Assignment 5

Miray Hozer

2025-11-10

Dataset

The `penguins` dataset (available through the `palmerpenguins` package and can be loaded after installing and loading the package) contains real-world data collected from three penguin species are Adélie, Chinstrap, and Gentoo living in the Palmer Archipelago of Antarctica. The measurements were taken between 2007 and 2009 consists many variables about the penguins such as `species`, `island`, `bill_length_mm`, and so on.
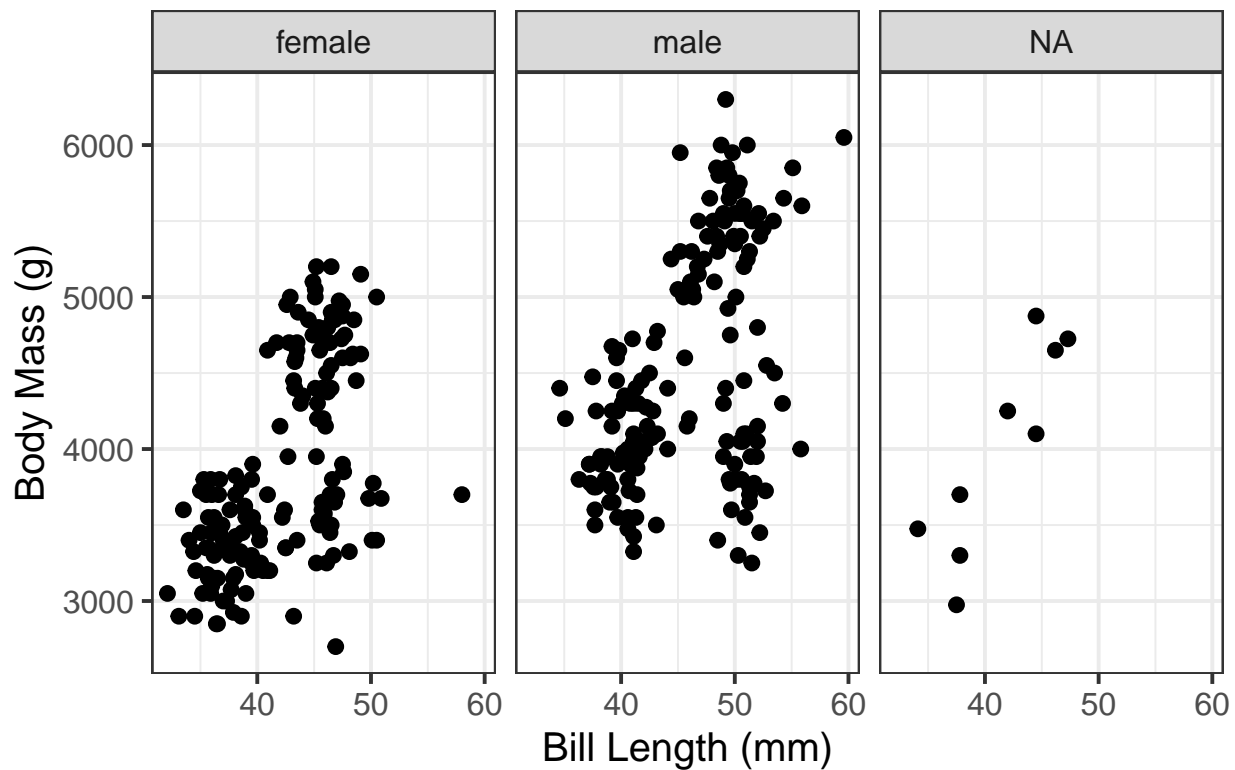
Drawing a plot for association

1. Draw a single plot showing association between the bill lengths and body mass of the penguins according to the their sex (60 pts).

```r
data("penguins",package ="palmerpenguins" )
dataset<-na.omit(penguins)
ggplot(penguins, aes(x = bill_length_mm, y = body_mass_g)) +
   geom_point() +
    labs(title = "Association between Bill Length and Body Mass by Sex",
         x = "Bill Length (mm)",
         y = "Body Mass (g)") +
    facet_grid(~sex) +
   theme_bw(base_size=15)
```

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

Association between Bill Length and Body Mass by

2. Interpret the plot (40 pts). The scatter plot shows the relationship between bill length and body mass. By comparing the two panels, we can also see that male penguins generally have longer bills and higher body mass than female penguins. Both panels show a positive relationship. This means that as bill length increases, body mass also tends to increase.

```r
install.packages("ggplot2")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.5'
## (as 'lib' is unspecified)
```

```r
install.packages("dplyr")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.5'
## (as 'lib' is unspecified)
```

```r
library(ggplot2)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
install.packages("GGally") #correlogram için
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.5'
## (as 'lib' is unspecified)
```

```r
library(GGally)
install.packages("ellipse")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.5'
## (as 'lib' is unspecified)
```

```r
library(ellipse)
```

```
##
## Attaching package: 'ellipse'

## The following object is masked from 'package:graphics':
##
##      pairs
```

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

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.5'
## (as 'lib' is unspecified)
```

```r
library(palmerpenguins)
```

```
##
## Attaching package: 'palmerpenguins'

## The following objects are masked from 'package:datasets':
##
##      penguins, penguins_raw
```

```r
install.packages("ggplot")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.5'
## (as 'lib' is unspecified)
```

```
## Warning: package 'ggplot' is not available for this version of R
##
## A version of this package for your version of R might be available elsewhere,
## see the ideas at
## https://cran.r-project.org/doc/manuals/r-patched/R-admin.html#Installing-packages
```