6. Worksheet: Among Site (Beta) Diversity – Part 2

Yongsoo Choi; Z620: Quantitative Biodiversity, Indiana University

12 February, 2025

OVERVIEW

In this worksheet, we continue to explore concepts, statistics, and visualizations related to β-diversity. Now that you know how to formally quantify β-diversity, we will learn how to 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 today, you should **push** this file to your GitHub repo, at whatever stage you are. This will enable you to pull your work onto your own computer.
6. When you have completed the worksheet, **Knit** the text and code into a single PDF file by pressing the `Knit` button in the RStudio scripting panel. This will save the PDF output in your Posit.cloud workspace: `/cloud/project/QB-2025/Week4-Beta/`
7. After Knitting, please submit the worksheet by making a **push** to your GitHub repo and then create a **pull request** via GitHub. Your pull request should include this file (**6.BetaDiversity_2_Worksheet.Rmd**) with all code blocks filled out and questions answered) and the PDF output of `Knitr` (**6.BetaDiversity_2_Worksheet.pdf**).

The completed exercise is due on **Wednesday, February 12th, 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, provide the code to:

1. clear your R environment,
2. print your current working directory,
3. set your working directory to your `Week4-Beta/` folder.
4. load the `vegan` R package (be sure to install if needed).

```
rm(list = ls())
getwd()
```

```
## [1] "/cloud/project/QB2025_Choi/Week4-Beta"
```

```r
setwd("/cloud/project/QB2025_Choi/Week4-Beta/")
library(vegan)
```

```
## Loading required package: permute
```

```
## Loading required package: lattice
```

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

```r
library(ade4)
```

2) LOADING DATA

Load dataset

In the R code chunk below, load the `doubs` dataset from the `ade4` package

```r
data(doubs)
fish <- doubs$fish
fish <- fish[-8, ]
```

3) HYPOTHESIS TESTING

A. Multivariate Procedures for Categorical Designs

Earlier work done in the Doubs River suggested that the river has four distinct regions of habitat quality: the first region (sites 1-14) of "high quality"; the second (sites 15 - 19) and fourth (sites 26 - 30) of "moderate quality"; and the third (sites 20 - 25) of "low quality".

In the code chunk below, test the hypothesis that fish community composition varies with river quality.

1. create a factor vector that categorizes habitat quality in the Doubs River,
2. use the multivariate analyses for categorical predictors to describe how fish community structure relates to habitat quality.

```r
quality <- c(rep("HQ", 13), rep("MQ", 5), rep("LQ", 6), rep("MQ", 5))
adonis2(fish ~ quality, method = "bray", permutations = 999)
```

```
## Permutation test for adonis under reduced model
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = fish ~ quality, permutations = 999, method = "bray")
##          Df SumOfSqs      R2     F Pr(>F)
## Model     2   3.0947 0.45765 10.97  0.001 ***
## Residual 26   3.6674 0.54235
## Total    28   6.7621 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
library(indicspecies)
indval <- multipatt(fish, cluster = quality, func = "IndVal.g", control = how(nperm = 999))
summary(indval)
```

```
##
##  Multilevel pattern analysis
##  ---------------------------
##
##  Association function: IndVal.g
```

```
##  Significance level (alpha): 0.05
##
##  Total number of species: 27
##  Selected number of species: 23
##  Number of species associated to 1 group: 1
##  Number of species associated to 2 groups: 22
##
##  List of species associated to each combination:
##
##  Group MQ  #sps.  1
##       stat p.value
## Teso 0.686   0.021 *
##
##  Group HQ+MQ  #sps.  2
##       stat p.value
## Satr 0.860   0.005 **
## Phph 0.859   0.011 *
##
##  Group LQ+MQ  #sps.  20
##       stat p.value
## Alal 0.935   0.001 ***
## Gogo 0.933   0.001 ***
## Ruru 0.916   0.001 ***
## Legi 0.901   0.001 ***
## Baba 0.895   0.001 ***
## Chna 0.866   0.001 ***
## Spbi 0.866   0.002 **
## Cyca 0.866   0.001 ***
## Acce 0.866   0.001 ***
## Lele 0.863   0.006 **
## Titi 0.853   0.005 **
## Chto 0.829   0.002 **
## Rham 0.829   0.003 **
## Anan 0.829   0.003 **
## Eslu 0.827   0.027 *
## Pefl 0.806   0.014 *
## Blbj 0.791   0.003 **
## Scer 0.766   0.009 **
## Abbr 0.750   0.006 **
## Icme 0.661   0.025 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
fish.rel <- decostand(fish, method = "total")
phi <- multipatt(fish.rel, cluster = quality, func = "r.g",
                 control = how(nperm = 999))
summary(phi)
```

```
##
##  Multilevel pattern analysis
##  ---------------------------
##
##  Association function: r.g
##  Significance level (alpha): 0.05
##
```

```
##  Total number of species: 27
##  Selected number of species: 18
##  Number of species associated to 1 group: 9
##  Number of species associated to 2 groups: 9
##
##  List of species associated to each combination:
##
##  Group HQ  #sps.  3
##       stat p.value
## Phph 0.802   0.001 ***
## Neba 0.734   0.002 **
## Satr 0.650   0.002 **
##
##  Group LQ  #sps.  2
##       stat p.value
## Alal 0.693   0.001 ***
## Ruru 0.473   0.029 *
##
##  Group MQ  #sps.  4
##       stat p.value
## Anan 0.571   0.012 *
## Spbi 0.557   0.013 *
## Chto 0.542   0.014 *
## Icme 0.475   0.029 *
##
##  Group LQ+MQ  #sps.  9
##       stat p.value
## Legi 0.658   0.002 **
## Baba 0.645   0.003 **
## Rham 0.600   0.012 *
## Acce 0.594   0.006 **
## Cyca 0.586   0.009 **
## Chna 0.571   0.005 **
## Blbj 0.571   0.008 **
## Gogo 0.523   0.019 *
## Abbr 0.499   0.027 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Question 1: Based on the PERMANOVA, IndVal, and phi coefficient analyses, what did you learn about the relationship between habitat quality and the fish species composition? Are the different analyses consistent with one another and do they agree with the visualizations (heat maps, cluster dendograms, ordinations) that you created?

> ***Answer 1***: According to the permutation test, we can know there is a significant difference in fish species composition by the quality of water (p=0.001). Additionally, our indval result suggests that Teso is an indicator species of moderate water quality and its value is 0.686 which I don't think very high. Except for Teso species, there is no speceis which indicated sngle group. However, other 22 species are associates with the quality of two groups. For example, Satr and Phph indicate the high and moderate water quality and Alal, Gogo, Ruru, Legi, ... , Icme indicate the low and moderate quality. Finally, our phi coefficient of association shows that Phph, Neba and Satr prefer high quality water, Alal and Ruru prefer low quality water, Anan, Spbi, Chto, Icme prefer moderate wuality water, and Legi, Baba, Rham, Acce, Cyna, Chna, Blbj, Gogo and Abbr prefer low and moderate quality water.

Overall, I think the results are consistent with each other. However, interestingly, Teso is the only species which indicates high quality water but this species does not prefer high quality water. Before conducting this analysis I asusmes that an indicator of high quality water should prefer high wuality water, but it wasn't true. Also, I think this results correspond to the visualization we created. Expecially, in our PCoA graph, many species are clustered at the (0,0) which is similar to our results (many species in MQ+LQ) and a few species related to high quality water, such as Satr, Phph, Neba, are located at the left side.

B. Multivariate Procedures for Continuous Designs

i. Mantel Test

In the R code chunk below, do the following:

1. create distance matrices for both fish communities and environmental factors, and
2. use a Mantel test to test the hypothesis that fish assemblages are correlated with stream environmental variables.

```
fish.dist <- vegdist(doubs$fish[-8, ], method = "bray")
env.dist <- vegdist(scale(doubs$env[-8, ]), method = "euclid")
mantel(fish.dist, env.dist)
```

```
##
## Mantel statistic based on Pearson's product-moment correlation
##
## Call:
## mantel(xdis = fish.dist, ydis = env.dist)
##
## Mantel statistic r: 0.604
##       Significance: 0.001
##
## Upper quantiles of permutations (null model):
##   90%   95% 97.5%   99%
## 0.109 0.144 0.181 0.209
## Permutation: free
## Number of permutations: 999
```

Question 2: What do the results from our Mantel test suggest about fish diversity and stream environmental conditions? How does this relate to your hypothesis about stream quality influencing fish communities?

> ***Answer 2***: It shows us that the correlation coefficient between fish diverity and environment is 0.6 which indicates they have a moderate positive relationship. Also, its p-value is 0.001, suggesting a significant relationship between two matrices. According to these results, it seems like both the quality of water and the environmental factors are related to the fish diversity. I think the quality of water is determined by the environmental factors and they are both affect to fish diversity.

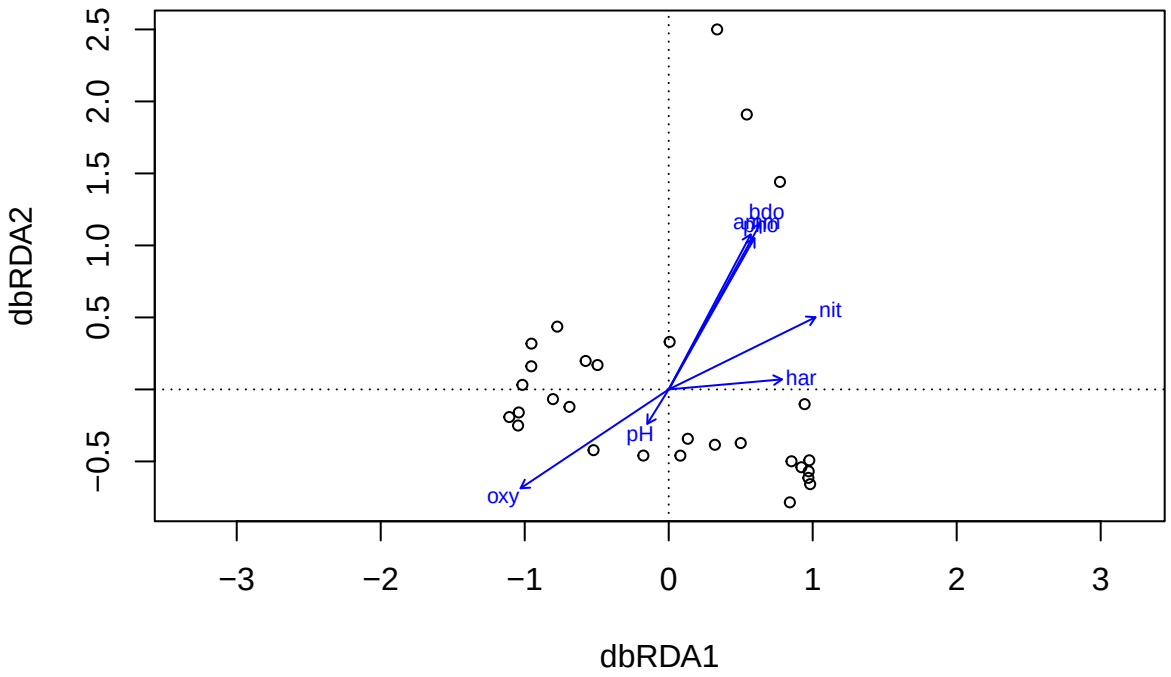
ii. Constrained Ordination

In the R code chunk below, do the following:

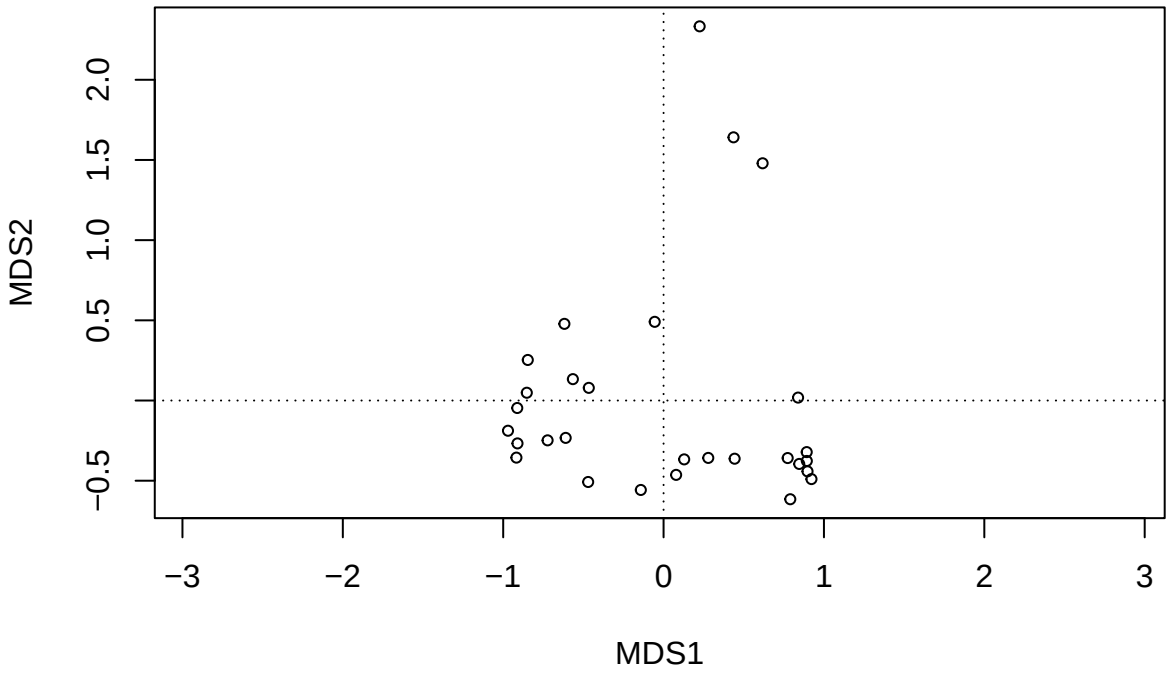
1. create an environmental matrix of the water chemistry data included in the `doubs` dataset using forward and reverse selection of variables,
2. conduct a redundancy analysis on the fish assemblages of the Doubs River,
3. use a permutation test to determine the significance of the constrained analysis,
4. use a permutation test to determine the correlation of each environmental factor on the constrained axes,
5. calculate the explained variation on the first and second constrained axes,

6. plot the constrained ordination results including labeled points for each site, and
7. add vectors that demonstrate the influence of each environmental factor the constrained ordination.

```
fish.db <- vegdist(fish, method = "bray", upper = TRUE, diag = TRUE)
env.chem <- as.matrix(doubs$env[-8, 5:11])
doubs.dbrda <- dbrda(fish.db ~., as.data.frame(env.chem))
ordiplot(doubs.dbrda)
```




```
doubs.dbrda.mod0 <- dbrda(fish.db ~ 1, as.data.frame(env.chem))
ordiplot(doubs.dbrda.mod0)
```



```r
doubs.dbrda.mod1 <- dbrda(fish.db ~., as.data.frame(env.chem))
doubs.dbrda <- ordiR2step(doubs.dbrda.mod0, doubs.dbrda.mod1, perm.max = 200)
```

```
## Step: R2.adj= 0
## Call: fish.db ~ 1
##
##                 R2.adjusted
## <All variables>  0.53032584
## + oxy            0.27727176
## + nit            0.25755208
## + bdo            0.17477787
## + pho            0.14568614
## + har            0.14174915
## + amm            0.14142804
## <none>           0.00000000
## + pH            -0.01827054
##
##       Df    AIC      F Pr(>F)
## + oxy  1 47.939 11.742  0.002 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.2772718
## Call: fish.db ~ oxy
##
##                 R2.adjusted
## <All variables>   0.5303258
## + bdo             0.4009000
## + amm             0.3474192
## + pho             0.3452702
## + har             0.3331357
## + nit             0.3316120
## <none>            0.2772718
## + pH              0.2586983
##
##       Df    AIC      F Pr(>F)
## + bdo  1 43.404 6.5716  0.002 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.4009
## Call: fish.db ~ oxy + bdo
##
##                 R2.adjusted
## <All variables>   0.5303258
## + nit             0.4980793
## + har             0.4695121
## <none>            0.4009000
## + pho             0.3938042
## + amm             0.3869134
## + pH              0.3865240
##
##       Df    AIC     F Pr(>F)
## + nit  1 39.134 6.034  0.002 **
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.4980793
## Call: fish.db ~ oxy + bdo + nit
##
##                 R2.adjusted
## + amm             0.5415705
## <All variables>   0.5303258
## + pho             0.5277128
## + har             0.5218852
## <none>            0.4980793
## + pH              0.4843267
```

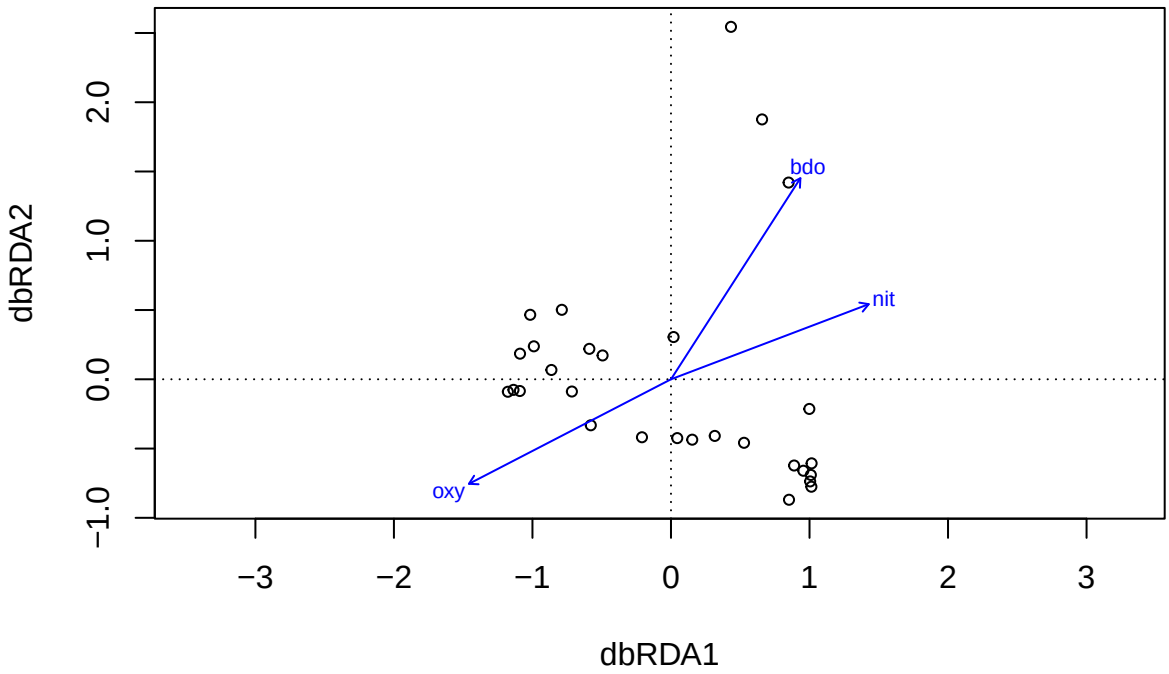
```
doubs.dbrda$call
```

```
## dbrda(formula = fish.db ~ oxy + bdo + nit, data = as.data.frame(env.chem))
```

```
doubs.dbrda$anova
```

```
##                  R2.adj Df    AIC       F Pr(>F)
## + oxy           0.27727  1 47.939 11.7421  0.002 **
## + bdo           0.40090  1 43.404  6.5716  0.002 **
## + nit           0.49808  1 39.134  6.0340  0.002 **
## <All variables> 0.53033
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
ordiplot(doubs.dbrda)
```




```
permutest(doubs.dbrda, permutations = 999)
```

```
##
## Permutation test for dbrda under reduced model
##
```

```
## Permutation: free
## Number of permutations: 999
##
## Model: dbrda(formula = fish.db ~ oxy + bdo + nit, data =
## as.data.frame(env.chem))
## Permutation test for all constrained eigenvalues
##          Df Inertia      F Pr(>F)
## Model     3  3.7317 10.262  0.001 ***
## Residual 25  3.0304
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
envfit(doubs.dbrda, env.chem[,c(4,6,7)], perm = 999)
```

```
##
## ***VECTORS
##
##       dbRDA1   dbRDA2     r2 Pr(>r)
## nit  0.87724  0.48005 0.6431  0.001 ***
## oxy -0.82864 -0.55979 0.7656  0.001 ***
## bdo  0.55603  0.83116 0.8939  0.001 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Permutation: free
## Number of permutations: 999
```

```
dbrda.explainvar1 <- round(doubs.dbrda$CCA$eig[1] /
                           sum(c(doubs.dbrda$CCA$eig, doubs.dbrda$CA$eig)), 3)*100
dbrda.explainvar2 <- round(doubs.dbrda$CCA$eig[2] /
                           sum(c(doubs.dbrda$CCA$eig, doubs.dbrda$CA$eig)), 3)*100

par(mar = c(5, 5, 4, 4) + 0.1)
plot(scores(doubs.dbrda, display = "wa"), xlim = c(-1.3, 1.1),
     ylim = c(-1.1, 2.7), xlab = paste("dbRDA 1 (", dbrda.explainvar1, "%)",
     sep = ""), ylab = paste("dbRDA 2 (", dbrda.explainvar2, "%)", sep = ""),
     pch = 16, cex = 2.0, tyep = "n", cex.lab = 1.5,
     cex.axis = 1.2, axe = FALSE)
```

```
## Warning in plot.window(...): "tyep" is not a graphical parameter
```

```
## Warning in plot.xy(xy, type, ...): "tyep" is not a graphical parameter
```

```
## Warning in title(...): "tyep" is not a graphical parameter
```

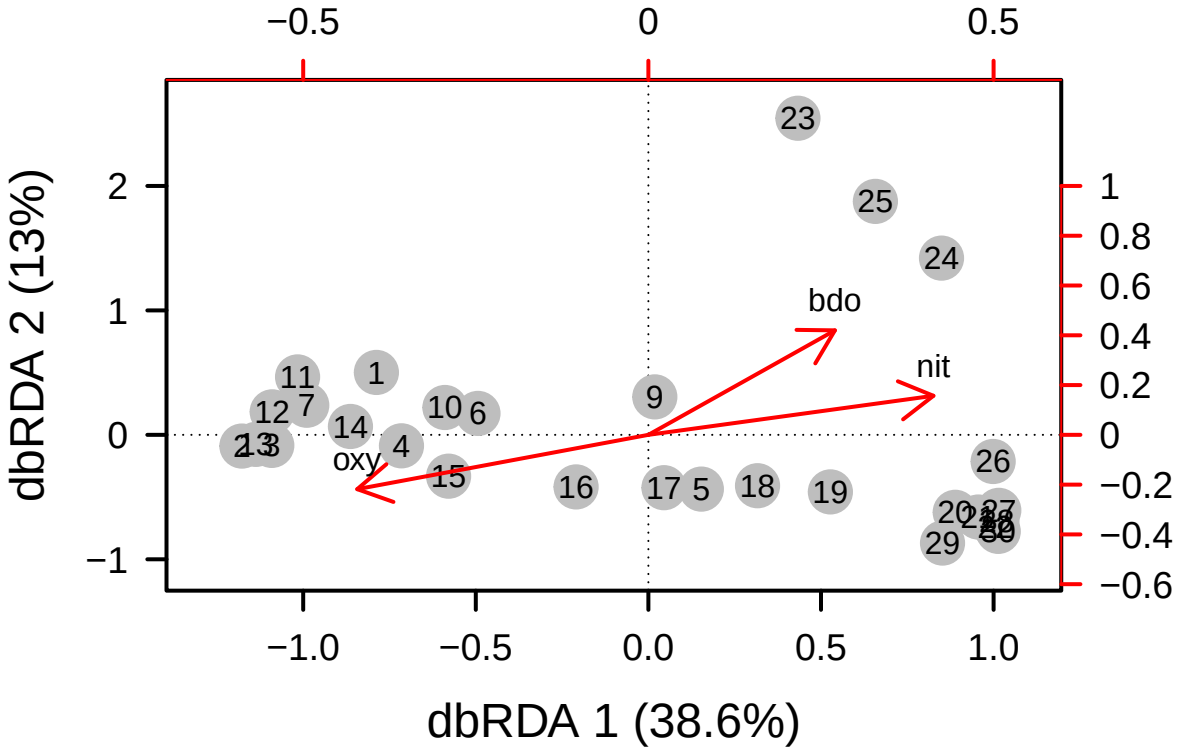
```
axis(side = 1, labels = T, lwd.ticks = 2, cex.axis = 1.2, las = 1)
axis(side = 2, labels = T, lwd.ticks = 2, cex.axis = 1.2, las = 1)
abline(h = 0, v = 0, lty = 3)
box(lwd = 2)

points(scores(doubs.dbrda, display = "wa"),
       pch = 19, cex = 3, bg = "gray", col = "gray")
text(scores(doubs.dbrda, display = "wa"),
     labels = row.names(scores(doubs.dbrda, display = "wa")))


vectors <- scores(doubs.dbrda, display = "bp")
arrows(0, 0, vectors[, 1], vectors[, 2],
```

```
       lwd = 2, lty = 1, length = 0.2, col = "red")
text(vectors[,1], vectors[,2], pos = 3,
     label = row.names(vectors))
axis(side = 3, lwd.ticks = 2, cex.axis = 1.2, las = 1, col = "red", ld = 2,2,
     at = pretty(range(vectors[, 1])) * 2, labels = pretty(range(vectors[,1])))
```

```
## Warning in axis(side = 3, lwd.ticks = 2, cex.axis = 1.2, las = 1, col = "red",
## : "ld" is not a graphical parameter
```

```
axis(side = 4, lwd.ticks = 2, cex.axis = 1.2, las = 1, col = "red", ld = 2,2,
     at = pretty(range(vectors[, 2])) * 2, labels = pretty(range(vectors[,2])))
```

```
## Warning in axis(side = 4, lwd.ticks = 2, cex.axis = 1.2, las = 1, col = "red",
## : "ld" is not a graphical parameter
```



Question 3: Based on the constrained ordination, what are the environmental variables (or groups of correlated variables) that seem to be contributing to variation in fish community structure?

> ***Answer 3***: Among many different environmental factors, oxygen, nitrogen and biochemical oxygen demand are the best explanatory variables (which hass the lowest AIC values). Also, it corresponds to the graph where oxygen primarily explains the variation in sites on the left half of the plot, biochemical oxygen demand is associated with sites in the upper right region, and nitrogen is linked to sites on the right side.

iii. Variation Partitioning

In the code chunk below,

1. Create a matrix model of the selected environmental variables,
2. Create a matrix model of the selected PCNM axes,
3. Perform constrained and partial constrained ordinations using the spatial and environmental models you just created,

4. Test the significance of each of your constrained ordinations using permutation tests,
5. Partition the variation among sites into the relative importance of space, environment, spatially structured environment, and residuals,
6. Plot the variation partitioning output to visualize it.

```
doubs.dbrda$anova
```

```
##                  R2.adj Df    AIC       F Pr(>F)
## + oxy           0.27727  1 47.939 11.7421  0.002 **
## + bdo           0.40090  1 43.404  6.5716  0.002 **
## + nit           0.49808  1 39.134  6.0340  0.002 **
## <All variables> 0.53033
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
env.mod <- model.matrix( ~ oxy + bdo + nit, as.data.frame(env.chem))[, -1]
rs <- rowSums(fish)/sum(fish)
doubs.pcnmw <- pcnm(dist(doubs$xy[-8, ]), w = rs, dist.ret = T)
doubs.pcnmw$values > 0
```

```
##  [1]  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE
## [13]  TRUE  TRUE  TRUE  TRUE  TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [25] FALSE FALSE
```

```
doubs.space <- as.data.frame(scores(doubs.pcnmw))
doubs.pcnm.mod0 <- dbrda(fish.db ~ 1, doubs.space)
doubs.pcnm.mod1 <- dbrda(fish.db ~ ., doubs.space)
step.pcnm <- ordiR2step(doubs.pcnm.mod0, doubs.pcnm.mod1, per.max = 200)
```

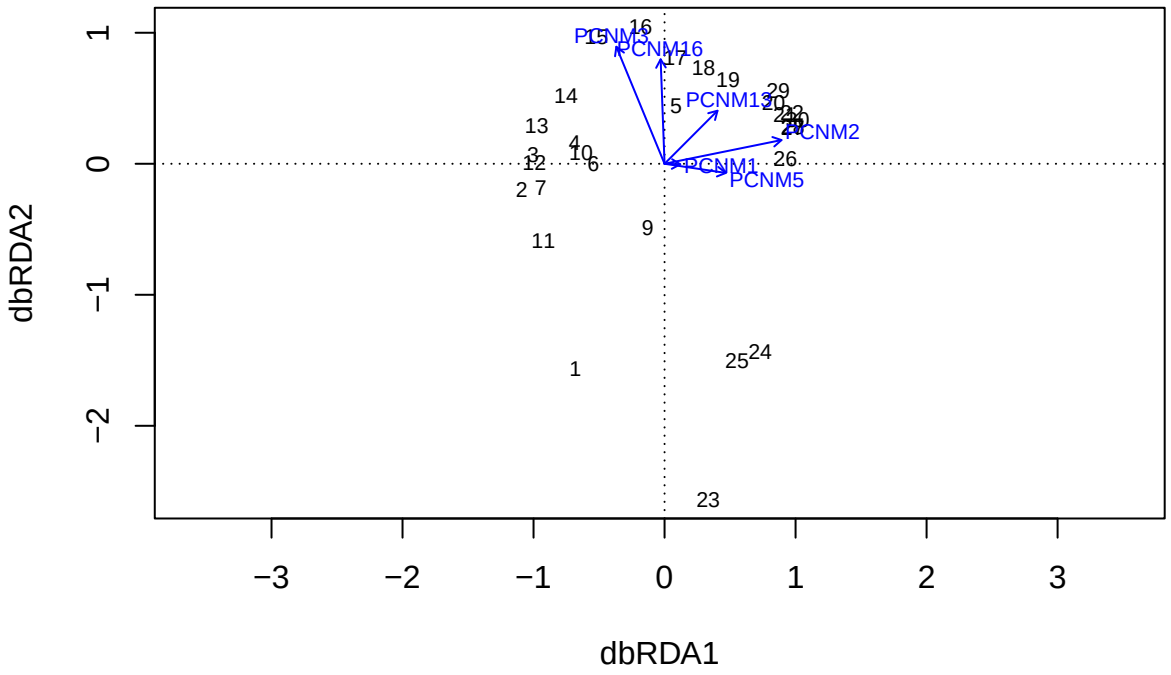
```
## Step: R2.adj= 0
## Call: fish.db ~ 1
##
##                   R2.adjusted
## <All variables>  0.626011301
## + PCNM2          0.235370423
## + PCNM3          0.078394885
## + PCNM13         0.065305668
## + PCNM5          0.046185074
## + PCNM6          0.032809156
## + PCNM16         0.030486700
## + PCNM14         0.029680999
## + PCNM9          0.020357410
## + PCNM15         0.013632610
## + PCNM8          0.009411968
## + PCNM1          0.003986221
## + PCNM17         0.002415012
## + PCNM10         0.001326442
## <none>           0.000000000
## + PCNM7         -0.001861430
## + PCNM11        -0.006841522
## + PCNM4         -0.007089863
## + PCNM12        -0.014396973
##
##         Df    AIC     F Pr(>F)
## + PCNM2  1 49.574 9.619  0.002 **
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.2353704
## Call: fish.db ~ PCNM2
##
##                 R2.adjusted
## <All variables>   0.6260113
## + PCNM3           0.3429270
## + PCNM5           0.3057368
## + PCNM1           0.2885396
## + PCNM16          0.2786746
## + PCNM14          0.2744520
## + PCNM15          0.2692809
## + PCNM6           0.2659866
## + PCNM13          0.2636194
## + PCNM9           0.2517847
## + PCNM8           0.2496240
## + PCNM10          0.2434688
## + PCNM7           0.2431476
## + PCNM17          0.2404343
## + PCNM11          0.2366833
## <none>            0.2353704
## + PCNM12          0.2288789
## + PCNM4           0.2189522
##
##         Df    AIC      F Pr(>F)
## + PCNM3  1 46.083 5.4196  0.002 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.342927
## Call: fish.db ~ PCNM2 + PCNM3
##
##                 R2.adjusted
## <All variables>   0.6260113
## + PCNM5           0.4076020
## + PCNM1           0.3970300
## + PCNM16          0.3853210
## + PCNM15          0.3828748
## + PCNM14          0.3781827
## + PCNM13          0.3770376
## + PCNM6           0.3595644
## + PCNM8           0.3556885
## + PCNM7           0.3541631
## + PCNM10          0.3526775
## + PCNM17          0.3513683
## + PCNM9           0.3433672
## <none>            0.3429270
## + PCNM11          0.3416399
## + PCNM12          0.3396547
## + PCNM4           0.3311509
##
##         Df    AIC      F Pr(>F)
## + PCNM5  1 43.941 3.8385   0.01 **
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.407602
## Call: fish.db ~ PCNM2 + PCNM3 + PCNM5
##
##                 R2.adjusted
## <All variables>   0.6260113
## + PCNM1           0.4721469
## + PCNM16          0.4631976
## + PCNM15          0.4589111
## + PCNM14          0.4535248
## + PCNM13          0.4511582
## + PCNM6           0.4305640
## + PCNM7           0.4261965
## + PCNM8           0.4224505
## + PCNM17          0.4181666
## + PCNM10          0.4154485
## + PCNM11          0.4112178
## + PCNM9           0.4111995
## + PCNM12          0.4087602
## <none>            0.4076020
## + PCNM4           0.3976526
##
##         Df    AIC     F Pr(>F)
## + PCNM1  1 41.411 4.057   0.01 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.4721469
## Call: fish.db ~ PCNM2 + PCNM3 + PCNM5 + PCNM1
##
##                 R2.adjusted
## <All variables>   0.6260113
## + PCNM13          0.5212427
## + PCNM16          0.5208668
## + PCNM15          0.5161770
## + PCNM14          0.5147355
## + PCNM6           0.4999020
## + PCNM7           0.4936559
## + PCNM8           0.4904113
## + PCNM17          0.4856884
## + PCNM10          0.4835952
## + PCNM11          0.4760087
## + PCNM9           0.4751424
## + PCNM12          0.4747221
## <none>            0.4721469
## + PCNM4           0.4651218
##
##          Df    AIC      F Pr(>F)
## + PCNM13  1 39.346 3.4612   0.01 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Step: R2.adj= 0.5212427
## Call: fish.db ~ PCNM2 + PCNM3 + PCNM5 + PCNM1 + PCNM13
##
##                 R2.adjusted
## <All variables>   0.6260113
## + PCNM16          0.5767968
## + PCNM15          0.5715331
## + PCNM14          0.5698343
## + PCNM6           0.5475140
## + PCNM7           0.5392074
## + PCNM8           0.5379134
## + PCNM11          0.5281106
## + PCNM9           0.5267003
## + PCNM10          0.5265029
## + PCNM12          0.5255581
## <none>            0.5212427
## + PCNM17          0.5171800
## + PCNM4           0.5152311
##
##          Df   AIC      F Pr(>F)
## + PCNM16  1 36.48 4.0192  0.006 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.5767968
## Call: fish.db ~ PCNM2 + PCNM3 + PCNM5 + PCNM1 + PCNM13 + PCNM16
##
##                 R2.adjusted
## <All variables>   0.6260113
## + PCNM6           0.6043089
## + PCNM8           0.5970286
## + PCNM12          0.5946888
## + PCNM7           0.5946475
## + PCNM9           0.5883735
## + PCNM10          0.5851333
## + PCNM15          0.5846468
## <none>            0.5767968
## + PCNM17          0.5748533
## + PCNM4           0.5733749
## + PCNM11          0.5711176
## + PCNM14          0.5652509
##
##         Df    AIC      F Pr(>F)
## + PCNM6  1 35.182 2.5296  0.066 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
plot(step.pcnm)
```




```
step.pcnm$anova
```

```
##                  R2.adj Df    AIC      F Pr(>F)
## + PCNM2         0.23537  1 49.574 9.6190  0.002 **
## + PCNM3         0.34293  1 46.083 5.4196  0.002 **
## + PCNM5         0.40760  1 43.941 3.8385  0.010 **
## + PCNM1         0.47215  1 41.411 4.0570  0.010 **
## + PCNM13        0.52124  1 39.346 3.4612  0.010 **
## + PCNM16        0.57680  1 36.480 4.0192  0.006 **
## <All variables> 0.62601
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
space.mod <- model.matrix( ~ PCNM2 + PCNM3 + PCNM5 + PCNM1 +
                             PCNM13 + PCNM16 + PCNM6 , doubs.space)[, -1]
doubs.total.env <- dbrda(fish.db ~ env.mod)
doubs.total.space <- dbrda(fish.db ~ space.mod)

doubs.env.cond.space <- dbrda(fish.db ~ env.mod + Condition(space.mod))
doubs.space.cond.env <- dbrda(fish.db ~ space.mod + Condition(env.mod))

permutest(doubs.env.cond.space, permutations = 999)
```

```
##
## Permutation test for dbrda under reduced model
##
## Permutation: free
## Number of permutations: 999
##
## Model: dbrda(formula = fish.db ~ env.mod + Condition(space.mod))
## Permutation test for all constrained eigenvalues
##          Df Inertia     F Pr(>F)
## Model     3 0.85158 4.423  0.001 ***
```

```
## Residual 18 1.15519
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
permutest(doubs.space.cond.env, permutations = 999)
```

```
##
## Permutation test for dbrda under reduced model
##
## Permutation: free
## Number of permutations: 999
##
## Model: dbrda(formula = fish.db ~ space.mod + Condition(env.mod))
## Permutation test for all constrained eigenvalues
##          Df Inertia      F Pr(>F)
## Model     7  1.8752 4.1741  0.001 ***
## Residual 18  1.1552
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
permutest(doubs.total.env, permutations = 999)
```

```
##
## Permutation test for dbrda under reduced model
##
## Permutation: free
## Number of permutations: 999
##
## Model: dbrda(formula = fish.db ~ env.mod)
## Permutation test for all constrained eigenvalues
##          Df Inertia      F Pr(>F)
## Model     3  3.7317 10.262  0.001 ***
## Residual 25  3.0304
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
permutest(doubs.total.space, permutations = 999)
```

```
##
## Permutation test for dbrda under reduced model
##
## Permutation: free
## Number of permutations: 999
##
## Model: dbrda(formula = fish.db ~ space.mod)
## Permutation test for all constrained eigenvalues
##          Df Inertia      F Pr(>F)
## Model     7  4.7553 7.1089  0.001 ***
## Residual 21  2.0068
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
doubs.varpart <- varpart(fish.db, env.mod, space.mod)
doubs.varpart
```

```
##
## Partition of squared Bray distance in dbRDA
```

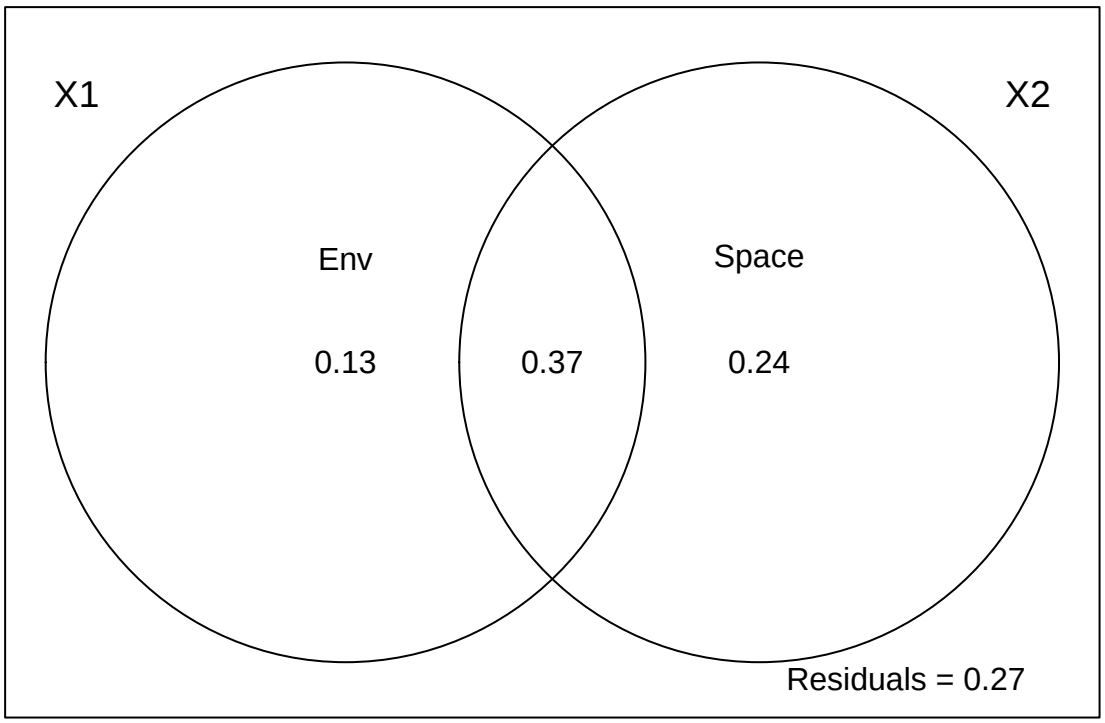
```
##
## Call: varpart(Y = fish.db, X = env.mod, space.mod)
##
## Explanatory tables:
## X1:  env.mod
## X2:  space.mod
##
## No. of explanatory tables: 2
## Total variation (SS): 6.7621
## No. of observations: 29
##
## Partition table:
##                       Df R.squared Adj.R.squared Testable
## [a+c] = X1             3   0.55186       0.49808     TRUE
## [b+c] = X2             7   0.70323       0.60431     TRUE
## [a+b+c] = X1+X2       10   0.82917       0.73426     TRUE
## Individual fractions
## [a] = X1|X2            3                 0.12995     TRUE
## [b] = X2|X1            7                 0.23618     TRUE
## [c]                    0                 0.36813    FALSE
## [d] = Residuals                          0.26574    FALSE
## ---
## Use function 'dbrda' to test significance of fractions of interest
```

```
par(mar = c(2,2,2,2))
plot(doubs.varpart)
text(1, 0.25, "Space")
text(0, 0.25, "Env")
text("Variation Partitioning of\nDoubs Fish Diversity", side = 3, line = -3)
```

```
## Warning in xy.coords(x, y, recycle = TRUE, setLab = FALSE): NAs introduced by
## coercion
```

```
## Warning in text.default("Variation Partitioning of\nDoubs Fish Diversity", :
## "side" is not a graphical parameter
```

```
## Warning in text.default("Variation Partitioning of\nDoubs Fish Diversity", :
## "line" is not a graphical parameter
```



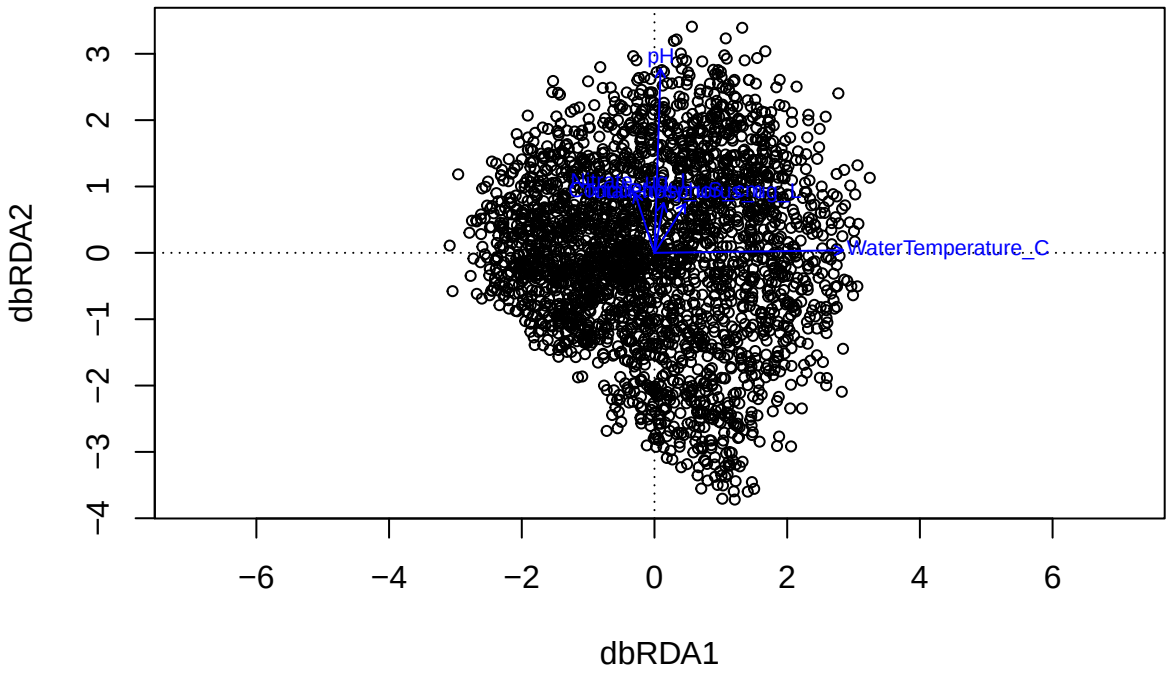
Question 4: Interpret the variation partitioning results.

> ***Answer 4***: Our results suggest that environment alone explain about 13 percent of variation, and space alone explains about 24 %, both environment and space explain about 37% and neither of them explain about 27 % of variation.

SYNTHESIS

Load the dataset from that you and your partner are using for the team project. Use one of the hypothesis-testing tools introduced in the beta diversity module. Interpret the findings of your data with respect to principles of biodiversity.

```
fish.team <- read.csv("/cloud/project/QB2025_Choi/Fish_Dataset.csv")
fish.team.var <- cbind(fish.team[, 2:3], fish.team[, 5], fish.team[, 7:9], fish.team[, 23:658])
colnames(fish.team.var)[3] = "water_temp"
fish.team.env <- fish.team[, 5:9]
only.species <- fish.team.var[,7:642]
fish.team.db <- vegdist(only.species, method = "bray", upper = TRUE, diag = TRUE)
fish.team.dbrda <- dbrda(fish.team.db ~ ., as.data.frame(fish.team.env))
ordiplot(fish.team.dbrda)
```



>

Here, we used dbRDA plot to show which environmental factor has a strong effect on fish diversity. Before this week, we divide our latitude data into 5 groups, but here if we divided them into 5 groups, we thought 5 groups are too little to do ordination analysis. Thus, here we used original data which have specific latitude and longgitude information. For the graph we obtained, I think the pH and water temperature are the most significant environmental factors. > Next, we tried to confirm which environmental factor has the lowest AIC value. However, when we tried to run "ordiR2step()" to get the AIC value, the r studio kept crashing. Thus, the code below is what we tried to run but it did not work. I think 5 groups are too small and using the whole latitude data is too large. I think we should discuss this further and have to find a good criterion for dividing our spatial dataset.

```
fish.team.dbrda.mod0 <- dbrda(fish.team.db ~ 1, as.data.frame(fish.team.env))
fish.team.dbrda.mod1 <- dbrda(fish.team.db ~ ., as.data.frame(fish.team.env))
fish.team.dbrda <- ordiR2step(fish.team.dbrda.mod0, fish.team.dbrda.mod1, perm.max = 0) #the data is to
fish.team.dbrda$call
fish.team.dbrda$anova
ordiplot(fish.team.dbrda)
permutest(fish.team.dbrda, permutations = 999)
envfit(fish.team.dbrda, env.chem[,c(1,3,4)], perm = 999)
fish.dbrda.explainvar1 <- round(fish.team.dbrda$CCA$eig[1] /
                            sum(c(fish.team.dbrda$CCA$eig, fish.team.dbrda$CA$eig)), 3)*100
fish.dbrda.explainvar2 <- round(fish.team.dbrda$CCA$eig[2] /
                            sum(c(fish.team.dbrda$CCA$eig, fish.team.dbrda$CA$eig)), 3)*100
par(mar = c(5, 5, 4, 4) + 0.1)
plot(scores(fish.team.dbrda, display = "wa"), xlim = c(-1.3, 1.1),
     ylim = c(-1.1, 2.7), xlab = paste("dbRDA 1 (", fish.dbrda.explainvar1, "%)",
     sep = ""), ylab = paste("dbRDA 2 (", fish.dbrda.explainvar2, "%)", sep = ""),
     pch = 16, cex = 2.0, tyep = "n", cex.lab = 1.5,
     cex.axis = 1.2, axe = FALSE)

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)
```

```
points(scores(fish.team.dbrda, display = "wa"),
       pch = 19, cex = 3, bg = "gray", col = "gray")
text(scores(fish.team.dbrda, display = "wa"),
     labels = row.names(scores(fish.team.dbrda, display = "wa")))


vectors <- scores(fish.team.dbrda, display = "bp")
arrows(0, 0, vectors[, 1], vectors[, 2],
       lwd = 2, lty = 1, length = 0.2, col = "red")
text(vectors[,1], vectors[,2], pos = 3,
     label = row.names(vectors))
axis(side = 3, lwd.ticks = 2, cex.axis = 1.2, las = 1, col = "red", ld = 2,2,
     at = pretty(range(vectors[, 1])) * 2, labels = pretty(range(vectors[,1])))

axis(side = 4, lwd.ticks = 2, cex.axis = 1.2, las = 1, col = "red", ld = 2,2,
     at = pretty(range(vectors[, 2])) * 2, labels = pretty(range(vectors[,2])))
```