Estatística Computacional I

Lupércio França Bessegato Dep. de Estatística/UFJF **A**fjf.

Roteiro Geral



- 1. Programando em R
- 2. Preparação, limpeza e manipulação de dados
- 3. Gráficos em R
- 4. Tópicos especiais
- 5. Referências

- 2

Programando em R

Codificação



Debbuging



256

258

- Em geral, cometemos erros na construção de códigos com muitas funções
- Exemplos:

√ Mensagem de erro

Error in "a" + 1 : argumento não-numérico para operador binário √ Aviso:

> 1:3 + 1:2 [1] 2 4 4 Warning message: In 1:3 + 1:2 :

comprimento do objeto maior não é múltiplo do comprimento do objeto menor

- Não ignore!

• Sinal de que algo não funciona como pretendido

Estatística Computacional I - 2020



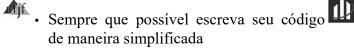
Prevenção de Erros

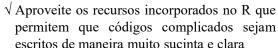


- Construa seu código em módulos
 - √ Não se repita
 - Crie funções para executar tarefas específicas e as execute quando necessário
 - Use o mesmo código, caso necessite realizar a mesma tarefa várias vezes
 - É mais fácil detectar e corrigir quaisquer erros.

Estatística Computacional I - 2020







- Ex.: família apply
- √Aprenda esses recursos e tente aplicá-los sempre

Estatística Computacional I - 2020



· Comente seu código



- √Para qualquer pessoa entender o que ele pretende fazer
 - Muitas vezes será você quem se esquecerá do que estava fazendo quando o escreveu
- √Dê nomes descritivos para as variáveis e funções
 - Contextualizados ao problema

259



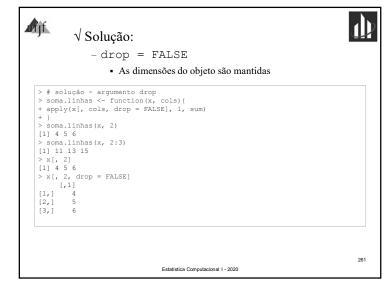
• Pense sempre nos casos especiais



√Por exemplo, um erro comum é não pensar sobre o que acontecerá quando sua entrada for de comprimento 1 ou 0

```
> # Função para somar subconjunto de colunas de uma matriz
> soma.linhas <- function(x, cols){
+ apply(x[, cols], 1, sum)
> (x <- matrix(1:9, ncol = 3))
  [,1] [,2] [,3]
[1,] 1 4
[2,] 2 5 8
> soma.linhas(x, 2:3)
[1] 11 13 15
> soma.linhas(x, 2)]
Erro: ']' inesperado in "soma.linhas(x, 2)]"
> dim(x[, 2])
NULL
> is.vector(x[, 2])
[1] TRUE
```

Estatística Computacional I - 2020





Codificação em módulos:



262

√R é uma linguagem funcional

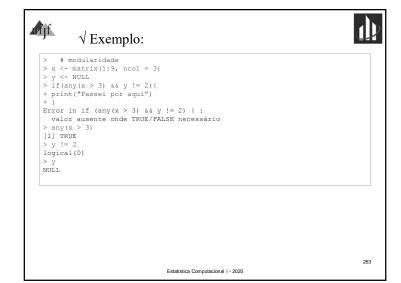
- As operações são realizadas por funções, que geralmente são independentes

 $\sqrt{\dot{E}}$ boa prática manter seu código modular:

- Ser composto de funções distintas que executam pequenas tarefas
- É possível testá-las (e depurá-las) individualmente

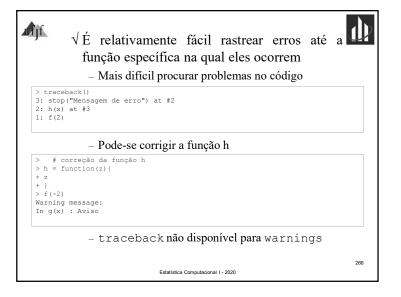
√ Instrução complexa que não funciona:

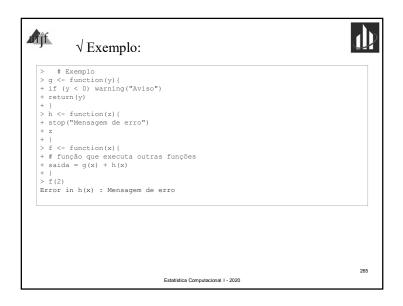
- Divida-a em partes e verifique se cada uma delas está fazendo o que você pretende



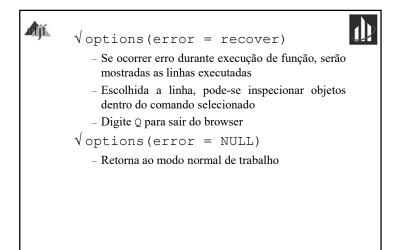


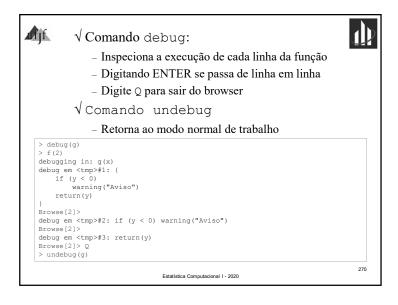
- √ Redução da complexidade de códigos mais complicados:
 - Escrever funções separadas, responsáveis por parcelas de seu programa
 - Testá-las separadamente com relação a resultados e erros Estes podem então ser testados para erros separadamente
 - Facilita sua reutilização em outras circunstâncias

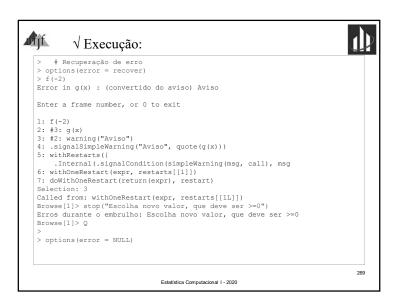


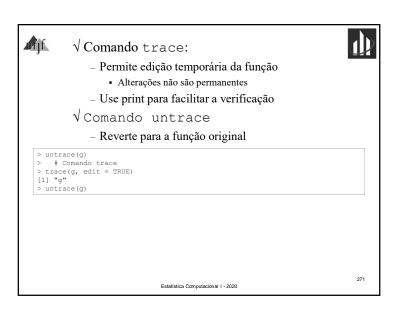


```
AFIF.
              - Pode-se transformar warning em error, com o
                 propósito de debugar a função
  > options(warn = 2)
  > f(-2)
  Error in g(x) : (convertido do aviso) Aviso
  > traceback()
  7: doWithOneRestart(return(expr), restart)
  6: withOneRestart(expr, restarts[[1L]])
  5: withRestarts({
         .Internal(.signalCondition(simpleWarning(msg, call), msg,
            call))
         .Internal(.dfltWarn(msg, call))
    ), muffleWarning = function() NULL)
  4: .signalSimpleWarning("Aviso", quote(g(x)))
  3: warning("Aviso") at #2
  2: g(x) at #3
  1: f(-2)
              - Recomendável inserção de stop na função g
                                                                           267
                               Estatística Computacional I - 2020
```









Referências



Bibliografia Recomendada



- ALBERT, J.; RIZZO, M. R by Example. Springer, 2012.
- CHRISTIAN, N. Basic Programming, Lecture Notes
- DALGAARD, P. *Introductory statistics with R*. Springer, 2008.
- KLEIBER, C.; ZEILEIS, A. Applied econometrics with R. Springer, 2008.
- GARDENER, M. Beginning R: The statistical programming language. John Wiley & Sons, 2012.