Plotting continous data in R with ggplot2

Nathan Brouwer

2017-09-11

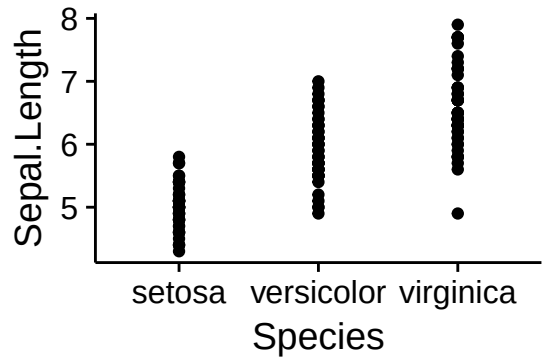
Introduction to ggplot using qplot

- We're going to plot some data using the qplot command
- We'll need to have 2 packages loaded
- ggplot2, which has the function qplot()
- cowplot, which provides some nice defaults
- We'll use the iris dataset that comes with R

A basic plot in ggplot using qplot()

- Unless told otherwise, qplot plots dots

```
qplot(y = Sepal.Length,
      x = Species,
        data = iris)
```



Box plot with labels

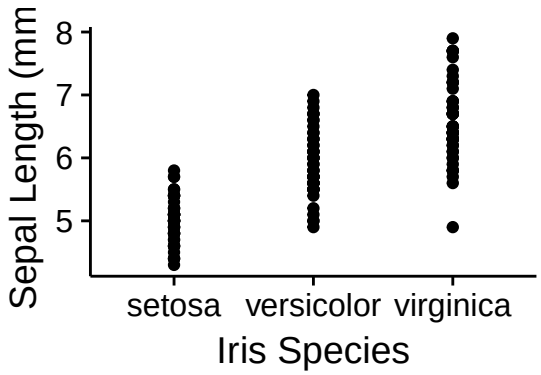
- R will usually generate labels for the x and y axes based on the command.

- These cab be changed by adding another command after the qplot() command
- Add The command + **xlab("...")** sets the labels for the x-axis, + **ylab("...")** for the y axis.
- Text for the labels goes in quotes (ie, "Iris species").
- THe use of the "+" is different than for most other R packages
- Forgetting the quotes will cause the code to fail.

- Note that units (mm) are included for the y axis.

```
qplot(y = Sepal.Length,
      x = Species,
        data = iris) +
```

```
  xlab("Iris Species") +
  ylab ("Sepal Length (mm)" )
```

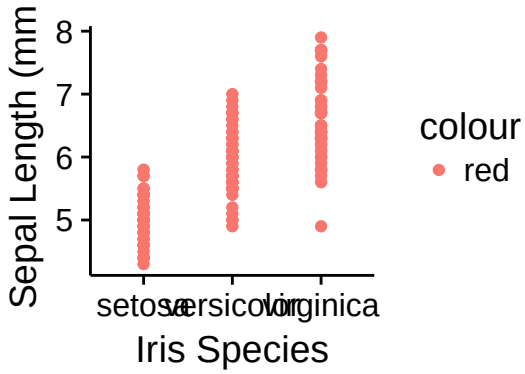


Changing colors in R plots

Changing colors in R plots part 1

- If we wanted we could change the color of the dots using the argument **"col ="**. This code can be used to change the color of most types of plots in R.
- This doesn't increase the information content of the figure but maybe makes it nicer to look at.

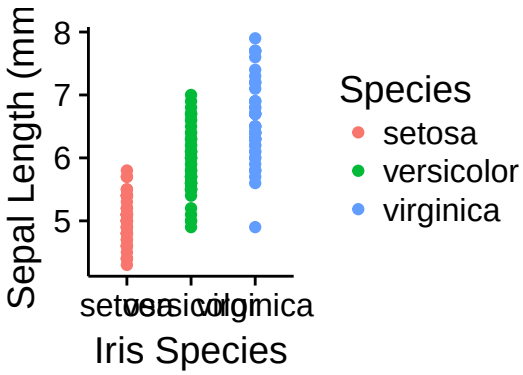
```
qplot(y = Sepal.Length,
      x = Species,
        data = iris,
      color = "red") +
  xlab("Iris Species") +
  ylab ("Sepal Length (mm)" )
```



Changing colors in R plots part 2

```
#dopt w/color changes
qplot(y = Sepal.Length,
      x = Species,
        data = iris,
      color = Species) +
```

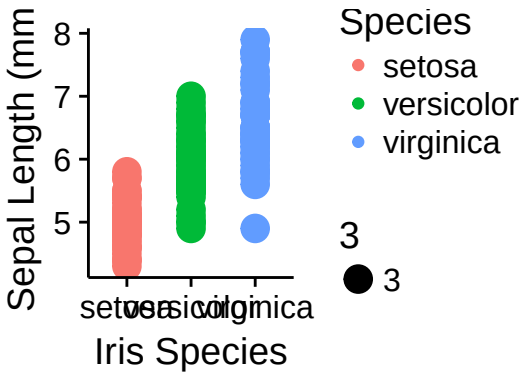
```
  xlab("Iris Species") +
  ylab ("Sepal Length (mm)")
```



Tweaking plots: changing the point size

Run the code below, Can you see what changed?

```
#dopt w/color changes
qplot(y = Sepal.Length,
      x = Species,
        data = iris,
      color = Species,
      size = 3) +
  xlab("Iris Species") +
  ylab ("Sepal Length (mm)")
```

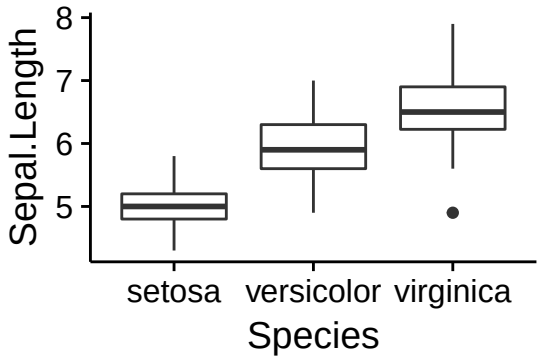


Boxplot with qplot

Basic boxplot with qplot

- note use of arguement "**geom** = ..."

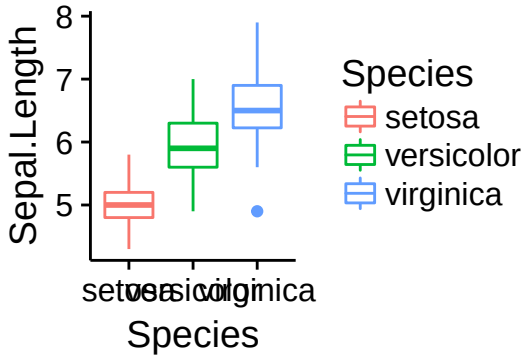
```
qplot(y = Sepal.Length,
      x = Species,
        data = iris,
      geom = "boxplot")
```



Basic boxplot with colors

- same as before, using "color ="

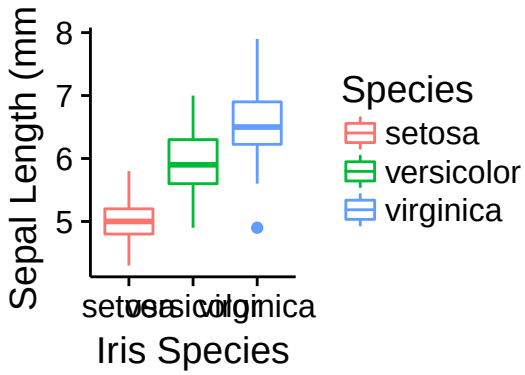
```
#dopt w/color changes
qplot(y = Sepal.Length,
      x = Species,
        data = iris,
      geom = "boxplot",
      color = Species)
```



Basic boxplot lables

- now use + xlab() and + ylab()

```
qplot(y = Sepal.Length,
      x = Species,
        data = iris,
      geom = "boxplot",
      color = Species) +
  xlab("Iris Species") +
  ylab ("Sepal Length (mm)")
```



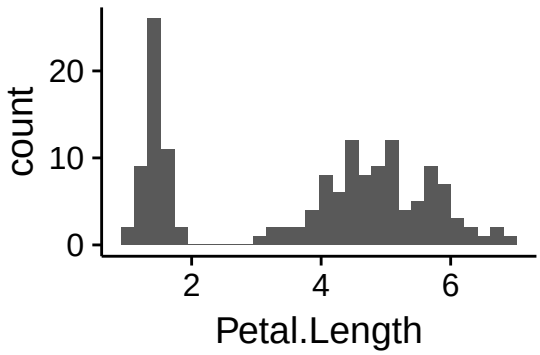
Histograms using qplot

- made with geom = "histogram" arguement
- very very easy to make in R with ggplot
- very very very hard to make in Excel
- You should make them all the time for you data!

Histograms of iris data

- This code makes a histogram of one of the iris species' Petal.Length.
- Note that you don't have "y =" or "x =" for a histogram!

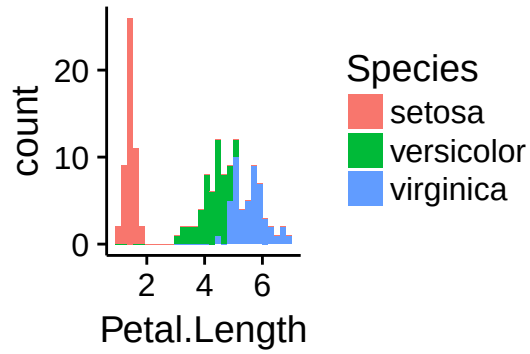
```
qplot(Petal.Length,
      data = iris)
```



Histogram with colors

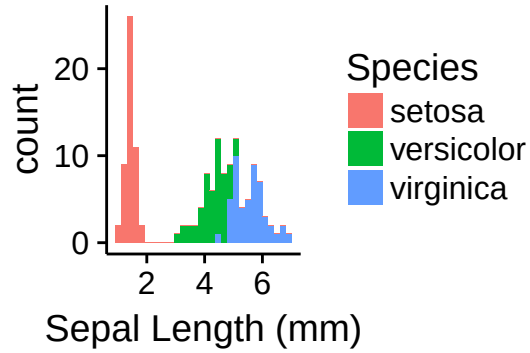
What does this show?

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



Histogram with axes labels

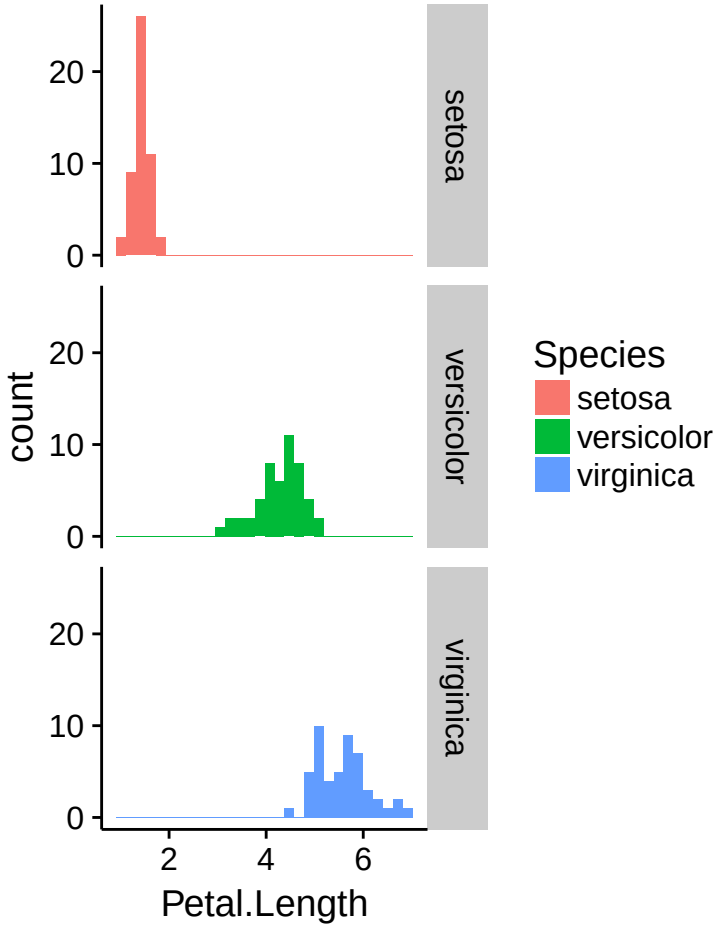
```
qplot(Petal.Length,
      data = iris,
      fill = Species) +
  xlab ("Sepal Length (mm)")
```



Multiple histograms: "Facets"

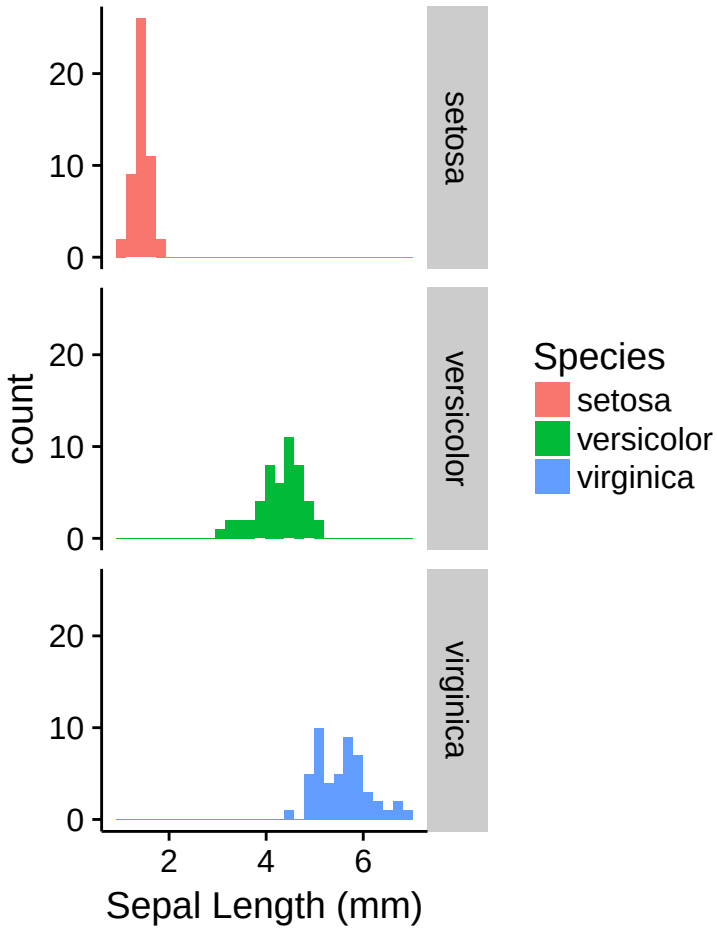
What does this show?

```
qplot(Petal.Length,
      data = iris,
      fill = Species,
      facets = Species ~.)
```



Add a label to x-axis

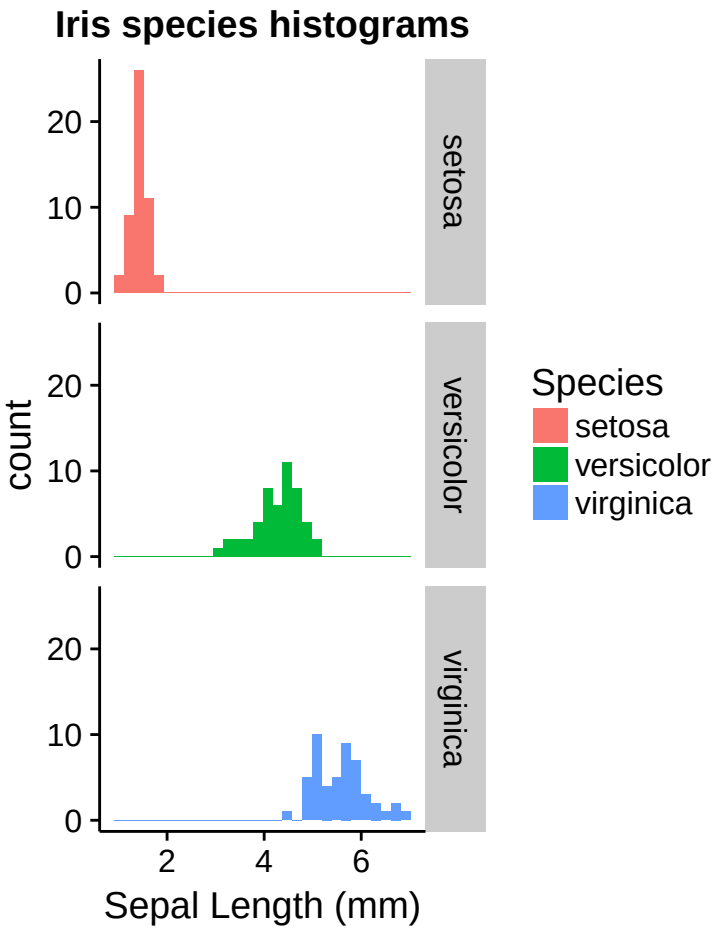
```
qplot(Petal.Length,
      data = iris,
      fill = Species,
      facets = Species ~.) +
  xlab("Sepal Length (mm)")
```



Modifying histograms: titles with the main = argument

- Titles are good for your own personal use but actually are almost never used in figures published in papers and books.
- We can add a title like this using the arguement "main ="

```
qplot(Petal.Length,
      data = iris,
      fill = Species,
      main = "Iris species histograms",
      facets = Species ~.) +
  xlab ("Sepal Length (mm)")
```


Iris species histograms
setosa
versicolor
virginica
count
20
10
0
2
4
6
Sepal Length (mm)
Species
setosa
versicolor
virginica