Regional_nMDS_SIMPER_PERMANOVA

Jackson Hoeke

July 28, 2021

```
library(knitr)
hook_output = knit_hooks$get('output')
knit_hooks$set(output = function(x, options) {
    # this hook is used only when the linewidth option is not NULL
    if (!is.null(n <- options$linewidth)) {
        x = knitr:::split_lines(x)
        # any lines wider than n should be wrapped
        if (any(nchar(x) > n)) x = strwrap(x, width = n)
        x = paste(x, collapse = '\n')
    }
    hook_output(x, options)
})
```

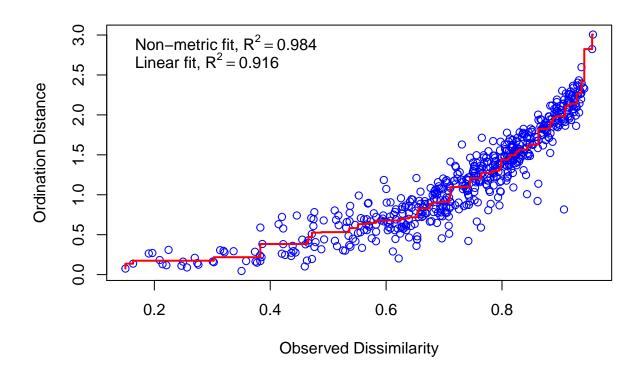
nMDS Plots

```
# Import data
density <- data.frame(read.csv("invert_matrix_transects.csv"))</pre>
density[is.na(density)] <- 0</pre>
# Create groups using 'newnames' to split data
grp.MPA <- density %>%
  group_by_all() %>%
  mutate(newnames=paste0(Region, Designation, Year))
# Split data into smaller frames
split.MPA <- split(grp.MPA, grp.MPA$newnames)</pre>
for (I in 1:length(split.MPA)) {assign(unique(split.MPA[[I]] newnames),
                                          split.MPA[[I]])}
# Determine average densities of organisms for each Region, Designation,
# and Year
take.avg <- lapply(split.MPA, function(x){</pre>
  avgs <- colMeans(x[,9:165])
  avgs <- as.data.frame(avgs)</pre>
  t.avgs <- as.data.frame(transpose(avgs))</pre>
  colnames(t.avgs) <- rownames(avgs)</pre>
  lab <- data.frame(Year=x[1,2], Region=x[1,3], MPA_Group=x[1,4],</pre>
```

```
Designation=x[1,7])
  region.unit <- cbind(lab,t.avgs)
})
# Bind rows of new data frames together from the split
only.avgs <- bind rows(take.avg)
# create data frame of labels (Year, Region, MPA, Designation)
reg.labels <- only.avgs[,1:4]
# create data frame of species matrix
reg.data <- only.avgs[,5:160]
# Generate MDS
set.seed(13)
reg.mds <- metaMDS(reg.data, distance = "bray", k = 2, trymax = 999,
              autotransform = FALSE)
## Run 0 stress 0.1275924
## Run 1 stress 0.1275924
## ... New best solution
## ... Procrustes: rmse 2.385989e-06 max resid 7.659247e-06
## ... Similar to previous best
## Run 2 stress 0.1275924
## ... New best solution
## ... Procrustes: rmse 1.844029e-06 max resid 7.034715e-06
## ... Similar to previous best
## Run 3 stress 0.1278416
## ... Procrustes: rmse 0.02874042 max resid 0.1086056
## Run 4 stress 0.2271201
## Run 5 stress 0.1275924
## ... Procrustes: rmse 1.010998e-05 max resid 4.161017e-05
## ... Similar to previous best
## Run 6 stress 0.1275924
## ... Procrustes: rmse 2.369381e-06 max resid 8.113587e-06
## ... Similar to previous best
## Run 7 stress 0.1278416
## ... Procrustes: rmse 0.02874176 max resid 0.1086152
## Run 8 stress 0.1278416
## ... Procrustes: rmse 0.02874434 max resid 0.1086339
## Run 9 stress 0.1275925
## ... Procrustes: rmse 1.607837e-05 max resid 5.72162e-05
## ... Similar to previous best
## Run 10 stress 0.1278416
## ... Procrustes: rmse 0.02874215 max resid 0.1086188
## Run 11 stress 0.1275924
## ... Procrustes: rmse 4.953412e-06 max resid 1.400328e-05
## ... Similar to previous best
## Run 12 stress 0.1275924
```

```
## ... Procrustes: rmse 1.944918e-06 max resid 6.397273e-06
## ... Similar to previous best
## Run 13 stress 0.1275924
## ... Procrustes: rmse 4.155899e-06 max resid 1.615037e-05
## ... Similar to previous best
## Run 14 stress 0.1275924
## ... New best solution
## ... Procrustes: rmse 1.795318e-06 max resid 6.614027e-06
## ... Similar to previous best
## Run 15 stress 0.1275924
## ... Procrustes: rmse 1.895525e-06 max resid 8.2185e-06
## ... Similar to previous best
## Run 16 stress 0.1946847
## Run 17 stress 0.1275924
## ... Procrustes: rmse 6.338998e-07 max resid 1.78383e-06
## ... Similar to previous best
## Run 18 stress 0.1959829
## Run 19 stress 0.1861935
## Run 20 stress 0.2014039
## *** Solution reached
\mbox{\# Divide MDS into data} and species MDS scores across 2 axes
data.scores <- as.data.frame(scores(reg.mds))</pre>
data.scores$site <- rownames(data.scores)</pre>
data.scores <- cbind(data.scores,reg.labels)</pre>
head(data.scores)
##
          NMDS1
                      NMDS2 site Year Region
                                                     MPA Group Designation
## 1 -0.7680514 -0.51461743 1 2007 Central
                                                   Point Lobos
## 2 -0.9460948 -0.35464770 2 2008 Central
                                                 Point Buchon
                                                                       MPA
## 3 -0.6779332 -1.32543000
                             3 2009 Central
                                                  Point Lobos
                                                                       MPA
## 4 -1.1065013 -0.03252102 4 2012 Central
                                                  Point Buchon
                                                                       MPA
## 5 -0.2274672 -0.28054402
                               5 2016 Central Piedras Blancas
                                                                       MPA
## 6 0.8785583 -0.30256414
                               6 2019 Central
                                                  Point Buchon
                                                                       MPA
species.scores<- as.data.frame(scores(reg.mds, "species"))</pre>
species.scores$species <- rownames(species.scores)</pre>
head(species.scores)
##
                                      NMDS1
                                                   NMDS2
## Acorn.barnacle
                                  1.1080684 -0.11131935
## Aggregated.nipple.sponge
                                 -2.1317519 -0.10495806
## Basket.star
                                  1.1059118 -1.16115624
## Bat.star
                                 -0.3149907 0.24852968
## Bat.star.red.sea.star.complex -1.3323847 -0.03665008
## Benthic.siphonophore
                                        NaN
                                                     NaN
                                                        species
## Acorn.barnacle
                                                 Acorn.barnacle
## Aggregated.nipple.sponge
                                      Aggregated.nipple.sponge
## Basket.star
                                                    Basket.star
## Bat.star
                                                       Bat.star
## Bat.star.red.sea.star.complex Bat.star.red.sea.star.complex
## Benthic.siphonophore
                                          Benthic.siphonophore
```

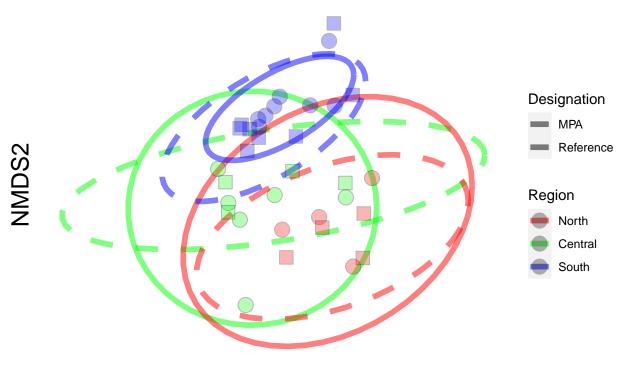
```
# Plot stress
stressplot(reg.mds)
```



```
# Plot nMDS in ggplot
ggplot() +
  stat_ellipse(data=data.scores,aes(x=NMDS1,y=NMDS2, color=Region,
                                    linetype=Designation), lwd=2, alpha=0.5) +
  geom_point(data=data.scores,aes(x=NMDS1,y=NMDS2, fill=Region, pch=Designation),
             size=5, alpha=0.3) +
  guides(pch = "none", size = "none") +
  ggtitle("MPA vs. Reference") +
  scale_linetype_manual(values=c(1,2), breaks = c("MPA", "Reference")) +
  scale_shape_manual(values=c(21,22)) +
  scale_color_manual(values=c("red","green","blue"),
                     breaks = c("North", "Central", "South")) +
  scale_fill_manual(values=c("red","green","blue"),
                   breaks = c("North", "Central", "South")) +
  theme(axis.text.x = element_blank(), # remove x-axis text
        axis.text.y = element_blank(), # remove y-axis text
        axis.ticks = element_blank(), # remove axis ticks
        axis.title.x = element_text(size=18), # remove x-axis labels
        axis.title.y = element_text(size=18), # remove y-axis labels
        panel.background = element_blank(),
       panel.grid.major = element_blank(), #remove major-grid labels
```

```
panel.grid.minor = element_blank(), #remove minor-grid labels
plot.background = element_blank(),
plot.title = element_text(size=20))
```

MPA vs. Reference



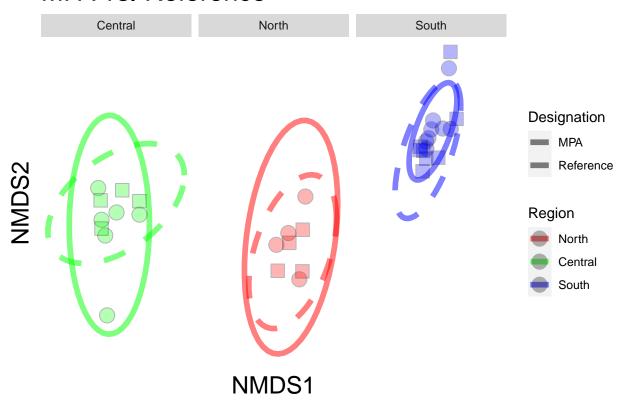
NMDS1

```
# Now to use facets to plot several NMDSs at once
 base <- ggplot() +</pre>
    stat_ellipse(data=data.scores,
                 aes(x=NMDS1,y=NMDS2, color=Region, linetype=Designation),
                 lwd=2, alpha=0.5) +
   geom_point(data=data.scores,
               aes(x=NMDS1,y=NMDS2, fill=Region, pch=Designation),
               size=5, alpha=0.3) +
   guides(pch = "none", size = "none") +
   ggtitle("MPA vs. Reference") +
   scale_linetype_manual(values=c(1,2),
                          breaks = c("MPA", "Reference")) +
   scale_shape_manual(values=c(21,22)) +
   scale_color_manual(values=c("red", "green", "blue"),
                       breaks = c("North", "Central", "South")) +
    scale_fill_manual(values=c("red","green","blue"),
                      breaks = c("North", "Central", "South")) +
   theme(axis.text.x = element_blank(), # remove x-axis text
          axis.text.y = element_blank(), # remove y-axis text
          axis.ticks = element_blank(), # remove axis ticks
          axis.title.x = element_text(size=18), # remove x-axis labels
```

```
axis.title.y = element_text(size=18), # remove y-axis labels
panel.background = element_blank(),
panel.grid.major = element_blank(), #remove major-grid labels
panel.grid.minor = element_blank(), #remove minor-grid labels
plot.background = element_blank(),
plot.title = element_text(size=20))

# First by region
base + facet_wrap(~Region)
```

MPA vs. Reference



```
# Then by Year

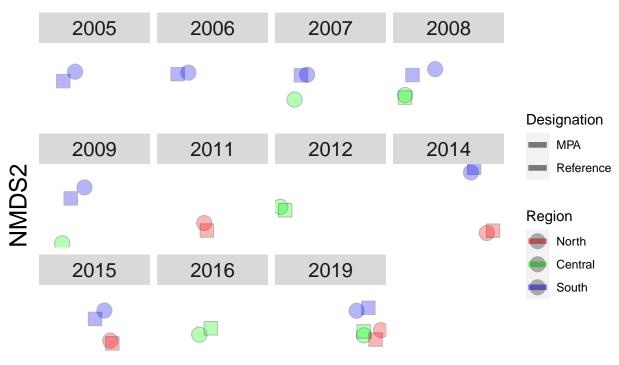
base + facet_wrap(~Year) +
    theme(strip.text.x = element_text(size = 15))

## Too few points to calculate an ellipse
```

```
## Too few points to calculate an ellipse
```

Warning: Removed 34 row(s) containing missing values (geom_path).

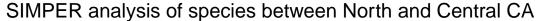
MPA vs. Reference

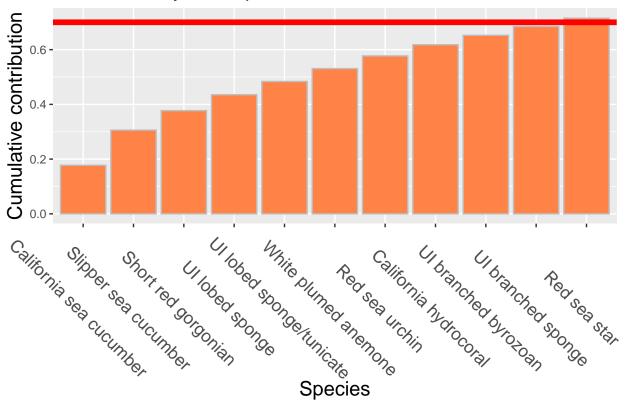


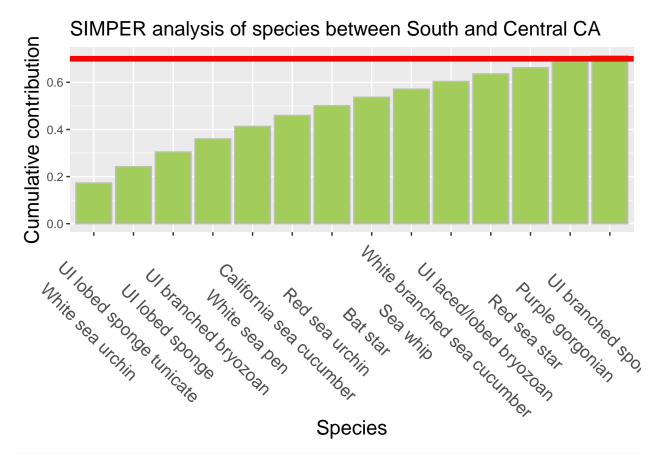
NMDS1

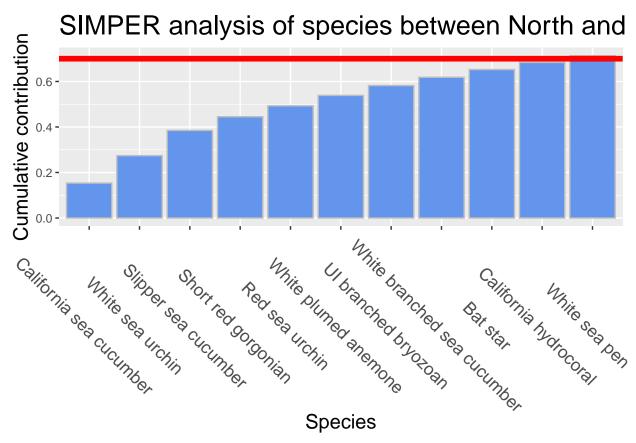
```
## SIMPER analysis
# Use data frames set up for nMDS to use in SIMPER analysis
# First, compare differences between regions
set.seed(13)
sim <- with(reg.labels, simper(reg.data, Region))</pre>
   cumulative contributions of most influential species:
##
## $Central_North
##
    California.sea.cucumber
                                 Slipper.sea.cucumber
                                                             Short.red.gorgonian
                                                                        0.3772839
                   0.1774812
                                             0.3060658
##
            UI.lobed.sponge UI.lobed.sponge.tunicate
                                                            White.plumed.anemone
##
                   0.4355861
                                             0.4839536
                                                                        0.5310859
##
             Red.sea.urchin
                                California.hydrocoral
                                                            UI.branched.bryozoan
##
                   0.5774277
                                             0.6182181
                                                                       0.6527903
##
         UI.branched.sponge
                                          Red.sea.star
##
                   0.6852496
                                             0.7154496
##
##
   $Central_South
##
              White.sea.urchin
                                   UI.lobed.sponge.tunicate
                                                    0.2420651
##
                      0.1726507
               UI.lobed.sponge
                                        UI.branched.bryozoan
##
                      0.3044175
                                                    0.3600304
##
##
                  White.sea.pen
                                    California.sea.cucumber
##
                      0.4129377
                                                   0.4594726
                Red.sea.urchin
##
                                                    Bat.star
                                                   0.5367493
##
                      0.5008071
##
                       Sea.whip White.branched.sea.cucumber
##
                      0.5711367
                                                    0.6037940
##
       UI.laced.lobed.bryozoan
                                                Red.sea.star
##
                      0.6358096
                                                   0.6622935
##
              Purple.gorgonian
                                          UI.branched.sponge
                      0.6880643
                                                   0.7127600
##
##
##
   $North_South
##
       California.sea.cucumber
                                            White.sea.urchin
##
                      0.1534472
                                                    0.2741373
##
          Slipper.sea.cucumber
                                         Short.red.gorgonian
##
                      0.3852217
                                                   0.4451593
##
                Red.sea.urchin
                                        White.plumed.anemone
                      0.4932306
##
                                                   0.5390345
##
          UI.branched.bryozoan White.branched.sea.cucumber
##
                      0.5824273
                                                    0.6187451
##
                       Bat.star
                                       California.hydrocoral
##
                      0.6526233
                                                    0.6841423
##
                  White.sea.pen
                      0.7133553
# Divide into species and cumulative contributions, then make data frames from these
# vectors
CN_sp <- c("California sea cucumber", "Slipper sea cucumber",</pre>
```

```
"Short red gorgonian", "UI lobed sponge",
           "UI lobed sponge/tunicate", "White plumed anemone",
           "Red sea urchin", "California hydrocoral",
           "UI branched byrozoan", "UI branched sponge", "Red sea star")
CN con \leftarrow c(0.1773808, 0.3059056, 0.3770985, 0.4353616, 0.4837151,
            0.5308375,0.5771436,0.6178814,0.6524418,0.6848954,0.7150882)
CN_data <- data.frame(CN_sp,CN_con)</pre>
CS_sp <- c("White sea urchin", "UI lobed sponge tunicate",
           "UI lobed sponge", "UI branched bryozoan",
           "White sea pen", "California sea cucumber",
           "Red sea urchin", "Bat star", "Sea whip",
           "White branched sea cucumber", "UI laced/lobed bryozoan",
           "Red sea star", "Purple gorgonian", "UI branched sponge")
CS_{con} \leftarrow c(0.1726507, 0.2420651, 0.3044175, 0.3600304, 0.4129377,
            0.4594726,0.5008071,0.5367493,0.5711367,0.6037940,
            0.6358096, 0.6622935, 0.6880643, 0.7127600)
CS_data <- data.frame(CS_sp,CS_con)</pre>
NS_sp <- c("California sea cucumber", "White sea urchin",
           "Slipper sea cucumber", "Short red gorgonian",
           "Red sea urchin", "White plumed anemone",
           "UI branched bryozoan", "White branched sea cucumber",
           "Bat star", "California hydrocoral", "White sea pen")
NS_{con} \leftarrow c(0.1533639, 0.2740252, 0.3850616, 0.444977, 0.4930209,
            0.5388162,0.5821998,0.6185102,0.6523830,0.6838560,
            0.7130625)
NS_data <- data.frame(NS_sp,NS_con)</pre>
# Plot contributions with one graph per region
CN_graph <- ggplot(CN_data, aes(y=CN_con,x=factor(CN_sp,level=CN_sp))) +</pre>
  geom_bar(position="dodge", stat="identity", fill='#ff8247',color='gray') +
  ggtitle("SIMPER analysis of species between North and Central CA") +
  geom_hline(yintercept = 0.7,color='red',size=2) +
  theme(axis.text.x = element_text(angle = 315, size=13),
        plot.title = element_text(size=15),
        axis.title = element_text(size=15)) +
  xlab("Species") +
  ylab("Cumulative contribution")
CN_graph
```





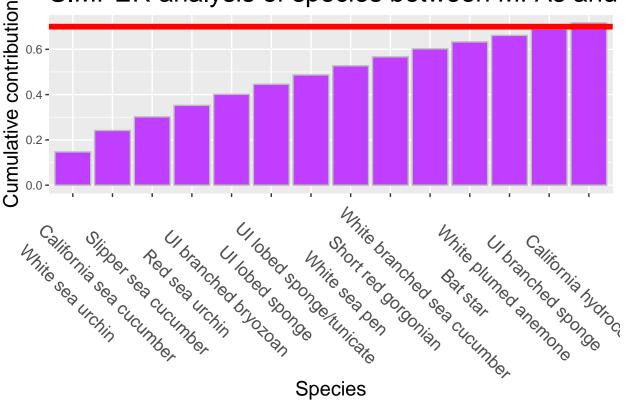




```
# Now, use SIMPER to analyze differences in MPA vs. Reference designation
set.seed(13)
sim.des <- with(reg.labels, simper(reg.data, Designation))</pre>
   cumulative contributions of most influential species:
##
##
   $MPA_Reference
                                     California.sea.cucumber
##
              White.sea.urchin
                                                   0.2417720
##
                      0.1469531
          Slipper.sea.cucumber
##
                                              Red.sea.urchin
##
                      0.3017131
                                                   0.3526541
##
          UI.branched.bryozoan
                                             UI.lobed.sponge
                      0.4012280
##
                                                   0.4459862
##
      UI.lobed.sponge.tunicate
                                               White.sea.pen
##
                      0.4868980
                                                    0.5273387
##
           Short.red.gorgonian White.branched.sea.cucumber
##
                      0.5666367
                                                   0.6021512
##
                       Bat.star
                                        White.plumed.anemone
##
                      0.6327757
                                                   0.6615762
##
            UI.branched.sponge
                                       California.hydrocoral
                      0.6896547
# Use the provided information to create a data frame
Des_sp <- c("White sea urchin", "California sea cucumber",</pre>
```

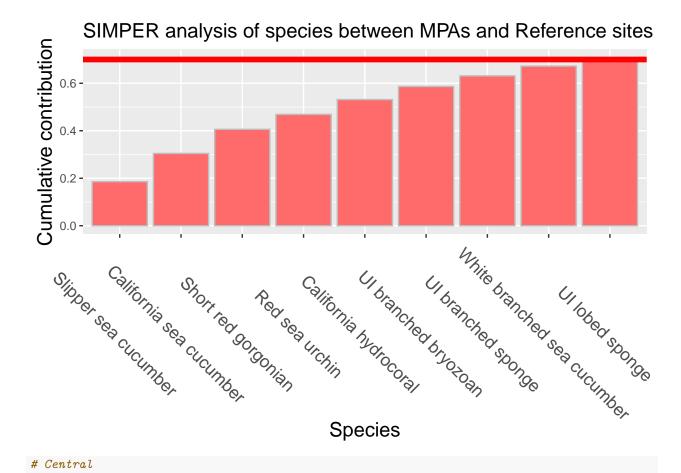
```
"Slipper sea cucumber", "Red sea urchin",
            "UI branched bryozoan", "UI lobed sponge",
            "UI lobed sponge/tunicate", "White sea pen",
            "Short red gorgonian", "White branched sea cucumber",
            "Bat star", "White plumed anemone", "UI branched sponge",
            "California hydrocoral")
Des_con <- c(0.1469395,0.2417168,0.3016337,0.3525581,0.4011254,
             0.4458663, 0.4867733, 0.5272089, 0.5664956, 0.6020061,
             0.6326269, 0.6614230, 0.6894982, 0.7165490)
Des_data <- data.frame(Des_sp,Des_con)</pre>
# And plot it
Des_graph <- ggplot(Des_data, aes(y=Des_con,x=factor(Des_sp,level=Des_sp))) +</pre>
  geom_bar(position="dodge", stat="identity", fill='#bf3eff',color='gray') +
  ggtitle("SIMPER analysis of species between MPAs and Reference sites") +
  geom_hline(yintercept = 0.7,color='red',size=2) +
  theme(axis.text.x = element_text(angle = 315, size=13),
        plot.title = element_text(size=20),
        axis.title = element_text(size=15)) +
  xlab("Species") +
  ylab("Cumulative contribution")
Des_graph
```

SIMPER analysis of species between MPAs and



Now, to subset by region & SIMPER by designation separately

```
North.avg <- subset(only.avgs, Region == "North")
Central.avg <- subset(only.avgs, Region == "Central")</pre>
South.avg <- subset(only.avgs, Region == "South")
# North
north.labels <- North.avg[,1:4]</pre>
north.data <- North.avg[,5:160]</pre>
set.seed(13)
north.sim <- with(north.labels, simper(north.data, Designation))</pre>
north.sim
## cumulative contributions of most influential species:
##
## $MPA_Reference
                                    California.sea.cucumber
##
          Slipper.sea.cucumber
                                                   0.3047077
##
                     0.1857117
                                             Red.sea.urchin
##
           Short.red.gorgonian
##
                     0.4064165
                                                   0.4693606
##
         California.hydrocoral
                                       UI.branched.bryozoan
                                                   0.5870058
##
                     0.5317557
##
            UI.branched.sponge White.branched.sea.cucumber
##
                     0.6316958
                                                   0.6728573
##
               UI.lobed.sponge
##
                     0.7081504
north_sp <- c("Slipper sea cucumber", "California sea cucumber",</pre>
              "Short red gorgonian", "Red sea urchin",
              "California hydrocoral", "UI branched bryozoan",
              "UI branched sponge", "White branched sea cucumber",
              "UI lobed sponge")
north_con <- c(0.1855598,0.3043917,0.4060153,0.4688773,0.5311454,
               0.5863474,0.6310009,0.6721305,0.7073433)
north_data <- data.frame(north_sp,north_con)</pre>
north_graph <- ggplot(north_data,</pre>
                       aes(y=north_con,x=factor(north_sp,level=north_sp))) +
  geom_bar(position="dodge", stat="identity", fill='#ff6a6a',color='gray') +
  ggtitle("SIMPER analysis of species between MPAs and Reference sites in Northern CA") +
  geom_hline(yintercept = 0.7,color='red',size=2) +
  theme(axis.text.x = element_text(angle = 315, size=13),
        plot.title = element_text(size=15),
        axis.title = element_text(size=15)) +
  xlab("Species") +
  ylab("Cumulative contribution")
north_graph
```

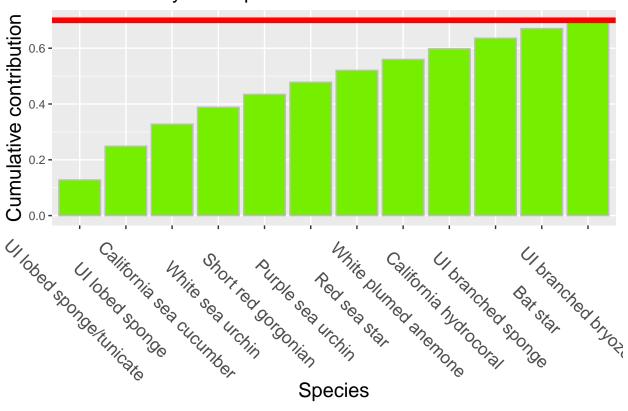


```
# Central
central.labels <- Central.avg[,1:4]</pre>
central.data <- Central.avg[,5:160]</pre>
set.seed(13)
central.sim <- with(central.labels, simper(central.data, Designation))</pre>
central.sim
## cumulative contributions of most influential species:
##
## $MPA_Reference
  UI.lobed.sponge.tunicate
                                       UI.lobed.sponge
                                                         California.sea.cucumber
##
                   0.1278430
                                             0.2486069
                                                                        0.3276743
##
           White.sea.urchin
                                   Short.red.gorgonian
                                                               Purple.sea.urchin
                                                                        0.4778682
                   0.3883098
                                             0.4344837
##
##
               Red.sea.star
                                  White.plumed.anemone
                                                           California.hydrocoral
                   0.5206175
##
                                             0.5596661
                                                                        0.5978635
##
         UI.branched.sponge
                                              Bat.star
                                                            UI.branched.bryozoan
##
                   0.6357187
                                             0.6703310
                                                                        0.7013164
central_sp <- c("UI lobed sponge/tunicate","UI lobed sponge",</pre>
                 "California sea cucumber", "White sea urchin",
                 "Short red gorgonian", "Purple sea urchin",
                 "Red sea star", "White plumed anemone",
                 "California hydrocoral", "UI branched sponge",
                 "Bat star", "UI branched bryozoan")
```

Species

central_con <- c(0.1278430,0.2486069,0.3276743,0.3883098,0.4344837,

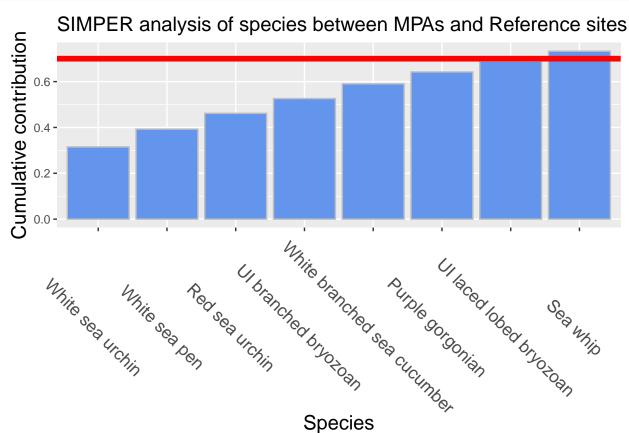
SIMPER analysis of species between MPAs and Reference sites



```
# South
south.labels <- South.avg[,1:4]
south.data <- South.avg[,5:160]
set.seed(13)
south.sim <- with(south.labels, simper(south.data, Designation))
south.sim

## cumulative contributions of most influential species:
##
## $MPA_Reference
## White.sea.urchin White.sea.pen
## 0.3150594 0.3927712</pre>
```

```
##
                Red.sea.urchin
                                        UI.branched.bryozoan
##
                      0.4616202
                                                   0.5259156
##
   White.branched.sea.cucumber
                                            Purple.gorgonian
                      0.5901728
                                                   0.6422199
##
##
       UI.laced.lobed.bryozoan
                                                    Sea.whip
##
                      0.6887477
                                                   0.7325898
south_sp <- c("White sea urchin","White sea pen","Red sea urchin",</pre>
              "UI branched bryozoan", "White branched sea cucumber",
              "Purple gorgonian", "UI laced lobed bryozoan", "Sea whip")
south_con \leftarrow c(0.3150594, 0.3927712, 0.4616202, 0.5259156, 0.5901728,
               0.6422199, 0.6887477, 0.7325898)
south_data <- data.frame(south_sp,south_con)</pre>
south graph <- ggplot(south data,</pre>
                       aes(y=south_con,x=factor(south_sp,level=south_sp))) +
  geom_bar(position="dodge", stat="identity", fill='#6495ed',color='gray') +
  ggtitle("SIMPER analysis of species between MPAs and Reference sites in Southern CA") +
  geom_hline(yintercept = 0.7,color='red',size=2) +
  theme(axis.text.x = element_text(angle = 315, size=13),
        plot.title = element_text(size=15),
        axis.title = element_text(size=15)) +
  xlab("Species") +
  ylab("Cumulative contribution")
south_graph
```



PERMANOVA

```
# PERMANOVA assumptions
# 1) Objects in the data set are exchangable under the
# null hypothesis
# 2) Exchangable objects (sites, samples, observations,
# etc.) are independent
# 3) Exchangable objects have similar multivariate
# dispension(i.e. each group has a similar degree of
# multivariate scatter)
# Basically, so long as the data is theoretically exchangable
# (one site COULD be identical to another) and the observations
# are independent, then the test meets assumptions and can be run.
# Run with Designation*Region
set.seed(13)
reg.div <- adonis2(reg.data ~ Region*Designation,</pre>
                   data = reg.labels, permutations = 999,
                   method = "bray")
reg.div
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
## adonis2(formula = reg.data ~ Region * Designation, data = reg.labels, permutations = 999,
method = "bray")
## Df SumOfSqs R2 F Pr(>F)
## Region 2 2.5605 0.28157 6.0456 0.001 ***
## Designation 1 0.2378 0.02615 1.1228 0.310
## Region:Designation 2 0.3658 0.04023 0.8637 0.618
## Residual 28 5.9295 0.65205
## Total 33 9.0936 1.00000
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
# Run with just Designation to see if there is a significant difference
set.seed(13)
reg.div2 <- adonis2(reg.data ~ Designation,</pre>
                    data = reg.labels, permutations = 999,
                    method = "bray")
reg.div2
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
## adonis2(formula = reg.data ~ Designation, data = reg.labels, permutations = 999, method =
"bray")
```

```
## Df SumOfSqs R2 F Pr(>F)
## Designation 1 0.2357 0.02592 0.8515 0.544
## Residual 32 8.8579 0.97408
## Total 33 9.0936 1.00000
```

No change in significance, use the interaction analysis in final results.