Exercise 1

Pham Thi Thai - T00727094

2024-01-22

```r
# The Iris Flower Dataset
head(iris)
```

```
##   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 1          5.1         3.5          1.4         0.2  setosa
## 2          4.9         3.0          1.4         0.2  setosa
## 3          4.7         3.2          1.3         0.2  setosa
## 4          4.6         3.1          1.5         0.2  setosa
## 5          5.0         3.6          1.4         0.2  setosa
## 6          5.4         3.9          1.7         0.4  setosa
```

```r
# Numeric summaries
mean <- lapply(iris[, 1:4], mean)
sd <- lapply(iris[, 1:4], sd)
mean(iris$Sepal.Length)
```

```
## [1] 5.843333
```

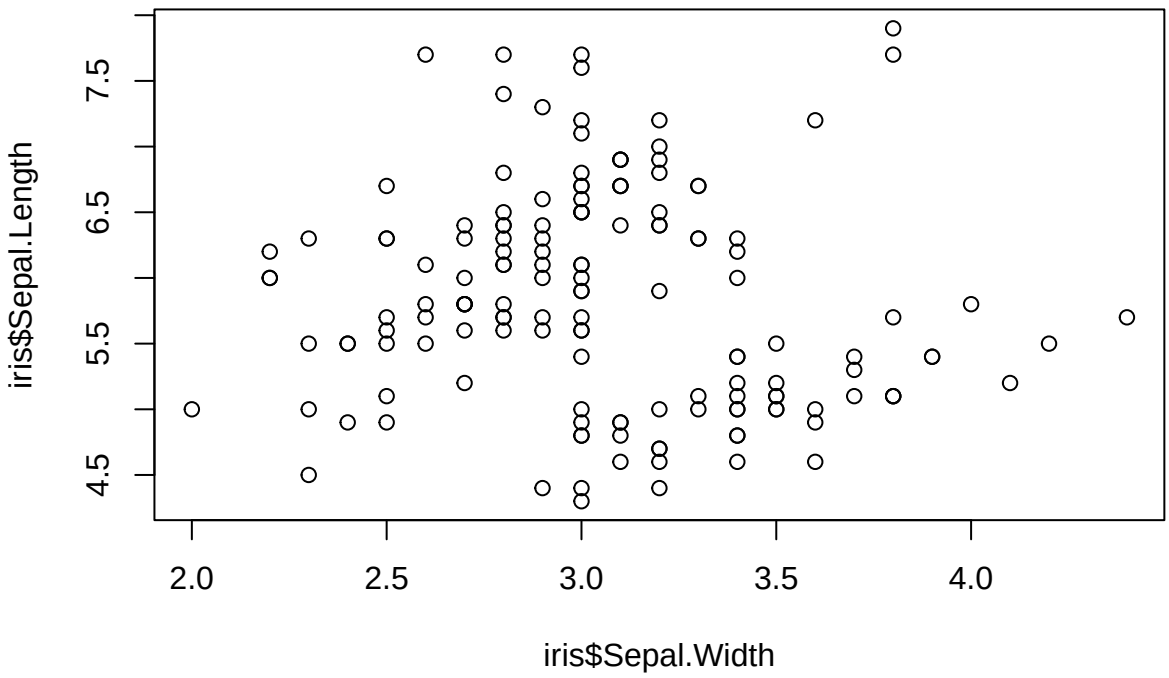
```r
cbind (mean, sd)
```

```
##              mean     sd
## Sepal.Length 5.843333 0.8280661
## Sepal.Width  3.057333 0.4358663
## Petal.Length 3.758    1.765298
## Petal.Width  1.199333 0.7622377
```

```r
summary(iris$Species)
```

```
##     setosa versicolor  virginica
##         50         50         50
```

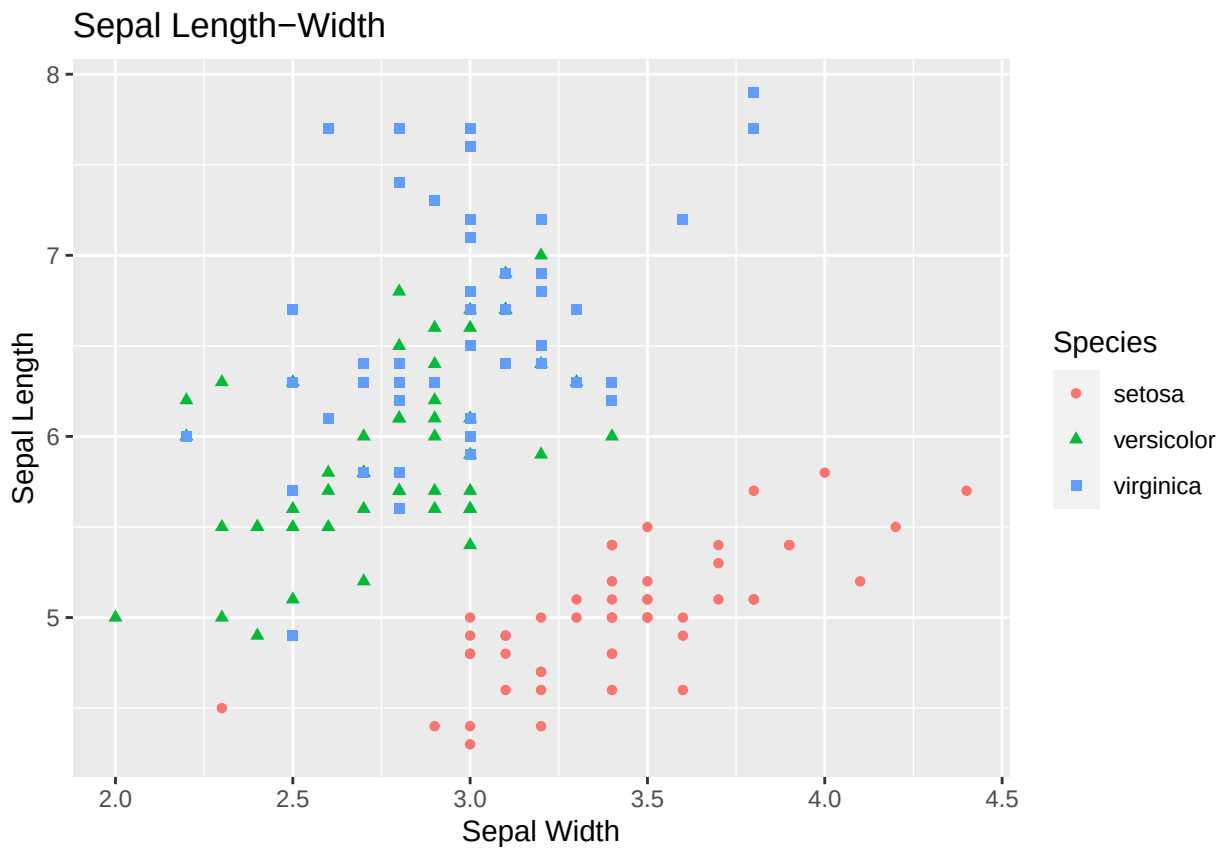
```r
# Simple graph
plot (iris$Sepal.Width, iris$Sepal.Length)
```



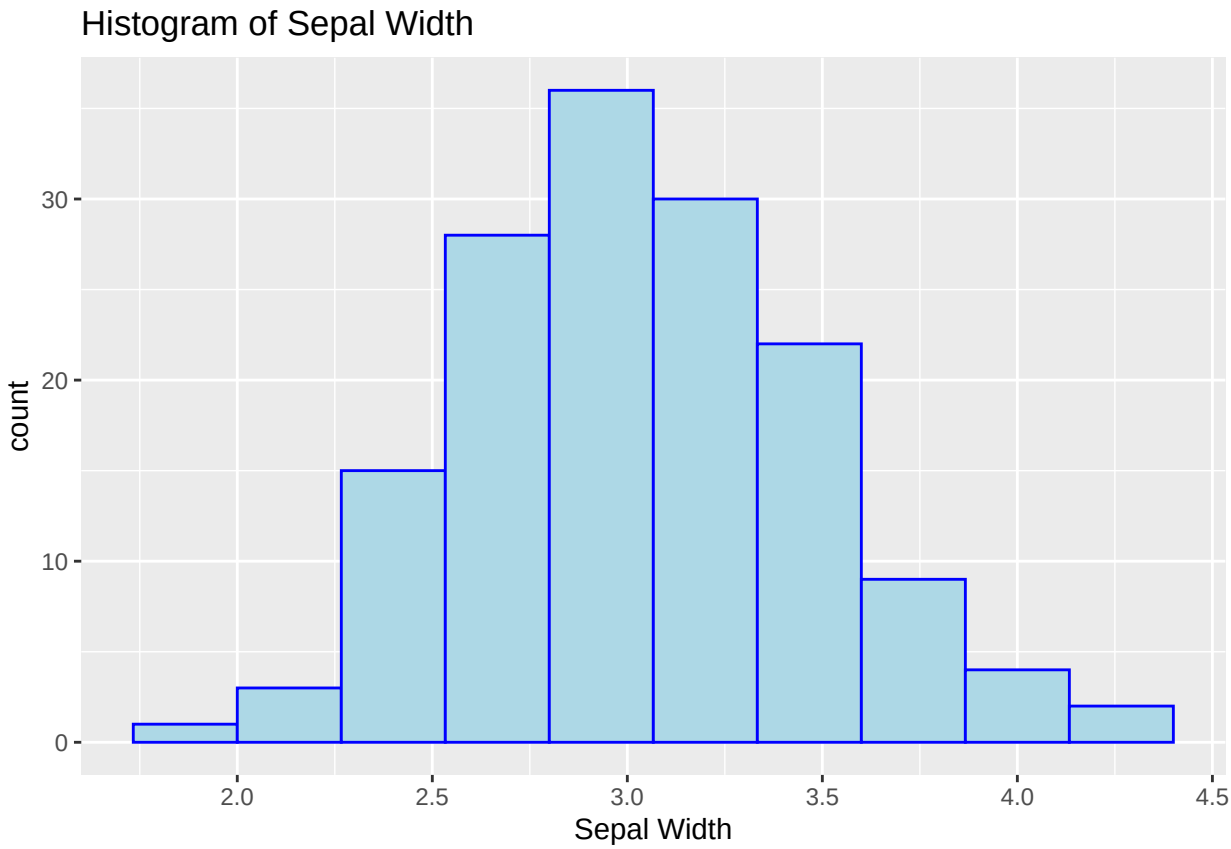

```
# Graph using ggplot2
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.3.2
```

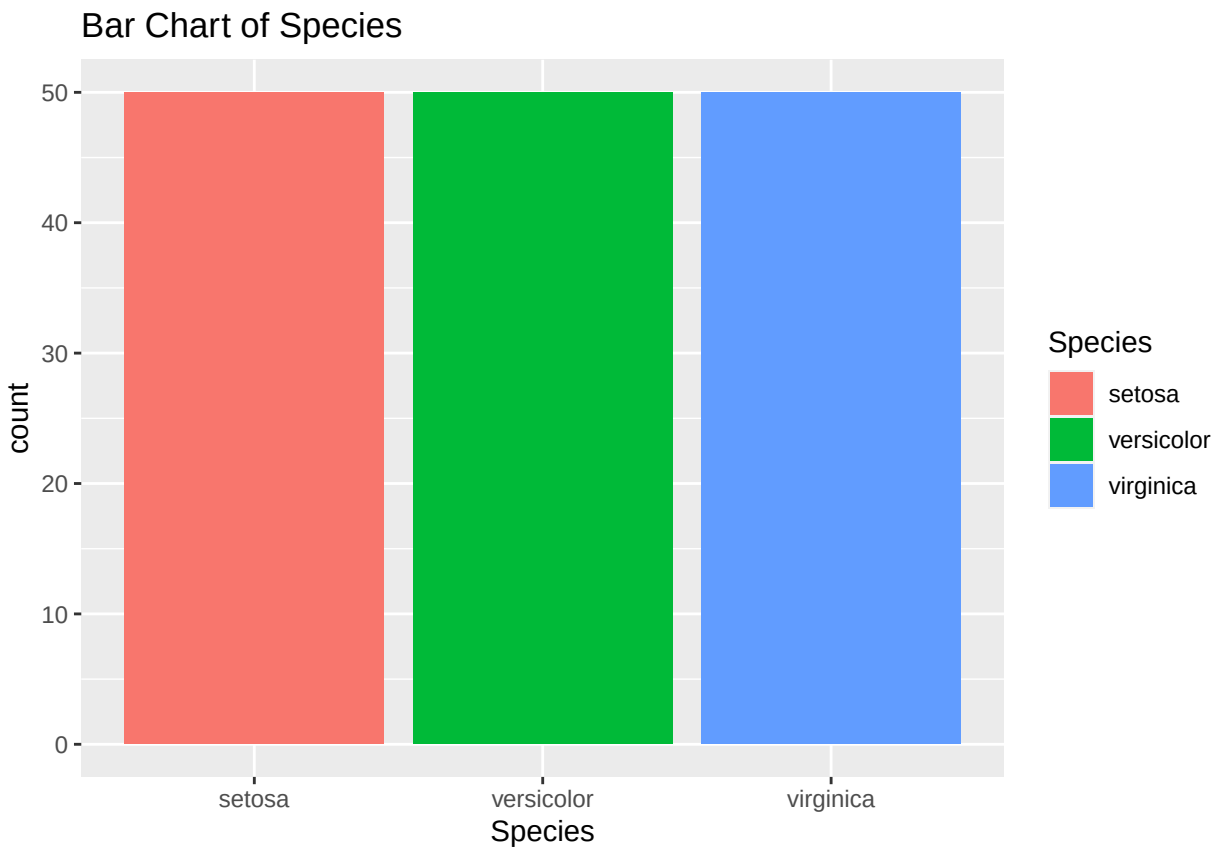
```
ggplot(data=iris) +
geom_point(aes(x = Sepal.Width, y = Sepal.Length,
               color=Species, shape=Species)) +
xlab("Sepal Width") +
ylab("Sepal Length") +
ggtitle("Sepal Length-Width")
```



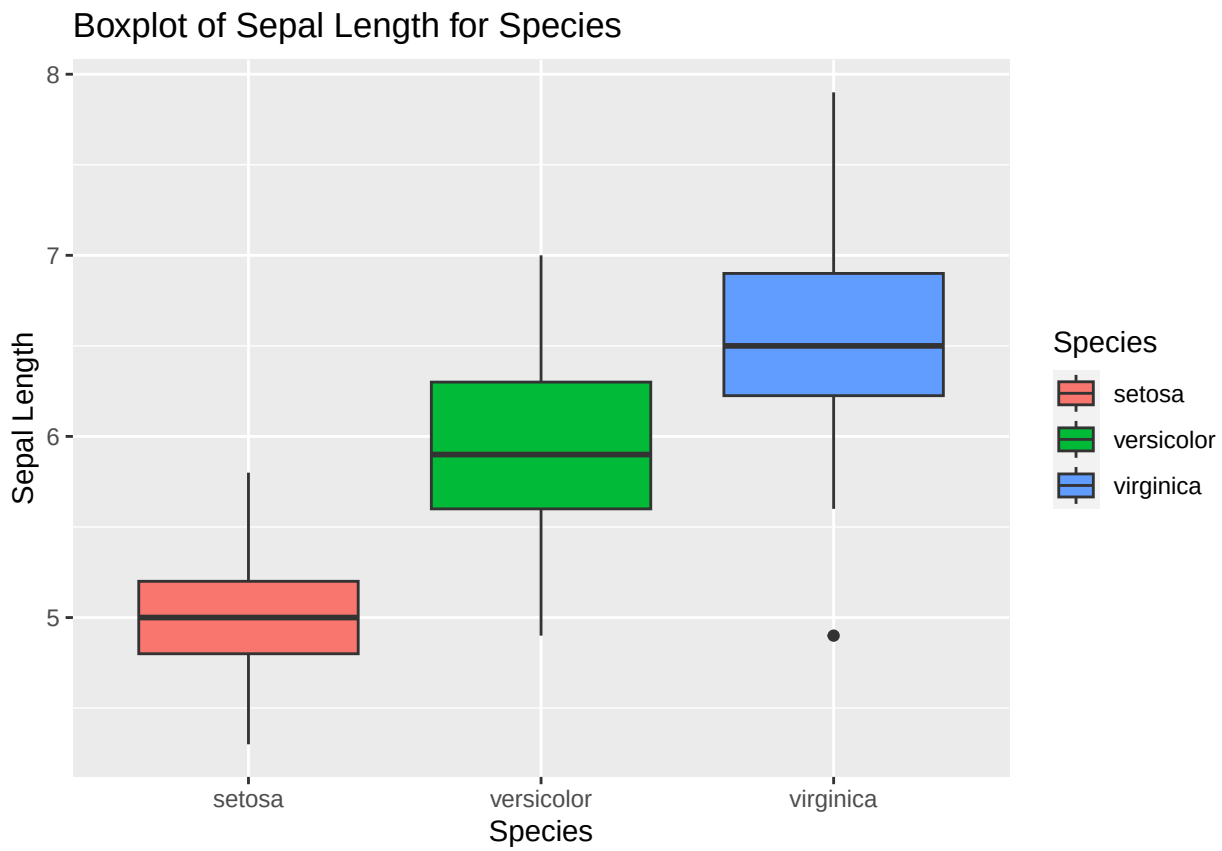

```
# Histogram
library (ggplot2)
ggplot (data = iris) +
  geom_histogram (aes(x = Sepal.Width), color = "blue",
                  fill = "light blue", bins = 10) +
  xlab("Sepal Width") +  ggtitle("Histogram of Sepal Width")
```




```
# Barchart
library (ggplot2)
ggplot (data = iris) +
  geom_bar (aes(x = Species, fill = Species)) +
  xlab("Species") + ggtitle("Bar Chart of Species")
```

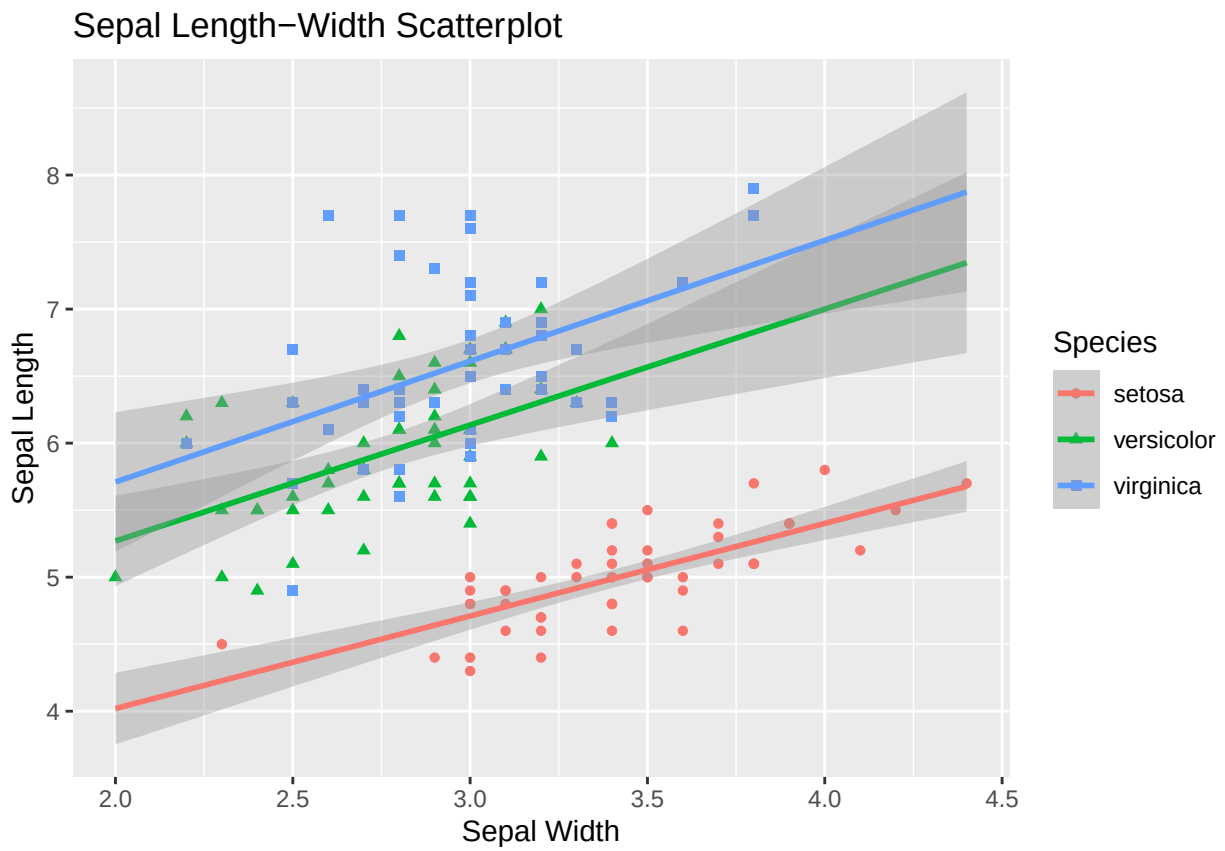



```
# Boxplot
library (ggplot2)
ggplot (data = iris) +
  geom_boxplot (aes(x = Species, y = Sepal.Length, fill = Species)) +
  xlab ("Species") + ylab("Sepal Length") +
  ggtitle("Boxplot of Sepal Length for Species")
```



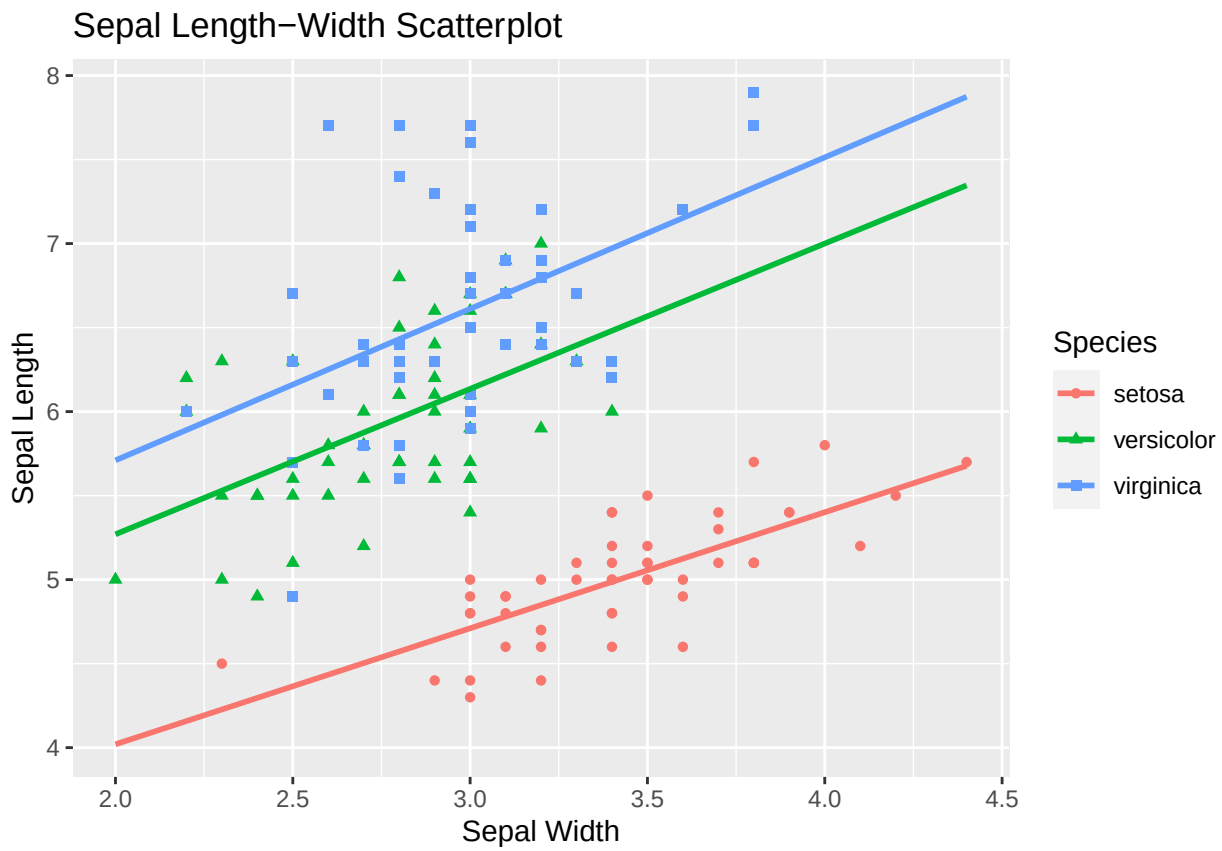

```
# Scatterplots
library(ggplot2)
ggplot(data=iris, aes(x = Sepal.Width, y = Sepal.Length,
                      color=Species, shape=Species)) +
geom_point() +
geom_smooth(method=lm, se=TRUE, fullrange=TRUE) +
xlab("Sepal Width") + ylab("Sepal Length") +
ggtitle("Sepal Length-Width Scatterplot")
```

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



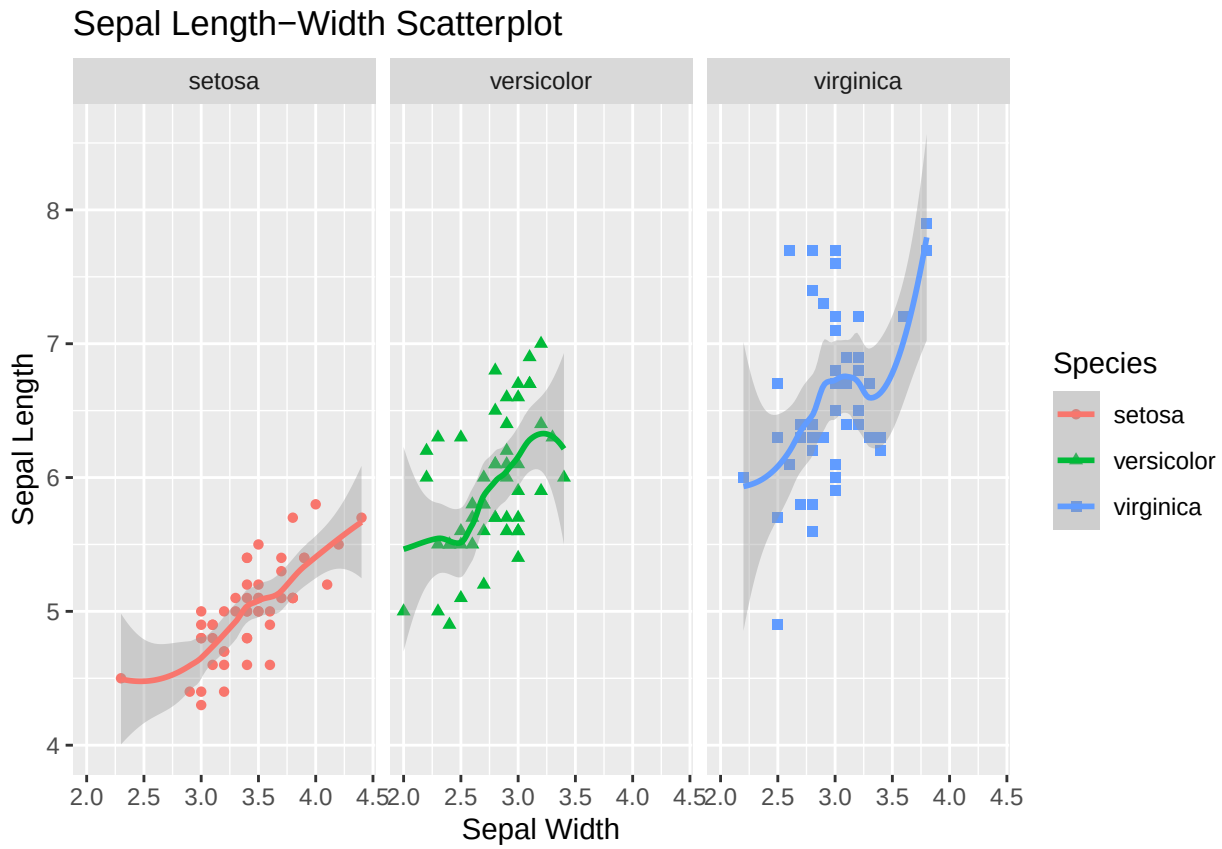

```
# Scatterplots
library(ggplot2)
ggplot(data=iris, aes(x = Sepal.Width, y = Sepal.Length,
                      color=Species, shape=Species)) +
geom_point() +
geom_smooth(method=lm, se=FALSE, fullrange=TRUE) +
xlab("Sepal Width") + ylab("Sepal Length") +
ggtitle("Sepal Length-Width Scatterplot")
```

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




```
library(ggplot2)
ggplot(data=iris, aes(x = Sepal.Width, y = Sepal.Length,
                      color=Species, shape=Species)) +
geom_point() +
xlab("Sepal Width") +
ylab("Sepal Length") +
ggtitle("Sepal Length-Width Scatterplot") +
facet_grid(cols = vars (Species))+
geom_smooth()
```

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




```r
# Date in R
xd <- as.Date("2012-07-27")
weekdays(xd)
```

```
## [1] "Friday"
```

```r
# Install and load package for financial data
pkg_list = c('quantmod', 'zoo', 'tseries', 'PEIP','gridExtra','moments')
# Install packages if needed
for (pkg in pkg_list)
{
  # Try loading the library.
  if ( ! library(pkg, logical.return=TRUE, character.only=TRUE) )
    {
         # If the library cannot be loaded, install it; then load.
        install.packages(pkg)
        library(pkg, character.only=TRUE)
  }
}
```

```
## Warning: package 'quantmod' was built under R version 4.3.2
```

```
## Loading required package: xts
```

```
## Loading required package: zoo
```

```
## Warning: package 'zoo' was built under R version 4.3.2
##
## Attaching package: 'zoo'
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
## Loading required package: TTR
## Warning: package 'TTR' was built under R version 4.3.2
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
## Warning: package 'tseries' was built under R version 4.3.2
## Warning: package 'PEIP' was built under R version 4.3.2
## Warning: package 'gridExtra' was built under R version 4.3.2
```

```
# Stock Prices
start.date = '2019-12-1' # starting date of stock
end.date = '2022-12-1' # ending date of stock
# Download the selected stocks from Yahoo finance
getSymbols(c('AAPL','MSFT','IBM', "GOOG","AMZN","TSLA"),
              src = "yahoo", from = start.date, to = end.date)
```

```
## [1] "AAPL" "MSFT" "IBM"  "GOOG" "AMZN" "TSLA"
```

```
stocks <- merge(AAPL = AAPL[, "AAPL.Adjusted"], MSFT = MSFT[, "MSFT.Adjusted"],
                IBM = IBM[, "IBM.Adjusted"], GOOG = GOOG[, "GOOG.Adjusted"],
                AMZN = AMZN[, "AMZN.Adjusted"], TSLA = TSLA[, "TSLA.Adjusted"])

names(stocks )<-c("Price.APPLE","Price.MSFT","Price.IBM","Price.GOOG",
                  "Price.AMZN","Price.TSLA")
```

```
# Data
options(width = 70)
head(stocks)
```

```
##            Price.APPLE Price.MSFT Price.IBM Price.GOOG Price.AMZN
## 2019-12-02    64.33828   144.1121  104.3364    64.4960    89.0800
## 2019-12-03    63.19112   143.8808  103.7162    64.7640    88.4980
## 2019-12-04    63.74888   144.4012  103.6142    66.0270    88.0345
## 2019-12-05    64.68410   144.4782  103.5514    66.4065    87.0240
## 2019-12-06    65.93359   146.2321  104.5797    67.0310    87.5800
## 2019-12-09    65.01053   145.8563  105.1292    67.1780    87.4755
##            Price.TSLA
## 2019-12-02   22.32467
## 2019-12-03   22.41333
## 2019-12-04   22.20200
## 2019-12-05   22.02467
## 2019-12-06   22.39267
## 2019-12-09   22.63533
```

```
tail(stocks)
```

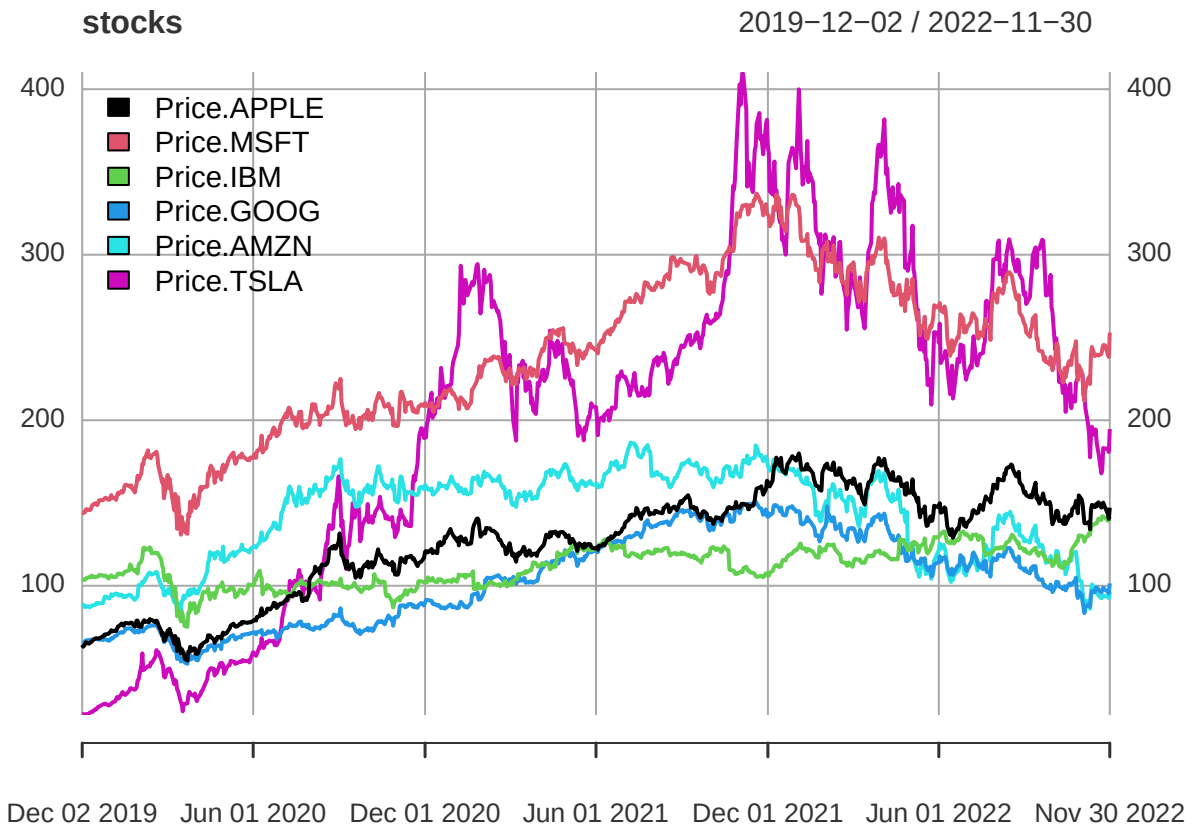
```
##            Price.APPLE Price.MSFT Price.IBM Price.GOOG Price.AMZN
```

```
## 2022-11-22    149.3452   242.8764  142.0453      97.33      93.20
## 2022-11-23    150.2303   245.4040  141.7119      98.82      94.13
## 2022-11-25    147.2867   245.3148  141.3498      97.60      93.41
## 2022-11-28    143.4184   239.6352  139.2635      96.25      93.95
## 2022-11-29    140.3853   238.2178  139.5588      95.44      92.42
## 2022-11-30    147.2072   252.8976  141.8548     101.45      96.54
##            Price.TSLA
## 2022-11-22     169.91
## 2022-11-23     183.20
## 2022-11-25     182.86
## 2022-11-28     182.92
## 2022-11-29     180.83
## 2022-11-30     194.70
```

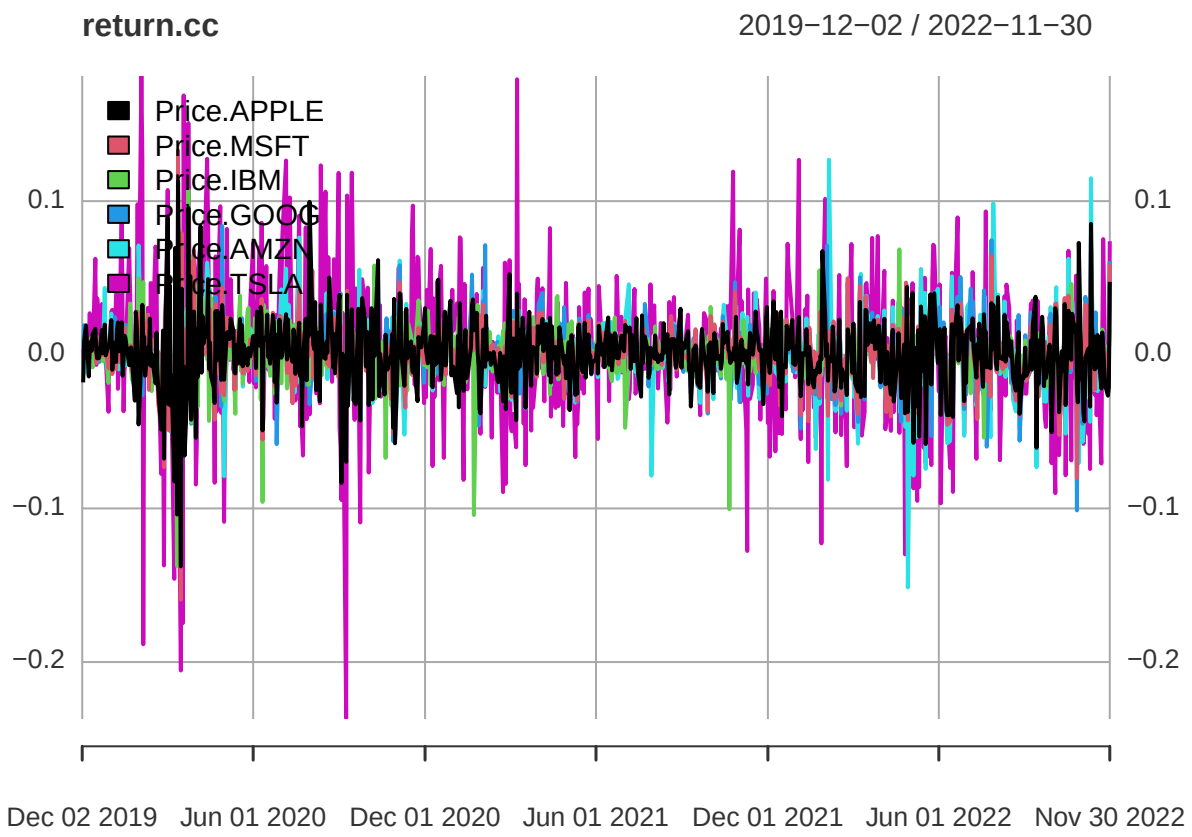
```
nrow(stocks)
```

```
## [1] 756
```

```
# Plot Stock Prices
plot(stocks, legend.loc=1)
```




```
# Log Returns
return.cc = diff(log(stocks)) #calculate log returns
plot(return.cc, legend.loc=1)
```


return.cc
2019−12−02 / 2022−11−30
Price.APPLE
Price.MSFT
Price.IBM
Price.GOOG
Price.AMZN
Price.TSLA
0.1
0.0
−0.1
−0.2
Dec 02 2019
Jun 01 2020
Dec 01 2020
Jun 01 2021
Dec 01 2021
Jun 01 2022
Nov 30 2022