



Machine Learning com R

Seja bem-vindo!!!

Estatística Descritiva com R

Tipos de Dados(objetos) no R

- Vetor;
- Matriz;
- Lista;
- Data Frame (tabelas);

Tipos de Dados(objetos) no R

- Vetor:

- coleção de elementos: strings, números, booleano;
- Todos os valores devem ser iguais.

```
> y <- c(TRUE, TRUE, FALSE, FALSE)
> y
[1] TRUE TRUE FALSE FALSE
> z <- c("Sarah", "Tracy", "Jon")
> z
[1] "Sarah" "Tracy" "Jon"
> |
```

- Crie um vetor de nome "n" do tipo numérico.

Tipos de Dados(objetos) no R

- Vetor: Inserindo novo elemento em um vetor

```
z <- c(z, "Annette")  
z
```

- Insira um novo elemento no vetor "n"

Tipos de Dados(objetos) no R

- Vetor: Acessando valores.

```
> y[1]
[1] TRUE
> z[3]
[1] "Jon"
> |
```

Tipos de Dados(objetos) no R

- Vetor: Acessando valores (sequencial).

```
> y[2:4]
[1] TRUE FALSE FALSE
>
> z[1:3]
[1] "SARAH" "Tracy" "Jon"
> |
```

Tipos de Dados(objetos) no R

- Vetor: Acessando multiplo valores (slice).

```
> y[c(1,3)]  
[1] TRUE FALSE  
> z[c(2,4)]  
[1] "Tracy" "annette"  
> |
```


Tipos de Dados(objetos) no R

- Vetor: Deletando elemento (mantendo o índice)

```
> y[3] <- NA
> y
[1] TRUE TRUE NA FALSE
> z[1] <- NA
> z
[1] NA "Tracy" "Jon" "annette"
```

Tipos de Dados(objetos) no R

- Vetor: Deletando elemento e indice

```
> y <- y[-3]
> y
[1] TRUE TRUE FALSE
> z <- z[-1]
> z
[1] "Tracy" "Jon" "annette"
> |
```

Tipos de Dados(objetos) no R

- Matriz:

- é um tipo de vetor multi-dimensional, linhas x colunas;
- Pode ser: string, números, booleanos.

```
m <- matrix(nrow = 2, ncol = 2)
m
```

Output

| | [,1] | [,2] |
|------|------|------|
| [1,] | NA | NA |
| [2,] | NA | NA |

Tipos de Dados(objetos) no R

- Matriz:

- Todos os dados devem ter o mesmo Tipo;

```
m <- matrix(data = c(TRUE, "dois", 4, "seis", FALSE, "gabriel",  
                      "IESB", 10, 2019, "machine learning",  
                      "IA", "Big data", 1, 2, 3, 5),  
            nrow = 4, ncol = 4)
```

```
> m
```

| | [,1] | [,2] | [,3] | [,4] |
|------|--------|-----------|--------------------|------|
| [1,] | "TRUE" | "FALSE" | "2019" | "1" |
| [2,] | "dois" | "gabriel" | "machine learning" | "2" |
| [3,] | "4" | "IESB" | "IA" | "3" |
| [4,] | "seis" | "10" | "Big data" | "5" |

```
> |
```

Tipos de Dados(objetos) no R

- Matriz: Acessando valores.

```
> m[3]
[1] "4"
> m[6]
[1] "gabriel"
> |
```

```
> m[1,2]
[1] "seis"
> m[2,1]
[1] "dois"
> |
```

Tipos de Dados(objetos) no R

- Matriz: Acessando sequência de valores.

```
> m[2:4,2] ## linhas 2 a 4, coluna 2  
[1] "gabriel" "IESB"    "10"  
>
```

```
> m[3,2:4] ## linha 3, colunas 2 a 4  
[1] "IESB" "IA"    "3"  
> |
```

Tipos de Dados(objetos) no R

- Matriz: Acessando valores específicos.

```
> m[c(1,3),c(1,3)]  
      [,1] [,2]  
[1,] "TRUE" "2019"  
[2,] "4"    "IA"  
>
```

```
> m[c(2,4),c(1,3,4)]  
      [,1] [,2] [,3]  
[1,] "dois" "machine learning" "2"  
[2,] "seis" "Big data" "5"  
> |
```

Tipos de Dados(objetos) no R

- Matriz: Acessando indices inexistentes.

```
> m[5,5]  
Error in m[5, 5] : subscript out of bounds
```


Tipos de Dados(objetos) no R

- Matriz: Alterando valores.

```
> m[1,1] <- "IESB"
> m
```

| | [,1] | [,2] | [,3] | [,4] |
|------|--------|-----------|--------------------|------|
| [1,] | "IESB" | "FALSE" | "2019" | "1" |
| [2,] | "dois" | "gabriel" | "machine learning" | "2" |
| [3,] | "4" | "IESB" | "IA" | "3" |
| [4,] | "seis" | "10" | "Big data" | "5" |

```
> |
```

Tipos de Dados(objetos) no R

- Matriz: Excluindo valor.

```
> m[3,2] <- NA
> m
```

| | [,1] | [,2] | [,3] | [,4] |
|------|--------|-----------|--------------------|------|
| [1,] | "IESB" | "FALSE" | "2019" | "1" |
| [2,] | "dois" | "gabriel" | "machine learning" | "2" |
| [3,] | "4" | NA | "IA" | "3" |
| [4,] | "seis" | "10" | "Big data" | "5" |

Tipos de Dados(objetos) no R

- Matriz: Excluindo indice.

```
m_2 <- m[-4,] ## excluindo linha inteira  
m_2
```

```
m_2 <- m[, -2] ## excluindo coluna inteira  
m_2
```

```
m_2 <- m[-4, -2] ## excluindo linha e coluna  
m_2
```

Tipos de Dados(objetos) no R

- Lista:

- Semelhante a um vetor, mas aceita diferentes tipos de dados juntos.
- Uma posição da lista aceita string, números, lista, vetor, matriz, tabelas

```
l <- list(1,"dois",FALSE)

[[1]]
[1] 1

[[2]]
[1] "dois"

[[3]]
[1] FALSE
```

Tipos de Dados(objetos) no R

- Lista: Acessando valores.

```
> l[[1]]  
[1] 1  
> l[[3]]  
[1] FALSE  
>
```

Tipos de Dados(objetos) no R

- Lista: alterando/Inserindo valores.

```
> l[[2]] <- "alteração lista"
> l
[[1]]
[1] 1

[[2]]
[1] "alteração lista"

[[3]]
[1] FALSE
```

```
[1] FALSE
[[3]]
```

```
> l[[2]][2] <- "novo valor"
> l
[[1]]
[1] 1

[[2]]
[1] "alteração lista" "novo valor"

[[3]]
[1] FALSE
```

```
[1] FALSE
[[3]]
```

Tipos de Dados(objetos) no R

- Lista: inserindo outros tipos de dados na lista. Inserindo um vetor.

```
> l[[1]] <- c('vetor_1', 'vetor_2', 'vetor_3')
> l
[[1]]
[1] "vetor_1" "vetor_2" "vetor_3"

[[2]]
[1] "alteração lista" "novo valor"

[[3]]
[1] FALSE
```

Tipos de Dados(objetos) no R

- Lista: Acessando uma posição específica do vetor dentro da lista.

```
> l[[1]][2]  
[1] "vetor_2"  
> |
```


Tipos de Dados(objetos) no R

- Lista: inserindo outros tipos de dados na lista. Lista dentro da lista

```
>
> l[[2]] <- list('lista_1','lista_2','lista_3','lista_4')
> l
[[1]]
[1] "vetor_1" "vetor_2" "vetor_3"

[[2]]
[[2]][[1]]
[1] "lista_1"

[[2]][[2]]
[1] "lista_2"

[[2]][[3]]
[1] "lista_3"

[[2]][[4]]
[1] "lista_4"

[[3]]
[1] FALSE
```

Tipos de Dados(objetos) no R

- Lista: Acessando uma posição específica da lista dentro da lista

```
>  
> l[[2]][4]  
[[1]]  
[1] "lista_4"
```

Tipos de Dados(objetos) no R

- Lista: inserindo outros tipos de dados na lista. Inserindo uma matriz.

```
> l[[3]] <- matrix(nrow = 2, ncol = 3, data = c(1,2,3,4,5,6))
> l
[[1]]
[1] "vetor_1" "vetor_2" "vetor_3"

[[2]]
[[2]][[1]]
[1] "lista_1"

[[2]][[2]]
[1] "lista_2"

[[2]][[3]]
[1] "lista_3"

[[2]][[4]]
[1] "lista_4"

[[3]]
      [,1] [,2] [,3]
[1,]    1    3    5
[2,]    2    4    6
```

Tipos de Dados(objetos) no R

- Lista: Acessando uma posição específica do matriz dentro da lista.

```
> l[[3]][1,3]  
[1] 5  
> |
```

Tipos de Dados(objetos) no R

- Lista: Deletando um índice da lista.

```
> l <- l[-2]
> l
[[1]]
[1] "vetor_1" "vetor_2" "vetor_3"
```

Tipos de Dados(objetos) no R

•Data Frame:

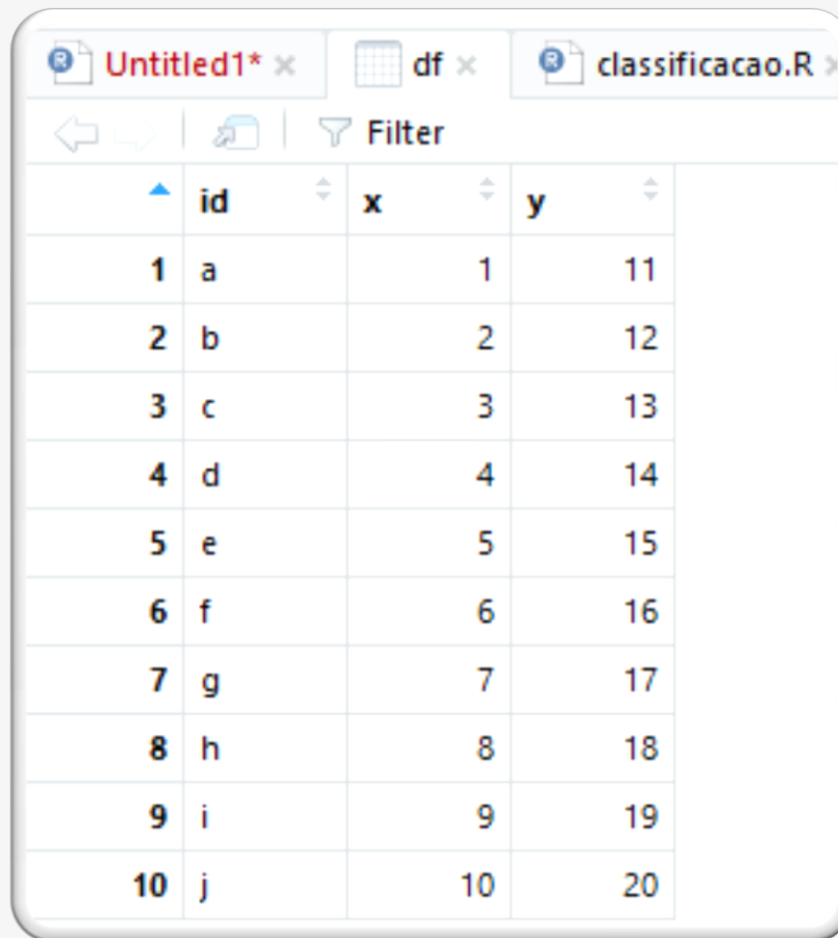
- Semelhante a uma matriz, tabela;
- Aceita diferentes tipos de dados juntos;
- **MAIS IMPORTANTE TIPO DE DADOS NO R.**

```
> df <- data.frame(id = letters[1:10], x = 1:10, y = 11:20)
> df
```

| | id | x | y |
|----|----|----|----|
| 1 | a | 1 | 11 |
| 2 | b | 2 | 12 |
| 3 | c | 3 | 13 |
| 4 | d | 4 | 14 |
| 5 | e | 5 | 15 |
| 6 | f | 6 | 16 |
| 7 | g | 7 | 17 |
| 8 | h | 8 | 18 |
| 9 | i | 9 | 19 |
| 10 | j | 10 | 20 |

Tipos de Dados(objetos) no R

- Data Frame:



The screenshot shows an R IDE window with three tabs: 'Untitled1*' (active), 'df', and 'classificacao.R'. Below the tabs is a toolbar with navigation arrows, a copy icon, and a 'Filter' button. The main area displays a data frame with the following data:

| | id | x | y | |
|----|----|---|----|----|
| 1 | a | | 1 | 11 |
| 2 | b | | 2 | 12 |
| 3 | c | | 3 | 13 |
| 4 | d | | 4 | 14 |
| 5 | e | | 5 | 15 |
| 6 | f | | 6 | 16 |
| 7 | g | | 7 | 17 |
| 8 | h | | 8 | 18 |
| 9 | i | | 9 | 19 |
| 10 | j | | 10 | 20 |

Tipos de Dados(objetos) no R

- Data Frame: acessando valores

- Uma coluna e TODAS as linhas

```
> df$x  
[1] 1 2 3 4 5 6 7 8 9 10  
> |
```

- Colunas especificas e todas as linhas

```
> df[,c("id","y")]  
  id y  
1  a 11  
2  b 12  
3  c 13  
4  d 14  
5  e 15  
6  f 16  
7  g 17  
8  h 18  
9  i 19  
10 j 20  
> |
```

- Colunas e Linhas específicas

```
> df[c(1,5,10),c("id","y")]  
  id y  
1  a 11  
5  e 15  
10 j 20  
> |
```

- Uma linha e todas as colunas

```
> df[3,]  
  id x  y  
3  c 3 13  
> |
```


Tipos de Dados(objetos) no R

- Data Frame: Deletando dados

- Deletando uma coluna

```
df_2 <- df[, -3]

df_2 <- df ## criando uma cópia
df_2$y <- NULL ## deletando a coluna y
```

- Deletando multiplicas columnas

```
df_2 <- df[, c(-1, -3)] ## deletando a primeira e a terceira coluna

df_2 <- df ## criando uma cópia
df_2[, c('id', 'y')] <- NULL ## deletando as colunas id e y
```

Tipos de Dados(objetos) no R

- Data Frame: Deletando dados

- Deletando uma linha

```
## excluindo linha específica  
df_2 <- df[-5,]
```

- Deletando multiplas linhas

```
## excluindo multiplas linhas  
df_2 <- df[c(-1,-5,-10),]
```

Tipos de Dados(objetos) no R

- DataFrame: Selecionando(excluindo) valores com condições.

- CONDIÇÃO DE IGUALDADE

```
> df[df$id == 'd',]  
  id x  y  
4  d 4 14  
> |
```

Tipos de Dados(objetos) no R

- DataFrame: Acessando valores com condições.

- CONDIÇÃO | (OR/OU)

```
> df[(df$id == 'd') | (df$x == 7),]  
  id x  y  
4  d 4 14  
7  g 7 17  
> |
```

Tipos de Dados(objetos) no R

- DataFrame: Acessando valores com condições.

- CONDIÇÃO & (AND/E)

```
> df[(df$id == 'd') & (df$x == 7),]  
[1] id x  y  
<0 linhas> (ou row.names de comprimento 0)  
> |
```

Tipos de Dados(objetos) no R

- VAMOS PRATICAR...



Referências

- <https://swcarpentry.github.io/r-novice-inflammation/13-supp-data-structures/index.html>

-