

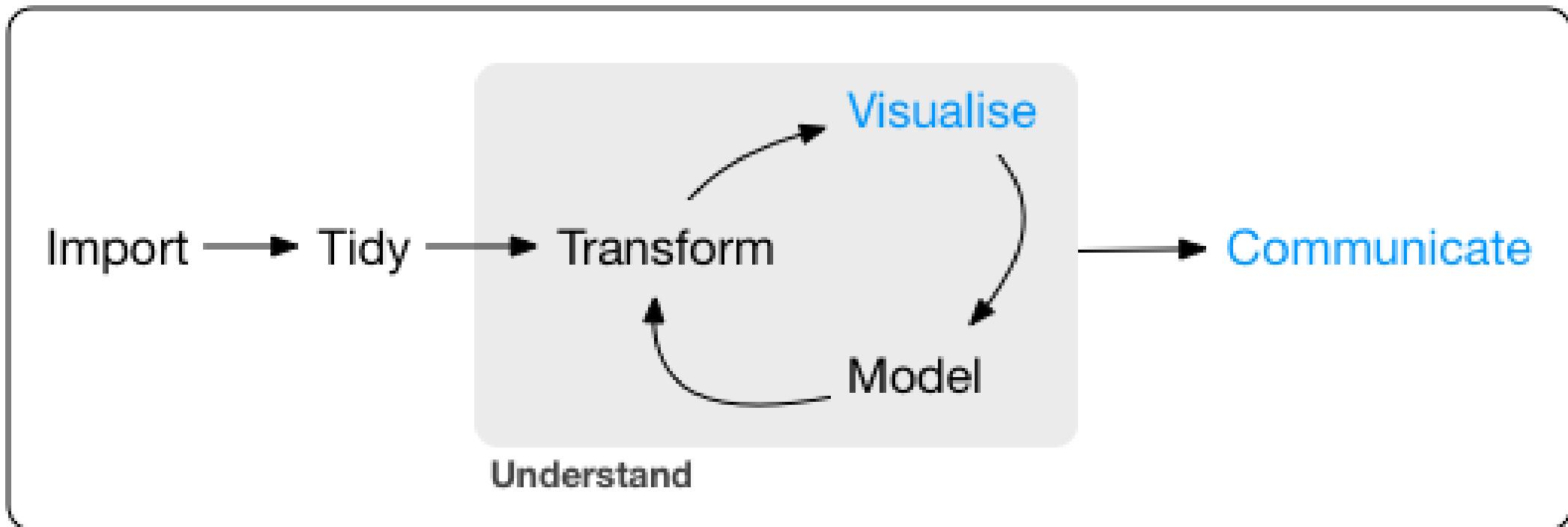
Data Visualisation

TOPIC - 4 | AJAY KOLI

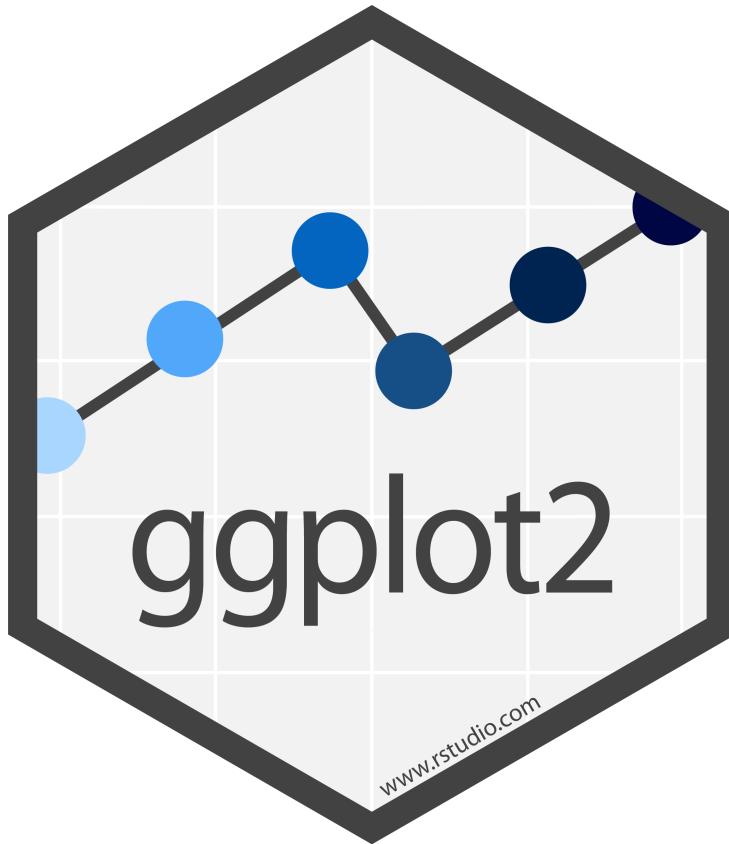
2020-10-20 (updated on 2021-02-15)



Course Progress



Program



ggplot2 package

Packages required:

```
library(palmerpenguins) # to access penguin data  
library(tidyverse) # to use ggplot2 pkg
```

- Packages recommended:

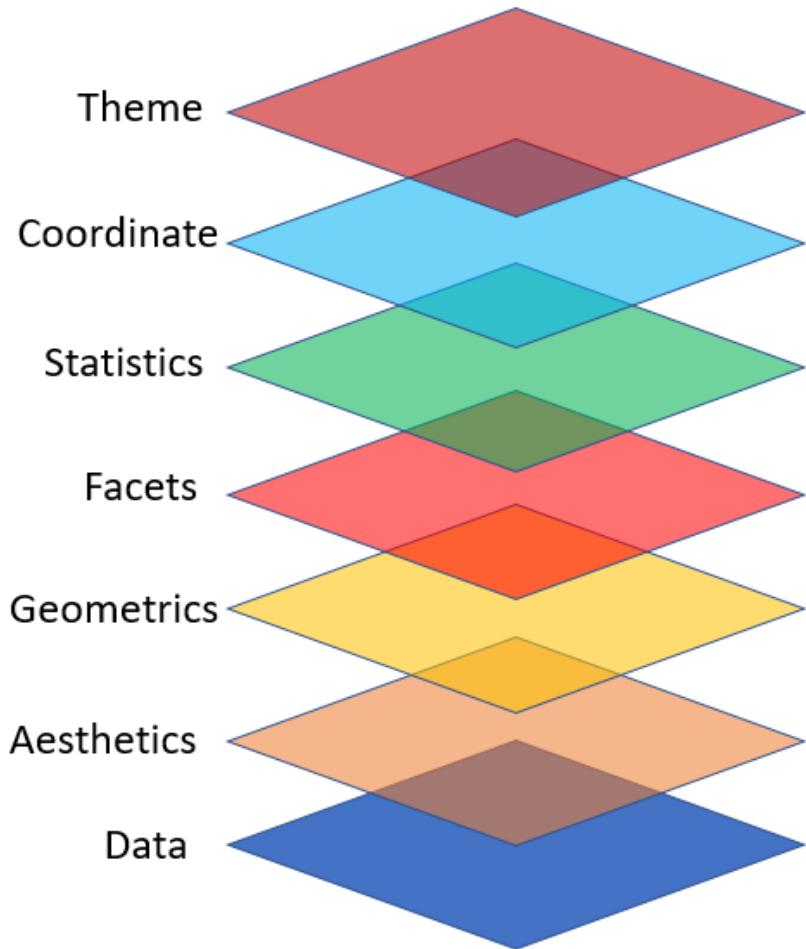
```
install.packages(c(  
  "directlabels", "dplyr", "gameofthrones", "ggforce", "gghighlight",  
  "ggnewscale", "ggplot2", "ggraph", "ggrepel", "ggtext", "ggthemes",  
  "hexbin", "mapproj", "maps", "munsell", "ozmaps", "paletteer",  
  "patchwork", "rmapshaper", "scico", "seriation", "sf", "stars",  
  "tidygraph", "tidyrr", "wesanderson"  
))
```

ggplot2:

- "is a system for declaratively creating graphics, based on [The Grammar of Graphics](#)"
- [Hadley Wickham](#)



Source: [tidyverse.org](#)

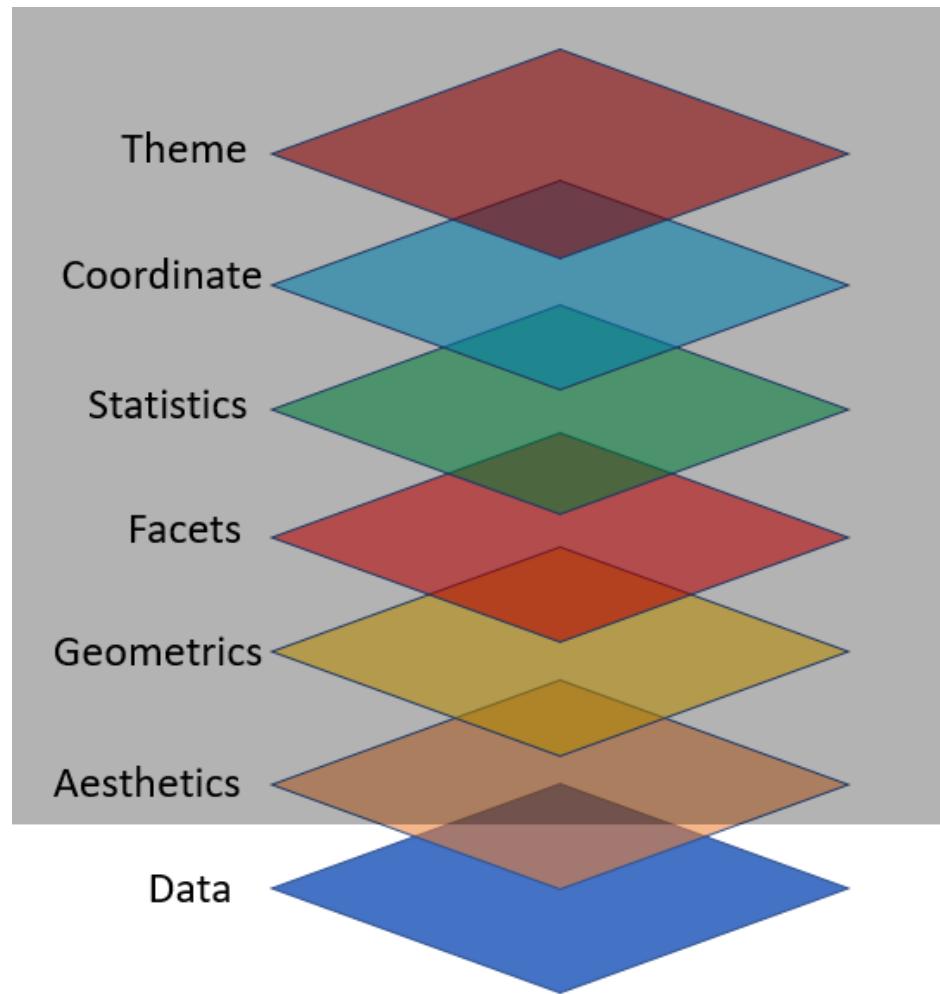


Artwork: [Allison Horst](#) & Image source: [Eric Chow](#)

Task

Codes

Output



Task

Codes

Output

```
ggplot(data = penguins)
```

Task

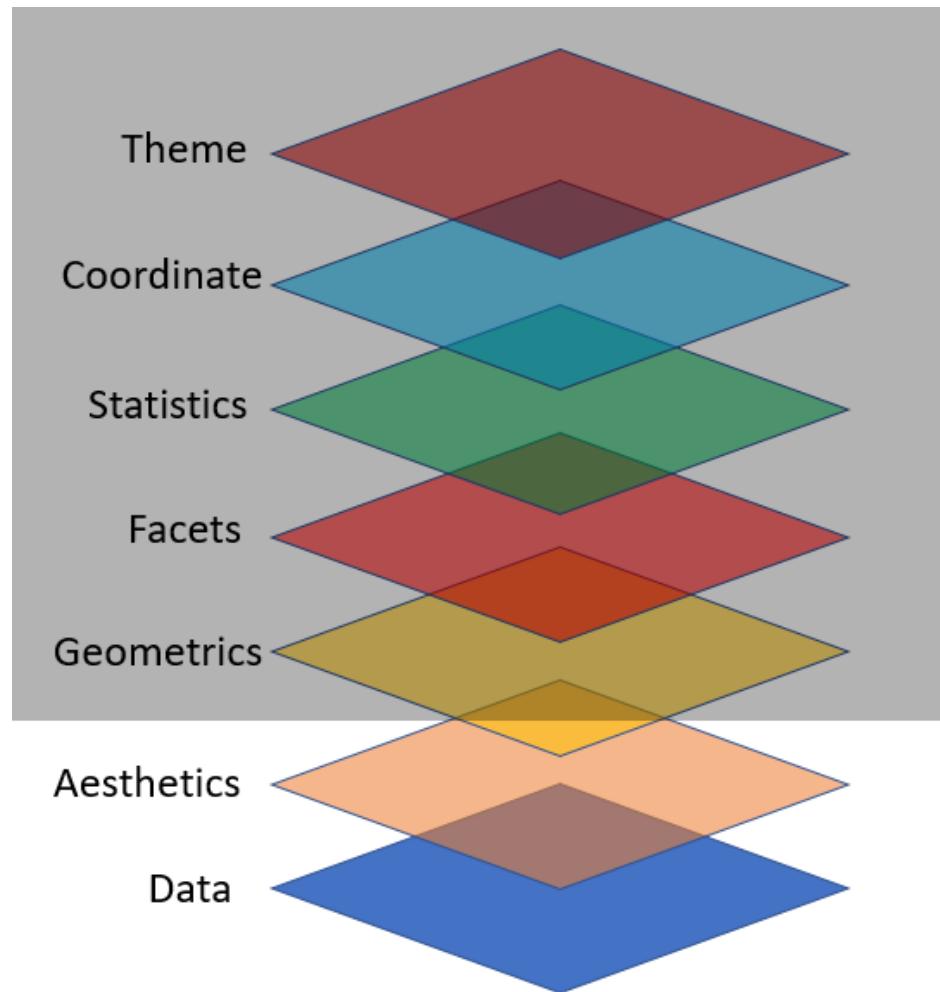
Codes

Output

Task

Codes

Output



Task

Codes

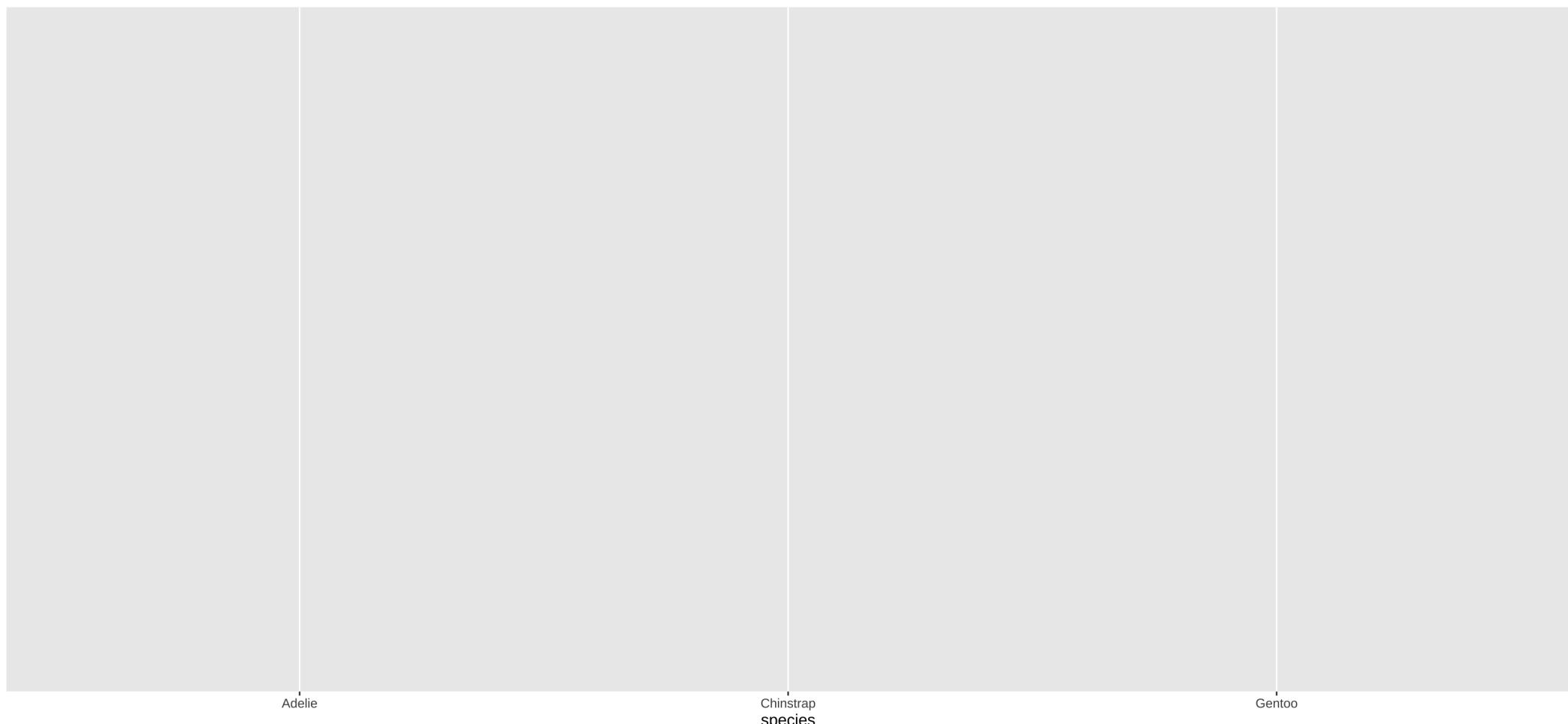
Output

```
ggplot(data = penguins,  
       mapping = aes(x = species))
```

Task

Codes

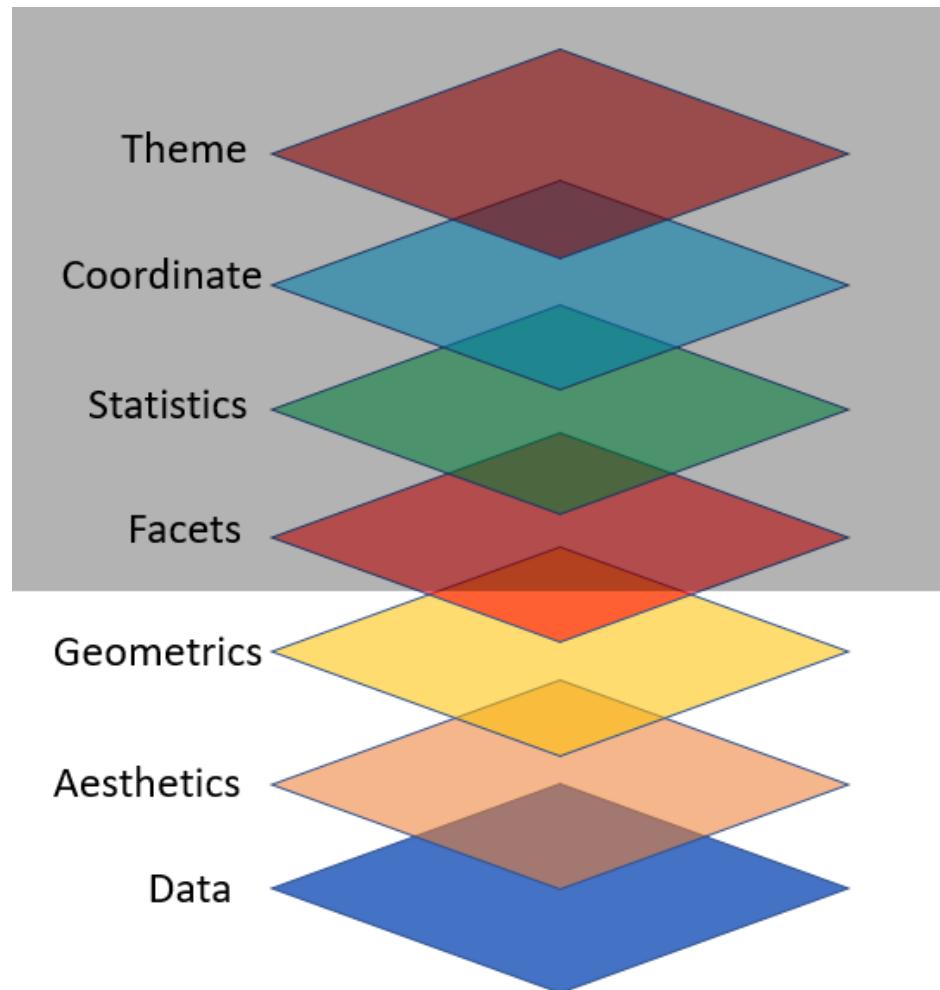
Output



Task

Codes

Output



Task

Codes

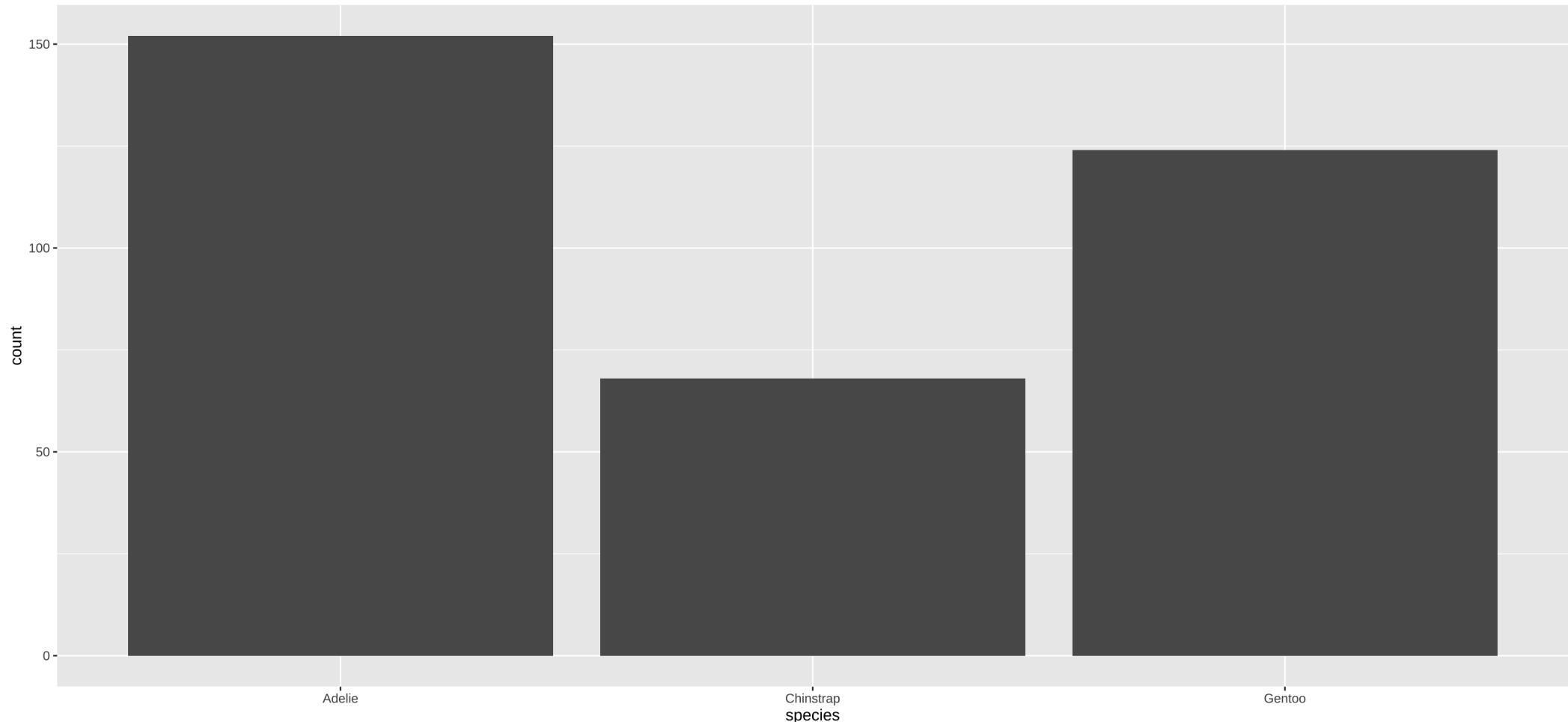
Output

```
ggplot(data = penguins,  
       mapping = aes(x = species)) +  
  geom_bar()
```

Task

Codes

Output



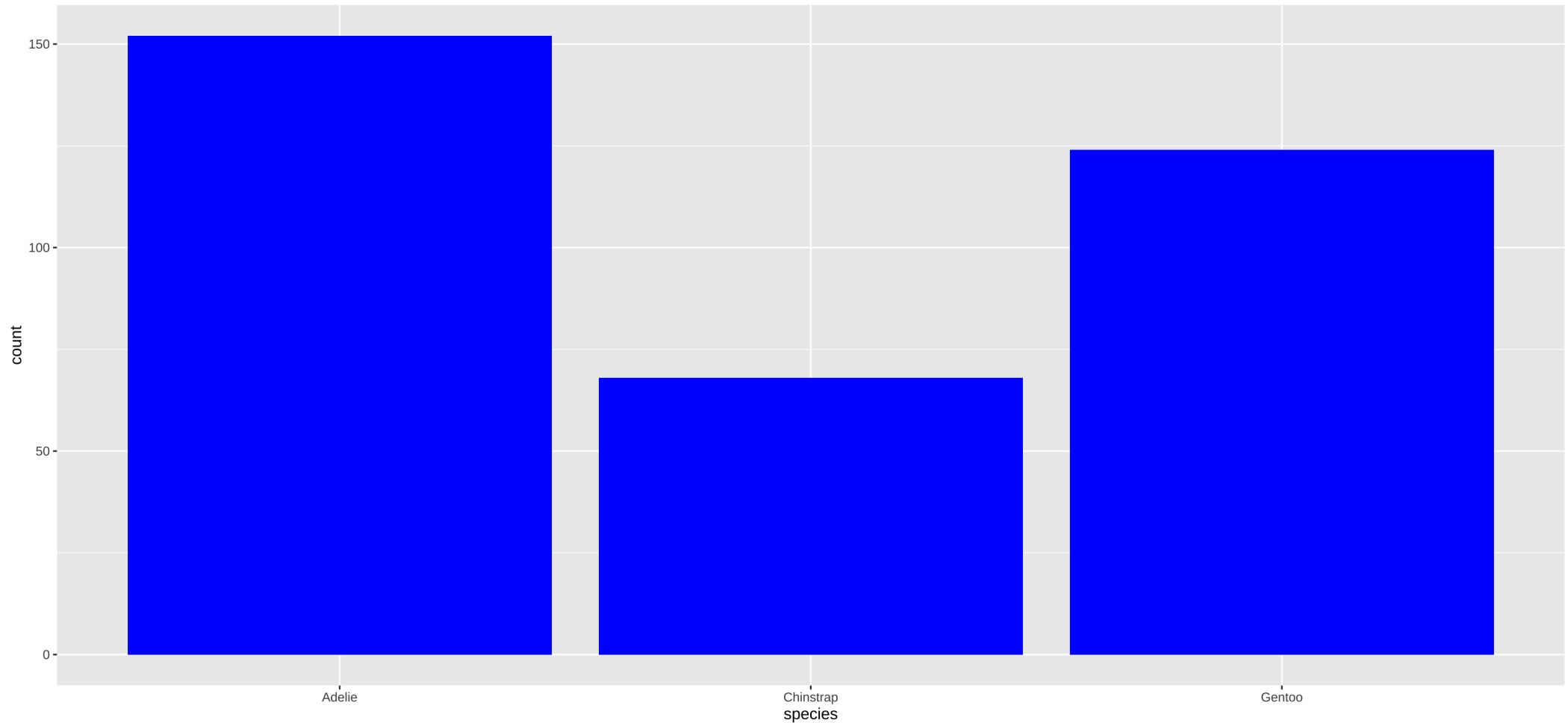
How to export plot to your computer?

```
ggplot(data = penguins,  
       mapping = aes(x = species)) +  
  geom_bar() +  
  ggsave("peng-species.pdf") # also try jpg/jpeg/png
```

Saving 15 x 7 in image

How to add color to bars?

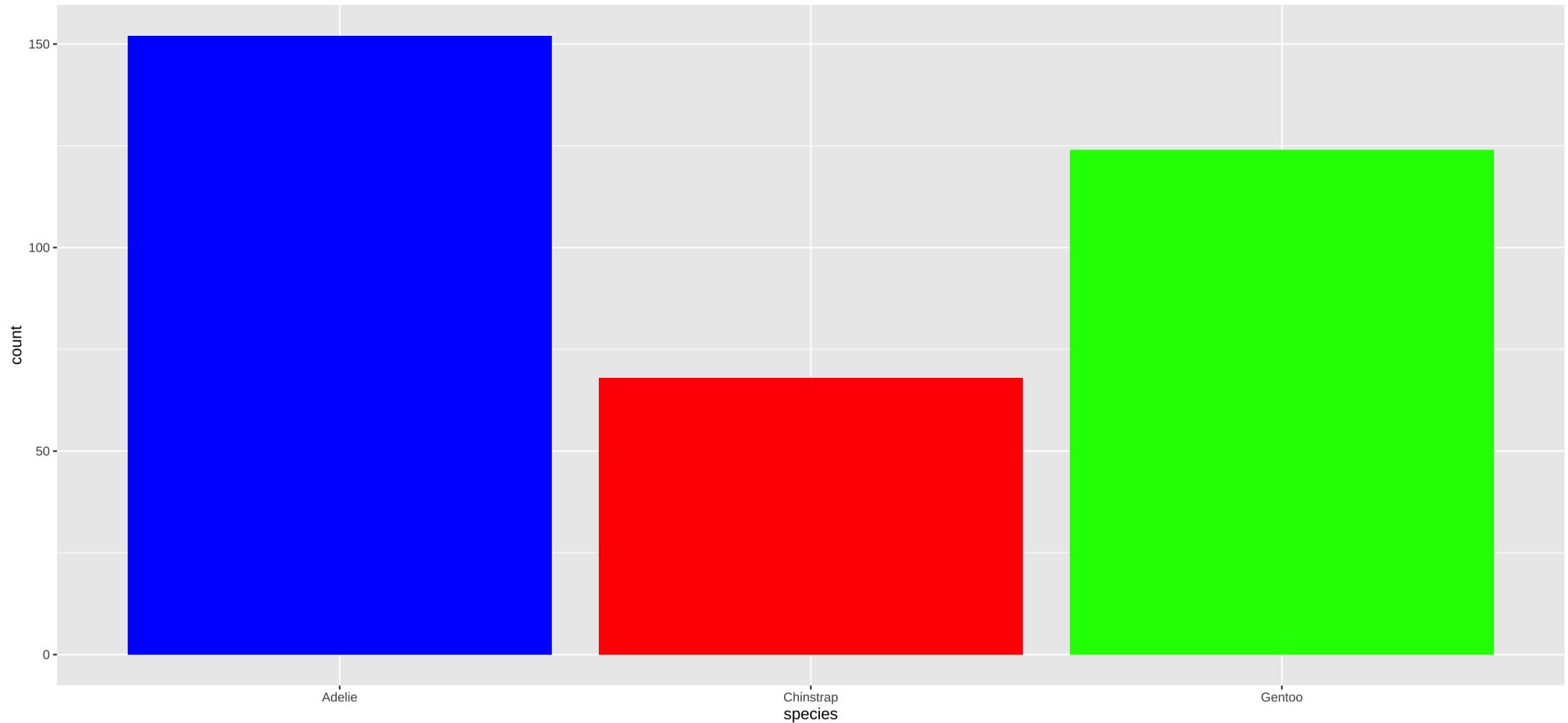
```
ggplot(data = penguins,  
       mapping = aes(x = species)) +  
  geom_bar(fill = "blue")
```



```
ggplot(data = penguins,  
       mapping = aes(x = species)) +  
  geom_bar(fill = c("blue", "red", "green"))  
  
# color names should be equal to the factor levels  
# in case of factor species levels are three  
# Adele, Chinstrap & Gentoo
```

Codes

Output



How to add text labels on the plot?

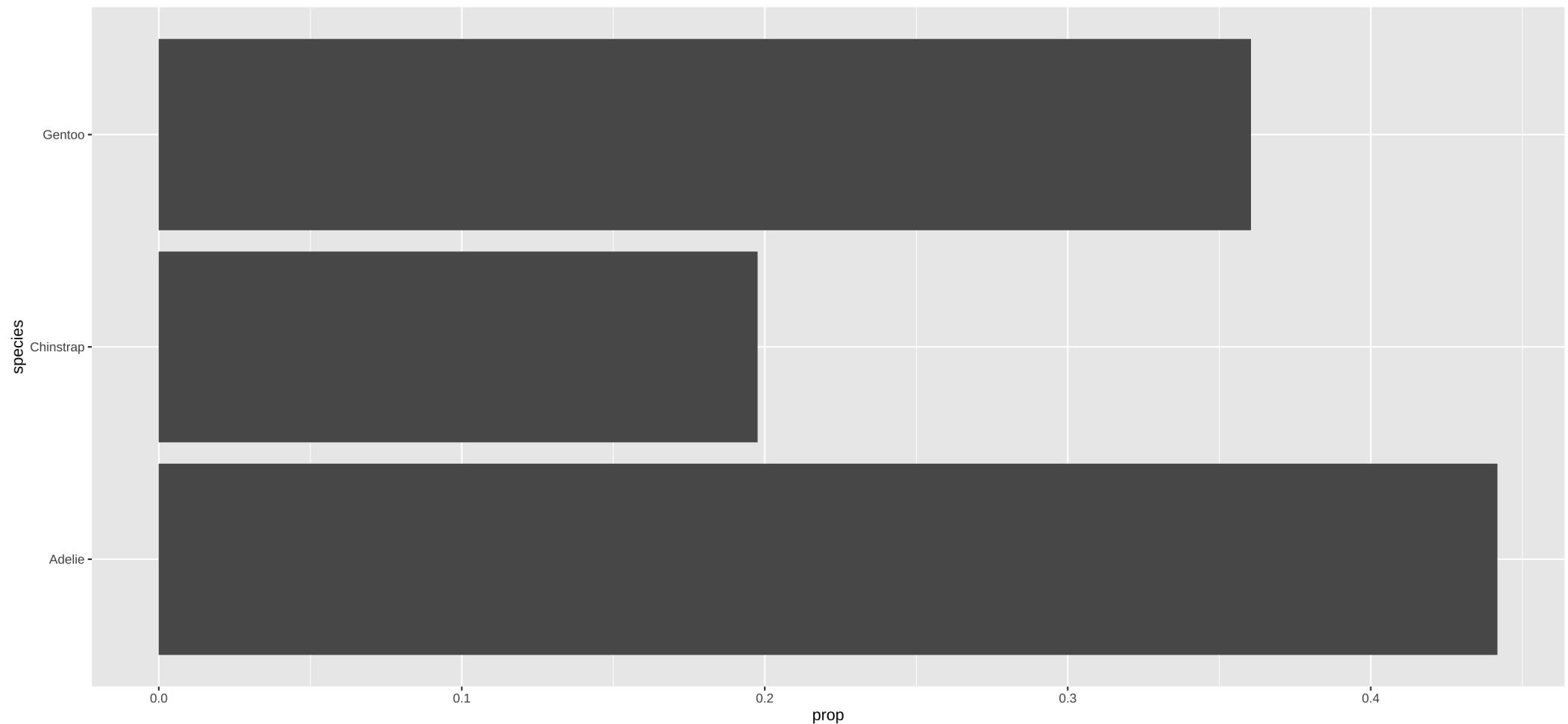
```
penguins %>%  
  count(species)
```

```
# A tibble: 3 x 2
  species      n
  <fct>    <int>
1 Adelie     152
2 Chinstrap   68
3 Gentoo     124
```

```
penguins %>%  
  count(species) %>%  
  mutate(prop = n / sum(n))
```

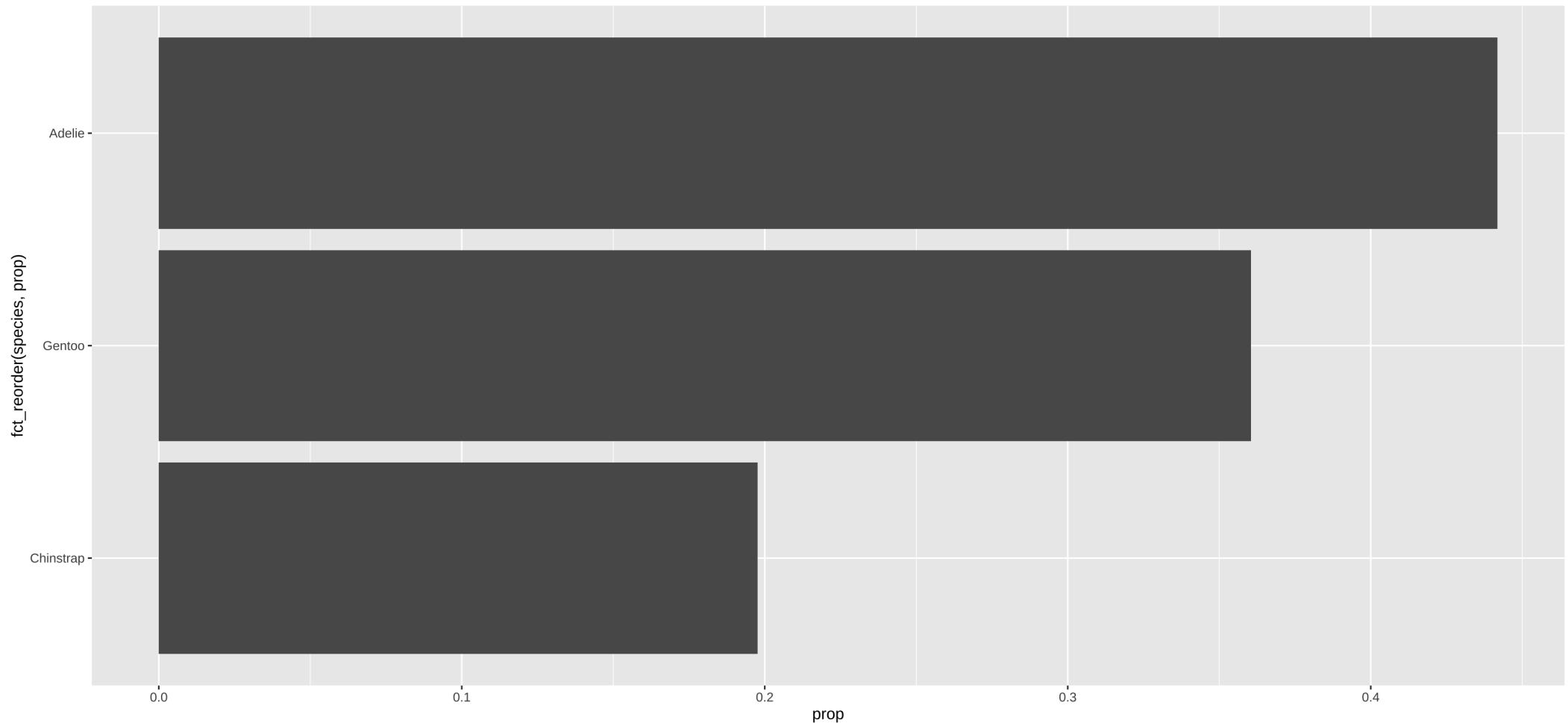
```
# A tibble: 3 x 3
  species      n    prop
  <fct>     <int> <dbl>
1 Adelie     152  0.442
2 Chinstrap   68  0.198
3 Gentoo     124  0.360
```

```
penguins %>%  
  count(species) %>%  
  mutate(prop = n / sum(n)) %>%  
  ggplot(aes(x = prop, y = species)) +  
    geom_col()
```



How to reorder the factor levels in the plot?

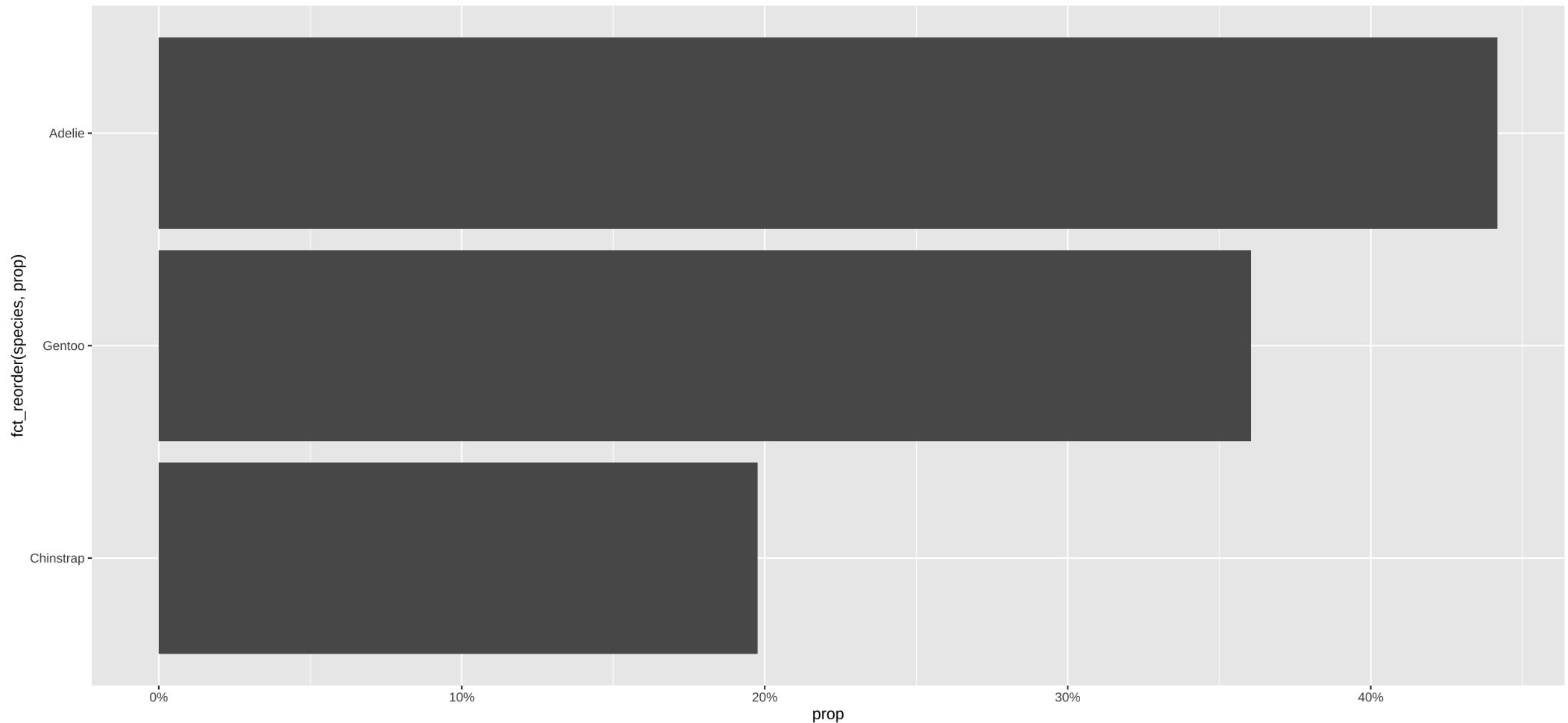
```
penguins %>%  
  count(species) %>%  
  mutate(prop = n / sum(n)) %>%  
  ggplot(aes(x = prop,  
             y = fct_reorder(species, prop))  
         )  
    ) +  
  geom_col()
```



How to add percent values on the plot?

```
library(scales)

penguins %>%
  count(species) %>%
  mutate(prop = n / sum(n)) %>%
  ggplot(aes(x = prop, y = fct_reorder(species, prop))) +
  geom_col() +
  scale_x_continuous(labels = label_percent(accuracy = 1))
```

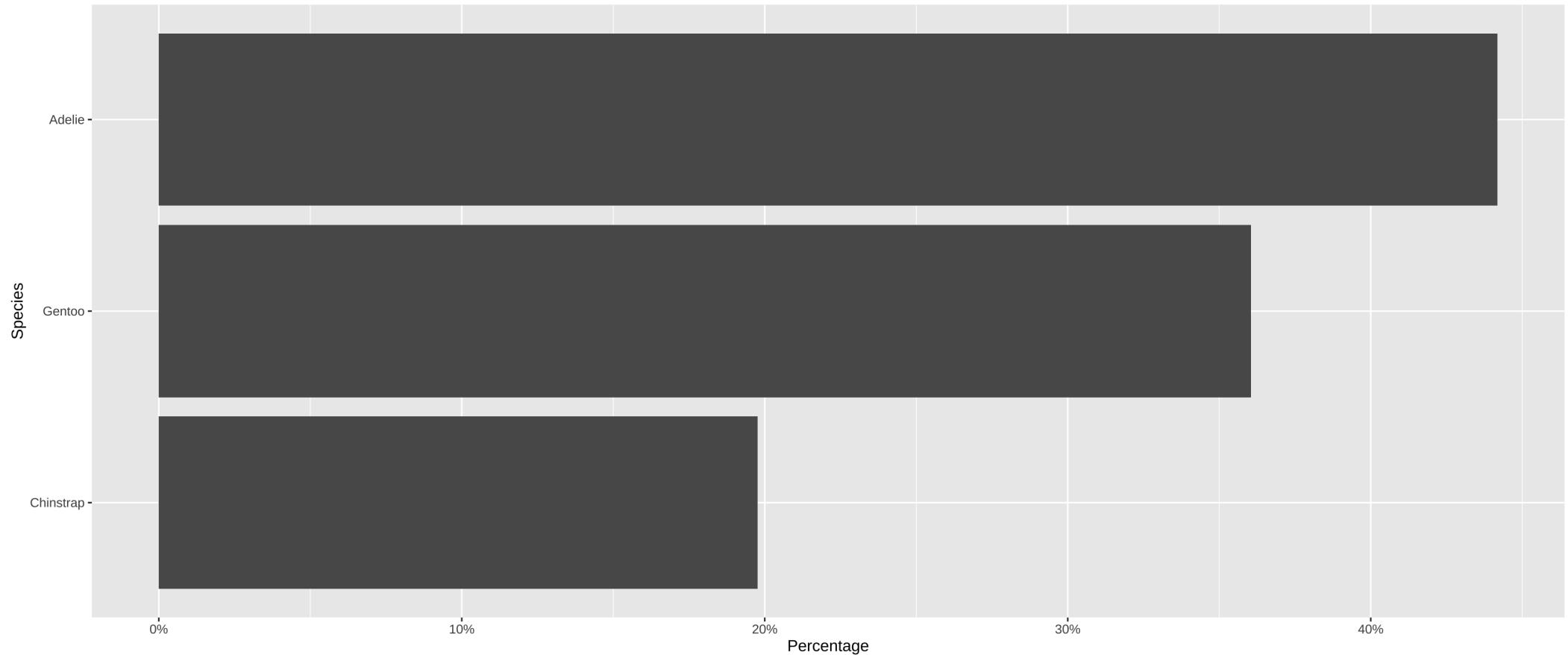


How to add plot title, subtitle,
comments, and axis labels?

```
penguins %>%  
  count(species) %>%  
  mutate(prop = n / sum(n)) %>%  
  ggplot(aes(x = prop, y = fct_reorder(species, prop))) +  
  geom_col() +  
  scale_x_continuous(labels = label_percent(accuracy = 1)) +  
  labs(  
    x = "Percentage",  
    y = "Species",  
    title = "Species distribution of penguins",  
    subtitle = "Adelie, Gentoo, and Chinstrap Penguins at Palmer Station",  
    caption = "Source: allisonhorst.github.io/palmerpenguins")
```

Species distribution of penguins

Adelie, Gentoo, and Chinstrap Penguins at Palmer Station LTER

Source: allisonhorst.github.io/palmerpenguins

How to add color in plot based upon
another variable?

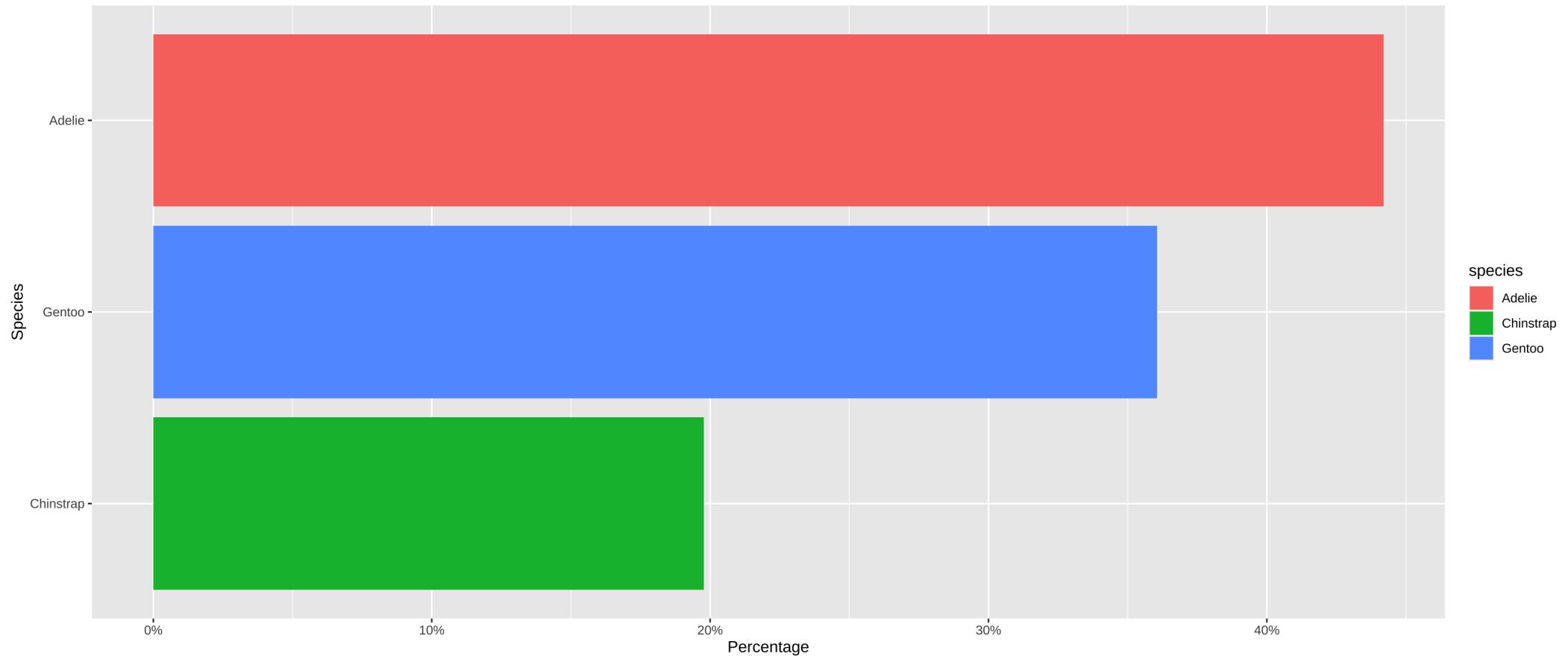
- `library(RColorBrewer)` use RColorBrewer package.
- `scale_fill_brewer()` function for box plot, bar plot, violin plot, dot plot, etc.
- `scale_color_brewer()` function for lines and points.

Source: [DATA NOVIA](#)

```
penguins %>%  
  count(species) %>%  
  mutate(prop = n / sum(n)) %>%  
  ggplot(aes(x = prop, y = fct_reorder(species, prop),  
             fill = species)) +  
  geom_col() +  
  scale_x_continuous(labels = label_percent(accuracy = 1)) +  
  labs(  
    x = "Percentage",  
    y = "Species",  
    title = "Species distribution of penguins",  
    subtitle = "Adelie, Gentoo, and Chinstrap Penguins at Palmer Stat",  
    caption = "Source: allisonhorst.github.io/palmerpenguins")
```

Species distribution of penguins

Adelie, Gentoo, and Chinstrap Penguins at Palmer Station LTER

Source: allisonhorst.github.io/palmerpenguins

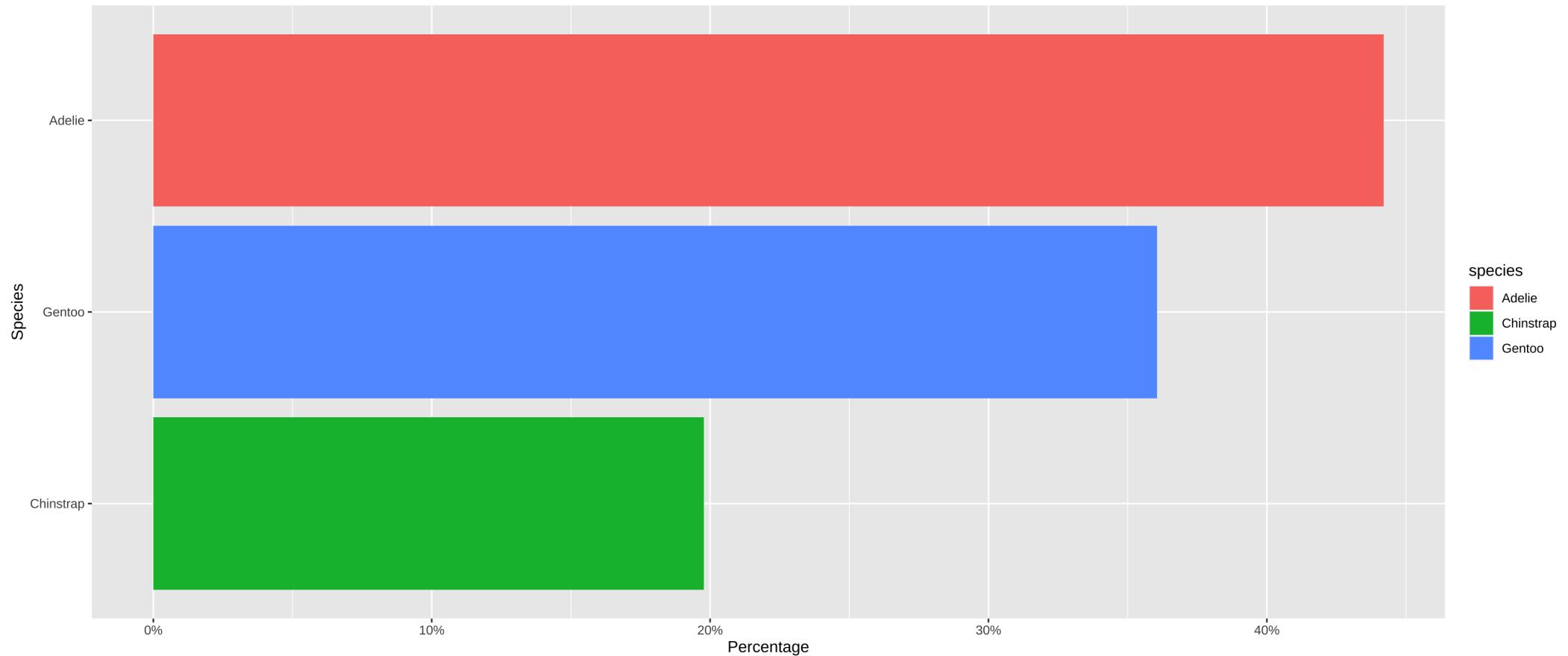
How to create a plot object?

```
species_plot <- penguins %>%  
  count(species) %>%  
  mutate(prop = (n / sum(n))) %>%  
  ggplot(aes(x = prop, y = fct_reorder(species, prop),  
             fill = species)) +  
  geom_col() +  
  scale_x_continuous(labels = label_percent(accuracy = 1)) +  
  labs(  
    x = "Percentage",  
    y = "Species",  
    title = "Species distribution of penguins",  
    subtitle = "Adelie, Gentoo, and Chinstrap Penguins at Palmer Stat",  
    caption = "Source: allisonhorst.github.io/palmerpenguins")
```

species_plot

Species distribution of penguins

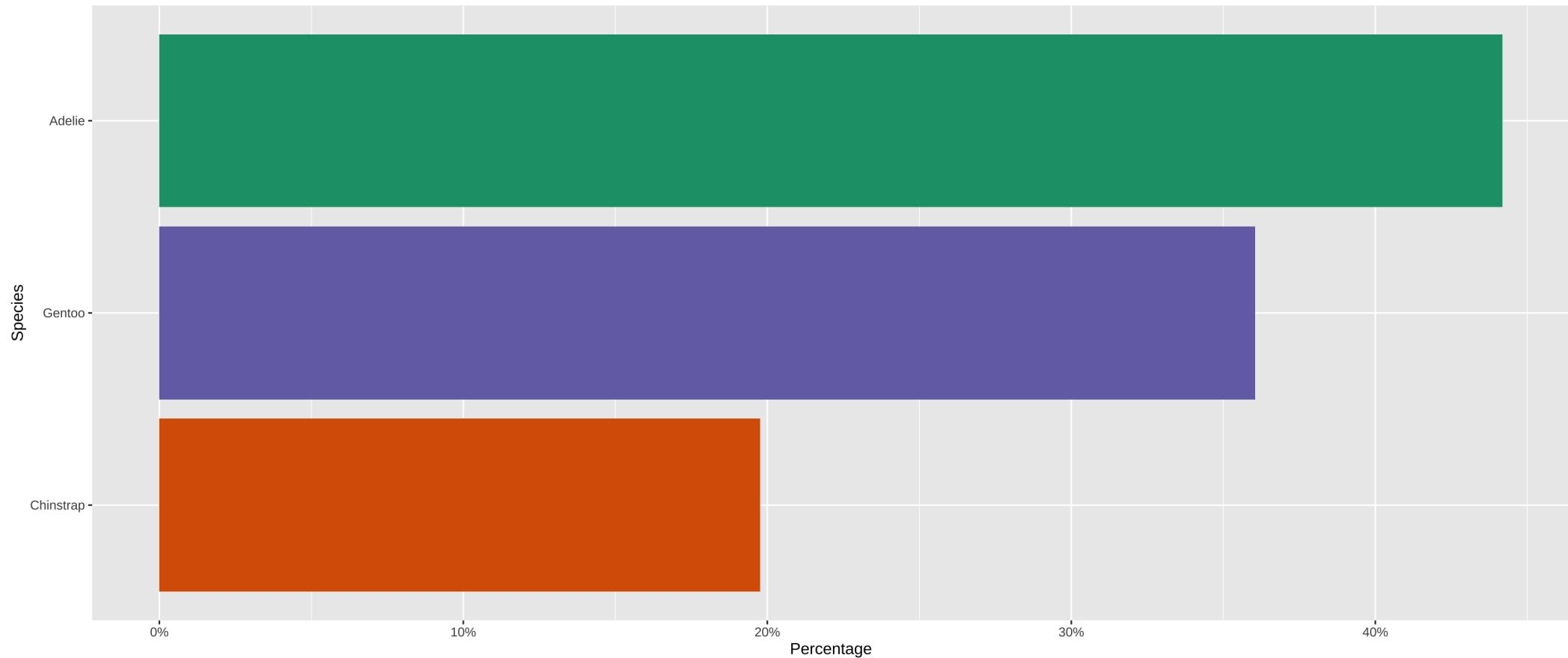
Adelie, Gentoo, and Chinstrap Penguins at Palmer Station LTER

Source: allisonhorst.github.io/palmerpenguins

```
species_plot +  
  scale_fill_brewer(palette = "Dark2") +  
  theme(legend.position = "none")
```

Species distribution of penguins

Adelie, Gentoo, and Chinstrap Penguins at Palmer Station LTER

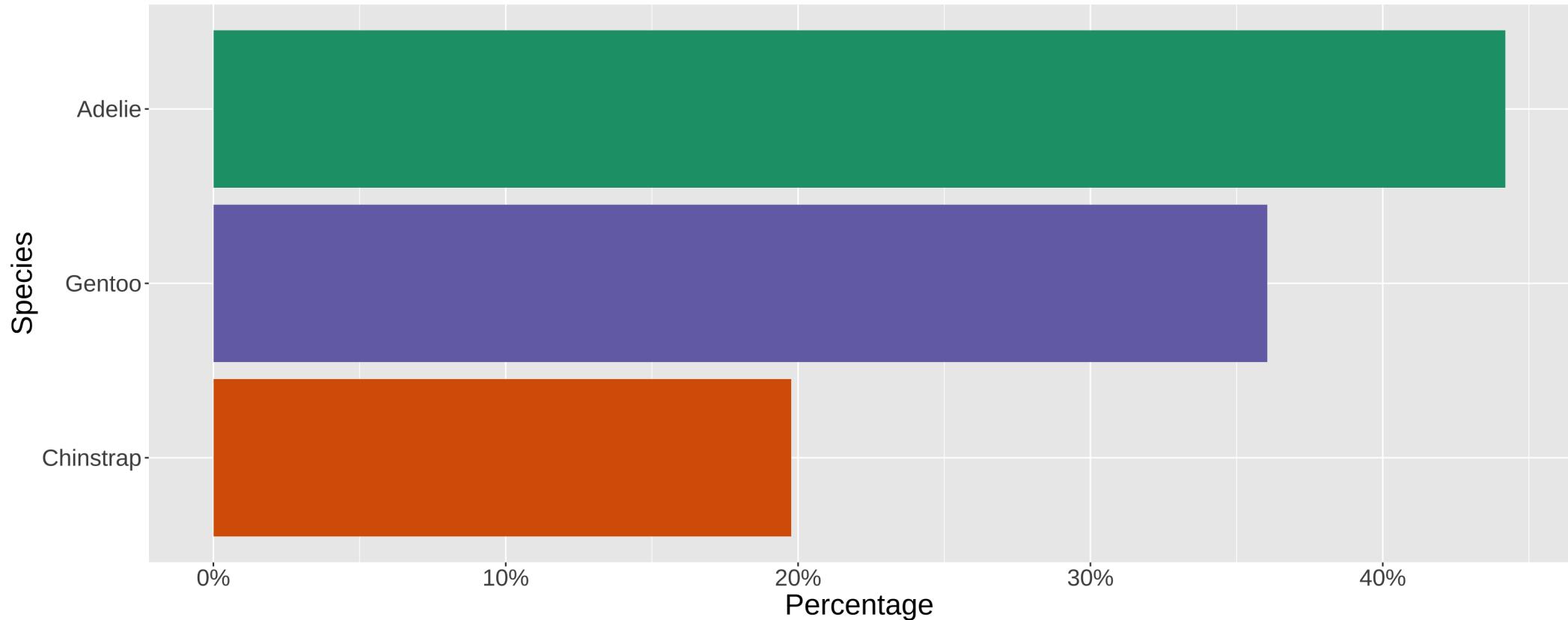
Source: allisonhorst.github.io/palmerpenguins

How to change font size of the text
labels?

```
species_plot +  
  scale_fill_brewer(palette = "Dark2") +  
  theme(legend.position = "none") +  
  theme(text = element_text(size = 20))
```

Species distribution of penguins

Adelie, Gentoo, and Chinstrap Penguins at Palmer Station LTER

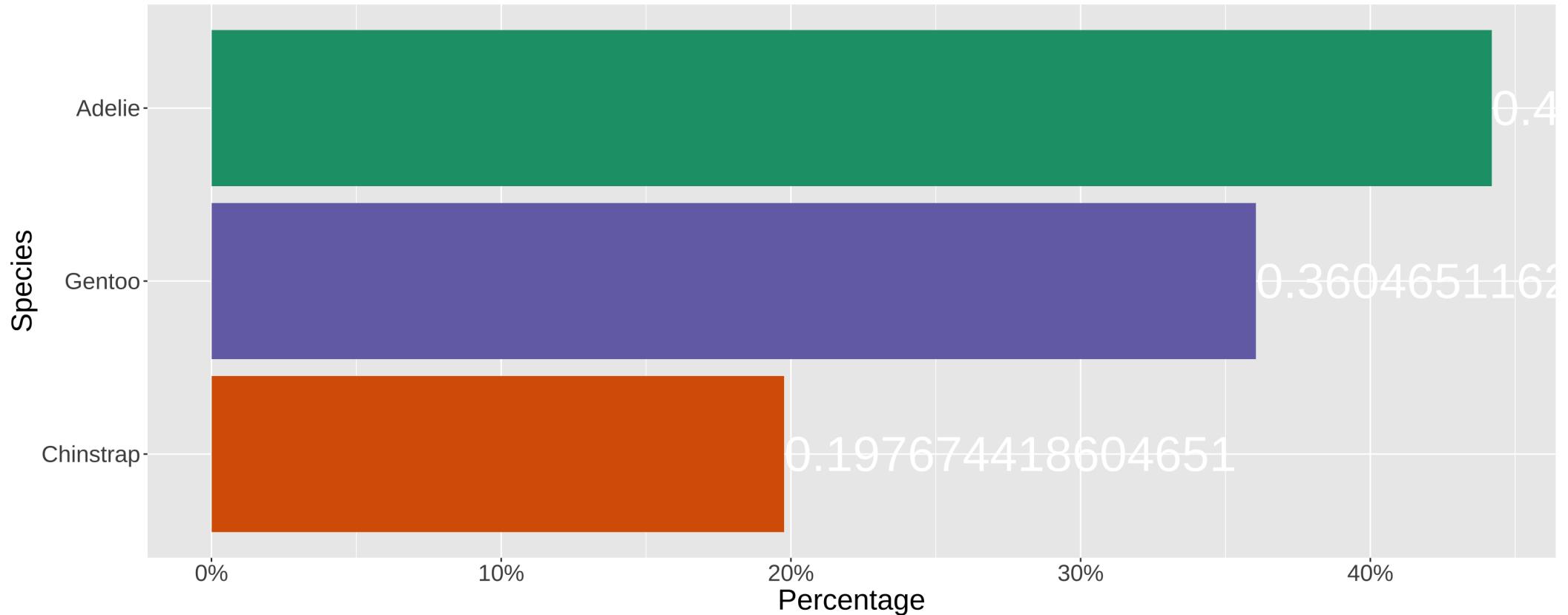
Source: [Statistics Globe](#)Source: allisonhorst.github.io/palmerpenguins

How to add count labels
on the bar plot?

```
species_plot +  
  scale_fill_brewer(palette = "Dark2") +  
  theme(text = element_text(size = 20),  
        legend.position = "none") +  
  geom_text(aes(label = prop), size = 12, hjust = 0, color = "white")
```

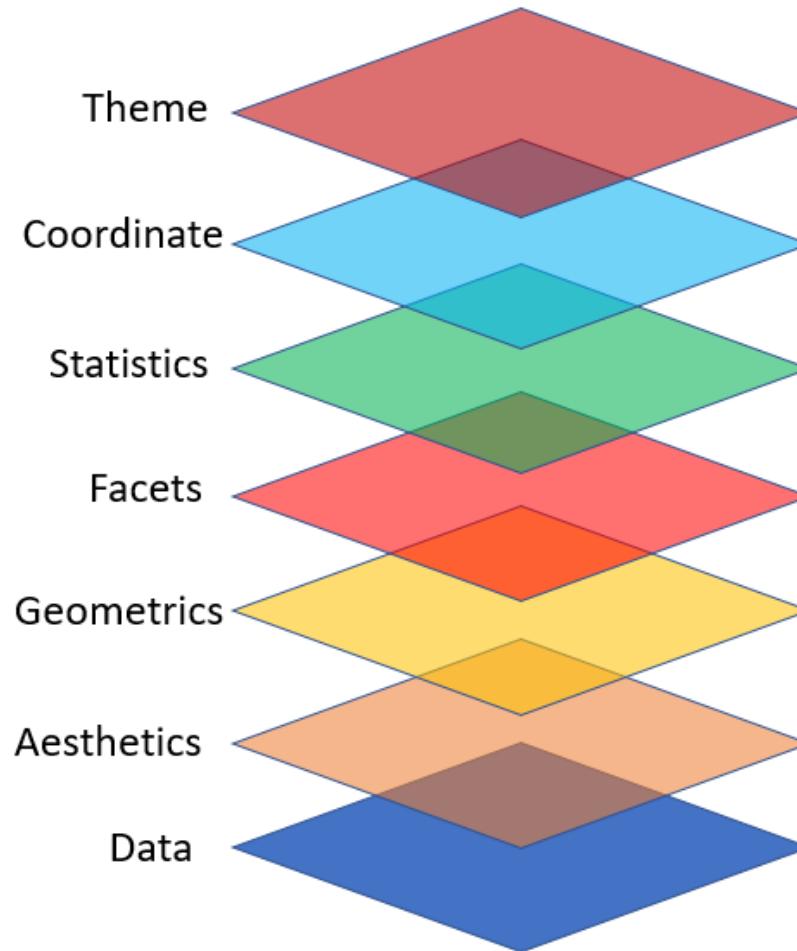
Species distribution of penguins

Adelie, Gentoo, and Chinstrap Penguins at Palmer Station LTER



Source: allisonhorst.github.io/palmerpenguins

How to add a different theme to plot?



```
library(ggthemes)

species_plot +
  scale_fill_brewer(palette = "Dark2") +
  geom_text(aes(label = n), size = 12, hjust = 2, color = "white") +
  theme_excel() +
  theme(text = element_text(size = 20),
        legend.position = "none")

# keep the order theme_bw() then theme()
```

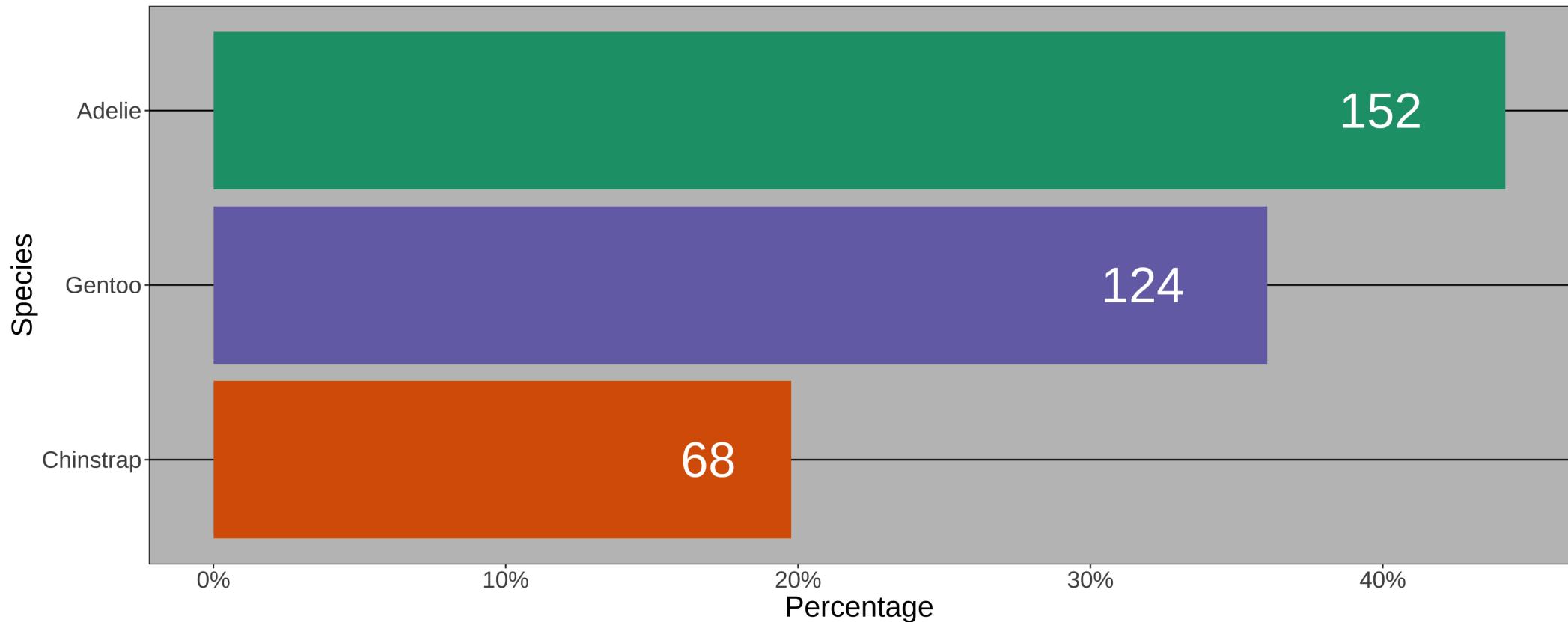
Task

Codes

Output

Species distribution of penguins

Adelie, Gentoo, and Chinstrap Penguins at Palmer Station LTER

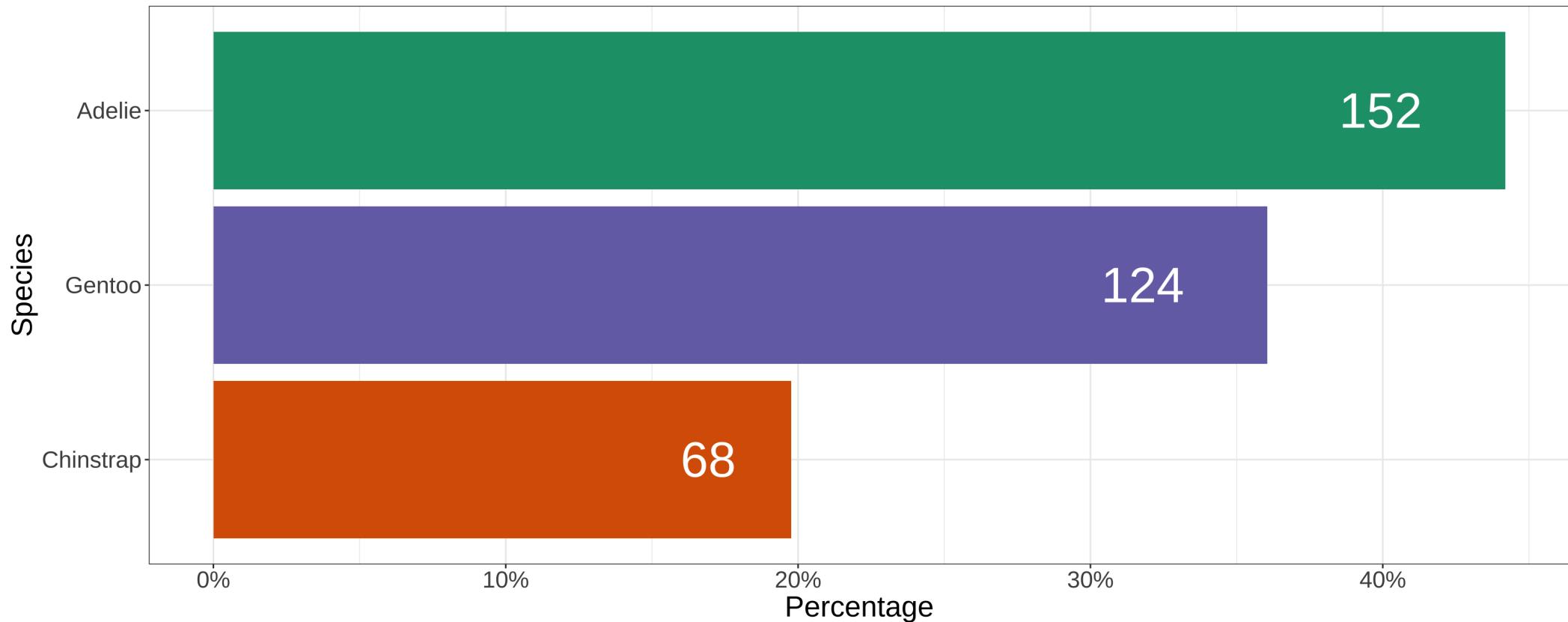


Source: allisonhorst.github.io/palmerpenguins

```
species_plot +  
  scale_fill_brewer(palette = "Dark2") +  
  geom_text(aes(label = n), size = 12, hjust = 2, color = "white") +  
  theme_bw() +  
  theme(text = element_text(size = 20),  
        legend.position = "none")
```

Species distribution of penguins

Adelie, Gentoo, and Chinstrap Penguins at Palmer Station LTER

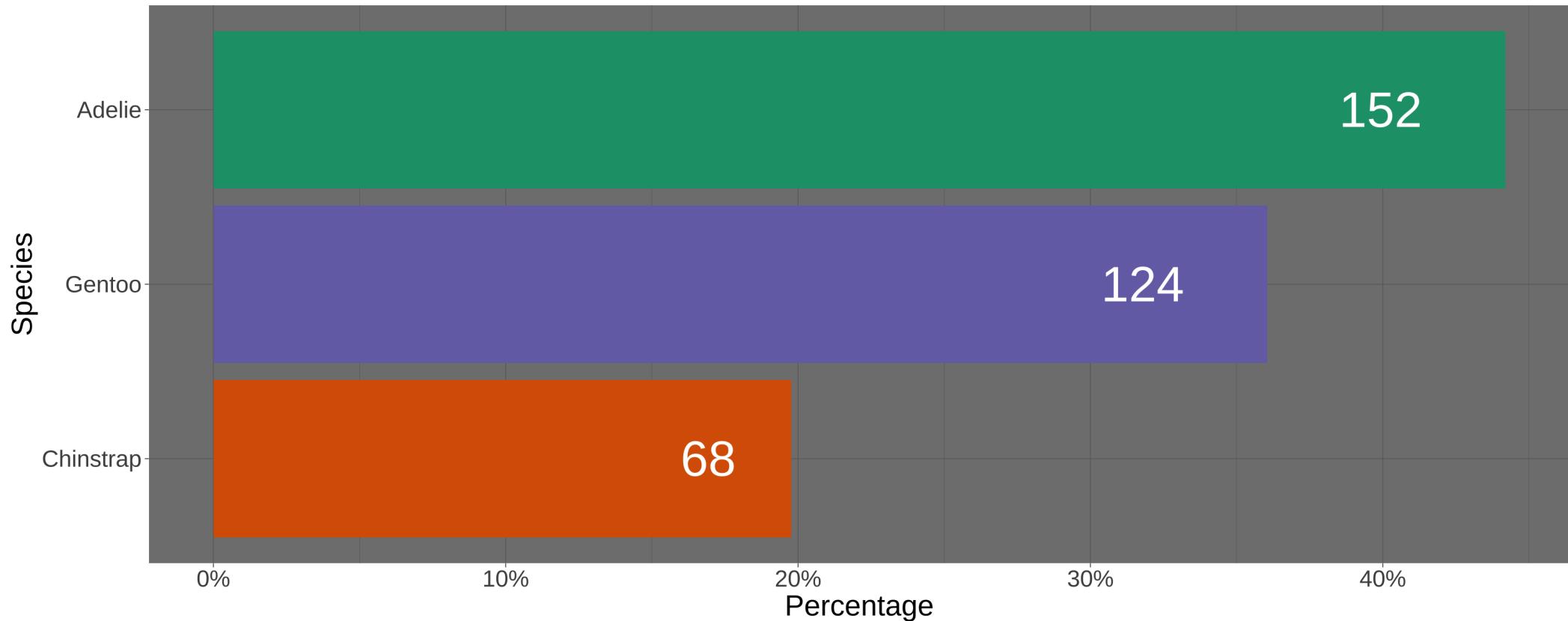


Source: allisonhorst.github.io/palmerpenguins

```
species_plot +  
  scale_fill_brewer(palette = "Dark2") +  
  geom_text(aes(label = n), size = 12, hjust = 2, color = "white") +  
  theme_dark() +  
  theme(text = element_text(size = 20),  
        legend.position = "none")  
  
# keep the order theme_bw() then theme()
```

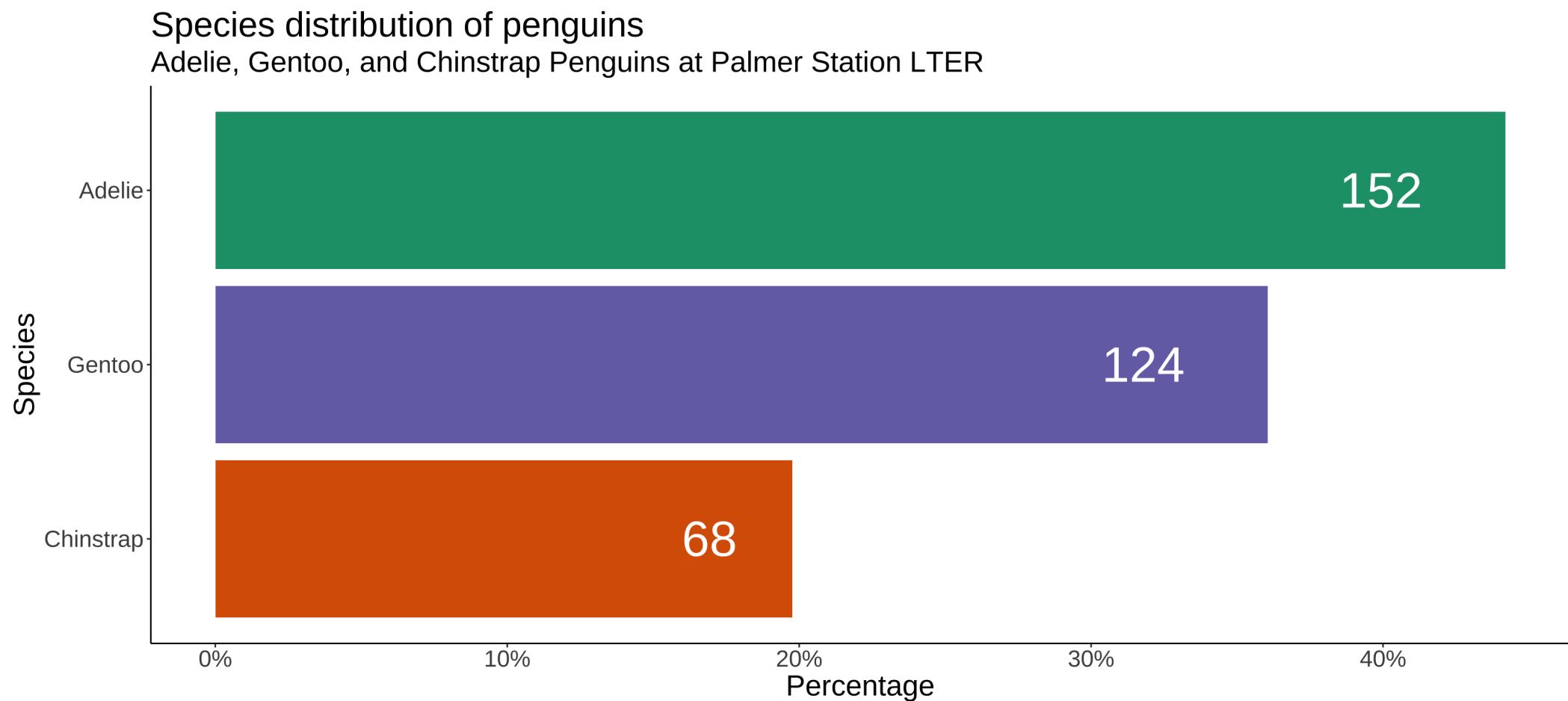
Species distribution of penguins

Adelie, Gentoo, and Chinstrap Penguins at Palmer Station LTER



Source: allisonhorst.github.io/palmerpenguins

```
species_plot +  
  scale_fill_brewer(palette = "Dark2") +  
  geom_text(aes(label = n), size = 12, hjust = 2, color = "white") +  
  theme_classic() +  
  theme(text = element_text(size = 20),  
        legend.position = "none")  
  
# keep the order theme_bw() then theme()
```

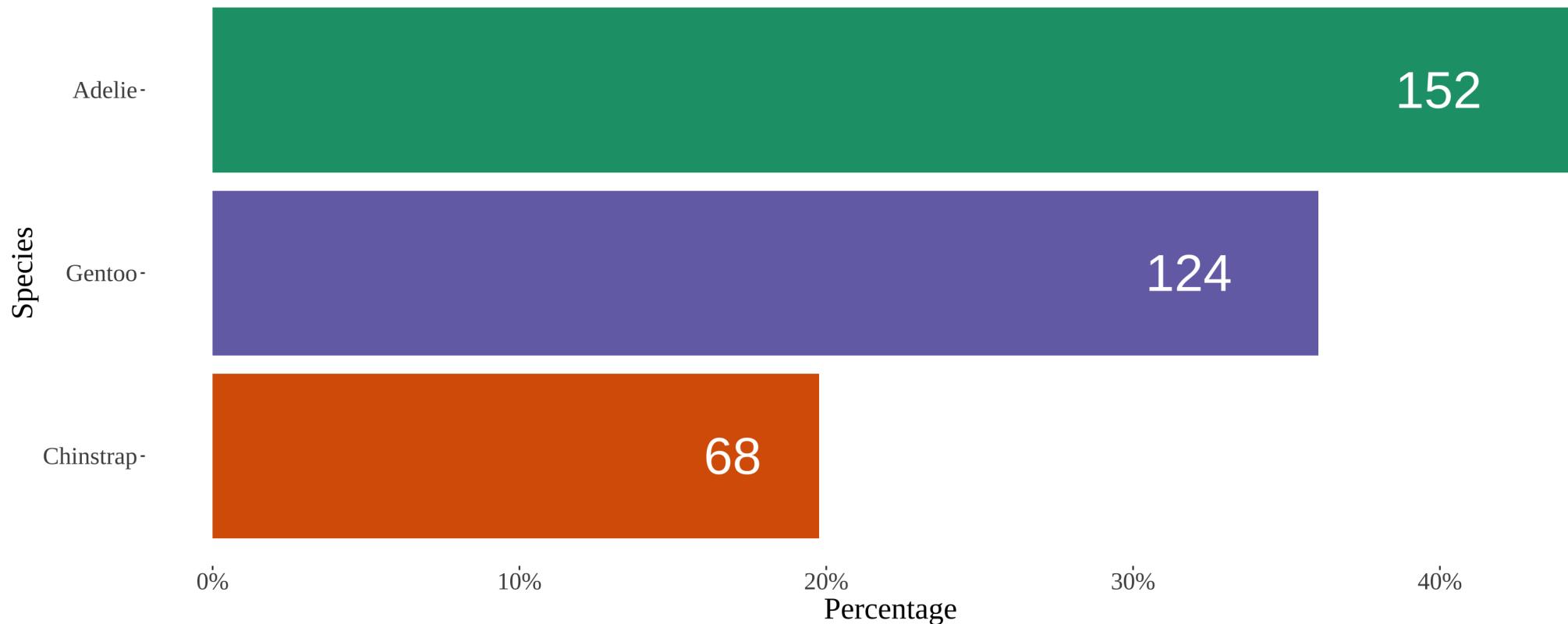


Source: allisonhorst.github.io/palmerpenguins

```
species_plot +  
  scale_fill_brewer(palette = "Dark2") +  
  geom_text(aes(label = n), size = 12, hjust = 2, color = "white") +  
  theme_tufte() +  
  theme(text = element_text(size = 20),  
        legend.position = "none")  
  
# keep the order theme_bw() then theme()
```

Species distribution of penguins

Adelie, Gentoo, and Chinstrap Penguins at Palmer Station LTER



Source: allisonhorst.github.io/palmerpenguins