Cheat Sheet Base R

Getting Help

help.search('weighted mean') Get help of a particular function.

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

Find help for a package.

More about an object

str(iris)

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

Find the class an object belongs to.

Using Libraries

install.packages('dplyr')

Download and install a package from CRAN

library(dplyr)

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

dplyr: select

Use a particular function from a package.

data(iris)

Load a build-in dataset into the environment.

Working Directory

getwd()

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

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

Vectors

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

Vector Functions

sort(x)

table(x) Return x sorted.

unique(x) Return x reversed.

See counts of values.

By Position

×[4]

×[-4]

x[2:4]

x[-(2:4)]

two to four.

x[c(1, 5)]

By Value

are equal to 10.

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

Element with

for (i in 1:4){

For Loop

Programming

for (variable in sequence){ Do something

while (condition) {

hile Loop

Do something

Example

while (i < 5){

Example

i <- i + 1

j <- i + 10 print(j)

rev(x)

See unique values

Selecting Vector Elements

The fourth element.

All but the fourth.

Elements two to four

All elements except

Elements one and

x[x == 10]Elements which

x[x < 0]All elements less than zero.

Elements in the set

Named Vectors

x['apple']

name 'apple'

f Statements

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

Example

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

function_name <- function(var){</pre> return(new_variable) Do something

Example

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

Reading and Writing Data

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

| a != b | a == b |
|-----------------------|--------------------------|
| Not equal | Are equal |
| a < b | a > b |
| Less than | Greater than |
| a <= b | a >= b |
| Less than or equal to | Greater than or equal to |
| is.null(a) | is.na(a) |
| ls null | Is missing |
| | |

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', 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). |

Maths Functions

| 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 quantiles. |
| round(x, n) | Round to n decimal places. | rank(x) | Rank of elements. |
| sig.fig(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'

ω

2

Н

The Environment

| <pre>rm(list = ls())</pre> | | 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 Studio to browse variables in your environment.

df[2, 2]

environment.

<- matrix(\times , nrow = 3, ncol = Create a matrix from x ω

3

m[2, m[, 1] - Select a column] - Select a row

m %*% n Transpose t(m)

Matrix Multiplication solve(m, n)

Find x in: m * x = n

Lists

m[2, 3] - Select an element

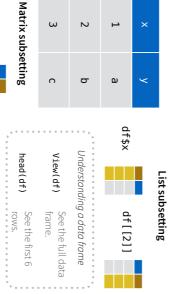
A list is collection of elements which can be of different types. $l \leftarrow list(x = 1:5, y = c('a',$ 'b'))

Second element ו[[2]] only the first New list with element. [1] Element named l\$x only element New list with named y. l['y']

dplyr library. 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.



Also see the **stringr** library.

paste(x, collapse = ' ') paste(x, y, sep = ' ')Join elements of a vector together. Join multiple vectors together.

gsub(pattern, replace, x) grep(pattern, x) Find regular expression matches in x

toupper(x) Replace matches in x with a string.

Convert to lowercase. Convert to uppercase.

tolower(x)

nchar(x)

Number of characters in a string.

Factors

factor(x)

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

> Turn a numeric vector into a cut(x, breaks = 4)factor but 'cutting' into

sections.

Statistics

 $lm(x \sim y, data=df)$ Linear model

difference between

difference

means.

proportions

between

Preform a t-test for t.test(x, y)

prop.test

Test for a

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

Get more detailed information out a model. summary

pairwise.t.test

Preform a t-test for

paired data.

Analysis of

aov

variance.

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

> Uniform Binomial

> > runif rbinom

dunif

qunif qbinom

dbinom dpois

pbinom punif

ppois pnorm

qpois

Normal

rnorm rpois

dnorm

qnorm

Variates Random

Hunction Density

Cumulative Distribution

Quantile

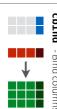
Distributions

Poison

df[, 2]

columns and Number of dim(df) columns. ncol(df Number of

df[2,



Plotting

Also see the **ggplot2** library.

rbind - Bind rows **+**

order.

plot(x)
Values of x in

against y.



Histogram of hist(x)