

February 2005: Peter Adler: NCEAS Postdoctoral Fellow and Power R user, provides this concise list of useful R functions.

For further details concerning any of these functions, consult the R Language help system by entering `?<function_name>` at the R system prompt (example: `>?read.table`)

These Are a Few of My Favorite R Things

Input/Output:

<code>read.table()</code>	Read in many kinds of data
<code>read.csv()</code>	Read in comma delimited data
<code>write.table()</code>	Write tabular data in many formats
<code>scan()</code>	Read in really huge datasets (faster but more complicated function)
<code>getwd()</code>	See the current working directory
<code>setwd()</code>	Set the current working directory

Create and manipulate data objects:

<code>c()</code>	Make a vector of specified elements, or add to a vector
<code>seq()</code>	Numeric sequences
<code>rep()</code>	Vectors of repeated values
<code>matrix()</code>	2-dimensional arrays
<code>array()</code>	Multidimensional arrays
<code>data.frame()</code>	Looks like a matrix, but some columns can be numeric, others character (actually this is a special type of list)
<code>list()</code>	Collections of different data objects (“cell arrays” in Matlab)
<code>rm()</code>	Remove (erase) objects
<code>ls()</code>	Print a list of the all objects currently in memory

Indexing:

<code>which()</code>	Find elements of a vector that satisfy a condition
<code>is.element()</code>	Given two vectors, find elements that occur in both
<code>length()</code>	Get the length of a vector
<code>dim()</code>	Get or assign the dimensions of a matrix, dataframe, or array
<code>ncol()</code>	The number of columns in a matrix
<code>nrow()</code>	The number of rows in a matrix
<code>colnames()</code>	Get or assign column names of a vector
<code>names()</code>	Get or assign the column names of a dataframe or object names of a list

Database operations:

<code>merge()</code>	Merge or join dataframes
<code>unique()</code>	List of each unique element in a vector (even if they repeat)
<code>subset()</code>	Pull out certain records in a dataframe
<code>aggregate()</code>	Do operations (means, sums) on columns of dataframe given grouping variables
<code>stack()</code>	Take repeated elements of a vector and put them in different columns
<code>reshape()</code>	Like <code>stack()</code> and <code>unstack()</code> but more flexible
<code>sort()</code>	Sort a vector

table() Count the number of times individual values occur

Take sums/means/variances of rows/columns/arrays:

rowSums() Just what it sounds like for 2-D matrices

colSums() Ditto

apply() Apply any function to specified dimensions of an array

Control structures:

for() Loops

break() Escape a for loop

if(), else() Self-explanatory

Probability:

runif() Draw random numbers from a uniform distribution

rnorm() Draw random numbers from a normal distribution

dnorm() Normal density function (many other distributions available)

sample() Sample elements from a vector

Text manipulation:

paste() Concatenate strings

grep() Find regular expressions in a character vector

substr() Extract or replace pieces of a character vector

General plotting functions:

plot() Scatter and line plots

matplot() Literally “matrix plot”, good for time series of many responses

hist() Histograms

barplot() Barplots

image() Good for “maps” of matrices

axis() Create a “custom” plot axis

text() Write text in a plot

mtext() Write text in the margins of a plot

pdf() Create the figure as a pdf file (also jpeg, postscript, etc)

My favorite arguments within par(), the graphics parameters list:

mfrow() Create a multipanel plot

new Add a new plot to an existing plot

mar() Lines of space in margin of each plotting panel

oma() Lines of space in the outer margin of the plot

tcl() Length of axis tick marks

cex Magnification of symbols and labels (also see cex.axis(), etc)

mgp() Controls distance between the axis and axis labels and title