Introduction to ggplot2

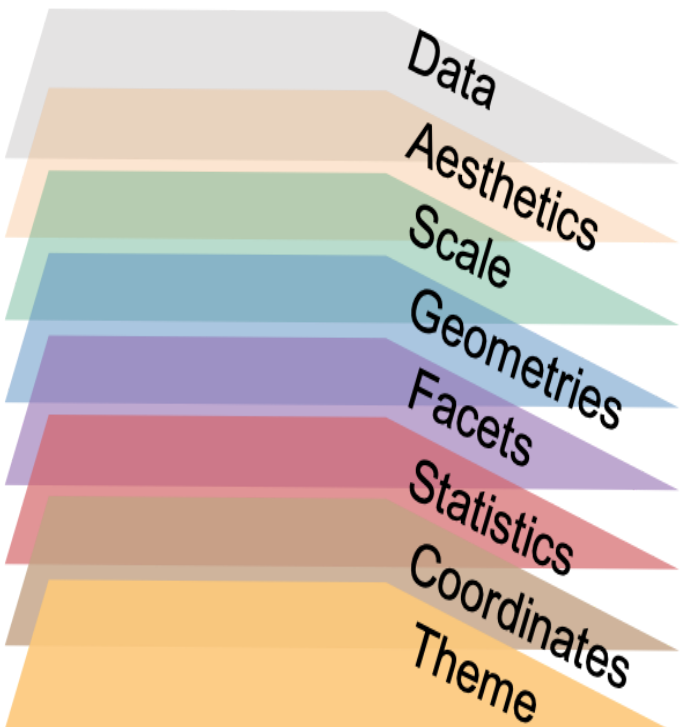
Sian Bladon

What is ggplot2?

ggplot2 is one of the *tidyverse* packages

The ‘gg’ stands for Grammar of Graphics, an approach to building plots using a combination of layers

Image from here



How to create a plot with ggplot2

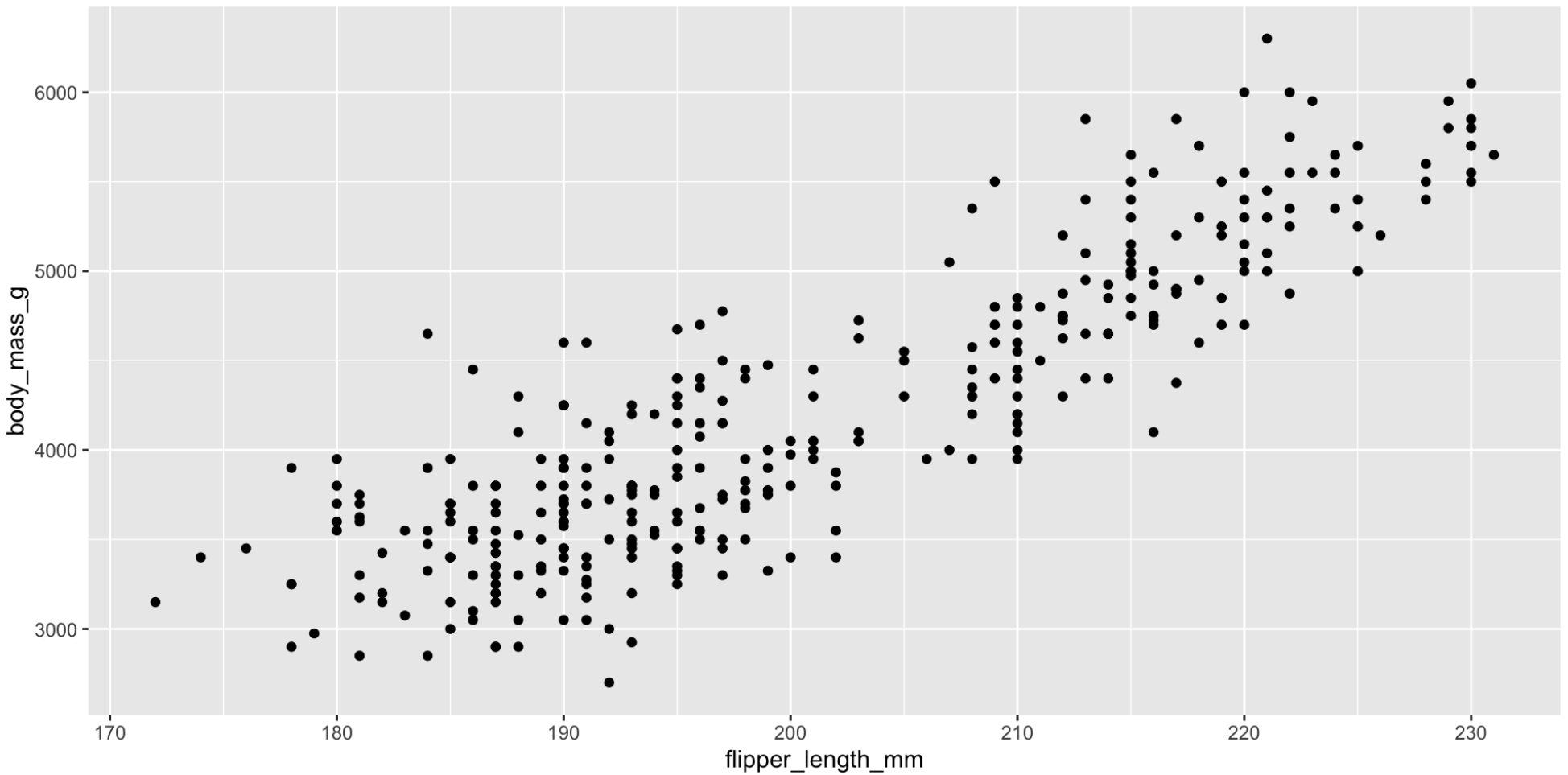
First, load your data.
We will use the penguins data set from the *palmerpenguins* package by Alison Horst

```
1 library (tidyverse)
2 library(palmerpenguins)
3 data(package = 'palmerpenguins')
```

If you have not already installed either of these packacges then you will need to `install.packages("palmerpenguins')` first

How to create a plot with ggplot2

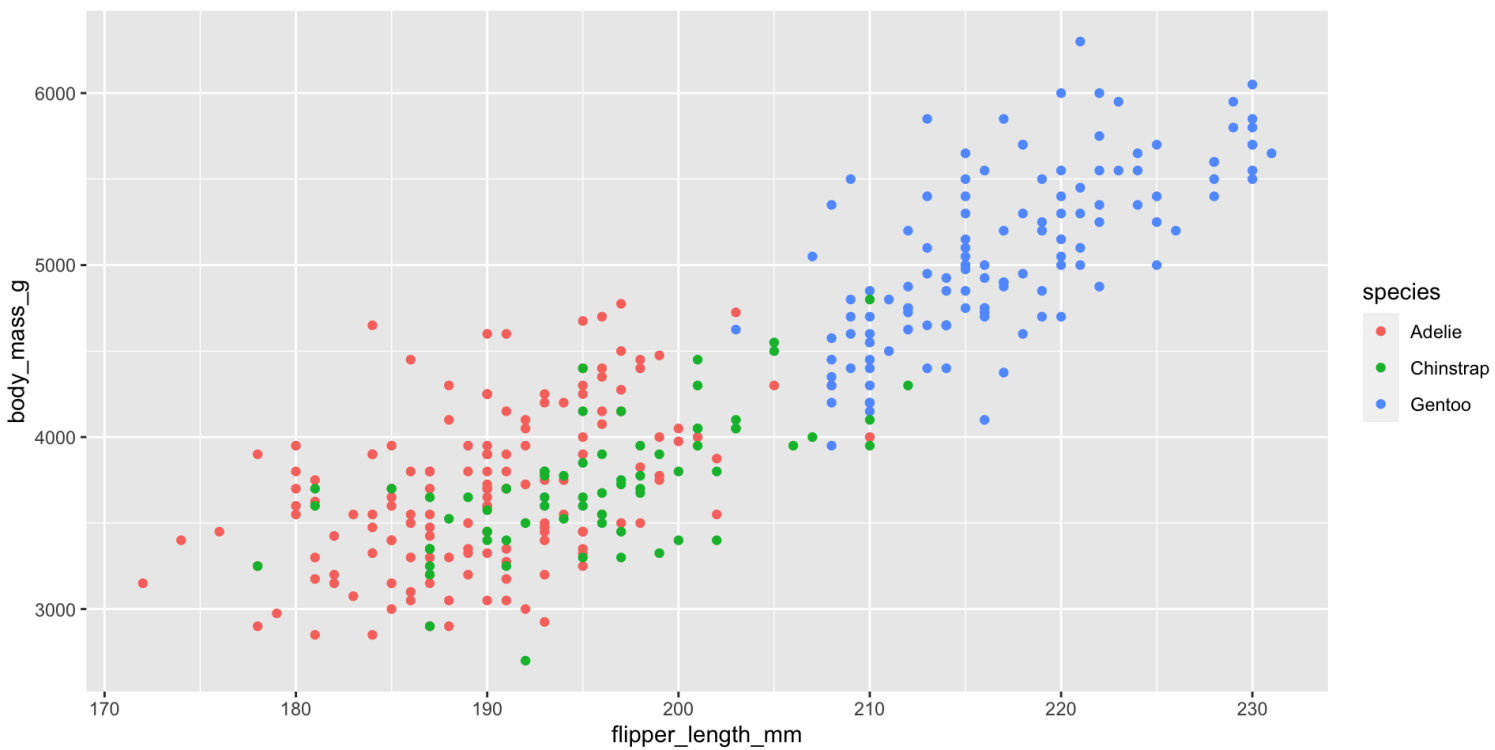
```
1 ggplot(data = penguins) +
2   geom_point(mapping = aes(x = flipper_length_mm, y = body_mass_g))
```



You can also map additional variables to aesthetics, including:

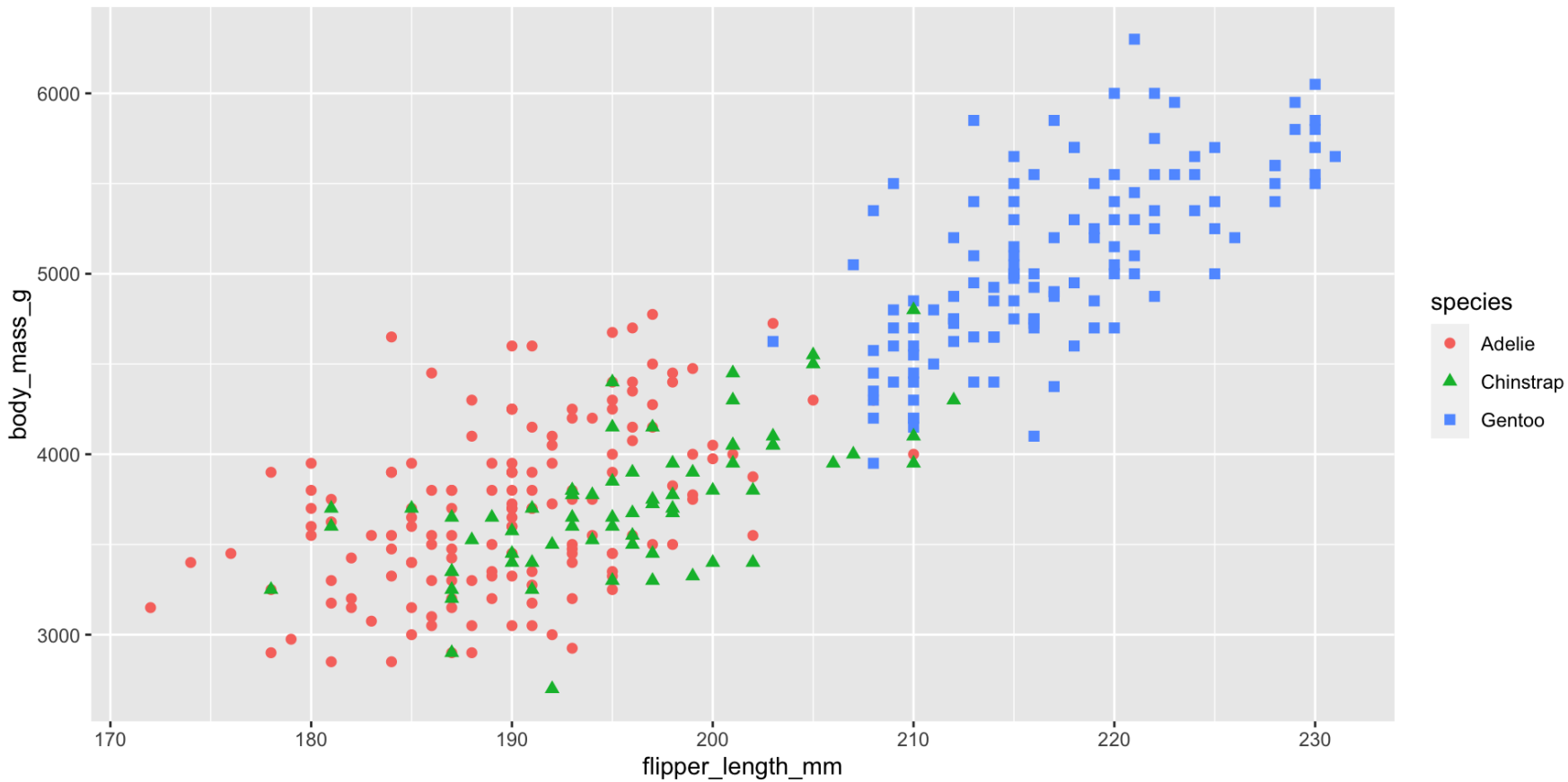
- colour

```
1  ggplot(data = penguins) +
2    geom_point(mapping = aes(x = flipper_length_mm, y = body_mass_g,
3                             colour = species))
```



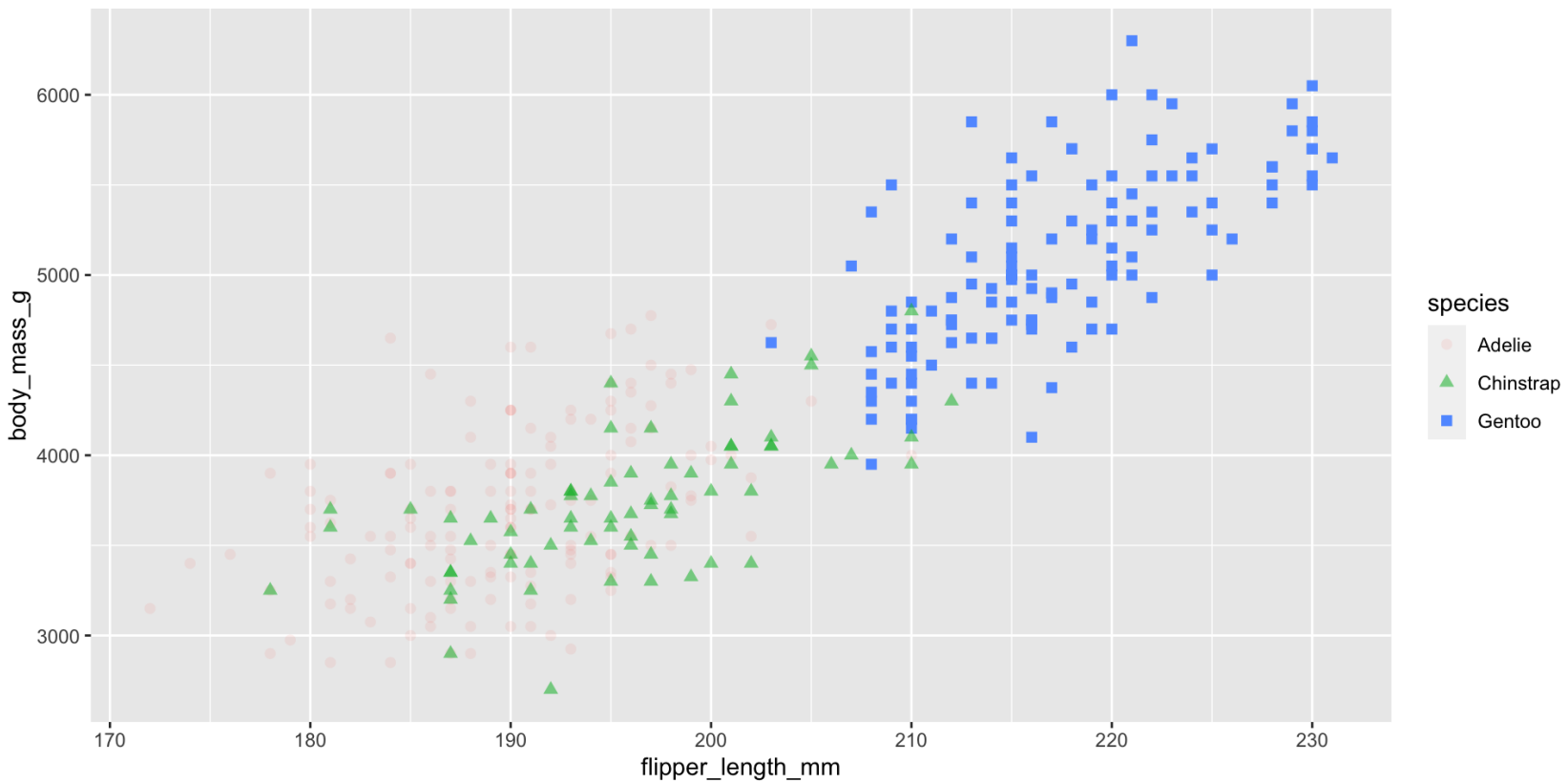
- shape

```
1  ggplot(data = penguins) +
2    geom_point(mapping = aes(x = flipper_length_mm, y = body_mass_g,
3                             colour = species, shape = species), size = 2)
```



- alpha (transparency)

```
ggplot(data = penguins) +
  geom_point(mapping = aes(x = flipper_length_mm, y = body_mass_g,
                           colour = species, shape = species, alpha = speci
             , size = 2)
```

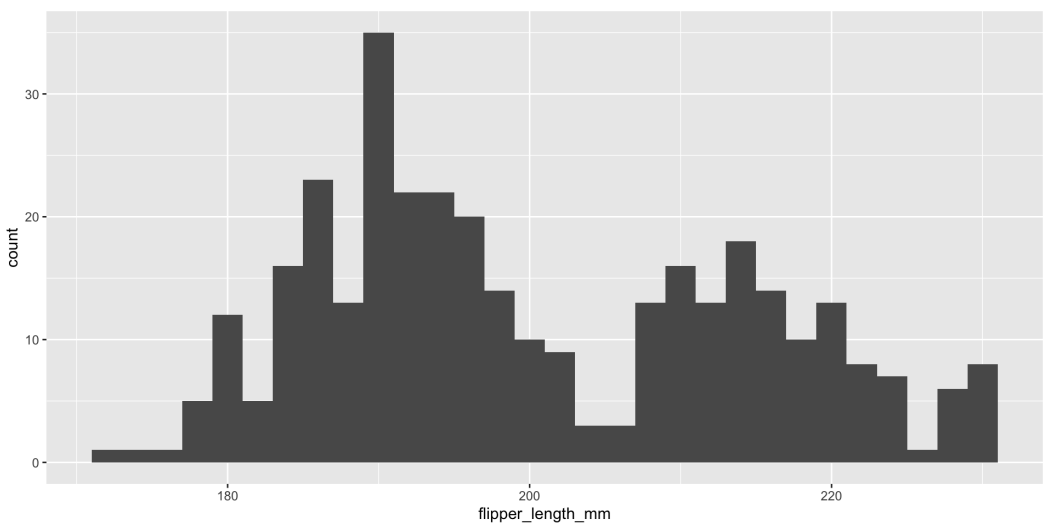


Geoms

There are many types of `geom`, here are a few examples:

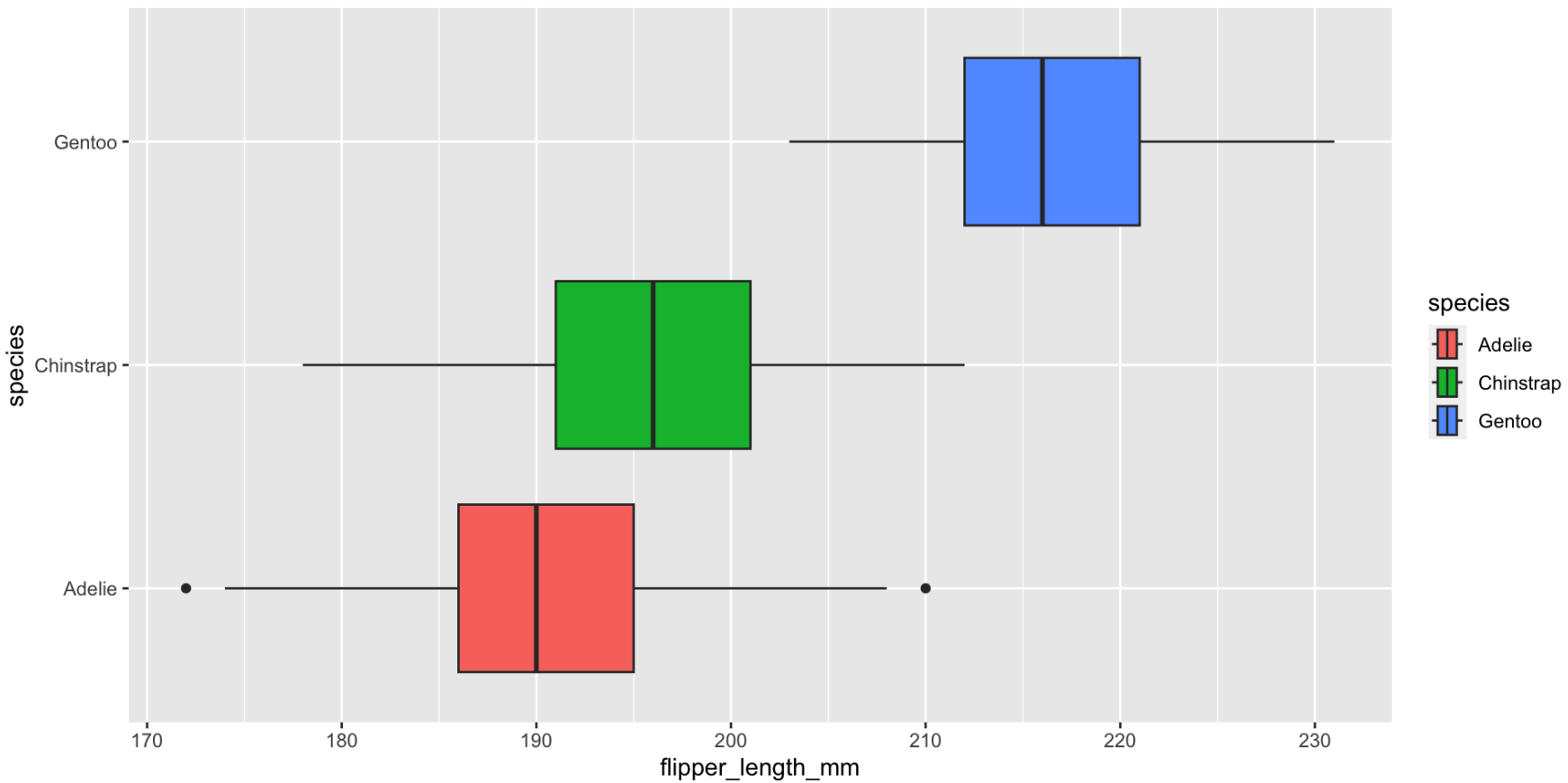
- `geom_histogram`

```
1  ggplot(data = penguins) +
2    geom_histogram(mapping = aes(x = flipper_length_mm), binwidth = 2)
```



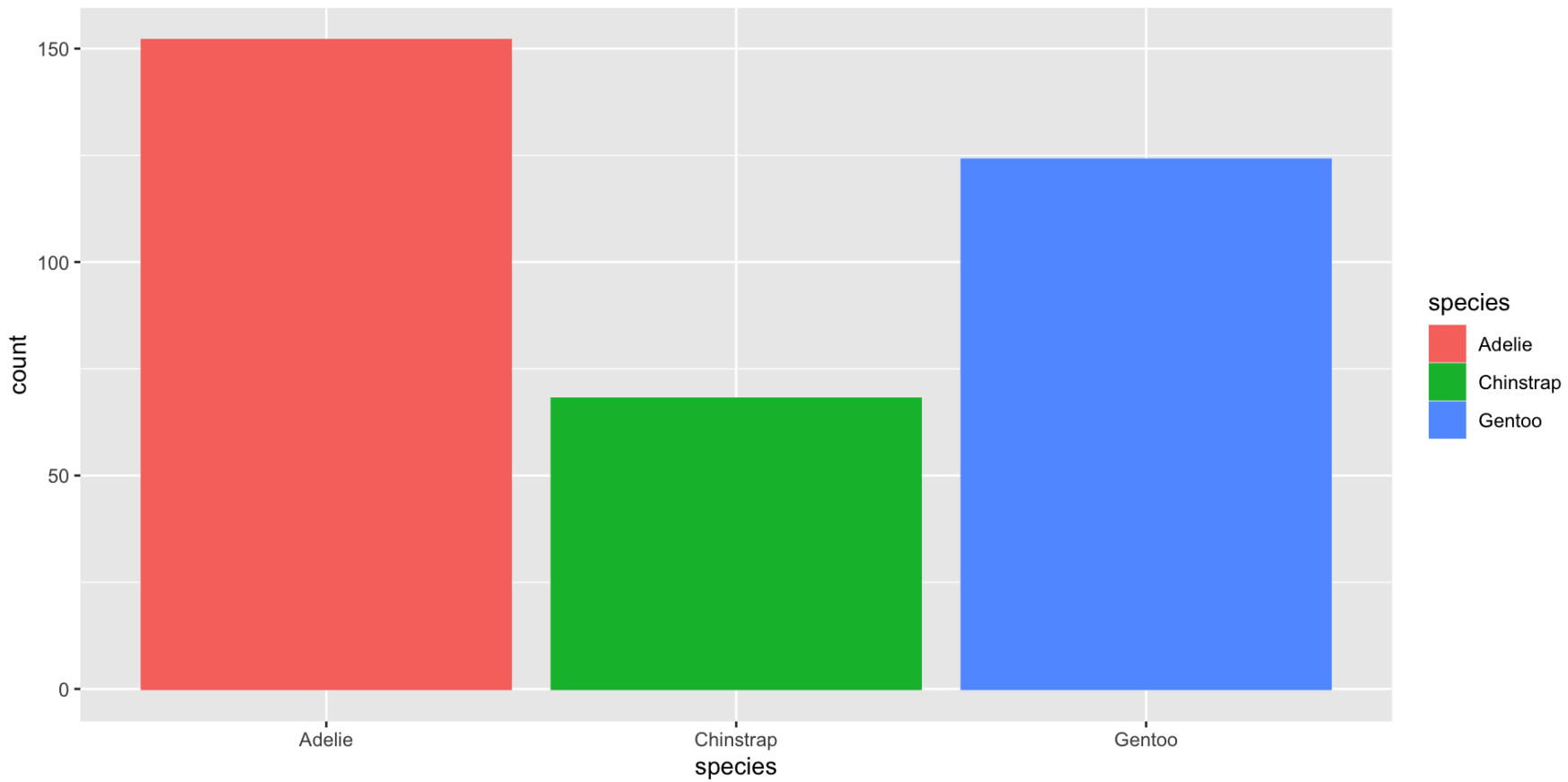
- geom_boxplot

```
ggplot(data = penguins) +
  geom_boxplot(mapping =
                 aes(x = flipper_length_mm, y = species, fill = species))
```



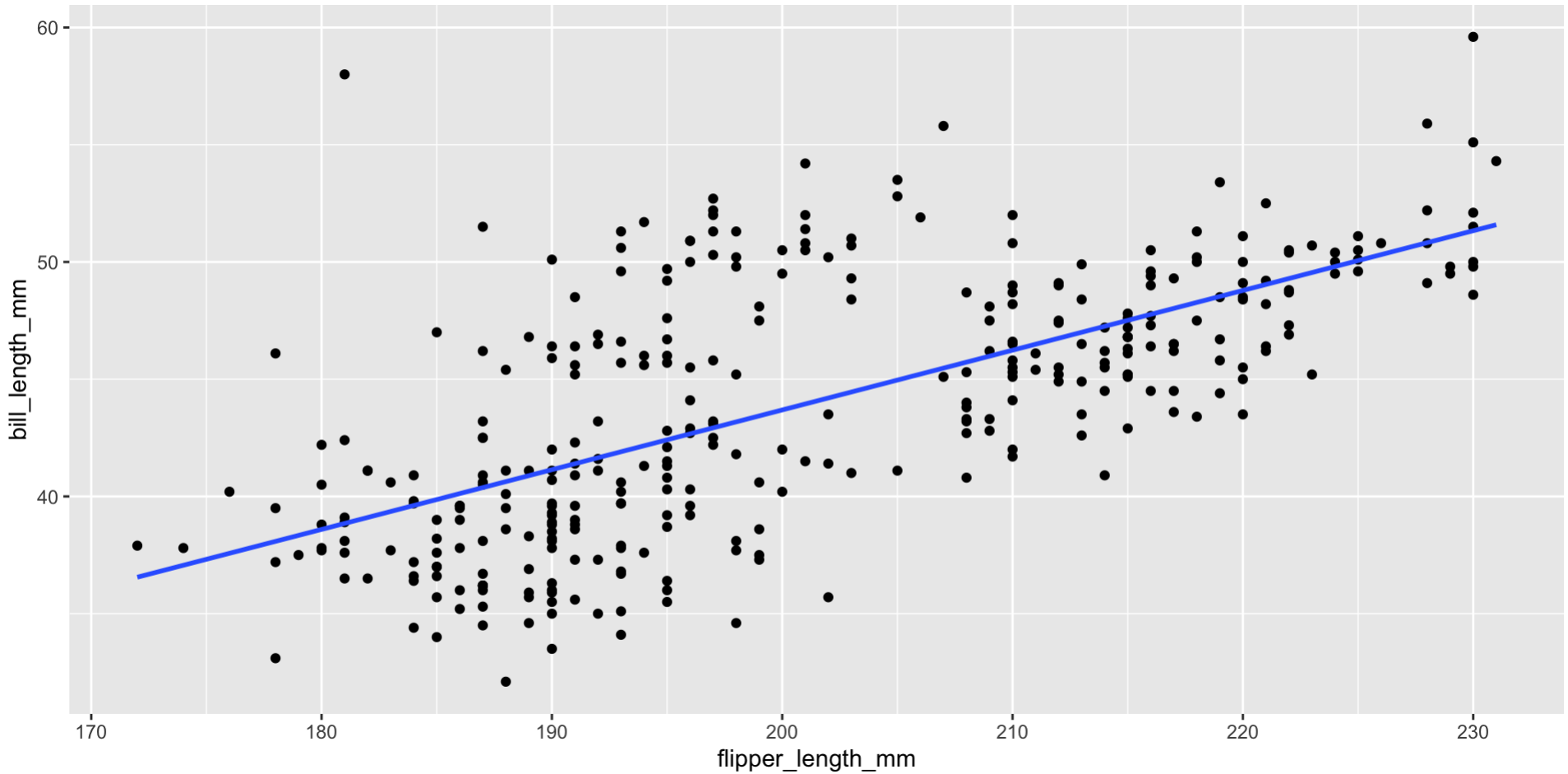
- geom_bar

```
1  ggplot(data = penguins) +
2    geom_bar(mapping =
3               aes(x = species, colour = species, fill = species))
```



You can also layer different `geoms`

```
ggplot(data = penguins) +
  geom_point(mapping = aes(x = flipper_length_mm, y = bill_length_mm)) +
  geom_smooth(mapping = aes(x = flipper_length_mm, y = bill_length_mm),
              method = "lm", se = FALSE)
```


60
50
40
bill_length_mm
170
180
190
200
210
220
230
flipper_length_mm

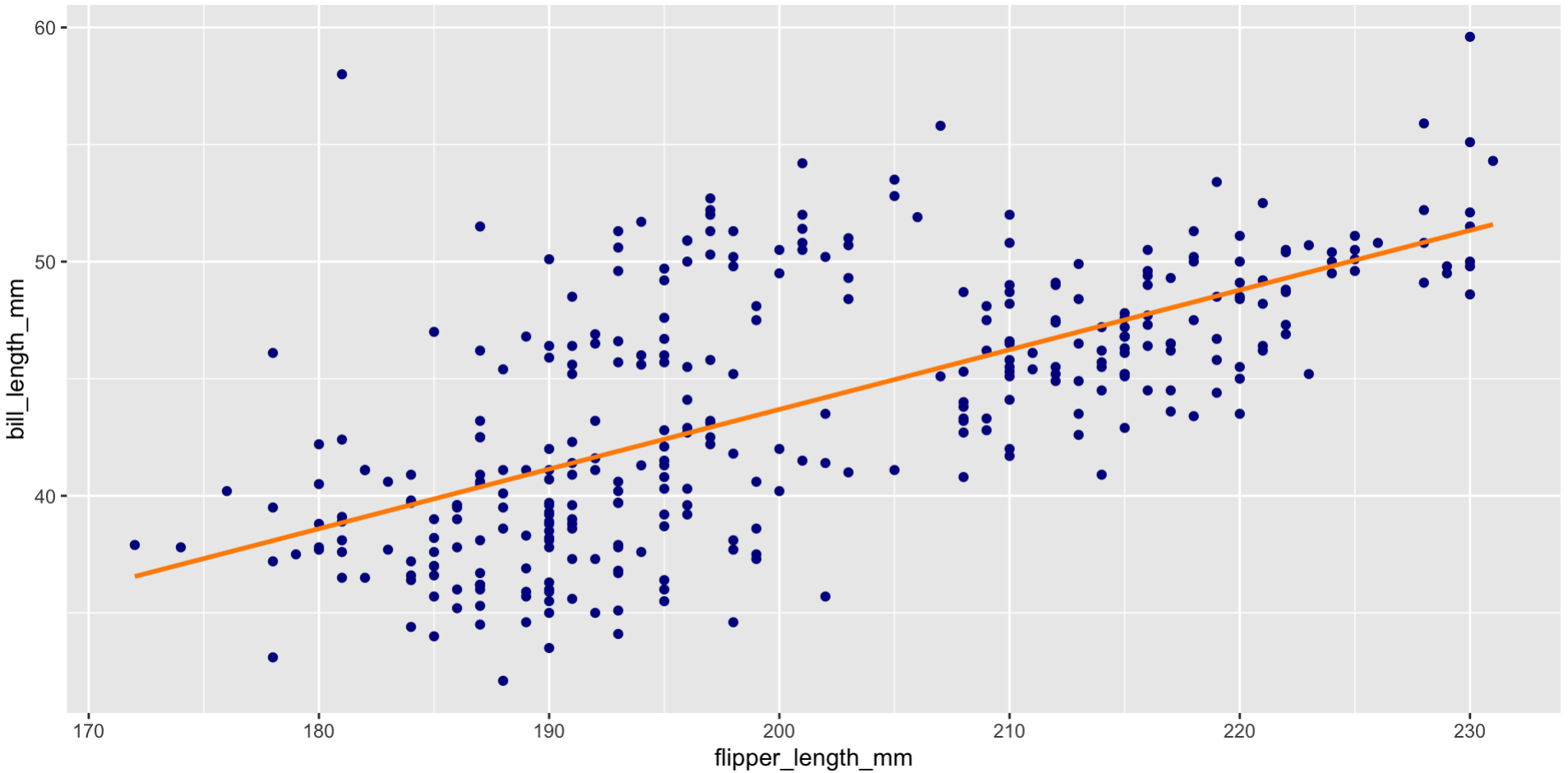
Customising plots

There are many ways to customise plots.

- changing colour

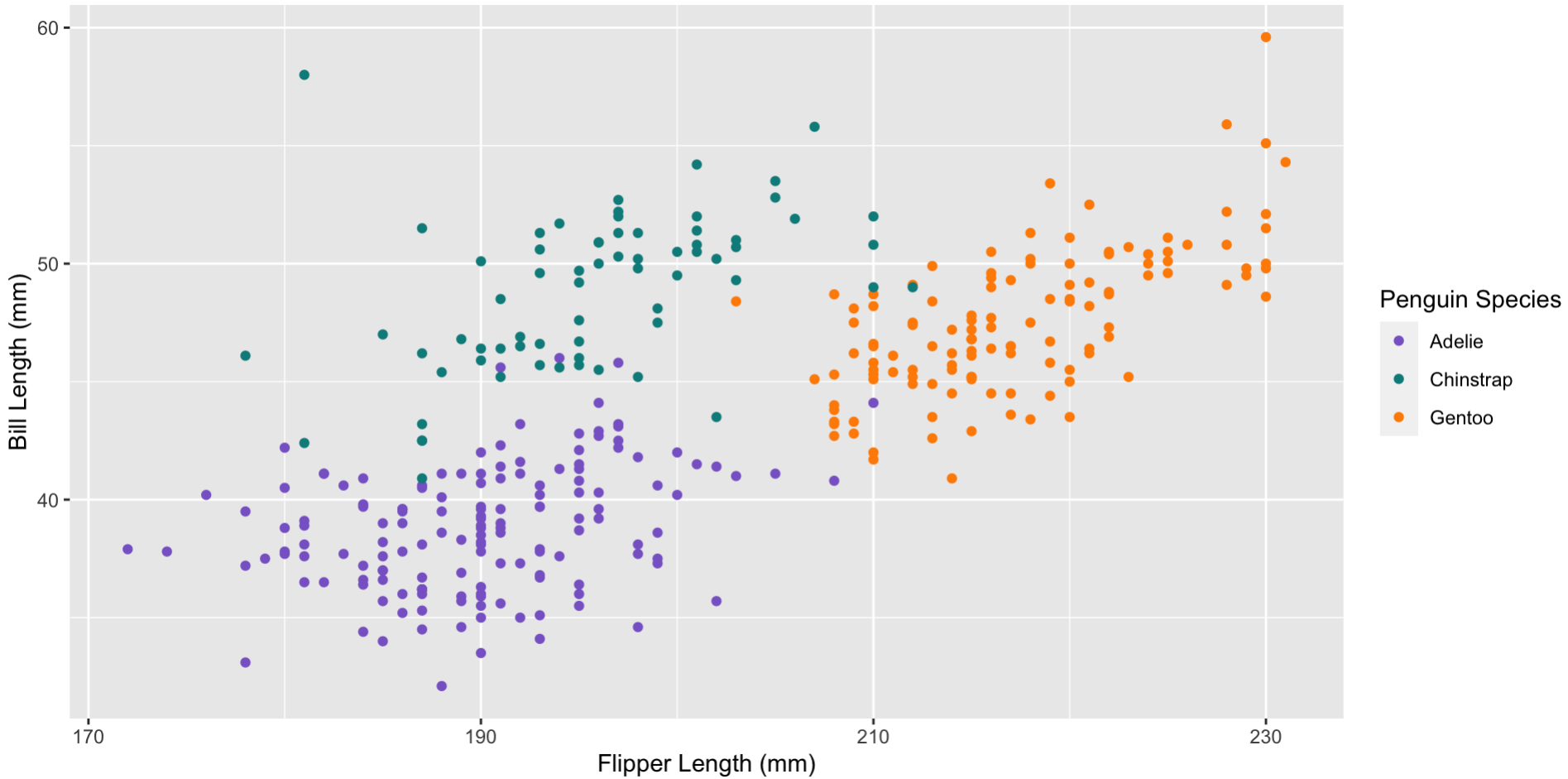
```
ggplot(data = penguins) +
  geom_point(mapping = aes(x = flipper_length_mm, y = bill_length_mm),
             colour = "darkblue") +
  geom_smooth(mapping = aes(x = flipper_length_mm, y = bill_length_mm),
              method = "lm", se = FALSE, colour = "darkorange")
```

Customising plots



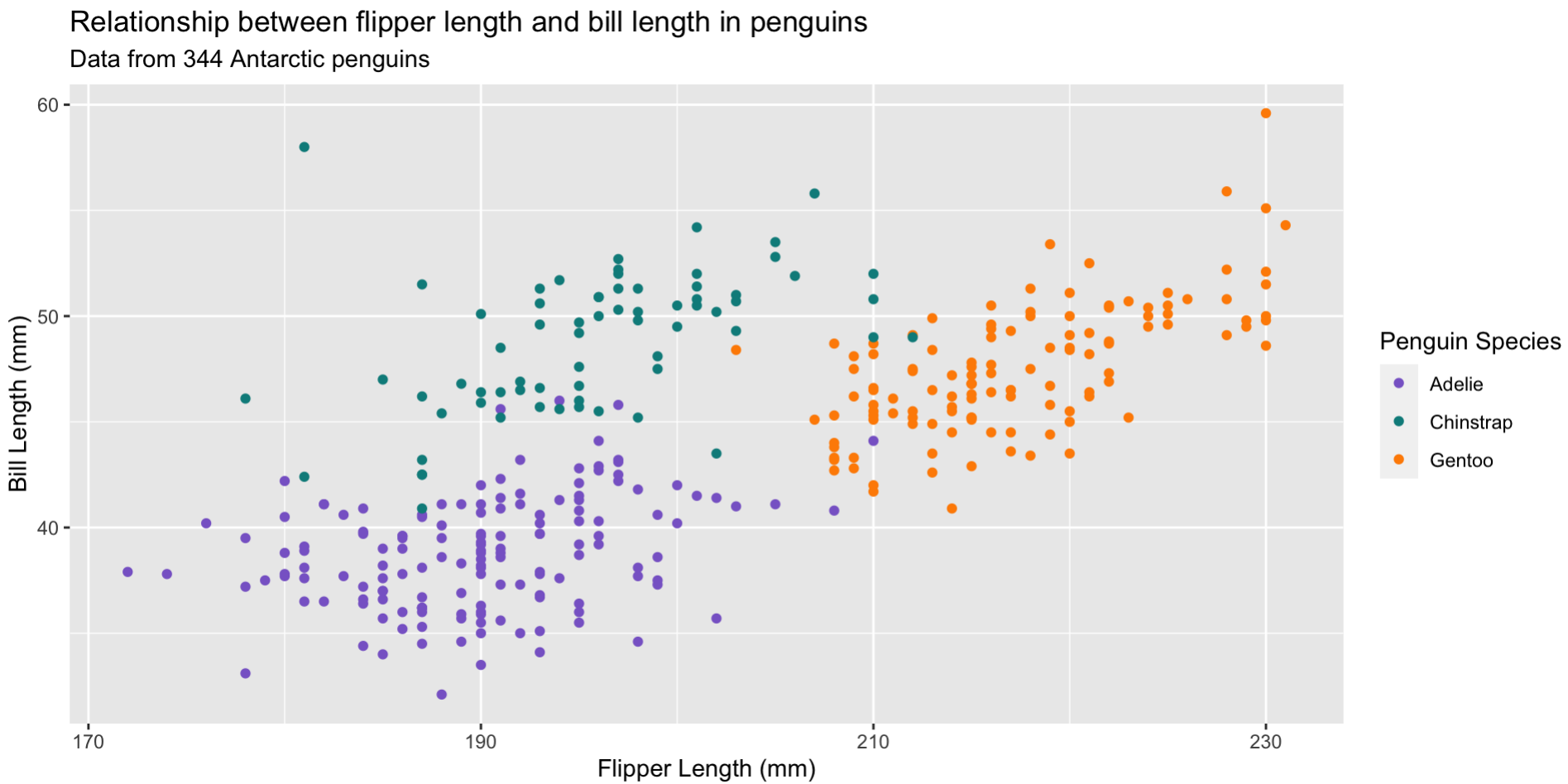
- scale

```
ggplot(data = penguins) +
  geom_point(mapping = aes(x = flipper_length_mm, y = bill_length_mm,
                           colour = species)) +
  scale_x_continuous(name = "Flipper Length (mm)",
                     breaks = c(170, 190, 210, 230)) +
  scale_y_continuous(name = "Bill Length (mm)") +
  scale_color_manual(name = "Penguin Species",
                     values = c("mediumpurple3", "darkcyan", "darkorange"))
```


60
50
40
Bill Length (mm)
170
190
210
230
Flipper Length (mm)
Penguin Species
Adelie
Chinstrap
Gentoo

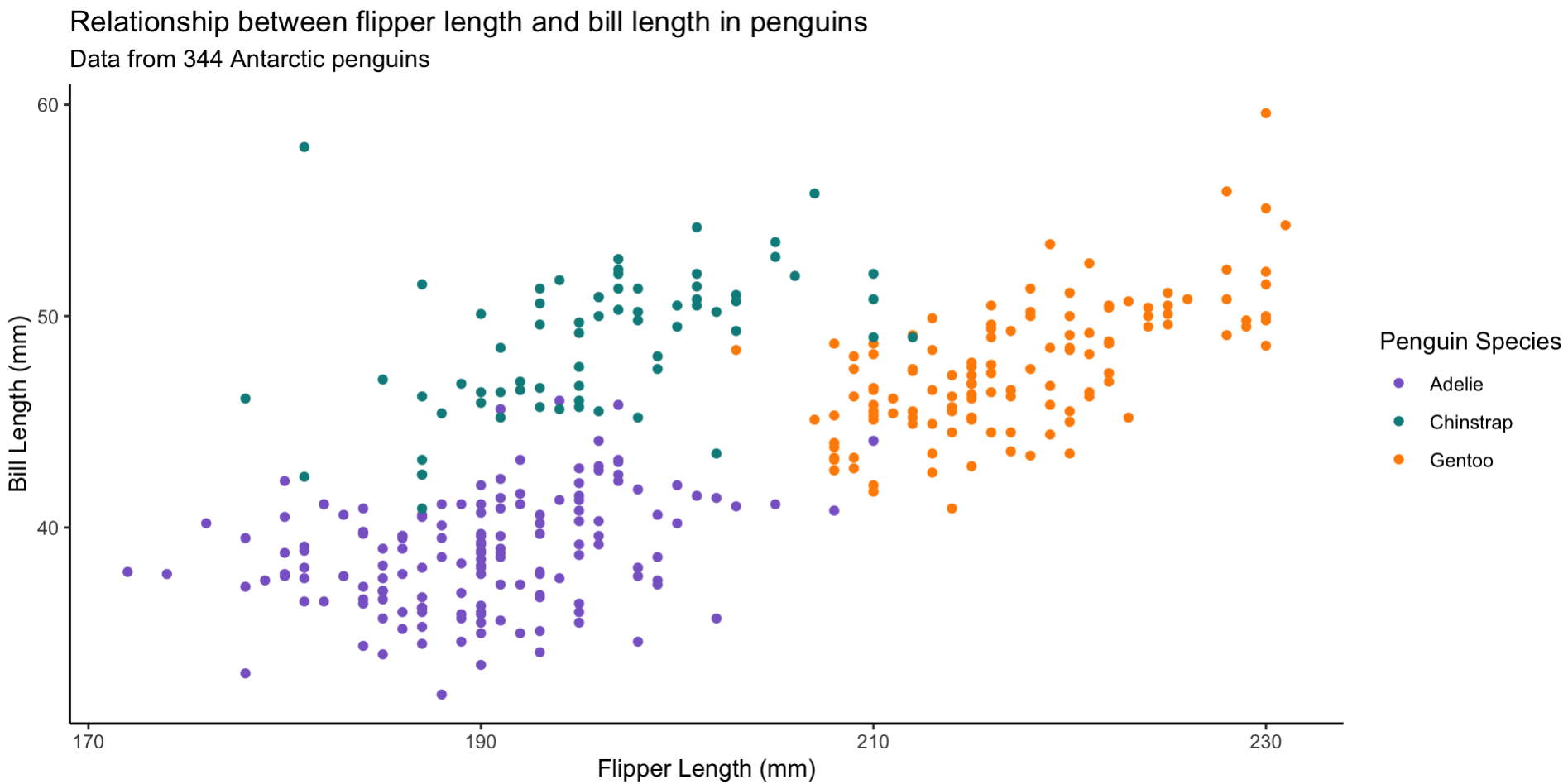
- add a title and subtitle

```
ggplot(data = penguins) +
  geom_point(mapping = aes(x = flipper_length_mm, y = bill_length_mm,
                           colour = species)) +
  scale_x_continuous(name = "Flipper Length (mm)",
                     breaks = c(170, 190, 210, 230)) +
  scale_y_continuous(name = "Bill Length (mm)") +
  scale_color_manual(name = "Penguin Species", values =
                       c("mediumpurple3", "darkcyan", "darkorange")) +
  labs(title = "Relationship between flipper length and bill length in peng
       subtitle = "Data from 344 Antarctic penguins")
```

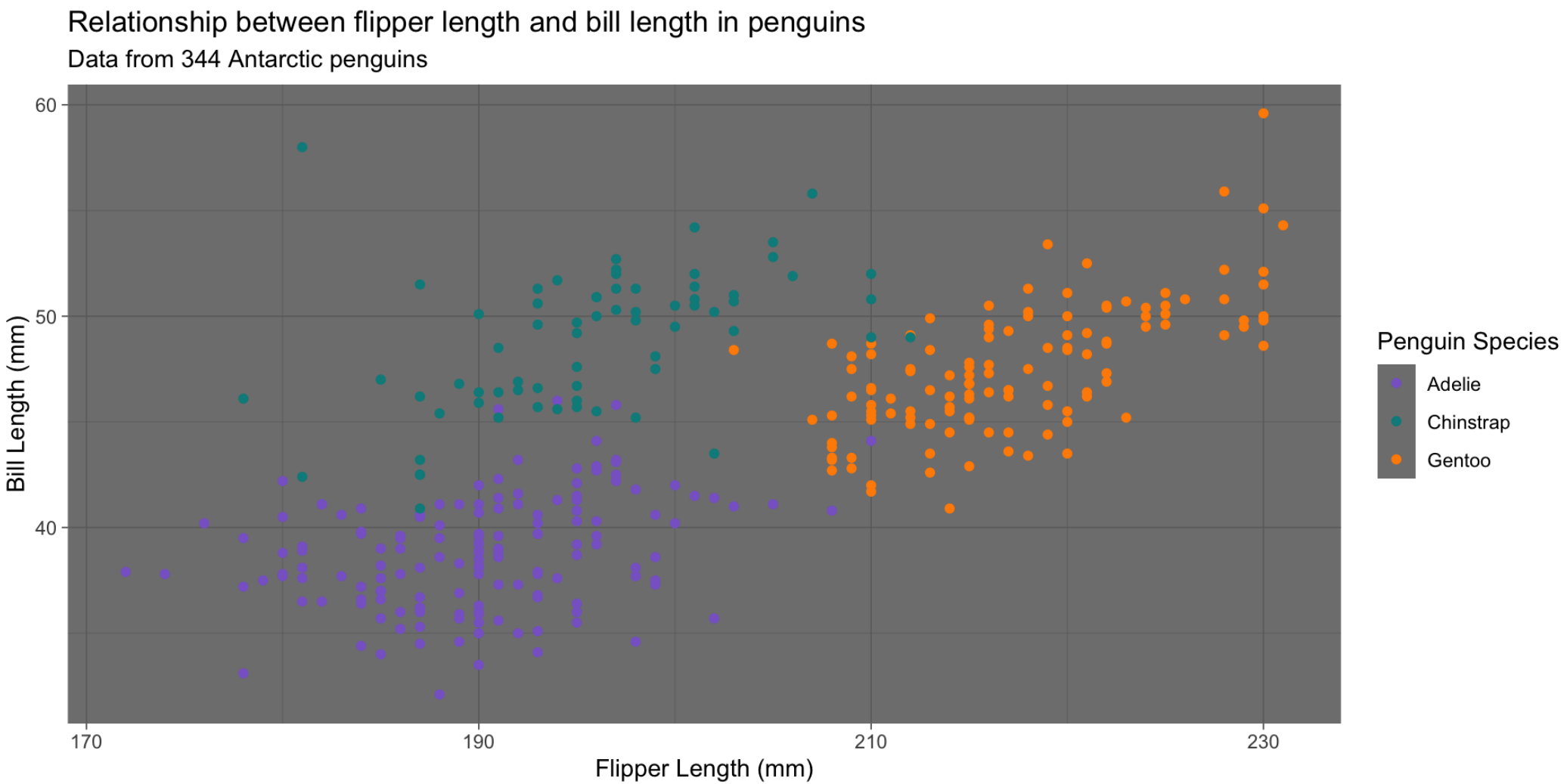

Relationship between flipper length and bill length in penguins
Data from 344 Antarctic penguins
Bill Length (mm)
60
50
40
170
190
210
230
Flipper Length (mm)
Penguin Species
Adelie
Chinstrap
Gentoo

- change the theme - `theme_classic`

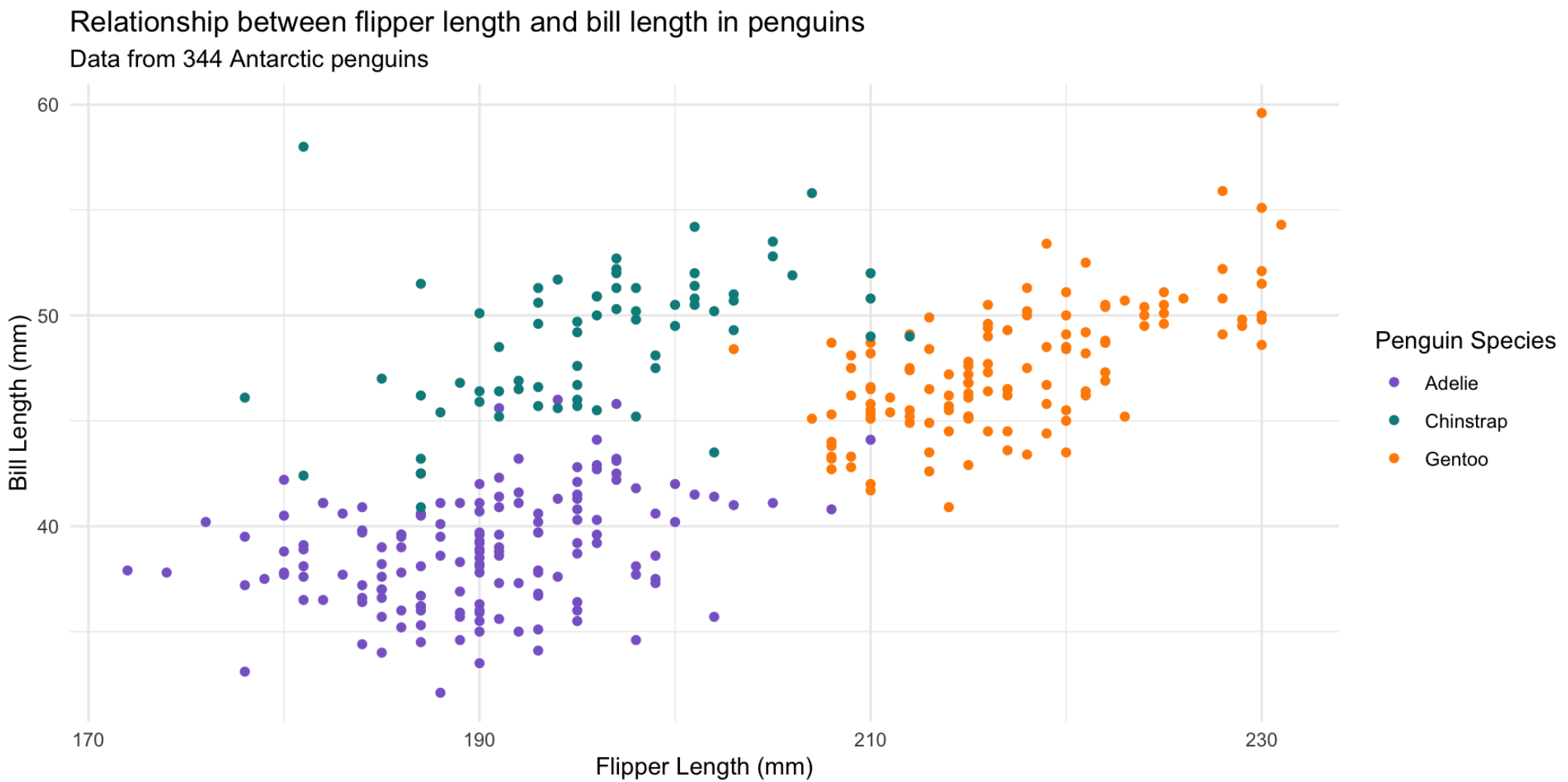
```
ggplot(data = penguins) +
  geom_point(mapping = aes(x = flipper_length_mm, y = bill_length_mm,
                           colour = species)) +
  scale_x_continuous(name = "Flipper Length (mm)",
                     breaks = c(170, 190, 210, 230)) +
  scale_y_continuous(name = "Bill Length (mm)") +
  scale_color_manual(name = "Penguin Species",
                     values = c("mediumpurple3", "darkcyan", "darkorange")) +
  labs(title = "Relationship between flipper length and bill length in peng
       subtitle = "Data from 344 Antarctic penguins") +
  theme_classic()
```


Relationship between flipper length and bill length in penguins
Data from 344 Antarctic penguins
Bill Length (mm)
60
50
40
170
190
210
230
Flipper Length (mm)
Penguin Species
Adelie
Chinstrap
Gentoo

- change the theme - `theme_dark`



- change the theme - `theme_minimal`

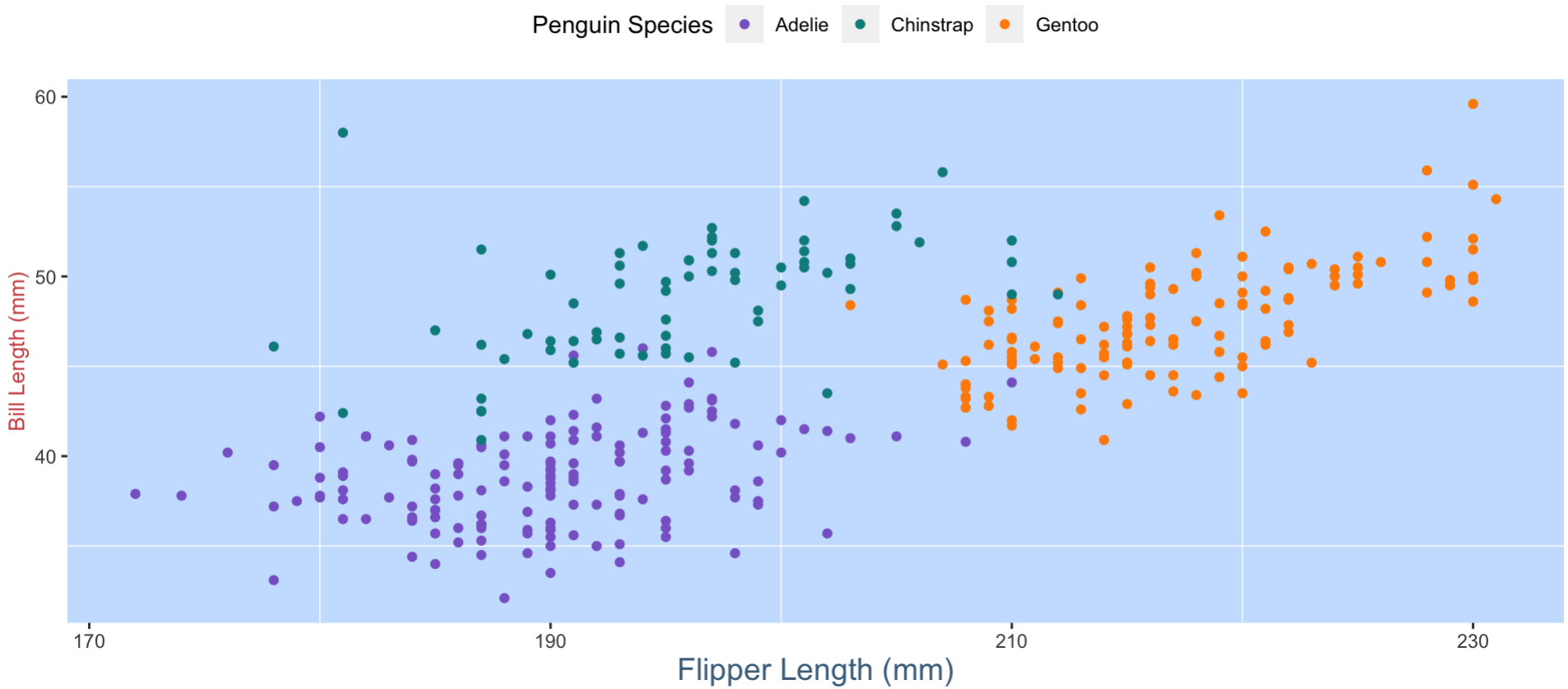


You can also manually change each theme element

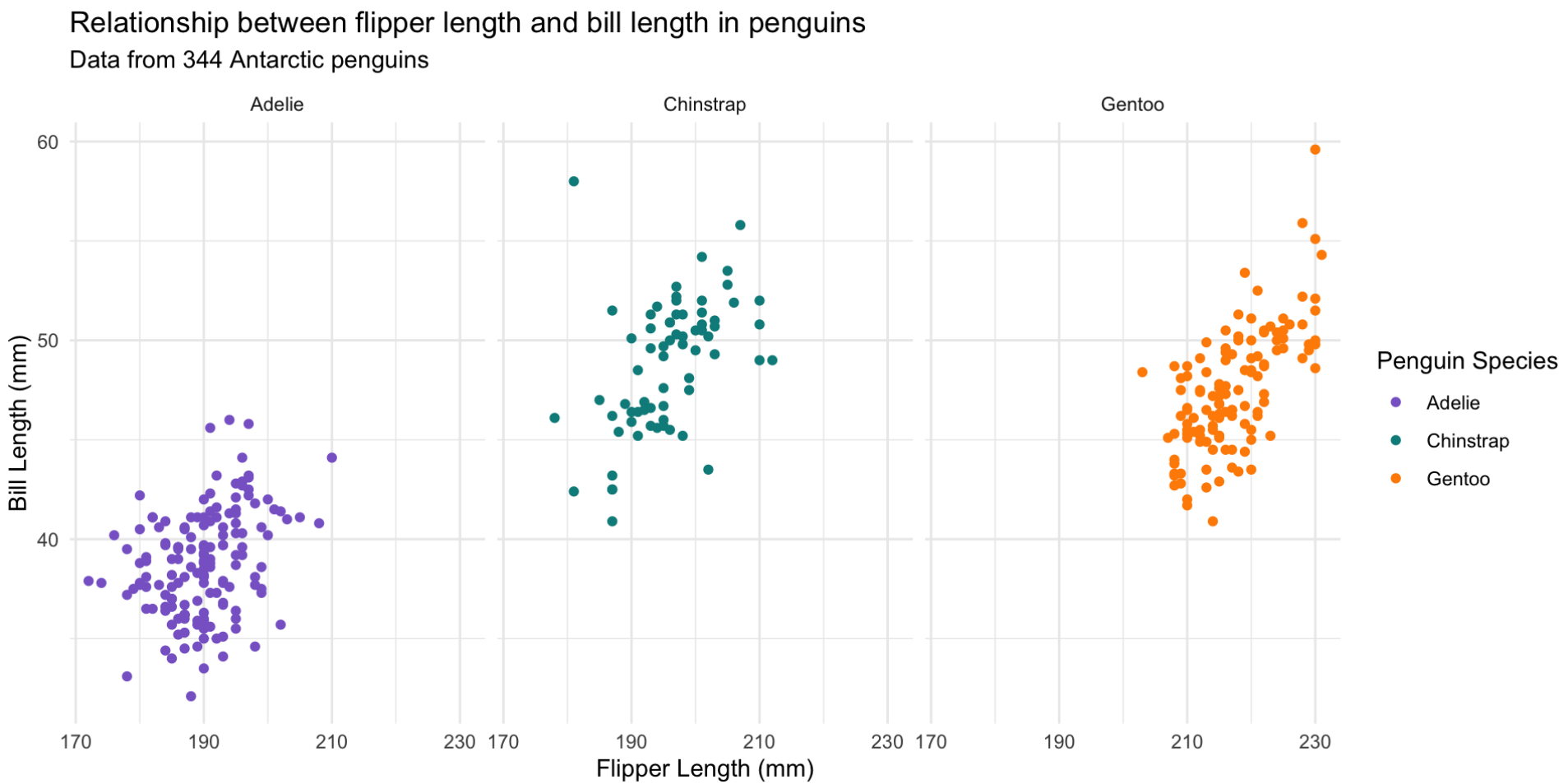
```
ggplot(data = penguins) +
  geom_point(mapping = aes(x = flipper_length_mm, y = bill_length_mm,
                           colour = species)) +
  scale_x_continuous(name = "Flipper Length (mm)",
                     breaks = c(170, 190, 210, 230)) +
  scale_y_continuous(name = "Bill Length (mm)") +
  scale_color_manual(name = "Penguin Species",
                     values = c("mediumpurple3", "darkcyan", "darkorange")) +
  labs(title = "Relationship between flipper length and bill length in peng
       subtitle = "Data from 344 Antarctic penguins") +
  theme(
    axis.title.x = element_text(colour = "skyblue4", size = 14),
    axis.title.y = element_text(colour = "indianred3", size = 10),
    panel.grid.major = element_blank(),
    panel.background = element_rect(fill = "lightsteelblue1"),
    legend.position = "top"
  )
```

Relationship between flipper length and bill length in penguins

Data from 344 Antarctic penguins



Use faceting to split your data into multiple plots



Useful resources

For a more detailed intro to ggplot2 with examples - R4DS Data Visualisation Chapter
Cheat sheet with examples of geoms - ggplot2 cheat sheet
List of all the theme elements - modify theme elements