

# Analiza danych dotyczących pingwinów

Marta Hałas

2024-12-18

## Wstęp

### Opis danych

Zbiór danych *penguins* z biblioteki *palmerpenguins* zawiera informacje dotyczące pingwinów z Wysp Palmera na Antarktyce. Przyjrzyjmy się zatem kolumnom w tym zbiorze danych.

- *species* - gatunek pingwina (Adelie, Gentoo, Chinstrap).
- *island* - wyspa (Torgersen, Biscoe, Dream), na której znaleziono pingwina.
- *bill\_length\_mm* - długość dzioba w milimetrach.
- *bill\_depth\_mm* - głębokość dzioba w milimetrach.
- *flipper\_length\_mm* - długość płetw w milimetrach.
- *body\_mass\_g* - masa pingwina w gramach.
- *sex* - płeć pingwina (male, female, NA).
- *year* - rok, w którym dane zostały zgromadzone (2007, 2008, 2009).

## Wygląd tabelki:

Liczba NA - braków danych w tabeli 19.

Wymiar tabelki: liczba wierszy to 344, liczba kolumn to 8.

## Opis analizy

Analizę rozpoczęłam od przedstawienia na wykresach danych związanych z płcią, gatunkiem i miejscem pochodzenia pingwinów.

### 1. Wykresy dotyczące liczby przebadanych pingwinów:

- **1.1** Rok, w którym dane zostały zgromadzone w zależności od liczby przebadanych pingwinów płci żeńskiej.
- **1.2** Rok, w którym dane zostały zgromadzone w zależności od liczby przebadanych pingwinów płci męskiej.
- **1.3** Rok, w którym dane zostały zgromadzone w zależności od liczby przebadanych pingwinów, u których występuje w wierszu w kolumnie *sex* wartość NA - brak danych.

### 2. Wykresy obrazujące dane wyspy w zależności od cech pingwinów:

- **2.1** Wyspy w zależności od liczby wystąpień pingwinów danej płci - żeńska, męska.
- **2.2** Wyspy w zależności od liczby wystąpień pingwinów danej płci - wartości NA (brak danych).
- **2.3** Wyspy w zależności od gatunków pingwina.

### 3. Wykres gatunków w zależności od płci pingwina:

- **3.1** Wykres gatunków w zależności od płci pingwina - męska, żeńska.
- **3.2** Wykres gatunków w zależności od płci pingwina - rozpatruję wartości NA - brak danych.

Kolejny etap poświęciłam policzeniu **odchylenia standardowego** masy oraz **wartości średniej** głębokości dzioba.

Na końcu stworzyłam tabelkę z wybranymi wskaźnikami sumarycznymi dla zmiennych zawierających:

- **1. Długość dzioba**
- **2. Głębokość dzioba**
- **3. Długość płetw**

## Wyniki

### Wykresy

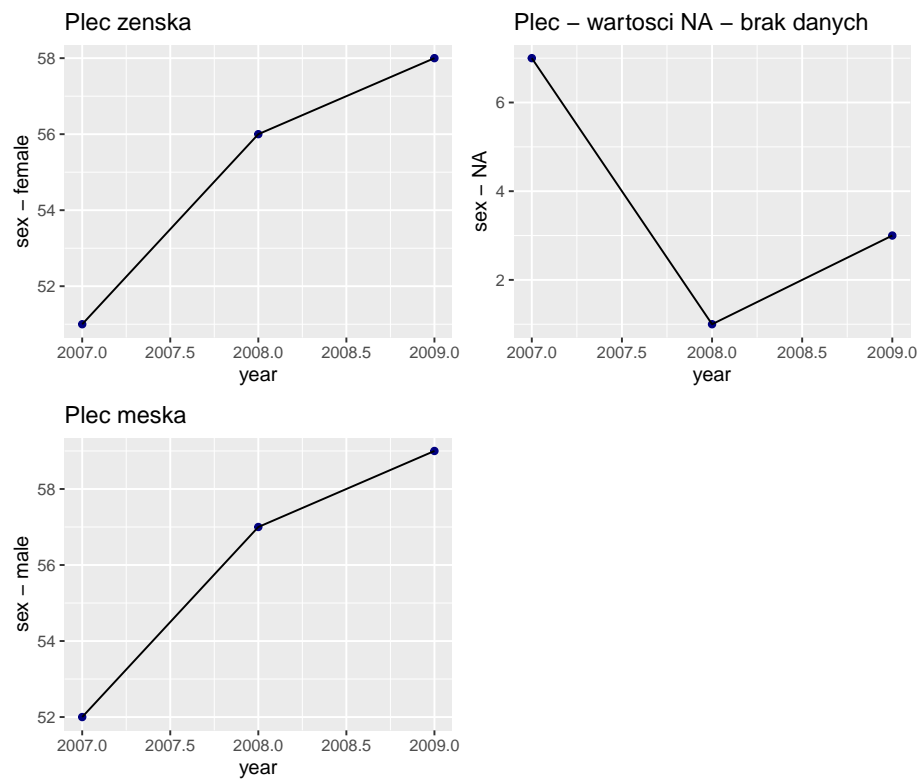


Figure 1: Wykresy dotyczące liczby przebadanych pingwinów

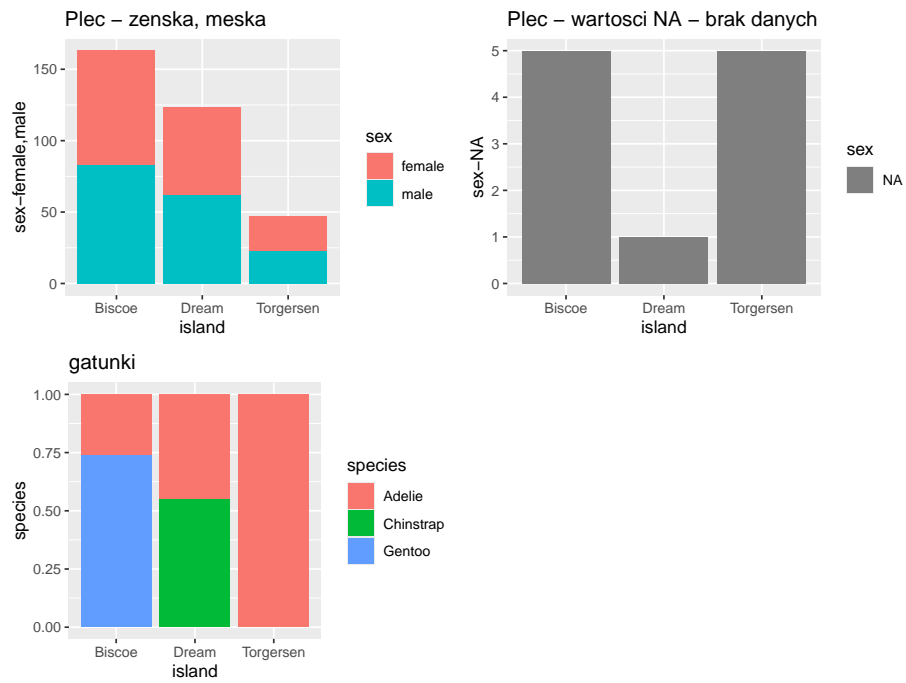


Figure 2: Wykres obrazujące dane wyspy w zależności od cech pingwinów

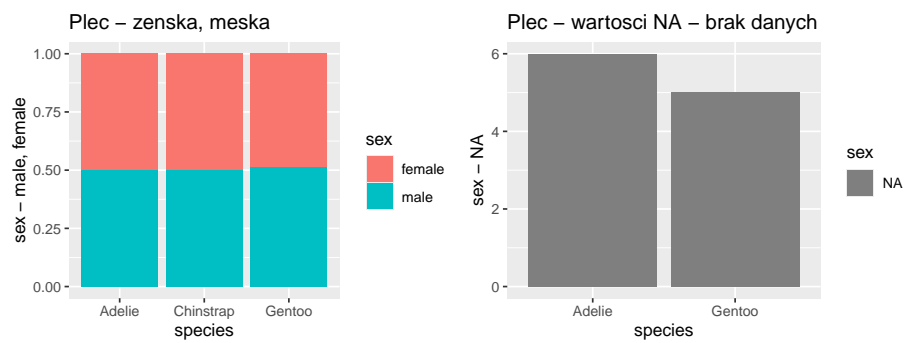


Figure 3: Wykres gatunków w zależności od płci pingwina

## Wskaźniki sumaryczne

### Wzór na odchylenie standardowe

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$$

Nasz  $\bar{x}$  to

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Wzór na wartość średnią

$$\frac{1}{n} \sum_{i=1}^n x_i$$

**Odchylenie standardowe** masy wynosi 801.9545357 g.

**Wartość** średnia głębokości dzioba wynosi 17.1511696 mm.

Table 1: Wybrane wskaźniki sumaryczne dla wybranych zmiennych

statystyka	długość dzioba	głębokość dzioba	długość płetw
Min.	32.10	13.10	172.00
1st Qu.	39.23	15.60	190.00
Median	44.45	17.30	197.00
Mean	43.92	17.15	200.92
3rd Qu.	48.50	18.70	213.00
Max.	59.60	21.50	231.00
NA's	2.00	2.00	2.00

## Wnioski

Po przeanalizowaniu pierwszych trzech wykresów widzimy, że **liczba badanych pingwinów** płci męskiej i żeńskiej wzrastała na przestrzeni lat, w których były gromadzone dane. Największy brak danych został zanotowany w 2007 roku, a najniższy w 2008 roku.

Kolejne trzy wykresy pokazują, że:

- **Wyspę Biscoe** zamieszkują dwa gatunki pingwinów: *Adelie* i *Gentoo*, z przewagą gatunku *Gentoo*. Na tej wyspie zebrano **najwięcej danych**. Nie widzimy również znaczącej różnicy w liczbie zebranych danych w zależności od płci. Z tej wyspy uzyskaliśmy 5 wartości NA w wierszach w kolumnie **sex**.
- Na wyspie **Dream** możemy znaleźć pingwiny gatunku: *Chinstrap* oraz *Adelie*. Nieznaczną przewagę ma gatunek *Chinstrap*. Na tej wyspie odnotowano **najmniejszy brak danych** – tylko 1 nieznana płeć. Również, jak w przypadku poprzedniej wyspy, liczba przebadanych pingwinów płci żeńskiej nie odbiega od liczby przebadanych pingwinów płci męskiej.
- Dane pochodzące z wyspy **Torgesen** dotyczą tylko jednego gatunku: *Adelie*. Jest to wyspa, z której otrzymaliśmy **najmniej danych**. Analogicznie jak powyżej, liczba przebadanych pingwinów ze względu na płeć jest sobie niemal równa. Liczba wartości NA - braku danych na temat płci pingwinów na wyspie Torgesen również wynosi 5.

**Ostatnie dwa wykresy** pokazują, że:

- Jedyne braki danych - nasze wartości NA, pochodzą z dwóch gatunków: *Adelie* (6 wartości NA) oraz *Gentoo* (5 wartości NA). Dane dotyczące gatunku *Chinstrap* są kompletne.
- Liczba przebadanych pingwinów ze względu na płeć jest sobie równa dla gatunku *Adelie* i *Chinstrap*. W przypadku gatunku *Gentoo* zauważamy lekką przewagę liczebną osobników płci żeńskiej.