

# Morro Bay Seine Analyses

Maurice Goodman

Updated October 2020

This document contains analyses of beach seine data from **O’Leary et al., “Effects of estuary-wide seagrass loss on fish populations”**, submitted to the journal *Estuaries and Coasts* and in review as of September 2020. Not all code is shown in document output, but code for plots or analyses which are not shown below can be viewed in the R markdown source code.

## Metadata

### Data Description

Because the post-decline trawl surveys did not include any eelgrass habitat, and to account for potential temporal variations in fish communities unrelated to eelgrass, we evaluated fish abundance, size, and species composition in eelgrass habitat and adjacent unvegetated mudflats at two sites in the spring (April-May) of 2017 and 2018 (Figure 1b). The sites are in the middle portion of the bay (Windy Cove) and forebay near the mouth (Coleman’s Beach). The beach seine used was 10.2 m long, 1.1 m high, and had a mesh size of 4 mm.

Seines were conducted at a mean tide height of  $0.43 \pm 0.07$  m MLLW). On any given day, we ensured that the number of replicates in each habitat type (eelgrass and unvegetated mudflat) was equal, and we alternated between habitat types. At Windy Cove, we conducted three replicate seines in each habitat type in each year. Seines were conducted 20-25 meters apart and we conducted seines in increasing distances from the edge of the eelgrass (into the bed for eelgrass seines and away from the bed for unvegetated seines). Seines were a minimum of 10 meters from the edge and maximum of 175 m from the edge, with the locations approximately consistent each year. Water depths of the seines varied with bottom rugosity, with a mean water depth of  $1.0 \pm 0.14$  m in unvegetated mudflats and  $1.6 \pm 0.11$  m in eelgrass. Although these depths differ, given that the height of the seine is 1 m, the depth should not have affected overall catch. Distance of haul depended on depth and eelgrass bed width and ranged from 5.6 to 23.5 m (mean of  $15.9 \pm 1.9$  m) in the eelgrass and from 6.4 to 34 m (mean  $19.8 \pm 3.6$  m) in the mudflats. At Coleman’s Beach, because of adjacent building structures and more continuous eelgrass beds, unvegetated habitat was limited and we therefore conducted two replicate seines per year in each habitat type. Further, because of the spacing of unvegetated areas at the site, one of the two eelgrass seines per year was directly adjacent to the unvegetated habitat, while the other was approximately 10 m from the edge. Both unvegetated sites were directly adjacent to eelgrass in each year. Depth in eelgrass habitat at Coleman’s Beach was  $1.5 \pm 0.04$  m, and in the unvegetated habitat the depth was  $1.6 \pm 0.1$  m. Distance of haul ranged from 8.5 to 12.6 m (mean  $9.4 \pm 0.75$  m) in the eelgrass and ranged from 5.4 to 8.5 m in the unvegetated mudflat (mean  $5.9 \pm 0.54$  m).

Because of differences in haul distance, we standardized fish biomass and abundance by meter of haul for all seines. All fish caught were identified to species (except for young of the year rockfish, which we identified to the genus *Sebastes*). Total fish length was measured to the nearest mm and fish were released alive at the collection site.

### Sample Size Breakdown

Below is a table of sample sizes for each site (Colemans and Windy Cove), broken down by year and habitat (seagrass or unvegetated). The data are nearly balanced, with either 2 or 3 observations per site, habitat,

and year. This yields a total of 20 observations.

```
seine_data <-
  read_csv(here("data", "seine_fish_data.csv")) %>%
  filter(year == 2017 | year == 2018) %>%
  mutate(year = factor(year, levels = c("2017", "2018"), ordered = TRUE)) %>%
  filter(site_name %in% c("Windy Cove", "Colemans")) %>%
  mutate(date = as.Date(date, format = "%m/%d/%Y"),
         spec_binomial = paste(genus, species, sep = " "),
         spec_binomial = dplyr::recode(spec_binomial, "Sebastes NA" = "Sebastes YOY"),
         month = map_chr(date, ~month.name[lubridate::month(.x)]),
         `seagrass/unvegetated` = map_chr(`seagrass/unvegetated`, capitalize)) %>%
  filter(month %in% c("April", "May"))

seine_data %>%
  group_by(`seagrass/unvegetated`, year, site_name) %>%
  summarise(count = length(unique(seine_id))) %>%
  spread(`seagrass/unvegetated`, count) %>%
  knitr::kable()
```

year	site_name	Seagrass	Unvegetated
2017	Colemans	2	2
2017	Windy Cove	3	3
2018	Colemans	2	2
2018	Windy Cove	3	3

## Abundance

### Bar Chart

Overall, abundance (measured as numbers density per meter of seine) does not seem to differ between eelgrass and unvegetated habitat, in either the forebay (Colemans) or backbay (Windy Cove).

```
seine_count_data <-
  seine_data %>%
  group_by(year, site_name, `seagrass/unvegetated`,
           seine_id, distance_m, spec_binomial) %>%
  summarise(count = n()) %>% ungroup() %>%
  complete(nesting(year, site_name, `seagrass/unvegetated`,
                  seine_id, distance_m), spec_binomial, fill = list(count = 0)) %>%
  separate(spec_binomial, c("genus", "species"), sep = " ") %>%
  mutate(density = count/distance_m) %>%
  rename(seagrass_unvegetated = `seagrass/unvegetated`)

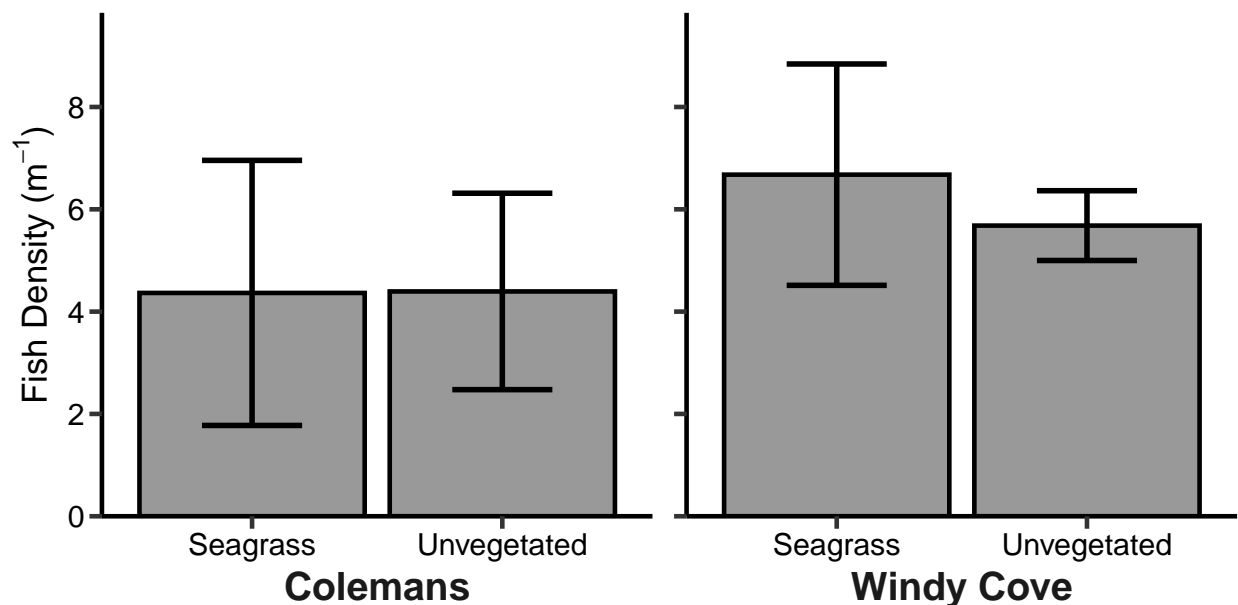
seine_summary <-
  seine_count_data %>%
  group_by(year, site_name, seagrass_unvegetated, seine_id) %>%
  summarize(density = sum(density)) %>% ungroup() %>%
  group_by(site_name, seagrass_unvegetated) %>%
  summarize(se = sd(density)/sqrt(n()),
           density = mean(density)) %>%
  complete(site_name, seagrass_unvegetated)

seine_summary %>%
```

```

ggplot(aes(seagrass_unvegetated, density)) +
  geom_bar(stat = "identity", position = "dodge", color = "black",
    fill = "grey60", size = 1) +
  geom_errorbar(aes(ymin = density - se, ymax = density + se),
    position = position_dodge(width = 0.8),
    width = 0.4, size = 1.2) +
  scale_y_continuous(expand = c(0,0), breaks = seq(0,12,2),
    limits = c(0, with(seine_summary, max(density + se + 1, na.rm = TRUE)))) +
  lemon::facet_rep_wrap(~site_name, strip.position = "bottom") +
  ylab(expression(paste("Fish Density (m-1)", "{-1}"))) +
  theme_trawls +
  coord_cartesian(clip = "off") +
  theme(axis.title.x = element_blank(),
    panel.spacing = unit(0, "lines"))

```



## Abundance Model

To assess differences in numbers density as a function of habitat, site, and their interaction, we fit a generalized linear model with a Gamma error distribution and a log link. The model confirms what is apparent from the bar chart above - numbers density is similar between both habitats and both sites.

```

seine_count_lm_data <-
  seine_count_data %>%
  group_by(year, site_name, seagrass_unvegetated, seine_id, distance_m) %>%
  summarize(count = sum(count)) %>%
  ungroup() %>%
  mutate(density = count/distance_m)

seine_count_model <- glm(density ~ site_name*seagrass_unvegetated,
  data = seine_count_lm_data,
  family = Gamma(link = "log"))

seine_count_model %>%
  car::Anova(test.statistic = "LR") %>% as.data.frame() %>%

```

```
rownames_to_column("term") %>%
mutate_if(is.numeric, round, digits = 3) %>%
knitr::kable()
```

term	LR Chisq	Df	Pr(>Chisq)
site_name	0.855	1	0.355
seagrass_unvegetated	0.070	1	0.791
site_name:seagrass_unvegetated	0.054	1	0.817

## Biomass

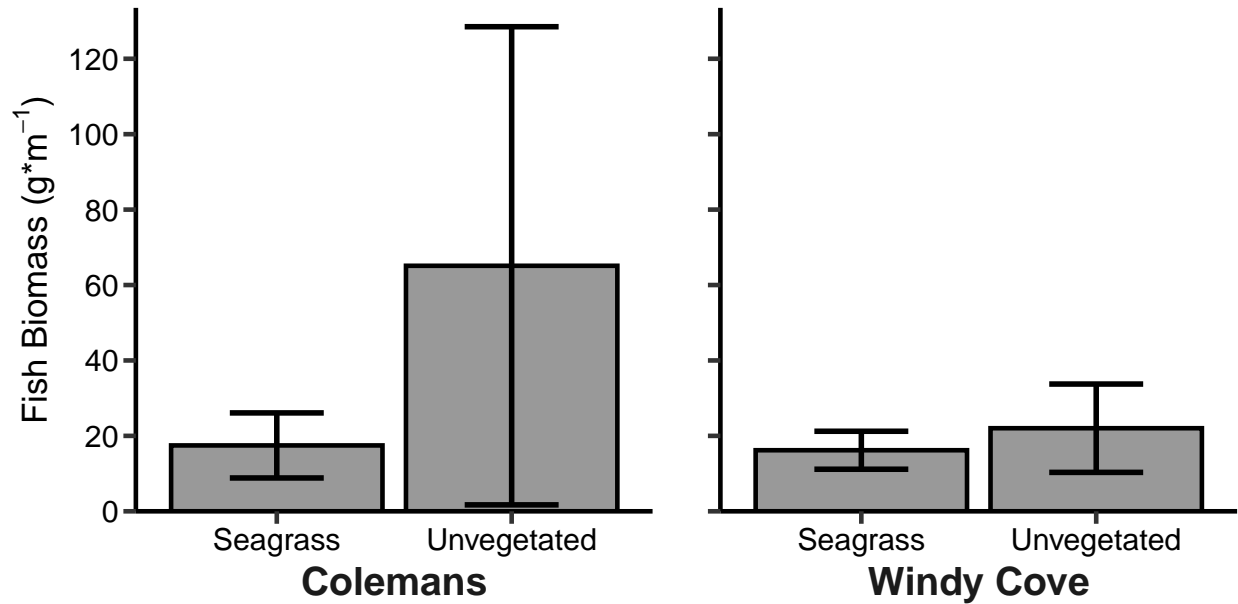
### Bar Chart

Here, biomass is estimated using individual fish lengths and length weight relationships from FishBase. Mean biomass is similar between the forebay seagrass habitat and the backbay seagrass and unvegetated habitat, but is markedly higher in the forebay unvegetated habitat, although all error bars overlap.

```
seine_biomass_lm_data <- seine_data %>%
  filter(!is.na(weight_g)) %>% # 1 bat ray (Myliobatis californica) observation
  group_by(year, site_name, `seagrass/unvegetated`, seine_id, distance_m) %>%
  summarize(biomass = sum(weight_g)) %>%
  ungroup() %>%
  mutate(scaled_biomass = biomass/distance_m) %>%
  rename("seagrass_unvegetated" = `seagrass/unvegetated`)

seine_biomass_summary <-
  seine_biomass_lm_data %>%
  group_by(site_name, seagrass_unvegetated) %>%
  summarize(mean = mean(scaled_biomass),
            se = sd(scaled_biomass)/sqrt(n()),
            lower = mean - se,
            upper = mean + se)

seine_biomass_summary %>%
  ggplot(aes(seagrass_unvegetated, mean)) +
  geom_bar(stat = "identity", position = "dodge", color = "black",
          fill = "grey60", size = 1) +
  geom_errorbar(aes(ymin = lower, ymax = upper),
               position = position_dodge(width = 0.8),
               width = 0.4, size = 1.2) +
  scale_y_continuous(expand = c(0,0), breaks = seq(0,130,20),
                     limits = c(0, max(seine_biomass_summary$upper + 5))) +
  lemon::facet_rep_wrap(~site_name, strip.position = "bottom") +
  ylab(expression(paste("Fish Biomass (g*m"^{-1}), "))) +
  theme_trawls +
  theme(axis.title.x = element_blank(),
        panel.spacing = unit(0, "lines"))
```



## Biomass Model

To assess differences in biomass as a function of site and habitat, we fit another generalized linear model with a Gamma error distribution and a log link. The model suggests that, as for numbers density, biomass is similar between both sites and habitats.

```
biomass_model <- glm(scaled_biomass ~ site_name*seagrass_unvegetated,
  data = seine_biomass_lm_data,
  family = Gamma(link = "log"))

biomass_model %>% car::Anova(test = "LR") %>% as.data.frame() %>%
  rownames_to_column("term") %>%
  mutate_if(is.numeric, round, digits = 3) %>%
  knitr::kable()
```

term	LR Chisq	Df	Pr(>Chisq)
site_name	1.057	1	0.304
seagrass_unvegetated	1.443	1	0.230
site_name:seagrass_unvegetated	0.750	1	0.386

## Species Composition

### NMDS Plot

We performed a 2-dimensional non-metric multi-dimensional scaling (NMDS) procedure on a Bray-Curtis dissimilarity matrix of the seine species number-density data, which resulted in a plot with a stress value of 0.13.

```
## Convert data to wide-format species counts
seine_sp_data <- seine_data %>%
  group_by(year, site_name, `seagrass/unvegetated`, seine_id, distance_m, spec_binomial) %>%
  summarise (Count = n()) %>% ungroup() %>%
  mutate(density = Count/distance_m) %>%
```

```

dplyr::select(-Count) %>%
spread(spec_binomial, density, fill = 0) %>%
rename(seagrass_unvegetated = `seagrass/unvegetated`)

## Species matrix
ordination_groups <- with(seine_sp_data, paste(seagrass_unvegetated, site_name))
seine_sp_matrix <- as.matrix(seine_sp_data[,6:23])
rownames(seine_sp_matrix) <- ordination_groups

## Bray-Curtis dissimilarity matrix
seine_bray <- vegdist(seine_sp_matrix, method = "bray")

## Fit NMDS
seine_mds <- metaMDS(seine_bray, k = 2)

## Format data for plotting
mds_plot_data <- seine_mds$points
mds_rownames <- rownames(mds_plot_data)
mds_plot_data <- as.data.frame(mds_plot_data, row.names = FALSE)
mds_plot_data$seagrass <- mds_rownames
nmbs_colors <- RColorBrewer::brewer.pal(6, "Paired")[c(2, 6)]

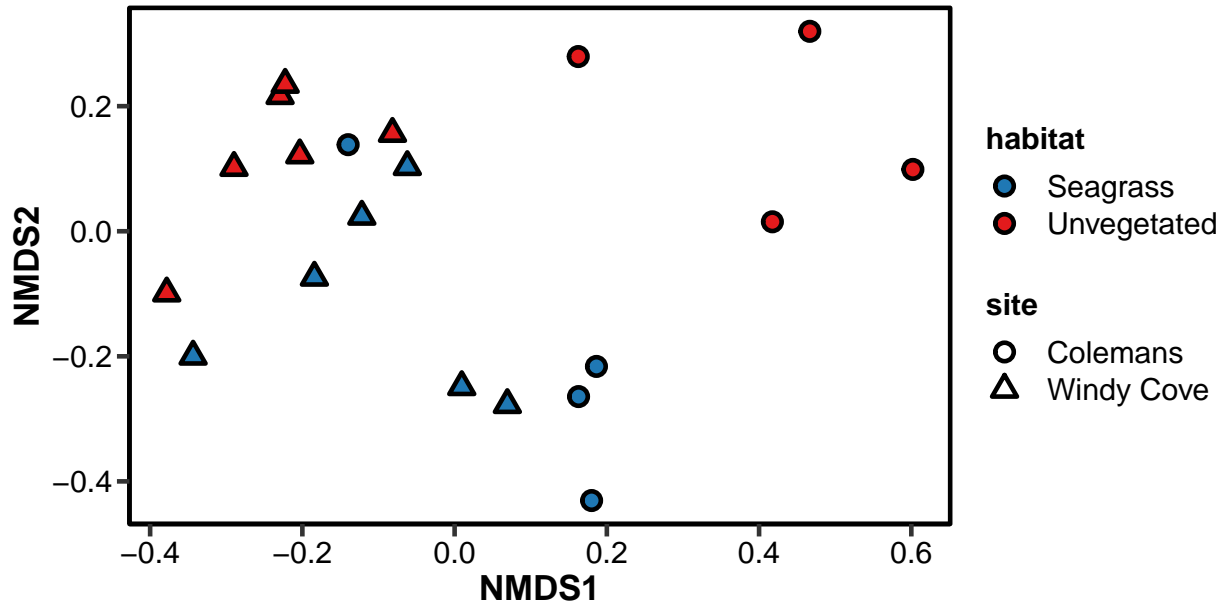
```

The plot shows marked, but non-monotonic, differences in species composition between seagrass and unvegetated habitats. Specifically, fish communities in unvegetated habitat are distinct between the forebay and the backbay, but there is considerable overlap between seagrass communities from the forebay and backbay. In short, unvegetated communities display greater  $\beta$ -diversity than seagrass communities.

```

mds_plot_data %>%
  separate(seagrass, c("seagrass", "site"), sep = " ", extra = "merge") %>%
  ggplot(aes(MDS1, MDS2, fill = seagrass, pch = site)) +
  geom_point(size = 3, stroke = 1.5) +
  scale_fill_manual(values = nmbs_colors) +
  scale_shape_manual(values=c(21, 24)) + theme_bw() +
  labs(x = "NMDS1", y = "NMDS2", fill = "habitat") +
  theme_trawls +
  cowplot::panel_border("black", 1) +
  coord_cartesian(clip = "off") +
  guides(fill=guide_legend(override.aes=list(shape=21)))

```



## PERMANOVA

We fit a permutational multivariate analysis of variance (PERMANOVA) on a Bray-Curtis dissimilarity matrix of species abundance with 10,000 permutations. The results agree with the qualitative results from the NMDS plot - fish communities are distinct between seagrass and unvegetated habitats, between the forebay and the backbay, and there are further differences not explained by habitat or site alone.

```
seine_permanova <- adonis2(
  seine_bray ~ site_name*seagrass_unvegetated,
  data = seine_sp_data,
  perm = 10000
)

seine_permanova %>% as.data.frame() %>%
  rownames_to_column("term") %>%
  mutate_if(is.numeric, round, digits = 3) %>%
  knitr::kable()
```

term	Df	SumOfSqs	R2	F	Pr(>F)
site_name	1	1.155	0.195	5.877	0.000
seagrass_unvegetated	1	0.979	0.165	4.982	0.000
site_name:seagrass_unvegetated	1	0.644	0.109	3.277	0.007
Residual	16	3.143	0.531		
Total	19	5.920	1.000		

## Pairwise Comparisons

Because the PERMANOVA displayed a significant interaction, we fit pairwise PERMANOVA models, each with 10,000 permutations, and adjusted p-values using the Benjamini & Hochberg method.

The results suggest that, as evident from the NMDS plot, seagrass communities are similar between the forebay and the backbay. However, in both habitats, seagrass communities are distinct from unvegetated communities, and the unvegetated communities from the forebay and the backbay are also distinct.

```
## Function to fit pairwise PERMANOVA models
source(here("R", "pairwise_permanova.R"))
```

```
seine_pairs <- pairwise_permanova(
  seine_sp_matrix,
  group_var = ordination_groups
)
```

```
knitr::kable(seine_pairs$contrasts)
```

group1	group2	R2	F_value	df1	df2	p_value
Seagrass Colemans	Unvegetated Colemans	0.404	4.068	1	7	0.037
Seagrass Colemans	Seagrass Windy Cove	0.146	1.367	1	9	0.179
Seagrass Colemans	Unvegetated Windy Cove	0.426	5.943	1	9	0.014
Unvegetated Colemans	Seagrass Windy Cove	0.399	5.304	1	9	0.014
Unvegetated Colemans	Unvegetated Windy Cove	0.556	10.015	1	9	0.014
Seagrass Windy Cove	Unvegetated Windy Cove	0.277	3.839	1	11	0.014

## SIMPER

For two pairwise comparisons of interest - that between the seagrass and unvegetated communities in the forebay, and that between the seagrass and unvegetated communities in the backbay, we fit a similarity percentages (SIMPER) analysis, reporting numbers density for species identified by the SIMPER as contributing to observed differences in community composition more than would be expected at random. For each SIMPER output, relevant species are listed, and an abundance plot is provided.

**How to interpret this output:** This output is ordered in terms of each species contribution to the difference between groups examined. The **average** variable is the average contribution to the overall dissimilarity, and the **sd** is the standard deviation of that contribution. The **cumsum** variable is the total percent of the variation accounted for by the species and all other species above it.

Code is shown in output for the first SIMPER only, as the code is near-id

### Coleman's Beach

```
colemans_simper <-
  seine_sp_matrix[seine_sp_data$site_name == "Colemans",] %>%
  filter_sp() %>%
  simper(seine_sp_data$seagrass_unvegetated[seine_sp_data$site_name == "Colemans"]) %>%
  summary %>% Seagrass_Unvegetated %>%
  format_simper() %>%
  rename(seagrass = ava, unvegetated = avb)
```

```
knitr::kable(colemans_simper)
```

Species	average	sd	ratio	seagrass	unvegetated	cumsum	n
Citharichthys stigmaeus	0.38	0.21	1.82	0.08	3.15	0.42	14
Leptocottus armatus	0.16	0.18	0.93	2.00	0.35	0.60	14
Syngnathus leptorhynchus	0.11	0.09	1.13	1.19	0.03	0.72	14
Heterostichus rostratus	0.08	0.11	0.76	0.42	0.00	0.81	14

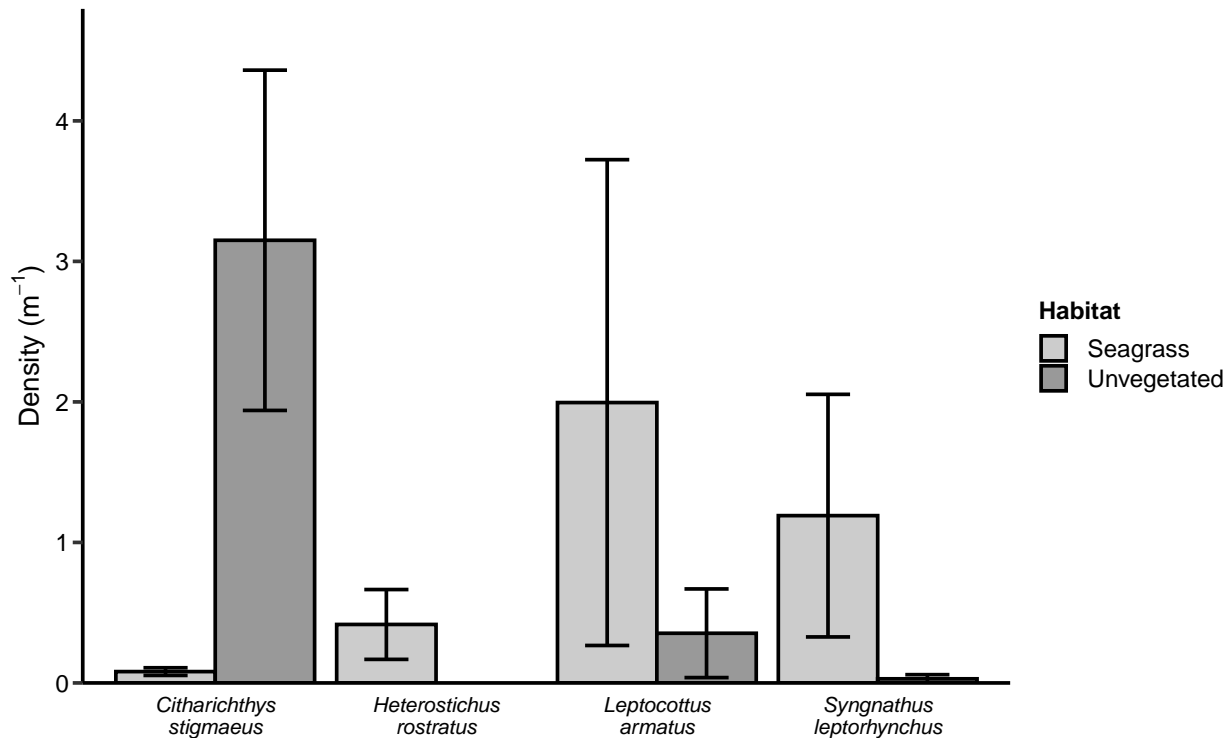


```

coleman_simper_summary <- seine_sp_data %>%
  filter(site_name == "Colemans") %>%
  sum_seine_abundance(colemans_simper$Species, "seagrass_unvegetated")

coleman_simper_summary %>%
  plot_seine_abundance("seagrass_unvegetated") +
  guides(fill = guide_legend("Habitat"))

```



## Windy Cove

```

windy_simper <-
  seine_sp_matrix[seine_sp_data$site_name == "Windy Cove",] %>%
  filter_sp() %>%
  simper(seine_sp_data$seagrass_unvegetated[seine_sp_data$site_name == "Windy Cove"]) %>%
  summary %$% Seagrass_Unvegetated %>%
  format_simper() %>%
  rename(seagrass = ava, unvegetated = avb)

knitr::kable(windy_simper)

```

Species	average	sd	ratio	seagrass	unvegetated	cumsum	n
Leptocottus armatus	0.28	0.22	1.25	1.65	4.15	0.40	13
Atherinops affinis	0.22	0.22	0.99	2.59	1.43	0.72	13
Syngnathus leptorhynchus	0.12	0.09	1.41	1.56	0.00	0.89	13

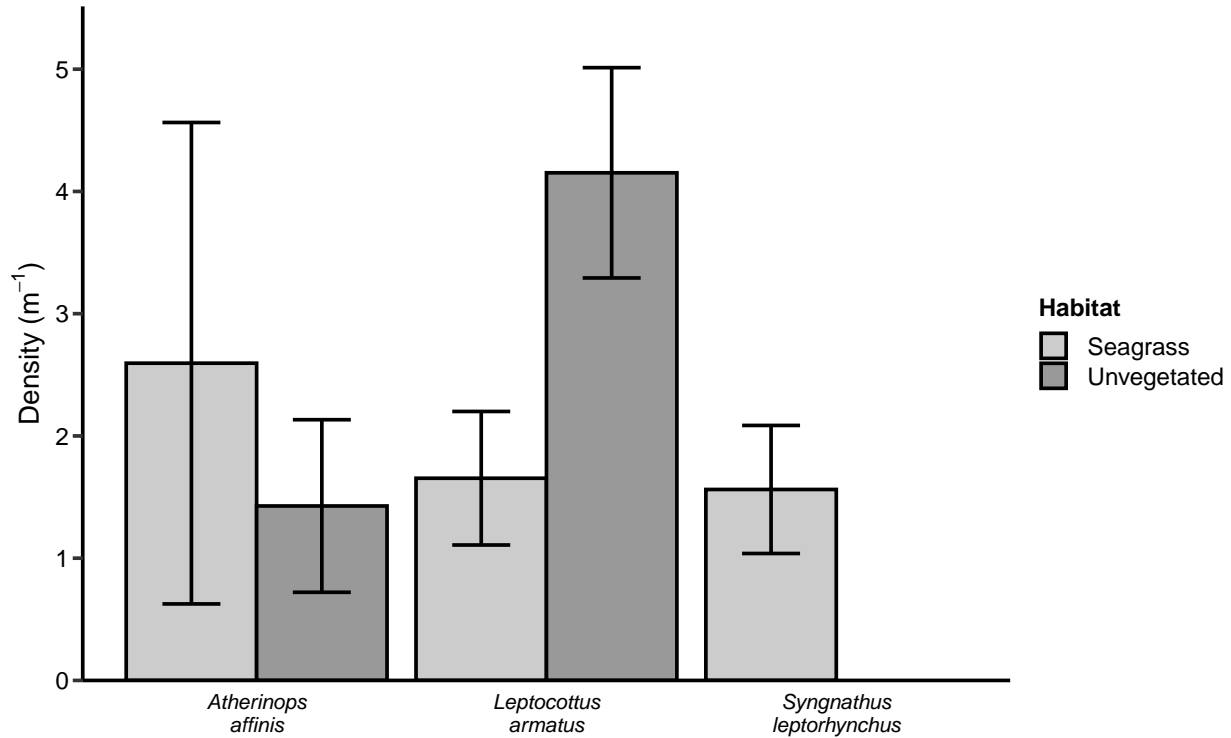
```

windy_simper_summary <- seine_sp_data %>%
  filter(site_name == "Windy Cove") %>%

```

```
sum_seine_abundance(windy_simper$Species, "seagrass_unvegetated")

windy_simper_summary %>%
  plot_seine_abundance("seagrass_unvegetated") +
  guides(fill = guide_legend("Habitat"))
```



### Combined SIMPER Plot

Abundance of species contributing to observed differences in community structure between unvegetated habitats (dark gray) and eelgrass (light gray) based on a SIMPER analysis from beach seine data for: Coleman Park Beach (mouth), and Windy Cove (mid-bay). Where there were no individuals of a species at a site, this is indicated with a zero.

