R Quick Reference Guide

Jeff Stanton

*Tips: Use R-Studio for everything. Save working code in the source window. Use comments liberally with the # character. Save the "workspace" at the end of each session so that data, variables, and history can be restored. Use websites for help:*

[*http://www.statmethods.net*](http://www.statmethods.net)

[*http://www.ats.ucla.edu/stat/r/*](http://www.ats.ucla.edu/stat/r/)

[*http://cran.r-project.org/doc/manuals/*](http://cran.r-project.org/doc/manuals/)

*In the material below: Italics = instructions*

**Bold = commands,** Plain = output

---------------

*Create a new vector/variable using the assignment operator “<-“ and the c() function (combine or concatenate):*

**year <- c(2015, 2014, 2013, 2012)**

*Create a factor:*

**oe <- factor(c("odd","even","odd","even"))**

*Reveal the structure and storage modality of a variable or any data object:*

**str(oe)**

Factor w/ 2 levels "even","odd": 2 1 2 1

**mode(oe)** or

**typeof(oe)**

[1] "numeric"

*Create a dataframe (a rectangular data structure with rows and columns):*

**yearDF <- data.frame(year, oe)**

*Show the first few elements/cases in a data structure (note row #s on left):*

**head(yearDF)**

year oe

1 2015 odd

2 2014 even

3 2013 odd

4 2012 even

*Show the dimensions (e.g., number of rows and columns) of a data object:*

**dim(yearDF)**

[1] 4 2

*Add a new variable onto a dataframe:*

**temp <- c(28,24,25,22)**

**yearDF <- data.frame(yearDF,temp)**

**dim(yearDF)**

[1] 4 3

*Access one variable in a data frame with the $ accessor and calculate descriptive statistics:*

**mean(yearDF$temp)**

[1] 24.75

**sd(yearDF$temp)**

[1] 2.5

**median(yearDF$temp)**

[1] 24.5

**var(yearDF$temp)**

[1] 6.25

*Correlate two variables:*

**cor(yearDF$year,yearDF$temp)**

[1] 0.8778762

*Access individual rows and columns or ranges of rows/columns:*

**yearDF[2,]**

year oe temp

2 2014 even 24

**yearDF[,2]**

[1] odd even odd even

Levels: even odd

**yearDF[,1:2]**

year oe

1 2015 odd

2 2014 even

3 2013 odd

4 2012 even

*Sort the contents of a dataframe:*

**yearDF[order(yearDF$year), ]**

year oe

4 2012 even

3 2013 odd

2 2014 even

1 2015 odd

*Set and get the working directory:*

**setwd("/Users/Jeff/DataMining")**

**getwd()**

[1] "/Users/Jeff/DataMining"

*Save a data object into a binary file on disk:*

**save(yearDF, file="Years.RData")**

*Load a data object from a binary file on disk:*

**newData <- load("Years.RData")**

*Save a data object to disk as a comman separated variables file:*

**write.csv(yearDF,file="Years.csv")**

*Read a data object from a comma separated file on disk:*

**newData <- read.csv(file="Years.csv")**

*Remove a variable or data object from the workspace:*

**rm(yearDF)**

*Creating a New Function:*

**# Run this code to define the function**

**sqrNum <- function(inputValue)**

**{**

**outputValue <- inputValue ^ 2**

**return(outputValue)**

**}**

**# After that, the function can be called**

**# from the command line or other code**

**sqrNum(5)**

*Install a package and make it available for immediate use:*

**install.packages("ggplot2")**

**library(ggplot2)**

*Convert one data type to another*

**thing <- 123.45**

**typeof(thing)**

[1] "double"

**newThing <- as.character(thing)**

**typeof(newThing)**

[1] "character"

**newThing <- as.numeric(newThing)**

**typeof(newThing)**

[1] "double"

*Show available datasets in the base package of R or any other package:*

**data()**

**data(package="ggplot2")**

*Installing Rattle:*

**# Before installing RGtk2 on a Mac**

**# You may need to update X11 at**

**# https://www.xquartz.org**

**# As well as the GTK toolkit at**

**#** [**http://r.research.att.com/#other**](http://r.research.att.com/#other)

**install.packages("RGtk2")**

**install.packages("rattle")**

**library("rattle")**

**rattle()**

*Stanton's R Tutorials on SlideShare:*

All URLS start with:

<http://www.slideshare.net/jmstanto>

/installing-r-and-rstudio

/getting-started-with-r

/jmstanto/analytics-movingdata

/jmstanto/basic-graphics-with-r

/jmstanto/rstudio-vs-rcmdr