Assignment: Temporal Diversity

RZ Moger-Reischer; Z620: Quantitative Biodiversity, Indiana University

13 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/rmoge/GitHub/QB2017_Moger-Reischer/Week5-Temporal"
```

```
#setwd("~/GitHub/QB-2017/Week5-Temporal/")
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
```

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.

```
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 1
##  $ 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
```

```
print(length(unique(portal$plot_id)))
```

```
## [1] 24
```

```
print(length(unique(portal$species_id)))
```

```
## [1] 48
```

```
print(length(unique(portal$species)))
```

```
## [1] 40
```

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:24 ***Answer 1b***: 48

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).

```
#'''attach(portal)
#str(portal)
#mysbs94=filter(portal,year==1994)
#str(mysbs94)
#mysbs94=select(mysbs94,plot_id,species_id)'''



portal2 <- unite(portal, col = date, c(year, month, day), sep = "-", remove = FALSE)
portal3 <- unite(portal2, col = species_id, c(genus, species), sep = "_", remove = FALSE)
time.by.species <- group_by(portal, year, plot_id) %>% count(species_id) %>% spread(key = species_id, va
```

```
mysbs82=filter(time.by.species,year==1982)
mysbs82strip<-mysbs82[-c(1,2)]
#dat %>% mutate_each_(funs(factor), l1) %>% mutate_each_(funs(as.numeric), l2)
mysbs82strip<- apply(mysbs82strip,2, as.numeric)
#2
my_type<-portal%>%filter(year==1982)%>%group_by(plot_id,plot_type)%>%count(species_id)%>%spread(key = sp

#3
S.obs <- function(x=""   ){
   rowSums(x>0) *1
}#for each row x, take the sum of columns for which x > 0


SimpE <- function(x = ""){
S <- S.obs(x)#obsvd richness
x = as.data.frame(x)
D <- diversity(x, "inv")
E <- (D)/S
return(E)
}
thuggy<-SimpE(mysbs82strip)
```

```
#apply(mysbs82strip,1,SimpE)
thugy<-apply(mysbs82strip,1,diversity)
#4
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
```

```
portal82.db <- vegdist(mysbs82strip, method = "bray", upper = TRUE, diag = TRUE)
portal82.pcoa <- cmdscale(portal82.db, eig = TRUE, k = 3)#do the PCoA
str(portal82.pcoa)
```

```
## List of 5
##  $ points: num [1:24, 1:3] -0.309 -0.139 0.336 -0.33 -0.276 ...
##   ..- attr(*, "dimnames")=List of 2
##   .. ..$ : NULL
##   .. ..$ : NULL
##  $ eig   : num [1:24] 2.234 0.661 0.557 0.384 0.282 ...
##  $ x     : NULL
##  $ ac    : num 0
##  $ GOF   : num [1:2] 0.668 0.705
```

```
explainvar1 <- round(portal82.pcoa$eig[1] / sum(portal82.pcoa$eig), 3) * 100#for each of the first thre
explainvar2 <- round(portal82.pcoa$eig[2] / sum(portal82.pcoa$eig), 3) * 100
explainvar3 <- round(portal82.pcoa$eig[3] / sum(portal82.pcoa$eig), 3) * 100
sum.eig <- sum(explainvar1, explainvar2, explainvar3)#total amt explained in these 3 dimensions

par(mar = c(5, 5, 1, 2) + 0.1)#structure the figure output
# Initiate Plot
plot(portal82.pcoa$points[ ,1], portal82.pcoa$points[ ,2], ylim = c(-0.2, 0.7), xlab = paste("PCoA 1 ("
# 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(portal82.pcoa$points[ ,1], portal82.pcoa$points[ ,2],
pch = 19, cex = 3, bg = "#bc8d4f", col = "#bc8d4f")
```

```r
text(portal82.pcoa$points[ ,1], portal82.pcoa$points[ ,2],
labels = row.names(portal82.pcoa$points))
#Now IDfy the most influential spp
portal82REL <- mysbs82strip
  for(i in 1:nrow(mysbs82strip)){
    portal82REL[i, ] = mysbs82strip[i, ] / sum(mysbs82strip[i, ])
    }

portal82.pcoa2 <- add.spec.scores(portal82.pcoa,portal82REL,method = "pcoa.scores")
```

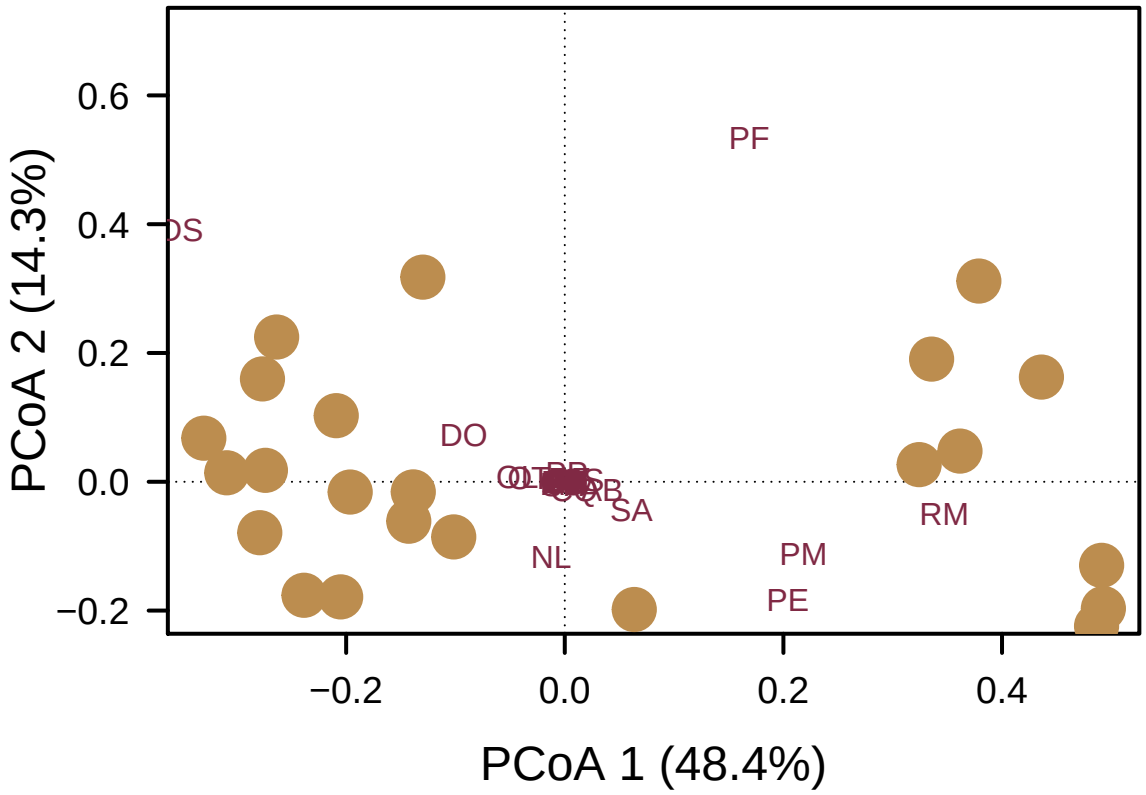
```
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
```

```
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
```

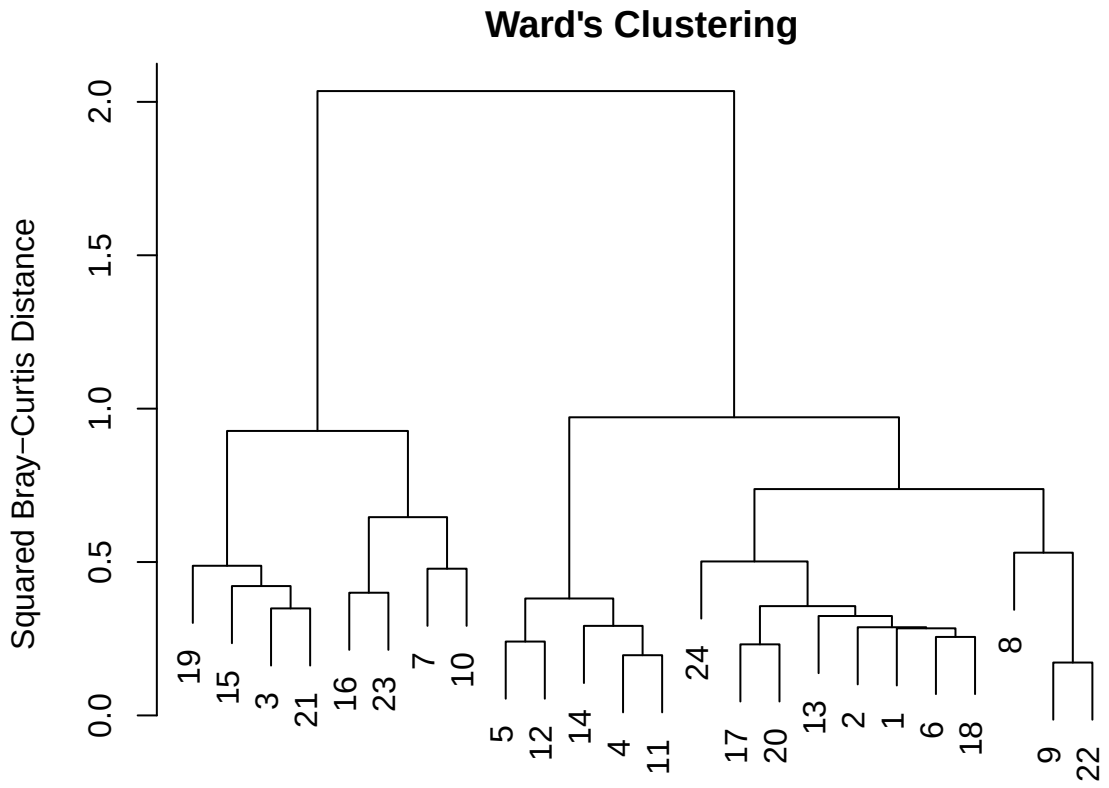
```
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
```

```
## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero

## Warning in cor(comm[, i], ordiscores[, j]): the standard deviation is zero
```

```r
###################################
#gives error: Error in is.data.frame(x) :
  #(list) object cannot be coerced to type 'double'
###############################

text(portal82.pcoa2$cproj[ ,1], portal82.pcoa2$cproj[ ,2],labels = row.names(portal82.pcoa2$cproj), col
```




```r
portal82.ward <- hclust(portal82.db, method = "ward.D2")#now visualize phylogenetically
par(mar = c(1, 5, 2, 2) + 0.1)#set up display settings
plot(portal82.ward, main = "Ward's Clustering",ylab = "Squared Bray-Curtis Distance")#plot the tree
```




```
#5
mytrtmnts<-c(my_type$plot_type)
adonis(mysbs82strip ~ mytrtmnts, method = "bray", permutations = 999)
```

```
##
## Call:
## adonis(formula = mysbs82strip ~ mytrtmnts, permutations = 999,      method = "bray")
##
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
##            Df SumsOfSqs MeanSqs F.Model      R2 Pr(>F)
## mytrtmnts   1    0.1006 0.10059 0.48973 0.02178  0.796
## Residuals  22    4.5190 0.20541         0.97822
## Total      23    4.6196                 1.00000
```

```
#P=0.786.
```

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:

Among the sites, Shannon's diversity ranges from 1.0364 to 2.1453.

Sites 16, 23, 7, 10 all cluster together, and all are Rodent Exclosure treatments.

Sites 19, 15, 3, 21 all cluster together, and all are LT K-ray Exclosures.

Sites 14, 4, 11, 12, 5 cluster together; and except 5 are controls.

Sites 6, 18, 1, 2, 13cluster together; 18, 6, and 13 are all Short-Term K-rat exclosures

Sites 17 and 20 cluster together, but they are not the same treatment.

Sites 22, 9, 8 cluster together, but only 22 and 8 are the same treatment.

Site 24, a Rodent Exclosure, is far distant from the other rodent exclosures.

Answer 2b:

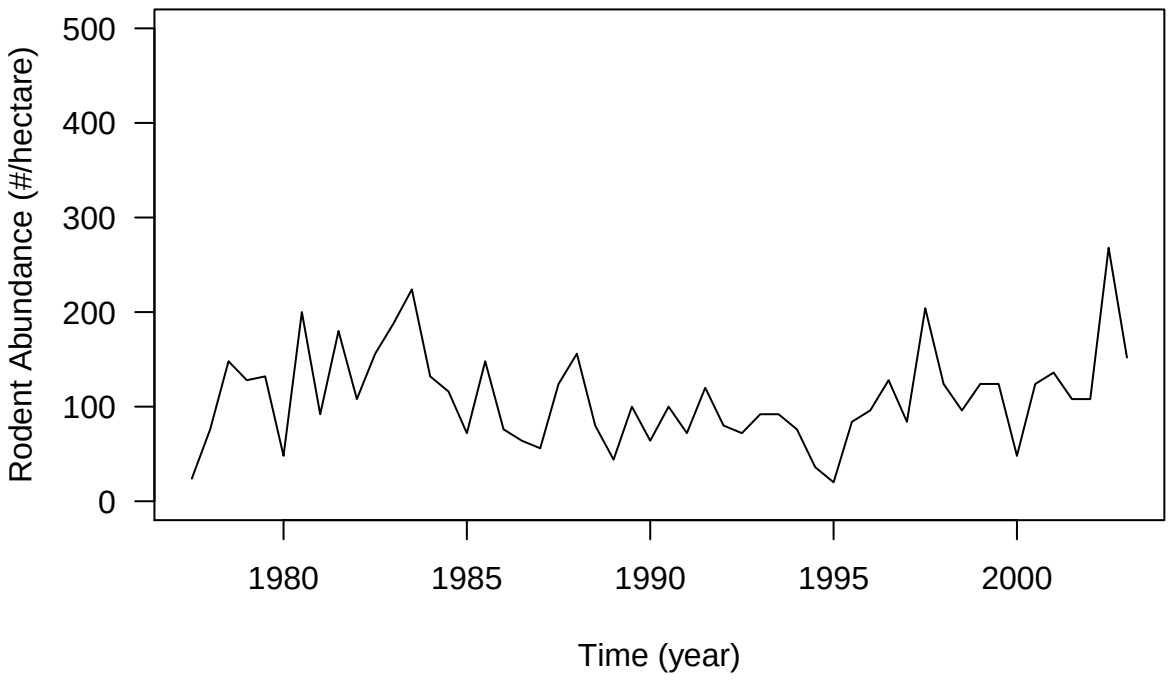
No, treatment is not a significant predictor of species composition (P = 0.786).

4) TIME SERIES ANALYSIS

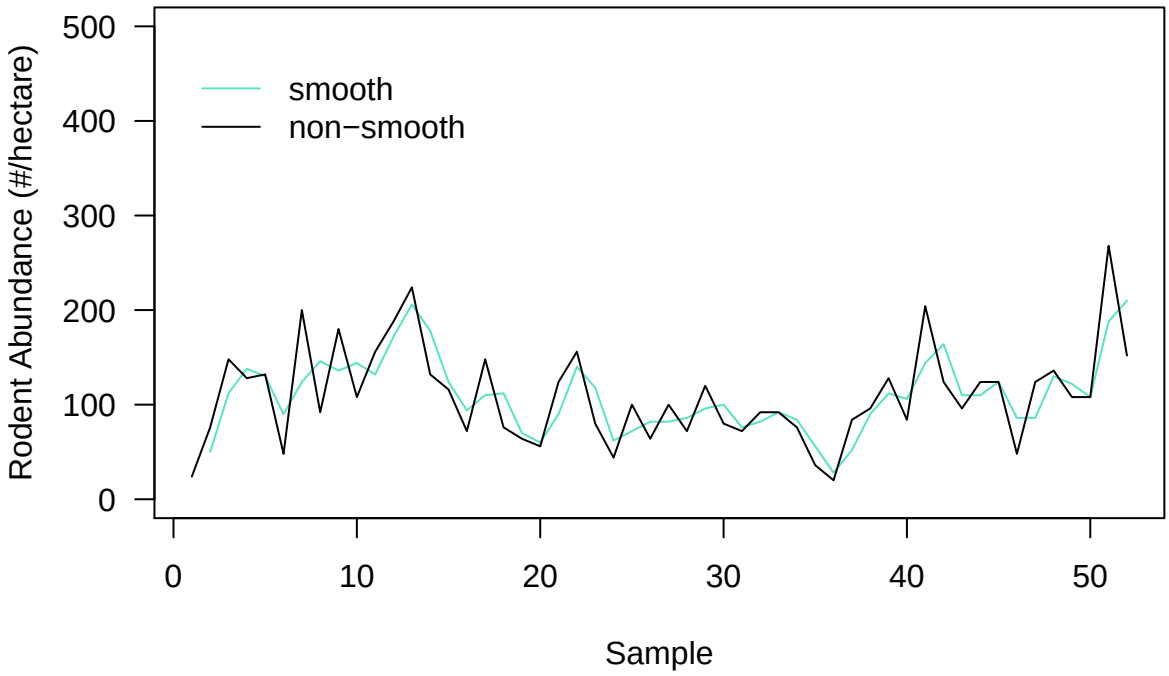
In the R code chunk below, do the following:

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 ways 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.

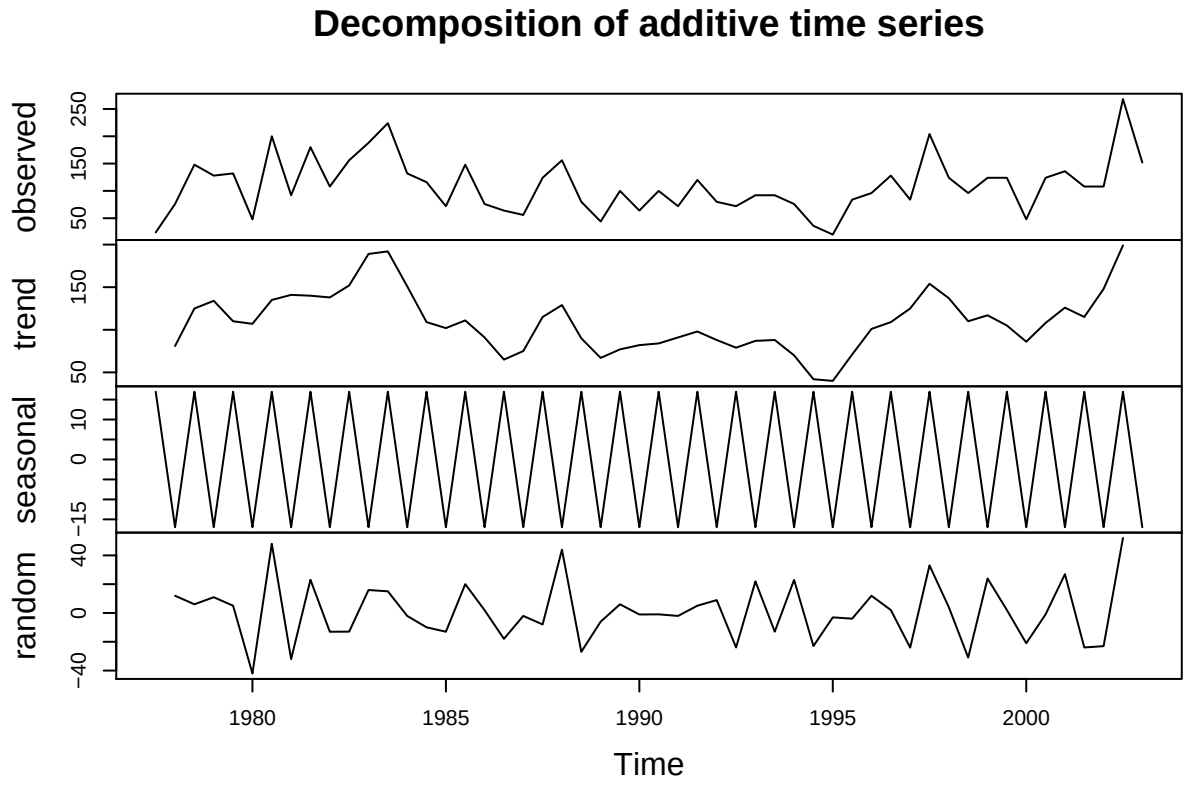
```
#1
time.by.spec.2 <- filter(portal, taxa=="Rodent") %>% filter(plot_id==18) %>% group_by(year, month, plot_
time.by.spec.2$season <- NA
time.by.spec.2$season <- time.by.spec.2$month %in% c(6:10)
time.by.spec.2$season <- ifelse(time.by.spec.2$season == TRUE, "rain", "norain")
time.by.spec.2<-group_by(time.by.spec.2, year, season)
abund<-count(time.by.spec.2,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
```




```
#2
abund.sm <- SMA(abund$nn, n = 2)#used 2 bc we have condensed to hexamonthly datapoints
plot(abund.sm, type = "l", col = "#56e4c6", ylab = "Rodent Abundance (#/hectare)",
xlab = "Sample", las = 1, ylim = c(0, 500))
lines(abund$nn, col = "black")
legend(0, 475, col = c("#56e4c6", "black"), lty = c(1,1),
c("smooth", "non-smooth"), bty = "n", cex = 1)
```




```
abund.comp <- decompose(abund.ts)
# decomposition categories
plot(abund.comp)
```




```
#3. Test whether your data meet the assumption of stationarity.
adf.raw <- adf.test(abund.ts, alternative = "stationary")
print(adf.raw$p.value)
```

```
## [1] 0.7758961
```

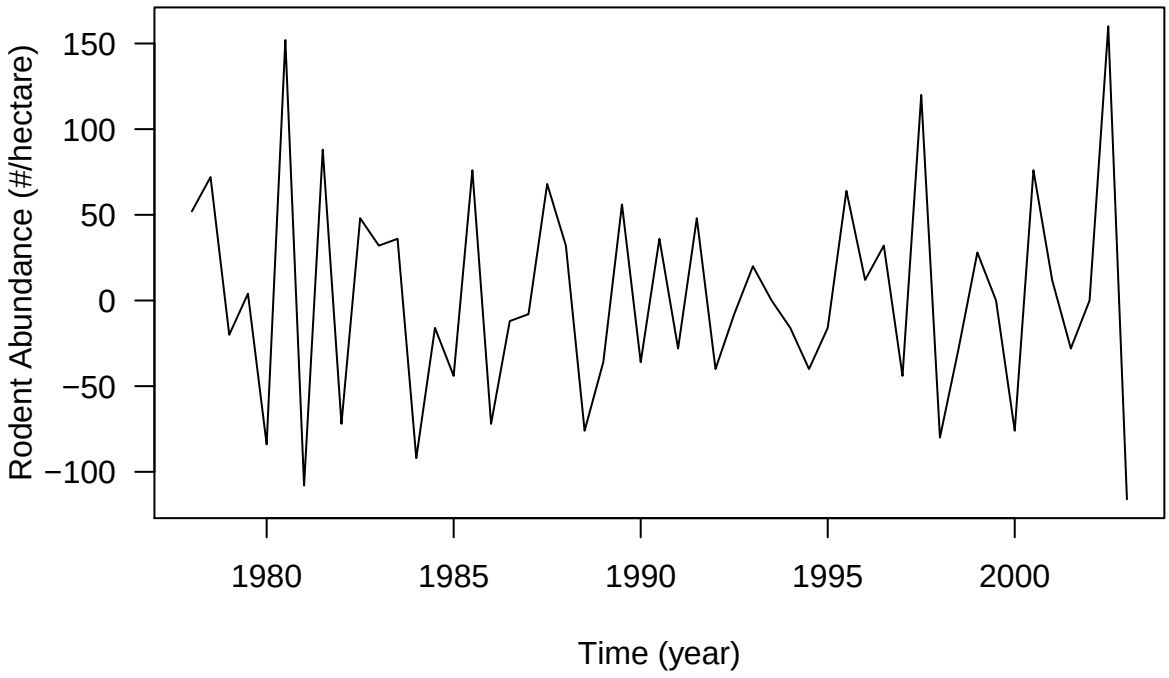
```
#for plot #18, the ts is NOT stationary: P=.77

#4. If it does not meet this assumption, explore ways to make your data stationary.
abund.ts.diff <- diff(abund.ts,lag=1)
adf.diff <- adf.test(abund.ts.diff, alternative = "stationary")
```

```
## Warning in adf.test(abund.ts.diff, alternative = "stationary"): p-value
## smaller than printed p-value
```

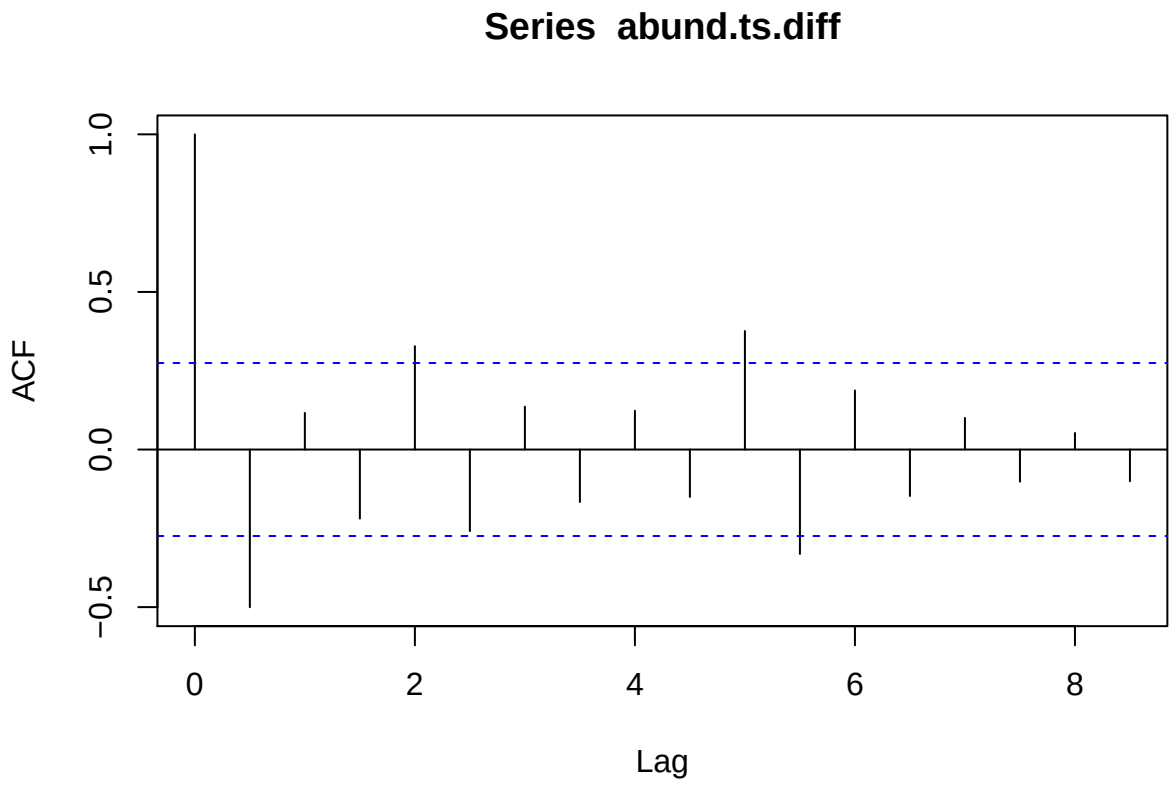
```
adf.diff$p.value
```

```
## [1] 0.01
```

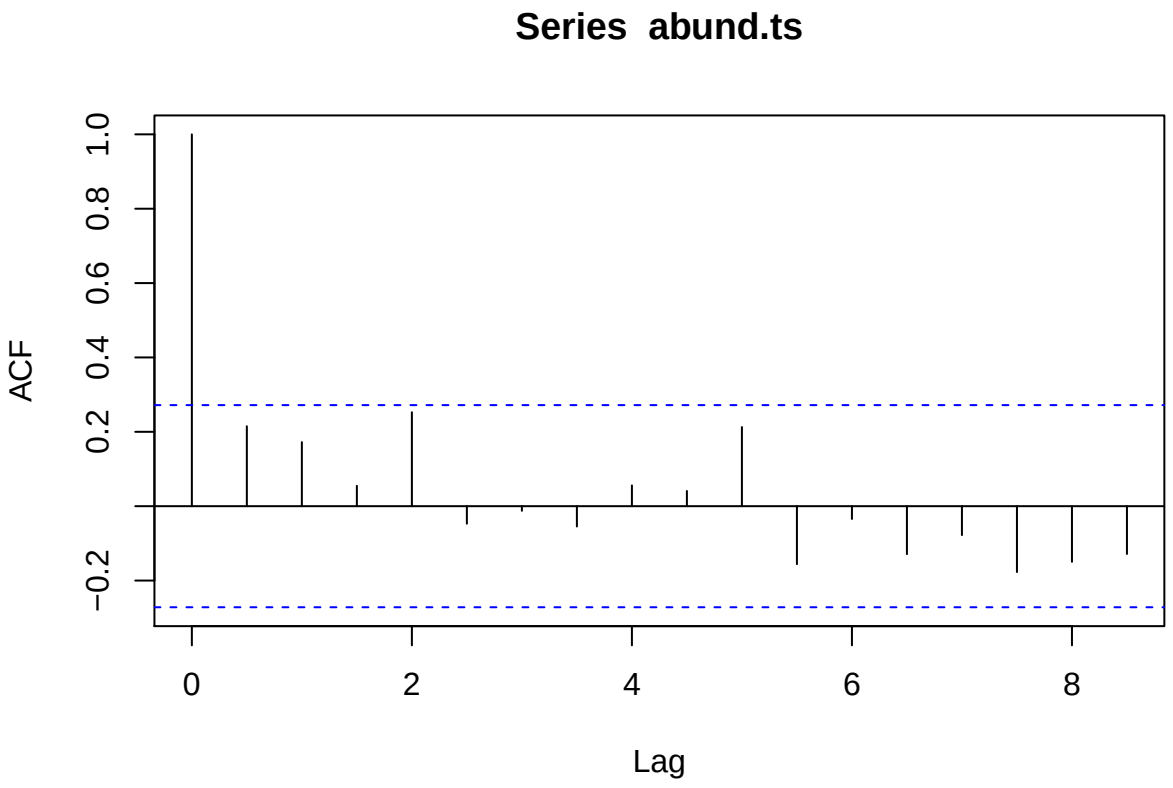
```
#now P=0.01, so null hypot of nonstationarity is rejected
plot.ts(abund.ts.diff, type = "l", ylab = "Rodent Abundance (#/hectare)", xlab = "Time (year)", las = 1
```



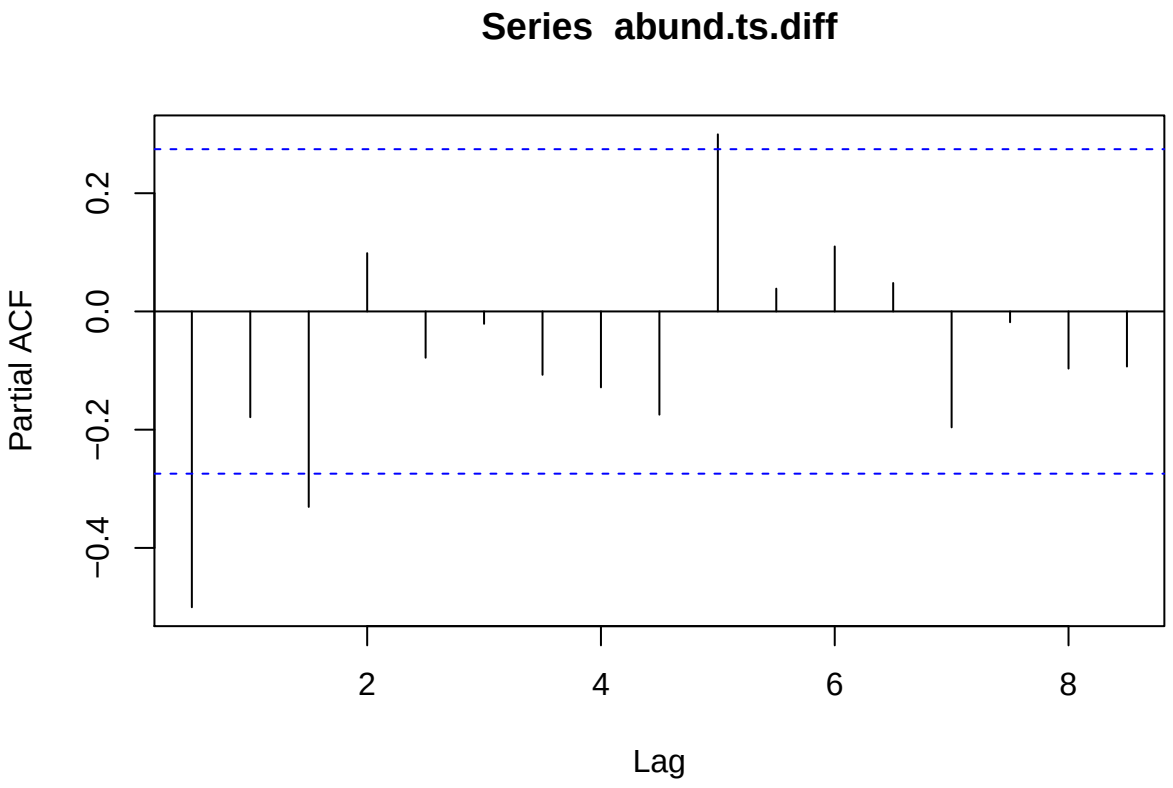

```
#5. Examine and plot time lags using the partial autocorrelation function (PACF) and autocorrelation fu
acf(abund.ts.diff)
```



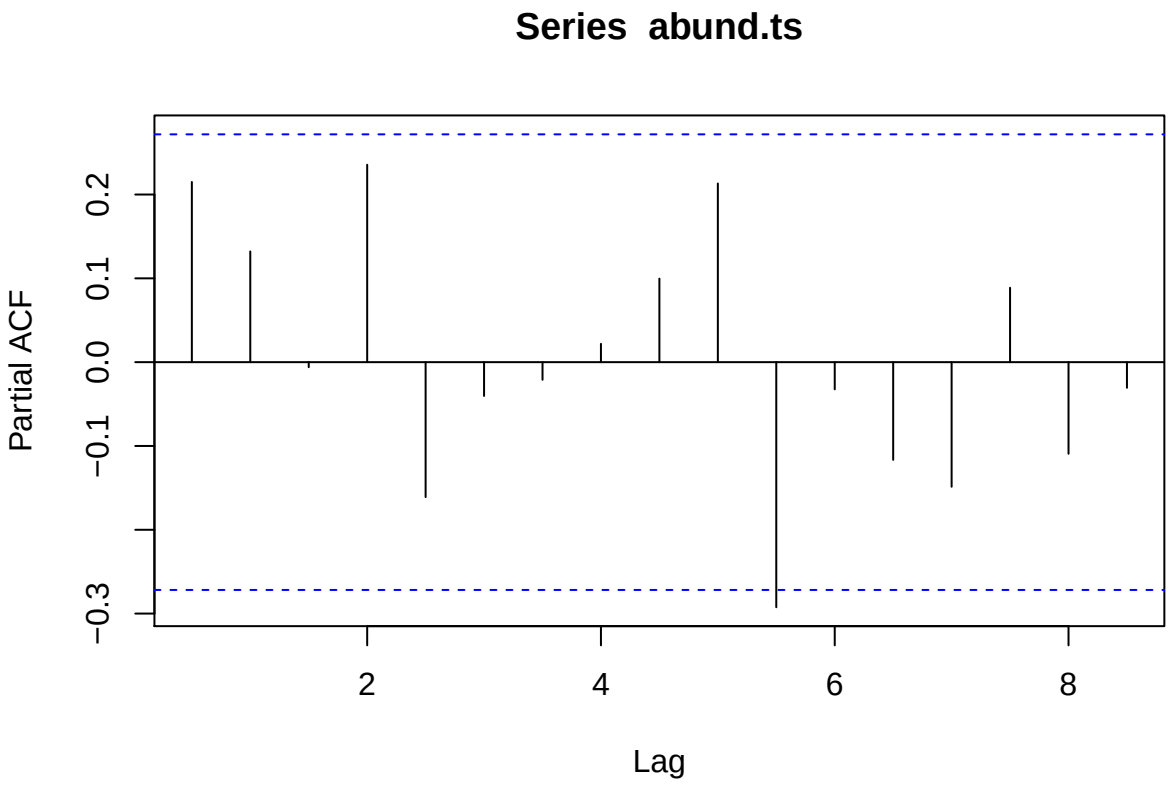

```
acf(abund.ts)
```




```
pacf(abund.ts.diff)
```



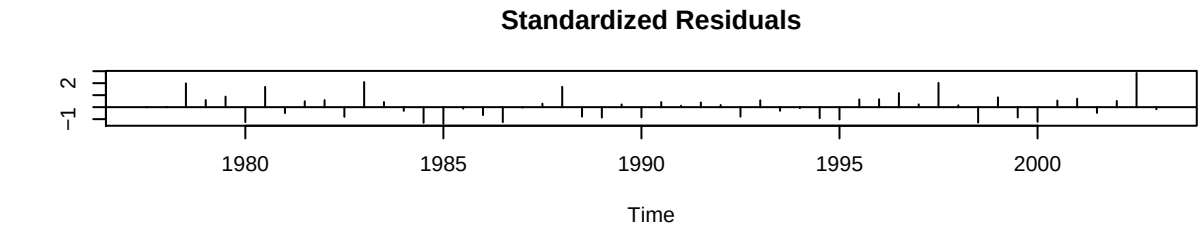

```
pacf(abund.ts)
```

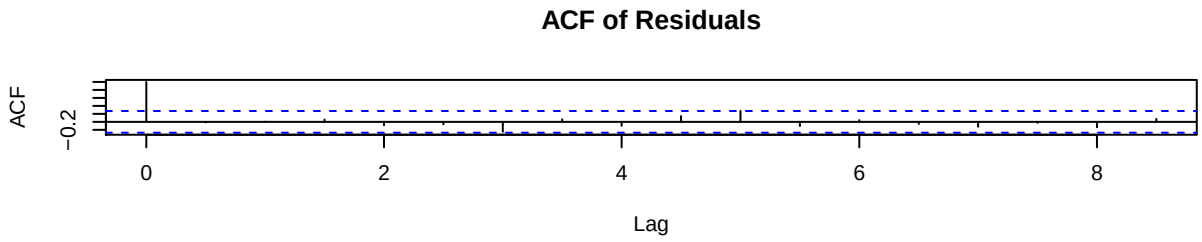


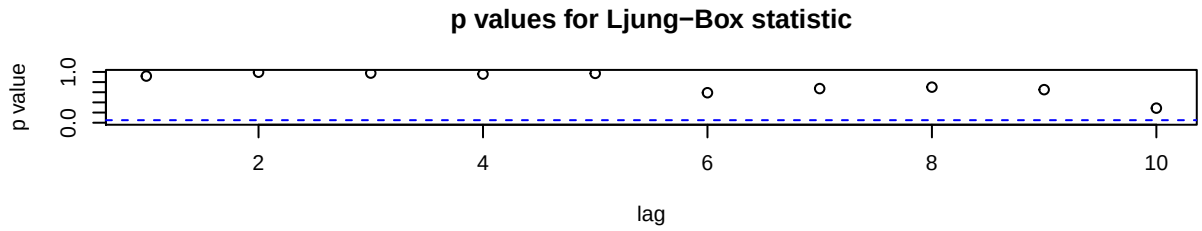

```
#6. Use the tools outlined in the Handout to create an ARMA model.
fit18<-auto.arima(abund.ts,ic="aic",trace=T)
```

```
##
##  ARIMA(1,0,1)(1,0,1)[2] with non-zero mean : 559.8672
##  ARIMA(0,0,0) with non-zero mean : 558.7625
##  ARIMA(1,0,0)(1,0,0)[2] with non-zero mean : 558.7181
##  ARIMA(0,0,1)(0,0,1)[2] with non-zero mean : 559.4827
##  ARIMA(0,0,0) with zero mean     : 648.0038
##  ARIMA(1,0,0) with non-zero mean : 558.1654
##  ARIMA(1,0,0)(0,0,1)[2] with non-zero mean : 559.3108
##  ARIMA(1,0,0)(1,0,1)[2] with non-zero mean : Inf
##  ARIMA(1,0,1) with non-zero mean : 558.7772
##  ARIMA(1,0,0) with zero mean     : 578.5296
##
##  Best model: ARIMA(1,0,0) with non-zero mean
```

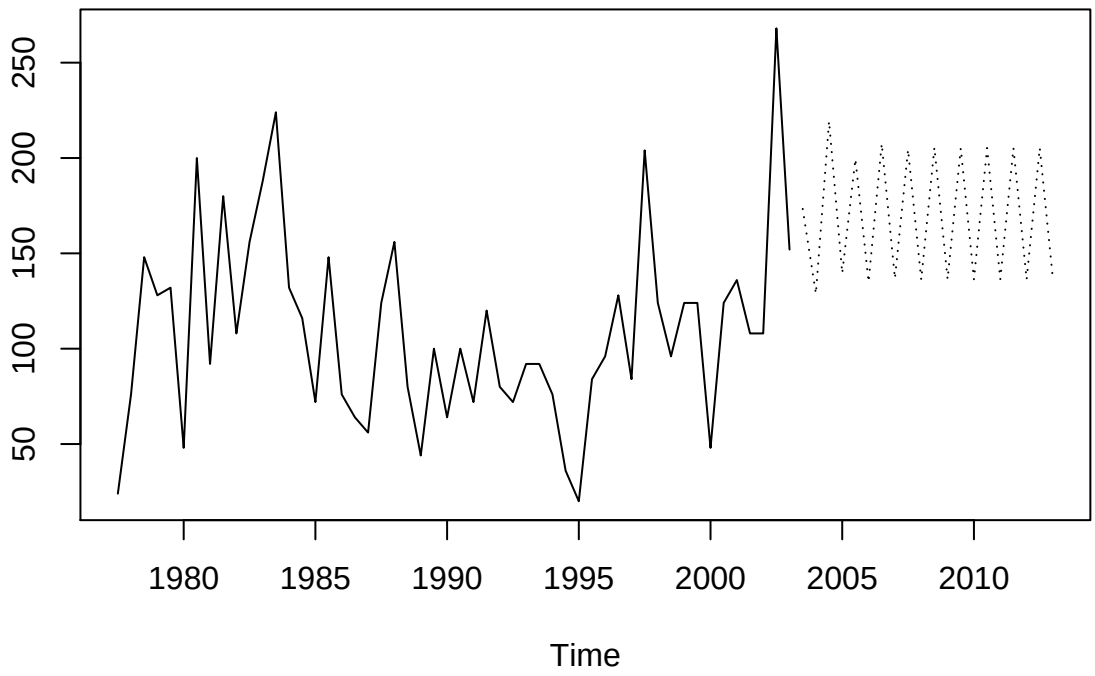
```
fit18<-arima(abund.ts,c(1,0,0),seasonal=list(order=c(2,1,0),period=2),include.mean=T)
#model is order (1,0,0) with nonzero mean
tsdiag(fit18)
```








```
pred18 <- predict(fit18, n.ahead = 20)
ts.plot(abund.ts, pred18$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:

No. There is a systematic change in species composition over time, not just effects of seasonality and stochastic noise.

Answer 3b:

The ACF function looks for correlations between measurements at different lags. That is, correlation when lag = 0 will always be 1. But at lag = 1, the ACF asks whether measurements at time t and at time t-1 are more similar to one another than would be expected if measurements at all time points were independent of one another. Similarly for lag = 2, ACF asks if measurements and time = t and time = t-2 are more similar than would be expected if all measurements were independent of one another. This would be tested for all valid values of t; the number of valid values will be lower for larger lags because the extent of the time series is not infinite. That said, this shouldn't be too much of a problem if the autocorrelation dies out as lags grow larger, and becomes nonsignificant.

In the ARMA model, an autocorrelation could be accounted for by adding a q term—a moving average.

I don't know how the ACF differs inside and outside of the ARMA.

Answer 3c:

There is a significant partial autocorrelation, what was accounted for in the lowest-AIC ARIMA model. At two lags there were ACF terms that were nearly, but not, significant. The lowest-AIC model did not include a q term. I also had to add a seasonal parameter and a nonzero mean; without these, the predictions especially would have been wrong.

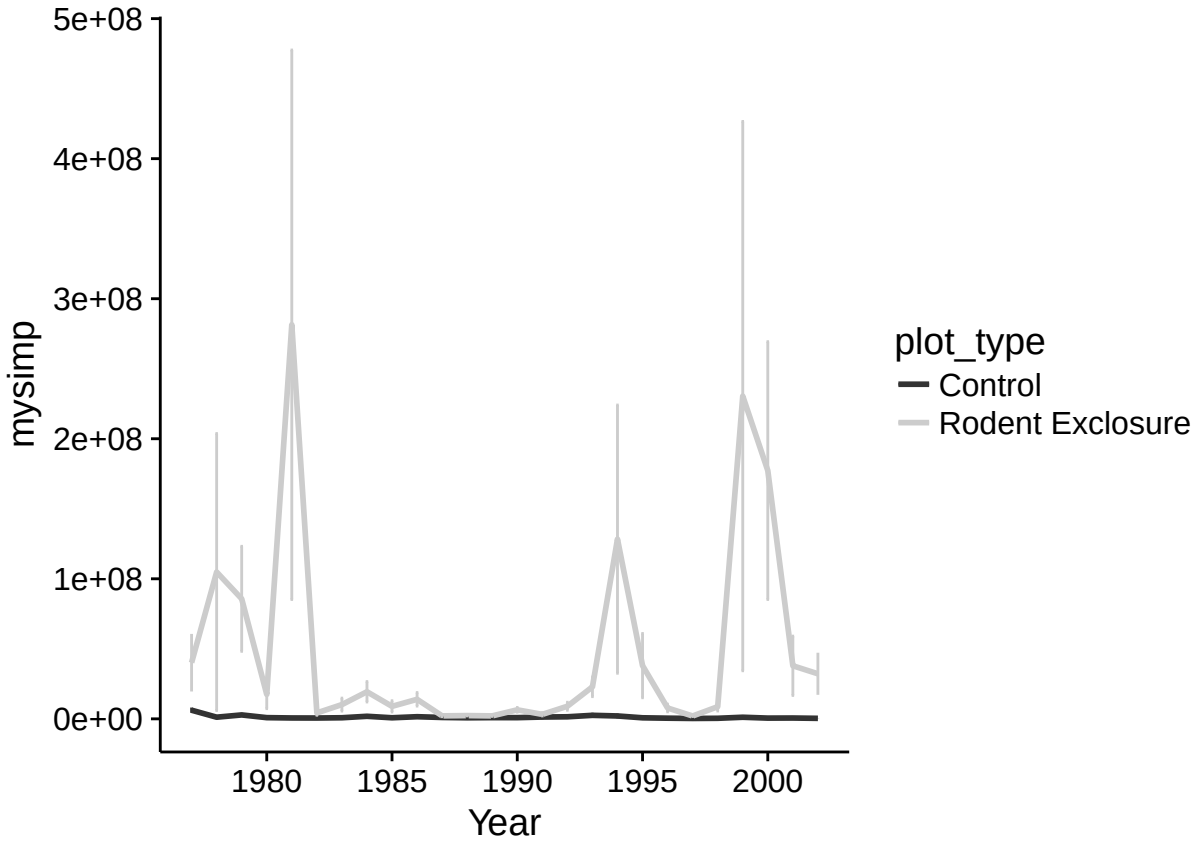
I used a simple moving average with a window of size 2. This smoothed the series a little, without oversmoothing. This made sense because the periodicity of timeseries should indeed have been every two measurements (they were collapsed into two measurements per year, one an aggregate of rainy season, one an aggregate of nonrainy season.

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.

```
#1. Create an appropriate data frame for RM-ANOVA (e.g., yearly species abundance values within plots).
# tbs matrix
time.by.speciesRM <- group_by(portal, year, plot_id, plot_type) %>% count(species_id) %>% spread(key = s
#time.by.speciesRM<- apply(time.by.speciesRM,2, as.numeric)
SimpD <- function(x = ""){
  D = 0
  N = sum(x)
  for (n_i in x){
    D = D + (n_i^2)/(N^2)
  }
  return(D)
}

mysimp <- as.data.frame(1/(SimpD((time.by.speciesRM[,-c(1:3)]))))

# Create df with experimental design and mysimp data
simp.all <- data.frame(time.by.speciesRM[,1:3,], mysimp)

names(simp.all)[4] <- "invD"
#choose two trtmnt types
simp.treat <- simp.all[which(simp.all$plot_type =="Control" | simp.all$plot_type == "Rodent Exclosure")

#2. Calculate the inverse of Simpson's diversity for each year, and plot it as a function of year for t



simp.treat.plot <- group_by(simp.treat, plot_type, year) %>% summarise(mean = mean(invD),sd = sd(invD),

simp.plot <- ggplot(simp.treat.plot, aes(x = year, y = mean, color = plot_type)) + geom_line(size = 1, s

plot(simp.plot)
```




```
#3. Perform an RM-ANOVA and construct a F-test using the AR(1), compound symmetery, and unstructured co

simp.rm <- lme(invD ~ plot_type * year, random = ~ 1 | plot_id,
correlation = corAR1(form = ~ 1 | plot_id),#I do not understand why the factor on this line is plot ID.
data = simp.treat)
#output
summary(simp.rm)
```

```
## Linear mixed-effects model fit by REML
##  Data: simp.treat
##        AIC      BIC    logLik
##   14139.42 14166.53 -7062.711
##
## Random effects:
##  Formula: ~1 | plot_id
##         (Intercept)  Residual
## StdDev:    15331779 102898180
##
## Correlation Structure: AR(1)
##  Formula: ~1 | plot_id
##  Parameter estimate(s):
##       Phi
## 0.1874974
## Fixed effects: invD ~ plot_type * year
##                                     Value  Std.Error  DF    t-value
## (Intercept)                     156026686 2228502236 343  0.0700141
```

```
## plot_typeRodent Exclosure      -1229150816 3484132014  12 -0.3527854
## year                                -77783    1120120 343 -0.0694415
## plot_typeRodent Exclosure:year      641619    1751120 343  0.3664052
##                                p-value
## (Intercept)                     0.9442
## plot_typeRodent Exclosure       0.7304
## year                            0.9447
## plot_typeRodent Exclosure:year  0.7143
##  Correlation:
##                                (Intr) plt_RE year
## plot_typeRodent Exclosure      -0.64
## year                           -1.00   0.64
## plot_typeRodent Exclosure:year  0.64  -1.00  -0.64
##
## Standardized Within-Group Residuals:
##           Min            Q1           Med            Q3           Max
## -0.6927097793 -0.3071160712 -0.0075071343  0.0002537696 11.2529386739
##
## Number of Observations: 359
## Number of Groups: 14
```

```
#BIC = 14166.53


# F-test
anova(simp.rm)
```

```
##                numDF denDF  F-value p-value
## (Intercept)        1   343 7.713627  0.0058
## plot_type          1    12 9.304367  0.0101
## year               1   343 0.046041  0.8302
## plot_type:year     1   343 0.134253  0.7143
```

```
#plot type is signif, but neither year nor intxn term is
set.caption("RMANOVA for Portal")
pander(anova(simp.rm))
```

Table 1: RMANOVA for Portal

	numDF	denDF	F-value	p-value
(Intercept)	1	343	7.714	0.005782
plot_type	1	12	9.304	0.01008
year	1	343	0.04604	0.8302
plot_type:year	1	343	0.1343	0.7143

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

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

```
##  plot_type          lsmean       SE df  lower.CL upper.CL
##  Control           1269955 10123999 13 -20601616 23141525
##  Rodent Exclosure 48684355 11827181 12  22915141 74453568
##
## Confidence level used: 0.95
```

```r
#now try other models of covariance structure


simp.rm <- lme(invD ~ plot_type * year, random = ~ 1 | plot_id,
correlation = corAR1(form = ~ "unstructured" | plot_id),#I do not understand why the factor on this lin
data = simp.treat)
#output
summary(simp.rm)
```

```
## Linear mixed-effects model fit by REML
##  Data: simp.treat
##        AIC      BIC    logLik
##   14139.42 14166.53 -7062.711
##
## Random effects:
##  Formula: ~1 | plot_id
##         (Intercept)  Residual
## StdDev:    15331779 102898180
##
## Correlation Structure: AR(1)
##  Formula: ~"unstructured" | plot_id
##  Parameter estimate(s):
##       Phi
## 0.1874974
## Fixed effects: invD ~ plot_type * year
##                                     Value  Std.Error  DF    t-value
## (Intercept)                     156026686 2228502236 343  0.0700141
## plot_typeRodent Exclosure      -1229150816 3484132014  12 -0.3527854
## year                               -77783    1120120 343 -0.0694415
## plot_typeRodent Exclosure:year      641619    1751120 343  0.3664052
##                                   p-value
## (Intercept)                        0.9442
## plot_typeRodent Exclosure          0.7304
## year                               0.9447
## plot_typeRodent Exclosure:year     0.7143
##  Correlation:
##                                   (Intr) plt_RE year
## plot_typeRodent Exclosure         -0.64
## year                              -1.00   0.64
## plot_typeRodent Exclosure:year     0.64  -1.00  -0.64
##
## Standardized Within-Group Residuals:
##           Min            Q1           Med            Q3           Max
## -0.6927097793 -0.3071160712 -0.0075071343  0.0002537696 11.2529386739
##
## Number of Observations: 359
## Number of Groups: 14
```

```r
#BIC = 14166.53


# F-test
anova(simp.rm)
```

```
##                numDF denDF  F-value p-value
```

```
## (Intercept)        1   343 7.713627  0.0058
## plot_type          1    12 9.304367  0.0101
## year               1   343 0.046041  0.8302
## plot_type:year     1   343 0.134253  0.7143
```

```
#plot type is signif, but neither year nor intxn term is
set.caption("RMANOVA for Portal")
pander(anova(simp.rm))
```

Table 2: RMANOVA for Portal

	numDF	denDF	F-value	p-value
(Intercept)	1	343	7.714	0.005782
plot_type	1	12	9.304	0.01008
year	1	343	0.04604	0.8302
plot_type:year	1	343	0.1343	0.7143

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

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

```
##  plot_type          lsmean       SE df  lower.CL upper.CL
##  Control           1269955 10123999 13 -20601616 23141525
##  Rodent Exclosure 48684355 11827181 12  22915141 74453568
##
## Confidence level used: 0.95
```

```
simp.rm <- lme(invD ~ plot_type * year, random = ~ 1 | plot_id,
correlation = corAR1(form = ~ "compound symmetry" | plot_id),#I do not understand why the factor on thi
data = simp.treat)
#output
summary(simp.rm)
```

```
## Linear mixed-effects model fit by REML
##  Data: simp.treat
##        AIC      BIC    logLik
##   14139.42 14166.53 -7062.711
##
## Random effects:
##  Formula: ~1 | plot_id
##         (Intercept)  Residual
## StdDev:    15331779 102898180
##
## Correlation Structure: AR(1)
##  Formula: ~"compound symmetry" | plot_id
##  Parameter estimate(s):
##       Phi
## 0.1874974
## Fixed effects: invD ~ plot_type * year
##                                     Value  Std.Error  DF    t-value
## (Intercept)                     156026686 2228502236 343  0.0700141
## plot_typeRodent Exclosure     -1229150816 3484132014  12 -0.3527854
## year                              -77783    1120120 343 -0.0694415
## plot_typeRodent Exclosure:year     641619    1751120 343  0.3664052
```

```
##                                p-value
## (Intercept)                     0.9442
## plot_typeRodent Exclosure       0.7304
## year                            0.9447
## plot_typeRodent Exclosure:year  0.7143
##  Correlation:
##                                (Intr) plt_RE year
## plot_typeRodent Exclosure      -0.64
## year                           -1.00   0.64
## plot_typeRodent Exclosure:year  0.64  -1.00  -0.64
##
## Standardized Within-Group Residuals:
##           Min            Q1           Med            Q3           Max
## -0.6927097793 -0.3071160712 -0.0075071343  0.0002537696 11.2529386739
##
## Number of Observations: 359
## Number of Groups: 14
```

```
#BIC = 14166.53


# F-test
anova(simp.rm)
```

```
##                numDF denDF  F-value p-value
## (Intercept)        1   343 7.713627  0.0058
## plot_type          1    12 9.304367  0.0101
## year               1   343 0.046041  0.8302
## plot_type:year     1   343 0.134253  0.7143
```

```
#plot type is signif, but neither year nor intxn term is
set.caption("RMANOVA for Portal")
pander(anova(simp.rm))
```

Table 3: RMANOVA for Portal

	numDF	denDF	F-value	p-value
(Intercept)	1	343	7.714	0.005782
plot_type	1	12	9.304	0.01008
year	1	343	0.04604	0.8302
plot_type:year	1	343	0.1343	0.7143

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

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

```
##  plot_type          lsmean       SE df  lower.CL upper.CL
##  Control           1269955 10123999 13 -20601616 23141525
##  Rodent Exclosure 48684355 11827181 12  22915141 74453568
##
## 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 intention of the RM-ANOVA is to acknowledge and account for the fact that, in a longitudinal study, datapoints represent the SAME piece of data, measured at a different time. There would be a big difference between measuring one human's blood pressure every 5 years for its entire lifetime, and measuring blood pressure of 14 different individuals aged 5 - 70 (at 5-year) intervals.

Answer 4b:

There is no apparent trend. There is much more noise and much wider error bars for the exclosure treatment than for the control treatment.

Answer 4c:

In ANOVA, the F stat is just a test statistic with a known distribution that can be used to assess above which value for F the area under that distribution's curve is small enough to reject the null hypothesis. That is, if F is above some critical value (determined by our choice of significance level alpha and degrees of freedom), the P will be below some critical value.

4d

The covariance structure did not seem to matter. The values for each information criterion, and for the phi parameter, were the same, irrespective of with covariance structure I indicated.

6) TEMPORAL BETA DIVERSITY

Turnover

In the R code chunk below, do the following:

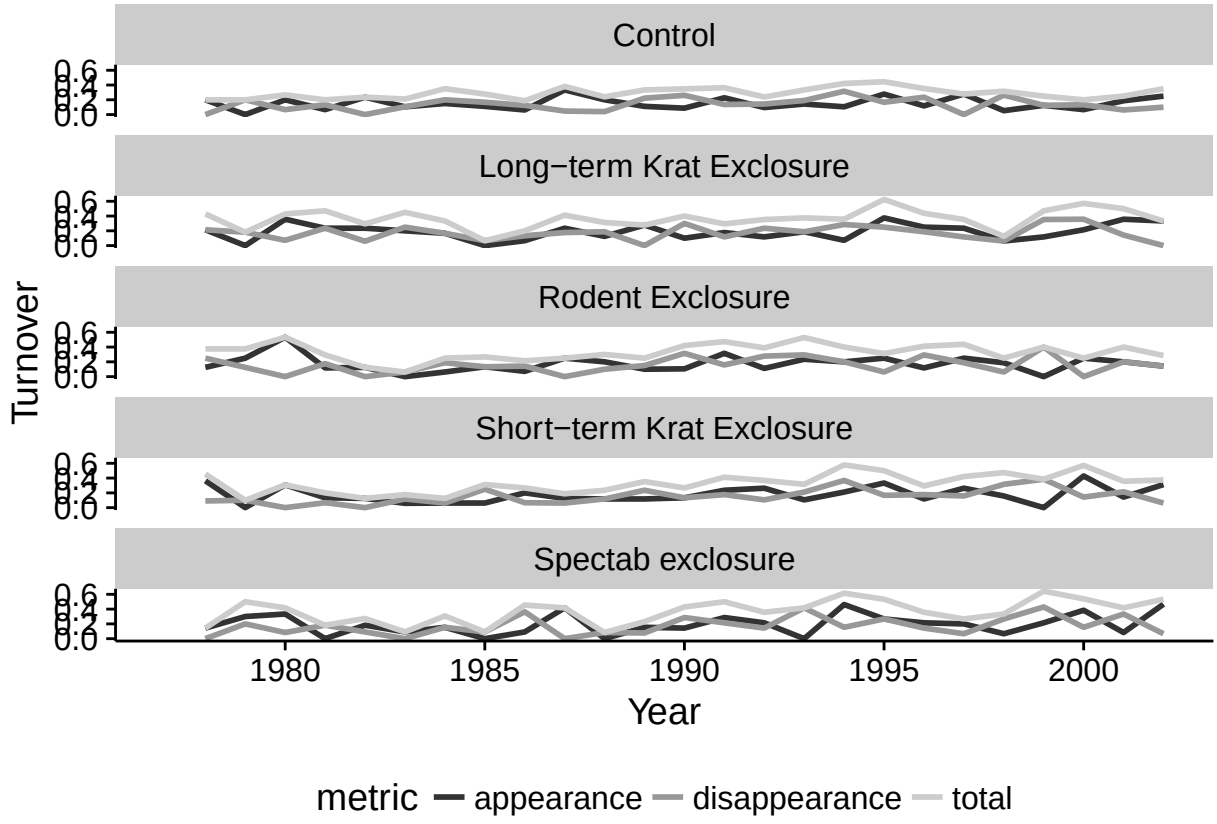
1. Calculate species abundances for each taxonomic group (the `taxa` column).
2. Calculate total turnover and turnover due to the gain/loss of species for each group.
3. Visualize turnover within each group

```
#1
portal.species.abunds <- group_by(portal, year, plot_type) %>% count(species_id)
#2
# total turnover
portal.total <- turnover(df = portal.species.abunds, time.var = "year", species.var = "species_id", abu
# species gained
portal.appearance <- turnover(df = portal.species.abunds,time.var = "year",species.var = "species_id",a
# species lost
portal.disappearance <- turnover(df = portal.species.abunds,time.var = "year",species.var = "species_id

portal.turnover <- full_join(portal.total, portal.disappearance) %>% full_join(portal.appearance)
```

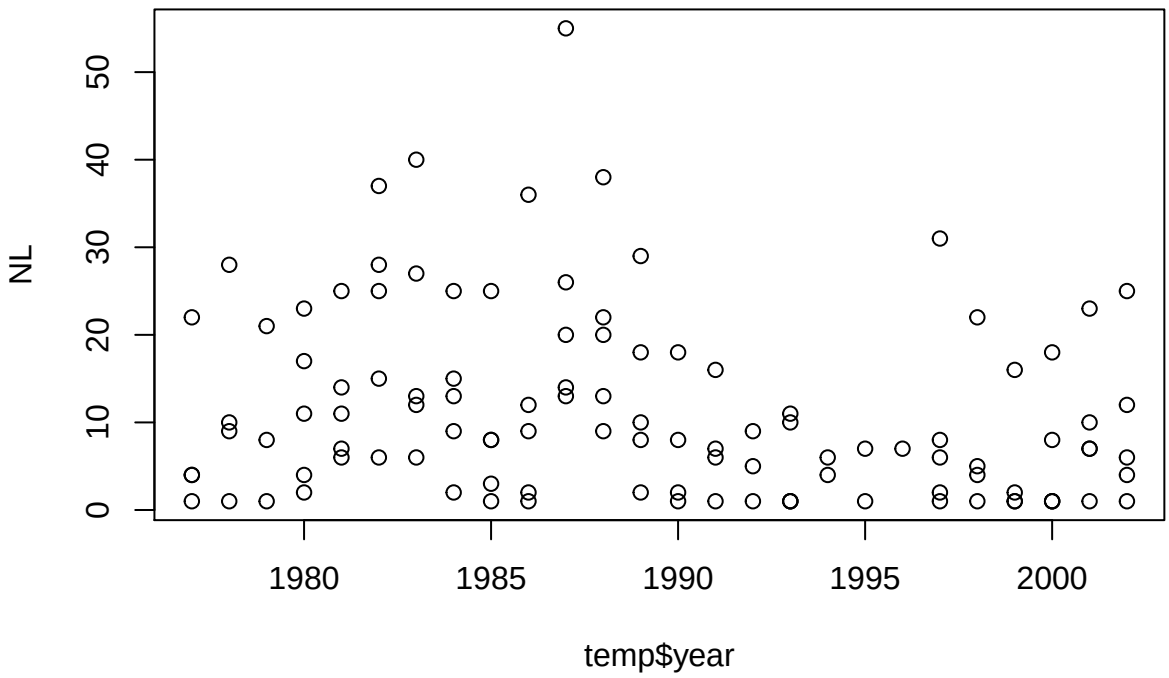
```
## Joining, by = c("year", "plot_type")
## Joining, by = c("year", "plot_type")
```

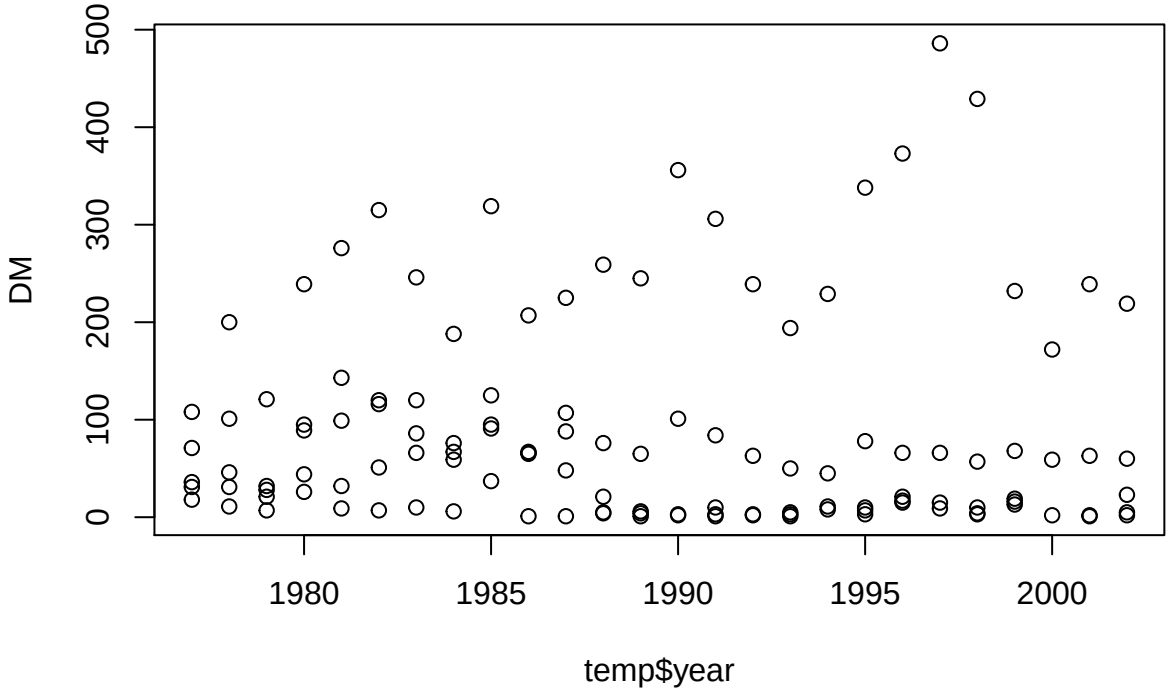
```
portal.turnover <- gather(portal.turnover, key = metric, value = turnover, total, appearance, disappear

#3
TO.plot <- ggplot(portal.turnover, aes(x = year, y = turnover, color = metric)) + geom_line(size = 1, s
```

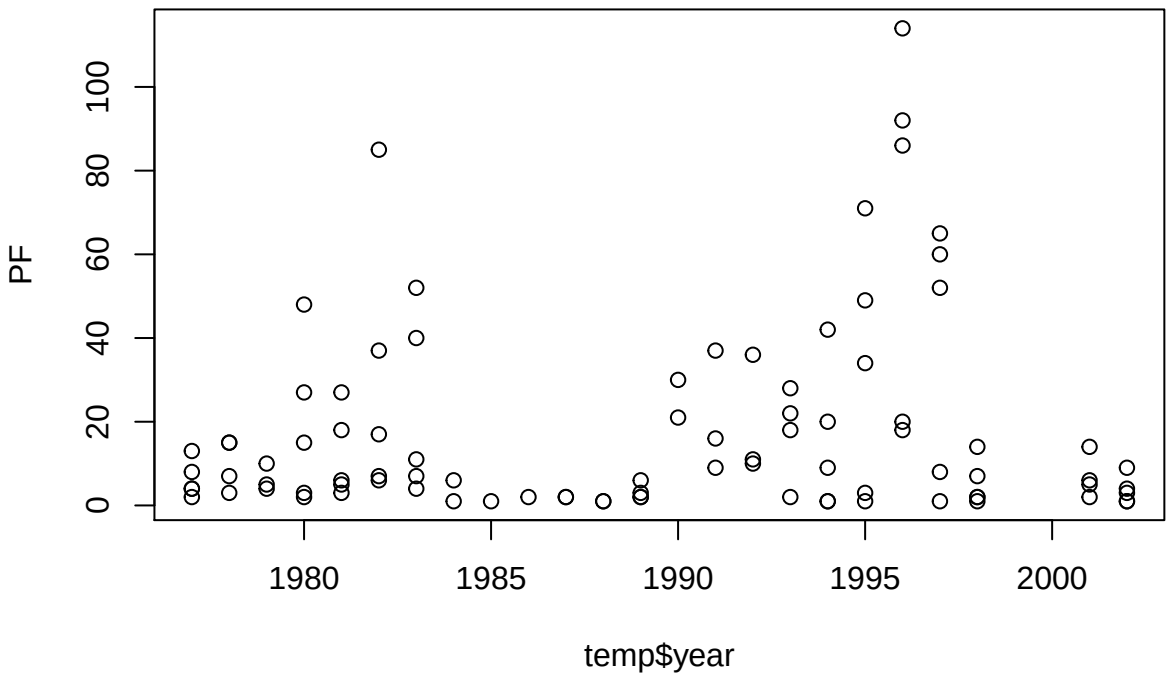
```
plot(TO.plot)
```

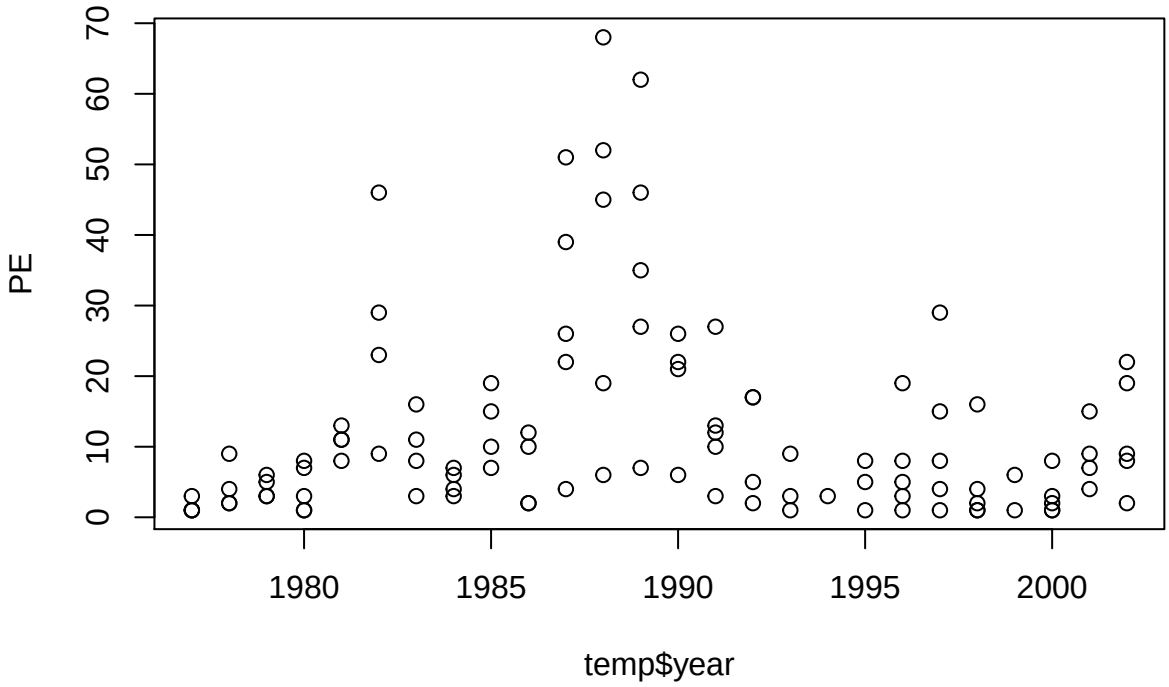


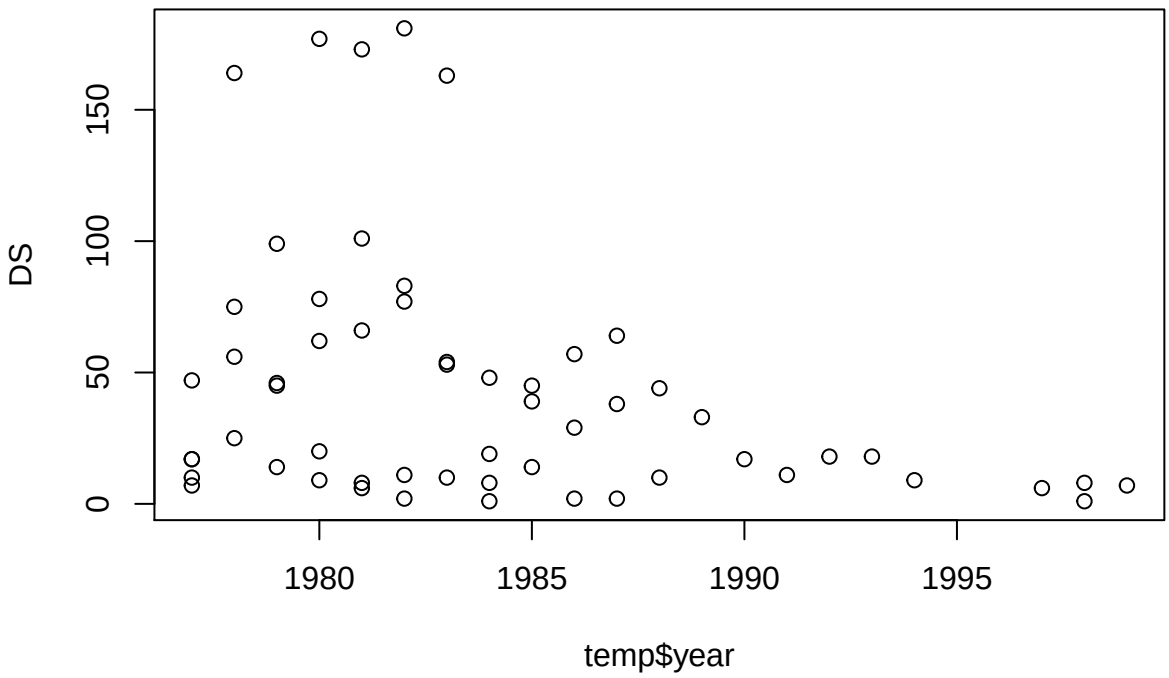

```
myt<-group_by(portal, species_id, year, plot_type) %>% count(species_id)
#for (id in myt$species_id){
#  print(id)
#}

plotea<-function(x = ""){
  for (item in x){
    temp<-filter(myt,species_id==item)
    plot(temp$n~temp$year, ylab=item)
  }
}

plotea((unique(portal$species_id)))
```

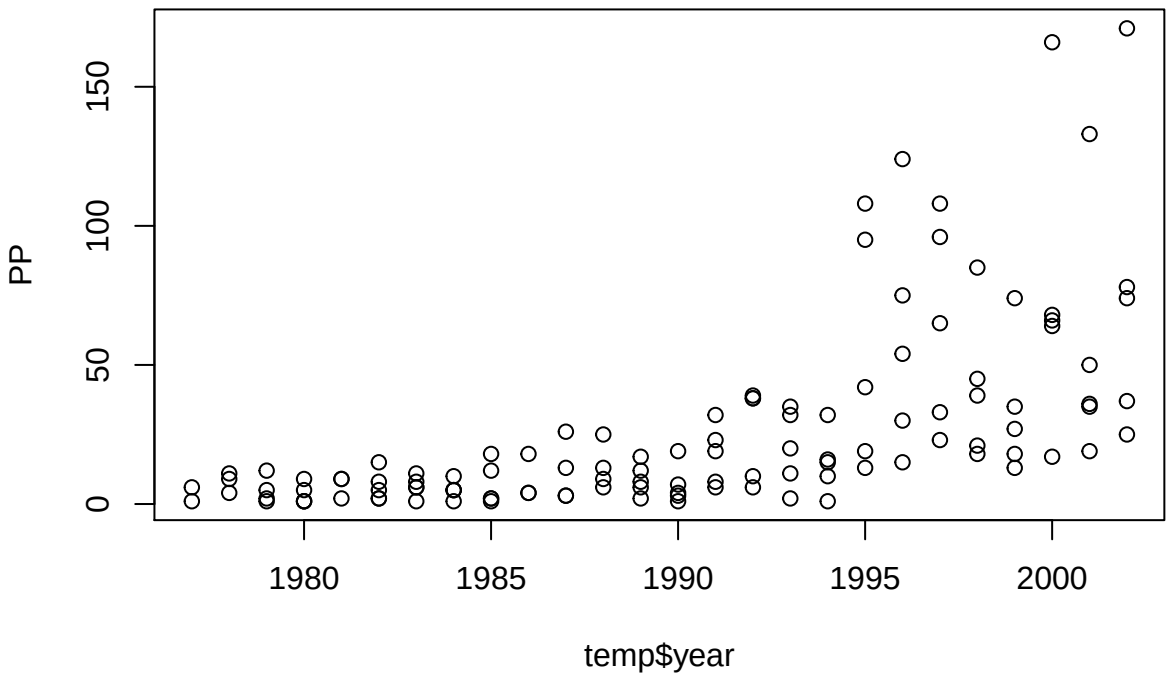

NL
0
10
20
30
40
50
1980
1985
1990
1995
2000
temp$year

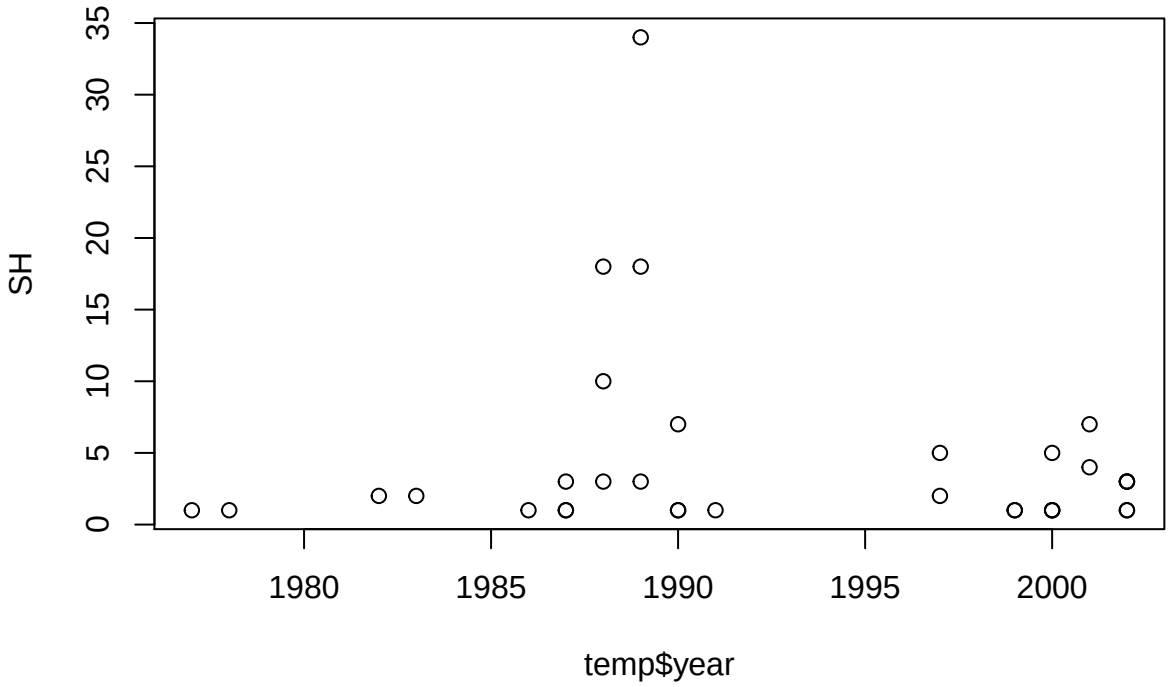

DM
0
100
200
300
400
500
1980
1985
1990
1995
2000
temp$year

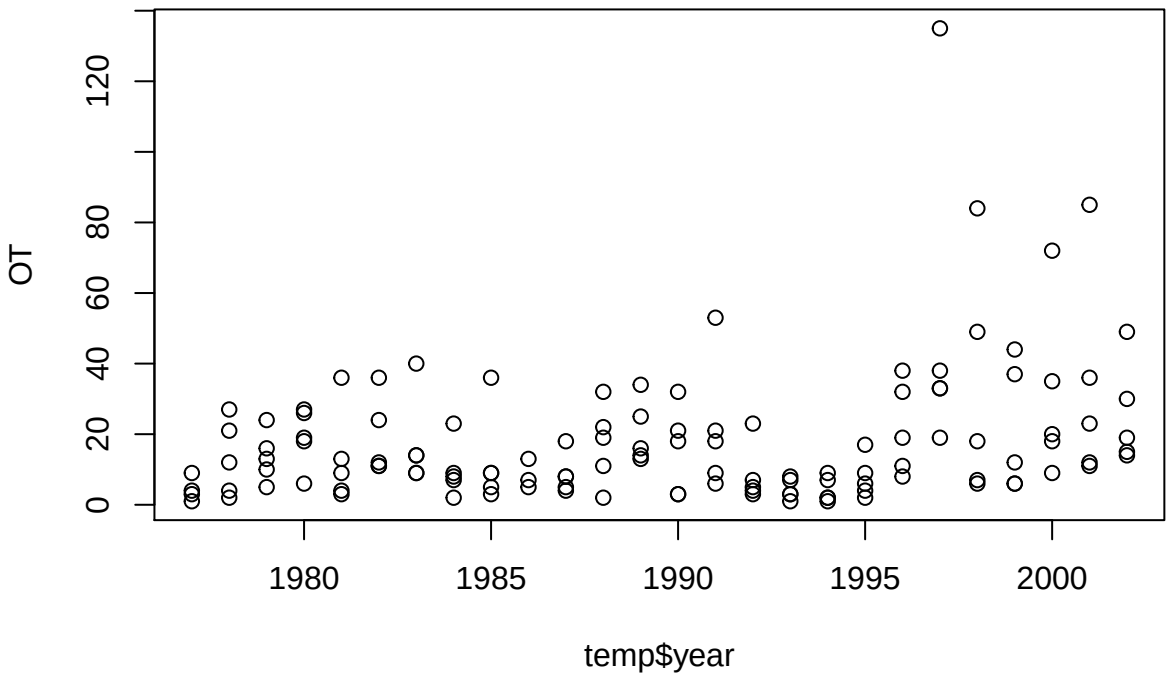

PF
0
20
40
60
80
100
1980
1985
1990
1995
2000
temp$year

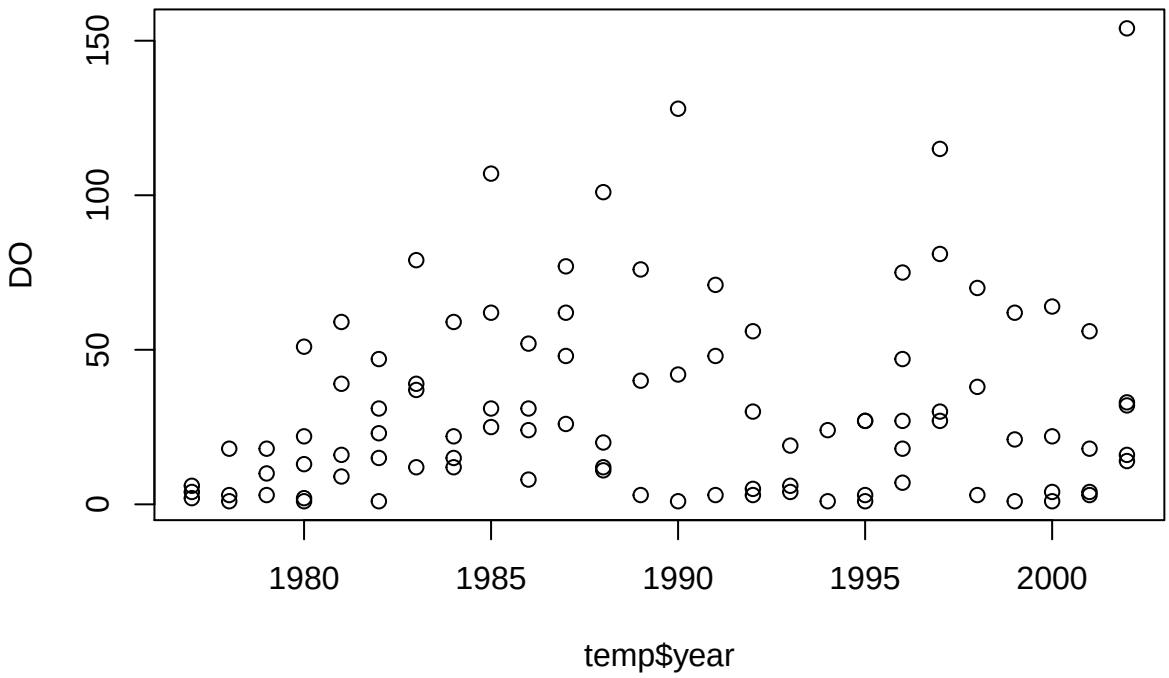

PE
0
10
20
30
40
50
60
70
1980
1985
1990
1995
2000
temp$year

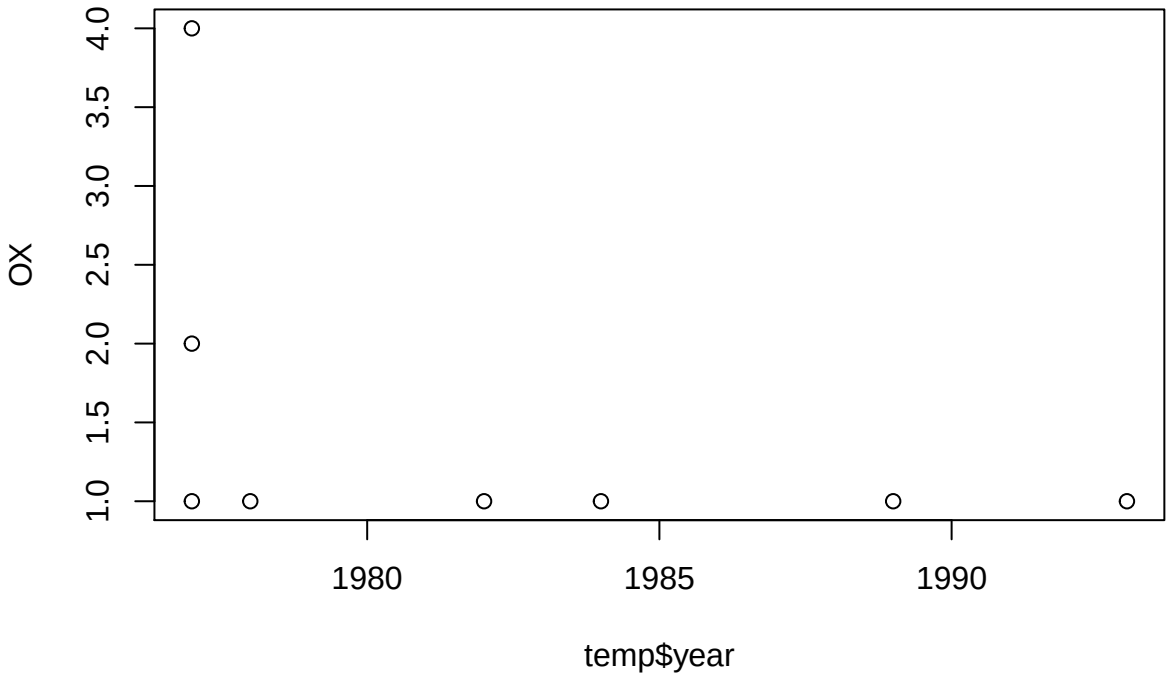

DS
0
50
100
150
1980
1985
1990
1995
temp$year

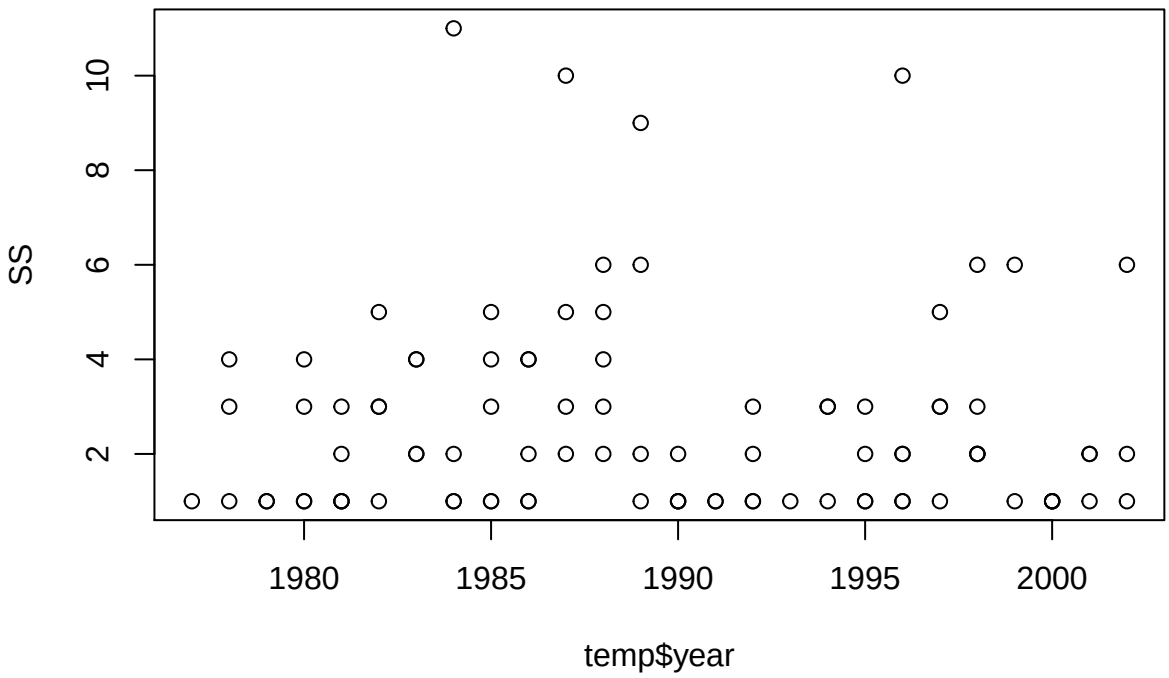

PP
0
50
100
150
1980
1985
1990
1995
2000
temp$year

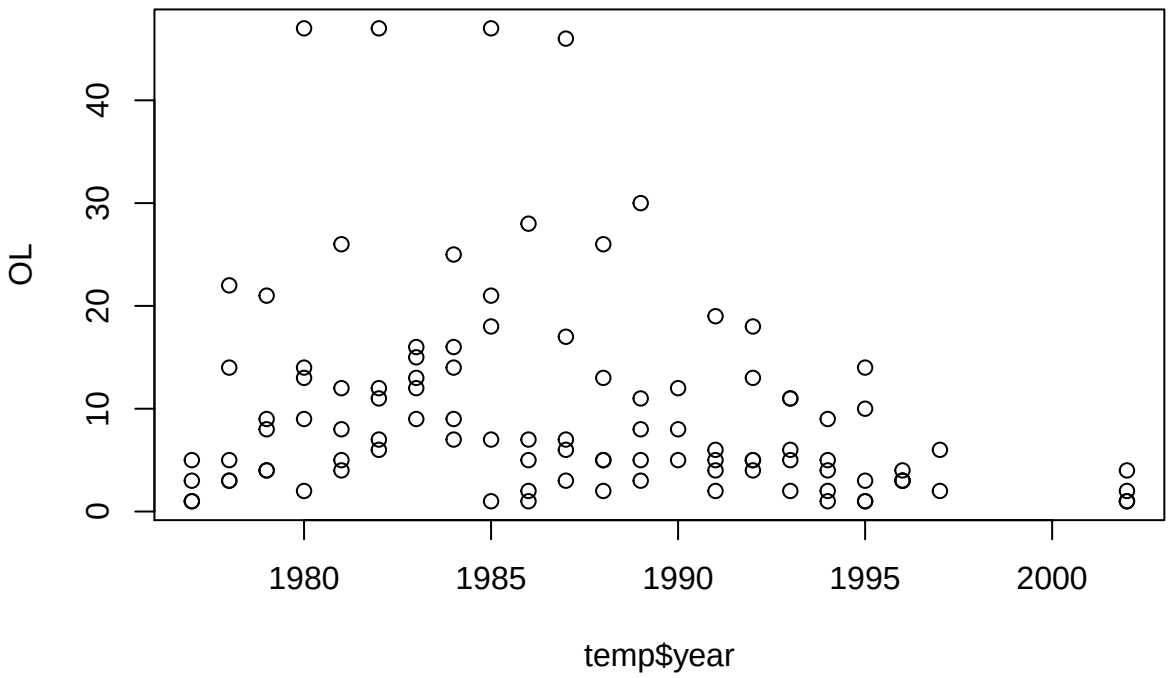

SH
0
5
10
15
20
25
30
35
1980
1985
1990
1995
2000
temp$year

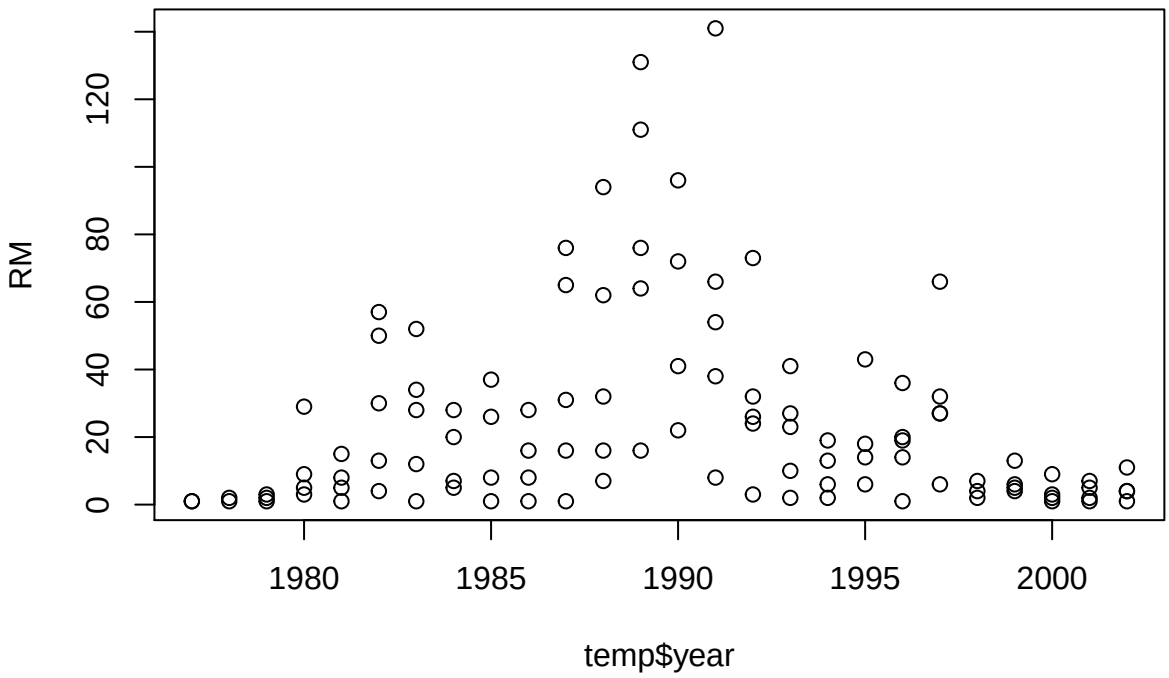

OT
0
20
40
60
80
120
1980
1985
1990
1995
2000
temp$year

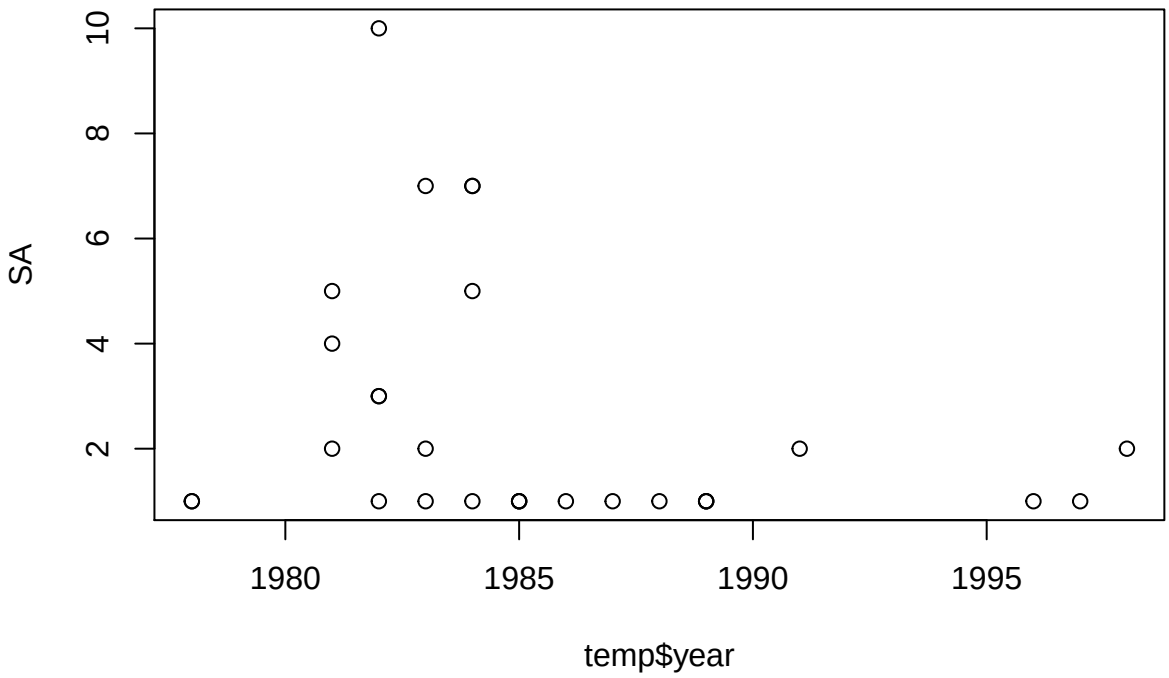

DO
0
50
100
150
1980
1985
1990
1995
2000
temp$year

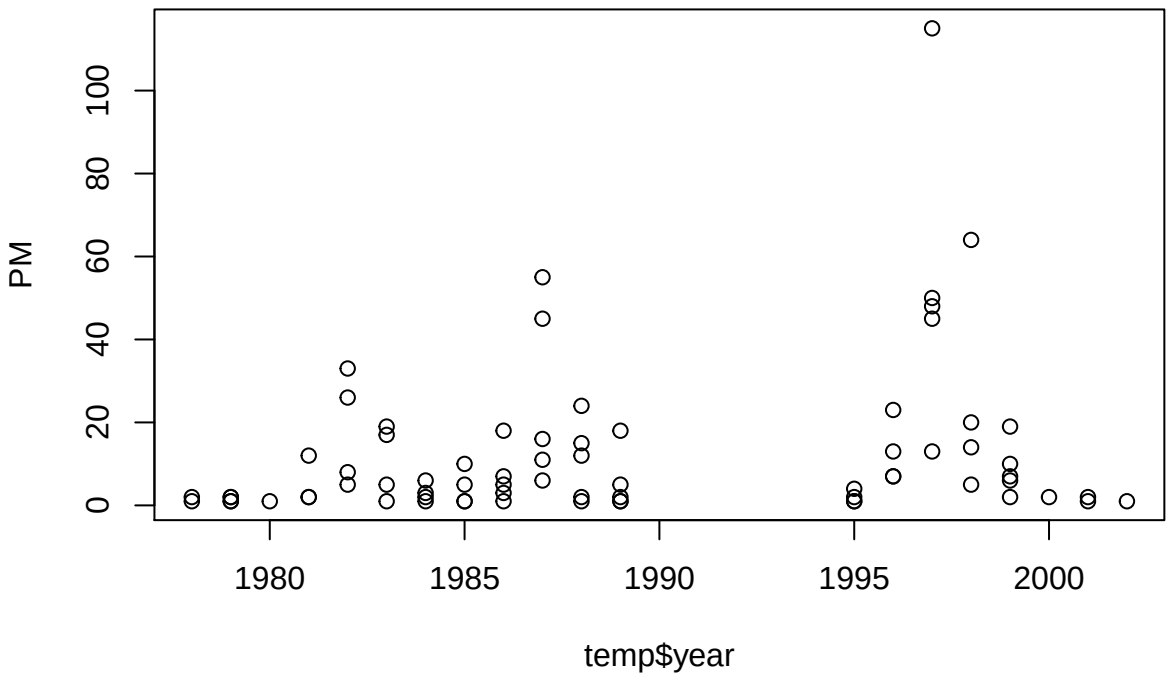

OX
1.0
1.5
2.0
2.5
3.0
3.5
4.0
1980
1985
1990
temp$year

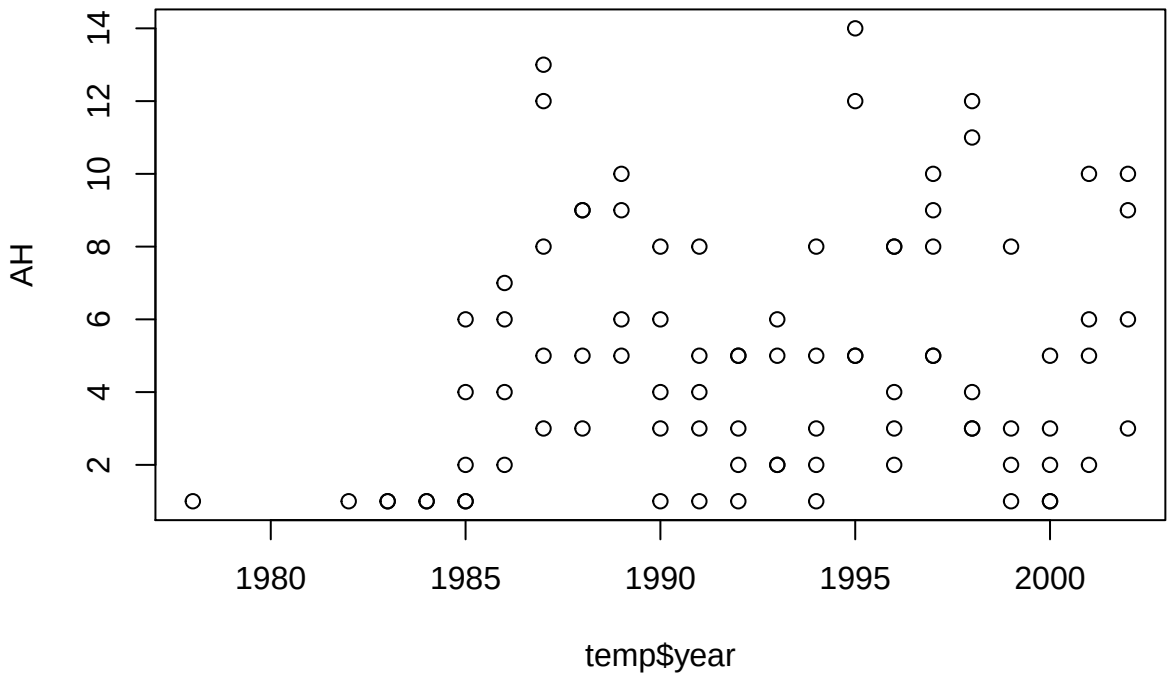

SS
2
4
6
8
10
1980
1985
1990
1995
2000
temp$year

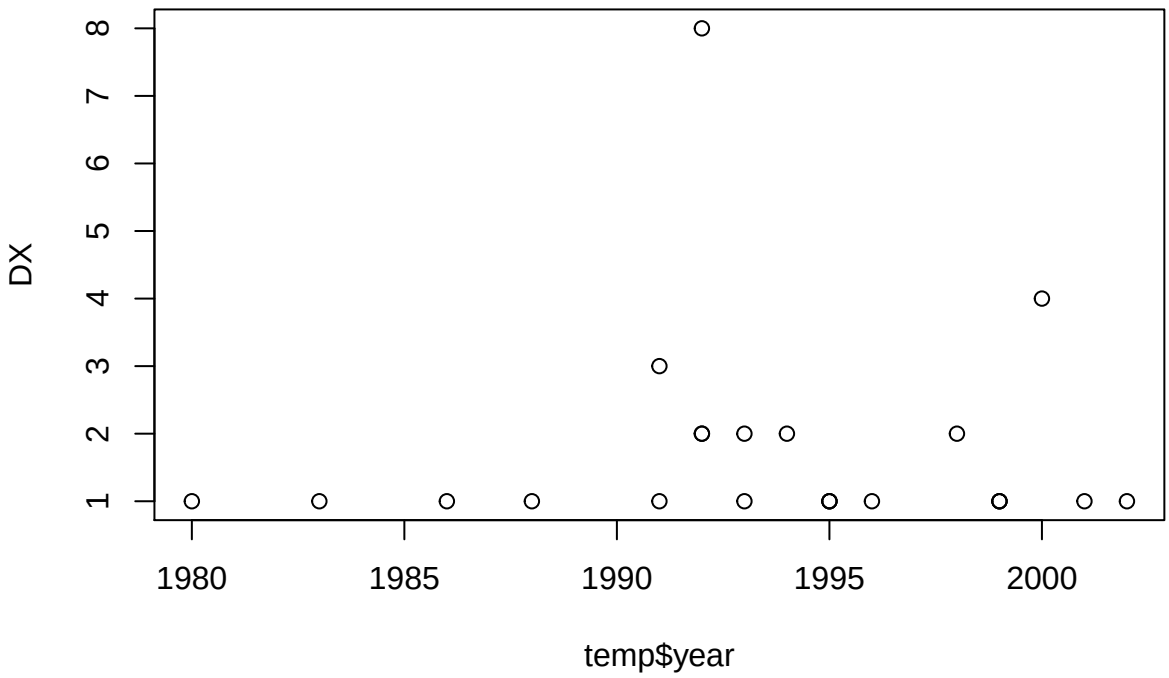

OL
0
10
20
30
40
1980
1985
1990
1995
2000
temp$year

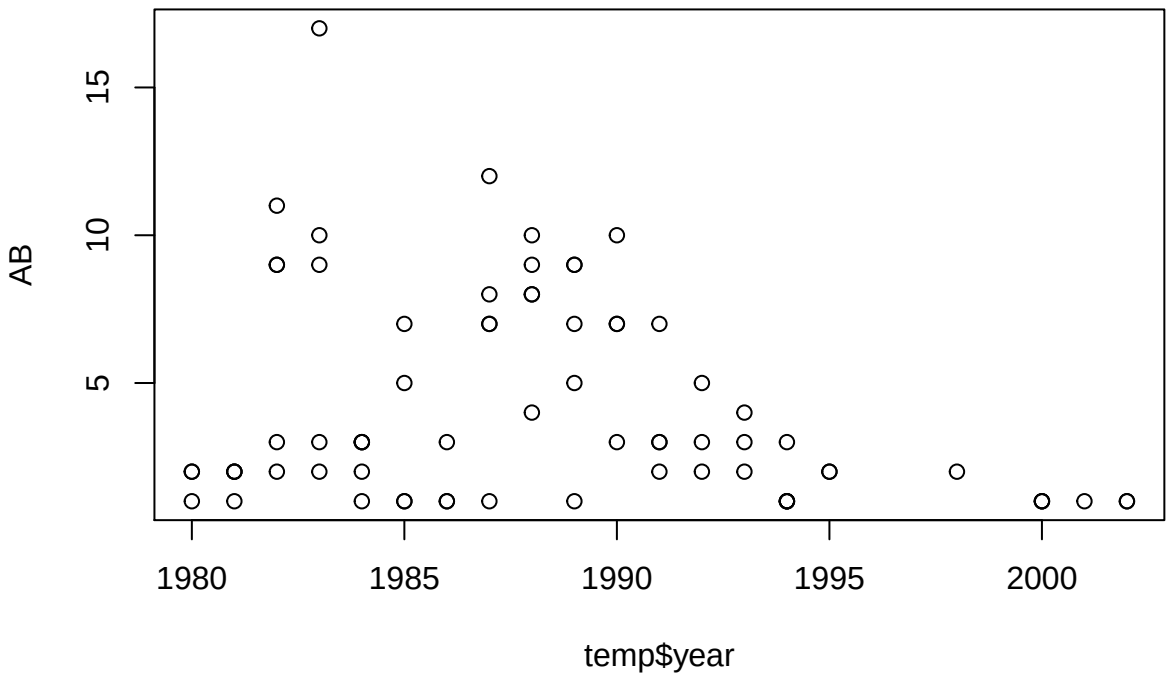

RM
0
20
40
60
80
120
1980
1985
1990
1995
2000
temp$year

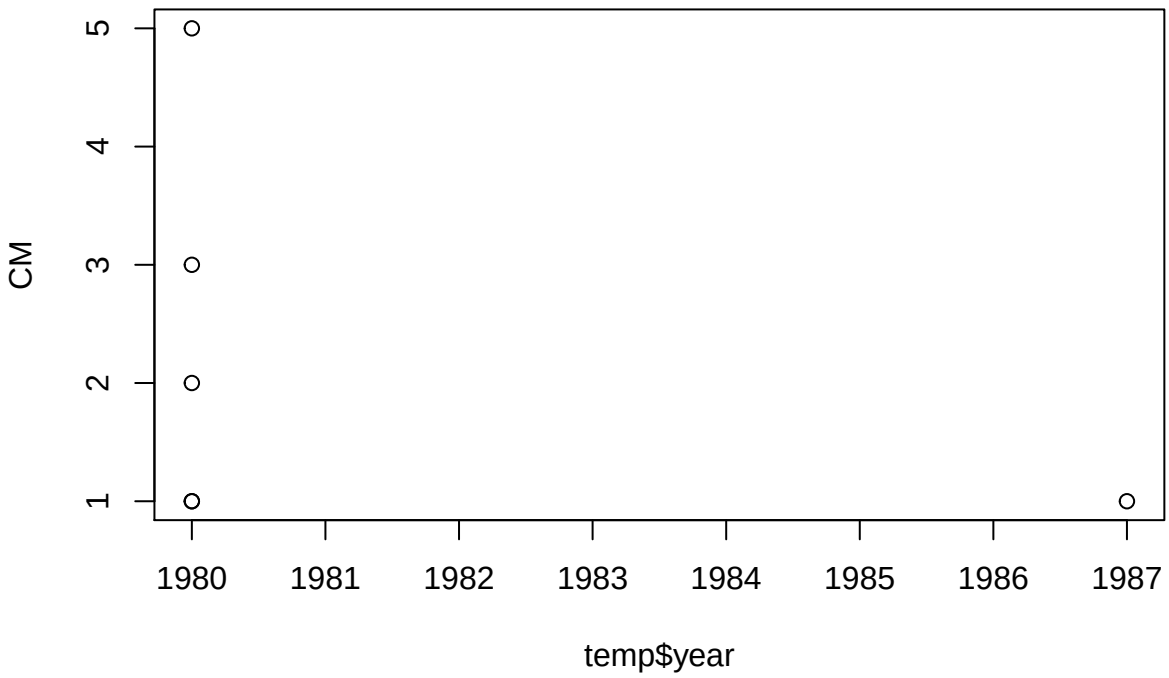

SA
2
4
6
8
10
1980
1985
1990
1995
temp$year

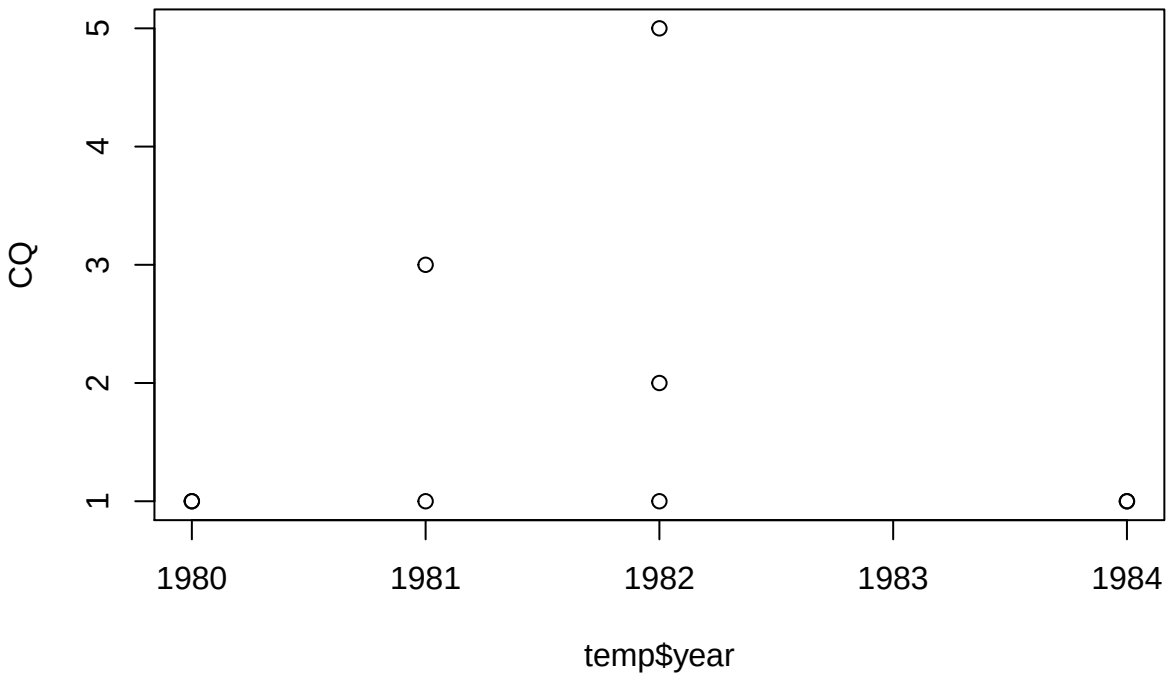

PM
0
20
40
60
80
100
1980
1985
1990
1995
2000
temp$year

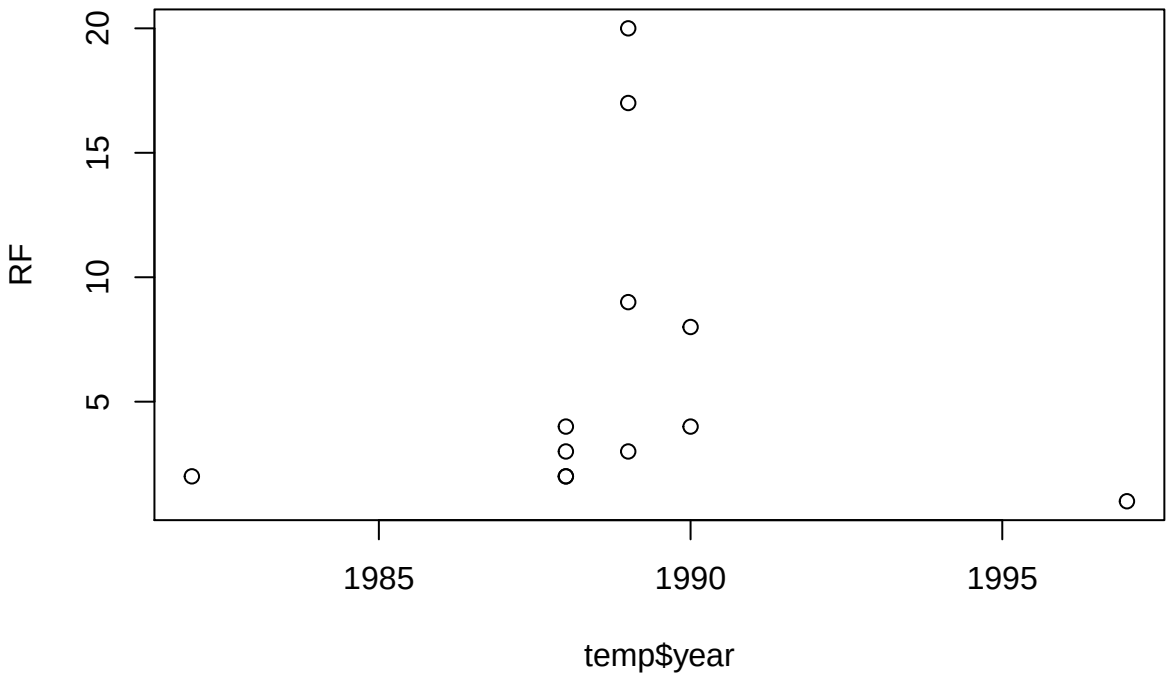

AH
2
4
6
8
10
12
14
1980
1985
1990
1995
2000
temp$year

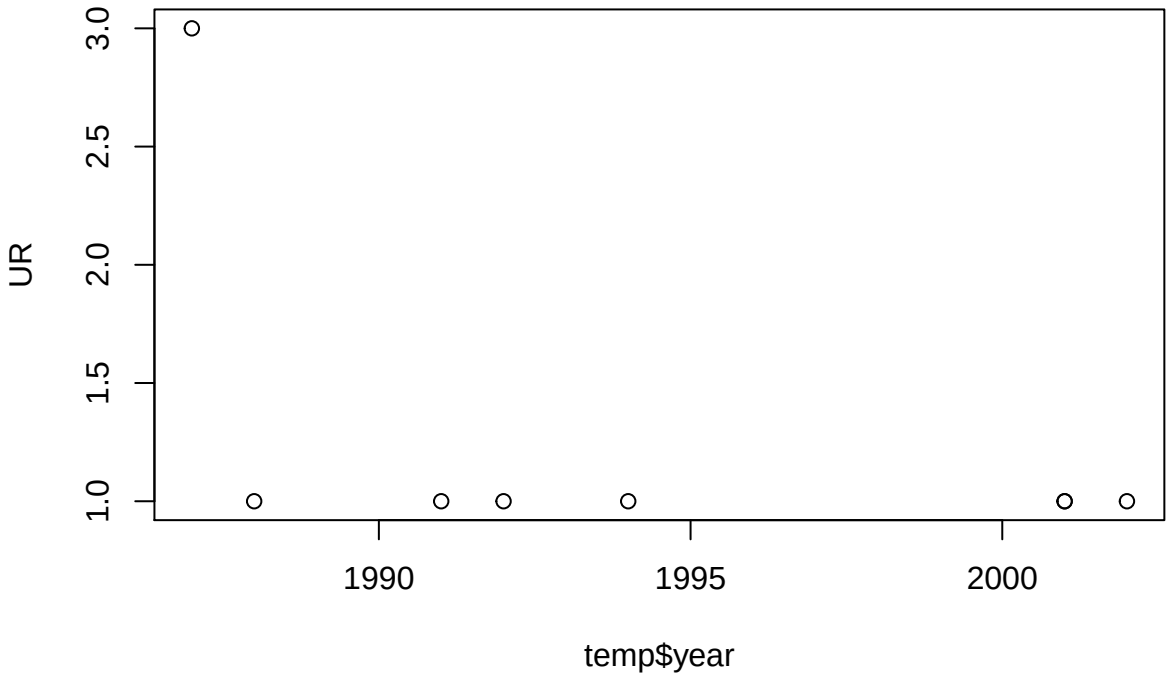

DX
1
2
3
4
5
6
7
8
1980
1985
1990
1995
2000
temp$year

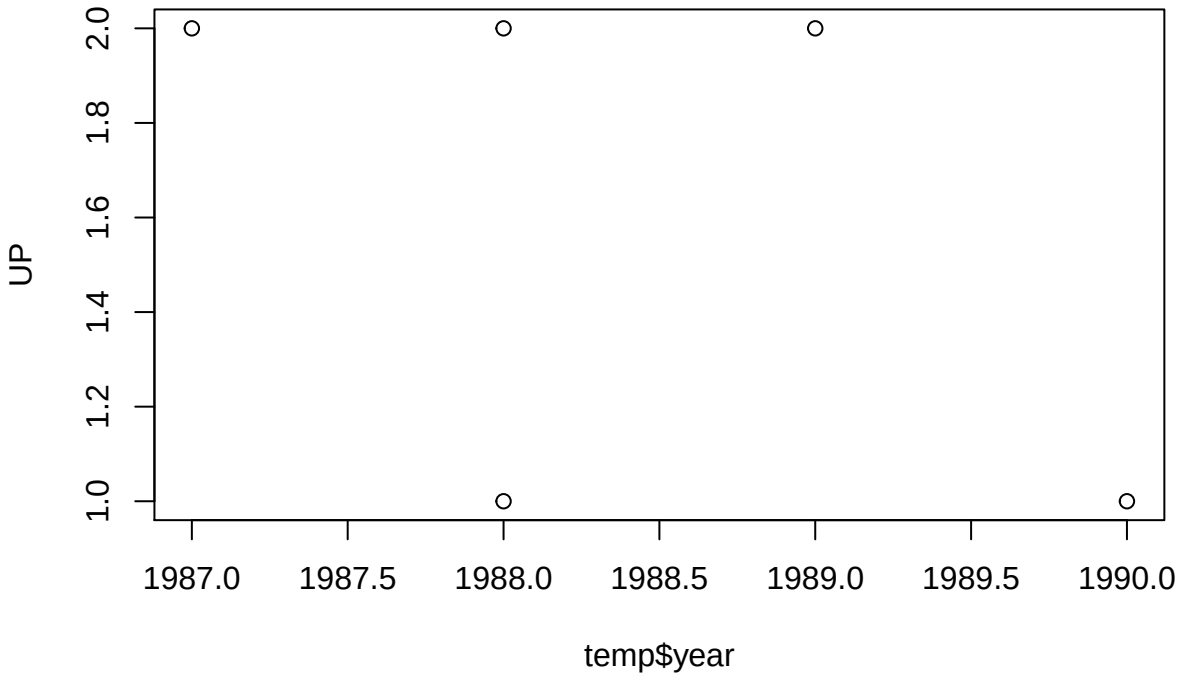

AB
5
10
15
1980
1985
1990
1995
2000
temp$year

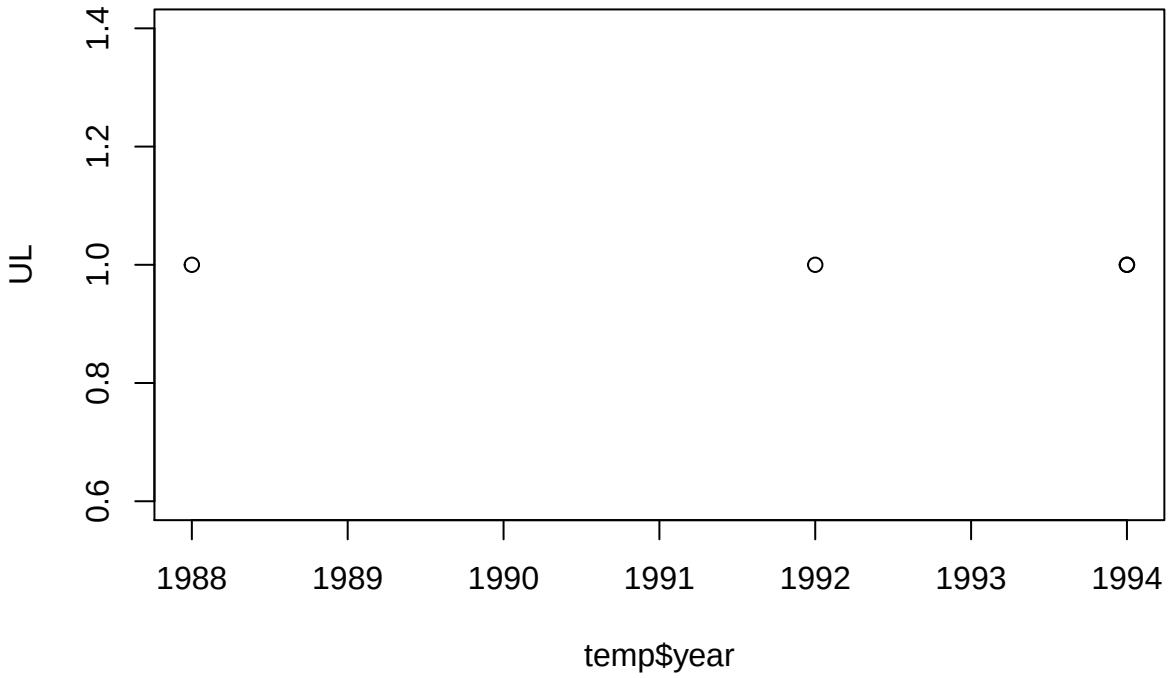

CM
1
2
3
4
5
1980
1981
1982
1983
1984
1985
1986
1987
temp$year

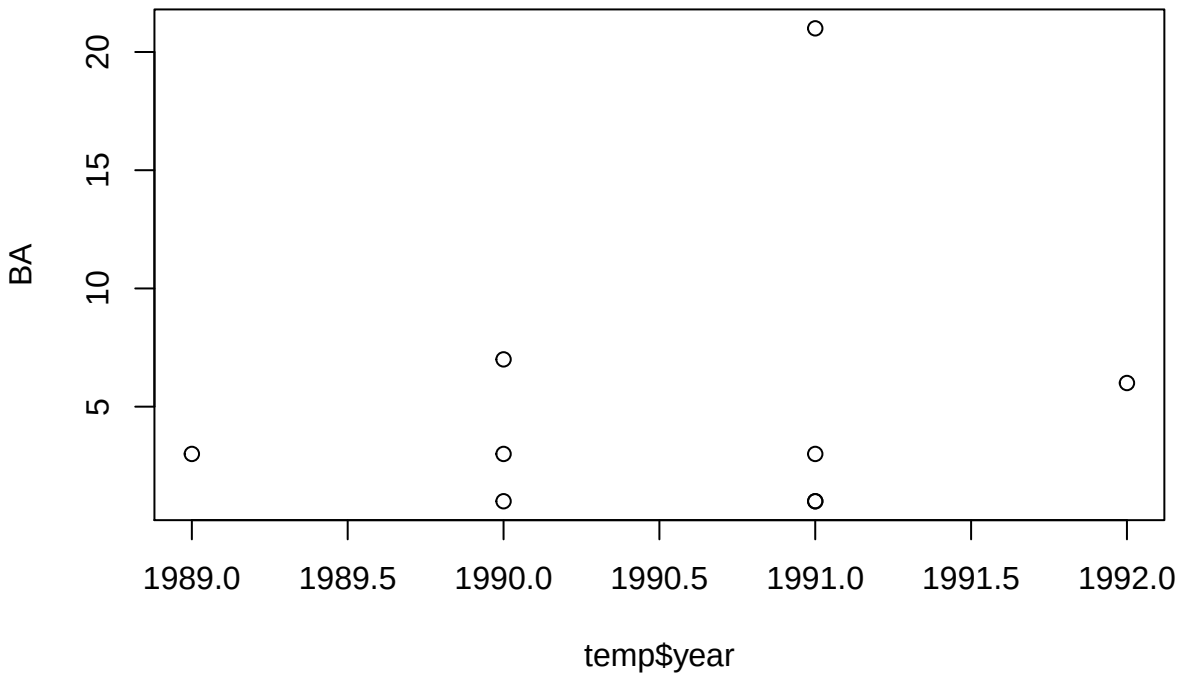

CQ
1
2
3
4
5
1980
1981
1982
1983
1984
temp$year

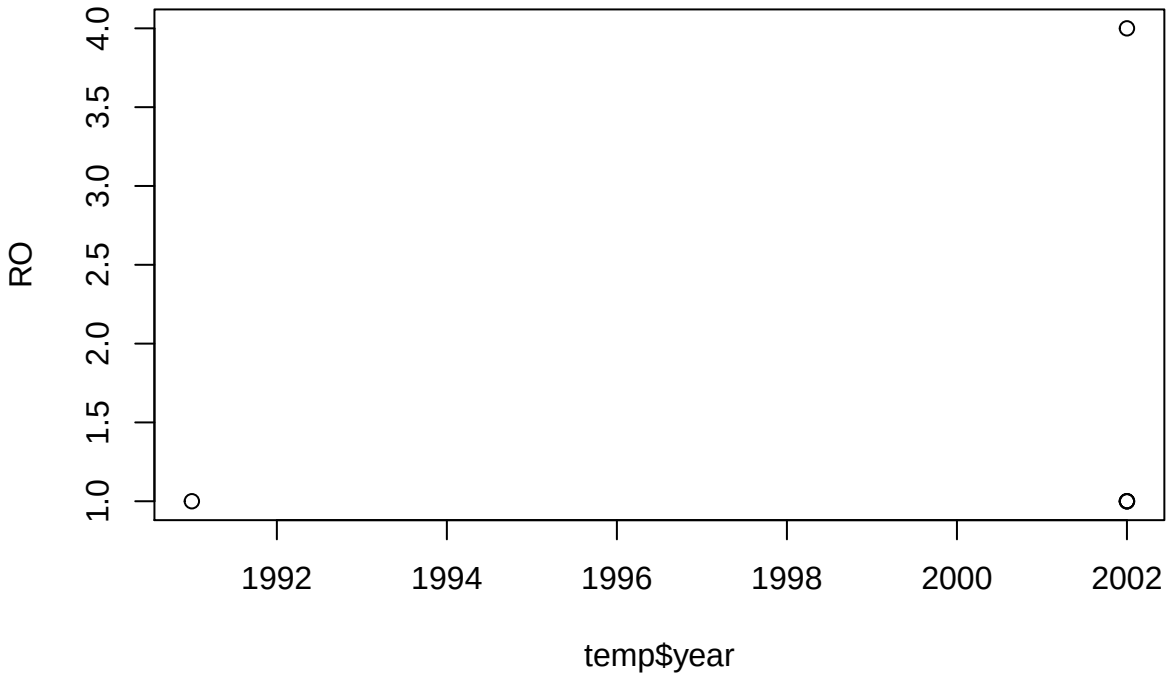

RF
5
10
15
20
1985
1990
1995
temp$year

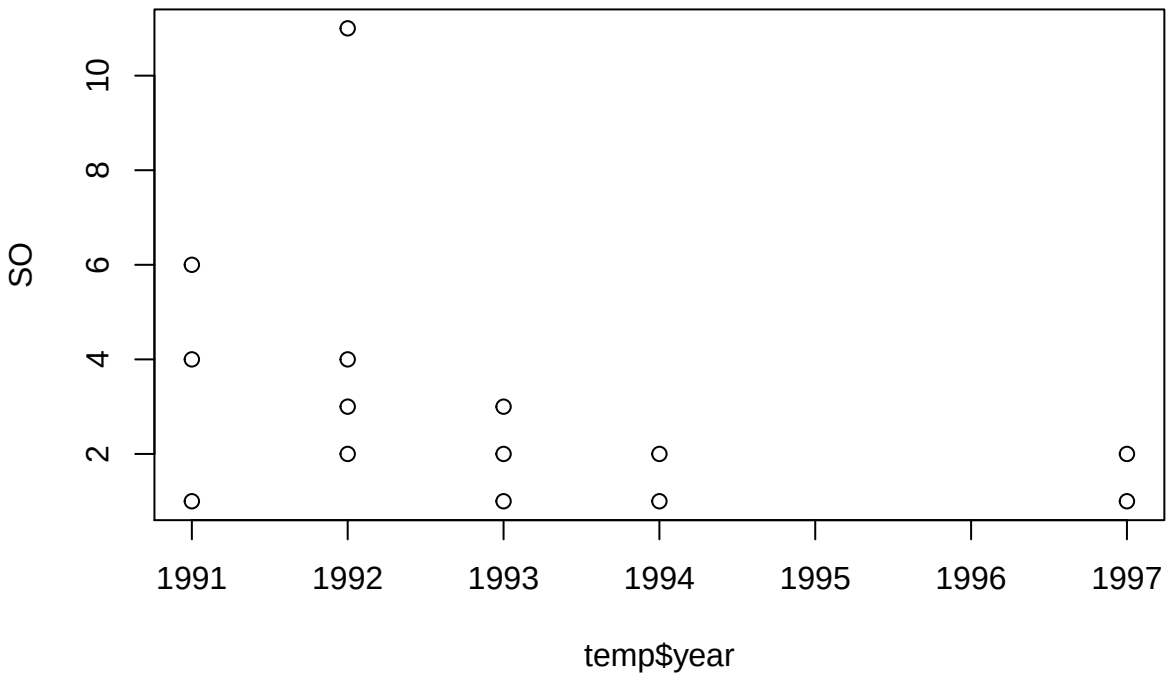

UR
3.0
2.5
2.0
1.5
1.0
1990
1995
2000
temp$year

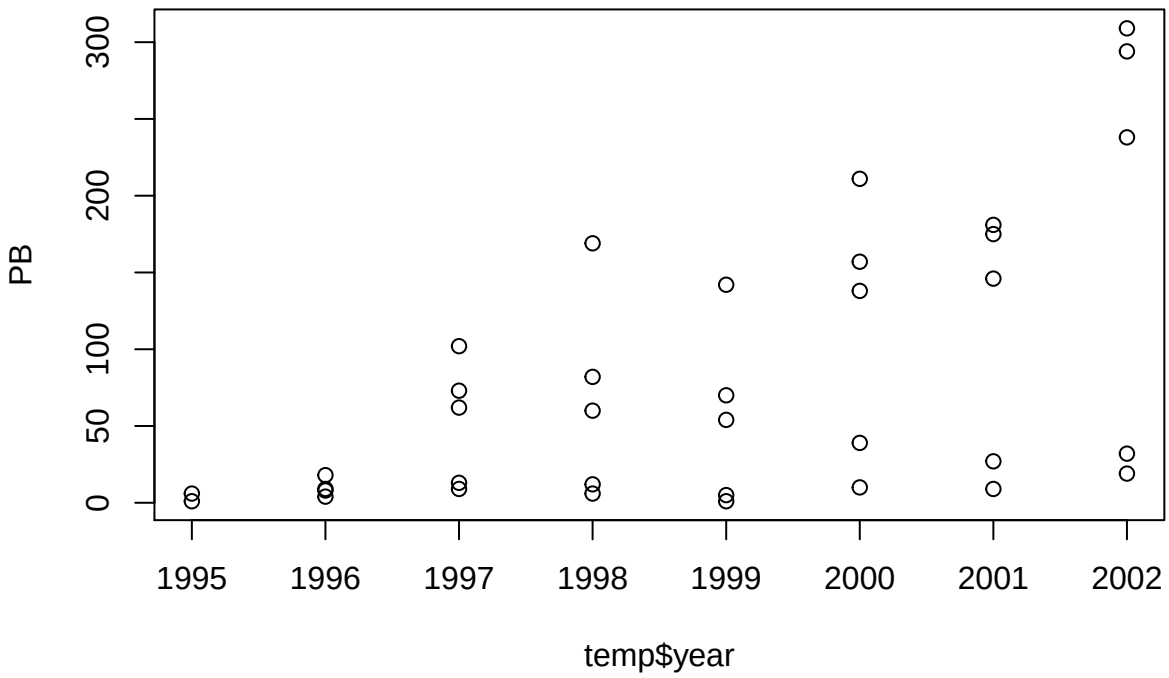

UP
1.0
1.2
1.4
1.6
1.8
2.0
1987.0
1987.5
1988.0
1988.5
1989.0
1989.5
1990.0
temp$year

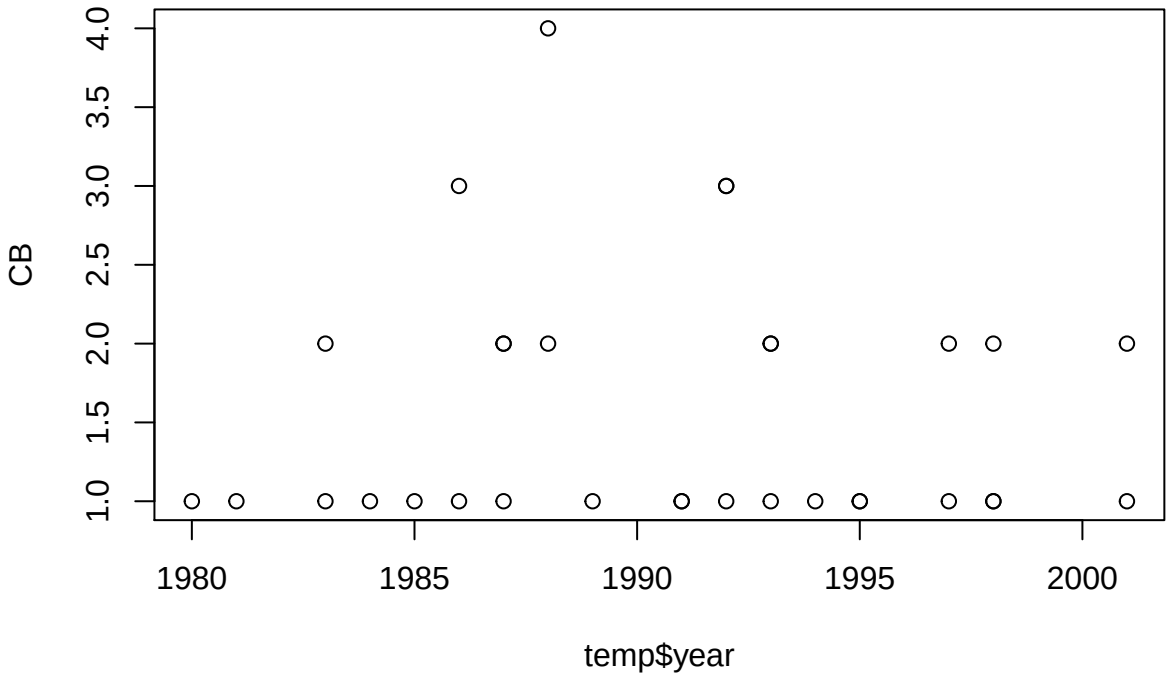

UL
0.6
0.8
1.0
1.2
1.4
1988
1989
1990
1991
1992
1993
1994
temp$year

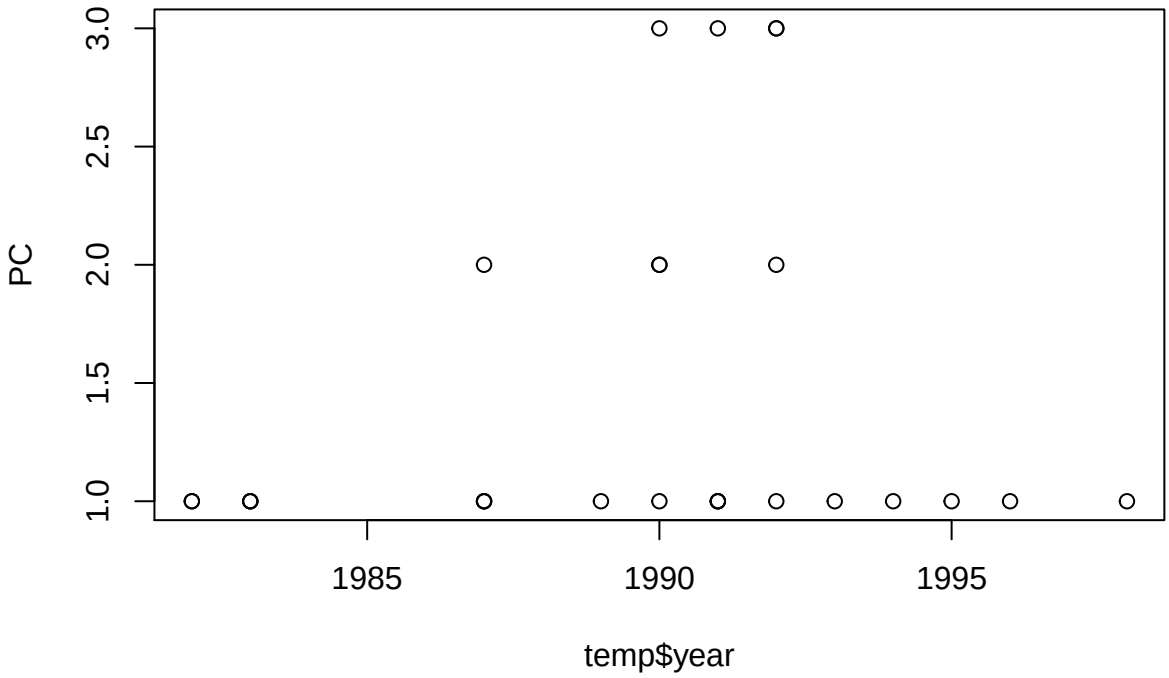

BA
5
10
15
20
1989.0
1989.5
1990.0
1990.5
1991.0
1991.5
1992.0
temp$year

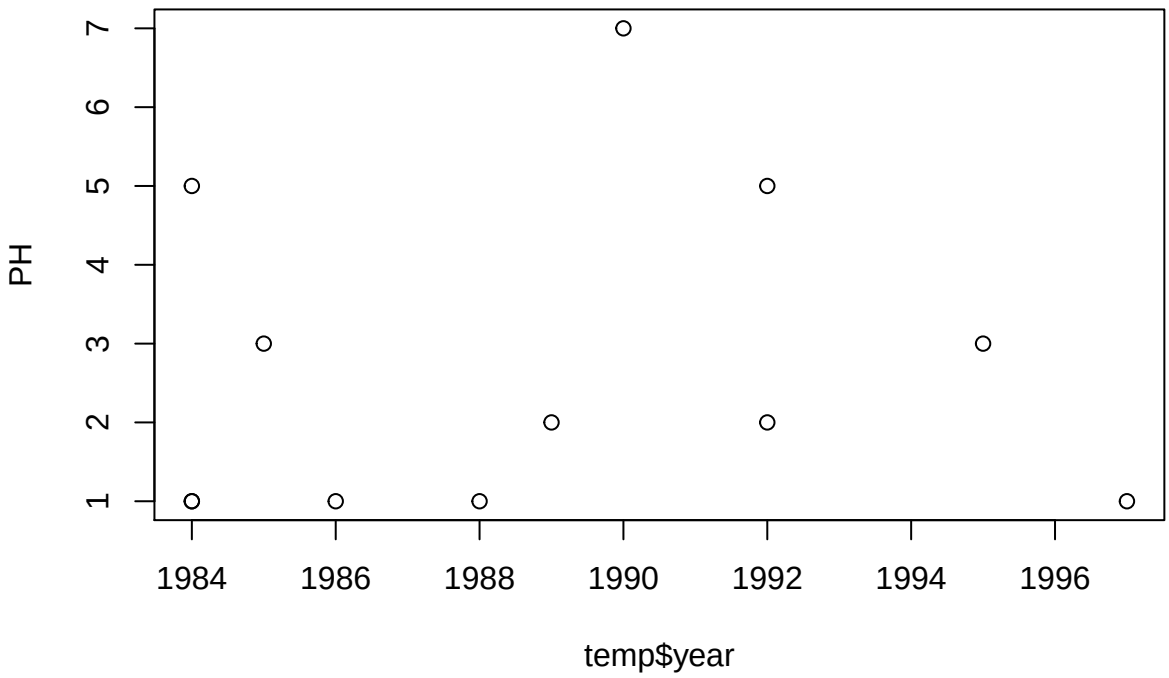

RO
1.0
1.5
2.0
2.5
3.0
3.5
4.0
1992
1994
1996
1998
2000
2002
temp$year

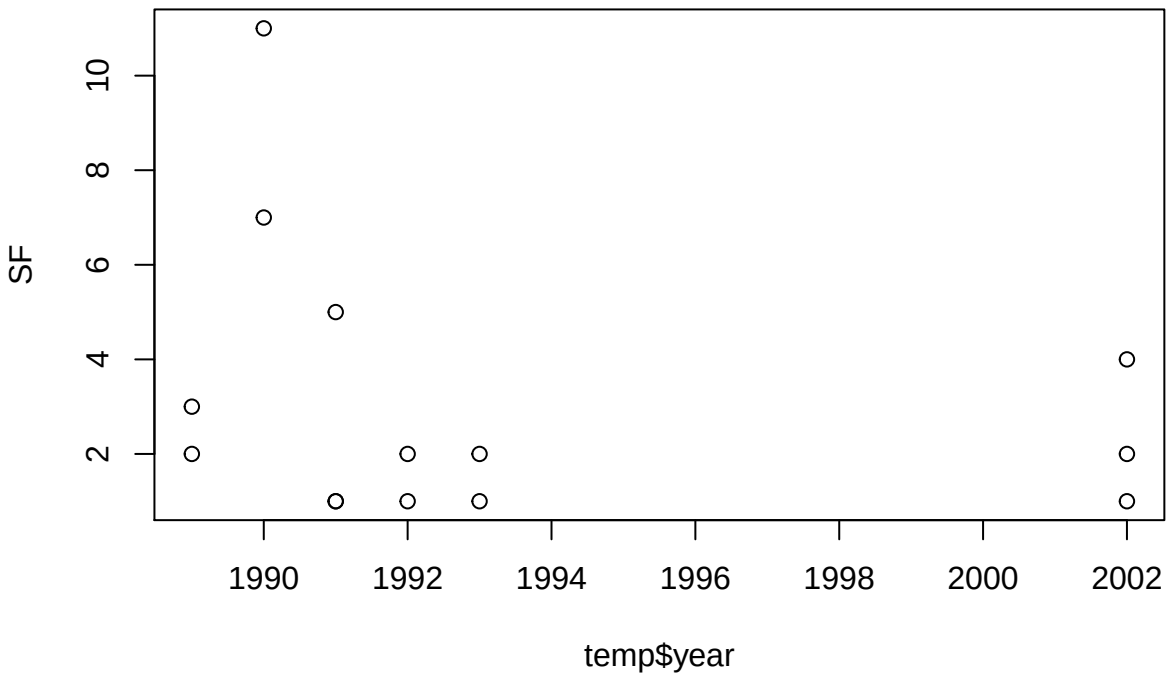

SO
2
4
6
8
10
1991
1992
1993
1994
1995
1996
1997
temp$year

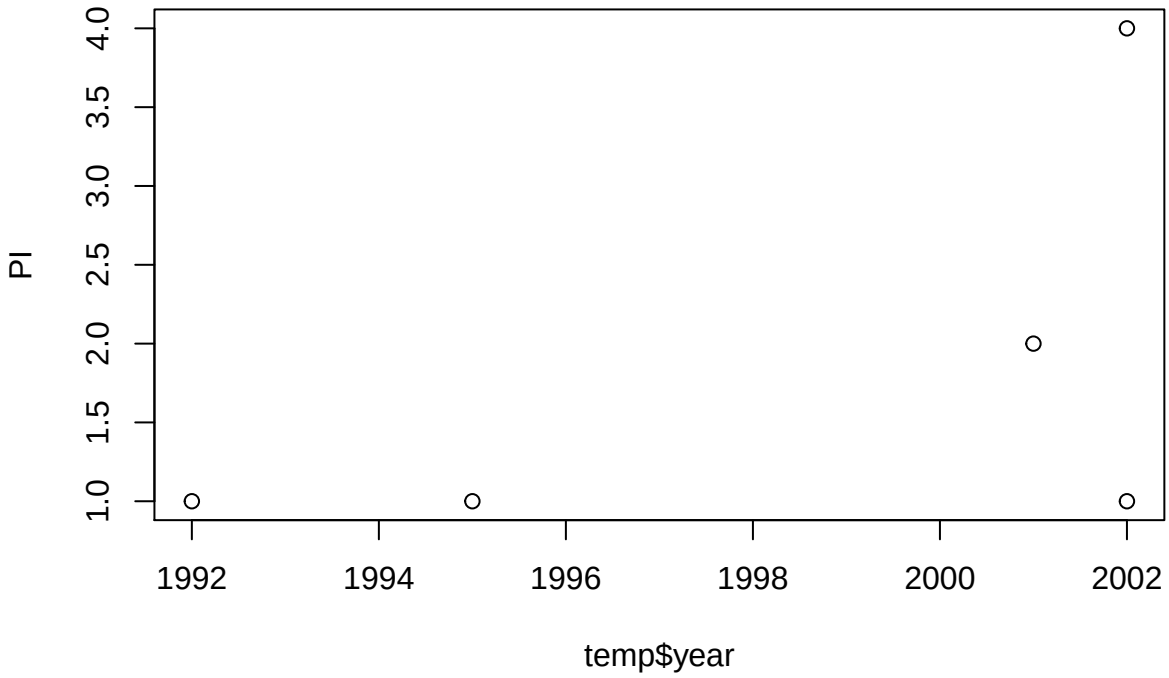

PB
0
50
100
200
300
1995
1996
1997
1998
1999
2000
2001
2002
temp$year

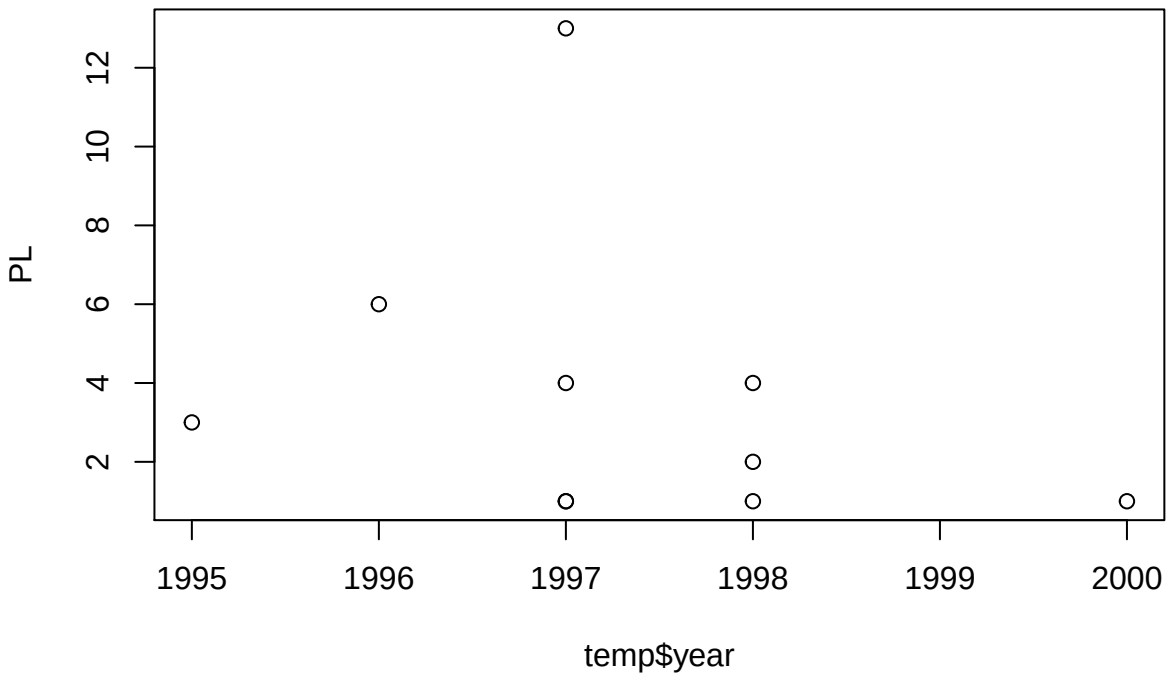

CB
1.0
1.5
2.0
2.5
3.0
3.5
4.0
1980
1985
1990
1995
2000
temp$year


PC
1.0
1.5
2.0
2.5
3.0
1985
1990
1995
temp$year

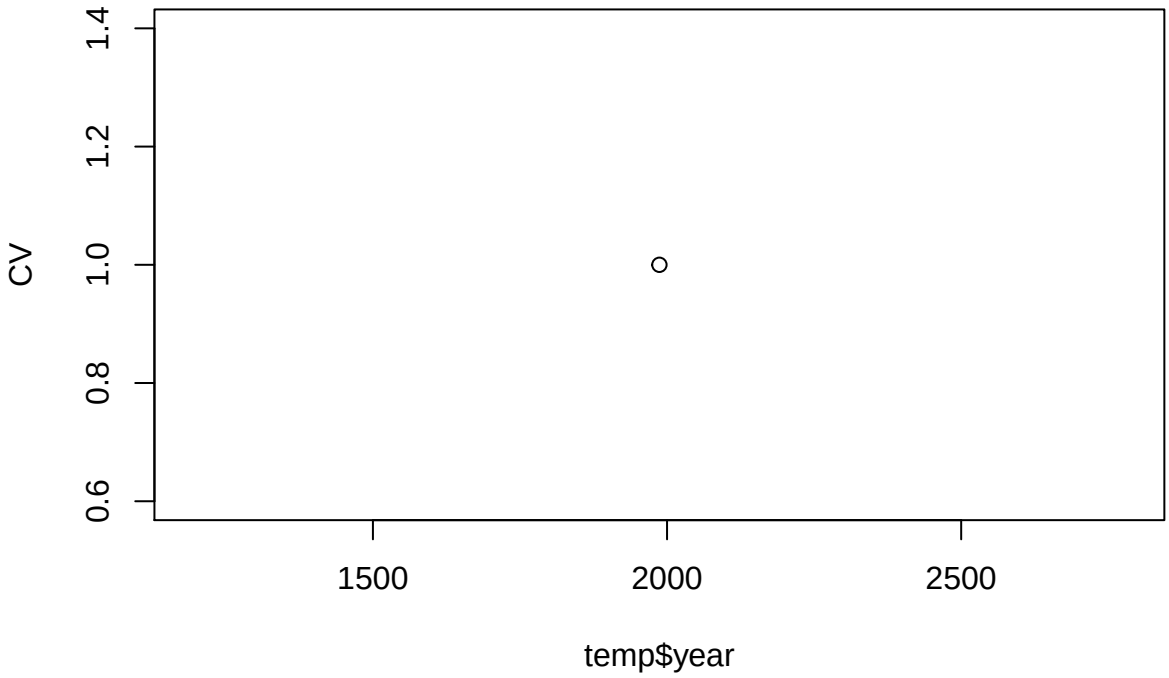

PH
1
2
3
4
5
6
7
1984
1986
1988
1990
1992
1994
1996
temp$year

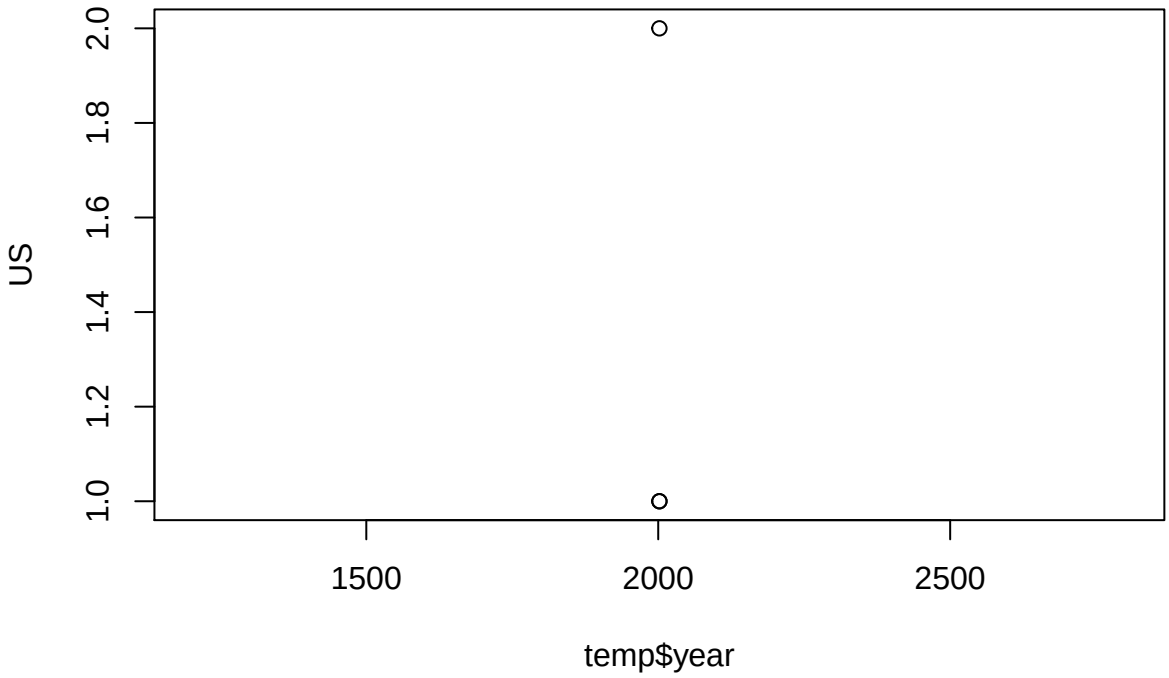

SF
2
4
6
8
10
1990
1992
1994
1996
1998
2000
2002
temp$year

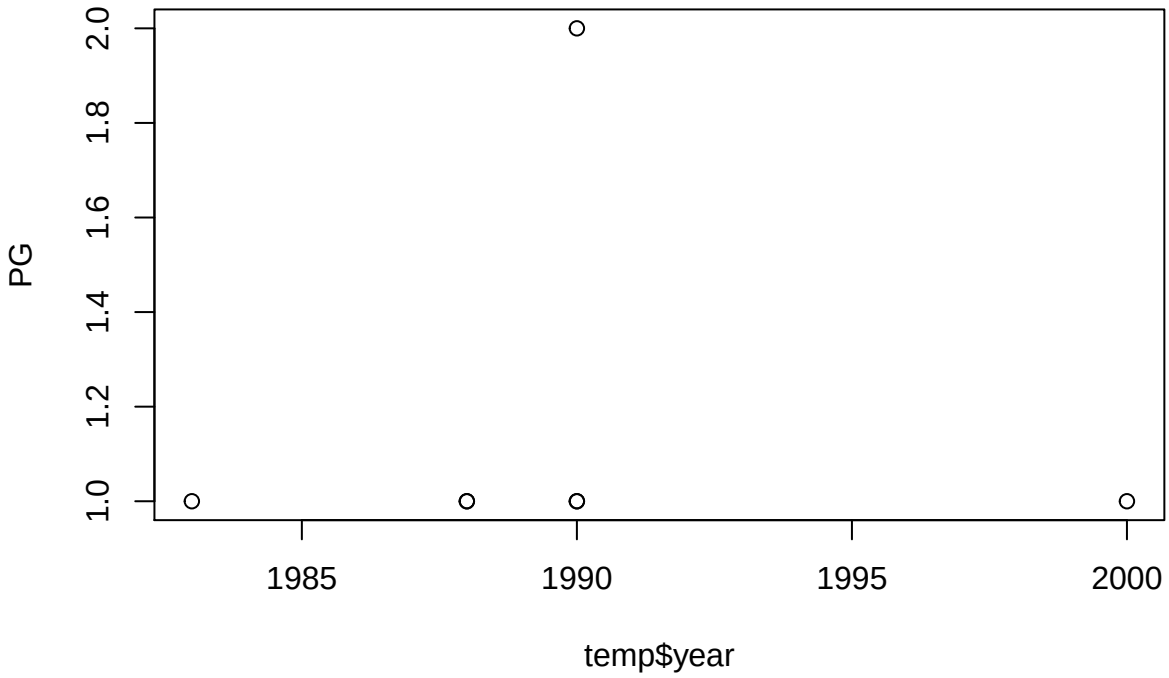

PI
1.0
1.5
2.0
2.5
3.0
3.5
4.0
1992
1994
1996
1998
2000
2002
temp$year

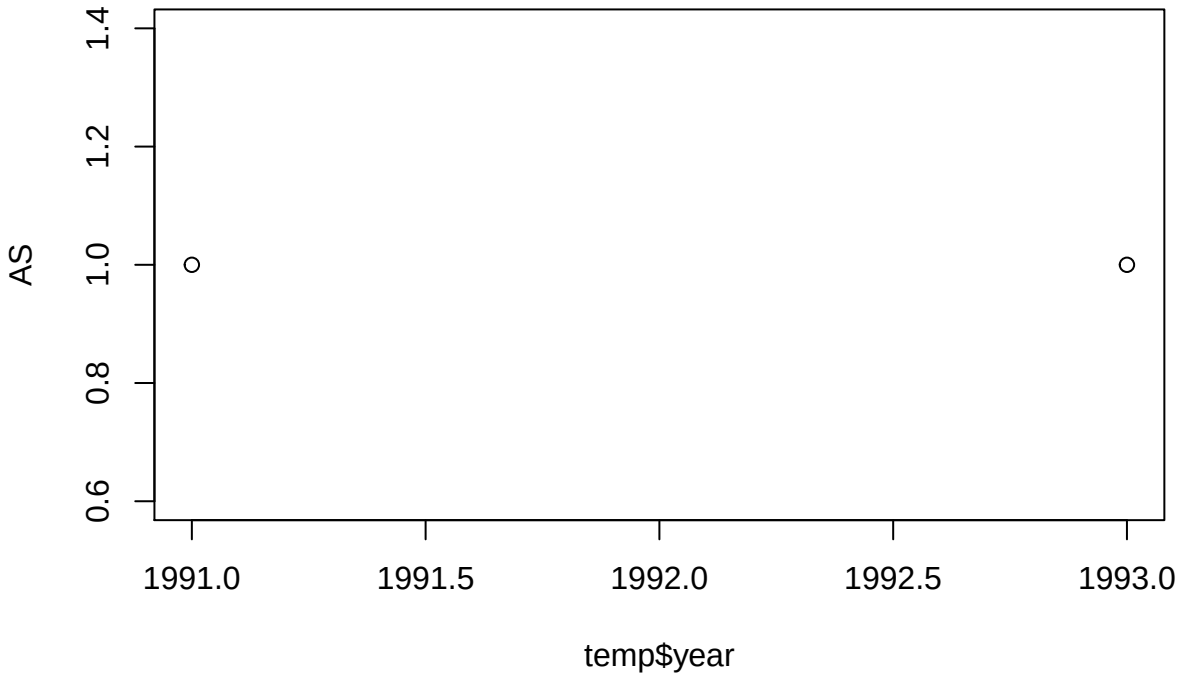

PL
2
4
6
8
10
12
1995
1996
1997
1998
1999
2000
temp$year

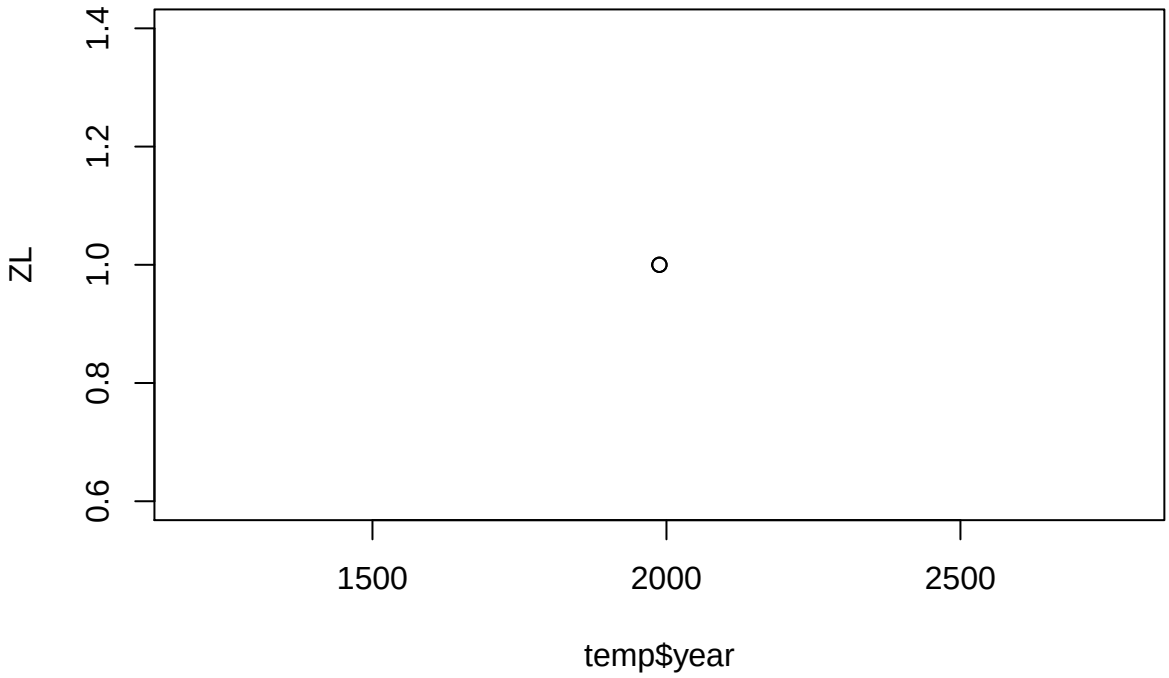

PX
0.6
0.8
1.0
1.2
1.4
1996
1997
1998
1999
2000
temp$year

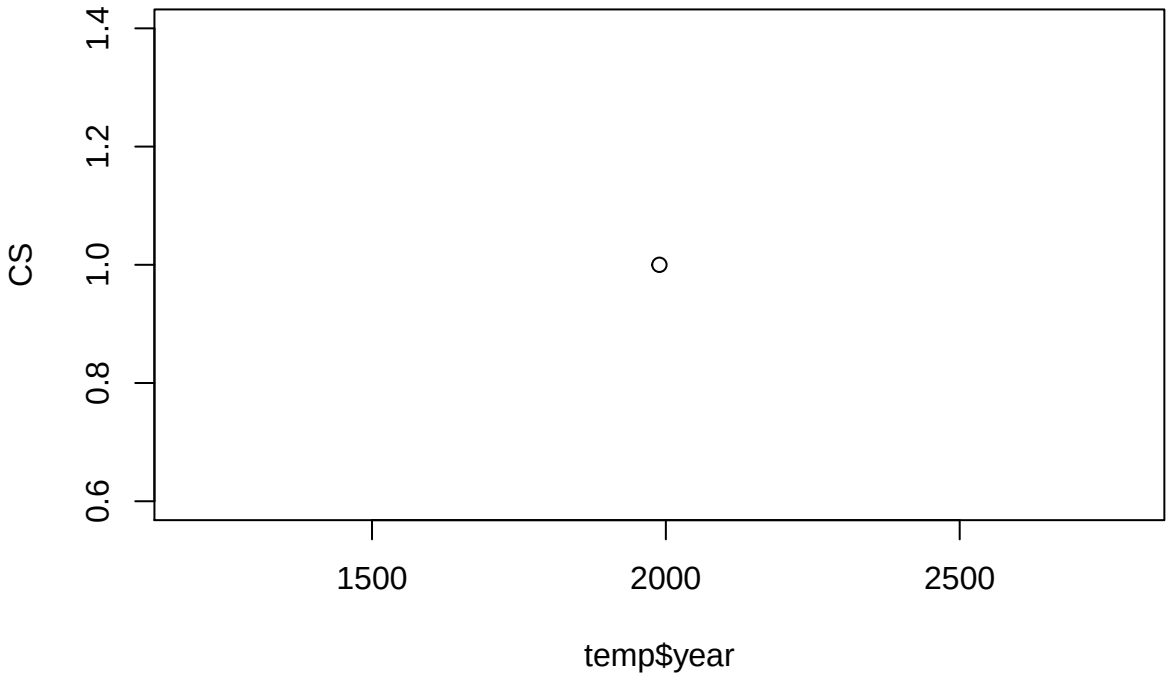

CV
0.6
0.8
1.0
1.2
1.4
1500
2000
2500
temp$year

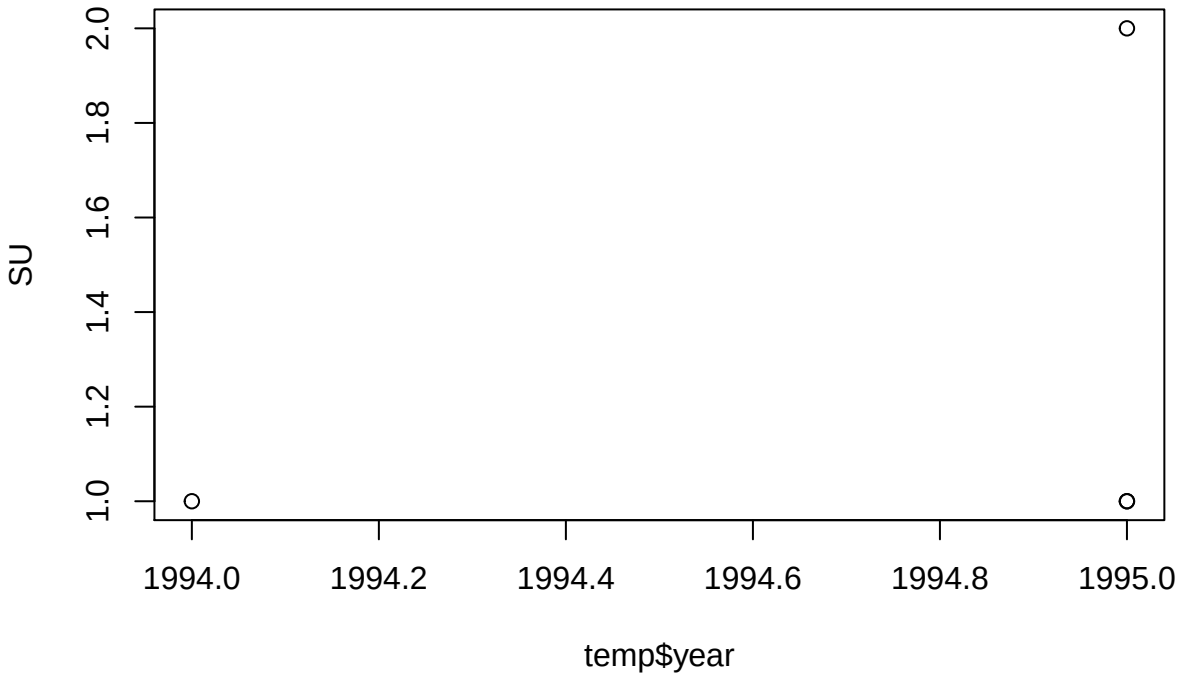

US
1.0
1.2
1.4
1.6
1.8
2.0
1500
2000
2500
temp$year

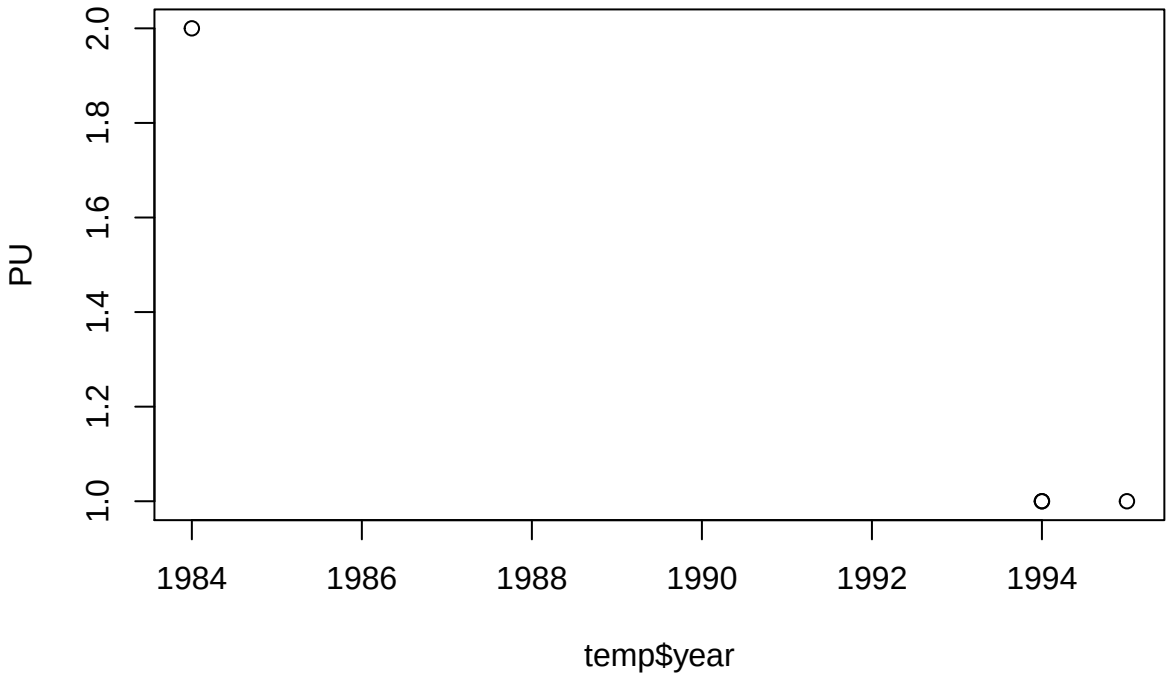

PG
1.0
1.2
1.4
1.6
1.8
2.0
1985
1990
1995
2000
temp$year

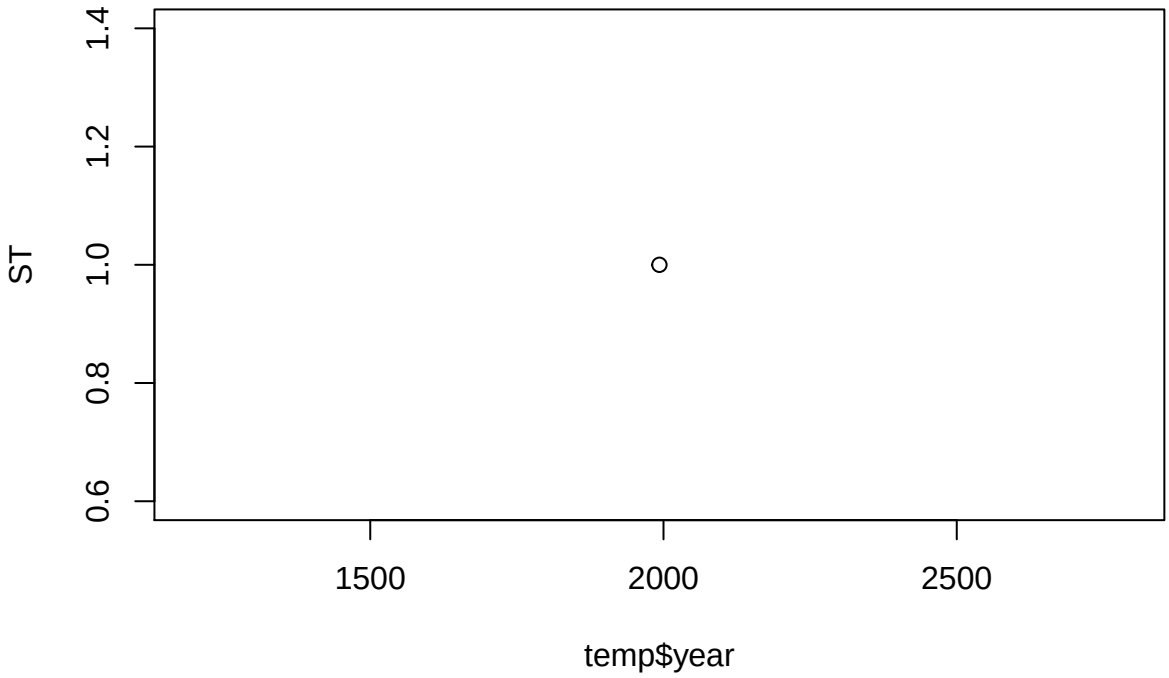

AS
0.6
0.8
1.0
1.2
1.4
1991.0
1991.5
1992.0
1992.5
1993.0
temp$year

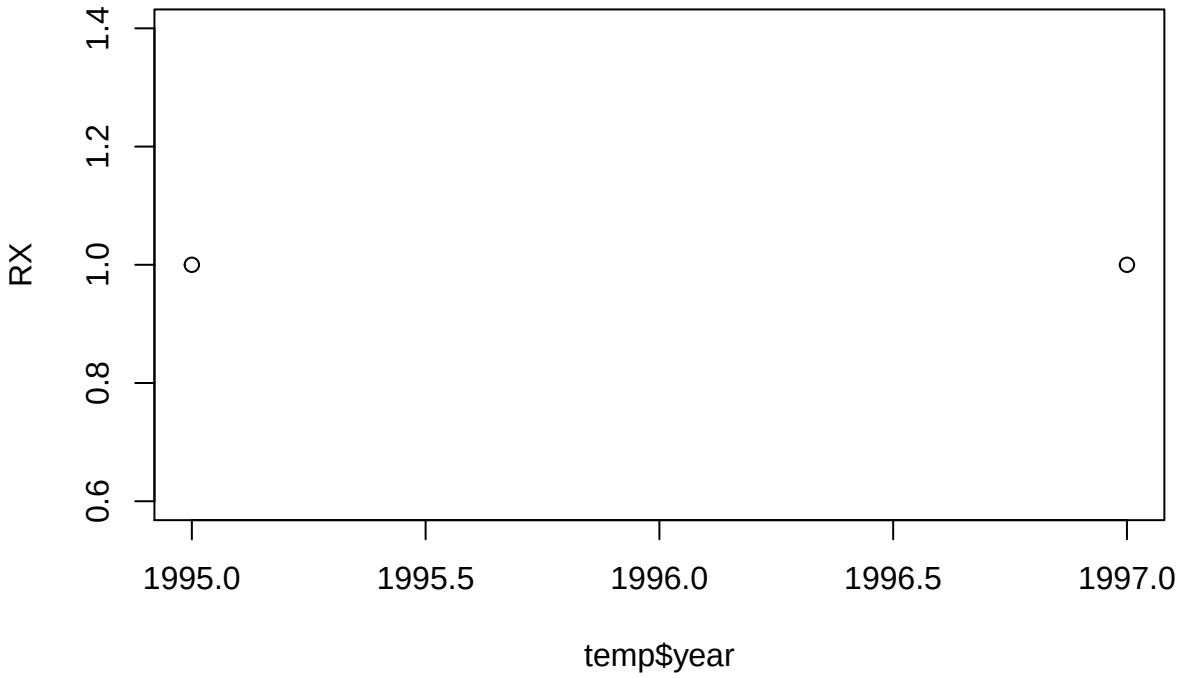

ZL
0.6
0.8
1.0
1.2
1.4
1500
2000
2500
temp$year

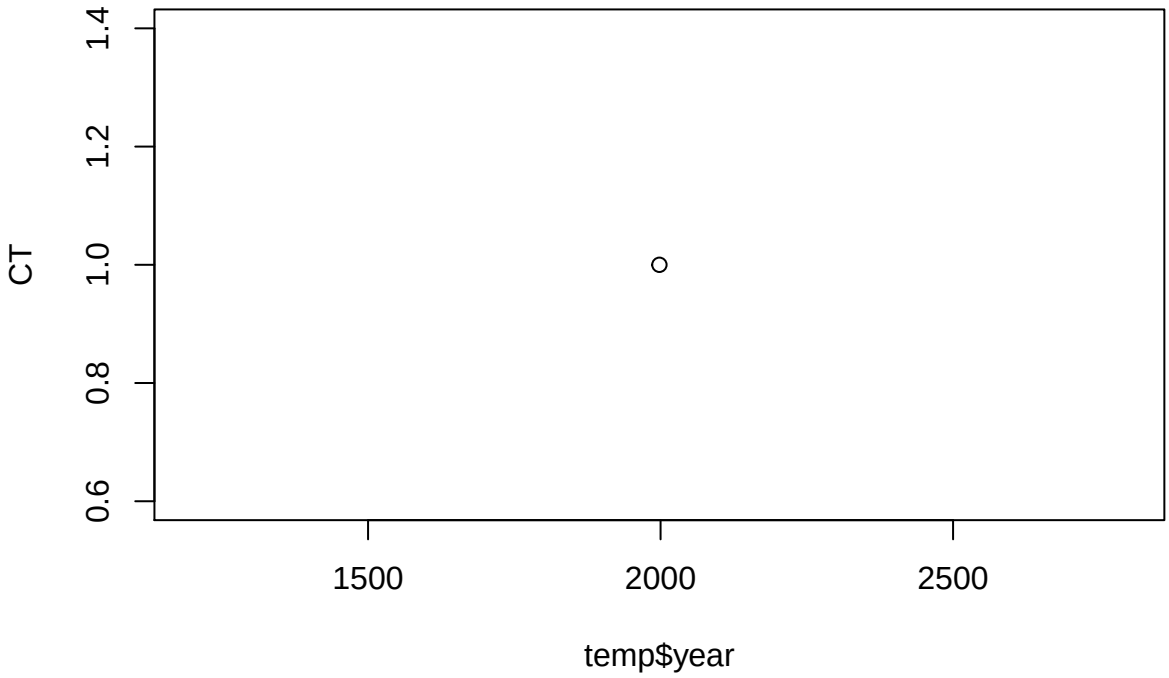

CS
0.6
0.8
1.0
1.2
1.4
1500
2000
2500
temp$year

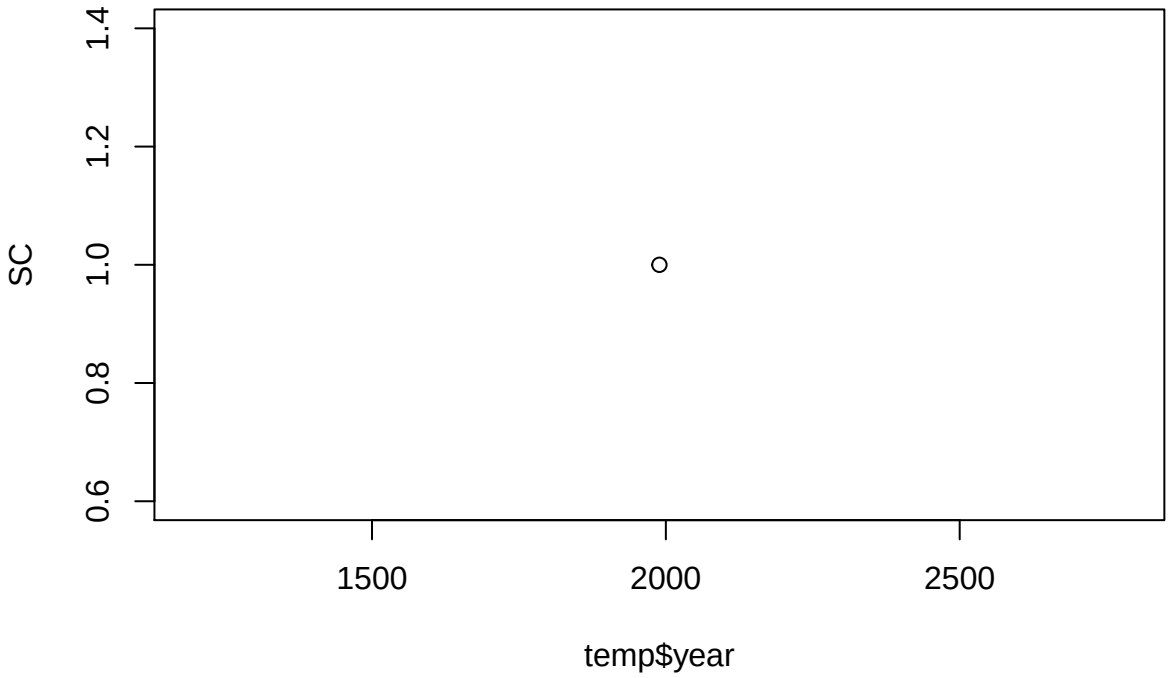

SU
1.0
1.2
1.4
1.6
1.8
2.0
1994.0
1994.2
1994.4
1994.6
1994.8
1995.0
temp$year

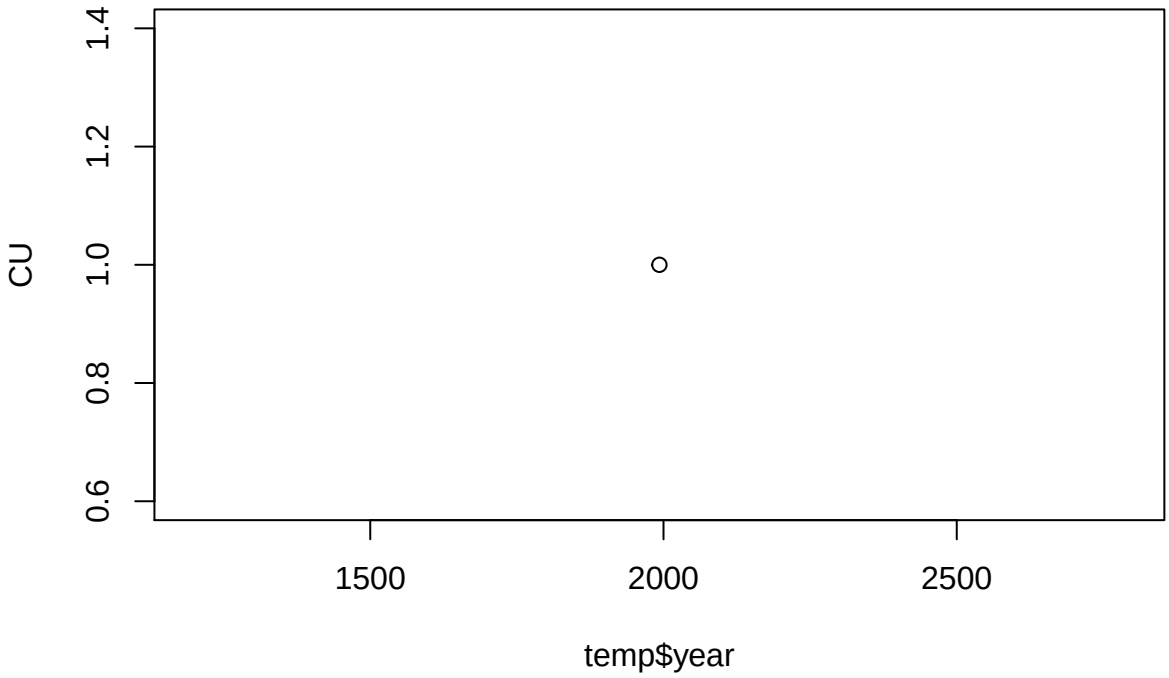

PU
1.0
1.2
1.4
1.6
1.8
2.0
1984
1986
1988
1990
1992
1994
temp$year


ST
0.6
0.8
1.0
1.2
1.4
1500
2000
2500
temp$year


RX
0.6
0.8
1.0
1.2
1.4
1995.0
1995.5
1996.0
1996.5
1997.0
temp$year


CT
0.6
0.8
1.0
1.2
1.4
1500
2000
2500
temp$year


SC
0.6
0.8
1.0
1.2
1.4
1500
2000
2500
temp$year



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?

a: Rather than asking which new species appear, and which species are no longer present, when a different site is considered, temporal turnover asks which new species appear, and which species are no longer present, at the same site but for different time steps.

b:

Several spp are represented by just one observation.

PU, SU, AS, PG, US, PX, PI, PC, CB, RO, UL, UP, UR, CQ, CM, OX were consistently rare or not present.

When present, PL was at a fairly stable level of rarity, but it was only observed between 1995 and 2000.

SF disappeared 1994 - 2001.

PH was stably rare 1984 to 1997.

CB was very variable, its abundances ranging from near 0 to over 300.

AB, RM, OL, SS, DO, OT, PP, DS, PE, PF, DM, NL, and AH were present in most years 1977 or 1980 to 2002 or 2003. They were not particularly variable. OL was absent 1998 through 2001. Abundances for PP rose in more recent years; DS abundances decreased in recent years (and after 1999 it was absent).

BA peaked at ~21 individuals in 1991, but was absent for all other years except 1992, 1989, 1990.

UR was observed in only 5 different years. 20 individuals were observed in 1990.

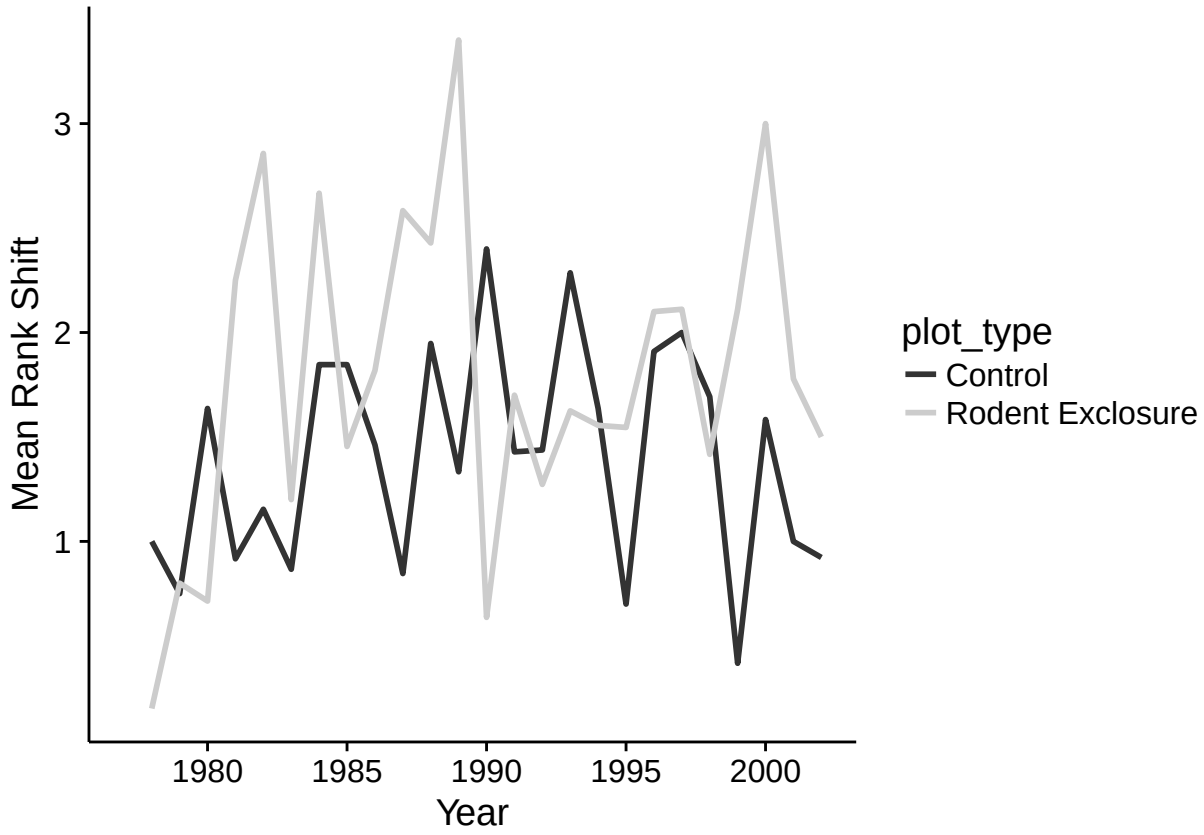
PM was variously rare and abundant 1978 to 2002. There were several observations of more than 40; however, during 1990 to 1994 there were no observations at all.

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.

```
portal.abunds.cont.rodent <- filter(portal.species.abunds, plot_type == "Rodent Exclosure" | plot_type =
# Calculate MRS
portal.rankshift <- rank_shift(df = as.data.frame(portal.abunds.cont.rodent), time.var = "year",
species.var = "species_id",
abundance.var = "n",
replicate.var = "plot_type")
# Replace year range w single value to plot
portal.rankshift$year <- as.numeric(substr(portal.rankshift$year_pair, 6, 9))
# now 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)
```




```
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:

It tells you about the community structure and whether it is more dynamic or more static.

Answer 6b:

From the figure alone, most of the time, it looks as though communities on control plots were somewhat more stable than communities on rodent exclosure plots.

The quantitative analysis indicates that this interpretation is correct, although there was no hypothesis test/ P-value associated with it.

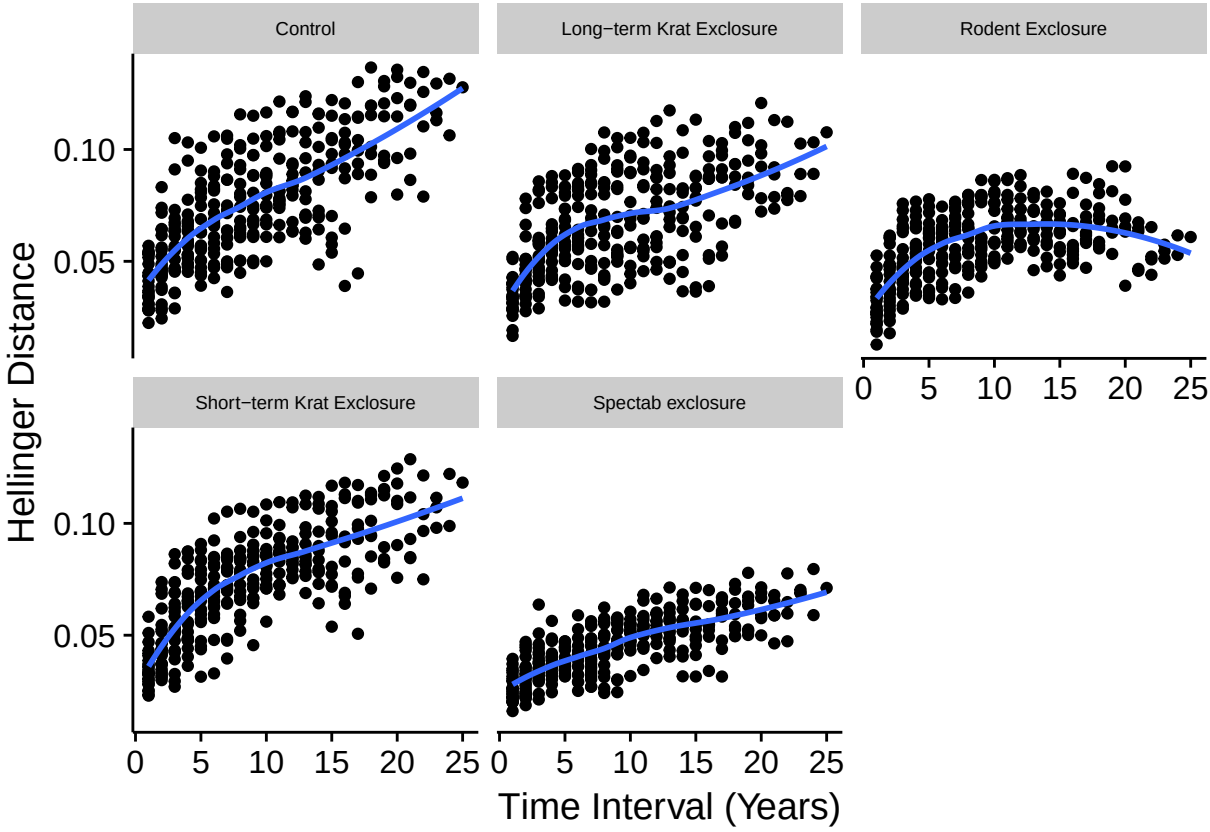
Rate Change Interval

In the R code chunk below, do the following:

1. Calculate the rate change interval using the Hellinger distance.

2. Plot the results.

```
#1
portal.species.abunds$tot.abund <- rep(sum(portal.species.abunds$n),length(portal.species.abunds$n))
# Now Hellinger transformation
portal.hellinger.transf <- portal.species.abunds %>%
mutate(hellinger.transf = sqrt(n / tot.abund))

portal.change.int <- rate_change_interval(portal.hellinger.transf,time.var = "year",
species.var = "species_id",
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)")

#2
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:

In the Mean Rank Shift analysis, we assessed how much a community changed between different censuses. Here, we want to be able to calculate a rate of change—if the change is directed/directional, how long will it take to achieve a particular un- or -desired outcome.

Answer 7b:

Let us assume that the Control represents is the expected community divergence over time. The longterm k-rat and short-term r-rat exclosure treatments exhibit fairly similar slopes (rates) of divergence, but perhaps slightly less than the control. Perhaps kangaroo rats are a keystone species that promotes biodiversity and turnover through some sort Lotka-Volterra-esque cycling. Perhaps it induces cycles of available vegetation, which in turn drive the rodent herbivore community structure. I could not find any metadata to indicate the specific meaning of 'spectab'; let us commit the presumption that it refers to excluding a very, very wide array of herbivores. Perhaps there are now even fewer keystone species driving turnover cycles, and a handful of specialists (or generalists?) become stably established. In the rodent exclosure treatment, the community structure begins to return toward its starting point after ~10 years. This is not straightforward to interpret. It is possible that initially the exclusion of rodents allowed for a diversification of ant species. However, after ~5 years, the most dominant ant species continued to outcompete less dominant ones, which became extirpated. By year ~10, the extirpated species were gone and stayed gone, while the dominant species remained.

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.

```
#1. Using total abundance as your focal variable, calculate stability (i.e., 1/CV) and synchrony for ea
portal.stab <- community_stability(df = as.data.frame(portal.species.abunds),
time.var = "year",
abundance.var = "n",
replicate.var = "plot_type")

pander(portal.stab)
```

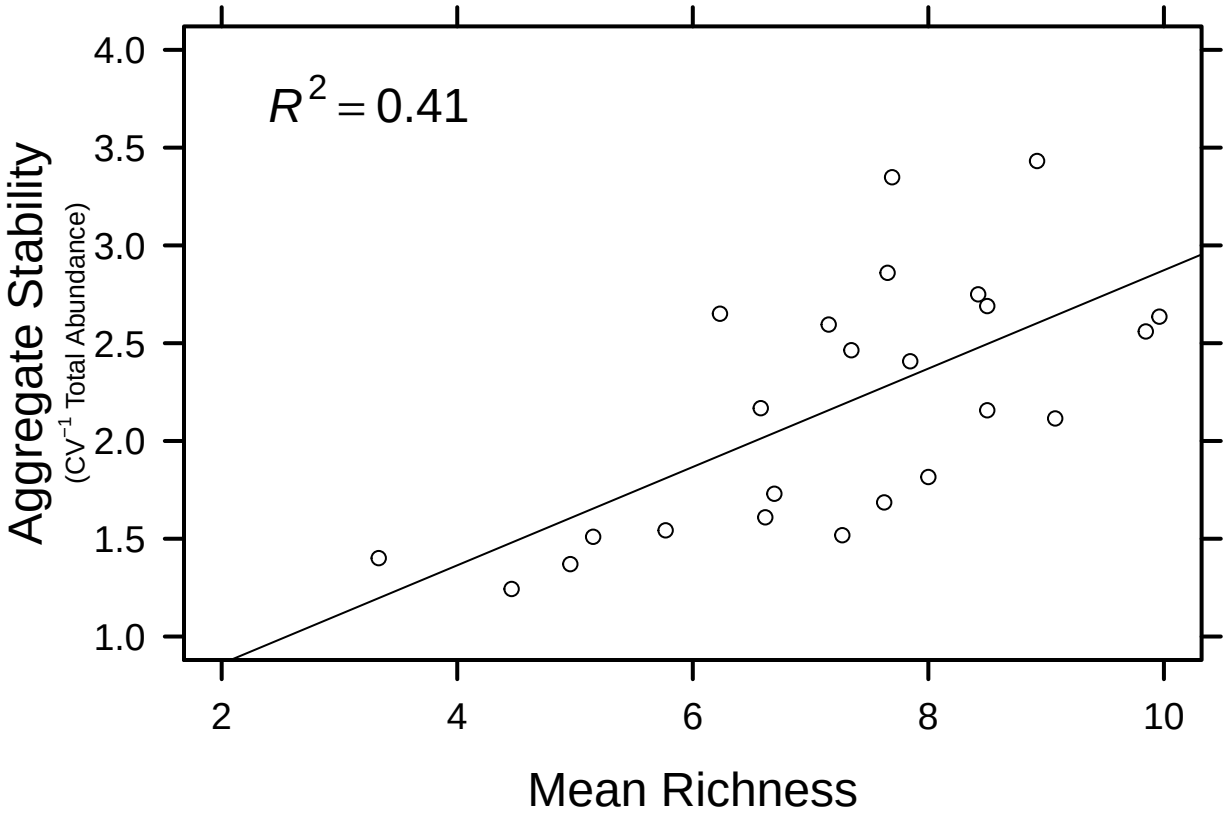
plot_type	stability
Control	3.044
Long-term Krat Exclosure	1.865
Rodent Exclosure	1.864
Short-term Krat Exclosure	2.462
Spectab exclosure	2.911

```
#2. Test for a biodiversity-stability relationship by regressing community stability on mean richness.
time.by.species <- group_by(portal, year, plot_id,
plot_type) %>% count(species_id) %>% spread(key = species_id, value = n, fill = 0)
richness <- as.data.frame(rowSums(time.by.species[,-c(1:3)] > 0))
rich.all <- data.frame(time.by.species[,1:3,], richness)
names(rich.all)[4] <- "richness"
```

```
rich.treat <- rich.all[which(rich.all$plot_type ==
"Control" | rich.all$plot_type == "Rodent Exclosure"), ]

portal.mean.rich.plot <- rich.all %>% group_by(plot_id) %>% summarise(mean.rich = mean(richness))
# Let's take a look at how stability metrics relate to mean richness
portal.plot.abunds <- as.data.frame(
group_by(portal, year, plot_id) %>% count(species_id))
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"
```

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

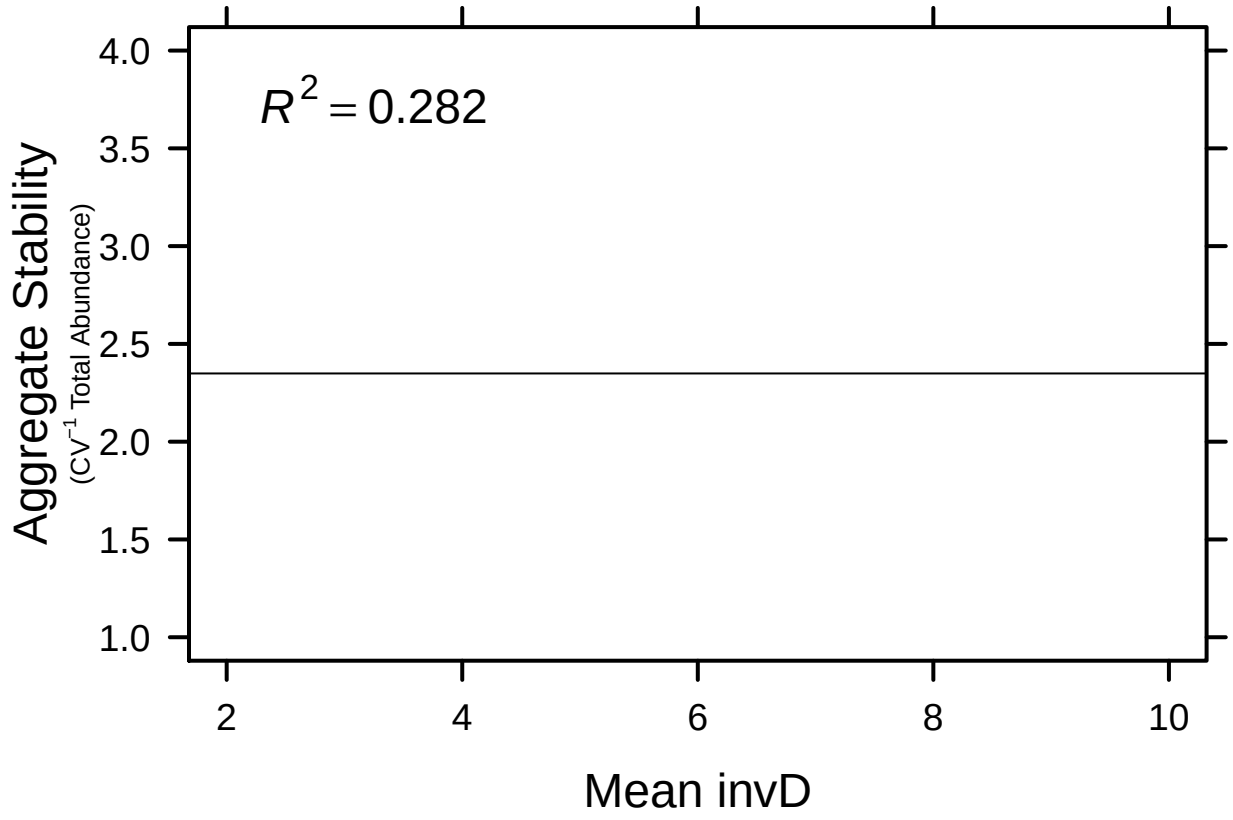



```
#3. Test for a biodiversity-stability relationship by regressing community stability on mean inverse Si
portal.mean.simp.plot <- simp.all %>% group_by(plot_id) %>% summarise(mean.simp = mean(invD))
# Let's take a look at how stability metrics relate to mean richness
portal.plot.abunds <- as.data.frame(
group_by(portal, year, plot_id) %>% count(species_id))
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.simp.plot %>%
inner_join(portal.stab.plot)
```

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

```
par(mar = c(5,5,1,1))
plot(portal.div.stab$stability ~ portal.div.stab$mean.simp,
xlab = "", ylab = "", yaxt = "n", xaxt = "n",
xlim = c(2,10), ylim = c(1,4))
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 invD", 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 = .8)
div.stab.lm <- lm(portal.div.stab$stability ~ portal.div.stab$mean.simp)
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:

Highest stability: Control. Stability is measured as the inverse of the relativized variance (that is, standard deviation normalized by the mean). Viz., it is ((CV)^(-1)).

Answer 8b:

Within the community, turnover will be driven by the dynamics of individual populations. If certain populations tend to be in (positively) correlated states of flux, they're in synchrony. This is important because if many populations fluctuate in tandem, then a single event/disaster/environmental change could alter the dynamics of many populations at once. On the other hand, if some populations tend to grow at the same time others shrink, that would be asynchrony.

Answer 8c:

There was a clear relationship between richness and stability. Richness was a good predictor of stability: as richness increased, so too did stability, and the slope (0.2515) was not insubstantial. It was highly sig-nificant (P=0.000452; R^2 = 0.410).

There was not a clear relationship between invD and stability. As invD changed, stability did not change much at all; the slope was essentially zero (-1.183e-08). That said, perhaps due to the power of a very large sample size, invD was a significant predictor of stability (P = 0.00449) even though the R^2 value was fairly low (R^2 = .282). Although the relationship was significant, I would not say that it was biologically meaningful/useful.

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:

Across space we expect to so spatial autocorrelation such that locations which are close together tend to be more similar in their environments than locations which are farther apart from one another. Similarly across time, timepoints are autocorrelated because the change in environmental makeup or community composition will tend to be less for timepoints 1 or 2 days/ weeks/ years apart than for timepoints many steps apart.

Across space we need to account for scale because it could make a big difference in the conclusion that we draw. For example, what may look like a tight, random cluster of points at a very large scale could look like a very evenly dispersed array of points at a finer grain. Across time, increasing the scale of the analysis could allow you to detect i) nonstationarity over longer time periods; ii) seasonality—for example, if you just measured 1 year, there would be no REPEATING pattern to call seasonality, but if you measured over 10 years, you might indeed pick up an annually-repeating pattern in your data.

Often it seems that challenges to measuring biodiversity come down to gathering the requisite amount of data at the requisite level of resolution. If you census a rodent community every week for 15 years, you will be very likely to pick up on trends and correlations in your data, if there are real biological phenomena causing them. But if you are only able to sample once each year, and only for 4 or 5 years, it will be much more difficult to use those data to delineate patterns and make predictions. So it can come down to whether the ecologist is able to make the requisite sampling effort required for visualizing and calculating statistically robust trends in the data. For sampling across space, the challenge may be similar: is the ecologist able to include a broad enough spatial array of sites in order to visualize and statically measure the trends that she is interested in? For some studies, the requisite spatial sampling effort might be small: perhaps the ecologist only wants to now the trends in mesopredator population in one forest during a trophic cascade. But if the ecologist wants learn about differences in mesopredator population dynamics during trophic cascades in different biomes, if there are any biologically meaningful trends, they will be more difficult to detect simply because the effort required to have enough data to measure them robustly will be comparatively vast.