Palmer Penguins Analysis

Daniel Cusimano

4/20/23

# 1. Summary/Abstract

*Write a summary of your project.*

# 2. Introduction

## 2.1 General Background Information

Penguins of the genus *Pygoscelis*, commonly known as “brush-tailed penguins”, occupy areas across Antartica and the surrounding islands. There are currently three known extant species within this genus: the Gentoo Penguin (*Pygoscelis papus*), the Chinstrap Penguin (*Pygoscelis antarctica*), and the Adelie Penguin (*Pygoscelis adeliae*). As environmentally sensitve species living in the Southern Pole, these penguins have become of interest to research studying the rate and effects of climate change as well as marine ecology and avian ecology. With such research potential, the *Pygoscelis* penguins have become key in our understanding of environmental, ecological, and evolutionary change in the South Pole.

For any productive research to occur by way of studying the brush-tailed penguins, it is first important to gather baseline data on behavior, ecology, and morphology. Such analyses then allow researchers to track changes in these qualities and discern potential influences thereof. The purpose of the current project is to focus on morphology and interspecific variation within the genus. We believe that such data will not only allow for an understanding of contemporary impacts of climate and ecological change on polar avian species but will also provide insight into the evolutionary mechanisms involved in the history of these species.

## 2.2 Data and data source

Data for this research comes from and was collected as a part of the Palmer Long-Term Ecological Research (LTER) study (**Gorman?**). The LTER is stationed West of the Antarctic Peninusla and extends North to South along the Palmer Basin.

The data collected from the LTER was produced by Dr. Kristen Gorman and made accessible as part of the project template provided by Dr. Margurite Butler (https://github.com/mbutler808/Project-template).

## 2.3 Questions/Hypotheses to be addressed

In this analysis, the research aims to investigate the distribution of size and shape across Pygoscelis penguins in Antarctica. Given the sympatric nature of different species within this genus, it is pertintent to consider questions of their evolution and divergence. Such investigations will potentially allow for predictive models of species response to environmental change. Prior to theoretical analyses, however, the current plan is to consider the extant morphological differnces.

Question to be addressed throughout this analysis consider:

1. How Body mass varies across the different species.
2. The role that sexual dimorphism plays in a species size variation.
3. Ways in which sympatric populations of Adelie and Chinstrap penguins differ.

# 3. Methods

The information used for this analysis was processed through the R statistical software. It was cleaned and organized prior to this work and through a process separate from this. The processing code used to clean and check the data can be found in the Supplement material accessible through the author’s open access github repository (https://github.com/cusisom/cusimano-rclass-project).

## 3.1 Data acquisition, import, and cleaning

Original data was imported from the Palmer Penguins dataset. The file used for this analysis was originally altered by Dr. Marguerite Butler and published as the penguins\_raw\_dirty.r script. After cleaning a separate script was generated for analysis. This processed data was imported through the penguins.rds file. All materials are accessible through the author’s repository for this project (https://github.com/cusisom/cusimano-rclass-project).

## 3.2 Statistical analysis

Statistal analyses performed for this project include Welch’s T-test for examination of Sexual Size Dimorphism, and Principal Components analysis and ANOVA tests for investigation into interspecific shape variation between sympatric species. All analyses were performed through the R statistical software.

# 4. Results

## 4.1 Exploratory/Descriptive analysis

The cleaned dataset used for this investigation focused on four morphological variables (Culmen Length, Culmen Depth, Flipper Length, and Body Mass) and two demographic variables (Species and Sex). It thus omitted information included in the original dataset pertaining to isotopic analysis and geographic categories. Table 1 summarizes the original dataset used. Visible here are all variables provided for this project.

[Table 1](#tbl-summary_table) shows a summary of the processed data.

Table 1: Data summary table.

| Variable | N | Mean | SE | Counts |
| --- | --- | --- | --- | --- |
| Species | 339 |  |  | Ade: 151, Gen: 120, Chi: 68 |
| Island | 339 |  |  | Bis: 164, Dre: 124, Tor: 51 |
| Sex | 330 |  |  | MAL: 168, FEM: 162 |
| Culmen Length | 339 | 43.90 | 0.30 |  |
| Culmen Depth | 339 | 17.17 | 0.11 |  |
| Flipper Length | 339 | 200.79 | 0.76 |  |
| Body Mass | 339 | 4196.17 | 43.62 |  |
| Delta 15 N | 327 | 8.74 | 0.03 |  |
| Delta 13 C | 328 | -25.68 | 0.04 |  |

This data was compared in earlier analyses to show the correlation of the different morphological variables available for investigation. These correlations, as well as depiction of Body Mass distribution, are illustrated in [Figure 1](#fig-cor_plots). The data is separated by species to show the relative distributions.

## 4.2 Body Mass Distribution

The first goal of this analysis was to determine factors contributing to body mass distribution in each species given that this variable seemed showed the greatest range in the dataset. The histogram above (Figure 1) highlights not only a significant range in body mass but considerable overlap in each of the three species. Replotted, however, this data clearly establishes differences in body mass range ([Figure 2](#fig-penguins_MassV)).

|  |
| --- |
| Figure 1: Body Mass Correlations |

Beyond the stark separation of Gentoo penguins from Adelie and Chinstrap penguins, it is interesting to notice the relatively narrow range in body mass of Chinstrap populations. Despite overlapping in overall size with Adelie, this pattern clearly differs from the others. While it is possible the the narrower range in body mass is a consequence of geography, the Adelie penguins (with a much greater range in mass) are sympatric with Chinstrap penguins. With this in mind, an alternative explanation may reflect differences in the degree of sexual dimorphism.

|  |
| --- |
| Figure 2: Species Body Mass Distribution |

Sexual size dimorphism, reflecting significant differences in the size of males and females of a species, theoretically account for a greater range of body mass across a species. Species with minimal dimorphism, therefore should reflect narrow ranges. [Figure 3](#fig-Dimorphism_violin) illustrates the body mass distributions for each species as separated by sex. The image suggests that 1) sexual size dimorphism exists in each species and 2) that the Chinstrap Penguins reflect relatively less dimorphism than Adelie and Gentoo Penguins.

|  |
| --- |
| Figure 3: Body Mass Dimorphism~Species |

## 4.3 Sexual Dimorphism

To determine the significance of sexual size dimorphism and the degree by which each varied by sex, I ran individual tests for each species.

A quick glimpse at Body Mass mean differences in male and female the palmer Penguins suggested that dimorphism was significant . This was confirmed by reviewing the t-tests for each species.

We found that Gentoo penguins are substantially larger than both Adelie and Chinstrap penguins, which are similar in size (**?@fig-mass\_species\_bars**, ANOVA P-value < 0.001 ).

Table 2: Mean body mass by sex: Adelie

| Variable | FEMALE (n=73) | MALE (n=73) | Total (n=146) |
| --- | --- | --- | --- |
| Mean Body Mass (SD) | 3368.84 (269.38) | 4043.49 (346.81) | 3706.16 (458.62) |

Table 3: Mean body mass by sex: Gentoo

| Variable | FEMALE (n=55) | MALE (n=61) | Total (n=116) |
| --- | --- | --- | --- |
| Mean Body Mass (SD) | 4671.36 (285.39) | 5484.84 (313.16) | 5099.14 (505.80) |

Table 4: Mean body mass by sex: Chinstrap

| Variable | FEMALE (n=34) | MALE (n=34) | Total (n=68) |
| --- | --- | --- | --- |
| Mean Body Mass (SD) | 3527.21 (285.33) | 3938.97 (362.14) | 3733.09 (384.34) |

The reported mean mass differences support the notion that Gentoo and Adelie Penguins reflect stronger Sexual Size Dimorphism but need to be confirmed statistically. Tables 5-7 present the results of individual t-tests for each species. In support of my hypothesis, Chinstrap Penguins show much lower t-values (-5.2077) than Adelie(-13.126) and Gentoo (-14.638) Penguins.

Table 5: Adelie Sexual Dimorphism T-test

|  | Females | Males | T.value | P.value |
| --- | --- | --- | --- | --- |
| mean of x | 3368.84 | 4043.49 | -13.13 | 0 |

Table 6: Gentoo Sexual Dimorphism T-test

|  | Females | Males | T.value | P.value |
| --- | --- | --- | --- | --- |
| mean of x | 4671.36 | 5484.84 | -14.64 | 0 |

Table 7: Chinstrap Sexual Dimorphism T-test

|  | Females | Males | T.value | P.value |
| --- | --- | --- | --- | --- |
| mean of x | 3527.21 | 3938.97 | -5.21 | 0 |

Table 8: Size differences between species by island.

|  | Df | Sum Sq | Mean Sq | F value | Pr(>F) |
| --- | --- | --- | --- | --- | --- |
| Species | 2 | 145838050.34 | 72919025.2 | 337.32 | 0.00 |
| Island | 2 | 13654.61 | 6827.3 | 0.03 | 0.97 |
| Residuals | 334 | 72202059.80 | 216173.8 |  |  |

## 4.4 Variance between Adelie and Chinstrap

The next goal was to see, besides by body mass and dimorphism, how to differentiate the Adelie and Chinstrap Penguins.

# 5. Discussion

## 5.1 Summary and Interpretation

*Summarize what you did, what you found and what it means.*

## 5.2 Strengths and Limitations

*Discuss what you perceive as strengths and limitations of your analysis.*

## 5.3 Conclusions

*What are the main take-home messages?*

*Include citations in your Rmd file using bibtex, the list of references will automatically be placed at the end*

This paper (Leek and Peng 2015) discusses types of analyses.

These papers (McKay et al. 2020a,b) are good examples of papers published using a fully reproducible setup similar to the one shown in this template.

Note that this cited reference will show up at the end of the document, the reference formatting is determined by the CSL file specified in the YAML header. Many more style files for almost any journal [are available](https://www.zotero.org/styles). You also specify the location of your bibtex reference file in the YAML. You can call your reference file anything you like, I just used the generic word references.bib but giving it a more descriptive name is probably better.

# 6. References

Horst, A. M., A. P. Hill, and K. B. Gorman. 2020. [Palmerpenguins: Palmer archipelago (antarctica) penguin data](https://doi.org/10.5281/zenodo.3960218).

Leek, J. T., and R. D. Peng. 2015. [Statistics. What is the question?](https://doi.org/10.1126/science.aaa6146) Science (New York, N.Y.) 347:1314–1315.

McKay, B., M. Ebell, W. Z. Billings, A. P. Dale, Y. Shen, and A. Handel. 2020a. [Associations Between Relative Viral Load at Diagnosis and Influenza A Symptoms and Recovery.](https://doi.org/10.1093/ofid/ofaa494) Open forum infectious diseases 7:ofaa494. United States.

McKay, B., M. Ebell, A. P. Dale, Y. Shen, and A. Handel. 2020b. [Virulence-mediated infectiousness and activity trade-offs and their impact on transmission potential of influenza patients.](https://doi.org/10.1098/rspb.2020.0496) Proceedings. Biological sciences 287:20200496. England.