

# Exercício 03 - EPI 90

Márcio Rodrigues

10/5/2020

```
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
```

```
library(caret)
```

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

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

```
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()
```

```
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 os dados extraídos do TABNET
# Carrega dados do SIM de 2018 salvado em Binário no HD
# Alterar esta pasta para a pasta
```

```
if(!exists("ext_raw")) {
  ext_raw <- loadData("~/Dropbox/Coding/R/mestrado/Epi90/epi90/bin/ext.rds")
}
glimpse(ext_raw)
```

```
## Rows: 3,003,052
## Columns: 99
## $ contador <fct> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 1...
## $ TIPOBITO <fct> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,...
## $ DTOBITO <fct> 19011997, 24011997, 19011997, 19011997, 19011997, 190119...
## $ NATURAL <fct> 851, 811, 811, 811, 850, 811, 811, 811, 811, 811, 811, 8...
## $ DTNASC <fct> 05011969, 15041985, 11011948, 25081977, 16041977, 270419...
## $ IDADE <fct> 428, 411, 449, 419, 419, 425, 409, 402, 409, 413, 417, 4...
## $ SEXO <fct> 1, 1, 1, 2, 2, 2, 2, 1, 1, 2, 1, 1, 2, 1, 1, 1, 2, 1, 1,...
## $ RACACOR <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N...
## $ ESTCIV <fct> 2, 1, 2, 1, 1, 2, 1, 1, 1, 1, 2, 1, 1, 1, 2, 2, 9, 2,...
## $ ESC <fct> NA, NA, 9, NA, NA, 9, 9, 1, NA, 9, 8, NA, 9, 9, 1, NA, N...
## $ OCUP <fct> 98500, 00500, 62100, 00600, 00600, 00800, 00600, NA, 006...
## $ CODBAIRES <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ CODMUNRES <fct> 1100155, 1100155, 1100155, 1100155, 1100155, 1100155, 11...
## $ LOCOR <fct> 4, 4, 4, 4, 4, 4, 4, 4, 4, 5, 4, 4, 4, 3, 4, 4, 3, 4,...
## $ CODMUNOCOR <fct> 1100155, 1100155, 1100155, 1100155, 1100155, 1100155, 11...
## $ IDADEMAE <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ ESCMAE <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ OCUPMAE <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ QTDFILVIVO <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ QTDFILMORT <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ GRAVIDEZ <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ GESTACAO <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ PARTO <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ OBITOPARTO <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ PESO <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ OBITOGRV <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ OBITOUPERP <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ ASSISTMED <fct> 1, 2, 1, 2, 1, 1, 1, 1, 1, 1, 2, 2, 1, 2, 2, 9, 2, 9, 1,...
## $ EXAME <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, 2, NA, NA, NA, N...
## $ CIRURGIA <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, 2, NA, NA, NA, N...
## $ NECROPSIA <fct> 1, 1, NA, 1, 2, 2, 1, 1, 1, 1, 1, 2, 2, NA, NA, NA, 2, N...
## $ CAUSABAS <fct> V892, V892, V892, V892, V892, V892, V892, V892, V892, V8...
## $ LINHAA <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ LINHAB <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
```

```

## $ LINHAC      <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ LINHAD      <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ LINHAI      <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ CIRCOBITO   <fct> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, NA, 1, 1, 1, 1, 1, 3, NA, ...
## $ ACIDTRAB    <fct> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, NA, NA, 2, 1, NA, 1, NA, N...
## $ FONTE       <fct> 2, 1, 2, 2, 2, 2, 2, 2, 2, 2, 3, NA, 9, NA, NA, 4, NA, N...
## $ CODESTAB    <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ ATESTANTE   <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ UFINFORM     <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ CONTADOR     <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ HORAOBITO    <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ CODBAIOCOR  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ TPASSINA     <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ DTATESTADO  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ TPPOS       <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ DTINVESTIG  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ CAUSABAS_0  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ DTCADASTRO  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ FONTEINV     <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ DTRECEBIM   <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ CB_PRE      <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ MORTEPARTO  <fct> NA, NA, 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, NA, NA, ...
## $ ORIGEM      <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ DTCADINF    <fct> NA, NA, 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, NA, NA, ...
## $ COMUNSVOIM  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ DTRECORIG   <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ DTRECORIGA  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ CAUSAMAT    <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ ESC2010     <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ ESCMAE2010  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ STDOEPIDEM  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ STDONOVA    <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ CODMUNCART  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ CODCART     <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ NUMREGCART  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ DTREGCART   <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ SERIESCFAL  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ ESCMAEAGR1  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ ESCFALAGR1  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ SERIESCMAE  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ SEMAGESTAC  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ TPMORTEOCO  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ EXPDIFDATA  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ DIFDATA     <fct> NA, NA, 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, NA, NA, ...
## $ DTCONCASO   <fct> NA, NA, 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, NA, NA, ...
## $ CODMUNNATU  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ ESTABDESCR  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ CRM         <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ NUMEROLOTE  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ STCODIFICA  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...

```

```
## $ CODIFICADO <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ VERSAOSIST <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ VERSAOSCB <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ ATESTADO <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ NUDIASOBCO <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ FONTES <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ TPRESGINFO <fct> NA, 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, NA, ...
## $ NUDIASINF <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ FONTESINF <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ ALTCAUSA <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
```

```
# Ajusta variáveis necessárias para o exercício
```

```
ext <- ext_raw %>%
  mutate(
    estado_ocorrendia = 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)
```

```
# 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(
    regioao = 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 <- ext %>%
  select(estado_ocorrencia, CODMUNOCOR, SEXO, CIRCOBITO, ano) %>%
  left_join(cidades, by = c("CODMUNOCOR" = "cod_munic6")) %>%
  ungroup()

glimpse(mort)

```

```

## Rows: 3,003,052
## Columns: 15
## $ estado_ocorrencia <chr> "11", "11", "11", "11", "11", "11", "11", "11", "...
## $ CODMUNOCOR          <int> 1100155, 1100155, 1100155, 1100155, 1100155, 1100...
## $ SEXO                <fct> Masculino, Masculino, Masculino, Feminino, Femini...
## $ CIRCOBITO           <fct> Acidente, Acidente, Acidente, Acidente, Acidente,...
## $ ano                 <dbl> 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1...
## $ uf                  <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N...
## $ codigo_uf           <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N...
## $ codigo_munic        <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N...
## $ nome_munic          <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N...
## $ populacao_str       <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N...
## $ populacao           <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N...
## $ cod_municipio       <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N...
## $ regiao              <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N...
## $ pop_regiao          <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N...
## $ pop_estado          <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N...

```

```

mort %>%
  filter(CIRCOBITO == "Homicídio") %>%
  group_by(ano, )

```

```

## # A tibble: 1,091,809 x 15
## # Groups:   ano [26]
##   estado_ocorrenc~ CODMUNOCOR SEXO  CIRCOBITO   ano uf   codigo_uf
##   <chr>             <int> <fct> <fct>     <dbl> <chr>   <dbl>
## 1 11                1100254 Femi~ Homicídio 1997 <NA>    NA
## 2 11                1100205 Masc~ Homicídio 1997 <NA>    NA
## 3 11                1100205 Masc~ Homicídio 1997 <NA>    NA

```

```
## 4 11          1100205 Masc~ Homicídio 1997 <NA>      NA
## 5 11          1100205 Masc~ Homicídio 1997 <NA>      NA
## 6 11          1100122 Masc~ Homicídio 1997 <NA>      NA
## 7 11          1100122 Masc~ Homicídio 1997 <NA>      NA
## 8 11          1100122 Masc~ Homicídio 1997 <NA>      NA
## 9 11          1100122 Masc~ Homicídio 1997 <NA>      NA
## 10 11         1101500 Masc~ Homicídio 1997 <NA>      NA
## # ... with 1,091,799 more rows, and 8 more variables: codigo_munic <chr>,
## #   nome_munic <chr>, populacao_str <chr>, populacao <dbl>,
## #   cod_municipio <chr>, regioao <chr>, pop_regiao <dbl>, pop_estado <dbl>
```