R Take-Home Assignment 2: Palmer Penguins (Functions & Visualization)

Your Name

2025-10-01

Overview

In this assignment you will explore the Palmer Penguins dataset, which contains measurements for three penguin species living in the Palmer Archipelago in Antarctica. Researchers collected data on their bill size, flipper length, body mass, sex, and island of origin. The dataset is a favourite in data science because it provides a simple but realistic setting to practise exploring biological variation. You will use this dataset to practise creating functions in R and visualising data in meaningful ways.

Source

`palmerpenguins` R package
Skills

`functions, grouped summaries, scatterplots, faceting`

Preparation

You will need: `dplyr`, `ggplot2`, `palmerpenguins`.

```
## Task: load packages (hint: library())
## library(dplyr)
## library(ggplot2)
## library(palmerpenguins)
```

1. Load and inspect the data

Before beginning any analysis, the first step is always to bring the data into R and take a first look. By knowing how many rows we have, which species are included, and what variables are measured, we can plan the kinds of questions we might answer.

Task: Load the `penguins` dataset from the `palmerpenguins` package. Then:

- Report the number of rows in the dataset.

- List the unique species included.

- Identify which variables are numeric.

```
## Task: load data (hint: data(penguins); or use penguins object)
## Task: inspect (hint: nrow(), dplyr::glimpse(), table(penguins$species))
```

2. Plot flipper length vs body mass

Biologists often wonder whether certain traits scale together. For penguins, larger flippers might be expected to support a heavier body. If this is true, we should see a clear relationship between flipper length and body mass. By colouring the points by species, we can also see whether the relationship holds consistently across Adelie, Chinstrap, and Gentoo penguins.

Task: Make a scatterplot of `flipper_length_mm` (x-axis) against `body_mass_g` (y-axis). Colour the points by `species`. Add axis labels and a title.

```
## Task: scatterplot (hint: ggplot() + geom_point())
```

3. Facet by island

Species sometimes live on different islands, and environmental conditions (like food availability) can influence body size. To check whether penguins differ not just by species but also by location, we can break the plot into separate panels for each island.

Task: Recreate the scatterplot and facet by `island`.

```
## Task: facet (hint: facet_wrap())
```

4. Write a summary function

Analyses are more powerful when we can repeat them easily. Suppose we want to compare average bill length, flipper length, and body mass for each species. Instead of writing separate code each time, we can create a function that computes these summaries for any species we provide.

Task: Write `species_summary(sp)` that returns mean bill length, flipper length, and body mass for species `sp`.

```
## Task: write function (hint: function() & summarise())
```

5. Apply your function

Now that you have written your function, you can quickly generate summaries for all species and combine them into one table. This gives a compact overview of how species differ.

Task: Apply your function to all species and bind results into one table.

```
## Task: apply (hint: unique(), lapply() & bind_rows())
```

6. Reflection

Finally, think about what these analyses reveal. Did larger flippers usually go hand in hand with heavier bodies? Were differences between species clear? Did location appear to matter?

Task: Write 3–4 sentences describing species differences and any island effects you observed.

Write your reflection here...