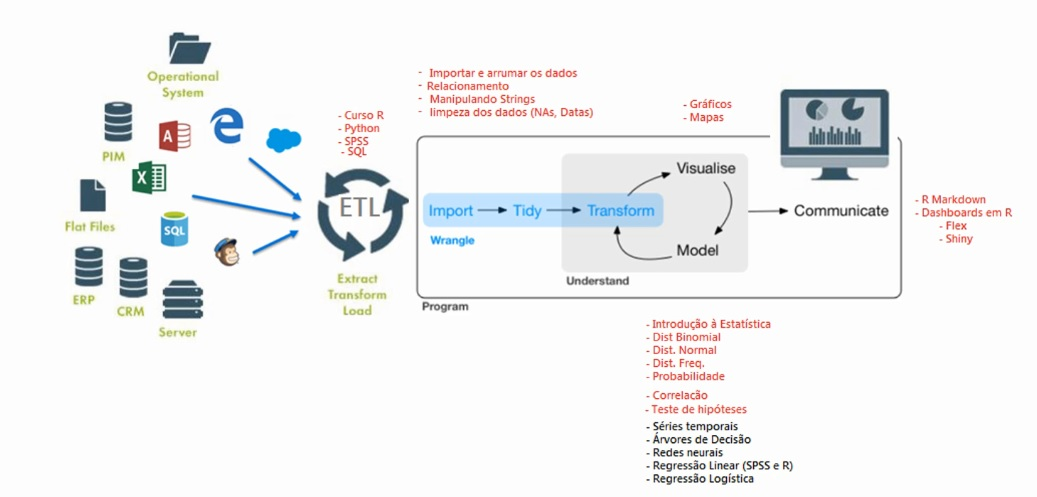
Linguagem R

Visão para Cientista de Dados



Vetores

|  |
| --- |
| # vetores  > idade <- c(10,20,40,50)  > idade  [1] 10 20 40 50  > idade2 <- scan()  1: 10  2: 20  3: 40  4: 50  5:  Read 4 items  > idade2  [1] 10 20 40 50  > length(idade)  [1] 4 |
|  |
| |  | | --- | | nome <- c("jose","Maria","Joao")  > nome  [1] "jose" "Maria" "Joao"  > length(nome)  [1] 3 | |

idade

[1] 10 20 40 50

> idade2

[1] 10 20 40 50

> nome

[1] "jose" "Maria" "Joao"

> idade[1]<- 14

> idade

[1] 14 20 40 50

> idade[1:3]<-20:22

> idade[1]

[1] 20

idade

[1] 20 21 22 50

idade

[1] 20 21 22 50

> nome

[1] "jose" "Maria" "Joao"

> idade[5]<-20

> idade

[1] 20 21 22 50 20

> idade[6:10]<- c(90,40,50,30,45)

> idade

[1] 20 21 22 50 20 90 40 50 30 45

> nome

[1] "jose" "Maria" "Joao"

> nome[1]<-"Wellington"

> nome[]

[1] "Wellington" "Maria" "Joao"

match("Joao", nome)

[1] 3

# Localiza em qual posição (índice) do vetor está o elemento "Joao".

idade

[1] 20 21 22 50 20 90 40 50 30 45

> xidade <- idade[-3]

> xidade

[1] 20 21 50 20 90 40 50 30 45

|  |
| --- |
| idade  [1] 20 21 22 50 20 90 40 50 30 45  > xidade <- idade[-3]  > xidade  [1] 20 21 50 20 90 40 50 30 45  > idade  [1] 20 21 22 50 20 90 40 50 30 45  > idade2  [1] 10 20 40 50  > idade2 <- idade+10  > idade2  [1] 30 31 32 60 30 100 50 60 40 55  > idade3 <- idade2-idade  > idade3  [1] 10 10 10 10 10 10 10 10 10 10  > idade4 <- idade2+idade  > idade4  [1] 50 52 54 110 50 190 90 110 70 100 |
|  |
| |  | | --- | | > | |

Sequência

1:8

[1] 1 2 3 4 5 6 7 8

> 3:10

[1] 3 4 5 6 7 8 9 10

seq(2,10)

[1] 2 3 4 5 6 7 8 9 10

seq(1,20,3) passo de 3

[1] 1 4 7 10 13 16 19

c <- seq(1,100,5)

length(c)

[1] 20

# Repetição

rep(5,10)

[1] 5 5 5 5 5 5 5 5 5 5

rep("Maria",20)

[1] "Maria" "Maria" "Maria" "Maria" "Maria" "Maria" "Maria" "Maria" "Maria" "Maria"

[11] "Maria" "Maria" "Maria" "Maria" "Maria" "Maria" "Maria" "Maria" "Maria" "Maria"

rep(1:4,3)

[1] 1 2 3 4 1 2 3 4 1 2 3 4

idades <- c(10,20,30)

rep(idades,3)

[1] 10 20 30 10 20 30 10 20 30

rep(1:4,3,each=2)

[1] 1 1 2 2 3 3 4 4 1 1 2 2 3 3 4 4 1 1 2 2 3 3 4 4

cidades <- c("Brasilia","Rio","São Paulo")

rep(cidades,2)

[1] "Brasilia" "Rio" "São Paulo" "Brasilia" "Rio" "São Paulo"

rep(cidades,3,each=2)

[1] "Brasilia" "Brasilia" "Rio" "Rio" "São Paulo" "São Paulo"

[7] "Brasilia" "Brasilia" "Rio" "Rio" "São Paulo" "São Paulo"

[13] "Brasilia" "Brasilia" "Rio" "Rio" "São Paulo" "São Paulo"

rep(cidades,c(1,2,3))

[1] "Brasilia" "Rio" "Rio" "São Paulo" "São Paulo" "São Paulo"

Nr Aleatórios

|  |
| --- |
| sample(1:60,6)  [1] 56 7 32 37 55 50  sample(1:60,6)  [1] 48 25 38 29 49 6 |
|  |
| |  | | --- | | sample(1:10,5,replace=TRUE)  [1] 8 4 4 6 1  sample(1:10,5,replace=TRUE)  [1] 8 4 2 7 4 | |

cidades

[1] "Brasilia" "Rio" "São Paulo"

sort(cidades)

[1] "Brasilia" "Rio" "São Paulo"

> sort(cidades)

[1] "Brasilia" "Rio" "São Paulo"

>

> sort(cidades, decreasing=TRUE)

[1] "São Paulo" "Rio" "Brasilia"

|  |
| --- |
| # Maior numero – maior rank  numero<-c(6,5,8,20)  rank(numero)  [1] 2 1 3 4  # Matrizes  v1<- c(10,20,30)  v2 <- c(20,30,40)  rbind(v1,v2)  [,1] [,2] [,3]  v1 10 20 30  v2 20 30 40  # r - row  > rbind(v1,v2)  [,1] [,2] [,3]  v1 10 20 30  v2 20 30 40  > # c - colunas  > cbind(v1,v2)  v1 v2  [1,] 10 20  [2,] 20 30  [3,] 30 40  a <- rbind(v1,v2)  b <- cbind(v1,v2)  > a  [,1] [,2] [,3]  v1 10 20 30  v2 20 30 40  # 2 linhas e 3 colunas  dim(a)  [1] 2 3  > b  v1 v2  [1,] 10 20  [2,] 20 30  [3,] 30 40  # 3 linhas e 2 colunas  dim(b)  [1] 3 2  matricula <- c(10,20,30)  nome <- c("Joao","Maria","Patricia")  >  cbind(matricula,nome)  matricula nome  [1,] "10" "Joao"  [2,] "20" "Maria"  [3,] "30" "Patricia"  rbind(matricula, nome)  [,1] [,2] [,3]  matricula "10" "20" "30"  nome "Joao" "Maria" "Patricia"  matrix(1:20,nrow=3,ncol=5,byrow=TRUE)  [,1] [,2] [,3] [,4] [,5]  [1,] 1 2 3 4 5  [2,] 6 7 8 9 10  [3,] 11 12 13 14 15  matrix(1:20,nrow=3,ncol=5,byrow=FALSE)  [,1] [,2] [,3] [,4] [,5]  [1,] 1 4 7 10 13  [2,] 2 5 8 11 14  [3,] 3 6 9 12 15 |
|  |
| |  | | --- | |  | |

matrix(c("a","b","c","d","e"),nrow=3,ncol=5,byrow=FALSE)

[,1] [,2] [,3] [,4] [,5]

[1,] "a" "d" "b" "e" "c"

[2,] "b" "e" "c" "a" "d"

[3,] "c" "a" "d" "b" "e"

> matrix(c("a","b","c","d","e"),nrow=3,ncol=5,byrow=TRUE)

[,1] [,2] [,3] [,4] [,5]

[1,] "a" "b" "c" "d" "e"

[2,] "a" "b" "c" "d" "e"

[3,] "a" "b" "c" "d" "e"

matA <- matrix(c("a","b","c","d","e"),nrow=3,ncol=5,byrow=FALSE)

matA

[,1] [,2] [,3] [,4] [,5]

[1,] "a" "d" "b" "e" "c"

[2,] "b" "e" "c" "a" "d"

[3,] "c" "a" "d" "b" "e"

matA[2,5]

[1] "d"

matA[2,]

[1] "b" "e" "c" "a" "d"

matA[,3]

[1] "b" "c" "d"

matA[1:2,4:5]

[,1] [,2]

[1,] "e" "c"

[2,] "a" "d"

ncol(matA)

[1] 5

nrow(matA)

[1] 3

dim(matA)

[1] 3 5

Modificação matriz

matA

[,1] [,2] [,3] [,4] [,5]

[1,] "a" "d" "b" "e" "c"

[2,] "b" "e" "c" "a" "d"

[3,] "c" "a" "d" "b" "e"

matA[1,3] <- "Well"

matA

[,1] [,2] [,3] [,4] [,5]

[1,] "a" "d" "Well" "e" "c"

[2,] "b" "e" "c" "a" "d"

[3,] "c" "a" "d" "b" "e"

matA[,4]<- c(10,20,30)

matA

[,1] [,2] [,3] [,4] [,5]

[1,] "a" "d" "Well" "10" "c"

[2,] "b" "e" "c" "20" "d"

[3,] "c" "a" "d" "30" "e"

Acrescentar uma coluna

matA

[,1] [,2] [,3] [,4] [,5]

[1,] "a" "d" "Well" "10" "c"

[2,] "b" "e" "c" "20" "d"

[3,] "c" "a" "d" "30" "e"

a<-c(10,20,30)

b <- cbind(matA,a)

b

a

[1,] "a" "d" "Well" "10" "c" "10"

[2,] "b" "e" "c" "20" "d" "20"

[3,] "c" "a" "d" "30" "e" "30"

matA

[,1] [,2] [,3] [,4] [,5]

[1,] "a" "d" "Well" "10" "c"

[2,] "b" "e" "c" "20" "d"

[3,] "c" "a" "d" "30" "e"

linha<-c(10,20,30,40,50)

R2<-rbind(matA,linha)

R2

[,1] [,2] [,3] [,4] [,5]

"a" "d" "Well" "10" "c"

"b" "e" "c" "20" "d"

"c" "a" "d" "30" "e"

linha "10" "20" "30" "40" "50"

R2

[,1] [,2] [,3] [,4] [,5]

"a" "d" "Well" "10" "c"

"b" "e" "c" "20" "d"

"c" "a" "d" "30" "e"

linha "10" "20" "30" "40" "50"

R2[-3,]

[,1] [,2] [,3] [,4] [,5]

"a" "d" "Well" "10" "c"

"b" "e" "c" "20" "d"

linha "10" "20" "30" "40" "50"

A<- matrix(0,5,5)

A

[,1] [,2] [,3] [,4] [,5]

[1,] 0 0 0 0 0

[2,] 0 0 0 0 0

[3,] 0 0 0 0 0

[4,] 0 0 0 0 0

[5,] 0 0 0 0 0

A+2

[,1] [,2] [,3] [,4] [,5]

[1,] 2 2 2 2 2

[2,] 2 2 2 2 2

[3,] 2 2 2 2 2

[4,] 2 2 2 2 2

[5,] 2 2 2 2 2

A-1

[,1] [,2] [,3] [,4] [,5]

[1,] -1 -1 -1 -1 -1

[2,] -1 -1 -1 -1 -1

[3,] -1 -1 -1 -1 -1

[4,] -1 -1 -1 -1 -1

[5,] -1 -1 -1 -1 -1

Adicionando nome para colunas e linhas

|  |
| --- |
| ncolunas<-c('r1','r2','r3','r4','r5')  nlinhas<-c('matricula','nome','endereco','telefone','sexo')  ncolunas  [1] "r1" "r2" "r3" "r4" "r5"  nlinhas  [1] "matricula" "nome" "endereco" "telefone" "sexo"  Dimnames= list(linha, coluna)    R2<-matrix(0,5,5,dimnames=list(nlinhas,ncolunas))  R2  r1 r2 r3 r4 r5  matricula 0 0 0 0 0  nome 0 0 0 0 0  endereco 0 0 0 0 0  telefone 0 0 0 0 0  sexo 0 0 0 0 0  R2<-matrix(0,5,5,dimnames=list(ncolunas,nlinhas))  R2  matricula nome endereco telefone sexo  r1 0 0 0 0 0  r2 0 0 0 0 0  r3 0 0 0 0 0  r4 0 0 0 0 0  r5 0 0 0 0 0  Determinante da Matrix Quadrada  A<-matrix(1:25,5,5)  > A  [,1] [,2] [,3] [,4] [,5]  [1,] 1 6 11 16 21  [2,] 2 7 12 17 22  [3,] 3 8 13 18 23  [4,] 4 9 14 19 24  [5,] 5 10 15 20 25  > det(A)  [1] 0  Diagonal  diag(A)  [1] 1 7 13 19 25 |
|  |
| |  | | --- | | > | |

Matrix Transporta – Cria uma nova matrix

Linha transforma em coluna e coluna transforma em linha

> A

[,1] [,2] [,3] [,4] [,5]

[1,] 1 6 11 16 21

[2,] 2 7 12 17 22

[3,] 3 8 13 18 23

[4,] 4 9 14 19 24

[5,] 5 10 15 20 25

> B<- t(A)

> B

[,1] [,2] [,3] [,4] [,5]

[1,] 1 2 3 4 5

[2,] 6 7 8 9 10

[3,] 11 12 13 14 15

[4,] 16 17 18 19 20

[5,] 21 22 23 24 25

Imagem da matrix em pixel

> B

[,1] [,2] [,3] [,4] [,5]

[1,] 1 2 3 4 5

[2,] 6 7 8 9 10

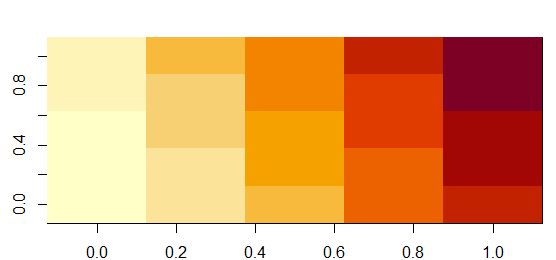
[3,] 11 12 13 14 15

[4,] 16 17 18 19 20

[5,] 21 22 23 24 25

> image(A)

> image(B)



Curva de Nível

B

[,1] [,2] [,3] [,4] [,5]

[1,] 1 2 3 4 5

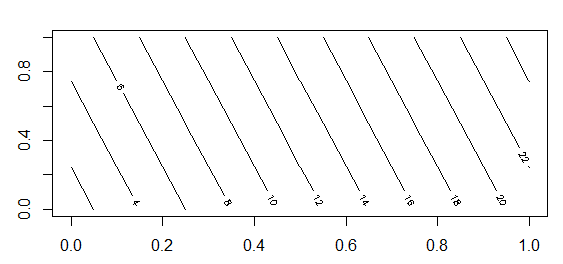
[2,] 6 7 8 9 10

[3,] 11 12 13 14 15

[4,] 16 17 18 19 20

[5,] 21 22 23 24 25

> contour(B)



Perspectiva

B

[,1] [,2] [,3] [,4] [,5]

[1,] 1 2 3 4 5

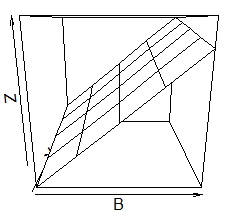
[2,] 6 7 8 9 10

[3,] 11 12 13 14 15

[4,] 16 17 18 19 20

[5,] 21 22 23 24 25

> persp(B)



ARRAY

Array é uma coleção de matrizes

A <- m1,m2,m3,m4

Criar array com 4 matrizes de 0 com 5 linhas x 5 colunas c(5,5,4)

a <- array(0,dim=c(5,5,4))

> a

, , 1

[,1] [,2] [,3] [,4] [,5]

[1,] 0 0 0 0 0

[2,] 0 0 0 0 0

[3,] 0 0 0 0 0

[4,] 0 0 0 0 0

[5,] 0 0 0 0 0

, , 2

[,1] [,2] [,3] [,4] [,5]

[1,] 0 0 0 0 0

[2,] 0 0 0 0 0

[3,] 0 0 0 0 0

[4,] 0 0 0 0 0

[5,] 0 0 0 0 0

, , 3

[,1] [,2] [,3] [,4] [,5]

[1,] 0 0 0 0 0

[2,] 0 0 0 0 0

[3,] 0 0 0 0 0

[4,] 0 0 0 0 0

[5,] 0 0 0 0 0

, , 4

[,1] [,2] [,3] [,4] [,5]

[1,] 0 0 0 0 0

[2,] 0 0 0 0 0

[3,] 0 0 0 0 0

[4,] 0 0 0 0 0

[5,] 0 0 0 0 0

B <- array(c("Vaneza","Wellington","Hugo","Pedro"),dim=c(2,4,3))

> B

, , 1

[,1] [,2] [,3] [,4]

[1,] "Vaneza" "Hugo" "Vaneza" "Hugo"

[2,] "Wellington" "Pedro" "Wellington" "Pedro"

, , 2

[,1] [,2] [,3] [,4]

[1,] "Vaneza" "Hugo" "Vaneza" "Hugo"

[2,] "Wellington" "Pedro" "Wellington" "Pedro"

, , 3

[,1] [,2] [,3] [,4]

[1,] "Vaneza" "Hugo" "Vaneza" "Hugo"

[2,] "Wellington" "Pedro" "Wellington" "Pedro"

dim(B)

[1] 2 4 3

v1 <- c("l1","l2","l3")

> v2 <- c(10,20,30)

> Ar01 <- array(c(v1,v2),dim=c(3,2,2))

> Ar01

, , 1

[,1] [,2]

[1,] "l1" "10"

[2,] "l2" "20"

[3,] "l3" "30"

, , 2

[,1] [,2]

[1,] "l1" "10"

[2,] "l2" "20"

[3,] "l3" "30"

> v1 <- c(10,20,40)

> v2 <- c(40,50,70)

> v3 <- c(60,40,80)

> r.names <- c("transporte","comida","estudo")

> c.names <- c("Joao","Maria","Carlos")

> matrizes.names <- c("Janeiro","Fevereiro","Março")

> dados <-array(c(v1,v2,v3),dim=c(3,3,3),dimnames=list(r.names,c.names,matrizes.names))

> dados

, , Janeiro

Joao Maria Carlos

transporte 10 40 60

comida 20 50 40

estudo 40 70 80

, , Fevereiro

Joao Maria Carlos

transporte 10 40 60

comida 20 50 40

estudo 40 70 80

, , Março

Joao Maria Carlos

transporte 10 40 60

comida 20 50 40

estudo 40 70 80

Nova forma

dados <-array(c(v1,v2,v3),dim=c(3,3,3),dimnames=list(c.names,r.names,matrizes.names))

> dados

, , Janeiro

transporte comida estudo

Joao 10 40 60

Maria 20 50 40

Carlos 40 70 80

, , Fevereiro

transporte comida estudo

Joao 10 40 60

Maria 20 50 40

Carlos 40 70 80

, , Março

transporte comida estudo

Joao 10 40 60

Maria 20 50 40

Carlos 40 70 80

Maria – gasto em todos meses

Linha 2, todas as colunas, todas as matrizes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| dados[2,,]  Janeiro Fevereiro Março  transporte 20 20 20  comida 50 50 50  estudo 40 40 40  gasto\_Maria <- dados[2,,]  > gasto\_Maria  Janeiro Fevereiro Março  transporte 20 20 20  comida 50 50 50  estudo 40 40 40  gasto\_maria\_alimentacao <- dados[2,2,]  > gasto\_maria\_alimentacao  Janeiro Fevereiro Março  50 50 50  # Alteração  > dados  , , Janeiro  transporte comida estudo  Joao 10 40 60  Maria 20 50 40  Carlos 40 70 80  , , Fevereiro  transporte comida estudo  Joao 10 40 60  Maria 20 50 40  Carlos 40 70 80  , , Março  transporte comida estudo  Joao 10 40 60  Maria 20 50 40  Carlos 40 70 80  > # alterar Maria - janeiro - estudo para 70  > dados[2,3,1]  [1] 40  > dados[2,3,1] <- 70  > dados[2,3,1]  [1] 70   |  | | --- | | Dados  , , Janeiro  transporte comida estudo  Joao 10 40 60  Maria 20 50 70  Carlos 40 70 80  , , Fevereiro  transporte comida estudo  Joao 10 40 60  Maria 20 50 40  Carlos 40 70 80  , , Março  transporte comida estudo  Joao 10 40 60  Maria 20 50 40  Carlos 40 70 80  > # Alterar dados de Joao - Janeiro - 40 60 80  > dados[1,,1]<- c(40,60,80)  > dados[,,1]  transporte comida estudo  Joao 40 60 80  Maria 20 50 70  Carlos 40 70 80 | |  | | |  | | --- | |  | |   # Somente a fevereiro e março - matriz 2 e 3  > dados[,,2:3]  , , Fevereiro  transporte comida estudo  Joao 10 40 60  Maria 20 50 40  Carlos 40 70 80  , , Março  transporte comida estudo  Joao 10 40 60  Maria 20 50 40  Carlos 40 70 80  Soma  ## Joao  > dados[1,,]  Janeiro Fevereiro Março  transporte 40 10 10  comida 60 40 40  estudo 80 60 60  Soma de Joao  > sum(dados[1,,])  [1] 400  ## transporte do Joao  > joao <- dados[1,,]  > joao  Janeiro Fevereiro Março  transporte 40 10 10  comida 60 40 40  estudo 80 60 60    > transporte\_joao <- joao[1,]  Média  > media\_trans\_joao <- mean(transporte\_joao)  > media\_trans\_joao  [1] 20  dados  , , Janeiro  transporte comida estudo  Joao 40 60 80  Maria 20 50 70  Carlos 40 70 80  , , Fevereiro  transporte comida estudo  Joao 10 40 60  Maria 20 50 40  Carlos 40 70 80  , , Março  transporte comida estudo  Joao 10 40 60  Maria 20 50 40  Carlos 40 70 80  > # soma do Joao nos 3 meses  > sum(dados[1,,])  [1] 400 |
| # aplicar uma função para toda planilha  > apply(dados, c(1), sum)  Joao Maria Carlos  400 360 570  apply(dados,c(1), mean)  Joao Maria Carlos  44.44444 40.00000 63.33333   |  | | --- | | dados  , , Janeiro  transporte comida estudo  Joao 40 60 80  Maria 20 50 70  Carlos 40 70 80  , , Fevereiro  transporte comida estudo  Joao 10 40 60  Maria 20 50 40  Carlos 40 70 80  , , Março  transporte comida estudo  Joao 10 40 60  Maria 20 50 40  Carlos 40 70 80  > dados[1,,]  Janeiro Fevereiro Março  transporte 40 10 10  comida 60 40 40  estudo 80 60 60  > dados[,2,]  Janeiro Fevereiro Março  Joao 60 40 40  Maria 50 50 50  Carlos 70 70 70  > # media de comida  > mean(dados[,2,])  [1] 55.55556  dados  , , Janeiro  transporte comida estudo  Joao 40 60 80  Maria 20 50 70  Carlos 40 70 80  , , Fevereiro  transporte comida estudo  Joao 10 40 60  Maria 20 50 40  Carlos 40 70 80  , , Março  transporte comida estudo  Joao 10 40 60  Maria 20 50 40  Carlos 40 70 80  > media\_transporte <- apply(dados,c(1),mean)  > media\_transporte  Joao Maria Carlos  44.44444 40.00000 63.33333 | | dados  , , Janeiro  transporte comida estudo  Joao 40 60 80  Maria 20 50 70  Carlos 40 70 80  , , Fevereiro  transporte comida estudo  Joao 10 40 60  Maria 20 50 40  Carlos 40 70 80  , , Março  transporte comida estudo  Joao 10 40 60  Maria 20 50 40  Carlos 40 70 80  > dados[1,,]  Janeiro Fevereiro Março  transporte 40 10 10  comida 60 40 40  estudo 80 60 60  > apply(dados,c(1),sum)  Joao Maria Carlos  400 360 570  > apply(dados[1,,],c(1), sum)  transporte comida estudo  60 140 200 | |  |   apply(dados[1,,],c(1), mean)  transporte comida estudo  20.00000 46.66667 66.66667  Gasto médio com comida  dados[,2,]  Janeiro Fevereiro Março  Joao 60 40 40  Maria 50 50 50  Carlos 70 70 70  > apply(dados[,2,],c(1),mean)  Joao Maria Carlos  46.66667 50.00000 70.00000 |
| |  | | --- | |  | |

FATORES

Variável quantitativa por fator?

Ex.: pesos por Estado Civil

Nível

Lista

Guarda uma coleção de objetos diferentes.

a <- list(4,c("joao","marcos", "maria"),matrix(0,4,5),mean)

> a

[[1]]

[1] 4

[[2]]

[1] "joao" "marcos" "maria"

...

> idades <- c(30,40,55)

> genero <- c("M","M","F")

> cidades <- c("BH","RJ","SP")

> r.names <- c("joao","maria","jose")

> c.names <- c("Alimentacao","Transporte","Moradia")

> m.names <- c("jan","fev","mar")

> gastos <- array(40:100, dim=c(3,3,3),dimnames= list(r.names,c.names,m.names))

> gastos

, , jan

Alimentacao Transporte Moradia

joao 40 43 46

maria 41 44 47

jose 42 45 48

, , fev

Alimentacao Transporte Moradia

joao 49 52 55

maria 50 53 56

jose 51 54 57

, , mar

Alimentacao Transporte Moradia

joao 58 61 64

maria 59 62 65

jose 60 63 66

lista01 <- list(idades,genero,cidades,gastos)

> lista01

[[1]]

[1] 30 40 55

[[2]]

[1] "M" "M" "F"

[[3]]

[1] "BH" "RJ" "SP"

[[4]]

, , jan

Alimentacao Transporte Moradia

joao 40 43 46

maria 41 44 47

jose 42 45 48

, , fev

Alimentacao Transporte Moradia

joao 49 52 55

maria 50 53 56

jose 51 54 57

, , mar

Alimentacao Transporte Moradia

joao 58 61 64

maria 59 62 65

jose 60 63 66

> names(lista01) <- c(" idades"," genero"," cidades", " Gastos")

> lista01

$` idades`

[1] 30 40 55

$` genero`

[1] "M" "M" "F"

$` cidades`

[1] "BH" "RJ" "SP"

$` Gastos`

, , jan

Alimentacao Transporte Moradia

joao 40 43 46

maria 41 44 47

jose 42 45 48

, , fev

Alimentacao Transporte Moradia

joao 49 52 55

maria 50 53 56

jose 51 54 57

, , mar

Alimentacao Transporte Moradia

joao 58 61 64

maria 59 62 65

jose 60 63 66

> lista01[1]

$` idades`

[1] 30 40 55

> lista01[2]

$` genero`

[1] "M" "M" "F"

> lista01$` idades`

[1] 30 40 55

> lista01$` idades`

[1] 30 40 55

> lista01$` idades`[2]

[1] 40

Alteração

> lista01$' idades'

[1] 30 40 55

> lista01$' idades'[1] <- 35

> lista01$' idades'

[1] 35 40 55

lista01$` Gastos`[,,1]

Alimentacao Transporte Moradia

joao 40 43 46

maria 41 44 47

jose 42 45 48

> lista01$` Gastos`[,,1]<- lista01$` Gastos`[,,1]+5

> lista01$` Gastos`[,,1]

Alimentacao Transporte Moradia

joao 45 48 51

maria 46 49 52

jose 47 50 53

Adicionando e removendo

> lista01[5]

$<NA>

NULL

> lista01[5]<- c("Mensagem de Atenção")

[[5]]

[1] "Mensagem de Atenção"

> names(lista01)

[1] " idades" " genero" " cidades" " Gastos" ""

> names(lista01)[5]<-"Msg"

> names(lista01)

[1] " idades" " genero" " cidades" " Gastos" "Msg"

names(lista01)

[1] " idades" " genero" " cidades" " Gastos" "Msg"

> lista01[5]<-NULL Elimina objeto

> names(lista01)

[1] " idades" " genero" " cidades" " Gastos"

Combinação de listas

> lista02 <- list(c("2017","2018","2019"),40,50)

> names(lista02)<-c("ano","nr1","nr2")

> lista02

$ano

[1] "2017" "2018" "2019"

$nr1

[1] 40

$nr2

[1] 50

> lista03 <- c(lista01,lista02)

> lista03

$` idades`

[1] 35 40 55

$` genero`

[1] "M" "M" "F"

$` cidades`

[1] "BH" "RJ" "SP"

$` Gastos`

, , jan

Alimentacao Transporte Moradia

joao 45 48 51

maria 46 49 52

jose 47 50 53

, , fev

Alimentacao Transporte Moradia

joao 49 52 55

maria 50 53 56

jose 51 54 57

, , mar

Alimentacao Transporte Moradia

joao 58 61 64

maria 59 62 65

jose 60 63 66

$ano

[1] "2017" "2018" "2019"

$nr1

[1] 40

$nr2

[1] 50

Dataframe

# Data.frame

> nome <- c("joao","maria","marcos")

> idade <- c(23,34,50)

> genero <- c("M", "F", "M")

> dados <- data.frame(nome,idade,genero)

> dados

nome idade genero

1 joao 23 M

2 maria 34 F

3 marcos 50 M

#informar a estrutura de dados

> str(dados)

'data.frame': 3 obs. of 3 variables:

$ nome : Factor w/ 3 levels "joao","marcos",..: 1 3 2

$ idade : num 23 34 50

$ genero: Factor w/ 2 levels "F","M": 2 1 2

# Posso usar os comandos da Matrix

>

> #saber a quantidade de linhas e colunas

> dim(dados)

[1] 3 3

nrow(dados)

[1] 3

> ncol(dados)

[1] 3

Nome das colunas

colnames(dados)

[1] "nome" "idade" "genero"

names(dados)

[1] "nome" "idade" "genero"

# Resumo de todas variaveis no data.frame

> summary(dados)

nome idade genero

joao :1 Min. :23.00 F:1

marcos:1 1st Qu.:28.50 M:2

maria :1 Median :34.00

Mean :35.67

3rd Qu.:42.00

Max. :50.00

Acessando as informações do data.frame

# acessando o dataframe

>

> dados[1,]

nome idade genero

1 joao 23 M

>

> dados[3,]

nome idade genero

3 marcos 50 M

> dados[1:2,]

nome idade genero

1 joao 23 M

2 maria 34 F

>

> dados

nome idade genero

1 joao 23 M

2 maria 34 F

3 marcos 50 M

>

> dados[,2]

[1] 23 34 50

>

> dados[2,2]

[1] 34

>

> dados[1:2,2:3]

idade genero

1 23 M

2 34 F

> dados[2:3,1:2]

nome idade

2 maria 34

3 marcos 50

**1.20 Usando search(), attach() e detach()**

A função search() mostra o caminho de procura do **R**. Quando executa-se essa função o **R**mostra o .GlobalEnv, pacotes, autoloads ou data.frames e listas adicionadas com o comando attach().

Para ver qual ou quais objetos estão “atachados", digite

  >  search()

   [1] ".GlobalEnv"        "package:MASS"      "package:stats"  
   [4] "package:graphics"  "package:grDevices" "package:utils"  
   [7] "package:datasets"  "package:methods"   "Autoloads"  
  [10] "package:base"

Com isso, aparecerá o nome do objeto ou dos objetos que estão “atachados".

Quando necessitamos fazer muitas referências para um objeto, podemos minimizar a digitação, colocando o data.frame ou lista no caminho de procura, através do comando attach().

Observe que o objeto *dados*não está atachado. Se você digitar

  >  nome

Aparecerá uma mensagem de erro, alertando que o objeto não foi encontrado. Nesse caso, o correto seria utilizar o seguinte comando:

  > dados$nome

  [1] "Adilson"

Para evitar ter que digitar dados$nome utiliza-se a função attach() da seguinte maneira:

  >  attach(dados)  
  >  search()

   [1] ".GlobalEnv"        "dados"             "package:MASS"  
   [4] "package:stats"     "package:graphics"  "package:grDevices"  
   [7] "package:utils"     "package:datasets"  "package:methods"  
  [10] "Autoloads"         "package:base"

Veja que o objeto "dados"aparece agora no caminho de procura do **R**!

Isso permite referenciar os componentes de uma lista , por exemplo, simplesmente digitando

  >  nome

  [1] "Adilson"

**Cuidado!**

Se você possui objetos com o mesmo nome em diferentes listas, isso pode causar alguma confusão. Antes de começar outra análise com outros dados, não esqueça de usar a função detach() para retirar o objeto do caminho de procura:

  > detach(dados)

# attache() - Facilitar depois na digitação

>

> attach(dados)

The following objects are masked \_by\_ .GlobalEnv:

genero, idade, nome

> dados$idade

[1] 23 34 50

> dados$nome

[1] joao maria marcos

Levels: joao marcos maria

> dados$genero

[1] M F M

Levels: F M

dados$idade[2:3]

[1] 34 50

Modificando o elemento do Data.Frame

#modificando as informações no data.frame

>

> dados$idade

[1] 23 34 50

>

> dados$idade <- c(40,44,39)

>

> dados

nome idade genero

1 joao 40 M

2 maria 44 F

3 marcos 39 M

dados

nome idade genero

1 joao 40 M

2 maria 44 F

3 marcos 39 M

> dados$idade[3]

[1] 39

> dados$idade[3]<-55

> dados

nome idade genero

1 joao 40 M

2 maria 44 F

3 marcos 55 M

Adicionando e removendo variáveis

# adicionando coluna

> # cidade

> dados

nome idade genero

1 joao 40 M

2 maria 44 F

3 marcos 55 M

> dados$cidade <- c("SP","RJ","BH")

> dados

nome idade genero cidade

1 joao 40 M SP

2 maria 44 F RJ

3 marcos 55 M BH

# incluir renda

> dados$renda <- c(8000, 9000, 15000)

> dados

nome idade genero cidade renda

1 joao 40 M SP 8000

2 maria 44 F RJ 9000

3 marcos 55 M BH 15000

# removendo coluna

> dados <- dados[,-5]

> dados

nome idade genero cidade

1 joao 40 M SP

2 maria 44 F RJ

3 marcos 55 M BH

|  |
| --- |
| # adicionando registro  > dados  nome idade genero cidade  1 joao 40 M SP  2 maria 44 F RJ  3 marcos 55 M BH  >  > dados[4,] <- c("Isabela",20,"F","BSB")  Warning message:  In `[<-.factor`(`\*tmp\*`, iseq, value = "Isabela") :  invalid factor level, NA generated  > # converter nome de fator para caractere  > dados$nome <- as.character(dados$nome)  > dados  nome idade genero cidade  1 joao 40 M SP  2 maria 44 F RJ  3 marcos 55 M BH  4 <NA> 20 F BSB    > dados$nome[4]  [1] NA  > dados$nome[4]<- "isabela"  >  > dados  nome idade genero cidade  1 joao 40 M SP  2 maria 44 F RJ  3 marcos 55 M BH  4 isabela 20 F BSB |
|  |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | |  | | --- | | str(dados)  'data.frame': 4 obs. of 4 variables:  $ nome : chr "joao" "maria" "marcos" "isabela"  $ idade : chr "40" "44" "55" "20"  $ genero: Factor w/ 2 levels "F","M": 2 1 2 1  $ cidade: chr "SP" "RJ" "BH" "BSB"    > # transformar idade de char para numeric    > dados$idade <- as.numeric(dados$idade)    > str(dados)  'data.frame': 4 obs. of 4 variables:  $ nome : chr "joao" "maria" "marcos" "isabela"  $ idade : num 40 44 55 20  $ genero: Factor w/ 2 levels "F","M": 2 1 2 1  $ cidade: chr "SP" "RJ" "BH" "BSB"  dados  nome idade genero cidade  1 joao 40 M SP  2 maria 44 F RJ  3 marcos 55 M BH  4 isabela 20 F BSB  # excluir a linha 3  > dados[-3,]  nome idade genero cidade  1 joao 40 M SP  2 maria 44 F RJ  4 isabela 20 F BSB  > dados  nome idade genero cidade  1 joao 40 M SP  2 maria 44 F RJ  3 marcos 55 M BH  4 isabela 20 F BSB  > dados <- dados[-3,] Definitivo  > dados  nome idade genero cidade  1 joao 40 M SP  2 maria 44 F RJ  4 isabela 20 F BSB  Funções | |  | | |  | | --- | |  | | | |

x<- c(10,20,30,40,50,60,70,80,90,100)

for (i in 1:length(x))

{

print(x[i])

}

x<- c(10,20,30,40,50,60,70,80,90,100)

>

> for (i in 1:length(x))

+ {

+ print(x[i])

+

+ }

[1] 10

[1] 20

[1] 30

[1] 40

[1] 50

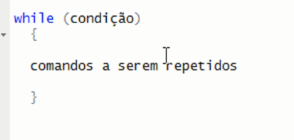
[1] 60

[1] 70

[1] 80

[1] 90

[1] 100



n<-1

> while (n<= 20)

+ {

+ print(n)

+ n <- n+1

+ }

[1] 1

[1] 2

[1] 3

[1] 4

[1] 5

[1] 6

[1] 7

[1] 8

[1] 9

[1] 10

[1] 11

[1] 12

[1] 13

[1] 14

[1] 15

[1] 16

[1] 17

[1] 18

[1] 19

[1] 20

n<-9

>

> if (n==9)

+ {

+ print(n)

+ print("vc digitou 9")

+ }

[1] 9

[1] "vc digitou 9"

Criação da Função

|  |
| --- |
| media<- function(x)  {  m<-mean(x)  cat("media - ", m,".")  }  valor <- c(10,20,30,40,50)  > media(valor)  media - 30 . |
|  |
| |  | | --- | | > | |

Aplicação Real

Bonus