# Data manipulation

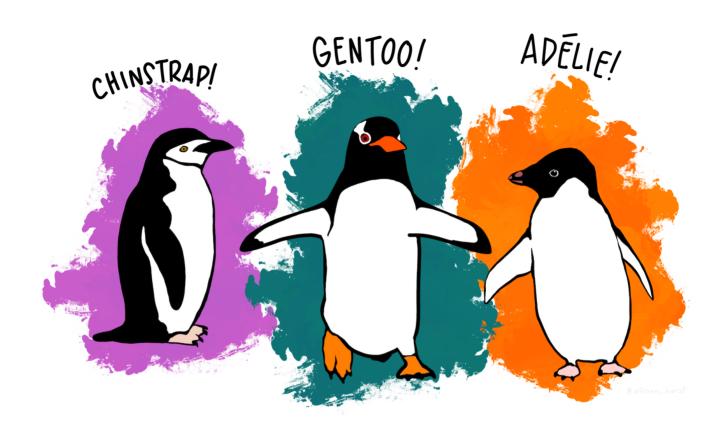
Gbadamassi G.O. Dossa

Updated on 2023-11-12 (created on 2021-09-13)

## Acknowledgements

The content of this module are based on materials from:

olivier gimenez's materials



#### **Motivation**

We use data from package palmerpenguins to familiarize ourselves with data manipulation with the tidyverse. We first load the packages we will need.

```
library(palmerpenguins)
library(tidyverse)
```

### Question 1

- a. Display the data penguins.
- b. Make sure you understand the columns we have in this dataset.
- c. Filter out penguins for which sex is missing.
- d. Select variables species, island, bill\_length\_mm and body\_mass\_g.
- e. Store the new dataset in a dat object.

### Question 2

- a. How many penguins do we have in the dataset?
- b. How many species?
- c. How many islands?
- d. Count the number of penguins per species.
- e. Count the number of penguins per species and per island.

#### Question 3

- a. Calculate the overall mean body mass.
- b. Calculate the mean body mass for each species.
- c. Calculate the mean of both traits bill length and body mass measured for each species.