

EEB 177 Final Project - Selachii Fossil Diversity Through the Mass Extinctions

Kelsey Jiang

Introduction

Selachii (Selachimorpha) is a taxonomic superorder that includes many modern-day species of sharks and rays and falls under the class Chondrichthyes. Members of this clade include species such as the “infamous” Great white shark, the world’s largest Whale shark, the striped Zebra shark, and even extinct species like the Megalodon. Sharks and their relatives have been in the ocean since 450 million years ago, showcasing their longevity as a species with the oldest shark fossil dating back to 455 million years ago during the Ordovician Period. Sharks are also one of the few species to have survived through the past five mass extinctions, having a lengthy life history that trumps even those of the dinosaurs. This is a rare feat and indicates the survivability of these species; however, with the recent rise in modern industrialization, sharks are beginning to suffer the effects of globalization. Shark fin soup is a common delicacy in many Southeast Asian countries, and the desire for shark fins has resulted in over 100 million shark killings each year, 73 million of those due entirely to shark finning (Cornelia 2006). This increase in overfishing has researchers labeling this modern era as the “6th Extinction,” and this project aims to analyze the extinction and speciation rates of the taxon Selachii during the mass extinctions and compare them to the modern era. This project also plans to dissect the diversity of Selachii through time by cross-referencing fossil occurrence data with the International Union of Conservation of Nature’s catalogue of endangered Chondrichthyes species and trying to find some correlation between areas of intense overfishing and increased endangered listings.

Methods

The fossil occurrence data utilized for this project was downloaded from the online Paleobiology Database as a CSV file. The resulting file was cleaned with shell commands, reformatted in Python, and visualized graphically in R Studio. The datasets that were downloaded from the PBDB included a lot of unnecessary info for this project so I had to cut out the headers and only include the columns of information that I needed for this project. In python, functions were run to include lists of information that would help me visualize the data sets I had collected. The graphs generated in R Studio organized fossil occurrences with the extant species at the top and plotted species diversity along the same axis. R Studio was also used to visually mark where the fossils were located on a world map and highlighted specific species that were classified as Threatened, Endangered, or Critically Endangered.

Finally, PyRate was run on the fossil occurrence data to generate significantly more accurate speciation, extinction, and net diversification rates.

Python Functions - Fossil Occurrences

The python function that reformatted my PBDB data set into one that could be used in ggplot to plot a fossil occurrences map.

```
infile = "/home/eeb177-student/Desktop/eeb-177/EEB-177-Final-Project/Selachii-Data/selachii_data.txt"
with open(infile, 'r') as ff:
    selachii_recs = ff.readlines()

from collections import defaultdict

species_ranges = defaultdict(list)

for line in selachii_recs:
    species = line.split(',')[0]
    minage = line.split(',')[3]
    maxage = line.split(',')[2]
    mean_age = (float(minage) + float(maxage))/2
    #print(species, mean_age)

    species_ranges[species].append(mean_age)

output = open("/home/eeb177-student/Desktop/eeb-177/EEB-177-Final-Project/Selachii-Data/output.txt", 'w')

for key in species_ranges.keys():
    ages = species_ranges[key]
    minage = min(ages)
    maxage = max(ages)
    genus = key.split(" ")[0].strip(' ')
    species = key.strip(' ')
    outline = "{},{},{},{}\n".format(genus, species, minage, maxage)
    output.write(outline)
    #print(genus, species, minage, maxage)
    #print(outline)
```

Output:

Cooperella, Cooperella typicalis, 305.35, 305.35

Sphyrna, Sphyrna americana, 14.1815, 14.1815

Odontaspis, Odontaspis vorax, 8.470500000000001, 14.895

Raja,Raja terhagenensis,25.565,25.565

Pristis,Pristis dubius,158.45,158.45

Moreyella,Moreyella aviculiformis,305.35,305.35

Hammondella,Hammondella globosa,301.29999999999995,301.29999999999995

Eotorpedo,Eotorpedo jaekeli,57.25,57.25

Carcharodon,Carcharodon rondeleti,8.470500000000001,8.470500000000001

Python Functions - Endangered Species

Using the downloaded data from IUCN, I cleaned up the file to include the genus, species, and two-letter code that represented its threat level. Once the information was formatted correctly, I created a diction to store all the species with their respective code. This data was eventually cross-referenced with the fossil occurrences data file to produce a list of 15 unique species ranging from 11 vulnerable species, 3 endangered species, and 1 critically endangered species. From here, I matched up the threatened species their coorelating coordinates in Python so that I could have a dataset ready for R.

```
infile = "/home/eeb177-student/Desktop/eeb-177/EEB-177-Final-Project/Selachii-Data/endangered-redlist.csv"
with open(infile, 'r') as ff:
    redlist_recs = ff.readlines()

def endangered_species(filename):
    EndangeredDict = {}
    fobj = open(filename,"r") # open the file
    all_records = fobj.readlines() [1:]
    for line in all_records:
        genus = line.split(',')[0]
        species = line.split(',')[1]
        code = line.split(',')[2].strip("\n")
        EndangeredDict[genus + " " + species] = code

    return EndangeredDict

endangered_species("endangered-redlist.csv")
```

Output: {'Aetobatus ocellatus': 'VU', 'Aetomylaeus maculatus': 'EN', 'Aetomylaeus nichofii': 'VU', 'Aetomylaeus vespertilio': 'EN', 'Alopias pelagicus': 'VU', 'Alopias superciliosus': 'VU', 'Alopias vulpinus': 'VU', 'Amblyraja radiata': 'VU', 'Anoxypristis cuspidata': 'EN', 'Aptychotrema timorensis': 'VU', 'Atelomycterus baliensis': 'VU', 'Atlantoraja castelnaui': 'EN', 'Atlantoraja cyclophora': 'VU', 'Atlantoraja platana': 'VU', 'Aulohalaelurus kanakorum': 'VU'}

```

infile = "/home/eeb177-student/Desktop/eeb-177/EEB-177-Final-Project/Selachii-Data/redli
with open(infile, 'r') as ff:
    redlist_coordinates_recs = ff.readlines()

def redlist_coordinates(filename):
    CoordinatesDict = {}
    fobj = open(filename, "r") # open the file
    all_records = fobj.readlines() [1:]
    for line in all_records:
        species = line.split(',')[0]
        taxon = line.split(',')[1]
        longitude = line.split(',')[2]
        latitude = line.split(',')[3].strip("\n")
        CoordinatesDict[species] = longitude + latitude

    return CoordinatesDict

redlist_coordinates("redlist-selachii-coordinates.csv")

```

The resulting dictionary yielded these results for the redlist-selachii-coordinates.csv file:

```

{ "Carcharhinus obscurus": "-80.044724" "9.281389", "Carcharhinus plumbeus":
"-80.044724" "9.281389", "Carcharias taurus": "-9.202222" "39.067780", "Carcharodon
carcharias": "-81.650002" "30.330000", "Cetorhinus maximus": "5.800000" "51.466667",
"Galeorhinus galeus": "-80.044724" "9.281389", "Isurus oxyrinchus": "-
122.036003" "37.053001", "Odontaspis ferox": "-76.800003" "35.400002", "Pristis
pristis": "-82.565277" "9.572500", "Pseudoginglymostoma brevicaudatum": "-
56.603333" "64.239441", "Rhincodon typus": "-82.565277" "9.572500", "Sphyrna
lewini": "-80.044724" "9.281389", "Sphyrna zygaena": "-9.202222" "39.067780",
"Squalus acanthias": "5.800000" "51.466667", "Stegostoma fasciatum": "-56.603333" "-
64.239441" }

```

Graphical Represenatation

1. Fossil Occurrences Data

The data was cleaned with shell commands to a CSV file with species, minimum age, and maximum age. It was then reorganized in Python to include genus, species, minimum age, and maximum age. The resulting data set was run in R Studio to produce the following graphs. Because the data set was so large, I split the data into two equal sets after alphabetically organizing the list. This was done to allow for better spacing and readability on the graphs.

Part one includes the first 173 species from the superorder Selachii and part two includes the later 173 species for a total of 346 species.

```
library(ggplot2)

setwd("~/Desktop/eeb-177/EEB-177-Final-Project/Selachii-Data")
sharks <- read.csv("~/Desktop/eeb-177/EEB-177-Final-Project/Selachii-Data/selachii_range")
names(sharks) <- c("genus", "species", "minage", "maxage")
head(sharks)
```

```
##           genus              species minage maxage
## 1   Carcharhinus Carcharhinus balochenisis  31.00  31.00
## 2   Synodontaspis Synodontaspis atlasi    44.55  51.90
## 3   Gladioserratus Gladioserratus magnus 101.40 101.40
## 4         Trygon      Trygon jaekeli   35.55  57.25
## 5 Pseudoginglymostoma Pseudoginglymostoma brevicaudatum 42.90 42.90
## 6         Mesiteia      Mesiteia emiliae   97.20  97.20
```

Fossil Occurrences Chart with Diversity Plot for Selachii Part 1

```
library(ggplot2)
library(tidyr)
library(dplyr)

##
## 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

library(forcats)

setwd("~/Desktop/eeb-177/EEB-177-Final-Project/Selachii-Data")
sharks1 <- read.csv("~/Desktop/eeb-177/EEB-177-Final-Project/Selachii-Data/selachii_spec")
names(sharks1) <- c("genus", "species", "minage", "maxage")
head(sharks1)
```

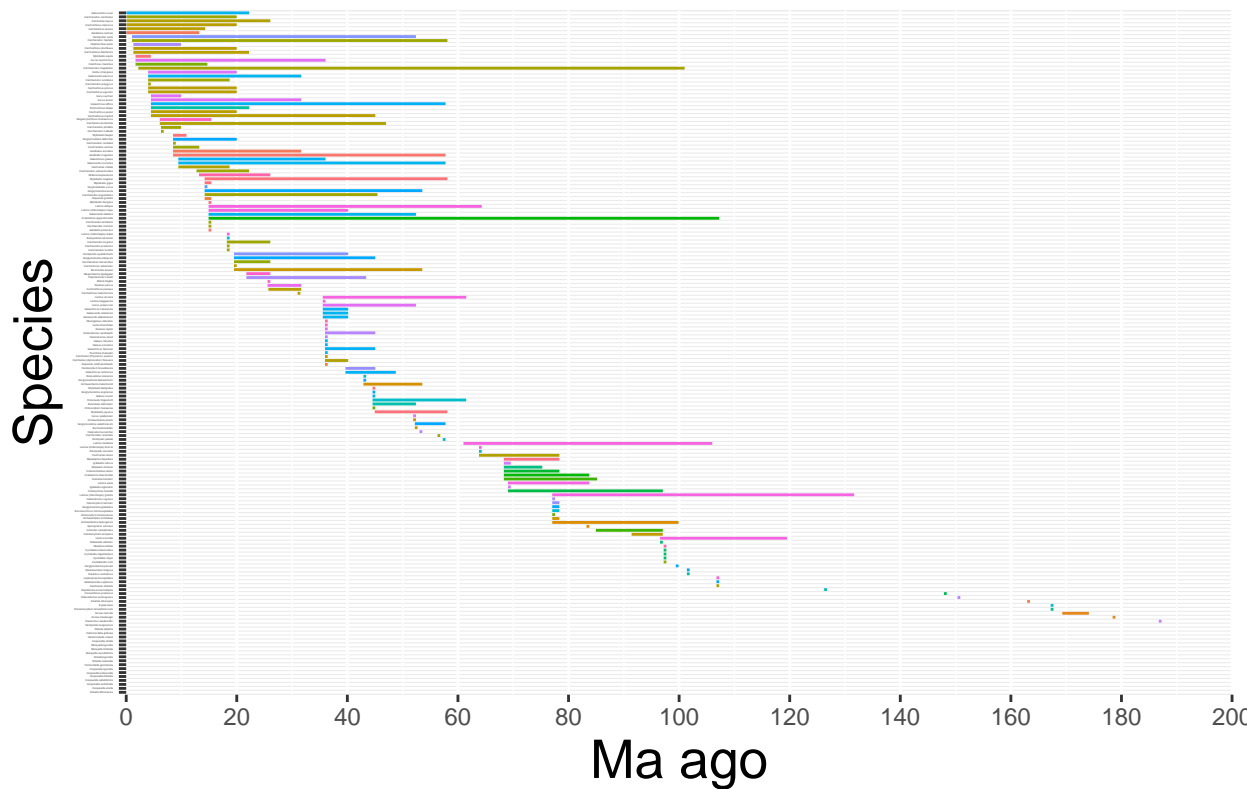
```
##           genus              species    minage    maxage
## 1 Aetobatis Aetobatis irregularis    8.47050   57.250
## 2 Aetobatis Aetobatis profundus   14.89500   14.895
## 3 Aetobatus Aetobatus arcuatus    8.47050   31.150
```

```
## 4 Aetobatus      Aetobatus narinari    0.06885  12.720
## 5 Akaimia        Akaimia altucuspis 162.95000 162.950
## 6 Alopecias      Alopecias grandis  14.18150  14.895
```

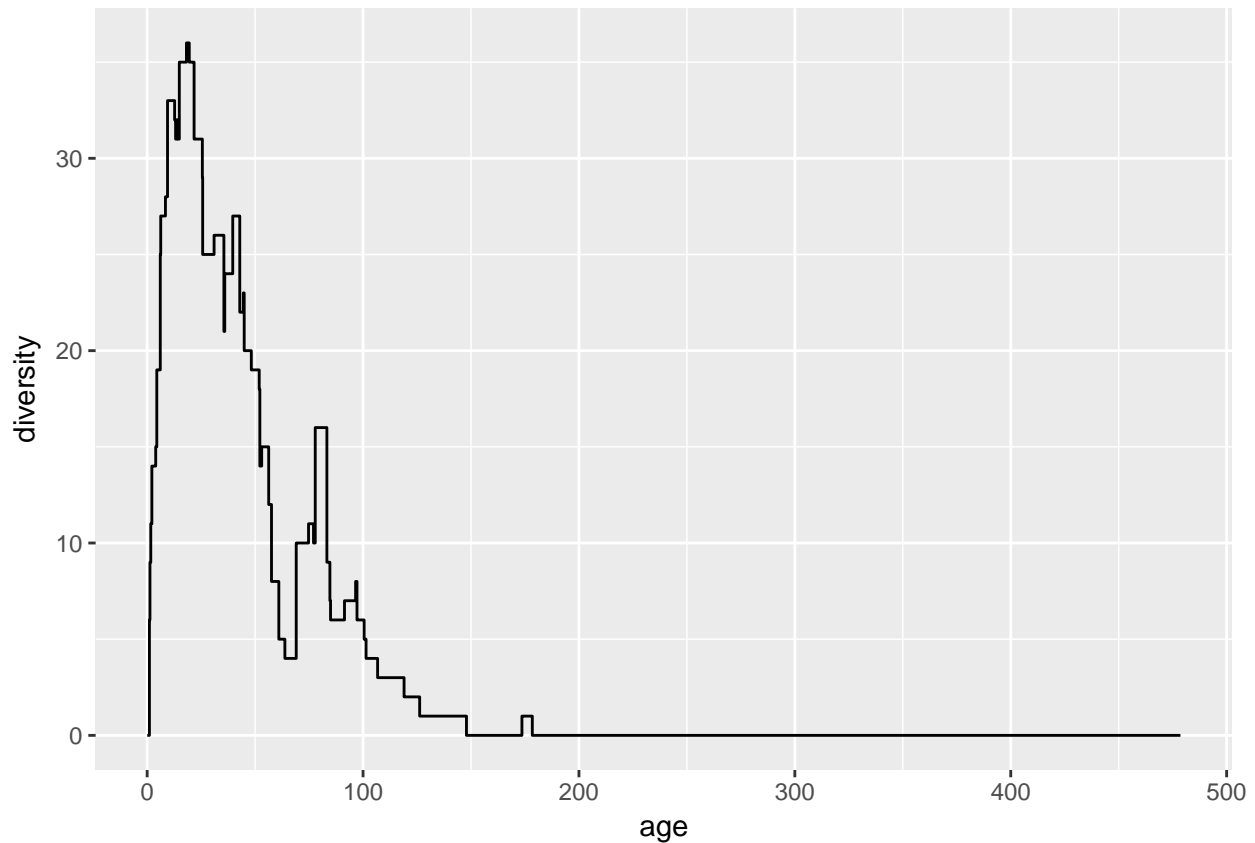
```
sharks1_occ <- ggplot(sharks1, aes( x = fct_reorder(species, minage, .desc = T), maxage))
sharks1_occ + geom_linerange(aes(ymin = minage, ymax = maxage + 0.5)) + theme(legend.position = "none")
```

```
## Warning: Removed 18 rows containing missing values (geom_linerange).
```

Selachii Part 1 Fossil Occurrences



```
diversity <- sharks1 %>% gather(key = type, value = age, minage, maxage) %>% mutate(count = 1)
ggplot(diversity, aes(x = age, y = diversity)) + geom_step()
```



```
#ggsave(filename = "selachii-occ-part-1.pdf")
```

Fossil Occurrences Chart with Diversity Plot for Selachii Part 2

```
library(ggplot2)
library(forcats)

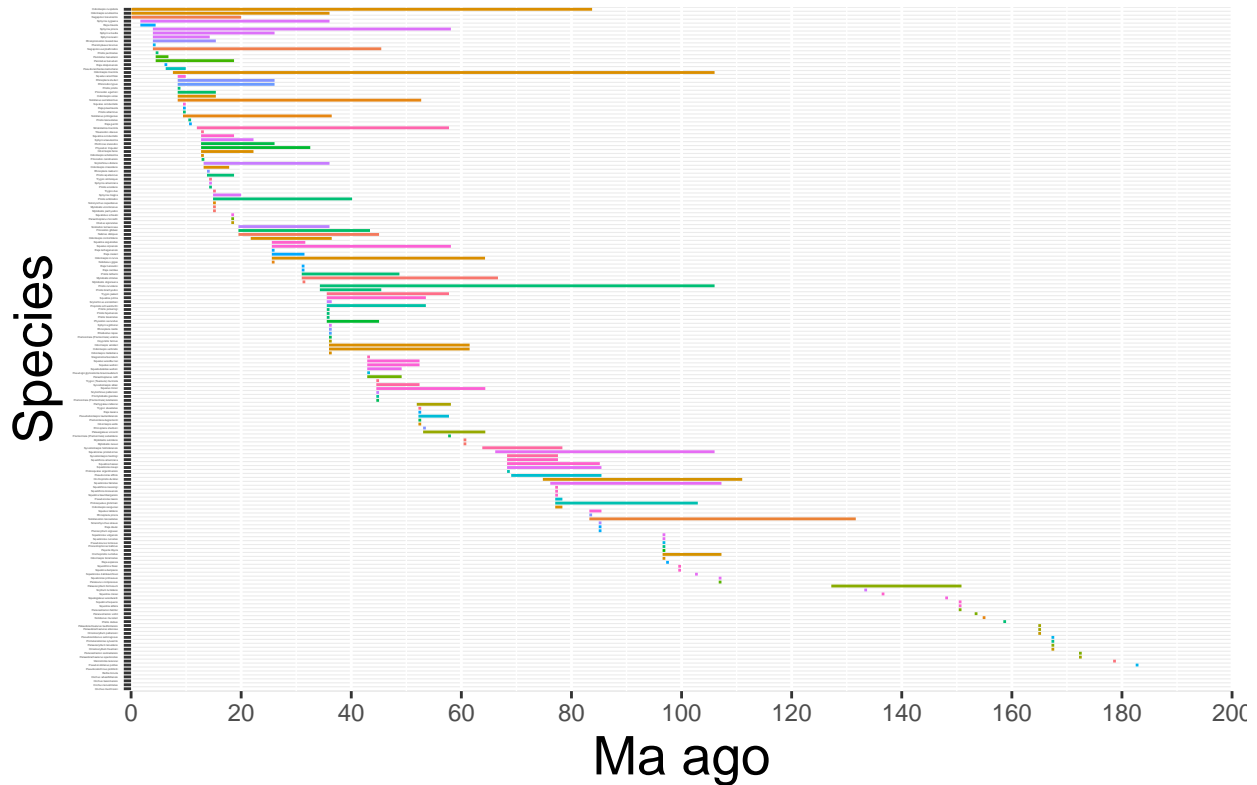
setwd("~/Desktop/eeb-177/EEB-177-Final-Project/Selachii-Data")
sharks2 <- read.csv("~/Desktop/eeb-177/EEB-177-Final-Project/Selachii-Data/selachii_spec")
names(sharks2) <- c("genus", "species", "minage", "maxage")
head(sharks2)
```

```
##      genus      species minage maxage
## 1 Myliobatis Myliobatis oligocaena 31.150 31.150
## 2 Myliobatis Myliobatis pachyodon 14.895 14.895
## 3 Myliobatis Myliobatis rauxi 60.400 60.400
## 4 Myliobatis Myliobatis striatus 31.000 66.150
## 5 Myliobatis Myliobatis sulcidens 60.400 60.400
## 6 Myliobatis Myliobatis vicomicanus 14.895 14.895
```

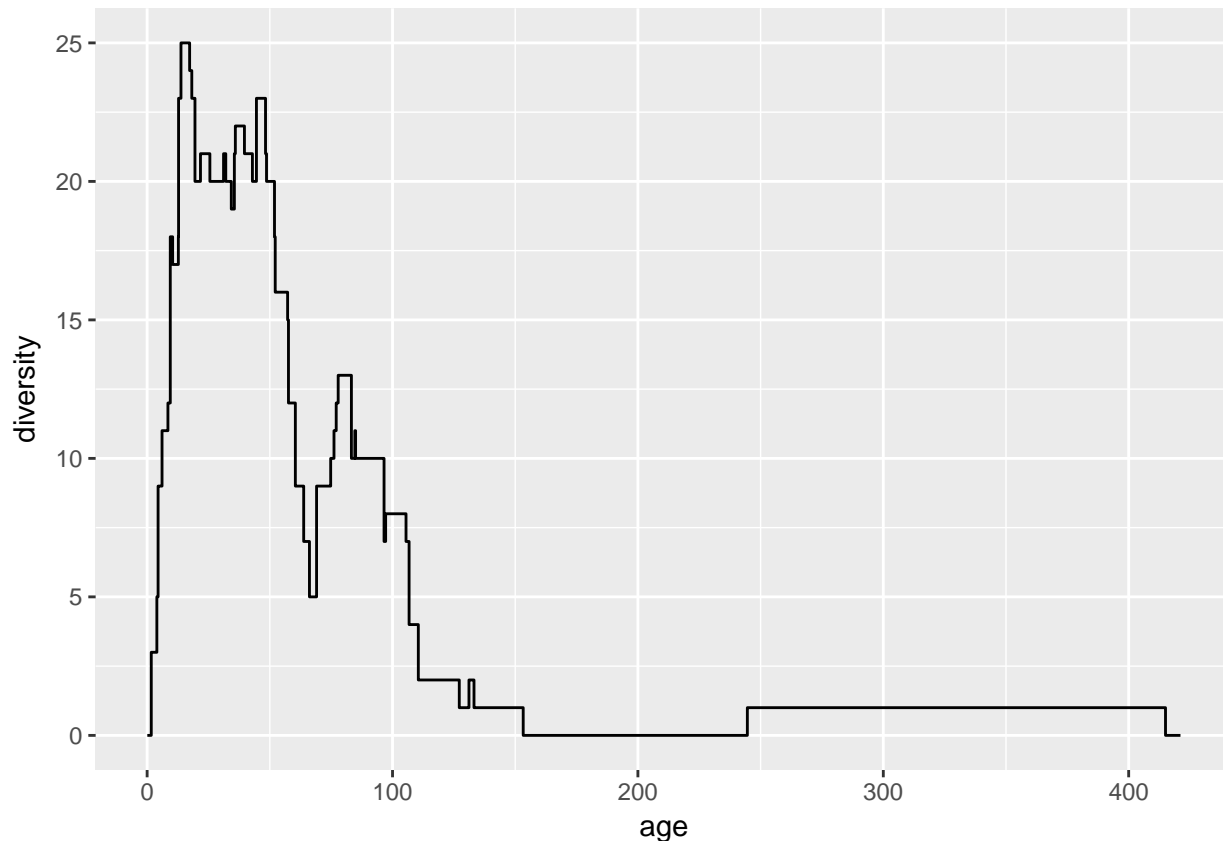
```
sharks2_occ <- ggplot(sharks2, aes( x = fct_reorder(species, minage, .desc = T), maxage,
sharks2_occ + geom_linerange(aes(ymin = minage, ymax = maxage + 0.5))) + theme(legend.pos
```

```
## Warning: Removed 6 rows containing missing values (geom_linerange).
```

Selachii Part 2 Fossil Occurrences



```
diversity <- sharks2 %>% gather(key = type, value = age, minage, maxage) %>% mutate(count
ggplot(diversity, aes(x = age, y = diversity)) + geom_step()
```

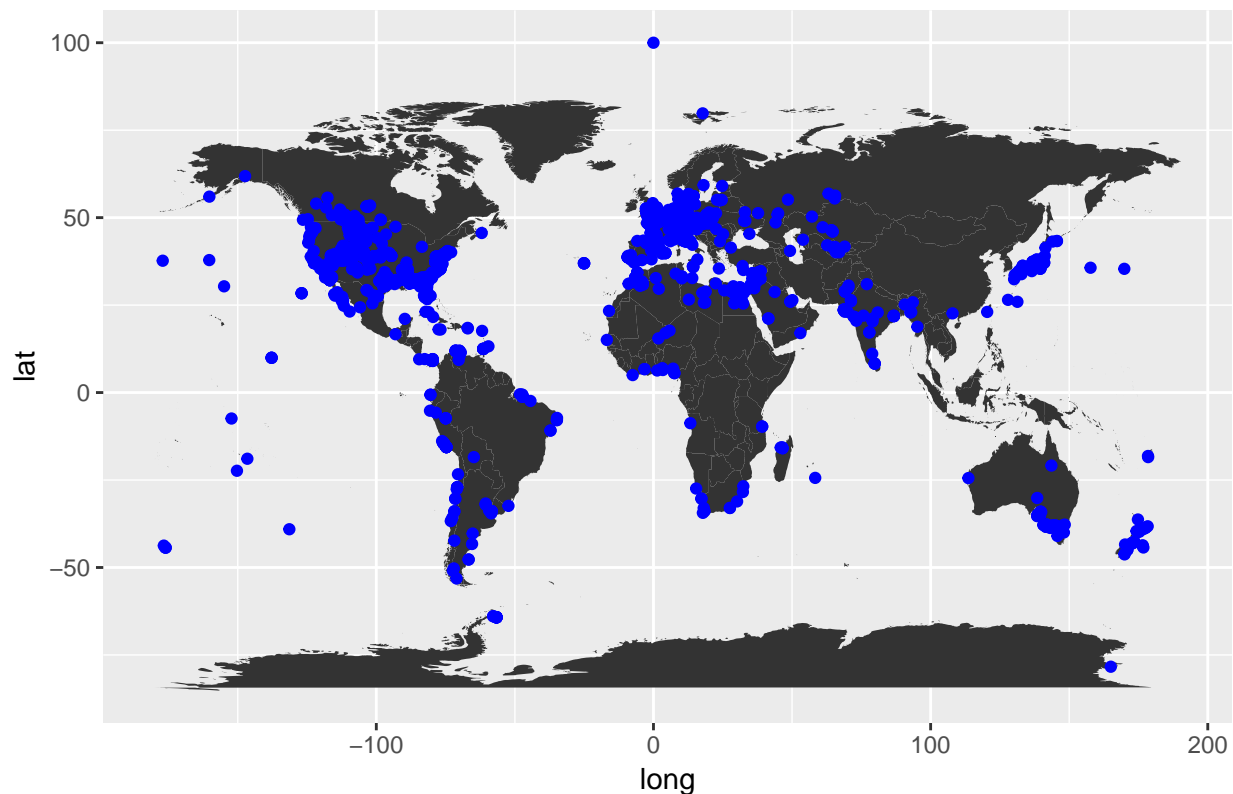
```
#ggsave(filename = "selachii-occ-part-2.pdf")
```

2. World Distribution of the Selachii Fossil Occurrences

This map was included to display where the fossils of many of the species were found. The data was downloaded from Paleobiology Database with coordinates of longitude and latitude so that they could be plotted onto a world map. Glancing at the resulting output, there seems to be a higher concentration throughout North America, Europe, and the coastline of South America.

```
library(ggmap)
library(maps)
library(mapdata)
world <- map_data("world")
selachii_distribution <- read.csv("/home/eeb177-student/Desktop/eeb-177/EEB-177-Final-Pr
names(selachii_distribution) <- c("longitude", "latitude")
selachii_distribution$longitude<-as.numeric(selachii_distribution$longitude)
ggplot() + geom_polygon(data = world, aes(x=long, y=lat, group = group)) + geom_point(d
```

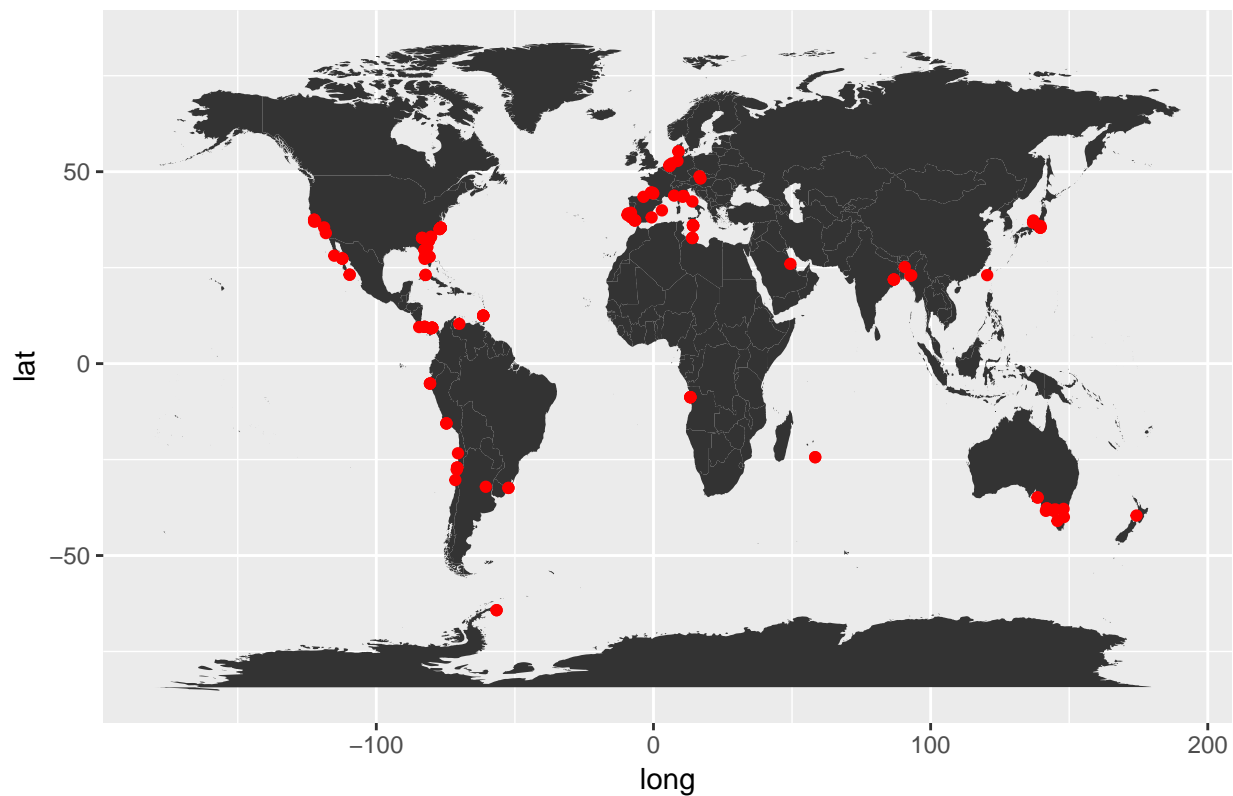
Selachii Fossil Occurrences on World Map



To organize the graphical representation even more, I cross-referenced my IUCN data set with my PBDB data set to see which extant species in my PBDB set were classified as Vulnerable (VU), Endangered (EN), or Critically Endangered (CR). I then made data sets with each individual threat-level category and graphed their coordinates onto a world map. Because the PBDB data set deals with fossil occurrences, there could be discrepancies in terms of abundance or location. The purpose of these world maps was to see if those species listed as VU, EN, and CR were more likely to show up in the South Pacific near countries like Japan, South Korea, China, and the Philippines. If this was the case, there could be a correlation between endangered species and the threat of overfishing since finning for sharks occurs at a higher rate in these waters compared to other countries.

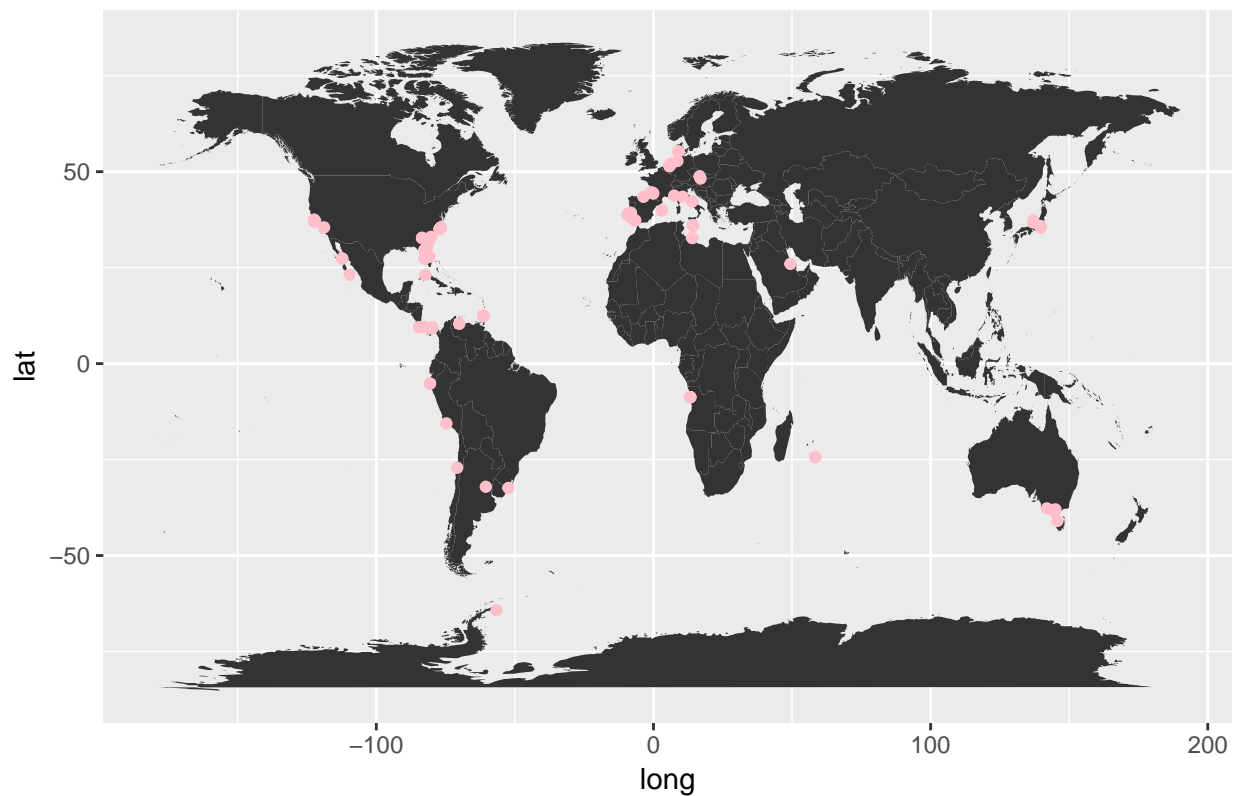
```
library(ggmap)
library(maps)
library(mapdata)
world <- map_data("world")
redlist_distribution <- read.csv("/home/eeb177-student/Desktop/eeb-177/EEB-177-Final-Pro
names(redlist_distribution) <- c("longitude", "latitude")
redlist_distribution$longitude<-as.numeric(redlist_distribution$longitude)
ggplot() + geom_polygon(data = world, aes(x=long, y=lat, group = group)) + geom_point(d
```

Threatened Selachii Species on World Map – Includes VU, EN, & CR



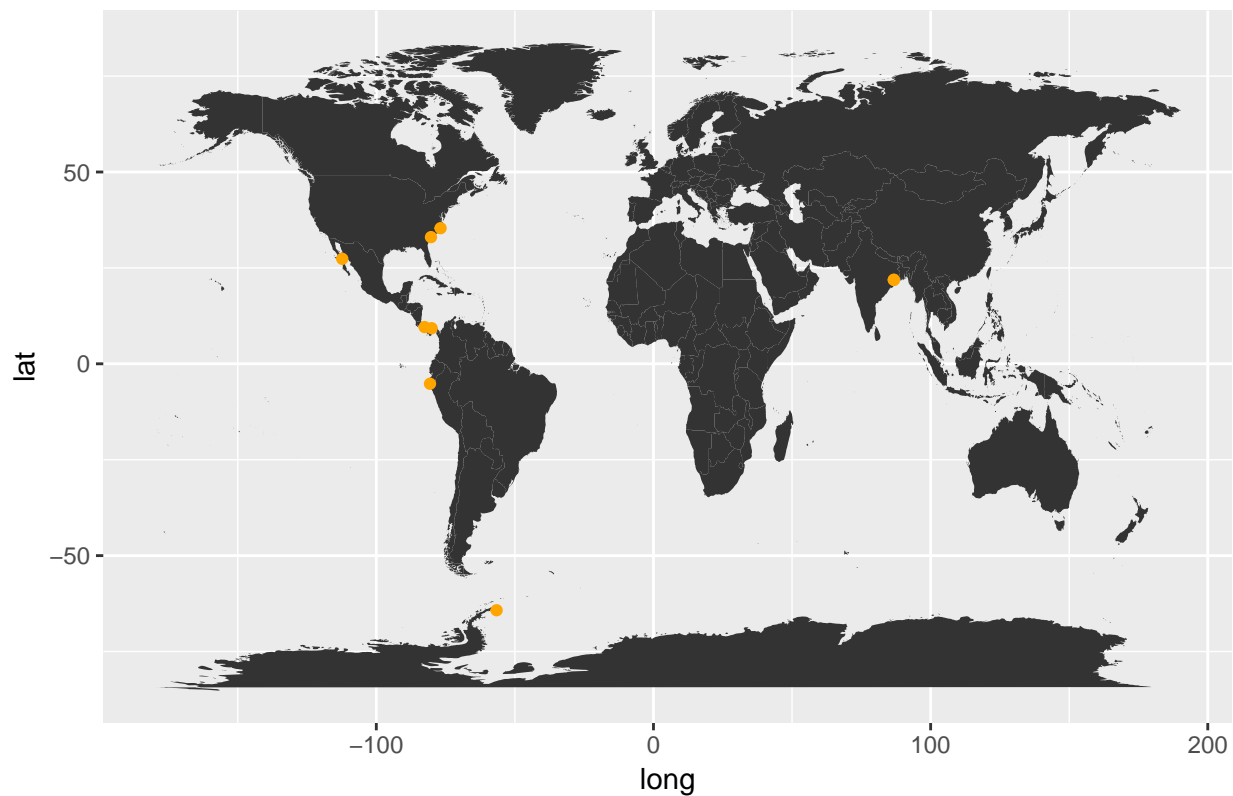
```
library(ggmap)
library(maps)
library(mapdata)
world <- map_data("world")
vulnerable_species <- read.csv("/home/eeb177-student/Desktop/eeb-177/EEB-177-Final-Project/Species/Vulnerable_Selachii.csv")
names(vulnerable_species) <- c("longitude", "latitude")
vulnerable_species$longitude <- as.numeric(vulnerable_species$longitude)
ggplot() + geom_polygon(data = world, aes(x=long, y=lat, group = group)) + geom_point(data = vulnerable_species, aes(x=longitude, y=latitude, color = status))
```

Vulnerable Selachii Species on World Map



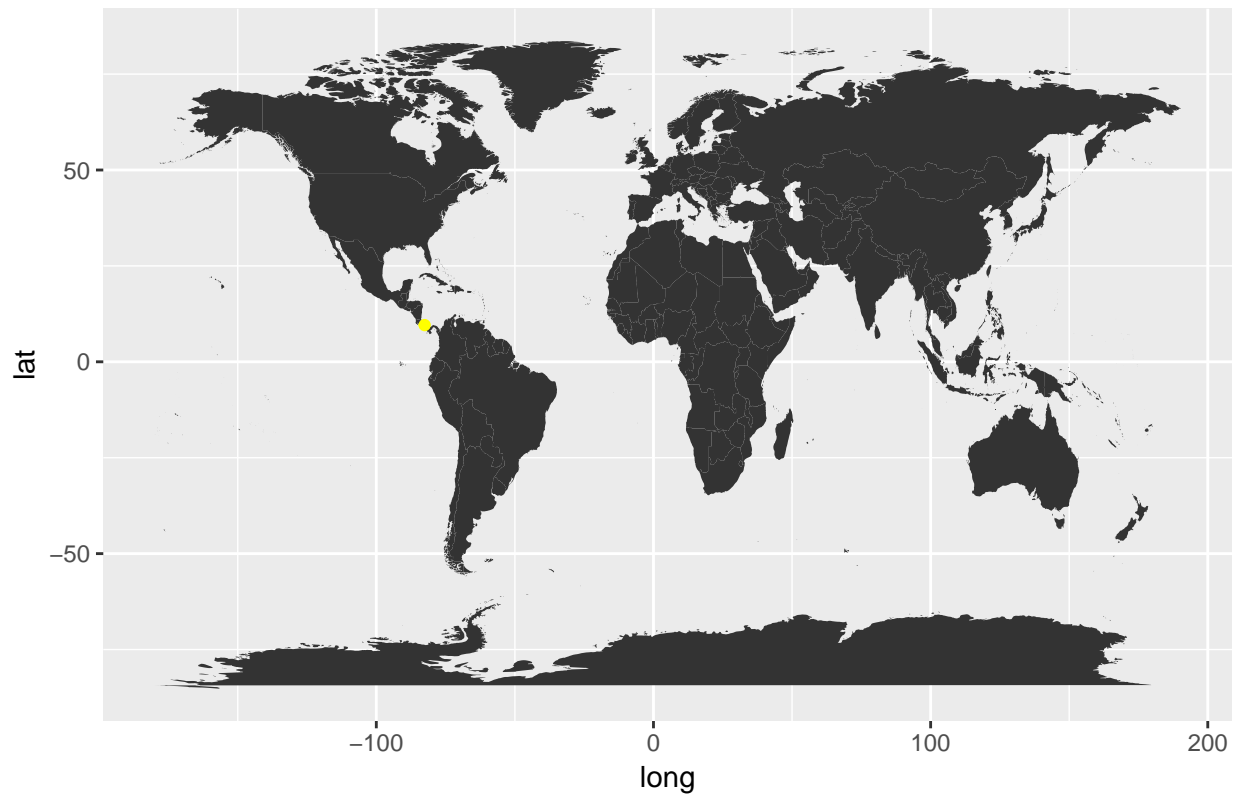
```
library(ggmap)
library(maps)
library(mapdata)
world <- map_data("world")
endangered_species <- read.csv("/home/eeb177-student/Desktop/eeb-177/EEB-177-Final-Project/Endangered_Species.csv")
names(endangered_species) <- c("longitude", "latitude")
endangered_species$longitude <- as.numeric(endangered_species$longitude)
ggplot() + geom_polygon(data = world, aes(x=long, y=lat, group = group)) + geom_point(data = endangered_species, aes(x=longitude, y=latitude, color = "red"))
```

Endangered Selachii Species on World Map



```
library(ggmap)
library(maps)
library(mapdata)
world <- map_data("world")
criticallyendangered_species <- read.csv("/home/eeb177-student/Desktop/eeb-177/EEB-177-F
names(criticallyendangered_species) <- c("longitude", "latitude")
criticallyendangered_species$longitude<-as.numeric(criticallyendangered_species$longitud
ggplot() + geom_polygon(data = world, aes(x=long, y=lat, group = group)) + geom_point(d
```

Critically Endangered Selachii Species on World Map



PyRate Graphs

Along with the other graphs produced in this project, I also ran PyRate through 1,000,000 iterations to estimate speciation and extinction rates through their temporal dynamics found in the fossil occurrence data. PyRate was used to explore different models of diversification by taking into account different fossilization rates and sampling patterns. The main model specifications included running PyRate with a -mG code that allows for the Gamma model of the heterogeneity of the preservation rate across lineages (Silvestro 2014).

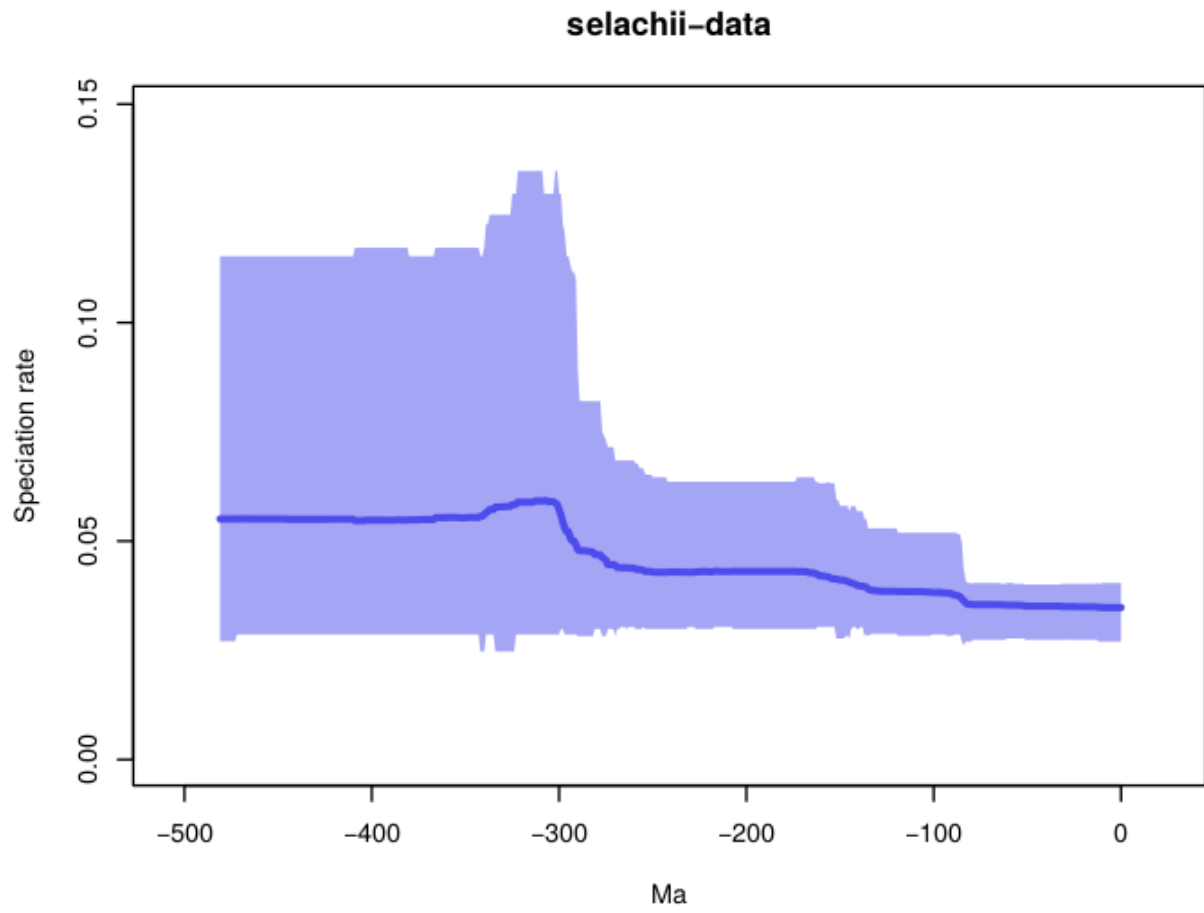


Figure 1: Speciation Rate v. Time

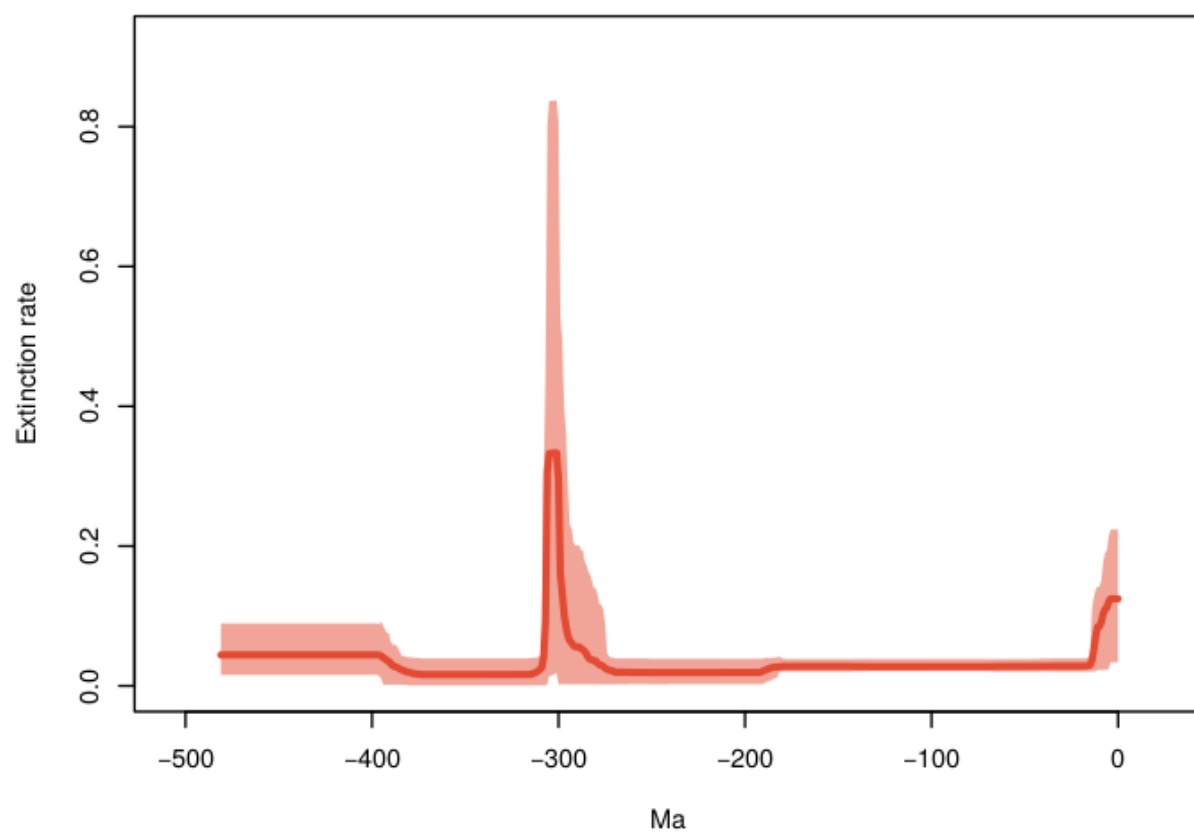


Figure 2: Extinction Rate v. Time

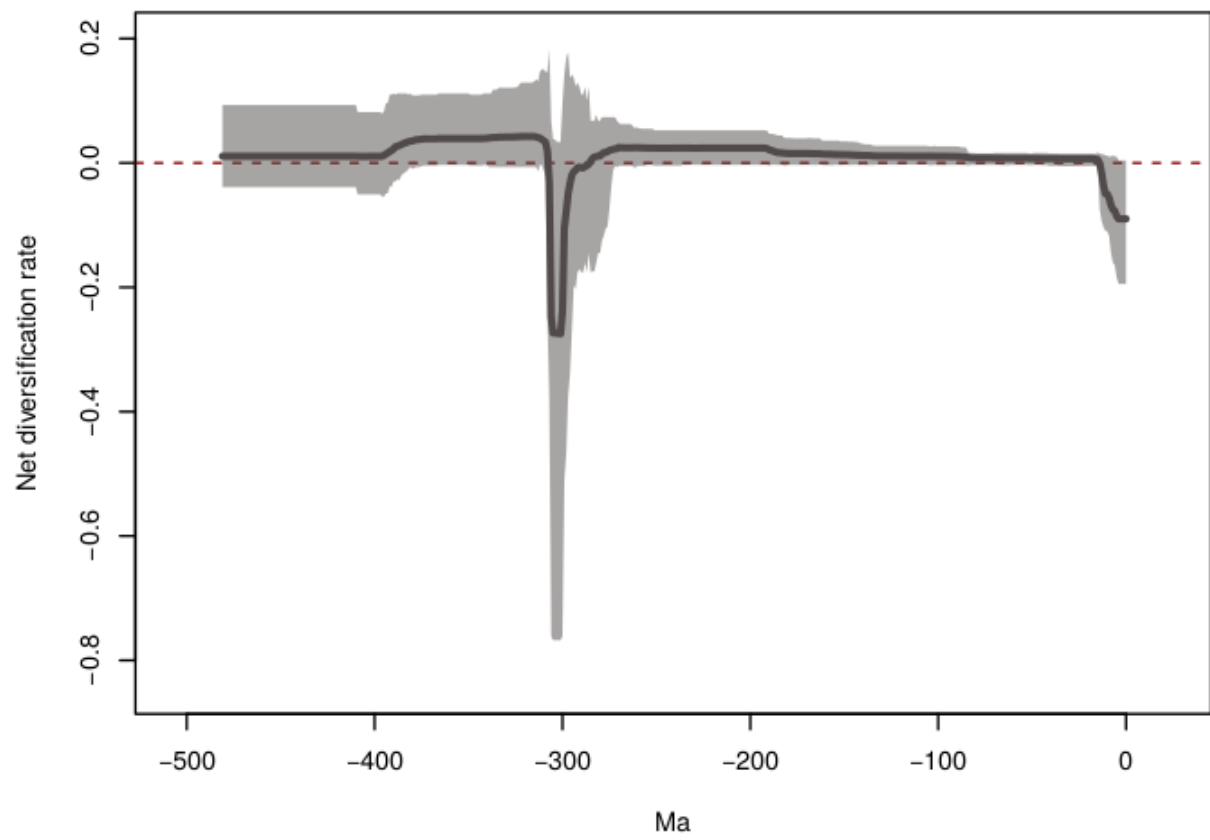


Figure 3: Net Diversification Rate v. Time

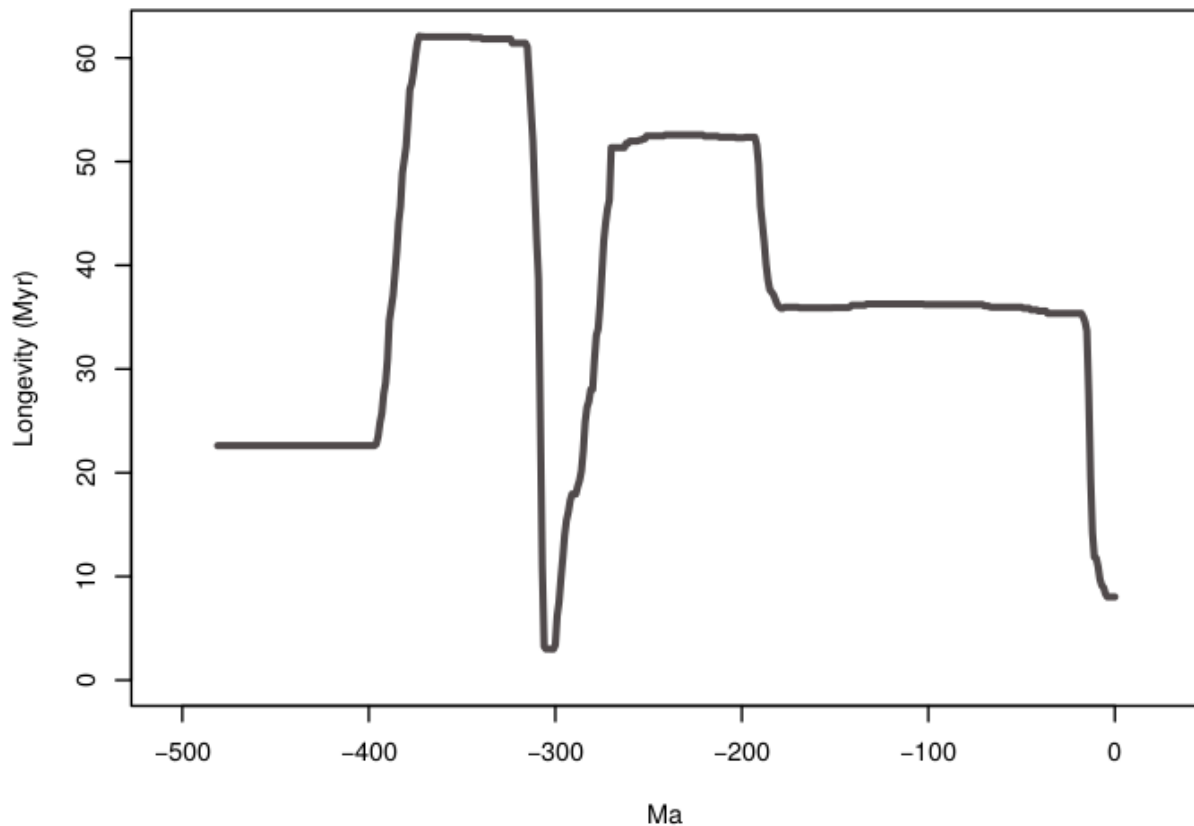


Figure 4: Longevity v. Time

Discussion

After completing all the data analysis, there were a few trends I noticed in each of the sets of data. Starting from the Fossil Occurrence Graphs, there was a clear increase in fossil occurrences nearing the 15 million years mark for both plots. This could be correlated to speciation rate as the collected fossils indicate an increase of species as we move into the present era. More unique fossils that are found could correlate to an increase of species during that time interval. Comparing the fossil occurrences data with the following diversity v. age plots, we also see a similar outcome; however, it does show a significant drop in diversity the closer time moves to the present.

Surprisingly, both the Selachii Part 1 and Part 2 graphs highlight a relative decline around 60 mya which could be a result of the K/T mass extinction (Kriwet and Benton 2004). Their overall diversification rate drops significantly, but this decline eventually leads to a rise in diversification which could be due to the shark's deep-ocean habitat having been a beneficial factor in surviving mass extinctions (Guillaume 2013). Cross-referencing the dates of the

other four mass extinctions did not yield any positive correlations as the diversification rate remained relatively flat from 400 mya to 150 mya. Although this neutral change could have been due to a lack of fossil records.

Moving to the PyRate Graphs, there was a much better depiction of speciation and extinction rates through time as it factors in fossil sampling and fossilization patterns. These graphs aligned nicely and pointed out a very dramatic increase in extinction rates around the 300 mya time mark. This rise in extinction and decrease in speciation rates does not line up with any of the mass extinction rates as the Late Devonian mass extinction began around 357 mya and the following Permian extinction did not start until 251 mya. However, looking at the colored areas of the PyRate graphs does indicate a large range of data occurred at 300 mya and the mean rate could have been offset by extra large data points in the graph. Another interesting trend was the slight increase in extinction rates nearing the modern era. Just like the diversity v. age plots, mentioned above, these also highlight an dramatic decline in species. One possible reasoning behind this could be due to the increase in industrialization, leading to an increase in fishing technology and ultimately a rise in overfishing. Not only that, shark fin soup in many East and South Asian countries is highly coveted and this demand for shark fin could have contributed to the massive decline in shark species over time.

To test this hypothesis that shark finning is contributing to this 6th mass extinction, I cross-referenced data from the IUCN with PBDB to make data sets that would make out the threatened species onto the world-map. Looking at my graphs, I wanted to see if those species that were considered threatened were found populated along the coastlines of East and South Asian, that overfishing could be a probable cause. However, looking at the results, there was no clear correlation between geography and species endangerment. What I did notice was a highly concentrated amount of fossils occurrences in Europe and North America. This lack of fossil data in the Asian countries could have been due to lack of scientific interest and therefore less fossils being discovered in that area. Overall, there was no clear correlation.

Finally, this project was a logical way of framing a biological question I had about shark overfishing, and how humanity might be contributing to this 6th mass extinction. I was able to use shell commands to -grep, -sort, -cut, and -uniq the data from PBDB and format it into a desirable data set. From there, I also learned how to call functions and organize information into a presentable format. R Studio was a great way to take that formatted information in Python and add a graphical representation. There were definitely challenges along the way with formatting axis and creating data sets, but the resources available were very helpful. Finally, PyRate was a great tool to pull all this information together and present a complete analysis on the extinction and speciation rates of the taxon Selachii.

Link to Github

<https://github.com/kelseydjiang/EEB-177-Final-Project>

References

- Cornelia Dean (Nyt). “National Briefing Science and Health: Shark Fin Trade A Threat to Species, Report Finds.” *New York Times*, Late Edition (East Coast) ed.: 15. Sep 26 2006. ProQuest. Web. 2 June 2016.
- Guinot, Guillaume, Sylvain Adnet, Lionel Cavin, and Henri Cappetta. “Cretaceous Stem Chondrichthyans Survived the End-Permian Mass Extinction.” *Nature News*. Nature Publishing Group, 29 Oct. 2013. Web. 14 Mar. 2017.
- Kriwet, Jürgen, and Stefanie Klug. “Diversity and biogeography patterns of Late Jurassic neoselachians(Chondrichthyes: Elasmobranchii).” *Geological Society, London, Special Publications* 295.1 (2008): 55-70.
- Kriwet, Jürgen, and Michael J. Benton. “Neoselachian (Chondrichthyes, Elasmobranchii) diversity across the Cretaceous–Tertiary boundary.” *Palaeogeography, Palaeoclimatology, Palaeoecology* 214.3 (2004): 181-194.
- Silvestro, Daniele, Nicolas Salamin, and Jan Schnitzler. “PyRate: A New Program to Estimate Speciation and Extinction Rates from Incomplete Fossil Data.” *Methods in Ecology and Evolution*. N.p., 26 Sept. 2014. Web. 24 Mar. 2017.