

Exercícios 2

James R. Hunter, Ph.D.

11 de novembro de 2024

Problema 1

Qualquer linguagem de programação é muito precisa. Detalhes contam. Aqui são três exemplos de código com erros. Quais são os erros e o que seria o código correto? Esses exemplos vêm do livro **R for Data Science**.

Exemplo A

```
my_variable <- 10  
my_variable
```

Porque este código não vai computar? Dica: uma dos autores é Turca.

Exemplo B

```
library(tidyverse)  
fliter(mpg, cyl = 8)
```

Aqui, quantos erros tem?

Exemplo C

```
filter(diamond, carat > 3)
```

“The devil is in the details.”

Problema 2

Utilize o código seguinte para calcular sua resposta.

Primeiro, faça o download de “einstein_soro.rds” para sua pasta de R.

```
soro <- readRDS("einstein_soro.rds")
```

- Esta conjunto de dados tem quantos casos?
- Esta conjunto de dados tem quantas variáveis?
- O que é o “class” e o “type” de soro?
- O que é a data de coleção (dt_collect) do terceiro caso?
- Imprima os quinta até décimo anos de nascimento (birth_yr).
- Quantas cidades diferentes existem nesta amostra?

Problema 3

O arquivo `trplasma.csv` mostra todas as mutações que uma amostra dos pacientes com HIV em falha virológica tiveram no gene transcriptase reversa do vírus. Um “1” na célula da planilha indica que a mutação estava presente e “0” demonstra ausência da mutação naquele momento.

A primeira coluna é o número de código do paciente e todos as outras colunas representam as mutações. Os nomes das colunas podem ser interpretados como:

- “tr” para *transcriptase reversa*,
 - o número seguinte como a posição do aminoácido (codon) com a mutação
 - o código para semana de exame de sangue (“bl” = *baseline* ou “_12” = 12 semanas).
- a. Primeira tarefa: importar o arquivo usando `readr::read_csv()` para um tibble chamada `trplas`.
 - b. Segunda tarefa: Transforme os dados para um conjunto *tidy*, usando as funções de `tidyr`.

Dica: planeje a transformação. Quais são as variáveis? Como vai dividir o código da mutação? Desenhe no papel como vai aparecer o tibble.

Problema 4

Carregue na memória o arquivo “`pac_demo.xlsx`” como `pac_data`, e responda à perguntas seguintes.

- a. Primeira tarefa: Quando falamos de carga viral de HIV, normalmente falamos de uma transformação das cópias do vírus em \log_{10} . Usando a função `log10()`, crie uma nova variável da valor logaritmico de carga viral, `log_cv` e salvar ele de volta para `pac_data`.
- b. Segunda tarefa: Com a variável de escolaridade (`esco1`), mude a valor para ser um factor com 4 níveis e mostra uma tabela com os níveis na ordem correta.