

# Linguagem R

### Aula 5

- Objetivo da aula:
- Introdução ao RStudio;
- Carregando um DataFrame;
- Tipos de dados e operadores;
- Estruturas de dados;
- Funções.

## Introdução a Linguagem R:

R é uma linguagem e ambiente para computação estatística e construção de gráficos.

R fornece uma ampla variedade de técnicas estatísticas e gráficas: modelagem linear e não linear, testes estatísticos, análise de séries temporais, classificação, agrupamento, etc.



## Instalação:

Antes da instalação do RStudio é necessário instalar o interpretador R.

O download do interpretador R e do RStudio estão na mesma página.



 $\mathbb{C}$ 

https://posit.co/download/rstudio-desktop/







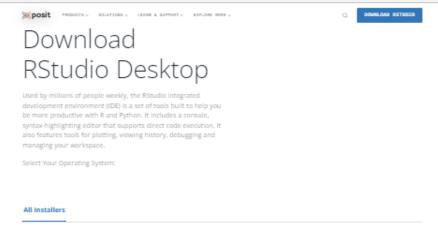












#### Step 1: Install R

RStudio requires R 3.3.0+. Choose a version of R that matches your computer's operating system.

DOWNLOAD AND INSTALL R

#### Step 2: Install RStudio Desktop

DOWNLOAD RETUDIO DESKTOP FOR WINDOWS

Size: 190.49MB | SWA-256: 82885925 | Version: 2022.07.2+576 | | Released: 2022-09-21



## Instalação do interpretador R:

Para a instalação completa do software (RStudio-IDE) é necessário a instalação do R.

Download R em: https://cran.rstudio.com de acordo com o sistema operacional do seu desktop.

## Download do intepretador R:



R

CRAN
Mirrors
What's new?
Search
CRAN Team

About R R Homepage The R Journal

Software
R Sources
R Binaries
Packages
Task Views
Other

Documentation
Manuals
FAQs
Contributed

#### The Comprehensive R Archive Network

#### Download and Install R

Precompiled binary distributions of the base system and contributed packages, Windows and Mac users most likely want one of these versions of R:

- Download R for Linux (Debian, Fedora/Redhat, Ubuntu)
- Download R for macOS
- Download R for Windows

R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

#### Source Code for all Platforms

Windows and Mac users most likely want to download the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled before you can use them. If you do not know what this means, you probably do not want to do it!

- The latest release (2022-10-31, Innocent and Trusting) R-4.2.2.tar.gz, read what's new in the latest version.
- Sources of R alpha and beta releases (daily snapshots, created only in time periods before a planned release).
- Daily snapshots of current patched and development versions are <u>available here</u>. Please read about <u>new features and bug fixes</u> before filing corresponding feature requests or bug reports.
- · Source code of older versions of R is available here.
- Contributed extension packages

#### Ouestions About R

If you have questions about R like how to download and install the software, or what the license terms are, please read
our answers to frequently asked questions before you send an email.

## Testando a Instalação do R:

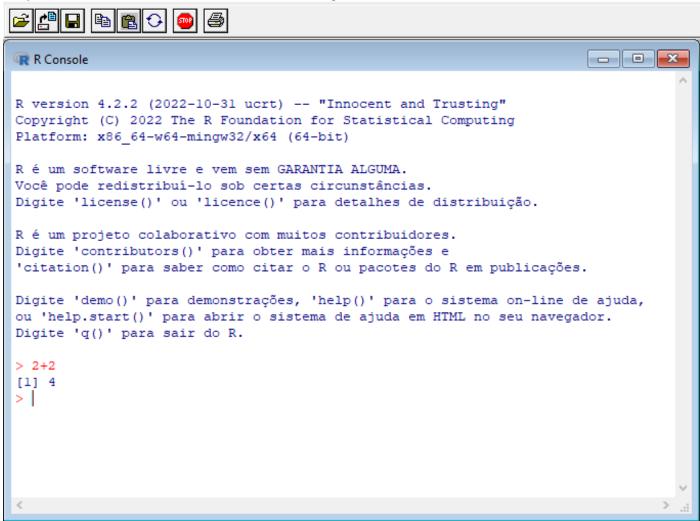
Após a instalação:

Abra o console do R e faça uma simples checagem digitando 2+2 no console.

Se a instalação estiver sido feita da maneira correta então o resultado será algo como mostrado a seguir.

#### RGui (64-bit)

Arquivo Editar Visualizar Misc Pacotes Janelas Ajuda



### **Rstudio IDE:**

RStudio é um ambiente de desenvolvimento integrado (IDE) para R.

RStudio inclui um console, um editor com destaque para sintaxe que suporta execução direta de código e ferramentas para plotagem, depuração e gerenciamento de espaço de trabalho.

Download RStudio em: https://posit.co/download/rstudio-desktop/ de acordo com o sistema operacional do seu desktop.







https://posit.co/download/rstudio-desktop/



O DOWNLOAD RSTUDIO

















Used by millions of people weekly, the RStudio integrated development environment (IDE) is a set of tools built to help you be more productive with R and Python. It includes a console, syntax-highlighting editor that supports direct code execution. It also features tools for plotting, viewing history, debugging and managing your workspace.

Select Your Operating System:

All Installers

#### Step 1: Install R

RStudio requires R 3.3.0+. Choose a version of R that matches your computer's operating system.

DOWNLOAD AND INSTALL R

#### Step 2: Install RStudio Desktop

DOWNLOAD RETUDIO DESKTOP FOR WINDOWS

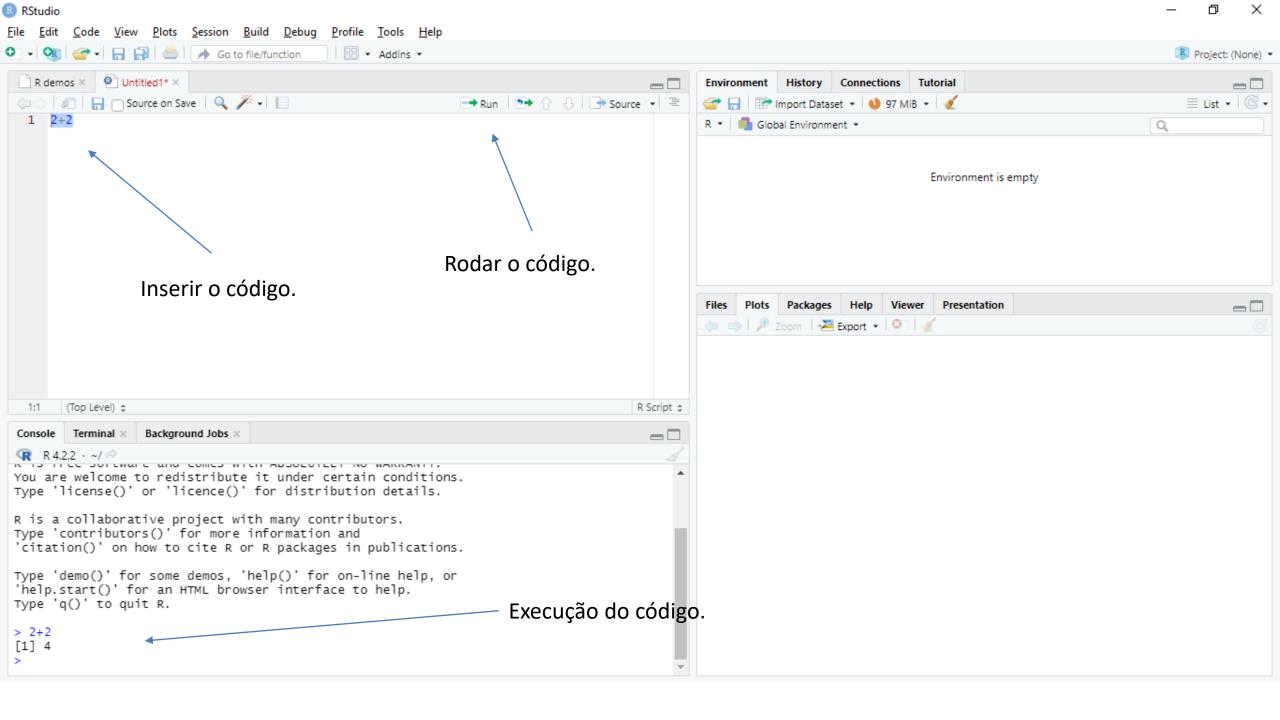
Size: 190.49MB | SWA-256: 8288F925 | Version: 2022.07.2+576 | | Released: 2022-09-21



### **RStudio:**

A IDE do RStudio deve estar configurada da maneira mostrada a seguir.

Execute o comando 2+2 no RStudio como mostrado a seguir.



## Começando com o RStudio:

Abra uma pasta em C:

Criando um arquivo com extensão .R

Abra a IDE do RStudio, acesse FILE, depois NEWFILE e por fim em RSCRIPT.

Crie um arquivo chamado vendas.R

Abra o Editor de textos e na pasta salve um arquivo com extensão .csv chamado vendas.csv

### Carregando o DataFrame:

No bloco de notas no arquivo vendas.csv digite e salve-o os dados abaixo:

```
*vendas - Bloco de Notas

Arquivo Editar Formatar Exibir Ajuda

ID, Data, Produto, Valor, Sede

14032022, 14/3/2022, ford, 50, Botafogo

15032022, 15/3/2022, bmw, 120, Maracanã

20032022, 20/3/2022, nissan, 90, Barra

220302022, 22/03/2022, jeep, 80, Ipanema

28032022, 28/03/2022, fiat, 60, Jacarepaguá
```

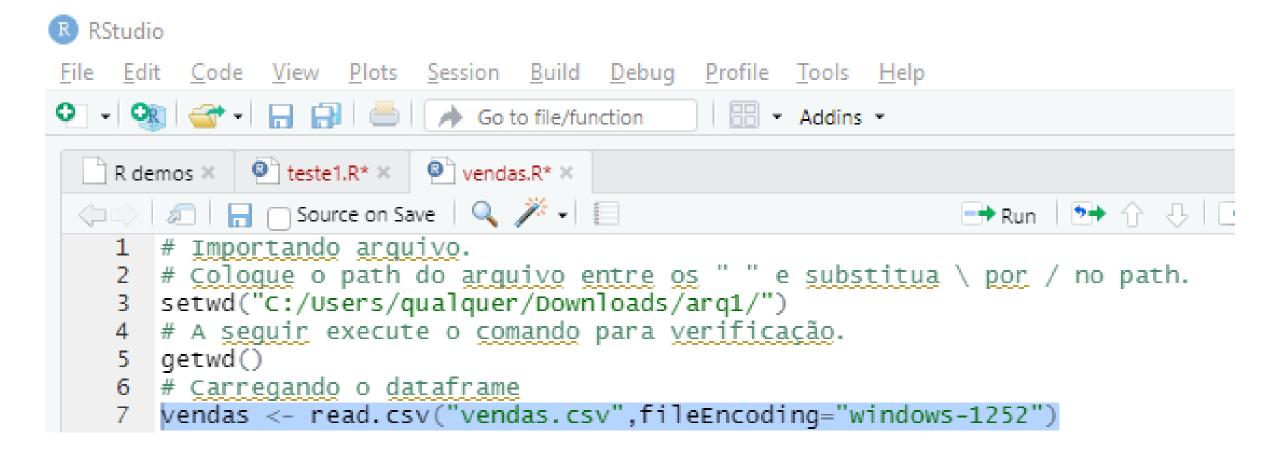
## Começando com o Rstudio:

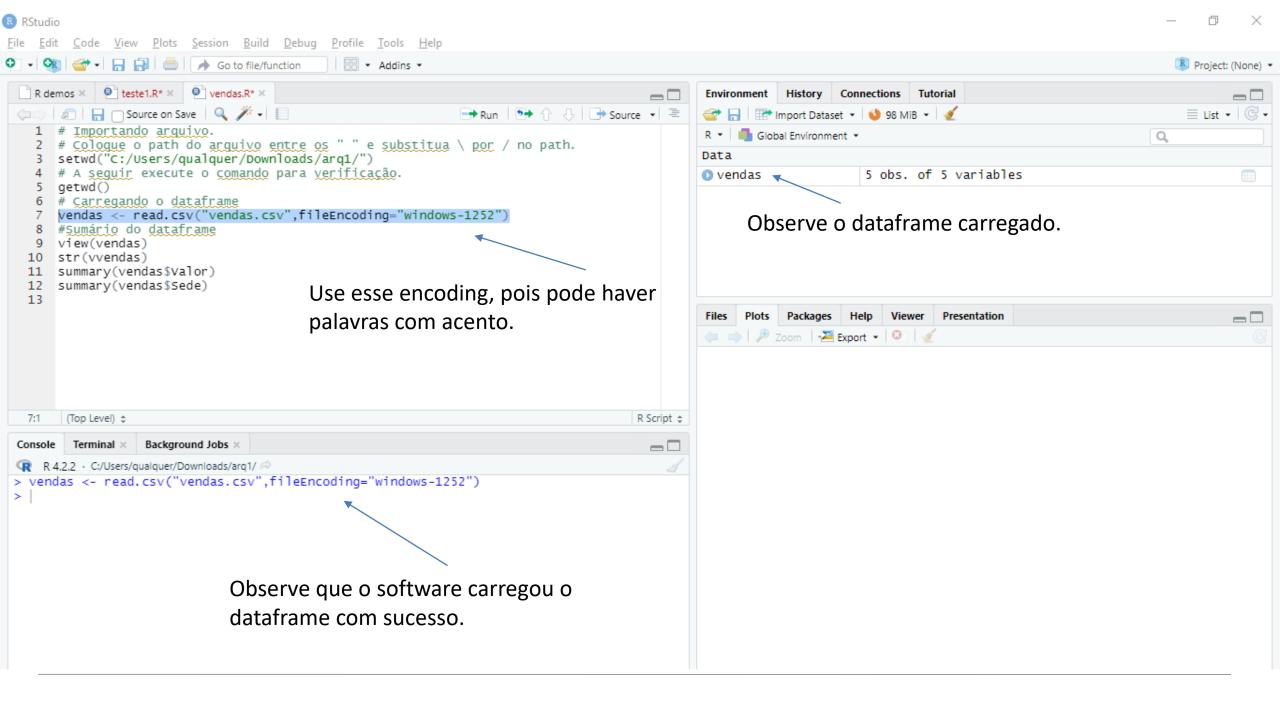
Fazendo a importação do arquivo.

Para isto usa-se a função setwd("path do arquivo") Atenção ao path pois usa-se / por exemplo: C:/usuário/aluno

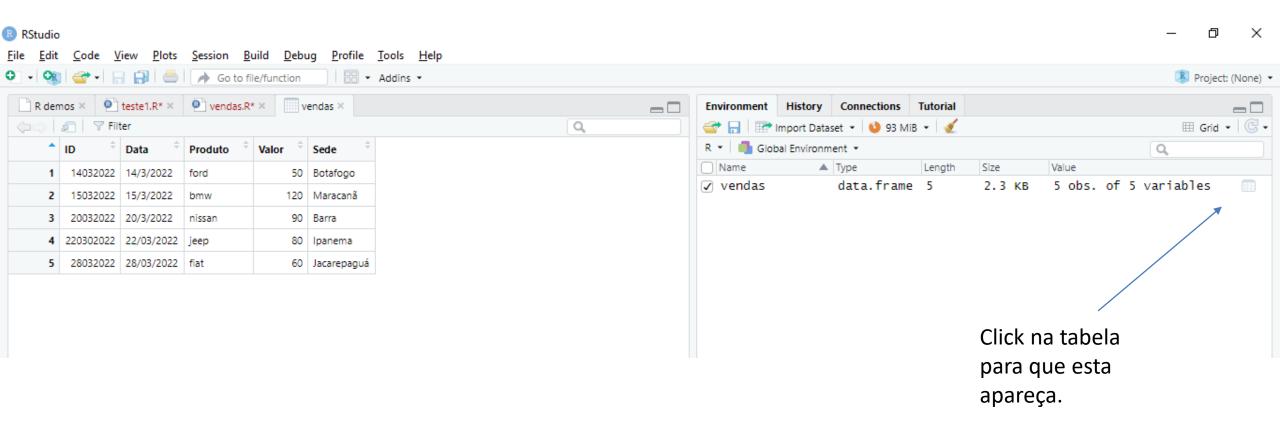
Depois usa-se a função getwd() para checar o procedimento anterior.

## Importando o arquivo .csv:





### Vista do DataFrame:



## Informações e Sumário do DataFrame:

Use str(nome do arquivo) para obter as informações.

Use summary(nome do arquivo) para obter um sumário dos dados.

Use summary(nome do arquivo\$nome da variável) para obter um sumário dos dados de uma coluna.

#### str:

```
8 #Sumário do dataframe
  9 str(vendas)
 10 summary(vendas$valor)
 11 summary(vendas$Sede)
 12
                                                                                   R Script $
 10:1 (Top Level) $
Console Terminal × Background Jobs ×
                                                                                     R 4,2,2 · C:/Users/qualquer/Downloads/arq1/
> view(vendas)
Error in view(vendas): could not find function "view"
> #Sumário do dataframe
> str(vendas)
'data.frame': 5 obs. of 5 variables:
 $ ID : int 14032022 15032022 20032022 220302022 28032022
$ Data : chr "14/3/2022" "15/3/2022" "20/3/2022" "22/03/2022" ...
 $ Produto: chr "ford" "bmw" "nissan" "jeep" ...
 $ valor : int 50 120 90 80 60
$ Sede : chr "Botafogo" "Maracanã" "Barra" "Ipanema" ...
>
```

## **Summary:**

```
#Sumário do dataframe
     str(vendas)
     summary(vendas)
     summary(vendas$valor)
     summary(vendas$5ede)
 13
                                                                                    R Script $
 10:1
      (Top Level) $
Console Terminal ×
                 Background Jobs ×
R 4,2,2 · C:/Users/qualquer/Downloads/arq1/
2 Sede : Cur.
                  BOLATOGO MANACANA
                                        Barra.
                                                трапеша ...
> summary(vendas)
                                          Produto
                                                               valor
       ID
                         Data
Min. : 14032022
                   Length:5
                                        Length:5
                                                           Min. : 50
1st Qu.: 15032022
                   Class :character Class :character
                                                           1st Qu.: 60
Median: 20032022
                     Mode :character
                                        Mode :character
                                                           Median: 80
Mean : 59486022
                                                           Mean: 80
                                                            3rd Qu.: 90
3rd Qu.: 28032022
        :220302022
                                                            Max.
                                                                   :120
Max.
     sede
Length:5
Class :character
Mode :character
```

## **Summary:**

```
#Sumário do dataframe
   9 str(vendas)
  10 summary(vendas)
  11 summary(vendas$valor)
      summary(vendas$Sede)
  12
 13
 11:1
       (Top Level) $
                                                                                       R Script $
                  Background Jobs ×
Console
        Terminal ×
R 4,2,2 · C:/Users/qualquer/Downloads/arq1/
                      Length: 5
                                          Length:5
 MIII. : 14U3ZUZZ
                                                              MILLI.
                                                                       20
                                          class :character
 1st Qu.: 15032022 Class :character
                                                              1st Qu.: 60
 Median: 20032022
                     Mode :character
                                          Mode :character
                                                              Median: 80
       : 59486022
                                                              Mean
                                                                     : 80
 Mean
 3rd ou.: 28032022
                                                              3rd Qu.: 90
        :220302022
                                                                     :120
 Max.
                                                              Max.
     Sede
 Length: 5
 class :character
 Mode :character
> summary(vendas$valor)
   Min. 1st Qu. Median
                            Mean 3rd Ou.
                                             мах.
             60
                      80
                                              120
     50
                              80
                                       90
```

### Tipos de Dados em R:

- •numeric (10.5, 55, 787.333)
- •integer (55L, 100L, onde "L" declara que o valor é inteiro)
- •complex (9 + 3i, onde "i" é o imaginário da função)
- •character (string) ("k", "R é show", "FALSE", "11.5")
- •logical (boolean) (TRUE ou FALSE)

## Tipos de dados e operadores:

O R tem as seguintes operações básicas:

sinal	operação
+	Adição
	Subtração
1	Divisão
580	Multiplicação
^	Potência
sqrt	Raiz quadrada

### Variáveis:

Para guardar os valores em uma variável em R, usamos <-.

Para imprimir a variável usa-se o print.

Class vai imprimir o tipo de variável.

Exemplo:

a <- 2.5^3

print(a)

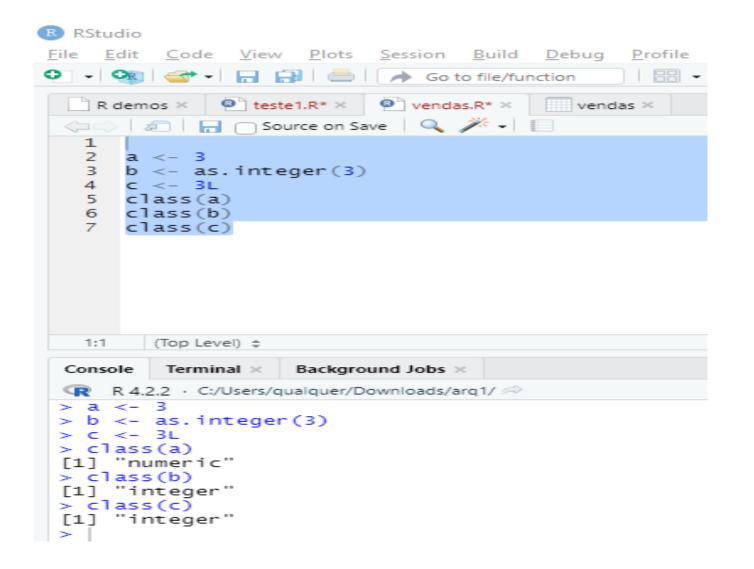
class(a)

```
RStudio
    Edit Code View Plots Session B
    R demos × 👂 teste1.R* × 👂 vendas.R
       a <- 2.5^3
      print(a)
      class(a)
        (Top Level) $
        Terminal ×
                   Background Jobs
     R 4,2,2 · C:/Users/qualquer/Downloads/arq1
 > a <- 2.5^3
 > print(a)
 [1] 15.625
 > class(a)
 [1] "numeric"
```

### Variáveis:

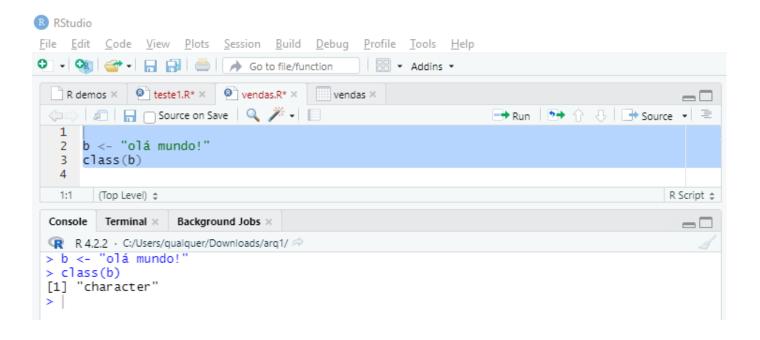
Se quiser trabalhar com número inteiro, deve-se definí-los como integer, através da função as.integer() ou L no final. Por Exemplo:

```
a <- 3
b <- as.integer(3)
c <- 3L
class(a)
class(b)
class(c)</pre>
```



## Variável (Character):

Em R, os dados também podem ser do tipo Character.



### Estruturas de dados no R:

#### **VETOR:**

Um vetor é um conjunto de dados do mesmo tipo.

Cria-se vetores com mais de um elemento dentro de um c().

Exemplo: vet <- c(2,3,4)

RStudio File Edit Code View Plots Session Build Debug Profile Tools Help R demos × teste1.R\* × vendas.R\* × vendas × ⟨□□⟩ | Ø□ | □ Source on Save | Q 
Ø▼ ▼ | □ | Run 1 3 C U Source - 3 1 vet <- c(3,6, 7.8, 332) print(vet) (Top Level) \$ R Script \$ 1:1 Console Terminal × Background Jobs × =R 4.2.2 · C:/Users/qualquer/Downloads/arq1/ > vet <- c(3,6, 7.8, 332) > print(vet) [1] 3.0 6.0 7.8 332.0 >

### Estruturas de dados no R:

#### **MATRIZ:**

É um conjunto bidimensional de dados.

Constrói-se uma matriz fornecendo os dados e o número de linhas e colunas.

Observe no exemplo a seguir a diferença da matriz 3x2 da matriz 2x3.

```
RStudio
                                             1:1
                                                  (Top Level) $
File Edit Code View Plots Session Build Debug
                                           Console
                                                 Terminal ×
                                                            Background Jobs
                                           R 4,2,2 · C:/Users/qualquer/Downloads/arc

→ Go to file/function

                                           > mat_32 <- matr1x(
                                              c(1,5,10,30,15,8)
  nrow=3,
 ncol=2,
                                               byrow=TRUE)
      mat_32 <- matrix(
                                           > print(mat_32)
       c(1,5,10,30,15,8),
                                                [,1] [,2]
       nrow=3.
                                           [1,]
       ncol=2,
                                           [2,] 10 30
                                                15 8
                                           [3,]
       byrow=TRUE)
      print(mat_32)
                                           > mat_23 <- matrix(
                                               c(1,5,10,30,15,8),
      mat_23 <- matrix(
                                              nrow=2,
        c(1,5,10,30,15,8),
                                              ncol=3,
  10
       nrow=2,
                                               byrow=TRUE)
  11
       ncol=3,
                                           > print(mat_23)
  12
       byrow=TRUE)
                                                [,1] [,2] [,3]
                                           \lceil 1, \rceil
      print(mat_23)
                                                  30 15
                                           [2,]
       (Top Level) $
```

### Outra forma de se construir uma matriz:

Pode-se construir uma matriz a partir da definição de vetores através da função rbind.

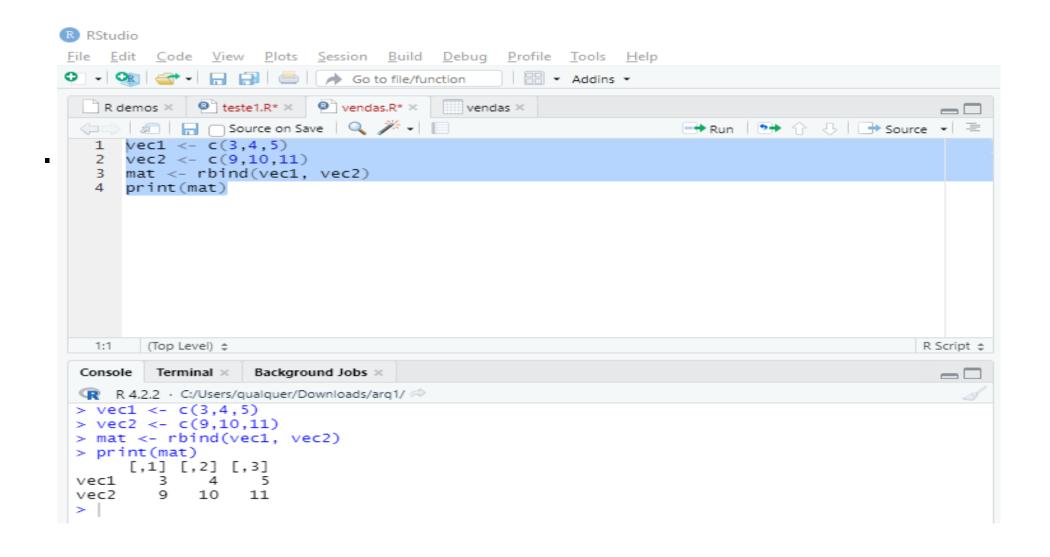
Para isto define-se as linhas e depois se contrói-se a matriz na sequência das linhas.

#### Exemplo:

```
vec1 <- c(1,2,3)
```

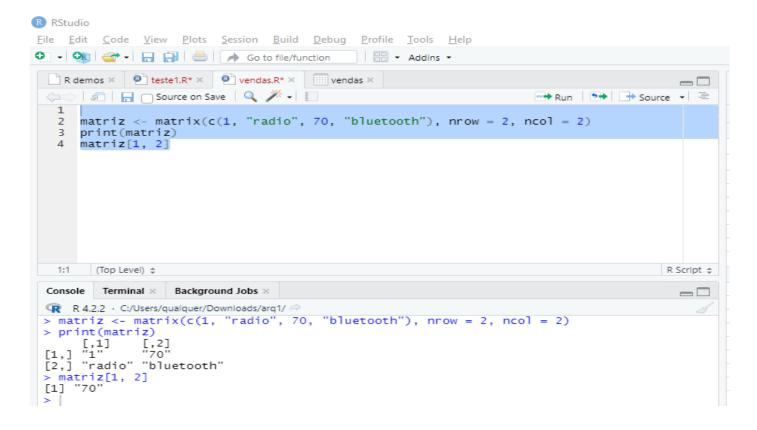
$$vec2 <- c(4,5,6)$$

Matriz <- rbind(vec1,vec2)</pre>



### Selecionando um elemento de uma matriz:

Basta usar-se matriz[i,j].

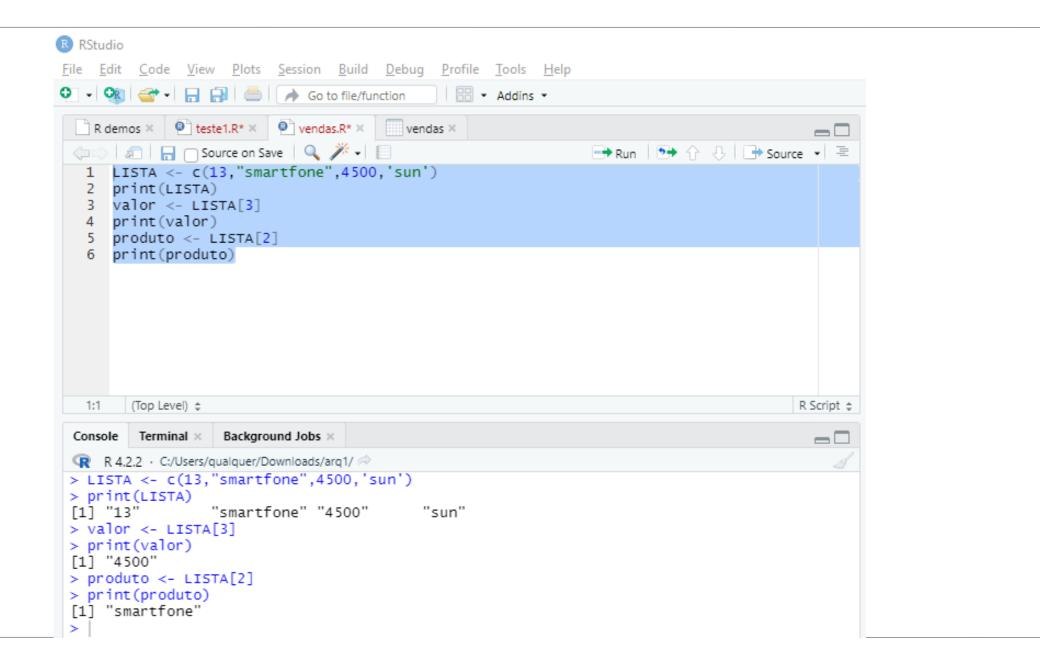


### Estruturas de dados no R:

#### LISTA:

São estruturas formadas por diversos tipos de elementos.

Para localizar um valor na lista usa-se: atributo <- nome\_da\_lista[índice]



#### Alterar um elemento de uma lista:

Usa-se: nome\_da\_lista[índice do elemento]<- novo valor

```
RStudio
<u>File Edit Code View Plots Session Build Debug Profile Tools Help</u>
◆ Go to file/function
 R demos ×  teste1.R* ×  vendas.R* ×  vendas ×
 Run Source - =
  3 list <- list("smart tv", 2500, "cheny")</pre>
     print(list)
    list[2] <- 3500
    print(list)
```



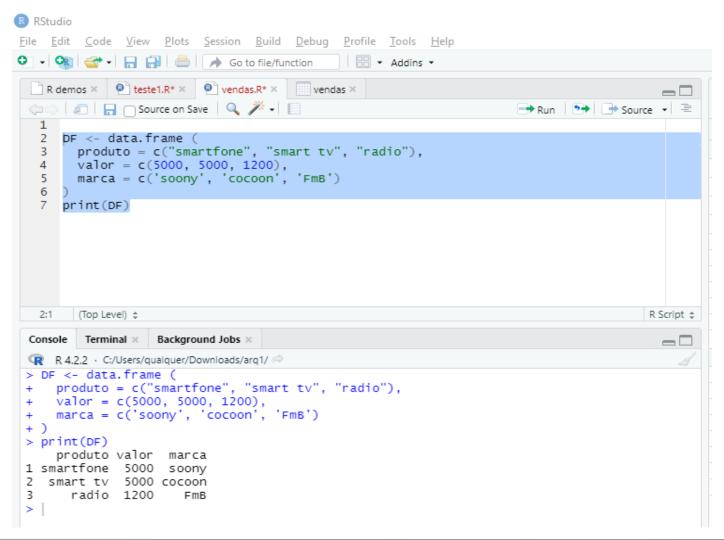
### Estruturas de dados no R:

#### **DataFrame:**

É uma tabela de dados.

É uma estrutura de dados muito usada em Ciência de Dados.

# Construção de um DataFrame:



#### **Fatores:**

FATORES são usadas para categorizar dados.

Exemplos de fatores:

- Sexo: Masculino/Feminino
- •Tipo de Música: Rock, Pop, Clássica, Jazz

Para criar um fator usa-se a função fator()

#### **Fatores:**

Criando um fator:

```
tipo_música <-
factor(c("Jazz", "Rock", "Classic", "Classic", "Pop", "Jazz", "Rock
", "Jazz"))
```

Imprimindo um fator: tipo\_música

# Funções Básicas em R:

#### A função:

- length(x) retorna o comprimento de um objeto x.
- sum(x) soma todos os elementos de um objeto x.
- rep(x,n) repete o número x, n vezes.
- seq(a, b, by=c) gera uma sequência de números contidos entre a e b, distantes c unidades um do outro.
- table(x) retorna uma tabela com as frequências absolutas de ocorrência da cada elemento de x.



## Funções em R:

Uma função é um bloco de código que só é executado quando este é chamado.

Pode-se passar dados (parâmetros) para uma função e esta função pode retornar dados como resultado.

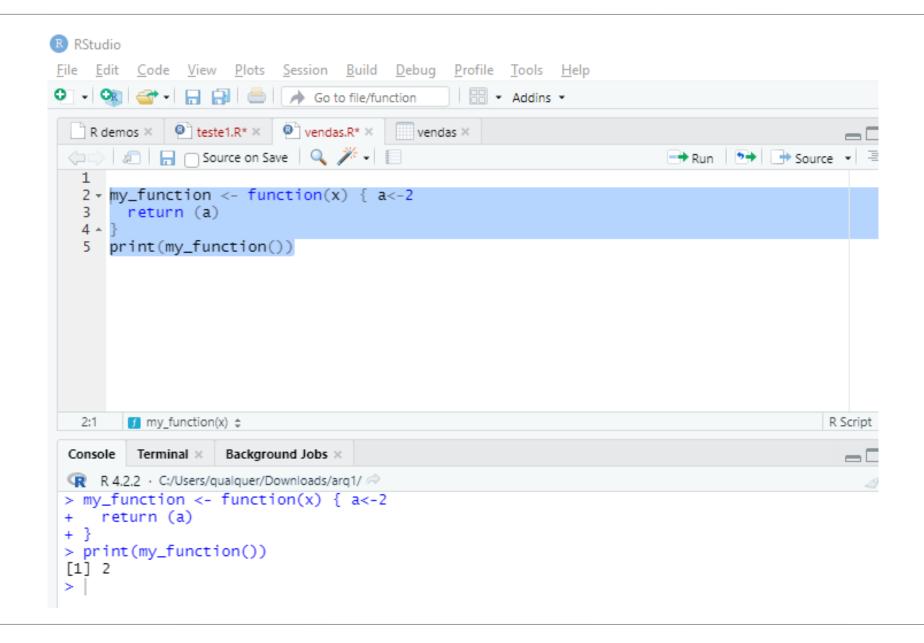
# Criando uma função:

Para criar uma função use: function()

my\_function <- function() { código }</pre>

## Chamar uma função:

```
Para chamar uma função use: my_function()
my_function <- function() { a <- 2
  return (a)
}
print(my_function())</pre>
```



### Outro exemplo:

```
my_function <- function(x, y) {a <- x + y
return(a)
}
print(my_function(2,5))</pre>
```

### Variáveis Globais:

São as variáveis criadas fora de uma função.

```
Exemplo:
txt <- "sensacional"
my_function <- function() {
   paste("R é", txt)
}
my_function()</pre>
```

### Variáveis Globais:

```
imposto <- 0.05 #imposto de 5%
my_function <- function() {
valor_imposto_produto=230*imposto
print(valor_imposto_produto)
}
my_function()</pre>
```