

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

Dustin Brewer; Z620: Quantitative Biodiversity, Indiana University

10 February, 2019

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**).

The completed exercise is due on **Wednesday, February 13th, 2019 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 “/8.BetaDiversity” folder, and
4. load the **vegan** R package (be sure to install if needed).

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

```
## [1] "C:/Users/dusti/GitHub/QB2019_Brewer/2.Worksheets/8.BetaDiversity"
```

```
library(vegan)

## Loading required package: permute

## Loading required package: lattice

## This is vegan 2.5-3

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

##
## Attaching package: 'gplots'

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

## BiodiversityR 2.11-1: Use command BiodiversityRGUI() to launch the Graphical User Interface;
## to see changes use BiodiversityRGUI(changeLog=TRUE, backward.compatibility.messages=TRUE)
```

2) LOADING DATA

Load dataset

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

```
# note, please do not print the dataset when submitting
library(ade4)
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 that will bin based on assumptions about habitat quality
```

```
quality <- c(rep("HQ", 13), rep("MQ", 5), rep("LQ", 6), rep("MQ", 5))
```

```
#Run a 'permanova,' which is a multivariate analog to univariate anova.
```

```
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
```

```
#determine if any species can be used as an indicator of a particular habitat type. Value close to 1 su
```

```
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 0.025 *
```

```
##
```

```
## Group HQ+MQ #sps. 2
```

```
##      stat p.value
```

```
## Satr 0.860    0.004 **
## Phph 0.859    0.007 **
##
## 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    0.004 **
## Titi 0.853    0.005 **
## Chto 0.829    0.003 **
## Rham 0.829    0.001 ***
## Anan 0.829    0.001 ***
## Eslu 0.827    0.022 *
## Pefl 0.806    0.016 *
## Blbj 0.791    0.003 **
## Scer 0.766    0.012 *
## Abbr 0.750    0.010 **
## Icme 0.661    0.029 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

additionally, can determine habitat preferences of each species, with 'phi.' -1 indicates strong avoidance

```
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  0.040 *
##
## Group MQ #sps. 4
##      stat p.value
## Anan 0.571  0.006 **
## Spbi 0.557  0.009 **
## Chto 0.542  0.008 **
## Icme 0.475  0.023 *
##
## Group LQ+MQ #sps. 9
##      stat p.value
## Legi 0.658  0.003 **
## Baba 0.645  0.003 **
## Rham 0.600  0.009 **
## Acce 0.594  0.004 **
## Cyca 0.586  0.005 **
## Chna 0.571  0.010 **
## Blbj 0.571  0.006 **
## Gogo 0.523  0.013 *
## Abbr 0.499  0.026 *
## ---
## 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, IndVal, and phi coefficient analyses all relate to understanding relations of fish species to the habitat categories that we specified. The PERMANOVA specified that there are differences with respect to species composition amongst the groups, the IndVal analysis showed which species indicate particular habitat categories and how strongly, and the phi analysis showed habitat preferences. So, all measures were related, but allowed for slightly different inferences (e.g., the IndVal analysis showed that Teso is a good indicator of medium quality habitat but the phi analysis does not show that this species prefers that habitat type).

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

```
#Compare the matrices with a 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%
## 0.109 0.141 0.174 0.206
## 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 Mantel test results suggest that fish diversity is correlated ($p = .001$) with environmental conditions. This provides support for the hypothesis that stream conditions affects fish communities.

ii. Constrained Ordination

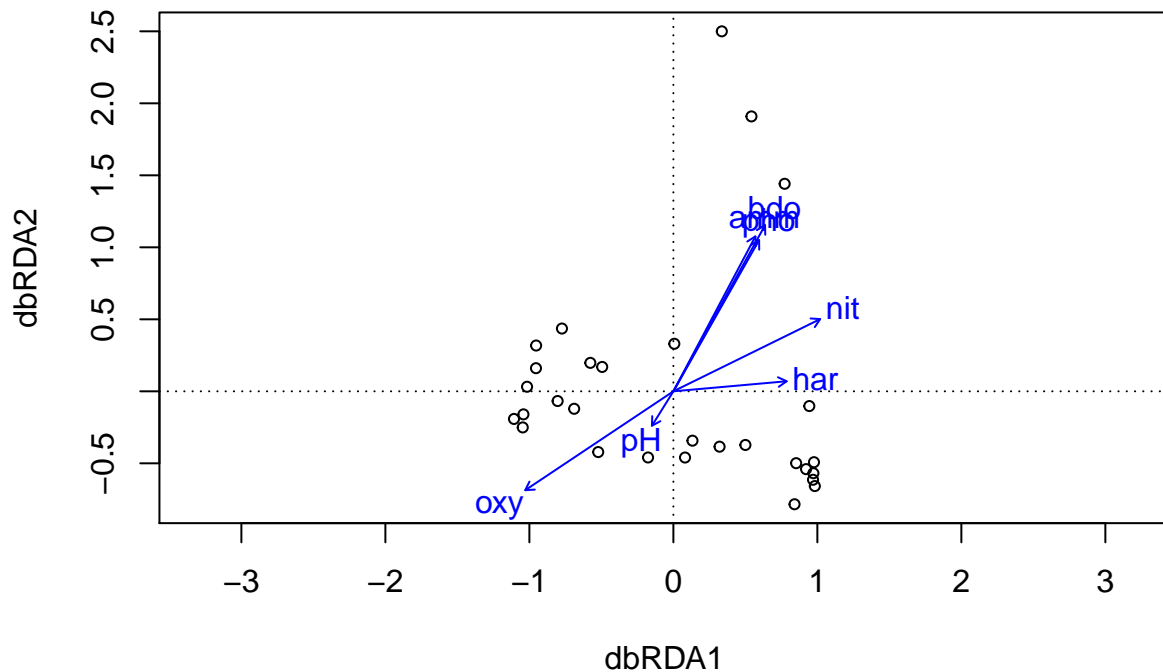
In the R code chunk below, do the following:

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 an environmental matrix
env.chm <- as.matrix(doubs$env[-8, 5:11])

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

#Perform dbRDA, which seems to take a PCoA further
doubs.dbrda <- dbrda(fish.db ~ ., as.data.frame(env.chm))
ordiplot(doubs.dbrda)
```



```
#check for pairwise correlations of environmental variables (should minimize use of correlated explanat
psych::corr.test(env.chm)
```

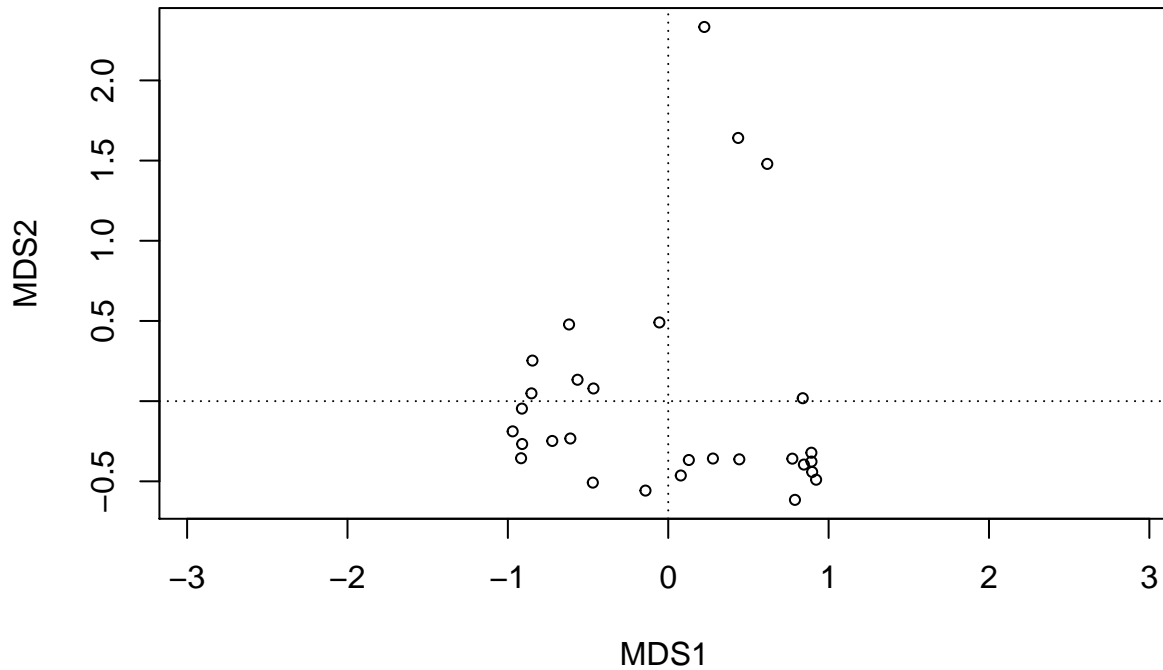
```
## Call:psych::corr.test(x = env.chm)
## Correlation matrix
##      pH  har  pho  nit  amm  oxy  bdo
## pH   1.00 0.08 -0.08 -0.04 -0.12 0.19 -0.16
## har  0.08 1.00 0.37 0.53 0.30 -0.37 0.34
## pho -0.08 0.37 1.00 0.80 0.97 -0.76 0.91
## nit -0.04 0.53 0.80 1.00 0.80 -0.69 0.68
## amm -0.12 0.30 0.97 0.80 1.00 -0.75 0.90
## oxy  0.19 -0.37 -0.76 -0.69 -0.75 1.00 -0.84
## bdo -0.16 0.34 0.91 0.68 0.90 -0.84 1.00
## Sample Size
## [1] 29
## Probability values (Entries above the diagonal are adjusted for multiple tests.)
##      pH  har  pho  nit  amm  oxy  bdo
## pH   0.00 1.00 1.00 1.00 1.00 1.00 1.00
## har  0.66 0.00 0.46 0.03 0.83 0.46 0.59
## pho  0.68 0.05 0.00 0.00 0.00 0.00 0.00
## nit  0.83 0.00 0.00 0.00 0.00 0.00 0.00
## amm  0.53 0.12 0.00 0.00 0.00 0.00 0.00
## oxy  0.32 0.05 0.00 0.00 0.00 0.00 0.00
## bdo  0.40 0.07 0.00 0.00 0.00 0.00 0.00
##
## To see confidence intervals of the correlations, print with the short=FALSE option
```

```

#Model selection, using AIC. First, model only the intercept (which does not take into account explanat
doubts.dbrda.mod0 <- dbrda(fish.db ~ 1, as.data.frame(env.chm))

#Have a look (note, no vectors)
ordiplot(doubts.dbrda.mod0)

```



```

#Now, model the full model with all explanatory variables
doubts.dbrda.mod1 <- dbrda(fish.db ~ ., as.data.frame(env.chm))

#Next, generate at all of the possible combinations of explanatory variables so that the best one can
doubts.dbrda <- ordiR2step(doubts.dbrda.mod0, doubts.dbrda.mod1, per.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

```

```

# Now, look at the model that was selected
doubts.dbrda$call

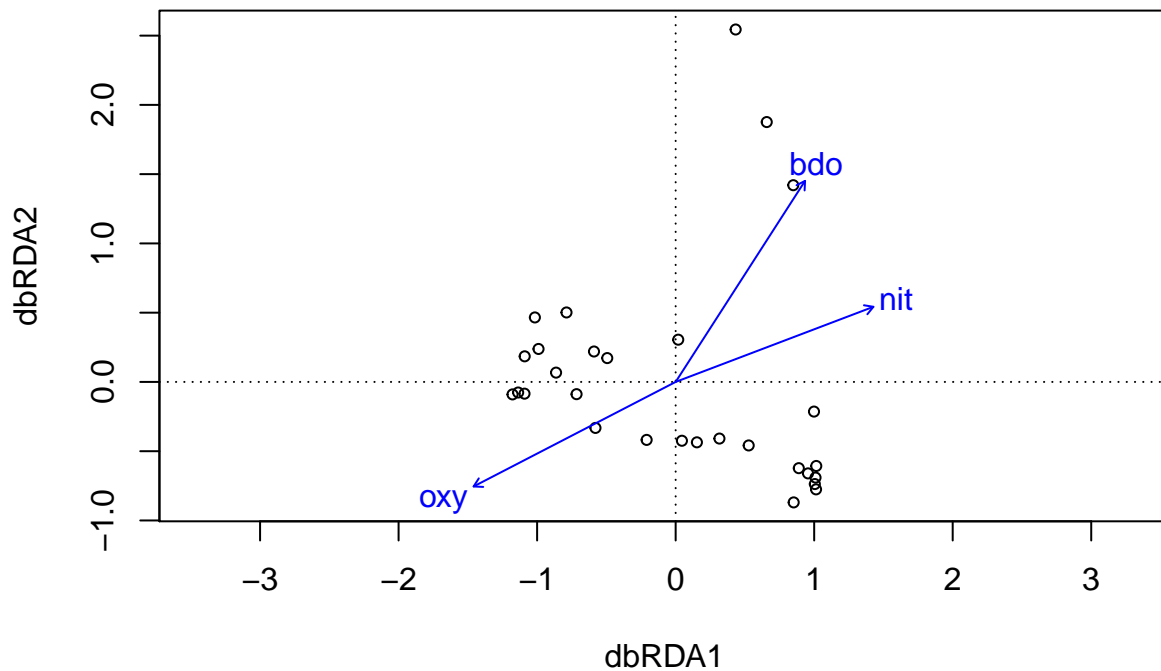
```

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

```
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)
```



```
#permutation test to evalulate 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.chm))
## 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
```

```
#permutation test to evaluate environmental variables
envfit(doubs.dbrda, env.chm[,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
```

```
#Calculate explained variation
dbrda.explainvar1 <- round(doubs.dbrda$CCA$eig[1] / sum(c(doubs.dbrda$CCA$eig, doubs.dbrda$CA$eig)), 3)
dbrda.expllainvar2 <- round(doubs.dbrda$CCA$eig[2] /
                           sum(c(doubs.dbrda$CCA$eig, doubs.dbrda$CA$eig)), 3) * 100
```

```
#Plot ordination
# 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.expllainvar2, "%)", 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)
```

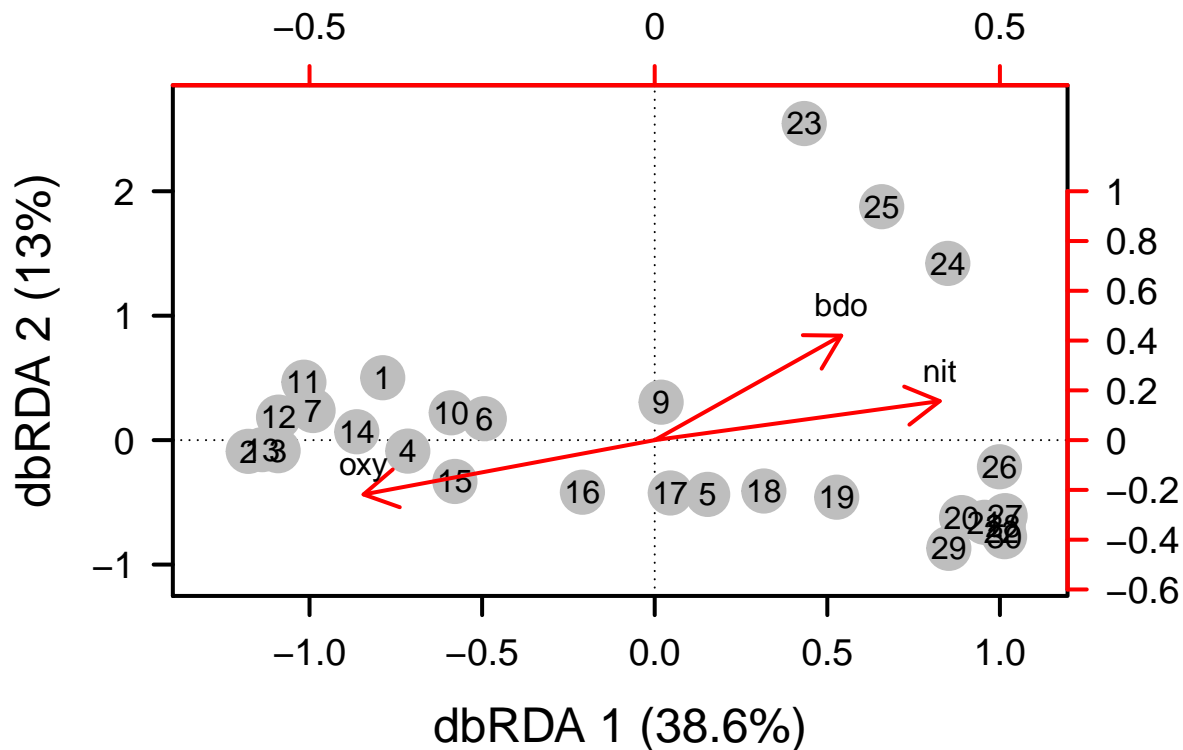
```
#Add points and labels
points(scores(doubs.dbrda, display = "wa"),
       pch = 19, cex = 3, bg = "gray", col = "gray")
text(scores(doubs.dbrda, display = "wa"),
     labels = row.names(scores(doubs.dbrda, display = "wa")))
```

```
#Add environmental vectors
vectors <- scores(doubs.dbrda, display = "bp")
```

```

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: It appears that bdo (highly correlated with pho), oxy, and nit are variables that seem to be contributing to 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.

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

```
#Create a spatial model (several steps)
#First, weight each site by it relative abundance
rs <- rowSums(fish)/sum(fish)
```

```
#perform a Principal Coordinates of Neighbor Matrices
doubts.pcnmw <- pcnm(dist(doubts$xy[-8,]), w = rs, dist.ret = T)
```

```
#Note that PCNM can return negative eigenvalues, but only positive values have meaning
doubts.pcnmw$values > 0
```

```
## [1] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## [12] TRUE TRUE TRUE TRUE TRUE TRUE FALSE FALSE FALSE FALSE FALSE
## [23] FALSE FALSE FALSE FALSE
```

```
#Need to perform model selection (forward and backwards) to determine which eigenvalues create the mo

doubts.space <- as.data.frame(scores(doubts.pcnmw))
doubts.pcnm.mod0 <- dbrda(fish.db ~ 1, doubts.space)
doubts.pcnm.mod1 <- dbrda(fish.db ~ ., doubts.space)
step.pcnm <- ordiR2step(doubts.pcnm.mod0, doubts.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 0.006 **
## ---
## 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 *
## ---

```

```

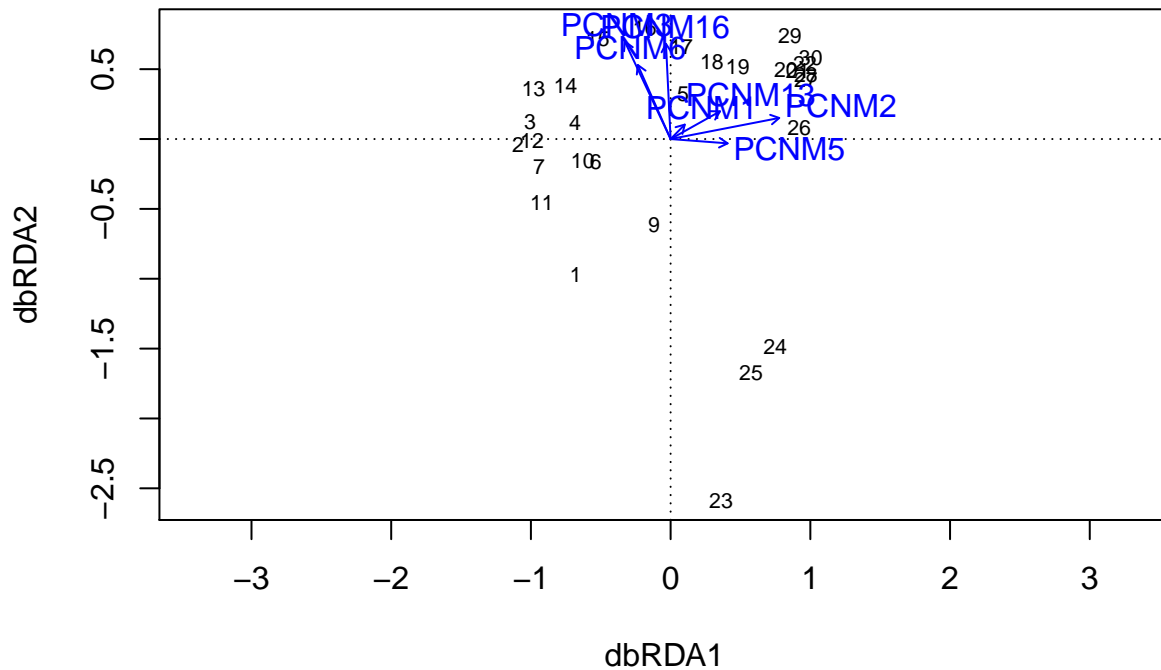
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.5212427
## Call: fish.db ~ PCNM2 + PCNM3 + PCNM5 + PCNM1 + PCNM13
##
##               R2.adjusted
## <All variables> 0.6260113
## + PCNM16       0.5767968
## + PCNM15       0.5715331
## + PCNM14       0.5698343
## + 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 0.042 *
## ---
## 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 0.078 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# visualize the dbRDA
plot(step.pcnm)
```



```
#Now, the object step.pcnm 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 0.006 **
## + 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 0.042 *
## <All variables> 0.62601
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

#Perform a partial constrained ordination
#conduct constrained ordinations
dubs.total.env <- dbrda(fish.db ~ env.mod)
dubs.total.space <- dbrda(fish.db ~ space.mod)

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

#Test for significance
permutest(dubs.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(dubs.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
```

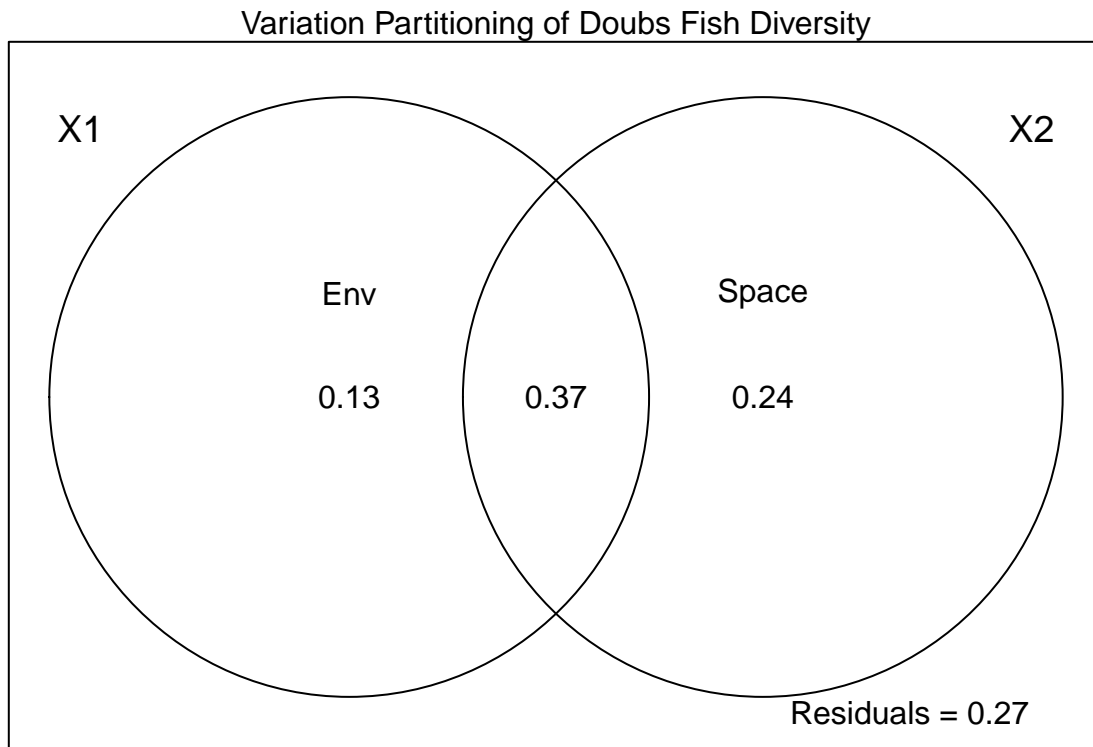
#Calculate fraction of variation explained by environment alone, space alone, environment and space, and

```
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
```

```
#plot
par(mar = c(2,2,2,2))
plot(doubs.varpart)
text(1,0.25, "Space")
text(0, 0.25, "Env")
mtext("Variation Partitioning of Doubs Fish Diversity")
```



Question 4: Interpret the variation partitioning results.

Answer 4: It appears that space alone explains more variability in fish community structure (24%) than environment alone (13%), and that together space and environment explain 37% of Doubs fish diversity.

SYNTHESIS

- 1) Using the jelly bean data from class (i.e., JellyBeans.txt), perform a PERMANOVA to test whether or not the vicariance event (random splitting of source community) had an effect on jelly bean composition. Based on your previous analyses with this data set, what are your thoughts about the importance of stochastic vs. deterministic factors on estimates of biodiversity?

Based on the PERMANOVA (see below), and previous analyses, it appears that a random split did in fact have an effect on jelly bean composition. This is a reminder that that stochastic factors are important to consider when estimating biodiversity, and that a purely deterministic estimate may not be sufficient.

```
Jelly <- read.csv("JellyBeanSampleData.csv")
Jelly <- as.data.frame(Jelly)

Site <- c(rep("A", 5), rep("B", 4))

adonis(Jelly[3:28] ~ Site, method = "bray", permutations = 999)

##
## Call:
## adonis(formula = Jelly[3:28] ~ Site, permutations = 999, method = "bray")
##
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
##          Df SumsOfSqs  MeanSqs F.Model    R2 Pr(>F)
## Site      1  0.08741 0.087413  1.9732 0.2199  0.061 .
## Residuals  7  0.31011 0.044301      0.7801
## Total     8  0.39752      1.0000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- 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.

As seen below, almost all of the sites that were affected the least by hemlock woolly adelgid (L's) were grouped at the bottom right to the ordination, and the sites that were affected the most (HA's and HB's) also appeared grouped. Therefore, I hypothesized that bird communities vary based on level of hemlock woolly adelgid impact. A PERMANOVA indicated that, indeed, there are differences between the sites affected at the low, medium, and high A/B levels with respect to avian community composition. This finding suggests that hemlock woolly adelgid is affecting bird communities.

```
Woolly <- read.csv("hf085-01-bird.csv")
Woolly <- as.data.frame(Woolly)

Woolly.db <- vegdist(Woolly[,3:49], method = "bray")
Woolly.pcoa <- cmdscale(Woolly.db, eig = TRUE, k = 3)
Woolly.pcoa
```

```

## $points
##           [,1]           [,2]           [,3]
## [1,] -0.33368920  0.129895661  0.093102653
## [2,] -0.10586899  0.069986593 -0.225276642
## [3,] -0.24094763  0.220733029  0.112657558
## [4,] -0.02641162  0.096215053  0.050319624
## [5,] -0.08886380 -0.411287157  0.097355215
## [6,]  0.20432689 -0.428625310  0.091270630
## [7,] -0.24767379 -0.228159700 -0.111731941
## [8,]  0.07180444 -0.033661685  0.035565269
## [9,] -0.23624721 -0.138494730 -0.012674971
## [10,] 0.13030991 -0.151249225 -0.251227010
## [11,] 0.11991373 -0.037593074 -0.393188786
## [12,] -0.33846639  0.149869407  0.110969999
## [13,] -0.38744417  0.183168658 -0.042262518
## [14,]  0.19878301 -0.324582509  0.170713510
## [15,] -0.27832171 -0.049178239 -0.028024589
## [16,] -0.25464095  0.217099447 -0.091691005
## [17,] -0.37837554  0.176233503 -0.033144159
## [18,] -0.06969224  0.116860540 -0.026296566
## [19,] -0.22452949 -0.140875461 -0.119712265
## [20,] -0.22806020 -0.345634770  0.024701722
## [21,] -0.08886380 -0.411287157  0.097355215
## [22,]  0.14861660  0.401432573 -0.088051497
## [23,]  0.33333034  0.077669035 -0.045282552
## [24,]  0.26144591 -0.169488964 -0.007110593
## [25,]  0.36846100  0.021574496 -0.099915380
## [26,]  0.14188830  0.288816149  0.419166842
## [27,]  0.37525307 -0.020181214  0.053878764
## [28,]  0.33374896  0.001800701 -0.005741973
## [29,]  0.12407830  0.234000640  0.211197320
## [30,]  0.20890781  0.096498426 -0.120313288
## [31,]  0.33412920  0.134258863 -0.220407688
## [32,]  0.37438332  0.095520427 -0.119691626
## [33,]  0.03212033  0.110200612  0.010542414
## [34,]  0.19727946 -0.251647006  0.108662141
## [35,]  0.28021160  0.267308737  0.446046034
## [36,] -0.10072084  0.184482698 -0.218438193
## [37,] -0.13983689 -0.041961789 -0.278205989
## [38,] -0.14932965  0.119788136 -0.044020488
## [39,] -0.05674154 -0.077198771  0.059482512
## [40,] -0.26426654 -0.132306623  0.389422296
##
## $eig
## [1] 2.253482e+00 1.685334e+00 1.216292e+00 9.641673e-01 8.572849e-01
## [6] 7.324546e-01 6.292752e-01 5.422719e-01 4.817173e-01 4.292873e-01
## [11] 4.193676e-01 3.532496e-01 2.906842e-01 2.261246e-01 2.224965e-01
## [16] 2.021990e-01 1.871534e-01 1.609909e-01 1.231629e-01 9.889673e-02
## [21] 8.369450e-02 6.774976e-02 3.825845e-02 1.834547e-02 1.944108e-03
## [26] 2.220446e-16 1.765081e-16 -1.524454e-02 -2.137711e-02 -3.284624e-02
## [31] -6.092746e-02 -8.394513e-02 -8.718989e-02 -1.072807e-01 -1.160433e-01
## [36] -1.225913e-01 -1.412258e-01 -1.472658e-01 -1.540401e-01 -1.929896e-01
##
## $x

```

```
## NULL
##
## $ac
## [1] 0
##
## $GOF
## [1] 0.3799222 0.4195960
```

```
plainvar1 <- round(Woolly.pcoa$eig[1] / sum(Woolly.pcoa$eig), 3) * 100
plainvar2 <- round(Woolly.pcoa$eig[2] / sum(Woolly.pcoa$eig), 3) * 100
plainvar3 <- round(Woolly.pcoa$eig[3] / sum(Woolly.pcoa$eig), 3) * 100
sum.EigA <- sum(plainvar1, plainvar2, plainvar3)
```

```
sum.EigA
```

```
## [1] 46.9
```

```
sites <- c(rep("HA", 10), rep("HB", 11), rep("L", 11), rep("M", 8))
```

```
#To do ordination, first define plot parameters
```

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

```
#Then, initiate plot
```

```
plot(Woolly.pcoa$points[,1], Woolly.pcoa$points[,2], ylim = c(-0.2, 0.7),
      xlab = paste("PCoA 1 (", plainvar1, "%)", sep = ""),
      ylab = paste("PCoA 2 (", plainvar2, "%)", 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)
```

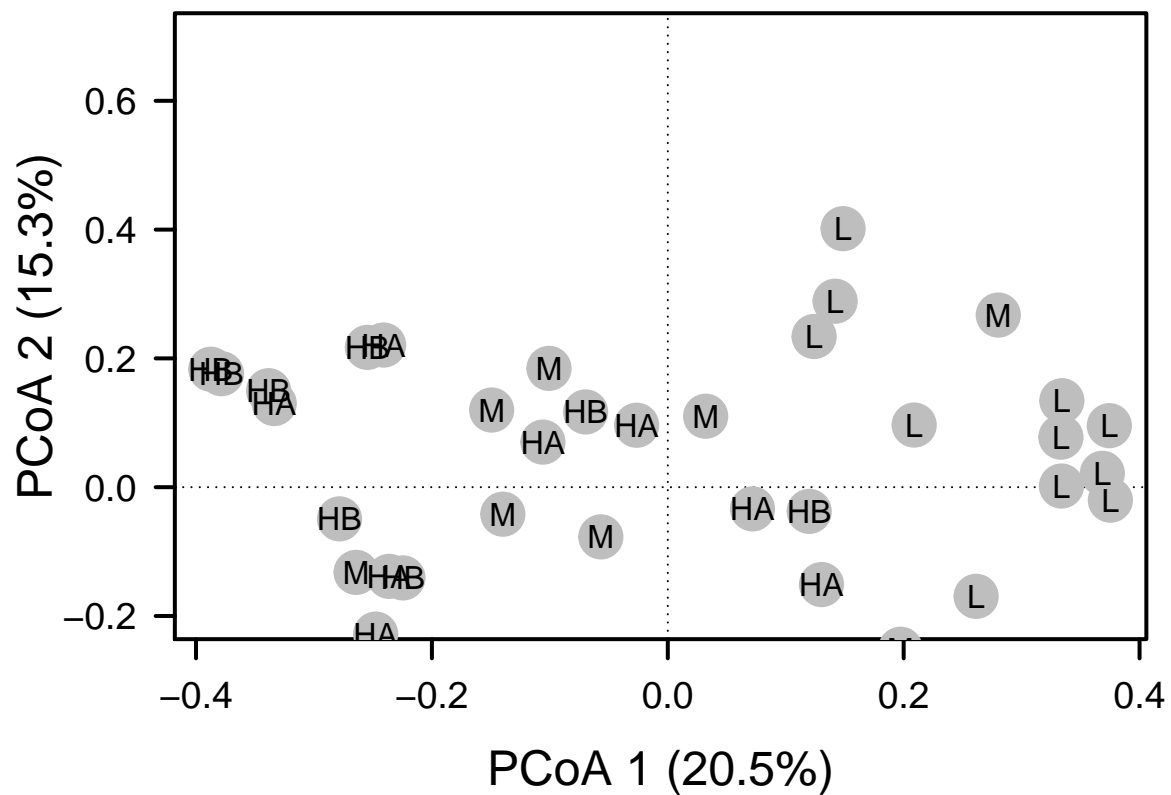
```
abline(h = 0, v = 0, lty = 3)
```

```
box(lwd = 2)
```

```
#Add points and labels
```

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

```
text(Woolly.pcoa$points[,1], Woolly.pcoa$points[,2],
     labels = sites, col = "black")
```



```
#create a factors vector
```

```
adonis(Woolly[,3:49] ~ sites, method = "bray", permutations = 999)
```

```
##
## Call:
## adonis(formula = Woolly[, 3:49] ~ sites, permutations = 999,      method = "bray")
##
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
##          Df SumsOfSqs MeanSqs F.Model    R2 Pr(>F)
## sites      3    1.9235  0.64118  2.5423 0.17482 0.001 ***
## Residuals 36    9.0794  0.25221      0.82518
## Total     39   11.0029      1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```