Module 2, Assignment 3

Ellen Bledsoe

2023-02-23

Assignment Details

Purpose

The goal of this assignment is to practice choosing the correct types of variables for certain plots and to practice writing code for those plots.

Task

Write R code which produces the correct plot style.

Criteria for Success

- Code is within the provided code chunks
- Code is commented with brief descriptions of what the code does
- Code chunks run without errors
- Code produces the correct result
    - Code that produces the correct answer will receive full credit
    - Code attempts with logical direction will receive partial credit
- Written answers address the questions in sufficient detail

Due Date

March 2 at midnight MST

Assignment Questions

For this assignment, we are going to be making even *more* plots using `ggplot2`! We are going to go back to our trusty data called `penguins` from the `palmerpenguins` package.

1. Load both the `palmerpenguins` package and the `tidyverse` package into the work space. (2 points)

```
library(palmerpenguins)
library(tidyverse)
```

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

```
## v ggplot2 3.3.5      v purrr    0.3.4
## v tibble  3.1.6      v dplyr    1.0.7
## v tidyr   1.1.4      v stringr 1.4.0
## v readr   2.1.1      v forcats 0.5.1


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

OPTIONAL When we use data from a data package, it doesn't automatically show up in our environment. Run this code chunk so it does show up in the environment.

```
penguins <- penguins
```

---

2. There are 8 columns in the `penguins` data frame. Determine if each variable is continuous or categorical and record your answer below.

   Note: the `year` column can be considered continuous or categorical. *For this assignment, we will consider it continuous.*

   Answer:

   - `species`:
   - `island`:
   - `bill_length_mm`:
   - `bill_depth_mm`:
   - `flipper_length`:
   - `body_mass_g`:
   - `sex`:
   - `year`: continuous

---

Important Instructions!

(a) For *all* of the plots below, be sure to add:

   - improved labels for the x-axis, y-axis, and the legend (if present)
   - one of the following themes to your code: `theme_bw`, `theme_classic`, or `theme_light`.

(b) For each question, you get to choose the variables you use for each plot! Pay attention to which variables are continuous vs. categorical and where they should go in the code to produce the correct type of plot.

(c) Because I'm having you choose your own variables, the examples in the provided Answer Key likely will not match yours (though they might). You should strive to produce a plot that looks *similar in style* to those in the Answer Key, but the number of categories and the actual data points will likely be different.

(d) You *cannot* use the same plots that you produced in Module 2 Assignment 1. You can use them for inspiration, but don't use the exact same variables.

(e) Each plot is worth 3 points. Your interpretation is worth 1 point.

_____

3. Produce a histogram plot (just one, not multiple on the same plot). Write 1-2 sentences describing what information you can gather from plot.

```
ggplot(penguins, aes(bill_depth_mm)) +
  geom_histogram(bins = 20) +
  labs(x = "Descriptive Axis Label",
       y = "Descriptive Axis Label") +
  theme_bw()
```
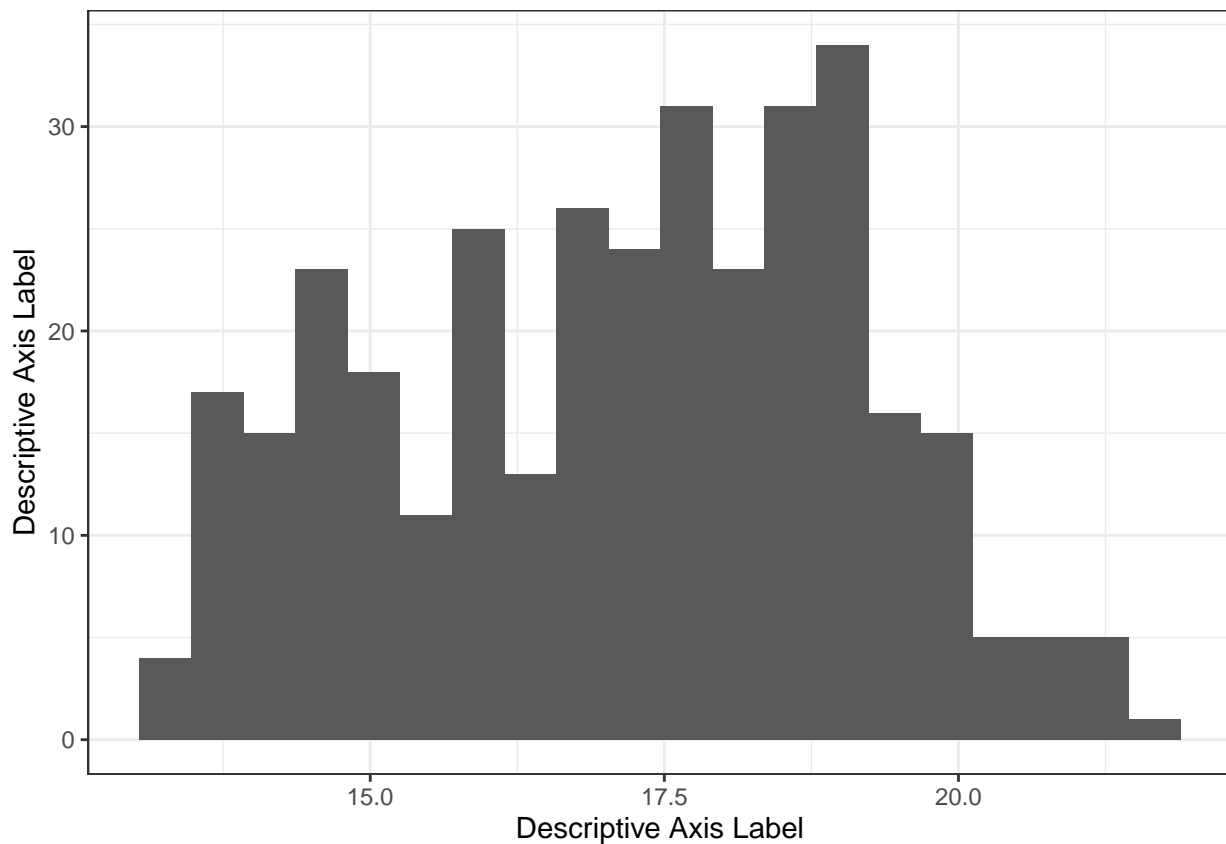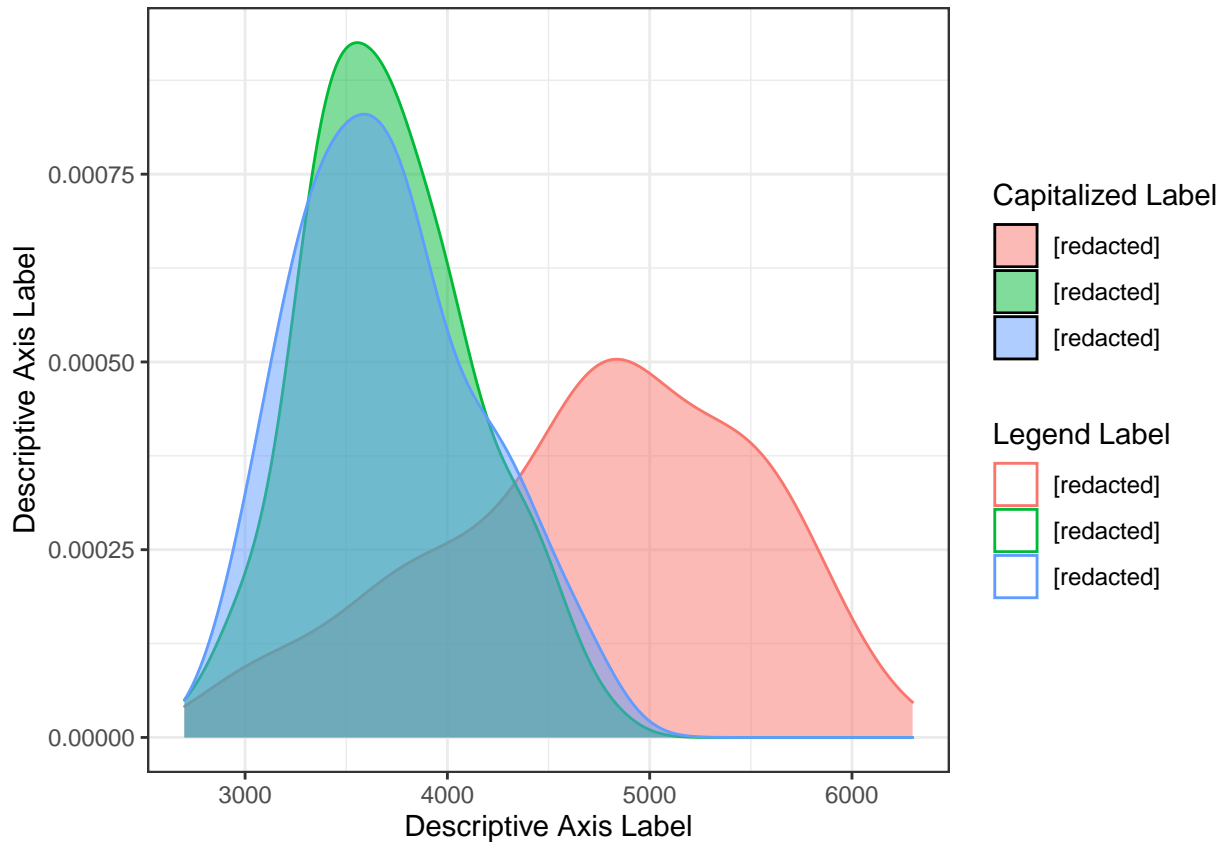

Answer: Dependent on their plot.

4. Create a multiple density plot. Make sure we can see each density plot (make sure one isn't blocking another)(3 points). Write 1-2 sentences describing what information you can gather from plot.

```
ggplot(penguins, aes(body_mass_g, color = island, fill = island)) +
  geom_density(alpha = 0.5) +
  labs(x = "Descriptive Axis Label",
       y = "Descriptive Axis Label",
       color = "Legend Label",
       fill = "Capitalized Label") +
  scale_fill_discrete(labels = c("[redacted]", "[redacted]", "[redacted]")) +
  scale_color_discrete(labels = c("[redacted]", "[redacted]", "[redacted]")) +
  theme_bw()
```

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


Answer: Dependent on their plot.

5. Create a box-and-whisker plot. Changing the color of the boxes is optional (if you want to get super fancy, you can change the color based on a different variable). Make sure to add points to the plot. Write 1-2 sentences describing what information you can gather from plot.
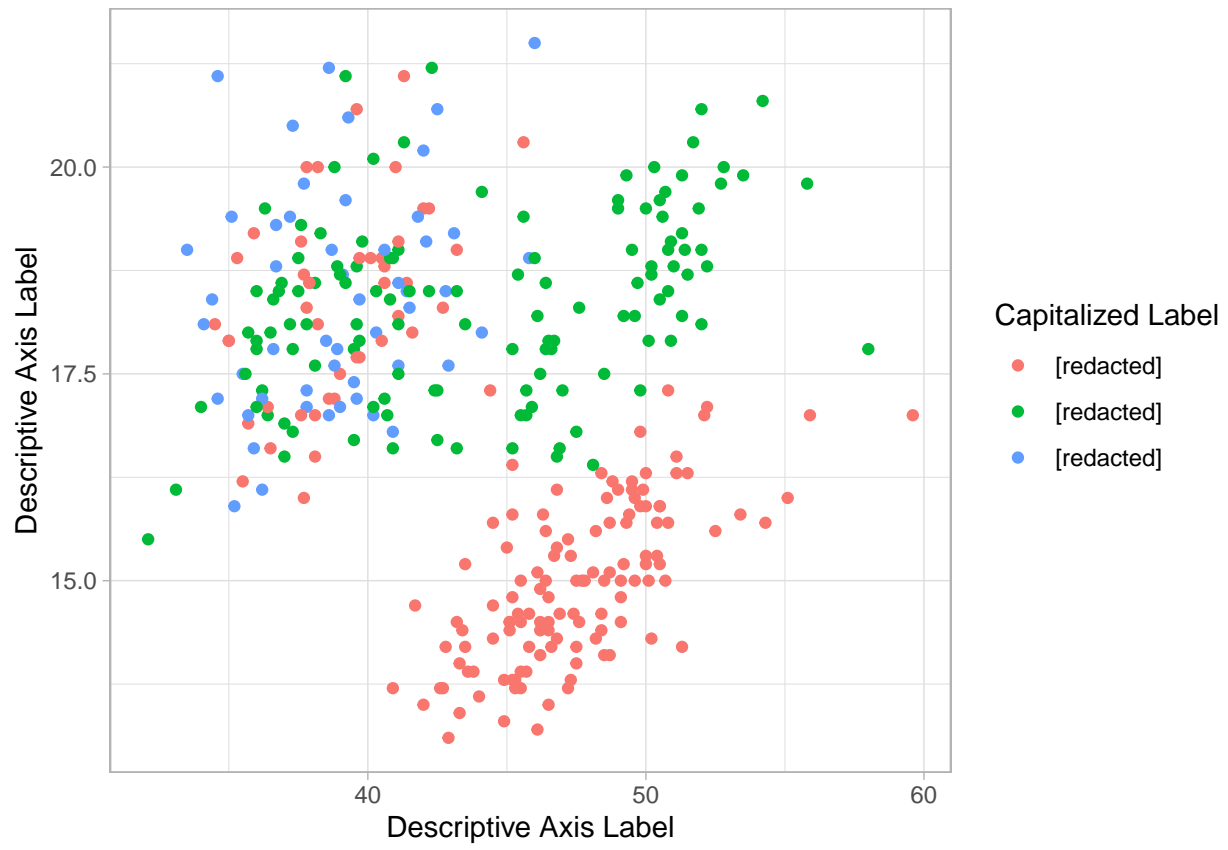
```r
ggplot(penguins, aes(species, flipper_length_mm, color = species)) +
  geom_boxplot() +
  geom_jitter(alpha = 0.5, width = 0.1) +
  labs(x = "Descriptive Axis Label",
       y = "Descriptive Axis Label",
       color = "Capitalized Label") +
  scale_x_discrete(labels = c("[redacted]", "[redacted]", "[redacted]")) +
  scale_color_discrete(labels = c("[redacted]", "[redacted]", "[redacted]")) +
  theme_classic()
```
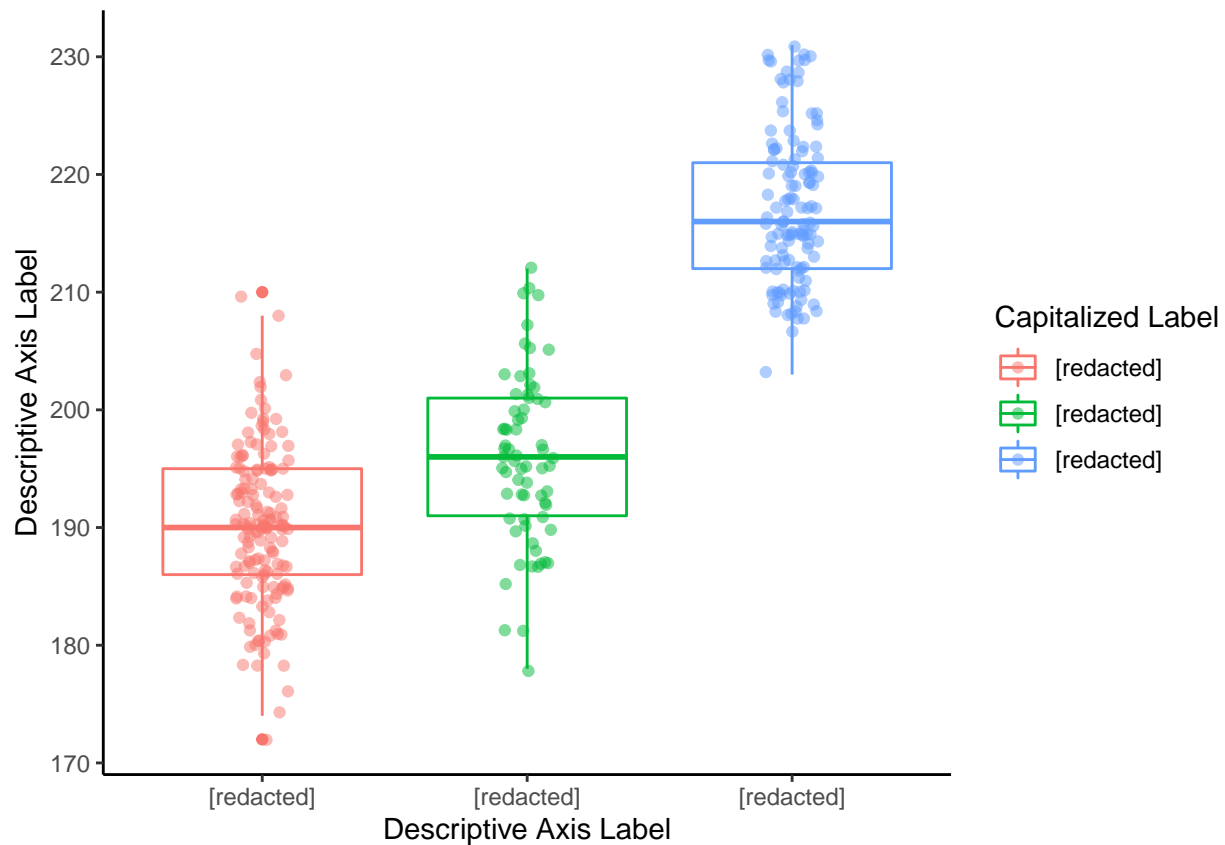
Answer: Dependent on their plot.

6. Create a scatter plot. Change the color of the points to create a "grouped" scatter plot. Write 1-2 sentences describing what information you can gather from plot. (4 points)

```
ggplot(penguins, aes(bill_length_mm, bill_depth_mm, color = island)) +
  geom_point() +
  labs(x = "Descriptive Axis Label",
       y = "Descriptive Axis Label",
       color = "Capitalized Label") +
  scale_color_discrete(labels = c("[redacted]", "[redacted]", "[redacted]")) +
  theme_light()
```

Answer: Dependent on their plot.