

The penguins of Antarctica

Selina Baldauf

25/3/2022

Contents

1	Introduction	1
2	Methods	1
2.1	The data	1
2.2	The analysis	2
3	Results	2
	References	4

1 Introduction

There are three main penguin species in Antarctica (*Chinstrap*, *Gentoo*, *Adelie*). You can see them in the following figure:

In this paper we want to answer the following questions

1. How bill depth depends on bill length?
2. Which penguin species has the highest body mass?

2 Methods

All analysis was done using R version 4.1.3 (R Core Team 2022) and the R markdown package (Allaire et al. 2021).

2.1 The data

The data was collected on islands in Antarctica and published by Gorman, Williams, and Fraser (2014). You can find the original paper with the title “Ecological sexual dimorphism and environmental variability within a community of Antarctic penguins (genus *Pygoscelis*)” (Gorman, Williams, and Fraser 2014) in PLoS ONE¹

The data is published via the `palmerpenguins` R package (Horst, Hill, and Gorman 2020) which you can find [on this website](#).

The data contains (among others) the following measurements:

- bill length
- bill depth
- body mass

¹paper available [here](#).

- sex
 - male
 - female

2.2 The analysis

We did some plots, calculated some summary statistics and a linear model of the form $y = ax + b + \epsilon$

3 Results

The mean weight of all penguin species is 4201.754386. *Gentoo* penguins have an average weight of 5076 g, *Adelie* penguins of 3701 g and *Chinstrap* penguins of 3733 g.

The figure below shows that *Gentoo* penguins have the highest body mass.

```
ggplot(penguins, aes(x = body_mass_g, fill = species)) +
  geom_histogram(alpha = 0.6) +
  scale_fill_manual(values = c("darkorange", "purple", "cyan4")) +
  theme_minimal()
```

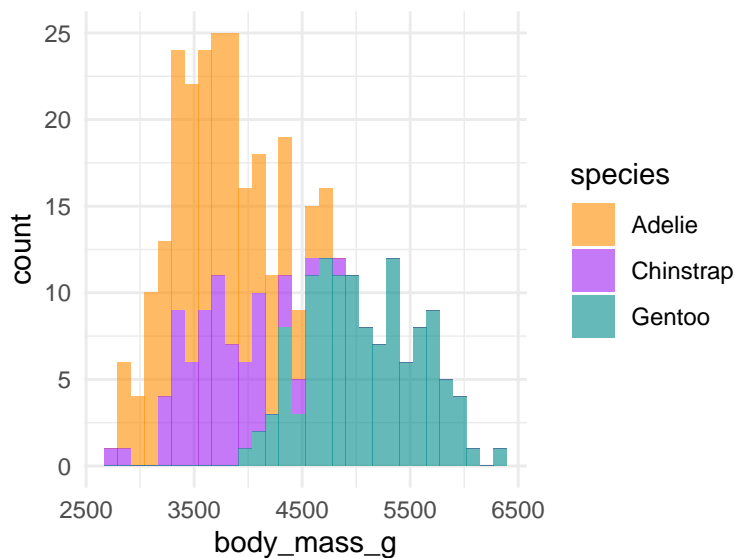


Figure 1: Histogram of weight of the three penguin species.

There is a positive relationship between bill length and bill depth for all 3 species, as the figure below shows.

```
ggplot(
  data = penguins,
  aes(
    x = bill_length_mm,
    y = bill_depth_mm,
    color = species,
    shape = species
  )
) +
  geom_point(size = 3, alpha = 0.8) +
  geom_smooth(method = "lm", se = FALSE) +
```

```
scale_color_manual(values = c("darkorange", "purple", "cyan4")) +
theme_minimal()
```

In general, it looks like the body characteristics differ between the sexes but also between the penguin species, as the table below illustrates:

```
penguins %>%
  filter(!is.na(sex)) %>%
  group_by(species, sex) %>%
  summarize(
    bill_length = mean(bill_length_mm, na.rm = TRUE),
    bill_depth = mean(bill_depth_mm, na.rm = TRUE),
    flipper_length = mean(flipper_length_mm, na.rm = TRUE),
    body_mass = mean(body_mass_g, na.rm = TRUE)
  )
```

species	sex	bill_length	bill_depth	flipper_length	body_mass
Adelie	female	37.25753	17.62192	187.7945	3368.836
Adelie	male	40.39041	19.07260	192.4110	4043.493
Chinstrap	female	46.57353	17.58824	191.7353	3527.206
Chinstrap	male	51.09412	19.25294	199.9118	3938.971
Gentoo	female	45.56379	14.23793	212.7069	4679.741
Gentoo	male	49.47377	15.71803	221.5410	5484.836

The linear model analysis shows that both bill length and species have a significant effect on the bill depth of the penguins.

```
# Calculate the linear model
lm_penguins <- lm(bill_depth_mm ~ bill_length_mm + species, data = penguins)
# Create the summary table
summary(lm_penguins)
```

```
##
## Call:
## lm(formula = bill_depth_mm ~ bill_length_mm + species, data = penguins)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.4529 -0.6864 -0.0508  0.5519  3.5915
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   10.59218    0.68302   15.508 < 2e-16 ***
## bill_length_mm    0.19989    0.01749   11.427 < 2e-16 ***
## speciesChinstrap -1.93319    0.22416  -8.624 2.55e-16 ***
## speciesGentoo   -5.10602    0.19142 -26.674 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9533 on 338 degrees of freedom
## (2 observations deleted due to missingness)
## Multiple R-squared:  0.769, Adjusted R-squared:  0.7669
## F-statistic: 375.1 on 3 and 338 DF, p-value: < 2.2e-16
```

References

- Allaire, JJ, Yihui Xie, Jonathan McPherson, Javier Luraschi, Kevin Ushey, Aron Atkins, Hadley Wickham, Joe Cheng, Winston Chang, and Richard Iannone. 2021. *Rmarkdown: Dynamic Documents for r*. <https://github.com/rstudio/rmarkdown>.
- Gorman, Kristen B., Tony D. Williams, and William R. Fraser. 2014. “Ecological Sexual Dimorphism and Environmental Variability Within a Community of Antarctic Penguins (Genus *Pygoscelis*).” Edited by André Chiaradia. *PLoS ONE* 9 (3): e90081. <https://doi.org/10.1371/journal.pone.0090081>.
- Horst, Allison Marie, Alison Presmanes Hill, and Kristen B Gorman. 2020. *Palmerpenguins: Palmer Archipelago (Antarctica) Penguin Data*. <https://doi.org/10.5281/zenodo.3960218>.
- R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.