R List of Important Stuff:

Asking R for help:
>help(FUNCTIONNAME) #pulls up the help for the specific function; this is R's equivalent of "man".
>?FUNCTIONNAME #this is the same as the above
>??WORD #this will search the help data for the word you typed
-Note: there is also a help search bar in the top right menu.

Saving:
There are two components to an R session, the console (i.e., the lists of commands) and the image (this is the data). I find it useful to save both separately.

- Go to File>Save As and save your console as a text file.
- Go to Workspace > Save Workspace File to save your workspace.

Note: once I have something well worked out I tend to copy from my workspace into a new text file (without the ">") so that I can then chuck that code directly into R to get my output. This code is also publishable as a workflow and can be run as a batch.

Setting your working directory:
-go to Misc. and set working director
OR
-at the command line time type what is below where PATH is the path for the directory you want to work in.
>setwd("PATH")

Importing Libraries:
Type the following at the prompt and substitute in the exact name of the package you want.
>install.packages("PACKAGENAME")
You will later be prompted to select a CRAN mirror. I always use Indiana, and this makes some sense and has served me well.

Loading Packages:
Type the following at the prompt:
>library(NAMEOFLIBRARY)

Importing Data:
The following is for a tab delim file where the first row has column names and the first column has row labels (as is the case in a matrix). "NAME" is important in that you need to give a name to the data so that you can call it up and use it…otherwise it will just print it to the screen.

> NAME = read.table(file="FILENAME", sep="\t", header =TRUE, row.names=1)

You can also make data by using the "c" (combine) function. That is, I want to make a vector that is a series from 1 to 5…
DATAVECTOR = c(1:5)

Basic Data Manipulation:

- Get the dimensions of your data table
>dim(DATANAME)

- Look at the headers of the data
>names(DATANAME)

- Look at a specific row, column, or cell in your table
>DATANAME[5,3] #refers to row 5, column 3
>DATANAME[5,] #refers to row 5
>DATANAME[,3] #refers to column 3
>DATANAME$VAR1 #refers to VAR1 (i.e., the name of a variable in your header)

- Transpose your data (i.e., flip it, what were rows are now columns, etc.)
>DATANAME2 = t(DATANAME)

- Add a variable/vector
>DATANAME$NEWVAR = c(1,2,3,4,5) # this added a new variable, called NEWVAR, that is a series from 1 to 5 (this could have alternatively been written "1:5".

- Combine columns or rows
>DATANAME3 = cbind(DATANAME2, DATANAME3) #binds columns (i.e., it puts them side by side in a new frame)
>DATANAME3 = rbind(DATANAME2, DATANAME3) #binds rows (i.e., it stacks them)
- Note: the "merge" function is useful for merging datasets and it will take arguments, such as to merge by a sample name.

- Subsetting data
>DATANAME2 = DATANAME[1:2] #creates a new table that is only the first two rows of your original data.
>DATANAME2 = DATANAME[,-2] #creates a new table that does not include the second column of your original data.
>DATANAME2 = subset(DATANAME, VAR1 == 0) # creates a new table that only includes rows where VAR1 had a value of 0. This function will take all logical operators and can work with combinations of operators.

Basic Graphing:

- the basic plot function is "plot". This accepts basic x and y data. This will print to the screen and can be saved using a function or using the menu.

>plot(DATANAME$VAR1, DATANAME$VAR2)

- you can color code by a variable by adding the optional argument "col=VAR"
- you can code variables by symbols by adding the optional argument "pch=VAR"
- Note: for colors and symbols it is often very useful to make new variables with the color names and/or the symbol numbers.

Colors: www.stat.columbia.edu/~tzheng/files/**Rcolor**.pdf

Symbols: http://www.endmemo.com/program/R/pchsymbols.php