Homework 1

Aditi 24205303

Introduction

The palmerpenguins dataset is a widely used dataset for teaching data visualization in R. It contains observations of penguins from three species—*Adelie*, *Chinstrap*, and *Gentoo*—collected in Antarctica.

This dataset is useful for understanding how physical characteristics such as **body mass**, **flipper length**, and **bill size** vary across species.

Dataset Description

The key variables we'll focus on are:

- **species**: Penguin species (*Adelie*, *Chinstrap*, *Gentoo*)
- **body_mass_g**: Body mass in grams
- **sex**: Penguin sex (*male* or *female*)

Summary Statistics

Let's briefly examine the distribution of body mass across species and sexes:

- Average body mass (g):
  - *Adelie*: ~3700 g
  - *Chinstrap*: ~3730 g
  - *Gentoo*: ~5075 g

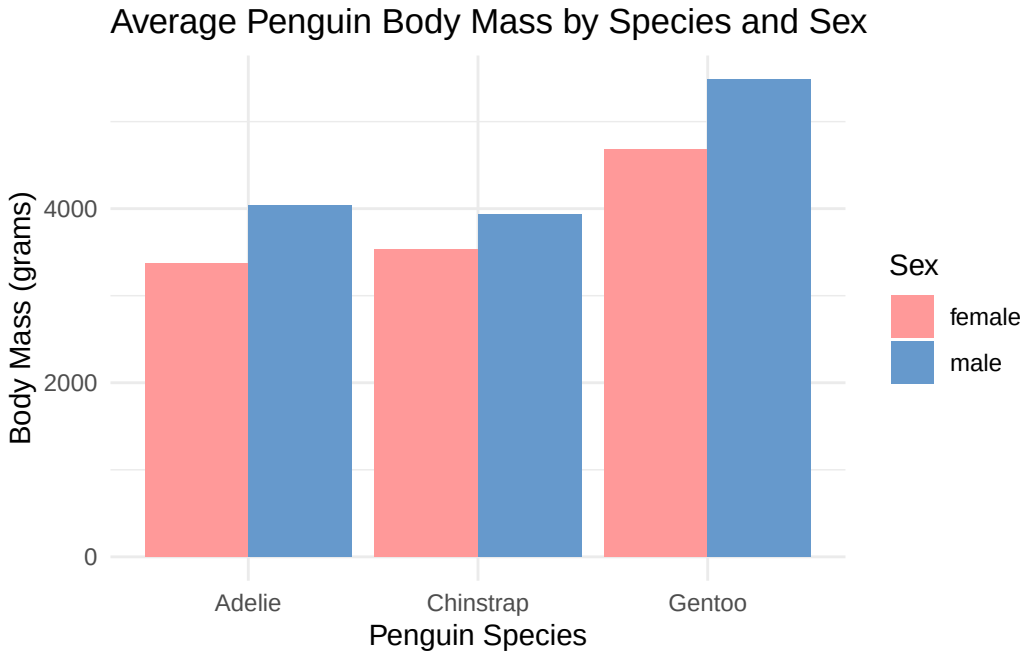
These numbers suggest that **Gentoo penguins are generally larger** than the other two species.

Body Mass by Species and Sex

The bar plot below shows the **average body mass** of male and female penguins across different species.

```
Warning: package 'ggplot2' was built under R version 4.3.3
```

```
Warning: package 'palmerpenguins' was built under R version 4.3.3
```



Explanation

This plot illustrates differences in body mass across penguin species, split by sex:

- **Male penguins** consistently weigh more than females.
- **Gentoo penguins** are the heaviest, regardless of sex.
- **Adelie** and **Chinstrap** penguins are smaller in comparison.

Interpretation

The visualized data clearly shows that **species** and **sex** are significant factors influencing penguin body mass. These patterns might relate to habitat, diet, or evolutionary traits specific to each species.