Summary of functions used by section *in*

An Introduction to Data Analysis and Graphics with R

Note: see help files for more information. Many of the summaries given here are shortened versions of the descriptions found in the help files.

2. Introduction to R

ls() – Lists the objects in your workspace.

rm() – Removes object(s) from workspace.

class() – Returns the class of an object (type of data structure).

matrix() – Creates a matrix.

array() – Creates an array.

mode() – Returns the mode of an object (type of data).

data.frame() – Creates a data frame.

list() – Creates a list.

sum() – Sums its arguments.

plot() – Generic plotting function.

args() – Returns the arguments of a function.

install.packages() – Installs a package on your computer.

detach() – Removes packages or data frame columns from the workspace.

is.na() – Returns TRUE if an argument is NA.

?function_name() – Opens help file for function "function_name".

rnorm() – Generates a vector of random (pseudo-random) numbers.

3. Vectors, matrices, and arrays

seq() – Generates a vector that contains a regular sequence of numbers.

rep() – Generates a vector with a repeated value or repeated sequence.

c() – Concatenate (or combine) arguments into one vector.

paste() – Combines data into a vector of character data.

sqrt() – Square root.

matrix() – Creates a matrix.

as.vector() – Coerces argument into a vector.

array() – Creates an array.

solve() – For solving matrix algebra equations (or inverting matrices).

outer() – Applies a function to all combinations of the elements of two vectors, result is a matrix.

4. Data frames, data input, and data output

read.table() – Read in data from a file and create a data frame with the data.

Summary of functions used by section *in*

An Introduction to Data Analysis and Graphics with R

edit() – Opens up a data frame in a editor.

data.frame() – Creates a data frame.

names() – For specifying or checking the names of an object.

attach() – Adds a database (typically data frame or library) to R's search path.

detach() – Removes a database from R's search path.

na.omit() – Strips NAs from an object.

is.na() – Returns TRUE if an argument is NA.

write.table() – Writes data to a file.

sink() – Sends output to a file.

plot() – Generic function for plotting R objects. To learn about R's default plot function, type ?plot in the R GUI.

5. Graphics, part I

plot() – Generic function for plotting R objects. To learn about R's default plot function, type ?plot.

text() – Draws a character string in a plotting region.

6. Manipulating data

mode() – Used to get or set the storage mode of an object.

length() – Used to get or set the length of vectors (including lists) and factors and of any other R object for which a method has been defined.

attributes() – Used to access an object's attributes. Usually returns a list of attributes.

rnorm() – Used to generate random numbers (or pseudo-random numbers) from a normal distribution.

class() – Used to view or set the class of an object.

matrix() – Used to create a matrix.

nrow() – Used to set or view the number of rows present in an array or matrix.

ncol() – Used to set or view the number of columns present in an array or matrix.

dim – Used to retrieve or set the dimensions of an object.

NROW() – Returns the number of rows in a vector.

as.character() – Creates or tests for objects of type "character".

as.data.frame() – Used to check if an object is a data frame, or to coerce to a data frame, if possible.

as.numeric() – Used to create a numeric object or to coerce an object to a numeric type.

as.vector() – Used to coerce (or attempt to coerce) its argument into a vector with a specified mode.

Summary of functions used by section *in*

An Introduction to Data Analysis and Graphics with R

read.table() – Reads a file in table format and creates a data frame from it.

paste() – Concatenates vectors after converting them to character.

subset() – Returns subsets of vectors, matrices or data frames subject to various constraints.

array() – Creates or tests for arrays.

lists() – Used to construct, coerce and check for both kinds of R lists (i.e. list and pairlist).

split() – Divides the data in the specified vector into groups defined by the argument f.

sort() – Used to sort a vector or factor in ascending (or descending) order.

order() – Returns a permutation which rearranges its first argument into ascending or descending order, breaking ties by further arguments. Can be used to sort an entire data frame.

match() – Returns a vector of the positions of (first) matches of its first argument in its second.

which() – Gives the true indices of a logical object, allowing for array indices.

is.na() – Returns a logical vector of the same length of its argument, containing TRUE for those elements marked NA or NaN, and FALSE otherwise.

rep() – Replicates the specified value, by the specified number of times.

factor() – Used to encode a vector as a factor.

levels() – Provides access to the levels attribute of a variable. Typically returns the value of the levels of its argument.

as.Date() – Converts between character representations and objects of class "Date" representing calendar dates.

mean() – Generic function for the arithmetic mean. Can also be used to trim the mean.

strptime() – Functions to convert between character representations and objects of classes "POSIXlt" and POSIXct" representing calendar dates and times.

weekdays() – Returns the weekday of a specified date.

as.POSIXct() – Converts an object to a date/time object.

as.POSIXlt() – Converts an object to a date/time object .

names() – Retrieves or sets the names of objects.

difftime() – Calculates the difference of two date/time objects and returns an object of class "difftime" with an attribute indicating the units.

data.frame() – Creates data frames (i.e. tightly couple collections of variables which share many of the properties f matrices and of lists), used as the fundamental data structure by most of R's modeling software.

Summary of functions used by section *in*

An Introduction to Data Analysis and Graphics with R

ifelse() – Returns yes or no depending on whether the elements of the first argument are TRUE or FALSE.

Sys.time() – Returns the system's idea of the current date with time.

for() – Used for setting up a loop with a counter variable.

cbind() – Combines a sequence of vector, matrix, or data frame arguments by columns.

rbind() – Combines a sequence of vector, matrix, or data frame arguments by rows.

merge() – Merges two data frames by common column or row names.

summary() – Produces summaries of objects.

table() – Uses cross-classifying factors to build a contingency table of the counts at each combination of factor levels.

lapply() – Returns a list of the same length as the specified vector, each element of which is the result of applying a function (FUN) to the corresponding elements of the specified vector.

sapply() – A user-friendly version of lapply that, by default, returns a vector or matrix if appropriate.

tapply() – Applies a function to each cell of a ragged array, that is to each non-empty group of values given by a unique combination of the levels of certain factors. Effectively applies a specified function to subsets of data that comprise groups within the levels of a column.

aggregate() – Splits the data into subsets, computes summary statistics for each, and returns the results in a convenient form.

write.table() – Prints its first argument (typically a data frame) to a file or connection.

unsplit() – Puts elements or rows back in the positions given by a list of factors or groups.

ave() – Used for averaging subsets of an object, where each subset consists of those observations with the same factor levels.

7. Exploratory data analysis

read.table() – Read in data from a file and create a data frame with the data.

names() – For specifying or checking the names of an object.

mean() – Returns the arithmetic mean.

median() – Returns the median.

sd() – Returns the standard deviation.

var() – Returns the variance.

summary() – Generates a summary of data or a model.

qqnorm() – Produces a normal quantile plot.

abline() – Adds a line(s) to a plot.

Summary of functions used by section *in*

An Introduction to Data Analysis and Graphics with R

ros() – Regression on order statistics for dealing with nodetects.

mean() – Generic function for the arithmetic mean. Can also be used to trim the mean.

plot() – Generic function for plotting R objects.

data.frame() – Creates a data frame.

read.table() – Read in data from a file and create a data frame with the data.

hist() – Generates a histogram.

boxplot() – Generates a boxplot.

qqnorm() – Produces a normal quantile plot.

qqline() – Adds a line to a normal quantile plot (through 1st and 3rd quartiles).

sort() – Sorts a vector.

ppoints() – Returns cumulative probabilities.

qnorm() – Returns quantiles for a normal distribution.

sd() – Returns the standard deviation .

abline() – Adds a line(s) to a plot.

axis() – Adds an axis to the current plot.

quantile() – Returns sample quantiles.

8. One- and two-sample tests (and the R approach to statistical output)

read.table() – Reads a file in table format and creates a data frame from it.

t.test() – Performs a *t* test.

library() – Loads a library (package).

names() – Retrieves or sets the names of objects.

boxplot() – Generates a boxplot.

class() – Returns the class of an object (type of data structure).

9. Classical linear models

read.table() – Reads a file in table format and creates a data frame from it.

names() – Retrieves or sets the names of objects.

summary() – Produces summaries of objects.

plot() – Generic function for plotting R objects.

lm() – Fits a linear model

coef() – Returns model coefficients.

resid() – Returns model residuals.

anova() – Returns an analysis of variance (deviance) table.

predict() – Returns model predictions.

Summary of functions used by section *in*

An Introduction to Data Analysis and Graphics with R

matlines() – Plots matrix columns.

cor() – Returns correlation coefficients.

pairs() – Generates a matrix of scatterplots.

update() – Removes or adds terms to a model.

I() – Treats an object "as is".

aov() – Performs an analysis of variance.

TukeyHSD() – Applies Tukey's Honestly Significant Different Test.

replications() – Returns the number of replicates for each term in a model. Can be used to check for balance in factorial designs.

with() – Applies an expression in a local environment that contains the specified data.

factor() – Encodes a vector as a factor.

par() – To specify or check graphical parameters.

10. Nonparametric analogs to t tests and ANOVA

wilcox.text() – Performs a Wilcoxon test.

kruskal.test() – Performs a Kruskal-Wallace rank sum test.

pairwise.wilcox.test() – Performs a pairwise Wilcoxon rank sum test.

11. Graphics II

par() – Can be used to set or query graphical parameters.

plot() – Generic function for plotting R objects. To learn about R's default plot function, type ?plot in the R GUI.

layout() – Allows the user to divide the plotting area into various regions specified as a matrix with the specified number of columns and rows. Allows for the specification of complex plotting arrangements.

layout.show() – Shows the outlines of the figures in the current layout. The only argument specifies the number of plots per page.

matrix() – Used to create a matrix from the given set of values. Is used to specify the arrangement of plots in the layout.

rnorm() – Used to generate random numbers (or pseudorandom numbers) from a normal distribution.

axis() – Used to add an axis to the current plot, allowing for the specification of the side, position, labels, and other options.

legend() – Used to add legends to plots, or a convenient way to add a simple textual description inside a plot region.

read.table() – Reads a file in table format and creates a data frame from it.

Summary of functions used by section *in*

An Introduction to Data Analysis and Graphics with R

na.omit() – Used to omit cases when an NA is present in an object.

as.Date() – Used to plot objects of classes "POSIXlt", "POSIXct", and "Date" representing calendar dates and times. Useful for plotting time series.

strptime() – Used to convert between character representations and objects of classes "POSIXlt" and "POSIXct" representing calendar dates and times.

text() – Used to draw character strings at specified coordinates in a plotting region.

arrows() – Used to draw arrows between pairs of coordinates in a plotting region.

lines() – Used to draw lines in a specified manner (i.e. either as a line joining coordinates or by using vectors of coordinates or with a specified function).

lowess() – A smoothing function that can be used to draw lines; it uses locally-weighted polynomial regression.

abline() – Used to add one or more straight lines through the current plot region, it can take a linear model (lm) as an argument.

lm() – A function used to fit linear models, it can be used to carry out regression, ANOVA, or ANCOVA.

demo() – A user-friendly interface for running some demonstration scripts; typically provides output for various utilities to demonstrate the available features.

expression() – Can be used to add mathematical equations and symbols to figures; often used within the text function for adding equations to figures.

demo(plotmath) – Can be used to demonstrate how to produce the desired mathematical equation.

mtext() – Used to write text into the margins of a plot.

rect() – Used to draw rectangles, typically within plot regions.

polygon() – Used to draw polygons in plot regions.

dotchart() – Draws a Cleveland dot plot.

hist() – Computes and draws a histogram of the given data values.

lists() – Used to construct, coerce and check for both kinds of R lists (i.e. list and pairlist).

boxplot() – Used to produce box-and-whisker plots of the given (grouped) values.

qqnorm() – Produces a normal quantile-quantile plot, which can be used to graphically assess normality.

qqline() – Adds a line to a normal quantile-quantile plot which passes through the first and third quartiles.

barplot() – Creates a bar plot with vertical or horizontal bars.

contour() – Creates a contour plot or adds contour lines to an existing plot.

Summary of functions used by section *in*

An Introduction to Data Analysis and Graphics with R

image() – Creates a grid of colored rectangles; can be used to display 3-dimensional or spatial data.

persp() – Draws a perspective plot of surfaces.

png() – Used to create png format bitmap files.

dev.off() – Used to shutdown the specified device (by default, the current device); typically used to signal the end of a graphics file (i.e. it shuts down the png device).

pdf() – Creates a pdf file.

12. Generalized linear models

read.table() – Reads a file in table format and creates a data frame from it.

glm() – Used to fit generalized linear models, uses a link function and a description of the error distribution.

summary() – Used to produce a summaries of various model fitting functions.

cbind() – Combines a sequence of vector, matrix, or data frame arguments by columns.

seq() – Produces regular sequences

predict() – Used to predict the results of various model fitting functions.

plot() – Generic function for plotting R objects. To learn about R's default plot function, type ?plot in the R GUI.

points() – Used to draw a sequence of points at specified coordinates (typically contained in vectors).

anova() – Used to compute an analysis of variance (or deviance) table for a fitted model object.

13. Generalized additive models

gam() – Fits a generalized additive model.

s() – Specifies that a smoothed function should be applied to a predictor.

14. Nonlinear regression

attach() – Used to attach a database to R's search path; objects in the database can be accessed by simply specifying their names.

plot() – Generic function for plotting R objects.

legend() – Used to add legends to plots, or a convenient way to add a simple textual description inside a plot region.

nls() – Used to determine the nonlinear least-squares estimate of the parameters of a nonlinear model.

summary() – Used to produce a summaries of various model fitting functions.

Summary of functions used by section *in*

An Introduction to Data Analysis and Graphics with R

coef() – A generic function that is used to extract model coefficients from fitted models.

SSmicmen() – A self-starting Michaelis-Menton function. Provides initial estimates of the parameters to be fitted using nls.

seq() – Produces regular sequences, from a given value to a given value, by a given value.

lines() – Used to draw lines in a specified manner (i.e. either as a line joining coordinates or by using vectors of coordinates or with a specified function).

15. Grouping, loops, and conditional execution

for() – Sets up a loop with a counter variable.

while() – Sets up a loop that continues until some criterion is met.

if() – Used for conditional execution.

ifelse() – Used for conditional execution with multiple elements.

16. Distributions and simulations

pnorm() – Returns the probability of a quantile or a vector of quantiles, given a mean and standard deviation.

plot() – Generic function for plotting R objects.

seq() – Produces regular sequences, from a given value to a given value, by a given value.

curve() – Used to produce a curve corresponding to a given function or expression as a function of x.

qnorm() – The inverse of pnorm; used to produce a quantile or quantiles, given a probability or a vector of probabilities and a mean and standard deviation.

rnorm() – Used to generate random numbers (or pseudorandom numbers) from a normal distribution, given a mean and standard deviation.

par() – Can be used to set or query graphical parameters. Parameters can be set by specifying arguments to par. Many arguments are available, to see them type ?par.

hist() – Computes and draws a histogram of the given data values.

qqnorm() – Produces a normal quantile-quantile plot, which can be used to graphically assess normality.

qqline() – Adds a line to a normal quantile-quantile plot which passes through the first and third quartiles.

qunif() – Used to provide the quantile or quantiles of a probability or a vector of probabilities from a uniform distribution.

punif() – The inverse of qunif; used to get the probability or probabilities of a quantile or a vector of quantiles from a uniform distribution.

Summary of functions used by section *in*

An Introduction to Data Analysis and Graphics with R

numeric() – Used to create a numeric object or coerce an object to a numeric type.

for () – Used to create for loops.

mean() – Returns the arithmetic mean of its arguments.

runif() – Used to give random numbers (or pseudo-random numbers) from a uniform distribution.

read.table() – Reads a file in table format and creates a data frame from it.

names() – Retrieves or sets the names of objects.

na.omit() – Used to omit cases when an NA is present in an object.

subset() – Returns subsets of vectors, matrices or data frames subject to various constraints.

t.test() – Performs a t test; by default, the assumption is that variances are unequal.

log10() – Computes the base 10 logarithm.

sqrt() – Computes the square root of a value or a vector of values.

sd() – Returns the standard deviation

mean() – Returns the arithmetic mean of its arguments.

abs() – Returns the absolute value of a value or a vector of values.

qt() – Provides a quantile or a vector of quantiles from a t distribution, given a probability or a vector of probabilities.

power.t.test() – Used to compute the power of a t test, or to determine the parameters to obtain a target power.

17. Functions

function() – Used to define a function.

debug() – Flags a function for debugging.

18. Batch processing

R CMD BATCH – Command used to run R non-interactively from a shell.