Exercício 01 - EPI 90

Márcio Rodrigues

10/5/2020

```r
knitr::opts_chunk$set(echo = TRUE, warning = FALSE)

library(openxlsx)
library(scales)
library(readxl)
library(WriteXLS)
library(ggthemes)
library(RColorBrewer)
library(lubridate)
```

```
##
## Attaching package: 'lubridate'
```

```
## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

```r
library(caret)
```

```
## Loading required package: lattice
```

```
## Loading required package: ggplot2
```

```r
library(tidyverse)
```

```
## -- Attaching packages -------------------------------------------- tidyverse 1.3.0 --
```

```
## v tibble  3.0.3     v dplyr   1.0.2
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0
## v purrr   0.3.4
```

```
## -- Conflicts ----------------------------------------------- tidyverse_conflicts() --
## x lubridate::as.difftime() masks base::as.difftime()
## x readr::col_factor()      masks scales::col_factor()
## x lubridate::date()        masks base::date()
## x purrr::discard()         masks scales::discard()
## x dplyr::filter()          masks stats::filter()
## x lubridate::intersect()   masks base::intersect()
## x dplyr::lag()             masks stats::lag()
## x purrr::lift()            masks caret::lift()
## x lubridate::setdiff()     masks base::setdiff()
## x lubridate::union()       masks base::union()
```

```r
library(here)
```

```
## here() starts at /Users/msrodrigues/Dropbox/Coding/R/mestrado/Epi90/epi90
```

```
library(usethis)
library(googlesheets4)
library(DescTools)
```

```
##
## Attaching package: 'DescTools'

## The following objects are masked from 'package:caret':
##
##      MAE, RMSE
```

```
library(obAnalytics)
library(collapse)
```

```
## collapse 1.3.2, see ?`collapse-package` or ?`collapse-documentation`
## Note: stats::D -> D.expression, D.call

##
## Attaching package: 'collapse'

## The following object is masked from 'package:DescTools':
##
##      Recode

## The following object is masked from 'package:lubridate':
##
##      is.Date

## The following object is masked from 'package:stats':
##
##      D
```

```
library(tictoc)
library(ribge)

Sys.setenv(TZ="Brazil/East")
options(tz="Brazil/East")
Sys.getenv("TZ")
```

```
## [1] "Brazil/East"
```

```
options(scipen = 999999)
Sys.setlocale("LC_TIME", "pt_BR")
```

```
## [1] "pt_BR"
```

```
source("~/Dropbox/Coding/R/funs/msrfun.R")

# Carrega dados do SIM de 2018 salva em Binário no HD
sim_raw <- loadData("~/Dropbox/Coding/R/mestrado/Epi90/epi90/bin/sim.rds")
glimpse(sim_raw)
```

```
## Rows: 2,629,382
## Columns: 88
## $ CONTADOR   <fct> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 1...
## $ ORIGEM     <fct> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,...
## $ TIPOBITO   <fct> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,...
## $ DTOBITO    <fct> 05092017, 11022017, 11022017, 11022017, 11022017, 110220...
## $ HORAOBITO  <fct> 0700, 1330, 0500, 0830, 0320, 1335, 0855, 1115, 1731, 15...
```

```
## $ NATURAL    <fct> 812, 812, 812, 812, 812, 812, 812, 812, 812, 812, 812, 8...
## $ CODMUNNATU <fct> 120039, 120040, 120010, 120040, 120070, 120040, 120010, ...
## $ DTNASC     <fct> 03031997, 09022017, 13071933, 06022002, 02061966, 281219...
## $ IDADE      <fct> 420, 202, 483, 415, 450, 470, 461, 473, 204, 226, 485, 1...
## $ SEXO       <fct> 1, 1, 2, 1, 2, 2, 1, 2, 1, 2, 2, 2, 1, 1, 2, 2, 2, 1, 2,...
## $ RACACOR    <fct> 5, 4, 4, 4, 4, 4, 4, 1, 4, NA, 4, 4, 4, 2, 4, 5, 4, 4, 4...
## $ ESTCIV     <fct> NA, NA, 3, 1, 1, 1, 2, 1, NA, NA, 3, NA, NA, 1, 3, 3, 3,...
## $ ESC        <fct> 3, NA, 2, 9, 1, 4, 9, 2, NA, NA, 2, NA, NA, 4, 1, 1, 2, ...
## $ ESC2010    <fct> 2, NA, 1, 9, 0, 3, 9, 1, NA, NA, 1, NA, NA, 3, 0, 0, 1, ...
## $ SERIESCFAL <fct> 6, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N...
## $ OCUP       <fct> NA, NA, 999993, NA, 622020, NA, 622020, NA, NA, NA, 9999...
## $ CODMUNRES  <fct> 120039, 120040, 120040, 120040, 120070, 120040, 120040, ...
## $ LOCOCOR    <fct> 5, 1, 1, 3, 1, 1, 1, 1, 1, 1, 1, 3, 1, 1, 1, 1, 1, 1,...
## $ CODESTAB   <fct> NA, 2000733, 2001578, NA, 2001578, 2001578, 2001586, 200...
## $ ESTABDESCR <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ CODMUNOCOR <fct> 120039, 120040, 120040, 120040, 120040, 120040, 120040, ...
## $ IDADEMAE   <fct> NA, 36, NA, NA, NA, NA, NA, NA, 20, 24, NA, 35, 22, NA, ...
## $ ESCMAE     <fct> NA, 4, NA, NA, NA, NA, NA, NA, 4, 2, NA, 3, 4, NA, NA, N...
## $ ESCMAE2010 <fct> NA, 3, NA, NA, NA, NA, NA, NA, 3, 1, NA, 2, 3, NA, NA, N...
## $ SERIESCMAE <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ OCUPMAE    <fct> NA, 999992, NA, NA, NA, NA, NA, NA, 999991, 999992, NA, ...
## $ QTDFILVIVO <fct> NA, 02, NA, NA, NA, NA, NA, NA, 00, 03, NA, 03, NA, NA, ...
## $ QTDFILMORT <fct> NA, 00, NA, NA, NA, NA, NA, NA, 02, 00, NA, 00, NA, NA, ...
## $ GRAVIDEZ   <fct> NA, 1, NA, NA, NA, NA, NA, NA, 1, 1, NA, 1, 1, NA, NA, N...
## $ SEMAGESTAC <fct> NA, 24, NA, NA, NA, NA, NA, NA, 30, 40, NA, 26, 40, NA, ...
## $ GESTACAO   <fct> NA, 2, NA, NA, NA, NA, NA, NA, 3, 5, NA, 2, 5, NA, NA, N...
## $ PARTO      <fct> NA, 2, NA, NA, NA, NA, NA, NA, 2, 1, NA, 2, 1, NA, NA, N...
## $ OBITOPARTO <fct> NA, 3, NA, NA, NA, NA, NA, NA, 3, 3, NA, 3, 3, NA, NA, N...
## $ PESO       <fct> NA, 0660, NA, NA, NA, NA, NA, NA, 0856, 3360, NA, 1182, ...
## $ TPMORTEOCO <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ OBITOGRAV  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ OBITOPUERP <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ ASSISTMED  <fct> 2, NA, 1, NA, 1, NA, 1, 1, 1, 1, NA, 1, NA, NA, NA, NA, ...
## $ EXAME      <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ CIRURGIA   <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ NECROPSIA  <fct> 2, NA, 2, NA, 2, NA, NA, NA, 2, NA, NA, 2, NA, NA, NA, N...
## $ LINHAA     <fct> *R58X, *P285, *A419, *G80X, *I64X, *R688, *A419, *A419, ...
## $ LINHAB     <fct> NA, *P220, *N390, NA, *I10X, *J189, *J159, *J159, *E872,...
## $ LINHAC     <fct> NA, *P000, *I694, NA, *E119, *C920, *C349, *C73X, *P369,...
## $ LINHAD     <fct> NA, NA, NA, NA, *D649, NA, *J440, *I10X, NA, NA, *J440, ...
## $ LINHAII    <fct> NA, NA, NA, NA, NA, NA, *E149, NA, *P000, NA, *E149, *E8...
## $ CAUSABAS   <fct> R58, P000, I694, G809, E119, C920, C349, C73, P369, Q870...
## $ CB_PRE     <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ CRM        <fct> 1809, 660, 1566, 1890, 1566, 1397, 928, 928, 1217, 679, ...
## $ COMUNSVOIM <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, 120040, ...
## $ DTATESTADO <fct> 06092017, 11022017, 11022017, 11022017, 11022017, 110220...
## $ CIRCOBITO  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ ACIDTRAB   <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ FONTE      <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ NUMEROLOTE <fct> NA, 20170010, 20170005, 20170018, 20170005, 20170005, 20...
## $ TPPOS      <fct> NA, S, N, S, N, N, N, N, S, N, N, S, S, N, N, N, NA, NA,...
## $ DTINVESTIG <fct> NA, 23032017, NA, 01062017, NA, NA, NA, NA, 30032017, NA...
## $ CAUSABAS_O <fct> NA, P000, I694, R99, E119, C920, C349, C73, P000, Q870, ...
## $ DTCADASTRO <fct> 22022018, 14032017, 14032017, 14032017, 14032017, 140320...
```

```
## $ ATESTANTE  <fct> 5, 1, 5, 5, 5, 5, 2, 2, 2, 1, 5, 1, 3, 1, 5, 5, 1, 1, 5,...
## $ STCODIFICA <fct> N, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S,...
## $ CODIFICADO <fct> N, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S,...
## $ VERSAOSIST <fct> 2...0, 3.2.00, 3.2.00, 3.2.00, 3.2.00, 3.2.00, 3.2.00, 3...
## $ VERSAOSCB  <fct> NA, 3.2, 3.2, 3.2, 3.2, 3.2, 3.2, 3.2, 3.2, 3.2, 3.2, 3....
## $ FONTEINV   <fct> NA, 2, NA, 2, NA, NA, NA, NA, 2, NA, NA, 2, 8, NA, NA, N...
## $ DTRECEBIM  <fct> NA, 30062017, 16032017, 16102017, 16032017, 16032017, 16...
## $ ATESTADO   <fct> R58, P285/P220/P000, A419/N390/I694, G80, I64/I10/E119/D...
## $ DTRECORIGA <fct> 22022018, 16032017, 16032017, 16032017, 16032017, 160320...
## $ CAUSAMAT   <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ ESCMAEAGR1 <fct> NA, 12, NA, NA, NA, NA, NA, NA, 12, 10, NA, 11, 12, NA, ...
## $ ESCFALAGR1 <fct> 03, NA, 10, 09, 00, 12, 09, 10, NA, NA, 10, NA, NA, 12, ...
## $ STDOEPIDEM <fct> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,...
## $ STDONOVA   <fct> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,...
## $ DIFDATA    <fct> 170, 139, 033, 247, 033, 033, 034, 034, 140, 434, 035, 1...
## $ NUDIASOBCO <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ NUDIASOBIN <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ DTCADINV   <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ TPOBITOCOR <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ DTCONINV   <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ FONTES     <fct> NA, SXSSXX, NA, NA, NA, NA, NA, NA, SXSSXX, SXSSXX, NA, ...
## $ TPRESGINFO <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ TPNIVELINV <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ NUDIASINF  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ DTCADINF   <fct> NA, 30052017, NA, NA, NA, NA, NA, NA, 29052017, 12052017...
## $ MORTEPARTO <fct> NA, 3, NA, NA, NA, NA, NA, NA, 3, 3, NA, 3, 3, NA, NA, N...
## $ DTCONCASO  <fct> NA, 23032017, NA, NA, NA, NA, NA, NA, 30032017, 12052017...
## $ FONTESINF  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ ALTCAUSA   <fct> NA, 2, NA, NA, NA, NA, NA, NA, 2, 2, NA, 2, 1, NA, NA, N...
```

```r
# Ajusta variáveis necessárias para o exercício
sim <- sim_raw %>%
  mutate(
    estado_ocorrencia = str_extract(CODMUNOCOR,"^.."),
    CODMUNOCOR = as.integer(as.character(CODMUNOCOR)),
    SEXO = recode(SEXO,
                      "0" = "Ignorado",
                      "1" = "Masculino",
                      "2" = "Feminino"),
    CIRCOBITO = recode(CIRCOBITO,
                      "9" = "Ignorado",
                      "1" = "Acidente",
                      "2" = "Suicídio",
                      "3" = "Homicídio"),
    ano = as.numeric(str_sub(DTOBITO, 5, 8))

  )




# Carrega dados das cidades do IBGE
cidades <- ribge::populacao_municipios(ano = 2020)
```

```r
# Regiões do Brasil
norte <- c(11:17)
nordeste <- c(21:29)
sudeste <- c(31:33, 35)
sul <- c(41:43)
centro_oeste <- c(50:53)


# Inclui a informação da região a qual a cidade pertence na tabela cidades
cidades <-  cidades %>%
  mutate(
    regiao = case_when(
      codigo_uf %in% norte ~ "Norte",
      codigo_uf %in% nordeste ~ "Nordeste",
      codigo_uf %in% sudeste ~ "Sudeste",
      codigo_uf %in% sul ~ "Sul",
      codigo_uf %in% centro_oeste ~ "Centro Oeste",
      TRUE ~ "NULL")
  ) %>%
  group_by(regiao) %>%
  mutate(
    pop_regiao = sum(populacao)
  ) %>% ungroup() %>%
  group_by(uf) %>%
  mutate(
    pop_estado = sum(populacao)
  ) %>%
  ungroup()


# Seleciona variáveis e Join da tabela do SIM
# com a Informação da população da cidade de ocorrência
mort <- sim %>%
  select(estado_ocorrencia, CODMUNOCOR, SEXO, CIRCOBITO, ano) %>%
  left_join(cidades, by = c("CODMUNOCOR" = "cod_munic6") ) %>%
  ungroup()

glimpse(mort)
```

```
## Rows: 2,629,382
## Columns: 15
## $ estado_ocorrencia <chr> "12", "12", "12", "12", "12", "12", "12", "12", "...
## $ CODMUNOCOR        <int> 120039, 120040, 120040, 120040, 120040, 120040, 1...
## $ SEXO              <fct> Masculino, Masculino, Feminino, Masculino, Femini...
## $ CIRCOBITO         <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N...
## $ ano               <dbl> 2017, 2017, 2017, 2017, 2017, 2017, 2017, 2017, 2...
## $ uf                <chr> "AC", "AC", "AC", "AC", "AC", "AC", "AC", "AC", "...
## $ codigo_uf         <dbl> 12, 12, 12, 12, 12, 12, 12, 12, 12, 12, 12, 12, 1...
## $ codigo_munic      <chr> "00393", "00401", "00401", "00401", "00401", "004...
## $ nome_munic        <chr> "Porto Walter", "Rio Branco", "Rio Branco", "Rio ...
## $ populacao_str     <chr> "12241", "413418", "413418", "413418", "413418", ...
## $ populacao         <dbl> 12241, 413418, 413418, 413418, 413418, 413418, 41...
```

```
## $ cod_municipio    <chr> "1200393", "1200401", "1200401", "1200401", "1200...
## $ regiao           <chr> "Norte", "Norte", "Norte", "Norte", "Norte", "Nor...
## $ pop_regiao       <dbl> 18672591, 18672591, 18672591, 18672591, 18672591,...
## $ pop_estado       <dbl> 894470, 894470, 894470, 894470, 894470, 894470, 8...
```

Exercício 1

Os dados da quantidade de habitantes referem-se a estimativa da população feita pelo IBGE para o ano de 2020 e provem do banco de dados de população por cidade do IBGE. Os dados de mortalidade são provenientes do SIM.

Mortalidade por Suicídio por 100k por região do Brasil em 2018

a) Quais as taxas de suicídio por 100.00 habitantes no Brasil, no ano de 2018, segundo a região do país?

```
# Mortalidade por 100k por região do Brasil em 2018
mort %>%
  filter(!regiao == "NULL") %>%
  filter(ano == 2018) %>%
  filter(CIRCOBITO == "Suicídio") %>%
  group_by(regiao) %>%
  summarise(
    suicidios = n(),
    pop_regiao = unique(pop_regiao),
    mort_100k = suicidios / pop_regiao * 10^5
    )
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
## # A tibble: 5 x 4
##   regiao       suicidios pop_regiao mort_100k
##   <chr>            <int>      <dbl>     <dbl>
## 1 Centro Oeste      1169   16504303      7.08
## 2 Nordeste          2836   57374243      4.94
## 3 Norte              936   18672591      5.01
## 4 Sudeste           4498   89012240      5.05
## 5 Sul               2822   30192315      9.35
```

Mortalidade por suicídio por unidade da federação

b) Quais as taxas de suicídio por 100.00 habitantes no Brasil, no ano de 2017, segundo a unidade da federação?

```
# Mortalidade por suicídio por 100k por estado Brasil em 2018
mortalidade_por_estado <- mort %>%
  filter(!regiao == "NULL") %>%
  filter(ano == 2017) %>%
  filter(CIRCOBITO == "Suicídio") %>%
  group_by(uf) %>%
  summarise(
    suicidios = n(),
    pop_estado = unique(pop_estado),
    mort_100k = suicidios / pop_estado * 10^5
    )
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
mortalidade_por_estado
```

```
## # A tibble: 27 x 4
##    uf    suicidios pop_estado mort_100k
##    <chr>     <int>      <dbl>     <dbl>
##  1 AC           61     894470      6.82
##  2 AL           98    3351543      2.92
##  3 AM          205    4207714      4.87
##  4 AP           40     861773      4.64
##  5 BA          580   14930634      3.88
##  6 CE          633    9187103      6.89
##  7 DF          174    3055149      5.70
##  8 ES          205    4064052      5.04
##  9 GO          480    7113540      6.75
## 10 MA          277    7114598      3.89
## # ... with 17 more rows
```

Mortalidade por suicídio por 100k habitantes por sexo no Brasil em 2017

c) Quais as taxas de suicídio por 100.00 habitantes no Brasil, no ano de 2017, segundo o sexo?

Segundo IBGE, no censo de 2010 a proporção da população é de 51,03% de mulheres, dado utilizado para calcular a população de 2020, na qual os dados estão baseados. Não foi feita a estimativa da correção desta proporção.

```r
# Mortalidade por suicídio por 100k habitantes por sexo no Brasil em 2017


# População por sexo
populacao_total <- sum(cidades$populacao)
populacao_F <- populacao_total * 0.5103

mort %>%
  filter(!regiao == "NULL") %>%
  filter(ano == 2017) %>%
  filter(CIRCOBITO == "Suicídio") %>%
  filter(!SEXO == "Ignorado") %>%
  group_by(SEXO) %>%
  summarise(
    suicidios = n(),
    ) %>%

  mutate(
    pop = c(populacao_total - populacao_F, populacao_F),
    mort_100k = suicidios / pop * 10^5
  ) %>%
  ungroup()
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
## # A tibble: 2 x 4
##   SEXO      suicidios        pop mort_100k
##   <fct>         <int>      <dbl>     <dbl>
## 1 Masculino      9293 103696762.      8.96
## 2 Feminino       2522 108058930.      2.33
```