

Capstone Project Fish Species Analysis

Introduction to Data Science

Kevin Tajkowski

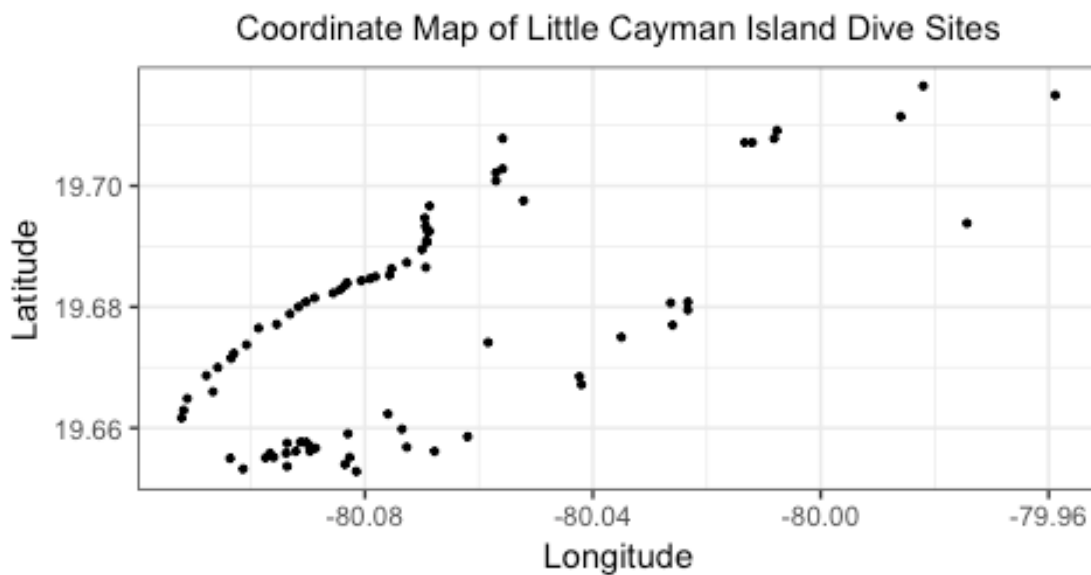
Library Calls

Load the necessary packages.

```
library(readr)
library(dplyr)
library(stringr)
library(ggplot2)
library(grid)
library(gridExtra)
```

Simple Visualization

Use the Zone_Codes file to generate a plot showing the latitude and longitude of all dive sites at Little Cayman.



Data Preparation

Load the data into a data frame. Remove all non-fish species by Family_Name. Non-fish species observed in the data set include sharks, rays, eels, turtles and a dolphin.

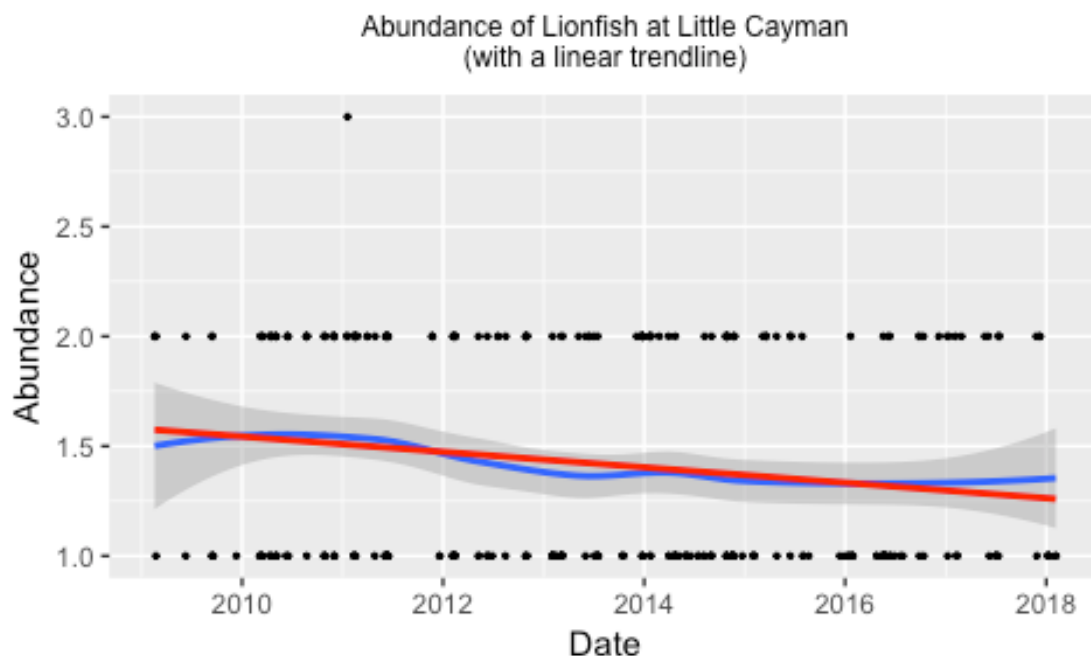
```
Little_Cayman <- read_csv("Little_Cayman.csv")
non_fish_family <- c("Requiem Shark", "Conger", "Dolphin", "Stingray",
                    "Manta", "Moray", "Eagle Ray", "Snake Eel", "Carpet Shark", "Hammerhead Shark",
                    "Electric Ray", "Round Stingrays", "Sea Turtles", "#N/A")
LC <- Little_Cayman %>% filter(!(Family_Name %in% non_fish_family))
```

Lionfish Exploration

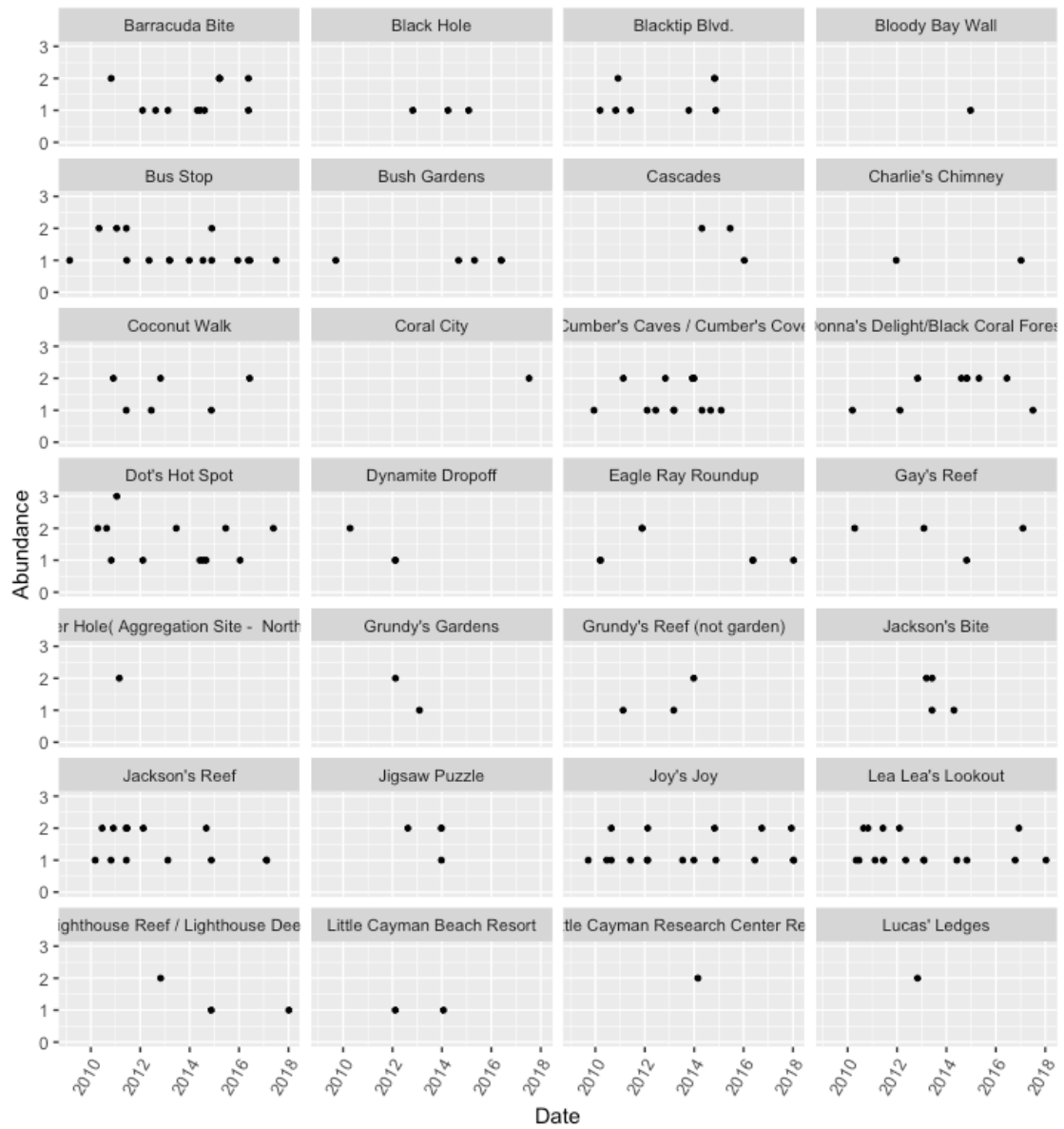
The red lionfish, otherwise known as the common lionfish, is species 683 and is a member of the scorpionfish family. Here is a list of the 5 species of scorpionfish observed at Little Cayman.

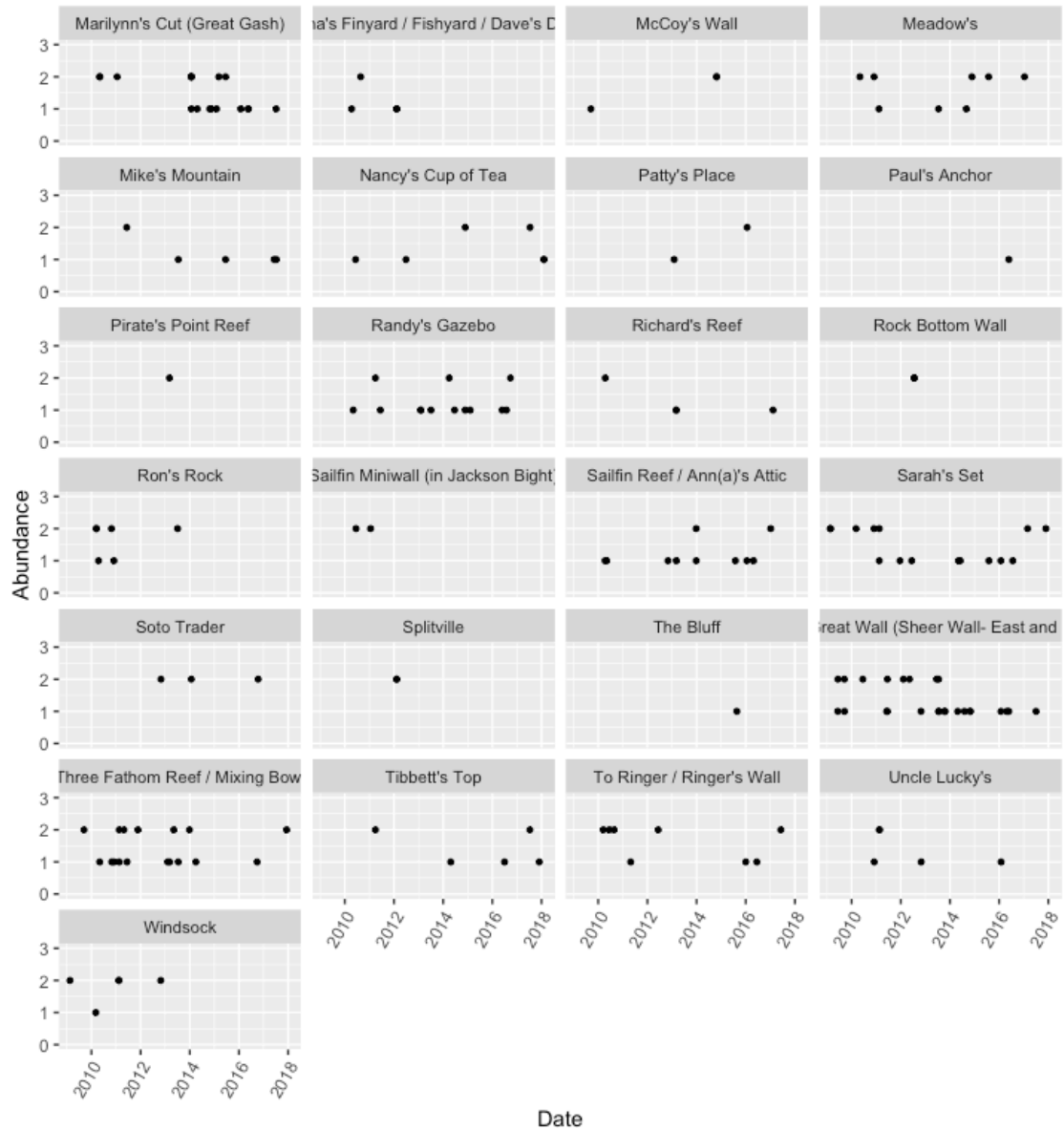
```
## # A tibble: 5 x 2
##           English_Name count
##           <chr> <int>
## 1 Mushroom Scorpionfish     2
## 2 Plumed Scorpionfish       2
## 3 Red Lionfish (aka Common Lionfish) 347
## 4 Reef Scorpionfish         1
## 5 Spotted Scorpionfish    144
```

Generate plots showing lionfish abundance as a function of time. The following figure shows the lionfish presence at all dive sites at Little Cayman. A smooth curve with corresponding confidence interval and a trendline were added to the plot.



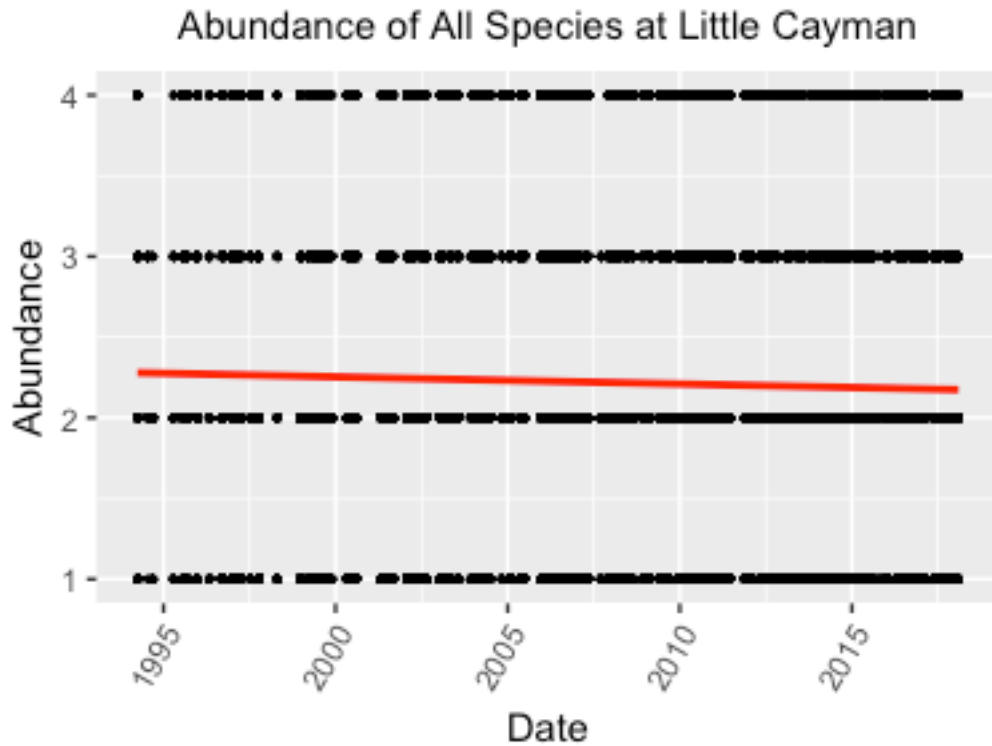
The following series of plots show the lionfish presence at numerous dive sites at Little Cayman.



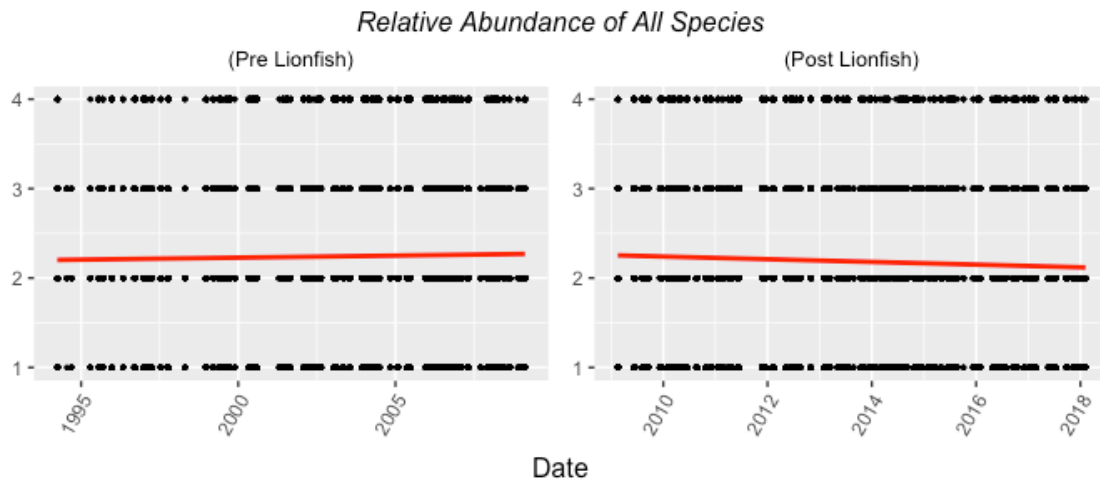


Fish Exploration at Little Cayman

This plot shows the abundance of all species in all families for all survey data at Little Cayman.



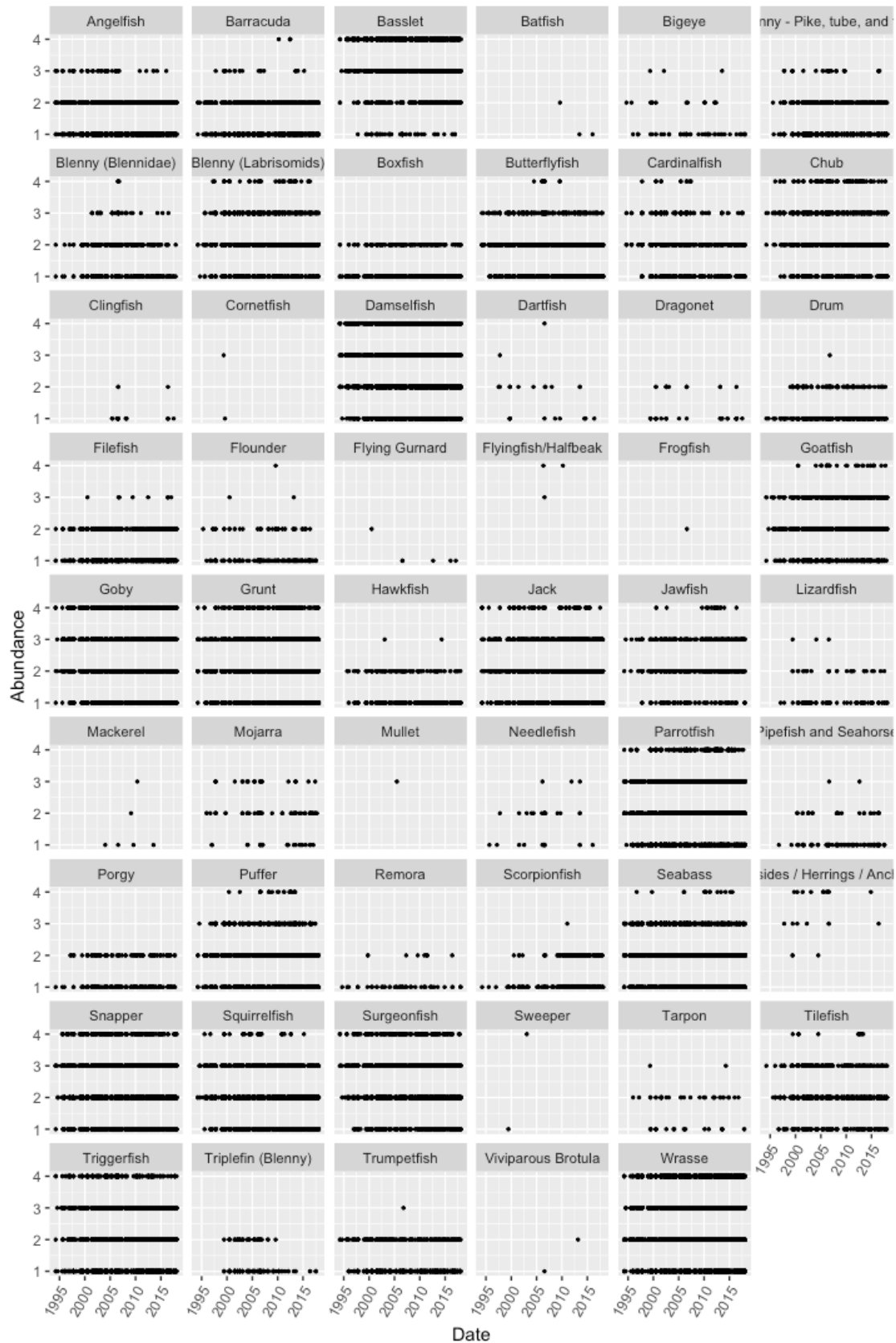
The next two plots show the species abundance for all species and corresponding trendlines separated into two groups: before the lionfish arrived and after.



Fish Family Exploration for all Sites

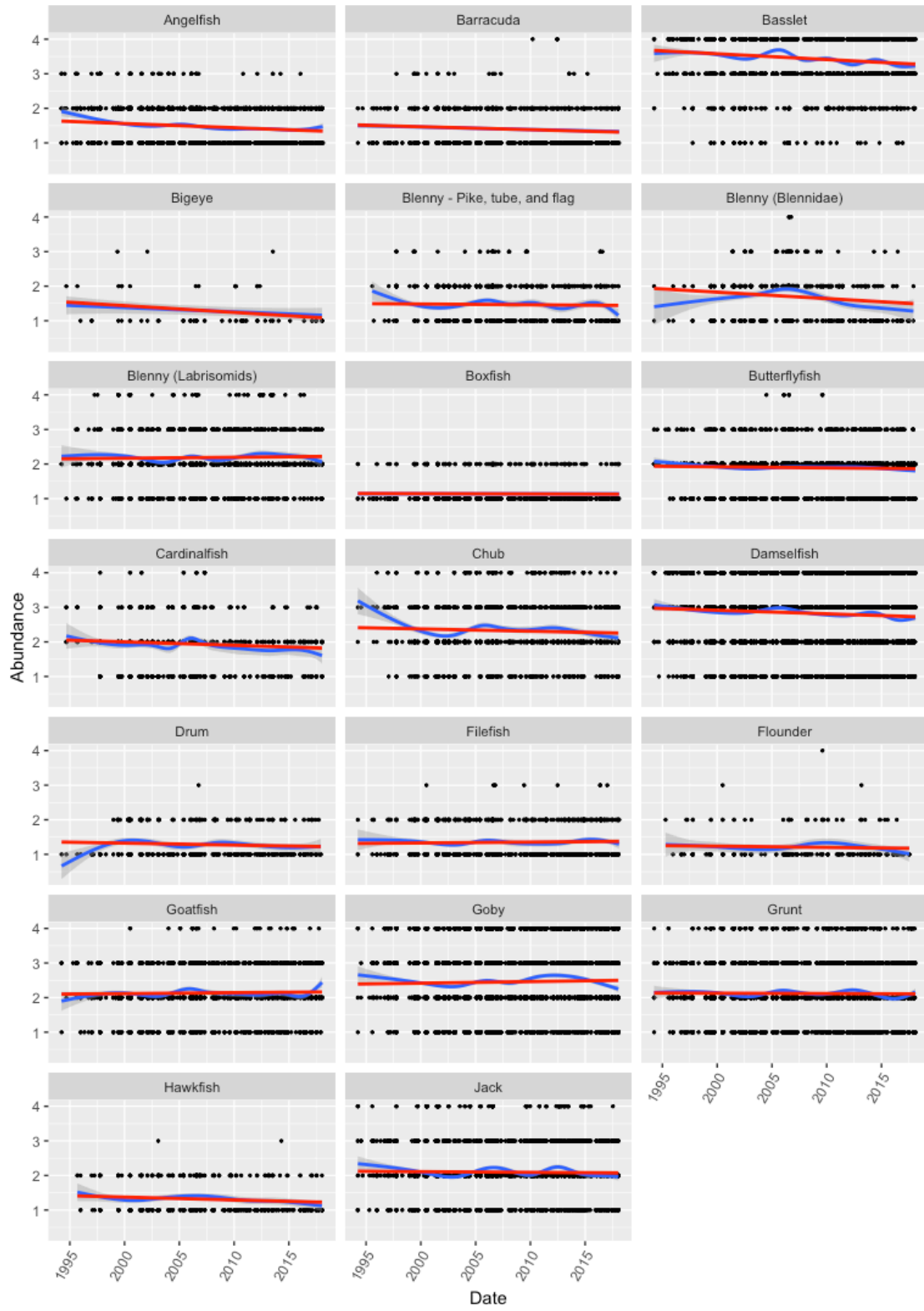
Generate plots of all fish families at all dive sites at Little Cayman using the Family_Name.

Abundance of Fish Families at Little Cayman

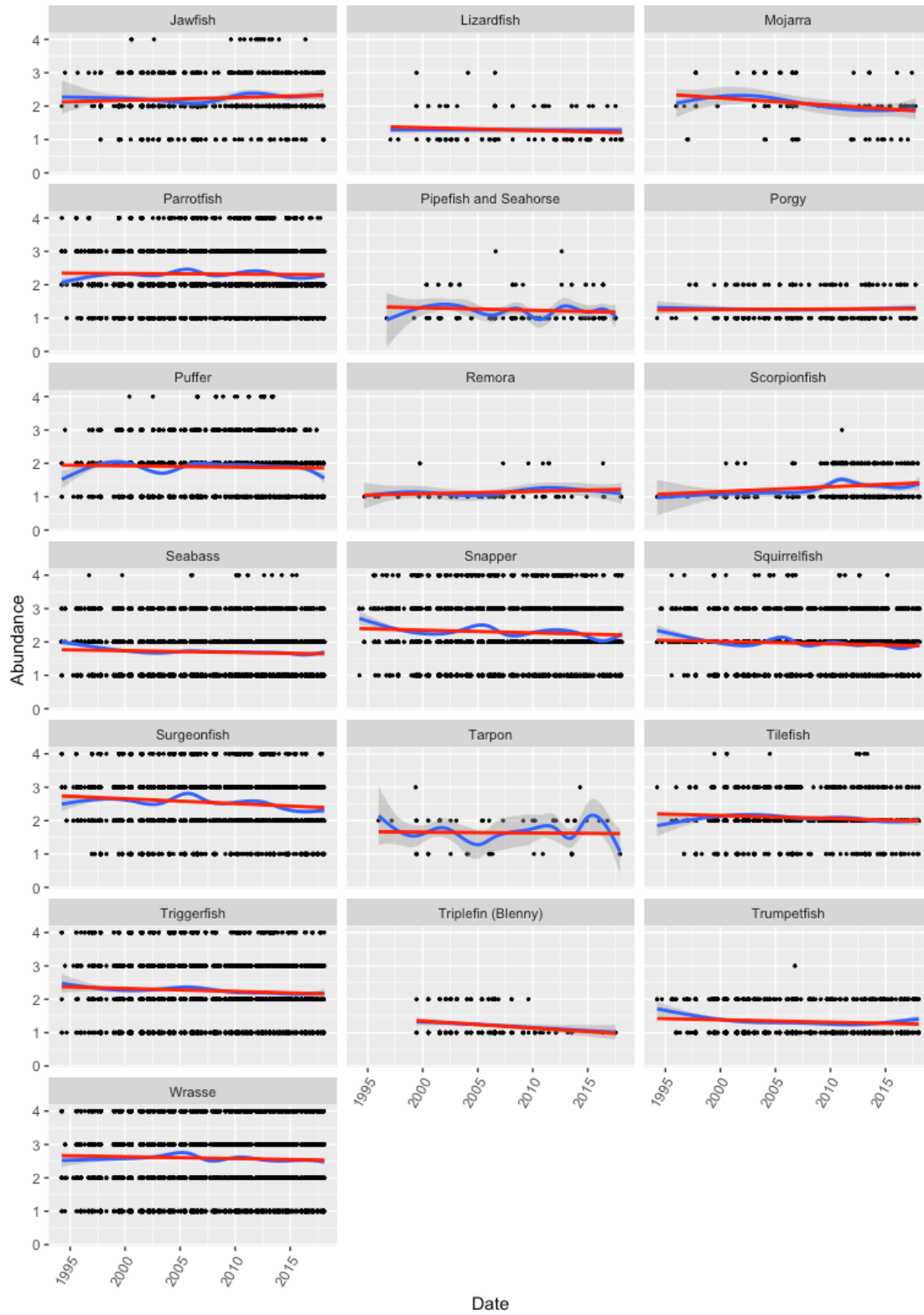


Fish families with less than thirty surveys for all species in the family are identified and removed from the data set. The abundance vs. time of numerous families is then plotted side-by-side with a smooth curve (including confidence interval) added to each family plot.

Abundance of Fish Families at Little Cayman



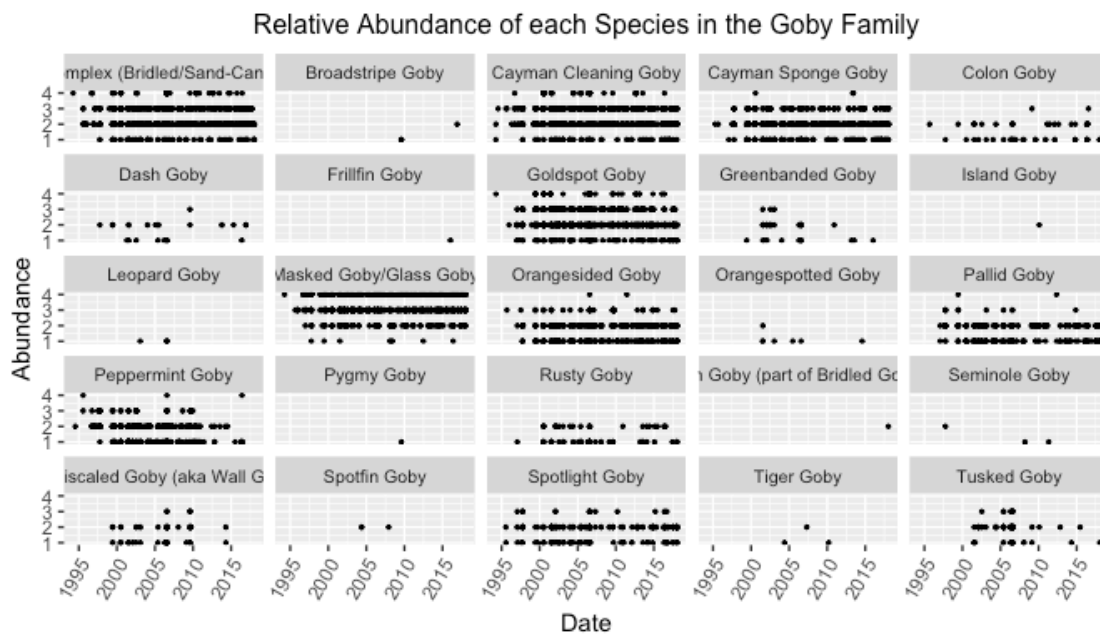
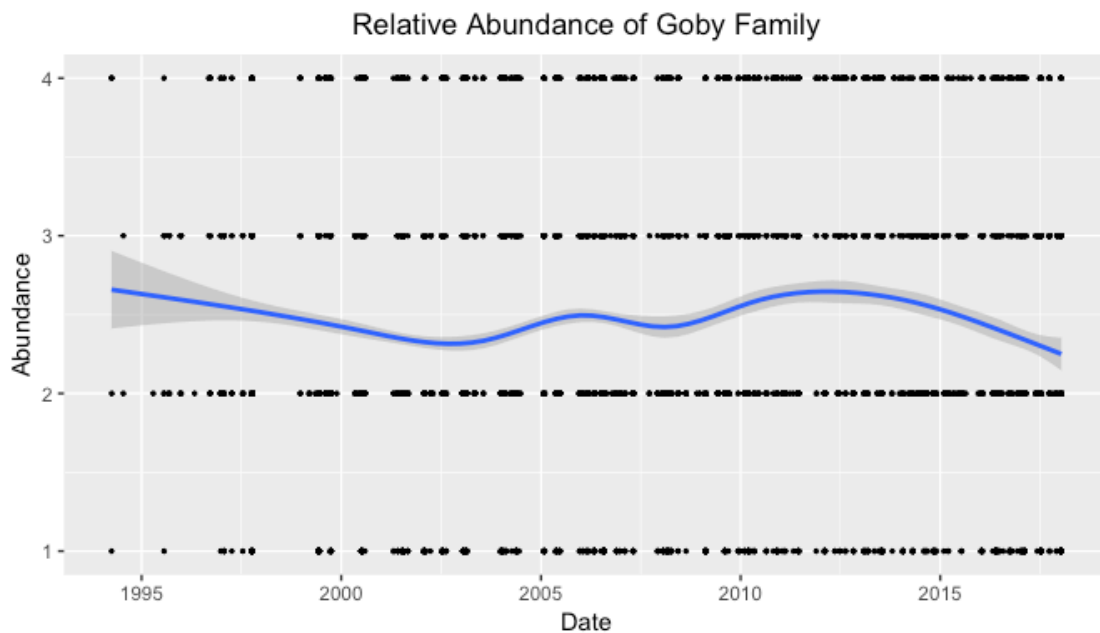
Abundance of Fish Families at Little Cayman



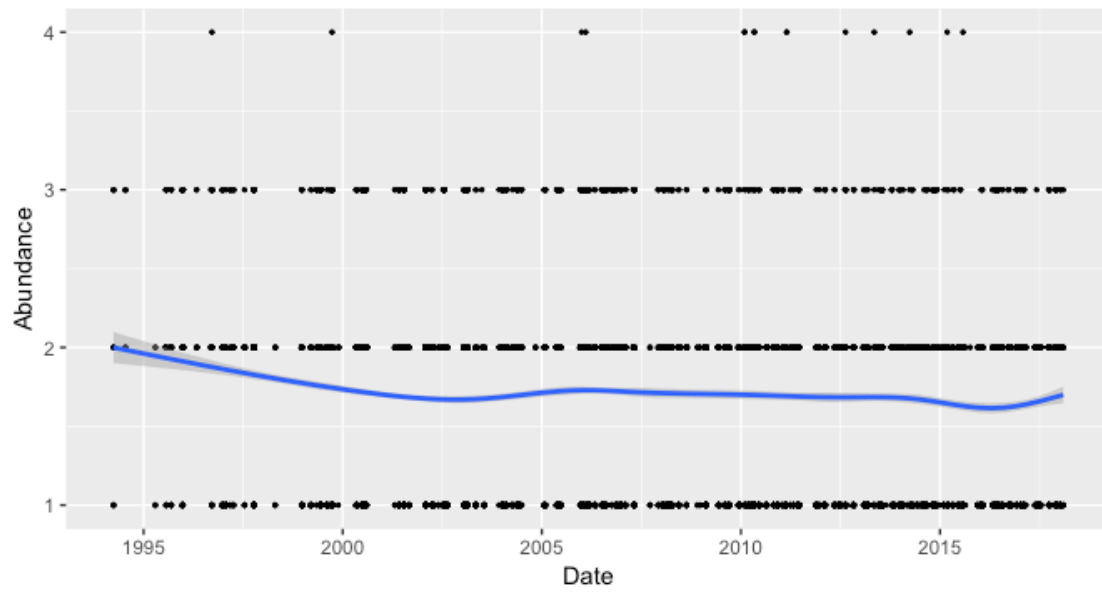
Fish Species Exploration for Specific Families

The following pairs of plots show the abundance vs. time for specific fish families. The first plot shows the entire family at all sites at Little Cayman. A smooth curve and confidence interval are added to this plot. The second plot shows all of the species within each family.

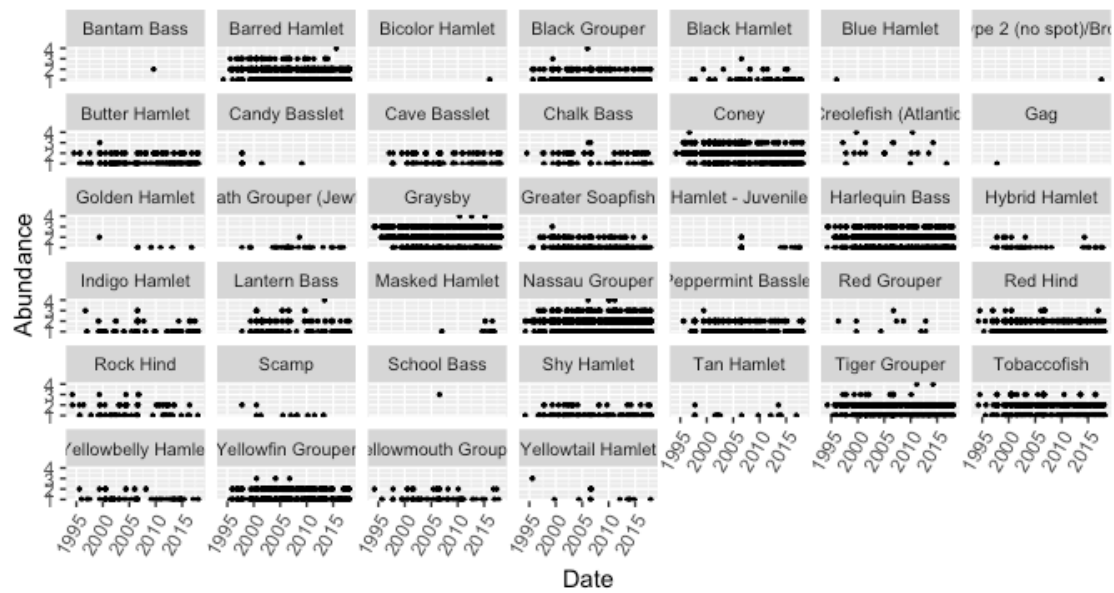
Grouper and snapper are prime predators in the Caribbean waters. Grouper are part of the seabass family. The juveniles of these species, the fish in the goby family and the fish in the wrasse family have small, cigar-shaped bodies, ideal prey for the lionfish.



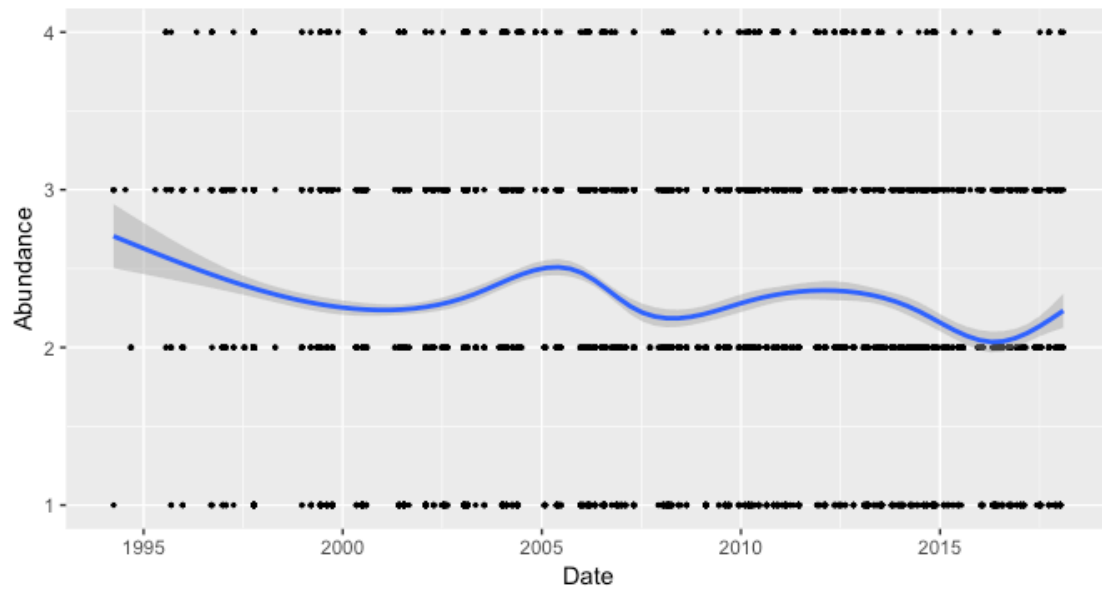
Relative Abundance of Seabass Family



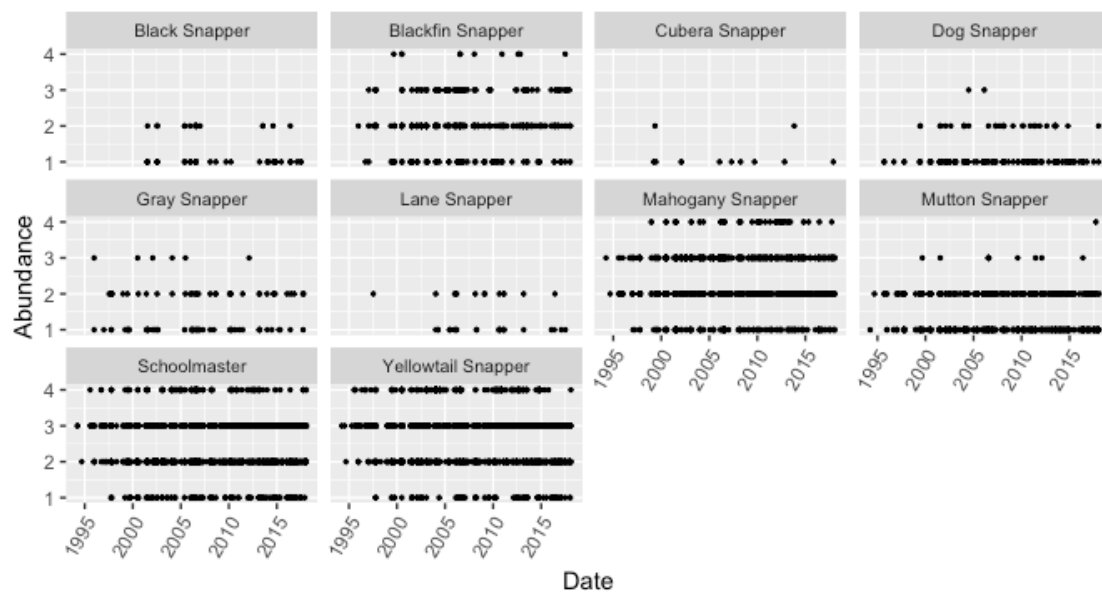
Relative Abundance of each Species in the Seabass Family

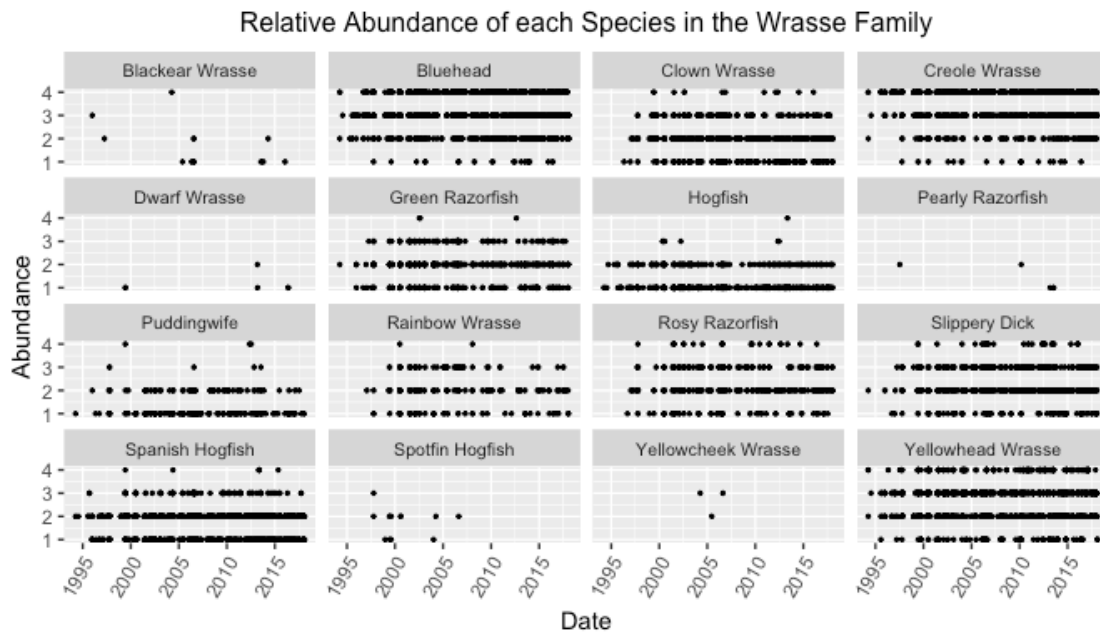
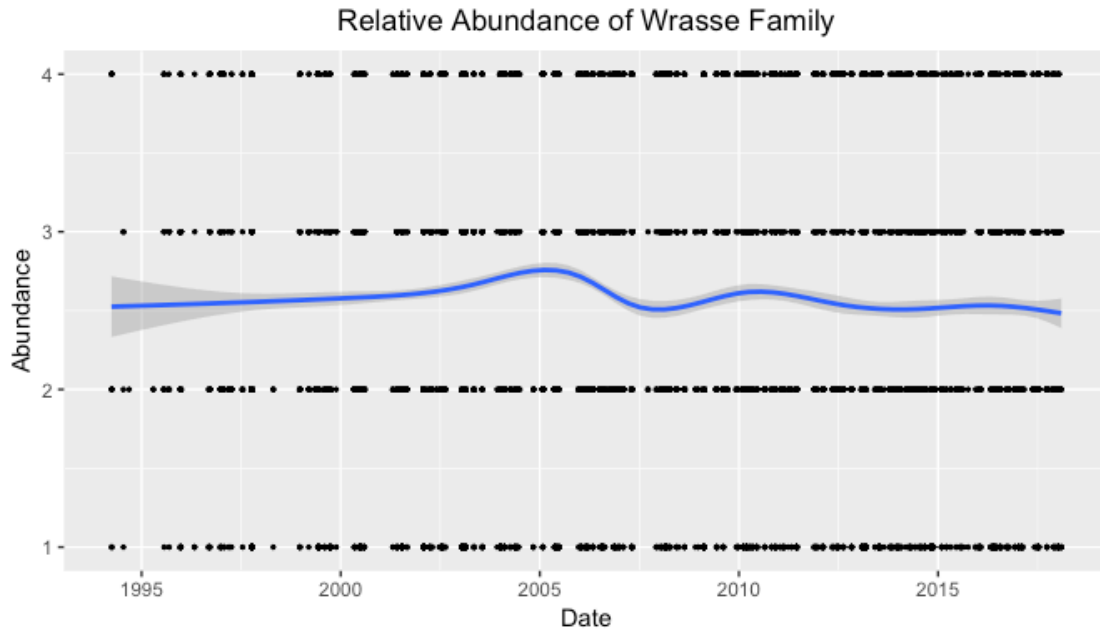


Relative Abundance of Snapper Family



Relative Abundance of each Species in the Snapper Family

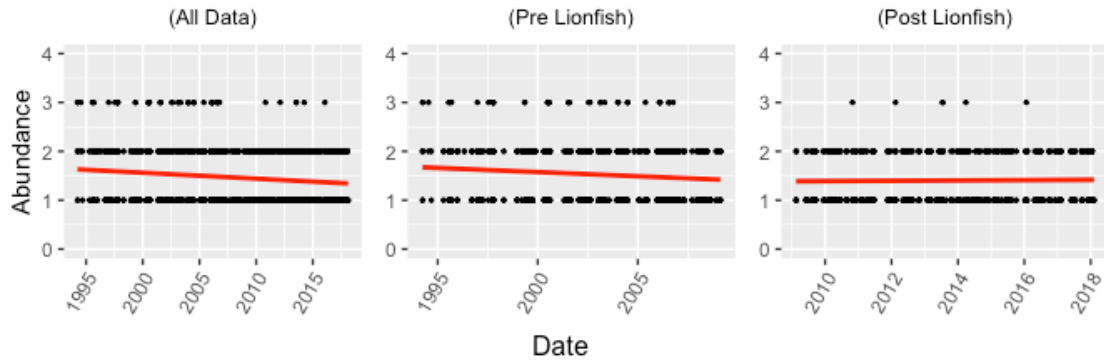




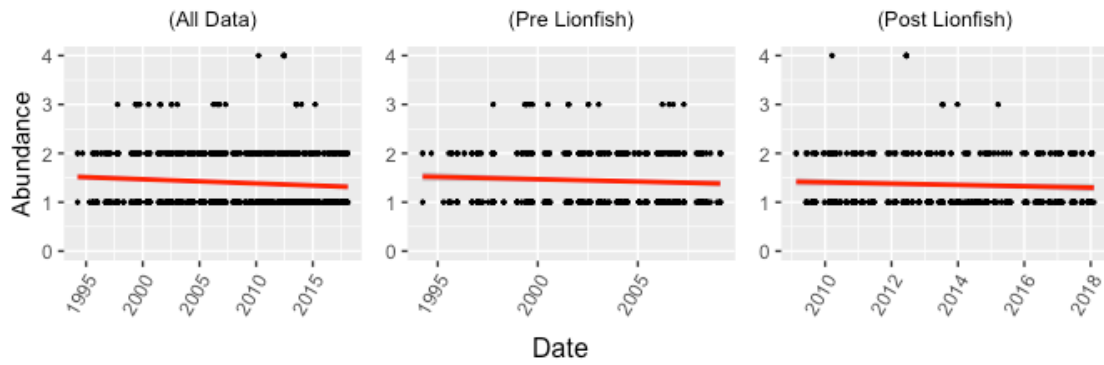
Exploration of All Families Before and After Lionfish Invasion

The following plots show the survey data for each fish family. The plot on the left shows all the data for the named family, the center plot shows the data from when the surveys began to when the first lionfish sighting occurred (2009-02-17) and the plot on the right shows the data from the first lionfish sighting to when the data was obtained. Each sub plot has a corresponding trendline added.

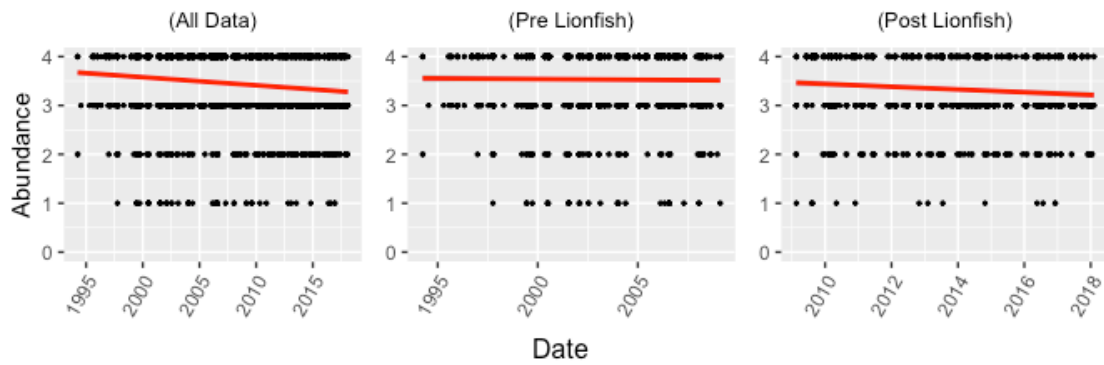
Relative Abundance of the Angelfish Family



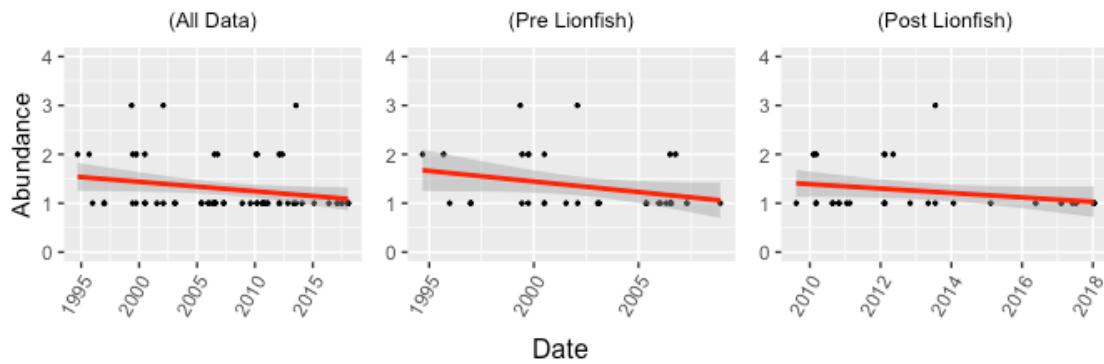
Relative Abundance of the Barracuda Family

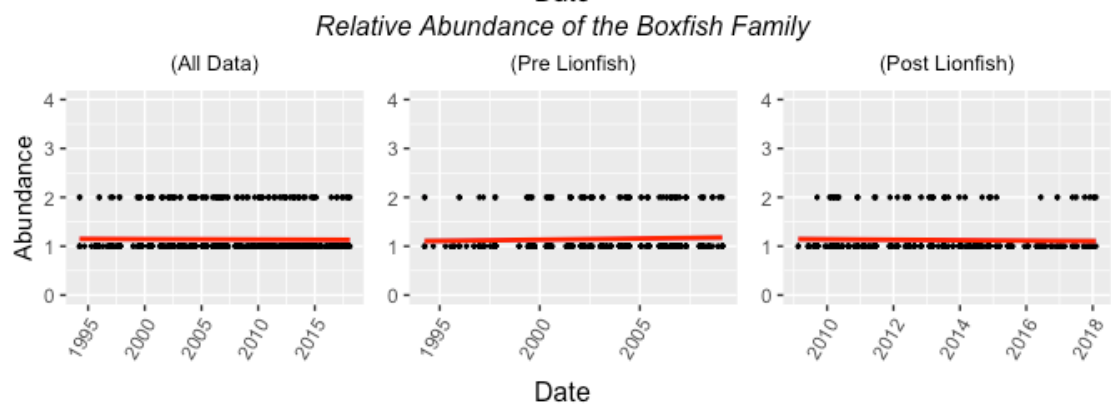
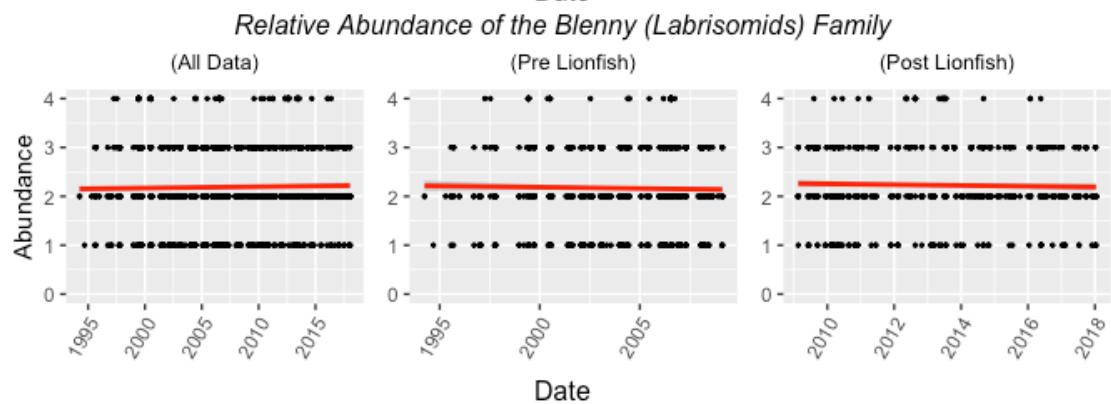
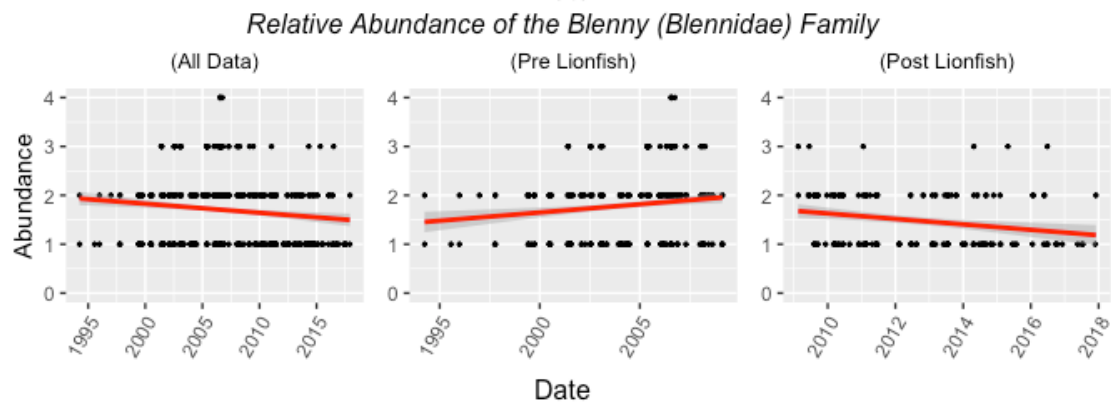
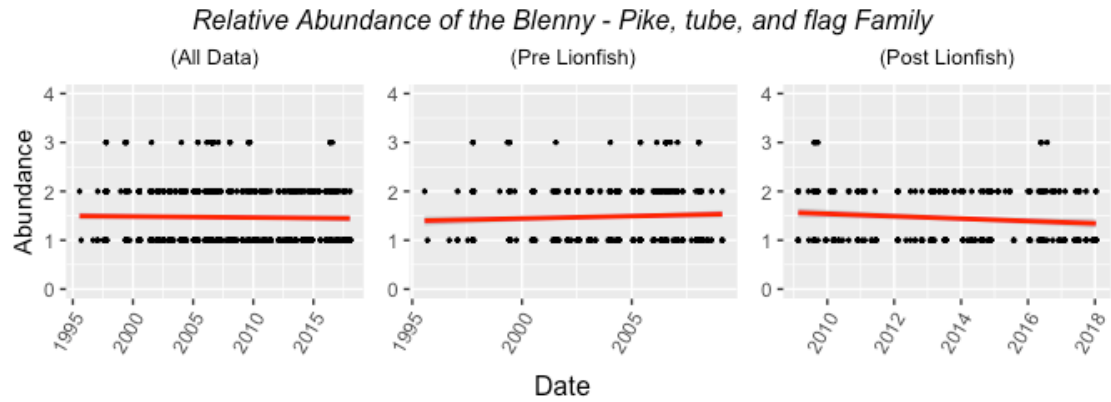


Relative Abundance of the Basslet Family

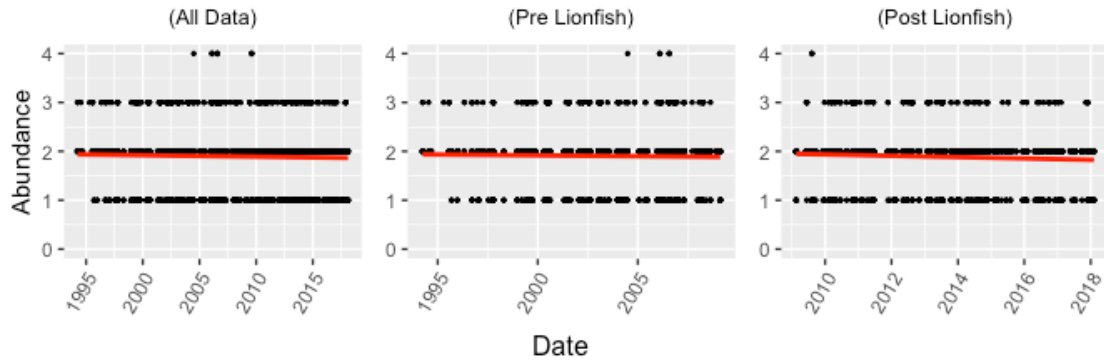


Relative Abundance of the Bigeye Family

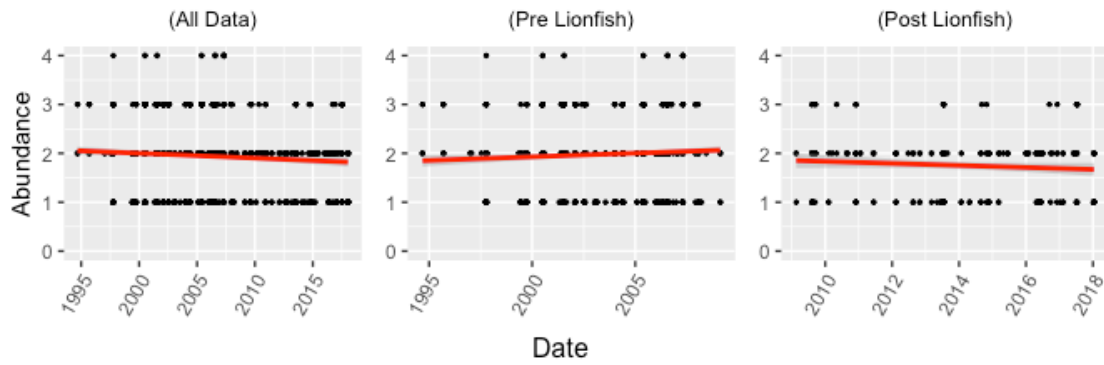




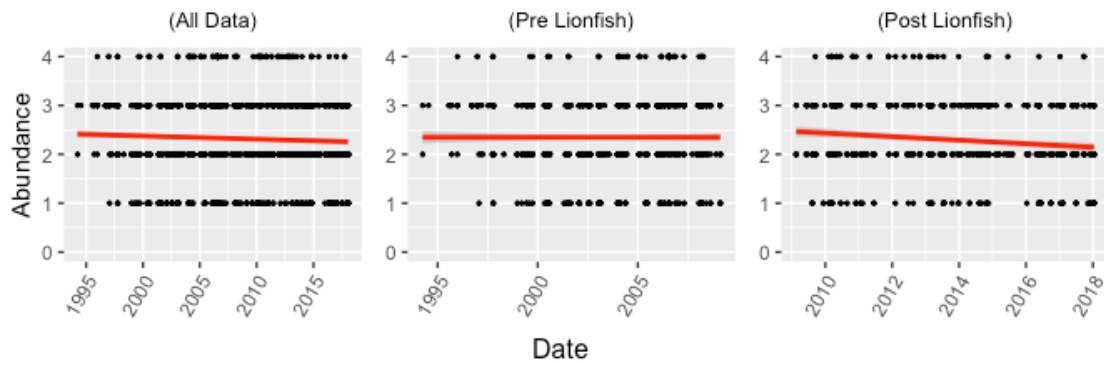
Relative Abundance of the Butterflyfish Family



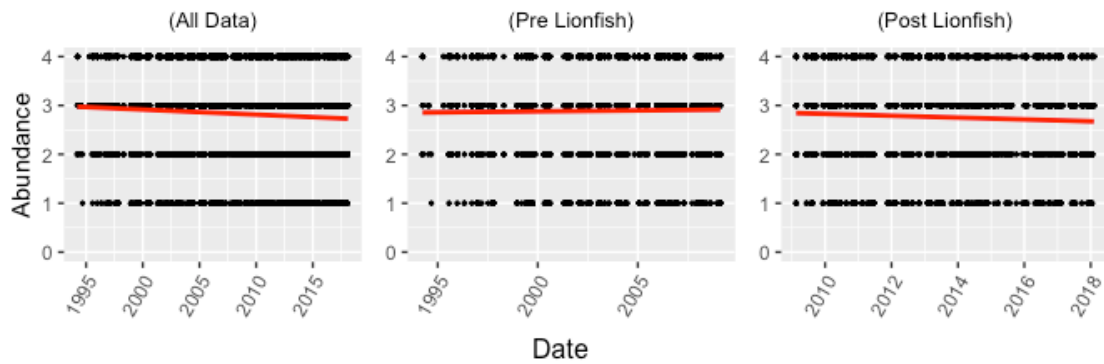
Relative Abundance of the Cardinalfish Family



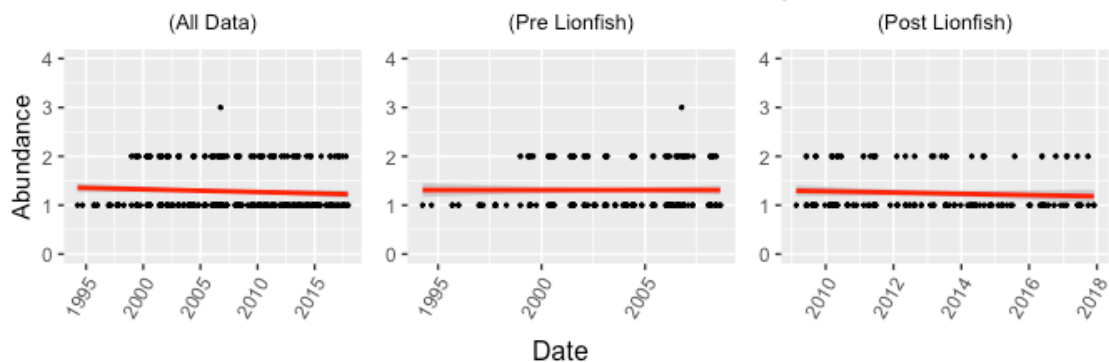
Relative Abundance of the Chub Family



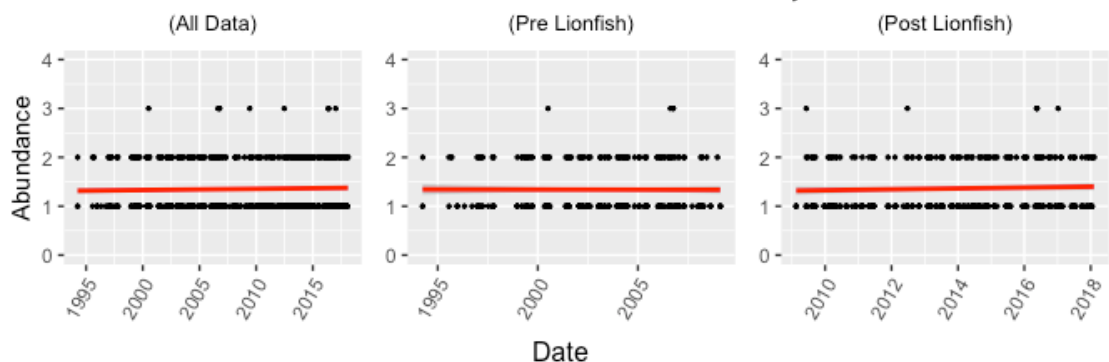
Relative Abundance of the Damselfish Family



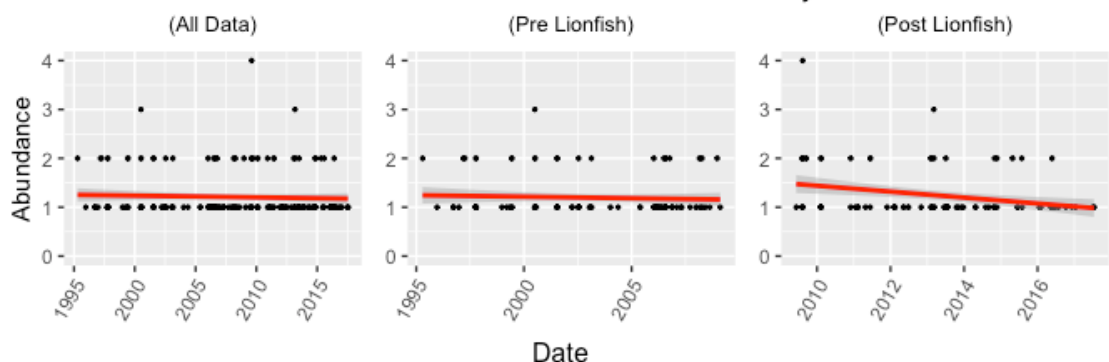
Relative Abundance of the Drum Family



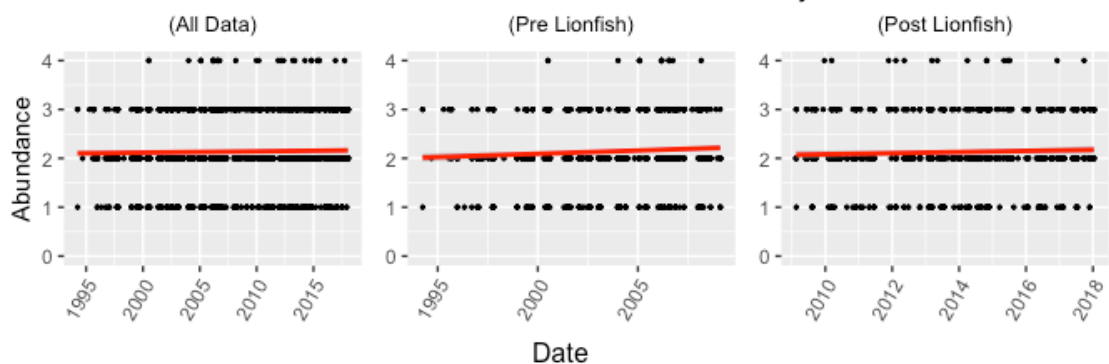
Relative Abundance of the Filefish Family



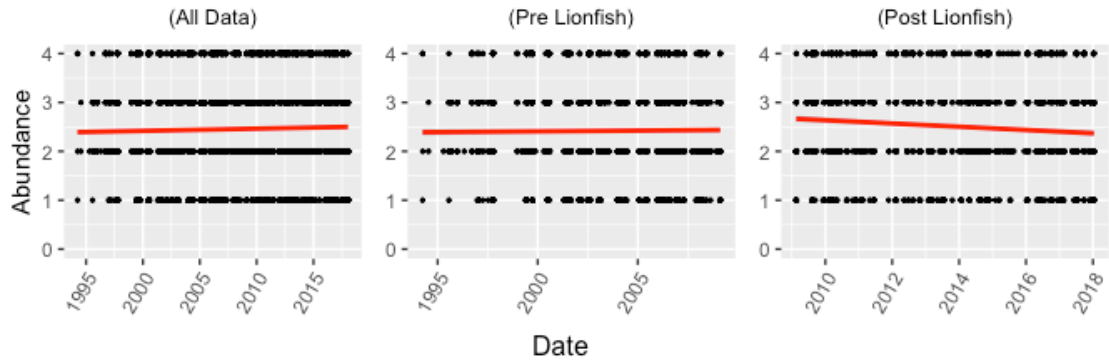
Relative Abundance of the Flounder Family



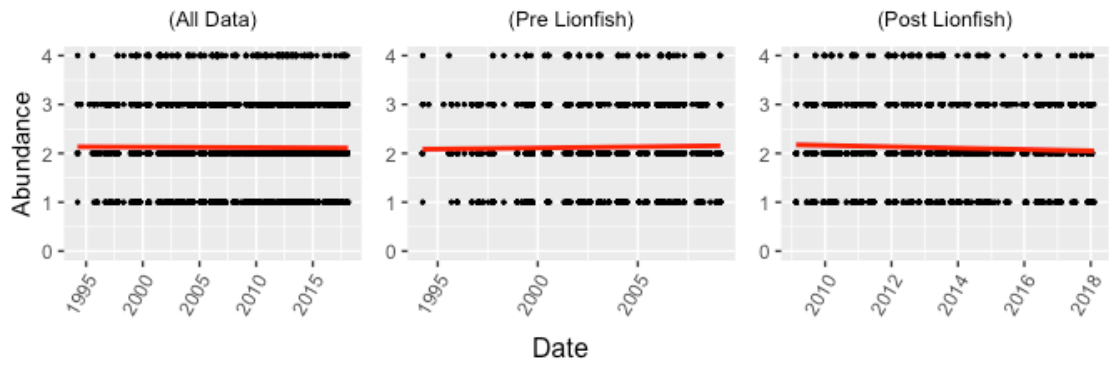
Relative Abundance of the Goatfish Family



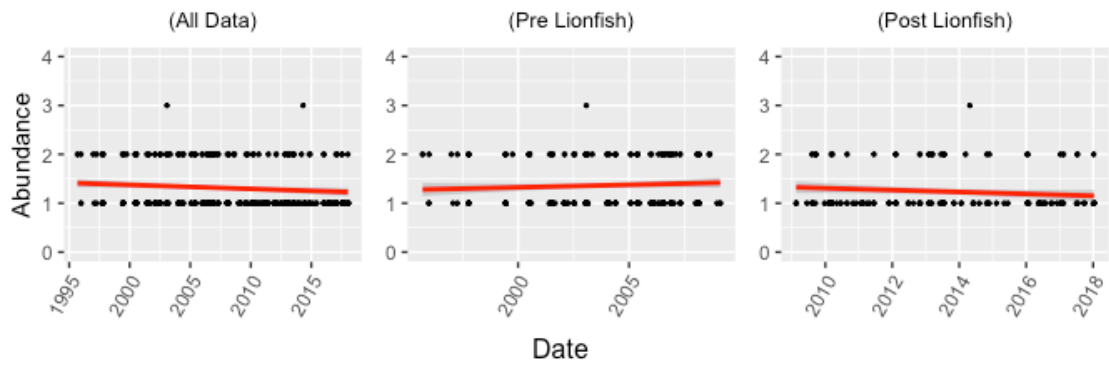
Relative Abundance of the Goby Family



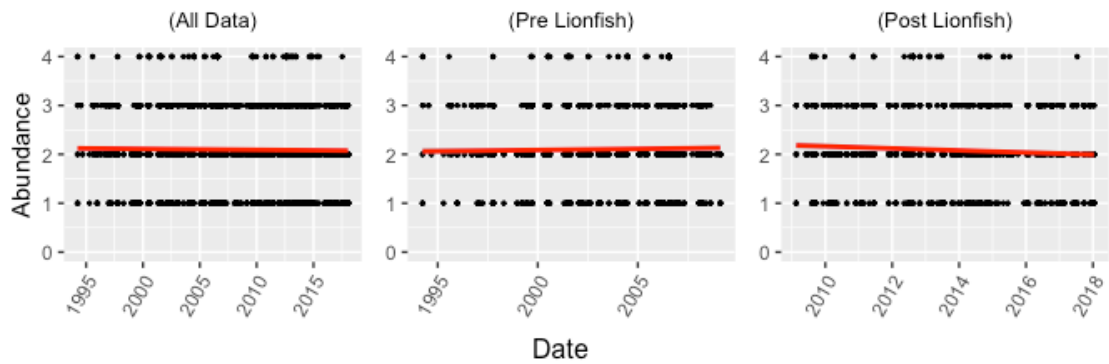
Relative Abundance of the Grunt Family



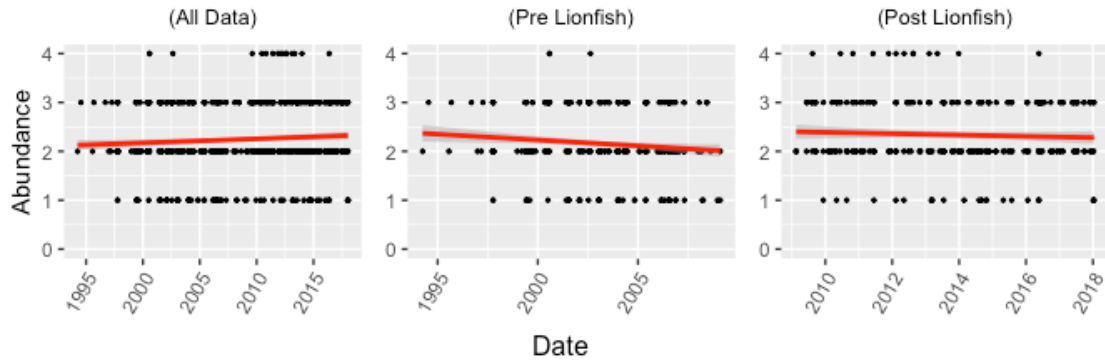
Relative Abundance of the Hawkfish Family



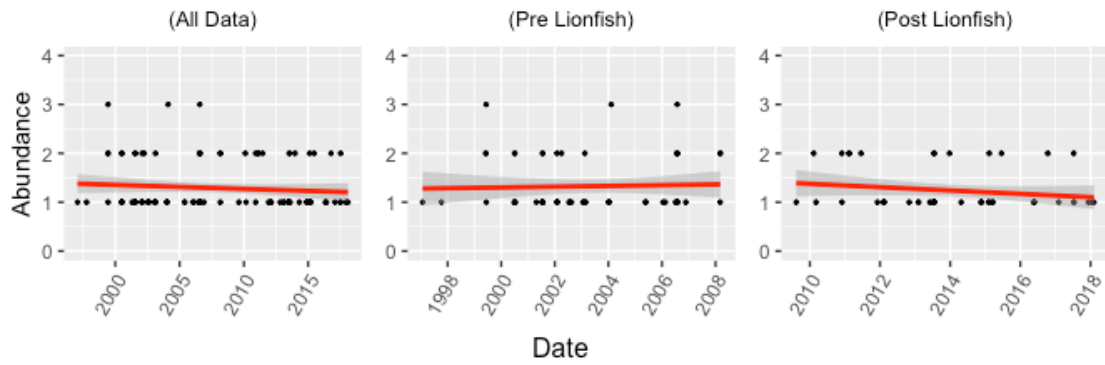
Relative Abundance of the Jack Family



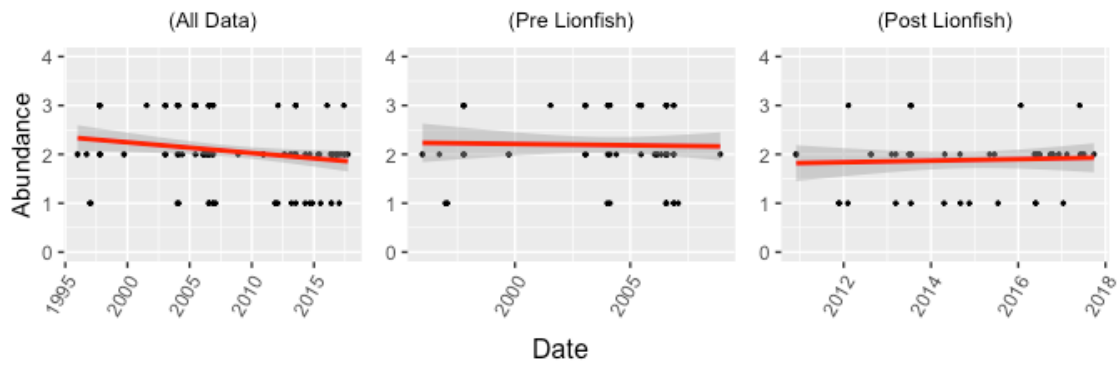
Relative Abundance of the Jawfish Family



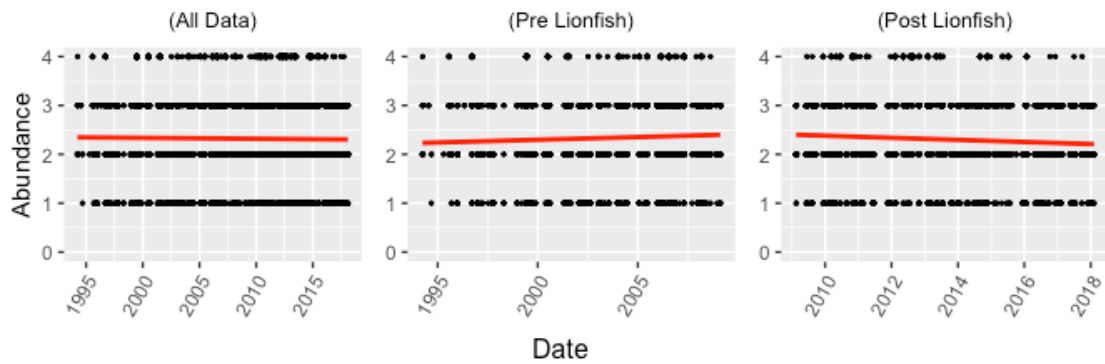
Relative Abundance of the Lizardfish Family



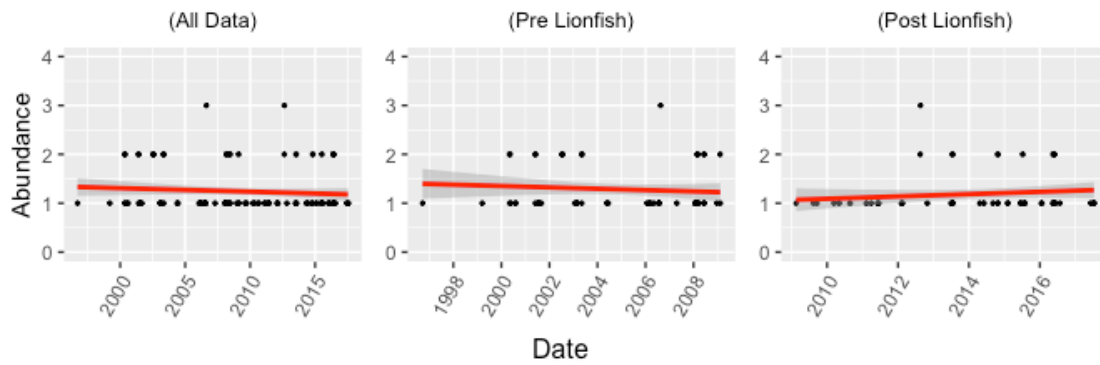
Relative Abundance of the Mojarra Family



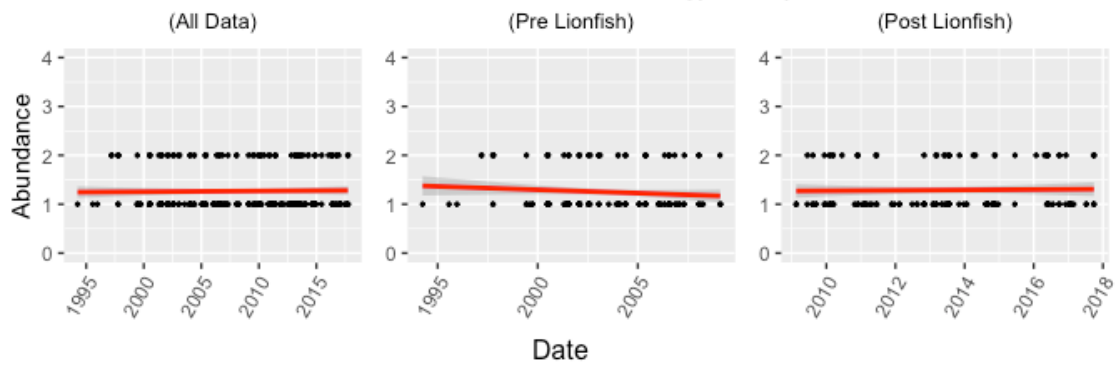
Relative Abundance of the Parrotfish Family



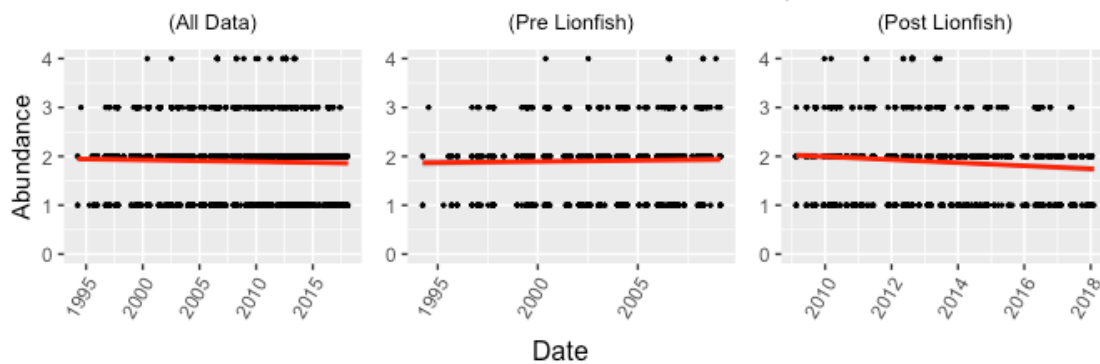
Relative Abundance of the Pipefish and Seahorse Family



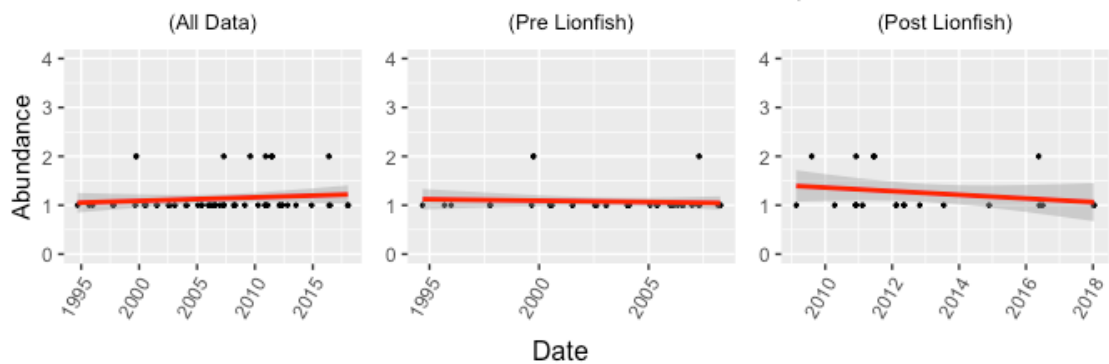
Relative Abundance of the Porgy Family



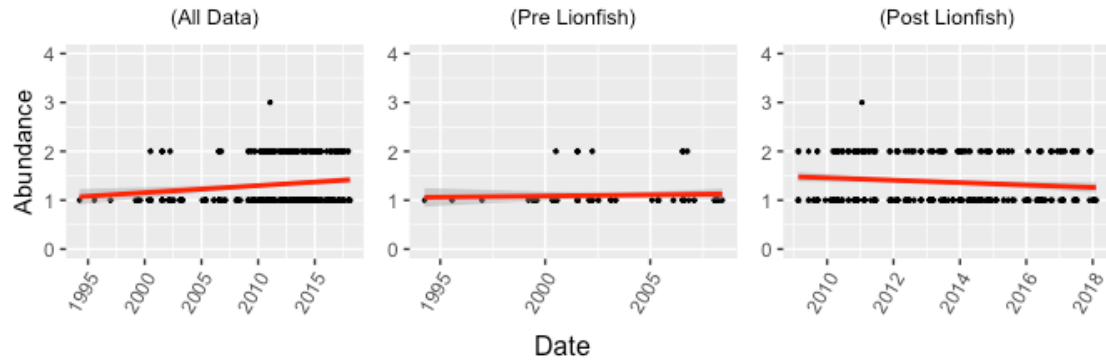
Relative Abundance of the Puffer Family



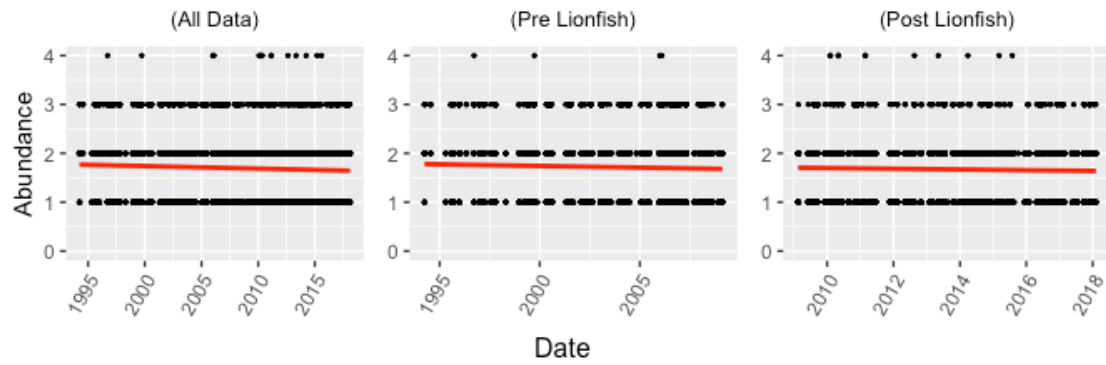
Relative Abundance of the Remora Family



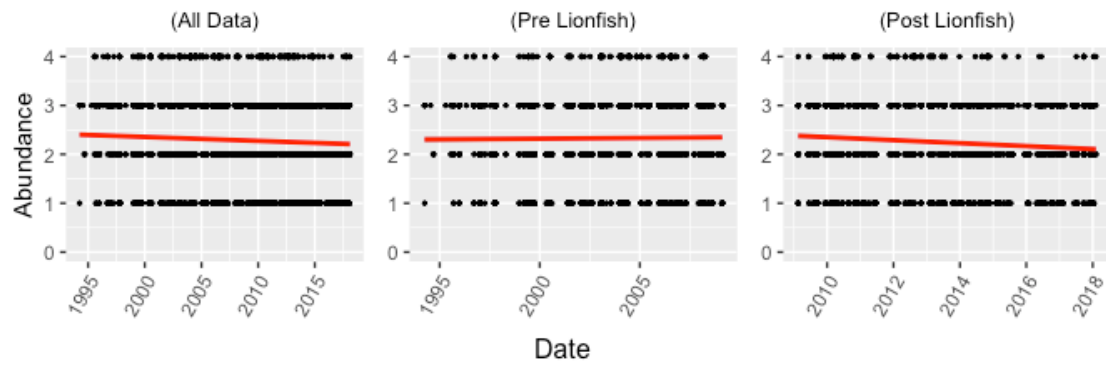
Relative Abundance of the Scorpionfish Family



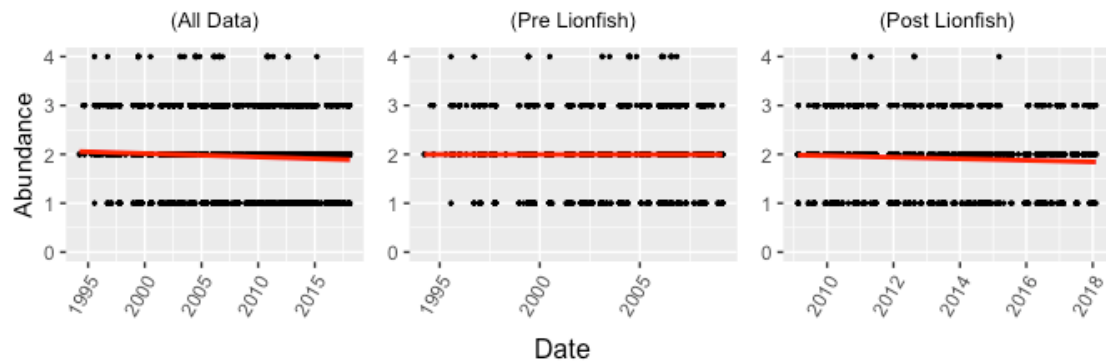
Relative Abundance of the Seabass Family



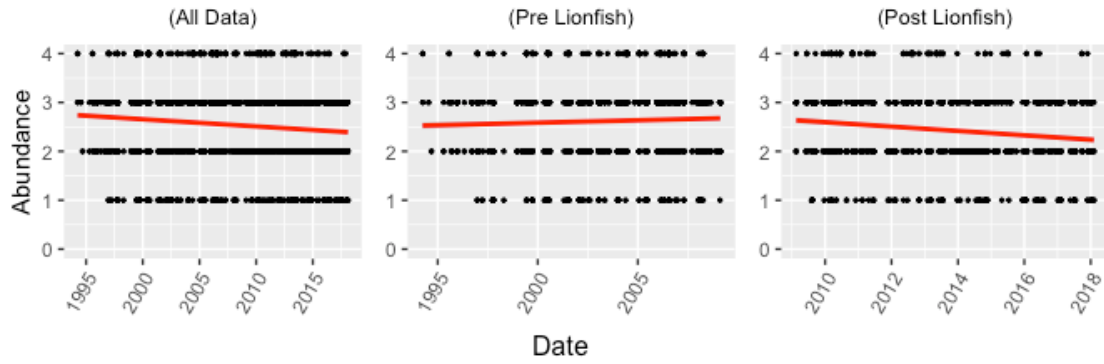
Relative Abundance of the Snapper Family



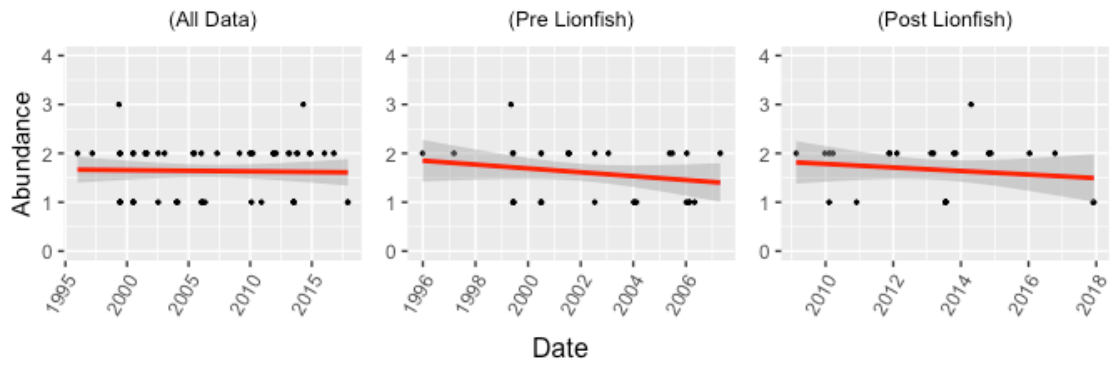
Relative Abundance of the Squirrelfish Family



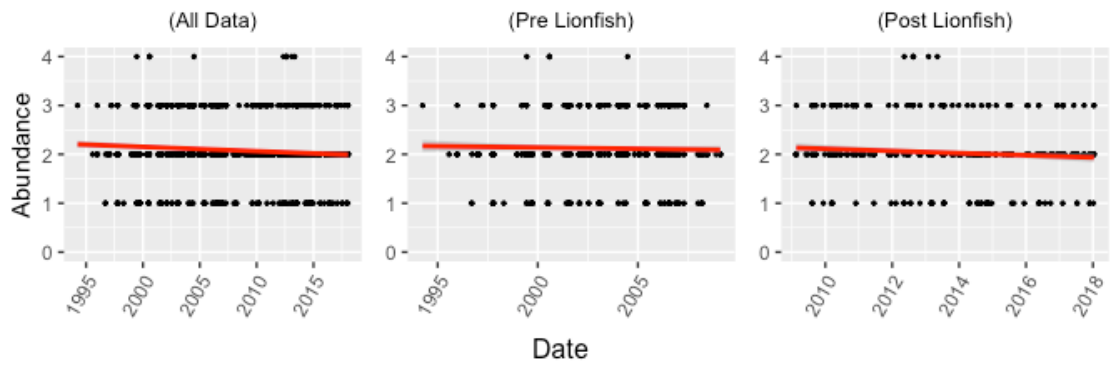
Relative Abundance of the Surgeonfish Family



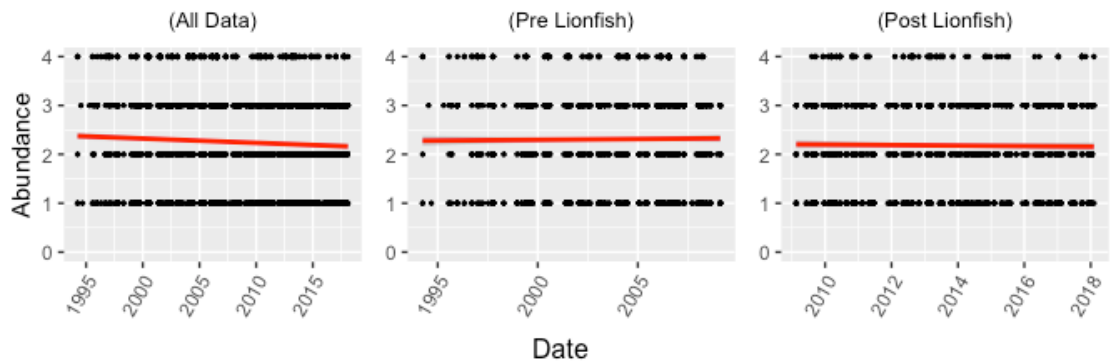
Relative Abundance of the Tarpon Family

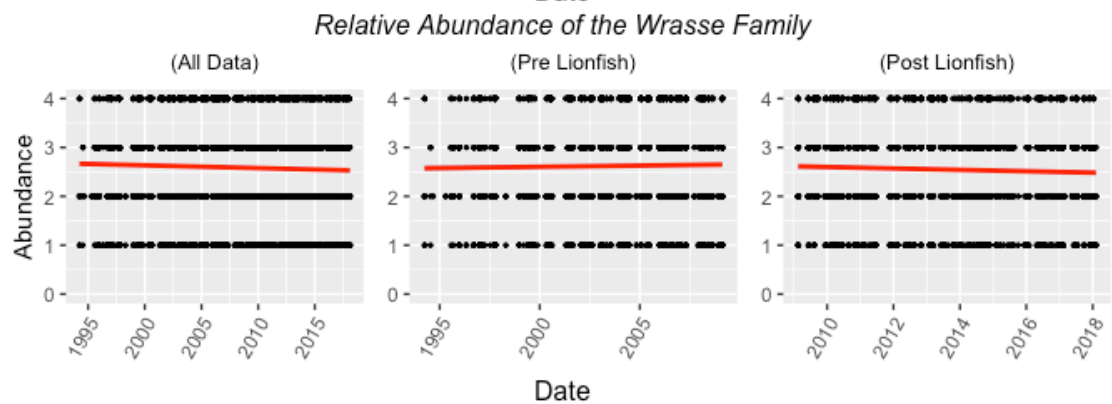
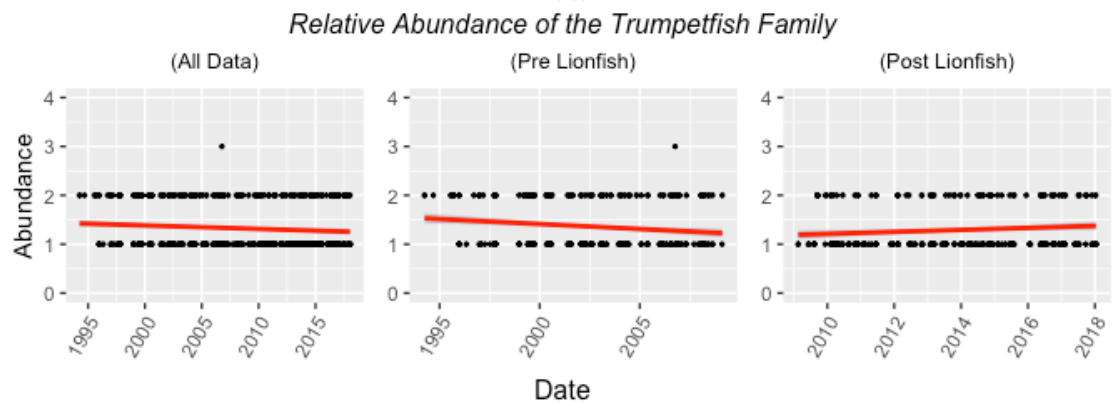
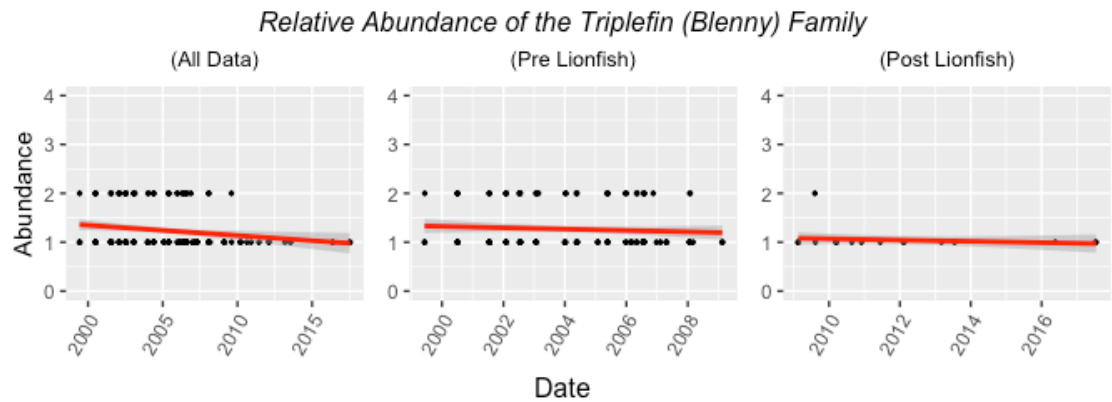


Relative Abundance of the Tilefish Family



Relative Abundance of the Triggerfish Family





Summary of Initial Findings

