## Base R

## Cheat Sheet

## **Getting Help**

Get help of a particular function.

# help.search('weighted mean')

Search the help files for a word or phrase help(package = 'dplyr')

Find help for a package.

### str(iris)

class(iris) Get a summary of an object's structure.

Find the class an object belongs to.

# Jsing Packages

Download and install a package from CRAN. install.packages('dplyr')

## library(dplyr)

Load the package into the session, making all its functions available to use

## dplyr: select

Use a particular function from a package

data(iris)

Load a built-in dataset into the environment

# **Working Directory**

inputs are found and outputs are sent). Find the current working directory (where

# setwd('C://file/path')

Change the current working directory.

directory to the folder you are working in. Use projects in RStudio to set the working

# **Creating Vectors**

ectors

### 2:6 seq(2, 3, by=0.5) c(2, 4, 6)rep(1:2, each=3) rep(1:2, times=3)2 4 6 121212 2.0 2.5 3.0 23456 Repeat elements of a vector Join elements into Repeat a vector A complex sequence An integer sequence a vector

## **/ector Functions**

### sort(x) rev(x)

table(x) Return x sorted. Return x reversed unique(x)

See counts of values. See unique values

# Selecting Vector Elements

By Position

×[4] The fourth element

×[-4] All but the fourth

x[2:4]

Elements two to four

x[-(2:4)]All elements except two to four.

=

x[c(1, 5)]Elements one and

۵

By Value

x[x == 10]are equal to 10. Elements which

x[x < 0] All elements less than zero

x[x %in% c(1, 2, 5)] Elements in the set

Named Vectors

മ

Are equal Not equal

a > b a < b

Greater than Less than

Greater than or equal to

Is missing ls null

a <= b a >= b

is.null(a) is.na(a)

:= b 

x['apple'] name 'apple' Element with

## for (variable in sequence){ Do something

for (i in 1:4){ j <- i + 10

## Programming

while (condition){

While Loop

Do something

Example

### while (i < 5){ i ^- i + 1

f Statements

if (condition) { else { Do something different Do something

Example

if (i > 3){ print('Yes') print('No')

### function\_name <- function(var){</pre> return(new\_variable) Do something

**Functions** 

Example

square <- function(x){</pre> squared <- x\*x

return(squared)

# **Reading and Writing Data**

Also see the **readr** package.

| 1put                                     | Ouput                                    | Description  |
|--|--|--|
| df <- read.table('file.txt')             | <pre>write.table(df, 'file.txt')</pre>   | Read and write a delimited text file.  |
| <pre>df &lt;- read.csv('file.csv')</pre> | write.csv(df, 'file.csv')                | Read and write a comma<br>separated value file. This is a<br>special case of read.table/<br>write.table. |
| <pre>load('file.RData')</pre>            | <pre>save(df, file = 'file.Rdata')</pre> | Read and write an R data file, a file type special for R.  |

Converting between common data types in R. Can always go from a higher value in the table to a lower value.

| as.factor   | as.character                                       | as.numeric                          | as.logical                      |
|---|--|-------------------------------------|---------------------------------|
| '1', '0', '1',<br>levels: '1', '0'  | '1', '0', '1'                                      | 1, 0, 1                             | TRUE, FALSE, TRUE               |
| Character strings with preset levels. Needed for some statistical models. | Character strings. Generally preferred to factors. | Integers or floating point numbers. | Boolean values (TRUE or FALSE). |

|              |                                 | •                      |                         |
|--------------|---------------------------------|------------------------|-------------------------|
| log(x)       | Natural log.                    | sum(x)                 | Sum.                    |
| exp(x)       | Exponential.                    | mean(x)                | Mean.                   |
| max(x)       | Largest element.                | median(x)              | Median.                 |
| min(x)       | Smallest element.               | <pre>quantile(x)</pre> | Percentage              |
| round(x, n)  | Round to n decimal              | rank(x)                | Rank of elements.       |
|              | places.                         |                        |                         |
| signif(x, n) | Round to n significant figures. | var(x)                 | The variance.           |
| cor(x, y)    | Correlation.                    | sd(x)                  | The standard deviation. |
|              |                                 |                        |                         |

# Variable Assignment

> a <- 'apple'

[1] 'apple'

# The Environment

| rm(list = ls())               |              | rm(x)             |              | ls()                      |
|-------------------------------|--------------|-------------------|--------------|---------------------------|
| Remove all variables from the | environment. | Remove x from the | environment. | List all variables in the |

# You can use the environment panel in RStudio to browse variables in your environment.

df[2,

2]

rows.

Number of

environment.

## atrices

| 5           | = ,                    | -       |           | m[2,             |  |
|-------------|------------------------|---------|-----------|------------------|--|
| ב           | E                      | _       |           | _                |  |
|             | , 1] - Select a column |         |           | ] - Select a row |  |
| solve(m, n) | Matrix Multiplication  | ≡ %*% ⊐ | Transpose | t(m)             |  |

### Lists

m[2, 3] - Selectan element

Find x in: m \* x = n

tolower(x)

nchar(x)

Number of characters in a string.

 $l \leftarrow list(x = 1:5, y = c('a', 'b'))$ be of different types.

|               | 1[[2]] | A list is a colle                               |
|---------------|--------|---|
| New list with | ا[1]   | A list is a collection of elements which can be |
|               | l\$x   | which can be                                    |

Element named

Second element

of [

only the first element.

## ו['y']

only element New list with named y.

### dplyr package. Also see the

# **Data Frames**

 $df \leftarrow data.frame(x = 1:3, y = c('a', 'b', 'c'))$ A special case of a list where all elements are the same length.

### Н 0 О മ df\$x Understanding a data frame View(df) List subsetting df[[2]] frame.

### df[ , 2] **Matrix subsetting**



**rbind** - Bind rows. **↓** 

| <b>+</b> |                  |
|----------|------------------|
|          | Dilla colaitiis. |
|          |                  |

### Strings

# Also see the **stringr** package.

Join multiple vectors together.

paste(x, collapse = ' ') paste(x, y, sep = ' ')

gsub(pattern, replace, x)

grep(pattern, x)

toupper(x)

Replace matches in x with a string. Convert to lowercase. Convert to uppercase. Find regular expression matches in x Join elements of a vector together.

### Factors

Turn a vector into a factor. Can set the levels of the factor and factor(x) the order.

Turn a numeric vector into a cut(x, breaks = 4)factor by 'cutting' into sections.

## **Statistics**

 $lm(y \sim x, data=df)$ Linear model. difference between Perform a t-test for t.test(x, y)

prop.test

difference

between Test for a

 $glm(y \sim x, data=df)$ Generalised linear model

Get more detailed information out a model. summary

pairwise.t.test

means.

proportions

Perform a t-test for

paired data.

Analysis of

aov

variance.

See the full data

ω

2

head(df) rows. See the first 6

nrow(df) Number of rows. cbind - Bind columns

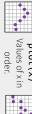
columns and

# Distributions

| qunif    | punif                      | dunif               | runif              | Uniform  |
|----------|----------------------------|---------------------|--------------------|----------|
| qbinon   | pbinom                     | dbinom              | rbinom             | Binomial |
| qpois    | ppois                      | dpois               | rpois              | Poisson  |
| qno rm   | pnorm                      | dnorm               | rnorm              | Normal   |
| Quantile | Cumulative<br>Distribution | Density<br>Function | Random<br>Variates |          |

## Plotting

Also see the **ggplot2** package.

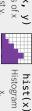


Dates

See the **lubridate** package.

## plot(x)







 $RStudio^{\bullet} is a trademark of RStudio, Inc. \bullet \underline{CCBY} Mhairi McNeill \bullet mhairihmcneill@gmail.com \bullet 844-448-1212 \bullet \underline{rstudio.com}$