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

Lauren Albert; Z620: Quantitative Biodiversity, Indiana University

08 February, 2023

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, it is *imperative* that you **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 '8.BetaDiversity' folder.
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 (**8.BetaDiversity_2_Worksheet.Rmd**) with all code blocks filled out and questions answered) and the PDF output of `Knitr` (**8.BetaDiversity_2_Worksheet.pdf**).

The completed exercise is due on **Wednesday, February 8th, 2023 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 "*/6.BetaDiversity*" folder, and

4. load the **vegan** R package (be sure to install if needed).

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

```
## [1] "/Users/laurenalbert/GitHub/QB2023_Albert/2.Worksheets/6.BetaDiversity"
```

```
library(vegan)
```

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

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

```
## This is vegan 2.6-4
```

```
library(ade4)
options(repos = list(CRAN="http://cran.rstudio.com/"))
```

2) LOADING DATA

Load dataset

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

```
# note, pleae do not print the dataset when submitting
data(doubs)
```

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.

```
#create a factors vector
fish <- doubs$fish
fish <- fish[-8,]
quality <- c(rep("HQ", 13), rep("MQ", 5), rep("LQ", 6), rep("MQ", 5))

#run PERMANOVA with adonis function
#adonis(fish ~ quality, method = "bray", permutations = 999)

install.packages("indicspecies")
```

```
##
## The downloaded binary packages are in
##  /var/folders/b_/hd3hcpsx0ln0dsvky9m34_c40000gn/T//RtmpABKy3S/downloaded_packages
```

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

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

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***: Working through these three analyses helps illustrate the story of habitat quality and fish species in the Doubs River. The PERMANOVA suggests that quality is a significant predictor, and then the IndVal points to groups of species with scores close to 1– suggesting species are a strong indicator of that respective quality site. Finally, the phi coefficient shows habitat preferences based on quality. These analyses support the initial visualizations in that there were species clusters forming (ordination) and similarities appearing (heat maps/dendograms). ### 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.

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

#mantel test
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%
## 0.0956 0.1321 0.1580 0.1905
## 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?

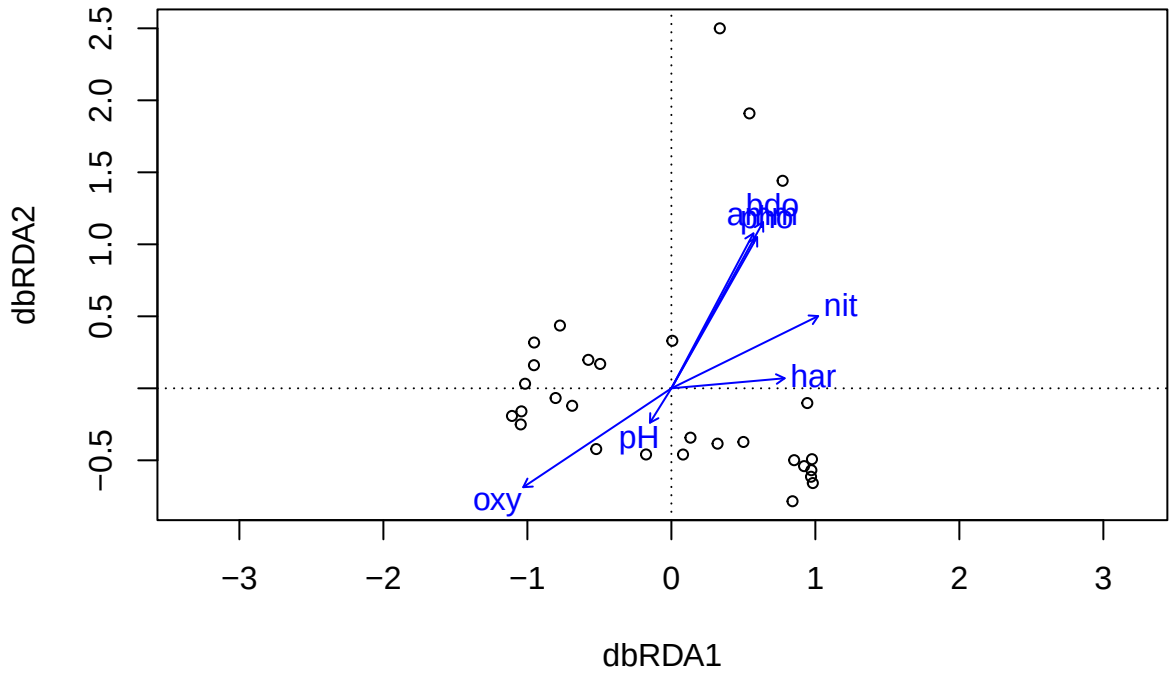
> ***Answer 2***: The results from the Mantel test suggest there is a positive correlation between fish diversity and stream conditions. If there is a correlation between these two variables, that suggests stream quality should have some influence fish community composition.

ii. Constrained Ordination

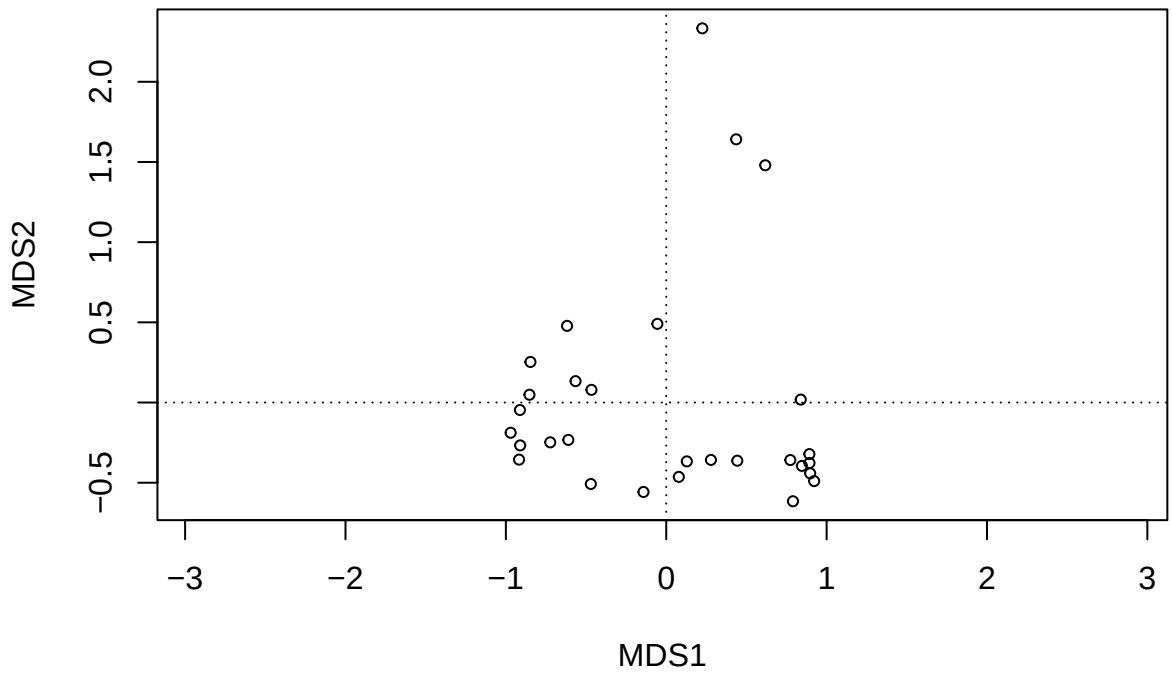
In the R code chunk below, do the following:

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.

```
#define environmental matrix
env.chem <- as.matrix(doubs$env[-8, 5:11])

fish.db <- vegdist(fish, method = "bray")

#perform dbRDA
doubs.dbrda <- dbrda(fish.db ~ ., as.data.frame(env.chem))
ordiplot(doubs.dbrda)
```




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




```
#model full model, with all explanatory variables
doubs.dbrda.mod1 <- dbrda(fish.db ~ ., as.data.frame(env.chem))

#all combinatns of explanatory variables in model-- function returns model with lowest AIC value
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  0.006 **
## ---
## 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
```

```
#which model was selected
doubs.dbrda$call
```

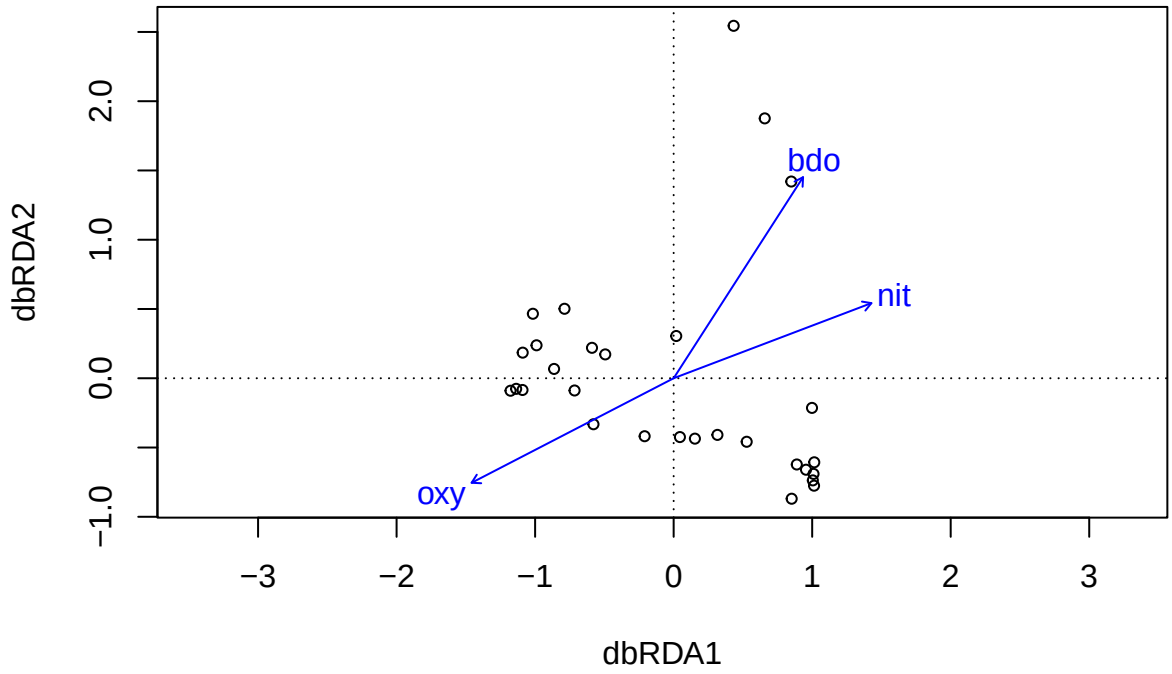
```
## dbrda(formula = fish.db ~ oxy + bdo + nit, data = as.data.frame(env.chem))
```

```
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  0.006 **
## <All variables> 0.53033
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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



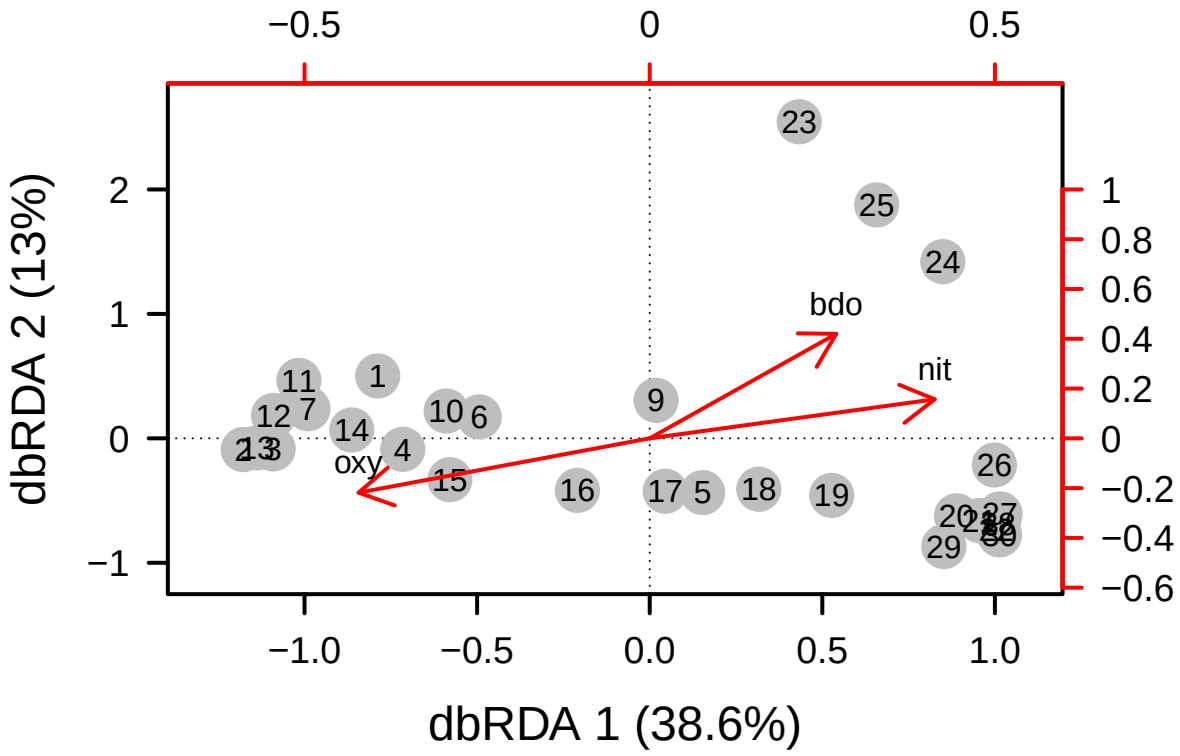

```
#permutation tests to evaluate significance
permutest(doubs.dbrda, permutations = 999)
```

```
##
## 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
```

```r
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
```

```r
#calculate explained variation
dbrda.explainvar1 <- round(doubs.dbrda$CCA$eig[1] / sum(c(doubs.dbrda$CCA$eig, doubs.dbrda$CA$eig)), 3)
dbrda.explainvar2 <- round(doubs.dbrda$CCA$eig[2] / sum(c(doubs.dbrda$CCA$eig, doubs.dbrda$CA$eig)), 3)

#define plot parameters
par(mar = c(5, 5, 4, 4) + 0.1)

#initiate plot
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, type = "n", cex.lab = 1.5, cex.axis = 1.2, 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 and labels
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")))

#add environmental vectors
vectors <- scores(doubs.dbrda, display = "bp")
#row.names(vectors) <- rownames(vectors)

arrows(0, 0, vectors[,1], vectors[,2],
       lwd = 2, lty = 1, length = 0.2, col = "red")
text(vectors[,1], vectors[,2], pos = 3,
     labels = row.names(vectors))
axis(side = 3, lwd.ticks = 2, cex.axis = 1.2, las = 1, col = "red", lwd = 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", lwd = 2.2,
     at = pretty(range(vectors[, 2])) * 2, labels = pretty(range(vectors[, 2])))
```



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***: The environmental variables that contribute to the variation in fish community are dissolved oxygen, biological oxygen demand, and nitrate concentration.

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.

```
#environmental models uses oxy, bdo, and nit and has r2 of 0.53
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  0.006 **
## <All variables> 0.53033
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#create matrix model for environmental data
env.mod <- model.matrix(~ oxy + bdo + nit, as.data.frame(env.chem))[, -1 ]

#weight each site by it's relative abundance
rs <- rowSums(fish)/ sum(fish)

#perform PCNM
doubs.pcnmw <- pcnm(dist(doubs$xy[-8,]), w = rs, dist.ret = T)

#PCNM can return negative eigenvalues, but onlty the eigenvectors associated with the positive eigenval
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
```

```
#perform model selection to determine which eigenvalues create the most informative model with the fewe
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, perm.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 **
```

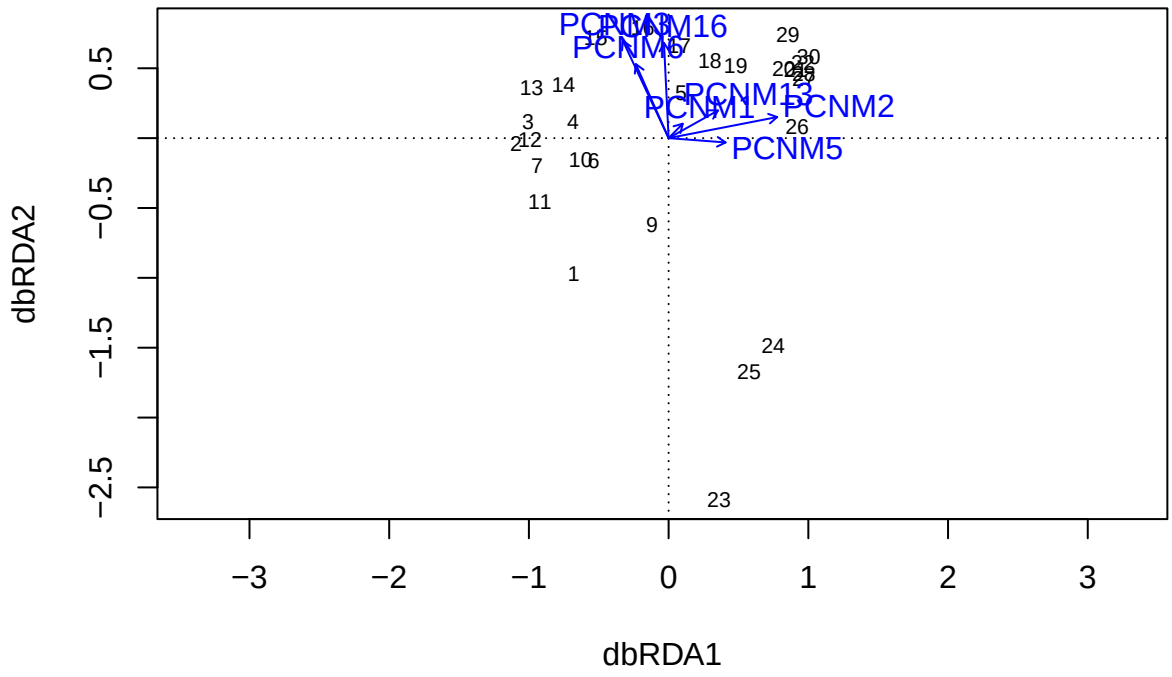
```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.2353704
## Call: fish.db ~ PCNM2
##
##                 R2.adjusted
## <All variables>   0.6260113
## + PCNM3           0.3429270
## + PCNM5           0.3057368
## + PCNM1           0.2885396
## + PCNM16          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  0.004 **
## ---
## 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  0.012 *
## ---
## 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  0.004 **
## ---
## 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  0.014 *
## ---
## 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   0.01 **
## ---
## 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   0.03 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.6043089
## Call: fish.db ~ PCNM2 + PCNM3 + PCNM5 + PCNM1 + PCNM13 + PCNM16 + PCNM6
##
##                 R2.adjusted
## <All variables>   0.6260113
## + PCNM8           0.6248697
```

```
## + PCNM12          0.6208788
## + PCNM10          0.6170988
## + PCNM7           0.6142419
## + PCNM15          0.6140369
## + PCNM9           0.6107110
## <none>            0.6043089
## + PCNM17          0.6037430
## + PCNM11          0.5978305
## + PCNM4           0.5963667
## + PCNM14          0.5932113
##
##         Df    AIC     F Pr(>F)
## + PCNM8  1 34.219 2.151  0.106
```

```
#visualize the biplot showing how each vector explains variation across sites
plot(step.pcnm)
```




```
#object 'step.pcnm' now contains the selected model
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  0.004 **
## + PCNM5         0.40760  1 43.941 3.8385  0.012 *
## + PCNM1         0.47215  1 41.411 4.0570  0.004 **
```

```
## + PCNM13        0.52124  1 39.346 3.4612  0.014 *
## + PCNM16        0.57680  1 36.480 4.0192  0.010 **
## + PCNM6         0.60431  1 35.182 2.5296  0.030 *
## <All variables> 0.62601
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#construct a spatial model using only the selected PCNM axes
space.mod <- model.matrix(~ PCNM2 + PCNM3 + PCNM5 + PCNM1 + PCNM13 + PCNM16 + PCNM6, doubs.space)[, -1]


#conduct constrained ordinations
doubs.total.env <- dbrda(fish.db ~ env.mod)
doubs.total.space <- dbrda(fish.db ~ space.mod)

#construct partial constrained ordinations
doubs.env.cond.space <- dbrda(fish.db ~ env.mod + Condition(space.mod))
doubs.space.cond.env <- dbrda(fish.db ~ space.mod + Condition(env.mod))

#test for significance of the dbRDA fractions
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)
```

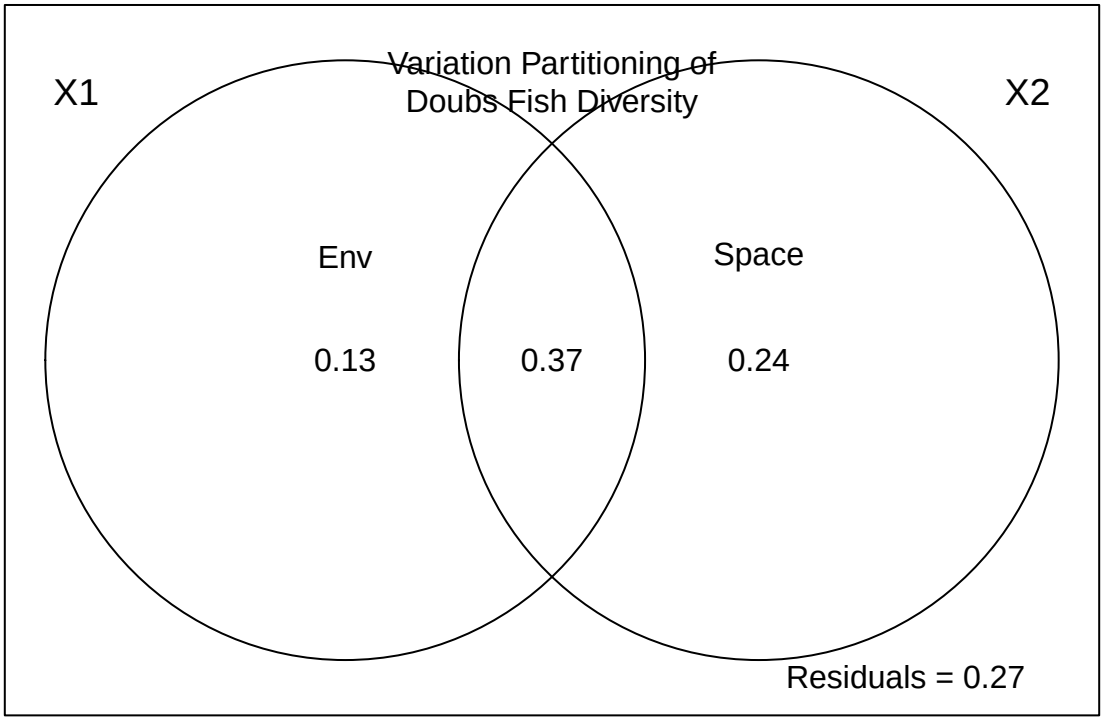
```
##
## Permutation test for dbrda under reduced model
##
## Permutation: free
## Number of permutations: 999
##
## Model: dbrda(formula = fish.db ~ space.mod)
## Permutation test for all constrained eigenvalues
##          Df Inertia      F Pr(>F)
## Model     7  4.7553 7.1089  0.001 ***
## Residual 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")
mtext("Variation Partitioning of\nDoubs Fish Diversity", side = 3, line = -3)
```



Question 4: Interpret the variation partitioning results.

> ***Answer 4*** : This diagram provides that 13% of the variation in the fish community is explained by the environment alone, 24% by space alone, and 37% by space and environment, simultaneously.

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.

```
library(readr)
library(ggplot2)
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.5.0
## v purrr   1.0.1      v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(plyr)
```

```
## ------------------------------------------------------------------------------
## You have loaded plyr after dplyr - this is likely to cause problems.
## If you need functions from both plyr and dplyr, please load plyr first, then dplyr:
## library(plyr); library(dplyr)
## ------------------------------------------------------------------------------
##
## Attaching package: 'plyr'
##
## The following objects are masked from 'package:dplyr':
##
##     arrange, count, desc, failwith, id, mutate, rename, summarise,
##     summarize
##
## The following object is masked from 'package:purrr':
##
##     compact
```

```
#import data
zoopfield <- read_csv("zoopfield_2009.csv")
```

```
## New names:
## Rows: 282 Columns: 88
## -- Column specification
## -------------------------------------------------------- Delimiter: "," chr
## (2): Lake_Name, date dbl (55): Round, JD, M_m_l, M_s_l, M_n_l, M_m_e, M_s_e,
## M_n_e, U_m_l, U_s_l,... num (11): dent, cerio, pulic, Bosmina, parv, ambig,
## diaph, scaph, alona, pre... lgl (20): M_m_l2, M_s_l2, U_m_l2, U_s_l2, O_m_l2,
## O_s_l2, E_m_l2, E_s_l2, os...
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## * `P_ug` -> `P_ug...66`
## * `P_ug` -> `P_ug...67`
```

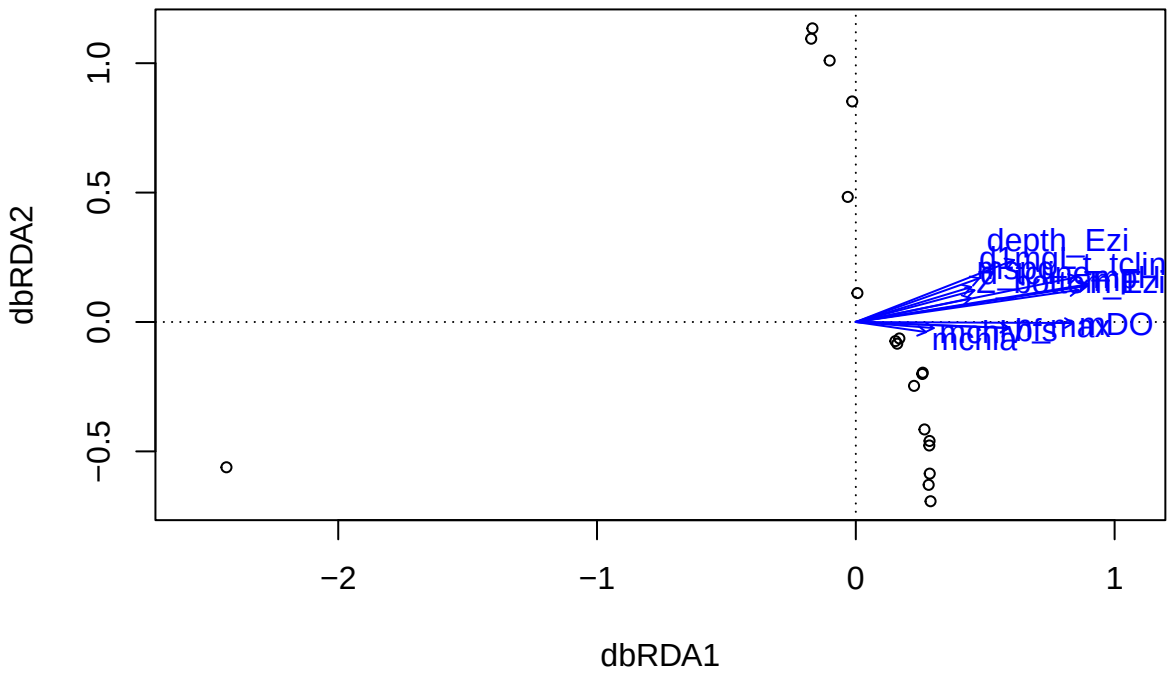
```
#check structure and change lake names to numeric values
#str(zoopfield)
zoopfield$Lake_Name <- revalue(zoopfield$Lake_Name, c("Airline" ="1", "Beaver Dam" ="2", "Beaver dam" =
```

```
## The following `from` values were not present in `x`: Beaver dam, T-lake, university, Willow, Wampler
```

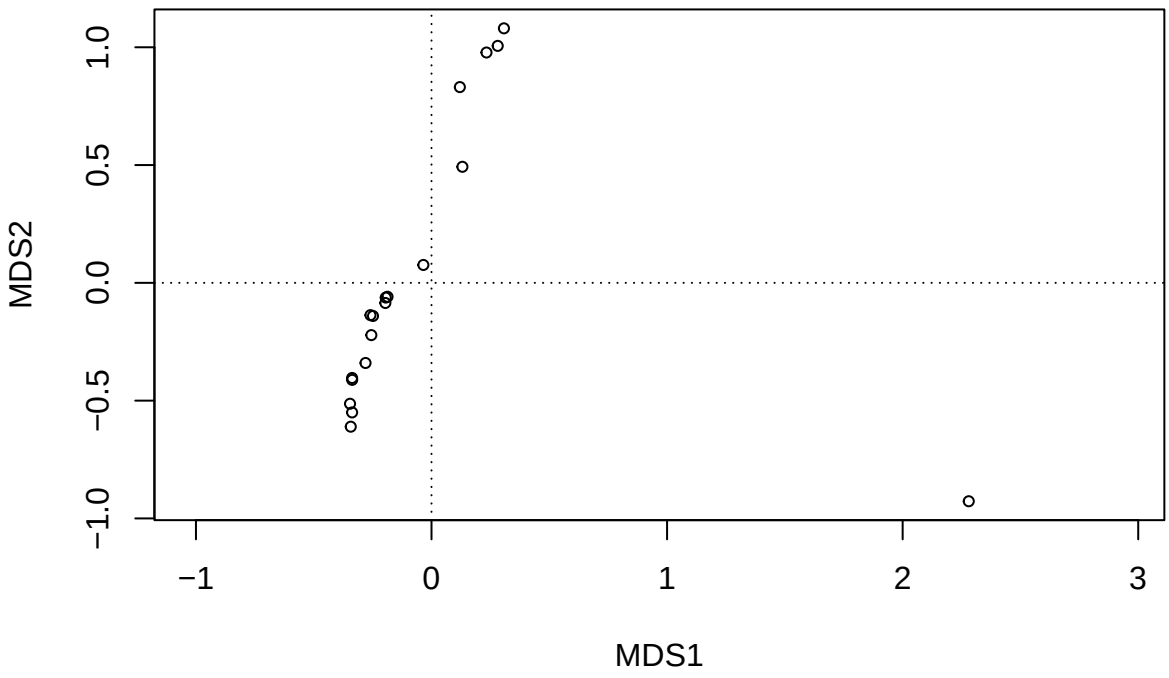
```
zoopfield$Lake_Name<- as.numeric(as.character(zoopfield$Lake_Name))


#subsetting data into a data frame with lake name, zooplankton species, and environmental characteristi
zoopwork <- zoopfield[, c(1, 37:49, 73:88)]

#aggregating all rounds for a lake (taking sum of species and environmental characteristics)
zoopagg <- zoopwork %>%
  group_by(Lake_Name)%>%
  summarize_all(sum, na.rm = TRUE)

#creating vectors for the factors to test hypothesis
species <- as.matrix(zoopagg[, c(2:14)])
env.lakes <- as.matrix(zoopagg[, c(15:30)])

library(vegan)
library(ade4)
#PERMANOVA
adonis2(species ~ env.lakes, method = "bray", permutations = 999)
```

```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = species ~ env.lakes, permutations = 999, method = "bray")
##           Df SumOfSqs      R2      F Pr(>F)
## env.lakes 14  2.14899 0.93746 4.2827  0.001 ***
## Residual   4  0.14337 0.06254
## Total     18  2.29235 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#bray-curtis dissimilarity
zoop.db <- vegdist(species, method = "bray", na.rm = TRUE)

#CONSTRAINED ORDINATION PLOT
zoop.dbrda <- dbrda(zoop.db ~ ., as.data.frame(env.lakes))
ordiplot(zoop.dbrda)
```




```
#model only the intercept
zoop.dbrda.mod0 <- dbrda(zoop.db ~ 1, as.data.frame(env.lakes))
ordiplot(zoop.dbrda.mod0)
```




```
#model full model, with all explanatory variables
zoop.dbrda.mod1 <- dbrda(zoop.db ~ ., as.data.frame(env.lakes))

#all combinations of explanatory variables in model-- function returns model with lowest AIC value
zoop.dbrda <- ordiR2step(zoop.dbrda.mod0, zoop.dbrda.mod1, perm.max = 200)
```

```
## Step: R2.adj= 0
## Call: zoop.db ~ 1
##
##                 R2.adjusted
## <All variables>  0.71856701
## + mT             0.28571569
## + mpH            0.28378726
## + t_tcline       0.26054910
## + T_Ezi          0.25800535
## + mDO            0.24120163
## + depth_Ezi      0.19510153
## + d1mgL          0.17580800
## + Z_bottom       0.16380010
## + d_tcline       0.15272119
## + n_max          0.10090019
## + bfs            0.10057823
## + mspc           0.04783092
## + mchlv          0.04563539
## + mchla          0.04145927
## <none>           0.00000000
```

```
## + wT             0.00000000
## + wAT            0.00000000
##
##      Df    AIC   F Pr(>F)
## + mT  1 11.256 8.2   0.04 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.2857157
## Call: zoop.db ~ mT
##
##                 R2.adjusted
## <All variables>   0.7185670
## + d1mgL           0.4334381
## + Z_bottom        0.4294195
## + d_tcline        0.4066537
## + depth_Ezi       0.3901793
## + T_Ezi           0.3327959
## + mchla           0.3235412
## + mchlv           0.3211948
## + t_tcline        0.3036682
## + mDO             0.2992198
## <none>            0.2857157
## + wT              0.2857157
## + wAT             0.2857157
## + mspc            0.2778121
## + n_max           0.2762490
## + bfs             0.2744745
## + mpH             0.2480732
##
##         Df    AIC      F Pr(>F)
## + d1mgL  1 7.7016 5.4325  0.008 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.4334381
## Call: zoop.db ~ mT + d1mgL
##
##                 R2.adjusted
## <All variables>   0.7185670
## + depth_Ezi       0.4663145
## + mDO             0.4614798
## <none>            0.4334381
## + wT              0.4334381
## + wAT             0.4334381
## + T_Ezi           0.4304369
## + Z_bottom        0.4294020
## + n_max           0.4292415
## + bfs             0.4278097
## + d_tcline        0.4270590
## + t_tcline        0.4254332
## + mspc            0.4153733
## + mchla           0.4152456
## + mchlv           0.4151559
```

```
## + mpH             0.4041366
##
##              Df    AIC      F Pr(>F)
## + depth_Ezi   1 7.3396 1.9856  0.148
```

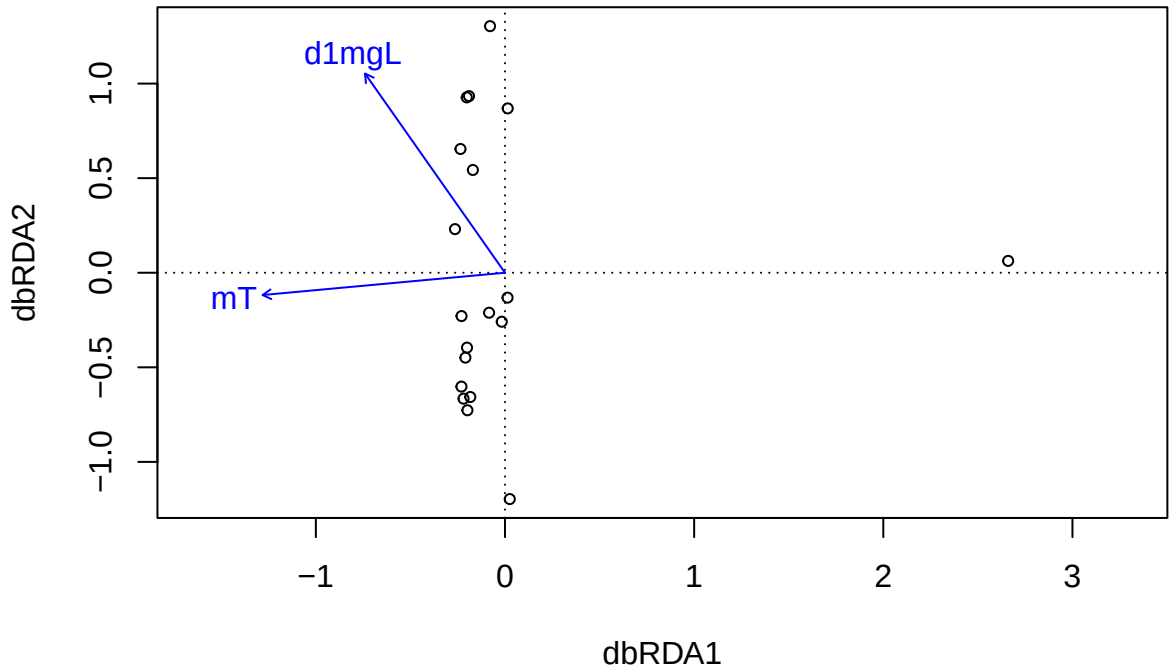
```
#which model was selected
zoop.dbrda$call
```

```
## dbrda(formula = zoop.db ~ mT + d1mgL, data = as.data.frame(env.lakes))
```

```
zoop.dbrda$anova
```

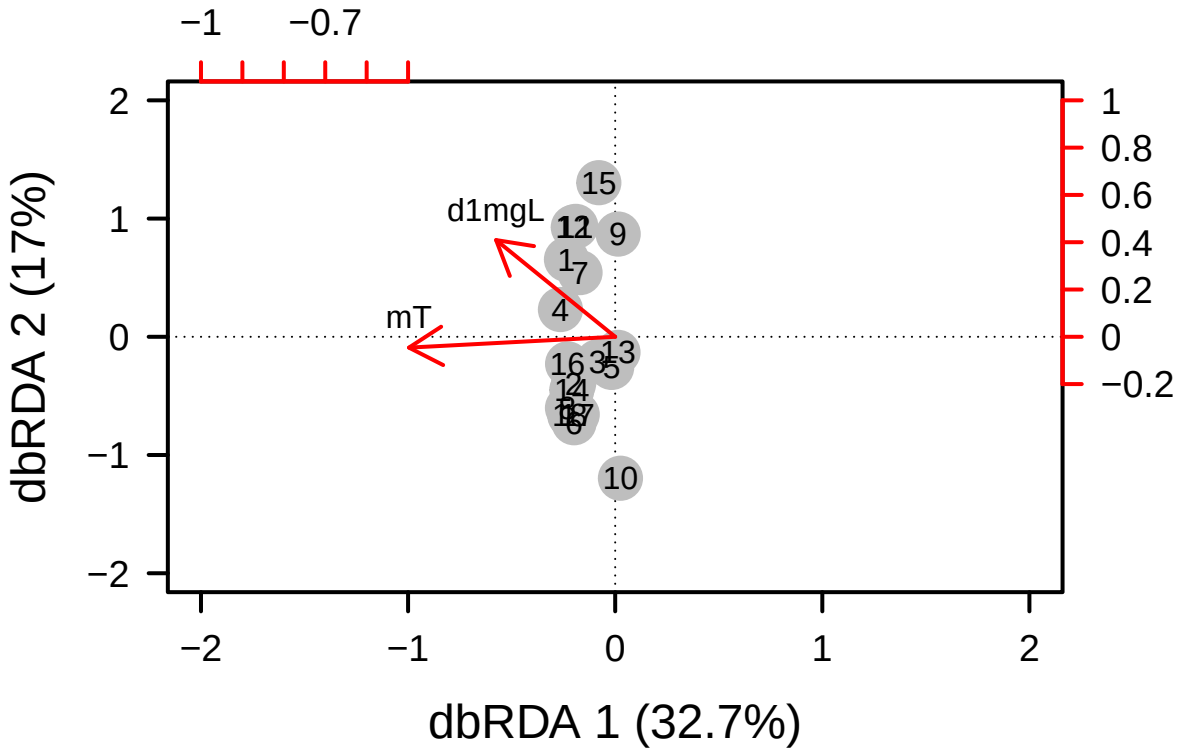
```
##                  R2.adj Df     AIC      F Pr(>F)
## + mT            0.28572  1 11.2557 8.2000  0.040 *
## + d1mgL         0.43344  1  7.7016 5.4325  0.008 **
## <All variables> 0.71857
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
ordiplot(zoop.dbrda)
```




```
#calculate explained variation
dbrda.explainvar1 <- round(zoop.dbrda$CCA$eig[1] / sum(c(zoop.dbrda$CCA$eig, zoop.dbrda$CA$eig)), 3) *
dbrda.explainvar2 <- round(zoop.dbrda$CCA$eig[2] / sum(c(zoop.dbrda$CCA$eig, zoop.dbrda$CA$eig)), 3) *
```

```
#define plot parameters
par(mar = c(5, 5, 4, 4) + 0.1)

#initiate plot
plot(scores(zoop.dbrda, display = "wa"), xlim = c(-2.0, 2.0),
     ylim = c(-2.0, 2.0), xlab = paste("dbRDA 1 (", dbrda.explainvar1, "%)",
                                       sep = ""), ylab = paste("dbRDA 2 (", dbrda.explainvar2, "%)", sep
     pch = 16, cex = 2.0, type = "n", cex.lab = 1.5, cex.axis = 1.2, 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 and labels
points(scores(zoop.dbrda, display = "wa"),
       pch = 19, cex = 3, bg = "gray", col = "gray")
text(scores(zoop.dbrda, display = "wa"),
     labels = row.names(scores(zoop.dbrda, display = "wa")))

#add environmental vectors
vectors <- scores(zoop.dbrda, display = "bp")
#row.names(vectors) <- rownames(vectors)

arrows(0, 0, vectors[,1], vectors[,2],
       lwd = 2, lty = 1, length = 0.2, col = "red")
text(vectors[,1], vectors[,2], pos = 3,
     labels = row.names(vectors))
axis(side = 3, lwd.ticks = 2, cex.axis = 1.2, las = 1, col = "red", lwd = 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", lwd = 2.2,
     at = pretty(range(vectors[, 2])) * 2, labels = pretty(range(vectors[, 2])))
```


−1
−0.7
2
1
0
−1
−2
dbRDA 2 (17%)
1
0.8
0.6
0.4
0.2
0
−0.2
d1mgL
mT
15
9
1
7
4
16
3
5
13
10
−2
−1
0
1
2
dbRDA 1 (32.7%)