

Penguins: Day 1

2024-06-13

Packages

Penguin Exploration

Data were collected and made available by Dr. Kristen Gorman and the Palmer Station, Antarctica LTER, a member of the Long Term Ecological Research Network. We are going to explore and visualize these data on day 1!

Hints: In RStudio, you will write code in chunks between the “`“` marks. To run code, you will click the green arrow in the far right of the code chunk. Lines 29 - 31 are a code chunk. Variable names may not include blank spaces.

1. Take a glimpse of the data set `penguins`.

```
glimpse() # insert data set name inside the ().
```

How many rows does this data set have? How many columns? What is the variable name in the output that will give you the penguin’s flipper length in millimeters?

Add responses

2. Let’s calculate a summary statistic! In the code chunk below, fill in the `...` to calculate the mean body mass (g) for the Chinstrap species of penguin.

Hint: Look at your question 1 output to see your variable names!

```
penguins |>
  filter(species == "...") |> # insert the name of the penguin species inside the " "
  summarize(mean_bm = mean(...)) # insert the variable name
```

3. Now, create a scatterplot of flipper length on the x-axis, body mass on the y-axis, using species to color the results, by filling in the `...` below with correctly formatted variable names.

Hint: Look at your question 1 output to see your variable names!

```
penguins |>
  ggplot(
    aes(x = flipper_length_mm , y = body_mass_g , color = species) # insert your variable names here
  ) +
  geom_point() +
  labs(x = "flipper length (mm)",
       y = "body mass (g)")
```

What are some characteristics you notice?

Add response

Looking forward

Let’s compare body mass and flipper length across a subset of penguins! What can we take away from this graph? The following code chunk will produce charts to help you answer this question.

```

penguins |>
  filter(island %in% c("Torgersen", "Biscoe"),
         !is.na(sex)) |>
  ggplot(
    aes(x = body_mass_g, y = flipper_length_mm, color = species)
  ) +
  geom_point(alpha = 0.4) +
  geom_smooth(method = "loess", se = F) +
  facet_wrap(~sex) +
  labs(
    title = "Penguins body mass vs flipper length",
    subtitle = "on Torgersen and Biscoe island",
    y = "flipper length (mm)",
    x = "body mass (g)",
    color = "species",
    caption = "data and more found here: https://allisonhorst.github.io/palmerpenguins/"
  ) +
  theme(axis.text.x=element_text(angle=45, hjust=1)) +
  theme_bw()

```

We will learn and practice with each of the following functions and arguments above as the semester goes on!
 Now click “Knit” above to produce a nice HTML document of your results.

Example Lessons

If you would like to see more instances of the container used in class, please see the following class websites and lab-0 instructions used to help students work with containers!

This is Prof. Mine Çetinkaya-Rundel’s Intro to Data Science course:

<https://sta199-s24.github.io/>

This is day 1 lab from that course:

<https://sta199-s24.github.io/labs/lab-0.html>

This is Dr. Elijah Meyer’s Intro to Data Science course:

<https://sta199-f23-1.github.io/>