Data Analysis: Programming

Institute of Technology Tallaght

Department of Computing

R

- R is an interpreted programming language for data processing and visualisation
- It incorporates many statistical methods, both simple and advanced
- Based on the S language that was developed by researchers at Bell Labs in New Jersey
- Developed by Ross Ihaka and Robert Gntleman at University of Auckland in New Zealand under the GNU public license (early 1980s)
- Today, maintained by a dedicated group called the R Core Team

General

Function/operator	Brief description
`options`	Set various R options
`#`	A comment (ignored by interpreter)
`getwd`	Print current working directory
`setwd`	Set current working directory
`library`	Load an installed package
`install.packages`	Download and install package
`update.packages`	Update installed packages
`help` or `?`	Function/object help file
`help.search` or `??`	Search help files
`q`	Quit R

Numerics, Arithmetic, Assignment and Vectors

Function/operator	Brief description
`+, *, -, /, ^`	Arithmetic
`sqrt`	Square root
`log`	Logarithm
`exp`	Exponential
`<-, =`	Object assignment
`c`	Vector creation
`:, seq`	Sequence creation
`rep`	Value/vector repetition
`sort`	Vector sorting
`length`	Determine vector length
`[ ]`	Vector subsetting/extraction
`sum`	Sum all vector elements
`prod`	Multiply all vector elements

Matrices and Arrays

Function/operator	Brief description
`matrix`	Create a matrix
`rbind`	Create a matrix (bind rows)
`cbind`	Create a matrix (bind columns)
`dim`	Get matrix dimensions
`nrow`	Get number of rows
`ncol`	Get number of columns
`[ , ]`	Matrix/array subsetting
`diag`	Diagonal elements/identity matrix
`t`	Matrix transpose
`*`	Scalar matrix multiple
`+, -`	Matrix addition/subtraction
`%*%`	Matrix multiplication
`solve`	Matrix inversion
`array`	Create an array

Non-Numeric Values

Function/operator	Brief description
`TRUE, FALSE`	Reserved logical values
`T, F`	Unreserved versions of above
`==, !=, >, <, >=, <=`	relational operators
`any`	Checks whether any entries are `TRUE`
`all`	Checks whether all entries are `TRUE`
`&&, &, ||, |, !`	logical operators
`which`	Determines indexes of `TRUE`s
`" "`	Creates a character string
`nchar`	Gets number of characters in a string
`cat`	Concatenates strings (no return)
`paste`	Pastes strings (returns a string)
`\`	String escape
`substr`	Subsets a string
`sub, gsub`	String matching and replacement
`factor`	Creates a factor vector
`levels`	Gets levels of a factor
`cut`	Creates factor from continuum

Lists and Data Frames

Function/operator	Brief description
`list`	Create a list
`[[ ]]`	Unnamed member reference
`[ ]`	List slicing (multiple members)
`$`	Get named member/variable
`data.frame`	Create a data frame
`[ , ]`	Extract data frame row/columns

Special Values, Classes and Coercion

Function/operator	Brief description
`Inf, -Inf`	Value for ±infinity
`is.infinite`	Element-wise check for `Inf`
`is.finite`	Element-wise check for finiteness
`NaN`	Value for invalid numerics
`is.nan`	Element-wise check for `NaN`
`NA`	Value for missing observation
`is.na`	Element-wise check for `NA` OR `NaN`
`na.omit`	Delete all `NA`s and `NaN`s
`NULL`	Value for "empty"
`is.null`	Check for `NULL`
`attributes`	List explicit attributes
`attr`	Obtain specific attribute
`dimnames`	Get array dimension names
`class`	Get object class (S3)
`is._`	Object-checking functions
`as._`	Object-coercion functions

Basic Plotting

Function/operator	Brief description
`plot`	Create/display base R plot
`type`	Set plot type
`main, xlab, ylab`	Set axis labels
`col`	Set point/line color
`pch, cex`	Set point type/size
`lty, lwd`	Set line type/width
`xlim, ylim`	Set plot region limits
`abline`	Add vertical/horizontal line
`segments`	Add specific line segments
`points`	Add points
`lines`	Add lines following coords
`arrows`	Add arrows
`text`	Add text
`legend`	Add/control legend
`qplot`	Create `ggplot2` "quick plot"
`geom_point`	Add points geom
`geom_line`	Add lines geom
`size, shape, color`	Set geom constants
`linetype`	Set geom line type
`mapping, aes`	Geom aesthetic mapping
`geom_hline`	Add horizontal lines geom
`geom_segment`	Add line segments geom