Exercises

June 12, 2019

```
[1]: # load ggplot module
     library(ggplot2)
```

```
[2]: # mtcars dataset
     head(mtcars, n=3)
```

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.0	6	160	110	3.90	2.620	16.46	0	1	4	4
Mazda RX4 Wag	21.0	6	160	110	3.90	2.875	17.02	0	1	4	4
Datsun 710	22.8	4	108	93	3.85	2.320	18.61	1	1	4	1

```
[3]: # diamonds dataset
     tail(diamonds, n=3)
```

carat	cut	color	clarity	depth	table	price	x	y	z
0.70	Very Good	D	SI1	62.8	60	2757	5.66	5.68	3.56
0.86	Premium	H	SI2	61.0	58	2757	6.15	6.12	3.74
0.75	Ideal	D	SI2	62.2	55	2757	5.83	5.87	3.64

```
[4]: # all posible buildin datasets
     head(data()$results, n =3)
```

Package	LibPath	Item	Title
ggplot2	/anaconda3/envs/AlgebraR/lib/R/library	diamonds	Prices of 50,000 round cut diam
ggplot2	/anaconda3/envs/AlgebraR/lib/R/library	economics	US economic time series
ggplot2	/anaconda3/envs/AlgebraR/lib/R/library	economics_long	US economic time series

```
[5]: iris.dataset.csv <- read.csv("data.csv", header = TRUE)
     iris.dataset.txt <- read.table("data.txt", header = TRUE)
```

```
[6]: head(iris.dataset.csv, n = 3)
```

sepal_length	sepal_width	petal_length	petal_width	species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
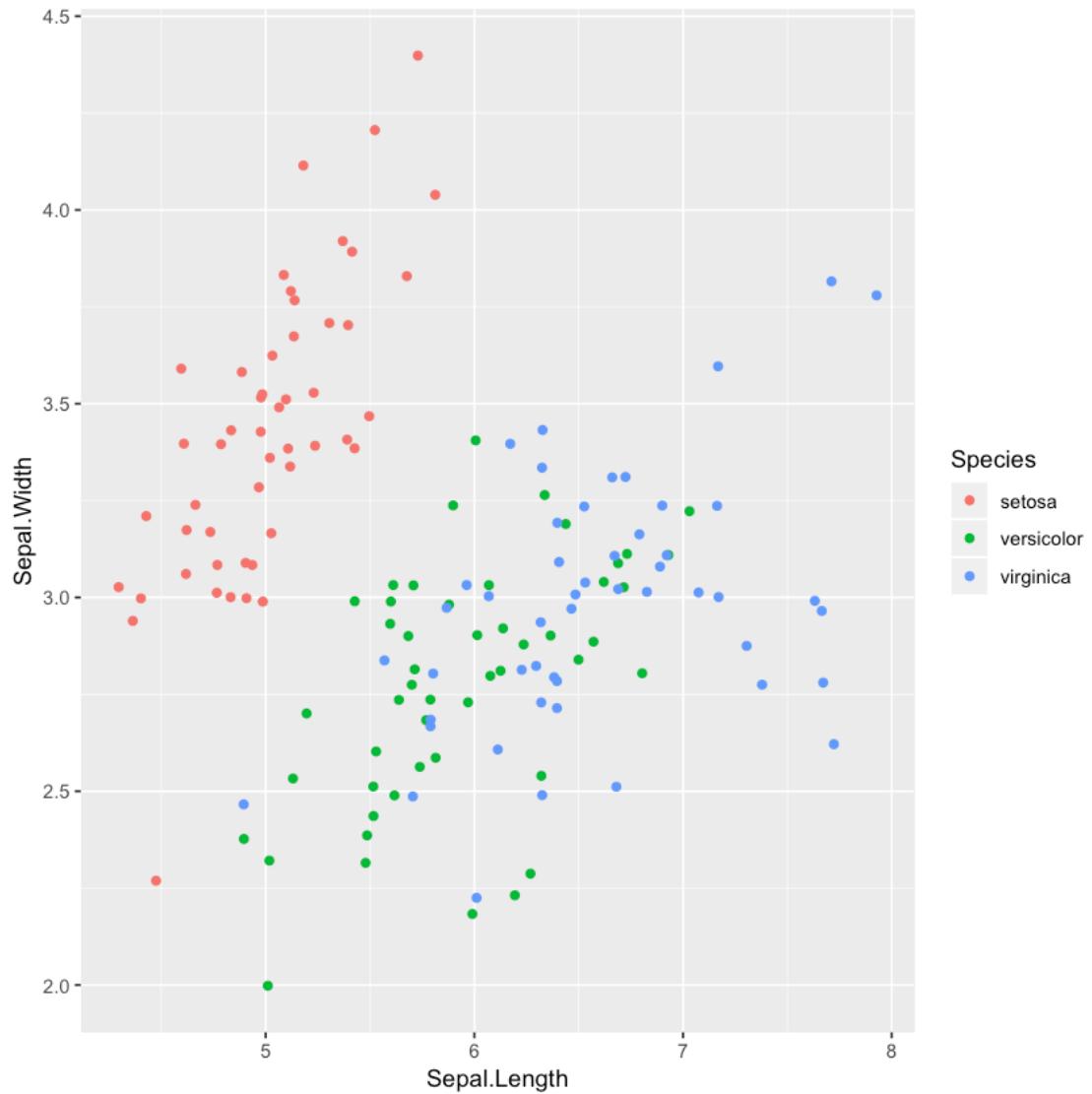
```
[7]: head(iris.dataset.txt, n = 3)
```

sepal_length	sepal_width	petal_length	petal_width	species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa

```
[8]: head(iris, n = 3)
```

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa

```
[9]: ggplot(iris, aes( x=Sepal.Length, y=Sepal.Width, color=Species)) + geom_jitter()
```



```
[11]: library(gridExtra)

# Sepal length
HisSl <- ggplot(data=iris, aes(x=Sepal.Length))+
  geom_histogram(binwidth=0.2, color = "black", aes(fill=Species)) +
  xlab("Sepal Length (cm)") +
  ylab("Frequency") +
```

```r
   theme(legend.position="left")+
   ggtitle("Histogram of Sepal Length")+
   geom_vline(data=iris, aes(xintercept = mean(Sepal.Length)),␣
 ↪linetype="dashed",color="grey")

# Sepal Width
HisSw <- ggplot(data=iris, aes(x=Sepal.Width)) +
    geom_histogram(binwidth = 0.2, color="blue", aes(fill=Species)) +
    xlab("Sepal Width") +
    ylab("Frequency") +
    theme(legend.position="none", text=element_text(size=5)) +
    ggtitle("Histiogram of Sepal Width")+
    geom_vline(data=iris, aes(xintercept = mean(Sepal.Width)),␣
 ↪linetype="dashed", color="black")

#combine both in one
grid.arrange(HisSl,
             HisSw,
             nrow=1, widths=c(100,70))
```
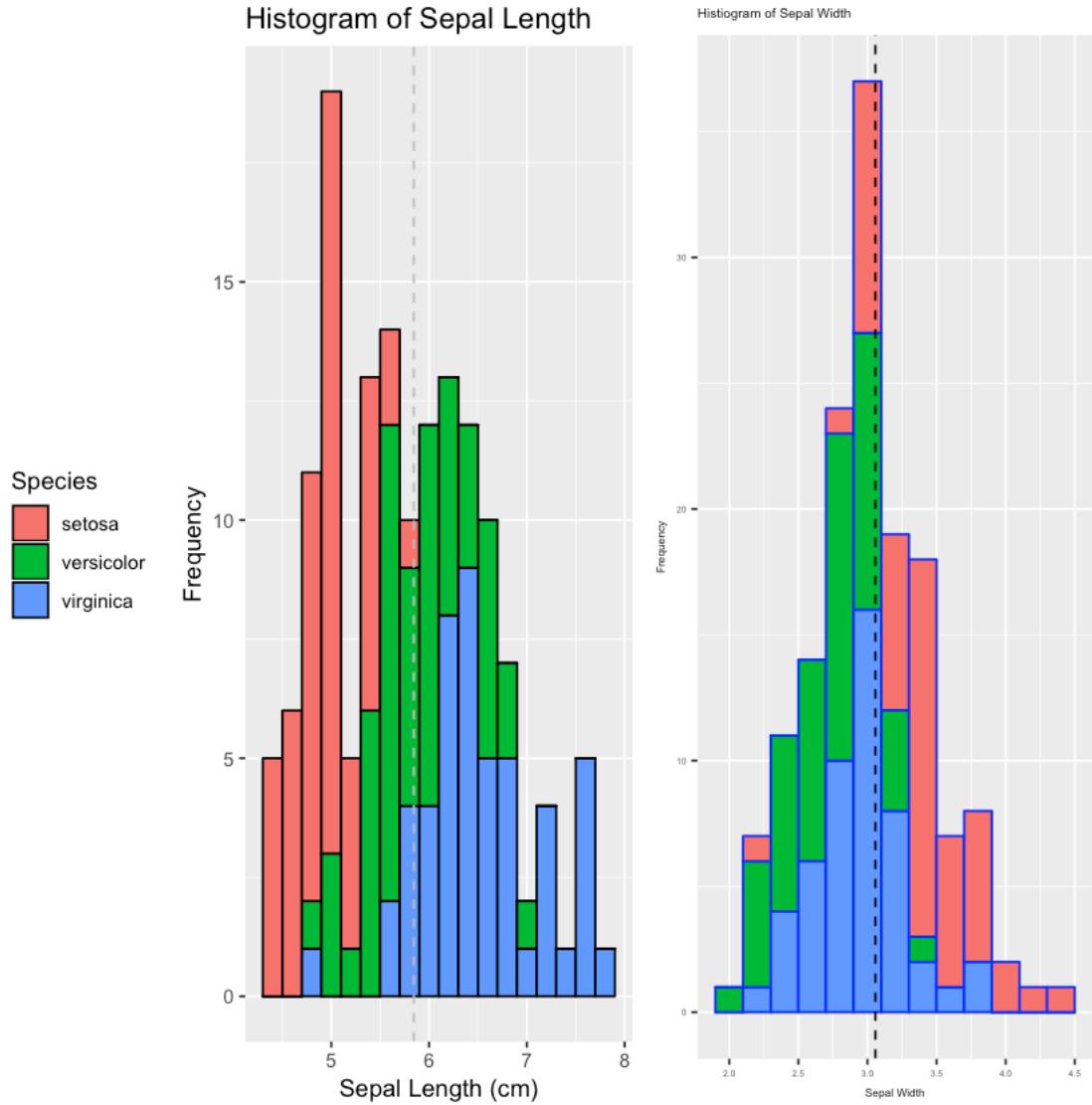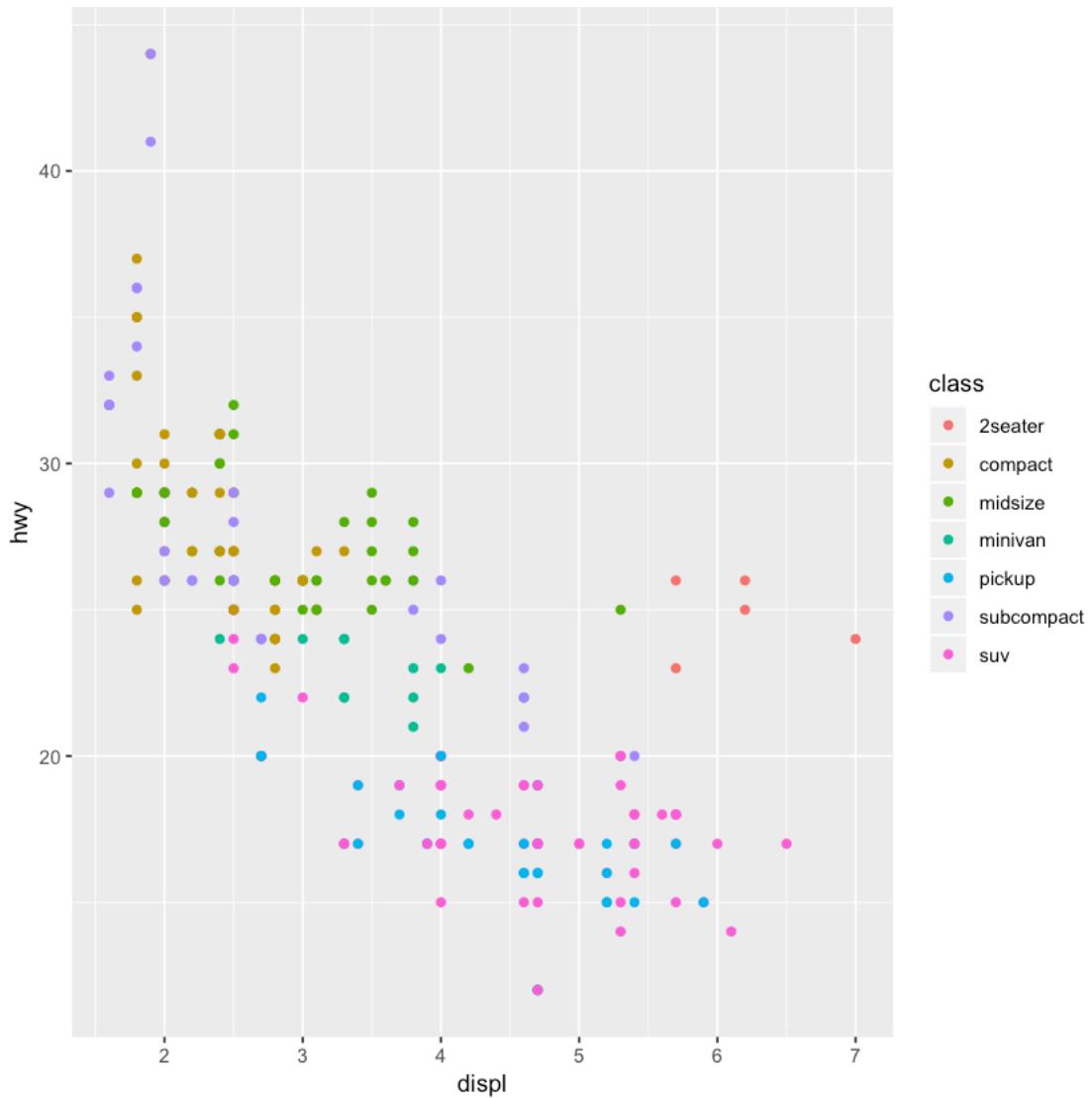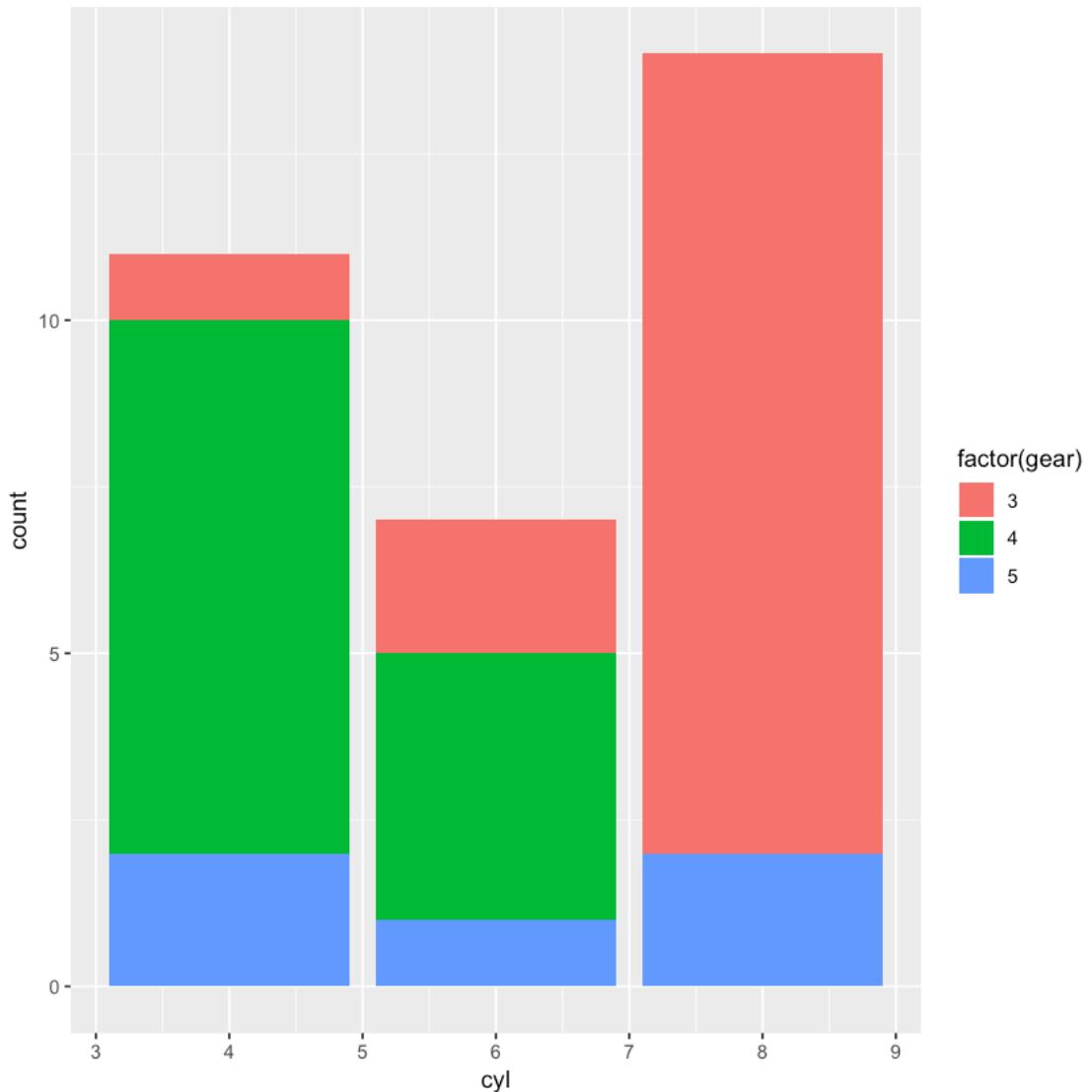
Histogram of Sepal Length

Histogram of Sepal Width
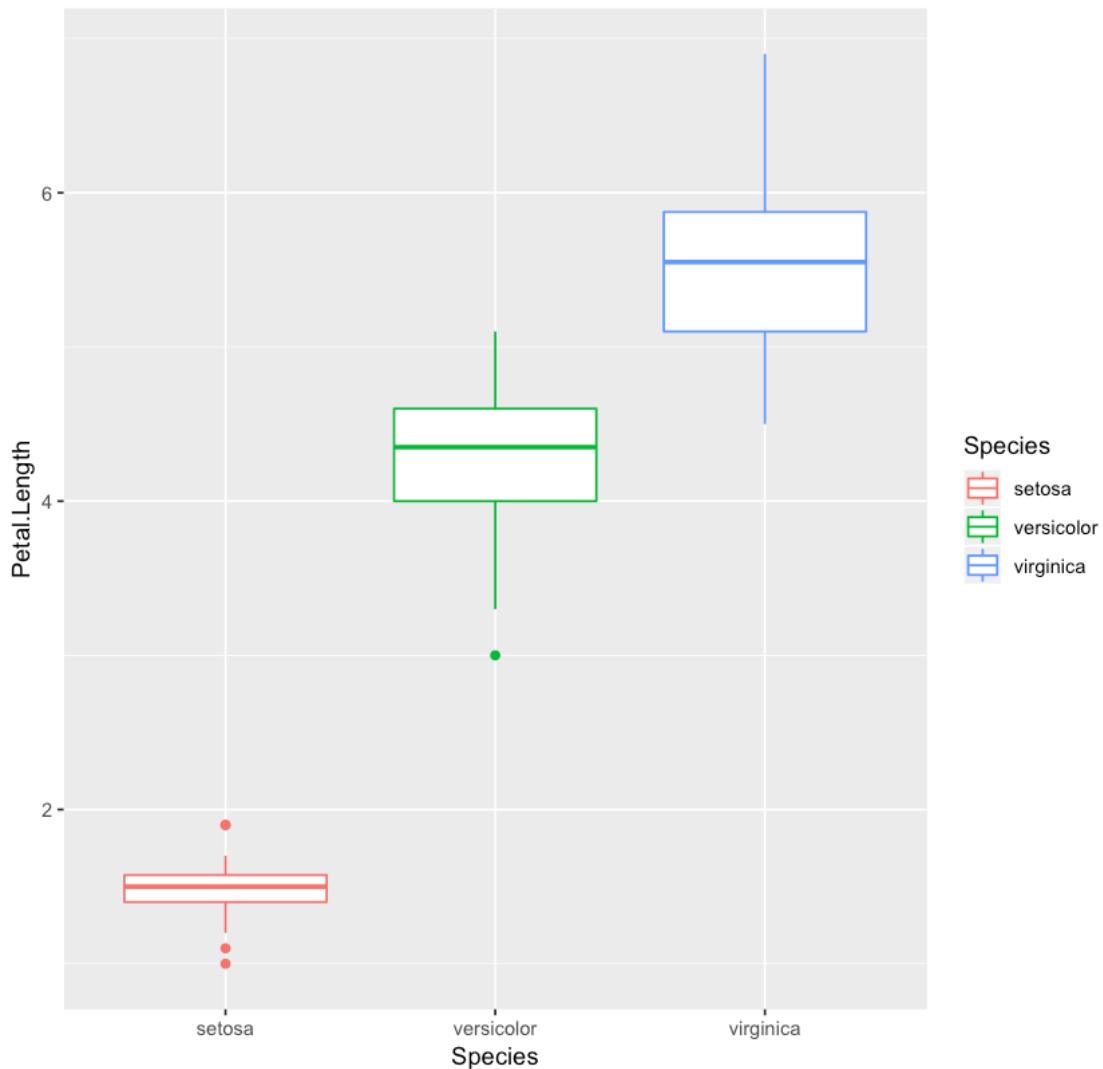
```
[12]: ggplot(data = mpg, aes(x = displ, y = hwy, color=class)) + geom_point()
      c <- ggplot(data=mtcars, aes(cyl, fill = factor(gear)))
      c + geom_bar()
```

```
[13]: ggplot(iris, aes(x=Species, y=Petal.Length, colour=Species)) +
      geom_boxplot() +
      labs(title = "Iris Petal Length Box Plot", x = "Species")
```

Iris Petal Length Box Plot


```
[14]: head(economics, n = 5)
      head(presidential, n = 5)
      b <- ggplot(economics, aes(date, unemploy)) + geom_line()
      yrng <- range(economics$unemploy)
      b <- b + geom_rect(aes(NULL, NULL, xmin = start, xmax = end, fill = party),␣
       →ymin = yrng[1], ymax = yrng[2], data = presidential) +␣
       →scale_fill_manual(values = alpha(c("blue", "red"), .5))
      b
```

date	pce	pop	psavert	uempmed	unemploy
1967-07-01	507.4	198712	12.5	4.5	2944
1967-08-01	510.5	198911	12.5	4.7	2945
1967-09-01	516.3	199113	11.7	4.6	2958
1967-10-01	512.9	199311	12.5	4.9	3143
1967-11-01	518.1	199498	12.5	4.7	3066

name	start	end	party
Eisenhower	1953-01-20	1961-01-20	Republican
Kennedy	1961-01-20	1963-11-22	Democratic
Johnson	1963-11-22	1969-01-20	Democratic
Nixon	1969-01-20	1974-08-09	Republican
Ford	1974-08-09	1977-01-20	Republican