

Lista 6

Marcos Prates

17 de maio de 2024

- Forma de entrega: Mandar pelo Moodle os arquivos necessários para resolução da lista de exercícios.
- Salvar arquivo com nome: “seunome_Lista6.zip”

1. Populações do Censo 2010 - Parte 1

- (a) **Leia** as bases de dados por estado(disponível na pasta dados)
- (b) Junte-as em um único arquivo
- (c) Observe que a base possui a população **segmentada**
- (d) Crie uma base auxiliar contendo as populações por município
- (e) Descubra quais são os municípios com mais de 500.000 habitantes
- (f) Restrinja a base inicial aos municípios encontrados no item 5
- (g) Crie 4 bases auxiliares
 - i. Uma contendo as populações dos município por sexo
 - ii. Uma contendo as populações dos município por etnia
 - iii. Uma contendo as populações dos município por analfabetismo
 - iv. Uma contendo as populações dos município por situação do domicílio
- (h) O item (a) deve ser realizado sequencialmente e em paralelo
- (i) Valeu a pena paralelizar o código?

2. Populações do Censo 2010 - Parte 2

- (a) Para cada município do item 5 da parte 1 crie um relatório
- (b) Os parâmetros do relatório são (para cada cidade):
 - i. **cidade** - Nome da cidade em análise
 - ii. **estado** - Sigla do estado em análise
 - iii. **pop_sexo** - Base contendo a população do sexo Feminino e Masculino
 - iv. **pop_etnia** - Base contendo a população Amarela, Branca, Ignorado, Indígena, Parda, Preta
 - v. **pop_analf** - Base contendo a população Analfabeta (Não), Alfabetizada (Sim) e NA
 - vi. **pop_situacao** - Base contendo a população Rural e Urbana
- (c) O item (a) deve ser realizado sequencialmente e em paralelo
- (d) Há um ganho em paralelizar o código?

Dicas: Para listar arquivos em um diretório pense em:

- `?list.files`

Sempre que o objetivo for calcular a **população agregada** por algum fator pense em:

- `?aggregate`
- `?dplyr::group_by`

Sempre que o objetivo for restringir a base pense em:

- `?subset`
- `?dplyr::filter`

3. Gerando um relatório

```

rmarkdown::render(input = "relatorio/relatorio.Rmd",
  output_dir = "relatorio/cidades/sequencial/",
  output_file = paste0(4322004, ".html"),
  intermediates_dir = tempfile(),
  params = list(estado = "RS",
    cidade = "Triunfo",
    pop_sexo = data_sexo,
    pop_etnia = data_etnia,
    pop_analf = data_analf,
    pop_situacao = data_situacao))

```

- `data_sexo` - data.frame contendo as populações por sexo da cidade em questão
- `data_etnia` - data.frame contendo as populações por etnia da cidade em questão
- `data_analf` - data.frame contendo as populações por níveis de analfabetismo da cidade em questão
- `data_situacao` - data.frame contendo as populações urbana e rural da cidade em questão

Exemplo de dado de entrada para o relatório (`data_sexo`):

```

##  COD_UF NO_UF COD_MUN  NO_MUN    SEXO    POP
## 1    23    CE 2304400 Fortaleza Feminino 1304267
## 2    23    CE 2304400 Fortaleza Masculino 1147918

```