

Desafio 09

Rafael Gomes Carneiro - RA185462

SQLite - Processamento de Dados em Lote

Benilton Carvalho, Guilherme Ludwig

SQLite

SQLite é uma aplicação de banco de dados de pouquíssimo impacto e com um potencial de aplicação bastante amplo. Habitualmente, em ambientes de produção que exijam alta performance ou operações mais complexas, SQLite não é a ferramenta padrão de escolha. Para aplicações de menor porte, o SQLite é uma ferramenta de grande importância, pois adequa-se aos padrões de SQL e é multiplataforma.

Objetivo

Ao fim deste laboratório, você deverá ser capaz de:

- Usar a estratégia de leitura por chunks implementada no pacote `readr` para leitura dos dados de vôos;
- Utilizar a função `SideEffectChunkCallback$new()` para depositar os dados do arquivo diretamente no banco de dados (sem devolver resultados para o usuário);

Recomendações

- *Não* utilize o servidor `jupyter.ime.unicamp.br` para executar esta atividade;
- Instale os pacotes `readr` e `RSQLite`, se necessário. A sugestão é empregar `install.packages(c("readr", "RSQLite"), dep=TRUE, type='win.binary')`.
- Atente para mensagens de erro adicionais (como outros pacotes faltantes) e instale-os conforme requerido.

Atividades

```
library(readr)
library(RSQLite)
library(DBI)
```

1. Crie um arquivo de banco de dados em SQLite chamado `voos.sqlite3`. (Dica: o comando `dbConnect()` se conecta num banco de dados se o arquivo apontado existir ou cria um novo, caso o arquivo não exista.)

```
voos <- dbConnect(SQLite(), '../dados/voos.sqlite3')

dbListTables(voos)
```

```
[1] "airlines" "airports" "flights"
```

2. Leia os arquivos `airlines.csv` e `airports.csv`. Deposite o conteúdo de cada um destes arquivos nas tabelas, respectivamente, `airlines` e `airports`. Utilize o comando `dbWriteTable()` para isso.

```
# Baixando os arquivos
airlines <- read_csv("../dados/airlines.csv")
airports <- read_csv("../dados/airports.csv")

# Aplicando como tabelas
dbWriteTable(voos, "airlines", airlines, overwrite = TRUE)
dbWriteTable(voos, "airports", airports, overwrite = TRUE)
```

3. Crie uma função chamada `lerDados` contendo 2 argumentos, `input` e `pos`. A função deve apresentar ao usuário uma mensagem de progresso da leitura do arquivo `flights.csv` (utilize o comando `message()`), aos moldes do apresentado abaixo. A função deve salvar apenas os vôos que partiram ou chegaram aos seguintes aeroportos BWI, MIA, SEA, SFO e JFK, numa tabela chamada `flights`. Observe que a função não deve retornar nada para o usuário, deve apenas gravar a tabela obtida do chunk no banco de dados. (Dica: utilize o comando `dbWriteTable()` e estude como o argumento `append` deve ser utilizado para permitir que os *chunks* intermediários sejam adicionados ao fim da tabela.)

```
# Simulando um valor para 'pos'
pos = 1000
message("Leitura atingiu a linha ", pos)
```

```

lerDados <- function(input, pos) {
  # Mostrar progresso
  message("Leitura atingiu a linha ", pos)

  # Filtrar voos de/para os aeroportos desejados
  aeroportos_interesse <- c("BWI", "MIA", "SEA", "SFO", "JFK")

  dados_filtrados <- subset(
    input,
    ORIGIN_AIRPORT %in% aeroportos_interesse | DESTINATION_AIRPORT %in% aeroportos_interesse
  )

  # Gravar no banco (acrescentando aos dados já existentes)
  dbWriteTable(voos, "flights", dados_filtrados, append = TRUE)

  # Não retorna nada
  invisible(NULL)
}

```

```

flights <- read_csv_chunked(
  "../dados/flights.csv",
  callback = SideEffectChunkCallback$new(lerDados),
  chunk_size = 10000
)

```

```

-- Column specification -----
cols(
  .default = col_double(),
  AIRLINE = col_character(),
  TAIL_NUMBER = col_character(),
  ORIGIN_AIRPORT = col_character(),
  DESTINATION_AIRPORT = col_character(),
  SCHEDULED_DEPARTURE = col_character(),
  DEPARTURE_TIME = col_character(),
  WHEELS_OFF = col_character(),
  WHEELS_ON = col_character(),
  SCHEDULED_ARRIVAL = col_character(),
  ARRIVAL_TIME = col_character(),
  CANCELLATION_REASON = col_character()
)
i Use `spec()` for the full column specifications.

```

Leitura atingiu a linha 1

Leitura atingiu a linha 10001

Leitura atingiu a linha 20001

Leitura atingiu a linha 30001

Leitura atingiu a linha 40001

Leitura atingiu a linha 50001

Leitura atingiu a linha 60001

Leitura atingiu a linha 70001

Leitura atingiu a linha 80001

Leitura atingiu a linha 90001

Leitura atingiu a linha 100001

Leitura atingiu a linha 110001

Leitura atingiu a linha 120001

Leitura atingiu a linha 130001

Leitura atingiu a linha 140001

Leitura atingiu a linha 150001

Leitura atingiu a linha 160001

Leitura atingiu a linha 170001

Leitura atingiu a linha 180001

Leitura atingiu a linha 190001

Leitura atingiu a linha 200001

Leitura atingiu a linha 210001

Leitura atingiu a linha 220001

Leitura atingiu a linha 230001

Leitura atingiu a linha 240001

Leitura atingiu a linha 250001

Leitura atingiu a linha 260001

Leitura atingiu a linha 270001

Leitura atingiu a linha 280001

Leitura atingiu a linha 290001

Leitura atingiu a linha 300001

Leitura atingiu a linha 310001

Leitura atingiu a linha 320001

Leitura atingiu a linha 330001

Leitura atingiu a linha 340001

Leitura atingiu a linha 350001

Leitura atingiu a linha 360001

Leitura atingiu a linha 370001

Leitura atingiu a linha 380001

Leitura atingiu a linha 390001

Leitura atingiu a linha 400001

Leitura atingiu a linha 410001

Leitura atingiu a linha 420001

Leitura atingiu a linha 430001

Leitura atingiu a linha 440001

Leitura atingiu a linha 450001

Leitura atingiu a linha 460001

Leitura atingiu a linha 470001

Leitura atingiu a linha 480001

Leitura atingiu a linha 490001

Leitura atingiu a linha 500001

Leitura atingiu a linha 510001

Leitura atingiu a linha 520001

Leitura atingiu a linha 530001

Leitura atingiu a linha 540001

Leitura atingiu a linha 550001

Leitura atingiu a linha 560001

Leitura atingiu a linha 570001

Leitura atingiu a linha 580001

Leitura atingiu a linha 590001

Leitura atingiu a linha 600001

Leitura atingiu a linha 610001

Leitura atingiu a linha 620001

Leitura atingiu a linha 630001

Leitura atingiu a linha 640001

Leitura atingiu a linha 650001

Leitura atingiu a linha 660001

Leitura atingiu a linha 670001

Leitura atingiu a linha 680001

Leitura atingiu a linha 690001

Leitura atingiu a linha 700001

Leitura atingiu a linha 710001

Leitura atingiu a linha 720001

Leitura atingiu a linha 730001

Leitura atingiu a linha 740001

Leitura atingiu a linha 750001

Leitura atingiu a linha 760001

Leitura atingiu a linha 770001

Leitura atingiu a linha 780001

Leitura atingiu a linha 790001

Leitura atingiu a linha 800001

Leitura atingiu a linha 810001

Leitura atingiu a linha 820001

Leitura atingiu a linha 830001

Leitura atingiu a linha 840001

Leitura atingiu a linha 850001

Leitura atingiu a linha 860001

Leitura atingiu a linha 870001

Leitura atingiu a linha 880001

Leitura atingiu a linha 890001

Leitura atingiu a linha 900001

Leitura atingiu a linha 910001

Leitura atingiu a linha 920001

Leitura atingiu a linha 930001

Leitura atingiu a linha 940001

Leitura atingiu a linha 950001

Leitura atingiu a linha 960001

Leitura atingiu a linha 970001

Leitura atingiu a linha 980001

Leitura atingiu a linha 990001

Leitura atingiu a linha 1000001

Leitura atingiu a linha 1010001

Leitura atingiu a linha 1020001

Leitura atingiu a linha 1030001

Leitura atingiu a linha 1040001

Leitura atingiu a linha 1050001

Leitura atingiu a linha 1060001

Leitura atingiu a linha 1070001

Leitura atingiu a linha 1080001

Leitura atingiu a linha 1090001

Leitura atingiu a linha 1100001

Leitura atingiu a linha 1110001

Leitura atingiu a linha 1120001

Leitura atingiu a linha 1130001

Leitura atingiu a linha 1140001

Leitura atingiu a linha 1150001

Leitura atingiu a linha 1160001

Leitura atingiu a linha 1170001

Leitura atingiu a linha 1180001

Leitura atingiu a linha 1190001

Leitura atingiu a linha 1200001

Leitura atingiu a linha 1210001

Leitura atingiu a linha 1220001

Leitura atingiu a linha 1230001

Leitura atingiu a linha 1240001

Leitura atingiu a linha 1250001

Leitura atingiu a linha 1260001

Leitura atingiu a linha 1270001

Leitura atingiu a linha 1280001

Leitura atingiu a linha 1290001

Leitura atingiu a linha 1300001

Leitura atingiu a linha 1310001

Leitura atingiu a linha 1320001

Leitura atingiu a linha 1330001

Leitura atingiu a linha 1340001

Leitura atingiu a linha 1350001

Leitura atingiu a linha 1360001

Leitura atingiu a linha 1370001

Leitura atingiu a linha 1380001

Leitura atingiu a linha 1390001

Leitura atingiu a linha 1400001

Leitura atingiu a linha 1410001

Leitura atingiu a linha 1420001

Leitura atingiu a linha 1430001

Leitura atingiu a linha 1440001

Leitura atingiu a linha 1450001

Leitura atingiu a linha 1460001

Leitura atingiu a linha 1470001

Leitura atingiu a linha 1480001

Leitura atingiu a linha 1490001

Leitura atingiu a linha 1500001

Leitura atingiu a linha 1510001

Leitura atingiu a linha 1520001

Leitura atingiu a linha 1530001

Leitura atingiu a linha 1540001

Leitura atingiu a linha 1550001

Leitura atingiu a linha 1560001

Leitura atingiu a linha 1570001

Leitura atingiu a linha 1580001

Leitura atingiu a linha 1590001

Leitura atingiu a linha 1600001

Leitura atingiu a linha 1610001

Leitura atingiu a linha 1620001

Leitura atingiu a linha 1630001

Leitura atingiu a linha 1640001

Leitura atingiu a linha 1650001

Leitura atingiu a linha 1660001

Leitura atingiu a linha 1670001

Leitura atingiu a linha 1680001

Leitura atingiu a linha 1690001

Leitura atingiu a linha 1700001

Leitura atingiu a linha 1710001

Leitura atingiu a linha 1720001

Leitura atingiu a linha 1730001

Leitura atingiu a linha 1740001

Leitura atingiu a linha 1750001

Leitura atingiu a linha 1760001

Leitura atingiu a linha 1770001

Leitura atingiu a linha 1780001

Leitura atingiu a linha 1790001

Leitura atingiu a linha 1800001

Leitura atingiu a linha 1810001

Leitura atingiu a linha 1820001

Leitura atingiu a linha 1830001

Leitura atingiu a linha 1840001

Leitura atingiu a linha 1850001

Leitura atingiu a linha 1860001

Leitura atingiu a linha 1870001

Leitura atingiu a linha 1880001

Leitura atingiu a linha 1890001

Leitura atingiu a linha 1900001

Leitura atingiu a linha 1910001

Leitura atingiu a linha 1920001

Leitura atingiu a linha 1930001

Leitura atingiu a linha 1940001

Leitura atingiu a linha 1950001

Leitura atingiu a linha 1960001

Leitura atingiu a linha 1970001

Leitura atingiu a linha 1980001

Leitura atingiu a linha 1990001

Leitura atingiu a linha 2000001

Leitura atingiu a linha 2010001

Leitura atingiu a linha 2020001

Leitura atingiu a linha 2030001

Leitura atingiu a linha 2040001

Leitura atingiu a linha 2050001

Leitura atingiu a linha 2060001

Leitura atingiu a linha 2070001

Leitura atingiu a linha 2080001

Leitura atingiu a linha 2090001

Leitura atingiu a linha 2100001

Leitura atingiu a linha 2110001

Leitura atingiu a linha 2120001

Leitura atingiu a linha 2130001

Leitura atingiu a linha 2140001

Leitura atingiu a linha 2150001

Leitura atingiu a linha 2160001

Leitura atingiu a linha 2170001

Leitura atingiu a linha 2180001

Leitura atingiu a linha 2190001

Leitura atingiu a linha 2200001

Leitura atingiu a linha 2210001

Leitura atingiu a linha 2220001

Leitura atingiu a linha 2230001

Leitura atingiu a linha 2240001

Leitura atingiu a linha 2250001

Leitura atingiu a linha 2260001

Leitura atingiu a linha 2270001

Leitura atingiu a linha 2280001

Leitura atingiu a linha 2290001

Leitura atingiu a linha 2300001

Leitura atingiu a linha 2310001

Leitura atingiu a linha 2320001

Leitura atingiu a linha 2330001

Leitura atingiu a linha 2340001

Leitura atingiu a linha 2350001

Leitura atingiu a linha 2360001

Leitura atingiu a linha 2370001

Leitura atingiu a linha 2380001

Leitura atingiu a linha 2390001

Leitura atingiu a linha 2400001

Leitura atingiu a linha 2410001

Leitura atingiu a linha 2420001

Leitura atingiu a linha 2430001

Leitura atingiu a linha 2440001

Leitura atingiu a linha 2450001

Leitura atingiu a linha 2460001

Leitura atingiu a linha 2470001

Leitura atingiu a linha 2480001

Leitura atingiu a linha 2490001

Leitura atingiu a linha 2500001

Leitura atingiu a linha 2510001

Leitura atingiu a linha 2520001

Leitura atingiu a linha 2530001

Leitura atingiu a linha 2540001

Leitura atingiu a linha 2550001

Leitura atingiu a linha 2560001

Leitura atingiu a linha 2570001

Leitura atingiu a linha 2580001

Leitura atingiu a linha 2590001

Leitura atingiu a linha 2600001

Leitura atingiu a linha 2610001

Leitura atingiu a linha 2620001

Leitura atingiu a linha 2630001

Leitura atingiu a linha 2640001

Leitura atingiu a linha 2650001

Leitura atingiu a linha 2660001

Leitura atingiu a linha 2670001

Leitura atingiu a linha 2680001

Leitura atingiu a linha 2690001

Leitura atingiu a linha 2700001

Leitura atingiu a linha 2710001

Leitura atingiu a linha 2720001

Leitura atingiu a linha 2730001

Leitura atingiu a linha 2740001

Leitura atingiu a linha 2750001

Leitura atingiu a linha 2760001

Leitura atingiu a linha 2770001

Leitura atingiu a linha 2780001

Leitura atingiu a linha 2790001

Leitura atingiu a linha 2800001

Leitura atingiu a linha 2810001

Leitura atingiu a linha 2820001

Leitura atingiu a linha 2830001

Leitura atingiu a linha 2840001

Leitura atingiu a linha 2850001

Leitura atingiu a linha 2860001

Leitura atingiu a linha 2870001

Leitura atingiu a linha 2880001

Leitura atingiu a linha 2890001

Leitura atingiu a linha 2900001

Leitura atingiu a linha 2910001

Leitura atingiu a linha 2920001

Leitura atingiu a linha 2930001

Leitura atingiu a linha 2940001

Leitura atingiu a linha 2950001

Leitura atingiu a linha 2960001

Leitura atingiu a linha 2970001

Leitura atingiu a linha 2980001

Leitura atingiu a linha 2990001

Leitura atingiu a linha 3000001

Leitura atingiu a linha 3010001

Leitura atingiu a linha 3020001

Leitura atingiu a linha 3030001

Leitura atingiu a linha 3040001

Leitura atingiu a linha 3050001

Leitura atingiu a linha 3060001

Leitura atingiu a linha 3070001

Leitura atingiu a linha 3080001

Leitura atingiu a linha 3090001

Leitura atingiu a linha 3100001

Leitura atingiu a linha 3110001

Leitura atingiu a linha 3120001

Leitura atingiu a linha 3130001

Leitura atingiu a linha 3140001

Leitura atingiu a linha 3150001

Leitura atingiu a linha 3160001

Leitura atingiu a linha 3170001

Leitura atingiu a linha 3180001

Leitura atingiu a linha 3190001

Leitura atingiu a linha 3200001

Leitura atingiu a linha 3210001

Leitura atingiu a linha 3220001

Leitura atingiu a linha 3230001

Leitura atingiu a linha 3240001

Leitura atingiu a linha 3250001

Leitura atingiu a linha 3260001

Leitura atingiu a linha 3270001

Leitura atingiu a linha 3280001

Leitura atingiu a linha 3290001

Leitura atingiu a linha 3300001

Leitura atingiu a linha 3310001

Leitura atingiu a linha 3320001

Leitura atingiu a linha 3330001

Leitura atingiu a linha 3340001

Leitura atingiu a linha 3350001

Leitura atingiu a linha 3360001

Leitura atingiu a linha 3370001

Leitura atingiu a linha 3380001

Leitura atingiu a linha 3390001

Leitura atingiu a linha 3400001

Leitura atingiu a linha 3410001

Leitura atingiu a linha 3420001

Leitura atingiu a linha 3430001

Leitura atingiu a linha 3440001

Leitura atingiu a linha 3450001

Leitura atingiu a linha 3460001

Leitura atingiu a linha 3470001

Leitura atingiu a linha 3480001

Leitura atingiu a linha 3490001

Leitura atingiu a linha 3500001

Leitura atingiu a linha 3510001

Leitura atingiu a linha 3520001

Leitura atingiu a linha 3530001

Leitura atingiu a linha 3540001

Leitura atingiu a linha 3550001

Leitura atingiu a linha 3560001

Leitura atingiu a linha 3570001

Leitura atingiu a linha 3580001

Leitura atingiu a linha 3590001

Leitura atingiu a linha 3600001

Leitura atingiu a linha 3610001

Leitura atingiu a linha 3620001

Leitura atingiu a linha 3630001

Leitura atingiu a linha 3640001

Leitura atingiu a linha 3650001

Leitura atingiu a linha 3660001

Leitura atingiu a linha 3670001

Leitura atingiu a linha 3680001

Leitura atingiu a linha 3690001

Leitura atingiu a linha 3700001

Leitura atingiu a linha 3710001

Leitura atingiu a linha 3720001

Leitura atingiu a linha 3730001

Leitura atingiu a linha 3740001

Leitura atingiu a linha 3750001

Leitura atingiu a linha 3760001

Leitura atingiu a linha 3770001

Leitura atingiu a linha 3780001

Leitura atingiu a linha 3790001

Leitura atingiu a linha 3800001

Leitura atingiu a linha 3810001

Leitura atingiu a linha 3820001

Leitura atingiu a linha 3830001

Leitura atingiu a linha 3840001

Leitura atingiu a linha 3850001

Leitura atingiu a linha 3860001

Leitura atingiu a linha 3870001

Leitura atingiu a linha 3880001

Leitura atingiu a linha 3890001

Leitura atingiu a linha 3900001

Leitura atingiu a linha 3910001

Leitura atingiu a linha 3920001

Leitura atingiu a linha 3930001

Leitura atingiu a linha 3940001

Leitura atingiu a linha 3950001

Leitura atingiu a linha 3960001

Leitura atingiu a linha 3970001

Leitura atingiu a linha 3980001

Leitura atingiu a linha 3990001

Leitura atingiu a linha 4000001

Leitura atingiu a linha 4010001

Leitura atingiu a linha 4020001

Leitura atingiu a linha 4030001

Leitura atingiu a linha 4040001

Leitura atingiu a linha 4050001

Leitura atingiu a linha 4060001

Leitura atingiu a linha 4070001

Leitura atingiu a linha 4080001

Leitura atingiu a linha 4090001

Leitura atingiu a linha 4100001

Leitura atingiu a linha 4110001

Leitura atingiu a linha 4120001

Leitura atingiu a linha 4130001

Leitura atingiu a linha 4140001

Leitura atingiu a linha 4150001

Leitura atingiu a linha 4160001

Leitura atingiu a linha 4170001

Leitura atingiu a linha 4180001

Leitura atingiu a linha 4190001

Leitura atingiu a linha 4200001

Leitura atingiu a linha 4210001

Leitura atingiu a linha 4220001

Leitura atingiu a linha 4230001

Leitura atingiu a linha 4240001

Leitura atingiu a linha 4250001

Leitura atingiu a linha 4260001

Leitura atingiu a linha 4270001

Leitura atingiu a linha 4280001

Leitura atingiu a linha 4290001

Leitura atingiu a linha 4300001

Leitura atingiu a linha 4310001

Leitura atingiu a linha 4320001

Leitura atingiu a linha 4330001

Leitura atingiu a linha 4340001

Leitura atingiu a linha 4350001

Leitura atingiu a linha 4360001

Leitura atingiu a linha 4370001

Leitura atingiu a linha 4380001

Leitura atingiu a linha 4390001

Leitura atingiu a linha 4400001

Leitura atingiu a linha 4410001

Leitura atingiu a linha 4420001

Leitura atingiu a linha 4430001

Leitura atingiu a linha 4440001

Leitura atingiu a linha 4450001

Leitura atingiu a linha 4460001

Leitura atingiu a linha 4470001

Leitura atingiu a linha 4480001

Leitura atingiu a linha 4490001

Leitura atingiu a linha 4500001

Leitura atingiu a linha 4510001

Leitura atingiu a linha 4520001

Leitura atingiu a linha 4530001

Leitura atingiu a linha 4540001

Leitura atingiu a linha 4550001

Leitura atingiu a linha 4560001

Leitura atingiu a linha 4570001

Leitura atingiu a linha 4580001

Leitura atingiu a linha 4590001

Leitura atingiu a linha 4600001

Leitura atingiu a linha 4610001

Leitura atingiu a linha 4620001

Leitura atingiu a linha 4630001

Leitura atingiu a linha 4640001

Leitura atingiu a linha 4650001

Leitura atingiu a linha 4660001

Leitura atingiu a linha 4670001

Leitura atingiu a linha 4680001

Leitura atingiu a linha 4690001

Leitura atingiu a linha 4700001

Leitura atingiu a linha 4710001

Leitura atingiu a linha 4720001

Leitura atingiu a linha 4730001

Leitura atingiu a linha 4740001

Leitura atingiu a linha 4750001

Leitura atingiu a linha 4760001

Leitura atingiu a linha 4770001

Leitura atingiu a linha 4780001

Leitura atingiu a linha 4790001

Leitura atingiu a linha 4800001

Leitura atingiu a linha 4810001

Leitura atingiu a linha 4820001

Leitura atingiu a linha 4830001

Leitura atingiu a linha 4840001

Leitura atingiu a linha 4850001

Leitura atingiu a linha 4860001

Leitura atingiu a linha 4870001

Leitura atingiu a linha 4880001

Leitura atingiu a linha 4890001

Leitura atingiu a linha 4900001

Leitura atingiu a linha 4910001

Leitura atingiu a linha 4920001

Leitura atingiu a linha 4930001

Leitura atingiu a linha 4940001

Leitura atingiu a linha 4950001

Leitura atingiu a linha 4960001

Leitura atingiu a linha 4970001

Leitura atingiu a linha 4980001

Leitura atingiu a linha 4990001

Leitura atingiu a linha 5000001

Leitura atingiu a linha 5010001

Leitura atingiu a linha 5020001

Leitura atingiu a linha 5030001

Leitura atingiu a linha 5040001

Leitura atingiu a linha 5050001

Leitura atingiu a linha 5060001

Leitura atingiu a linha 5070001

Leitura atingiu a linha 5080001

Leitura atingiu a linha 5090001

Leitura atingiu a linha 5100001

Leitura atingiu a linha 5110001

Leitura atingiu a linha 5120001

Leitura atingiu a linha 5130001

Leitura atingiu a linha 5140001

Leitura atingiu a linha 5150001

Leitura atingiu a linha 5160001

Leitura atingiu a linha 5170001

Leitura atingiu a linha 5180001

Leitura atingiu a linha 5190001

Leitura atingiu a linha 5200001

Leitura atingiu a linha 5210001

Leitura atingiu a linha 5220001

Leitura atingiu a linha 5230001

Leitura atingiu a linha 5240001

Leitura atingiu a linha 5250001

Leitura atingiu a linha 5260001

Leitura atingiu a linha 5270001

Leitura atingiu a linha 5280001

Leitura atingiu a linha 5290001

Leitura atingiu a linha 5300001

Leitura atingiu a linha 5310001

Leitura atingiu a linha 5320001

Leitura atingiu a linha 5330001

Leitura atingiu a linha 5340001

Leitura atingiu a linha 5350001

Leitura atingiu a linha 5360001

Leitura atingiu a linha 5370001

Leitura atingiu a linha 5380001

Leitura atingiu a linha 5390001

Leitura atingiu a linha 5400001

Leitura atingiu a linha 5410001

Leitura atingiu a linha 5420001

Leitura atingiu a linha 5430001

Leitura atingiu a linha 5440001

Leitura atingiu a linha 5450001

Leitura atingiu a linha 5460001

Leitura atingiu a linha 5470001

Leitura atingiu a linha 5480001

Leitura atingiu a linha 5490001

Leitura atingiu a linha 5500001

Leitura atingiu a linha 5510001

Leitura atingiu a linha 5520001

Leitura atingiu a linha 5530001

Leitura atingiu a linha 5540001

Leitura atingiu a linha 5550001

Leitura atingiu a linha 5560001

Leitura atingiu a linha 5570001

Leitura atingiu a linha 5580001

Leitura atingiu a linha 5590001

Leitura atingiu a linha 5600001

Leitura atingiu a linha 5610001

Leitura atingiu a linha 5620001

Leitura atingiu a linha 5630001

Leitura atingiu a linha 5640001

Leitura atingiu a linha 5650001

Leitura atingiu a linha 5660001

Leitura atingiu a linha 5670001

Leitura atingiu a linha 5680001

Leitura atingiu a linha 5690001

Leitura atingiu a linha 5700001

Leitura atingiu a linha 5710001

Leitura atingiu a linha 5720001

Leitura atingiu a linha 5730001

Leitura atingiu a linha 5740001

Leitura atingiu a linha 5750001

Leitura atingiu a linha 5760001

Leitura atingiu a linha 5770001

Leitura atingiu a linha 5780001

Leitura atingiu a linha 5790001

Leitura atingiu a linha 5800001

Leitura atingiu a linha 5810001

4. Leia o arquivo `flights.csv`, restringindo-se às colunas `YEAR`, `MONTH`, `DAY`, `AIRLINE`, `FLIGHT_NUMBER`, `ORIGIN_AIRPORT`, `DESTINATION_AIRPORT` e `ARRIVAL_DELAY`, e aplique a função `lerDados()` criada acima. Observe, novamente, que a função `lerDados()` não retorna nada para o usuário. Por isso, a função de callback a ser utilizada é `SideEffectChunkCallback$new()`. Leia 100 mil registros por vez.

```
read_csv_chunked(  
  "../dados/flights.csv",  
  callback = SideEffectChunkCallback$new(lerDados),  
  chunk_size = 100000,  
  col_types = cols(  
    YEAR = col_double(),  
    MONTH = col_double(),  
    DAY = col_double(),  
    AIRLINE = col_character(),
```

```
FLIGHT_NUMBER = col_double(),  
ORIGIN_AIRPORT = col_character(),  
DESTINATION_AIRPORT = col_character(),  
ARRIVAL_DELAY = col_double(),  
  .default = col_skip() # ignora todas as outras colunas  
)  
)
```

Leitura atingiu a linha 1

Leitura atingiu a linha 100001

Leitura atingiu a linha 200001

Leitura atingiu a linha 300001

Leitura atingiu a linha 400001

Leitura atingiu a linha 500001

Leitura atingiu a linha 600001

Leitura atingiu a linha 700001

Leitura atingiu a linha 800001

Leitura atingiu a linha 900001

Leitura atingiu a linha 1000001

Leitura atingiu a linha 1100001

Leitura atingiu a linha 1200001

Leitura atingiu a linha 1300001

Leitura atingiu a linha 1400001

Leitura atingiu a linha 1500001

Leitura atingiu a linha 1600001

Leitura atingiu a linha 1700001

Leitura atingiu a linha 1800001

Leitura atingiu a linha 1900001

Leitura atingiu a linha 2000001

Leitura atingiu a linha 2100001

Leitura atingiu a linha 2200001

Leitura atingiu a linha 2300001

Leitura atingiu a linha 2400001

Leitura atingiu a linha 2500001

Leitura atingiu a linha 2600001

Leitura atingiu a linha 2700001

Leitura atingiu a linha 2800001

Leitura atingiu a linha 2900001

Leitura atingiu a linha 3000001

Leitura atingiu a linha 3100001

Leitura atingiu a linha 3200001

Leitura atingiu a linha 3300001

Leitura atingiu a linha 3400001

Leitura atingiu a linha 3500001

Leitura atingiu a linha 3600001

Leitura atingiu a linha 3700001

Leitura atingiu a linha 3800001

Leitura atingiu a linha 3900001

Leitura atingiu a linha 4000001

Leitura atingiu a linha 4100001

Leitura atingiu a linha 4200001

Leitura atingiu a linha 4300001

Leitura atingiu a linha 4400001

Leitura atingiu a linha 4500001

Leitura atingiu a linha 4600001

Leitura atingiu a linha 4700001

Leitura atingiu a linha 4800001

Leitura atingiu a linha 4900001

Leitura atingiu a linha 5000001

Leitura atingiu a linha 5100001

Leitura atingiu a linha 5200001

Leitura atingiu a linha 5300001

Leitura atingiu a linha 5400001

Leitura atingiu a linha 5500001

Leitura atingiu a linha 5600001

Leitura atingiu a linha 5700001

Leitura atingiu a linha 5800001

NULL

5. Acesse o banco de dados e, por meio de uma chamada em SQL, apresente o tempo médio de atraso de chegada por aeroporto de destino, a sigla do aeroporto, o nome completo do aeroporto e o nome completo da companhia aérea. Ordene o resultado (na mesma chamada de SQL) por ordem decrescente deste atraso médio (i.e., o primeiro registro deve ser o aeroporto que tem o maior tempo de atraso na chegada). Atente para o fato de que o mesmo nome de coluna pode acontecer em diferentes tabelas.

```
# Consulta SQL
comando <- "
SELECT
  f.DESTINATION_AIRPORT AS sigla_aeroporto,
  a.AIRPORT AS nome_aeroporto,
  al.AIRLINE AS nome_companhia,
  AVG(f.ARRIVAL_DELAY) AS atraso_medio
FROM flights f
JOIN airports a
  ON f.DESTINATION_AIRPORT = a.IATA_CODE
JOIN airlines al
  ON f.AIRLINE = al.IATA_CODE
GROUP BY f.DESTINATION_AIRPORT, a.AIRPORT, al.AIRLINE
ORDER BY atraso_medio DESC
"

# Executar a consulta e armazenar resultado
resultado <- dbGetQuery(voos, comando)

# Mostrar resultado
head(resultado, 10) # por exemplo, os 10 primeiros
```

| | sigla_aeroporto | nome_aeroporto |
|----|-----------------|--|
| 1 | DTW | Detroit Metropolitan Airport |
| 2 | JAX | Jacksonville International Airport |
| 3 | IAH | George Bush Intercontinental Airport |
| 4 | MSN | Dane County Regional Airport |
| 5 | IAH | George Bush Intercontinental Airport |
| 6 | ATL | Hartsfield-Jackson Atlanta International Airport |
| 7 | LGA | LaGuardia Airport (Marine Air Terminal) |
| 8 | ORD | Chicago O'Hare International Airport |
| 9 | MTJ | Montrose Regional Airport |
| 10 | MSP | Minneapolis-Saint Paul International Airport |

| | nome_companhia | atraso_medio |
|----|------------------------|--------------|
| 1 | Skywest Airlines Inc. | 143.00000 |
| 2 | Delta Air Lines Inc. | 41.00000 |
| 3 | US Airways Inc. | 36.20000 |
| 4 | United Air Lines Inc. | 34.50000 |
| 5 | Skywest Airlines Inc. | 27.81749 |
| 6 | Spirit Air Lines | 26.14202 |
| 7 | Frontier Airlines Inc. | 24.43499 |
| 8 | Spirit Air Lines | 22.70203 |
| 9 | Skywest Airlines Inc. | 22.34328 |
| 10 | Spirit Air Lines | 22.24610 |

```
dbDisconnect(voos)

# Função para converter Sys.time() em horário legível
horario_normal <- function(timestamp = Sys.time()) {
  format(timestamp, "%d/%m/%Y %H:%M:%S")
}

cat("Arquivo compilado em:", horario_normal(), "\n")
```

Arquivo compilado em: 30/09/2025 12:16:59