

Base R

Folha de referência

Obtendo ajuda

Acessando ajuda

?mean

Obtenha ajuda para uma função específica

help.search('weighted mean')

Busque palavra ou frase nos arquivos de ajuda

help(package = 'dplyr')

Encontre a Ajuda de um pacote

Mais sobre um objeto

str(iris)

Veja o sumário da estrutura de um objeto

class(iris)

Veja a que classe um objeto pertence

Usando pacotes

install.packages('dplyr')

Baixe e instale um pacote do CRAN

library(dplyr)

Carregue um pacote na sessão atual, deixando todas as suas funções acessíveis a uso

dplyr::select

Use uma função específica de dado pacote

data(iris)

Carregue um conjunto de dados pré-definido no ambiente de trabalho

Diretório de trabalho

getwd()

Obtenha o diretório de trabalho (de onde dados são carregados e resultados são salvos)

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

Mude o diretório de trabalho

Use projetos em RStudio para definir o diretório de trabalho a pasta que você está trabalhando

Vetores

Criando vetores

c(2, 4, 6)	2 4 6	Junte elementos em um vetor
2:6	2 3 4 5 6	Uma sequência de integrais
seq(2, 3, by=0.5)	2.0 2.5 3.0	Uma sequência complexa
rep(1:2, times=3)	1 2 1 2 1 2	Repetir um vetor
rep(1:2, each=3)	1 1 1 2 2 2	Repetir elementos de um vetor

Funções de vetores

sort(x) Ordena x	rev(x) Reverte x
table(x) Conta valores	unique(x) Valores únicos de x

Selecionando elementos de vetores

Por Posição

x[4]	O quarto elemento
x[-4]	Todos exceto o quarto
x[2:4]	Os elementos segundo ao quarto
x[-(2:4)]	Todos elementos exceto o segundo e o quarto
x[c(1, 5)]	Os elementos primeiro e quinto

Por Valor

x[x == 10]	Os elementos iguais a 10
x[x < 0]	Os elementos menores que 0
x[x %in% c(1, 2, 5)]	Os elementos do conjunto 1,2,5

Vetores Nomeados

x['apple']	O elemento de nome "apple"
-------------------	----------------------------

Programando

For Loop

```
for (variável in sequência) {
  Fazer algo
}
```

Exemplo

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

While Loop

```
while (condição) {
  Fazer algo
}
```

Exemplo

```
while ( i < 5 ) {
  print(i)
  i <- i + 1
}
```

Declarações if

```
if (condição) {
  Fazer algo
} else {
  Fazer algo diferente
}
```

Exemplo

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

Funções

```
nome_da_funcao <- function(var) {
  Fazer algo
  return(nova_var)
}
```

Exemplo

```
square <- function(x){
  squared <- x*x
  return(squared)
}
```

Lendo e escrevendo dados

Veja também o pacote **readr**

Entrada	Saída	Descrição
df <- read.table('file.txt')	write.table(df, 'file.txt')	Ler e escrever um arquivo de texto delimitado
df <- read.csv('file.csv')	write.csv(df, 'file.csv')	Ler e escrever um arquivo com separações por vírgula. Este é um caso especial de read.table/write.table
load('file.RData')	save(df, file = 'file.Rdata')	Ler e escrever um arquivo tipo R data (especial do R)

Condições	a == b	Igual a	a > b	Maior que	a >= b	Maior que ou igual a	is.na(a)	É faltante
a != b	Não é igual a	a < b	Menor que	a <= b	Menor que ou igual a	is.null(a)	É nulo	

Tipos de variáveis

Convertendo tipos comuns de dados em R.
Sempre possível do valor superior ao inferior da tabela

as.logical	TRUE, FALSE, TRUE	Valores booleanos (TRUE ou FALSE)
as.numeric	1, 0, 1	Inteiros ou números em ponto flutuantes
as.character	'1', '0', '1'	Sequência de caracteres. Geralmente preferidos a fatores
as.factor	'1', '0', '1', levels: '1', '0'	Sequência de caracteres com níveis. Necessário para alguns modelos estatísticos

Funções matemáticas

log(x)	Log natural	sum(x)	Soma
exp(x)	Exponencial	mean(x)	Média
max(x)	Elemento maior	median(x)	Mediana
min(x)	Elemento menor	quantile(x)	Quantis
round(x, n)	Arredondar a n unidades decimais	rank(x)	Ranque de elementos
signif(x, n)	Arredondar a n unidades decimais significativas	var(x)	Variância
cor(x, y)	Correlação	sd(x)	Desvio padrão

Designar variável

```
> a <- 'apple'
> a
[1] 'apple'
```

O ambiente de R

ls()	Lista todos os elementos presentes no ambiente
rm(x)	Remove o elemento x do ambiente
rm(list = ls())	Remove todos os elementos presentes no ambiente

Você pode usar o painel em RStudio para navegar pelas variáveis em seu ambiente de trabalho

Matrizes

```
m <- matrix(x, nrow = 3, ncol = 3)
```

Cria uma matriz desde x

	m[2,]	Seleciona uma linha		m[, 1]	Seleciona uma coluna
	m[2, 3]	Seleciona um elemento		t(m)	Transpor
				m %*% n	Multiplica matrizes
				solve(m, n)	Encontra x em: m * x = n

Listas

```
l <- list(x = 1:5, y = c('a', 'b'))
```

Uma lista é uma coleção de elementos que podem ser de diferentes tipos

2º elemento da lista l	Nova lista só com o 1º elemento	Elemento de nome x	Nova lista só com o elemento y

Veja também o pacote **dplyr**

Data Frames

```
df <- data.frame(x = 1:3, y = c('a', 'b', 'c'))
```

Um caso especial de lista onde todos elementos tem mesmo comprimento

x	y
1	a
2	b
3	c

Subconjunto de lista

df\$x		df[[2]]	
-------	--	---------	--

Entendendo um data.frame

View(df)	Ver o data.frame inteiro
head(df)	Ver as primeiras seis linhas

Subconjunto de matriz

df[, 2]		nrow(df)	Número de linhas
df[2,]		ncol(df)	Número de colunas
df[2, 2]		dim(df)	Número de linhas e colunas

cbind

	Colar colunas	
--	---------------	--

rbind

	Colar linhas	
--	--------------	--

sequência de caracteres

Veja também o pacote **stringr**

paste(x, y, sep = ' ')	Junta múltiplos vetores
paste(x, collapse = ' ')	Junta elementos de um único vetor
grep(pattern, x)	Encontra expressão regular em x
gsub(pattern, replace, x)	Substitui expressão regular em x por sequência de caracteres
toupper(x)	Converte a maiúsculas
tolower(x)	Converte a minúsculas
nchar(x)	Número de caracteres numa sequência

Fatores

factor(x)	Transforma vetor numérico em categórico (fator), podendo definir níveis e ordem do último
cut(x, breaks = 4)	Transforma vetor numérico em categórico (fator) ao "partir" o primeiro em porções

Estatística

lm(y ~ x, data=df)	Modelo linear	t.test(x, y)	Conduz teste t para diferença entre médias	prop.test	Testa diferença entre proporções
glm(y ~ x, data=df)	Modelo linear generalizado	pairwise.t.test	Conduz teste t para dados pareados	aov	Análise de variância
summary(modelo)	Obtém informação detalhada do modelo				

Distribuições

	Variáveis aleatórias	Funções de densidade	Distribuição cumulativa	Quantis
Normal	rnorm	dnorm	pnorm	qnorm
Poisson	rpois	dpois	ppois	qpois
Binomial	rbinom	dbinom	pbinom	qbinom
Uniforme	runif	dunif	punif	qunif

Plotar

Veja também o pacote **ggplot2**

	plot(x)	Valores de x em ordem		plot(x, y)	Valores de x contra y		hist(x)	Histograma de x
--	---------	-----------------------	--	------------	-----------------------	--	---------	-----------------

Datas

Veja também o pacote **lubridate**