Winter Habitat Use by Common Loons (Gavia Immer)

on Mount Desert Island, Maine

Ellie Gabrielson, Will Draxler, Autumn Pauly

03/12/2024

Contents

Background Information

Introduction

Although Common Loons (Gavia immer) have been studied extensively during the breeding season not much is known of their behavior on their ocean wintering grounds, where Loons spend a significant part of their lives (Bent, 2009). The few studies that do look at winter behavior often yield conflicting results (McIntyre, 1978; Daub, 1989; Ford and Gieg, 1995). The possible use of different shoreline habitats for activities such as feeding, sleeping, or taking shelter is worthy of study. McIntyre (1978) found common loons were semi-social and formed loose rafts and that loon behavior varied with tides, weather, and time of day. Daub (1989) and later Ford and Gieg (1995) found refuting evidence suggesting loons were neither social nor territorial, and did not change their behavior much by tide. Holm and Burger (2002) found that the general pattern for loons was that they foraged during slack and main-flow water. Additionally, the possible difference in dive times as a result of varying tidal heights is worthy of study. Thompson and Price (2006) observed that dive times were longer during low tide in comparison to other tidal stages, though more studies should be conducted to assess the use of tidal stage and behavior.

Our general hypothesis when analyzing this data is that environmental conditions affect the behavior of the *G. immer*. The null hypothesis to this statement is that environmental conditions do not affect the behavior of the *G. immer*.

Methods

In 2023, observations were conducted at nine established observation sites along the coastline of Mount Desert Island in Hancock County (Maine). Observations occurred from January 2023 to March 2023 on a biweekly basis, where each waterfowl individual was recorded as well as their behaviors, referencing Daub's (1989) list of behaviors, and coordinate location. Locational coordinates were obtained by estimating the individual's rough location digitally (accurate to + or - 10 m). Each transect site was visited for roughly 5-10 minutes, where data collection followed initial observation. Utilizing a Kestrel 2500, weather data was recorded accurately to + or - 1°C of temperatures and within 3% of wind speed. Referencing Daub's (1989) list of behaviors, each species' behavior during the observation was noted. The weather and environmental data recorded included wind speed and direction, barometer, cloud cover, precipitation, humidity, temperature, tidal percentage, and wave class (Beaufort Wind Speed). Tidal data was gathered from NOAA's digital tidal charts.

In 2024, this study was repeated with the addition of observational variables including species dive times and tidal height. Dive times were recorded by using binoculars and a handheld timing watch to observe and time an individual several times. Tidal height was gathered from NOAA's digital tidal chart.

Data Analysis

Load Packages and Data

```r
#install.packages("tidyverse")
library(tidyverse)
library(rmarkdown)
library(scales)
library(ggridges)
library(rstatix)
library(parsnip)
```

```r
loons_2023 <- read_csv("loon_2023_tidy.csv")
loons_dive_2024 <- read_csv("loons_dive_2024.csv")
loons_2024 <- read_csv("loons_2024.csv")
```

Diving Loons 2024

```
# loons_dive_2024 <- loons_2024 %>%
#   filter(species == "common_loon",
#          behavior == "diving")
```

```
# loons_dive_2024 <- loons_dive_2024 %>%
#   separate_longer_delim(dive_time_obs, delim = ", ") %>%
#   separate_longer_delim(dive_time_obs, delim = ",") %>%


# loons_dive_2024 <- loons_dive_2024
#   mutate(dive_time_obs = as.numeric(dive_time_obs))
```

General Statistics

The mean dive time of loons during this study was 41.95 seconds with a standard error of 1.53.

```
#summary of dive times in 2024
summary(loons_dive_2024)
```

```
##      tide             behavior          dive_time_obs    tidal_height
##  Length:176         Length:176         Min.   : 8.00   Min.   :-0.010
##  Class :character   Class :character   1st Qu.:26.00   1st Qu.: 1.530
##  Mode  :character   Mode  :character   Median :36.00   Median : 2.750
##                                        Mean   :36.87   Mean   : 2.269
##                                        3rd Qu.:47.25   3rd Qu.: 3.300
##                                        Max.   :78.00   Max.   : 3.710
##
##     latitude       longitude      meters_offshore    location
##  Min.   :44.39   Min.   :-68.25   Min.   :  2.00   Length:176
##  1st Qu.:44.40   1st Qu.:-68.25   1st Qu.: 15.00   Class :character
##  Median :44.41   Median :-68.24   Median : 30.00   Mode  :character
##  Mean   :44.41   Mean   :-68.23   Mean   : 53.01
##  3rd Qu.:44.42   3rd Qu.:-68.22   3rd Qu.: 60.00
##  Max.   :44.42   Max.   :-68.19   Max.   :300.00
##                                   NA's   :2
##  general_tide
##  Length:176
##  Class :character
##  Mode  :character
##
##
##
##
```

```
#standard error function
std.error <- function(x) sd(x)/sqrt(length(x))

std.error(loons_dive_2024$dive_time_obs)
```

```
## [1] 1.094932
```

There was an average of 9.55 loons observed per day during this study, with there being 2.12 being observed at each site.

```r
loons_2023 %>%
  summary()
```

```
##       year          month            day              date           
##  Min.   :2023   Min.   :1.000   Min.   : 4.00   Min.   :2023-01-04  
##  1st Qu.:2023   1st Qu.:1.000   1st Qu.:12.00   1st Qu.:2023-01-14  
##  Median :2023   Median :1.000   Median :17.00   Median :2023-01-27  
##  Mean   :2023   Mean   :1.267   Mean   :18.08   Mean   :2023-01-26  
##  3rd Qu.:2023   3rd Qu.:2.000   3rd Qu.:27.00   3rd Qu.:2023-02-05  
##  Max.   :2023   Max.   :2.000   Max.   :31.00   Max.   :2023-02-20  
##                                                                     
##      time          tide_percentage      tide               cal           
##  Length:131        Min.   :  0.00   Length:131         Length:131        
##  Class1:hms        1st Qu.: 10.00   Class :character   Class :character  
##  Class2:difftime   Median : 20.00   Mode  :character   Mode  :character  
##  Mode  :numeric    Mean   : 38.08                                        
##                    3rd Qu.: 62.50                                        
##                    Max.   :100.00                                        
##                                                                          
##    location           species              number     latitude    
##  Length:131         Length:131         Min.   :1   Min.   :44.22  
##  Class :character   Class :character   1st Qu.:1   1st Qu.:44.28  
##  Mode  :character   Mode  :character   Median :1   Median :44.29  
##                                        Mean   :1   Mean   :44.31  
##                                        3rd Qu.:1   3rd Qu.:44.33  
##                                        Max.   :1   Max.   :44.40  
##                                                                   
##    longitude      meters_offshore   behavior         behavior_notes    
##  Min.   :-68.41   Min.   :  0.0   Length:131         Length:131        
##  1st Qu.:-68.31   1st Qu.: 30.0   Class :character   Class :character  
##  Median :-68.25   Median : 85.0   Mode  :character   Mode  :character  
##  Mean   :-68.27   Mean   :117.9                                        
##  3rd Qu.:-68.20   3rd Qu.:164.5                                        
##  Max.   :-68.18   Max.   :450.0                                        
##                                                                        
##  sky_condition      weather_notes      precipitation       temperature   
##  Length:131         Length:131         Length:131         Min.   :25.10  
##  Class :character   Class :character   Class :character   1st Qu.:33.20  
##  Mode  :character   Mode  :character   Mode  :character   Median :39.40  
##                                                           Mean   :38.08  
##                                                           3rd Qu.:42.15  
##                                                           Max.   :49.50  
##                                                                          
##    wind_speed     wind_direction       barometer        humidity    
##  Min.   : 0.000   Length:131         Min.   :28.85   Min.   :38.00  
##  1st Qu.: 2.000   Class :character   1st Qu.:29.70   1st Qu.:64.00  
##  Median : 3.100   Mode  :character   Median :29.88   Median :73.00  
##  Mean   : 5.047                      Mean   :29.81   Mean   :70.45  
```

```
##  3rd Qu.: 8.850                      3rd Qu.:30.05   3rd Qu.:82.00
##  Max.   :15.000                      Max.   :30.22   Max.   :97.00
##                                      NA's   :4
##  cloud_coverage     wave_class    human_activity     water_activity
##  Min.   :  0.00   Min.   :0.500   Length:131         Length:131
##  1st Qu.: 40.00   1st Qu.:1.000   Class :character   Class :character
##  Median : 85.00   Median :1.500   Mode  :character   Mode  :character
##  Mean   : 69.39   Mean   :1.698
##  3rd Qu.:100.00   3rd Qu.:2.000
##  Max.   :100.00   Max.   :3.500
##
##  shelter_gradient      notes            moon_phase        overall_abundance
##  Length:131         Length:131         Length:131         Min.   : 4.00
##  Class :character   Class :character   Class :character   1st Qu.: 7.00
##  Mode  :character   Mode  :character   Mode  :character   Median : 8.00
##                                                           Mean   : 9.55
##                                                           3rd Qu.:12.00
##                                                           Max.   :15.00
##
##  specific_abundance shelter_factor     general_tide
##  Min.   :0.000      Length:131         Length:131
##  1st Qu.:1.000      Class :character   Class :character
##  Median :2.000      Mode  :character   Mode  :character
##  Mean   :2.122
##  3rd Qu.:3.000
##  Max.   :6.000
##
```

General Tidal Stage and Behavior for 2023

The first analysis that will be conducted will be investigating the potential relationship between general tidal stages and *G. immer* behavior. As stated above, the behaviors chosen are referencing Daub's (1989) list of behaviors. The tidal condition was sorted into three categories: `high`, `mid`, and `low`.

It was hypothesized that tidal height would significantly affect the behavior of *G. immer* The null hypothesis stated that tidal height would not have a significant effect on *G. immer* behavior.

Statistics

Chi-Squared Test of Independence The Chi-Squared Test of Independence examines whether there is a relationship between two categorical variables. Our null hypothesis, that tidal height will not have a significant effect on *G. immer* behavior, assumes that tidal condition and behavior are independent. Our hypothesis, that tidal height will have a significant effect on *G. immer* behavior, assumes that tidal condition and behavior are dependent.

Expected Frequencies and Contingency Table Before performing the chi-square test, we need to calculate the expected frequencies for each cell in the contingency table. These expected frequencies represent what you would expect to observe if there were no association between tidal condition and behavior.

```
##
##          ashore diving drifting maintenance peering
##   high        0    111        7           2       1
```

```
##   low           1     65        26            1       2
##   low_ebb       0     13         0            0       0
##   mid           0     77         0            0       1
```

Chi-square Test Now that we have the contingency table of expected frequencies, we can perform the statistical test.

```
##
##  Pearson's Chi-squared test
##
## data:  contingency_table1
## X-squared = 47.09, df = 12, p-value = 4.497e-06
```

REPORT: A Chi-squared test indicated a significant relationship between *G. immer* behavior and general tidal stage (Xˆ2(12, n=307) = 75.8, df = 12, p-value > 0.005). This indicates that there is a relationship between general tidal stage and *G. immer* behavior, allowing the null hypothesis to be rejected.

General Tidal Stage and Abundance at each site

The next analysis that will be conducted will be investigating the potential relationship between general tidal stages and *G. immer* abundance at each site.

The hypothesis is that tidal height would significantly affect the abundance of *G. immer* The null hypothesis stated that tidal height would not have a significant effect on *G. immer* abundance at each site.

Statistics

Chi-Squared Test of Independence The Chi-Squared Test of Independence examines whether there is a relationship between two categorical variables.

Expected Frequencies and Contingency Table Before performing the chi-square test, we need to calculate the expected frequencies for each cell in the contingency table. These expected frequencies represent what you would expect to observe if there were no association between tidal condition and behavior.

```
##
##        BAR HARBOR PIER BRACY HARBOR NORTHEAST HARBOR SAND BEACH SEAL COVE
##   HIGH               6            1                2          0         4
##   LOW               12            7                6          8        15
##   MID               12            1                4          3         3
##
##        SEAL HARBOR BEACH SEAWALL SOMES SOUND SOUTHWEST HARBOR
##   HIGH                 7      10           1                2
##   LOW                  9       6           3                2
##   MID                  4       2           0                1
```

```
##
##              bar_harbor_bar bar_harbor_pier coa_pier hulls_cove sonogee
##   high_ebb                0               0       99         32      86
##   high_flood              0              29        0         78       8
##   high_slack              0               0        0         25       0
##   low_ebb                 0               0        9          2       0
```

```
##   low_flood              0               2        29          67        0
##   low_slack            121              16         1           4       17
##   mid_ebb                0               0        27           6        0
##   mid_flood              0              24         0          56        0
##
##             the_hotel_motel_holiday_inn
##   high_ebb                            0
##   high_flood                          0
##   high_slack                         31
##   low_ebb                             0
##   low_flood                          67
##   low_slack                           0
##   mid_ebb                            22
##   mid_flood                          69
```

Chi-square Test Now that we have the contingency table of expected frequencies, we can perform the statistical test.

REPORT: A Chi-squared test indicated no significant relationship between *G. immer* abundance and tidal condition at each site (X^2(16, n=131) = 26.28, df = 16, p-value = 0.05017). This indicates that there is no relationship between tidal height and *G. immer* abundance at each site, allowing us to reject the hypothesis.

Tidal Conditions and Dive Times in 2024

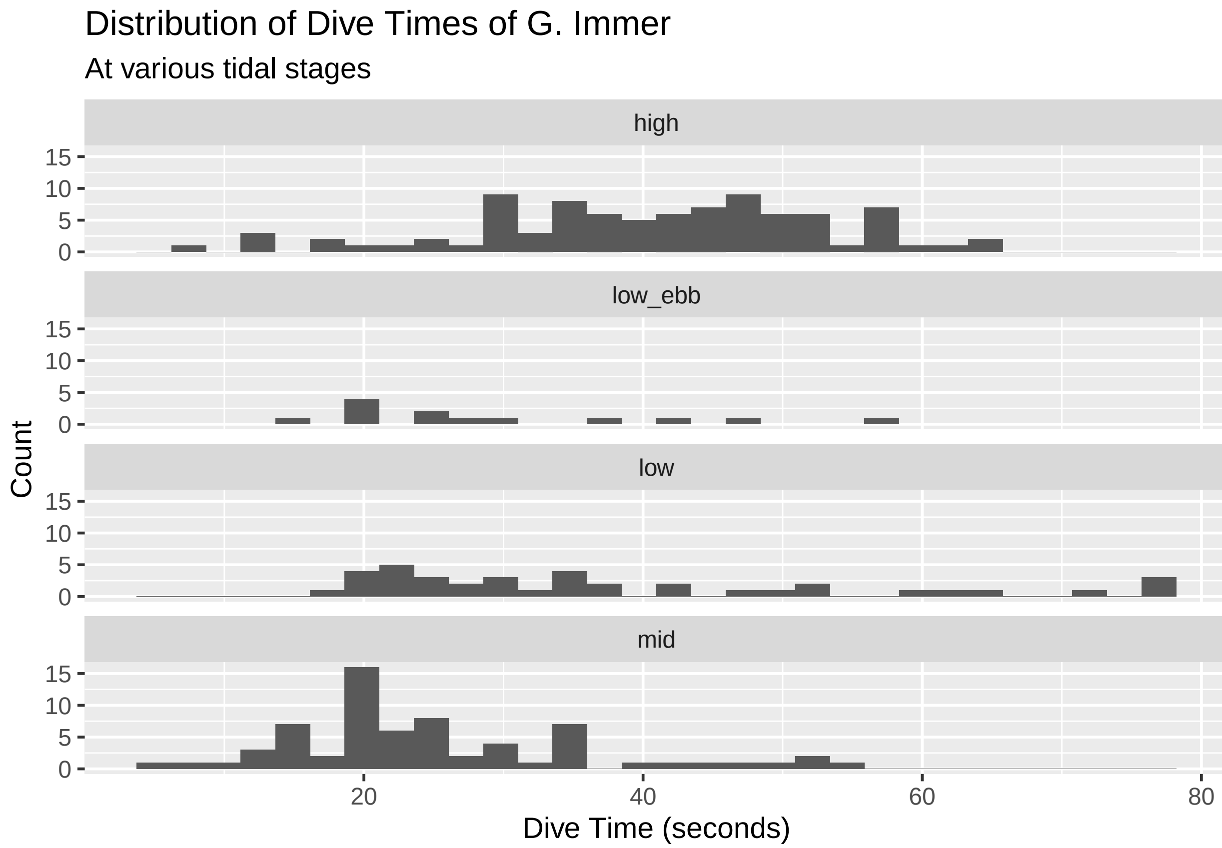
This study aimed to investigate the potential effect of tidal stage on the dive times of *G. immer*. The hypothesis posited that low tide would result in significantly longer dive times for *G. immer*. Conversely, the null hypothesis stated that tidal stage would not have a significant impact on *G. immer* dive times.

```r
#This is specifically to export coordinates, I know it's big but I didn't have a better solution

loons_dive_2024_1 <- loons_dive_2024 %>%
  filter(is.na(longitude2) == F ) %>%
  mutate(longitude = longitude2) %>%
  select(-longitude2)

loons_dive_2024_2 <- loons_dive_2024 %>%
  filter(is.na(longitude) == F ) %>%
  select(- longitude2)

loons_dive_2024 <- loons_dive_2024_1 %>%
  bind_rows(loons_dive_2024_2) %>%
  write_csv("loons_dive_2024_GPS.csv")
```

Visualizations

As can be seen below, we do not have a normal distribution for the variable `dive_time_obs`. We will need to perform a non-parametric test.

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Statistics The non-parametric nature of our dataset suggests that we need to use the Kruskal-Wallis H test (the "one-way ANOVA on ranks"), which is the non-parametric alternative to the one-way ANOVA test.

```
##
##  Kruskal-Wallis rank sum test
##
## data:  dive_time_obs by general_tide
## Kruskal-Wallis chi-squared = 51.221, df = 3, p-value = 4.389e-11
```

```
## # A tibble: 6 x 9
##   .y.         group1 group2    n1    n2 statistic        p    p.adj p.adj.signif
## * <chr>       <chr>  <chr>  <int> <int>     <dbl>    <dbl>    <dbl> <chr>
## 1 dive_time_~ high   low_e~    88    13    -2.81  4.93e- 3 2.96e- 2 *
## 2 dive_time_~ high   low       88    38    -1.75  8.07e- 2 4.84e- 1 ns
## 3 dive_time_~ high   mid       88    67    -6.99  2.82e-12 1.69e-11 ****
## 4 dive_time_~ low_e~ low       13    38     1.54  1.22e- 1 7.34e- 1 ns
## 5 dive_time_~ low_e~ mid       13    67    -0.981 3.27e- 1 1   e+ 0 ns
## 6 dive_time_~ low    mid       38    67    -3.91  9.29e- 5 5.57e- 4 ***
```

REPORT: The Kruskal-Wallis test revealed that there was no significant difference in dive times among the various tidal stages (Kruskal-Wallis rank sum test; $H(2) = 5.82$, $P = 0.054$). This suggests that tidal stage does not have a statistically significant effect on *G. immer* dive times, allowing us to reject our hypothesis. This implies that other factors besides tidal stage could be influencing dive behavior in *G. immer*.

Site Location and Dive Times in 2024

This study aimed to investigate the potential effect of site locations on the dive times of *G. immer*. The hypothesis posited that sites would result in significantly different dive times for *G. immer*. Conversely, the null hypothesis stated that location would not have a significant impact on *G. immer* dive times.

```r
loons_2024_loon <- loons_2024_loon %>%
  separate_longer_delim(dive_time_obs, delim = ",") %>%
  mutate(dive_time_obs = as.numeric(dive_time_obs))
```

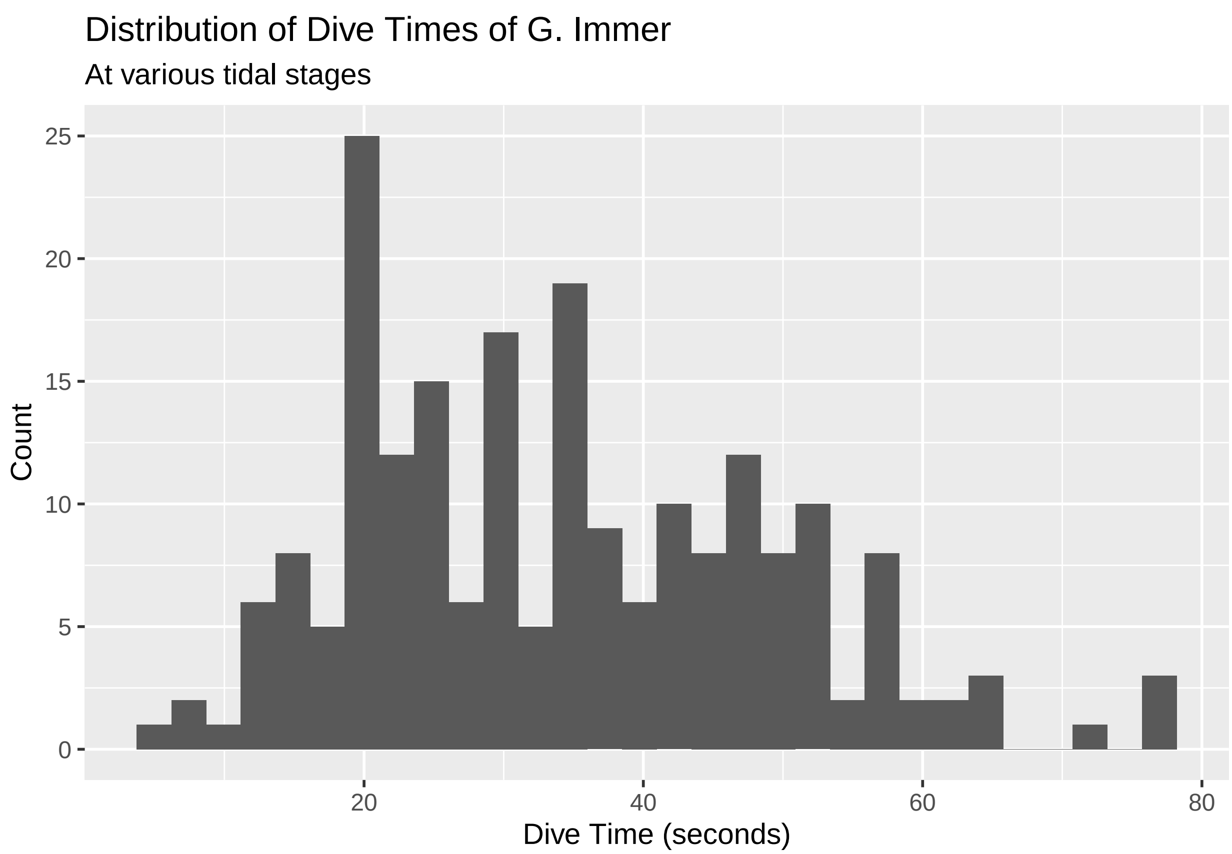
```
## Warning: There was 1 warning in 'mutate()'.
## i In argument: 'dive_time_obs = as.numeric(dive_time_obs)'.
## Caused by warning:
## ! NAs introduced by coercion
```

Visualizations

As can be seen below, we do not have a normal distribution for the variable `dive_time_obs`. We will need to perform a non-parametric test.

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```

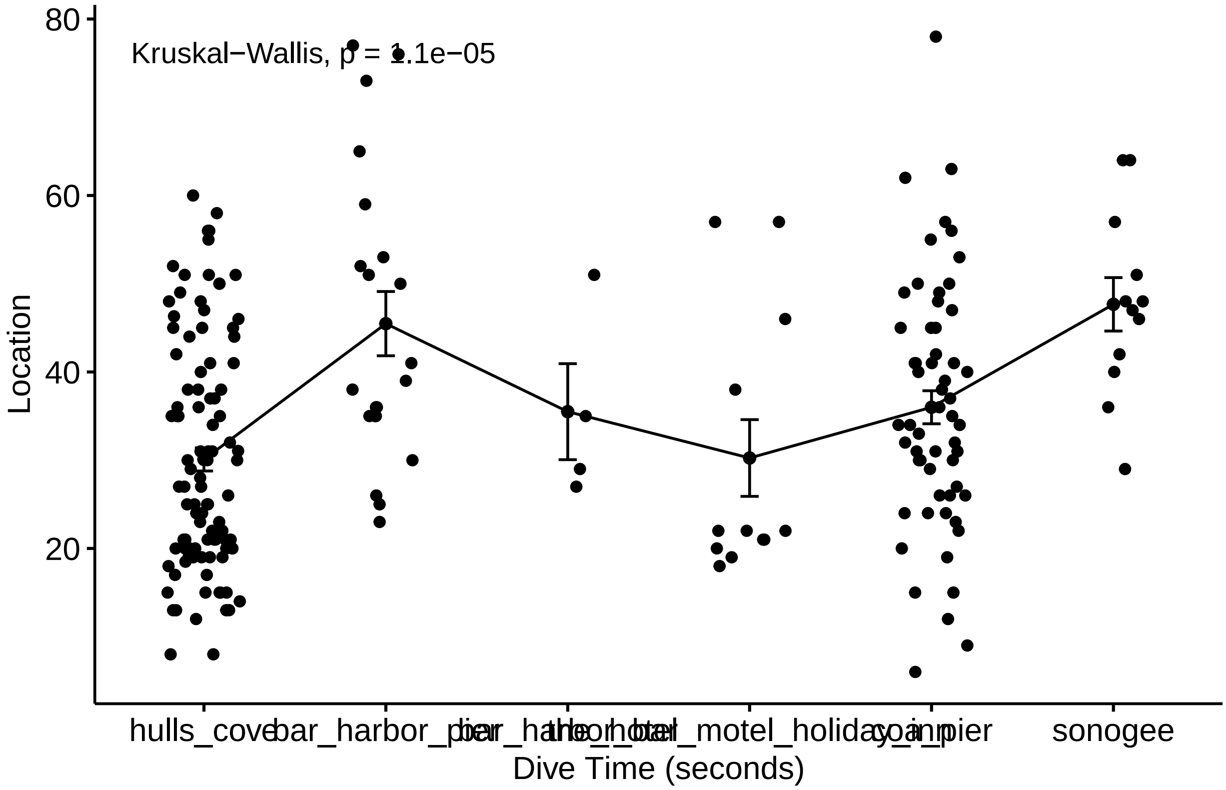
```
## Warning: Removed 864 rows containing non-finite outside the scale range
## ('stat_bin()').
```

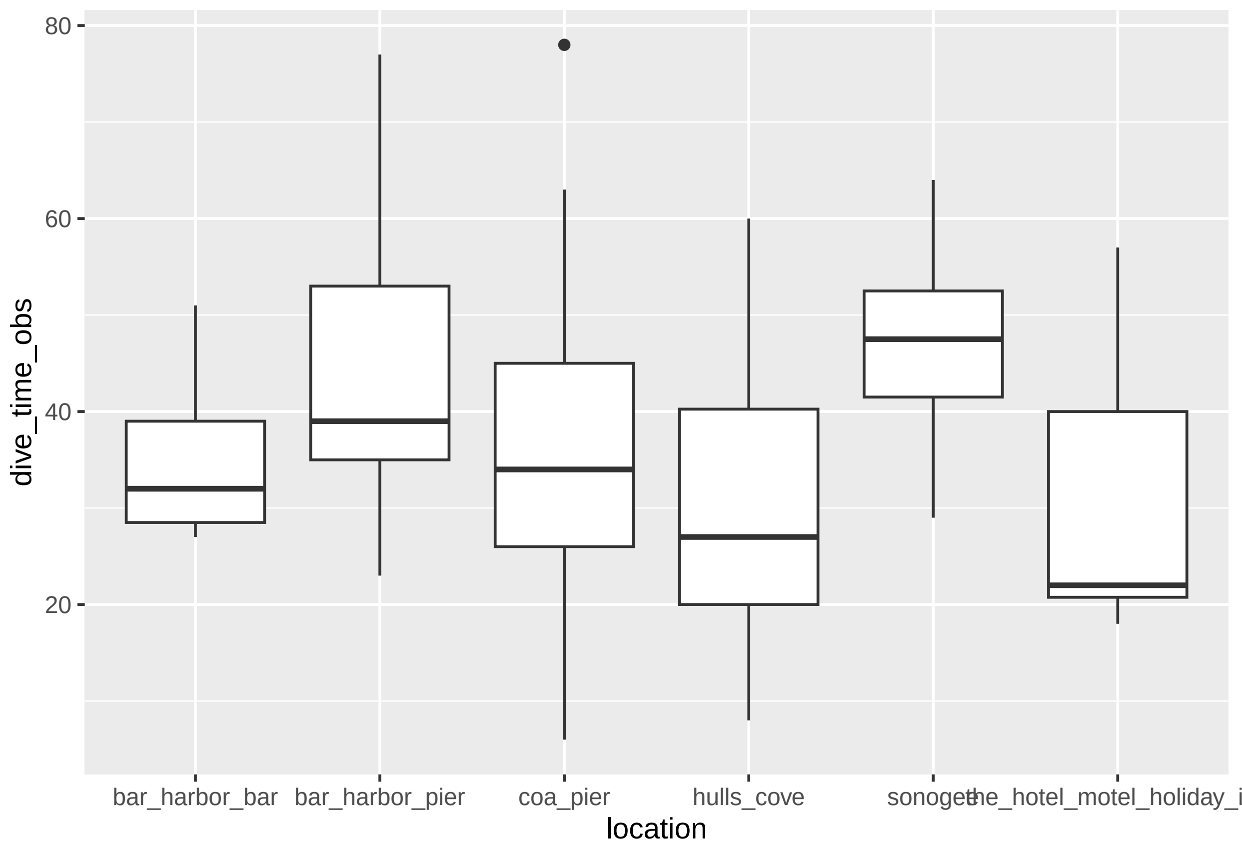


Statistics The non-parametric nature of our dataset suggests that we need to use the Kruskal-Wallis H test (the "one-way ANOVA on ranks"), which is the non-parametric alternative to the one-way ANOVA test.

```
##
##  Kruskal-Wallis rank sum test
##
## data:  dive_time_obs by location
## Kruskal-Wallis chi-squared = 30.615, df = 5, p-value = 1.116e-05
```

```
## # A tibble: 15 x 9
##    .y.          group1 group2    n1    n2 statistic       p   p.adj p.adj.signif
##  * <chr>        <chr>  <chr>  <int> <int>     <dbl>   <dbl>   <dbl> <chr>
##  1 dive_time_o~ bar_h~ bar_h~     4    21    0.910  3.63e-1 1   e+0 ns
##  2 dive_time_o~ bar_h~ coa_p~     4    57   -0.0359 9.71e-1 1   e+0 ns
##  3 dive_time_o~ bar_h~ hulls~     4   100   -0.891  3.73e-1 1   e+0 ns
##  4 dive_time_o~ bar_h~ sonog~     4    12    1.33   1.84e-1 1   e+0 ns
##  5 dive_time_o~ bar_h~ the_h~     4    12   -0.864  3.88e-1 1   e+0 ns
##  6 dive_time_o~ bar_h~ coa_p~    21    57   -2.02   4.37e-2 6.56e-1 ns
##  7 dive_time_o~ bar_h~ hulls~    21   100   -3.96   7.51e-5 1.13e-3 **
##  8 dive_time_o~ bar_h~ sonog~    21    12    0.746  4.55e-1 1   e+0 ns
##  9 dive_time_o~ bar_h~ the_h~    21    12   -2.75   5.98e-3 8.97e-2 ns
## 10 dive_time_o~ coa_p~ hulls~    57   100   -2.62   8.67e-3 1.30e-1 ns
## 11 dive_time_o~ coa_p~ sonog~    57    12    2.47   1.35e-2 2.02e-1 ns
## 12 dive_time_o~ coa_p~ the_h~    57    12   -1.51   1.31e-1 1   e+0 ns
## 13 dive_time_o~ hulls~ sonog~   100    12    4.00   6.46e-5 9.70e-4 ***
## 14 dive_time_o~ hulls~ the_h~   100    12   -0.145  8.85e-1 1   e+0 ns
## 15 dive_time_o~ sonog~ the_h~    12    12   -3.10   1.95e-3 2.92e-2 *
```


Kruskal−Wallis, p = 1.1e−05
80
60
40
20
Location
hulls_cove
bar_harbor_pier
bar_harbor_hotel
the_motel
holiday_inn
coal_pier
sonogee
Dive Time (seconds)



REPORT: The Kruskal-Wallis test revealed that there was no significant difference in dive times among the various tidal stages (Kruskal-Wallis rank sum test; H(2) = 5.82, P = 0.054). This suggests that tidal stage does not have a statistically significant effect on *G. immer* dive times, allowing us to reject our hypothesis. This implies that other factors besides tidal stage could be influencing dive behavior in *G. immer*.

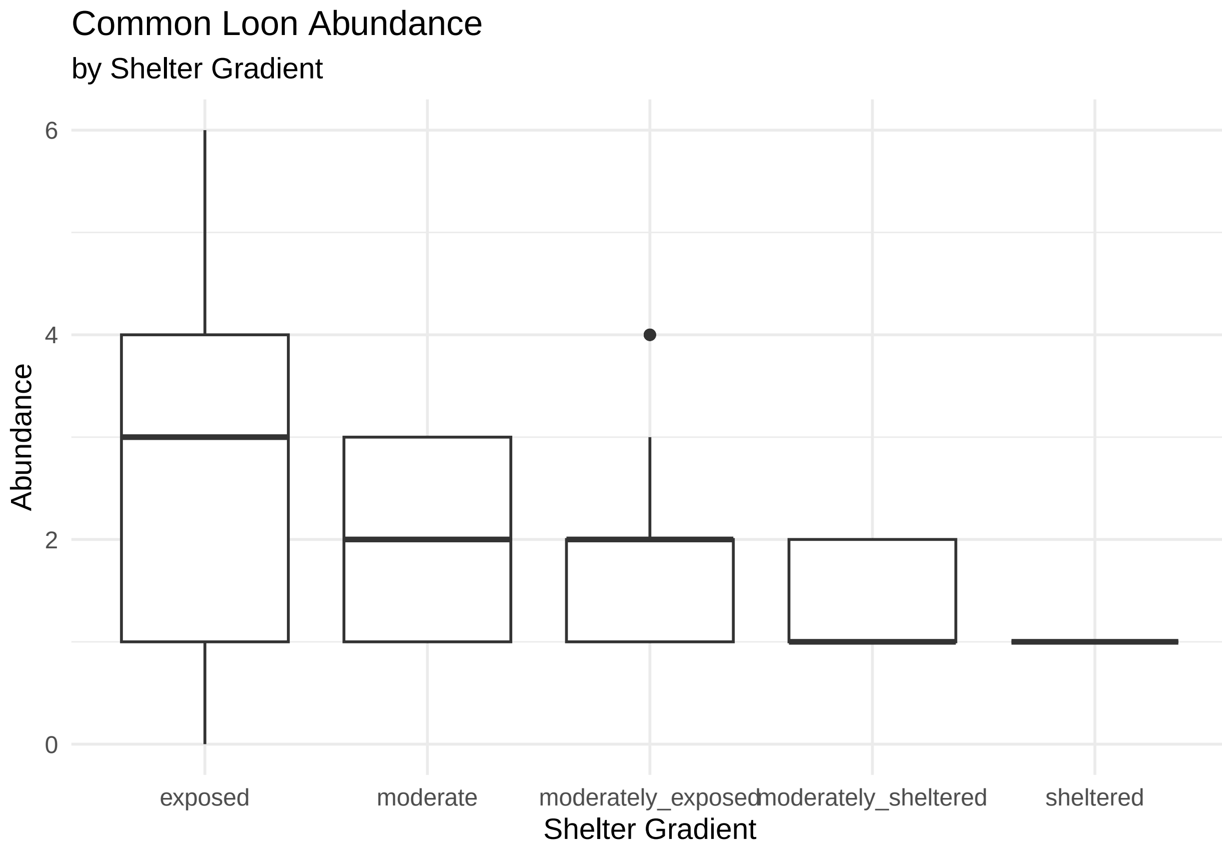
Exposure Level and Abundance in 2023

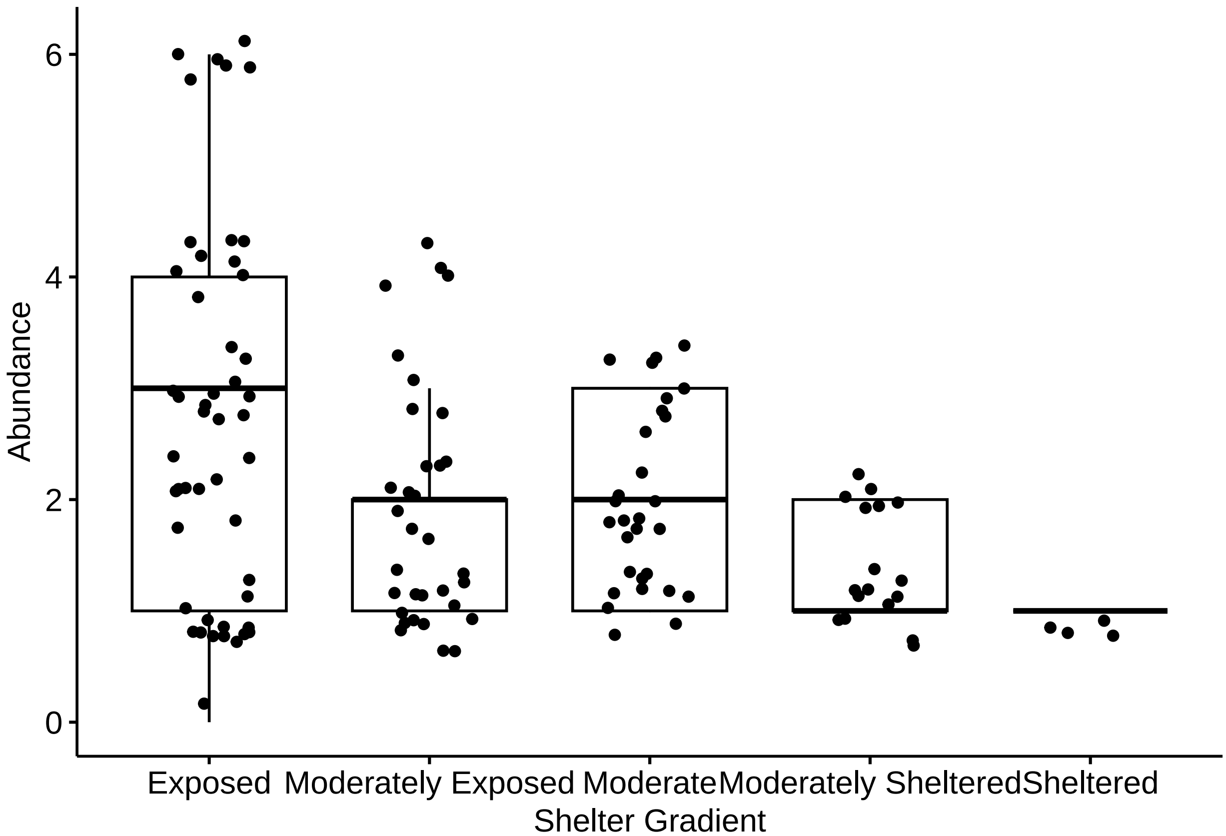
This study aimed to explore the relationship between coastal exposure level and the presence of *G. immer*. The hypothesis suggested that higher exposure levels of coastal locations would positively affect the presence of *G. immer*. Conversely, the null hypothesis posited that coastal exposure level would not significantly influence *G. immer* presence.

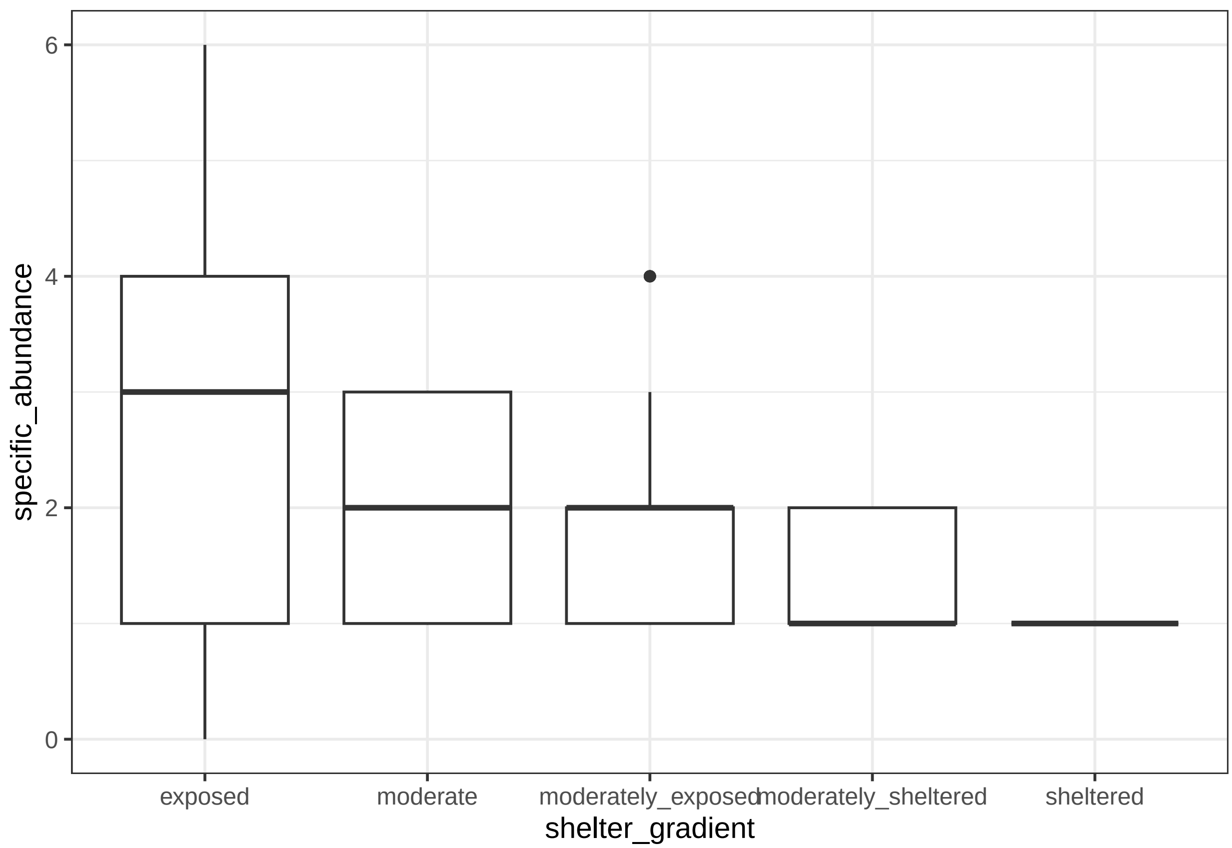
Visualizations

Below we are visualizing abundance in the form of boxplots.

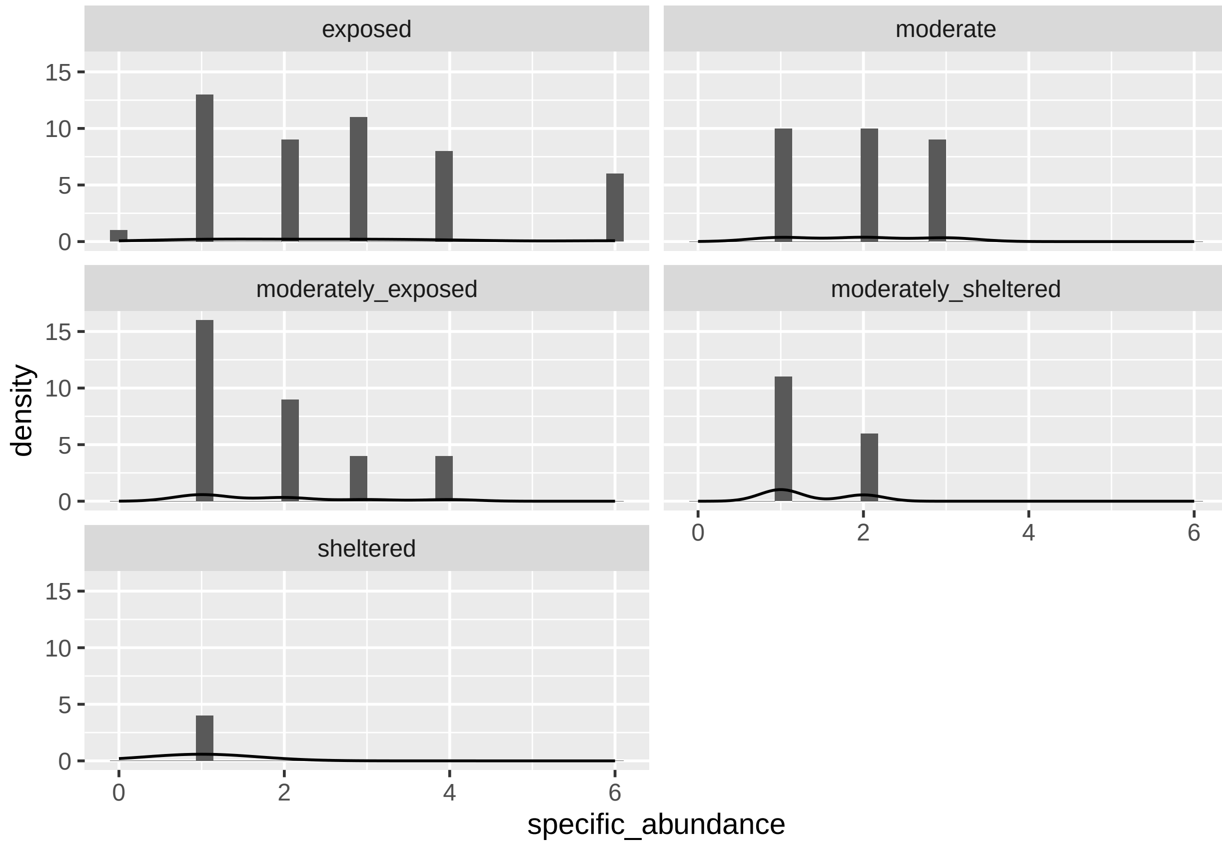
As we can see below, it appears as though there is the highest abundance at locations that are exposed and the lowest abundance at locations that are sheltered. Statistical tests will need to be performed to see if there is a significant difference in the abundances at different exposure levels.


Common Loon Abundance
by Shelter Gradient
Abundance
6
4
2
0
exposed
moderate
moderately_exposed
moderately_sheltered
sheltered
Shelter Gradient

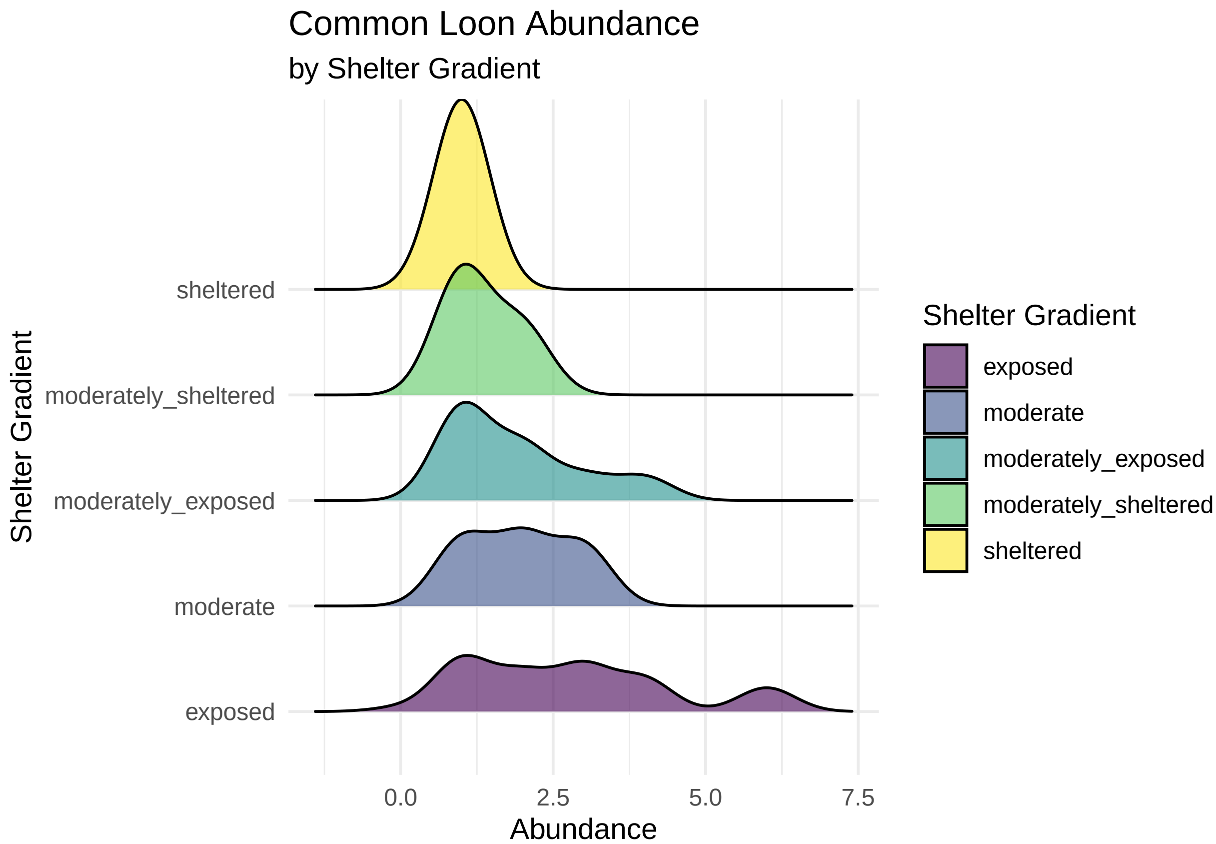

Abundance
0
2
4
6
Exposed
Moderately Exposed
Moderate
Moderately Sheltered
Sheltered
Shelter Gradient




```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
## Picking joint bandwidth of 0.467
```



Statistics

To determine if there is a significant difference in the abundance of *G. immer* in locations with varying exposure levels, where the exposure levels are categorical variable with five levels (exposed, moderately exposed, moderate, moderately sheltered, and sheltered), and the abundance is a continuous numerical variable, a one-way analysis of variance (ANOVA) will be used.

```
abundanceexposureaov = aov(specific_abundance ~ shelter_gradient, data = loons_2023)
summary(abundanceexposureaov)
```

```
##                    Df Sum Sq Mean Sq F value   Pr(>F)
## shelter_gradient   4  36.68   9.171   6.167 0.000145 ***
## Residuals        126 187.36   1.487
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
TukeyHSD(abundanceexposureaov)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = specific_abundance ~ shelter_gradient, data = loons_2023)
##
## $shelter_gradient
```

```
##                                                diff        lwr          upr
## moderate-exposed                       -0.78448276 -1.5782588  0.009293261
## moderately_exposed-exposed             -0.87121212 -1.6344100 -0.108014250
## moderately_sheltered-exposed           -1.39705882 -2.3496002 -0.444517493
## sheltered-exposed                      -1.75000000 -3.5063994  0.006399428
## moderately_exposed-moderate            -0.08672936 -0.9457668  0.772308103
## moderately_sheltered-moderate          -0.61257606 -1.6435025  0.418350393
## sheltered-moderate                     -0.96551724 -2.7656321  0.834597568
## moderately_sheltered-moderately_exposed -0.52584670 -1.5334180  0.481724627
## sheltered-moderately_exposed           -0.87878788 -2.6656298  0.908054014
## sheltered-moderately_sheltered         -0.35294118 -2.2284859  1.522603529
##                                            p adj
## moderate-exposed                       0.0543970
## moderately_exposed-exposed             0.0166513
## moderately_sheltered-exposed           0.0008059
## sheltered-exposed                      0.0513347
## moderately_exposed-moderate            0.9986485
## moderately_sheltered-moderate          0.4720480
## sheltered-moderate                     0.5744390
## moderately_sheltered-moderately_exposed 0.6002795
## sheltered-moderately_exposed           0.6535007
## sheltered-moderately_sheltered         0.9851163
```

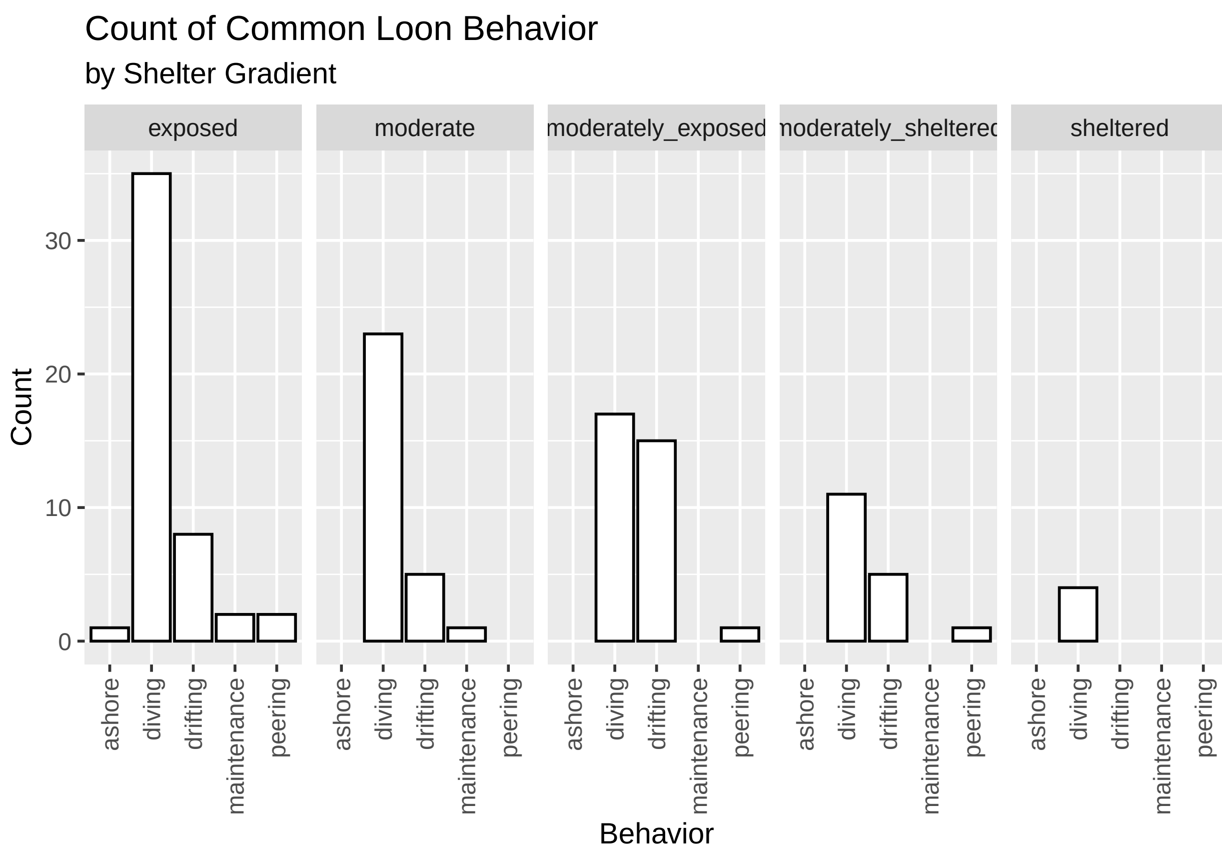
REPORT: The One-Way ANOVA revealed a significant effect of coastal exposure level on the abundance of *G. immer* (ANOVA, F_4,126 = 6.167; p = 0.000145). A Tukey HSD post-hoc comparison demonstrated that there was a significantly higher abundance of *G. immer* at exposed locations in comparison to moderately exposed locations. Additionally, there was a significantly higher abundance of *G. immer* at exposed locations in comparision to moderately sheltered locations. This indicates that coastal exposure level does indeed have a statistically significant influence on G. immer presence, allowing us to reject the null hypothesis.

Exposure Level and Behavior

This study aimed to investigate the potential influence of location exposure level on the behavior of Gavia immer (common loon). The hypothesis proposed that exposure level of the location significantly impacts G. immer behavior. In contrast, the null hypothesis stated that there would be no significant impact of location exposure level on G. immer behavior.

Visualizations

```
loons_2023 %>%
  count(behavior, shelter_gradient) %>%
  ggplot(aes(x = behavior, y = n)) +
  geom_col(fill = "white", color = "black") +
  facet_wrap(~ shelter_gradient, nrow = 1) +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1)) +
  labs(title = "Count of Common Loon Behavior",
       subtitle = "by Shelter Gradient",
       x = "Behavior",
       y = "Count")
```



Statistics

Chi-Squared Test of Independence The Chi-Squared Test of Independence examines whether there is a relationship between two categorical variables. Our null hypothesis is that exposure will not have a significant effect on *G. immer* behavior assumes that exposure level and behavior are independent.

Expected Frequencies and Contingency Table Before performing the chi-square test, we need to calculate the expected frequencies for each cell in the contingency table. These expected frequencies represent what you would expect to observe if there were no association between exposure and behavior.

```
##
##                         ashore diving drifting maintenance peering
##   exposed                    1     35        8           2       2
##   moderate                   0     23        5           1       0
##   moderately_exposed         0     17       15           0       1
##   moderately_sheltered       0     11        5           0       1
##   sheltered                  0      4        0           0       0
```

Chi-square Test Now that we have the contingency table of expected frequencies, we can perform the statistical test.

```
##
##  Pearson's Chi-squared test
```

```
##
## data:  contingency_table2
## X-squared = 16.769, df = 16, p-value = 0.4007
```

REPORT: The Chi-Squared test revealed no significant association between exposure level of the location and *G. immer* behavior (X^2(16,n=131) = 16.77, p = 0.40). This suggests that exposure level of the location does not have a statistically significant impact on *G. immer* behavior, allowing us to reject the hypothesis. This implies that other factors may be more influential in determining behavioral patterns in *G. immer*.

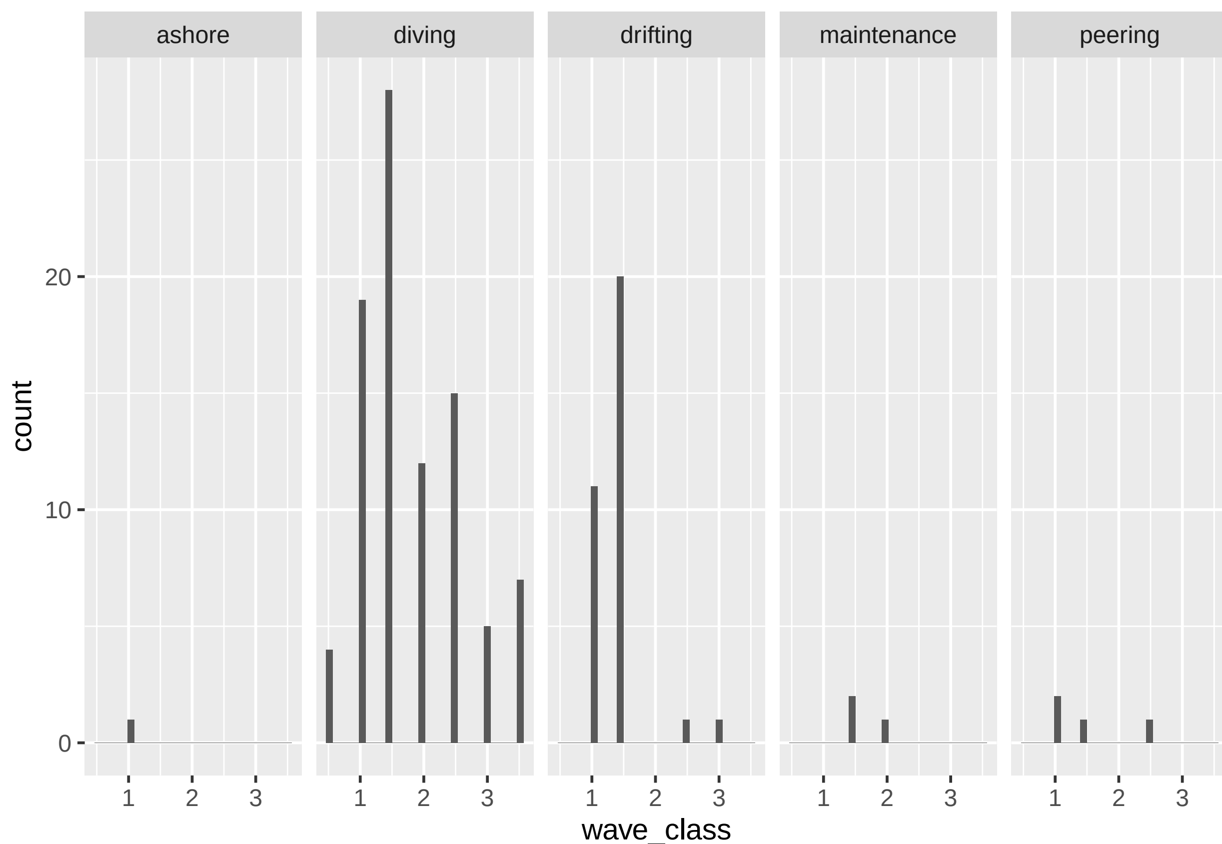
Wave Class and Behavior

This study aimed to investigate the potential influence of wave class on the behavior of *G. immer*. The hypothesis proposed that wave class significantly affects *G. immer* behavior. Conversely, the null hypothesis stated that there would be no significant effect of wave class on *G. immer* behavior.

Visualizations

```
loons_2023 %>%
 ggplot(aes(x = wave_class)) +
  geom_histogram() +
  facet_wrap(~ behavior, nrow = 1)
```

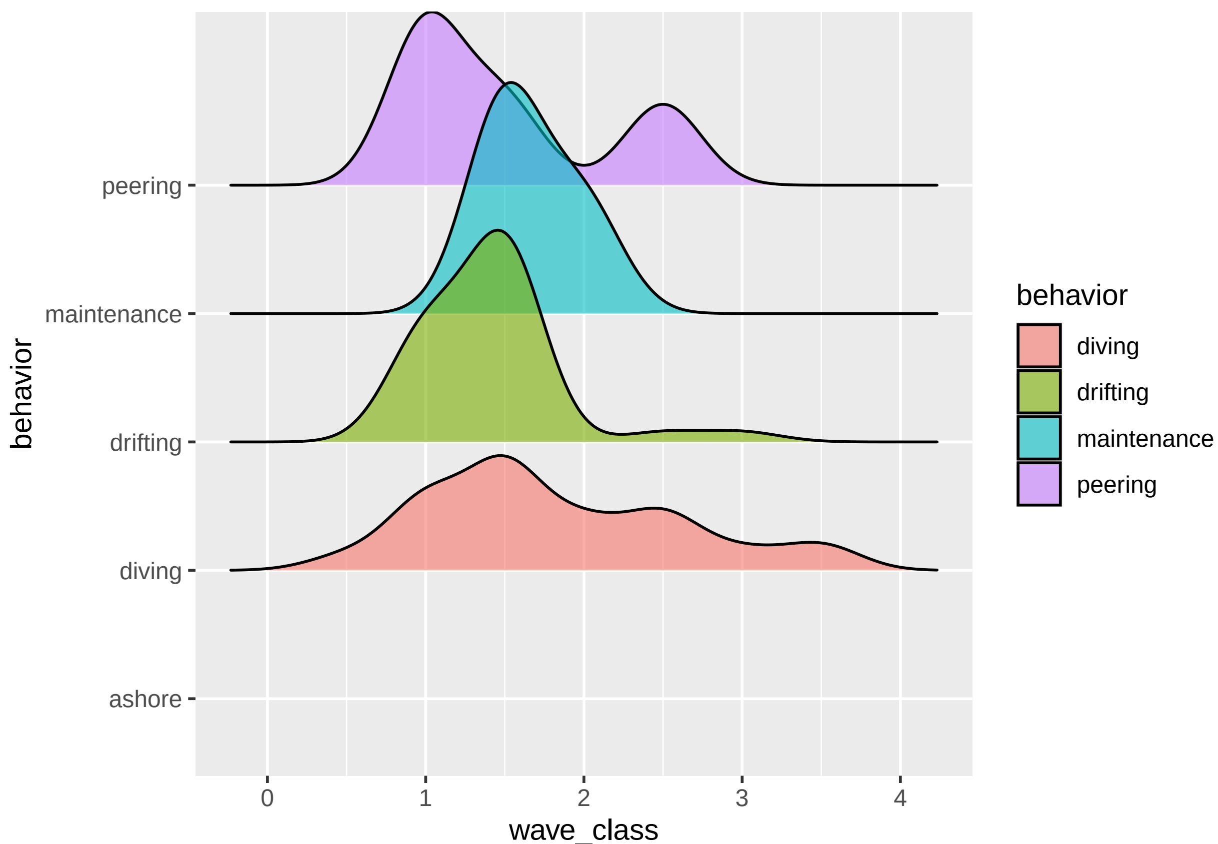
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```r
loons_2023 %>%
  ggplot(aes(x = wave_class, y = behavior, fill = behavior)) +
  geom_density_ridges(alpha = 0.6, bins = 1)
```

```
## Warning in geom_density_ridges(alpha = 0.6, bins = 1): Ignoring unknown
## parameters: `bins`
```

```
## Picking joint bandwidth of 0.244
```



Statistics

```r
waveclassaov = aov(wave_class ~ behavior, data = loons_2023)
summary(waveclassaov)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## behavior      4   4.79  1.1975   2.356 0.0572 .
## Residuals   126  64.05  0.5083
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

REPORT: Wave class did not significantly affect *G. immer* behavior (one-way ANOVA, F_4,126 = 2.356, P = 0.0572).

Statistics

Chi-Squared Test of Independence The Chi-Squared Test of Independence examines whether there is a relationship between two categorical variables.

Expected Frequencies and Contingency what is Table Before performing the chi-square test, we need to calculate the expected frequencies for each cell in the contingency table. These expected frequencies represent what you would expect to observe if there were no association between wave class and behavior.

```
##
##       ashore diving drifting maintenance peering
##   0.5      0      4        0           0       0
##   1        1     19       11           0       2
##   1.5      0     28       20           2       1
##   2        0     12        0           1       0
##   2.5      0     15        1           0       1
##   3        0      5        1           0       0
##   3.5      0      7        0           0       0
```

Chi-square Test Now that we have the contingency table of expected frequencies, we can perform the statistical test.

```
##
##  Pearson's Chi-squared test
##
## data:  contingency_table3
## X-squared = 28.64, df = 24, p-value = 0.234
```

REPORT: The Chi-Squared test revealed that wave class did not significantly affect *G. immer* behavior in our sample (Chi-squared(24, n=131) = 28.64, p = 0.23). This suggests that there is no statistically significant relationship between wave class and *G. immer* behavior, allowing us to reject the hypothesis. It is possible that other factors not accounted for in this study may have a greater influence on *G. immer* behavior.

Meters Offshore and Dive Times

This study aimed to explore the potential effect of distance from shore on the dive times of *G. immer*. The hypothesis suggested that distance from shore would significantly influence *G. immer* dive times. Conversely, the null hypothesis posited that there would be no significant effect of distance from shore on *G. immer* dive times.

Visualizations

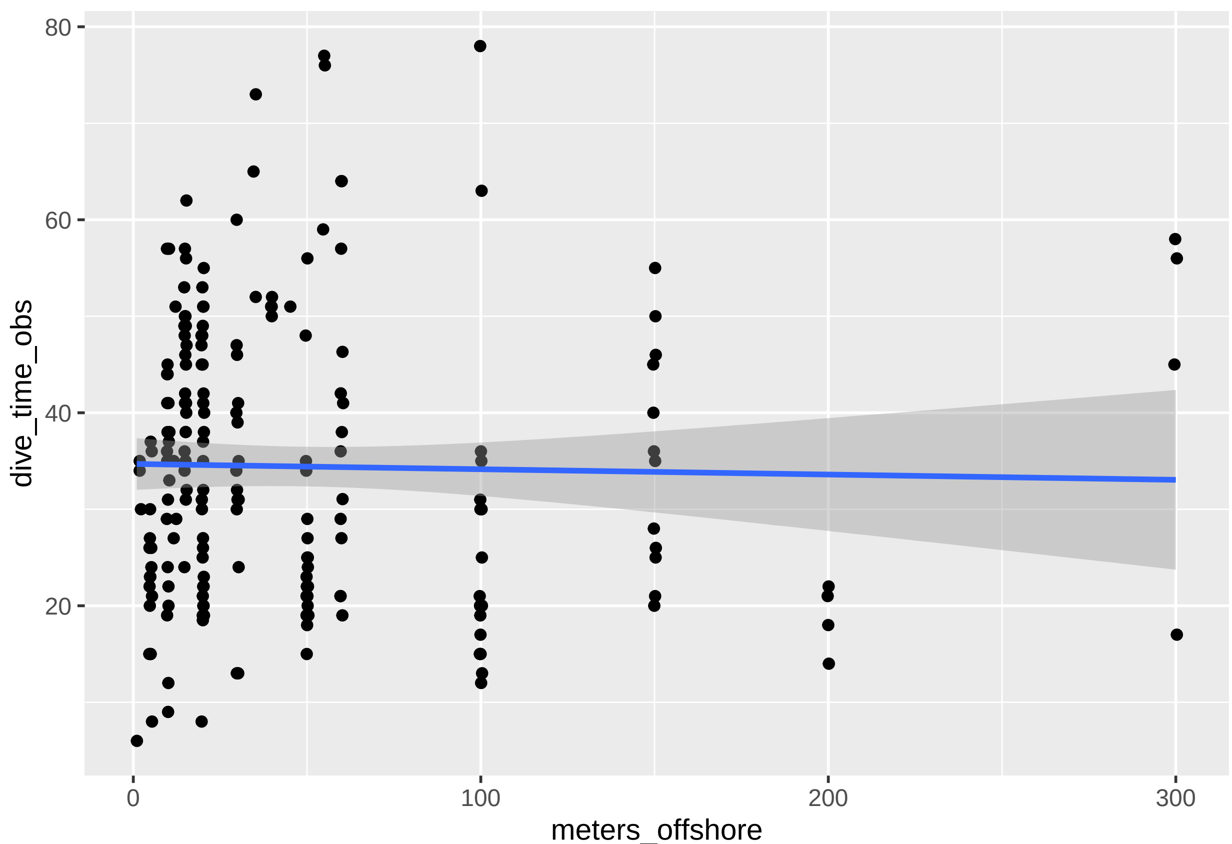
```
loons_dive_2024 <- loons_2024 %>%
  select(tide, behavior, dive_time_obs, tidal_height, meters_offshore) %>%
  filter(dive_time_obs != "n/a") %>%
  filter(dive_time_obs != "na")

loons_dive_2024 <- loons_dive_2024 %>%
```

```r
  separate_longer_delim(dive_time_obs, delim = ",") %>%
  mutate(dive_time_obs = as.numeric(dive_time_obs))

loons_dive_2024 %>%
  ggplot(aes(x = meters_offshore, y = dive_time_obs)) +
  geom_jitter() +
  geom_smooth(method = "lm")
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

```
## Warning: Removed 2 rows containing non-finite outside the scale range
## (`stat_smooth()`).
```

```
## Warning: Removed 2 rows containing missing values or values outside the scale range
## (`geom_point()`).
```

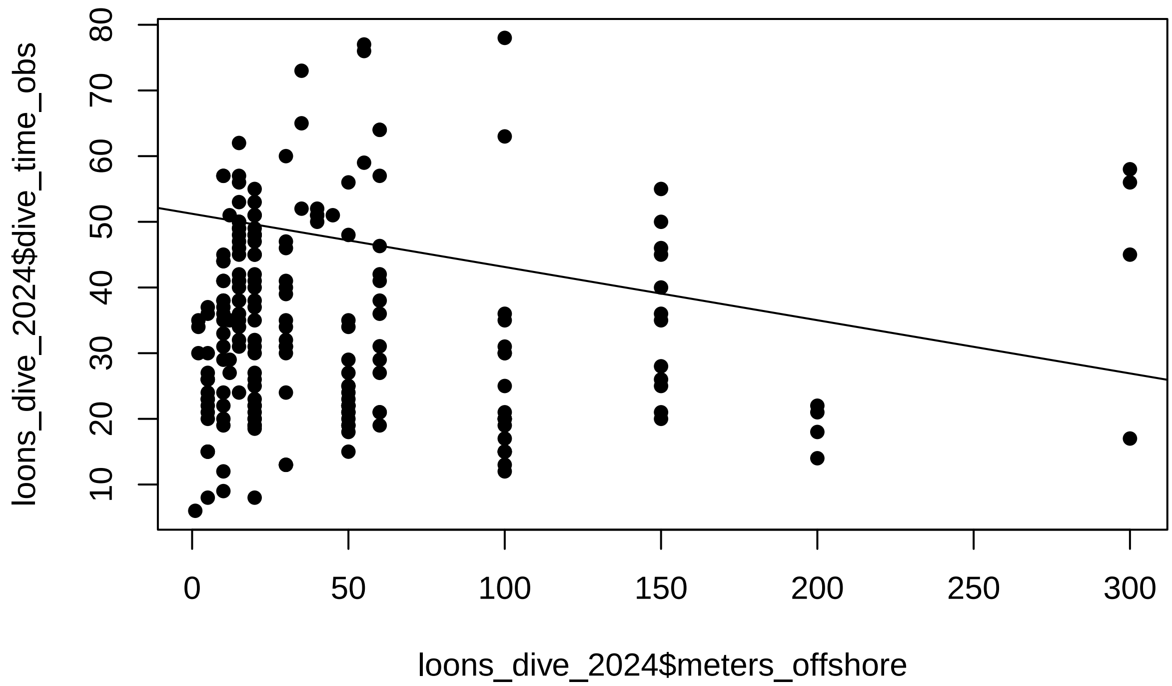



```r
#lm
lm3=lm(loons_dive_2024$meters_offshore~loons_dive_2024$dive_time_obs)
lm3
```

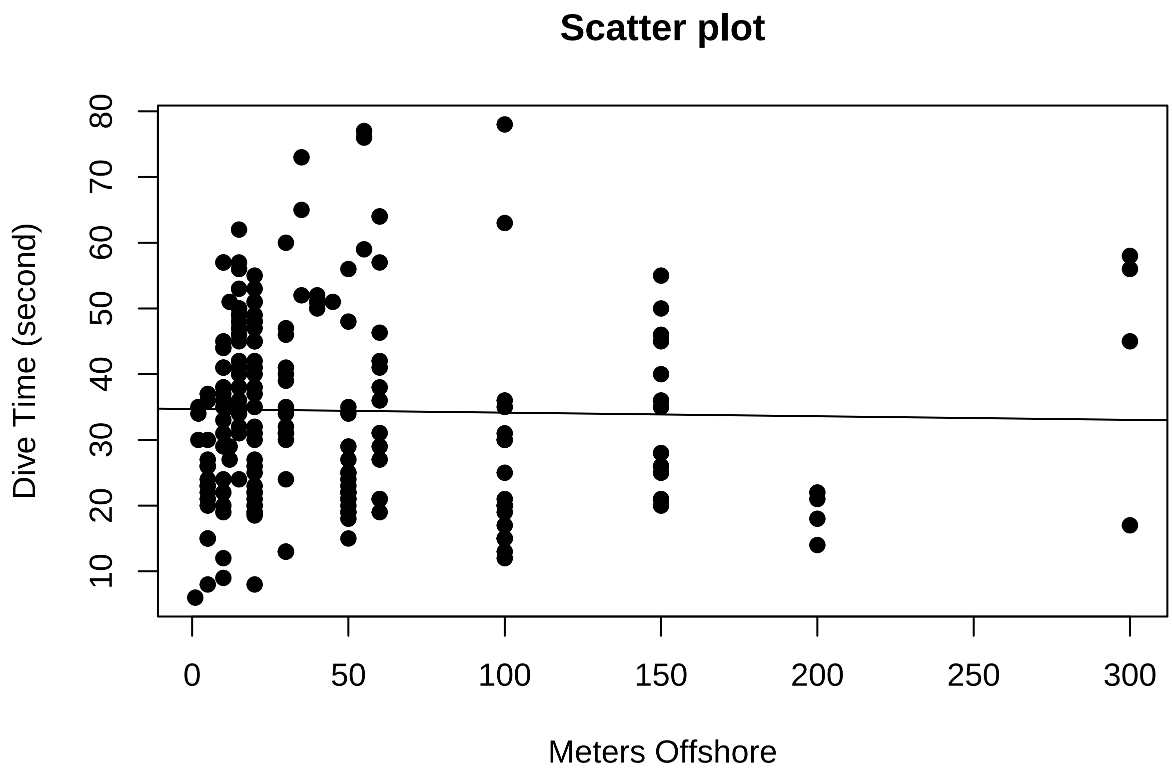
```
##
## Call:
## lm(formula = loons_dive_2024$meters_offshore ~ loons_dive_2024$dive_time_obs)
##
```

```
## Coefficients:
##                   (Intercept)  loons_dive_2024$dive_time_obs
##                      51.22103                       -0.08102
```

```
plot(loons_dive_2024$meters_offshore,loons_dive_2024$dive_time_obs,pch=16)
lm3=lm(loons_dive_2024$meters_offshore~loons_dive_2024$dive_time_obs)
abline(lm3)
```




```
#scatterplot
meters_offshore <- loons_dive_2024$meters_offshore
dive_time_obs <- loons_dive_2024$dive_time_obs

plot(meters_offshore, dive_time_obs, main="Scatter plot", xlab="Meters Offshore", ylab="Dive Time (seco
abline(lm(dive_time_obs~meters_offshore), col="black")
```



Statistics

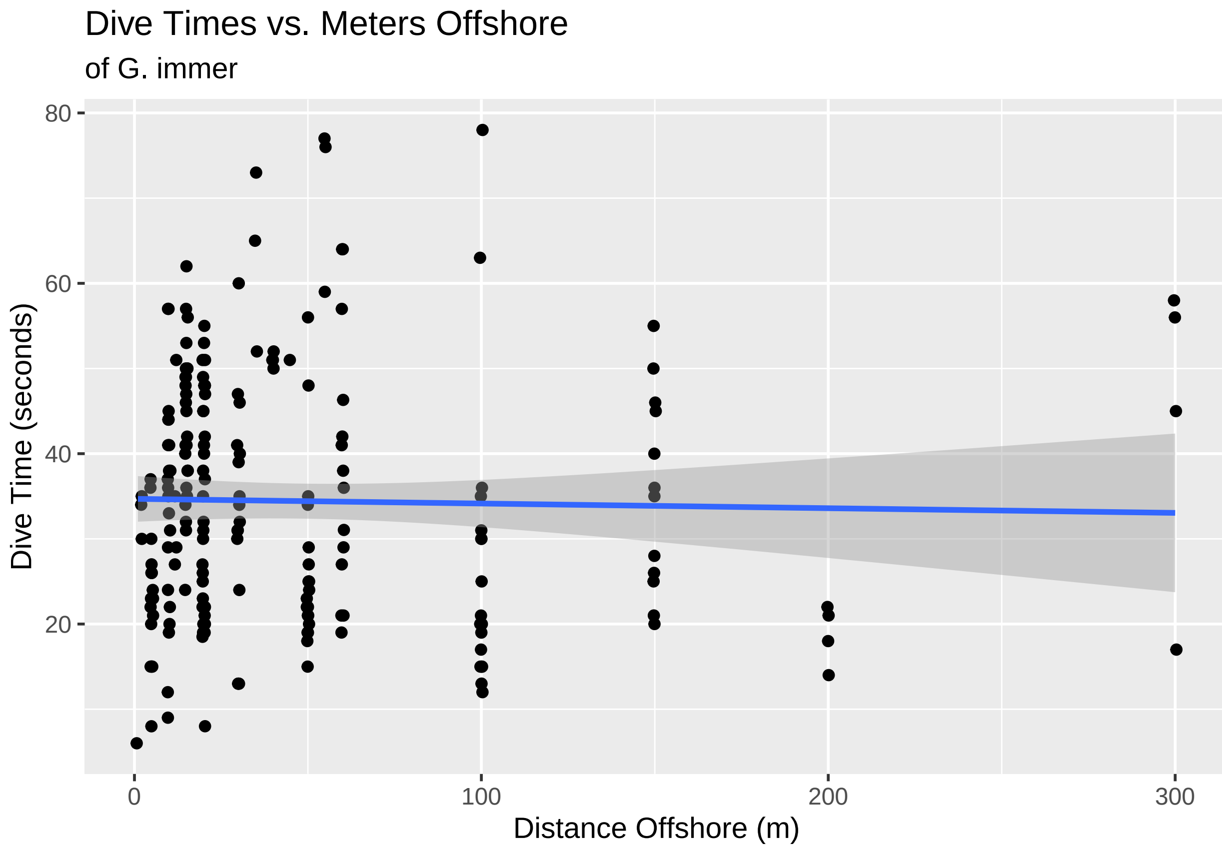
With both of our variables being continuous and numerical, we need to perform a regression test. First, we need to determine if there is a linear relationship between the variables.

```
loons_dive_2024 %>%
  ggplot(aes(x = meters_offshore, y = dive_time_obs)) +
  geom_jitter() +
  geom_smooth(method = "lm") +
  labs(title = "Dive Times vs. Meters Offshore",
       subtitle = "of G. immer",
       x = "Distance Offshore (m)",
       y = "Dive Time (seconds)")
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

```
## Warning: Removed 2 rows containing non-finite outside the scale range
## (`stat_smooth()`).
```

```
## Warning: Removed 2 rows containing missing values or values outside the scale range
## (`geom_point()`).
```

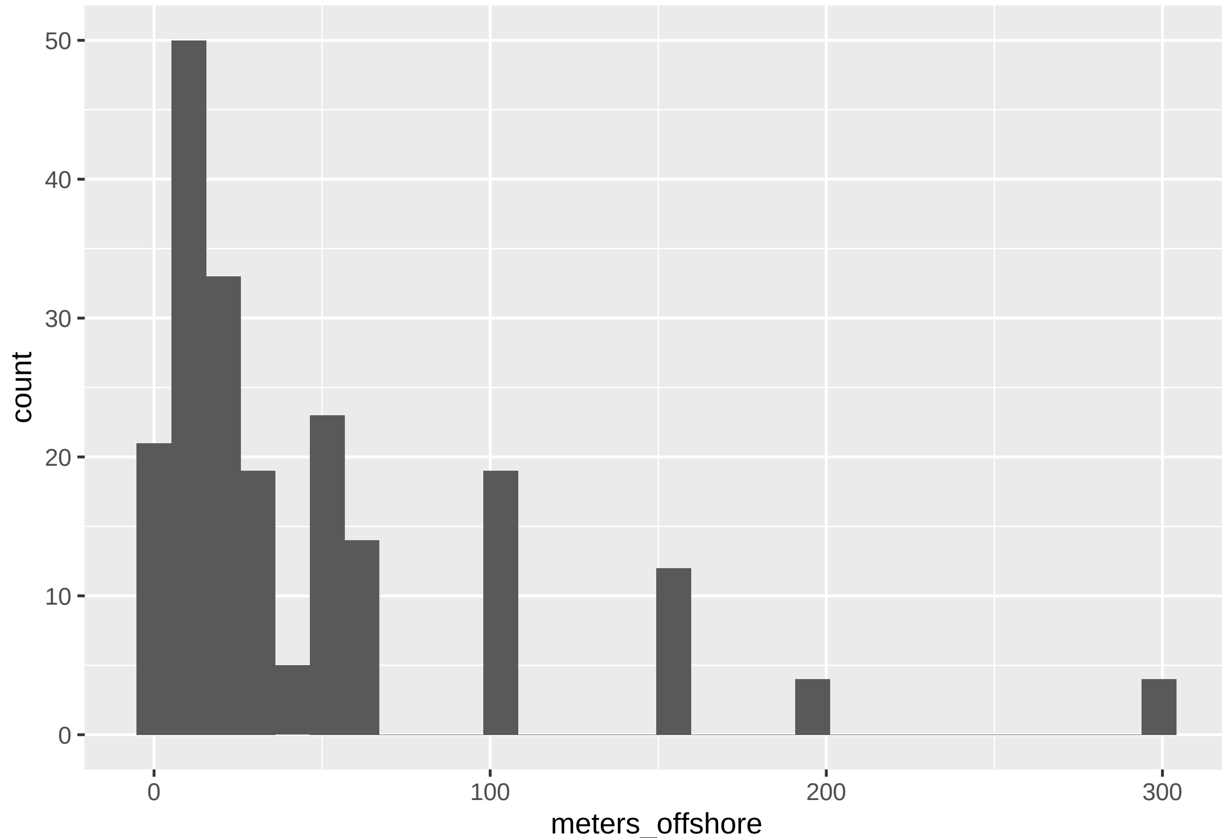


Next, we need to determine if both variables are normally distributed.

```r
#Meters Offshore
ggplot(data = loons_dive_2024, mapping = aes(x = meters_offshore)) +
  geom_histogram()
```

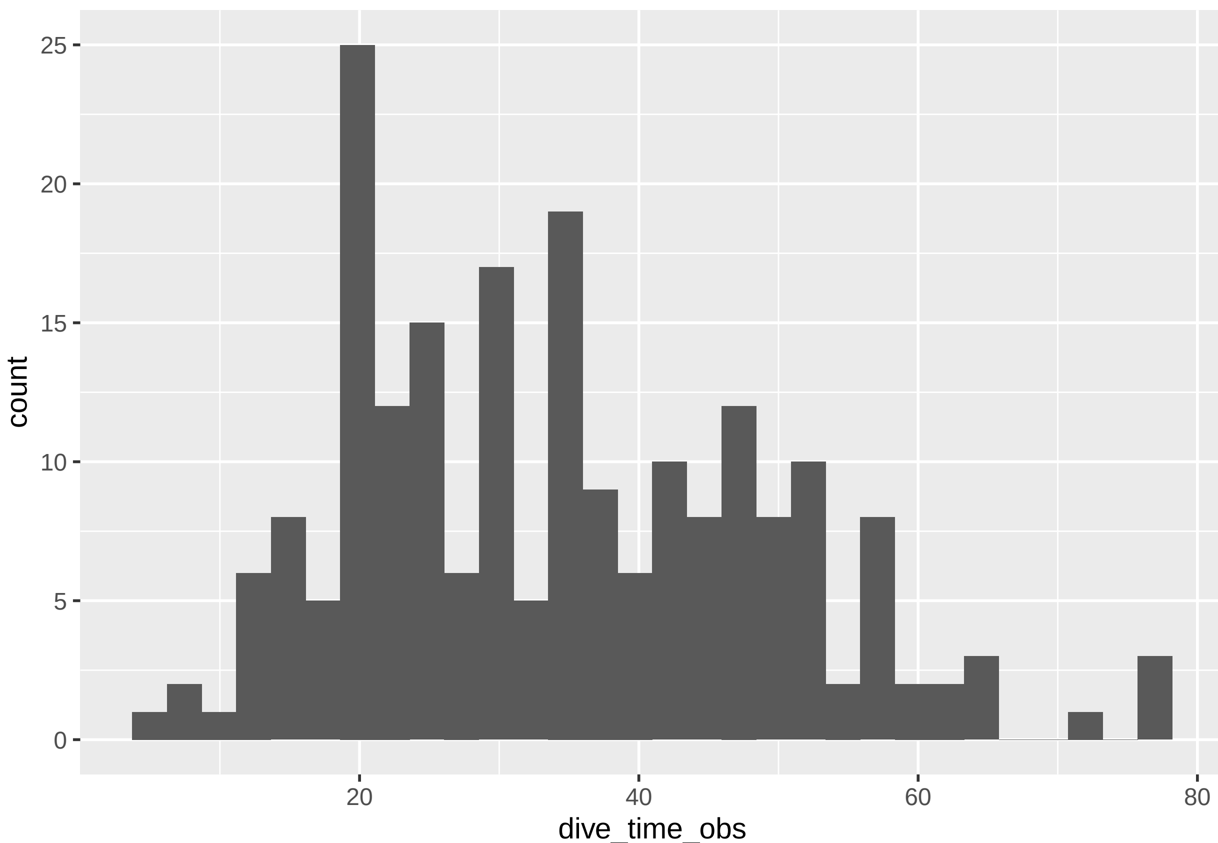
```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```

```
## Warning: Removed 2 rows containing non-finite outside the scale range
## ('stat_bin()').
```




```r
#Dive Times
ggplot(data = loons_dive_2024, mapping = aes(x = dive_time_obs)) +
  geom_histogram()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Statistics With the variable `meters_offshore` having a non-normal distribution, we will want to use a non-parametric version of a linear correlation test (i.e. Pearson's R). One such test is Spearman's Rho.

```
dive_wave_corr <- cor.test(x=loons_dive_2024$meters_offshore, y = loons_dive_2024$dive_time_obs, method

dive_wave_corr
```

```
##
##  Spearman's rank correlation rho
##
## data:  loons_dive_2024$meters_offshore and loons_dive_2024$dive_time_obs
## S = 1447337, p-value = 0.7449
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.02291793
```

REPORT: A Spearman's correlation revealed a moderate positive monotonic relationship between the variables `meters_offshore` and `dive_time_obs` (rs[93] = .50, $p < 0.005$). This indicates that there is a statistically significant association between distance from shore and G. immer dive times, allowing us to reject the null hypothesis. This implies that G. immer may exhibit longer dive times when further from the shore.

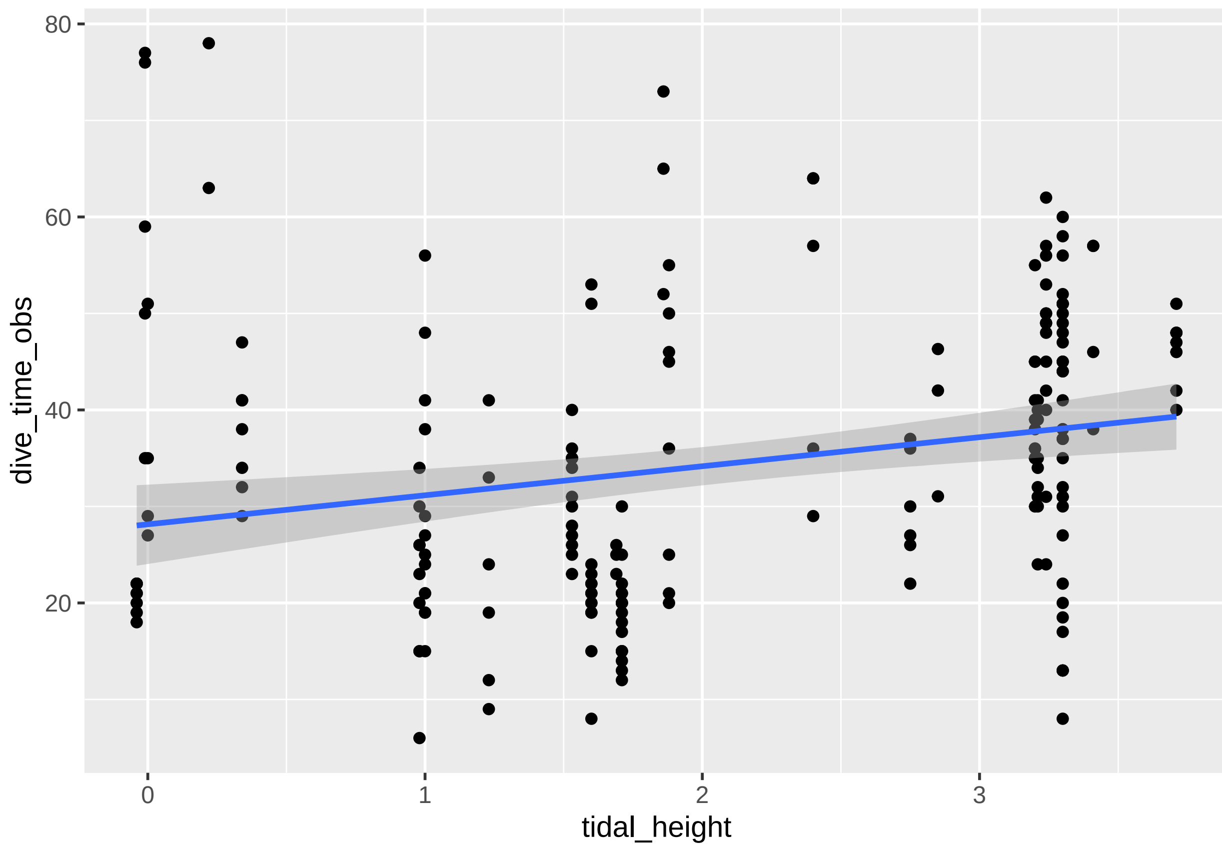
Tidal height and dive times

This study aimed to investigate the potential effect of tidal height on the dive times of *G. immer*. The hypothesis proposed that tidal height would significantly influence *G. immer* dive times. Conversely, the null hypothesis stated that there would be no significant effect of tidal height on *G. immer* dive times.

Visualizations

```
loons_dive_2024 %>%
  ggplot(aes(x = tidal_height, y = dive_time_obs)) +
  geom_point() +
  geom_smooth(method = "lm")
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

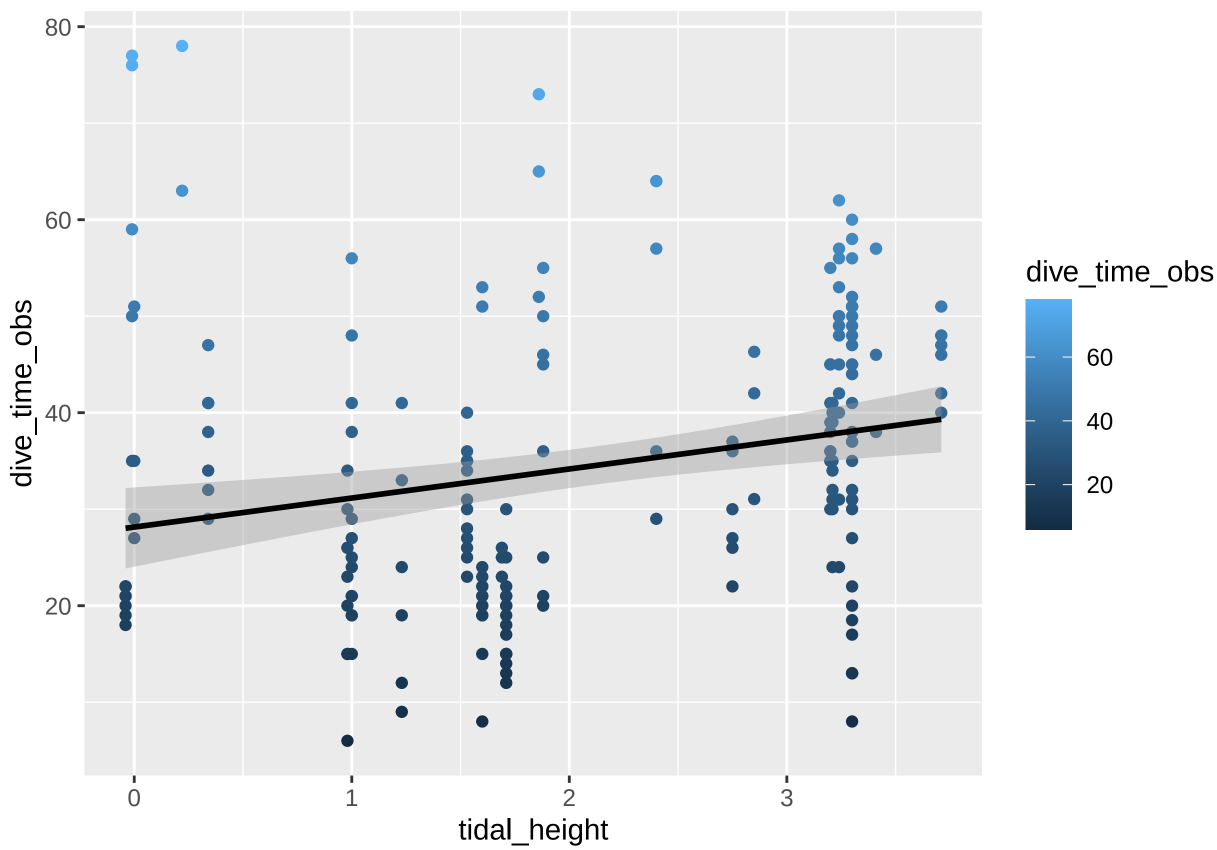


Statistics

First, we need to see if our variables have a linear relationship

```
loons_dive_2024 %>%
  ggplot(aes(x = tidal_height, y = dive_time_obs, color = dive_time_obs)) +
  geom_point() +
  geom_smooth(method = "lm", col = "black")
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



Statistics With the variable `tidal_height` having a non-normal distribution, we will want to use a non-parametric version of a linear correlation test, such as the Spearman's correlation.

```
dive_height_corr <- cor.test(x=loons_dive_2024$tidal_height, y = loons_dive_2024$dive_time_obs, method =

dive_height_corr
```

```
##
##  Spearman's rank correlation rho
##
## data:  loons_dive_2024$tidal_height and loons_dive_2024$dive_time_obs
## S = 988003, p-value = 2.386e-06
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## 0.321862
```

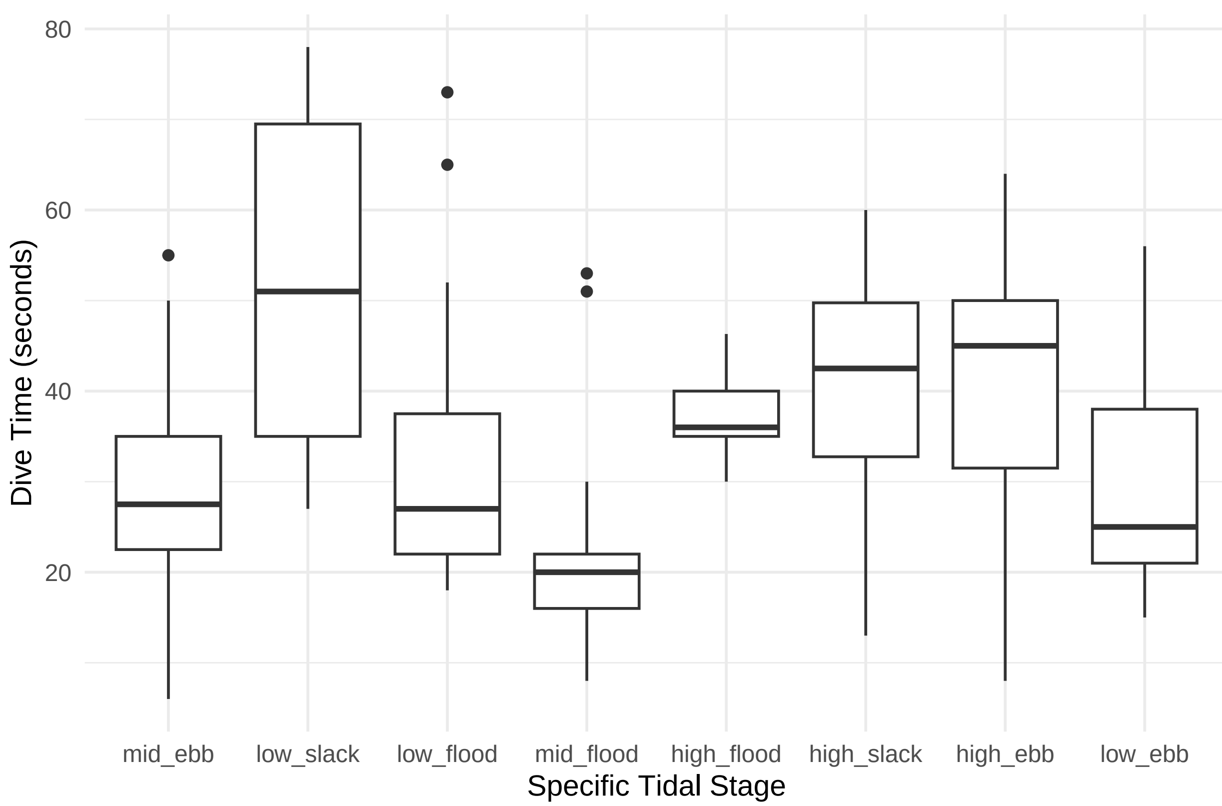
REPORT: There was a weak negative monotonic relationship between tidal height and G. immer dive times (Spearman's rho; rs[176] = 0.216, $p > 0.05$). This suggests that tidal height does have a statistically significant impact on *G. immer* dive times, while also implying that other factors may have a greater influence on dive behavior in *G. immer*.

Tidal classification and dive times

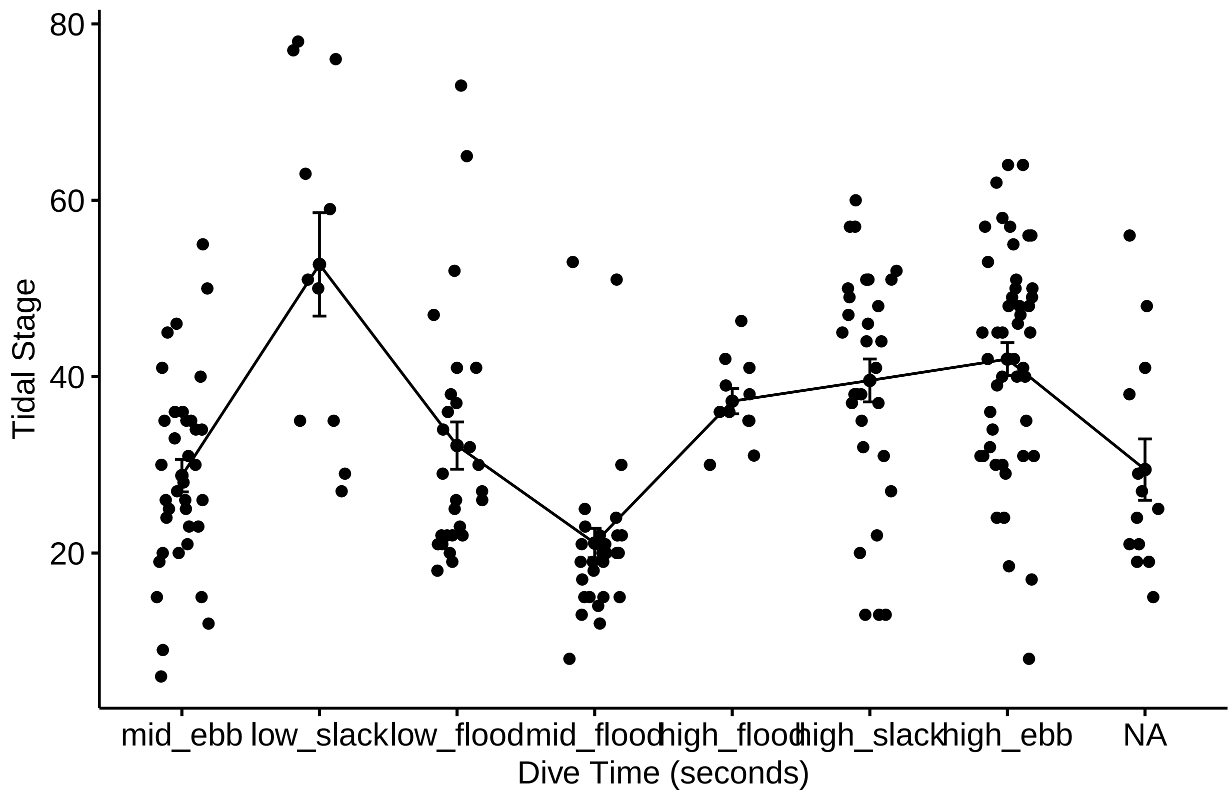
This study aimed to explore the potential effect of tidal class on the dive times of *G. immer*. The hypothesis posited that tidal class would significantly influence *G. immer* dive times. Conversely, the null hypothesis stated that there would be no significant effect of tidal class on *G. immer* dive times.

Visualizations

```
loons_dive_2024$tide <- fct_relevel(loons_dive_2024$tide, c("mid_ebb",
                                                            "low_slack",
                                                            "low_flood",
                                                            "mid_flood",
                                                            "high_flood",
                                                            "high_slack",
                                                            "high_ebb"))

loons_dive_2024 %>%
ggplot(mapping = aes(x = tide, y = dive_time_obs)) +
  geom_boxplot() +
  theme_minimal() +
  labs(title = "",
       x = "Specific Tidal Stage",
       y = "Dive Time (seconds)")
```



```r
library(ggpubr)
divetimeagainggline <- ggline(loons_dive_2024, x = "tide", y = "dive_time_obs",
       add = c(add = "mean_se", "jitter"),
       order = c("mid_ebb", "low_slack", "low_flood", "mid_flood", "high_flood", "high_slack", "high_ebb
       ylab = "Tidal Stage", xlab = "Dive Time (seconds)", title = "")
divetimeagainggline
```



Statistics

Below are the statistics tests that test against all tidal stages.

```r
full_tide_aov = aov(dive_time_obs ~ tide, data = loons_dive_2024)
summary(full_tide_aov)
```

```
##              Df Sum Sq Mean Sq F value   Pr(>F)    
## tide          7  14337  2048.1   13.35 4.55e-14 ***
## Residuals   198  30388   153.5                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
TukeyHSD(full_tide_aov)
```

```
##   Tukey multiple comparisons of means
```

```
##     95% family-wise confidence level
##
## Fit: aov(formula = dive_time_obs ~ tide, data = loons_dive_2024)
##
## $tide
##                               diff        lwr         upr     p adj
## low_slack-mid_ebb      23.9494949  10.873934  37.0250564 0.0000019
## low_flood-mid_ebb       3.4074074  -6.255240  13.0700552 0.9602388
## mid_flood-mid_ebb      -7.6487455 -16.948346   1.6508554 0.1932755
## high_flood-mid_ebb      8.4376768  -4.637885  21.5132382 0.5000738
## high_slack-mid_ebb     10.7888889   1.406378  20.1714002 0.0121891
## high_ebb-mid_ebb       13.2115839   4.805420  21.6177483 0.0000786
## low_ebb-mid_ebb         0.6837607 -11.597250  12.9647717 0.9999998
## low_flood-low_slack   -20.5420875 -34.118111  -6.9660640 0.0001719
## mid_flood-low_slack   -31.5982405 -44.918309 -18.2781718 0.0000000
## high_flood-low_slack  -15.5118182 -31.695512   0.6718760 0.0708862
## high_slack-low_slack  -13.1606061 -26.538692   0.2174794 0.0574536
## high_ebb-low_slack    -10.7379110 -23.450316   1.9744944 0.1666312
## low_ebb-low_slack     -23.2657343 -38.814525  -7.7169438 0.0002143
## mid_flood-low_flood   -11.0561529 -21.047183  -1.0651226 0.0187539
## high_flood-low_flood    5.0302694  -8.545754  18.6062929 0.9482767
## high_slack-low_flood    7.3814815  -2.686767  17.4497300 0.3291577
## high_ebb-low_flood      9.8041765   0.638932  18.9694211 0.0266837
## low_ebb-low_flood      -2.7236467 -15.536193  10.0888995 0.9980503
## high_flood-mid_flood   16.0864223   2.766354  29.4064910 0.0066700
## high_slack-mid_flood   18.4376344   8.717272  28.1579967 0.0000007
## high_ebb-mid_flood     20.8603294  12.078671  29.6419877 0.0000000
## low_ebb-mid_flood       8.3325062  -4.208513  20.8735251 0.4607501
## high_slack-high_flood   2.3512121 -11.026873  15.7292976 0.9994319
## high_ebb-high_flood     4.7739072  -7.938498  17.4863125 0.9445258
## low_ebb-high_flood     -7.7539161 -23.302707   7.7948744 0.7915954
## high_ebb-high_slack     2.4226950  -6.446717  11.2921067 0.9907647
## low_ebb-high_slack    -10.1051282 -22.707751   2.4974944 0.2206316
## low_ebb-high_ebb      -12.5278232 -24.421442  -0.6342041 0.0310709
```

REPORT: The ANOVA test demonstrated that tidal class has a significant effect on the length of G. immer dive times (ANOVA, F_6,86 = 4.599; p = 0.000427). An unplanned analytical comparison demonstrated that *G. immer* dive times at low slack tide were significantly different than those at mid ebb tide (TukeyB, $p < 0.005$), and dive lengths at high ebb tide were significantly different than those at mid ebb tide (TukeyB, $p < 0.05$). Otherwise, there were no significant differences in dive times between other tidal classes (TukeyB, $p > 0.05$). These findings allow us to reject the null hypothesis.

Below are the statistics tests that test against all tidal stages.

```r
loons_dive_2024_2 <- loons_dive_2024 %>%
  filter(tide == "low_slack"|tide == "high_flood"|tide == "low_flood") %>%
  mutate(flood = if_else(tide == "high_flood"|tide == "low_flood", "flood", "low_slack"))

loons_dive_2024_2
```

```
## # A tibble: 49 x 6
##    tide      behavior dive_time_obs tidal_height meters_offshore flood
##    <fct>     <chr>            <dbl>        <dbl>           <dbl> <chr>
##  1 low_flood diving              23         1.69              20 flood
```

```
##  2 low_flood diving              25         1.69               20 flood
##  3 low_flood diving              26         1.69               20 flood
##  4 low_flood diving              52         1.86               35 flood
##  5 low_flood diving              65         1.86               35 flood
##  6 low_flood diving              73         1.86               35 flood
##  7 low_slack diving              51         0                  12 low_slack
##  8 low_slack diving              35         0                  12 low_slack
##  9 low_slack diving              29         0                  12 low_slack
## 10 low_slack diving              27         0                  12 low_slack
## # i 39 more rows
```

```r
t.test(dive_time_obs ~ flood, data = loons_dive_2024_2)
```

```
##
##  Welch Two Sample t-test
##
## data:  dive_time_obs by flood
## t = -3.087, df = 12.343, p-value = 0.009139
## alternative hypothesis: true difference in means between group flood and group low_slack is not equal
## 95 percent confidence interval:
##  -32.515378  -5.656536
## sample estimates:
##     mean in group flood mean in group low_slack
##                33.64132                52.72727
```

REPORT: A Welch's Two Sample t-test demonstrated that there was a significant difference in dive times during low-slack periods and flooding periods (df = 12.743, p = 0.04036).

Discussion / Conclusion

This study provides insights into the winter ecology and behavioral dynamics of the Common Loon (Gavia immer) along the coastline of Mount Desert Island, Maine. While much of the existing research has focused on the breeding season ecology of this species, this study sheds light on their non-breeding season behaviors, particularly in coastal wintering habitats.

Our findings reveal relationships between environmental factors and loon behavior, highlighting the multifaceted nature of ecological interactions shaping the dynamics of G. immer populations. Furthermore, our study highlights the significant role of habitat characteristics, particularly coastal exposure level, in shaping the spatial distribution and abundance of loon populations along the coastline. The observed preference of loons for exposed coastal environments underscores the importance of habitat structure and resource availability in mediating population dynamics and ecological interactions.

Further studies are warranted to determine more comprehensive ecological dynamics of loon populations in diverse coastal environments to ensure the long-term conservation and sustainability of this species.

Citations

Bent, A. C. 2009. Life Histories of North American Diving Birds. Cornell University Press: Ithaca, New York.

Daub, B. C. 1989. Behavior of Common Loons in Winter. Journal of Field Ornithology 60: 305-311.

Ford, T. B. and Gieg, J. A. 1995. Winter Behavior of the Common Loon. Journal of Field Ornithology 66: 22-29.

McIntyre, J. W. 1978. Wintering Behavior of Common Loons. The Auk 95: 396-403.

Holm, K. and Burger, A. 2002. Foraging Behavior and Resource Partitioning by Diving Birds During Winter in Areas of Strong Tidal Currents. Waterbirds 25(3): 312-325.

Thompson, Stephanie A. & Price, J. J. (2006). Water Clarity and Diving Behavior in Wintering Common Loons. Waterbirds: The International Journal of Waterbird Biology, 29(2), 169–175.

Randomly Sampling Data

Now, we can see that we have a normal distribution of dive times during all tidal cycles, which now makes an ANOVA an appropriate statistical test to use for this data frame.

To determine if there is a significant difference in dive times of Common loons during different tidal stages, where the tidal stages are categorical with three levels (high tide, mid tide, and low tide), and the dive time is a continuous numerical variable, a one-way analysis of variance (ANOVA) will be used.