Exploration des données palmerpenguins

Olivier Gimenez

10/25/2020

# Exploration des données

## Motivation

Dans cette section, on se propose d’**explorer** les données du package palmerpenguins. Une publication récente de la chercheuse, Dr Kristen Gorman, qui a partagé ces données est Connors et al. (2020).

## Données

Les données se présentent sous la forme suivante (10 premières lignes) :

penguins %>%   
 slice(1:10) %>%  
 knitr::kable()

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| species | island | bill\_length\_mm | bill\_depth\_mm | flipper\_length\_mm | body\_mass\_g | sex | year |
| Adelie | Torgersen | 39.1 | 18.7 | 181 | 3750 | male | 2007 |
| Adelie | Torgersen | 39.5 | 17.4 | 186 | 3800 | female | 2007 |
| Adelie | Torgersen | 40.3 | 18.0 | 195 | 3250 | female | 2007 |
| Adelie | Torgersen | NA | NA | NA | NA | NA | 2007 |
| Adelie | Torgersen | 36.7 | 19.3 | 193 | 3450 | female | 2007 |
| Adelie | Torgersen | 39.3 | 20.6 | 190 | 3650 | male | 2007 |
| Adelie | Torgersen | 38.9 | 17.8 | 181 | 3625 | female | 2007 |
| Adelie | Torgersen | 39.2 | 19.6 | 195 | 4675 | male | 2007 |
| Adelie | Torgersen | 34.1 | 18.1 | 193 | 3475 | NA | 2007 |
| Adelie | Torgersen | 42.0 | 20.2 | 190 | 4250 | NA | 2007 |

## Exploration numérique

Il y a 344 manchots dans le jeu de données, et 3 espèces différentes. Les données ont été récoltées dans 3 îles de l’archipel Palmer en Antarctique.

La moyenne par espèce des traits mesurés sur les manchots est la suivante :

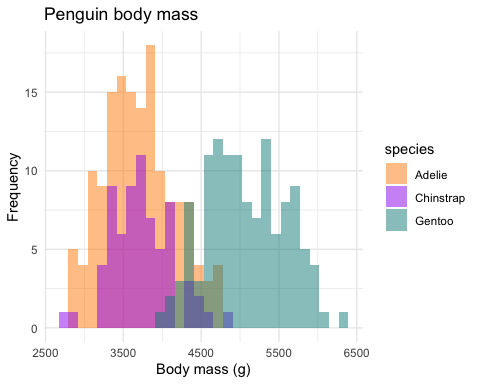
## # A tibble: 3 x 6  
## species bill\_length\_mm bill\_depth\_mm flipper\_length\_mm body\_mass\_g year  
## <fct> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 Adelie 38.8 18.3 190. 3701. 2008.  
## 2 Chinstrap 48.8 18.4 196. 3733. 2008.  
## 3 Gentoo 47.5 15.0 217. 5076. 2008.

## Exploration graphique

Un histogramme de la masse des manchots par espèce :

penguins %>%  
 ggplot() +  
 aes(x = body\_mass\_g) +   
 geom\_histogram(aes(fill = species),   
 alpha = 0.5,   
 position = "identity") +  
 scale\_fill\_manual(values = c("darkorange","purple","cyan4")) +  
 theme\_minimal() +  
 labs(x = "Body mass (g)",  
 y = "Frequency",  
 title = "Penguin body mass")

## Warning: Removed 2 rows containing non-finite values (stat\_bin).



## Image de la fin

Les trois espèces de manchots considérées :



## Références

Connors, Brendan, Michael J. Malick, Gregory T. Ruggerone, Pete Rand, Milo Adkison, James R. Irvine, Robert Campbell, and Kristen Gorman. 2020. “Climate and Competition Influence Sockeye Salmon Population Dynamics Across the Northeast Pacific Ocean.” *Canadian Journal of Fisheries and Aquatic Sciences* 77 (6): 943–49. <https://doi.org/10.1139/cjfas-2019-0422>.