EDS222_Final_Project

Julia Parish

2022/01/15

Contents


```
library(tidyverse)
library(here)
library(janitor)
library(grid)
library(sf)
library(ggmap)
library(ggrepel)
library(kableExtra)
library(sjPlot)
```

This statistical analysis was completed as an assignment for my Master's program course, Environmental Data Science 222: Statistics for Environmental Data Science.

Research Question

Were albatross populations impacted by the 2011 Tohoku tsunami event?

Researchers estimated the Tohoku tsunami flooded between 26% - 52% of all Black-footed albatross nests and impacted more than 275,000 albatross nests throughout Papahānaumokuākea [@Reynolds:2017]. This post will attempt to analyze and quantify the impact of the Tohoku tsunami on two Hawaiian albatross species, the Laysan albatross and the Black-footed albatross.

Background

The Hawaiian archipelago is home to thousands of albatrosses. Albatrosses are an incredible bird species that have inspired authors, the fashion industry, and birders across the globe. They live to be over 65 years old, have the largest wingspan of any bird, are monogamous rearers, and have nest-site fidelity (meaning they return to the same nest location every year). Three species of albatross breed in Hawaii, the Laysan (*Phoebastria immutabilis*), Black-footed (*Phoebastria nigripes*), and Short-tailed (*Phoebastria albatrus*) albatross. Laysan albatrosses are listed as near threatened due to threats from climate change to their habitat and breeding grounds and long-line fishing operations [@Arata:2009]. The IUCN Red List of Threatened Species assessed the Black-footed albatross as near threatened in 2020 [@iucn:2020]. The Short-tailed albatross almost went extinct in the early 1900s due to feather hunting and is currently listed as an endangered

Table 1: Hawaiian Albatross Species Names

Common_name	Hawaiian_name	Species_code	Scientific_n
Black-foot albatross	Ka upu	BFAL	Phoebastria nig
Laysan albatross	Mōlī	LAAL	Phoebastria im
Short-tailed albatross	Makalena or Ka upuakea	STAL	Phoebastria alb

species in the United States [@usfws:stal].

```
data.frame("Common_name" = c("Black-foot albatross", "Laysan albatross", "Short-tailed albatross"),
           "Hawaiian_name" = c("Ka upu", "Mōlī", "Makalena or Ka upuakea"),
           "Species_code" = c("BFAL","LAAL", "STAL"),
           "Scientific_name" = c("Phoebastria nigripes", "Phoebastria immutabilis", "Phoebastria albatru
  kable(caption = "Hawaiian Albatross Species Names") %>%
  kable_paper(full_width = FALSE) %>%
  kable_styling(latex_options = "striped",
                font_size = 15) %>%
  column_spec(1, bold = T) %>%
  row_spec(0, bold = T, color = "black")
```

Papahanaumokuakea, also known as the Northwestern Hawaiian Islands, are comprised of atolls, reefs, and pinnacles, and are where 95% of all Black-footed albatross and 99% of Laysan albatross nest [@Arata:2009]. These low-lying islands are at extreme risk of inundation from tsunamis [@Reynolds:2017]. On March 11, 2011, a 9.0 earthquake hit the Tōhoku region of Japan. The earthquake lasted over 6 minutes, creating a tsunami that impacted coastal areas and island nations throughout the Pacific region. Approximately 20,000 people lost their lives from the earthquake and resulting tsunami. The tsunami also killed or injured thousands of marine and terrestrial species. Many wildlife species found in the Papahanaumokuakea Marine National Monument (PMNM) were impacted by the Tohoku tsunami.

***Note**: The Short-tailed albatross was not included in this analysis as there is only one breeding pair in Hawaii.*

Analysis Plan

This statistical analysis will be testing the hypothesis if the Tohoku tsunami impacted albatross populations in Hawaii. The table below outlines the phases of the analysis.

$$H_0 : There\ was\ no\ impact\ of\ the\ Tohoku\ tsunami\ on\ albatross\ populations\ in\ Hawaii.$$

$$H_1 : There\ was\ an\ impact\ of\ the\ Tohoku\ tsunami\ on\ albatross\ populations\ in\ Hawaii.$$

```
data.frame("Phase" = c(1:5),
           "Description" = c("Identify research question",
                             "Collect data",
                             "Visualize data",
                             "Conduct regression analysis",
                             "Conclusion & Future Research")) %>%
  kable(caption = "Tohoku Tsunami Impact Analysis Plan Outline") %>%
  kable_paper(full_width = FALSE) %>%
  kable_styling(latex_options = "striped",
                font_size = 15) %>%
```

Table 2: Tohoku Tsunami Impact Analysis Plan Outline

Phase	Description
1	Identify research question
2	Collect data
3	Visualize data
4	Conduct regression analysis
5	Conclusion & Future Research

```
  column_spec(1, bold = T) %>%
  row_spec(0, bold = T, color = "black")
```

Collect Data

After researching several data sources for albatross populations, I retrieved banding data for both Laysan and Black-footed albatross from the USGS Bird Banding Laboratory (BBL). The BBL has data on bird species for the past 100 years. The data contains information about an individual bird's sex, age, health condition, and coordinates where the bird was banded. To access data from the BBL, it requires establishing an account on the USGS Bird Banding Lab Bander Portal website. Once you submit a data request, files will be available for download within 24 - 48 hours. I requested data for Black-footed and Laysan albatrosses in Hawaii between the years 1996 and 2020.

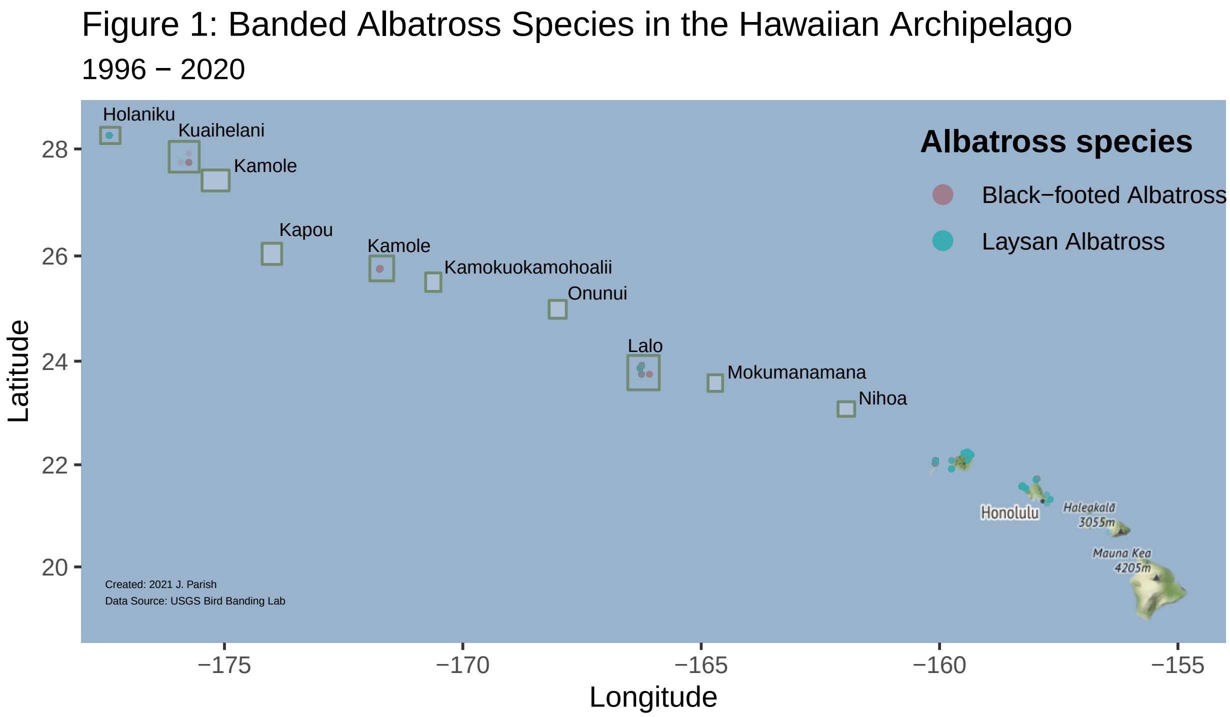
```
# read in banding data
albie_band <- read_csv(here("data", "laal_bfal_ca_hi_bbl2021.csv")) #both bfal and laal banding data fo
```

Visualize Data

Upon completion of data tidying and transformation, the next phase of analysis is to visualize the data.

Distribution Map The initial visualization was to map the banding data on to a map of Hawaii using the ggmap() package. The banding data point distribution accurately reflects the known species distribution of both species of albatross in Hawaii. The only two main Hawaiian Islands where albatross nest are Kauai and Oahu. This map shows most albatross populations reside with the PMNM.

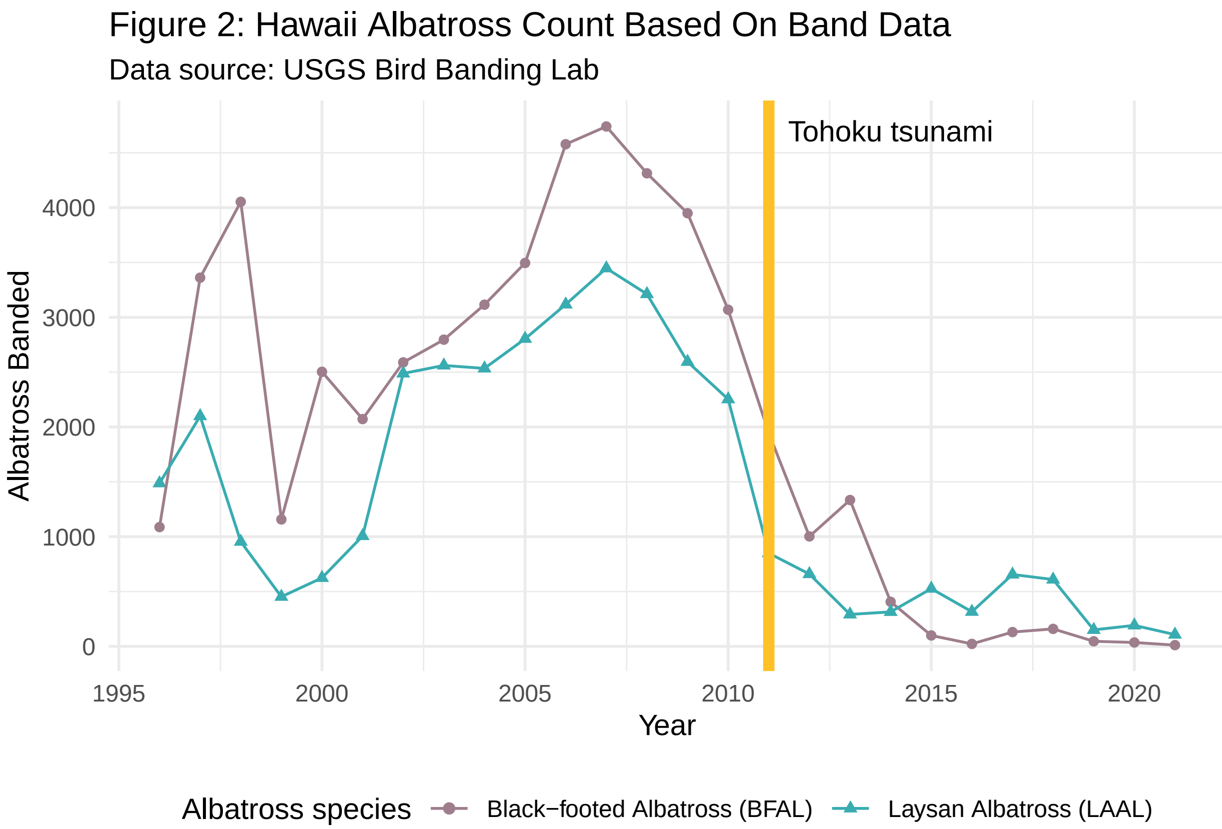
```
albie_map
```




```
band_count <- albie_band %>%
  group_by(species, year) %>%
  summarise(total = sum(count))
```

Scatterplot To conduct analysis on the albatross banding data, it is necessary to create a total count of albatross banded for each year from 1996 to 2020. Once a total count was summed, I used a line and point plot to visualize albatross count annually and added a line to indicate with the Tohoku tsunami occurred (2011). This plot does suggest that there may be a negative effect from the Tohoku tsunami on albatross populations as the counts drop significantly after 2011. It also shows that the number of banded Black-footed albatrosses declined more than Laysan albatrosses after the tsunami. This trend may reflect the data that Reynolds et al. found as Black-footed albatrosses' nest along coastal areas whereas Laysan albatrosses tend to nest more inland or at higher elevations on islands [@Reynolds:2017].

```
albie_count_plot
```



Population Means I calculated the mean of banded albatross species between 1996 and 2020 as well as comparing the population means of each species pre- and post-tsunami event. These means indicate that more black-foot albatross is banded than Laysan albatross in Hawaii. This is an interesting data point as the total Laysan albatross population is larger than the Black-footed albatross population [@plan:2012]. The Laysan albatross actually holds the honor of having the largest population of all albatross species in the world (out of 21 species). Reviewing the two population mean tables also indicated that the number of banded albatrosses declined after the Tohoku tsunami.

```
pop_params_table <- knitr::kable(pop_params_summary,
                                 digits = 0,
                                 col.names = c('Albatross Species', 'Population Mean', 'Population Medi
                                 align = "lccc",
                                 caption = "Hawaii Albatross Population Summary 1996 - 2021") %>%
  kable_paper(full_width = FALSE) %>%
  kable_styling(latex_options = "striped",
                font_size = 15) %>%
  column_spec(1, bold = T) %>%
  row_spec(0, bold = T, color = "black") #%>%
  # save_kable(here("images/pop_params.png"))
```

```
pop_params_table
```

```
pop_mean_table
```

Conduct and Interpret Regression Analysis

To determine the distribution of the count data for both albatross species, I created Q-Q plots. The regression analysis used on the band count data was linear regression.

Table 3: Hawaii Albatross Population Summary 1996 - 2021

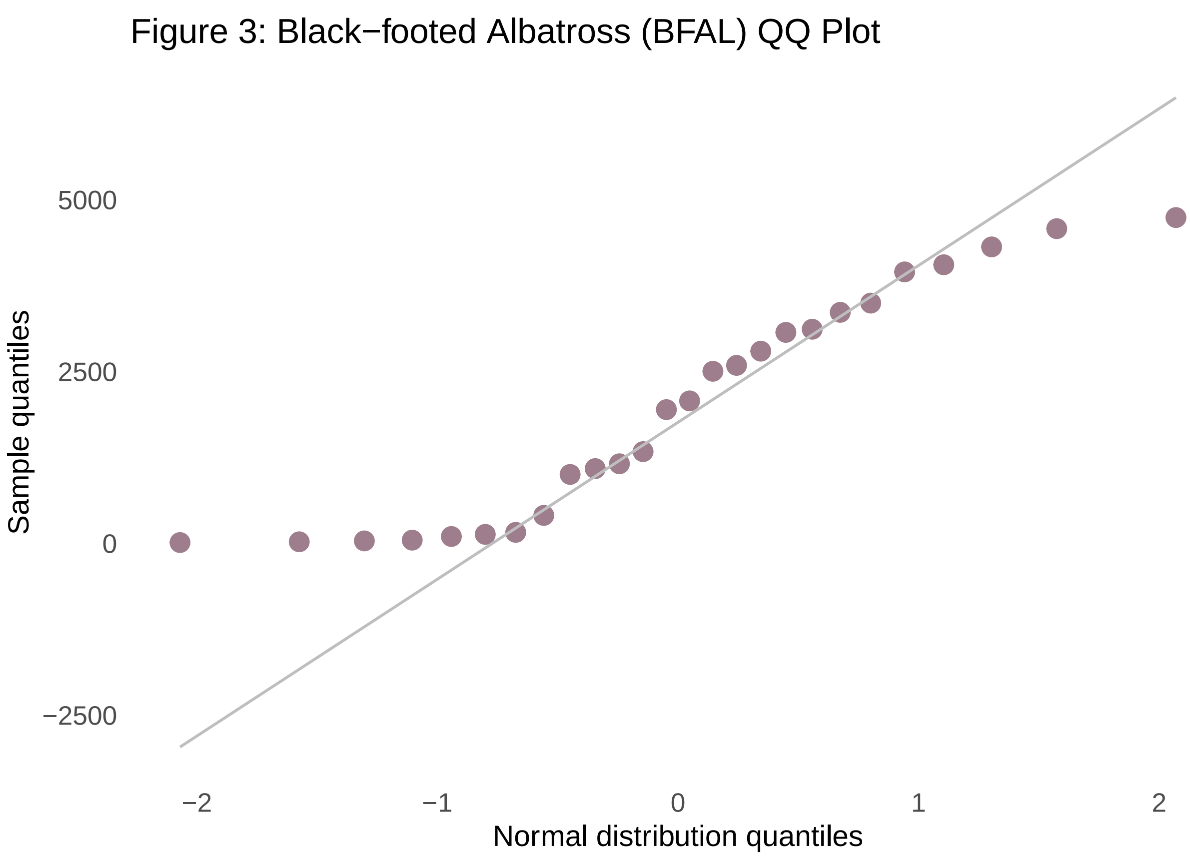
Albatross Species	Population Mean	Population Median	Population M
BFAL	2003	2008	4740
LAAL	1396	902	3447

Table 4: Hawaii Albatross Count Means

Species	Pre_Pop_Mean	Post_Pop_Mean
BFAL	3627.11	517.8
LAAL	2778.89	455.9

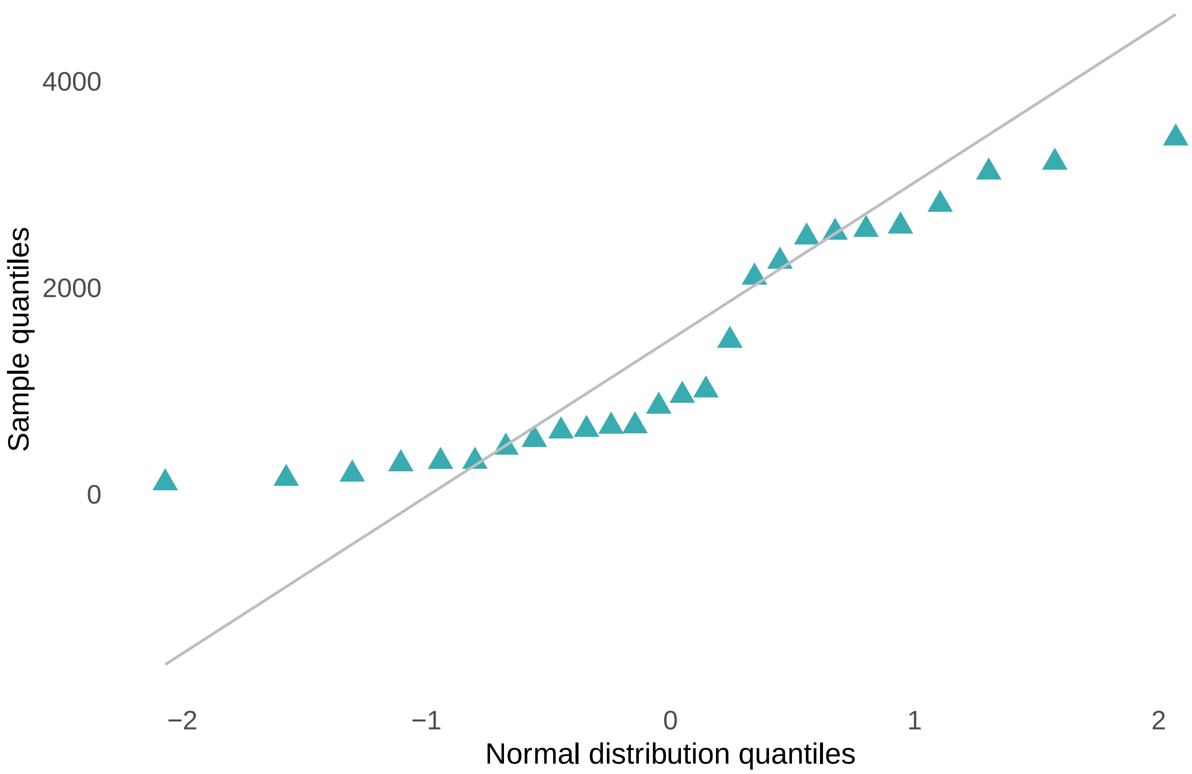
QQ Plot The QQ plots for both the Black-footed and Laysan albatross show that the distribution is kurtosis, or they have heavy tails. Neither count data for the albatross species have a normal distribution based on the QQ plots.

```
bfal_qqplot
```




```
laal_qqplot
```

Figure 4: Laysan Albatross (LAAL) QQ Plot



Simple Linear Regression The first regression I conducted on the albatross banding data was a simple linear regression.

$$\text{Albatross count}_i = \beta_0 + \beta_1 \text{tsunami event} + \varepsilon_i$$

```
tab_model(post_tsunami_simple_mod ,
          pred.labels = c("Intercept (mean albatross count pre-tsunami)", "Avg. banded albatross count p
          string.ci = "Conf. Int (95%)",
          string.p = "P-value",
          title = "Table 4: Simple Linear Regression Model for Banded Albatross in Hawaii",
          digits = 4)
```

Table 4: Simple Linear Regression Model for Banded Albatross in Hawaii

total

Predictors

Estimates

Conf. Int (95%)

P-value

Intercept (mean albatross count pre-tsunami)

2616.9667

2274.2624 – 2959.6710

<0.001

Avg. banded albatross count post-tsunami

-2169.0121

-2695.8899 – -1642.1343

<0.001

Observations

52

R2 / R2 adjusted

0.578 / 0.569

The result of the simple regression shows that the intercept, 2616, is the mean number of banded albatrosses in Hawaii prior to the Tohoku tsunami event in 2011. The second predictor, or **1**, is the mean banded albatrosses less than the intercept for any year after the tsunami. For any year after the tsunami, the mean count is 447. There is a 5% chance that the average banded albatross will be between 2,2274 and 2,959 prior to the Tohoku tsunami. The confidence interval post-tsunami shows that the mean will be -2,695 and -1,642 of the total albatrosses. Because the p-value is smaller than 0.05, the null hypothesis that there was no impact from the Tohoku tsunami on albatross counts is rejected. The data provides convincing evidence that there is a negative difference in mean number of albatrosses post-tsunami. There is a statistically significant difference, at the 5% significance level, in the count of albatrosses in Hawaii. This model explains 57% of the variation in the albatross count data around its mean.

Multiple Linear Regression The second regression I conducted on the albatross banding data was a multiple linear regression with an interaction term of the tsunami event on each albatross species.

$$\text{Albatross count}_i = \beta_0 + \beta_1 \text{tsunami event} + \beta_2 \text{albatross species} + \beta_3 \text{tsunami event:species} + \varepsilon_i$$

```
tab_model(interaction_model,
          pred.labels = c("Intercept (mean BFAL count pre-tsunami)", "Mean BFAL count post-tsunami", "Me
          string.ci = "Conf. Int (95%)",
          string.p = "P-value",
          title = "Table 5: Multiple Linear Regression Model Results for Banded Albatross in Hawaii",
          digits = 4)
```

Table 5: Multiple Linear Regression Model Results for Banded Albatross in Hawaii

total

Predictors

Estimates

Conf. Int (95%)

P-value

Intercept (mean BFAL count pre-tsunami)

3125.1333

2676.1144 – 3574.1523

<0.001

Mean BFAL count post-tsunami

-2653.4061

-3343.7333 – -1963.0788

<0.001

Mean LAAL count pre-tsunami

-1016.3333

-1651.3420 – -381.3246

0.002

Mean LAAL count compared to BFAL count post-tsunami

968.7879

-7.4823 – 1945.0580

0.052

Observations

52

R2 / R2 adjusted

0.653 / 0.631

The result of the simple regression shows that the intercept, 3125, is the mean number of banded Black-footed albatrosses in Hawaii prior to the Tohoku tsunami event in 2011. The second predictor, or **1**, is the mean banded Black-footed albatrosses less than the intercept for any year after the tsunami. For any year after the tsunami, the mean count is 472. The **2** is the mean banded Laysan albatrosses less than the mean number of Black-footed albatross before the tsunami. The mean Laysan albatross count prior to the tsunami was 2,109. The **3** is the mean Laysan albatrosses count post-tsunami compared to Black-footed albatrosses' post-tsunami. There are, on average, 969 more Laysan albatrosses than Black-footed albatrosses for any year after the Tohoku tsunami.

Since the p-values are smaller than 0.05, the null hypothesis that there was no impact from the Tohoku tsunami on albatross counts is rejected. The data provides convincing evidence that there is a negative difference in mean number of both Black-footed and Laysan albatrosses post-tsunami. There is a statistically significant difference, at the 5% significance level, in the count of albatross in Hawaii. The p-value is greater than 0.05 when comparing the mean Laysan albatross count post-tsunami to the Black-footed albatross. This predictor may indicate that Laysan albatross counts were not as significantly impacted by the tsunami as Black-footed albatross. This model explains 63% of the variation in the albatross count data around its mean.

Conclusion and Future Research

In conclusion, the linear regression model predicts that there are significant negative relationships between the independent variable (years before and after tsunami) and the dependent variable (count of banded albatross). It was also found that a multiple regression with an interaction term based on the tsunami event impact of each albatross species is a better model fit. This analysis indicates that the Tohoku tsunami had a negative impact on albatross populations in Hawaii.

For future analysis on the impact of the Tohoku tsunami, I would like to research what other variables may be influencing the number of albatrosses being banded by biologist in Hawaii each year. I have a suspension that there has been reduced banding effort throughout the archipelago since the tsunami, especially in 2020 due to the pandemic. I do not know why there would be a reduction in banding effort, but this may be a significant influence on the count data used for this analysis. I would also like to contact researchers in Hawaii to determine if there is more comprehensive census data for albatross.