

New RMD file DH

Dylan

2025-06-12

Setting up my environment

Notes: setting up my R environment by loading the 'tidyverse' and 'palmerpenguins' packages

```
install.packages("tidyverse")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.3'  
## (as 'lib' is unspecified)
```

```
install.packages("palmerpenguins")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.3'  
## (as 'lib' is unspecified)
```

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --  
## v dplyr      1.1.4      v readr      2.1.4  
## v forcats    1.0.0      v stringr    1.5.0  
## v ggplot2     3.5.2      v tibble     3.2.1  
## v lubridate  1.9.2      v tidyr      1.3.1  
## v purrr      1.0.1
```

```
## -- Conflicts ----- tidyverse_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag()     masks stats::lag()  
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(palmerpenguins)
```

```
# Load packages
```

```
library(tidyverse)
```

```
library(palmerpenguins)
```

```
# Preview dataset
```

```
glimpse(penguins)
```

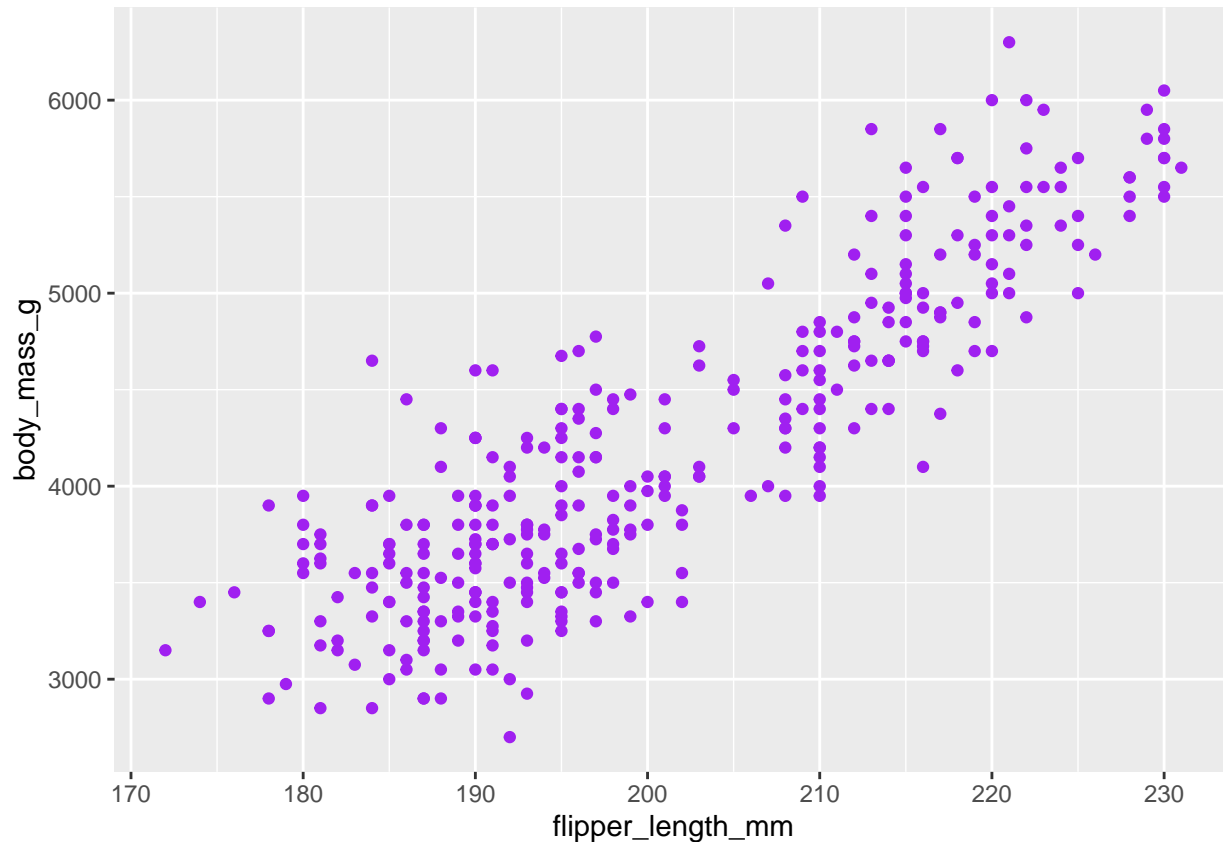
```
## Rows: 344  
## Columns: 8  
## $ species      <fct> Adelie, Adelie, Adelie, Adelie, Adelie, Adelie, Adel~  
## $ island        <fct> Torgersen, Torgersen, Torgersen, Torgersen, Torgersen, Torgersen~  
## $ bill_length_mm <dbl> 39.1, 39.5, 40.3, NA, 36.7, 39.3, 38.9, 39.2, 34.1, ~  
## $ bill_depth_mm <dbl> 18.7, 17.4, 18.0, NA, 19.3, 20.6, 17.8, 19.6, 18.1, ~  
## $ flipper_length_mm <int> 181, 186, 195, NA, 193, 190, 181, 195, 193, 190, 186~  
## $ body_mass_g    <int> 3750, 3800, 3250, NA, 3450, 3650, 3625, 4675, 3475, ~  
## $ sex            <fct> male, female, female, NA, female, male, female, male~
```

```
## $ year                <int> 2007, 2007, 2007, 2007, 2007, 2007, 2007, 2007, 2007~
```

```
# Plot 1: Basic scatterplot with purple points
```

```
ggplot(data = penguins, aes(x = flipper_length_mm, y = body_mass_g)) +  
  geom_point(color = "purple")
```

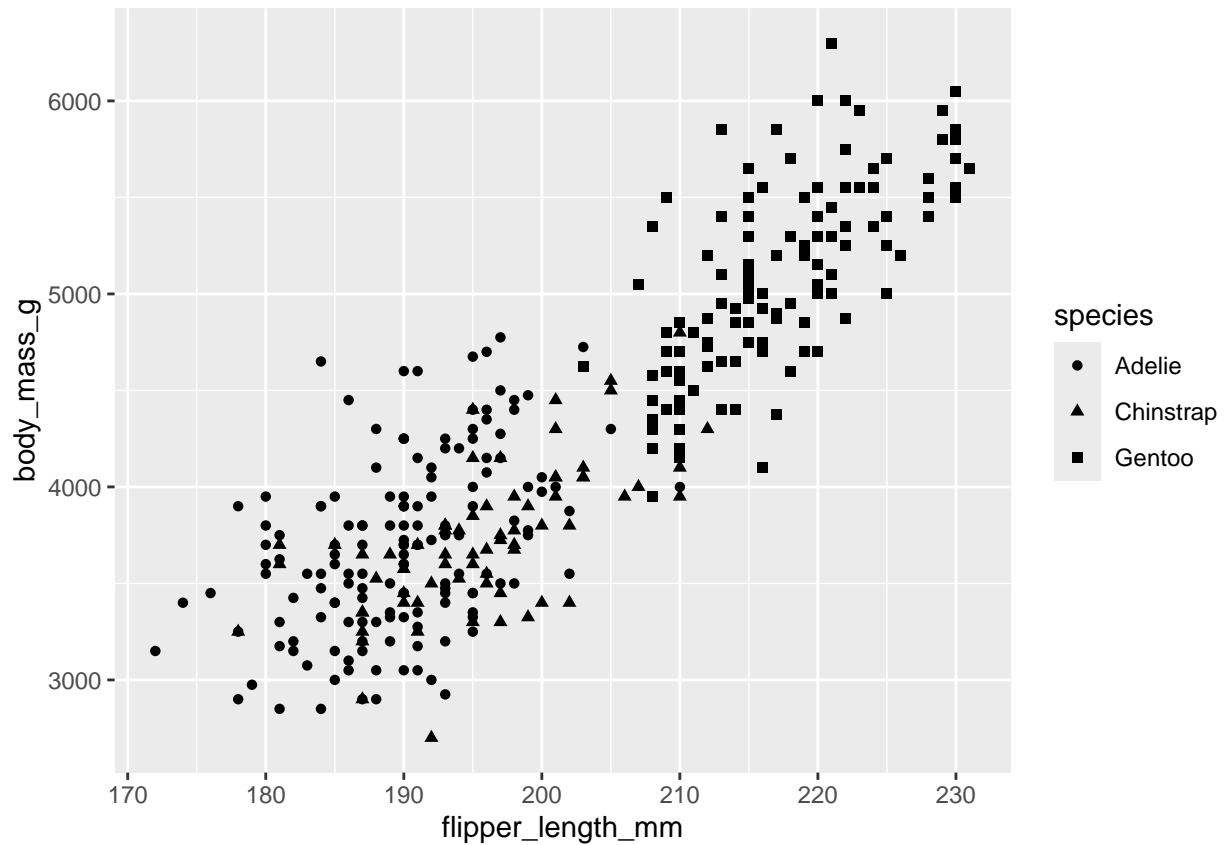
```
## Warning: Removed 2 rows containing missing values or values outside the scale range  
## (`geom_point()`).
```



```
# Plot 2: Scatterplot with shape mapped to species
```

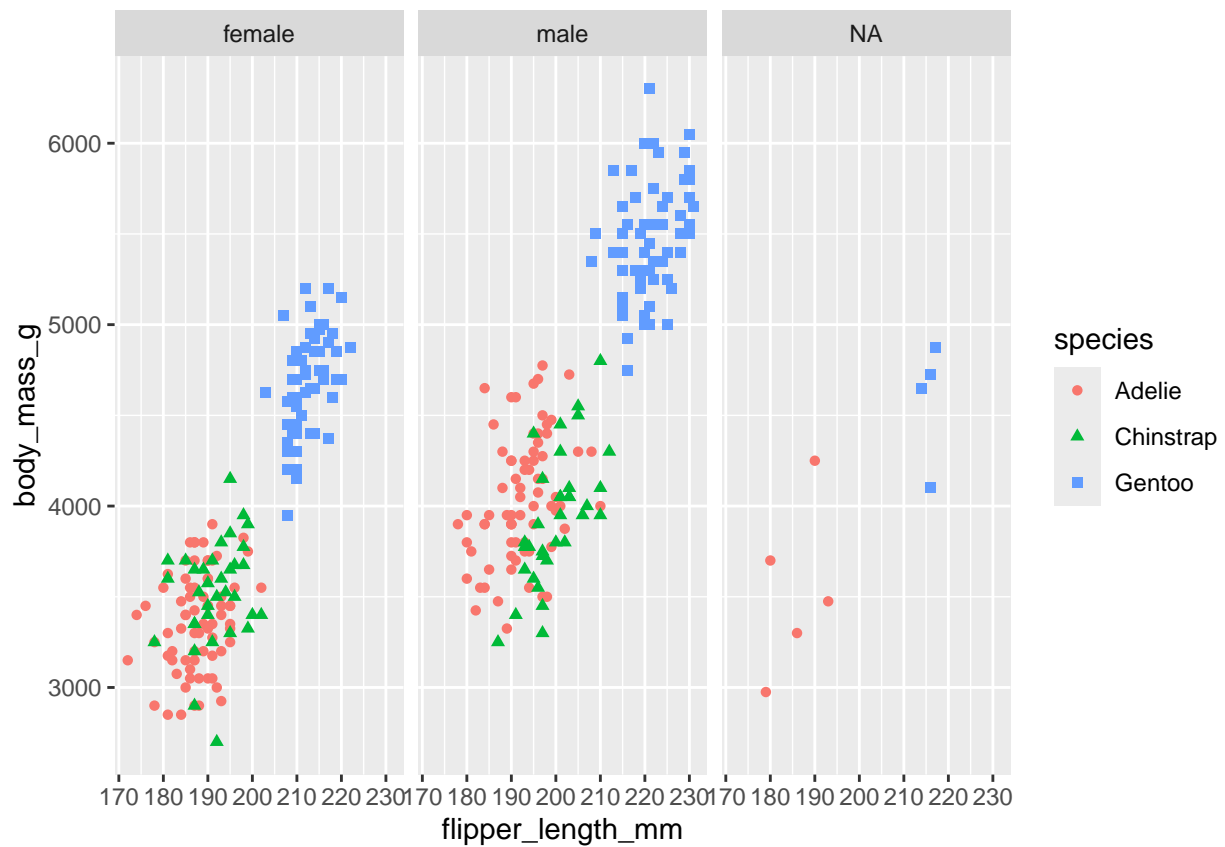
```
ggplot(data = penguins, aes(x = flipper_length_mm, y = body_mass_g)) +  
  geom_point(aes(shape = species))
```

```
## Warning: Removed 2 rows containing missing values or values outside the scale range  
## (`geom_point()`).
```



```
# Plot 3: Scatterplot with color and shape mapped to species, faceted by sex  
ggplot(data = penguins, aes(x = flipper_length_mm, y = body_mass_g)) +  
  geom_point(aes(color = species, shape = species)) +  
  facet_wrap(~sex)
```

```
## Warning: Removed 2 rows containing missing values or values outside the scale range  
## (`geom_point()`).
```



```
# Plot 4: Same as above, but with missing 'sex' values dropped
penguins %>%
  drop_na(sex) %>%
  ggplot(aes(x = flipper_length_mm, y = body_mass_g)) +
  geom_point(aes(color = species, shape = species)) +
  facet_wrap(~sex)
```

