Data visualisation - solutions

Gbadamassi G.O. Dossa

Updated on 2023-11-12 (created on 2021-09-13)

Acknowledgements

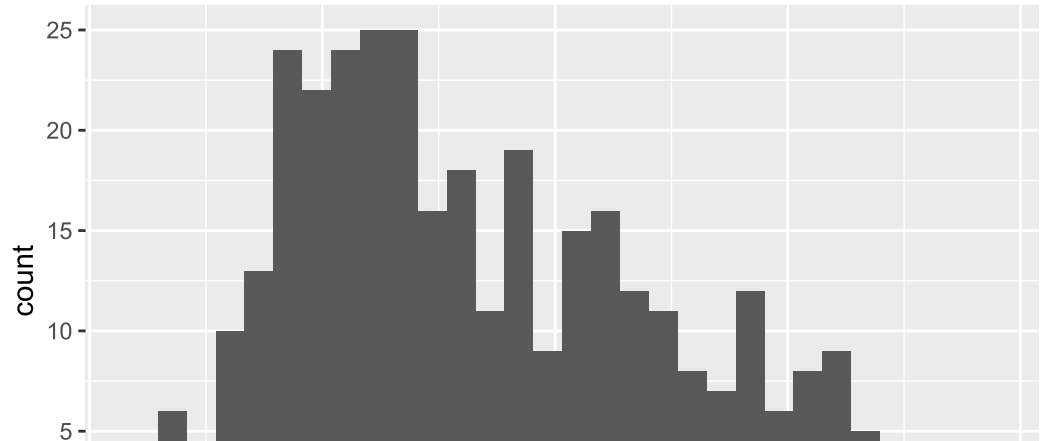
The content of this module are based on materials from:

olivier gimenez's materials

Question 1a: histogram of body mass

a. Build a histogram of body mass.

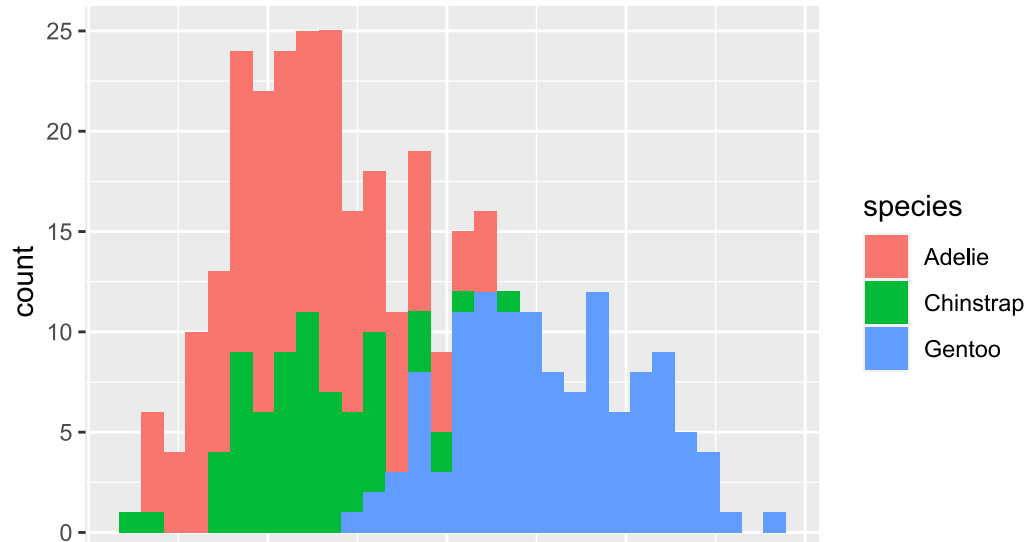
```
library(palmerpenguins)
library(tidyverse) # instead reading or load tidyverse you can load direct
penguins %>%
  ggplot() +
  aes(x = body_mass_g) +
  geom_histogram()
```



Question 1b: a color per species

b. Fill in histogram with a color per species.

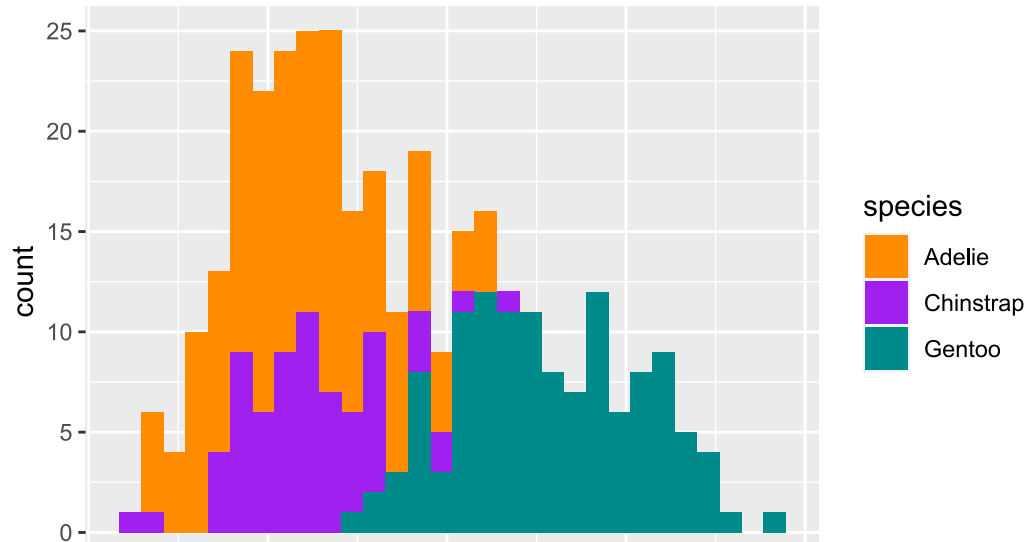
```
dossa1<-penguins %>%
  ggplot() +
  aes(x = body_mass_g) +
  geom_histogram(aes(fill = species))
dossa1
```



Question 1c: darkorange, purple and cyan4

c. Change the color by default to darkorange, purple and cyan4.

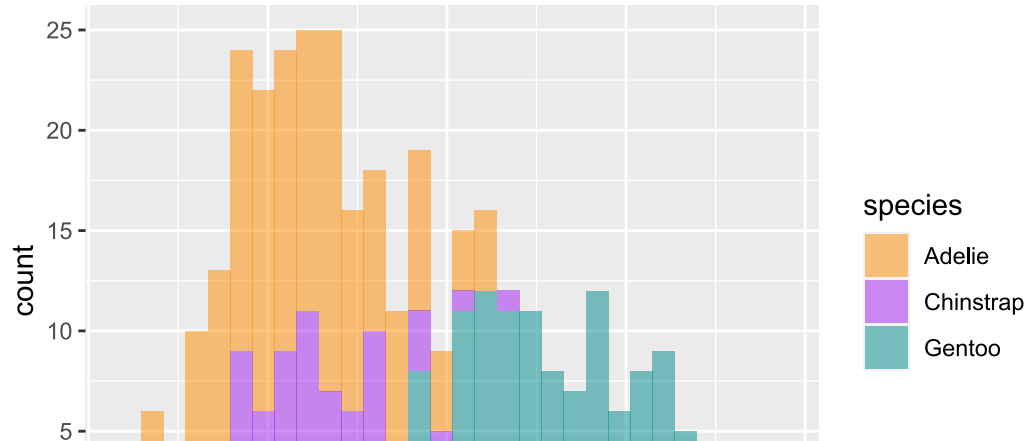
```
penguins %>%
  ggplot() +
  aes(x = body_mass_g) +
  geom_histogram(aes(fill = species)) +
  scale_fill_manual(values = c("darkorange","purple","cyan4"))
```



Question 1d: adjust transparency

d. Adjust transparency for all three histograms.

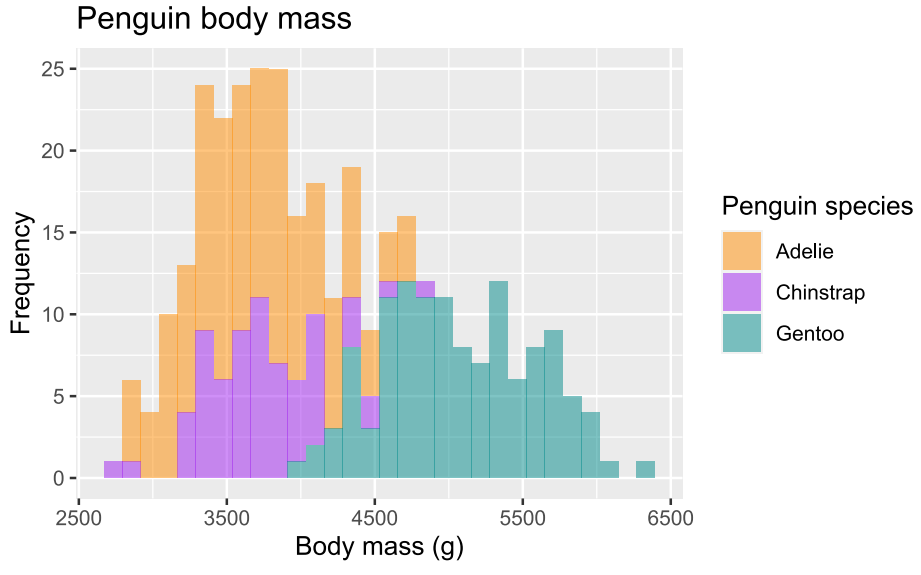
```
penguins %>%
  ggplot() +
  aes(x = body_mass_g) +
  geom_histogram(aes(fill = species), alpha = 0.5) +
  scale_fill_manual(values = c("darkorange","purple","cyan4"))
# Note that transparency does not work on histogram the way we expected on
```



Question 1e: add titles

e. Add a title to the axes, the legend and the figure.

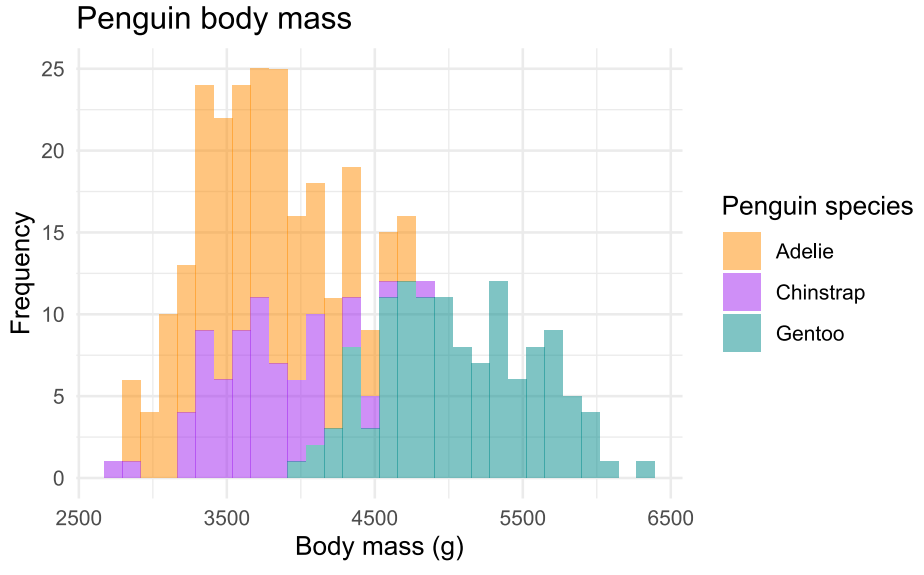
```
penguins %>%
  ggplot() +
  aes(x = body_mass_g) +
  geom_histogram(aes(fill = species),
                 alpha = 0.5) +
  scale_fill_manual(values = c("darkorange",
                               "purple",
                               "cyan4")) +
  labs(x = "Body mass (g)",
       y = "Frequency",
       title = "Penguin body mass",
       fill = "Penguin species")
```



Question 1f: change theme

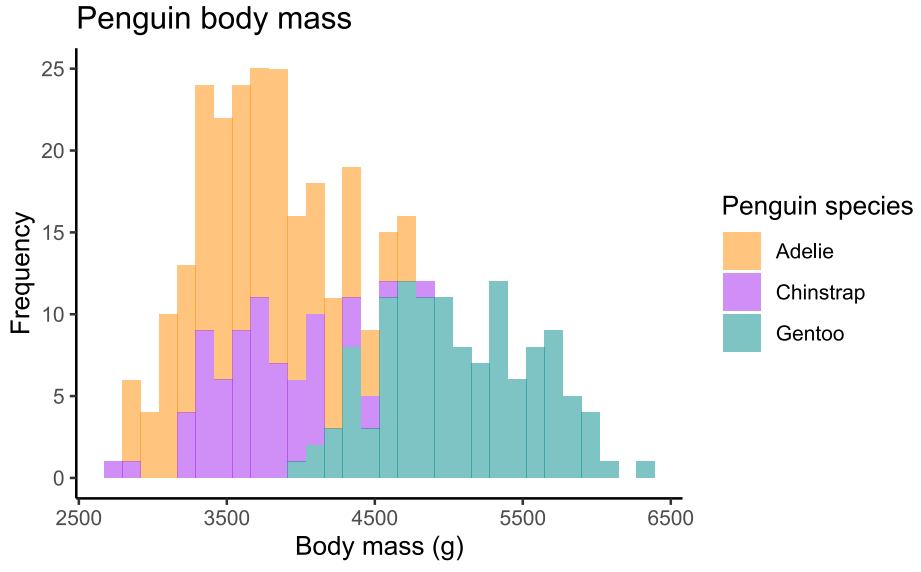
f. Change the theme used by default.

```
penguins %>%
  ggplot() +
  aes(x = body_mass_g) +
  geom_histogram(aes(fill = species),
                 alpha = 0.5) +
  scale_fill_manual(values = c("darkorange",
                               "purple",
                               "cyan4")) +
  labs(x = "Body mass (g)",
       y = "Frequency",
       title = "Penguin body mass",
       fill = "Penguin species") +
  theme_minimal() # here I used theme_minimal
```



f. Change the theme used by default.

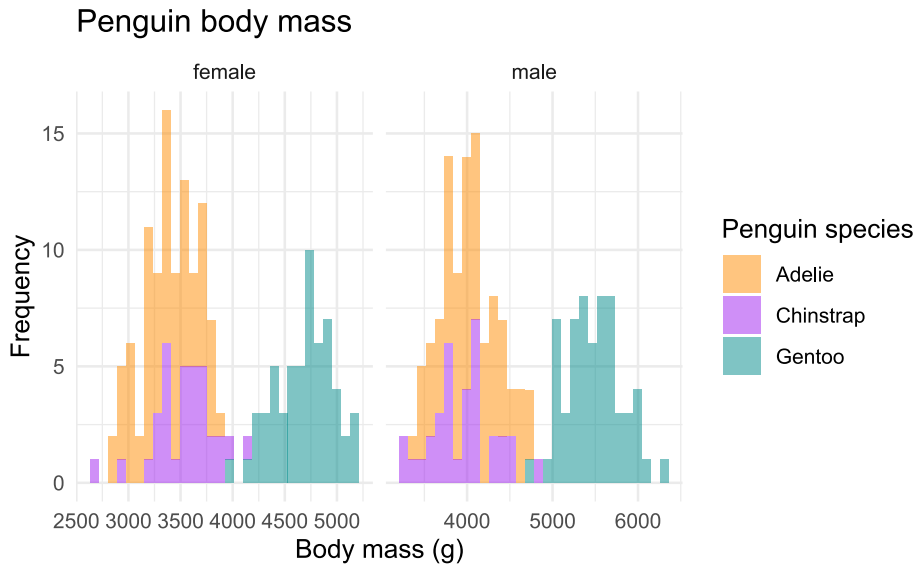
```
penguins %>%
  ggplot() +
  aes(x = body_mass_g) +
  geom_histogram(aes(fill = species),
                 alpha = 0.5) +
  scale_fill_manual(values = c("darkorange",
                               "purple",
                               "cyan4")) +
  labs(x = "Body mass (g)",
       y = "Frequency",
       title = "Penguin body mass",
       fill = "Penguin species") +
  theme_classic()
```



Question 1g: split by sex

g. Do a-f for males and females and display on same figure.

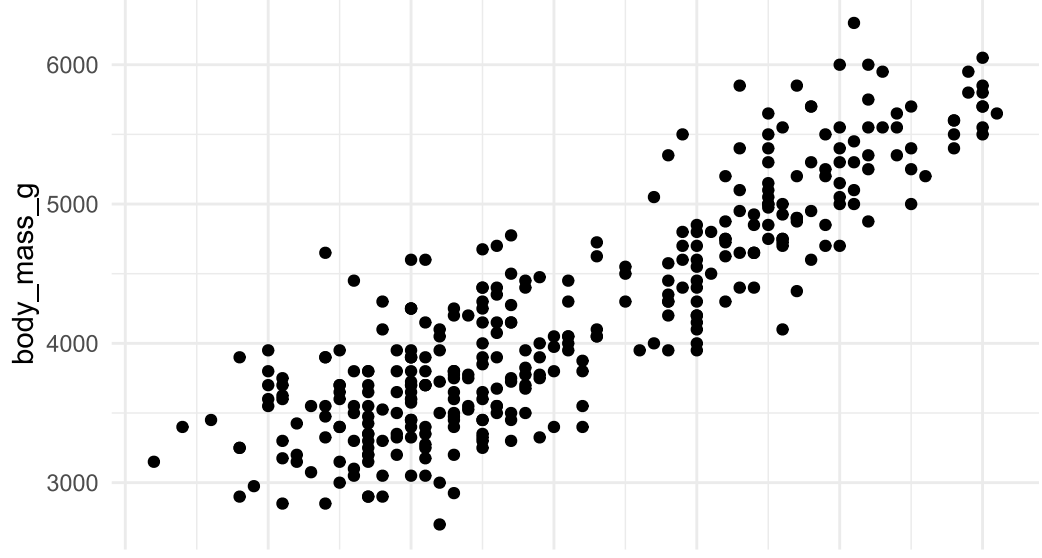
```
penguins %>%
# filter out penguins w/ missing sex
  filter(!is.na(sex)) %>%
  ggplot() +
  aes(x = body_mass_g) +
  geom_histogram(aes(fill = species),
# fill histogram, adjust transparency
                 alpha = 0.5) +
# change colour
  scale_fill_manual(values = c("darkorange",
                               "purple",
                               "cyan4")) +
  labs(x = "Body mass (g)", # x lab
       y = "Frequency", # y lab
       title = "Penguin body mass",
       fill = "Penguin species") +
  theme_minimal() + # change theme
# hist by sex, w/ diff X scale
  facet_wrap(~sex, scales = "free_x")
```



Question 2a: scatter plot

a. Build a scatter plot of body mass with respect to flipper length.

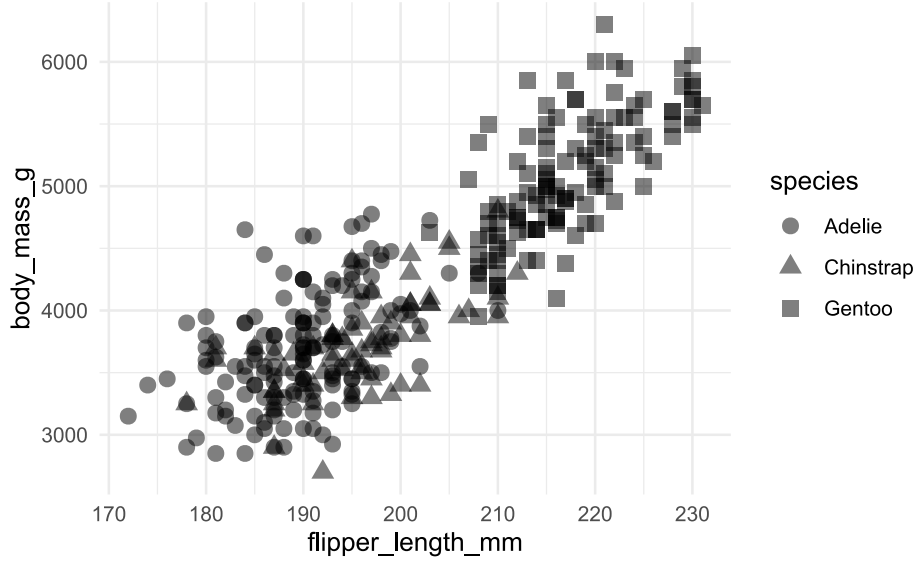
```
penguins %>%
  ggplot() +
  aes(x = flipper_length_mm,
      y = body_mass_g) +
  geom_point() +
  theme_minimal()
```



Question 2b: species-specific shapes

b. Consider a shape for the points different for each species. Increase point size and adjust transparency.

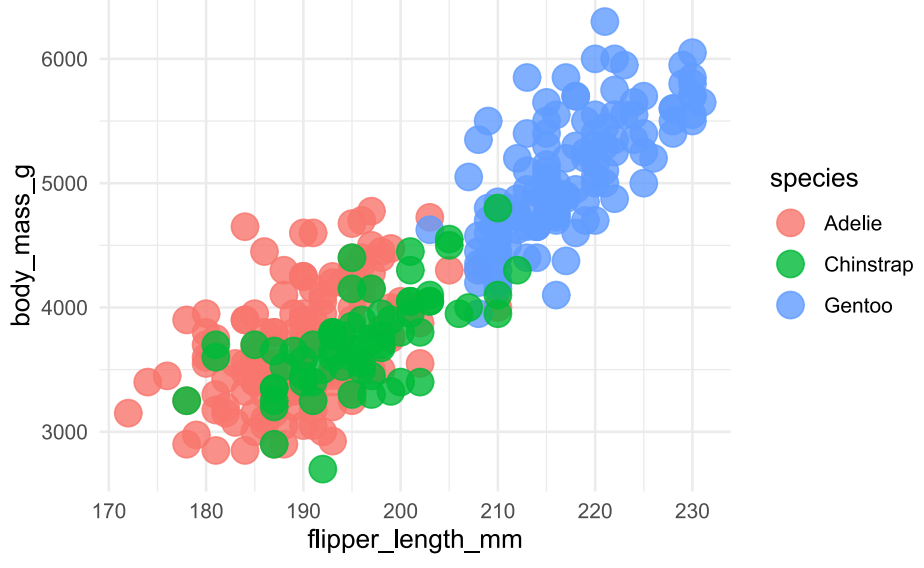
```
penguins %>%
  ggplot() +
  aes(x = flipper_length_mm,
      y = body_mass_g) +
  geom_point(aes(shape = species),
             size = 3,
             alpha = 0.5) +
  theme_minimal()
```



Question 2c: species-specific colors

c. Consider a color for the points different for each species. Increase point size and adjust transparency.

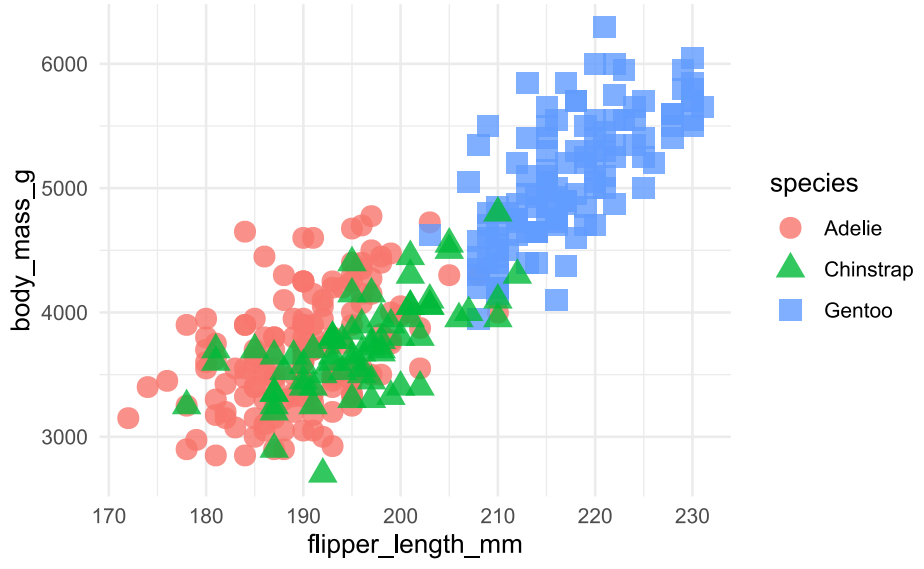
```
penguins %>%
  ggplot() +
  aes(x = flipper_length_mm,
      y = body_mass_g) +
  geom_point(aes(color = species),
             size = 5,
             alpha = 0.8) +
  theme_minimal()
```



Question 2d: species-specific shapes and colors

d. Combine b-c.

```
penguins %>%
  ggplot() +
  aes(x = flipper_length_mm,
      y = body_mass_g) +
  geom_point(aes(color = species,
                 shape = species),
             size = 4,
             alpha = 0.8) +
  theme_minimal()
```

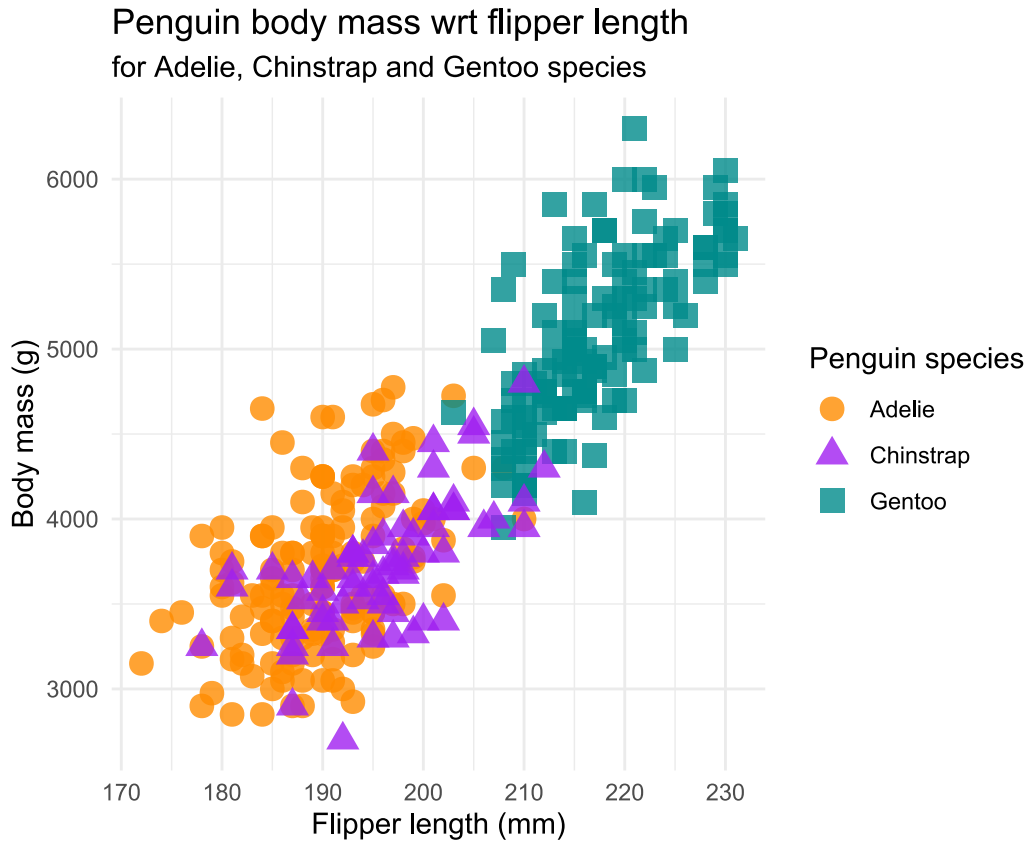


Question 2e: add titles

e. Change colors to darkorange, purple and cyan4. Add titles to axes, legend and figure.

```
penguins %>%
  ggplot() +
  aes(x = flipper_length_mm,
      y = body_mass_g) +
  geom_point(aes(color = species,
                 shape = species),
                 size= 4,
                 alpha = 0.8) +
  scale_color_manual(values = c("darkorange","purple","cyan4")) +
  labs(title = "Penguin body mass wrt flipper length",
       subtitle = "for Adelie, Chinstrap and Gentoo species",
       x = "Flipper length (mm)",
       y = "Body mass (g)",
       color = "Penguin species",
       shape = "Penguin species") +
  theme_minimal()
```

Question 2e


Penguin body mass wrt flipper length
for Adelie, Chinstrap and Gentoo species
Body mass (g)
6000
5000
4000
3000
170
180
190
200
210
220
230
Flipper length (mm)
Penguin species
Adelie
Chinstrap
Gentoo