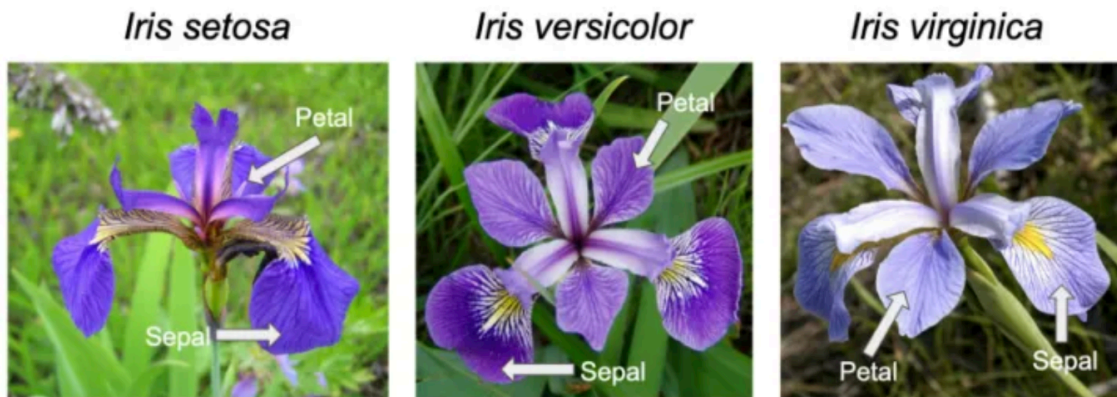# Iris

Figure 1: iris-image

Iris Dataset

Iris dataset is found 'natively' in BaseR. It gets loaded when you start R.

```
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
```

```
str(iris)
```

```
'data.frame':    150 obs. of  5 variables:
 $ Sepal.Length: num  5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
 $ Sepal.Width : num  3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
 $ Petal.Length: num  1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
 $ Petal.Width : num  0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
 $ Species     : Factor w/ 3 levels "setosa","versicolor",..: 1 1 1 1 1 1 1 1 1 1 ...
```

```
library(tidyverse)
```

Plot Sepal and Petal measurements for the 3 species

```
s1 <- ggplot(iris, aes(x=Sepal.Length, y=Sepal.Width,color=Species))+
        geom_point(show.legend=FALSE)+
        xlim(0,10)+
        ylim(0,5.0)
p1 <- ggplot(iris, aes(x=Petal.Length, y=Petal.Width,color=Species))+
        geom_point()

library(patchwork)
  s1 + p1 +
    theme(legend.position="bottom")
```
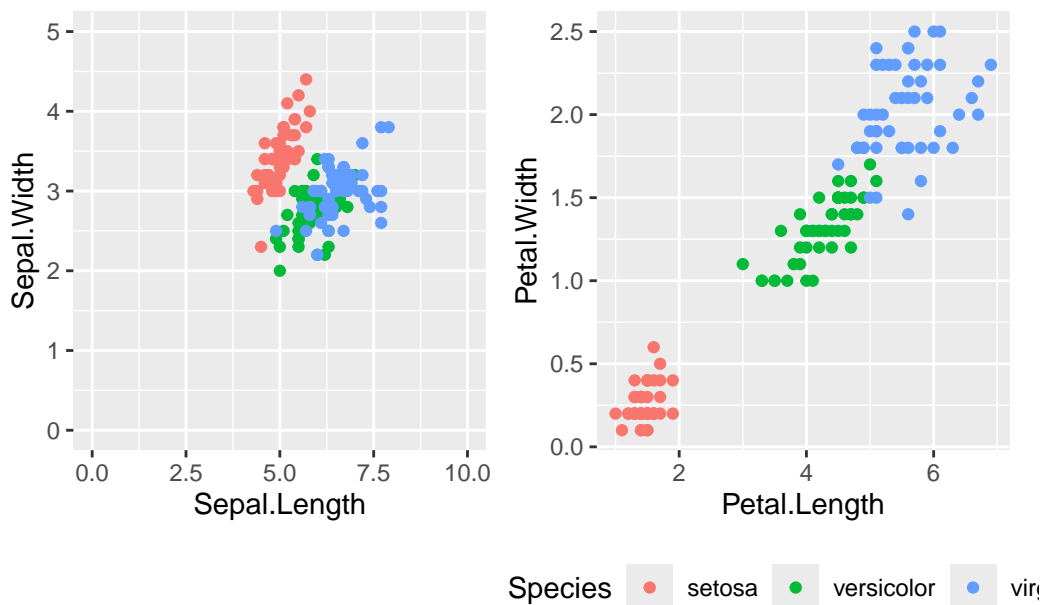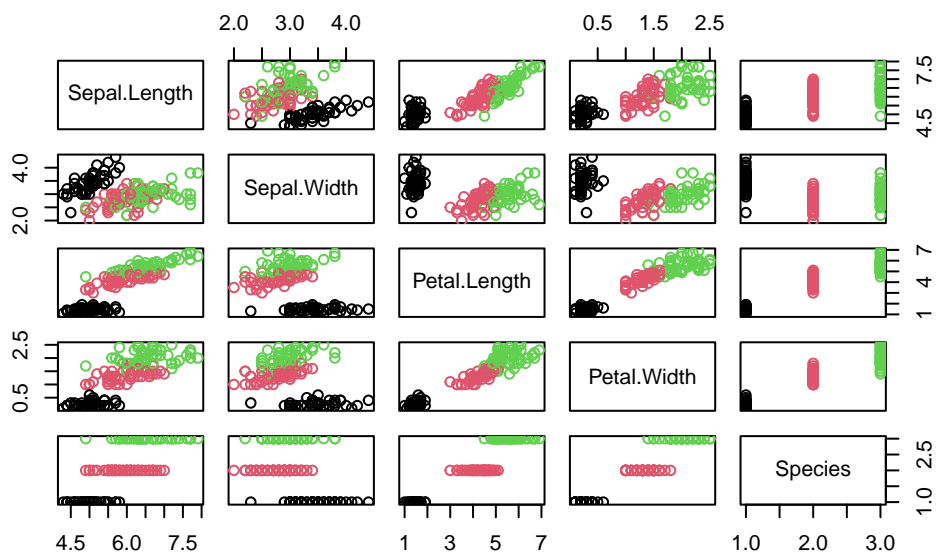
Boxplot of the `Petal.Length` and `Petal.Width` across `Species`

```
pl <- ggplot(iris, aes(x=Species,y=Petal.Length, color=Species))+
        geom_boxplot(show.legend=FALSE)+
        theme_bw()+
        ylim(0,8)
pw <- ggplot(iris, aes(x=Species,y=Petal.Width, color=Species))+
        geom_boxplot(show.legend=FALSE)+
        theme_bw()+
        ylim(0,8)

pl+pw
```
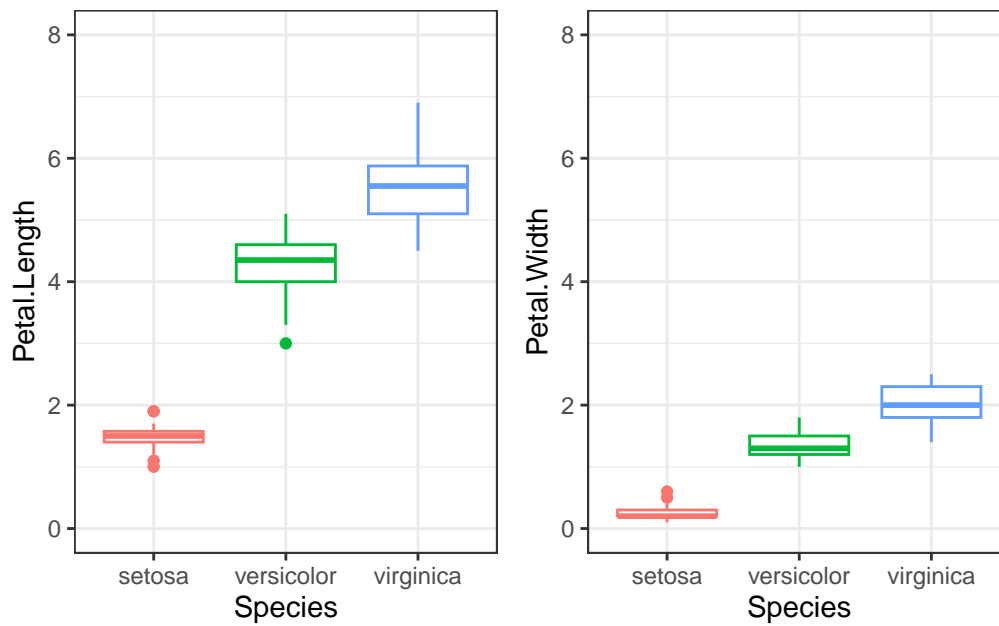
Pairwise Comparisons

```r
iris <- iris[sample(1:nrow(iris)),]
plot(iris, col = iris$Species)
```



```r
summary(iris)
```

```
  Sepal.Length    Sepal.Width     Petal.Length    Petal.Width
 Min.   :4.300   Min.   :2.000   Min.   :1.000   Min.   :0.100
 1st Qu.:5.100   1st Qu.:2.800   1st Qu.:1.600   1st Qu.:0.300
 Median :5.800   Median :3.000   Median :4.350   Median :1.300
 Mean   :5.843   Mean   :3.057   Mean   :3.758   Mean   :1.199
 3rd Qu.:6.400   3rd Qu.:3.300   3rd Qu.:5.100   3rd Qu.:1.800
 Max.   :7.900   Max.   :4.400   Max.   :6.900   Max.   :2.500
       Species
 setosa    :50
 versicolor:50
 virginica :50
```

ggRidges for density plots

```r
library(ggridges)
```

```r
ggplot(iris, aes(x=Petal.Length, y=Species,fill=Species))+
  geom_density_ridges()+
  xlim(0,8)+
  theme_bw()
```
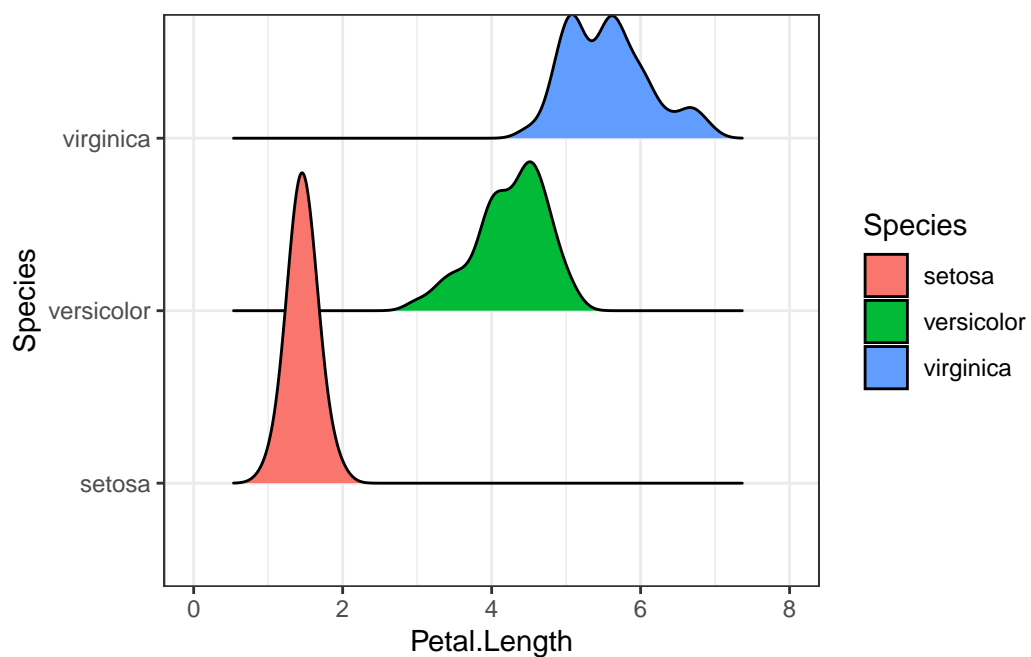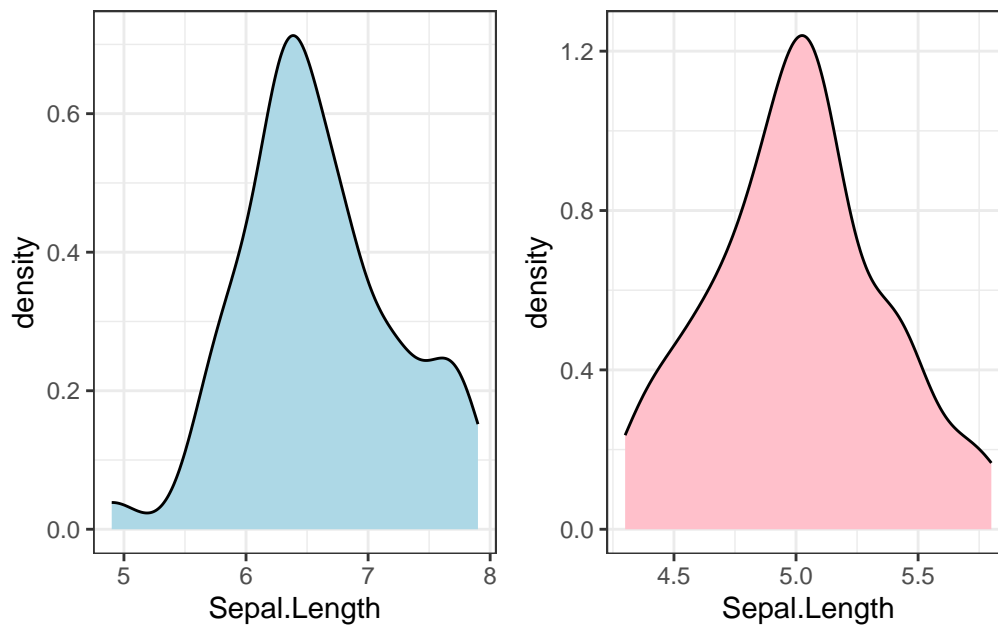

> vline for median
>
> ```r
> geom_density_ridges(quantile_lines=TRUE,quantile=2)
> ```

_____

_____

_____

or just use `geom_density()`

```
irisvi <- filter(iris, Species == "virginica")
irisse <- filter(iris, Species == 'setosa')

dvi <- ggplot(irisvi,aes(Sepal.Length))+
  geom_density(fill="lightblue")+
  theme_bw()
dse <- ggplot(irisse,aes(Sepal.Length))+
  geom_density(fill='pink')+
  theme_bw()
dvi+dse
```


_____

_____

_____

_____

Regression for `Species=Versicolor`

```
irisV <- filter(iris, Species=="versicolor")
head(irisV)
```

```
  Sepal.Length Sepal.Width Petal.Length Petal.Width    Species
1          6.6         2.9          4.6         1.3 versicolor
2          5.5         2.6          4.4         1.2 versicolor
3          6.2         2.9          4.3         1.3 versicolor
4          6.4         2.9          4.3         1.3 versicolor
5          6.7         3.1          4.4         1.4 versicolor
6          6.6         3.0          4.4         1.4 versicolor
```

RegressionPlot for `versicolor`

```
ggplot(irisV, aes(x=Sepal.Width, y=Sepal.Length))+
  geom_point(color="brown")+
  geom_smooth(method="lm",se=FALSE)+
  ggtitle("Simple Regression")
```

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