

# Esquenta para o curso: Introdução à Estatística com R

## 🔥 Aula 1 🔥



Profª Edneide Ramalho

10 de Julho de 2021

# Quem sou eu e o que é Stat-Tech?

## Edneide Ramalho

- Doutora em Biometria e Estatística Aplicada - UFRPE
- Pós-doc no Laboratório de Immunopaltologia Keizo Asami (LIKA-UFPE)
- Resposta de Emergência para COVID-19 (IRR-PE)
- Consultora na Stat-Tech
- Trabalha com o R desde 2015
- Principal interesse: Modelagem matemática e computacional aplicada à epidemiologia e análise de dados em saúde
- Instagram: [https://www.instagram.com/edneide\\_ramalho/](https://www.instagram.com/edneide_ramalho/)
- LinkedIn: <https://www.linkedin.com/in/edneide-ramalho-05054ba2/>

# Stat-Tech

**Startup** de consultoria estatística, mentoria acadêmica, cursos on-line e projetos de ciência de dados



- Instagram: [https://www.instagram.com/stat\\_tech/](https://www.instagram.com/stat_tech/)
- YouTube: <https://www.youtube.com/channel/UCRmGVuVRNgUtX0Onei1k82g>
- E-mail: [contact@stat-techbr.com.br](mailto:contact@stat-techbr.com.br)

# O é R e RStudio?

- Foi criado por Ross Ihaka e por Robert Gentleman no departamento de Estatística da Universidade de Auckland, Nova Zelândia.
- O nome R é dado a partir das iniciais de seus criadores.
- Voltada à manipulação, análise e visualização de dados, tem grande popularidade entre estatísticos e analistas de dados.

## Vantagens

- É um software livre, e atualmente mantido por uma comunidade de colaboradores voluntários que contribuem com código fonte da linguagem e com a expansão de funcionalidades por bibliotecas.

# RStudio

Principal IDE (Integrated Development Environment) utilizada para programação em R.



# Por que aprender R?

- É uma ferramenta livre e gratuita
- Usada em quase todas as áreas do conhecimento
- Tem uma grande comunidade de usuários
- É uma plataforma independente
- Tem uma biblioteca robusta de visualização de dados
- Usada para Estatística e Ciência de Dados

# Curso



**INTRODUÇÃO À  
ESTATÍSTICA COM**

**CURSO ONLINE (AO VIVO)**

**De 19 a 30 de Julho de 2021**

**Segundas, Quartas e Sextas  
das 19:30h as 21:30h**

**VAGAS  
LIMITADAS!!!**

**Stat-Tech**

- Sorteio de 4 bolsas de 50% no dia 17/07/2021 no instagram.

A close-up photograph of a person's hands writing in a spiral-bound notebook. The person is wearing an orange ribbed sweater. They are holding a yellow pencil and writing on a lined page. In the background, a silver laptop is open on a white desk. Another orange notebook is visible to the right. The text "Conteúdos do curso" is overlaid in white on the person's hand and the notebook.

# Conteúdos do curso





# Parte 1: Conceitos Básicos

- Nesta parte vamos entender alguns conceitos básicos como:
  - Estatística
  - População
  - Amostra
  - Inferência Estatística
  - Parâmetros e estatísticas



## Parte 2: Estatística Descritiva

- Medidas de resumo
- Medidas de tendência central
- Medidas de dispersão
- Quantis e quartis



## Parte 3: Gráficos com ggplot2

- Gráfico de barras
- Gráfico de barras agrupadas
- Gráfico de linhas
- Histograma
- Box-plot



# Parte 4: Distribuição Normal

- O que é a distribuição normal e porque ela é importante
- Distribuição Normal padrão
- Calculando probabilidades a partir de uma distribuição Normal
- Teste de Normalidade
- Q-Q plot
- Gráfico de densidade



# Parte 5: Teste de Hipóteses

- O que é um teste de hipóteses
- P-valor
- Suposições para a realização dos testes
- Normalidade
- Homogeneidade de variância (homocedasticidade)
- Principais testes de hipóteses, interpretação e como reportar os resultados:
  - Qui-quadrado
  - Teste Exato de Fisher
  - Teste t de Student
  - Teste de Mann-Whitney
  - Teste t pareado
  - Teste de Wilcoxon
  - ANOVA (de um fator)
  - Teste de Kruskal-Wallis
  - ANOVA de medidas repetidas
  - Teste de Friedman

A close-up photograph of a person's hands writing in a spiral-bound notebook. The person is wearing an orange ribbed sweater. They are holding a yellow pencil and writing on a lined page. In the background, a silver laptop is open on a white desk. Another orange notebook is partially visible on the right side of the frame. The text "Primeiros passos no R" is overlaid in white on the center of the image.

# Primeiros passos no R

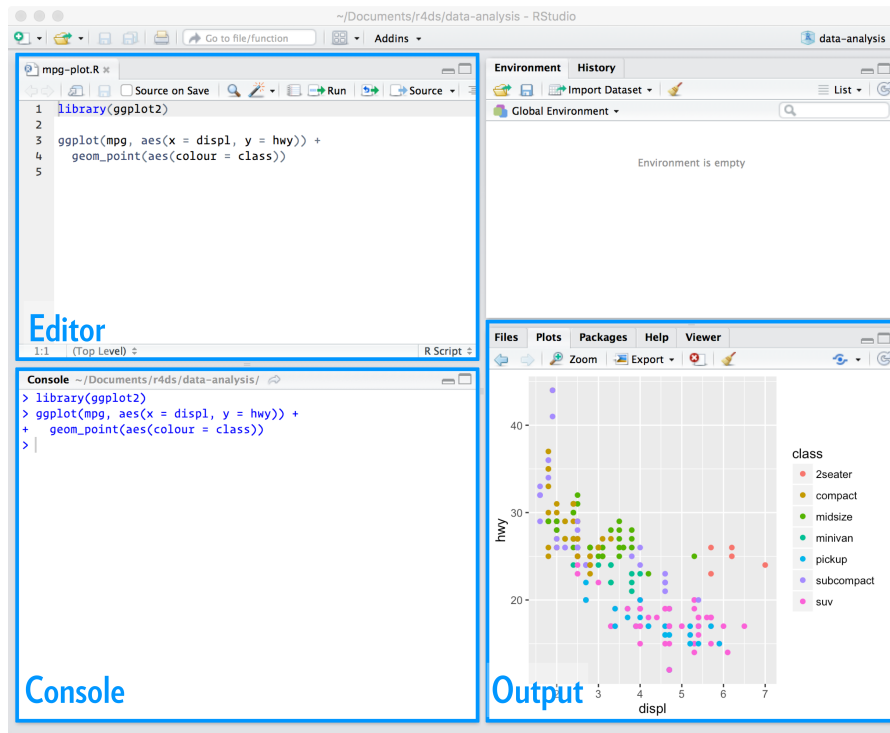


# Instalação do R e RStudio



- R para Windows: <https://cran.r-project.org/bin/windows/base/>
- R para Mac OS: <https://cran.r-project.org/bin/macosx/>
- R para Linux: <https://cran.r-project.org/>
- RStudio: <https://rstudio.com/products/rstudio/download/>
- RStudio Cloud: <https://rstudio.cloud/>

# Telas do RStudio

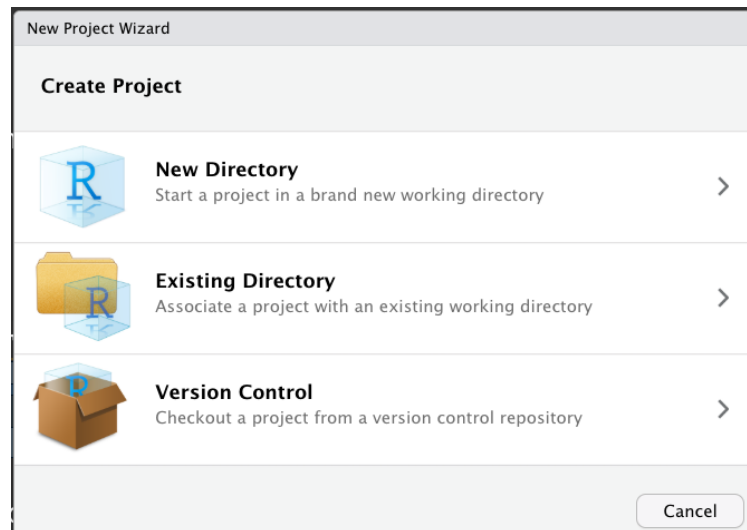


Fonte: Livro da Curso-R (<https://livro.curso-r.com/index.html>)



# Criando projetos no R

- Vamos aprender a criar projetos no R, antes de começar nossos primeiros comandos.
- `File -> New Project`
- ou Clicar no segundo ícone na parte superior esquerda: "Creat a project"



# Usando R como calculadora

- Primeiro, vamos criar um script no R
- Vamos fazer algumas operações matemáticas simples:

```
# Adição
```

```
2 + 2
```

```
# Subtração
```

```
4 - 2
```

```
# Multiplicação
```

```
4 * 35
```

```
# Divisão
```

```
165 / 5
```

```
# Potência
```

```
4 ^ 3
```

```
# Resto da divisão
```

```
15 %% 2
```

```
# Parte inteira da divisão
```

```
15 %/% 2
```

# Usando R como calculadora

```
# Raiz quadrada  
sqrt(4)
```

```
# Logaritmo  
log(10)
```

```
# Exponencial  
exp(12)
```

# Objetos

- Você pode armazenar valores em um **objeto**, usando o operador de atribuição `<-` (atalho: **ALT + -** ou **option + -**).

```
x <- 10
```

```
x
```

```
## [1] 10
```

# Objetos

- O R é case-sensitive, isto é, diferencia entre letras maiúsculas e minúsculas:

```
z <- 37  
Z <- 45
```

```
z
```

```
## [1] 37
```

```
Z
```

```
## [1] 45
```

- Também podemos criar objetos com palavras:

```
nome <- "Arthur"  
nome
```

```
## [1] "Arthur"
```

# Classes dos objetos

```
class(nome)
```

```
## [1] "character"
```

```
class(x)
```

```
## [1] "numeric"
```

```
class(z)
```

```
## [1] "numeric"
```

- Existe também a classe lógica: **TRUE** ou **FALSE**:

```
matriculado = TRUE
```

```
class(matriculado)
```

```
## [1] "logical"
```

# Criando vetores no R

- Um vetor é criado usando a função `c()`, que concatena os elementos separados por vírgulas inseridos dentro dos parênteses.

```
idades <- c(15, 18, 35, 24, 17, 80, 67, 45)
```

```
idades
```

```
## [1] 15 18 35 24 17 80 67 45
```

```
class(idades)
```

```
## [1] "numeric"
```

# Criando vetores no R

- Vetor de caracteres:

```
idades <- c("Recife", "Olinda", "Rio de Janeiro", "Belo Horizonte",  
            "Campinas", "Londrina")
```

```
idades
```

```
## [1] "Recife"          "Olinda"          "Rio de Janeiro" "Belo Horizonte"  
## [5] "Campinas"        "Londrina"
```

```
class(idades)
```

```
## [1] "character"
```



# Criando vetores no R

- Podemos fazer operações com vetores também:

```
idades_novas <- idades + 1
```

```
idades
```

```
## [1] 15 18 35 24 17 80 67 45
```

```
idades_novas
```

```
## [1] 16 19 36 25 18 81 68 46
```

# Criando vetores no R

```
idades_novas2 <- idades * 1.5  
idades_novas2
```

```
## [1] 22.5 27.0 52.5 36.0 25.5 120.0 100.5 67.5
```

# Criando vetores no R

```
idades_novas3 <- idades - 1  
idades_novas3
```

```
## [1] 14 17 34 23 16 79 66 44
```

```
idades + idades_novas2
```

```
## [1] 37.5 45.0 87.5 60.0 42.5 200.0 167.5 112.5
```

```
idades_novas3 - idades_novas2
```

```
## [1] -8.5 -10.0 -18.5 -13.0 -9.5 -41.0 -34.5 -23.5
```

# Criando vetores no R

- Sequências numéricas

```
seq1 <- 1:10  
seq1
```

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

```
seq2 <- seq(1, 10, by = 2)  
seq2
```

```
## [1] 1 3 5 7 9
```

# Índices dos vetores

- Usamos colchetes `[]` para identificar o valor do vetor numa determinada posição:

```
idades[3]
```

```
## [1] 35
```

```
idades[3]
```

```
## [1] "Rio de Janeiro"
```

# Testes lógicos

- Um teste lógico vai retornar verdadeiro (**TRUE**) ou falso (**FALSE**). Vejamos alguns exemplos:

```
num <- 35
```

```
# Igualdade  
num == 10
```

```
## [1] FALSE
```

```
# Menor que  
num < 10
```

```
## [1] FALSE
```

```
# Maior que  
num > 10
```

```
## [1] TRUE
```

# Testes lógicos

```
# Menor ou igual que  
num <= 50
```

```
## [1] TRUE
```

```
# Maior ou igual que  
num >= 35
```

```
## [1] TRUE
```

```
# Diferente  
num != 35
```

```
## [1] FALSE
```

# Testes lógicos

- Estes testes também funcionam com vetores:

```
idades
```

```
## [1] 15 18 35 24 17 80 67 45
```

```
idades > 15
```

```
## [1] FALSE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
```

## Usando mais de uma condição

- E: &
- OU: |

```
idades > 15 & idades < 18
```

```
idades <= 35 | idades < 12
```



# Exercícios

- (1) Crie e imprima um objeto chamado `nome` que recebe seu nome.
- (2) Crie e imprima um objeto chamado `idade` que recebe sua idade.
- (3) Crie e imprima um objeto chamado `cidade` que recebe a cidade onde você nasceu.
- (4) Crie e imprima um objeto chamado `matriculado` que receba o valor lógico `TRUE` se você estiver matriculado no **Curso de Introdução à Estatística com R**, ou `FALSE`, caso contrário.
- (5) Crie dois vetores (`nome` e `idade`): um com 10 nomes de amigos ou familiares, e outro com as idades correspondentes a cada um deles. Teste a condição da idade ser maior que 35 anos. Depois, identifique quais são estas pessoas imprimindo seus nomes.
- (6) Crie dois vetores (`peso` (Kg) e `altura` (m)): referente aos nomes criados anteriormente, ou seja, ambos os vetores precisam ter tamanho 10. Calcule o IMC usando estes dois vetores. Lembre que o IMC é calculado como:

$$IMC = \frac{peso}{altura^2}$$

Quantas e quais pessoas estão com **Sobrepeso**? (IMC entre 25 e 29.9).