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

Timothy Biewer-Heisler; Z620: Quantitative Biodiversity, Indiana University

22 April, 2021

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, April 23rd, 2021 before 09:00 AM**.

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 "*/8.BetaDiversity*" folder, and
4. load the `vegan` R package (be sure to install if needed).

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

```
## [1] "/Users/tbiewerh/GitHub/QB2021_Biewer-Heisler/2.Worksheets/8.BetaDiversity"
```

```
setwd("~/GitHub/QB2021_Biewer-Heisler/2.Worksheets/8.BetaDiversity")
#install.packages("vegan")
require("vegan")
```

```
## Loading required package: vegan
```

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

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

```
## This is vegan 2.5-7
```

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
# note, please do not print the dataset when submitting
package.list <- c('vegan', 'ade4','viridis', 'gplots', 'BiodiversityR', 'indicspecies')
for(package in package.list){
  if(!require(package, character.only = T, quietly = T)) {
    install.packages(package)
    library(package, character.only = T)
  }
}
```

```
##
## Attaching package: 'gplots'
```

```
## The following object is masked from 'package:stats':
##
##     lowess
```

```
## Registered S3 methods overwritten by 'lme4':
##   method                          from
##   cooks.distance.influence.merMod car
##   influence.merMod                car
##   dfbeta.influence.merMod         car
##   dfbetas.influence.merMod        car
```

```
## BiodiversityR 2.12-3: Use command BiodiversityRGUI() to launch the Graphical User Interface;
## to see changes use BiodiversityRGUI(changeLog=TRUE, backward.compatibility.messages=TRUE)
```

```
data(doubs)

str(doubs, max.level = 1)
```

```
## List of 4
##  $ env    :'data.frame': 30 obs. of  11 variables:
##  $ fish   :'data.frame': 30 obs. of  27 variables:
##  $ xy     :'data.frame': 30 obs. of  2 variables:
##  $ species:'data.frame': 27 obs. of  4 variables:
```

```
head(doubs$env)
```

```
##   dfs alt   slo flo pH har pho nit amm oxy bdo
## 1   3 934 6.176  84 79  45   1  20   0 122  27
## 2  22 932 3.434 100 80  40   2  20  10 103  19
```

```
## 3 102 914 3.638 180 83  52   5  22   5 105  35
## 4 185 854 3.497 253 80  72  10  21   0 110  13
## 5 215 849 3.178 264 81  84  38  52  20  80  62
## 6 324 846 3.497 286 79  60  20  15   0 102  53
```

```
e = 1
```

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.

```
quality <- c(rep("HQ", 13), rep("MQ", 5), rep("LQ", 6), rep("MQ", 5))
fish <- doubs$fish
fish <- fish[-8, ]
adonis(fish ~ quality, method = "bray", permutations = 999)
```

```
##
## Call:
## adonis(formula = fish ~ quality, permutations = 999, method = "bray")
##
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
##           Df SumsOfSqs MeanSqs F.Model      R2 Pr(>F)
## quality    2    3.0947 1.54733   10.97 0.45765  0.001 ***
## Residuals 26    3.6674 0.14105         0.54235
## Total     28    6.7621                 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
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   0.017 *
##
##  Group HQ+MQ  #sps.  2
##       stat p.value
## Satr 0.860   0.004 **
## Phph 0.859   0.009 **
##
##  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   0.001 ***
## Cyca 0.866   0.001 ***
## Acce 0.866   0.001 ***
## Lele 0.863   0.008 **
## Titi 0.853   0.003 **
## Chto 0.829   0.003 **
## Rham 0.829   0.001 ***
## Anan 0.829   0.002 **
## Eslu 0.827   0.018 *
## Pefl 0.806   0.009 **
## Blbj 0.791   0.002 **
## Scer 0.766   0.008 **
## Abbr 0.750   0.007 **
## 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   0.001 ***
## Satr 0.650   0.001 ***
##
##  Group LQ  #sps.  2
##       stat p.value
## Alal 0.693   0.001 ***
## Ruru 0.473   0.025 *
##
##  Group MQ  #sps.  4
##       stat p.value
## Anan 0.571   0.007 **
## Spbi 0.557   0.012 *
## Chto 0.542   0.011 *
## Icme 0.475   0.021 *
##
##  Group LQ+MQ  #sps.  9
##       stat p.value
## Legi 0.658   0.001 ***
## Baba 0.645   0.003 **
## Rham 0.600   0.003 **
## Acce 0.594   0.003 **
## Cyca 0.586   0.003 **
## Chna 0.571   0.003 **
## Blbj 0.571   0.009 **
## Gogo 0.523   0.013 *
## Abbr 0.499   0.017 *
## ---
## 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***: The PERMANOVA shows that overall fish species composition is associated for habitat quality. In terms of IndVal and phi it seems that all species are significantly associated with habitat quality. They all seems to agree with one another. The data here does seem to agree with our previous visualizations of quality. Satr and Alal, like previously mentioned, both seem to be significantly impacted by environment quality.

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%
## 0.107 0.144 0.181 0.220
## 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?

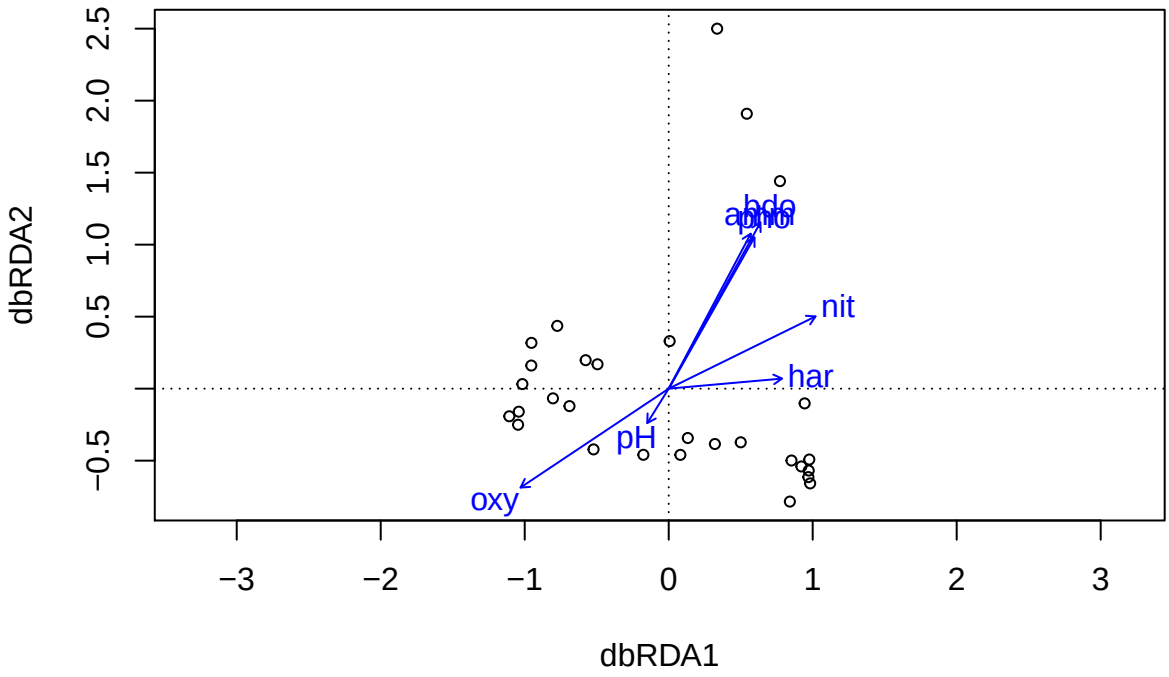
> ***Answer 2***: The test seems to show that fish diversity is positively correlated with stream environmental conditions with a correlation of 0.6 and a significance of 0.001. This supports previous data showing that environment quality is significantly correlated with fish diversity.

ii. Constrained Ordination

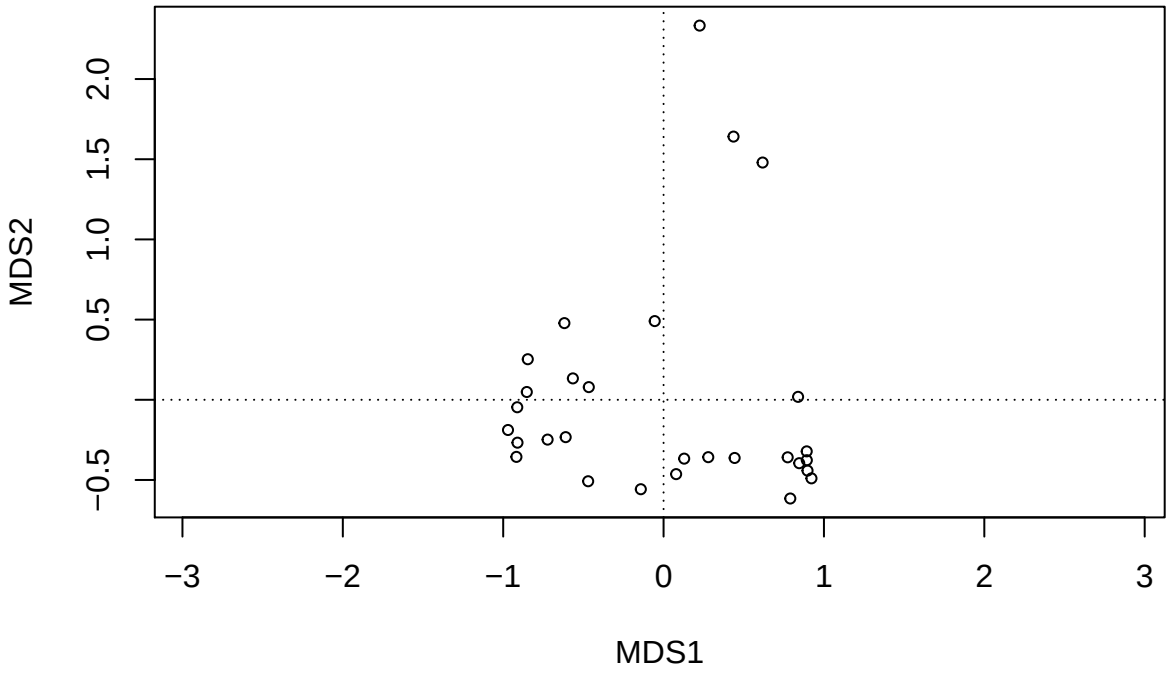
In the R code chunk below, do the following:

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 <- doubs$fish
fish <- fish[-8, ]

fish.dj <- vegdist(fish, methods = "jaccard", binary = TRUE)

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




```
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
```

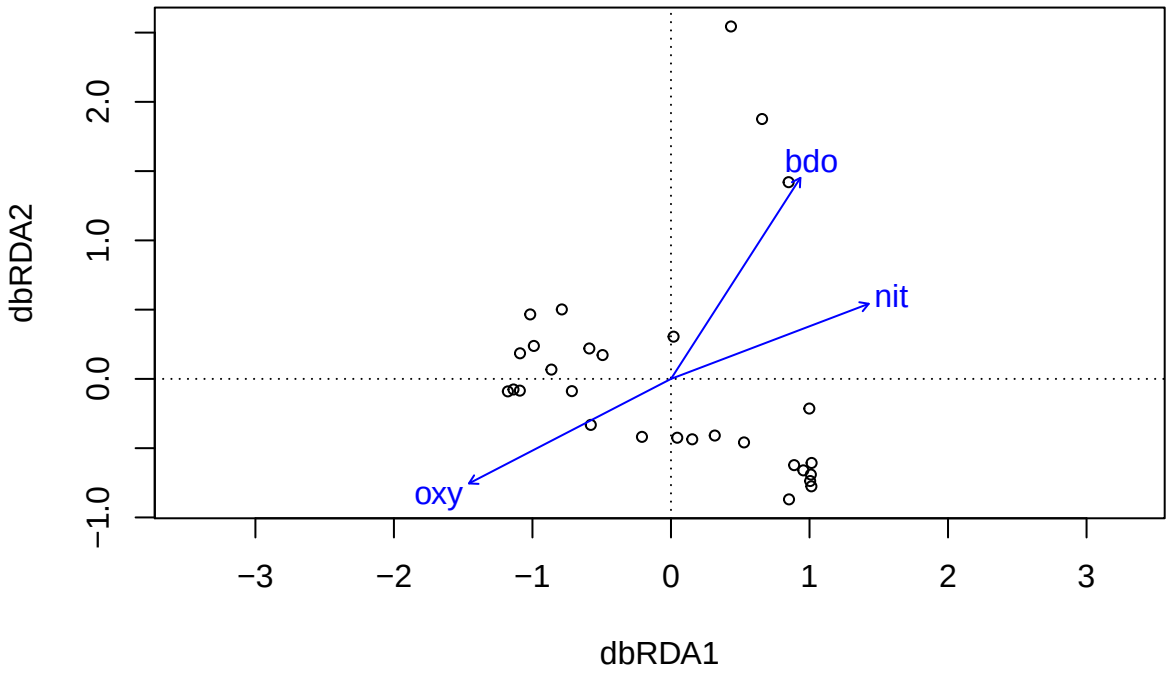
```
doubs.dbrda$call
```

```
## 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)
```

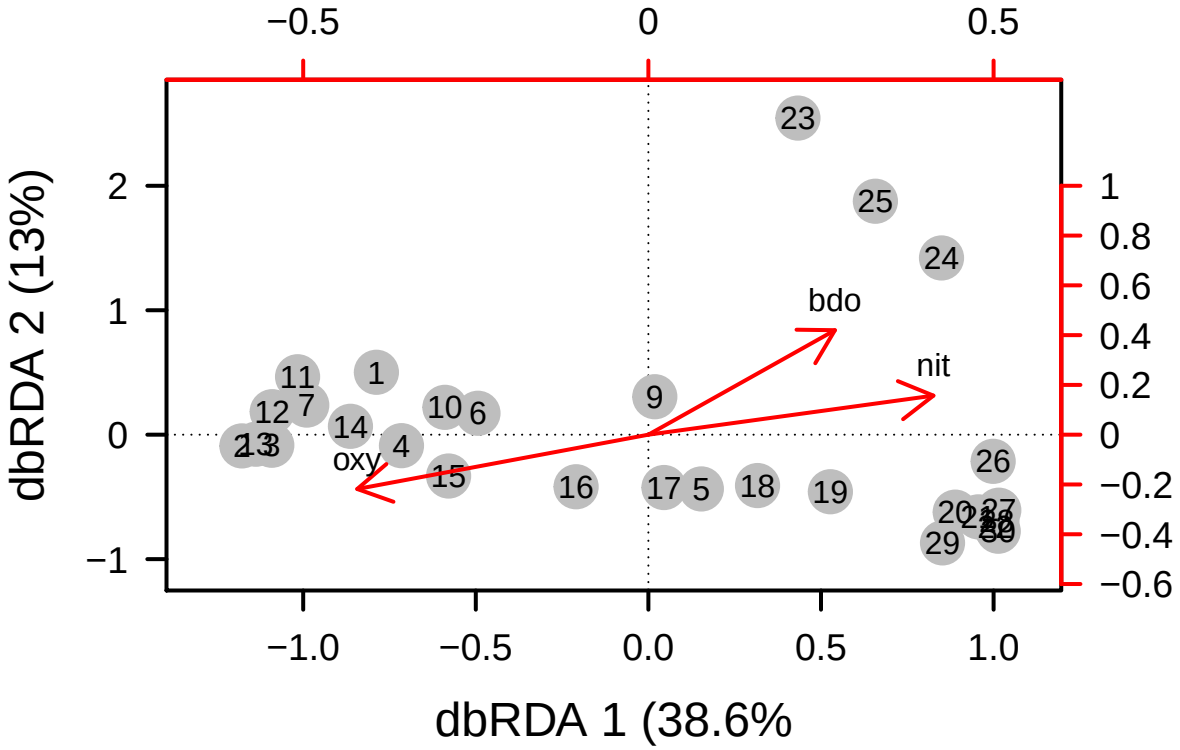
```
##
## Permutation test for dbrda under reduced model
##
## Permutation: free
## Number of permutations: 999
##
## Model: dbrda(formula = fish.db ~ oxy + bdo + nit, data =
## as.data.frame(env.chem))
## Permutation test for all constrained eigenvalues
##          Df Inertia      F Pr(>F)
```

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

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

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, 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

axis(side = 4, lwd.ticks = 2, cex.axis = 1.2, las = 1, col = "red", lwd = 2.2, at = pretty(range(vectors
```



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***: From the constrained ordination it seems like oxy, nit, and bdo seem to have an effect on variation in fish community structure.

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, perm.max = 200)
```

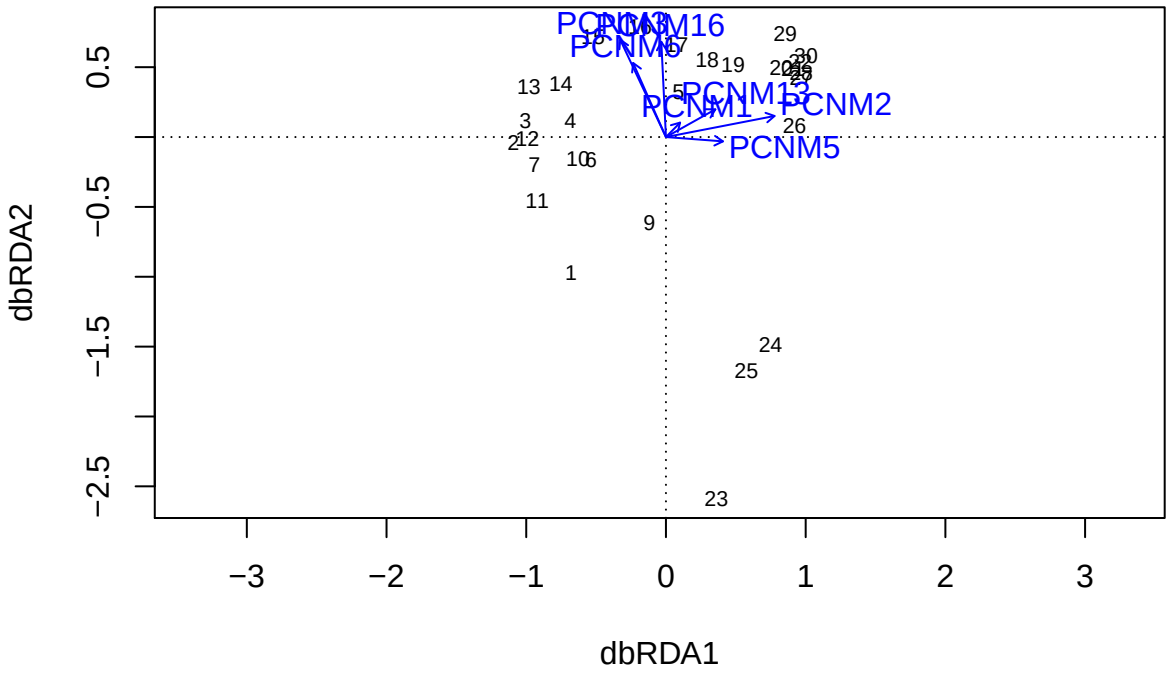
```
## Step: R2.adj= 0
## Call: fish.db ~ 1
##
##                   R2.adjusted
## <All variables>  0.626011301
## + PCNM2          0.235370423
## + PCNM3          0.078394885
## + PCNM13         0.065305668
## + PCNM5          0.046185074
## + PCNM6          0.032809156
## + PCNM16         0.030486700
## + 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  0.006 **
## ---
## 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  0.016 *
## ---
## 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  0.008 **
## ---
## 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  0.012 *
## ---
## 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  0.048 *
## ---
## 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  0.094 .
## ---
## 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  0.006 **
## + PCNM5         0.40760  1 43.941 3.8385  0.016 *
## + PCNM1         0.47215  1 41.411 4.0570  0.008 **
## + PCNM13        0.52124  1 39.346 3.4612  0.014 *
## + PCNM16        0.57680  1 36.480 4.0192  0.012 *
## + PCNM6         0.60431  1 35.182 2.5296  0.048 *
## <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]

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, permutation = 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
```

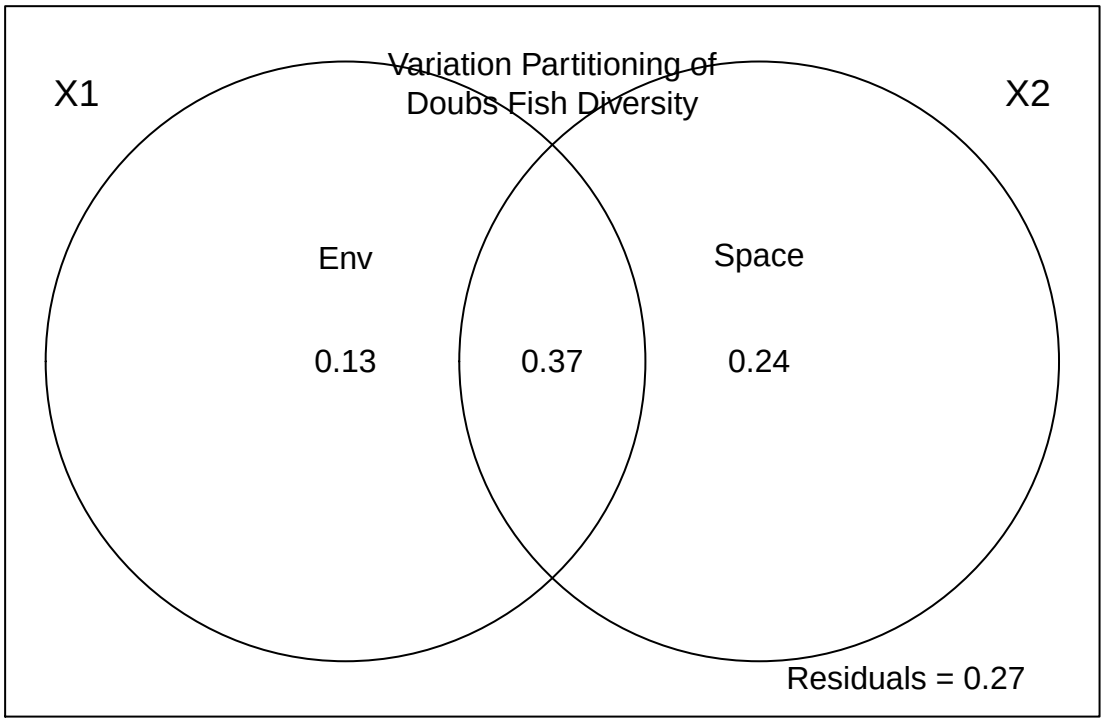
```
permutest(doubs.total.space, permutations = 999)
```

```
##
## 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+b] = 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]                    0                 0.36813    FALSE
## [c] = X2|X1            7                 0.23618     TRUE
## [d] = Residuals                          0.26574    FALSE
## ---
## Use function 'dbrda' to test significance of fractions of interest
```

```r
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***: 73% of the variation is due to environment, space, and environmnet and space, with 27% due to anything else. If you want the most fish distribution variation explained you would take the combination of environment and space. These together have the best chance at explaining the most of the variation. Environment taken alone only explains 13% of the variation, and so should not be considered without spatial variation.

SYNTHESIS

As in the previous worksheet, use the `mobsim` package from the DataWrangling module to simulate two local communities each containing 1000 individuals (N) and 25 species (S), but with one having a random spatial distribution and the other having a patchy spatial distribution. Take ten (10) subsamples from each site using the quadrat function and answer the following questions:

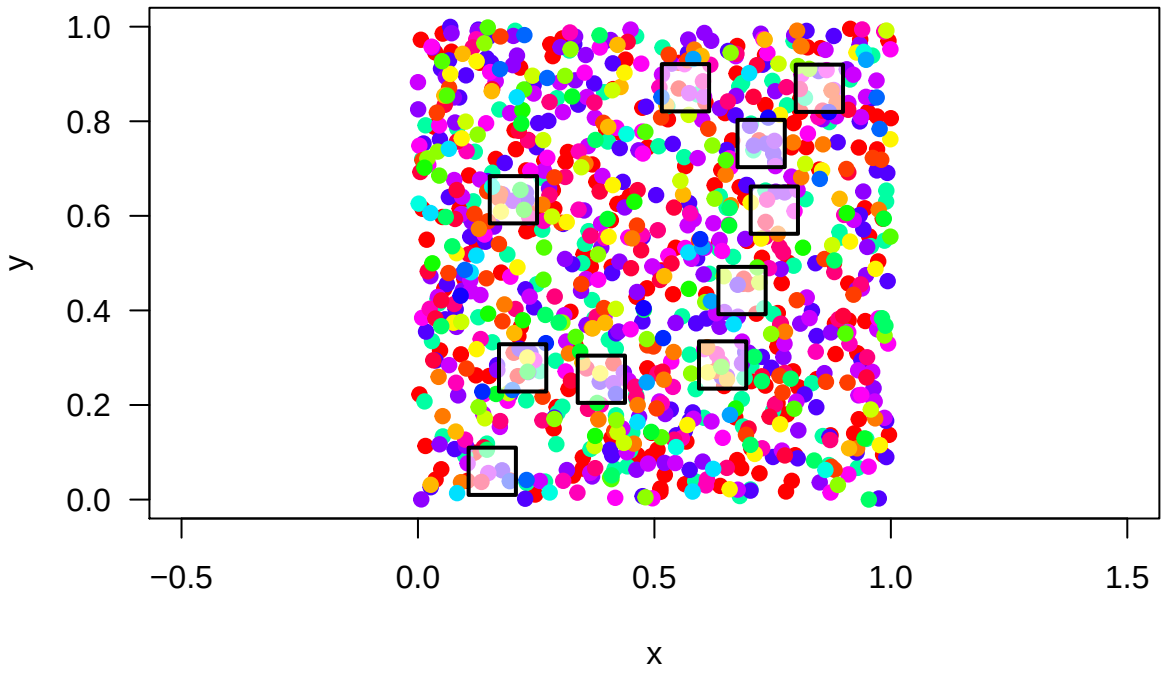
```
package.list <- c('mobsim', 'knitr', 'vegan', 'tidyr', 'dplyr', 'ggplot2', 'formatR')
for (package in package.list) {
  if (!require(package, character.only = TRUE, quietly = TRUE)) {
    install.packages(package)
    library(package, character.only = TRUE)
  }
}
```

```
##
## Attaching package: 'dplyr'
```

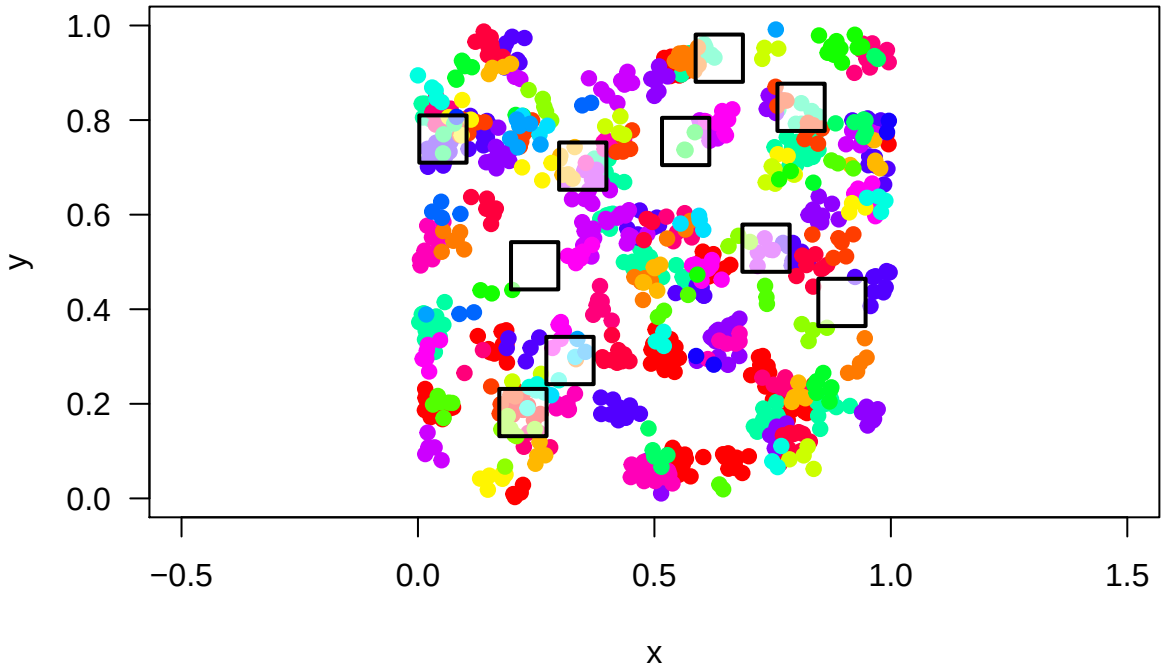
```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
# set page dimensions for printing
opts_chunk$set(tidy.opts = list(width.cutoff = 70),
        tidy = TRUE, fig.width = 5, fig.height = 5)

com <- sim_poisson_community(s_pool = 25, n_sim = 1000, sad_type = "lnorm",
        sad_coef = list("meanlog" = 2, "sdlog" = 1))
comm_mat <- sample_quadrats(com, n_quadrats = 10, quadrat_area = 0.01,
             method = "random", avoid_overlap = T)
```




```
comRan <- sim_thomas_community(s_pool = 25, n_sim = 1000, sad_type = "lnorm",
        sad_coef = list("meanlog" = 2, "sdlog" = 1))
comRan_mat <- sample_quadrats(comRan, n_quadrats = 10, quadrat_area = 0.01,
             method = "random", avoid_overlap = T)
```



1) Perform a PERMANOVA to test whether or not the spatial distribution of species affects species composition.

```
distrib <- c(rep("Thomas/Patchy", 10), rep("Poisson/Random", 10))
all_sites <- bind_rows(comRan_mat$spec_dat, comm_mat$spec_dat)

adonis(all_sites ~ distrib, method = "bray", permutations = 999)
```

```
## Warning in vegdist(lhs, method = method, ...): you have empty rows: their
```

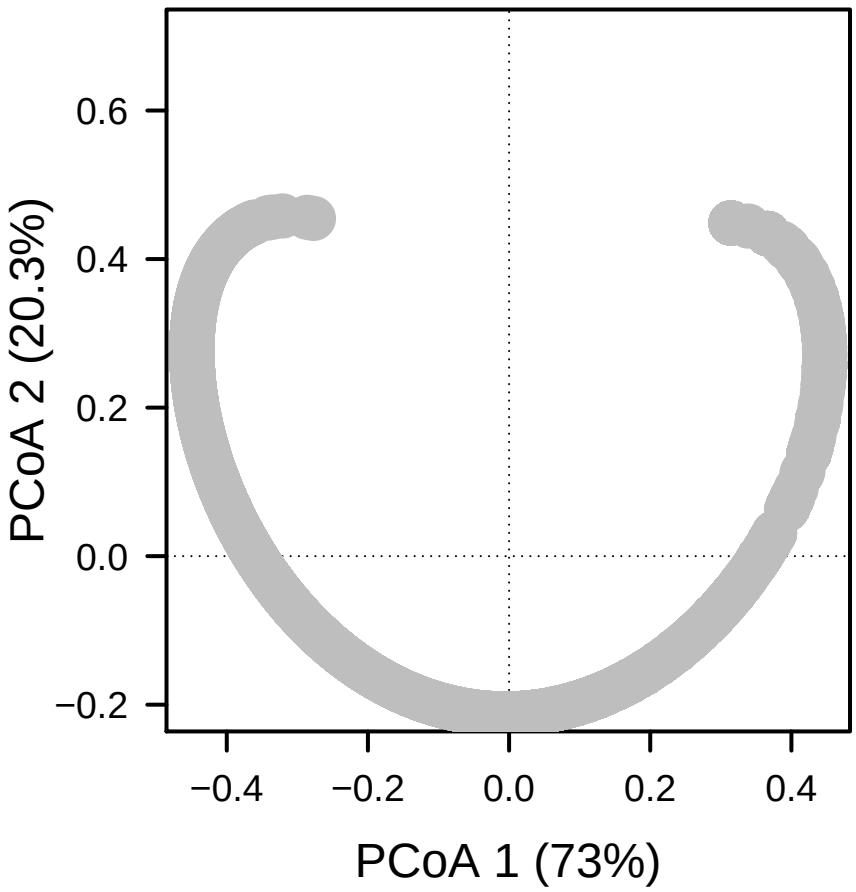
```
## dissimilarities may be meaningless in method "bray"

## Warning in vegdist(lhs, method = method, ...): missing values in results

##
## Call:
## adonis(formula = all_sites ~ distrib, permutations = 999, method = "bray")
##
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
##           Df SumsOfSqs MeanSqs F.Model R2 Pr(>F)
## distrib    1
## Residuals 18
## Total     19
```

This ANOVA shows that there is a statistically significant affect spacial distribution of species between random and patchy distributions.

2) Load the dataset you are using for your Team Project. Perform an ordination to visualize your dataset. Using this ordination, develop some hypotheses relevant to β-diversity. Use a statistic to test one of these hypotheses. Succinctly explain the finding and its relevance to your system.

```
setwd("~/GitHub/QB2021_Team3")

foodWebs <- read.table("135FoodWebs.txt", sep = "\t", header = TRUE, row.names = 1)

foodWebs95 <- subset(foodWebs, Web.ID %in% c(95))
exponentiated <- exp(foodWebs$Log.Abundance.)

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



pcoa <- cmdscale(abund.db, eig = TRUE, k = 3)

explainvar1 <- round(pcoa$eig[1]/sum(pcoa$eig), 3) * 100
explainvar2 <- round(pcoa$eig[2]/sum(pcoa$eig), 3) * 100
explainvar3 <- round(pcoa$eig[3]/sum(pcoa$eig), 3) * 100
sum.eig <- sum(explainvar1, explainvar2, explainvar3)

par(mar = c(5, 5, 1, 2) + 0.1)

plot(pcoa$points[, 1], pcoa$points[, 2], ylim = c(-0.2, 0.7), xlab = paste("PCoA 1 (",
    explainvar1, "%)", sep = ""), ylab = paste("PCoA 2 (", explainvar2,
    "%)", sep = ""), pch = 16, cex = 2, type = "n", cex.lab = 1.5, cex.axis = 1.2,
    axes = 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)
```

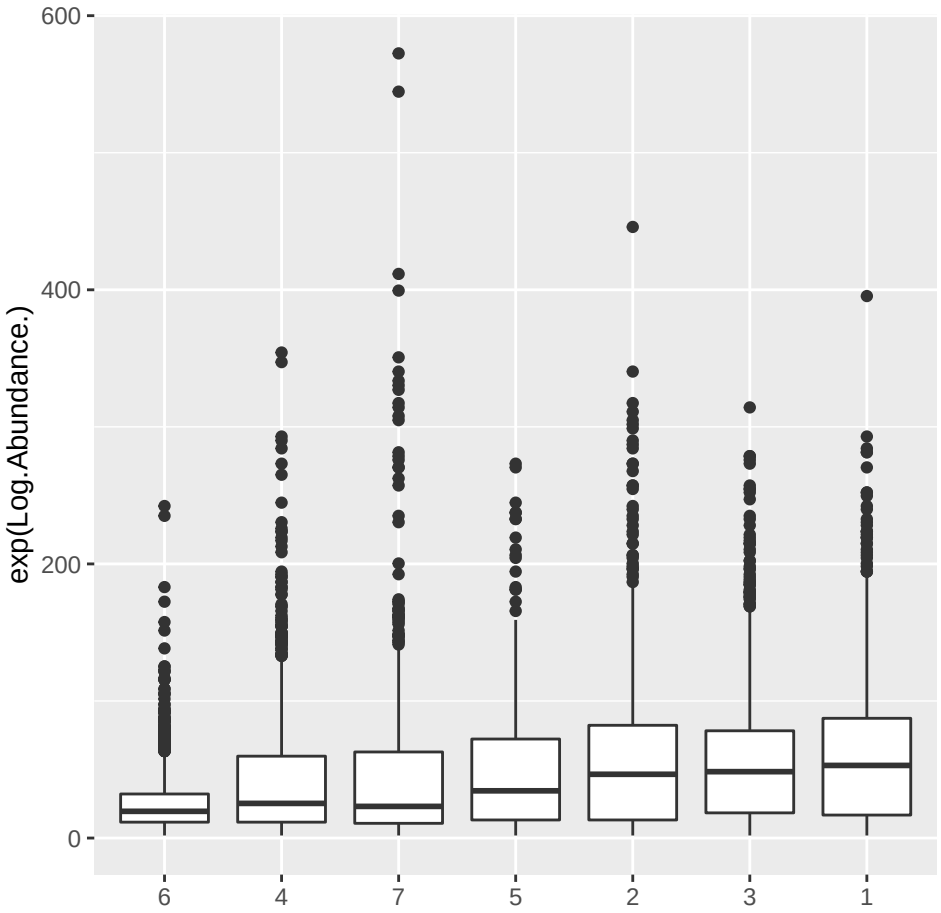
```
points(pcoa$points[, 1], pcoa$points[, 2], pch = 19, cex = 3, bg = "gray",
    col = "gray")
text(pcoa$points[, 1], pcoa$points[, 2], labels = row.names(pcoa$points))
```



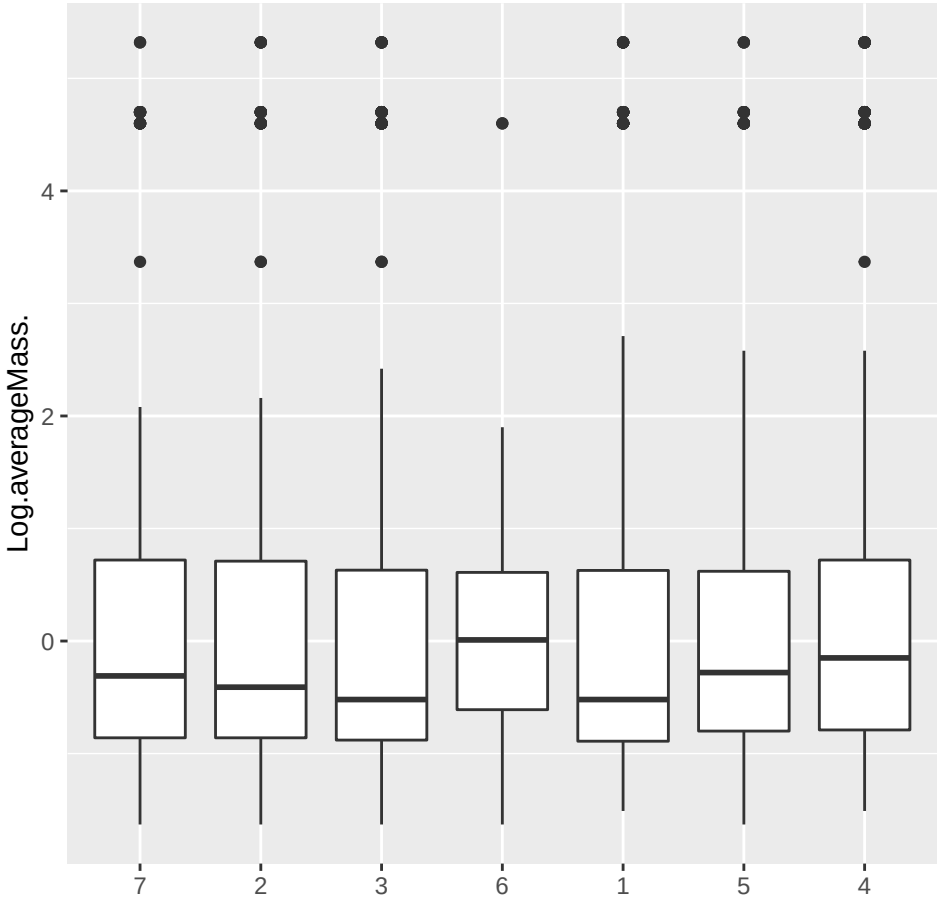

```
webTaxon <- data.frame(foodWebs$Web.ID, foodWebs$Taxon.ID)
```

We hypothesize that from the initial ordination plot the differences in abundance that we have are mainly clustered by human development effects. To do this we wil try to show a trend for abundance by environment. Below, using ANOVA we have shown that there is indeed a statistically significant difference in abundance between environments. Whether its due to greater human interaction with the environment is unknown.

```
setwd("~/GitHub/QB2021_Team3")

library(dplyr)
library(ggplot2)

dafra <- foodWebs %>% group_by(Ecosystem.Type.ID) %>% summarise(me = mean(exp(Log.Abundance.)))

library(forcats)

foodWebs %>% mutate(Ecosystem.Type.ID = fct_reorder(as.factor(Ecosystem.Type.ID),
    exp(Log.Abundance.), .fun = "mean")) %>% ggplot(aes(x = as.factor(Ecosystem.Type.ID),
    y = exp(Log.Abundance.))) + geom_boxplot() + xlab("Ecosystem Type ID") +
    theme(legend.position = "none") + xlab("")
```




```
foodWebs %>% mutate(Ecosystem.Type.ID = fct_reorder(as.factor(Ecosystem.Type.ID),
    Log.averageMass., .fun = "mean")) %>% ggplot(aes(x = as.factor(Ecosystem.Type.ID),
    y = Log.averageMass.)) + geom_boxplot() + xlab("Ecosystem Type ID") +
    theme(legend.position = "none") + xlab("")
```




```
model = lm(I(Log.Abundance.) ~ Ecosystem.Type.ID, data = foodWebs)
summary(model)
```

```
##
## Call:
## lm(formula = I(Log.Abundance.) ~ Ecosystem.Type.ID, data = foodWebs)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.0392 -0.8027  0.0173  0.7816  3.2438
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        3.751328   0.027234  137.75   <2e-16 ***
## Ecosystem.Type.ID -0.092160   0.005804  -15.88   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.01 on 7323 degrees of freedom
## Multiple R-squared:  0.03329,    Adjusted R-squared:  0.03316
## F-statistic: 252.2 on 1 and 7323 DF,  p-value: < 2.2e-16
```