Class 11 - Data Visualization with ggplot2

Md.Mahfujul Karim Sheikh

2025-02-04 12:24:30.114574

Contents


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

Packages

Loading required packages:

Setting global theme:

```
theme_set(theme_classic())
```

Data

Loading the built in iris data:

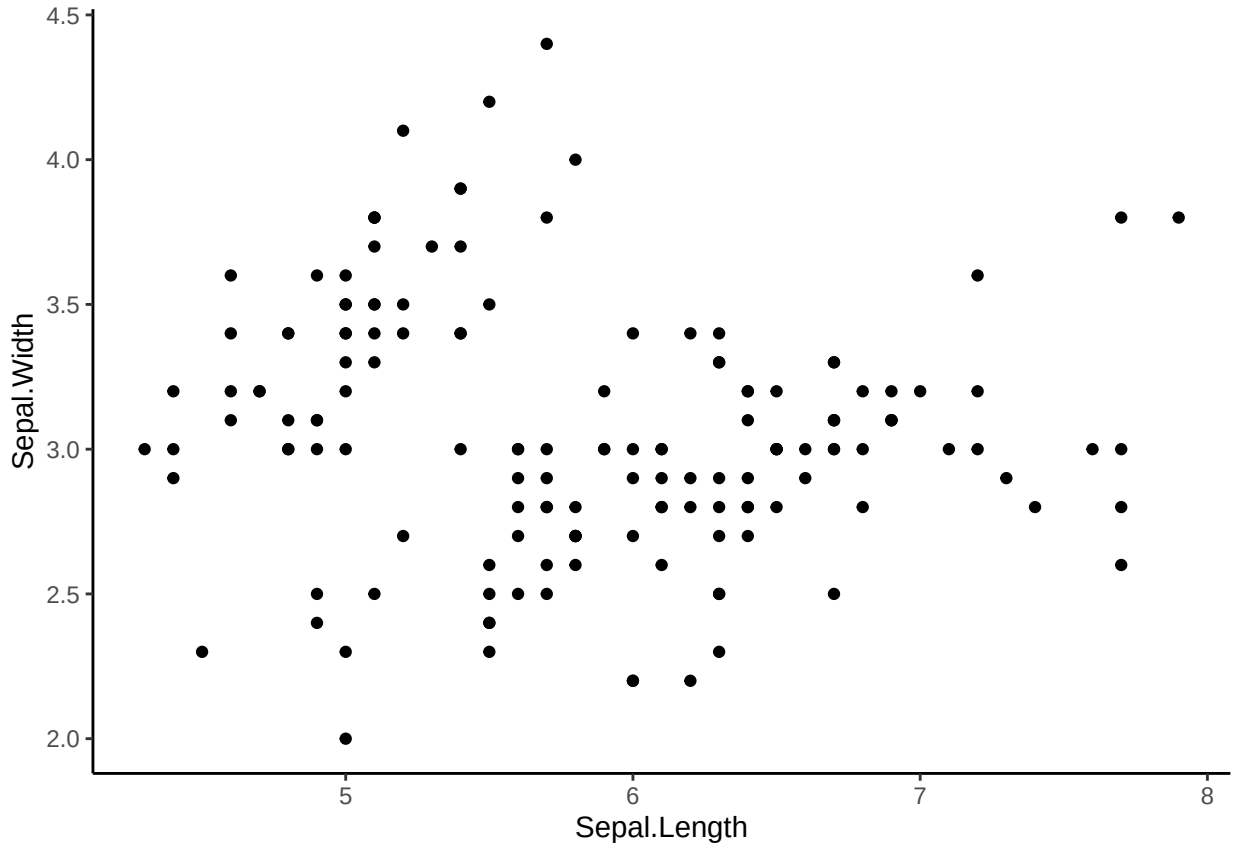
```
data(iris)
```

Loading the student survey data:

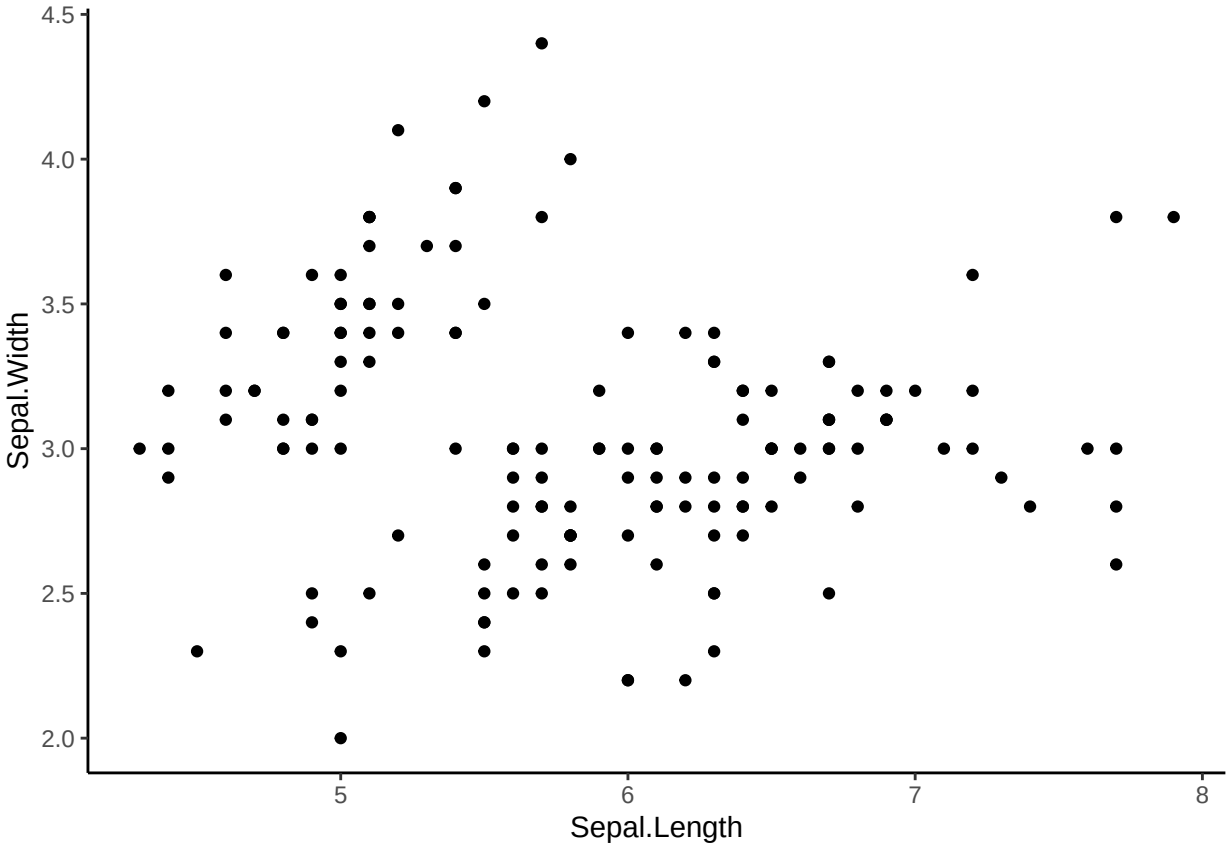
```
student <- readxl::read_excel("D:\\RProgramming\\Class11\\Self\\StudentSurveyData.xlsx")
```

Scatter Plot

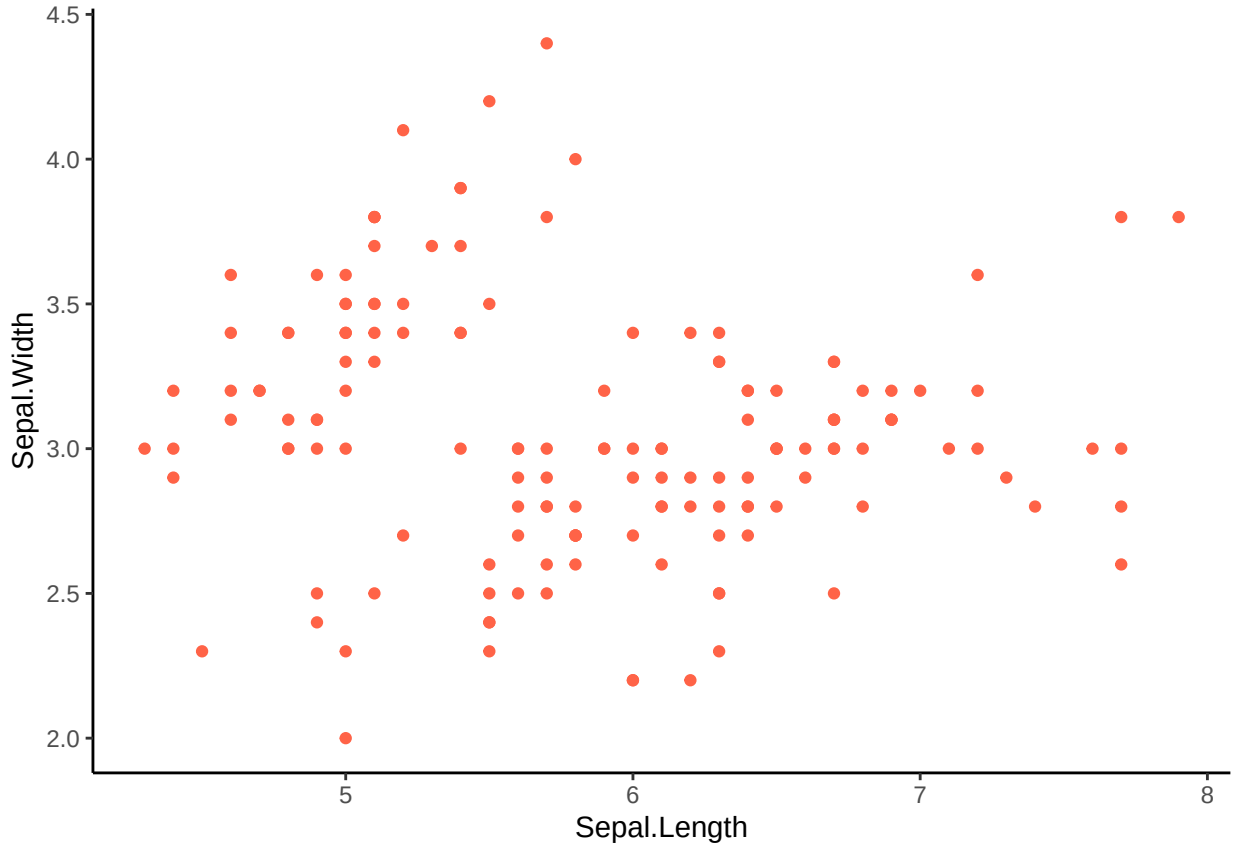
```
ggplot(data = iris) +
  geom_point(mapping = aes(x = Sepal.Length, y = Sepal.Width)) +
  theme_classic()
```



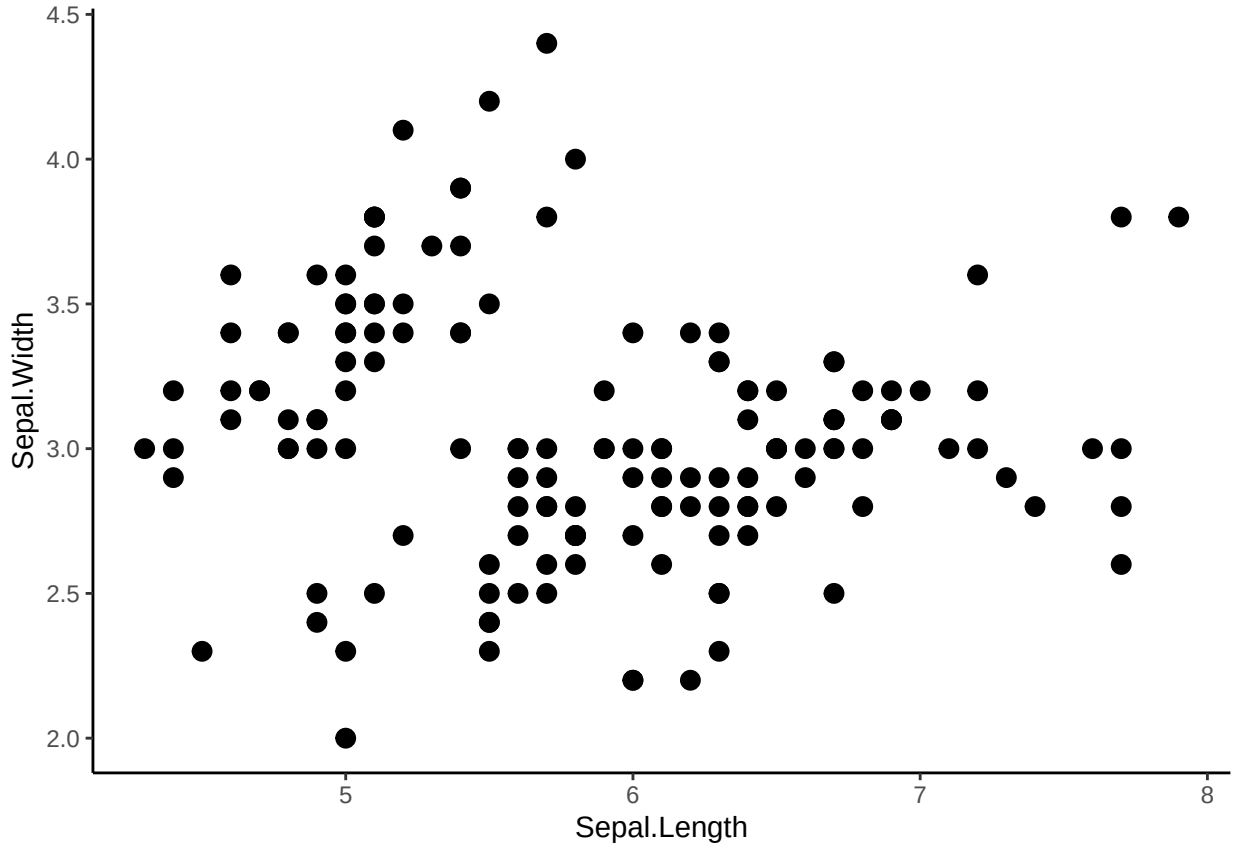

```
ggplot(data = iris) +
  aes(x = Sepal.Length, y = Sepal.Width) +
  geom_point()
```



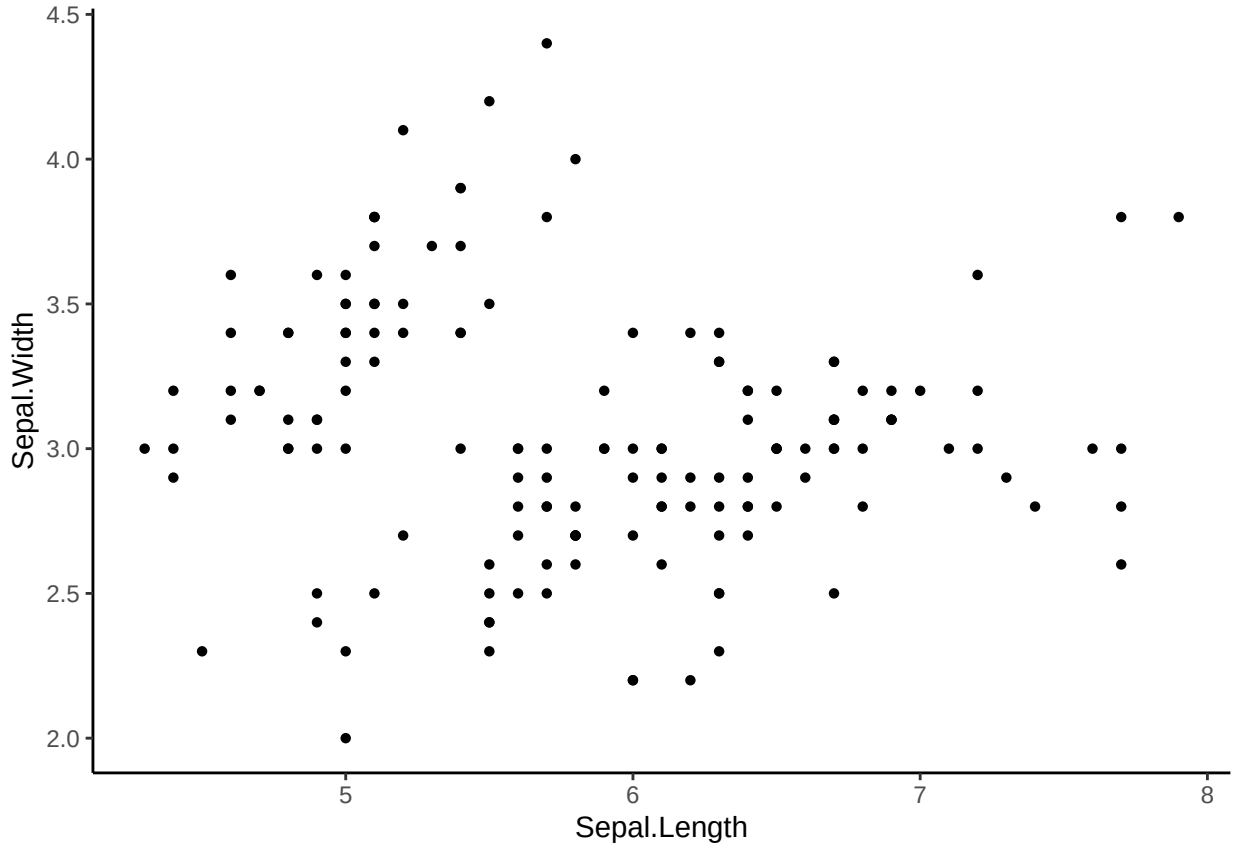

```
ggplot(data = iris, aes(x = Sepal.Length, y = Sepal.Width)) +
  geom_point(color = "tomato1")
```



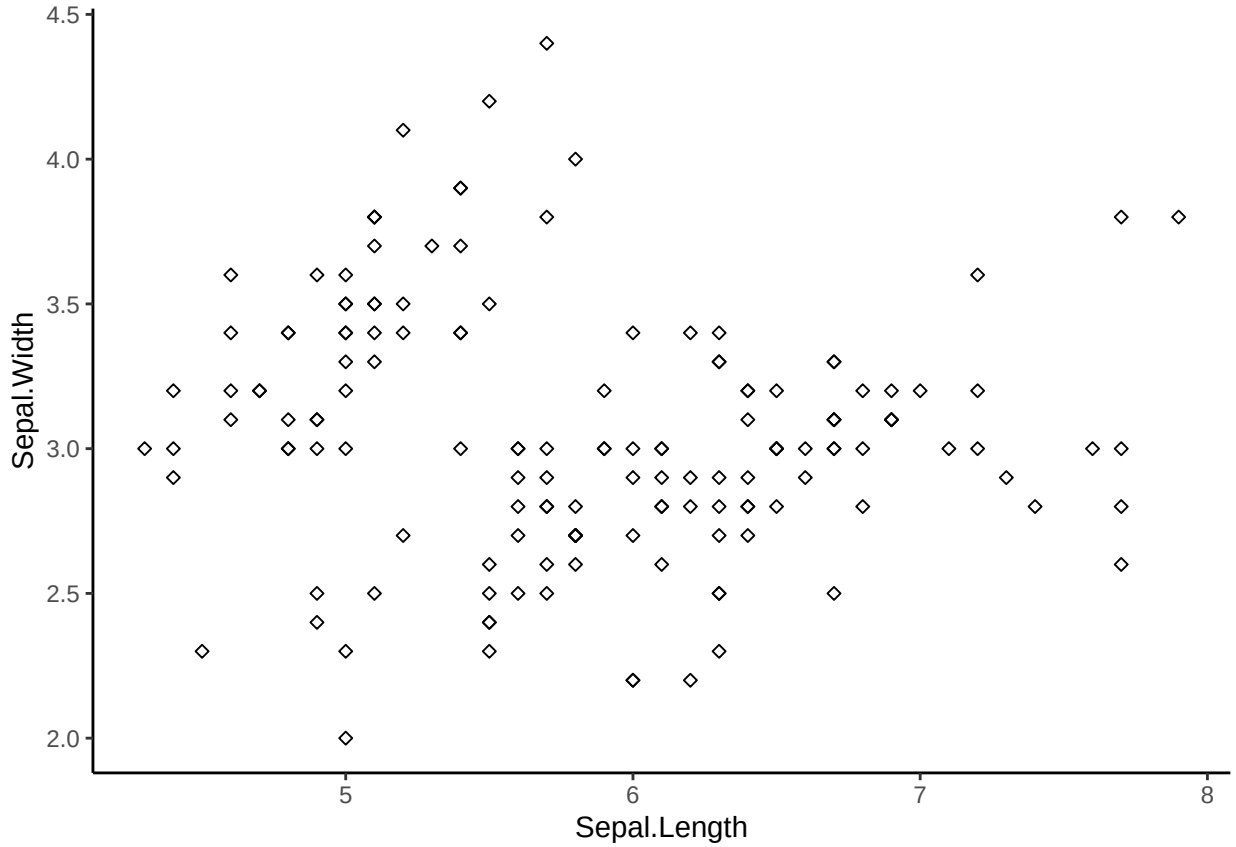

```
ggplot(data = iris, aes(x = Sepal.Length, y = Sepal.Width)) +
  geom_point(size = 3)
```



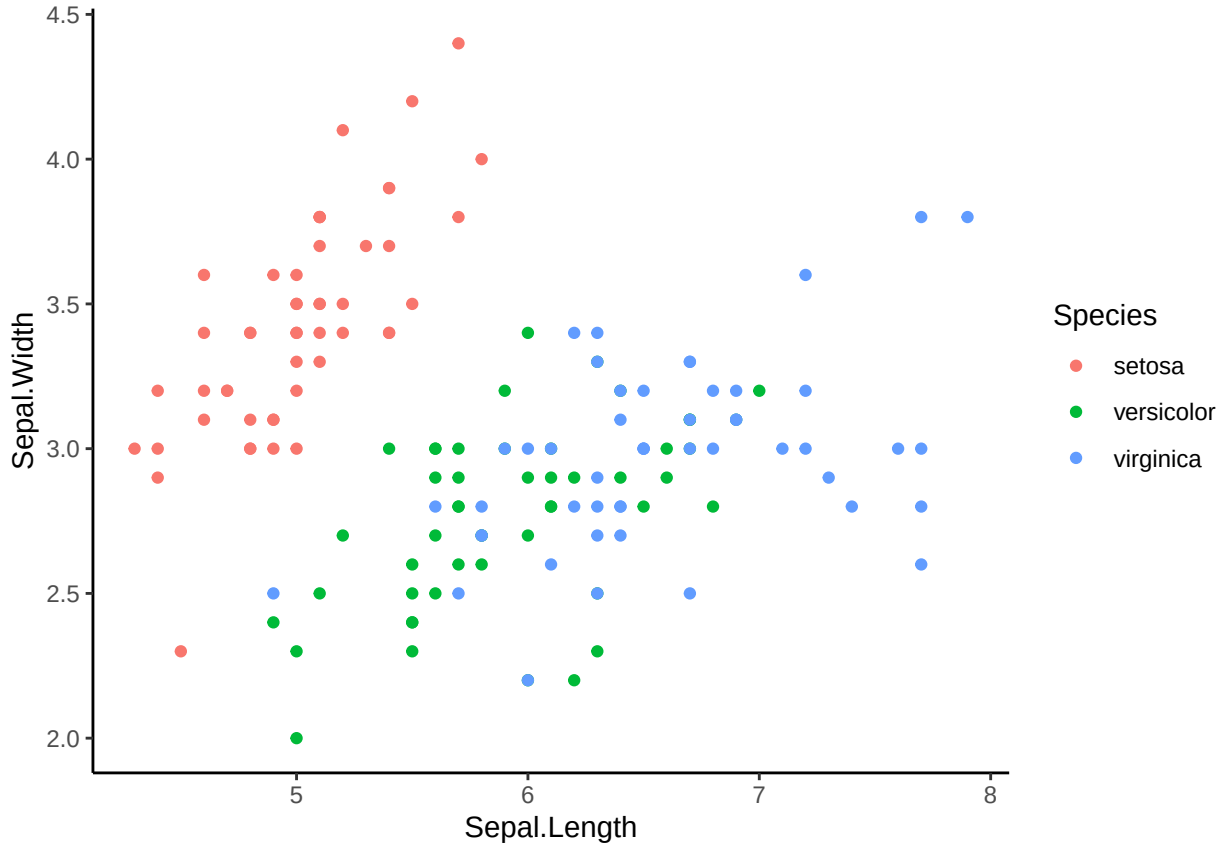

```
ggplot(data = iris, aes(x = Sepal.Length, y = Sepal.Width)) +
  geom_point(shape = 16)
```



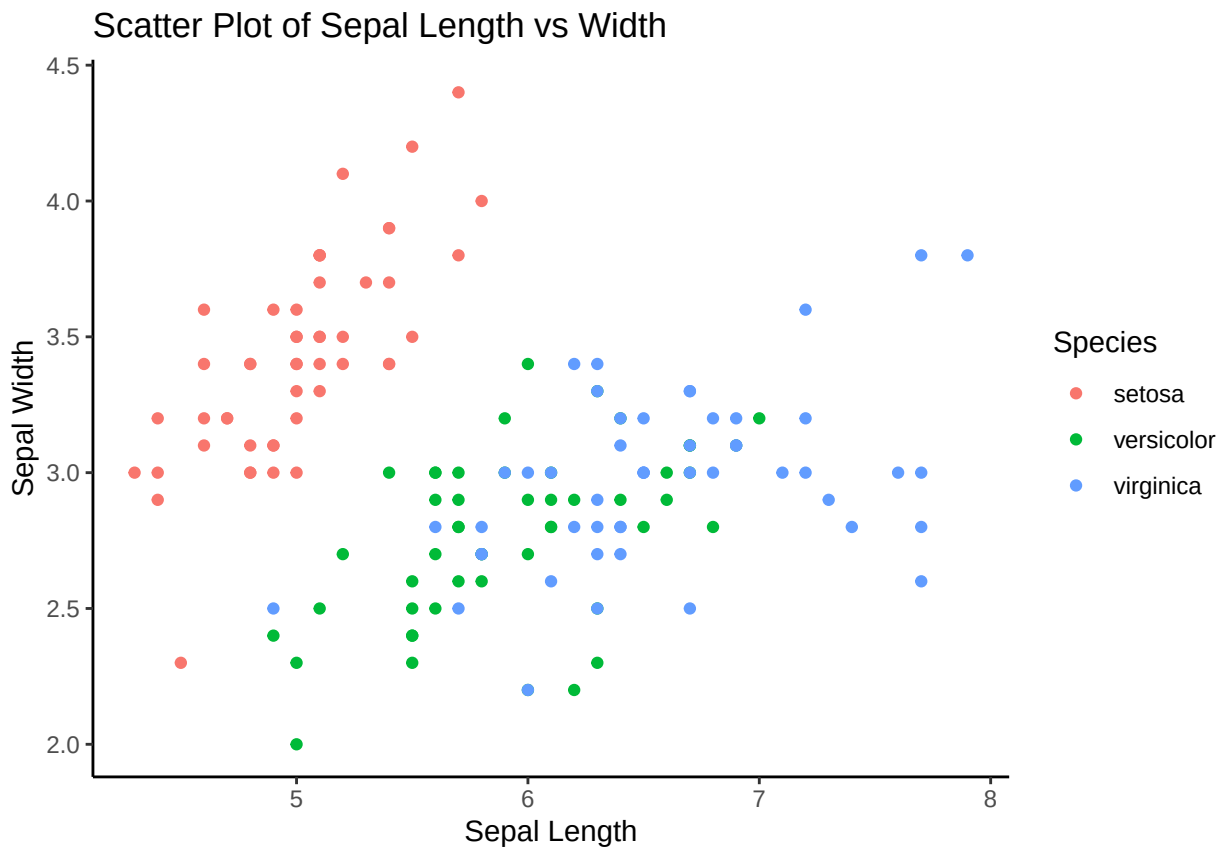

```
ggplot(data = iris, aes(x = Sepal.Length, y = Sepal.Width)) +
  geom_point(shape = "diamond filled")
```




```
ggplot(data = iris, aes(x = Sepal.Length, y = Sepal.Width, col = Species)) +
  geom_point()
```

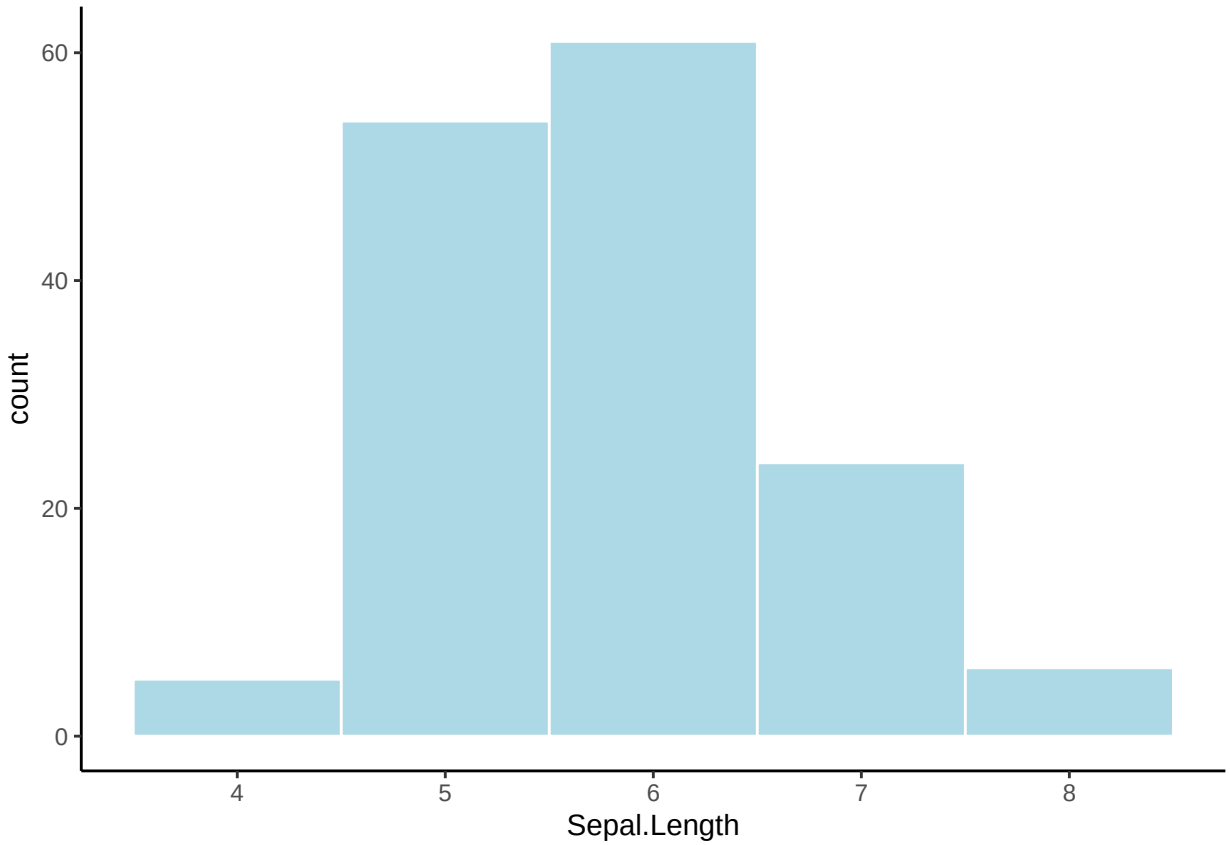



```
ggplot(data = iris, aes(x = Sepal.Length, y = Sepal.Width, col = Species)) +
  geom_point() +
  labs(x = "Sepal Length", y = "Sepal Width", col = "Species",
       title = "Scatter Plot of Sepal Length vs Width")
```

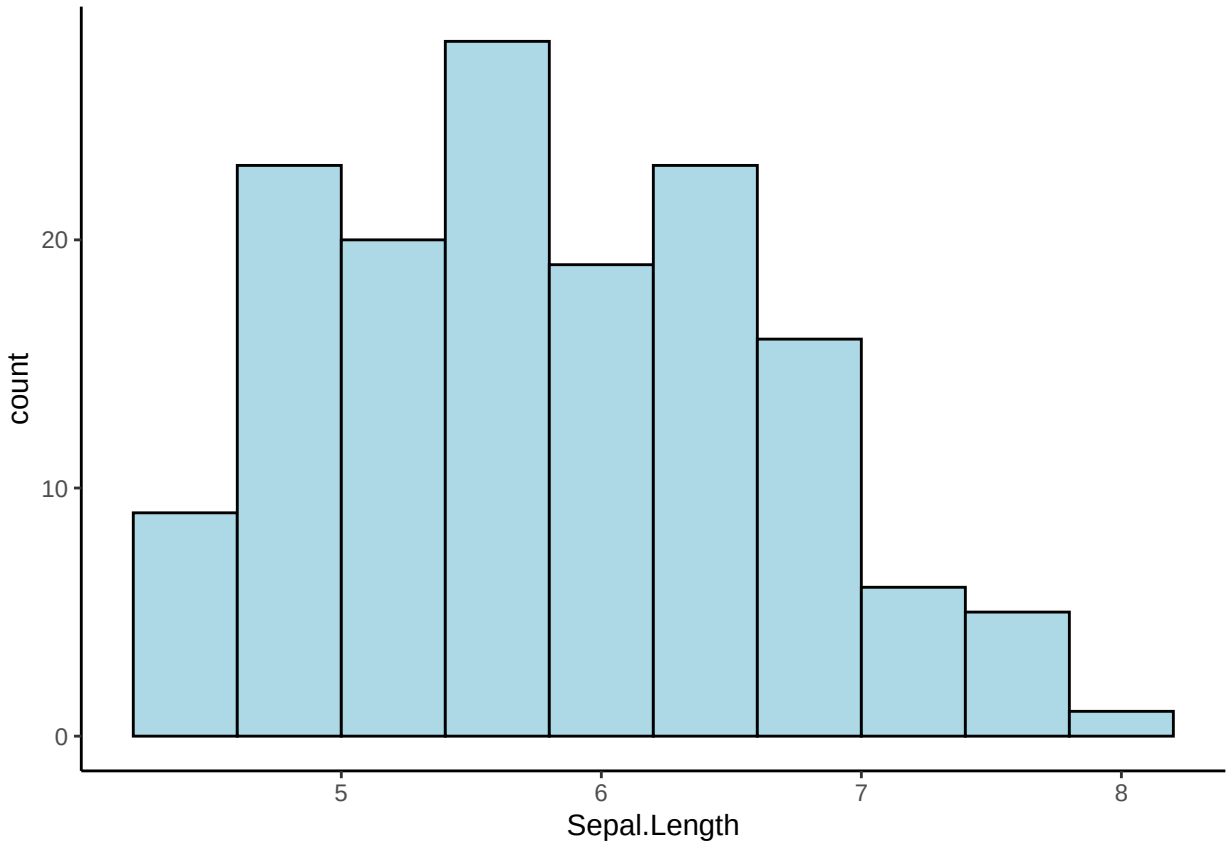


Histogram

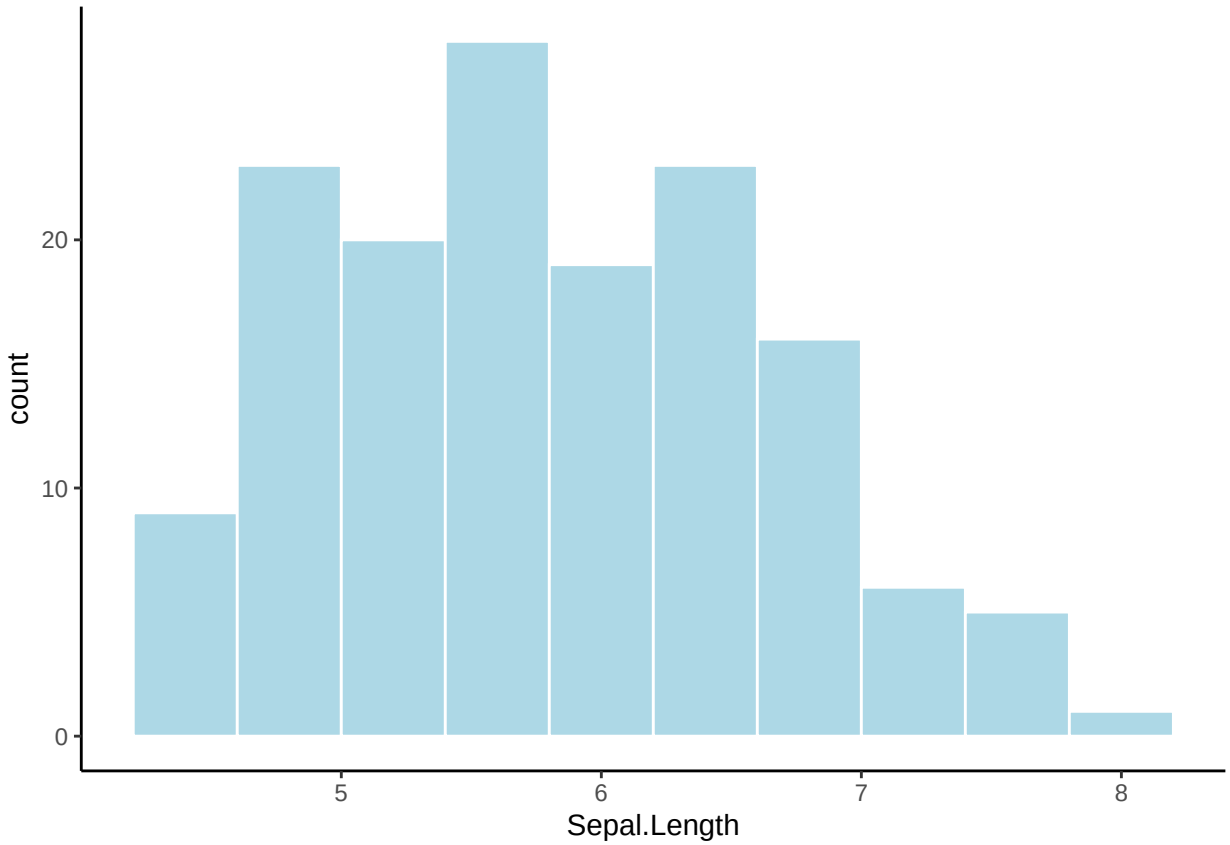
```
ggplot(iris, aes(x = Sepal.Length)) +
  geom_histogram(binwidth = 1, fill = "lightblue", col = "white")
```



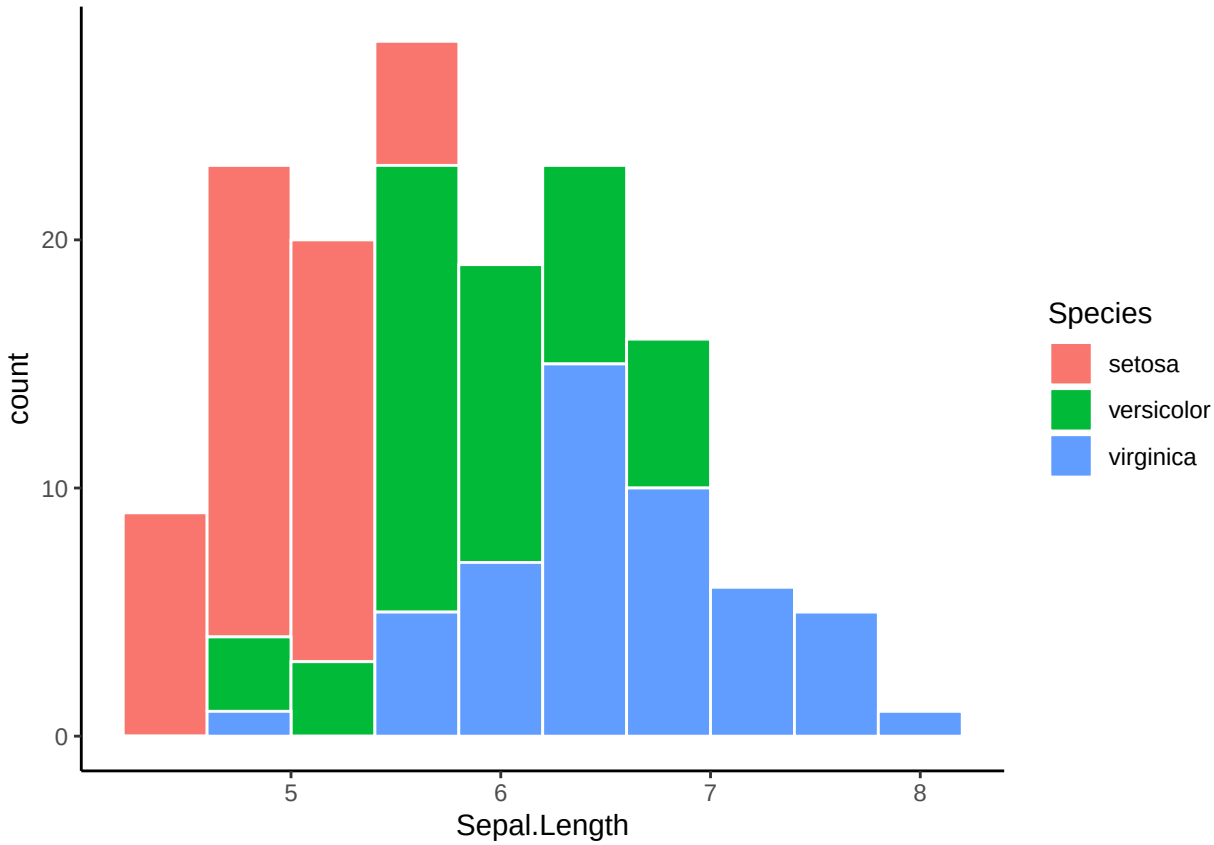

```
ggplot(data = iris) +
  geom_histogram(aes(x = Sepal.Length), bins = 10, fill = "lightblue", col = "black")
```



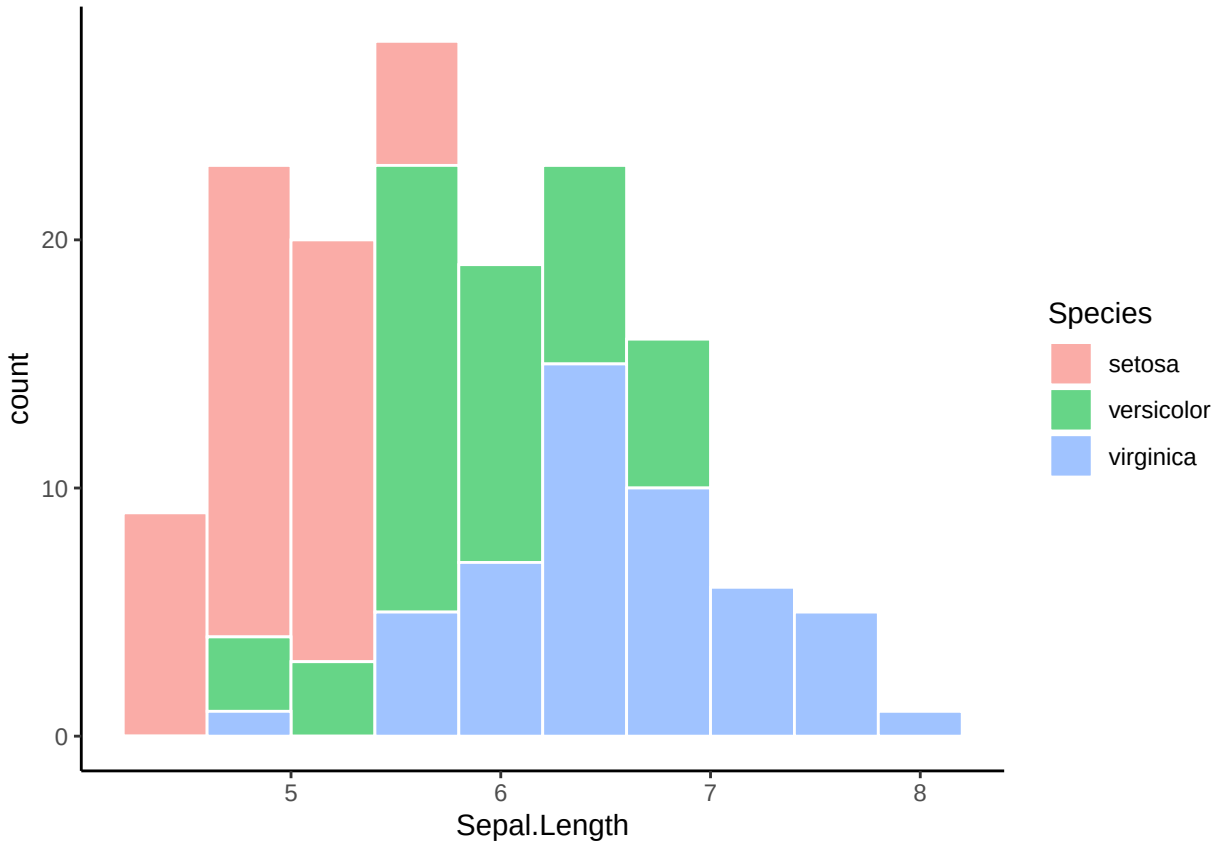

```
ggplot(data = iris) +
  geom_histogram(aes(x = Sepal.Length), bins = 10, fill = "lightblue", col = "white")
```




```
ggplot(data = iris) +
  geom_histogram(aes(x = Sepal.Length, fill = Species), bins = 10, col = "white")
```




```
ggplot(data = iris) +
  geom_histogram(aes(x = Sepal.Length, fill = Species), bins = 10, col = "white", alpha = 0.6)
```



Extra

```
res <- sample(1:100, 10)
res
```

```
##  [1] 12 90 72 52 87 63  5 10 28 59
```

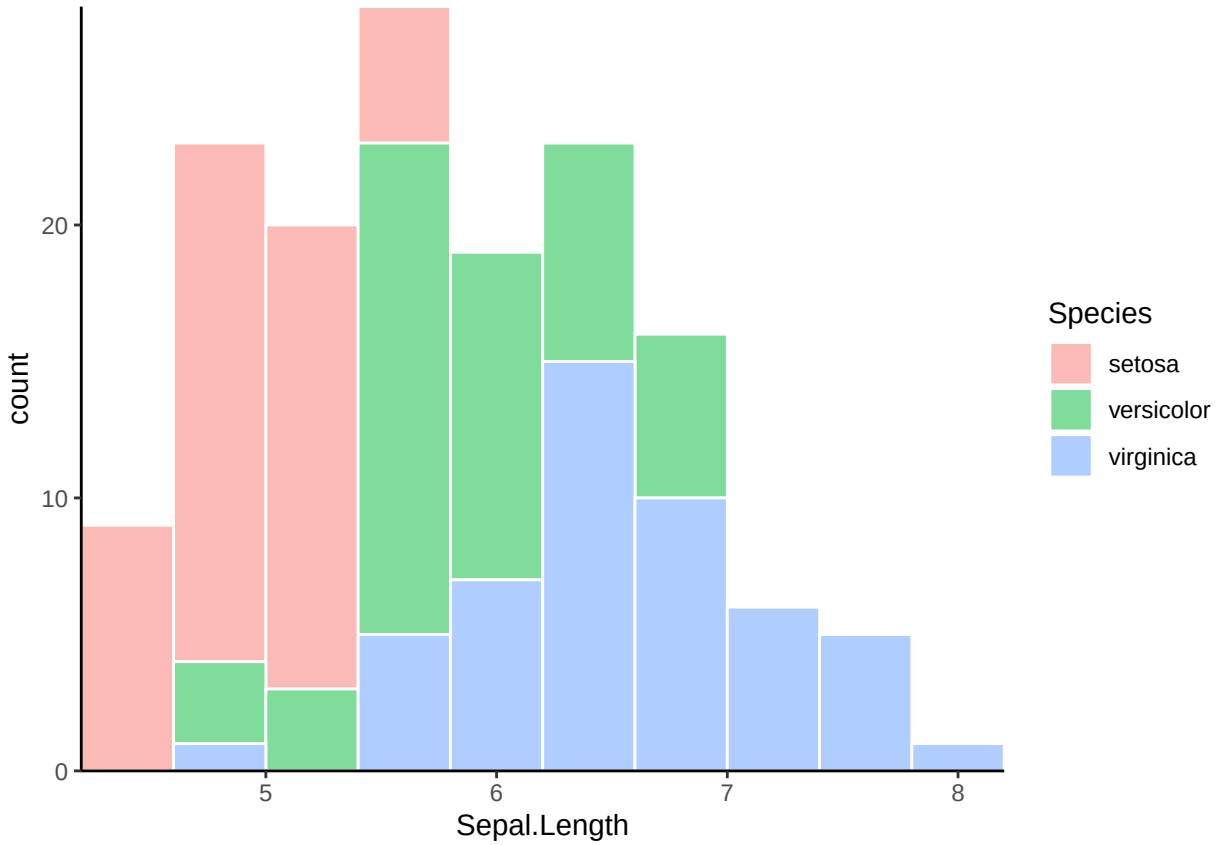
```
sum(res)
```

```
## [1] 478
```

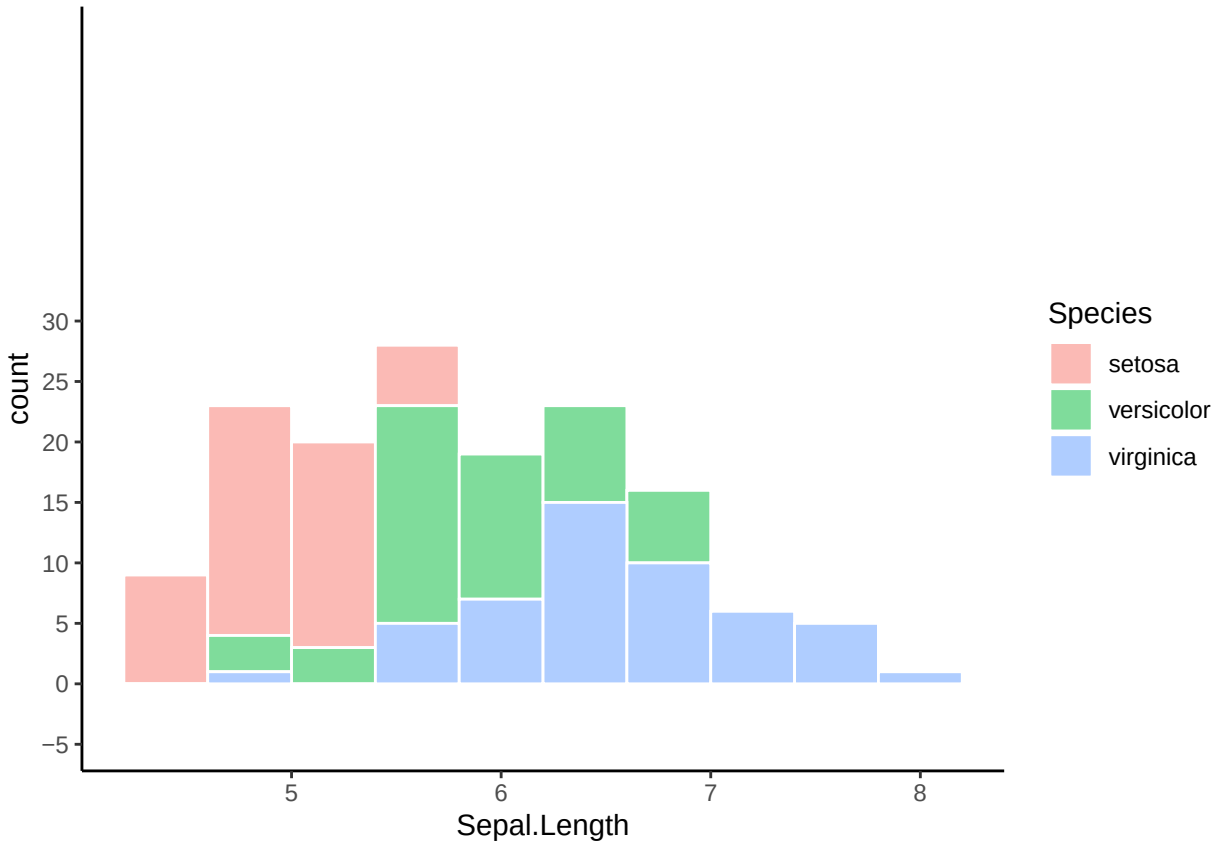
The total of the two numbers are 478.

Reduce Gap between Plot and Axis

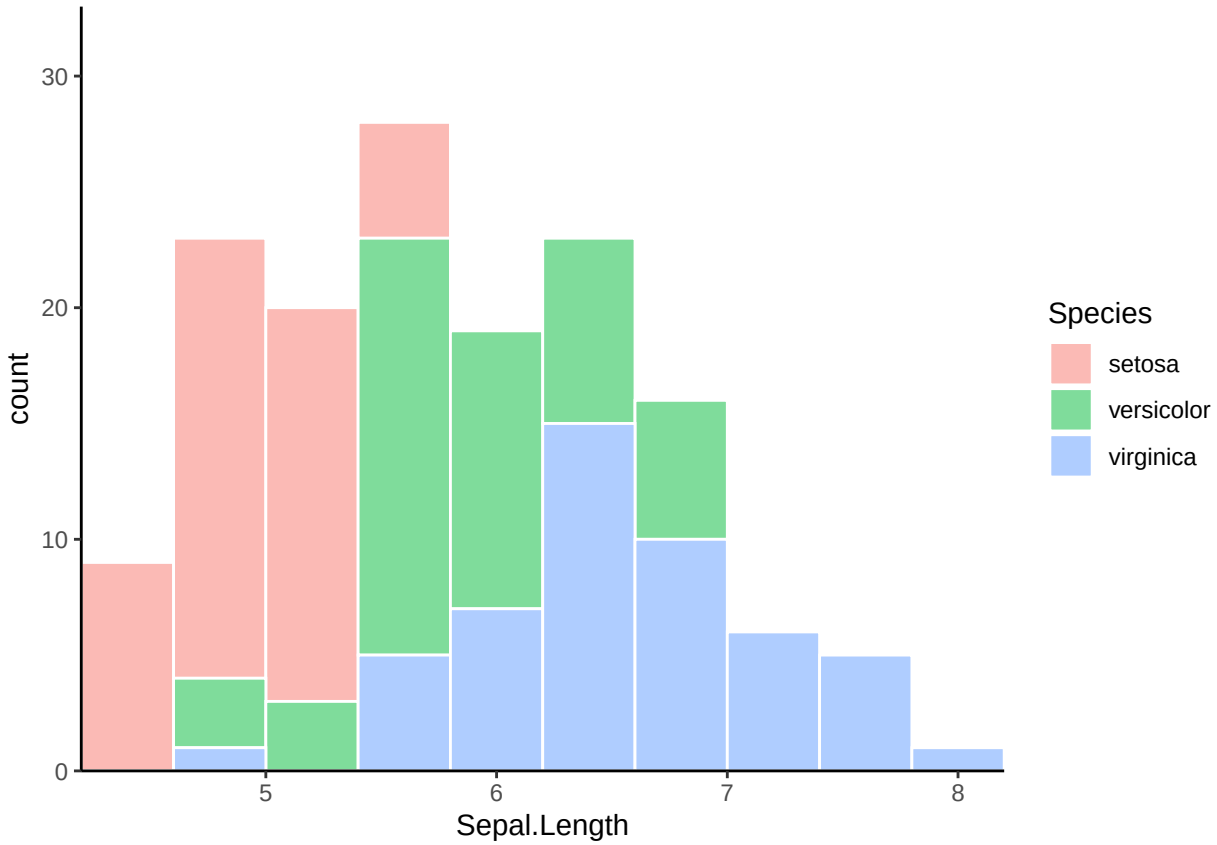
```
ggplot(data = iris) +
  geom_histogram(aes(x = Sepal.Length, fill = Species),
                 bins = 10, col = "white", alpha = 0.5) +
  coord_cartesian(expand = FALSE)
```




```
ggplot(data = iris) +
  geom_histogram(aes(x = Sepal.Length, fill = Species),
                 bins = 10, col = "white", alpha = 0.5) +
  scale_y_continuous(
    breaks = seq(-10, 30, by=5),
    expand = expansion(
      mult = c(-0.1, 1),      # expands upper portion of the plot by 20%
      add = c(10, 0)
      )        # increases gap at the bottom portion by 10 unit
    )
```



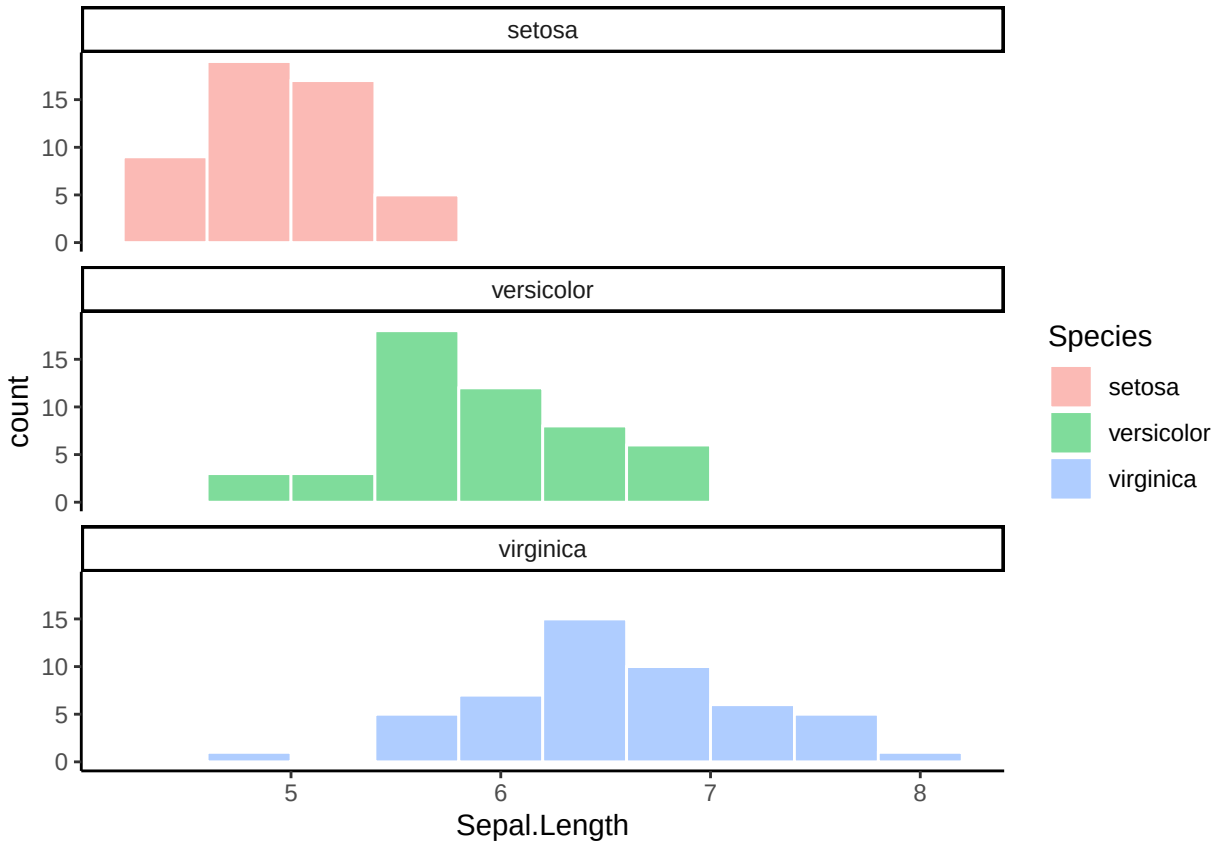

```
ggplot(data = iris) +
  geom_histogram(aes(x = Sepal.Length, fill = Species),
                 bins = 10, col = "white", alpha = 0.5) +
  scale_y_continuous(expand = expansion(add = c(0, 5))) +
  scale_x_continuous(expand = expansion(add = c(0, 0)))
```



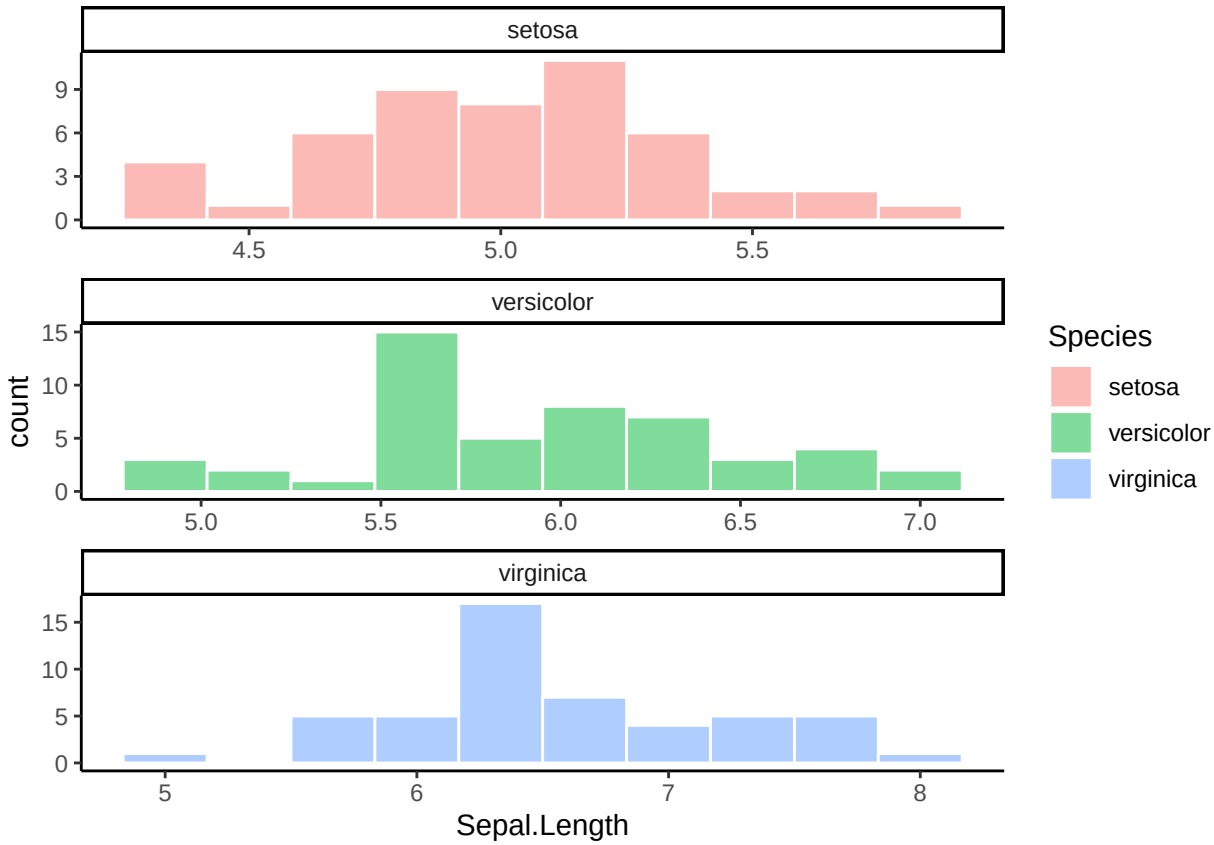
Facet

facet_wrap

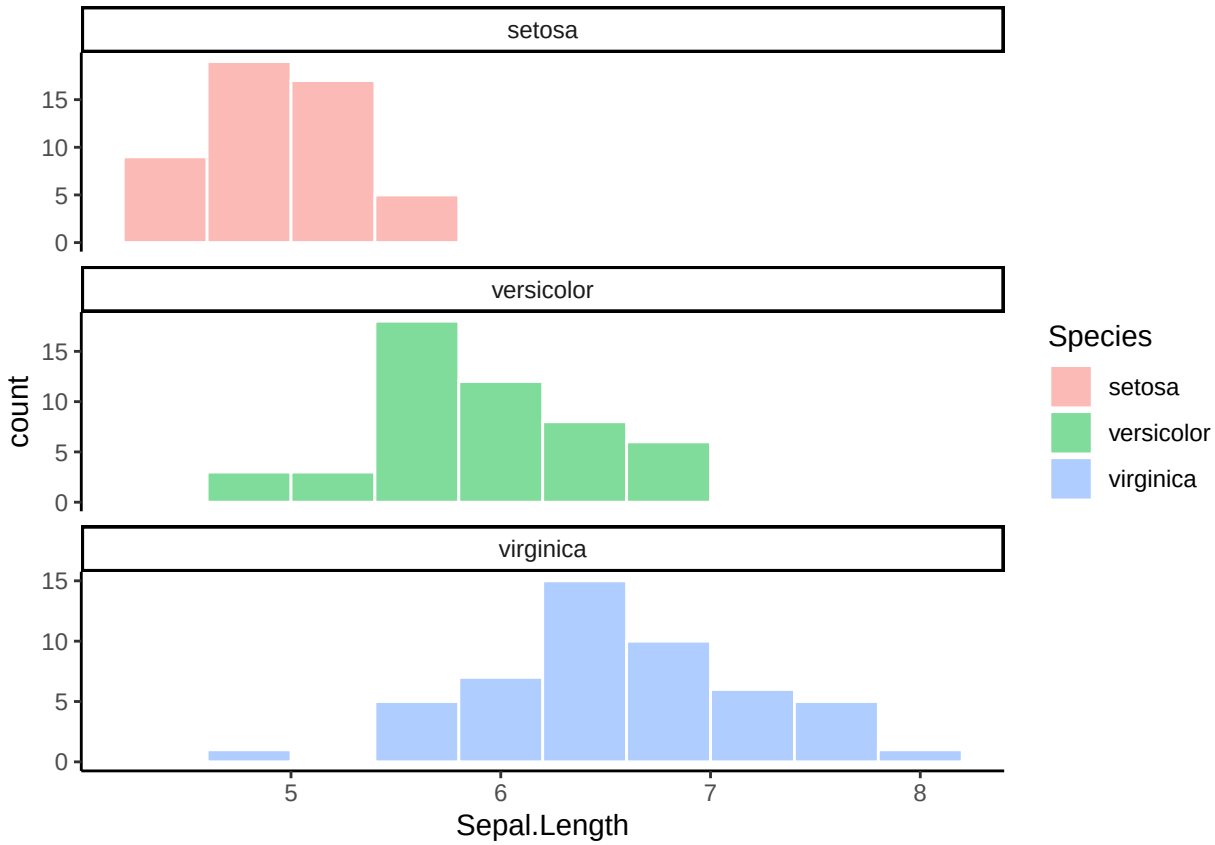
```
ggplot(data = iris) +
  geom_histogram(aes(x = Sepal.Length, fill = Species),
                 bins = 10, col = "white", alpha = 0.5) +
  facet_wrap(vars(Species), ncol = 1)
```




```
ggplot(data = iris) +
  geom_histogram(aes(x = Sepal.Length, fill = Species),
                 bins = 10, col = "white", alpha = 0.5) +
  facet_wrap(vars(Species), ncol = 1, scales = "free")
```

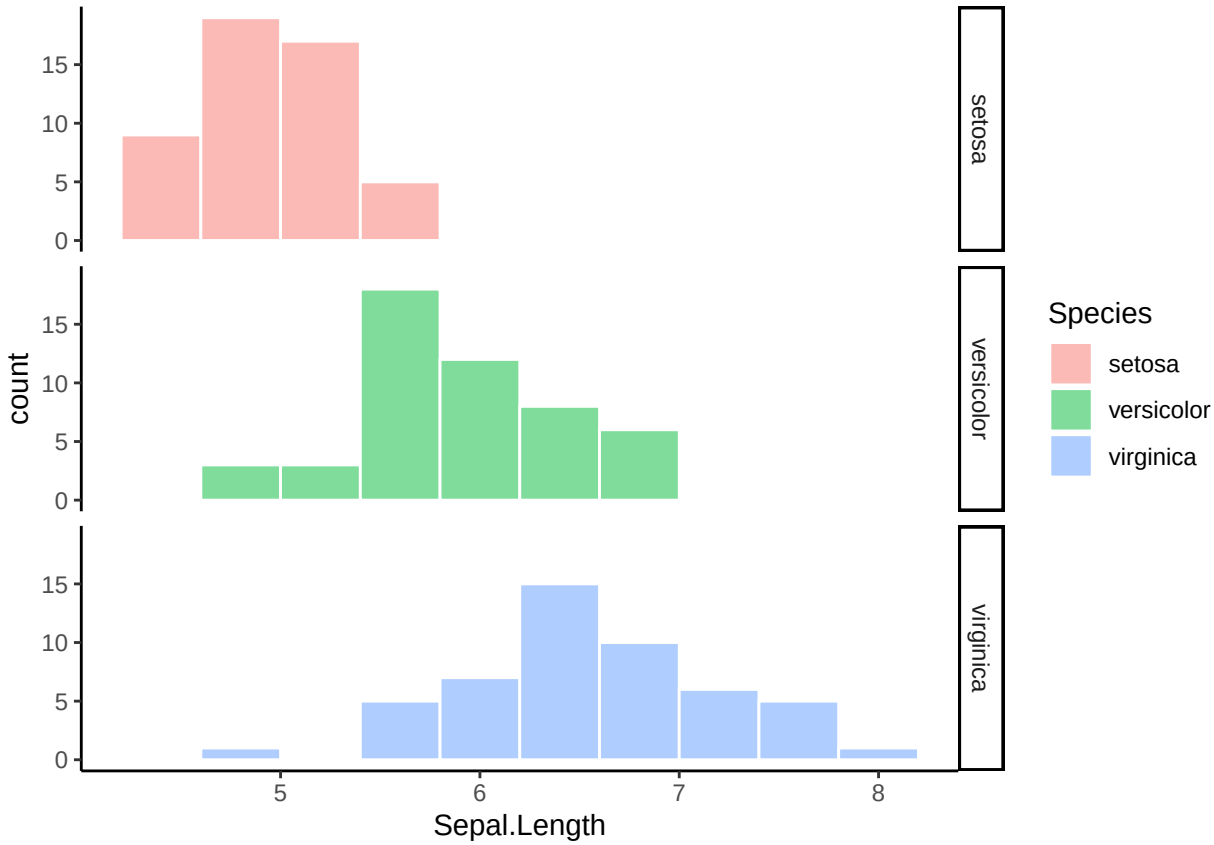



```
ggplot(data = iris) +
  geom_histogram(aes(x = Sepal.Length, fill = Species),
                 bins = 10, col = "white", alpha = 0.5) +
  facet_wrap(vars(Species), ncol = 1, scales = "free_y")
```

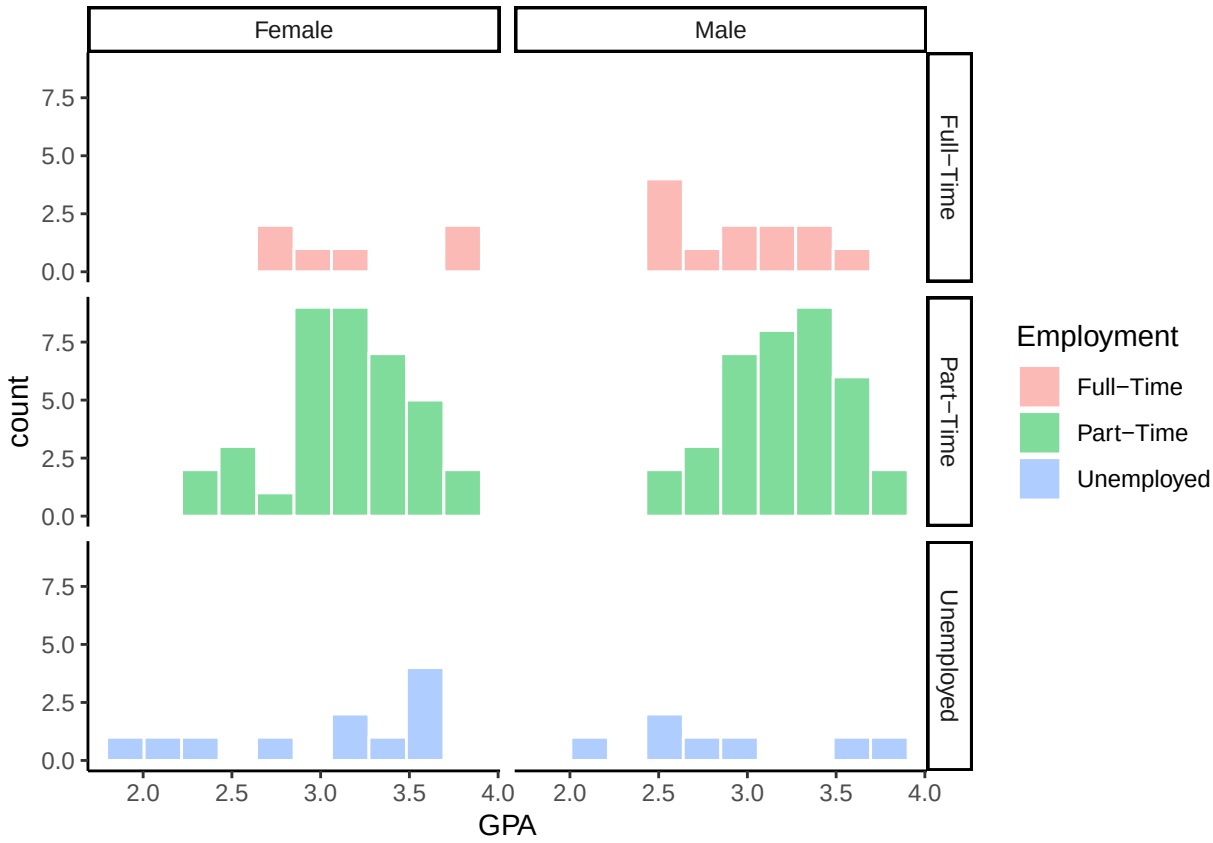


facet_grid

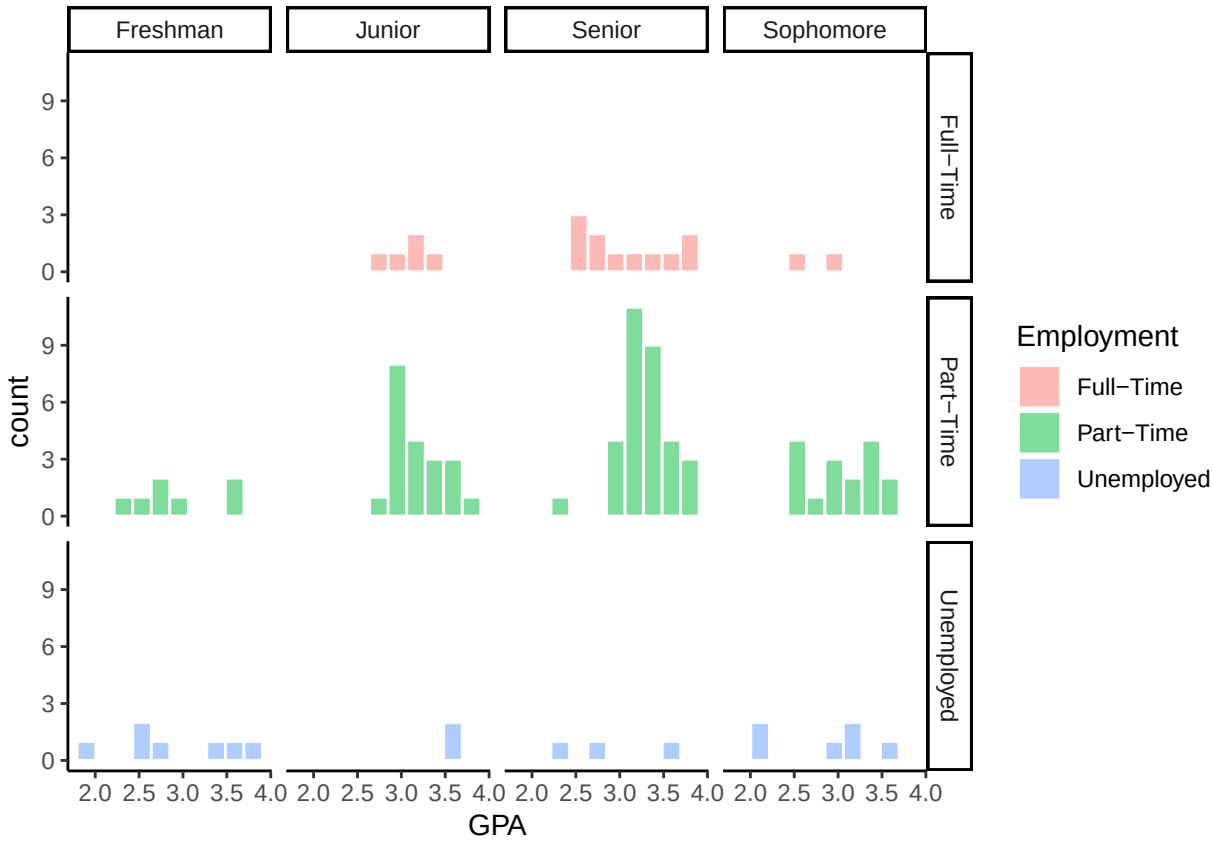
```
ggplot(data = iris) +
  geom_histogram(aes(x = Sepal.Length, fill = Species),
                 bins = 10, col = "white", alpha = 0.5) +
  facet_grid(rows = vars(Species))
```



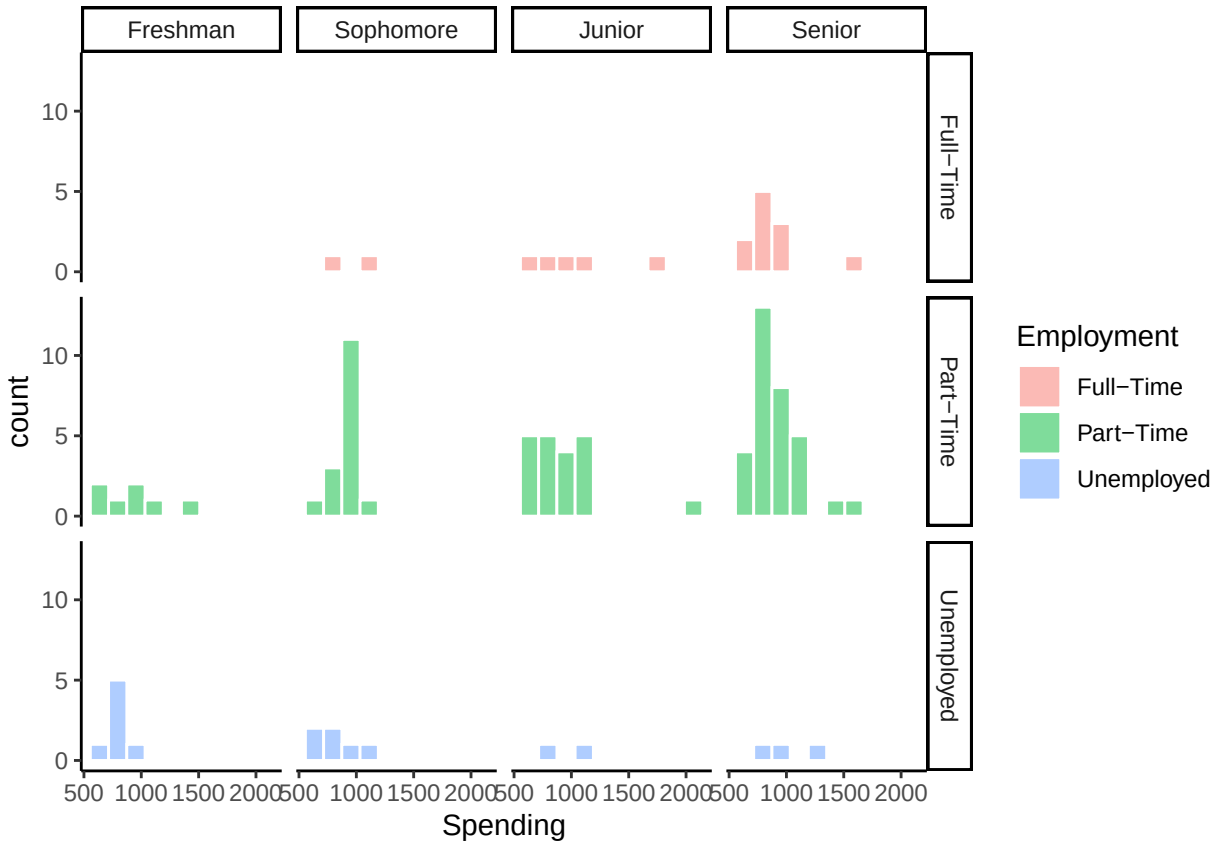

```
ggplot(data = student) +
  geom_histogram(aes(x = GPA, fill = Employment),
                 bins = 10, col = "white", alpha = 0.5) +
  facet_grid(rows =  vars(Employment), cols = vars(Gender))
```




```
ggplot(data = student) +
  geom_histogram(aes(x = GPA, fill = Employment),
                 bins = 10, col = "white", alpha = 0.5) +
  facet_grid(rows =  vars(Employment), cols = vars(Class))
```



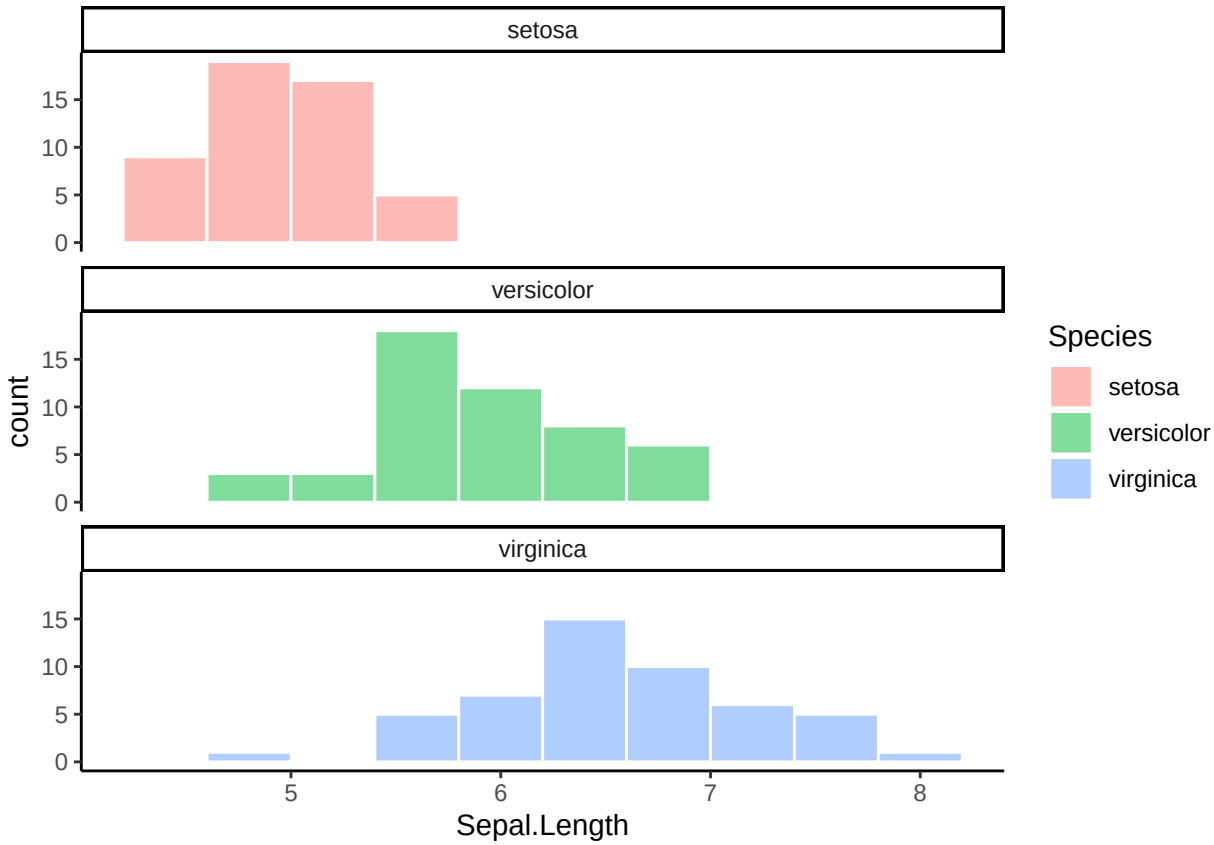

```
student %>%
  mutate(Class = factor(Class,
                        levels = c("Freshman","Sophomore", "Junior","Senior"))) %>%
  ggplot() +
  geom_histogram(aes(x = Spending, fill = Employment), bins = 10, col = "white", alpha = 0.5) +
  facet_grid(rows =  vars(Employment), cols = vars(Class))
```



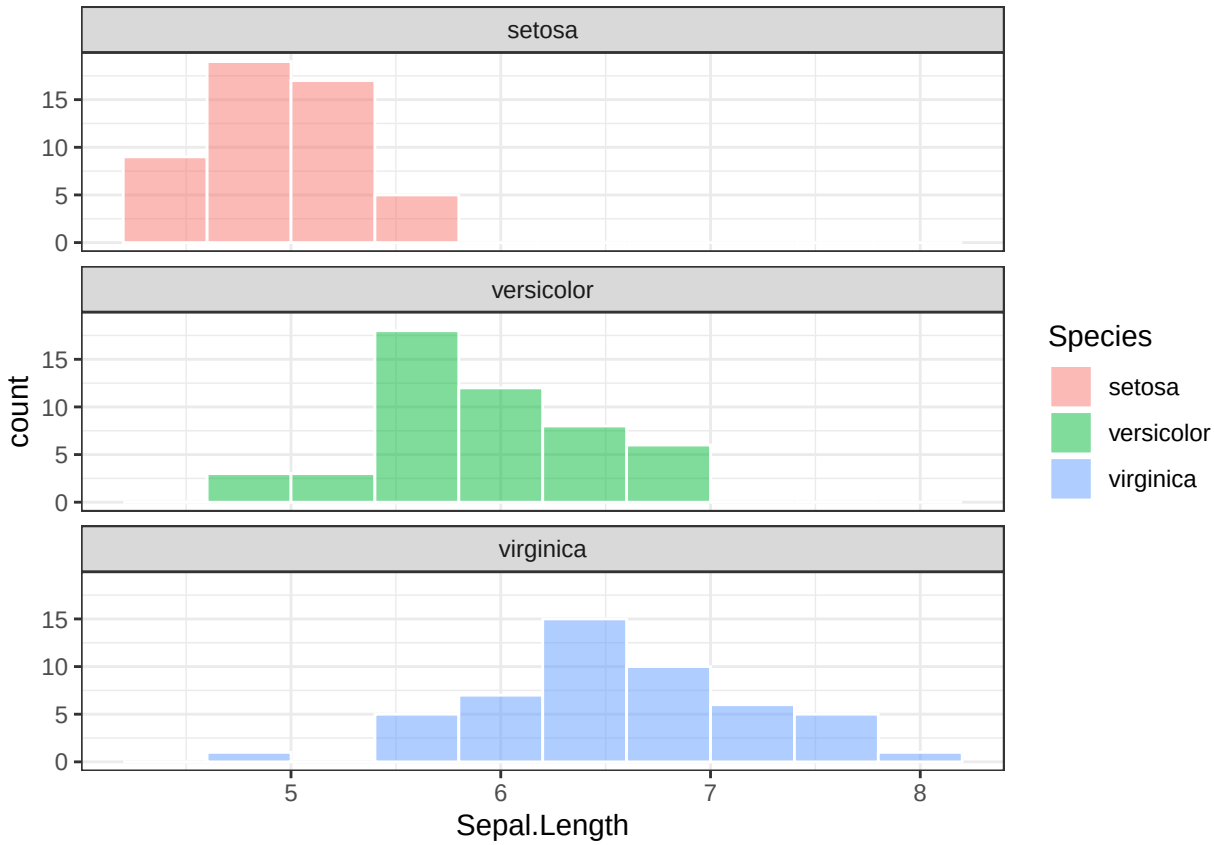
Theme

Built in Themes

```
ggplot(data = iris) +
  geom_histogram(aes(x = Sepal.Length, fill = Species),
                 bins = 10, col = "white", alpha = 0.5) +
  facet_wrap(vars(Species), ncol = 1) +
  theme_classic()
```

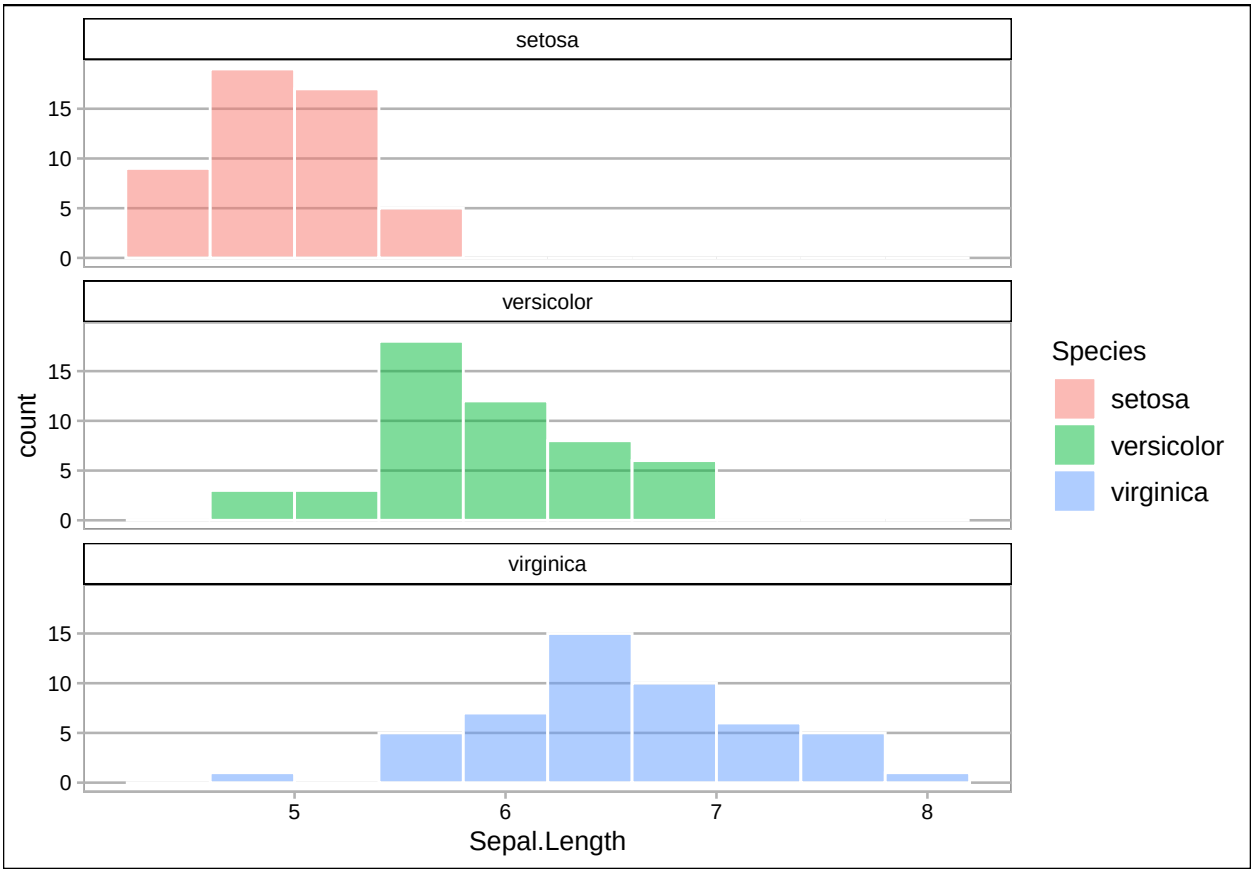



```
ggplot(data = iris) +
  geom_histogram(aes(x = Sepal.Length, fill = Species), bins = 10, col = "white", alpha = 0.5) +
  facet_wrap(vars(Species), ncol = 1) +
  theme_bw()
```



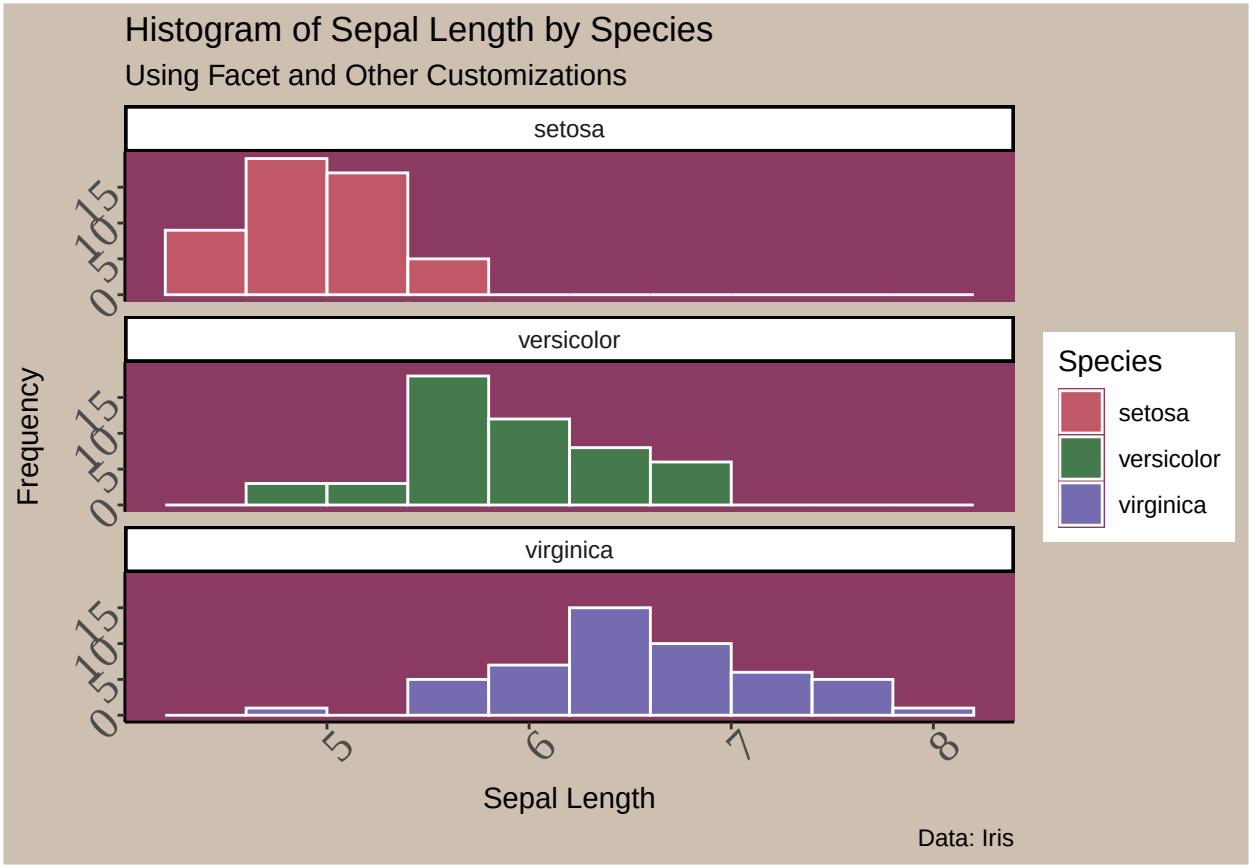
Themes from other packages

```
ggplot(data = iris) +
  geom_histogram(aes(x = Sepal.Length, fill = Species),
                 bins = 10, col = "white", alpha = 0.5) +
  facet_wrap(vars(Species), ncol = 1) +
  theme_calc()
```



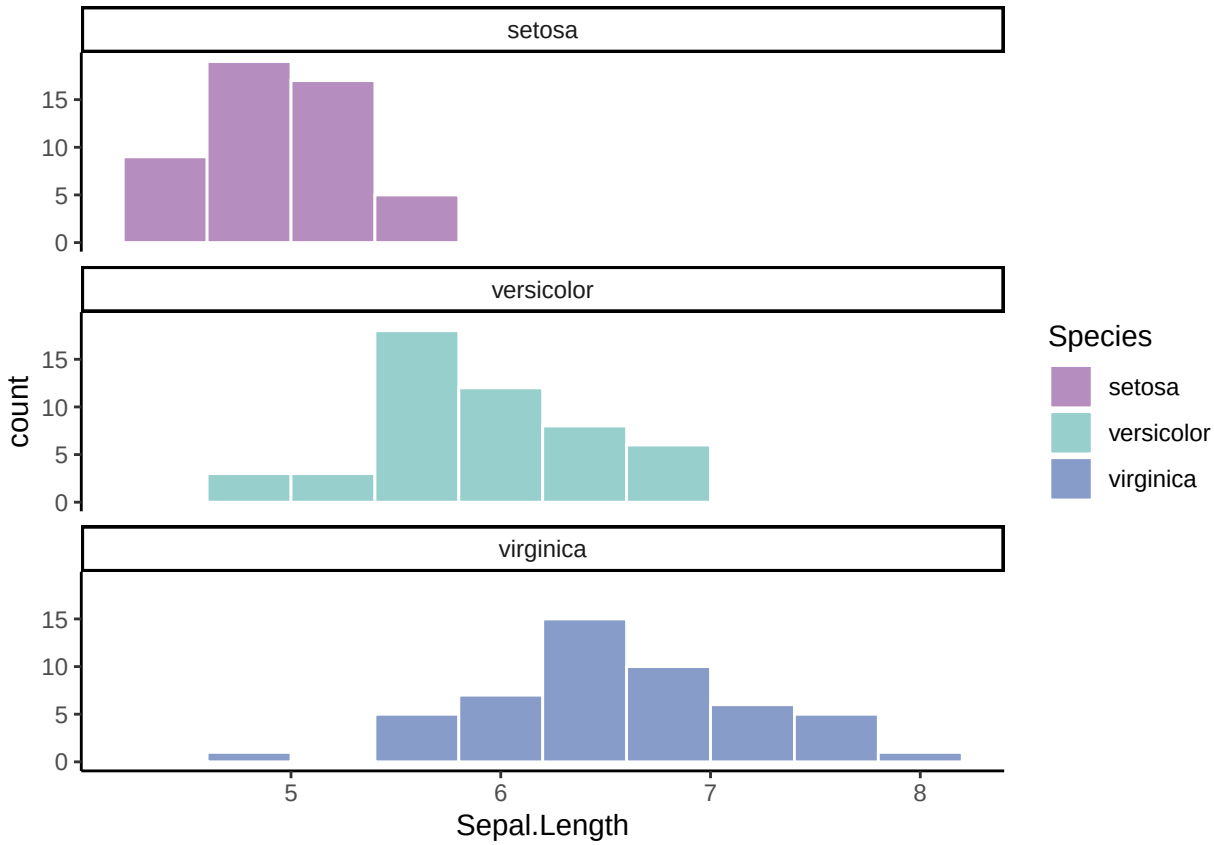

```r
p1 <- ggplot(data = iris) +
  geom_histogram(aes(x = Sepal.Length, fill = Species),
                 bins = 10, col = "white", alpha = 0.5) +
  facet_wrap(vars(Species), ncol = 1) +
  labs(
    title = "Histogram of Sepal Length by Species",
    x = "Sepal Length",
    y = "Frequency",
    fill = "Species",
    subtitle = "Using Facet and Other Customizations",
    caption = "Data: Iris"
  )

# ggThemeAssistGadget(p1)
p1 + theme(
  axis.text = element_text(family = "Times",
    size = 17, angle = 45),
  panel.background = element_rect(fill = "hotpink4"),
    plot.background = element_rect(fill = "antiquewhite3"))
```

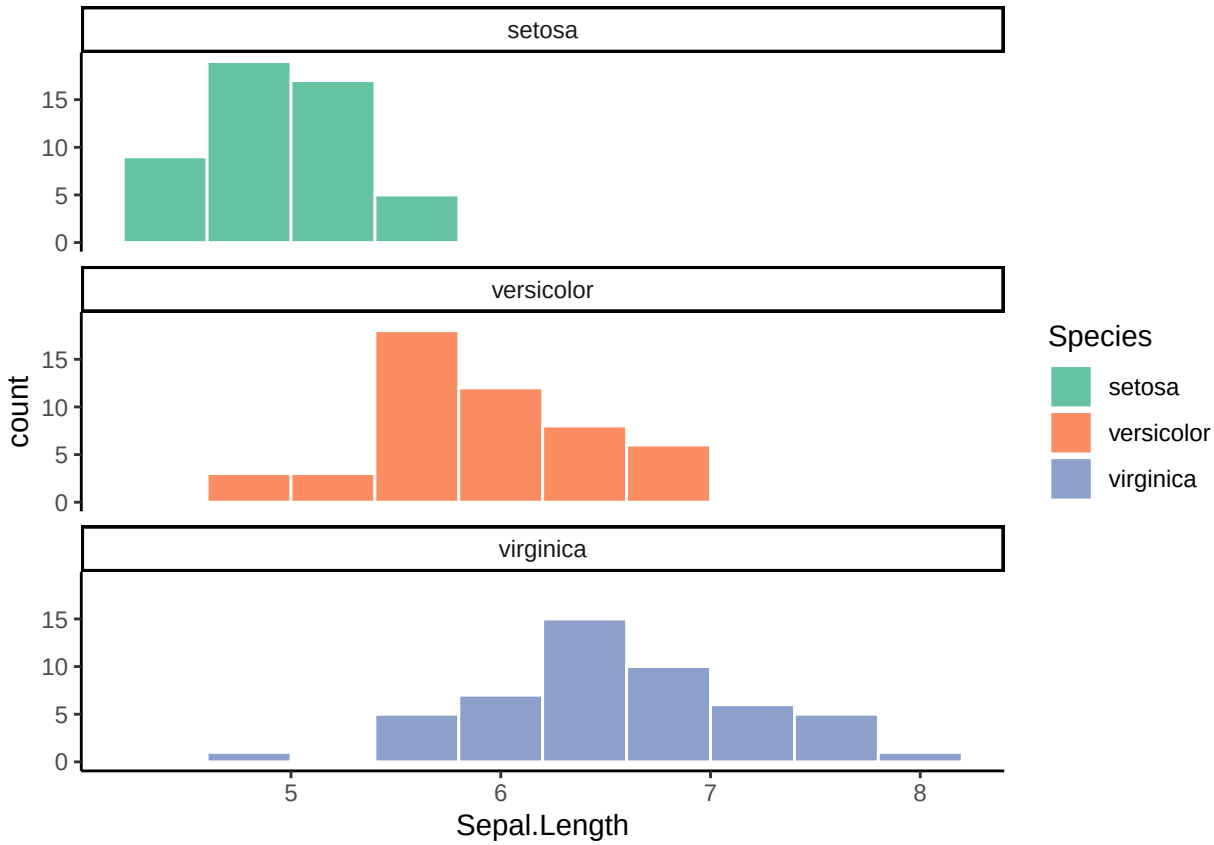


Manually Changing Color

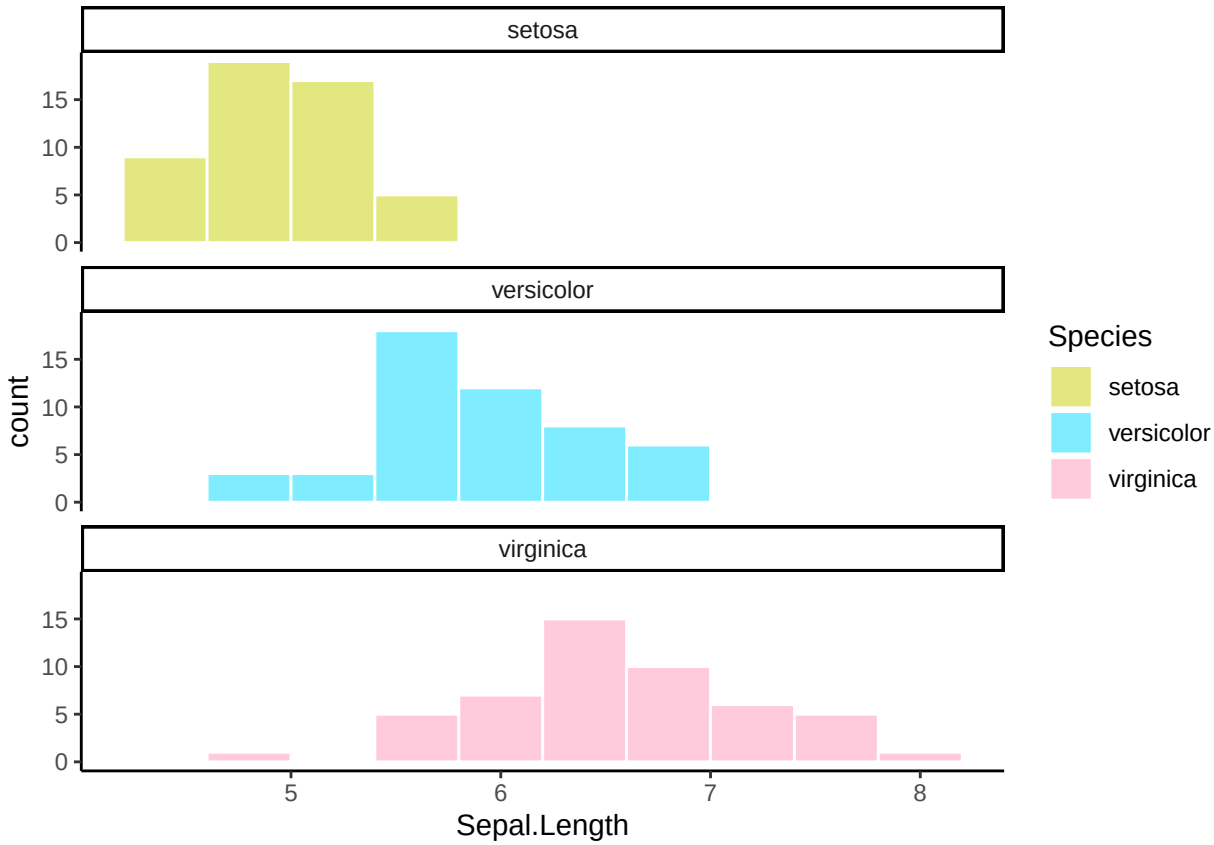
```
ggplot(data = iris) +
  geom_histogram(aes(x = Sepal.Length, fill = Species),
                 bins = 10, col = "white", alpha = 0.5) +
  facet_wrap(vars(Species), ncol = 1) +
  scale_fill_manual(values = c("setosa" = "#6C1C80", "versicolor" = "#30A19C", "virginica" = "#123B96"))
```




```
ggplot(data = iris) +
  geom_histogram(aes(x = Sepal.Length, fill = Species), bins = 10, col = "white", alpha = 1) +
  facet_wrap(vars(Species), ncol = 1) +
  scale_fill_brewer(palette = "Set2")
```




```
ggplot(data = iris) +
  geom_histogram(aes(x = Sepal.Length, fill = Species), bins = 10, col = "white", alpha = 0.5) +
  facet_wrap(vars(Species), ncol = 1) +
  scale_fill_hue(
    l = 80, c = 100,    # adjust luminosity and chroma
    h = c(90, 360)      # adjust range of hues
  )
```


setosa
versicolor
virginica
count
0
5
10
15
5
6
7
8
Sepal.Length
Species
setosa
versicolor
virginica