Homework 9

Cristian Sigala Crs4565

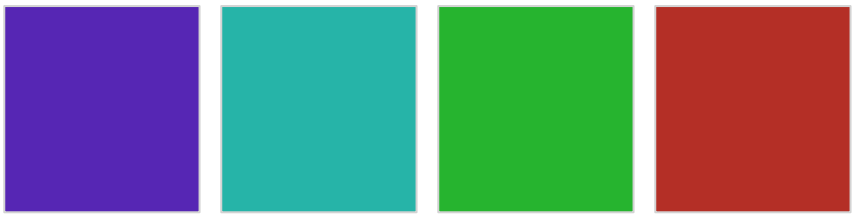
This homework is due on April 26, 2021 at 11:00pm. Please submit as a pdf file on Canvas.

Problem 1: (2 pts)

Use the color picker app from the **colorspace** package (`colorspace::choose_color()`) to create a qualitative color scale containing four colors. One of the four colors should be `#5626B4`, so you need to find three additional colors that go with this one.

```
# replace "#FFFFFF" with your own colors
colors <- c("#5626B4", "#26b4a8", "#26b42f", "#b42f26")

swatchplot(colors)
```

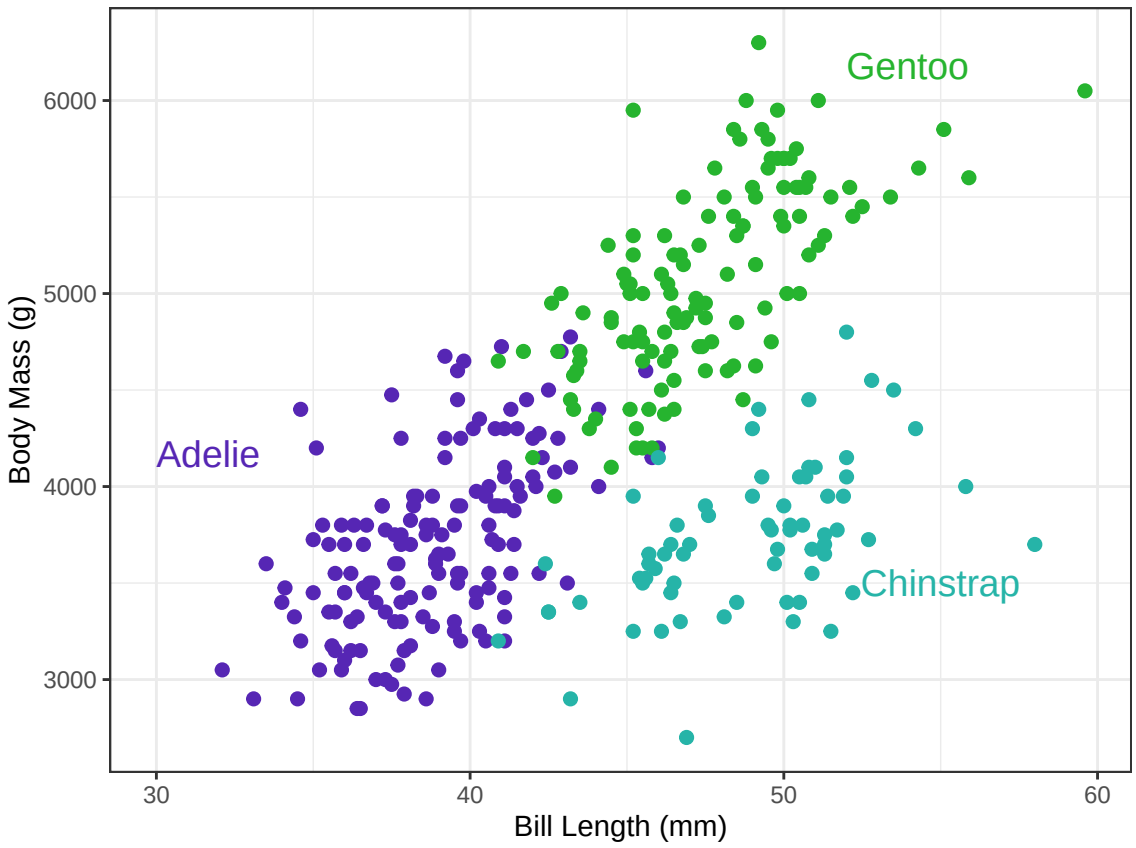


Problem 2: (4 pts) Take the following scatter plot of the penguins dataset and make three modifications:

1. Use the colors you chose in Problem 1.
2. Improve the visual appearance by choosing a theme and cleaning up axis labels.
3. Remove the need for a legend by direct-labeling the points.

```
penguin_labels <- tibble(
  species = c("Adelie", "Chinstrap", "Gentoo"),
  bill_length_mm = c(30, 55, 52),
  body_mass_g = c(4100, 3500, 6250),
  label = c("Adelie", "Chinstrap", "Gentoo"),
  hjust = c(0, 0.5, 0),
  vjust = c(0, 0.5, 1)
)
```

```r
ggplot(penguins, aes(bill_length_mm, body_mass_g, color = species)) +
  geom_point(size = 2, na.rm = TRUE) +
  theme_bw()+
  scale_color_manual(values = c("#5626B4", "#26b4a8", "#26b42f", "#b42f26")) +
  xlab("Bill Length (mm)") +
  ylab("Body Mass (g)") +
  geom_text(
    data = penguin_labels,
    aes(
      label = label,
      hjust = hjust, vjust = vjust
    ),
    size = 14/.pt
  ) +
  guides(color = "none", shape = "none")
```



Problem 3: (4 pts) The following scatter plot shows per-capita income versus number of inhabitants in all Texas counties in 2010. Use `geom_text_repel()` to label a subset of the counties by name. You can choose the counties to subset as you wish. Also, choose a theme and clean up the axis labeling, and make any other improvements to the plot design you consider appropriate.

Hint: If you're not sure how to select a subset of counties to label, check out the examples on the **ggrepel** website for some inspiration: https://ggrepel.slowkow.com/articles/examples.html#examples-1

```r
tx_census <- read_csv("https://wilkelab.org/SDS375/datasets/US_census.csv") %>%
  filter(state == "Texas") %>%
```

```
  select(county = name, pop2010, per_capita_income)




tx_census %>%
  mutate(
    label = ifelse(county %in%
                     c("Travis County", "Tarrant County",
                       "Harris County","Anderson County" ,
                      "Baylor County"), county, "")) %>%
  ggplot(aes(pop2010, per_capita_income)) +
  geom_point(size = 1) +
  scale_x_log10() +
  geom_text_repel(
    aes(label = label),
    max.overlaps = Inf,
    box.padding = 3) +
  xlab("Population") +
  ylab("Per Capita Income") +
  ggtitle("Per Capita Income and Pop of Texas Counties") +
  theme_classic()
```