Assignment: Temporal Diversity

Savannah Bennett; Z620: Quantitative Biodiversity, Indiana University

15 February, 2017

OVERVIEW

In this Assignment, we extend our understanding of diversity from the spatial dimension to the temporal dimension.

After completing this exercise you will know how to:

1. wrangle a large dataset to visualize and analyze time series data
2. test hypotheses from experiments with temporal data
3. quantify temporal β-diversity and stability

Directions:

1. Change "Student Name" on line 3 (above) with your name.
2. Complete as much of the exercise as possible during class; what you do not complete in class will need to be done on your own outside of class.
3. Use the Handout as a guide; it contains a more complete description of data sets along with the proper scripting needed to carry out the exercise.
4. Be sure to **answer the questions** in this exercise document; they also correspond to the Handout. Space for your answer is provided in this document and indicated by the ">" character. If you need a second paragraph be sure to start the first line with ">".
5. Before you leave the classroom, **push** this file to your GitHub repo.
6. When you are done with the Assignment, **Knit** the text and code into a html file.
7. After Knitting, please submit the completed Assignment by creating a **pull request** via GitHub. Your pull request should include this file *temporal_assignment.Rmd* and the html output of `Knitr` (*temporal_assignment.html*).

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 "*/Week5-Temporal*" folder, and
4. load any packages you need to complete the assignment.

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

```
## [1] "C:/Users/Savannah/GitHub/QB2017_Bennett/Week5-Temporal"
```

```
setwd("C:/Users/Savannah/GitHub/QB2017_Bennett/Week5-Temporal")
```

2) LOADING DATA

Load dataset

In the R code chunk below, do the following:

1. load the `portal` dataset from in the "*/Week5/data*" folder, and
2. explore the structure of the dataset.

```
package.list <- c('vegan', 'tidyr', 'dplyr', 'codyn', 'ggplot2',
'cowplot', 'MullerPlot', 'RColorBrewer', 'reshape2', 'lubridate',
'TTR', 'xtable', 'multcomp', 'pander', 'png', 'grid', 'tseries', 'nlme', 'forecast', 'lsmeans')

for (package in package.list) {
  if (!require(package, character.only = TRUE, quietly = TRUE)) {
    install.packages(package,repos='http://cran.us.r-project.org')
    library(package, character.only = TRUE)
  }
}
```

```
## This is vegan 2.4-2
```

```
##
## 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
```

```
##
## Attaching package: 'cowplot'
```

```
## The following object is masked from 'package:ggplot2':
##
##     ggsave
```

```
##
## Attaching package: 'reshape2'
```

```
## The following object is masked from 'package:tidyr':
##
##     smiths
```

```
##
## Attaching package: 'lubridate'
```

```
## The following object is masked from 'package:base':
##
##     date
```

```
##
## Attaching package: 'MASS'
```

```
## The following object is masked from 'package:dplyr':
##
##     select
```

```
##
## Attaching package: 'TH.data'

## The following object is masked from 'package:MASS':
##
##     geyser

##
## Attaching package: 'nlme'

## The following object is masked from 'package:dplyr':
##
##     collapse

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

##
## Attaching package: 'timeDate'

## The following object is masked from 'package:xtable':
##
##     align

## This is forecast 7.3

##
## Attaching package: 'forecast'

## The following object is masked from 'package:nlme':
##
##     getResponse
```

```
portal <- read.table("data/combined.csv", sep = ",", header = TRUE)

str(portal)
```

```
## 'data.frame':    34786 obs. of  13 variables:
##  $ record_id      : int  1 72 224 266 349 363 435 506 588 661 ...
##  $ month          : int  7 8 9 10 11 11 12 1 2 3 ...
##  $ day            : int  16 19 13 16 12 12 10 8 18 11 ...
##  $ year           : int  1977 1977 1977 1977 1977 1977 1977 1978 1978 1978 ...
##  $ plot_id        : int  2 2 2 2 2 2 2 2 2 2 ...
##  $ species_id     : Factor w/ 48 levels "AB","AH","AS",..: 16 16 16 16 16 16 16 16 16 16 ...
##  $ sex            : Factor w/ 3 levels "","F","M": 3 3 1 1 1 1 1 1 3 1 ...
##  $ hindfoot_length: int  32 31 NA NA NA NA NA NA NA NA ...
##  $ weight         : int  NA NA NA NA NA NA NA NA 218 NA ...
##  $ genus          : Factor w/ 26 levels "Ammodramus","Ammospermophilus",..: 13 13 13 13 13 13 13 13
##  $ species        : Factor w/ 40 levels "albigula","audubonii",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ taxa           : Factor w/ 4 levels "Bird","Rabbit",..: 4 4 4 4 4 4 4 4 4 4 ...
##  $ plot_type      : Factor w/ 5 levels "Control","Long-term Krat Exclosure",..: 1 1 1 1 1 1 1 1 1
```

```
rodent <- portal[portal$taxa == "Rodent",]
str(rodent)
```

```
## 'data.frame':    34247 obs. of  13 variables:
```

```
##  $ record_id      : int  1 72 224 266 349 363 435 506 588 661 ...
##  $ month          : int  7 8 9 10 11 11 12 1 2 3 ...
##  $ day            : int  16 19 13 16 12 12 10 8 18 11 ...
##  $ year           : int  1977 1977 1977 1977 1977 1977 1977 1978 1978 1978 ...
##  $ plot_id        : int  2 2 2 2 2 2 2 2 2 2 ...
##  $ species_id     : Factor w/ 48 levels "AB","AH","AS",..: 16 16 16 16 16 16 16 16 16 16 ...
##  $ sex            : Factor w/ 3 levels "","F","M": 3 3 1 1 1 1 1 3 1 ...
##  $ hindfoot_length: int  32 31 NA NA NA NA NA NA NA NA ...
##  $ weight         : int  NA NA NA NA NA NA NA NA 218 NA ...
##  $ genus          : Factor w/ 26 levels "Ammodramus","Ammospermophilus",..: 13 13 13 13 13 13 13 13
##  $ species        : Factor w/ 40 levels "albigula","audubonii",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ taxa           : Factor w/ 4 levels "Bird","Rabbit",..: 4 4 4 4 4 4 4 4 4 4 ...
##  $ plot_type      : Factor w/ 5 levels "Control","Long-term Krat Exclosure",..: 1 1 1 1 1 1 1 1 1 1
```

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

a. How many plots are in `portal`?
b. How many rodent species are there in the `portal` dataset?

Answer 1a: There are twenty four plots in this dataset. ***Answer 1b***: There are forty species in the 'portal' dataset.

3) WRANGLING THE PORTAL DATASET

In the R code chunk below, do the following:

1. Create a site-by-species matrix for any year of your choosing.
2. Create a vector of plot_type for sites in the site-by-species matrix.
3. Analyze alpha diversity (e.g., Shannon/Simpson) across the sites for that year.
4. Create a PCoA ordination of your site-by-species matrix.
5. Using the hypothesis testing tools you learned in the beta-diversity module, test the hypothesis that species abundances across sites vary as a factor of treatment type (i.e., plot_type).

```
portal <- read.table("data/combined.csv", sep = ",", header = TRUE)

# Date vector
portal <- unite(portal,col =date , c(year, month, day),sep ="-",remove =FALSE)

# Taxon vector
portal <- unite(portal,col =taxon, c(genus, species),sep ="_",remove =FALSE)

#1. Making time by species matrix #
time.by.species <- group_by(portal, year, plot_id, plot_type) %>% count(taxon) %>% spread(key =taxon,val
                                                                                          # Filter by year
dplyr::filter(time.by.species, year ==1984)
```

```
## Source: local data frame [24 x 51]
## Groups: year, plot_id, plot_type [24]
##
##     year plot_id                plot_type Ammodramus_savannarum
##    <int>   <int>                    <fctr>                 <dbl>
## 1   1984       1         Spectab exclosure                     0
## 2   1984       2                   Control                     0
## 3   1984       3  Long-term Krat Exclosure                     0
## 4   1984       4                   Control                     0
## 5   1984       5          Rodent Exclosure                     0
```

```
## 6   1984       6 Short-term Krat Exclosure                     0
## 7   1984       7          Rodent Exclosure                     0
## 8   1984       8                 Control                       0
## 9   1984       9         Spectab exclosure                     0
## 10  1984      10          Rodent Exclosure                     0
## # ... with 14 more rows, and 47 more variables:
## #   Ammospermophilus_harrisi <dbl>, Amphispiza_bilineata <dbl>,
## #   Baiomys_taylori <dbl>, Calamospiza_melanocorys <dbl>,
## #   Callipepla_squamata <dbl>, Campylorhynchus_brunneicapillus <dbl>,
## #   Chaetodipus_baileyi <dbl>, Chaetodipus_intermedius <dbl>,
## #   Chaetodipus_penicillatus <dbl>, Chaetodipus_sp. <dbl>,
## #   Cnemidophorus_tigris <dbl>, Cnemidophorus_uniparens <dbl>,
## #   Crotalus_scutalatus <dbl>, Crotalus_viridis <dbl>,
## #   Dipodomys_merriami <dbl>, Dipodomys_ordii <dbl>, Dipodomys_sp. <dbl>,
## #   Dipodomys_spectabilis <dbl>, Lizard_sp. <dbl>, Neotoma_albigula <dbl>,
## #   Onychomys_leucogaster <dbl>, Onychomys_sp. <dbl>,
## #   Onychomys_torridus <dbl>, Perognathus_flavus <dbl>,
## #   Perognathus_hispidus <dbl>, Peromyscus_eremicus <dbl>,
## #   Peromyscus_leucopus <dbl>, Peromyscus_maniculatus <dbl>,
## #   Pipilo_chlorurus <dbl>, Pipilo_fuscus <dbl>, Pipilo_sp. <dbl>,
## #   Pooecetes_gramineus <dbl>, Reithrodontomys_fulvescens <dbl>,
## #   Reithrodontomys_megalotis <dbl>, Reithrodontomys_montanus <dbl>,
## #   Reithrodontomys_sp. <dbl>, Rodent_sp. <dbl>, Sceloporus_clarki <dbl>,
## #   Sceloporus_undulatus <dbl>, Sigmodon_fulviventer <dbl>,
## #   Sigmodon_hispidus <dbl>, Sigmodon_ochrognathus <dbl>,
## #   Sparrow_sp. <dbl>, Spermophilus_spilosoma <dbl>,
## #   Spermophilus_tereticaudus <dbl>, Sylvilagus_audubonii <dbl>,
## #   Zonotrichia_leucophrys <dbl>
```

```
#Data frame
time.by.species <- as.data.frame(time.by.species)

#vector for plot type
time.by.species <- group_by(portal, year, plot_id, plot_type) %>%
  count(taxon) %>% spread(key = taxon, value = n, fill = 0)
dplyr:: filter(time.by.species, year == 1984)
```

```
## Source: local data frame [24 x 51]
## Groups: year, plot_id, plot_type [24]
##
##     year plot_id                plot_type Ammodramus_savannarum
##    <int>   <int>                    <fctr>                 <dbl>
## 1   1984       1         Spectab exclosure                     0
## 2   1984       2                   Control                     0
## 3   1984       3  Long-term Krat Exclosure                     0
## 4   1984       4                   Control                     0
## 5   1984       5          Rodent Exclosure                     0
## 6   1984       6 Short-term Krat Exclosure                     0
## 7   1984       7          Rodent Exclosure                     0
## 8   1984       8                   Control                     0
## 9   1984       9         Spectab exclosure                     0
## 10  1984      10          Rodent Exclosure                     0
## # ... with 14 more rows, and 47 more variables:
## #   Ammospermophilus_harrisi <dbl>, Amphispiza_bilineata <dbl>,
## #   Baiomys_taylori <dbl>, Calamospiza_melanocorys <dbl>,
```

```
## #   Callipepla_squamata <dbl>, Campylorhynchus_brunneicapillus <dbl>,
## #   Chaetodipus_baileyi <dbl>, Chaetodipus_intermedius <dbl>,
## #   Chaetodipus_penicillatus <dbl>, Chaetodipus_sp. <dbl>,
## #   Cnemidophorus_tigris <dbl>, Cnemidophorus_uniparens <dbl>,
## #   Crotalus_scutalatus <dbl>, Crotalus_viridis <dbl>,
## #   Dipodomys_merriami <dbl>, Dipodomys_ordii <dbl>, Dipodomys_sp. <dbl>,
## #   Dipodomys_spectabilis <dbl>, Lizard_sp. <dbl>, Neotoma_albigula <dbl>,
## #   Onychomys_leucogaster <dbl>, Onychomys_sp. <dbl>,
## #   Onychomys_torridus <dbl>, Perognathus_flavus <dbl>,
## #   Perognathus_hispidus <dbl>, Peromyscus_eremicus <dbl>,
## #   Peromyscus_leucopus <dbl>, Peromyscus_maniculatus <dbl>,
## #   Pipilo_chlorurus <dbl>, Pipilo_fuscus <dbl>, Pipilo_sp. <dbl>,
## #   Pooecetes_gramineus <dbl>, Reithrodontomys_fulvescens <dbl>,
## #   Reithrodontomys_megalotis <dbl>, Reithrodontomys_montanus <dbl>,
## #   Reithrodontomys_sp. <dbl>, Rodent_sp. <dbl>, Sceloporus_clarki <dbl>,
## #   Sceloporus_undulatus <dbl>, Sigmodon_fulviventer <dbl>,
## #   Sigmodon_hispidus <dbl>, Sigmodon_ochrognathus <dbl>,
## #   Sparrow_sp. <dbl>, Spermophilus_spilosoma <dbl>,
## #   Spermophilus_tereticaudus <dbl>, Sylvilagus_audubonii <dbl>,
## #   Zonotrichia_leucophrys <dbl>
```

```
#2. Site by species for 1984
sitebyspecies1984 <- subset(time.by.species, year == 1984)
plot.type <- sitebyspecies1984$plot_type
sitebyspecies1984 <- sitebyspecies1984[ ,-c(1:3)]

#3. Alpha diversity
simpsons <- diversity(sitebyspecies1984, "simp")
shannon <- diversity(sitebyspecies1984, "shannon")
str(simpsons)
```

```
##  num [1:24] 0.597 0.818 0.838 0.459 0.599 ...
```

```
str(shannon)
```

```
##  num [1:24] 1.365 2.016 2.105 0.886 1.355 ...
```

```
max(simpsons)
```

```
## [1] 0.8381392
```

```
min(simpsons)
```

```
## [1] 0.4489796
```

```
max(shannon)
```

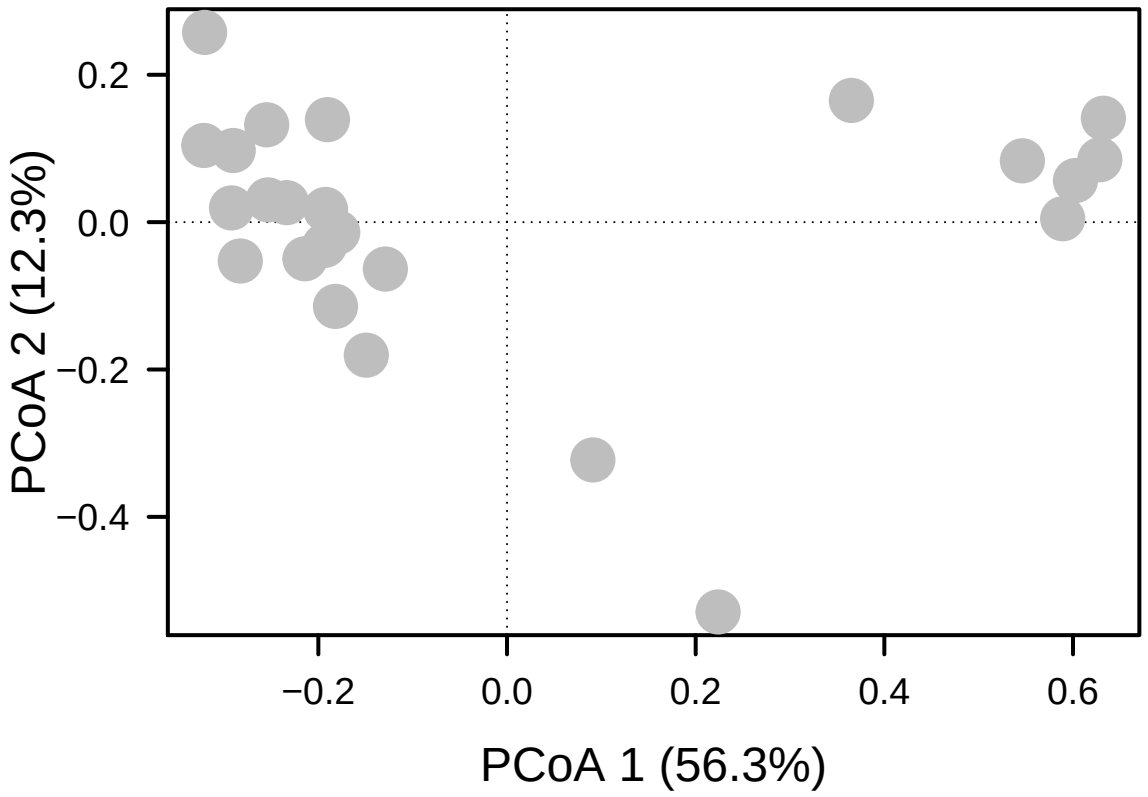
```
## [1] 2.104875
```

```
min(shannon)
```

```
## [1] 0.6615632
```

```
#4. PCoA ordination

portal.db <- vegdist(sitebyspecies1984,method ="bray",upper =TRUE,diag =TRUE)
pcoa <- cmdscale(portal.db,eig =TRUE,k =3)

# Interpreting PCoA output #
```

```
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)

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

# Initiate Plot
plot(pcoa$points[ ,1], pcoa$points[ ,2],
xlab = paste("PCoA 1 (", explainvar1,"%)",sep =""),
ylab = paste("PCoA 2 (", 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 & Labels
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))

#Relative Abundances
REL <-sitebyspecies1984
for(i in 1:nrow(sitebyspecies1984)){
REL[i, ] =sitebyspecies1984[i, ] / sum(sitebyspecies1984[i, ])
}

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




```
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: 'ade4'
## The following object is masked from 'package:vegan':
##
##     cca
##
## Attaching package: 'gplots'
## The following object is masked from 'package:stats':
##
##     lowess
## BiodiversityR 2.8-0: Use command BiodiversityRGUI() to launch the Graphical User Interface and to lea
```

```
#Beta Diversity- PERMANOVA

adonis(sitebyspecies1984 ~plot.type,method ="bray",permutations =999)
```

```
##
```

```
## Call:
## adonis(formula = sitebyspecies1984 ~ plot.type, permutations = 999,      method = "bray")
##
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
##           Df SumsOfSqs MeanSqs F.Model     R2 Pr(>F)
## plot.type  4    2.2725 0.56811  3.7564 0.4416  0.001 ***
## Residuals 19    2.8735 0.15124         0.5584
## Total     23    5.1460                 1.0000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
indval <- multipatt(sitebyspecies1984,cluster =plot.type,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: 48
##  Selected number of species: 5
##  Number of species associated to 1 group: 0
##  Number of species associated to 2 groups: 1
##  Number of species associated to 3 groups: 3
##  Number of species associated to 4 groups: 1
##
##  List of species associated to each combination:
##
##  Group Long-term Krat Exclosure+Rodent Exclosure  #sps.  1
##                             stat p.value
## Reithrodontomys_megalotis 0.905    0.01 **
##
##  Group Control+Short-term Krat Exclosure+Spectab exclosure  #sps.  3
##                           stat p.value
## Dipodomys_spectabilis  0.958   0.001 ***
## Dipodomys_ordii        0.923   0.020 *
## Spermophilus_spilosoma 0.804   0.031 *
##
##  Group Control+Rodent Exclosure+Short-term Krat Exclosure+Spectab exclosure  #sps.  1
##                      stat p.value
## Dipodomys_merriami 0.94    0.05 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
portal1 <- decostand(sitebyspecies1984,method ="total")
phi <- multipatt(portal1,cluster =plot.type,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: 48
##  Selected number of species: 3
##  Number of species associated to 1 group: 0
##  Number of species associated to 2 groups: 1
##  Number of species associated to 3 groups: 2
##  Number of species associated to 4 groups: 0
##
##  List of species associated to each combination:
##
##  Group Long-term Krat Exclosure+Rodent Exclosure  #sps.  1
##                             stat p.value
## Reithrodontomys_megalotis  0.7    0.02 *
##
##  Group Control+Short-term Krat Exclosure+Spectab exclosure  #sps.  2
##                          stat p.value
## Dipodomys_merriami    0.741   0.003 **
## Dipodomys_spectabilis 0.638   0.038 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Question 2: Describe how different biodiversity estimates vary among sites.

a. Does diversity vary among sites? Does this correspond to treatment type?
b. Is treatment type a significant predictor of site dissimilarity?

> ***Answer 2a***: The alpha diversity values seemed to vary by site. Simpson's index values ranged from 0.44 to 0.83, and Shannon values ranged from 0.66 to 2.10. To examine beta diversity, I performed a permanova, which suggested that the treatments had a significant impact on species diversity/distribution at the sites. ***Answer 2b***: Based on the permanova, treatment type is a significant predictor of site dissimilarity.

4) TIME SERIES ANALYSIS

In the R code chunk below, do the following:

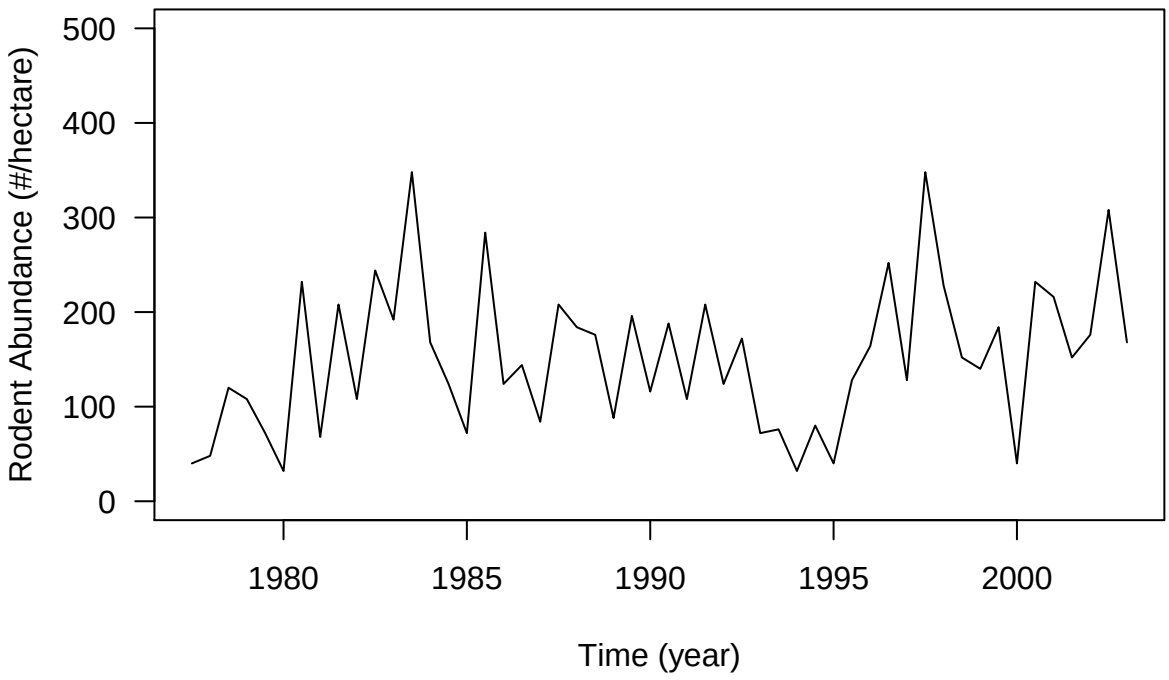
1. Create a time-by-species matrix that includes year, month, and plot_id for a site other than plot_id 2.
2. Examine per-hectare rodent abundance using simple moving average smoothing.
3. Test whether your data meets the assumption of stationarity.
4. If it does not meet this asumption, explore wasy to make your data stationary.
5. Examine and plot time lags using the partial autocorrelation function (PACF) and autocorrelation function (ACR).
6. Use the tools outlined in the Handout to create an ARMA model.

```
# 1. Time-by-species matrix with year, month, and plot_id
time.by.spec.2<- filter(portal, taxa =="Rodent") %>%
group_by(year, month, plot_id) %>% count(taxon)

# Create seasonality variable using month number (6 = June;
time.by.spec.2$season <- NA
time.by.spec.2$season <- time.by.spec.2$month %in% c(6:10)
```

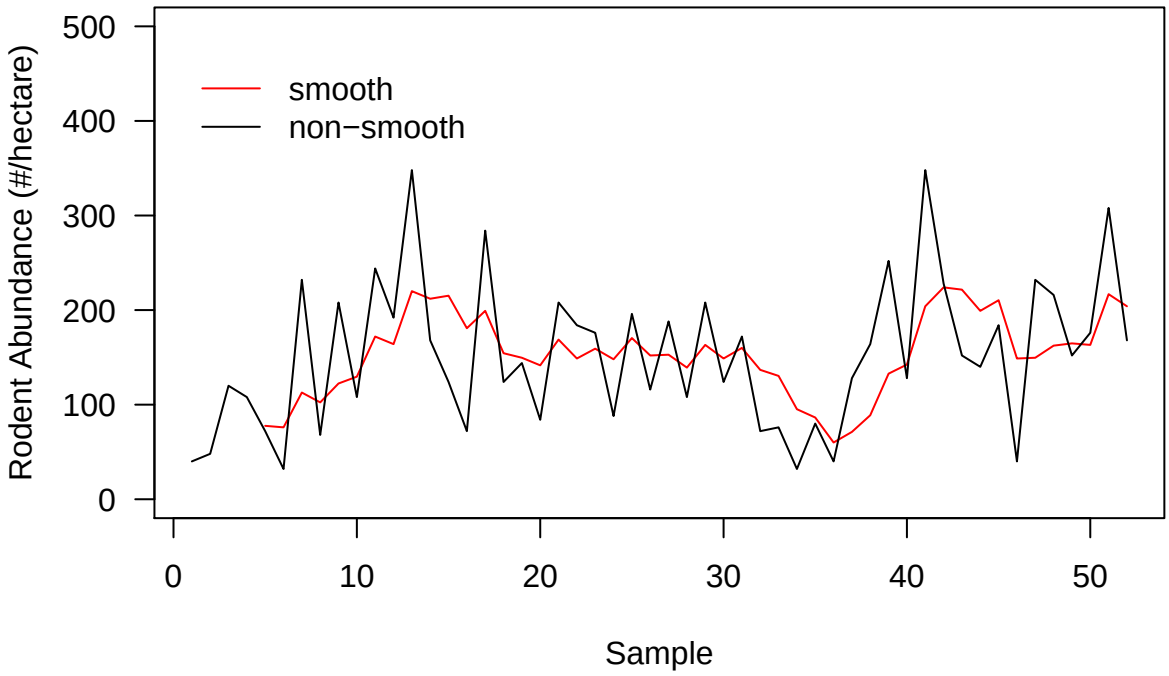
```r
# Rainy seasons are June - October
time.by.spec.2$season <- ifelse(time.by.spec.2$season == TRUE, "rain", "norain")

# Group the data by year and season
group_by(time.by.spec.2, year, season)
```

```
## Source: local data frame [16,391 x 6]
## Groups: year, season [52]
##
##     year month plot_id                   taxon     n season
##    <int> <int>   <int>                   <chr> <int>  <chr>
## 1   1977     7       1       Dipodomys_merriami     2   rain
## 2   1977     7       1       Perognathus_flavus     1   rain
## 3   1977     7       2 Chaetodipus_penicillatus     1   rain
## 4   1977     7       2       Dipodomys_merriami     1   rain
## 5   1977     7       2         Neotoma_albigula     1   rain
## 6   1977     7       2      Peromyscus_eremicus     1   rain
## 7   1977     7       3       Dipodomys_merriami     2   rain
## 8   1977     7       3    Dipodomys_spectabilis     1   rain
## 9   1977     7       3         Neotoma_albigula     1   rain
## 10  1977     7       4       Dipodomys_merriami     1   rain
## # ... with 16,381 more rows
```

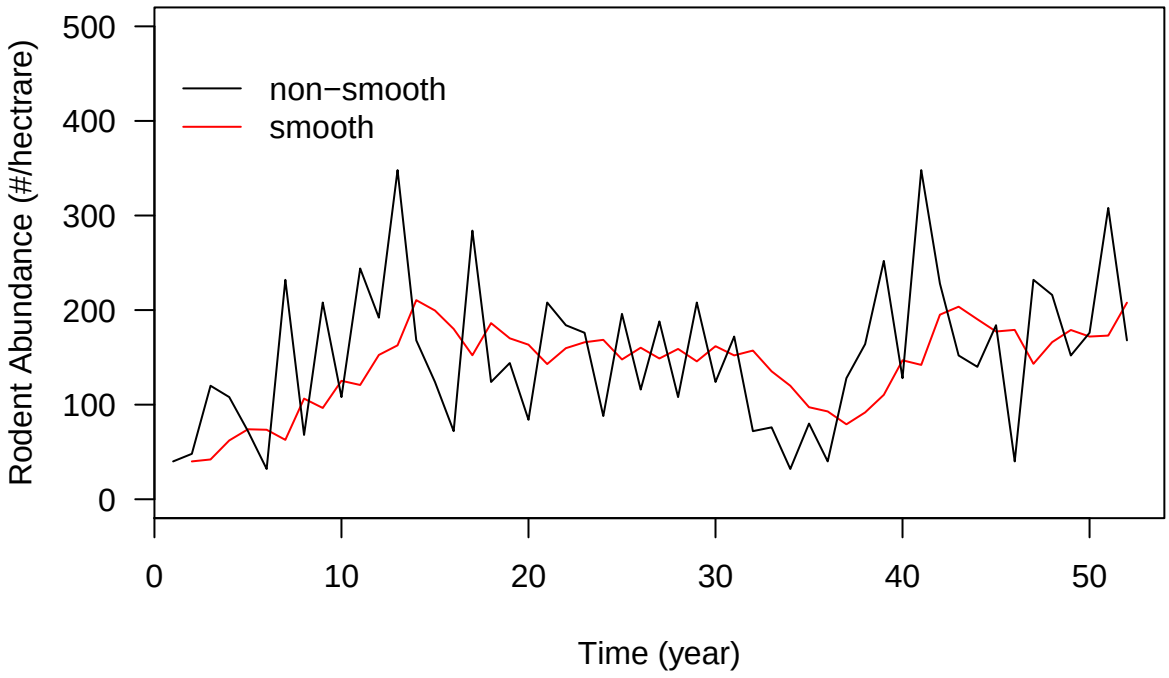
```r
#Filter
abund <- filter(time.by.spec.2, plot_id == 1) %>%
  group_by(year, season) %>%
  count(wt = n)

abund$nn <- abund$nn * 4

abund.ts <- ts(abund$nn, frequency = 2, start = c(1977, 2))

plot.ts(abund.ts, type = "l", ylab = "Rodent Abundance (#/hectare)",
        xlab = "Time (year)", las = 1, ylim = c(0, 500))
```




```
#2. Simple moving average smoothing
abund.sm <- SMA(abund$nn, n = 5)

plot(abund.sm, type = "l", col = "red", ylab = "Rodent Abundance (#/hectare)",
     xlab = "Sample", las = 1, ylim = c(0, 500))

lines(abund$nn, col = "black")

legend(0, 475, col = c("red", "black"), lty = c(1,1),
       c("smooth", "non-smooth"), bty = "n", cex = 1)
```




```
abund.hw <- HoltWinters(abund$nn, beta = FALSE, gamma = FALSE)
# abund.hw$fitted

plot(abund.hw, xlab = "Time (year)", ylim = c(0, 500),
     ylab = "Rodent Abundance (#/hectrare)", las = 1, main = NA)

legend(0, 475, col = c("black", "red"), lty = c(1,1),
       c("non-smooth", "smooth"), bty = "n", cex = 1)
```



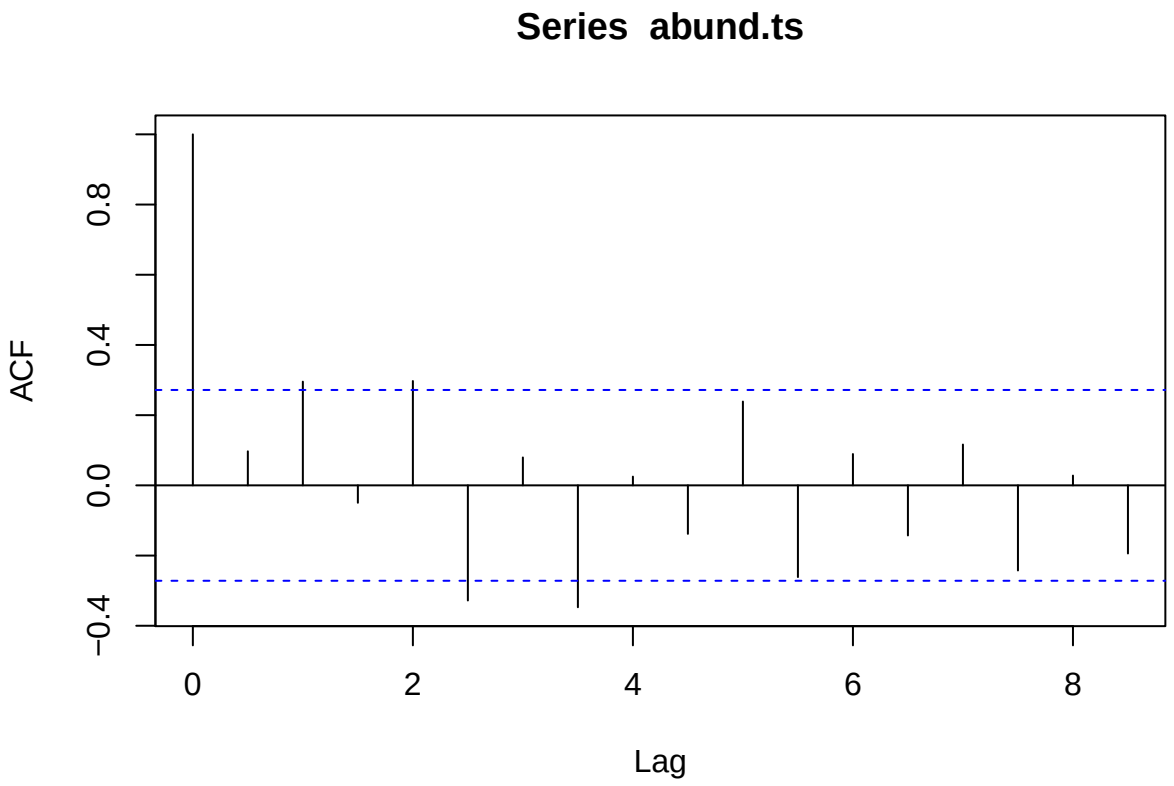

```
#3. Stationarity
adf.raw <- adf.test(abund.ts, alternative = "stationary")
adf.raw$p.value
```

```
## [1] 0.4469051
```

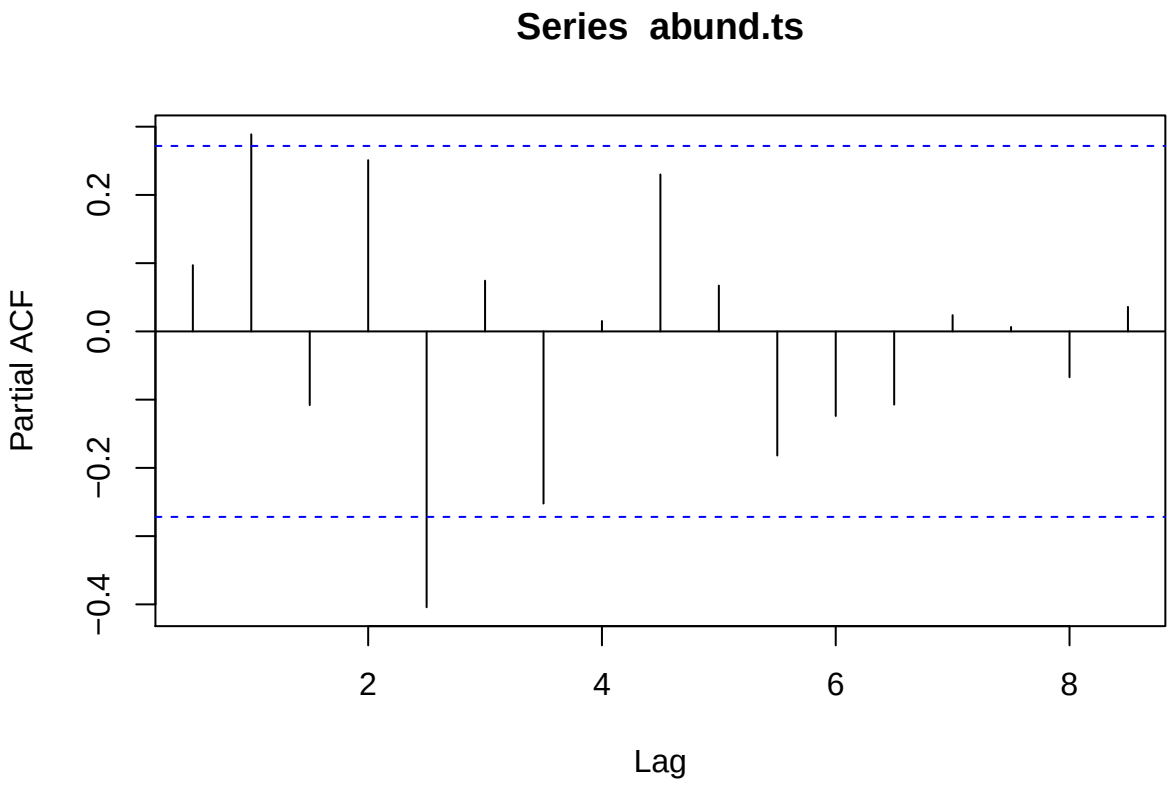
```
#4. Differencing time series function because does not meet stationarity
abund.ts.diff <- diff(abund.ts)
adf.diff <- adf.test(abund.ts.diff, alternative = "stationary")
adf.diff$p.value
```

```
## [1] 0.04273821
```

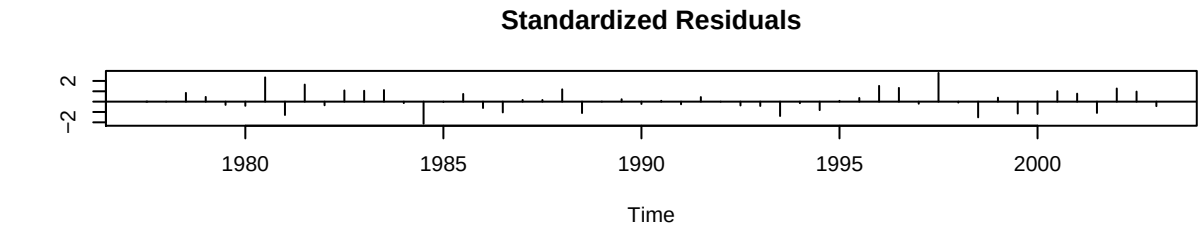
```
#5. PACF and ACR

#ACR
acf(abund.ts)
```

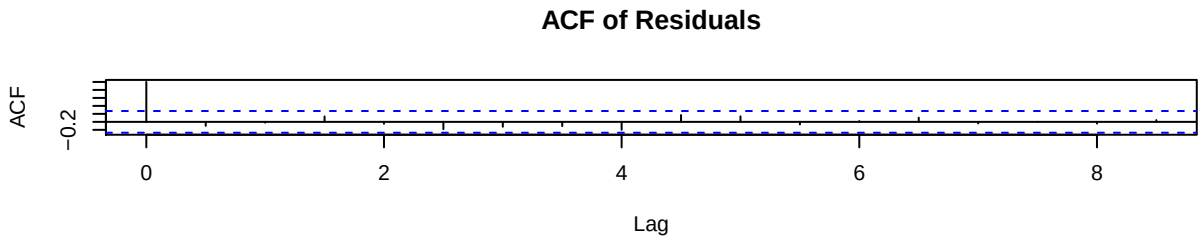


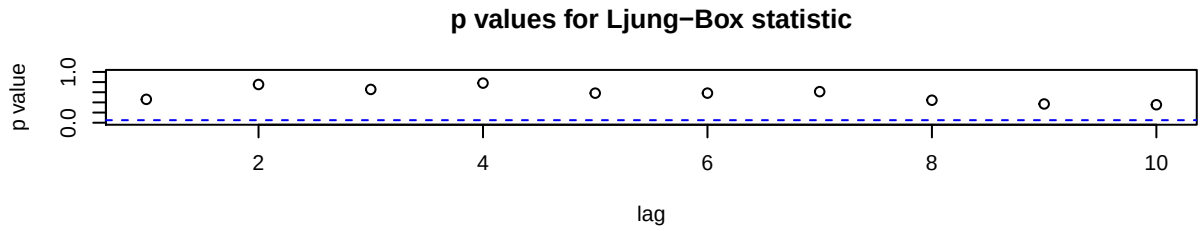

```
#PACF
pacf(abund.ts)
```



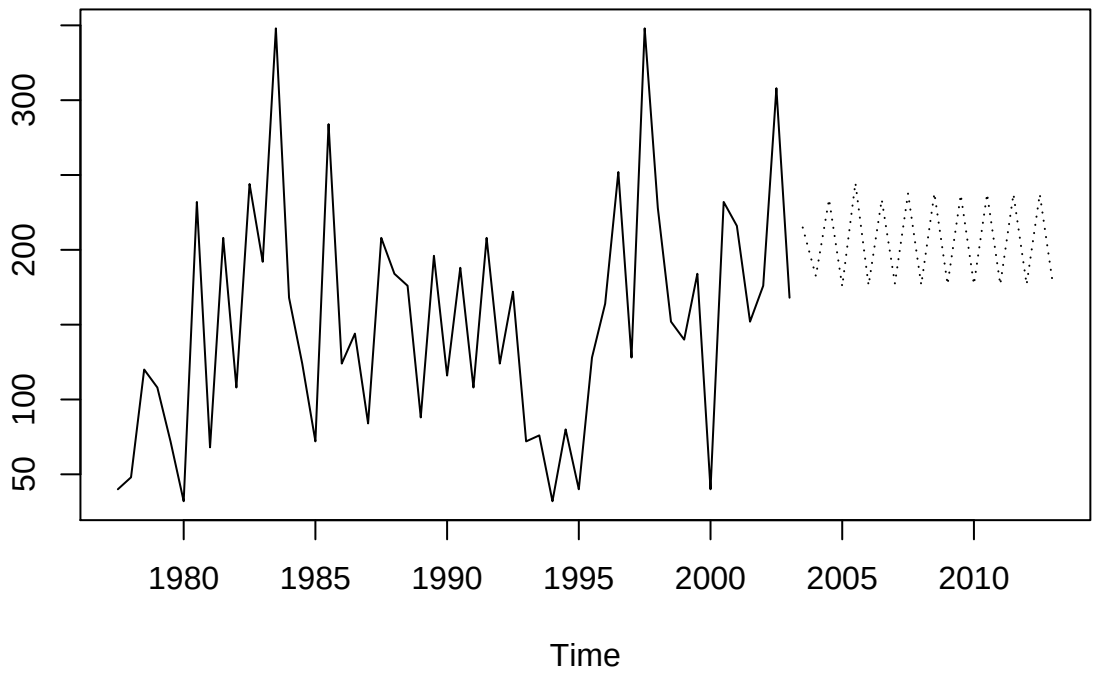

```
#6. ARMA Model

abund.arm <- auto.arima(abund.ts)

abund.arm <- arima((abund.ts), c(0, 0, 1), seasonal = list(order = c(2, 1, 0),
period = 2), include.mean = TRUE)

tsdiag(abund.arm)
```








```
pred.arm <- predict(abund.arm, n.ahead = 20)

ts.plot(abund.ts, pred.arm$pred, lty = c(1,3))
```



Question 3: Describe the results from your time series analysis.

a. Does your data meet the assumption of stationarity? If not, what does this violation imply?
b. What does the ACF function do and how does it relate to the ARMA model? How does this differ from the autocorrelation function (ACF)?
c. What results can you conclude from your full ARMA model along with other methods outlined in the time series setcion of the Handout?

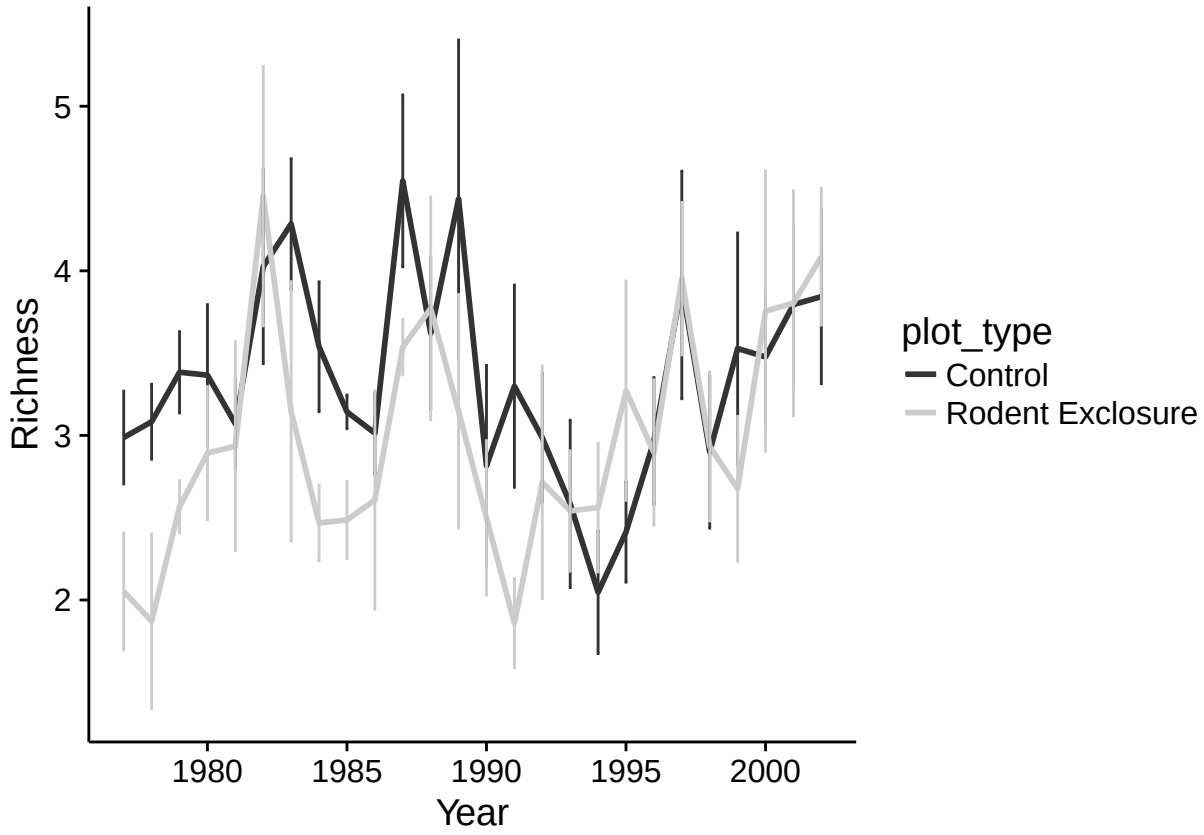
Answer 3a: The data does not meet the assumption of stationarity (p=0.44), which implies that the mean, variance, and correlation structure change over time. ***Answer 3b***: The ACF function examimes the coelation between the lagged intervals in a particular time series. It is used to visualize the data and parameterize AMRA models. PACF, on the other hand, shows lags that can be addressed with AR, or the autoregressive part of the ARMA model. ***Answer 3c***: Based on these plots, rodent abundance generally increased in the mid 1980s, and exhibited relatively sharp declines around 1995 and 2000. Abundance is predicted to stabilize or exhibit fluctuations to a lesser extent after 2005.

5) REPEATED MEASURES ANALYSIS OF VARIANCE (RM-ANOVA)

In the R code chunk below, do the following:

1. Create an appropriate data frame for RM-ANOVA (e.g., yearly species abundance values within plots).
2. Calculate the inverse of Simpson's diversity for each year, and plot it as a function of year for the Control and Rodent Exclosure plots.
3. Perform an RM-ANOVA and construct a F-test using the AR(1), compound symmetery, and unstructured covariance structures.

```r
#1. Creating data frame for RM-ANOVA
# Constructing time-vy-species matrix
time.by.species <- group_by(portal, year, plot_id, plot_type) %>% count(taxon) %>% spread(key =taxon,va

# Calculate observed richness from time.by.species matrix
# richness <- as.data.frame(rowSums(time.by.species[,-c(1:3)]))
# Calulate inverse Simpsons diversity from time.by.species matrix
richness <- as.data.frame(diversity(time.by.species[,-c(1:3)],index="invsimpson"))

# Create data frame with experimental design and richness data #
rich.all <- data.frame(time.by.species[,1:3,], richness)

# Rename Column
names(rich.all)[4] <-"richness"

# Pull out two of the five Portal treatments #
rich.treat <-rich.all[which(rich.all$plot_type =="Control"|rich.all$plot_type =="Rodent Exclosure"), ]

#Plot Data

rich.treat.plot <- group_by(rich.treat, plot_type, year) %>%
  summarise(
    mean = mean(richness), # avg. richness per group
    sd = sd(richness), # stand. dev. per group
    n = n(), # num. obs. per group
    sem = sd/sqrt(n)) # calc. std. err. mean.

rich.plot <- ggplot(rich.treat.plot, aes(x = year, y = mean, color = plot_type)) +
  geom_line(size = 1, show.legend = T) +
  geom_errorbar(aes(ymin = mean - sem, ymax = mean + sem), width = .1) +
  xlim(1977, 2002) +
  xlab("Year") +
  ylab("Richness")+
  scale_color_grey()

plot(rich.plot)
```




```
rich.rm <- lme(richness ~plot_type*year,random =~1|plot_id,
correlation = corAR1(form =~1|plot_id),
data =rich.treat)
summary(rich.rm)
```

```
## Linear mixed-effects model fit by REML
##  Data: rich.treat
##        AIC      BIC    logLik
##   1153.217 1180.322 -569.6087
##
## Random effects:
##  Formula: ~1 | plot_id
##         (Intercept) Residual
## StdDev:   0.6469478 1.250206
##
## Correlation Structure: AR(1)
##  Formula: ~1 | plot_id
##  Parameter estimate(s):
##       Phi
## 0.4250398
## Fixed effects: richness ~ plot_type * year
##                                   Value Std.Error  DF    t-value p-value
## (Intercept)                     1.36695  33.32814 343  0.0410150  0.9673
## plot_typeRodent Exclosure     -73.37575  52.01883  12 -1.4105613  0.1838
## year                            0.00100   0.01675 343  0.0595739  0.9525
## plot_typeRodent Exclosure:year  0.03670   0.02614 343  1.4039132  0.1612
```

```
##  Correlation:
##                                (Intr) plt_RE year
## plot_typeRodent Exclosure      -0.641
## year                           -1.000  0.641
## plot_typeRodent Exclosure:year  0.641 -1.000 -0.641
##
## Standardized Within-Group Residuals:
##         Min          Q1         Med          Q3         Max
## -1.93162654 -0.70434296 -0.08293334  0.45856267  5.06139168
##
## Number of Observations: 359
## Number of Groups: 14
```

```
rich.cmp <- lme(richness ~plot_type*year,random =~1|plot_id,
correlation = corCompSymm(form =~1|plot_id),
data =rich.treat)
summary(rich.cmp)
```

```
## Linear mixed-effects model fit by REML
##  Data: rich.treat
##        AIC      BIC    logLik
##   1215.368 1242.473 -600.6842
##
## Random effects:
##  Formula: ~1 | plot_id
##         (Intercept) Residual
## StdDev:   0.7104222 1.214275
##
## Correlation Structure: Compound symmetry
##  Formula: ~1 | plot_id
##  Parameter estimate(s):
##          Rho
## 6.672013e-18
## Fixed effects: richness ~ plot_type * year
##                                    Value Std.Error  DF    t-value p-value
## (Intercept)                     13.26069  22.33564 343  0.5937008  0.5531
## plot_typeRodent Exclosure      -68.60423  34.97126  12 -1.9617316  0.0734
## year                            -0.00498   0.01123 343 -0.4438008  0.6575
## plot_typeRodent Exclosure:year   0.03431   0.01758 343  1.9518712  0.0518
##  Correlation:
##                                (Intr) plt_RE year
## plot_typeRodent Exclosure      -0.639
## year                           -1.000  0.639
## plot_typeRodent Exclosure:year  0.639 -1.000 -0.639
##
## Standardized Within-Group Residuals:
##        Min         Q1        Med         Q3        Max
## -2.0346517 -0.6823806 -0.1001395  0.4790022  5.1406695
##
## Number of Observations: 359
## Number of Groups: 14
```

```
#rich.unstr <- lme(richness ~ plot_type*year, random = ~ 1 | plot_id,correlation =
  #corSymm(form = ~1 | plot_id),
  #data = rich.treat)
```

```
#summary(rich.unstr)

# Obtain F-test
anova(rich.rm)
```

```
##                numDF denDF   F-value p-value
## (Intercept)        1   343 256.22009  <.0001
## plot_type          1    12   0.72274  0.4119
## year               1   343   1.56060  0.2124
## plot_type:year     1   343   1.97097  0.1612
```

```
anova(rich.cmp)
```

```
##                numDF denDF   F-value p-value
## (Intercept)        1   343 255.02615  <.0001
## plot_type          1    12   0.73483  0.4081
## year               1   343   1.08893  0.2974
## plot_type:year     1   343   3.80980  0.0518
```

```
# Make cleaner ANOVA table #
set.caption("RMANOVA for Portal")
pander(anova(rich.rm))
```

Table 1: RMANOVA for Portal

	numDF	denDF	F-value	p-value
(Intercept)	1	343	256.2	0
plot_type	1	12	0.7227	0.4119
year	1	343	1.561	0.2124
plot_type:year	1	343	1.971	0.1612

```
pander(anova(rich.cmp))
```

	numDF	denDF	F-value	p-value
(Intercept)	1	343	255	0
plot_type	1	12	0.7348	0.4081
year	1	343	1.089	0.2974
plot_type:year	1	343	3.81	0.05177

```
# Use `lsmeans` package for time-corrected marginal means #
lsmeans(rich.rm, ~plot_type)
```

```
## NOTE: Results may be misleading due to involvement in interactions
```

```
##  plot_type          lsmean        SE df lower.CL upper.CL
##  Control          3.352478 0.2644674 13 2.781131 3.923825
##  Rodent Exclosure 3.002943 0.3066216 12 2.334872 3.671014
##
## Confidence level used: 0.95
```

```
lsmeans(rich.cmp, ~plot_type)
```

```
## NOTE: Results may be misleading due to involvement in interactions
```

```
##  plot_type          lsmean        SE df lower.CL upper.CL
##  Control          3.348317 0.2649103 13 2.776013 3.920621
##  Rodent Exclosure 2.997377 0.3064250 12 2.329735 3.665020
##
## Confidence level used: 0.95
```

Question 4: Describe the results from your RM-ANOVA.

a. In your own words describe what a RM-ANOVA test is doing
b. Is there a noticeable trend in the inverse of Simpson's diversity over time?
c. What does the result of your F-test tell you?
d. Of the three RM-ANOVA models with different covariance structures, which one is best? How does this affect the interpretation of your data?

Answer 4a: The RM-ANOVA is used to determine the effect of time, treatment, and time and treatment. It is useful when data is collected from the same location, site, ect. because it accounts for non-independence. In our dataset, it is determining whether richness is affected by time, exposure to rodents, or time x rodent exposure. ***Answer 4b***: the F-test indicates that the plot type, or treatment, has a significant effect on richness (p=0.00435). ***Answer 4c***: The autoregressive covariance structure is the best, and this suggests that the sum of the abundance values from past dates will influence abundance values in future dates.

6) TEMPORAL BETA DIVERSITY

Turnover

In the R code chunk below, do the following:

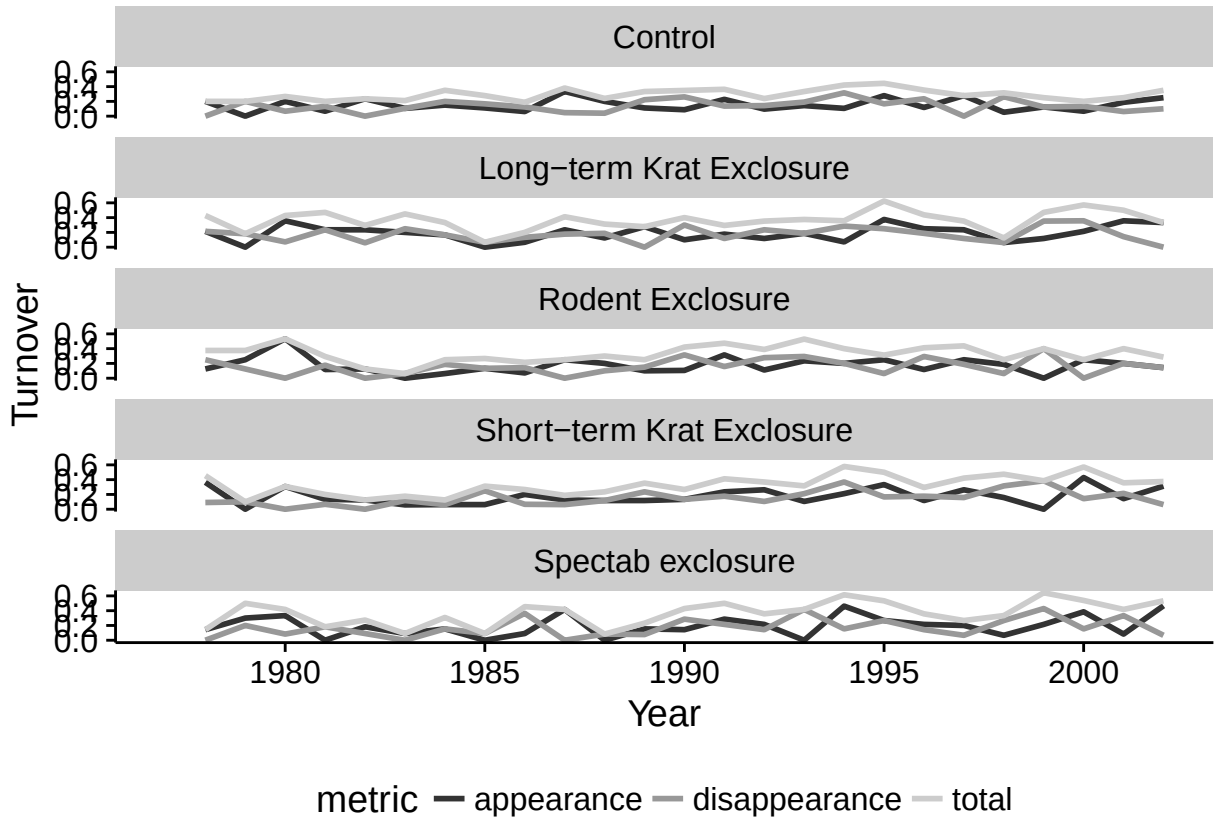
1. Calculate species abundances for each taxonomic group (the `taxa` column).
2. Calculate total turnover and turnover due to the gain/loss of species for each group.
3. Visualize turnover within each group

```
# First, calculate the species abundances from each site over time
portal.species.abunds <- group_by(portal, year, plot_type) %>% count(taxon)

# Calculate total turnover
portal.total <- turnover(df = portal.species.abunds,
  time.var = "year",
  species.var = "taxon",
  abundance.var = "n",
  replicate.var = "plot_type",
  metric = "total")

# Calculate species gained
portal.appearance <- turnover(df = portal.species.abunds,
  time.var = "year",
  species.var = "taxon",
  abundance.var = "n",
  replicate.var = "plot_type",
  metric = "appearance")

# Calculate species lost
portal.disappearance <- turnover(df = portal.species.abunds,
  time.var = "year",
  species.var = "taxon",
```

```
  abundance.var = "n",
  replicate.var = "plot_type",
  metric = "disappearance")

# Use `join()` from `dplyr` to join the columns by shared year & plot type columns
portal.turnover <- full_join(portal.total, portal.disappearance) %>%
  full_join(portal.appearance)
```

```
## Joining, by = c("year", "plot_type")
## Joining, by = c("year", "plot_type")
```

```
# Use `gather()` from `tidyr` to convert back to long-form
portal.turnover <- gather(portal.turnover, key = metric, value = turnover,
  total, appearance, disappearance)

turn.plot <- ggplot(
  portal.turnover, aes(x = year, y = turnover, color = metric)) +
  geom_line(size = 1, show.legend = T) +
  facet_wrap(~plot_type, ncol = 1) +
  xlim(1977, 2002) +
  xlab("Year") +
  ylab("Turnover") +
  theme(legend.position = "bottom") +
  scale_color_grey()

plot(turn.plot)
```



```r
max(portal.total[,1])
```

```
## [1] 0.6428571
```

```r
min(portal.total[,1])
```

```
## [1] 0.0625
```

Question 5:

a. How does temporal turnover relate to spatial turnover?
b. Which taxonomic group appears to be the most variable? Which group appears to be the least variable?

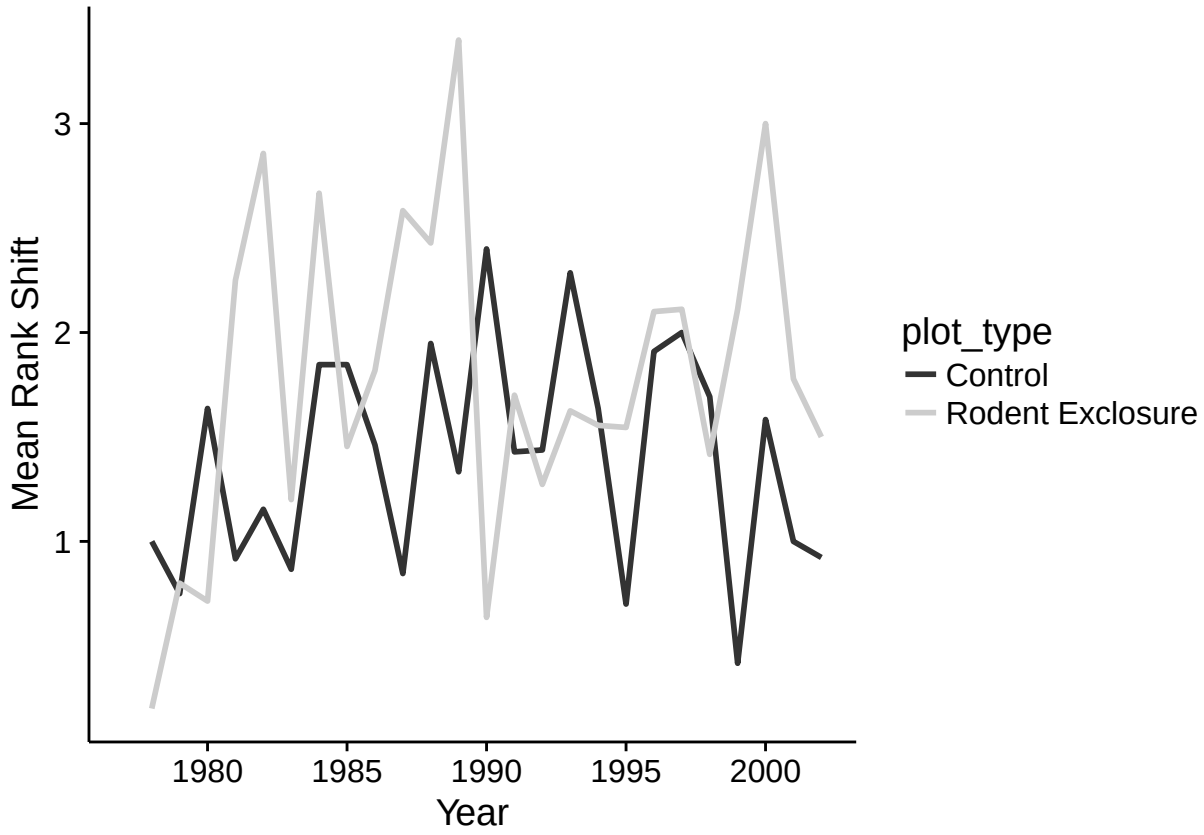
> 5a. Temporal turnover refers to changes in species composition over time in a particular area, while spatial turnover is the gains and losses of species from one place to another. Spatial turnover occurs when organisms move from one location to another, which can lead to temporal turnover because it causes species composition in one area to change over time. 5b. Spectab seems to be the most variable (highest turnover value of 0.642), whereas the rodent exclosure group has the lowest turnover (lowest turnover value of 0.0625).

Mean Rank Shift

In the code chunk below, do the following:

1. Choose two plot_types or two plot_ids and compare the mean rank shift (MRS) between them.
2. Plot MRS for each through time.

```r
# Pull out the two treatments we analyzed earlier
portal.abunds.cont.rodent <- filter(portal.species.abunds,
  plot_type == "Control" | plot_type == "Rodent Exclosure")

# Calculate MRS
portal.rankshift <- rank_shift(
  df = as.data.frame(portal.abunds.cont.rodent),
  time.var = "year",
  species.var = "taxon",
  abundance.var = "n",
  replicate.var = "plot_type")

# Replace the year range with a single value to plot
portal.rankshift$year <- as.numeric(substr(portal.rankshift$year_pair, 6, 9))

# Create ggplot
rankshift.plot <- ggplot(portal.rankshift, aes(x = year, y = MRS, color = plot_type)) +
  geom_line(size = 1) +
  xlim(1977, 2002) +
  xlab("Year") +
  ylab("Mean Rank Shift") +
  scale_color_grey()

plot(rankshift.plot)
```




```
# Does one plot type show higher or lower MRS, on average?
group_by(portal.rankshift, plot_type) %>%
  summarise(
    mean = mean(MRS),
    cv = sd(MRS)/mean)
```

```
## # A tibble: 2 × 3
##          plot_type     mean        cv
##              <chr>    <dbl>     <dbl>
## 1          Control 1.400675 0.3759483
## 2 Rodent Exclosure 1.788980 0.4361809
```

Question 6:

a. What does a change in the rank shift tell you about the community?
b. Interpret the analysis and figure you just made.

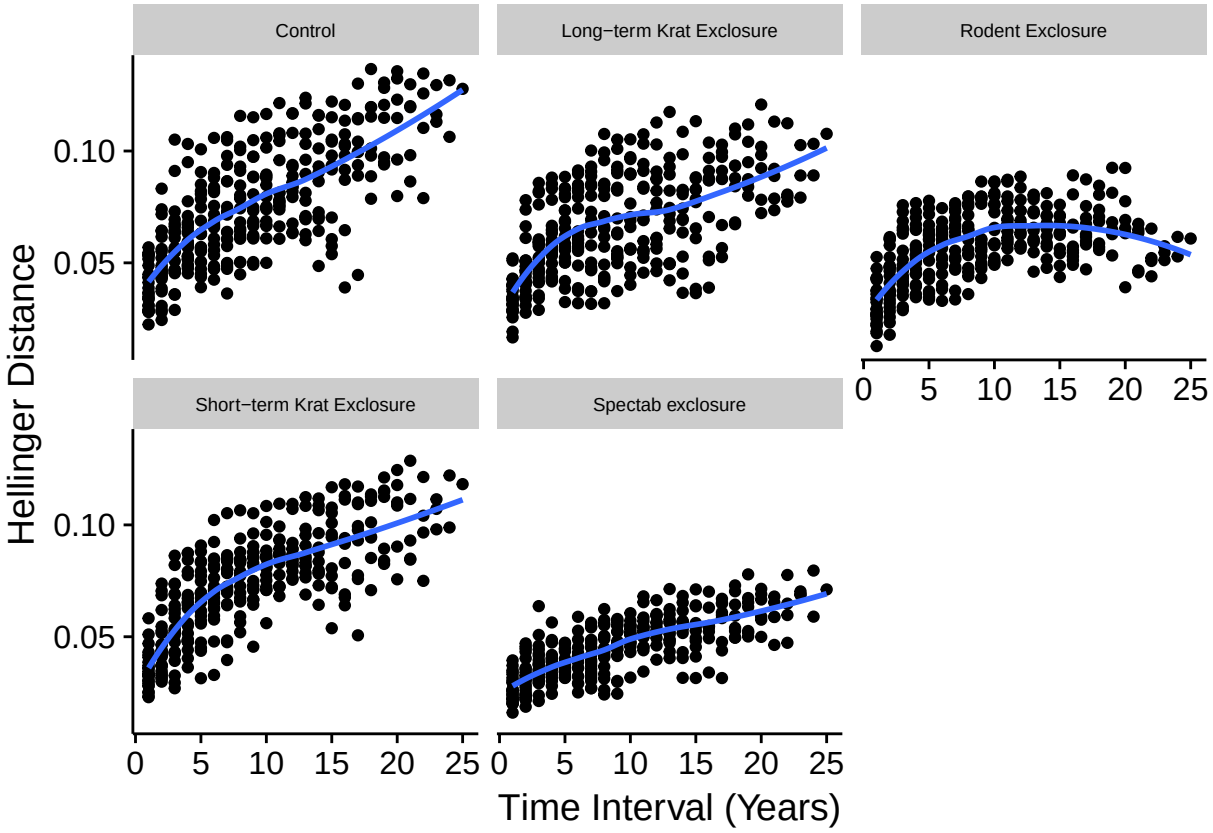
> ***Answer 6a***: A change in rank shift tells you about the change in rare and/or abundant species in a specific area over time. ***Answer 6b***: The rodent exposure plots seem to have greater mean rank shifts over time, which could indicate that it exhibited changes in the composition of rare and abundant species to a greater extent than the control plots.

Rate Change Interval

In the R code chunk below, do the following:

1. Calculate the rate change interval using the Hellinger distance.
2. Plot the results.

```r
# In order to calculate relative abundances, we need total abundances
# Let's add a column for total abundances
# We will relativize species abundances across the whole dataset so
# the transformed distances are preserved
portal.species.abunds$tot.abund <- rep(sum(portal.species.abunds$n),
                                       length(portal.species.abunds$n))

# Now, apply Hellinger transformation
portal.hellinger.transf <- portal.species.abunds %>%
  mutate(hellinger.transf = sqrt(n / tot.abund))

# The mutate function creates a new column "hellinger.transf"
# by taking the square root of species relative abundance
# We can use this new column as our "abundance" vector
portal.change.int <- rate_change_interval(portal.hellinger.transf,
  time.var = "year",
  species.var = "taxon",
  abundance.var = "hellinger.transf",
  replicate.var = "plot_type")

rate.plot <- ggplot(portal.change.int, aes(interval, distance)) +
  geom_point() +
  facet_wrap(~plot_type) +
  theme(strip.text.x = element_text(size = 7)) +
  stat_smooth(method = "loess", se = F, size = 1) +
  ylab("Hellinger Distance") +
  xlab("Time Interval (Years)")

rate.plot
```



Question 7:

a. What does it mean to calculate a distance metric across varying time intervals?
b. Interpret the overall results. Develop a hypothesis based on the different responses of each treatment.

Answer 7a: Calculating a distance metric across varying time intervals is measuring the distance among sites in terms of community dissimilarity over time. ***Answer 7b***: The control and short-term Krat exclosure plots exhibited an increase in helinger distance values over time, which indicates that community dissimilarity increasead over time. The long-term Krat exclosure and the Spectab exclosure also exhibited an increase in community dissimilarity over time, but the rate of change does not appear as fast as in the control and short-term Krat exclosure groups (as seen by the lower slopes). The rodent exclosure group exhibited an increase, followed by a decrease in community dissimilarity. From these results, I hypothesize that rodent exclosures limit changes in species composition over time because fewer species are able to enter and leave these plots. I also hypothesize that due to rodent exclosure, other species could perhaps become more abundant, leading to increased competition. This could explain the decrease in dissimilarity in the rodent exclosure treatment over time.

7) STABILITY

In the R code chunk below, do the following:

1. Using total abundance as your focal variable, calculate stability (i.e., 1/CV) and synchrony for each plot type.
2. Test for a biodiversity-stability relationship by regressing community stability on mean richness.
3. Test for a biodiversity-stability relationship by regressing community stability on mean inverse Simpson's diversity.

```r
#1. Calculate stability with each plot type
portal.stab <- community_stability(df = as.data.frame(portal.species.abunds),
  time.var = "year",
  abundance.var = "n",
  replicate.var = "plot_type")

pander(portal.stab)
```

plot_type	stability
Control	3.044
Long-term Krat Exclosure	1.865
Rodent Exclosure	1.864
Short-term Krat Exclosure	2.462
Spectab exclosure	2.911

```r
#Calculate synchrony for each plot type
portal.loreau <- synchrony(df = as.data.frame(portal.species.abunds),
  time.var = "year",
  species.var = "taxon",
  abundance.var = "n",
  replicate.var = "plot_type",
  metric = "Loreau")

names(portal.loreau)[2] <- "loreau"

portal.gross <- synchrony(df = as.data.frame(portal.species.abunds),
  time.var = "year",
  species.var = "taxon",
  abundance.var = "n",
  replicate.var = "plot_type",
  metric = "Gross")

names(portal.gross)[2] <- "gross"

pander(full_join(portal.loreau, portal.gross))
```

```
## Joining, by = "plot_type"
```

plot_type	loreau	gross
Control	0.1869	0.1263
Long-term Krat Exclosure	0.1578	0.07418
Rodent Exclosure	0.187	0.1423
Short-term Krat Exclosure	0.08773	0.001546
Spectab exclosure	0.1542	0.1393

```r
# Biodiversity-Stability

# Avg annual richness in each plot type
portal.mean.rich.plot <-rich.all %>%
  group_by(plot_id) %>%
  summarise(mean.rich = mean(richness))
```

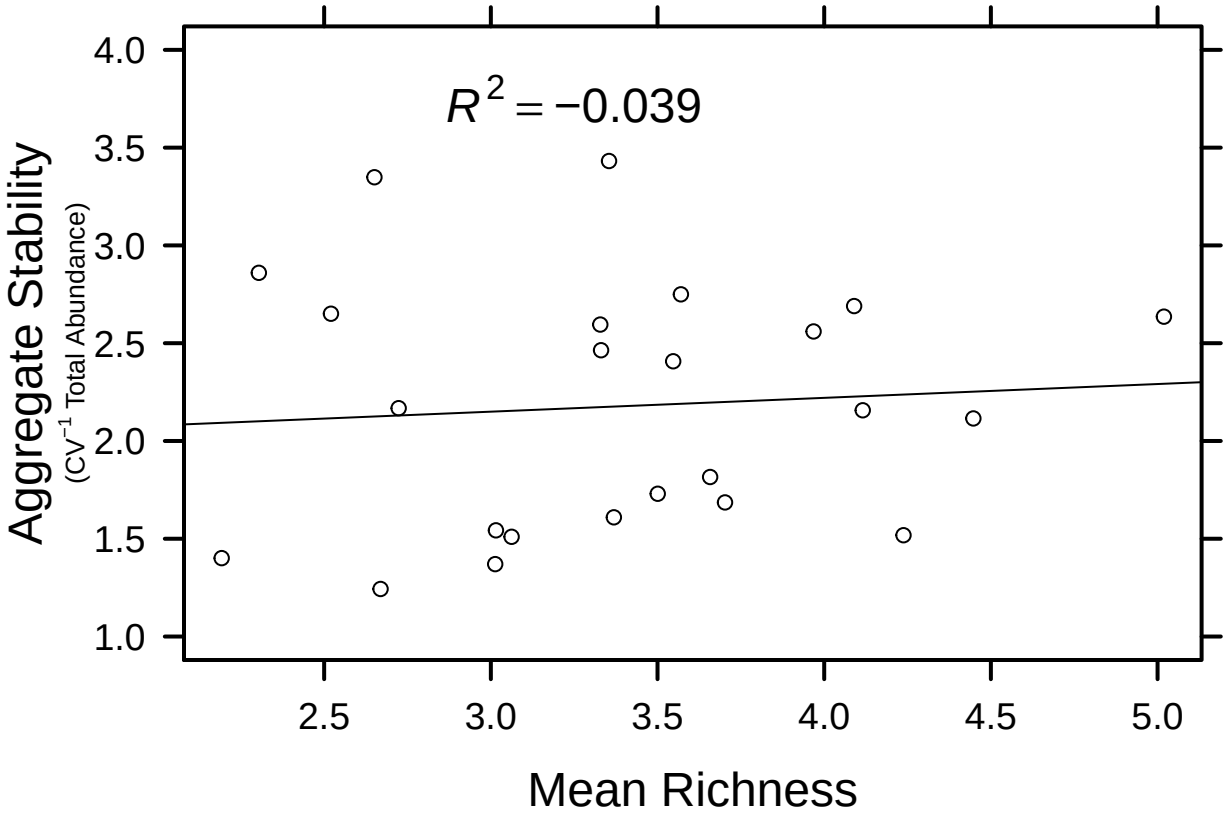
```
# How does stability metrics relate to mean richness?
portal.plot.abunds <- as.data.frame(
  group_by(portal, year, plot_id) %>% count(taxon))

portal.stab.plot <- community_stability(df =portal.plot.abunds,
   time.var ="year",
  abundance.var ="n",
  replicate.var ="plot_id")

# Join richness and stability
portal.div.stab <-portal.mean.rich.plot %>%
  inner_join(portal.stab.plot)
```

```
## Joining, by = "plot_id"
```

```
# Plot
par(mar = c(5,5,1,1))
plot(portal.div.stab$stability ~portal.div.stab$mean.rich,
xlab ="",ylab="",yaxt ="n",xaxt ="n",
lim = c(2,10),ylim = c(1,4))
```

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

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

## Warning in axis(side = side, at = at, labels = labels, ...): "lim" is not a
## graphical parameter

## Warning in axis(side = side, at = at, labels = labels, ...): "lim" is not a
## graphical parameter

## Warning in box(...): "lim" is not a graphical parameter

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

```
axis(side =1,cex.axis =1.2,lwd.ticks =2,las =1)
axis(side =2,cex.axis =1.2,lwd.ticks =2,las =1)
axis(side =3,lwd.ticks =2,las =1,labels =F)
axis(side =4,lwd.ticks =2,las =1,labels =F)
box(lwd =2)
mtext("Mean Richness",side =1,line =3,cex =1.5)
mtext("Aggregate Stability",side =2,line =3.5,cex =1.5)
mtext(expression(paste("(CV"^"-1"," Total Abundance)")),
      side =2,line =2.25,cex =0.8)

div.stab.lm <- lm(portal.div.stab$stability ~portal.div.stab$mean.rich)
abline(div.stab.lm)
r2 <- bquote(italic(R)^2==.(format(
  summary(div.stab.lm)$adj.r.square,digits =3)))
text(3.25,3.75,cex =1.5,labels =r2)
```



Question 8:

a. Which plot type has the highest stability in total abundance? How is stability of total abundance measured with the function you learned? How does this measure of stability relate to the coefficient of variation?

b. In your own words, describe the concept of synchrony

c. Interpret the results from the biodiversity-stability relationships you analyzed.

Answer 8a: The Control plot has the highest stability in total abundance. The stability of total abundance is measured by taking the variance divided by the average value of the community, and getting the inverse value of this. The stability essentially examines the lack of variation, while the coefficient of variation measures the variation. ***Answer 8b***: Synchrony is used to examine whether population densities change independently. ***Answer 8c***: Based on the plot obtained in this analysis, as species richness increases, community stability slightly increases as well.

SYNTHESIS

Compare and contrast the core concepts from temporal and spatial diversity (e.g., autocorrelation, scale, variability, etc.). Identify a few of the major challenges associated with studying biodiversity through time and across space.

Answer: Spatial and temporal analyses can show changes in biodiversity over various scales, and can be used to predict or forcast future changes in diversity. It is important to identify autocorrelation when conducting both spatial and temporal analyses, even though autocorrelation is determined differently in analysis. Spatial and temporal analyses can both be performed at small and large spatial scales as well. It is important to analyze variablity for spatial and temporal

analyses as well. Spatial and temporal changes in biodiversity are also related to one another in that biodiversity changes over spatial scales leads to changes over temporal scales.

There are several other factors to consider when conducting spatial and temporal analyses. Experiments must be conducted at the appropriate spatial and temporal scales in order to obtain meaningful, reliable results. Samples/data can be collected at fairly small spatial and time scales, but these samples may not be independent from each other. This must be taken into consideration when analyzing data from small spatial and time scales.