# RStudio Datasets

Nicholas Moore

2022-07-29

**Background:** Creating standard plots with RStudio's built-in datasets.

### Iris Data Overview

## Loading Data / Basic Plotting Tutorial

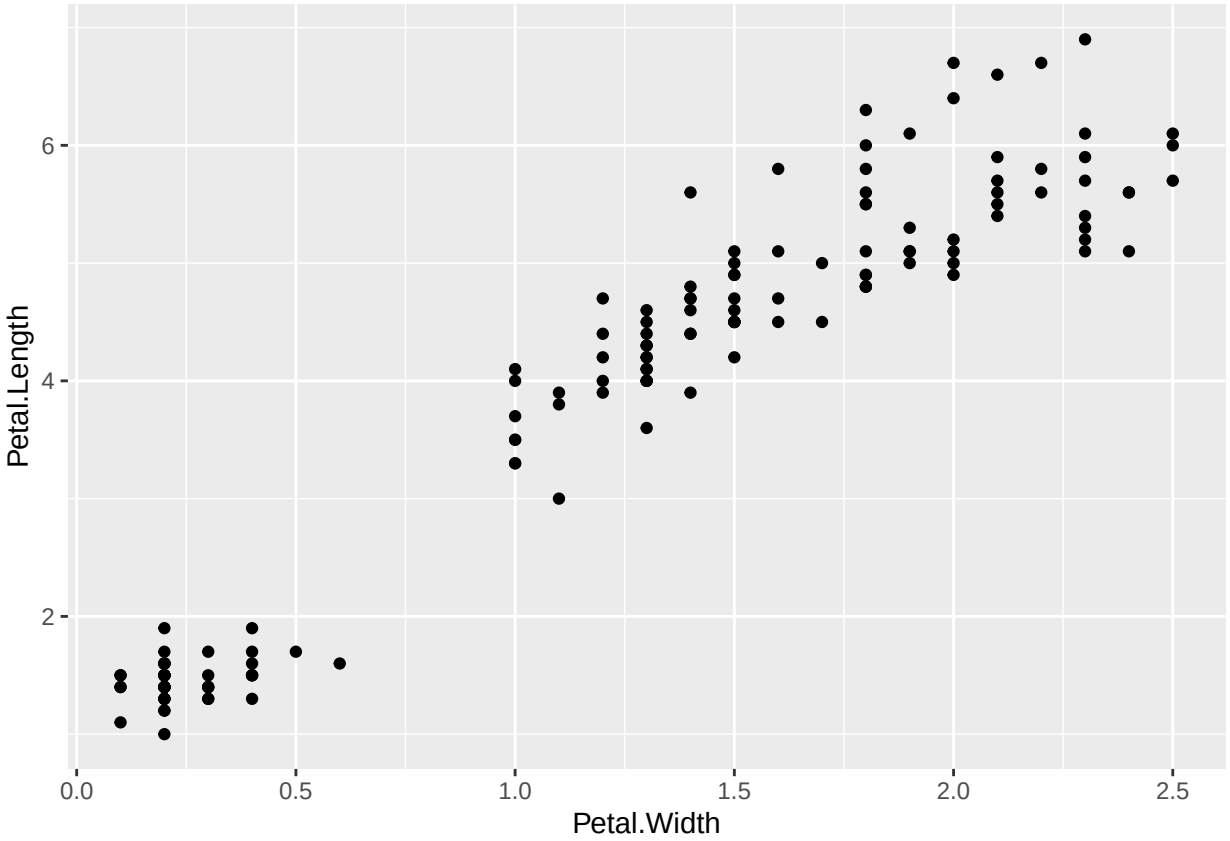
```
library(ggplot2) #qplot (basic scatterplot)
library(datasets)
library(tidyverse)
library(ggthemes)
library(tinytex)
library(prettydoc)
windowsFonts(Times = windowsFont("Times New Roman"))
library(ggthemes)
```

?datasets library(help = "datasets") ?iris

iris head(iris)

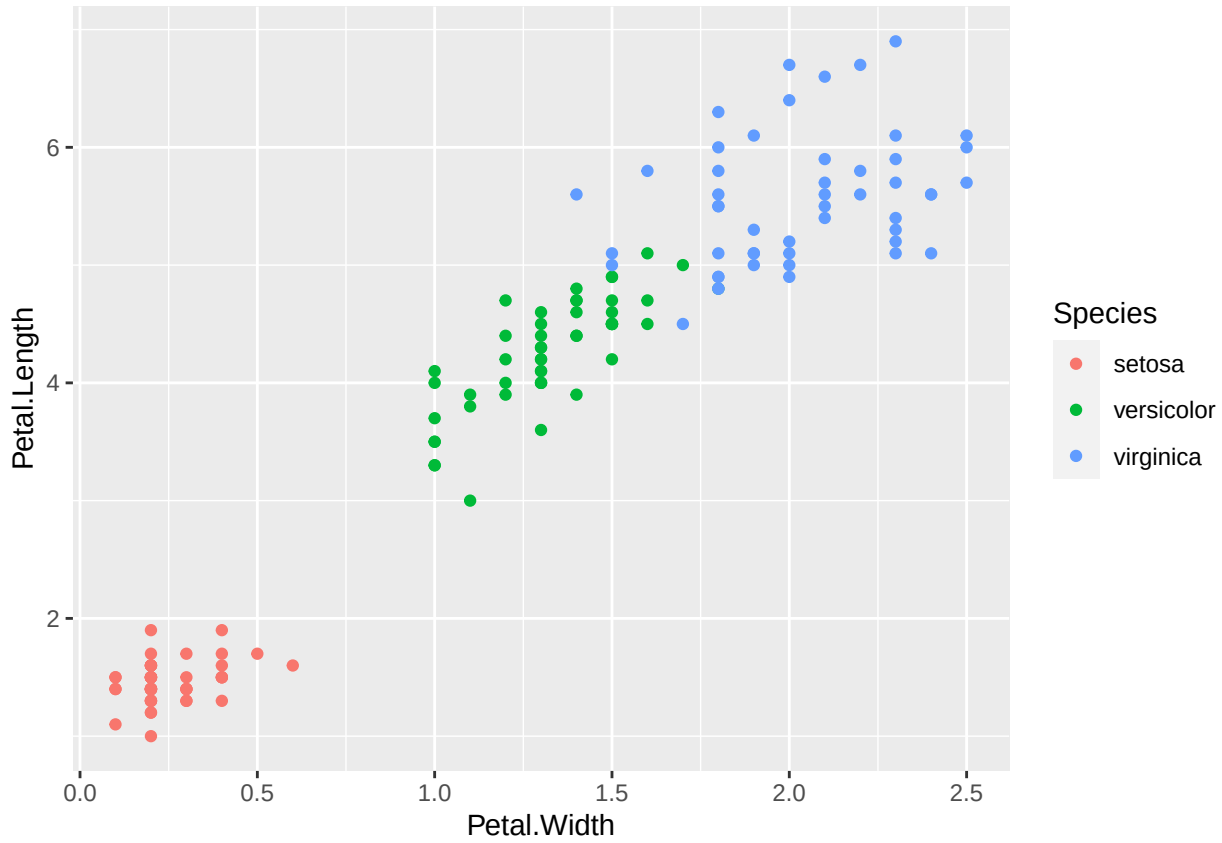
#basic scatterplot

```
qplot(Petal.Width, Petal.Length, data=iris)
```



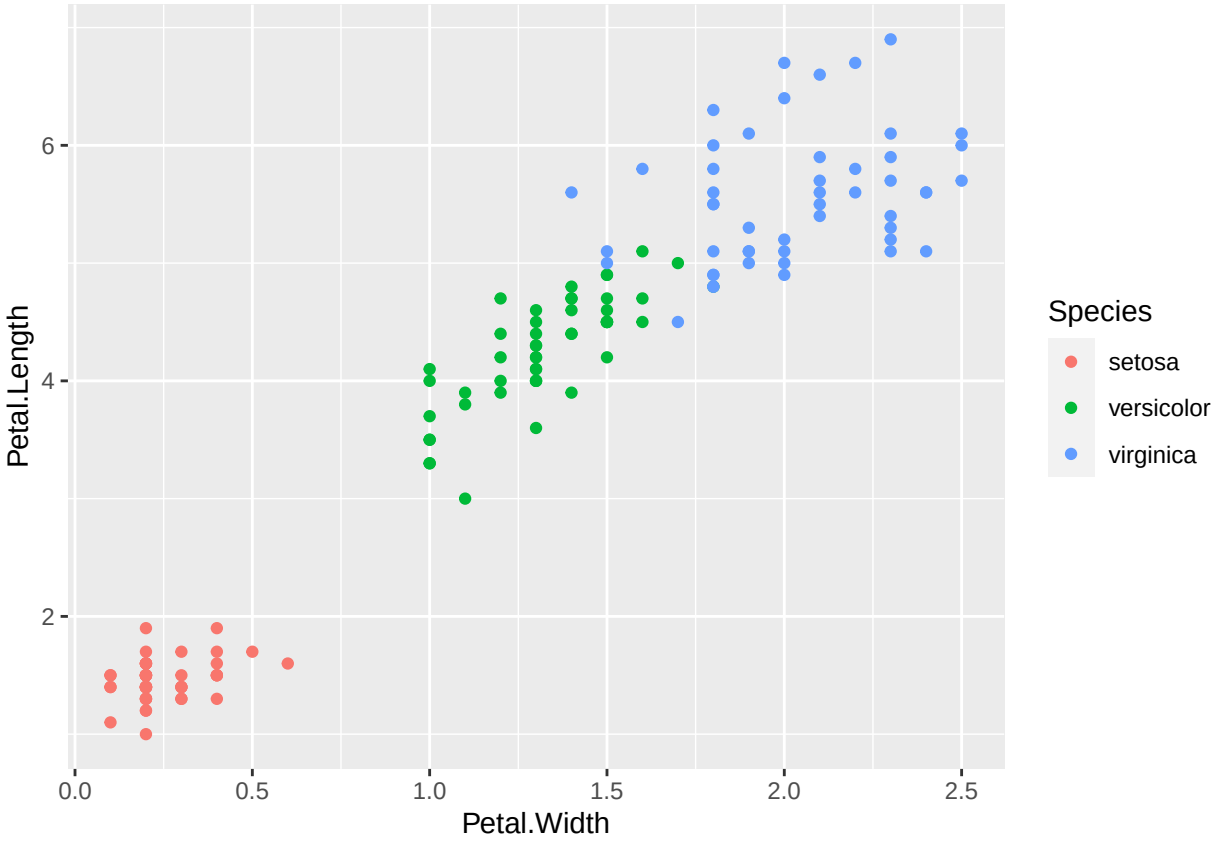

#basic scatterplot w/ species colored

```
qplot(Petal.Width, Petal.Length, color=Species, data=iris)
```



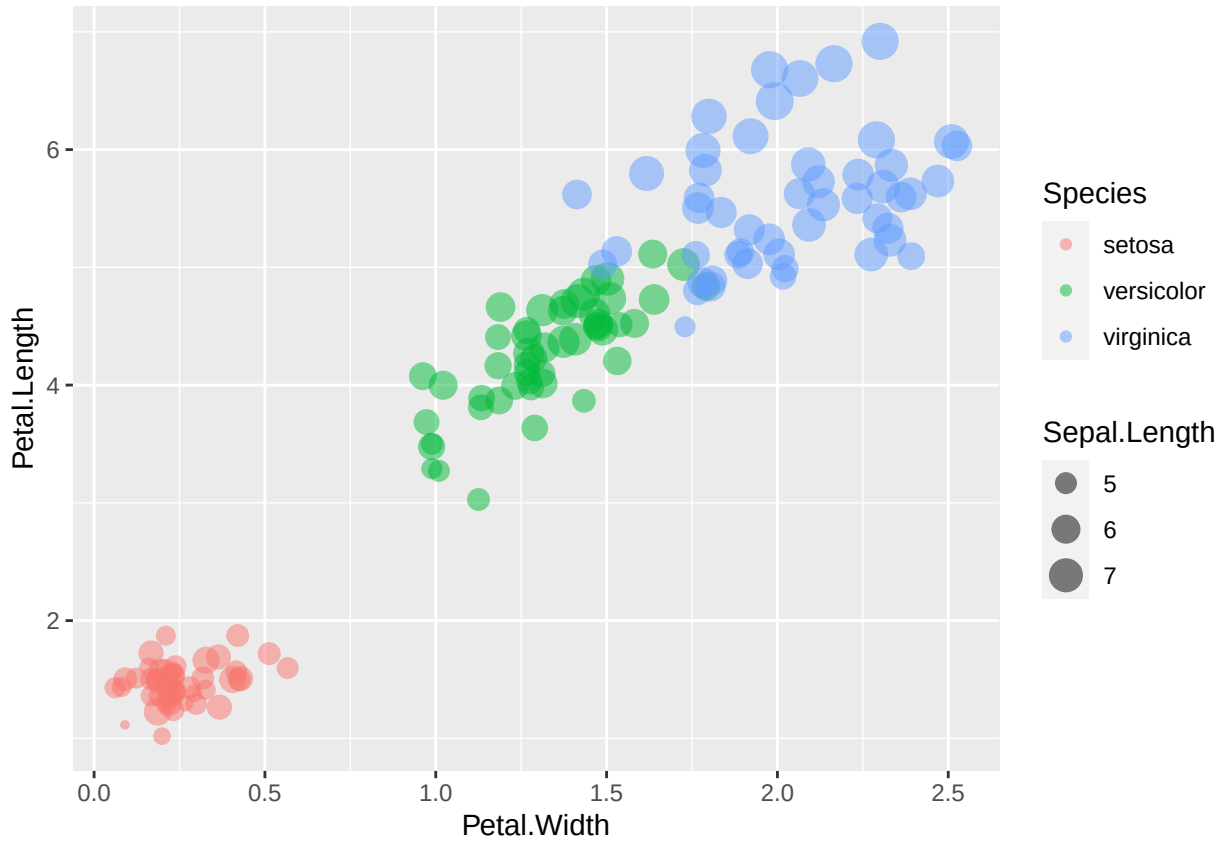

#same thing as above, but strictly ggplot2

```
ggplot(iris, aes(Petal.Width, Petal.Length, color=Species))+
      geom_point()
```



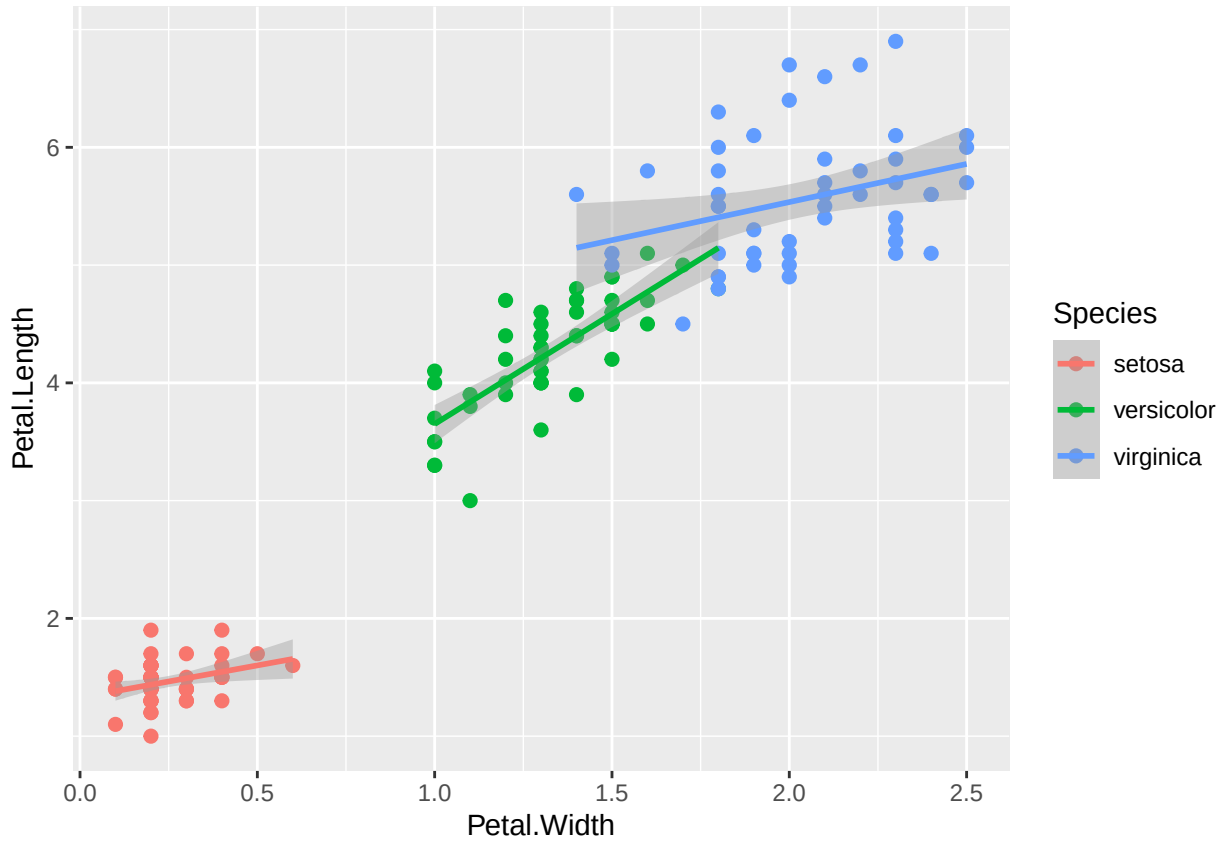

#size matches sepal length. jitter separates dots slightly for viewing. alpha makes dots more transparent

```
ggplot(iris, aes(Petal.Width, Petal.Length, color=Species, size=Sepal.Length))+
  geom_jitter(alpha=0.5)
```



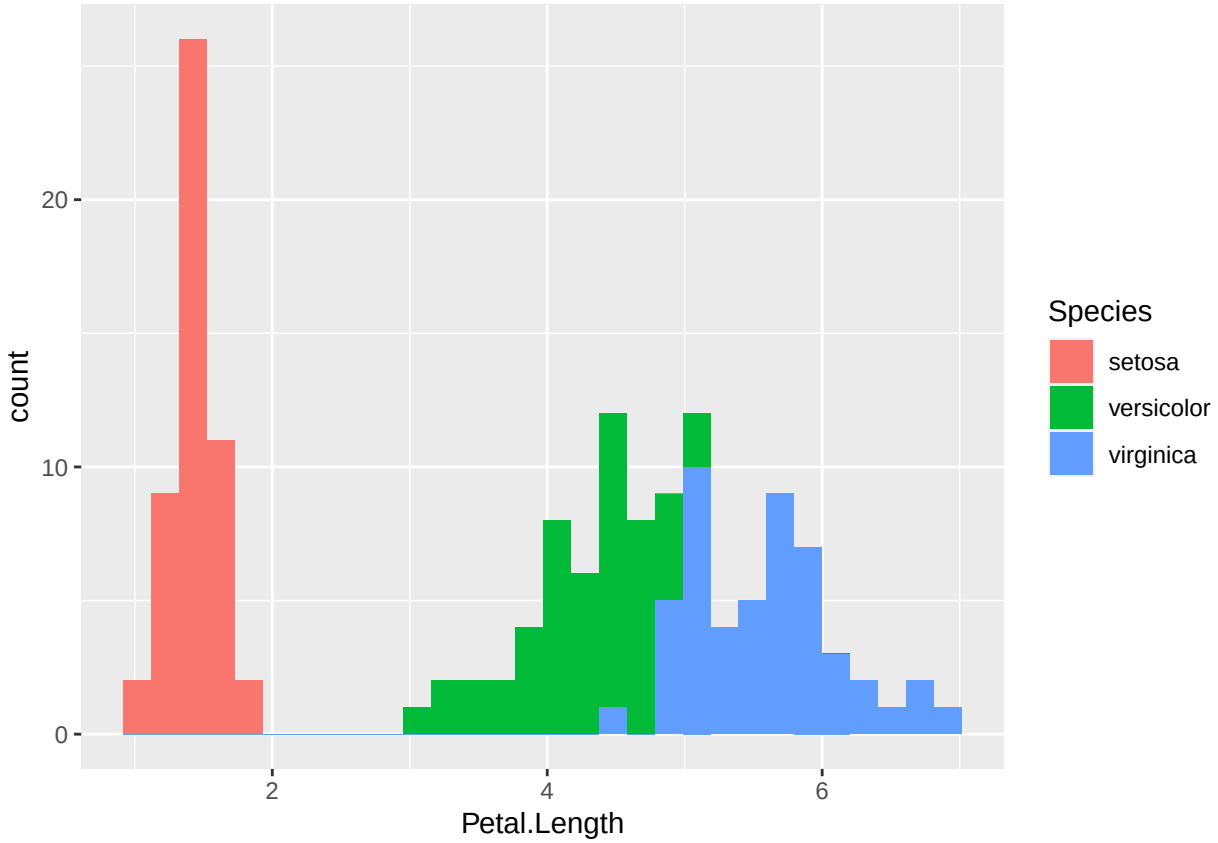

#size matches sepal length. jitter separates dots slightly for viewing. alpha makes dots more transparent. geom_smooth makes regression line with std. error

```
ggplot(iris, aes(Petal.Width, Petal.Length, color=Species, size=Sepal.Length))+
  geom_point(size=2)+
  geom_smooth(method=lm)+
  theme(legend.position = "right")
```



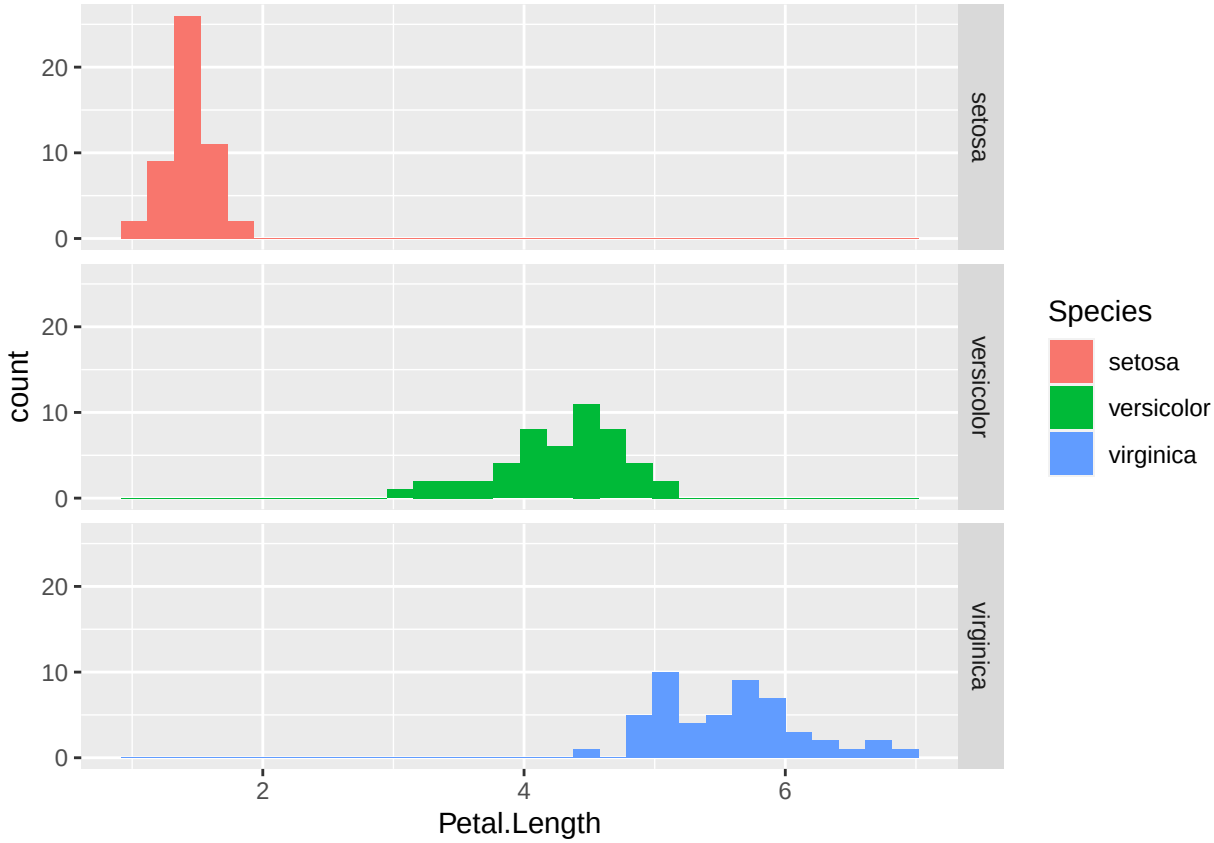

#histogram

```
ggplot(iris, aes(Petal.Length, fill=Species))+
  geom_histogram()+
  theme(legend.position = "right")
```




#histogram w/ indiv. plots for species w/ facet_grid

```
ggplot(iris, aes(Petal.Length, fill=Species))+
  geom_histogram()+
  facet_grid(Species ~ .)
```

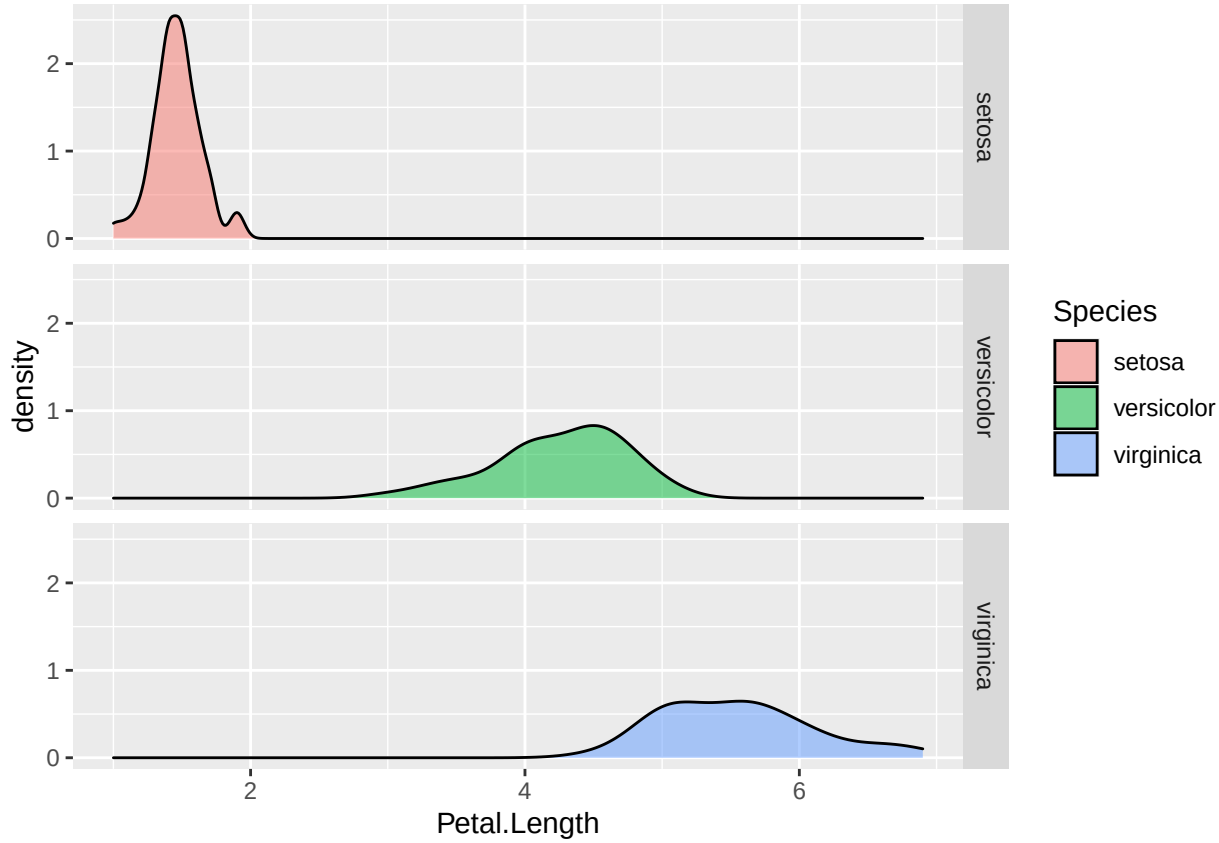



```
  theme(legend.position = "none")
```

```
## List of 1
##  $ legend.position: chr "none"
##  - attr(*, "class")= chr [1:2] "theme" "gg"
##  - attr(*, "complete")= logi FALSE
##  - attr(*, "validate")= logi TRUE
```

#histogram w/ indiv. plots for species w/ facet_grid (geom_density is cleaner, less blocky)

```
ggplot(iris, aes(Petal.Length, fill=Species))+
  geom_density(alpha=0.5)+
  facet_grid(Species ~ .)
```

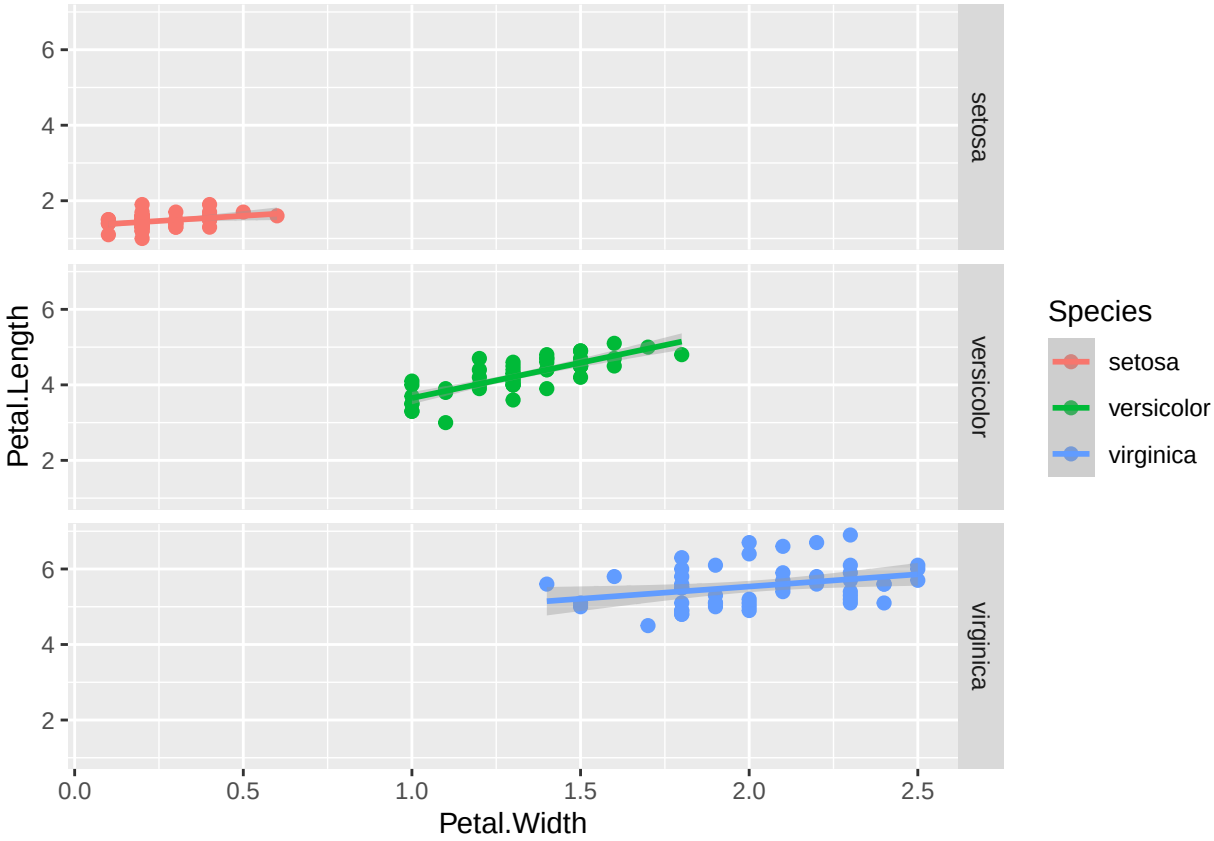



```
theme(legend.position = "none")
```

```
## List of 1
##  $ legend.position: chr "none"
##  - attr(*, "class")= chr [1:2] "theme" "gg"
##  - attr(*, "complete")= logi FALSE
##  - attr(*, "validate")= logi TRUE
```

#facet_grid with previous scatterplots

```
ggplot(iris, aes(Petal.Width, Petal.Length, color=Species, size=Sepal.Length))+
  geom_point(size=2)+
  geom_smooth(method=lm)+
  facet_grid(Species ~ .)
```



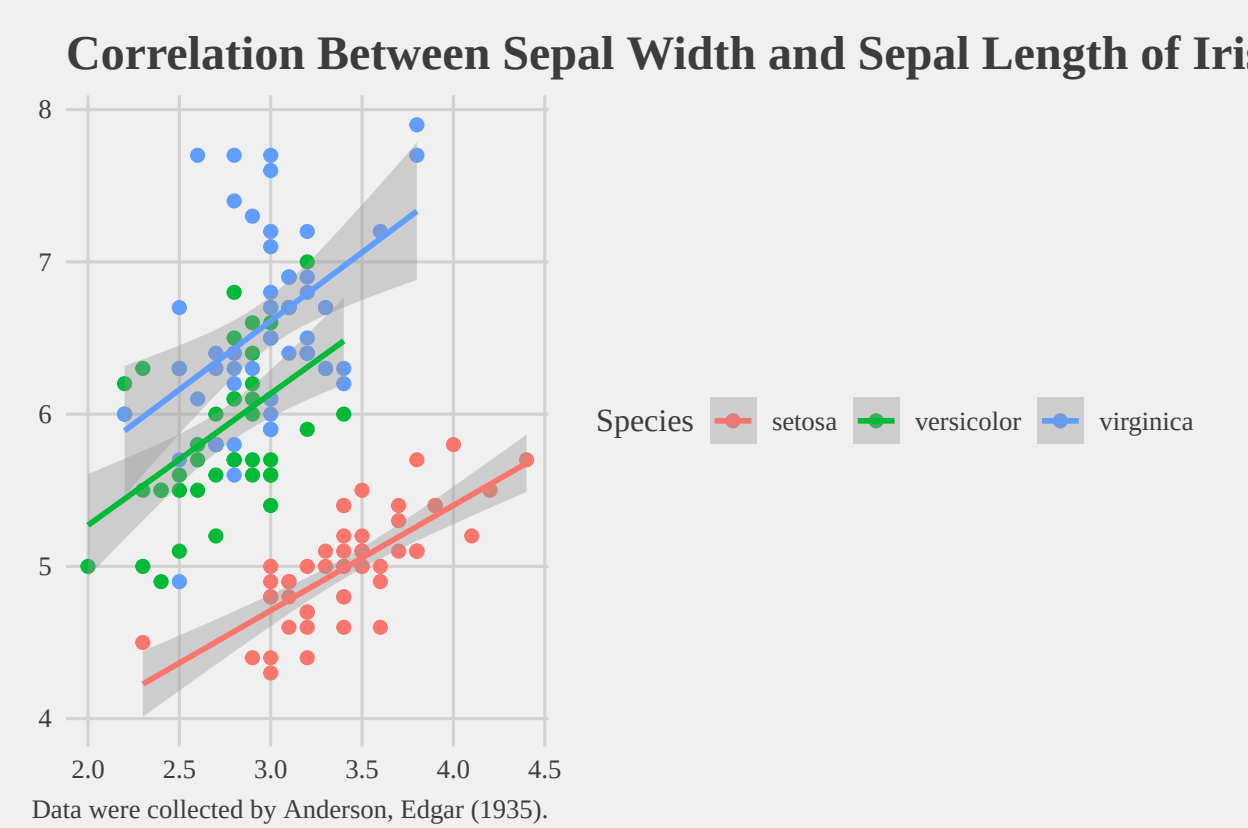

```
  theme(legend.position = "right")
```

```
## List of 1
##  $ legend.position: chr "right"
##  - attr(*, "class")= chr [1:2] "theme" "gg"
##  - attr(*, "complete")= logi FALSE
##  - attr(*, "validate")= logi TRUE
```

## Key Figures

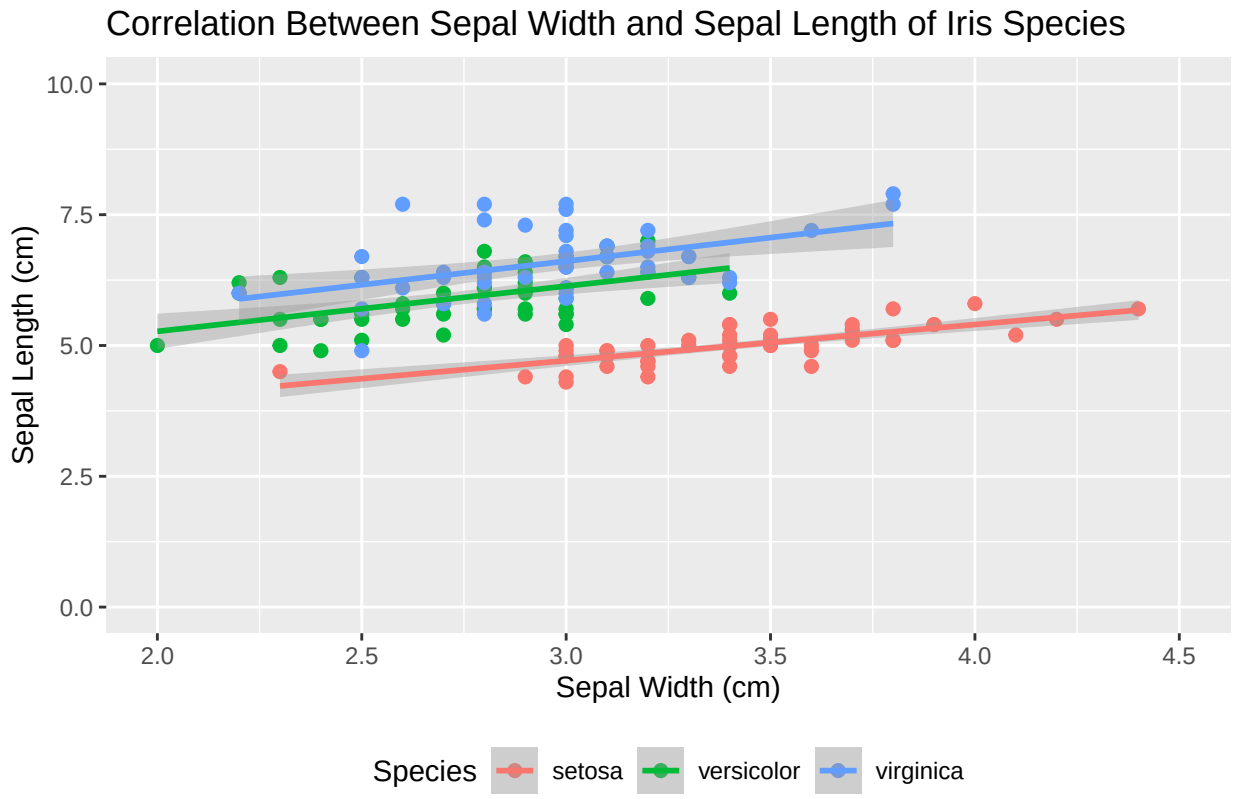
windowsFonts(Times = windowsFont("Times New Roman")) library(ggthemes)

```
ggplot(iris, aes(Sepal.Width, Sepal.Length, color=Species, size=Sepal.Length))+
  geom_point(size=2)+
  geom_smooth(method=lm)+
  theme_fivethirtyeight(base_size = 12, base_family = "Times")+
  labs(title="Correlation Between Sepal Width and Sepal Length of Iris Species",
       caption="Data were collected by Anderson, Edgar (1935).")+
  labs(x = "Sepal Width")+
  labs(y = "Sepal Length")+
  theme(legend.position = "right")
```



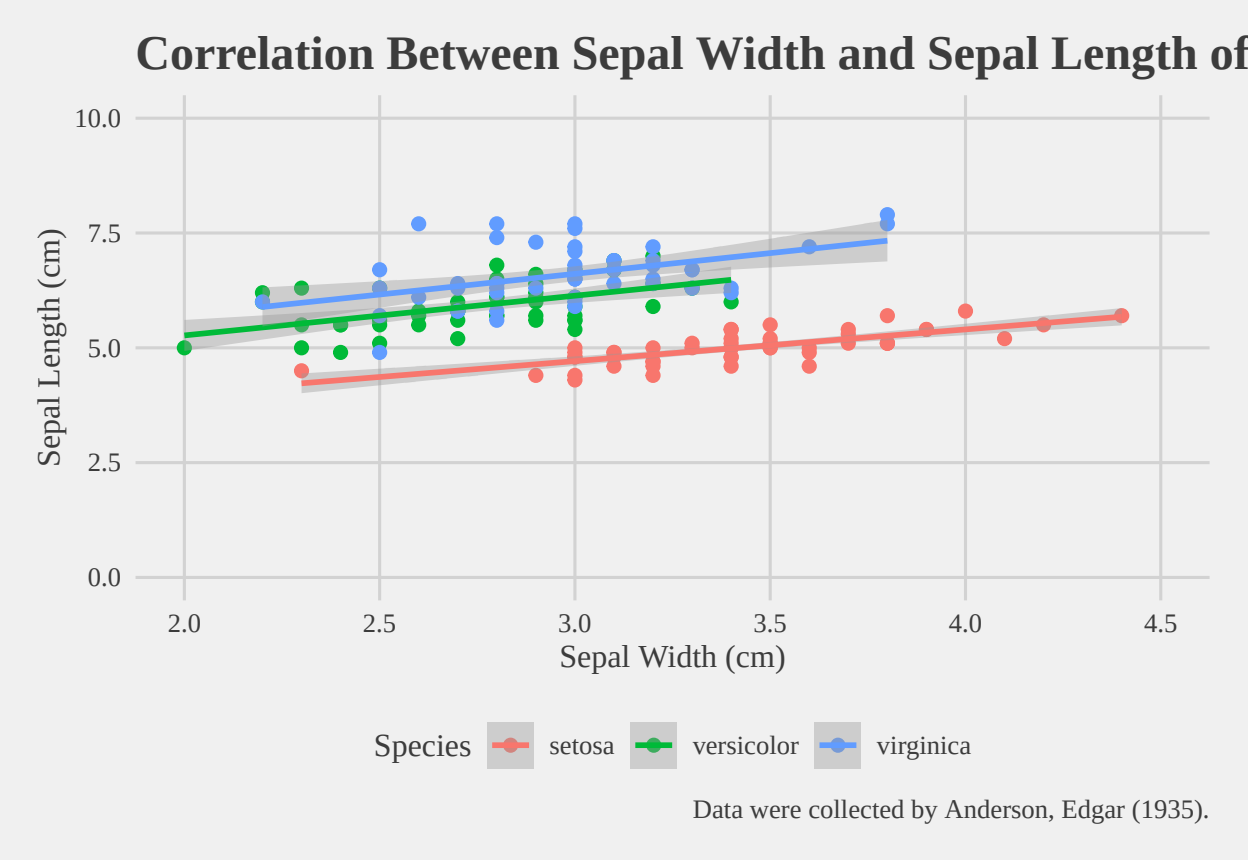

#SEPAL WIDTH VS. SEPAL LENGTH w/out theme

```
ggplot(iris, aes(Sepal.Width, Sepal.Length, color=Species, size=Sepal.Length))+
  geom_point(size=2)+
  geom_smooth(method=lm)+
  labs(title="Correlation Between Sepal Width and Sepal Length of Iris Species",
       caption="Data were collected by Anderson, Edgar (1935).")+
  labs(x = "Sepal Width (cm)")+
  labs(y = "Sepal Length (cm)")+
  xlim(2.0, 4.5)+
  ylim(0.0, 10.0)+
  theme(legend.position = "bottom")
```



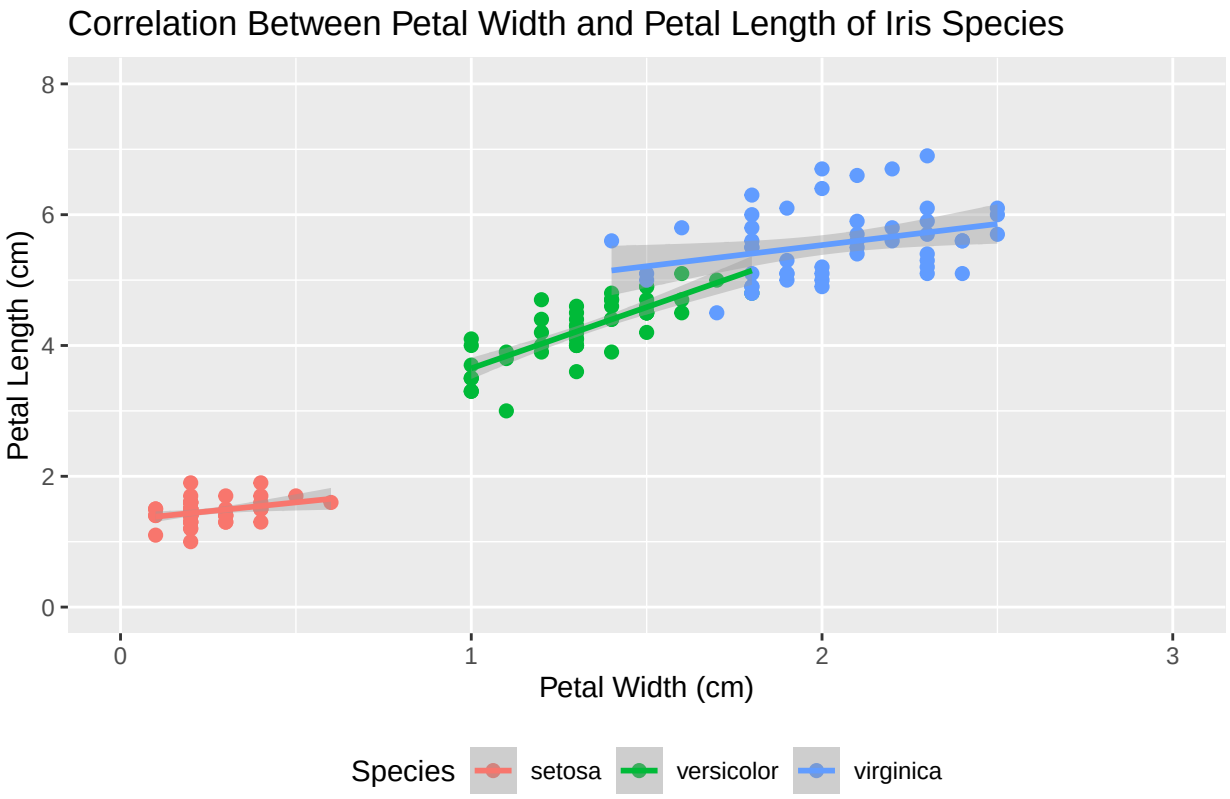

#SEPAL WIDTH VS. SEPAL LENGTH w/ theme

```
ggplot(iris, aes(Sepal.Width, Sepal.Length, color=Species, size=Sepal.Length))+
  geom_point(size=2)+
  geom_smooth(method=lm)+
  theme_fivethirtyeight(base_size = 12, base_family = "Times")+
  labs(title="Correlation Between Sepal Width and Sepal Length of Iris Species",
       caption="Data were collected by Anderson, Edgar (1935).")+
  theme(axis.title = element_text()) + ylab('Sepal Length (cm)') + xlab('Sepal Width (cm)')+
  xlim(2.0, 4.5)+
  ylim(0.0, 10.0)+
  theme(legend.position = "bottom")
```



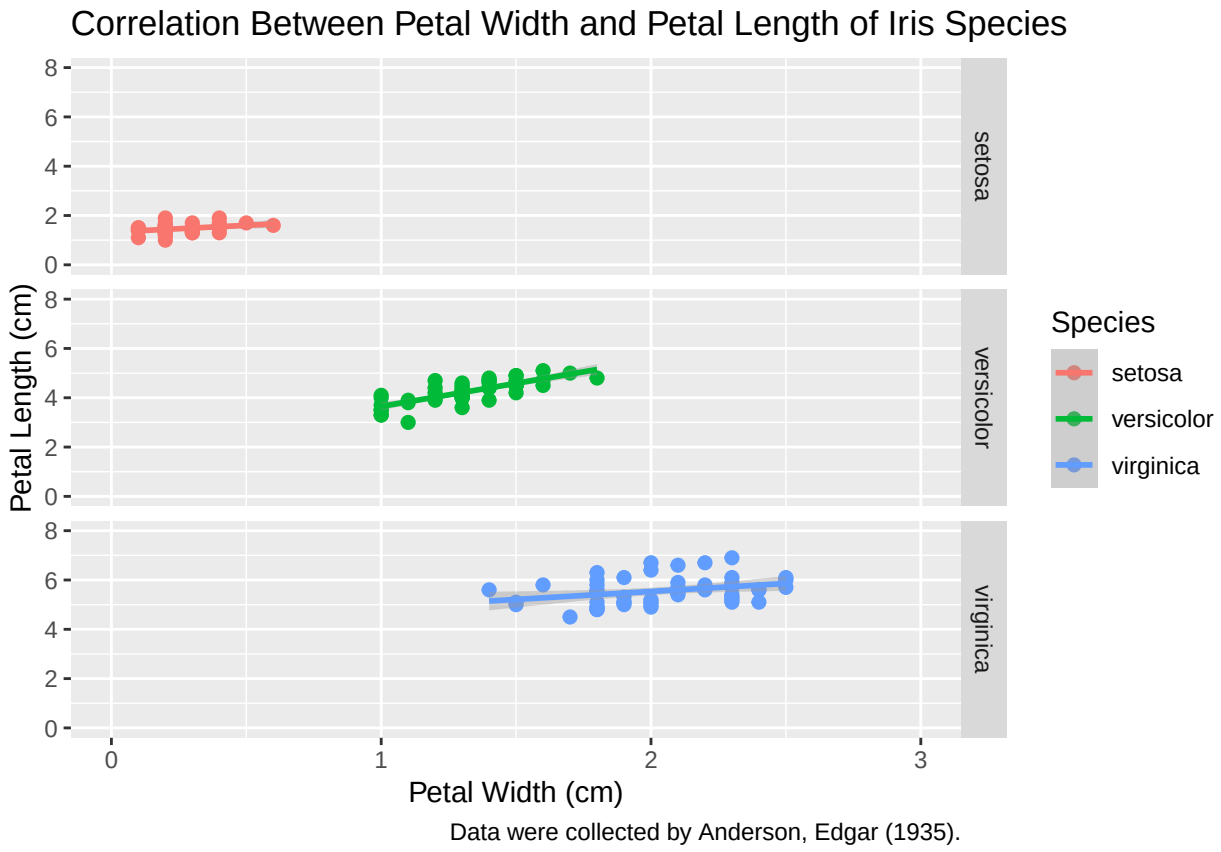

#PETAL WIDTH VS. PETAL LENGTH w/out theme

```
ggplot(iris, aes(Petal.Width, Petal.Length, color=Species, size=Petal.Length))+
  geom_point(size=2)+
  geom_smooth(method=lm)+
  labs(title="Correlation Between Petal Width and Petal Length of Iris Species",
       caption="Data were collected by Anderson, Edgar (1935).")+
  labs(x = "Petal Width (cm)")+
  labs(y = "Petal Length (cm)")+
  xlim(0.0, 3.0)+
  ylim(0.0, 8.0)+
  theme(legend.position = "bottom")
```



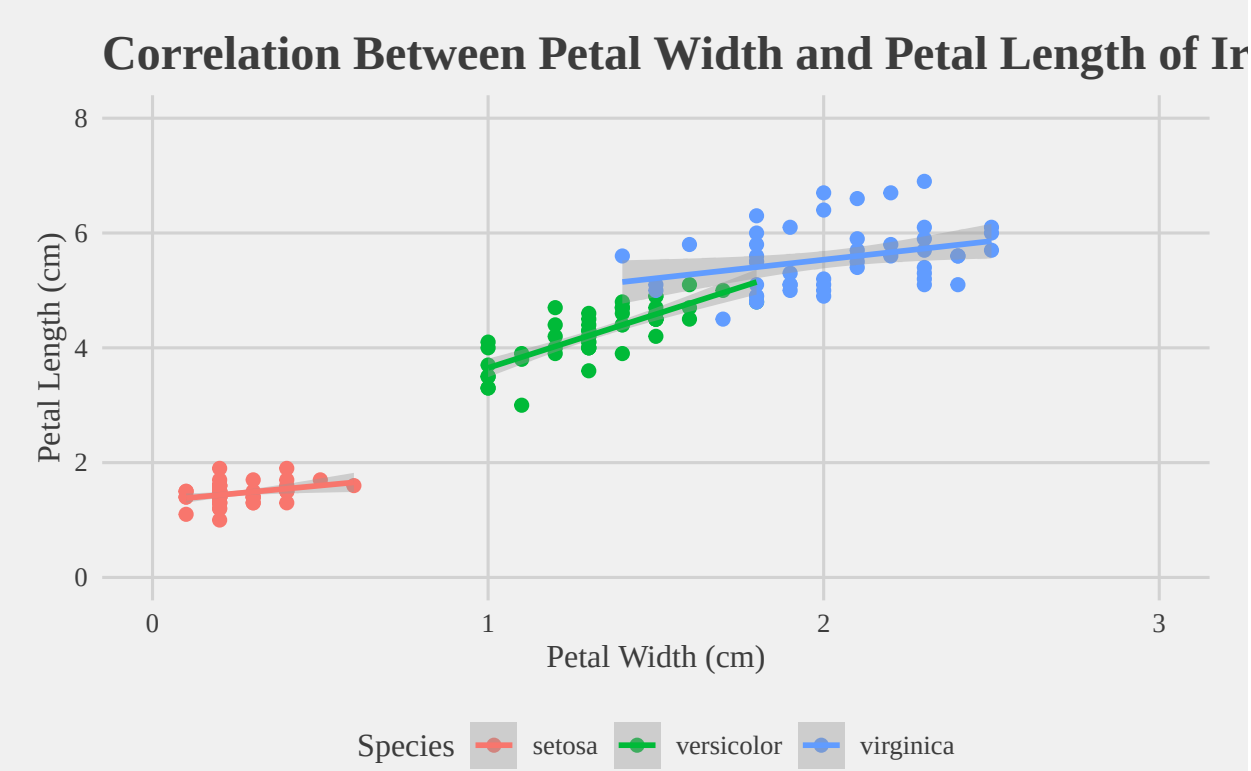

#PETAL WIDTH VS. PETAL LENGTH w/ theme & facet

```
ggplot(iris, aes(Petal.Width, Petal.Length, color=Species, size=Petal.Length))+
  geom_point(size=2)+
  geom_smooth(method=lm)+
  labs(title="Correlation Between Petal Width and Petal Length of Iris Species",
       caption="Data were collected by Anderson, Edgar (1935).")+
  labs(x = "Petal Width (cm)")+
  labs(y = "Petal Length (cm)")+
  xlim(0.0, 3.0)+
  ylim(0.0, 8.0)+
  facet_grid(Species ~ .)+
  theme(legend.position = "right")
```



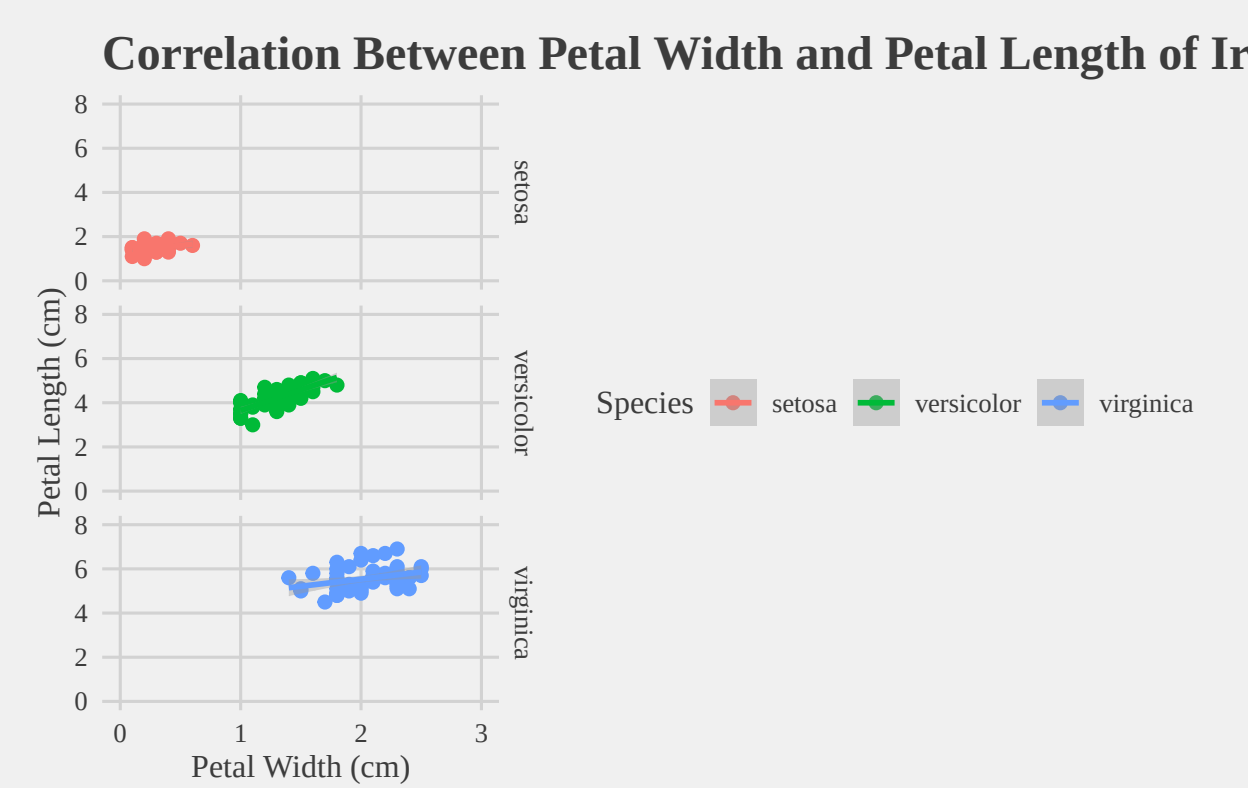

#PETAL WIDTH VS. PETAL LENGTH w/ theme

```
ggplot(iris, aes(Petal.Width, Petal.Length, color=Species, size=Petal.Length))+
  geom_point(size=2)+
  geom_smooth(method=lm)+
  theme_fivethirtyeight(base_size = 12, base_family = "Times")+
  labs(title="Correlation Between Petal Width and Petal Length of Iris Species",
       caption="Data were collected by Anderson, Edgar (1935).")+
  theme(axis.title = element_text()) + ylab('Petal Length (cm)') + xlab('Petal Width (cm)')+
  xlim(0.0, 3.0)+
  ylim(0.0, 8.0)+
  theme(legend.position = "bottom")
```



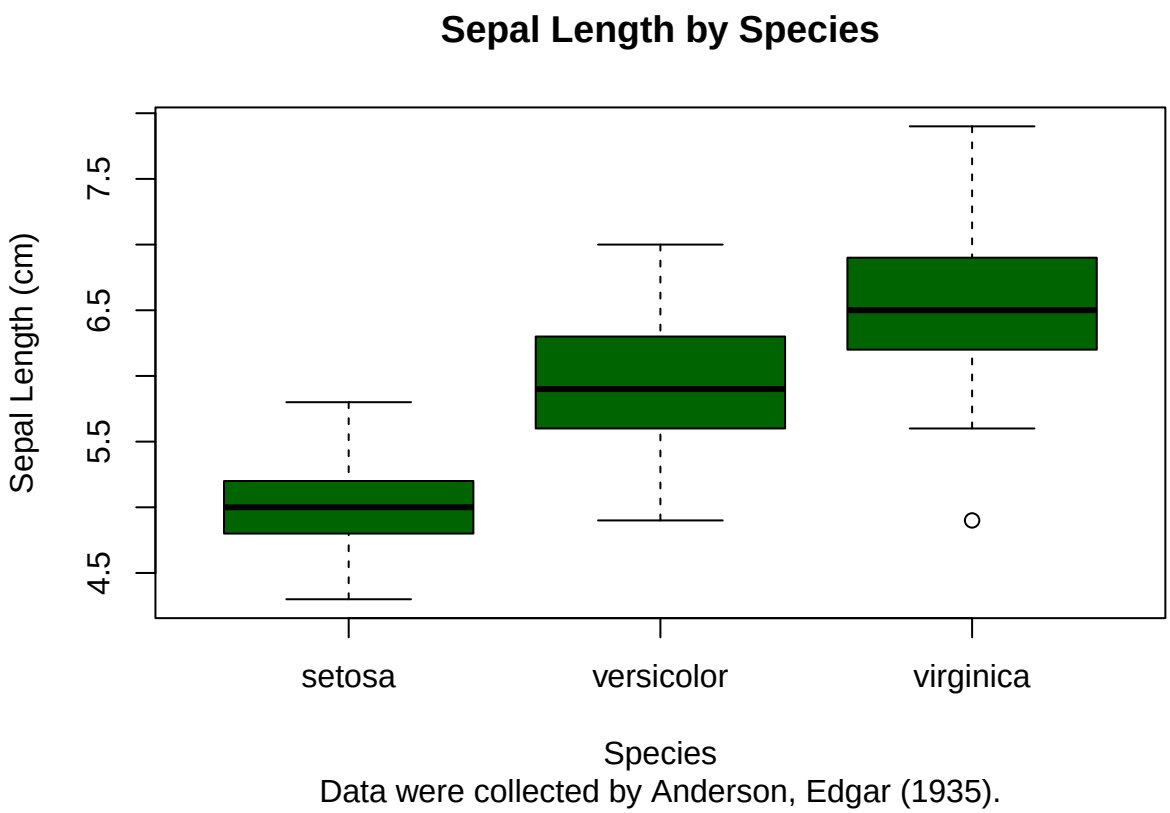

#PETAL WIDTH VS. PETAL LENGTH w/ theme & facet

```
ggplot(iris, aes(Petal.Width, Petal.Length, color=Species, size=Petal.Length))+
  geom_point(size=2)+
  geom_smooth(method=lm)+
  theme_fivethirtyeight(base_size = 12, base_family = "Times")+
  labs(title="Correlation Between Petal Width and Petal Length of Iris Species",
       caption="Data were collected by Anderson, Edgar (1935).")+
  theme(axis.title = element_text()) + ylab('Petal Length (cm)') + xlab('Petal Width (cm)')+
  xlim(0.0, 3.0)+
  ylim(0.0, 8.0)+
  facet_grid(Species ~ .)+
  theme(legend.position = "right")
```



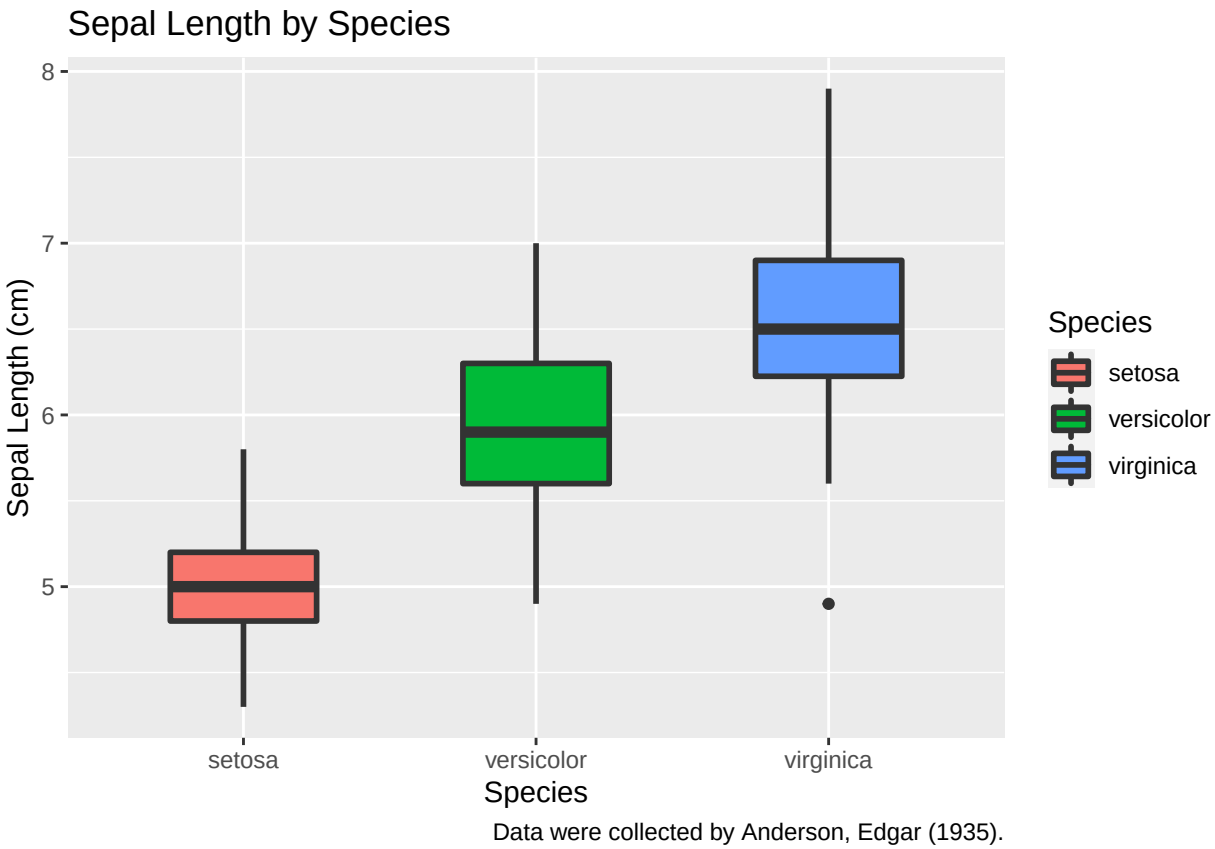

#Boxplot NO ggplot2 Sepal Length

```
boxplot(Sepal.Length~Species,
        data=iris,
        main='Sepal Length by Species',
        sub='Data were collected by Anderson, Edgar (1935).',
        xlab='Species',
        ylab='Sepal Length (cm)',
        col='darkgreen',
        border='black')
```



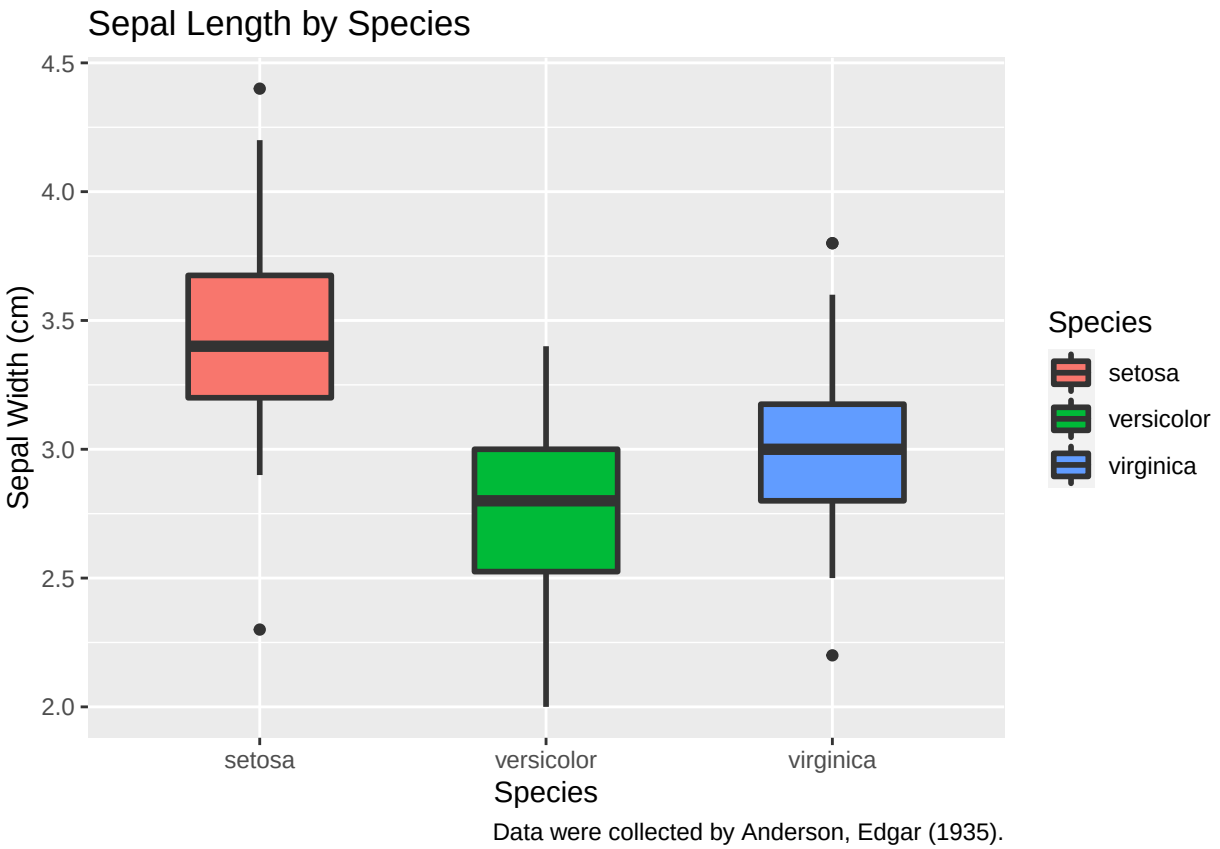

#Boxplot w/ ggplot2 Sepal Length

```
ggplot(iris, aes(x=Species, y=Sepal.Length, fill=Species)) +
  geom_boxplot(width=0.5,lwd=1)+
  labs(title="Sepal Length by Species", x = "Species", y = "Sepal Length (cm)",
       caption="Data were collected by Anderson, Edgar (1935).")
```



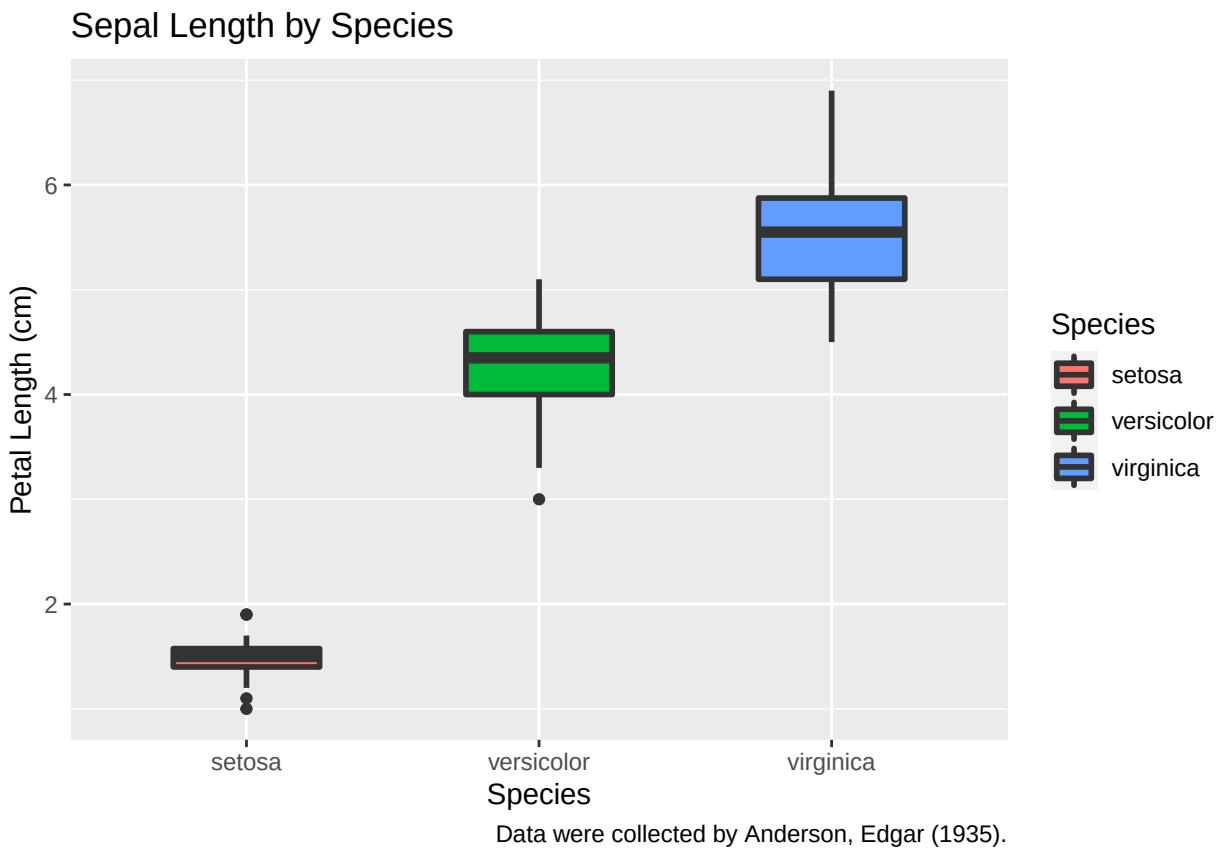

#Boxplot w/ ggplot2 Sepal Width

```
ggplot(iris, aes(x=Species, y=Sepal.Width, fill=Species)) +
  geom_boxplot(width=0.5,lwd=1)+
  labs(title="Sepal Length by Species", x = "Species", y = "Sepal Width (cm)",
       caption="Data were collected by Anderson, Edgar (1935).")
```



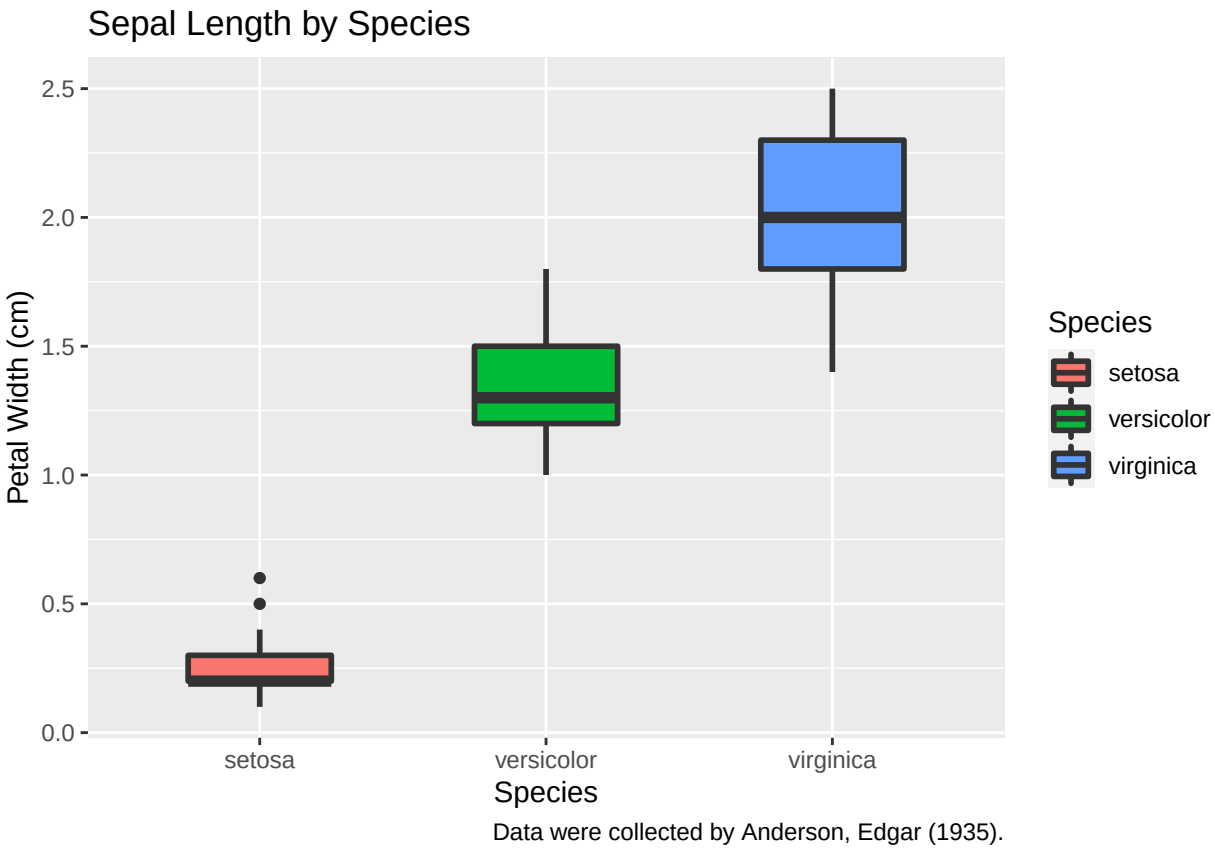

#Boxplot w/ ggplot2 Petal Length

```
ggplot(iris, aes(x=Species, y=Petal.Length, fill=Species)) +
  geom_boxplot(width=0.5,lwd=1)+
  labs(title="Sepal Length by Species", x = "Species", y = "Petal Length (cm)",
       caption="Data were collected by Anderson, Edgar (1935).")
```




#Boxplot w/ ggplot2 Petal Width

```
ggplot(iris, aes(x=Species, y=Petal.Width, fill=Species)) +
  geom_boxplot(width=0.5,lwd=1)+
  labs(title="Sepal Length by Species", x = "Species", y = "Petal Width (cm)",
       caption="Data were collected by Anderson, Edgar (1935).")
```


Sepal Length by Species
Petal Width (cm)
2.5
2.0
1.5
1.0
0.5
0.0
setosa
versicolor
virginica
Species
Species
setosa
versicolor
virginica
Data were collected by Anderson, Edgar (1935).