DataViz2 Customizing Graph

Ellie Kuskie

1) Penguins!

Setup

```r
library(ggplot2)
library(here)
```

here() starts at C:/Users/conno/OneDrive/Documents/UCSC Undergrad/DataScience4EEB/DataViz2-cu

```r
library(palmerpenguins)
```

```r
# Read in the palmer penguins data
penguins <- penguins
```
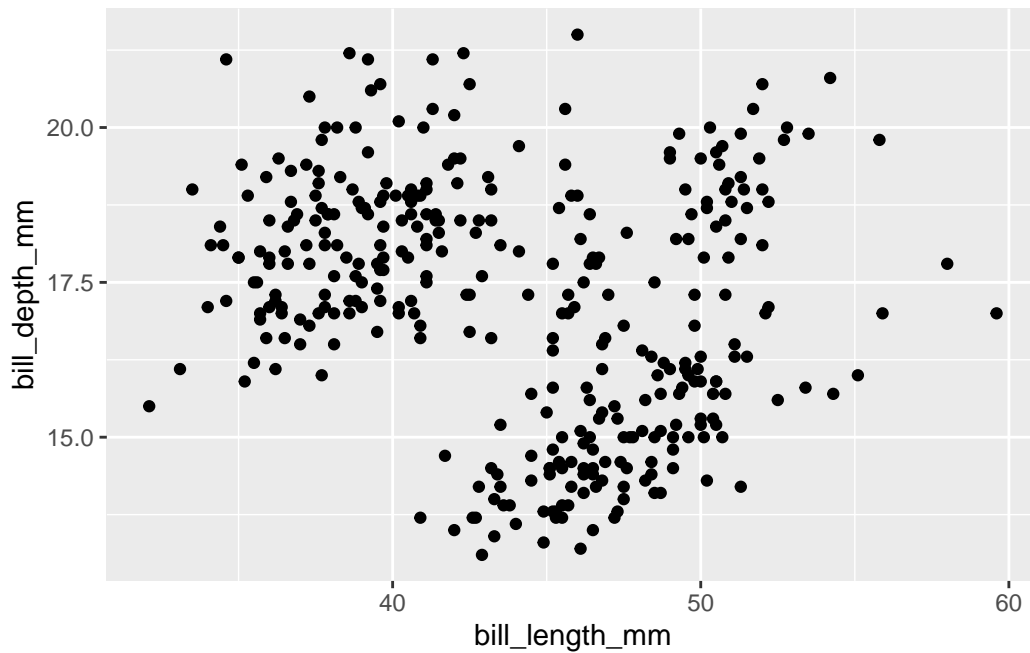
```r
# Putting a "?" in front of a function (or built-in dataset) will automatically bring up its
?penguins
```

starting httpd help server ... done

```r
# creating initial plot, no color

ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +
  geom_point()
```

Warning: Removed 2 rows containing missing values or values outside the scale range
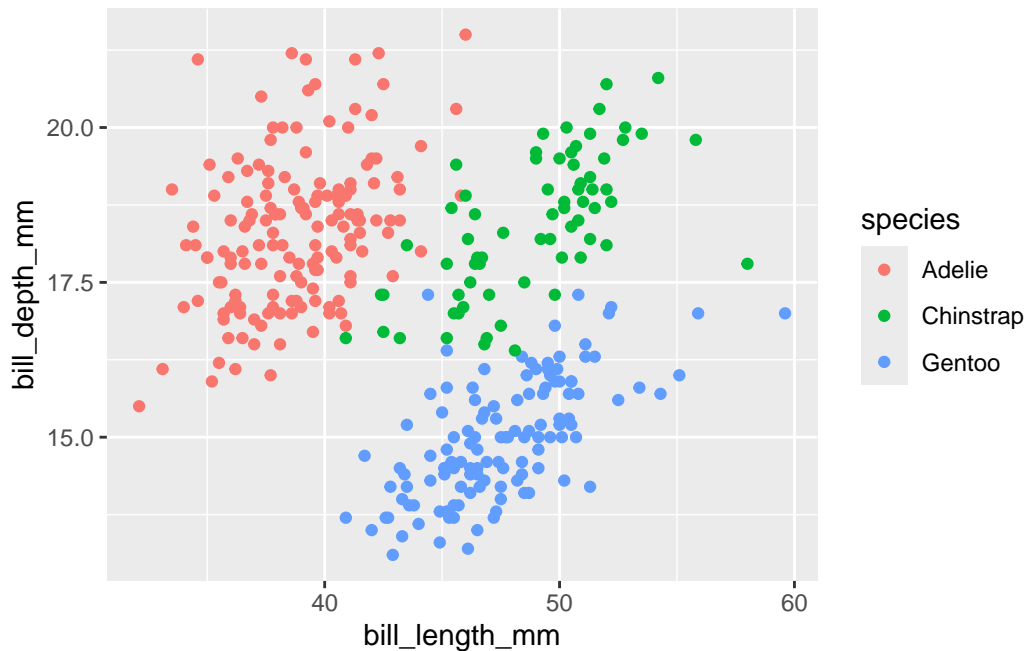(`geom_point()`).

Q1.1)

It does not look like there is a relationship between the variables, at most it looks 2 clumps. But if the 2 clumps are the same species then there is no relationship.

Q1.2)

```
# added in color coding for each species

ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +
  geom_point(aes(color = species))
```

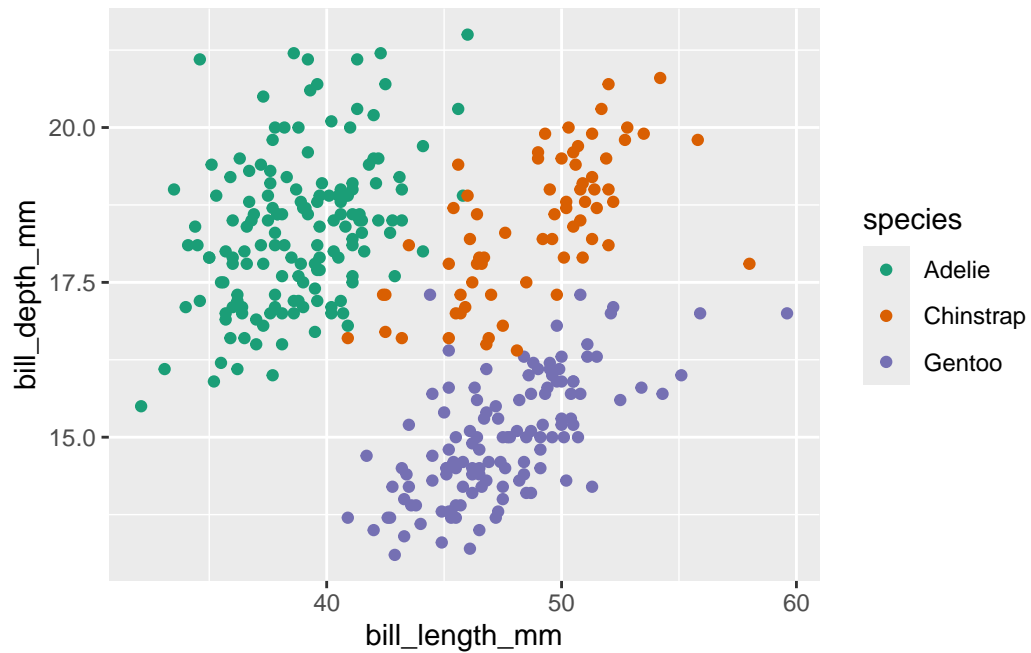Warning: Removed 2 rows containing missing values or values outside the scale range (`geom_point()`).

There now appears to be much more of a pattern between bill length and bill depth. For each species there is an increasing trend.

Q1.3)

```
# changing the color pallete to brewer palette Dark2
# (outside of geom_point)

ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +
  geom_point(aes(color = species)) +
  scale_color_brewer(palette = "Dark2")
```
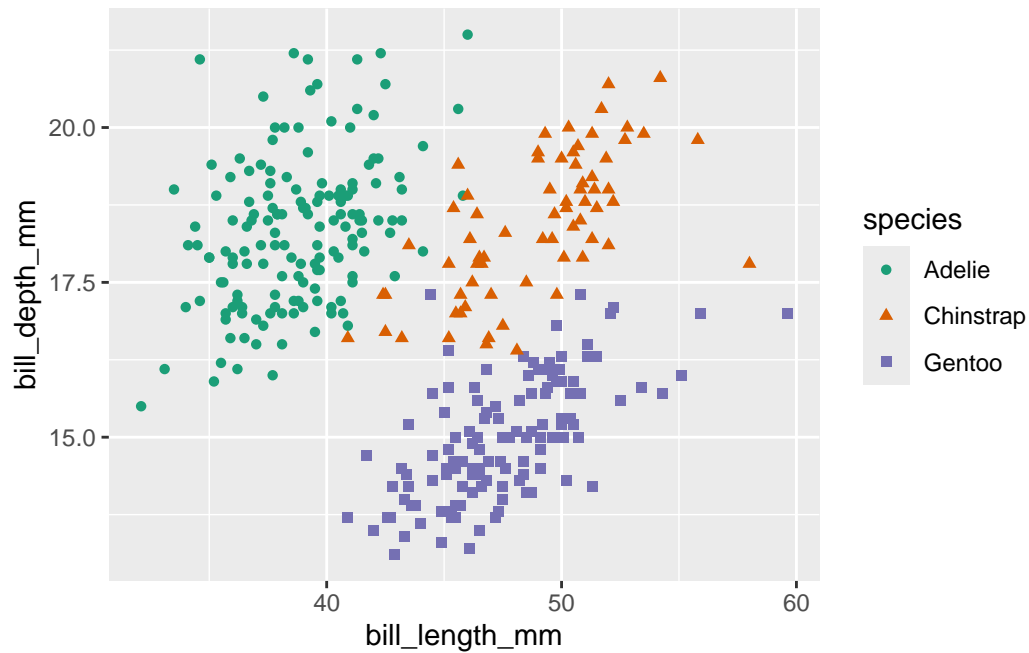
Warning: Removed 2 rows containing missing values or values outside the scale range
(`geom_point()`).

```
# added a shape based on species (within geom_point())

ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +
  geom_point(aes(color = species,
                 shape = species)) +
  scale_color_brewer(palette = "Dark2")
```

Warning: Removed 2 rows containing missing values or values outside the scale range
(`geom_point()`).

```
# changed shapes away from default (outside of geom_point())

ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +
  geom_point(aes(color = species,
                 shape = species)) +
  scale_color_brewer(palette = "Dark2") +
  scale_shape_manual(values = c(8, 11, 14))
```
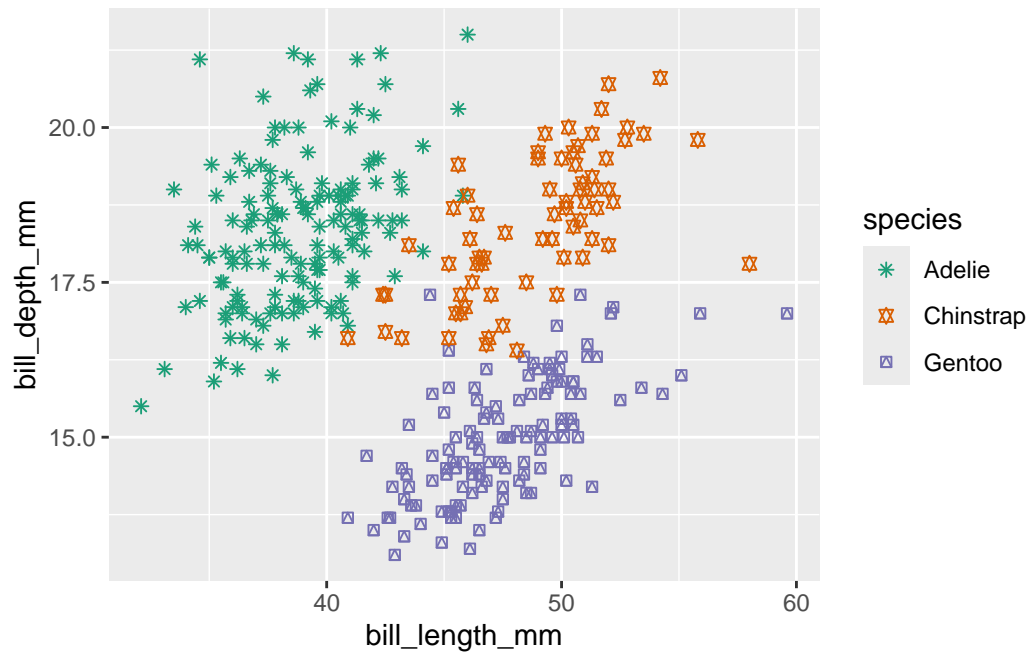
Warning: Removed 2 rows containing missing values or values outside the scale range
(`geom_point()`).

```
# Changed shape codes

ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +
  geom_point(aes(color = species,
                 shape = species)) +
  scale_color_brewer(palette = "Dark2") +
  scale_shape_manual(values = c(21, 22, 24))
```
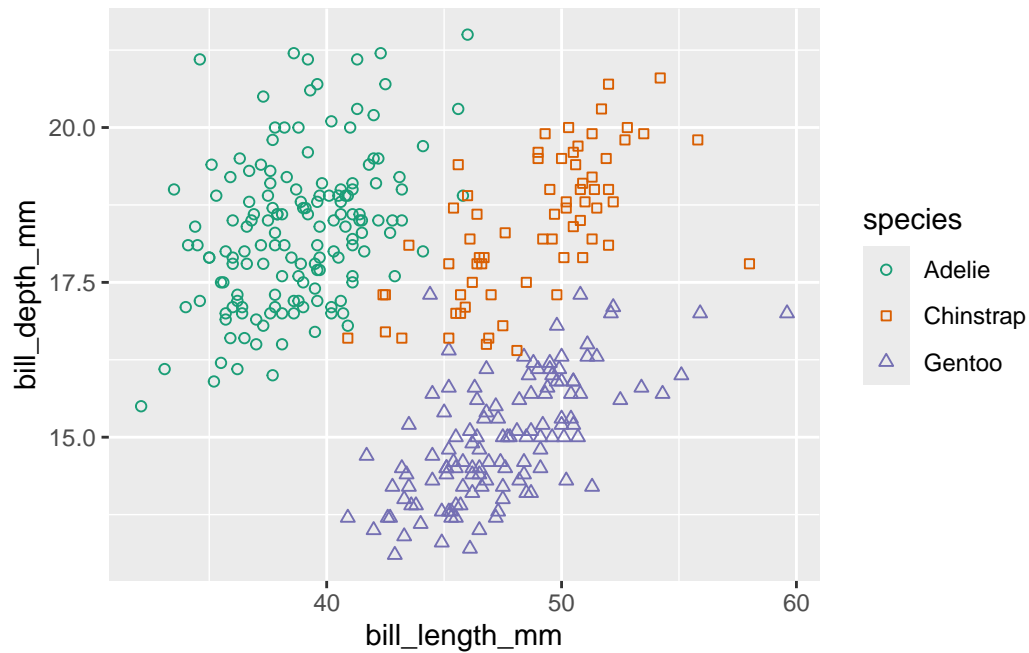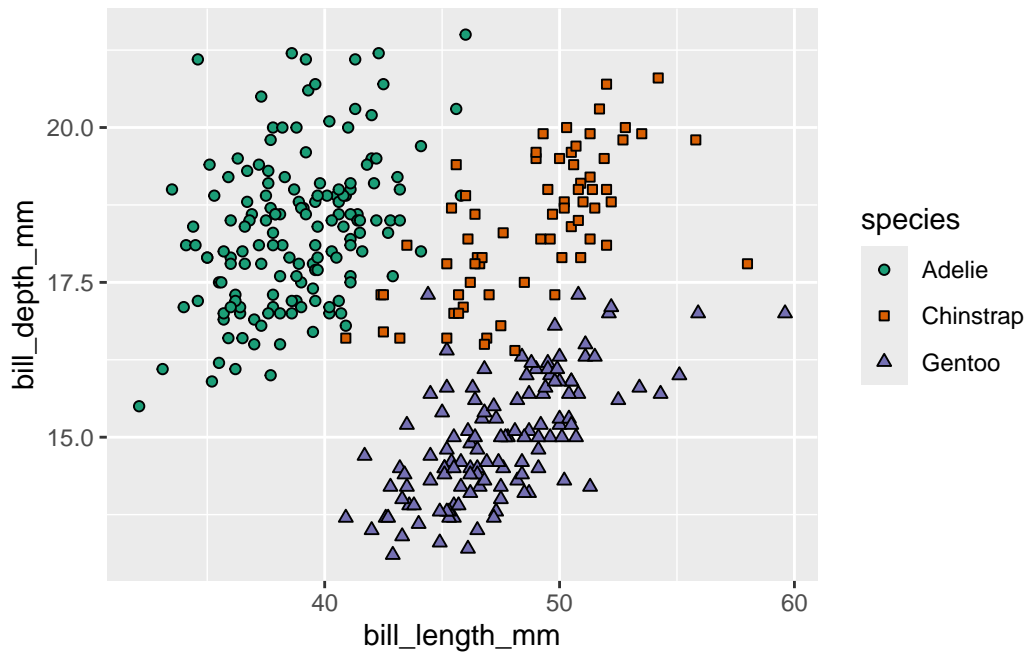
Warning: Removed 2 rows containing missing values or values outside the scale range
(`geom_point()`).

```
# Changed "color = species" to "fill = species" to make shapes solid
# with black outline

ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +
  geom_point(aes(fill = species,
                 shape = species)) +
  scale_fill_brewer(palette = "Dark2") +
  scale_shape_manual(values = c(21, 22, 24))
```
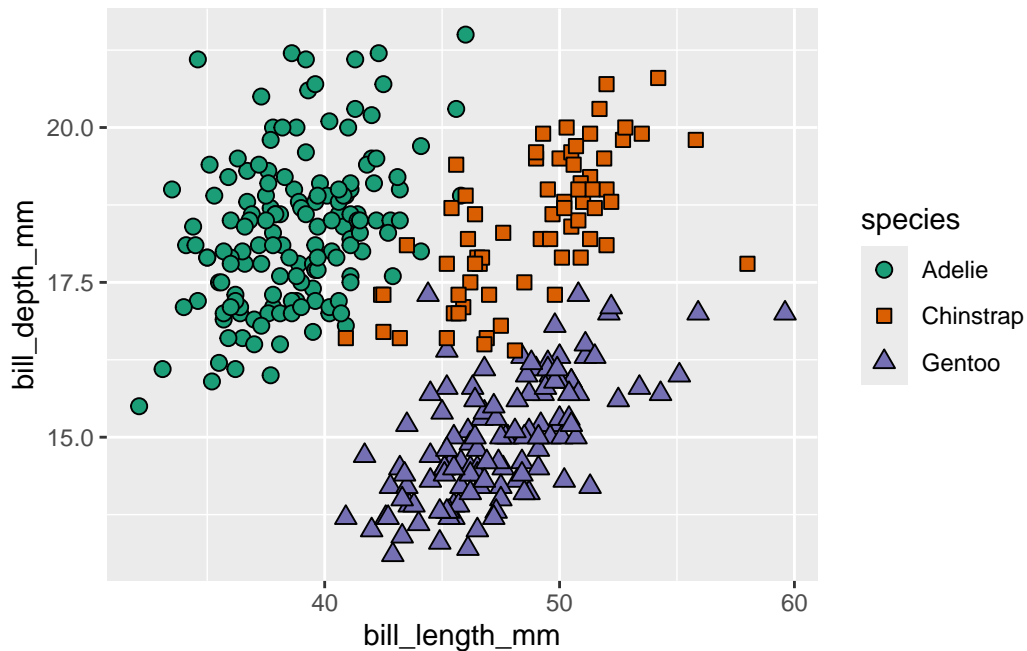
Warning: Removed 2 rows containing missing values or values outside the scale range
(`geom_point()`).

```
# Changed size of points not based on any values
# inside of geom_point() but outside of aes()

ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +
  geom_point(aes(fill = species,
                 shape = species),
             size = 2.5) +
  scale_fill_brewer(palette = "Dark2") +
  scale_shape_manual(values = c(21, 22, 24))
```

Warning: Removed 2 rows containing missing values or values outside the scale range
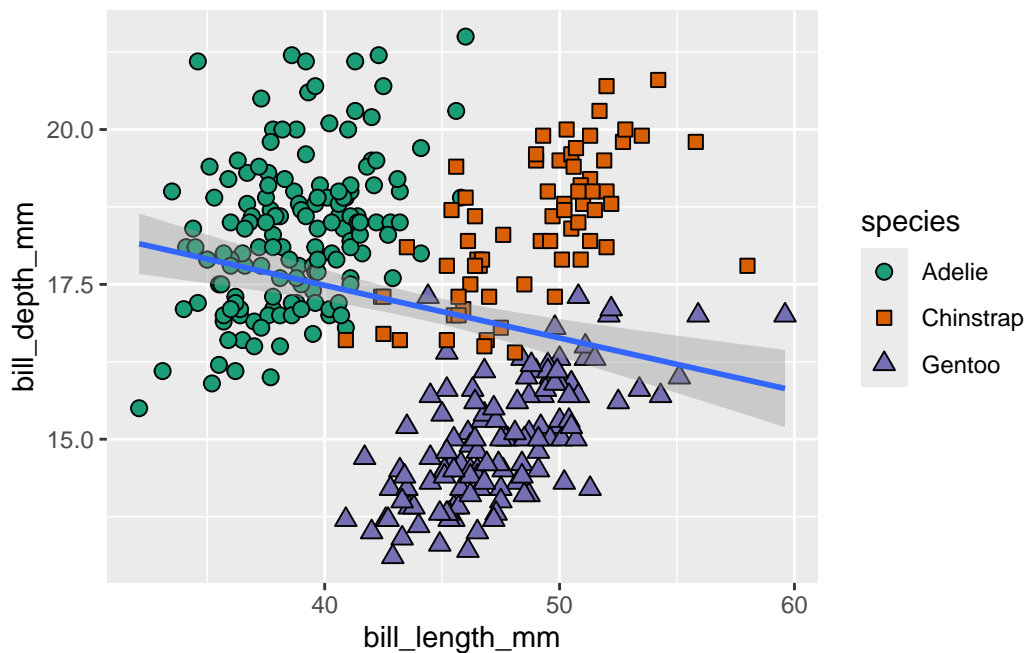(`geom_point()`).

```
# Added geom_smooth() to make line
# independent of color so its for whole group

ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +
  geom_point(aes(fill = species,
                 shape = species),
             size = 2.5) +
  geom_smooth(method = "lm") +
  scale_fill_brewer(palette = "Dark2") +
  scale_shape_manual(values = c(21, 22, 24))
```

`geom_smooth()` using formula = 'y ~ x'

Warning: Removed 2 rows containing non-finite outside the scale range
(`stat_smooth()`).

Warning: Removed 2 rows containing missing values or values outside the scale range
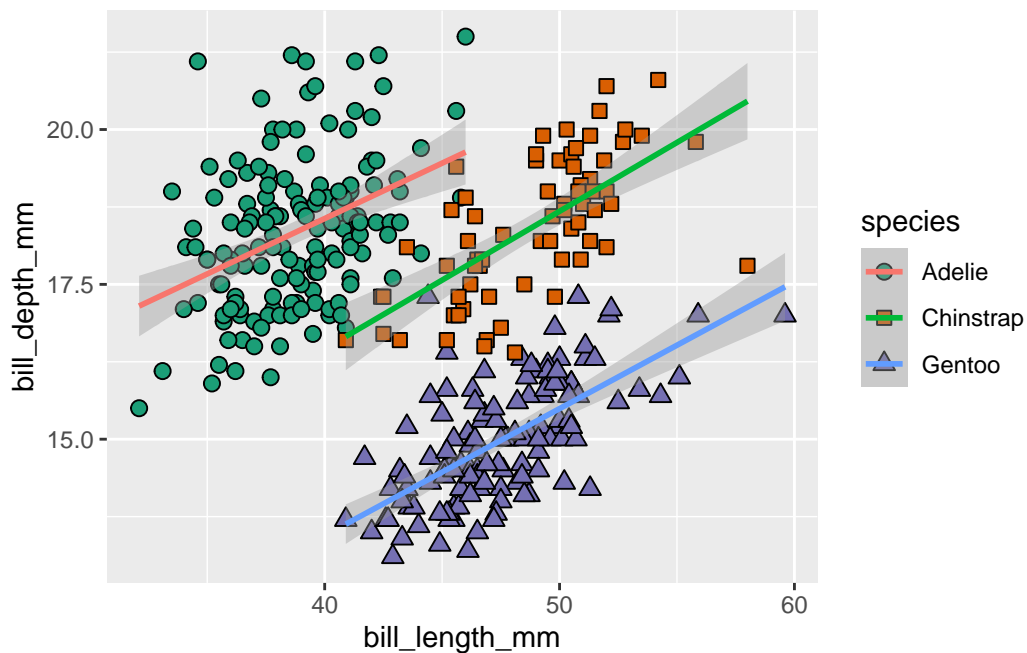(`geom_point()`).

```
# added aes(color = species) to geom_smooth() to make multiple lines
# they are the wrong color though

ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +
  geom_point(aes(fill = species,
                 shape = species),
             size = 2.5) +
  geom_smooth(method = "lm",
              aes(color = species)) +
  scale_fill_brewer(palette = "Dark2") +
  scale_shape_manual(values = c(21, 22, 24))
```

`geom_smooth()` using formula = 'y ~ x'

Warning: Removed 2 rows containing non-finite outside the scale range
(`stat_smooth()`).

Warning: Removed 2 rows containing missing values or values outside the scale range
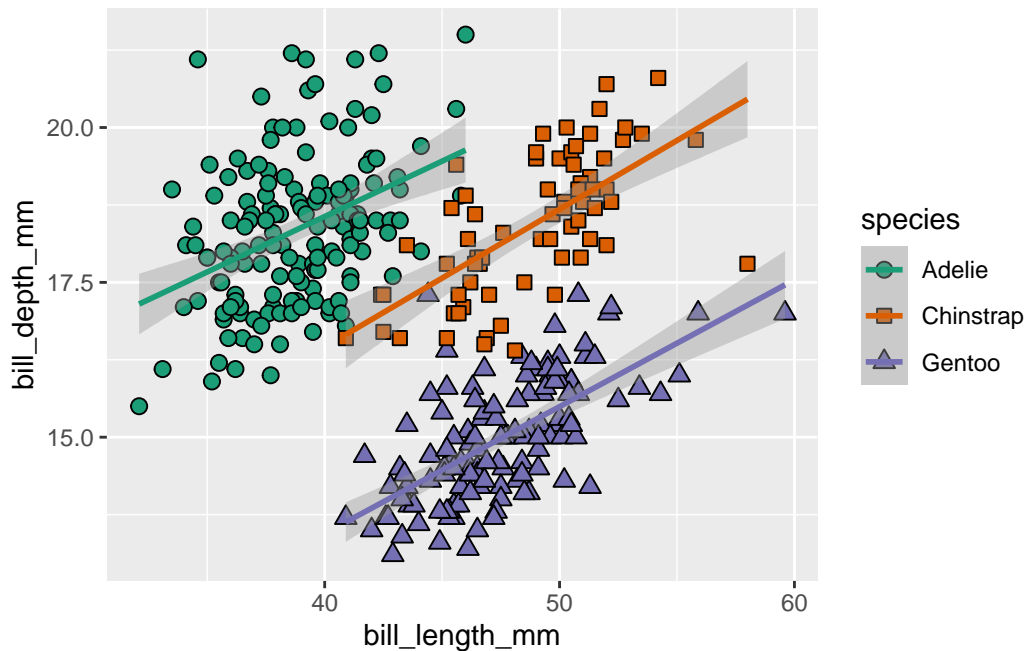(`geom_point()`).

```
# added scale_color_brewer so the new lines would be the right color
# scale_fill_brewer only worked for the points

ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +
  geom_point(aes(fill = species,
                 shape = species),
             size = 2.5) +
  geom_smooth(method = "lm",
              aes(color = species)) +
  scale_fill_brewer(palette = "Dark2") +
  scale_color_brewer(palette = "Dark2") +
  scale_shape_manual(values = c(21, 22, 24))
```

`geom_smooth()` using formula = 'y ~ x'

Warning: Removed 2 rows containing non-finite outside the scale range
(`stat_smooth()`).

Warning: Removed 2 rows containing missing values or values outside the scale range
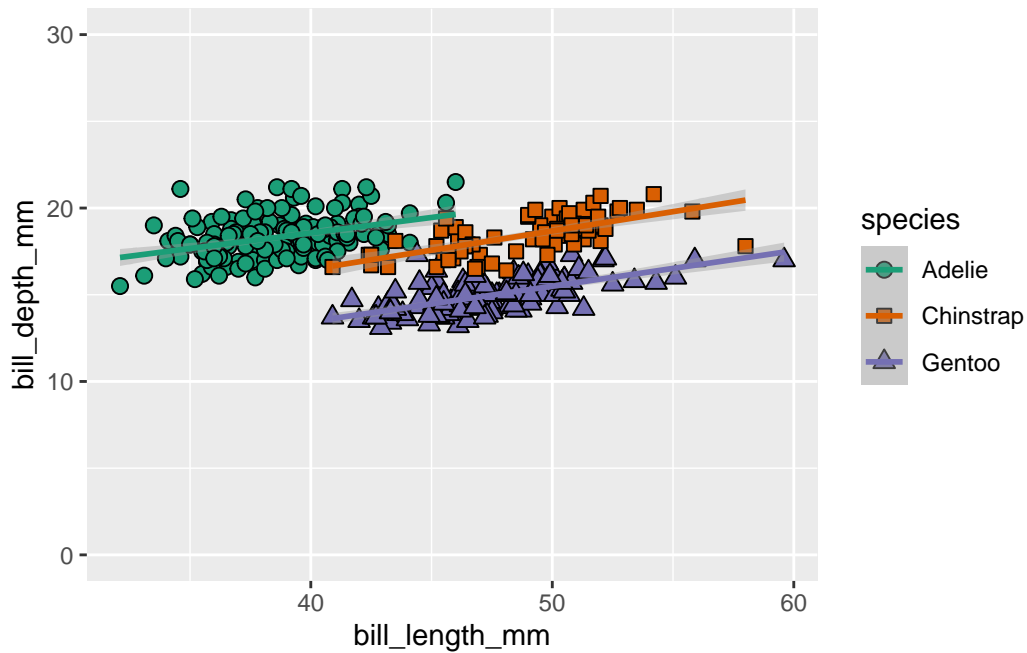(`geom_point()`).

```
# changed the y-axis scale (scrunched the graph)

ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +
  geom_point(aes(fill = species,
                 shape = species),
             size = 2.5) +
  geom_smooth(method = "lm",
              aes(color = species)) +
  scale_fill_brewer(palette = "Dark2") +
  scale_color_brewer(palette = "Dark2") +
  scale_shape_manual(values = c(21, 22, 24)) +
  scale_y_continuous(limits = c(0, 30))
```

`geom_smooth()` using formula = 'y ~ x'

Warning: Removed 2 rows containing non-finite outside the scale range
(`stat_smooth()`).

Warning: Removed 2 rows containing missing values or values outside the scale range
(`geom_point()`).
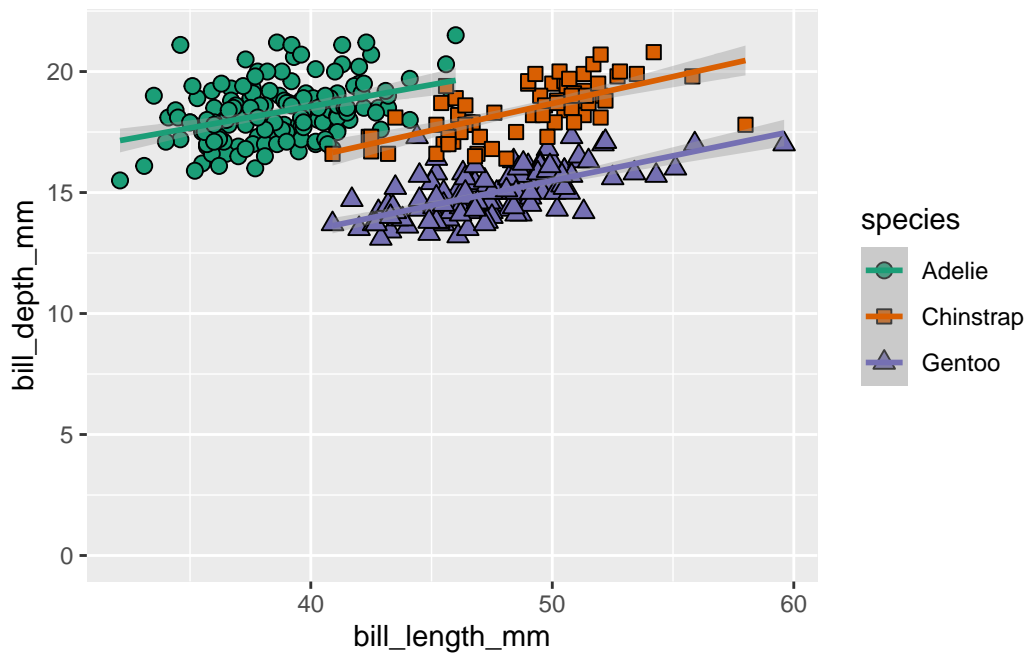
Q1.4)

```
?scale_y_continuous()
```

```
# Changed the y limits to 0 min and r default for max

ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +
  geom_point(aes(fill = species,
                 shape = species),
             size = 2.5) +
  geom_smooth(method = "lm",
              aes(color = species)) +
  scale_fill_brewer(palette = "Dark2") +
  scale_color_brewer(palette = "Dark2") +
  scale_shape_manual(values = c(21, 22, 24)) +
  scale_y_continuous(limits = c(0, NA))
```

`geom_smooth()` using formula = 'y ~ x'

Warning: Removed 2 rows containing non-finite outside the scale range
(`stat_smooth()`).

Warning: Removed 2 rows containing missing values or values outside the scale range
(`geom_point()`).


```
# Change y-axis to be log scaled

ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +
  geom_point(aes(fill = species,
                 shape = species),
             size = 2.5) +
  geom_smooth(method = "lm",
              aes(color = species)) +
  scale_fill_brewer(palette = "Dark2") +
  scale_color_brewer(palette = "Dark2") +
  scale_shape_manual(values = c(21, 22, 24)) +
  scale_y_log10()
```

`geom_smooth()` using formula = 'y ~ x'

Warning: Removed 2 rows containing non-finite outside the scale range
(`stat_smooth()`).

```
Warning: Removed 2 rows containing missing values or values outside the scale range
(`geom_point()`).
```
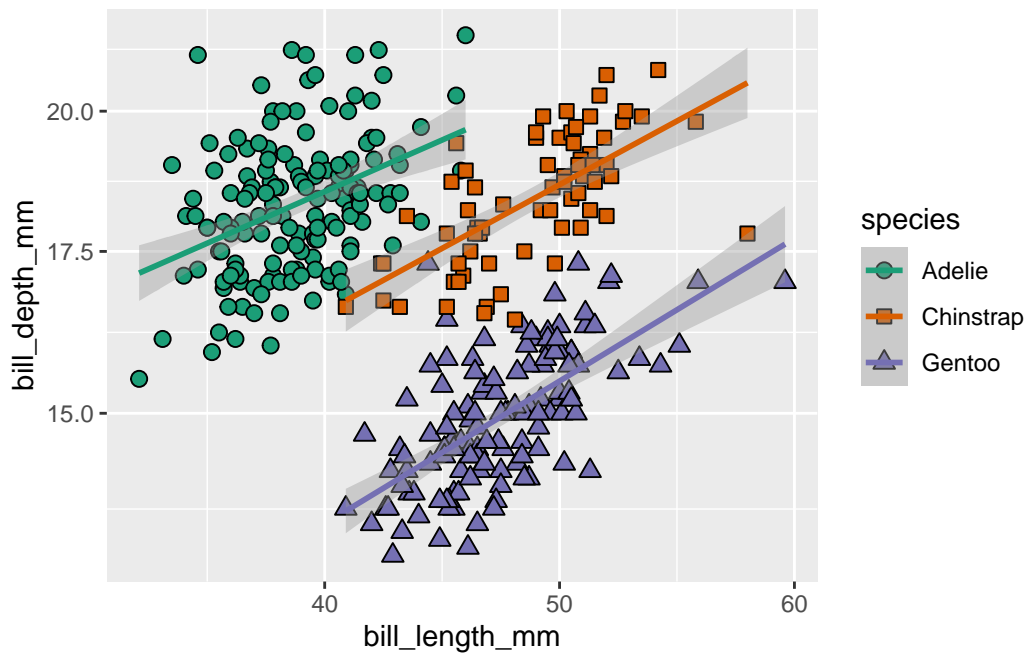


```
# Changed the axis titles

ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +
  geom_point(aes(fill = species,
                 shape = species),
             size = 2.5) +
  geom_smooth(method = "lm",
              aes(color = species)) +
  scale_fill_brewer(palette = "Dark2") +
  scale_color_brewer(palette = "Dark2") +
  scale_shape_manual(values = c(21, 22, 24)) +
  labs(x = "Bill length (mm)",
       y = "Bill depth (mm)")
```

```
`geom_smooth()` using formula = 'y ~ x'
```

```
Warning: Removed 2 rows containing non-finite outside the scale range
(`stat_smooth()`).
```

```
Warning: Removed 2 rows containing missing values or values outside the scale range
(`geom_point()`).
```
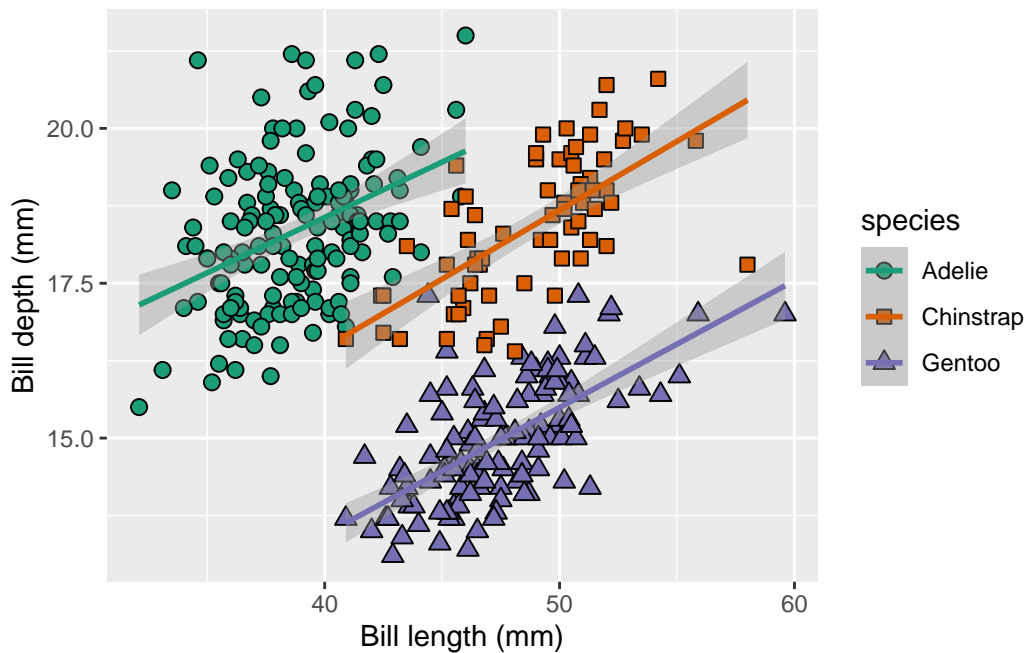

Q1.5)

```
#changed the legend title to be capitalized properly using labs()

ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +
  geom_point(aes(fill = species,
                 shape = species),
             size = 2.5) +
  geom_smooth(method = "lm",
              aes(color = species)) +
  scale_fill_brewer(palette = "Dark2") +
  scale_color_brewer(palette = "Dark2") +
  scale_shape_manual(values = c(21, 22, 24)) +
  labs(x = "Bill length (mm)",
       y = "Bill depth (mm)",
       fill = "Species",
       shape = "Species",
       color = "Species")
```

```
`geom_smooth()` using formula = 'y ~ x'
```

```
Warning: Removed 2 rows containing non-finite outside the scale range
(`stat_smooth()`).
```

```
Warning: Removed 2 rows containing missing values or values outside the scale range
(`geom_point()`).
```



```r
?theme()
```

```r
# Used the theme() function to change x-axis font color and size

ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +
  geom_point(aes(fill = species,
                 shape = species),
             size = 2.5) +
  geom_smooth(method = "lm",
              aes(color = species)) +
  scale_fill_brewer(palette = "Dark2") +
  scale_color_brewer(palette = "Dark2") +
  scale_shape_manual(values = c(21, 22, 24)) +
```
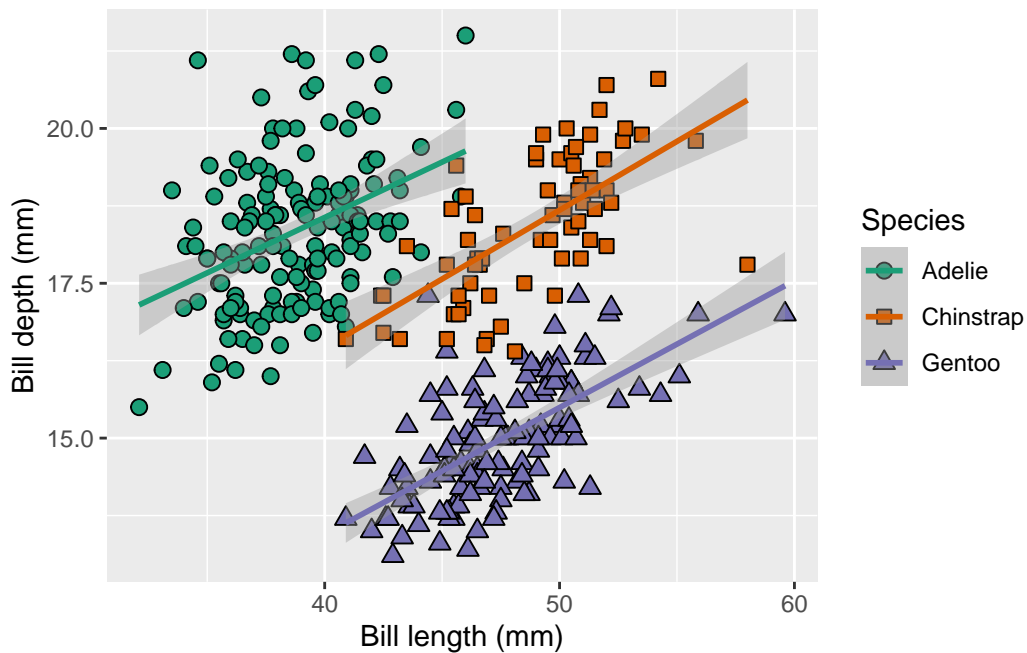
```
  labs(x = "Bill length (mm)",
       y = "Bill depth (mm)",
       fill = "Species",
       color = "Species",
       shape = "Species") +
  theme(axis.title.x = element_text(color = "red",
                                    size = 16))
```

`geom_smooth()` using formula = 'y ~ x'

Warning: Removed 2 rows containing non-finite outside the scale range
(`stat_smooth()`).

Warning: Removed 2 rows containing missing values or values outside the scale range
(`geom_point()`).

Q1.6)

```
?element_text()
```

the options for the face of the text are: "plain", "italic", "bold", and "bold.italic"

```
# added theme_classic() at the bottom

ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +
  geom_point(aes(fill = species,
                 shape = species),
             size = 2.5) +
  geom_smooth(method = "lm",
              aes(color = species)) +
  scale_fill_brewer(palette = "Dark2") +
  scale_color_brewer(palette = "Dark2") +
  scale_shape_manual(values = c(21, 22, 24)) +
  labs(x = "Bill length (mm)",
       y = "Bill depth (mm)",
       fill = "Species",
       color = "Species",
       shape = "Species") +
  theme(axis.title.x = element_text(color = "red",
                                    size = 16)) +
  theme_classic()
```
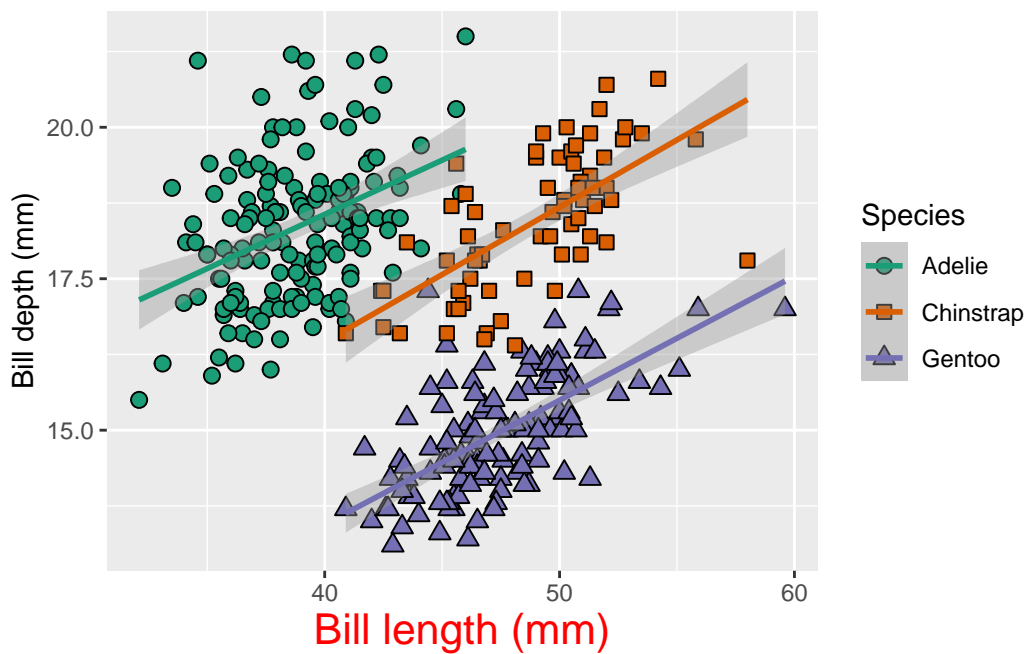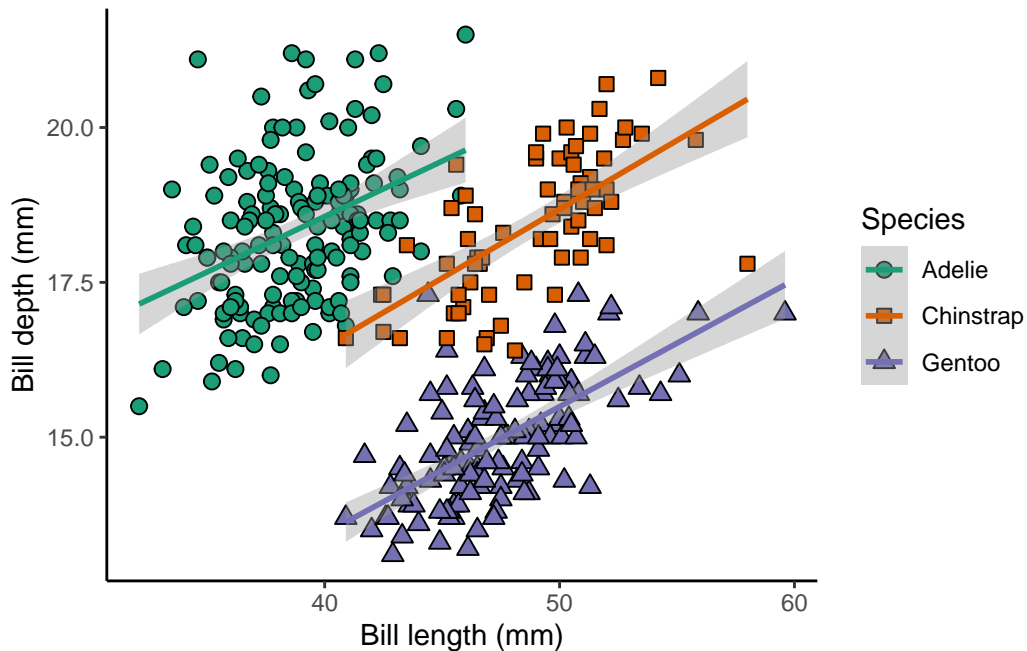
```
`geom_smooth()` using formula = 'y ~ x'
```

```
Warning: Removed 2 rows containing non-finite outside the scale range
(`stat_smooth()`).
```

```
Warning: Removed 2 rows containing missing values or values outside the scale range
(`geom_point()`).
```

```r
# Changed the location of theme_classic() to happen before theme()
# Allowed the larger red text of the x-axis to stay (order matters)

# Also changed around the themes to try other options
# I enjoy classic the most though (bw is good too)

ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +
  geom_point(aes(fill = species,
                 shape = species),
             size = 2.5) +
  geom_smooth(method = "lm",
              aes(color = species)) +
  scale_fill_brewer(palette = "Dark2") +
  scale_color_brewer(palette = "Dark2") +
  scale_shape_manual(values = c(21, 22, 24)) +
  labs(x = "Bill length (mm)",
       y = "Bill depth (mm)",
       fill = "Species",
       color = "Species",
       shape = "Species") +
  theme_classic() +
  theme(axis.title.x = element_text(color = "red",
                                    size = 16))
```
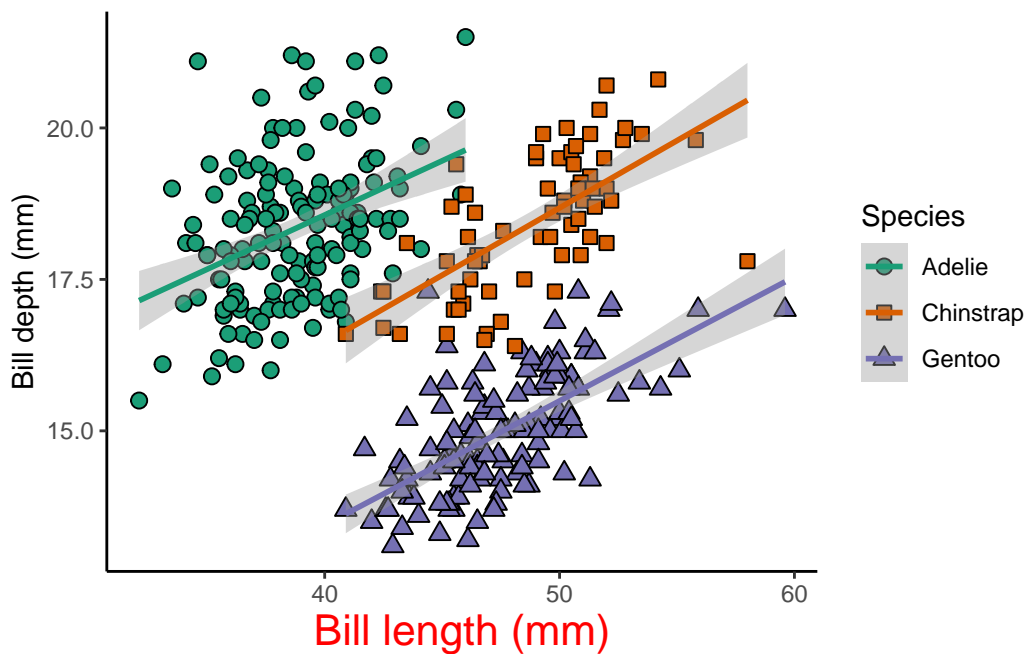
```
`geom_smooth()` using formula = 'y ~ x'
```

```
Warning: Removed 2 rows containing non-finite outside the scale range
(`stat_smooth()`).
```

```
Warning: Removed 2 rows containing missing values or values outside the scale range
(`geom_point()`).
```



```
# Graph bill length vs depth
ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +
  # Add points separated by species
  geom_point(aes(fill = species,
                 shape = species),
             # Change the size of the points
             size = 2.5) +
  # Add a lm line for each species
  geom_smooth(aes(color = species),
              method = "lm") +
  # Customize the fill and color of the geoms
  scale_fill_brewer(palette = "Dark2") +
  scale_color_brewer(palette = "Dark2") +
  # Manually select the shape of the points
```

```
  scale_shape_manual(values = c(21, 22, 24)) +
  # Make nicer axis and legend titles based on the aes() specified above
  labs(x = "Bill length (mm)",
       y = "Bill depth (mm)",
       fill = "Species",
       color = "Species",
       shape = "Species") +
  # Add in a plot title
  ggtitle("Bill depth vs length of three penguin species") +
  # Change the theme to theme_minimal
  theme_minimal() +
  theme(
    # Change the size of axis titles and text (both x and y at once!)
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 12),
    # Change the size of legend title and text
    legend.text = element_text(size = 12),
    legend.title = element_text(size = 12),
    # Place the legend inside the plot, instead of the default outside
    legend.position = "inside",
    # Specify exactly where in x/y space (from 0 to 1) the legend should sit
    legend.position.inside = c(0.9, 0.15),
    # Color (not fill!) the legend - color refers to the outline here
    legend.background = element_rect(color = "black")
  )
```
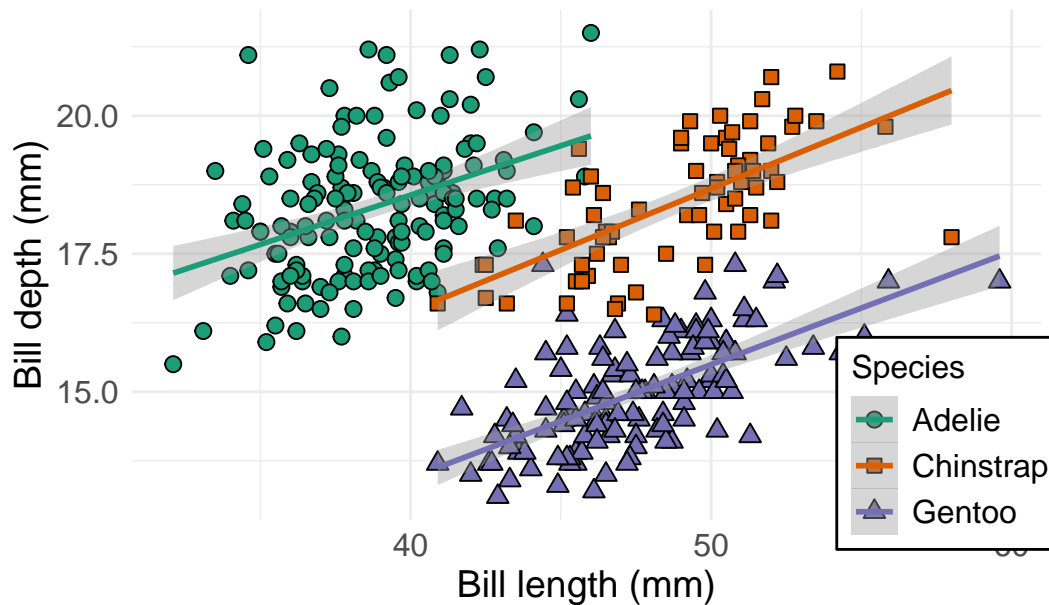
`geom_smooth()` using formula = 'y ~ x'

Warning: Removed 2 rows containing non-finite outside the scale range
(`stat_smooth()`).

Warning: Removed 2 rows containing missing values or values outside the scale range
(`geom_point()`).

Bill depth vs length of three penguin species

```
# Graph bill length vs depth
ggplot(data = penguins, aes(x = species, y = body_mass_g)) +
  # Put violin plots, which show the distribution of points like a histogram/density plot, u
  geom_violin(aes(fill = species),
              # make the violin plots quite transparent
              alpha = 0.25
              ) +
  # Add points separated by species
  geom_jitter(aes(fill = species,
                  shape = species),
              # Change the size of the points
              size = 2.5,
              # Change the default amount that the points are "jittered" in the x (width) and
              width = .1,
              height = 0,
              # Make the points slightly transparent
              alpha = 0.75
              ) +
  # Customize the fill and color of the geoms
  scale_fill_brewer(palette = "Dark2") +
  scale_color_brewer(palette = "Dark2") +
  # Manually select the shape of the points
  scale_shape_manual(values = c(21, 22, 24)) +
```
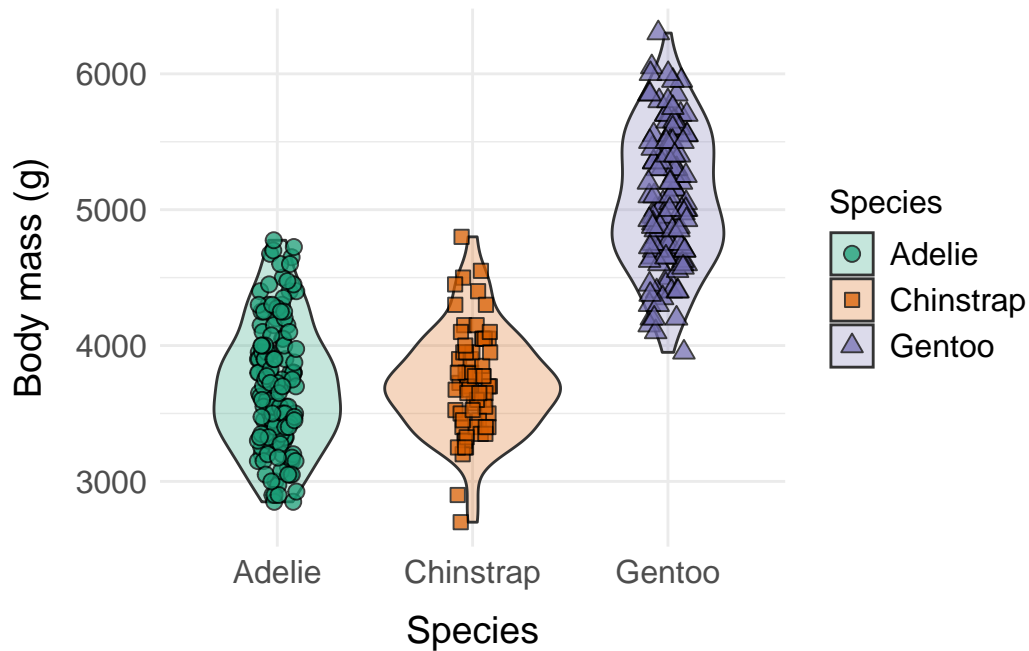
```
  # Make nicer axis and legend titles based on the aes() specified above
  labs(x = "Species",
       y = "Body mass (g)",
       fill = "Species",
       color = "Species",
       shape = "Species") +
  # Change the theme to theme_minimal
  theme_minimal() +
  theme(
    # Change the size of axis titles and text (both x and y at once!)
    axis.title = element_text(size = 14),
    # Add a little bit of space on the top (t) and the right (r) of the x and y axis titles
    # This separates them from the axis text a bit
    axis.title.x = element_text(margin = margin(t = 10)),
    axis.title.y = element_text(margin = margin(r = 10)),
    axis.text = element_text(size = 12),
    # Change the size of legend title and text
    legend.text = element_text(size = 12),
    legend.title = element_text(size = 12)
  )
```

Warning: Removed 2 rows containing non-finite outside the scale range
(`stat_ydensity()`).

Warning: Removed 2 rows containing missing values or values outside the scale range
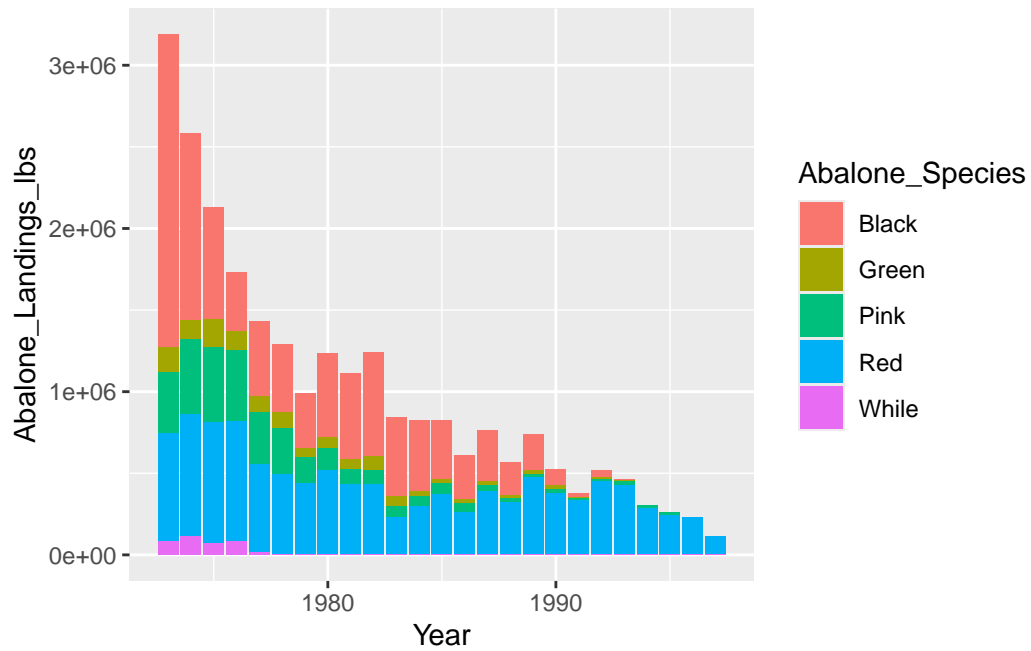(`geom_point()`).

Q2.1)

The audience for our new graph are policymakers. They are in charge of policy relating to the fishing and harvesting of marine organisms.
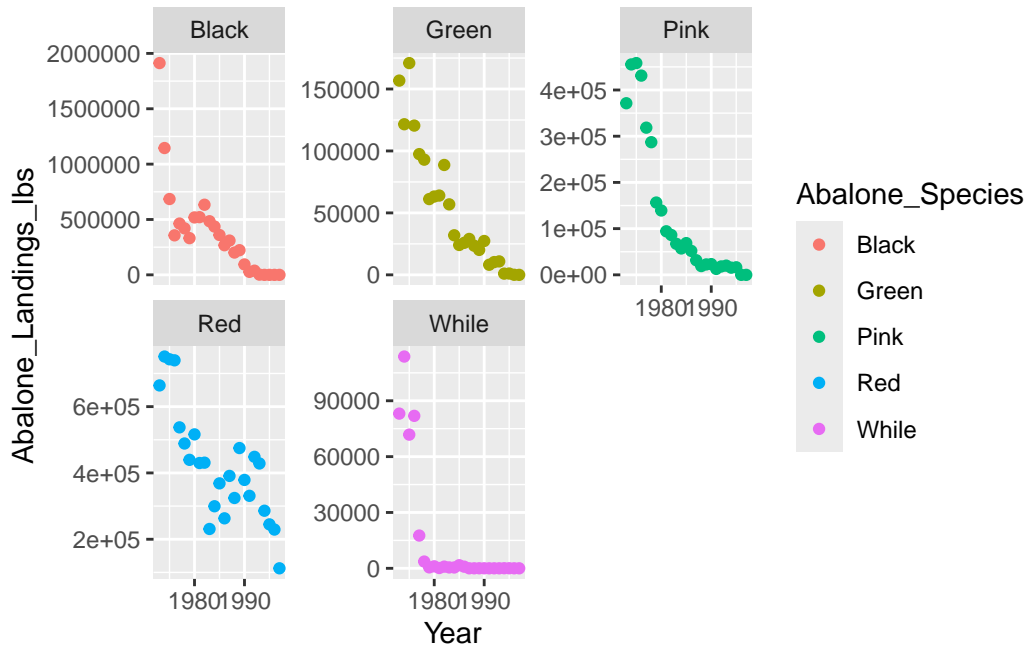
Q2.2

```
# stacked bar graph from previous activity

abalone_landings <- read.csv("abalone_landings.csv")
ggplot(abalone_landings, aes(fill = Abalone_Species , x = Year, y = Abalone_Landings_lbs)) +
  geom_bar(position = "stack", stat = "identity")
```

```
# faceted plot from previous activity
ggplot(abalone_landings, aes(x = Year, y = Abalone_Landings_lbs))+
  geom_point(aes(color =  Abalone_Species)) +
  facet_wrap(Abalone_Species ~ ., scales = "free_y")
```
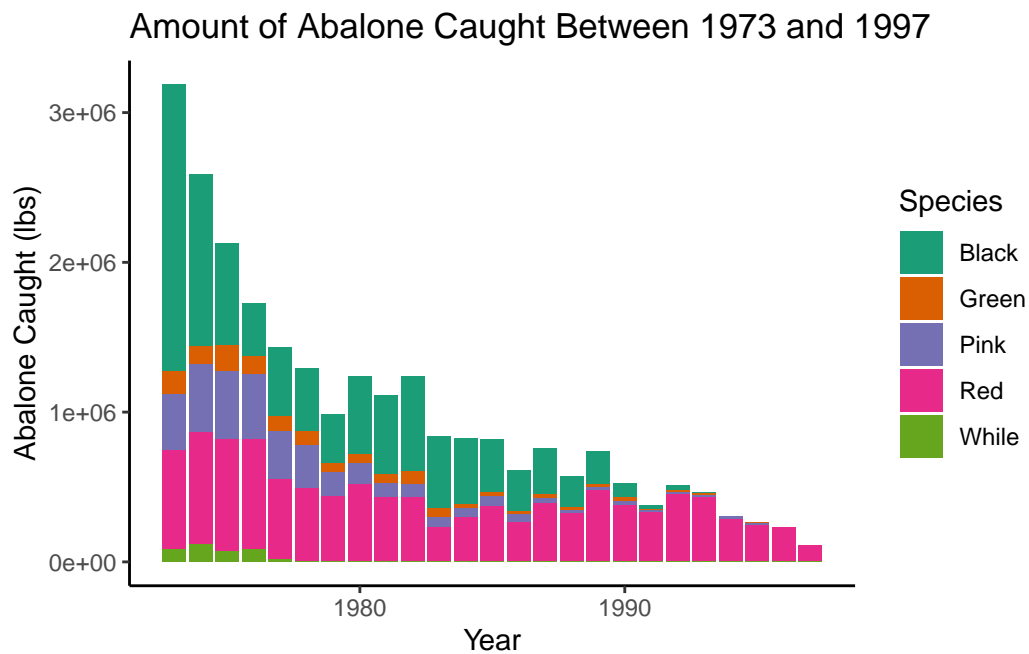
Q2.3

We want to adjust the colors of our first graph in order to better differentiate between the different abalone species. For the second graphs we want to add trendlines in order to better show the general trend for each species. For both we want to change the theme to make the graph less confusing and we want to change the axis and legend titles.

Q2.4

```
# Updated stacked bar chart wich colors changed, title added,
# axes titles changed, legend title changed, and theme changed

ggplot(abalone_landings, aes(fill = Abalone_Species , x = Year, y = Abalone_Landings_lbs)) +
  geom_bar(position = "stack", stat = "identity") +
  scale_fill_brewer(palette = "Dark2") +
  labs(x = "Year",
       y = "Abalone Caught (lbs)",
       fill = "Species",
       color = "Species") +
  ggtitle("Amount of Abalone Caught Between 1973 and 1997") +
  theme_classic()
```

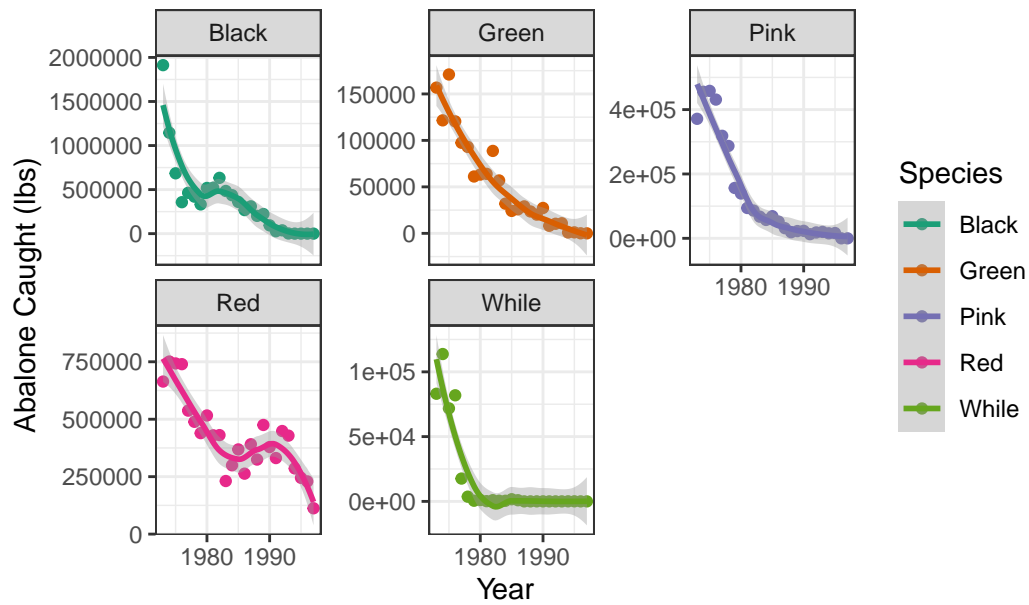## Amount of Abalone Caught Between 1973 and 1997


```
# Updated faced graphs with colors changed, smooth trendline,
# added, axes title changed, legend title changed, and theme added

ggplot(abalone_landings, aes(x = Year, y = Abalone_Landings_lbs))+
  geom_point(aes(color =  Abalone_Species)) +
  facet_wrap(Abalone_Species ~ ., scales = "free_y") +
  geom_smooth(aes(color = Abalone_Species)) +
  scale_color_brewer(palette = "Dark2") +
   labs(x = "Year",
       y = "Abalone Caught (lbs)",
       color = "Species") +
  ggtitle("Amount of Abalone Caught By Species") +
  theme_bw()
```

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

Amount of Abalone Caught By Species

Q2.5)

We hope that our audience will see that abalone as a whole have decreased at a rapid rate and that different species have been effected more strongly than others. We hope that this will lead to policy being put in place to protect abalone and allow their populations to recover.