Penguins Dataset

Gen

2024-02-26

```
# let's start by nstalling and loading the required packages
if (!requireNamespace("ggplot2", quietly = TRUE)) {
  install.packages("ggplot2")
}
if (!requireNamespace("palmerpenguins", quietly = TRUE)) {
  install.packages("palmerpenguins")
}
library(ggplot2)
library(palmerpenguins)
```

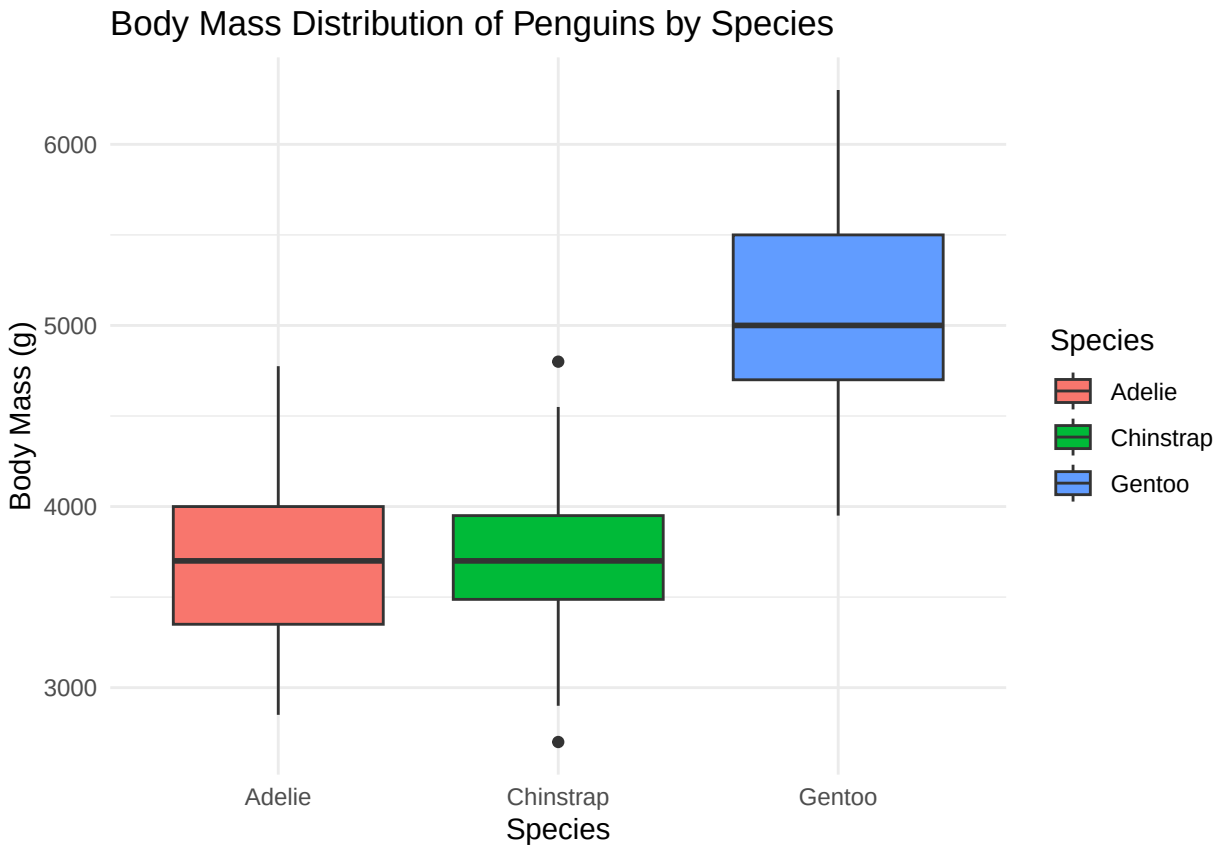
Now, I will load the penguins dataset

```
data("penguins")
```

Nex, I want to create a GGplot - Boxplot of body mass distribution of penguins by species

```
ggplot(penguins, aes(x = species, y = body_mass_g, fill = species)) +
  geom_boxplot() +
  labs(title = "Body Mass Distribution of Penguins by Species",
       x = "Species",
       y = "Body Mass (g)",
       fill = "Species") +
  theme_minimal()
```

```
## Warning: Removed 2 rows containing non-finite values (`stat_boxplot()`).
```

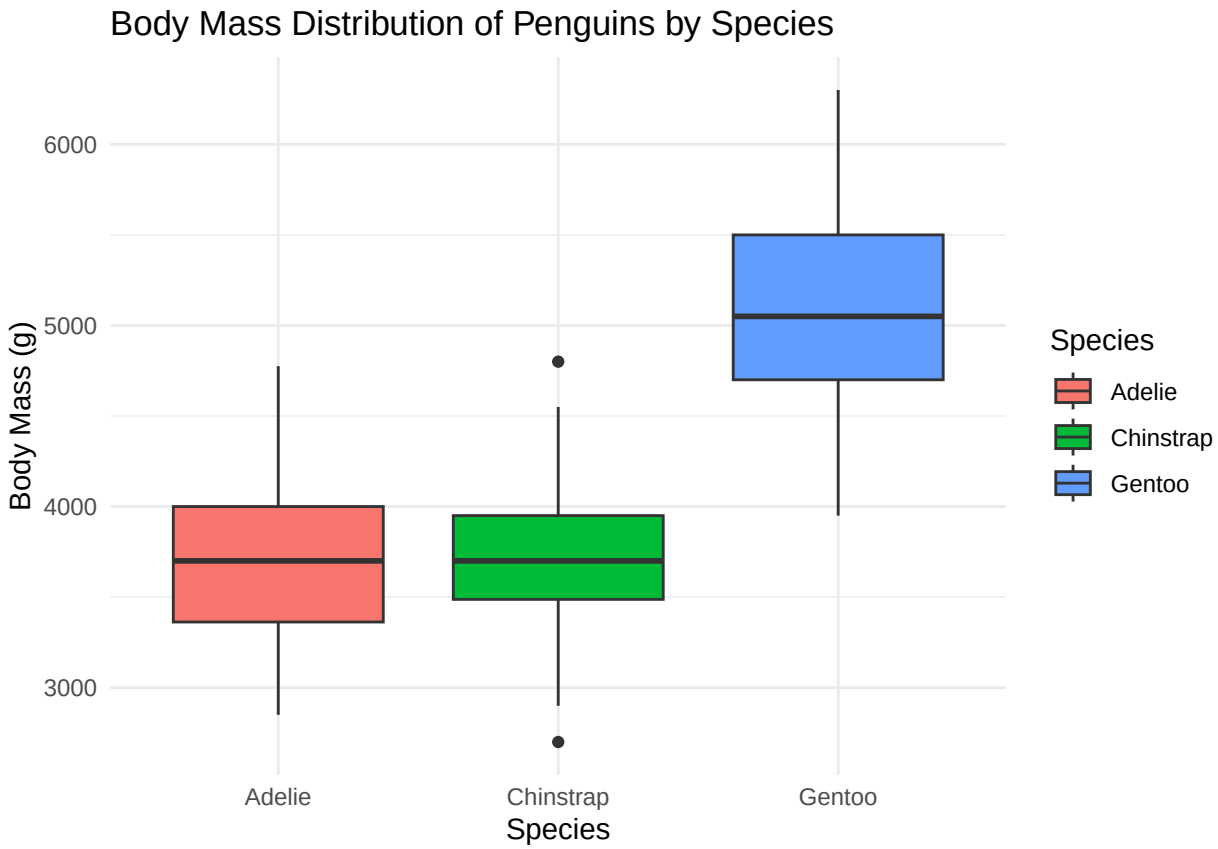


Oops, we got a warning! I will remove the rows with missing values and see what happens.

```
penguins_clean <- na.omit(penguins)
```

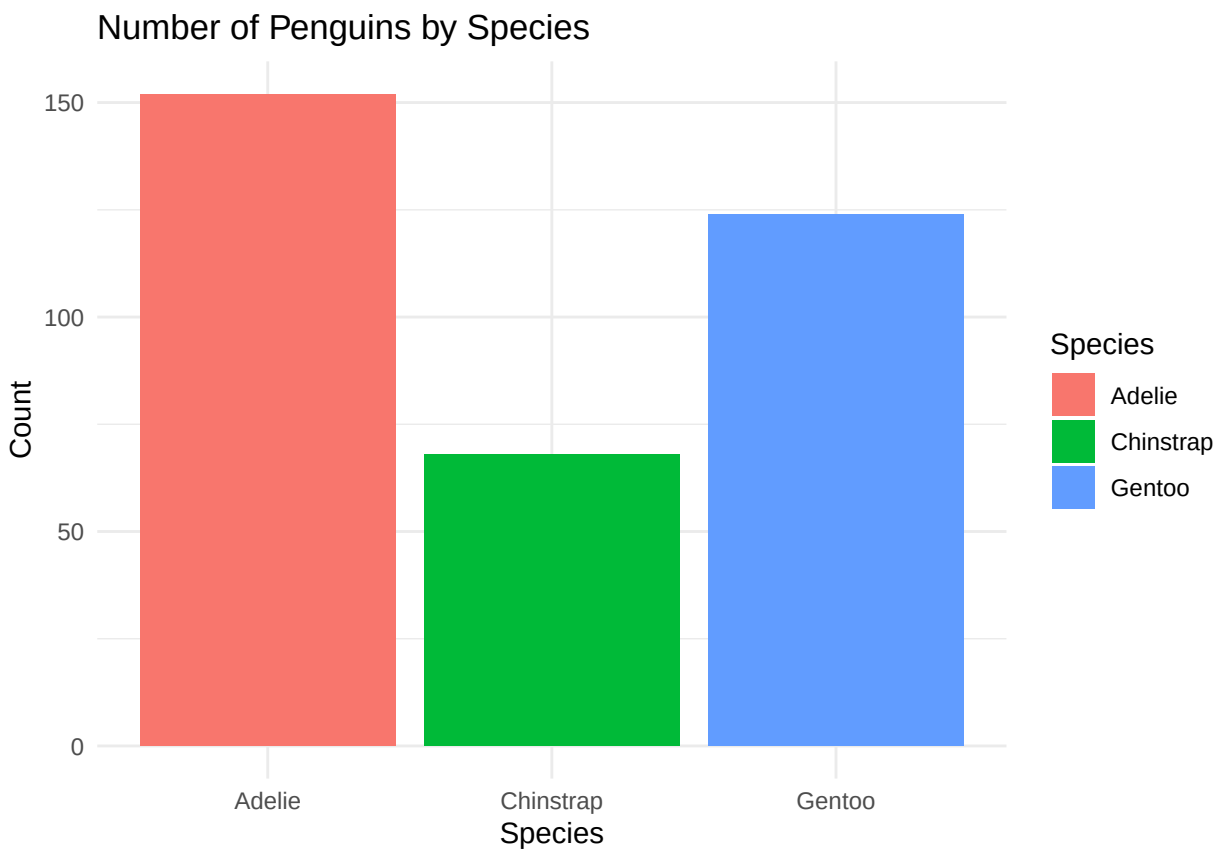
Now, I will re-create a GGplot - Boxplot of body mass distribution of penguins by species

```
ggplot(penguins_clean, aes(x = species, y = body_mass_g, fill = species)) +
  geom_boxplot() +
  labs(title = "Body Mass Distribution of Penguins by Species",
       x = "Species",
       y = "Body Mass (g)",
       fill = "Species") +
  theme_minimal()
```



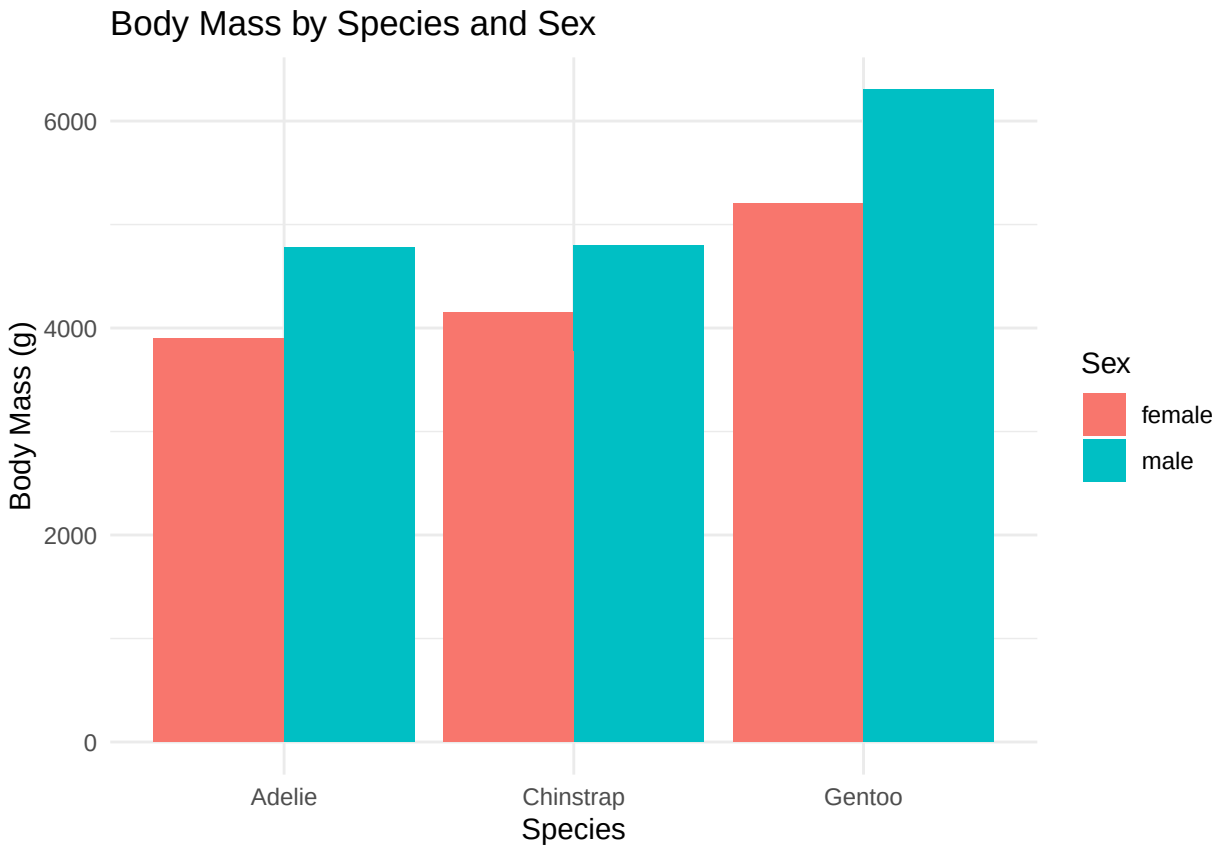
Much better!!! :) Now. I will create a bar chart of Number of penguins by species. I would like to know how many species we have in this dataset.

```
ggplot(penguins, aes(x = species, fill = species)) +
  geom_bar() +
  labs(title = "Number of Penguins by Species",
       x = "Species",
       y = "Count",
       fill = "Species") +
  theme_minimal()
```

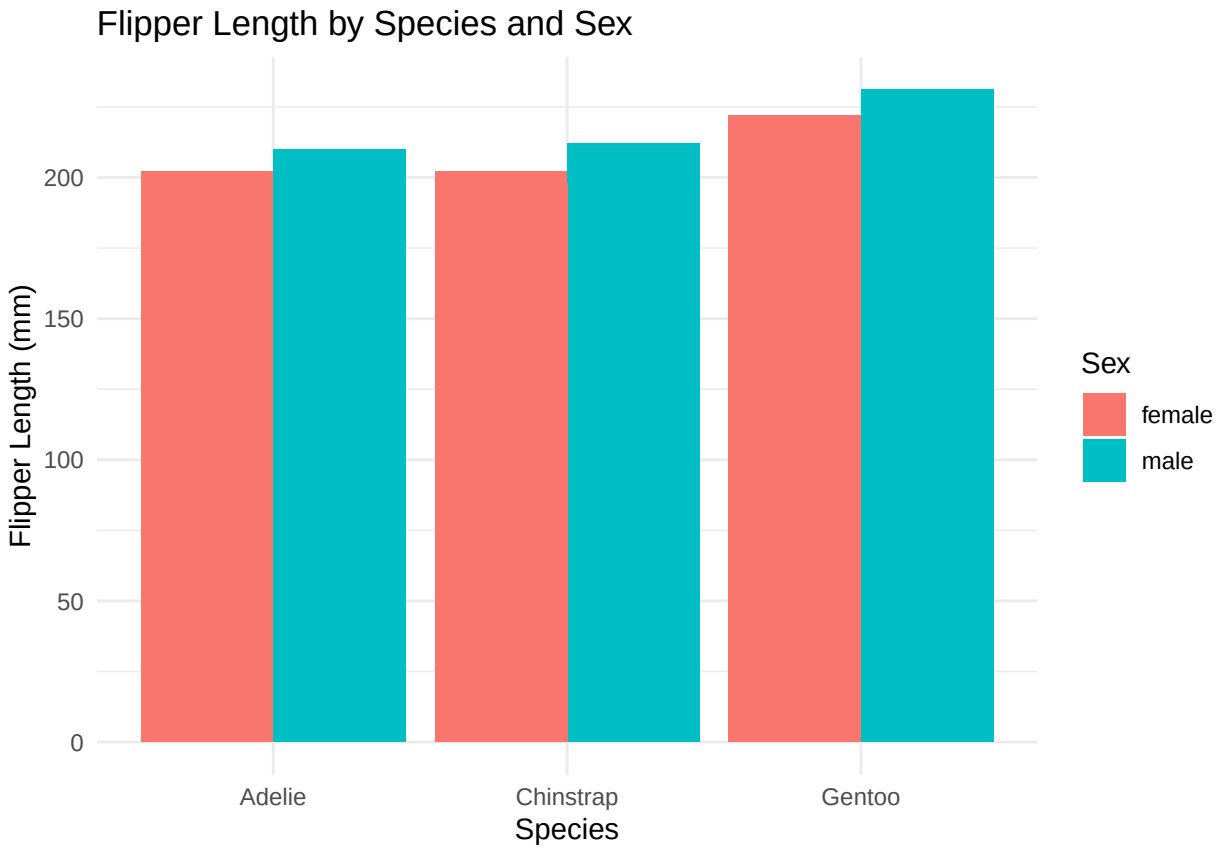


Sweet!! Now, I would like to see a bar chart of body mass by species & sex.

```
ggplot(penguins_clean, aes(x = species, y = body_mass_g, fill = sex)) +
  geom_bar(stat = "identity", position = "dodge") +
  labs(title = "Body Mass by Species and Sex",
       x = "Species",
       y = "Body Mass (g)",
       fill = "Sex") +
  theme_minimal()
```




```
#Lastly, I want to create a bar chart of flipper length grouped by species and sex
ggplot(penguins_clean, aes(x = species, y = flipper_length_mm, fill = sex)) +
  geom_bar(stat = "identity", position = "dodge") +
  labs(title = "Flipper Length by Species and Sex",
       x = "Species",
       y = "Flipper Length (mm)",
       fill = "Sex") +
  theme_minimal()
```


Flipper Length by Species and Sex
Flipper Length (mm)
200
150
100
50
0
Adelie
Chinstrap
Gentoo
Species
Sex
female
male