Mini-Tutorials for Creating Graphs in R
Boxplots

The examples below are intended to instruct you to create statistical graphs in R with minimal initial training in R. You should be able to follow the example codes to obtain graphs by modifying the included code. Some examples (and a key) will be included at the end of the document for practice. The screenshots I will show of the environment use R Studio, which is a free program you can find online. Other R environments will look different, but probably have similar functionality.

When you open up a new project environment in R Studio, it looks like this.
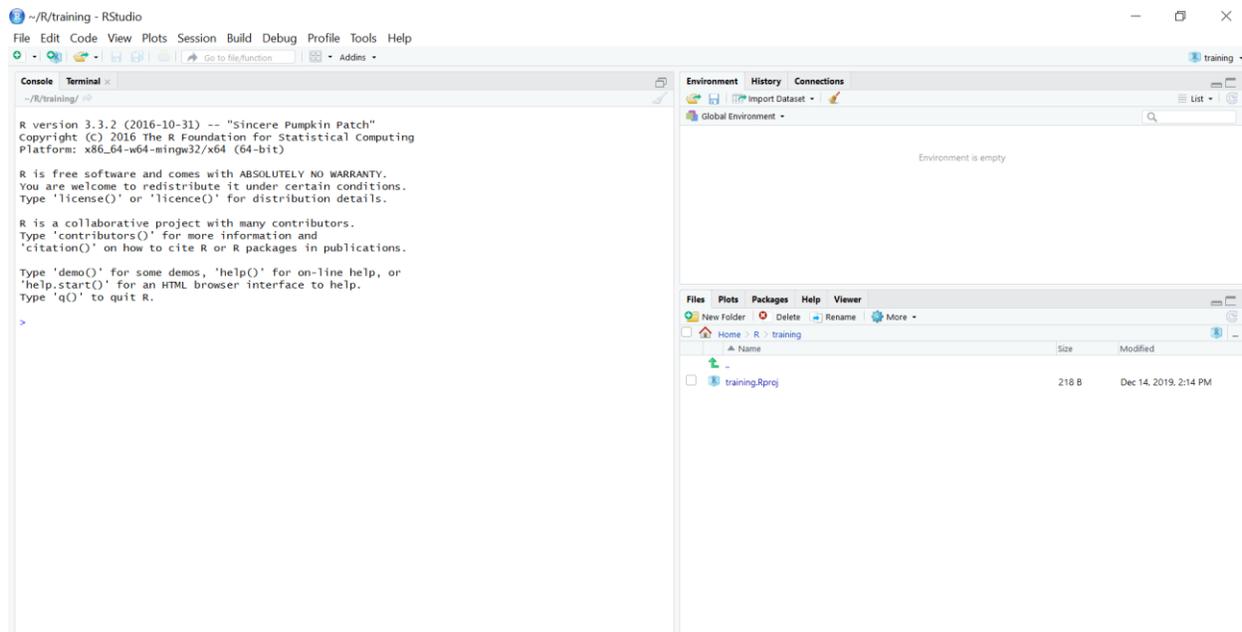
The command line environment is on the left. Images when we construct them will appear on the bottom right. As we add variables, they will appear in the list at the top right (name, dimensions and samples will display, which is useful for checking that you didn't skip entries when entering data by hand).

We are going to create a simple boxplot from raw data. Comparative boxplots will be dealt with at a later point.

Copy the commands shown into the command line.

Step 1. Enter the data to be plotted in vector form.

InflationRate=c(1.59, 3.01, 2.78, 4.27, 5.46, 5.84, 4.30, 3.27, 6.16, 11.03, 9.20, 5.75, 6.50, 7.62, 11.22, 13.58, 10.35, 6.16, 3.22, 4.30, 3.55, 1.91, 3.66, 4.08, 4.83, 5.39, 4.25, 3.03, 2

Betsy McCall

.96, 2.61, 2.81, 2.93, 2.34, 1.55, 2.19, 3.38, 2.83, 1.59, 2.27, 2.68, 3.39, 3.24, 2.85, 3.85, -0.34, 1.64, 3.16, 2.07, 1.47, 1.62, 0.12, 1.26, 2.13)

Boxplots are quite simple to create in R.

Step 2. Graphing the data.

boxplot(InflationRate, ylab="Inflation Rate")
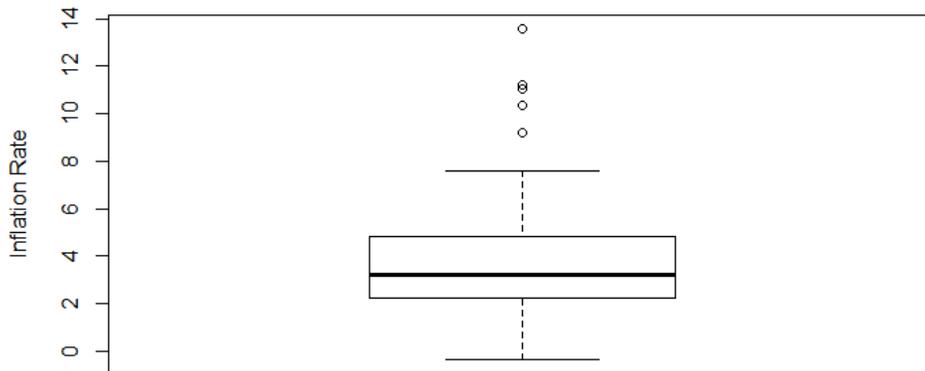
The graph we get is shown. The default option is a vertical boxplot, but you can make it horizontal, change the coloring and so forth.

boxplot(InflationRate, main="Boxplot of Inflation Rate", xlab="Inflation Rate", col = "green", border = "brown", horizontal = TRUE)

Boxplot of Inflation Rate

Betsy McCall
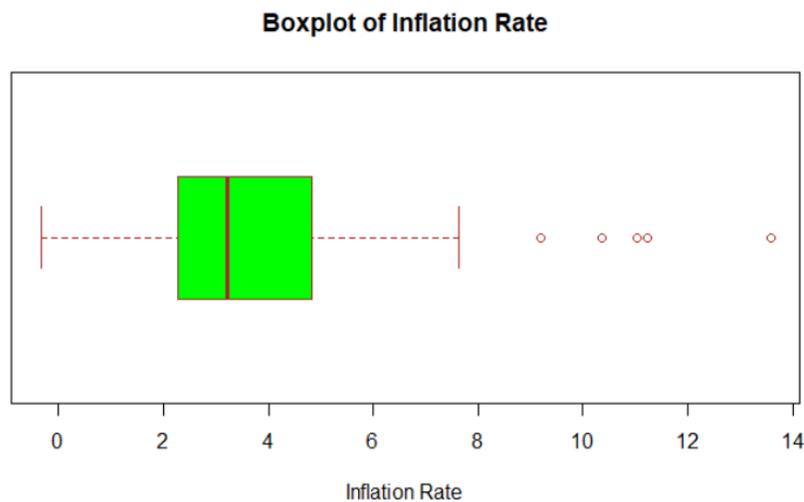
Note: if you change the orientation, you need to change the location of the axis label. Use y when the boxplot is vertical and x when the boxplot is horizontal.

Comparative Boxplots

Step 1. We can also create a comparative boxplot with two or more data sets. Let's create some random data to look at.

To get some random data to practice with, consider the following commands:

1. test1=c(runif(100, min=80, max=100))

runif produces a list of uniformly distributed random numbers between the specified min and max values. In this case, 100 such numbers are generated.

2. test2=c(rnorm(150, mean=64, sd=3.1))

rnorm generates a list of normally distributed numbers with the specified mean and standard deviation. In this case 150 such numbers are generated.

Step 2. To create the boxplot with both datasets, it looks something like this.

boxplot(test1, test2, main = "Multiple boxplots for comparision", at = c(1,2), names=c(1, 2), xlab = "test data sets", ylab="values", las = 2, col = c("red","blue"), border = "black", horizontal = FALSE, notch = FALSE)

Multiple boxplots for comparision
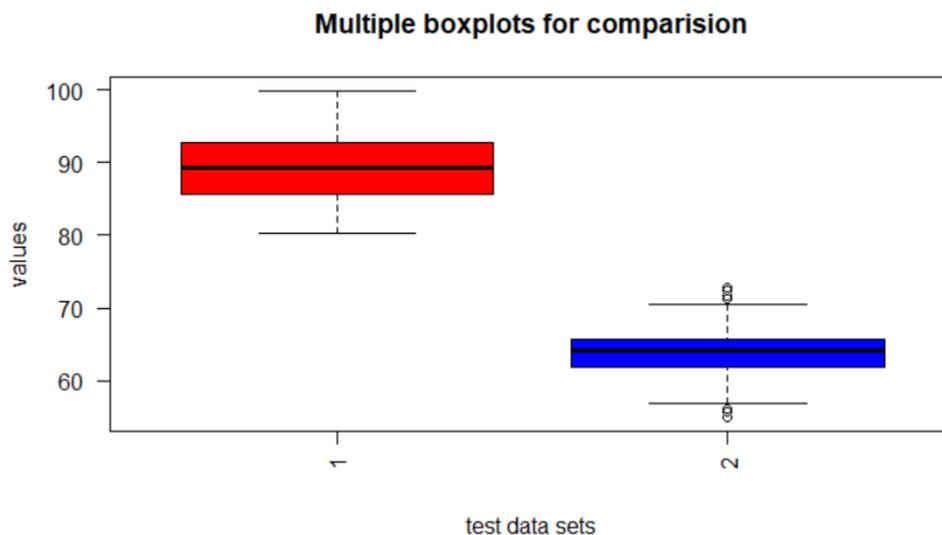
Practice.
To get some random data to practice with, use the commands above to obtain sample data with different ranges and distributions. Create any number of sets you like and experiment

Betsy McCall

with boxplots of various sizes, comparing multiple datasets, and using different types of features.

Solutions.
Some example plots.

You can also add the means to the boxplots. The example below is for vertical plots.

means=mean(test1,test2)

points(means, pch = 4, col = "white", lwd = 2)
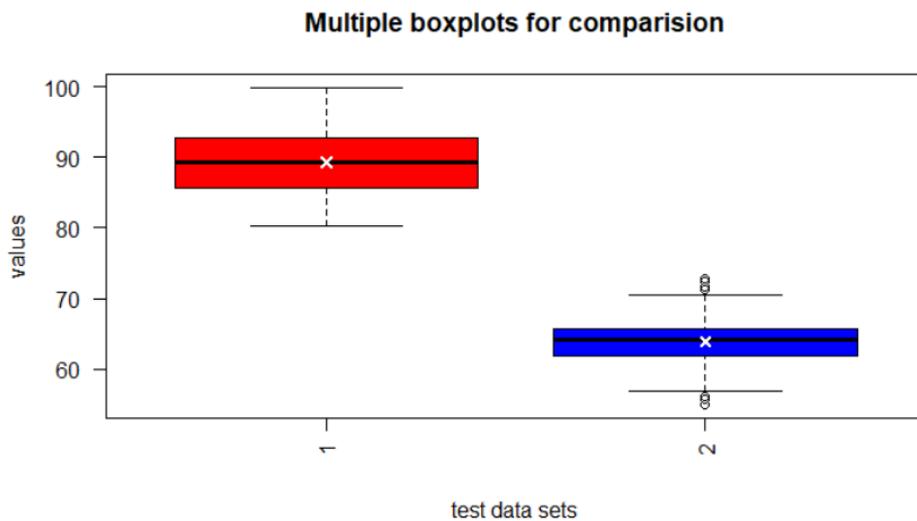
Here are 5 plots compared.

test4=c(rnorm(40, mean=100, sd=20))

test5=c(rnorm(40, mean=100, sd=15))

test6=c(rnorm(40, mean=97, sd=10))

test7=c(rnorm(40, mean=105, sd=18))

test8=c(rnorm(40, mean=120, sd=5))

boxplot(test4, test5, test6, test7, test8, main = "Multiple boxplots for comparision", at = c(1,2,3,4,5), names=c(4,5,6,7,8), xlab = "test data sets", ylab="values", las = 2, col = c("pink","skyblue", "darkgreen","yellow","orange"), border = "black", horizontal = FALSE, notch = FALSE)

Betsy McCall

Multiple boxplots for comparision

values

test data sets