kittiwake populations

An in-depth study of kittiwake populations: insights combining observational, historical, measurement, and location data.

Introduction

This analysis explores the intricate world of kittiwakes, with particular attention on two notable subspecies: the red-legged and black-legged kittiwakes.

Four main datasets—**Observational, Historical, Measurement, and Location**—have all been thoroughly analyzed to provide the foundation of our research.

The Observational data offers insights on the daily behaviors of kittiwakes at dawn, noon, mid-afternoon, and nightfall, as well as a window into the daily sighting patterns of these birds at a chosen observation point over several weeks.

The number of breeding pairs counted at four distinct locations over a six-year period, which makes up the historical data set, provides a longitudinal view of kittiwake breeding habits.

The weight, wingspan, and culmen length of individual birds from both subspecies give us important information about the physical distinctions and similarities between black-legged and red-legged kittiwakes.This comparative analysis helps identify traits exclusive to subspecies.

Sandeel concentration, cliff height, mean summer temperature, and coastal direction are among the geographical and environmental characteristics included in this dataset, which examines breeding pairs in 29 colonies. Analyzing these data facilitates an understanding of the ecological factors driving kittiwake breeding and habitat choices.

Observation Data

Exploratory Data Analysis

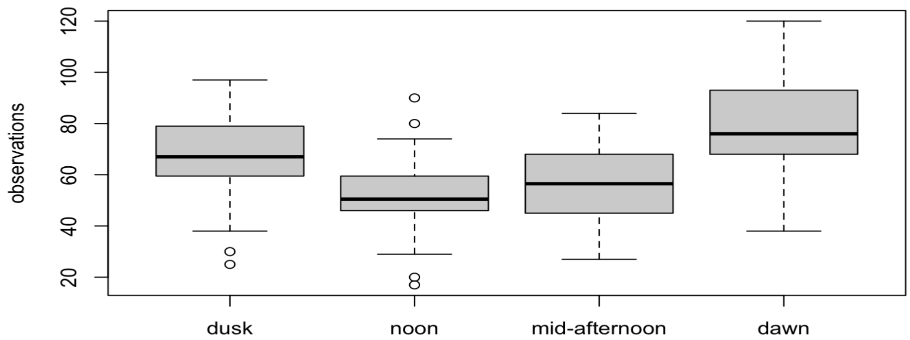
Time of Day	Mean	Median	Mode	SD	Variance	Min	Max
Dusk	78.10714	76.0	74	19.79401	391.8029	38	120
Noon	52.50000	50.5	48	15.95480	254.5556	17.00	90.00
Mid-afternoon	55.03571	56.5	71	14.96782	224.0357	27.00	84.00
Dawn	67.07143	67.0	62	17.45031	304.5132	25	97.00

Mean and Median: The average number of sightings is 52.50 at noon and reaches its peak at dusk (78.11). This implies that dusk is when kittiwakes are most commonly seen. The above observation is supported by the median number of sightings, which is similar to the mean and peaks at dusk (76.0) and lows at noon (50.5).

Mode: The most frequent number of sightings fluctuates over time, representing the average number of birds seen at each point in time.It is again highest at dusk (74) and lowest at Noon(48)

Standard Deviation (SD) and Variance: The data indicates that there is greater variability in the number of sightings during dusk (SD = 19.79, Variance = 391.80) when compared to other times of the day. At other times of the day, the variance and standard deviation are smaller, suggesting a higher degree of consistency in the number of sightings.

Minimum and Maximum: These values indicate the range of sightings. For instance, sightings at dusk vary from 38 at minimum to 120 at a maximum, a wider range than at other times, suggesting that dusk may have unusual patterns of bird presence.
The range of sightings is smaller at noon and mid-afternoon, indicating more steady numbers during these times.

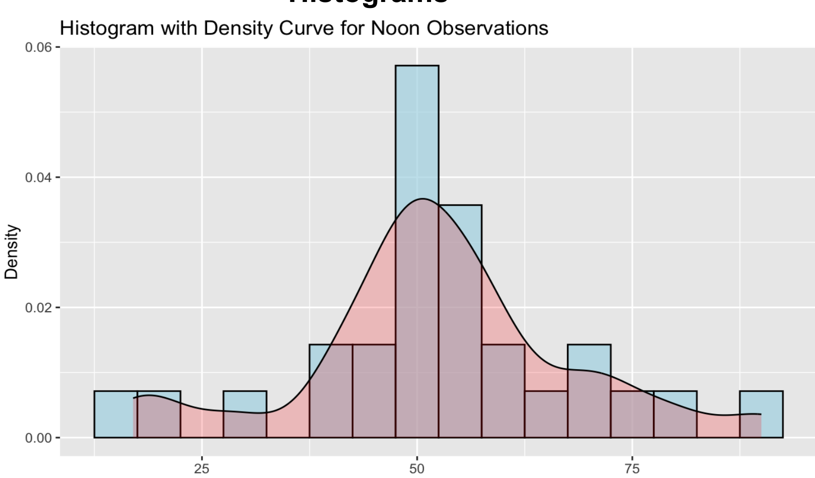


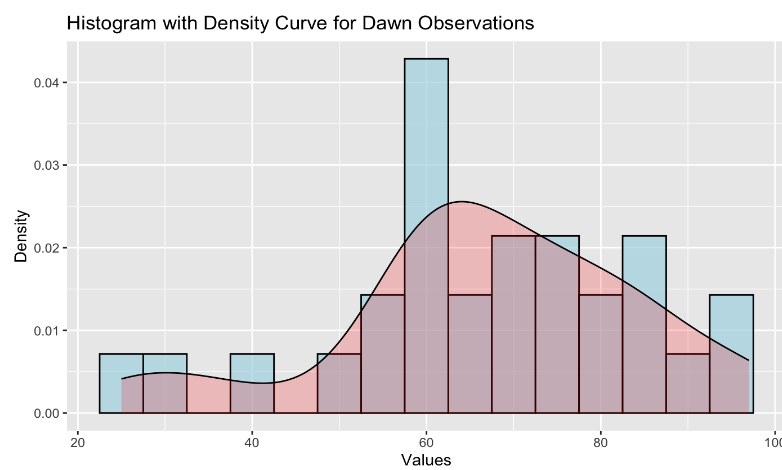
To investigate the distribution of kittiwake observations throughout the day, box plots were made.

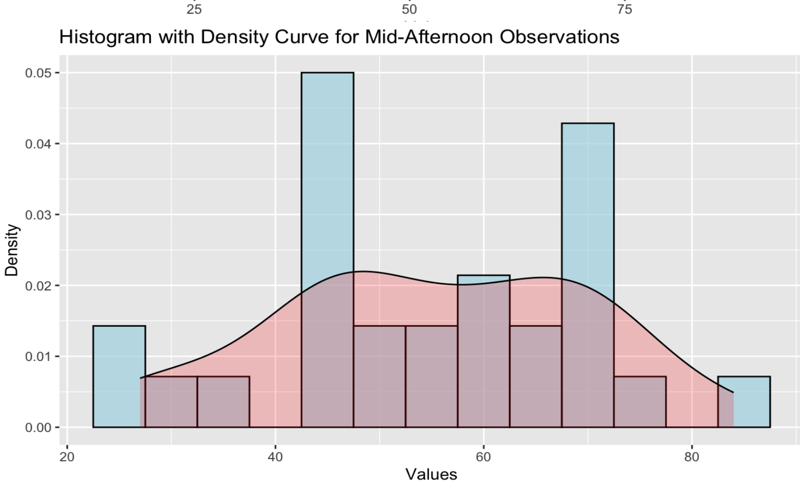
Notably, at dawn, there are two outliers that fall below the first quartile. These outliers show days on which the number of sightings was much fewer than the average range for this period of time.

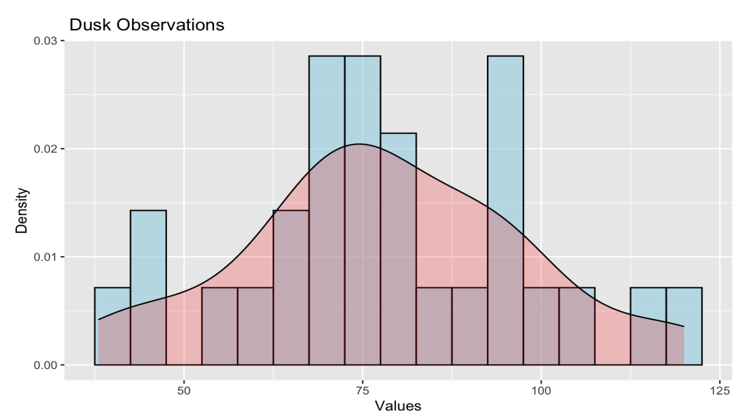
The noon data contains four outliers: two that fall below the first quartile and two that fall above the third. This represents sporadic days with noticeably fewer or noticeably more sightings than typical.

Histograms









Noon and Dusk: Data show normal distribution, suitable for standard statistical analysis.
Dawn: Right-skewed distribution, suggesting the need for non-standard analytical methods.
Mid-Afternoon: Flat and bimodal distribution.

2) Confidence Interval

For the average number of kittiwake sightings at noon, we will derive a 90% confidence interval. This will provide us a range that, within our confidence level, likely contains the true mean.

2.1) Results

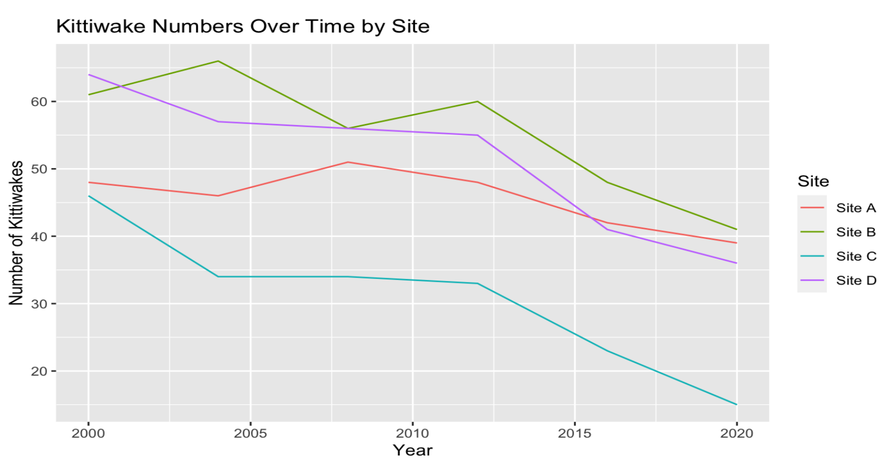
- Mean Noon Observations: 52.5
- 90 % Confidence Interval for Mean Noon Observations: [47.36 , 57.64]

2.2) Conclusion

- At noon, there were 52.5 sightings of kittiwakes on average, with a 15.95 standard deviation. Given the sample variability, the 90% confidence interval for this mean, which spans 47.36 to 57.64, shows where the true mean most likely lies.

Historical Data

As part of the study of the historical data on kittiwake populations, we first look at the trends of kittiwake numbers over time, broken down by site.



1) **Methods**

1.1) Analysis of Independence of Population Decline from Site:

Approach- We used a Paired T-test with bonferroni correction at a 95% confidence level to test these hypotheses.

Null Hypothesis (H0): the decline in kittiwake populations is not affected by the interaction of time and site.

Alternative Hypothesis (H1): The decline in kittiwake populations is significantly influenced by the interaction of time and site.

Formula: t=∑d√n(∑d2)−(∑d)2n−1

1.2) Breeding Pair Count at Site A in 2014:

Approach - We perform linear interpolation to estimate the number of breeding pairs at Site A in 2014, taking into account the data trends that had been observed up until that point.

Formula: $y0 + (y1 - y0/(x1 - x0)) * (x - x0)$

2) **Results**

Analysis of Independence of Population Decline from Site (P-Values):

Site A vs Site B: 0.31
Site A vs Site C: 0.102
Site A vs Site D: 1.00
Site B vs Site C: 0.010
Site B vs Site D: 1.00
Site C vs Site D: 0.043

Breeding Pair Count at Site A in 2014:

Estimated Breeding Pairs at Site A in 2014: 45

3) **Conclusion**

- Historical data analysis reveals significant changes in the kittiwake population decline over time at various sites. We reject the null hypothesis in light of the substantial interaction between time and site. This implies that factors specific to a particular site are possibly involved in the decline in kittiwake populations.
- The Estimated Breeding Pairs at Site A in 2014 using linear interpolation between the year 2012 and 2016 is 45.

Measurement Data

Exploratory Data Analysis

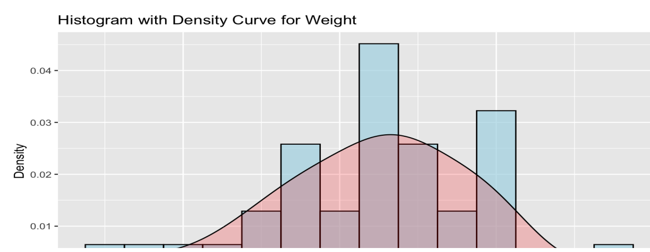
Measurement	Mean	Median	Mode	Min	Max	SD	Variance
Weight	343.58065	345	347	308	377	14.777853	218.38495
Wingspan	103.41935	105	105	87	116	8.357728	69.85161
Culmen	32.25806	32	31	24	39	13.33118	3.651189

Weight: The majority of birds (median) weigh 345g, while the average (mean) weighs roughly 343.58g. There is a lot of variation in the weight, ranging from 308g to 377g (variance: 218.38, standard deviation: 14.78).

Wingspan: The median wingspan is approximately 105 cm, with an average of 103.42 cm. With moderate variability (variance: 69.85, standard deviation: 8.36), the wingspan ranges from 87 to 116 cm.

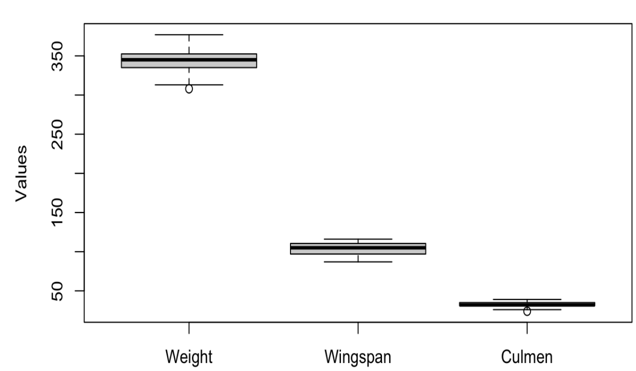
Culmen: The median length of culmen is 32 mm, and the average length is approximately 32.26 mm. Less variability is indicated by the range, which is 24 to 39 mm (variance: 13.33, standard deviation: 3.65).

Histogram



The histogram shows that
Weight follows normal distribution

BoxPlot



Culmen: Below the first quartile, there is an outlier, indicating that at least one culmen length observation is significantly shorter than the majority of the data.

Weight: Below the first quartile, there is also an outlier,suggesting that at least one bird in the dataset weighs substantially less than the majority of the other birds.

1) Methods

1.1) **Are wing span and culmen length independent?**

Approach - We use a correlation test within each subspecies to test this theory. We state A null hypothesis:

(H0): Within each subspecies, there is no correlation between culmen length and wingspan.
(H1): Within each subspecies, there is a correlation between culmen length and wingspan.

Formula: $\rho\ (x, y)\ =\ cov(x, y)\ /\ \sigma x\ *\ \sigma y$

1.2) **Variations in Weights Among Subspecies**

Approach - We conduct a two sample T-test because the weight has a normal distribution and the subspecies variances are comparable.

H0: The weights of the two subspecies of kittiwakes do not differ significantly.
H1: The weights of the two subspecies of kittiwakes differ significantly.

Formula: $(x1 - x2)\ /\ sp(\sqrt{}1/n1\ +\ 1/n2)$

1.3) **The Two Sub-Species Overall Differences**

Approach- The differences between the two subspecies were carefully evaluated using a two sample T-test with bonferroni correction across a variety of physical characteristics.

H0: The two subspecies measured traits do not significantly differ from one another.
H1: The two sub-species measured traits differ significantly from one another.

We use a 95% confidence interval to test these methods.

2) **Results**

2.1) **Correlation Analysis between culmen length and wingspan**

a) Black-legged Kittiwakes:
   - Correlation Coefficient- 0.2826 , P-value: 0.3074

b) Red-legged Kittiwakes:
   - Correlation coefficient- 0.5011, P-value: 0.04798

2.2) **Comparison of Kittiwake Sub-Species Average Weights**

T-Value: - 0.81522
P-Value: 0.4216

2.3) **Two sample T test Results for Comparing Physical Characteristics between Subspecies**

Weight P-Value: 1
Wingspan P-Value: 8.176454e-06
Culmen P-Value: 1.960265e-03

3) **Conclusions**

For **Black-legged Kittiwakes**, the relationship between culmen length and wing span is not statistically significant and has a weak correlation ($p = 0.3074$). This implies that culmen length and wingspan are independent in Black-legged kittiwakes.

Red-legged Kittiwakes: A statistically significant ($p = 0.04798$) moderately a positive correlation has been found between wing span and culmen length.This suggests that culmen length and wingspan are not independent in Red-legged Kittiwakes.

Weight Differences Between Subspecies: The t-test ($p = 0.4216$) indicates that there is no statistically significant difference between the Black-legged and Red-legged kittiwakes weights. This implies that the two subspecies average weights are comparable.

Overall Difference Between Subspecies: The results of the two-sample t-tests show that, although the weight of the two subspecies of kittiwakes does not differ significantly (p-value > 0.05), the wingspan and culmen length do ($p < 0.05$). This suggests that although there is no significant difference in weight between Red-legged and Black-legged kittiwakes, there are significant differences in their culmen and wingspan lengths.

Location Data

1) Methods

1.1) **Linear Model for Breeding Pairs:**
Created a linear regression model involving ecological variables to forecast the number of breeding pairs.

1.2) **Linear Model with Log-Transformation:**
Developed a linear model using the same predictors and the response as the logarithm of the number of breeding pairs.

1.3) **Model Selection:**
Based on fit statistics and model assumptions, the best model was selected.

1.4) **Analysis of Model Fit and Covariates:**
Evaluated the fit of the chosen model and examined the effect of each covariate on the number of breeding pairs.

1.5) **Confidence Interval for Prediction:**
Computed a 98% confidence interval for breeding pairs at a specified site, using the selected model.

2) Results

a) **Linear Model for Breeding Pairs**
- T-value: 10.317, P-value: 6.82e-10
- Multiple R-squared: 0.8679 , Adjusted R-squared: 0.8319
- AIC: 245.8843
- Residual standard error: 14.62

b) **Linear Model with Log-Transformation:**
- T-value : 19.964 , P-value : 1.38e-15
- Multiple R-squared: 0.9585 , Adjusted R-squared: 0.9472
- Residual standard error: 0.1453
- AIC: -21.60552

c) **Confidence Interval for Prediction:**
- Predicted Breeding Pairs: 93.55
- 98% Confidence Interval: [71.16, 122.98]

3) **Conclusions**

3.1) **Linear Model for Breeding Pairs:**

Breeding pairs were used as the response variable in a linear model that took consideration of cliff height, summer temperature, sandeel concentration, and coastal direction.

3.2) **Linear Model with Log-Transformation:**

A second model was developed using the logarithm of breeding pairs as the response variable

3.3) **Selection of Appropriate Model:**

The log-transformed model outperformed the other one in terms of fit, as indicated by its lower Akaike information criterion (AIC), higher R-squared values and less residual standard error. As a result, it was chosen as the better model for predicting breeding pairs.

3.4) **Impact of Covariates and Model Fit:**

The relationship between the selected covariates and breeding pairs was well-represented by the log-transformed model. The numbers of breeding pairs were notably impacted by sandeel concentration and cliff height, indicating their critical role in kittiwake breeding behavior.

3.5) **Confidence Interval for Prediction:**

A 98% confidence interval was computed using the log-transformed model for the expected number of breeding pairs at a site with particular parameters (west coast direction, 2.41 sandeel concentration, 25.1% summer temperature, and 3.12 cliff height). This interval gave a statistically significant estimate of the number of breeding pairs under these conditions, ranging from 71.16 to 122.98.