

Species

```
#Set up Code
```

```
knitr::opts_chunk$set(echo = TRUE)
```

In this section of code, I am loading the packages needed and looking at the structure of the dataset. Then, I am grouping the data by sex and adding two new columns called mean_hindfoot and mean_weight to the table.

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.3     v dplyr   1.0.7
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   2.0.1     v forcats 0.5.1
```

```
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()   masks stats::lag()
```

```
library(dplyr)
library(readr)
surveys_complete <- read_csv("data/surveys_complete.csv")
```

```
## New names:
## * ` ` -> ...1
```

```
## Rows: 30718 Columns: 10
```

```
## -- Column specification -----
## Delimiter: ","
## chr (2): species_id, sex
## dbl (8): ...1, record_id, month, day, year, plot_id, hindfoot_length, weight
```

```
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
str(surveys_complete)
```

```

## spec_tbl_df [30,718 x 10] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
## $ ...1 : num [1:30718] 1 2 3 4 5 6 7 8 9 10 ...
## $ record_id : num [1:30718] 63 64 65 66 67 68 69 70 71 74 ...
## $ month : num [1:30718] 8 8 8 8 8 8 8 8 8 8 ...
## $ day : num [1:30718] 19 19 19 19 19 19 19 19 19 19 ...
## $ year : num [1:30718] 1977 1977 1977 1977 1977 1977 ...
## $ plot_id : num [1:30718] 3 7 4 4 7 8 2 3 7 8 ...
## $ species_id : chr [1:30718] "DM" "DM" "DM" "DM" ...
## $ sex : chr [1:30718] "M" "M" "F" "F" ...
## $ hindfoot_length: num [1:30718] 35 37 34 35 35 32 15 21 36 12 ...
## $ weight : num [1:30718] 40 48 29 46 36 52 8 22 35 7 ...
## - attr(*, "spec")=
##   .. cols(
##     ... .1 = col_double(),
##     ... record_id = col_double(),
##     ... month = col_double(),
##     ... day = col_double(),
##     ... year = col_double(),
##     ... plot_id = col_double(),
##     ... species_id = col_character(),
##     ... sex = col_character(),
##     ... hindfoot_length = col_double(),
##     ... weight = col_double()
##   ... )
## - attr(*, "problems")=<externalptr>

```

```

by_sex <- surveys_complete %>%
  group_by(sex) %>%
  summarize(mean_hindfoot = mean(hindfoot_length, na.rm = TRUE),
            mean_weight = mean(weight, na.rm = TRUE))

```

In this section of code, I am grouping the data by species ID and adding the mean hindfoot length and mean weight columns.

```

by_species <- surveys_complete %>%
  group_by(species_id) %>%
  summarize(mean_hindfoot = mean(hindfoot_length, na.rm = TRUE),
            mean_weight = mean(weight, na.rm = TRUE))

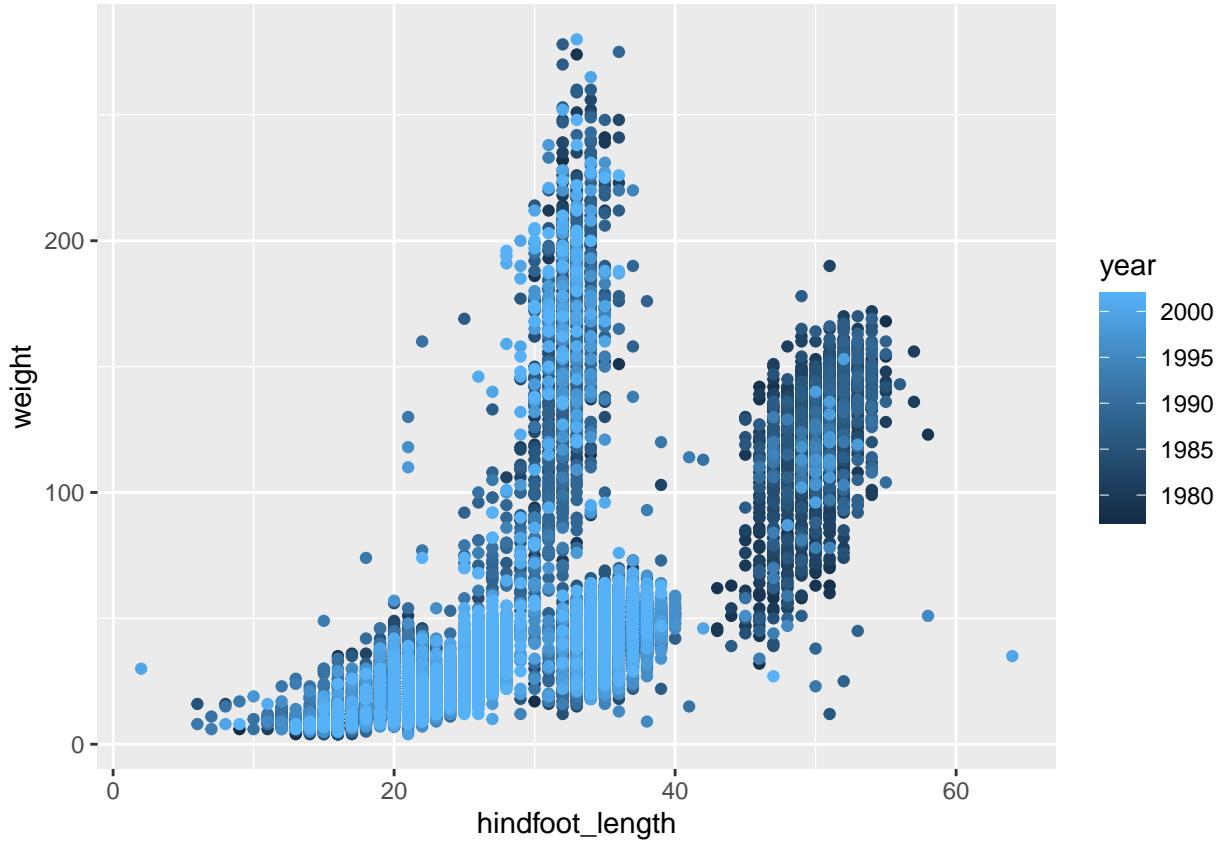
```

In this section, I am creating a scatterplot that looks at the relationship between hindfoot length and weight as well as looking at the year, which is mapped to the color aesthetic. The darker the point, the earlier the year.

```

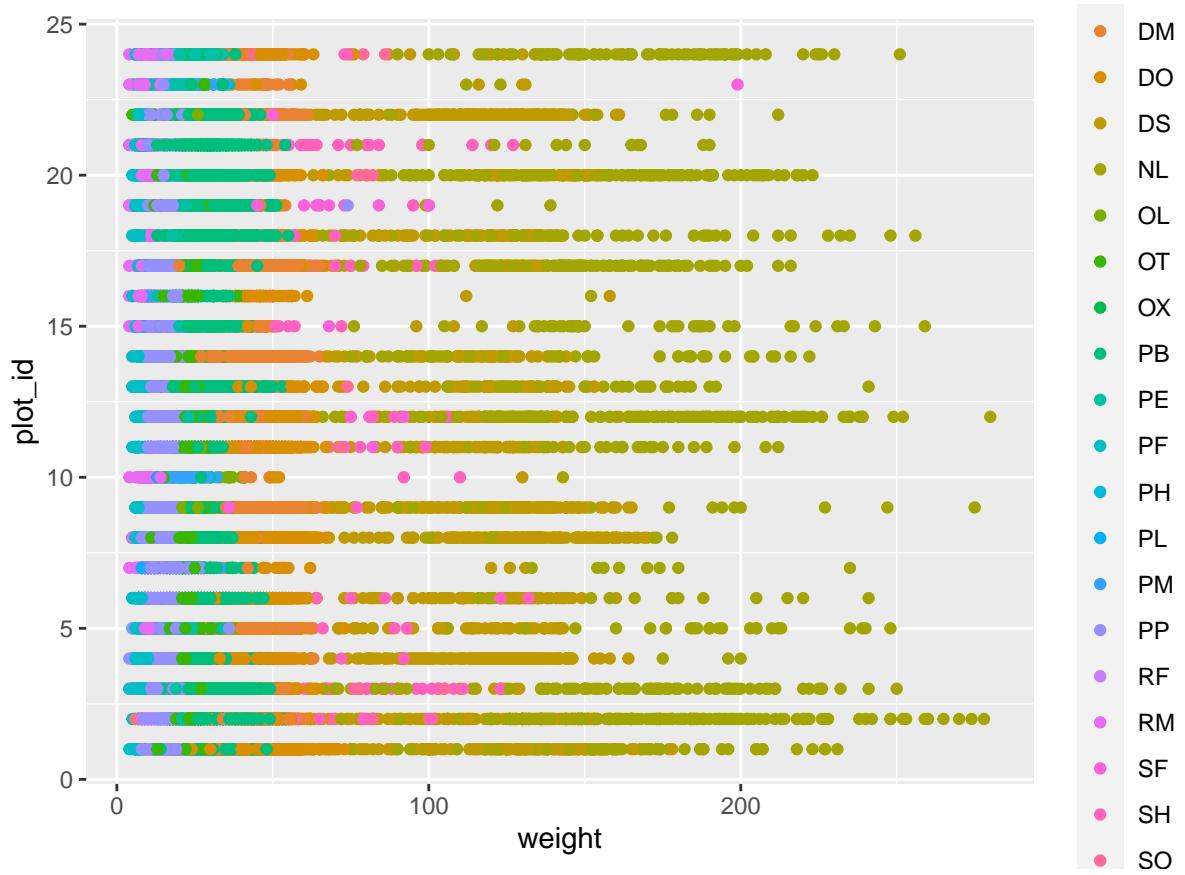
ggplot(data = surveys_complete) +
  geom_point(mapping = aes(x = hindfoot_length, y = weight, color = year))

```



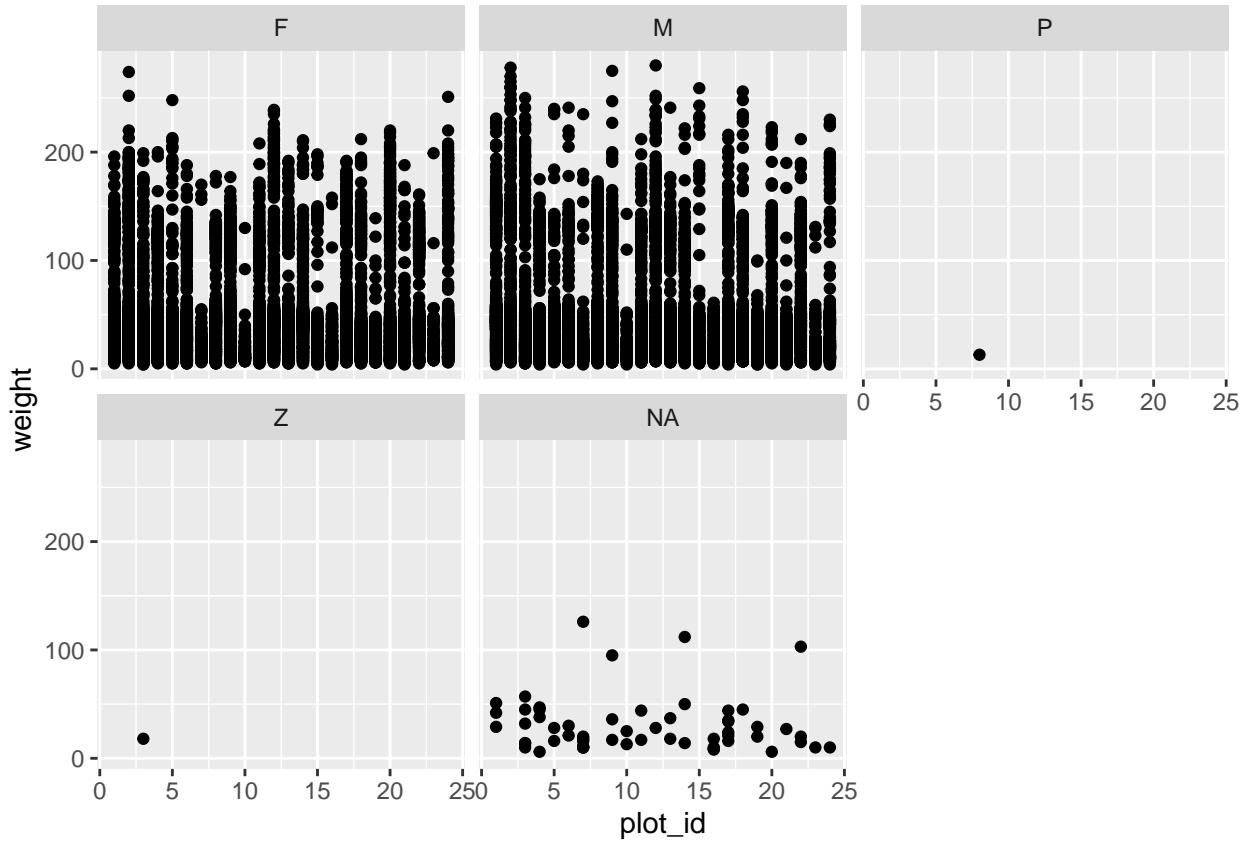
In this section, I am creating a visualization that shows the relationship between plot ID, weight, and species ID.

```
ggplot(data = surveys_complete) +  
  geom_point(mapping = aes(x = plot_id, y = weight, color = species_id)) +  
  coord_flip()
```



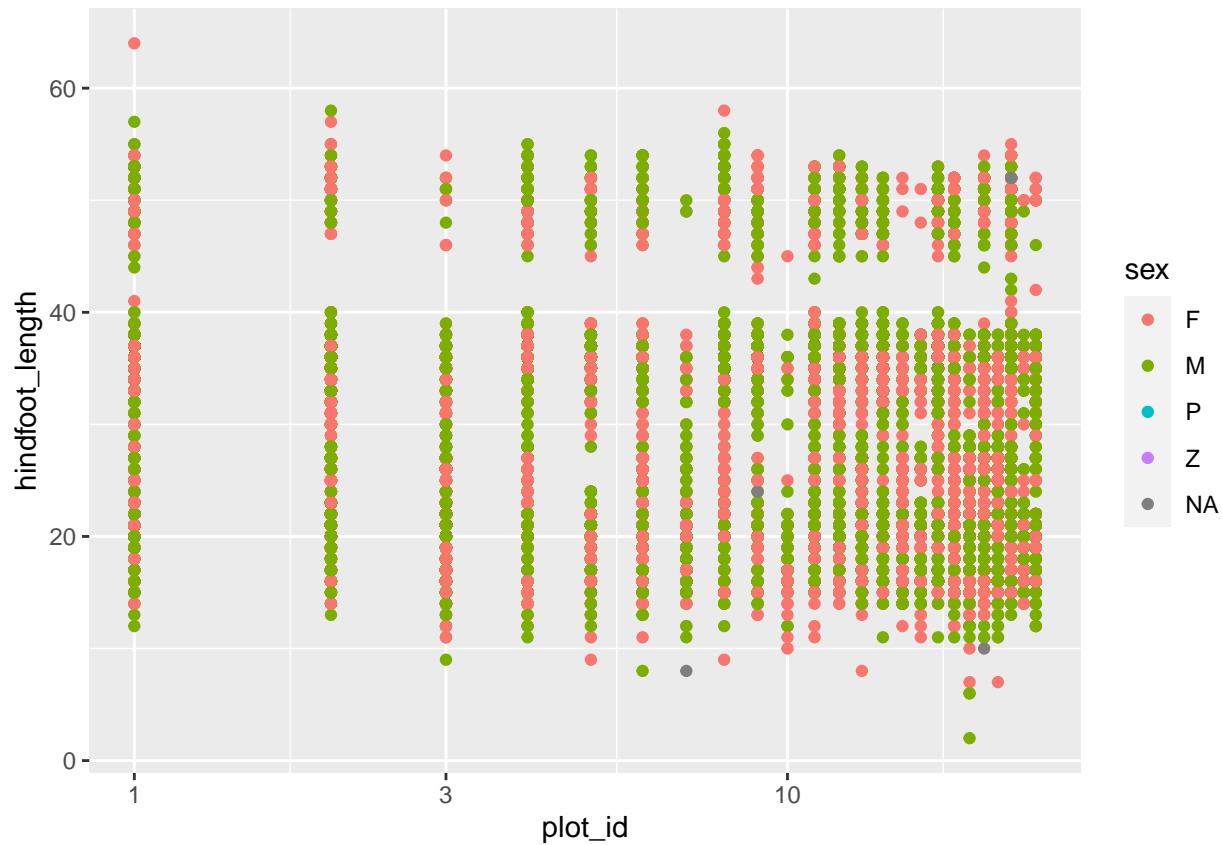
In this section, I am creating a set of scatter plots in which plot ID is mapped to the x axis and weight is mapped to the y axis. These graphs are faceted by sex.

```
ggplot(data = surveys_complete) +
  geom_point(mapping = aes(x = plot_id, y = weight)) +
  facet_wrap(~sex)
```



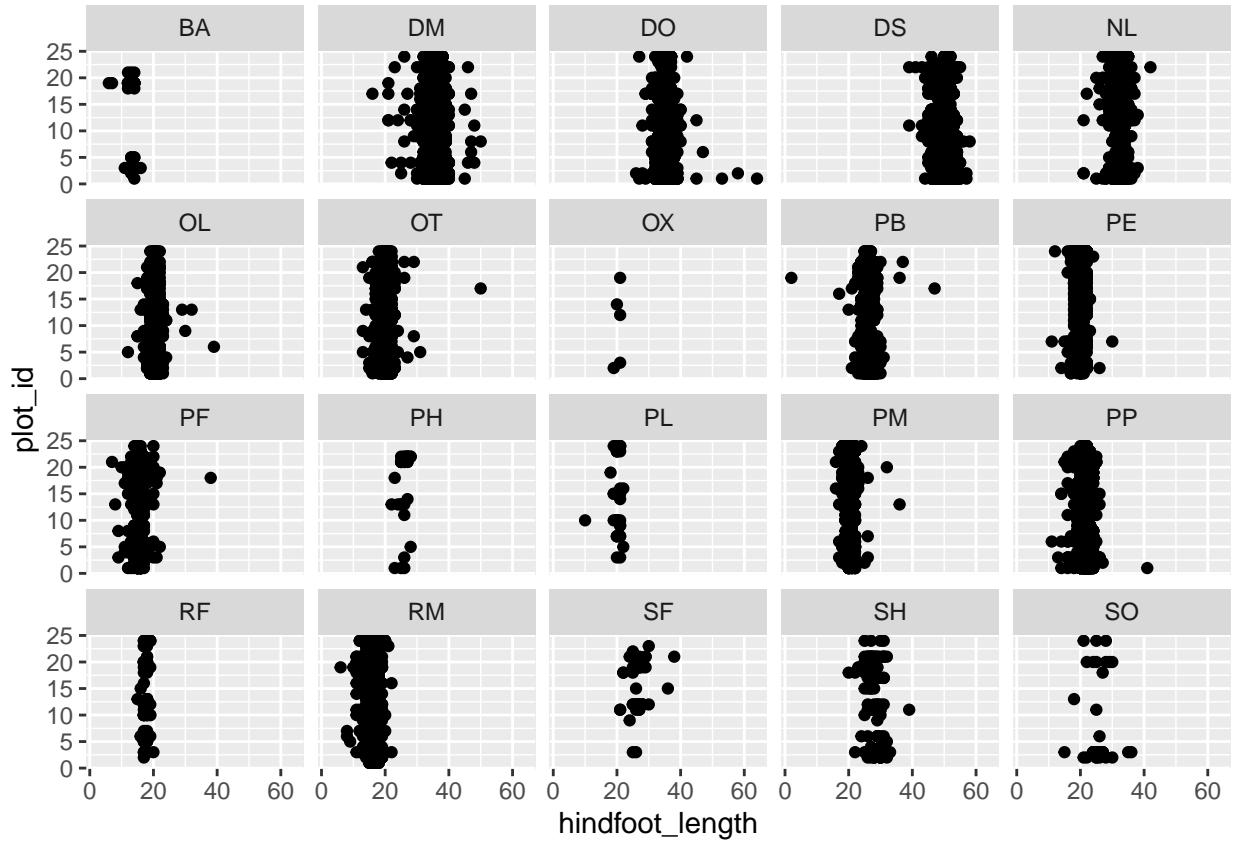
This code section creates a visualization that shows the relationship between plot ID, hindfoot length, and sex, which is mapped to the color aesthetic.

```
ggplot(data = surveys_complete) +
  geom_point(mapping = aes(x = plot_id, y = hindfoot_length, color = sex)) +
  scale_x_log10()
```



This section looks at the relationship between hindfoot length and plot ID and is faceted by species ID.

```
ggplot(data = surveys_complete) +
  geom_point(mapping = aes(x = plot_id, y = hindfoot_length)) +
  facet_wrap(~species_id) +
  coord_flip()
```



This code section created a scatterplot similar to the first one. It looks at the relationship between weight and hindfoot length, but also looks at how this data is broken down by species ID, which is mapped to the color aesthetic.

```
ggplot(data = surveys_complete) +
  geom_point(mapping = aes(x = weight, y = hindfoot_length, color = species_id)) +
  annotate("text", x = 115, y = 50, label = "these have long feet")
```

