

# Palmer\_Penguins

Madhav Mishra

2023-06-16

## Loading Packages

```
library(tidyverse)
```

```
## — Attaching core tidyverse packages — tidyverse 2.0.0 —  
## ✓ dplyr      1.1.2      ✓ readr      2.1.4  
## ✓ forcats   1.0.0      ✓ stringr   1.5.0  
## ✓ ggplot2    3.4.2      ✓ tibble    3.2.1  
## ✓ lubridate 1.9.2      ✓ tidyr     1.3.0  
## ✓ purrr     1.0.1  
## — Conflicts — tidyverse_conflicts() —  
## ✗ dplyr::filter() masks stats::filter()  
## ✗ dplyr::lag()    masks stats::lag()  
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(palmerpenguins)  
library(ggplot2)
```

## Dataset

```
View(penguins)
```

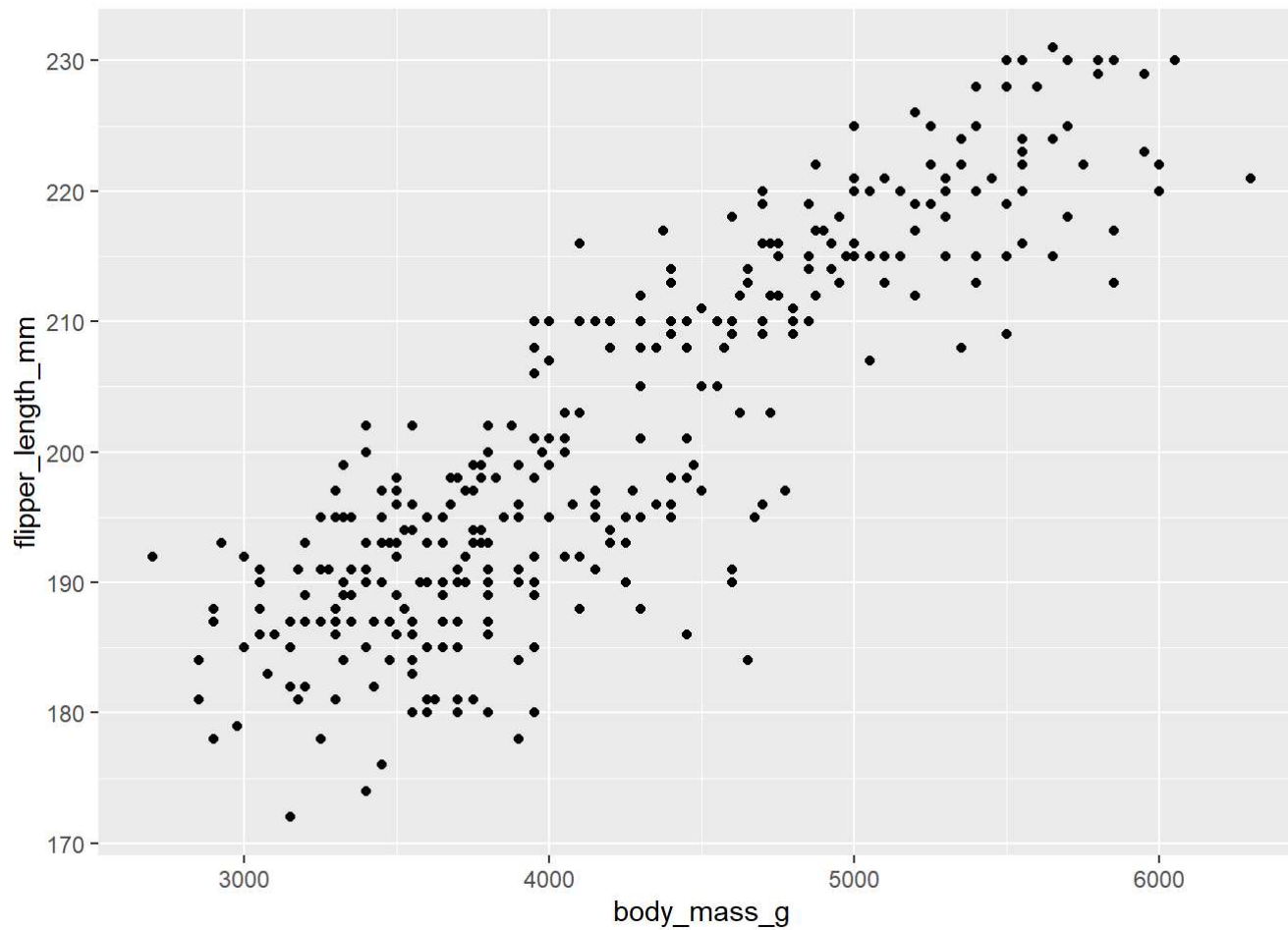
## Visualization

Here we will look over a series of visualizations

## Presenting a relation between Flipper and Body mass

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

```
## Warning: Removed 2 rows containing missing values (`geom_point()`).
```



## Presenting the above plot in purple

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

```
## Warning: Removed 2 rows containing missing values (`geom_point()`).
```



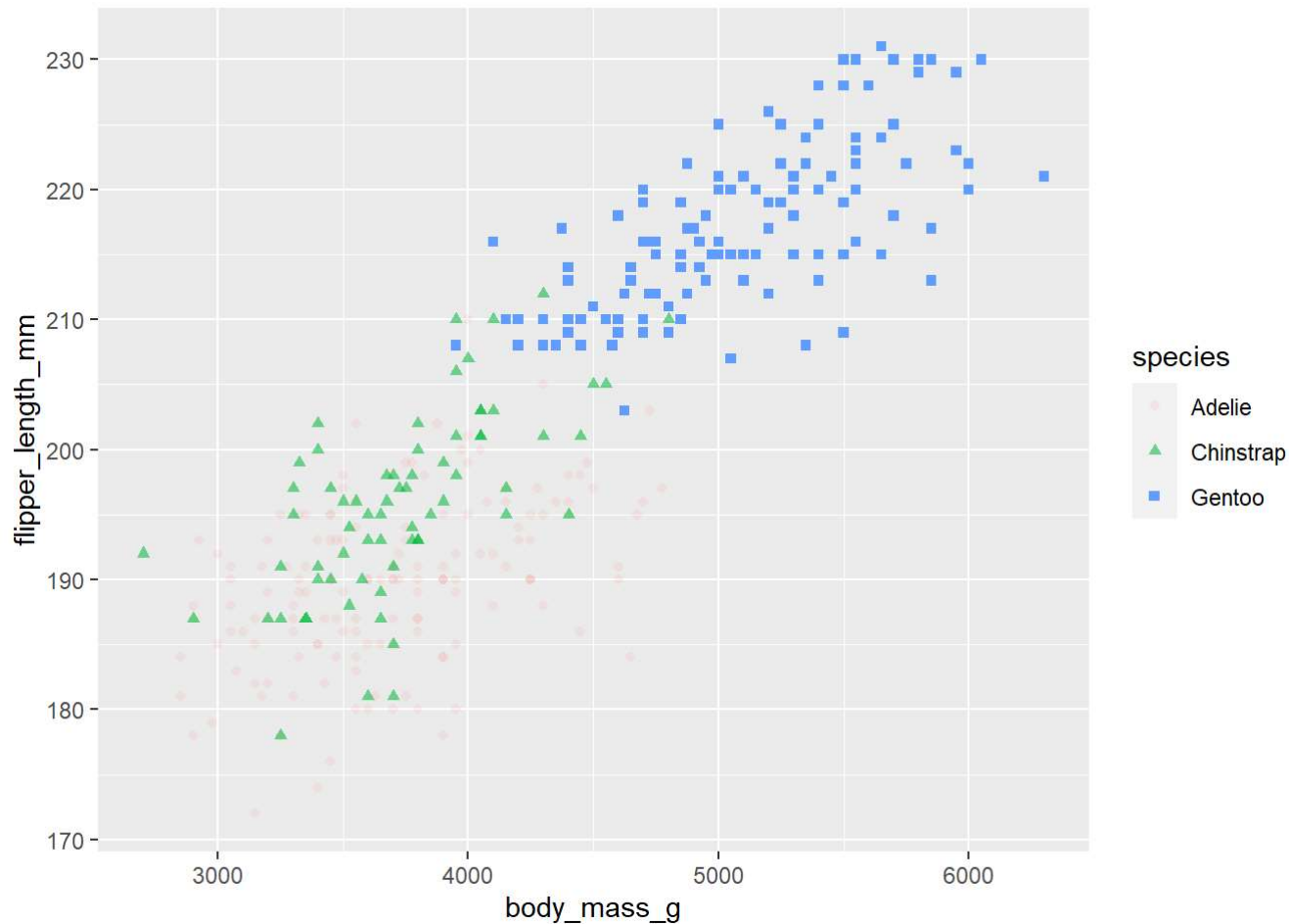
## Flipper and Body Mass by Species

Representing above relation as a plot between Flipper against Body Mass through Species based on color, shape & alpha

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

```
## Warning: Using alpha for a discrete variable is not advised.
```

```
## Warning: Removed 2 rows containing missing values (`geom_point()`).
```



## Flipper Length and Body Mass through various plots

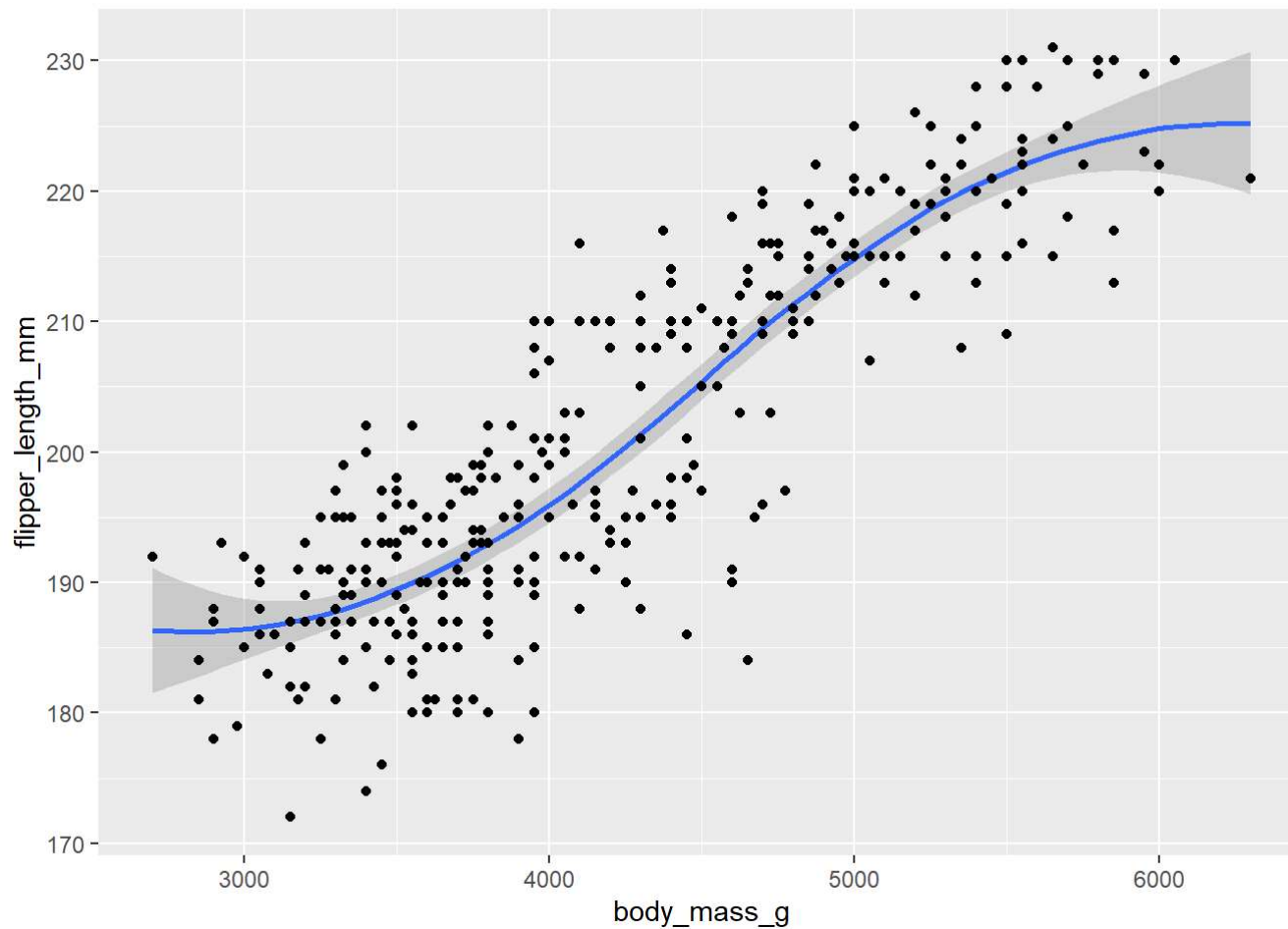
Here, I have made a visualizations using two plots i.e, Scatter and Smooth on Flipper Length against Body mass

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

```
## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```

```
## Warning: Removed 2 rows containing non-finite values (`stat_smooth()`).
```

```
## Warning: Removed 2 rows containing missing values (`geom_point()`).
```



# Flipper and Body mass using Shape, Color and Alpha

Plotting smooth and Scatter plot as done above but Representing each species with various shapes and colors. Also, using Alpha for maintaining Transparency.

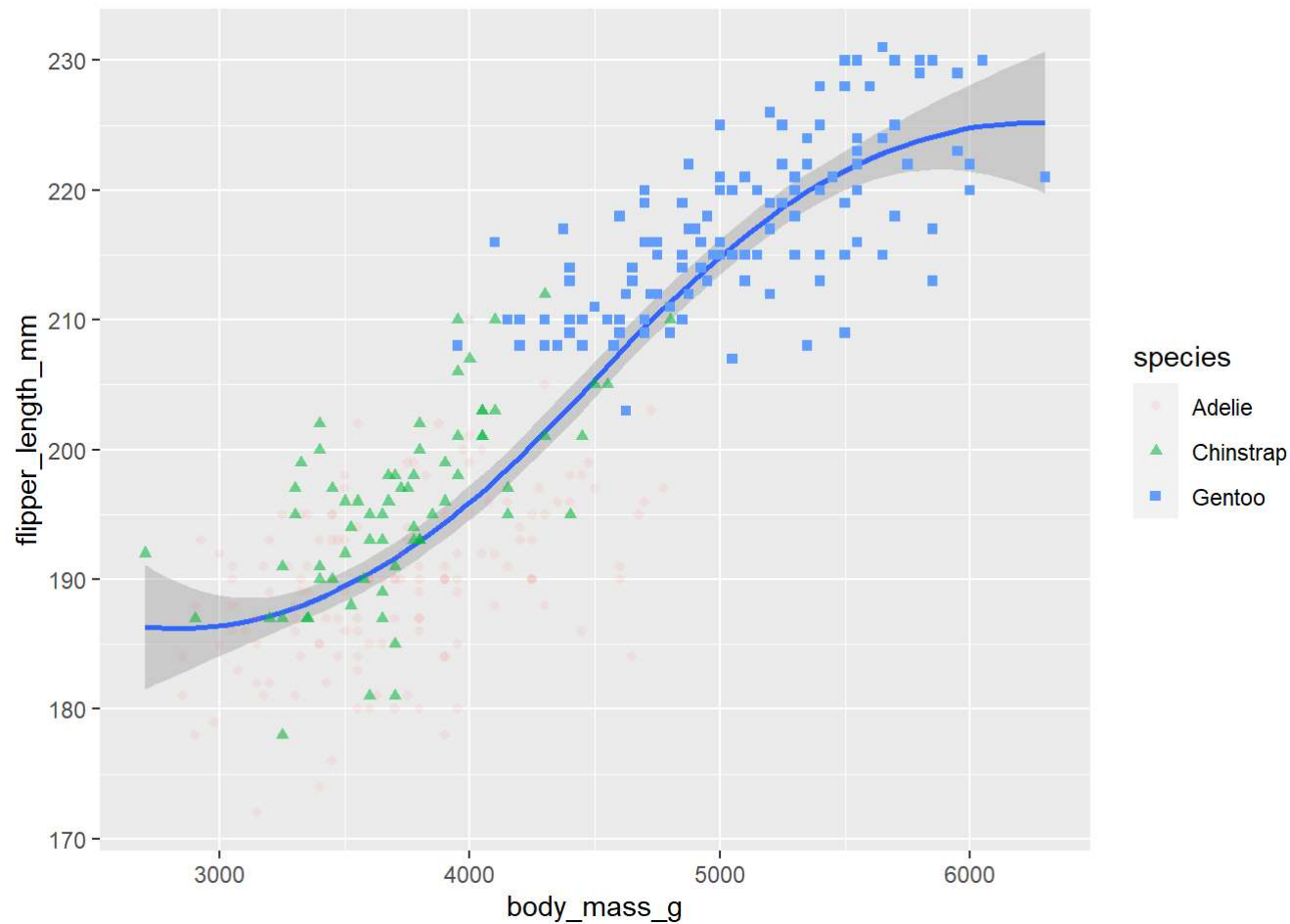
```
ggplot(data = penguins)+  
  geom_smooth(mapping = aes(x =body_mass_g, y = flipper_length_mm))+  
  geom_point(mapping = aes(x =body_mass_g, y = flipper_length_mm,color =species, shape = species,alpha = species))
```

```
## Warning: Using alpha for a discrete variable is not advised.
```

```
## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```

```
## Warning: Removed 2 rows containing non-finite values (`stat_smooth()`).
```

```
## Warning: Removed 2 rows containing missing values (`geom_point()`).
```



## Flipper and Body mass by species using Facet function

Here, we used facet function to plot flipper length against body mass, looking over breakdown by species

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

```
## Warning: Removed 2 rows containing missing values (`geom_point()`).
```



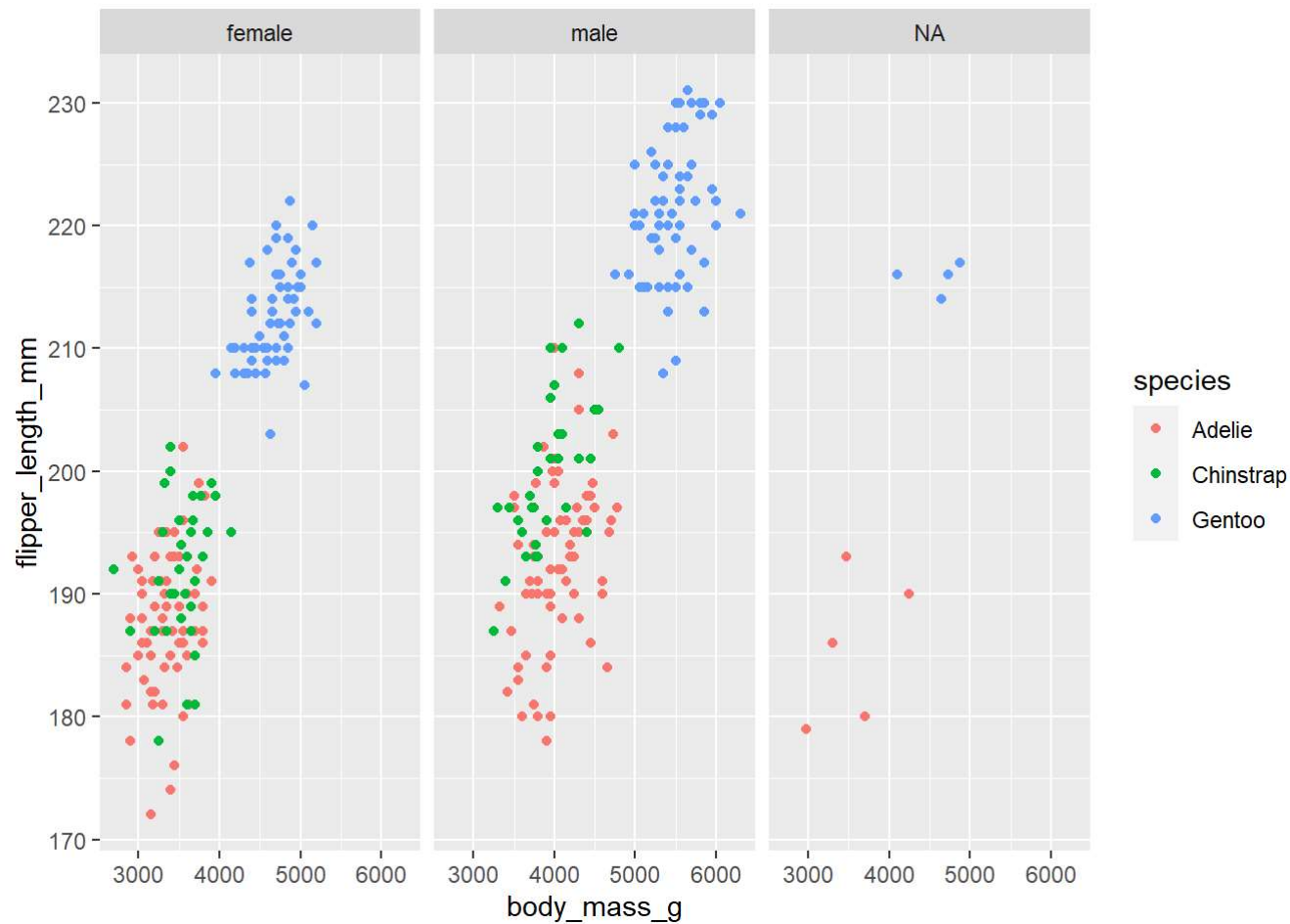
## Flipper and Body mass by sex using Facet function

Here, we used facet function to plot flipper length against body mass, looking over breakdown by sex

```
ggplot(data=penguins)+  
  geom_point(mapping = aes(x= body_mass_g , y =flipper_length_mm,color =species))+  
  facet_grid(~sex)
```

```
## Warning: Removed 2 rows containing missing values (`geom_point()`).
```





## Title, Subtitle and Caption

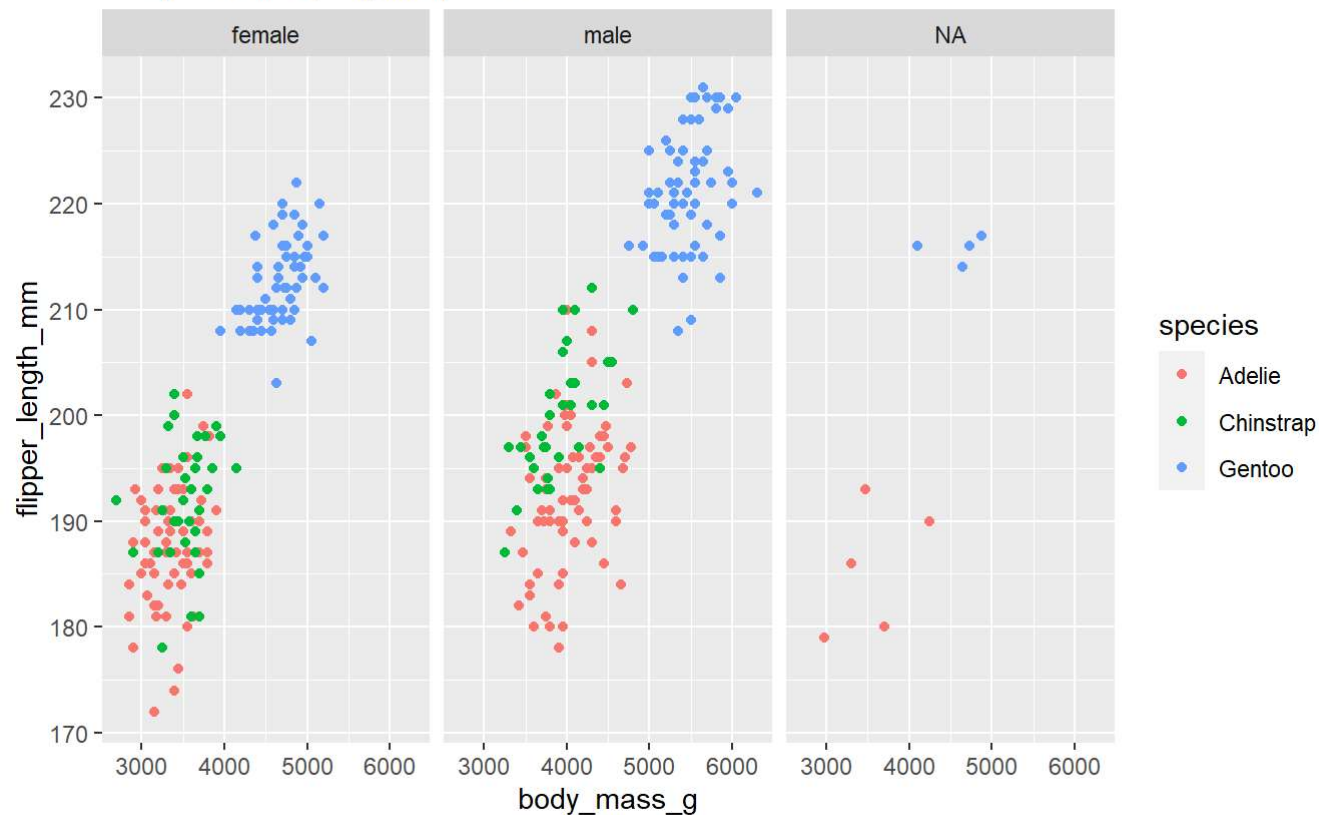
Here, we have included title, subtitle and caption for the above plot

```
ggplot(data=penguins)+
  geom_point(mapping = aes(x= body_mass_g , y =flipper_length_mm,color =species))+
  facet_grid(~sex)+
  labs(title = "Palmer Penguins: Data vs Flipper Length",subtitle = "Sample of three penguin species", caption = "Data Collected by Dr. Kristen Gorman")
```

```
## Warning: Removed 2 rows containing missing values (`geom_point()`).
```

## Palmer Penguins: Data vs Flipper Length

Sample of three penguin species



Data Collected by Dr. Kristen Gorman

## Assigning a variable p

Here, we have created a new variable p and assigned it the value of above plotted graph

```
p <- ggplot(data=penguins)+  
  geom_point(mapping = aes(x= body_mass_g , y =flipper_length_mm,color =species))
```

## Adding an annotation in plot

We are adding an Annotation in plot created few chunks above to specifically drew attention towards our result

```
p +labs(title = "Palmer Penguins: Data vs Flipper Length",subtitle = "Sample of three penguin species", caption = "Data Collected by Dr. Kristen Gorman")+  
  annotate("text",x = 5500,y=205,label = "Gentoos are the largest")
```

```
## Warning: Removed 2 rows containing missing values (`geom_point()`).
```

