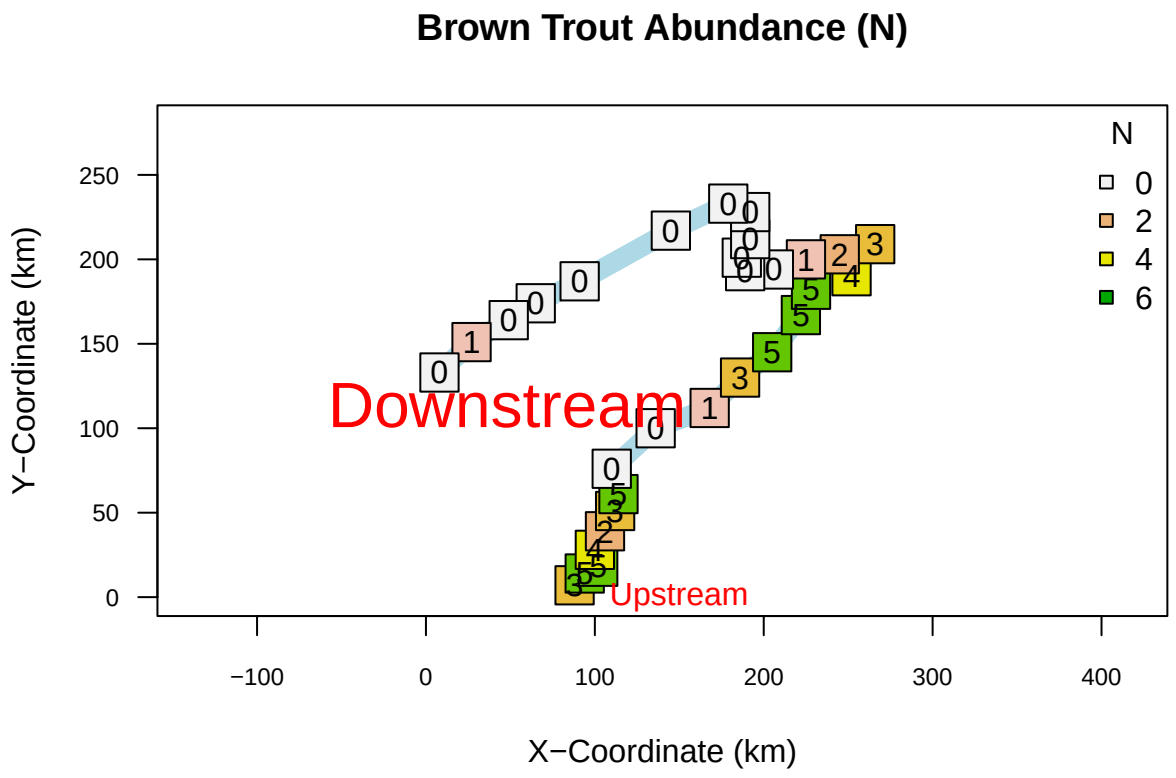
```
# Define Plot Parameters
opar <- par(no.readonly = TRUE)
par(mfrow = c(1,  1), mar = c(4, 4, 3, 4) + 0.1, xpd = TRUE)

# Species Richness
plot.new()
plot(doubs$xy, asp = 2, type = 'l', lwd = 10, col = "light blue", xlim = c(0,280), ylim = c(0,280),
     main = "Fish Richness (S)",
     xlab = "X-Coordinate (km)", ylab = "Y-Coordinate (km)")
points(doubs$xy, pch = 22, cex=3, bg = spa.S.color[spa.S + 1])
text(doubs$xy, as.character(spa.S), cex = 1, col = "black")
text(150, 0, "Upstream", cex = 1, col = "red")
text(48, 114, "Downstream", cex = 1, col = "red")
legend("topright", pch = 22, pt.cex = 1, bty = 'n',
       title = "S", legend = seq(0, 30, 10),
       pt.bg = spa.S.color[seq(1, 31, 10)])
```




```
par = opar
```

Answer 2a: How does fish richness vary along the sampled reach of the Doubs River?: It seems like there are more species encounterd in the downstream of the Doubs River than in the upstream. Based on what I found, closer to the Suisse border, the richness is also higher.

```
# Brown Trout Abundance
plot.new()
plot(doubs$xy, asp = 1, type = 'l', col = "light blue",
    lwd = 10,
     las = 1, cex.axis = 0.75,
     xlim = c(0, 280), ylim = c(0, 280),
     main = "Brown Trout Abundance (N)",
     xlab = "X-Coordinate (km)", ylab = "Y-Coordinate (km)")
points(doubs$xy, pch = 22, cex = 3, bg = spa.N.color[doubs$fish$Satr + 1])
text(doubs$xy, as.character(doubs$fish$Satr), cex = 1, col="black")
text(150, 0, "Upstream", cex = 1, col = "red")
text(48, 114, "Downstream", cex

 = 2, col = "red")
legend("topright", pch = 22, pt.cex = 1, bty = 'n',
       title = "N", legend = seq(0, 6, 2), pt.bg = spa.N.color[seq(1, 7, 2)])
```




```
par <- opar
```

Answer 2b How does Brown Trout (*Salmo trutta*) abundance vary along the sampled reach of the Doubs River?: The abundance of Brown Trout is higher in the upstream of the Doubs River. There is almost no presence of the Brown Trouts in the downstream . The highest abundance is found in the middle of the river, closer to the Suisse border.

Answer 2c: What do these patterns say about the limitations of using richness when examining patterns of biodiversity? I think the richness is not enough to explain

the biodiversity of the river as this measure does not reflecg how specific species are distributed along the river. The abundance of a specific species can be higher in a specific areas, which can be a limitation of using richness to examine the biodiversity of the river.

3) QUANTIFYING BETA-DIVERSITY

In the R code chunk below, do the following:

1. write a function (`beta.w()`) to calculate Whittaker's β-diversity (i.e., β_w) that accepts a site-by-species matrix with optional arguments to specify pairwise turnover between two sites, and
2. use this function to analyze various aspects of β-diversity in the Doubs River.

```
beta = function(site.by_species = ""){
  SbyS.pa = decostand(site.by_species, method="pa") # convert to presence-absence matrix
  S = ncol(SbyS.pa[, which(colSums(SbyS.pa) > 0)]) # number of species in the region
  a.bar = mean(specnumber(SbyS.pa)) # average species richness
  b.w = round(S/a.bar, 3)
  return (b.w)
}
```

Question 3: Using your `beta.w()` function above, answer the following questions: > ***Answer 3a: Describe how local richness (α) and turnover (β) contribute to regional (γ) fish diversity in the Doubs.***:

```
# Compute species richness for each site
alpha_richness = specnumber(doubs$fish)

# Compute beta diversity for each pair of sites
beta_diversity = beta(doubs$fish)

# Compute Gamma Diversity (Total number of unique species across all sites)
gamma_diversity = specnumber(colSums(doubs$fish > 0))

# Print out answers
print("Alpha Richness:")
```

```
## [1] "Alpha Richness:"
```

```
print(alpha_richness)
```

```
##  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
##  1  3  4  8 11 10  5  0  5  6  6  6  6 10 11 17 22 23 23 22 23 22  3  8  8 21
## 27 28 29 30
## 22 22 26 21
```

```
print(paste("Beta Diversity: ", beta_diversity))
```

```
## [1] "Beta Diversity:  2.16"
```

```r
print(paste("Gamma Diversity: ", gamma_diversity))
```

```
## [1] "Gamma Diversity:  27"
```

Answer 3b. Is the fish assemblage at site 1 more similar to the one at site 2 or site 10?

```r
# Compute beta diversity between site 1 and site 2

beta_diversity_1_2 = beta(doubs$fish[c(1, 2), ])
beta_diversity_1_10 = beta(doubs$fish[c(1, 10), ])
beta_diversity_2_10 = beta(doubs$fish[c(2, 10), ])
# Check which pair of sites has the lowest beta diversity
print("The pair of sites with the lowest beta diversity is:")
```

```
## [1] "The pair of sites with the lowest beta diversity is:"
```

```r
print(paste("Site 1 and Site 2: ", beta_diversity_1_2))
```

```
## [1] "Site 1 and Site 2:  1.5"
```

```r
print(paste("Site 1 and Site 10: ", beta_diversity_1_10))
```

```
## [1] "Site 1 and Site 10:  1.714"
```

The fish assemblage at site 1 is more similar to the one at site 2 than site 10.

Answer 3c: Using your understanding of the equation $\beta_w = \gamma/\alpha$, how would your interpretation of β change if we instead defined beta additively (i.e., $\beta = \gamma - \alpha$)? If we define beta additively, the beta diversity would be the absolute difference between total species richness and average local richness. In this case, we measure how many species are exclusive to different sites, rather than how many times local diversity fits into the total diversity.

The Resemblance Matrix

In order to quantify β-diversity for more than two samples, we need to introduce a new primary ecological data structure: the **Resemblance Matrix**.

Question 4: How do incidence- and abundance-based metrics differ in their treatment of rare species?

Answer 4:
Incidence-based metrics treat all species as equal, regardless of their count. They focus on the presence or absence of species in different sites. Abundance-based metrics, on the other hand, actually consider the abundance of species in addition to their presence or absence. They take into account the relative abundance of species in different sites.

In the R code chunk below, do the following:

1. make a new object, `fish`, containing the fish abundance data for the Doubs River, 2. remove any sites where no fish were observed (i.e., rows with sum of zero), 3. construct a resemblance matrix based on Sørensen's Similarity ("fish.ds"), and 4. construct a resemblance matrix based on Bray-Curtis Distance ("fish.db").

```
# Q4.a
fish = doubs$fish
# Q4.b
fish = fish[rowSums(fish) > 0, ]
# Q4.c
fish.ds = vegdist(fish, method = "bray", binary=TRUE)
fish.db = vegdist(fish, method = "bray")
print("Resemblance Matrices using Sorensen's Similarity")
```

```
## [1] "Resemblance Matrices using Sorensen's Similarity"
```

```
fish.ds
```

```
##            1          2          3          4          5          6          7
## 2  0.50000000
## 3  0.60000000 0.14285714
## 4  0.77777778 0.45454545 0.33333333
## 5  0.83333333 0.57142857 0.46666667 0.15789474
## 6  0.81818182 0.53846154 0.42857143 0.11111111 0.04761905
## 7  0.66666667 0.25000000 0.33333333 0.38461538 0.37500000 0.33333333
## 9  1.00000000 0.50000000 0.55555556 0.38461538 0.37500000 0.33333333 0.40000000
## 10 0.71428571 0.33333333 0.40000000 0.28571429 0.29411765 0.25000000 0.09090909
## 11 0.71428571 0.33333333 0.40000000 0.42857143 0.52941176 0.50000000 0.27272727
## 12 0.71428571 0.33333333 0.40000000 0.42857143 0.52941176 0.50000000 0.27272727
## 13 0.71428571 0.33333333 0.40000000 0.57142857 0.64705882 0.62500000 0.45454545
## 14 0.81818182 0.53846154 0.42857143 0.33333333 0.42857143 0.40000000 0.46666667
## 15 0.83333333 0.57142857 0.60000000 0.36842105 0.36363636 0.33333333 0.37500000
## 16 0.88888889 0.70000000 0.61904762 0.36000000 0.28571429 0.25925926 0.54545455
## 17 0.91304348 0.76000000 0.69230769 0.46666667 0.39393939 0.37500000 0.62962963
## 18 0.91666667 0.76923077 0.70370370 0.48387097 0.41176471 0.39393939 0.64285714
## 19 1.00000000 0.84615385 0.77777778 0.54838710 0.41176471 0.45454545 0.71428571
## 20 1.00000000 0.84000000 0.76923077 0.53333333 0.39393939 0.43750000 0.70370370
## 21 1.00000000 0.84615385 0.77777778 0.54838710 0.41176471 0.45454545 0.71428571
## 22 1.00000000 0.92000000 0.84615385 0.60000000 0.45454545 0.50000000 0.77777778
## 23 1.00000000 1.00000000 1.00000000 0.81818182 0.71428571 0.69230769 0.75000000
## 24 1.00000000 1.00000000 1.00000000 0.75000000 0.68421053 0.66666667 0.84615385
## 25 1.00000000 1.00000000 0.83333333 0.62500000 0.36842105 0.44444444 0.69230769
## 26 1.00000000 0.91666667 0.84000000 0.58620690 0.43750000 0.48387097 0.76923077
## 27 1.00000000 0.92000000 0.84615385 0.60000000 0.45454545 0.50000000 0.77777778
## 28 1.00000000 0.92000000 0.84615385 0.60000000 0.45454545 0.50000000 0.77777778
## 29 0.92592593 0.79310345 0.73333333 0.52941176 0.40540541 0.44444444 0.67741935
## 30 1.00000000 1.00000000 0.92000000 0.65517241 0.50000000 0.54838710 0.84615385
##            9         10         11         12         13         14         15
## 2
## 3
## 4
## 5
## 6
```

```
## 7
## 9
## 10 0.45454545
## 11 0.45454545 0.33333333
## 12 0.45454545 0.33333333 0.00000000
## 13 0.63636364 0.50000000 0.16666667 0.16666667
## 14 0.60000000 0.37500000 0.25000000 0.25000000 0.25000000
## 15 0.50000000 0.29411765 0.29411765 0.29411765 0.29411765 0.14285714
## 16 0.54545455 0.47826087 0.56521739 0.56521739 0.56521739 0.33333333 0.28571429
## 17 0.62962963 0.57142857 0.57142857 0.57142857 0.57142857 0.37500000 0.33333333
## 18 0.64285714 0.58620690 0.58620690 0.58620690 0.58620690 0.39393939 0.35294118
## 19 0.64285714 0.65517241 0.79310345 0.79310345 0.79310345 0.57575758 0.52941176
## 20 0.62962963 0.64285714 0.78571429 0.78571429 0.85714286 0.62500000 0.57575758
## 21 0.64285714 0.65517241 0.79310345 0.79310345 0.86206897 0.63636364 0.58823529
## 22 0.70370370 0.71428571 0.85714286 0.85714286 0.92857143 0.68750000 0.63636364
## 23 0.50000000 0.77777778 0.77777778 0.77777778 1.00000000 0.84615385 0.85714286
## 24 0.69230769 0.71428571 0.85714286 0.85714286 1.00000000 0.77777778 0.78947368
## 25 0.69230769 0.57142857 0.85714286 0.85714286 1.00000000 0.66666667 0.68421053
## 26 0.69230769 0.70370370 0.85185185 0.85185185 0.92592593 0.67741935 0.62500000
## 27 0.70370370 0.71428571 0.85714286 0.85714286 0.92857143 0.68750000 0.63636364
## 28 0.70370370 0.71428571 0.85714286 0.85714286 0.92857143 0.68750000 0.63636364
## 29 0.67741935 0.62500000 0.68750000 0.68750000 0.68750000 0.50000000 0.45945946
## 30 0.76923077 0.77777778 0.92592593 0.92592593 1.00000000 0.74193548 0.68750000
##            16         17         18         19         20         21         22
## 2
## 3
## 4
## 5
## 6
## 7
## 9
## 10
## 11
## 12
## 13
## 14
## 15
## 16
## 17 0.12820513
## 18 0.15000000 0.02222222
## 19 0.25000000 0.15555556 0.13043478
## 20 0.28205128 0.18181818 0.15555556 0.02222222
## 21 0.30000000 0.20000000 0.17391304 0.04347826 0.02222222
## 22 0.33333333 0.22727273 0.20000000 0.06666667 0.04545455 0.02222222
## 23 0.80000000 0.76000000 0.76923077 0.76923077 0.76000000 0.76923077 0.76000000
## 24 0.68000000 0.60000000 0.54838710 0.48387097 0.46666667 0.48387097 0.46666667
## 25 0.60000000 0.60000000 0.54838710 0.48387097 0.46666667 0.48387097 0.46666667
## 26 0.36842105 0.25581395 0.22727273 0.09090909 0.06976744 0.04545455 0.02325581
## 27 0.33333333 0.22727273 0.20000000 0.06666667 0.04545455 0.02222222 0.00000000
## 28 0.33333333 0.22727273 0.20000000 0.06666667 0.04545455 0.02222222 0.00000000
## 29 0.25581395 0.12500000 0.10204082 0.06122449 0.08333333 0.06122449 0.08333333
## 30 0.36842105 0.25581395 0.22727273 0.09090909 0.06976744 0.04545455 0.02325581
##            23         24         25         26         27         28         29
## 2
```

```
## 3
## 4
## 5
## 6
## 7
## 9
## 10
## 11
## 12
## 13
## 14
## 15
## 16
## 17
## 18
## 19
## 20
## 21
## 22
## 23
## 24 0.45454545
## 25 0.45454545 0.37500000
## 26 0.75000000 0.44827586 0.44827586
## 27 0.76000000 0.46666667 0.46666667 0.02325581
## 28 0.76000000 0.46666667 0.46666667 0.02325581 0.00000000
## 29 0.79310345 0.52941176 0.52941176 0.10638298 0.08333333 0.08333333
## 30 0.75000000 0.44827586 0.44827586 0.04761905 0.02325581 0.02325581 0.10638298
```

```
print("Resemblance Matrices using Bray-Curtis Distance")
```

```
## [1] "Resemblance Matrices using Bray-Curtis Distance"
```

```
fish.db
```

```
##             1          2          3          4          5          6          7
## 2  0.60000000
## 3  0.68421053 0.14285714
## 4  0.75000000 0.33333333 0.18918919
## 5  0.89189189 0.69565217 0.68000000 0.49090909
## 6  0.75000000 0.39393939 0.29729730 0.19047619 0.41818182
## 7  0.68421053 0.14285714 0.12500000 0.24324324 0.64000000 0.24324324
## 9  1.00000000 0.69230769 0.73333333 0.65714286 0.58333333 0.54285714 0.66666667
## 10 0.88235294 0.38461538 0.40000000 0.37142857 0.54166667 0.25714286 0.26666667
## 11 0.57142857 0.30434783 0.40740741 0.43750000 0.68888889 0.43750000 0.33333333
## 12 0.71428571 0.20000000 0.23529412 0.33333333 0.69230769 0.38461538 0.17647059
## 13 0.72727273 0.29032258 0.31428571 0.45000000 0.73584906 0.55000000 0.37142857
## 14 0.80645161 0.40000000 0.31818182 0.34693878 0.67741935 0.42857143 0.36363636
## 15 0.83333333 0.51111111 0.46938776 0.40740741 0.55223881 0.37037037 0.38775510
## 16 0.86046512 0.65384615 0.57142857 0.47540984 0.45945946 0.37704918 0.53571429
## 17 0.91489362 0.67857143 0.63333333 0.50769231 0.51282051 0.44615385 0.60000000
## 18 0.95555556 0.74074074 0.72413793 0.58730159 0.50000000 0.52380952 0.68965517
## 19 1.00000000 0.79310345 0.70967742 0.61194030 0.50000000 0.52238806 0.67741935
## 20 1.00000000 0.91176471 0.88888889 0.74025974 0.48888889 0.68831169 0.86111111
```

```
## 21 1.00000000 0.94594595 0.92307692 0.78313253 0.50000000 0.73493976 0.89743590
## 22 1.00000000 0.97619048 0.95454545 0.82795699 0.52830189 0.78494624 0.93181818
## 23 1.00000000 1.00000000 1.00000000 0.92000000 0.89473684 0.84000000 0.90000000
## 24 1.00000000 1.00000000 1.00000000 0.88888889 0.79591837 0.77777778 0.93548387
## 25 1.00000000 1.00000000 0.92592593 0.81250000 0.68888889 0.68750000 0.85185185
## 26 1.00000000 0.96363636 0.93220339 0.78125000 0.55844156 0.68750000 0.89830508
## 27 1.00000000 0.97333333 0.94936709 0.83333333 0.56701031 0.76190476 0.92405063
## 28 1.00000000 0.97560976 0.95348837 0.82417582 0.57692308 0.78021978 0.93023256
## 29 0.97777778 0.93939394 0.92233010 0.81481481 0.53719008 0.77777778 0.90291262
## 30 1.00000000 1.00000000 0.98095238 0.87272727 0.59349593 0.83636364 0.96190476
##             9         10         11         12         13         14         15
## 2
## 3
## 4
## 5
## 6
## 7
## 9
## 10 0.57142857
## 11 0.76000000 0.44000000
## 12 0.68750000 0.37500000 0.24137931
## 13 0.81818182 0.57575758 0.33333333 0.18918919
## 14 0.76190476 0.47619048 0.43589744 0.21739130 0.19148936
## 15 0.65957447 0.40425532 0.50000000 0.33333333 0.38461538 0.24590164
## 16 0.70370370 0.51851852 0.64705882 0.55172414 0.59322034 0.44117647 0.26027397
## 17 0.68965517 0.51724138 0.63636364 0.58064516 0.61904762 0.50000000 0.40259740
## 18 0.64285714 0.57142857 0.69811321 0.66666667 0.70491803 0.60000000 0.46666667
## 19 0.66666667 0.63333333 0.82456140 0.75000000 0.81538462 0.67567568 0.56962025
## 20 0.68571429 0.77142857 0.91044776 0.89189189 0.92000000 0.83333333 0.70786517
## 21 0.76315789 0.81578947 0.91780822 0.92500000 0.95061728 0.86666667 0.76842105
## 22 0.76744186 0.86046512 0.95180723 0.95555556 0.97802198 0.90000000 0.77142857
## 23 0.77777778 0.88888889 0.86666667 0.90909091 1.00000000 0.93750000 0.94594595
## 24 0.72413793 0.79310345 0.92307692 0.93939394 1.00000000 0.90697674 0.87500000
## 25 0.84000000 0.76000000 0.90909091 0.93103448 1.00000000 0.84615385 0.81818182
## 26 0.71929825 0.82456140 0.92592593 0.93442623 0.96774194 0.85915493 0.76315789
## 27 0.76623377 0.84415584 0.94594595 0.95061728 0.97560976 0.89010989 0.77083333
## 28 0.76190476 0.85714286 0.95061728 0.95454545 0.97752809 0.89795918 0.78640777
## 29 0.78217822 0.84158416 0.89795918 0.90476190 0.90566038 0.84347826 0.73333333
## 30 0.84466019 0.90291262 0.98000000 0.98130841 1.00000000 0.93162393 0.81967213
##            16         17         18         19         20         21         22
## 2
## 3
## 4
## 5
## 6
## 7
## 9
## 10
## 11
## 12
## 13
## 14
## 15
## 16
```

```
## 17 0.26190476
## 18 0.34146341 0.13953488
## 19 0.39534884 0.31111111 0.25000000
## 20 0.58333333 0.42000000 0.32653061 0.23529412
## 21 0.62745098 0.49056604 0.40384615 0.29629630 0.10169492
## 22 0.66071429 0.55172414 0.47368421 0.38983051 0.18750000 0.10447761
## 23 0.90909091 0.83333333 0.82608696 0.84000000 0.86666667 0.87878788 0.89473684
## 24 0.81818182 0.69491525 0.64912281 0.63934426 0.57746479 0.61038961 0.65517241
## 25 0.76470588 0.74545455 0.66037736 0.61403509 0.67164179 0.69863014 0.73493976
## 26 0.63855422 0.54022989 0.45882353 0.32584270 0.21212121 0.20000000 0.25217391
## 27 0.66990291 0.57009346 0.48571429 0.37614679 0.19327731 0.13600000 0.12592593
## 28 0.69090909 0.57894737 0.50000000 0.41379310 0.22222222 0.16666667 0.12676056
## 29 0.65354331 0.51145038 0.44186047 0.41353383 0.24475524 0.18120805 0.11949686
## 30 0.72093023 0.57894737 0.52671756 0.48148148 0.29655172 0.23178808 0.18012422
##            23         24         25         26         27         28         29
## 2
## 3
## 4
## 5
## 6
## 7
## 9
## 10
## 11
## 12
## 13
## 14
## 15
## 16
## 17
## 18
## 19
## 20
## 21
## 22
## 23
## 24 0.57894737
## 25 0.46666667 0.46153846
## 26 0.82978723 0.48275862 0.59259259
## 27 0.88059701 0.61538462 0.70270270 0.18867925
## 28 0.89189189 0.64705882 0.72839506 0.23893805 0.09774436
## 29 0.91208791 0.70588235 0.77551020 0.33846154 0.18666667 0.14649682
## 30 0.91397849 0.71153846 0.78000000 0.36363636 0.19736842 0.15723270 0.14772727
```

Question 5: Using the distance matrices from above, answer the following questions:

a. Does the resemblance matrix (`fish.db`) represent similarity or dissimilarity? What information in the resemblance matrix led you to arrive at your answer?

b. Compare the resemblance matrices (`fish.db` or `fish.ds`) you just created. How does the choice of the Sørensen or Bray-Curtis distance influence your interpretation of site (dis)similarity?

Answer 5a:
The resemblance matrix `fish.db` represents dissimilarity. The values in the matrix are between

0 and 1, where 0 indicates that two sites are similar and 1 indicates that two sites are completely different. If we look at the formula of Bray-Curtis, we can see that the formula is a sum of the absolute differences between the species abundances of two sites. If two sites are similar in terms of species composition (not only presence of that species but also similar in the count), the difference would be 0, or at least close to 0, the Bray-Curtis distance will be low. If two sites are different, the Bray-Curtis distance will be high.

Answer 5b: The choice of the Sørensen or Bray-Curtis distance influences the interpretation of site similarity. Sørensen's similarity is a binary measure that only considers the presence or absence of species. It is only sensitive to the number of shared species between two sites. Bray-Curtis distance, on the other hand, considers the abundance of species in addition to their presence or absence. It is sensitive to both the number of shared species and their abundnance in two sites. We can see for instance the Sorensen's similarity index between sites 1 and 4 is 0.77, which means that sites 1 and 2 share 77% of the common species. However, there's a big difference in the species abundance between the two sites. The Bray-Curtis distance between sites 1 and 4 is 0.75, which indicates that the two sites are very different in terms of species abundance.

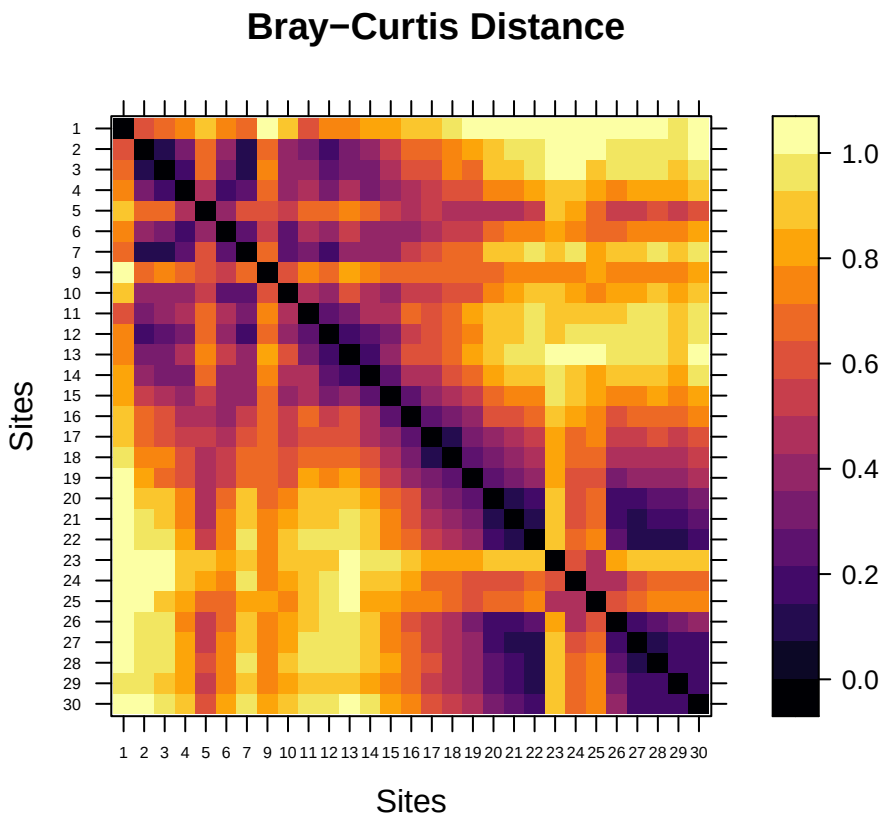
4) VISUALIZING BETA-DIVERSITY

A. Heatmaps

In the R code chunk below, do the following:

1. define a color palette,
2. define the order of sites in the Doubs River, and
3. use the `levelplot()` function to create a heatmap of fish abundances in the Doubs River.

```
# Define the order of the sites
order = rev(attr(fish.db, "Labels"))

# Heatmap
levelplot(as.matrix(fish.db)[, order],
  col.regions = inferno,
  aspect="iso",
  xlab = "Sites", ylab = "Sites",
  scales = list(cex=0.5),
  main = "Bray-Curtis Distance")
```

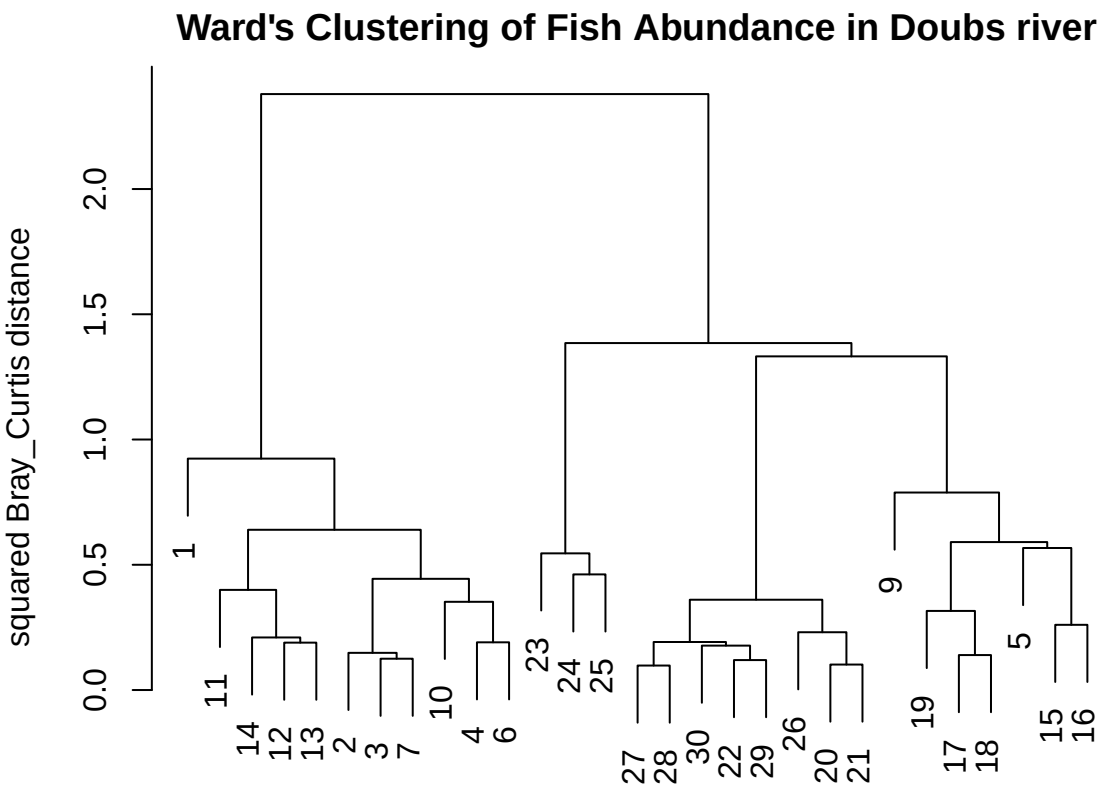


B. Cluster Analysis

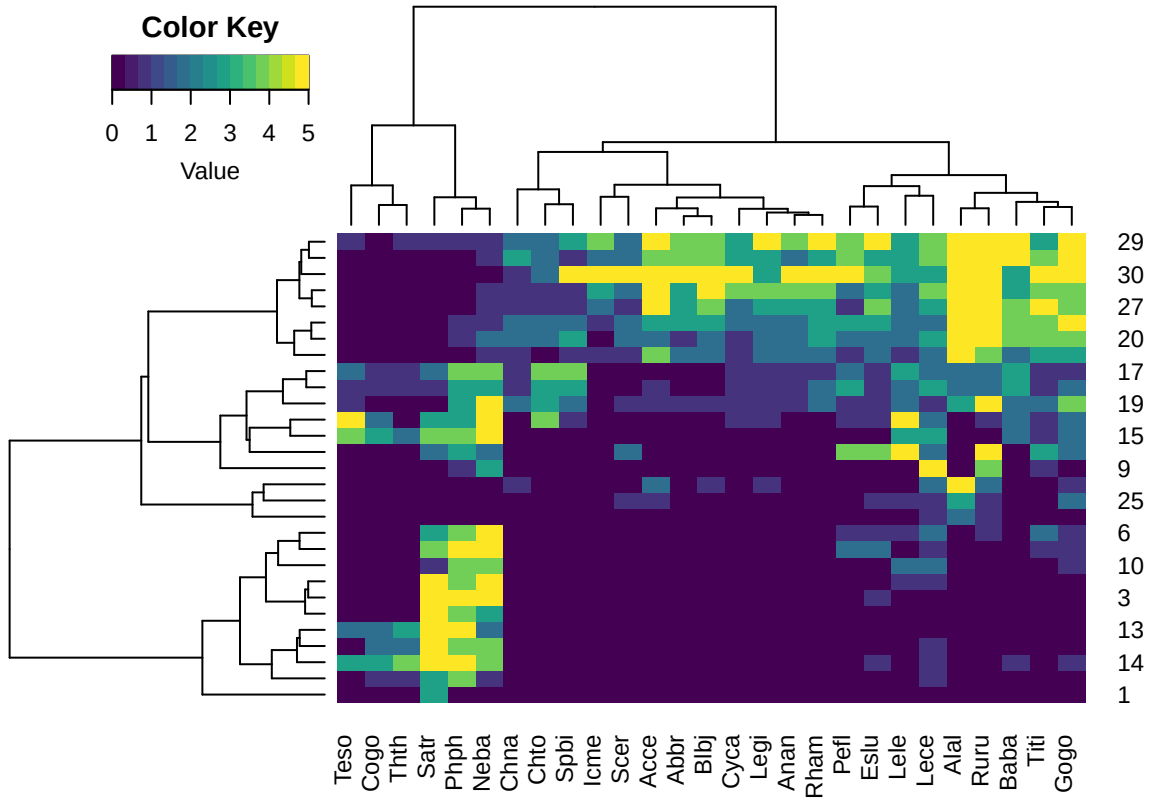
In the R code chunk below, do the following:

1. perform a cluster analysis using Ward's Clustering, and
2. plot your cluster analysis (use either `hclust` or `heatmap.2`).

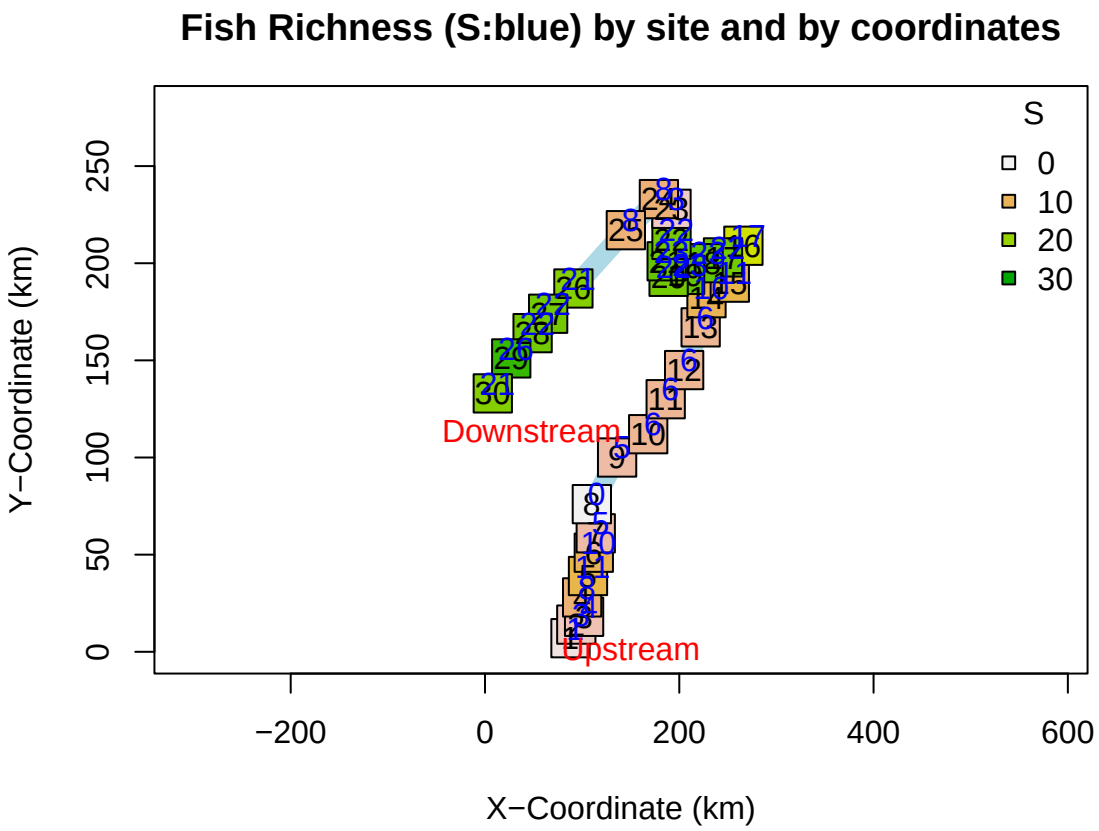
```
# Perform cluster analysis
fish.ward = hclust(fish.db, method = "ward.D2")

# Plot cluster
par(mar=c(1,5,2,2) + 0.1)
plot(fish.ward, main = "Ward's Clustering of Fish Abundance in Doubs river",
  ylab = "squared Bray_Curtis distance")
```




```
gplots::heatmap.2(as.matrix(fish),
   distfun = function(x) vegdist(x, method = "bray"),
   hclustfun = function(x) hclust(x, method = "ward.D2"),
   col = viridis, trace = "none", density.info = "none")
```




```
# ----------------------------------------
# Replot the distribution of species richness by location
opar <- par(no.readonly = TRUE)
par(mfrow = c(1,  1), mar = c(4, 4, 3, 4) + 0.1, xpd = TRUE)

# Species Richness
plot.new()
plot(doubs$xy, asp = 2, type = 'l', lwd = 10, col = "light blue", xlim = c(0,280), ylim = c(0,280),
     main = "Fish Richness (S:blue) by site and by coordinates",
     xlab = "X-Coordinate (km)", ylab = "Y-Coordinate (km)")
points(doubs$xy, pch = 22, cex=3, bg = spa.S.color[spa.S + 1])
text(doubs$xy, as.character(rownames(doubs$xy)), cex = 1, col = "black")
text(doubs$xy + 5, as.character(spa.S), cex = 1, col = "blue")
text(150, 0, "Upstream", cex = 1, col = "red")
text(48, 114, "Downstream", cex = 1, col = "red")
legend("topright", pch = 22, pt.cex = 1, bty = 'n',
       title = "S", legend = seq(0, 30, 10),
       pt.bg = spa.S.color[seq(1, 31, 10)])
```




```
par = opar
```

Question 6: Based on cluster analyses and the introductory plots that we generated after loading the data, develop an ecological hypothesis for fish diversity the `doubs` data set?

> ***Answer 6***: In this clustering plot, we can see that nearer to the downstream, we have more species. This suggests that the fish community composition is different between the upstream and downstream of the Doubs River. This could be due maybe to differences in habitat quality, water quality, or other environmental factors. Based on this, we can hypothesize that the fish community composition in the Doubs River changes along the river gradient.

C. Ordination

Principal Coordinates Analysis (PCoA)

In the R code chunk below, do the following:

1. perform a Principal Coordinates Analysis to visualize beta-diversity
2. calculate the variation explained by the first three axes in your ordination
3. plot the PCoA ordination,
4. label the sites as points using the Doubs River site number, and

```r
# Calculate PCoA
fish.pcoa = cmdscale(fish.db, eig = TRUE, k = 3)
fish.pcoa
```

```
## $points
##           [,1]         [,2]         [,3]
## 1  -0.32043291  0.135094526  0.610444135
## 2  -0.50271272 -0.053204485  0.115084582
## 3  -0.47207593 -0.075509864  0.025191639
## 4  -0.31629787 -0.065793925 -0.054743001
## 5   0.06659050 -0.103685307 -0.061632012
## 6  -0.24164704  0.022330383 -0.117689415
## 7  -0.44197058  0.013721347 -0.019578989
## 9  -0.02864444  0.138633466 -0.189715336
## 10 -0.29305689  0.037682458 -0.202079336
## 11 -0.43875065  0.071482520  0.153768001
## 12 -0.47279103 -0.013160075 -0.006637071
## 13 -0.47550783 -0.100573921  0.079082050
## 14 -0.37472705 -0.070111200 -0.076607528
## 15 -0.24370479 -0.143518516 -0.131459333
## 16 -0.07348942 -0.157547743 -0.149621439
## 17  0.04061679 -0.130989703 -0.133635222
## 18  0.14445084 -0.101088883 -0.146573604
## 19  0.22960480 -0.102509909 -0.156586724
## 20  0.40100791 -0.101317037 -0.022574780
## 21  0.43837569 -0.111705508  0.048669123
## 22  0.46516232 -0.124748909  0.094770962
## 23  0.11656630  0.659234875 -0.006196177
## 24  0.32004280  0.418009171 -0.030531755
## 25  0.22601066  0.463713151 -0.110983530
## 26  0.43466174  0.005166762  0.006723518
## 27  0.46255346 -0.090842861  0.086824252
## 28  0.46286016 -0.106465286  0.108084941
## 29  0.40942763 -0.173891275  0.123971141
## 30  0.47787757 -0.138404251  0.164230911
##
## $eig
##  [1]  3.695331e+00  1.098472e+00  7.104740e-01  4.149729e-01  3.045604e-01
##  [6]  1.917884e-01  1.569703e-01  1.319099e-01  1.294251e-01  8.667896e-02
## [11]  4.615780e-02  3.864487e-02  2.745800e-02  1.306508e-02  7.087896e-03
## [16]  4.039469e-03  1.300594e-03  2.570914e-17 -3.534426e-05 -3.940676e-03
## [21] -8.956051e-03 -1.461149e-02 -1.598905e-02 -2.145686e-02 -3.017013e-02
## [26] -3.432671e-02 -3.760052e-02 -6.037087e-02 -6.880061e-02
##
## $x
## NULL
##
## $ac
## [1] 0
##
## $GOF
## [1] 0.7484133 0.7798263
```

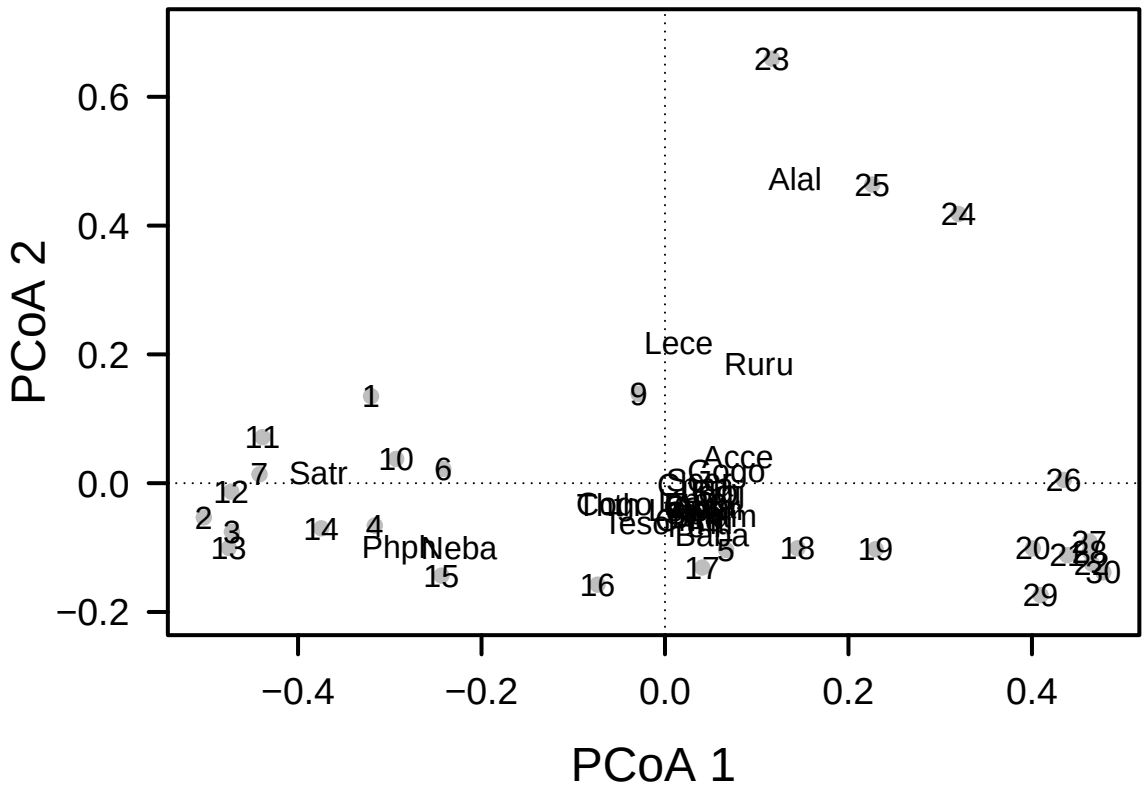
```r
# Explained variance for first 3 axes
explained_variance = round(sum(fish.pcoa$eig[1:3]/sum(fish.pcoa$eig)) * 100, 2)
print("Explained variance of the 3 first axes:")
```

```
## [1] "Explained variance of the 3 first axes:"
```

```r
print(explained_variance)
```

```
## [1] 81.4
```

```r
# Define Plot Parameters
par(mar = c(5, 5, 1, 2) + 0.1)

# Initiate Plot
plot(fish.pcoa$points[ ,1], fish.pcoa$points[ ,2], ylim = c(-0.2, 0.7),
     xlab = "PCoA 1",
     ylab = "PCoA 2",
     pch = 16, cex = 2.0, type = "n", cex.lab = 1.5,
     cex.axis = 0.5, axes = FALSE)

# # Add Axes
axis(side = 1, labels = T, lwd.ticks = 2, cex.axis = 1.2, las = 1)
axis(side = 2, labels = T, lwd.ticks = 2, cex.axis = 1.2, las = 1)
abline(h = 0, v = 0, lty = 3)
box(lwd = 2)

# # Add Points & Labels
points(fish.pcoa$points[ ,1], fish.pcoa$points[ ,2],
       pch = 19, cex = 1, bg = "gray", col = "gray")
text(fish.pcoa$points[ ,1], fish.pcoa$points[ ,2],
     labels = row.names(fish.pcoa$points))


## Layer the fishes on the PCA plot
## ================================================
# Relative abundance of each fsih
fishREL =fish
  for(i in 1:nrow(fish)){
    fishREL[i, ] = fish[i, ] / sum(fish[i, ])
    }

# Now, we use this information to calculate and add species scores
fish.pcoa = add.spec.scores(
  fish.pcoa,
  fishREL,
  method = "pcoa.scores")
# fish.pcoa
text(
  fish.pcoa$cproj[ ,1],
  fish.pcoa$cproj[ ,2],
  labels = row.names(fish.pcoa$cproj), col = "black")
```




```
## ================================================
# Quantitiatively determine the names of the influential species
spe.corr = add.spec.scores(fish.pcoa, fishREL, method = "cor.scores")$cproj
thres = 0.7
spe.corr # correlation of each species with the first two axes
```

```
##             Dim1        Dim2         Dim3
## Cogo -0.51482765 -0.18215901 -0.072493771
## Satr -0.65755705  0.01568556  0.673590464
## Phph -0.86746400 -0.16993159 -0.124630984
## Neba -0.76741139 -0.18556782 -0.369638303
## Thth -0.49177919 -0.14695731  0.008482829
## Teso -0.27862754 -0.30397390 -0.256885094
## Chna  0.67852376 -0.02836068 -0.036473404
## Chto  0.32879529 -0.38699081 -0.269705339
## Lele  0.12018082 -0.18717507 -0.448352524
## Lece  0.07290236  0.57840522 -0.396459547
## Baba  0.62696172 -0.53115226 -0.100687700
## Spbi  0.50746311 -0.39531008 -0.133488708
## Gogo  0.60796871  0.07542782 -0.296228387
## Eslu  0.36212467 -0.19005923 -0.085718516
## Pefl  0.36765239 -0.41186011 -0.108247401
## Rham  0.80887508 -0.41925666  0.141363010
## Legi  0.82017586 -0.17018026  0.124239408
## Scer  0.59181176  0.05518457  0.029158367
```

```
## Cyca  0.75951224 -0.44429261  0.173136584
## Titi  0.56832579 -0.36831053 -0.138049853
## Abbr  0.77047437 -0.34527139  0.292778034
## Icme  0.65659923 -0.30970901  0.351493014
## Acce  0.76351955  0.21557646  0.102881794
## Ruru  0.51227569  0.50525351 -0.241837628
## Blbj  0.81184827 -0.13246979  0.255811784
## Alal  0.44712826  0.81198434 -0.051671308
## Anan  0.79741215 -0.39189718  0.209449676
```

```
# Identify influential species
imp.spp = spe.corr[abs(spe.corr[,1]) >= thres | abs(spe.corr[,2]) >= thres, ]
imp.spp
```

```
##            Dim1       Dim2        Dim3
## Phph -0.8674640 -0.1699316 -0.12463098
## Neba -0.7674114 -0.1855678 -0.36963830
## Rham  0.8088751 -0.4192567  0.14136301
## Legi  0.8201759 -0.1701803  0.12423941
## Cyca  0.7595122 -0.4442926  0.17313658
## Abbr  0.7704744 -0.3452714  0.29277803
## Acce  0.7635195  0.2155765  0.10288179
## Blbj  0.8118483 -0.1324698  0.25581178
## Alal  0.4471283  0.8119843 -0.05167131
## Anan  0.7974122 -0.3918972  0.20944968
```

```
# Check by permutation test
fit = envfit(fish.pcoa, fishREL, perm=999)
fit
```

```
##
## ***VECTORS
##
##          Dim1     Dim2     r2 Pr(>r)
## Cogo -0.83884 -0.54438 0.2982  0.011 *
## Satr -0.99904  0.04371 0.4326  0.003 **
## Phph -0.94110 -0.33813 0.7814  0.001 ***
## Neba -0.91413 -0.40543 0.6234  0.001 ***
## Thth -0.87692 -0.48063 0.2634  0.024 *
## Teso -0.44704 -0.89452 0.1700  0.087 .
## Chna  0.99707 -0.07644 0.4612  0.003 **
## Chto  0.42032 -0.90738 0.2579  0.027 *
## Lele  0.33041 -0.94384 0.0495  0.539
## Lece  0.06856  0.99765 0.3399  0.012 *
## Baba  0.54118 -0.84091 0.6752  0.001 ***
## Spbi  0.57341 -0.81927 0.4138  0.002 **
## Gogo  0.97507  0.22188 0.3753  0.004 **
## Eslu  0.72044 -0.69352 0.1673  0.107
## Pefl  0.43762 -0.89916 0.3048  0.010 **
## Rham  0.72476 -0.68901 0.8301  0.001 ***
## Legi  0.93461 -0.35568 0.7016  0.001 ***
## Scer  0.98569  0.16858 0.3533  0.006 **
## Cyca  0.68181 -0.73153 0.7743  0.001 ***
```

```
## Titi  0.64378 -0.76521 0.4586  0.003 **
## Abbr  0.77254 -0.63497 0.7128  0.001 ***
## Icme  0.75626 -0.65427 0.5270  0.001 ***
## Acce  0.88799  0.45986 0.6294  0.001 ***
## Ruru  0.48379  0.87518 0.5177  0.002 **
## Blbj  0.95802 -0.28671 0.6766  0.001 ***
## Alal  0.28755  0.95777 0.8592  0.001 ***
## Anan  0.74277 -0.66954 0.7894  0.001 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Permutation: free
## Number of permutations: 999
```

5. identify influential species and add species coordinates to PCoA plot. > Influential species are `Satr`, `Phph`, `Neba`, `Alal`, `Lece`, `Ruru`.

In the R code chunk below, do the following:

1. identify influential species based on correlations along each PCoA axis (use a cutoff of 0.70), and
2. use a permutation test (999 permutations) to test the correlations of each species along each axis.

Question 7: Address the following questions about the ordination results of the `doubs` data set:

a. Describe the grouping of sites in the Doubs River based on fish community composition.
b. Generate a hypothesis about which fish species are potential indicators of river quality.

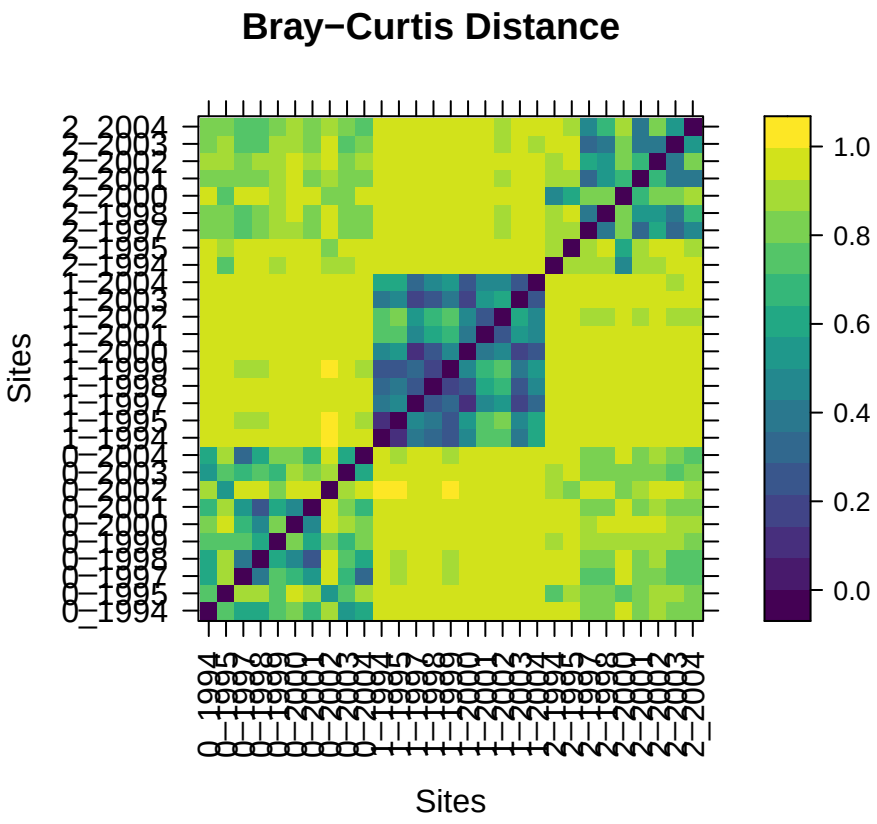
> ***Answer 7a***:
> We can see that the sites are grouped into approximately three main clusters. Top right quadrant, bottom right quadrant, and the left side of the plot. The sites in the top right quadrant (23, 25, 24) are more similar to each other in terms of fish community composition. These sites are positioned far from the central (0,0), suggesting that their fish communities are significantly different from the rest. In these sites, the main species is `Alal`. The sites in the bottom left quadrant are also similar to each other in terms of fish community composition. The main species in these sites is `Satr`, `Phph`, `Neba`. The sites in the bottom right quadrant are also similar to each other in terms of fish community composition. The main species in these sites is `Lece`, `Ruru`. ***Answer 7b***:
> Due to the number of distinct species that are present in the lower part of the plot, compared to the lower number of species in the higher part of the plot, I think that the quality of the water in the site 25, 24, 23 may be poorer than that of the sites in the lower quadrants.

SYNTHESIS

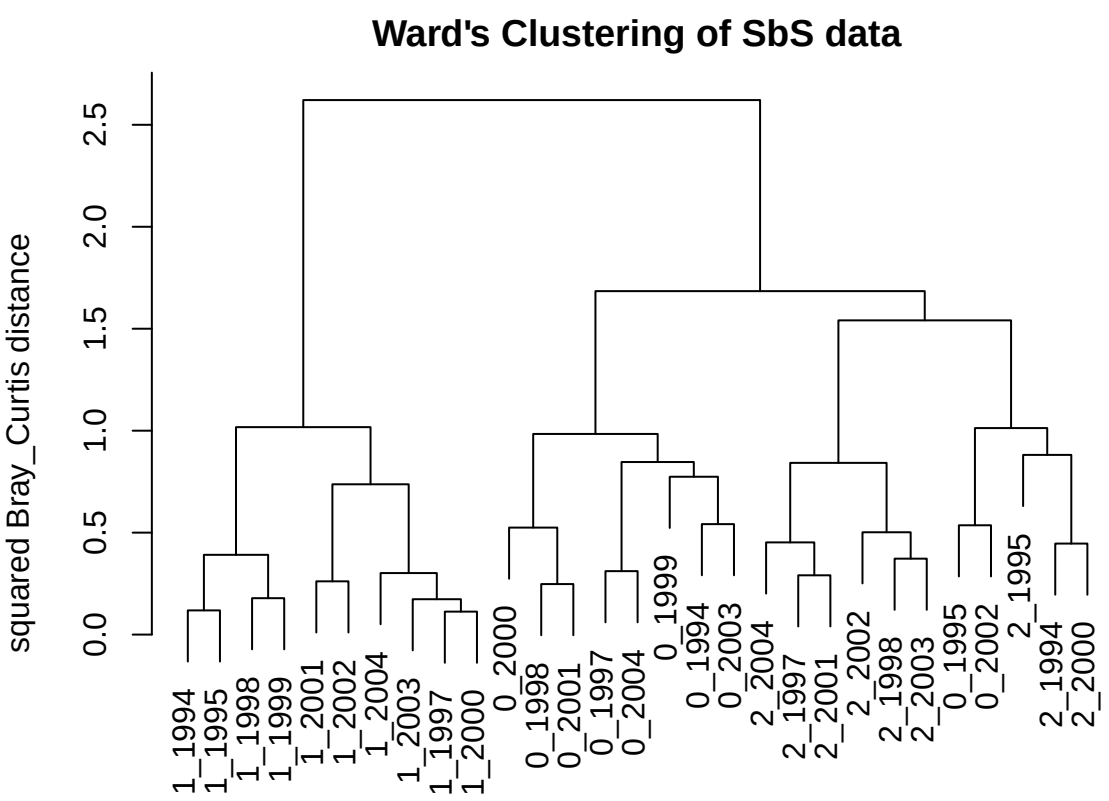
Load the dataset from that you and your partner are using for the team project. Use one of the tools introduced in the beta diversity module to visualize your data. Describe any interesting patterns and identify a hypothesis is relevant to the principles of biodiversity.

```r
# Load SbS dataset
data = read.csv("./SbS_full.csv", row.names= 1)

# Calculate Bray-Curtis distance
data.db = vegdist(data, method = "bray")
```

```
site_labels = rownames(data)

################################################3
# Heatmap of dissimilarities
levelplot(as.matrix(data.db),
  col.regions = viridis,
  aspect="iso",
  xlab = "Sites", ylab = "Sites",
  scales = list(cex=1,
  x=list(labels=site_labels, rot=90),
  y=list(labels=site_labels)),

  main = "Bray-Curtis Distance")
```




```
#################################################################3
# Perform cluster analysis
data.ward = hclust(data.db, method = "ward.D2")

# Plot cluster
par(mar=c(1,5,2,2) + 0.1)
plot(data.ward, main = "Ward's Clustering of SbS data",
  ylab = "squared Bray_Curtis distance")
```




```
##########################################################
# PCOA analysis for all sites
data.pcoa = cmdscale(data.db, eig = TRUE, k = 3)
data.pcoa
```

```
## $points
##              [,1]          [,2]         [,3]
## 0_1994 -0.2621034 -0.2699149247  0.074385720
## 0_1995 -0.2359875  0.1595108010  0.344080961
## 0_1997 -0.2932861 -0.3466560108 -0.101197617
## 0_1998 -0.2843073 -0.4276652457 -0.022757800
## 0_1999 -0.2787567 -0.1875842322  0.214850543
## 0_2000 -0.1968760 -0.4002278443  0.074979920
## 0_2001 -0.2911126 -0.4241363107 -0.002957168
## 0_2002 -0.1810289  0.1689894425  0.429592708
## 0_2003 -0.2715348 -0.0947729159  0.083963107
## 0_2004 -0.2536514 -0.3176398880 -0.032678583
## 1_1994  0.4785794 -0.0342757754 -0.006622986
## 1_1995  0.4496789 -0.0833756088 -0.004812516
## 1_1997  0.5275647 -0.0091488844 -0.044266100
## 1_1998  0.5232734 -0.0211786424 -0.016732378
## 1_1999  0.4782747 -0.0766809969 -0.014181681
## 1_2000  0.5258441  0.0031358443 -0.027164795
## 1_2001  0.4180193  0.0467956632 -0.006448752
## 1_2002  0.3339251  0.0615793860 -0.046432316
```

```
## 1_2003  0.5299988 -0.0001139364 -0.018876707
## 1_2004  0.4694560  0.0145119790 -0.046857413
## 2_1994 -0.1607187  0.2876104811  0.344095705
## 2_1995 -0.1294679  0.2737598386  0.310008007
## 2_1997 -0.3067167  0.2245664276 -0.351108294
## 2_1998 -0.2760078  0.2372780730 -0.320075152
## 2_2000 -0.2232449  0.3901000122  0.309488417
## 2_2001 -0.3004076  0.2578313078 -0.253731622
## 2_2002 -0.2414980  0.2521335308 -0.248891083
## 2_2003 -0.2830456  0.1993900669 -0.405471601
## 2_2004 -0.2648624  0.1161783626 -0.214180522
##
## $eig
##  [1]  3.500990e+00  1.535678e+00  1.259760e+00  7.007179e-01  5.987093e-01
##  [6]  5.557639e-01  4.598719e-01  4.169859e-01  2.993459e-01  2.604848e-01
## [11]  1.599484e-01  1.426802e-01  1.146544e-01  1.023091e-01  7.670019e-02
## [16]  7.110455e-02  4.697578e-02  4.041310e-02  2.855480e-02  1.651146e-02
## [21]  1.075181e-02  7.489826e-03  4.448349e-03  1.493358e-03  2.436551e-04
## [26]  1.561251e-16 -6.788577e-03 -1.486261e-02 -1.617749e-02
##
## $x
## NULL
##
## $ac
## [1] 0
##
## $GOF
## [1] 0.6025050 0.6046939
```

```
# Explained variance for first 3 axes
explained_variance = round(sum(data.pcoa$eig[1:3]/sum(data.pcoa$eig)) * 100, 2)
print("Explained variance of the 3 first axes:")
```

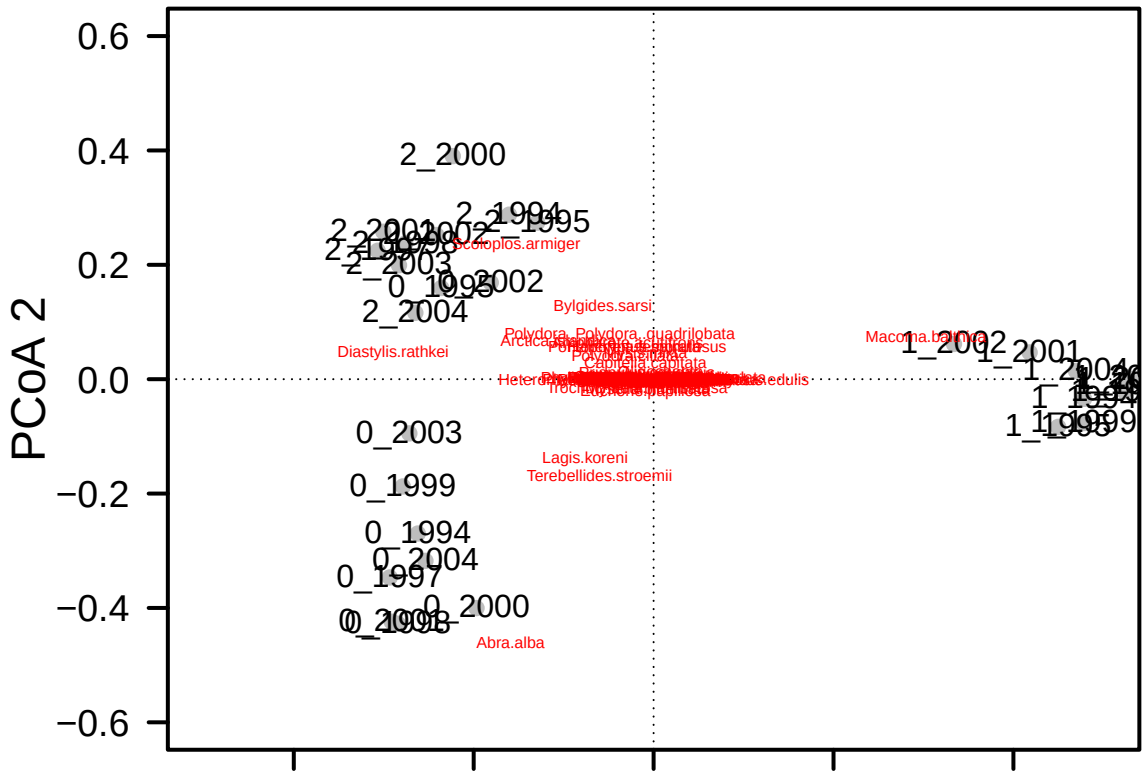
```
## [1] "Explained variance of the 3 first axes:"
```

```
print(explained_variance)
```

```
## [1] 60.69
```

```
# Define Plot Parameters
par(mar=c(1,5,2,2) + 0.1)

# Initiate Plot
plot(data.pcoa$points[ ,1],
  data.pcoa$points[ ,2],
     xlab = "PCoA 1",
     ylab = "PCoA 2",
     xlim = c(-0.5, 0.5),
     ylim = c(-0.6, .6),
     pch = 16, cex = 0.5,  type = "n", cex.lab = 1.5,
     cex.axis = 0.5, axes = FALSE)
```

```r
# # Add Axes
axis(side = 1, labels = T, lwd.ticks = 2, cex.axis = 1.2, las = 1)
axis(side = 2, labels = T, lwd.ticks = 2, cex.axis = 1.2, las = 1)
abline(h = 0, v = 0, lty = 3)
box(lwd = 2)

# # Add Points & Labels
points(data.pcoa$points[ ,1],
  data.pcoa$points[ ,2],
  pch = 19, cex = 1, bg = "gray", col = "gray")
text(data.pcoa$points[ ,1], data.pcoa$points[ ,2],
     labels = row.names(data.pcoa$points))

# Add species on top
dataREL <- data
  for(i in 1:nrow(data)){
    dataREL[i, ] = data[i, ] / sum(data[i, ])
    }

# Now, we use this information to calculate and add species scores
data.pcoa <- add.spec.scores(
  data.pcoa,
  dataREL,
  method = "pcoa.scores")

text(
  data.pcoa$cproj[ ,1],
  data.pcoa$cproj[ ,2],
  labels = row.names(data.pcoa$cproj), col = "red",
  cex=0.5)
```



Answer Cluster Analysis We see that some sites cluster consistently by location over different years. For example, site 1 contains all the years sampled, this suggests that the zoobenthos community composition is stable over time. Howeverm some other sites (0 and 2) have some temporal shifts. For instance, site 0_1995 and 0_2002 are clustered closer to sites 2_1994 2_1995, 2_2000, suggesting that there might be environmental changes, disturbances, or shifts in the zoobenthos community composition over time. ***Bray_Curtis Distance Heatmap*** We can see shows distinct blocks of similarity between the same site being sampled over years. This was expected. However, we find again that in the early years of site 2 samples (1994,95), site 2 had a zoobenthos biodiversity composition more similar to site 0 of the same year. In conclusion, for sites that show changes over time, we can think that they are locaed in more disturbed areas or are more exposed to environmental changes (pollutionm climate changes, invasive species?) ***PCOA analysis*** Here, we see that the sites 0, 1, 2 parated across the PCoA space, suggesting they have distinct community compositions. Site 1 (e.g., 1_1994, 1_1995, 1_1999, etc.) appears to form a tight cluster, which means that they have a relatively stable species composition over time. Sites 2 and 0(e.g., 2_1994, 2_1995, etc.) are more dispersed across multiple quadrants. This suggests that "these sites" have bigger variation in community composition.

I think species nearer to the center are more present in most sites. While species far from the origin are more specialized or strongly associated with particular site conditions. There are several patterns : Macoma balthica is strongly associated with site 1. This suggest that this species may adapt well in this site' specific condition. Scoloplos armiger and Bylgides sarsi are more abundant in these site 2. Abra alba, Lagis koreni, and Terebellides stroemii are characteristic site 0.