

# projeto final-analise de sobrevivencia

joseferson da silva barreto

2022-07-08

## Objetivo

ajustar um modelo de sobrevivência capaz de explicar as relações entre o tempo até o aparecimento de cancer

## Metodologia

aqui vou colocar as informações sobre as metodologias utilizadas

## Análises

### Carregando o Banco de Dados

```
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.1 --

## v ggplot2 3.3.6      v purrr   0.3.4
## v tibble  3.1.7      v dplyr   1.0.9
## v tidyr   1.2.0      v stringr 1.4.0
## v readr   2.1.2      v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

library(plotly)

##
## Attaching package: 'plotly'

## The following object is masked from 'package:ggplot2':
##
##   last_plot

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

```
## The following object is masked from 'package:graphics':
##
## layout
```

```
library(readxl)

dados<-read.csv("base_nao_identificada_1832.csv",sep = ";",encoding = "latin1")

# dados$Data.do.Úbito-dados$Data.de.Diagnostico

library(lubridate)
```

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

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

```
# (day( as.Date(dados$Data.do.Úbito)) - day(as.Date( dados$Data.de.Diagnostico)))
#
#
# (( as.Date(dados$Data.do.Úbito)) - (as.Date( dados$Data.de.Diagnostico)))

# (month( as.Date(dados$Data.do.Úbito)) - month(as.Date( dados$Data.de.Diagnostico)))
#
# floor_date(as.Date(dados$Data.do.Úbito),"month")
#
# mdy(dados$Data.do.Úbito)
```

```
dados %>%
# dplyr::select(where(is.factor)) %>%
glimpse()
```

```
## Rows: 160,743
## Columns: 38
## $ Código.do.Paciente      <int> 475186, 474801, 474803, 474809, 47481~
## $ Nome.do.RCBP            <chr> "RCBP BELO HORIZONTE", "RCBP BELO HOR~
## $ Sexo                    <chr> "FEMININO", "MASCULINO", "MASCULINO",~
## $ Data.de.Nascimento       <chr> "28/12/1950", "10/02/1919", "14/11/19~
## $ Idade                   <int> 49, 81, 60, 79, 81, 78, 86, 44, 51, 7~
## $ Raca.Cor                 <chr> "SEM INFORMAÇÃO", "SEM INFORMAÇÃO", "~
## $ Nacionalidade           <chr> "Não informado", "Não informado", "Nã~
## $ Naturalidade.Estado     <chr> "Não informado", "Não informado", "Nã~
## $ Naturalidade            <chr> "", "", "", "", "", "Não Informado", ~
## $ Grau.de.Instrução        <chr> "SEM INFORMAÇÃO", "SEM INFORMAÇÃO", "~
## $ Estado.Civil            <chr> "SEM INFORMAÇÃO", "SEM INFORMAÇÃO", "~
## $ Código.Profissão        <int> 0, 723, 0, 0, 621, 0, 0, 0, 0, 10,~
## $ Nome.Profissão          <chr> "SEM INFORMAÇÃO\t", "TRABALHADORES DE~
```

```
## $ Estado.Endereço      <chr> "MINAS GERAIS", "MINAS GERAIS", "MINA~
## $ Cidade.Endereço      <chr> "BELO HORIZONTE", "BELO HORIZONTE", "~
## $ Descrição.da.Topografia <chr> "UTERO, SOE", "BEXIGA, SOE", "ESTOMAG~
## $ Código.da.Topografia  <chr> "C559", "C679", "C169", "C250", "C169~
## $ Descrição.da.Morfologia <chr> "NEOPLASMA MALIGNO", "NEOPLASMA MALIG~
## $ Código.da.Morfologia  <int> 80003, 80003, 80003, 80003, 80003, 80~
## $ Descrição.da.Doença   <chr> "UTERO, SOE", "BEXIGA, SOE", "ESTOMAG~
## $ Código.da.Doença      <chr> "C55", "C679", "C169", "C250", "C169~
## $ Descrição.da.Doença.Infantil <chr> "", "", "", "", "", "", "", "", "", ""~
## $ Código.da.Doença.Infantil <chr> "", "", "", "", "", "", "", "", "", ""~
## $ Descrição.da.Doença.Adulto.Jovem <chr> "", "", "", "", "", "", "", "", "", ""~
## $ Código.da.Doença.Adulto.Jovem <chr> "", "", "", "", "", "", "", "", "", ""~
## $ Indicador.de.Caso.Raro <chr> "false", "false", "false", "false", ""~
## $ Meio.de.Diagnostico   <chr> "SDO", "SDO", "SDO", "SDO", "SEM INFO~
## $ Extensão              <chr> "SEM INFORMAÇÃO", "SEM INFORMAÇÃO", ""~
## $ Lateralidade          <chr> "", "", "", "", "", "", "", "", "", ""~
## $ Estadiamento         <chr> "", "", "", "", "", "", "", "", "", ""~
## $ TNM                   <chr> "", "", "", "", "", "", "", "", "", ""~
## $ Status.Vital          <chr> "", "", "", "", "", "", "", "", "", ""~
## $ Tipo.do.Obito         <chr> "CÂNCER", "CÂNCER", "CÂNCER", "CÂNCER~
## $ Data.do.Óbito         <chr> "11/02/2000", "22/04/2000", "12/04/20~
## $ Data.de.Último.Contato <chr> "", "", "", "", "", "", "", "", "", ""~
## $ Data.de.Diagnostico   <chr> "11/02/2000", "22/04/2000", "12/04/20~
## $ Metástase.à.distância <chr> "", "", "", "", "", "", "", "", "", ""~
## $ X                     <lg1> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
```

```
pulmao<-dados %>%
  dplyr::filter(Descrição.da.Topografia == "PULMAO, SOE")

pulmao$falhas<-pulmao$Tipo.do.Obito
pulmao$falhas[pulmao$falhas == "CÂNCER"] <- "1"
pulmao$falhas[pulmao$falhas == "NÃO CÂNCER"] <- "0"
pulmao$falhas[pulmao$falhas == "SEM INFORMAÇÃO"] <- "0"
pulmao$falhas[pulmao$falhas == ""] <- "0"

table(pulmao$falhas)
```

```
##
##      0      1
## 916 3511
```

## fazendo a análise descritiva

```
pulmao %>%
  dplyr::select(where(is.character)) %>%
  glimpse()
```

```
## Rows: 4,427
## Columns: 34
## $ Nome.do.RCBP      <chr> "RCBP BELO HORIZONTE", "RCBP BELO HOR~
```

```

## $ Sexo <chr> "MASCULINO", "MASCULINO", "MASCULINO"~
## $ Data.de.Nascimento <chr> "11/05/1933", "01/12/1950", "16/05/19~
## $ Raca.Cor <chr> "SEM INFORMAÇÃO", "SEM INFORMAÇÃO", "~
## $ Nacionalidade <chr> "Não informado", "Não informado", "Nã~
## $ Naturalidade.Estado <chr> "Não informado", "Não informado", "Nã~
## $ Naturalidade <chr> "", "", "", "", "", "", "", "", "", "", "~
## $ Grau.de.Instrução <chr> "SEM INFORMAÇÃO", "SEM INFORMAÇÃO", "~
## $ Estado.Civil <chr> "SEM INFORMAÇÃO", "SEM INFORMAÇÃO", "~
## $ Nome.Profissão <chr> "MEMBROS DAS FORÇAS ARMADAS\t", "SEM ~
## $ Estado.Endereço <chr> "MINAS GERAIS", "MINAS GERAIS", "MINA~
## $ Cidade.Endereço <chr> "BELO HORIZONTE", "BELO HORIZONTE", "~
## $ Descrição.da.Topografia <chr> "PULMAO, SOE", "PULMAO, SOE", "PULMAO~
## $ Código.da.Topografia <chr> "C349", "C349", "C349", "C349", "C349~
## $ Descrição.da.Morfologia <chr> "NEOPLASMA MALIGNO", "CARCINOMA SOLID~
## $ Descrição.da.Doença <chr> "BRONQUIOS OU PULMOES, SOE", "BRONQUI~
## $ Código.da.Doença <chr> "C349", "C349", "C349", "C349", "C349~
## $ Descrição.da.Doença.Infantil <chr> "", "", "", "", "", "", "", "", "", "", "~
## $ Código.da.Doença.Infantil <chr> "", "", "", "", "", "", "", "", "", "", "~
## $ Descrição.da.Doença.Adulto.Jovem <chr> "", "", "", "", "", "", "", "", "", "", "~
## $ Código.da.Doença.Adulto.Jovem <chr> "", "", "", "", "", "", "", "", "", "", "~
## $ Indicador.de.Caso.Raro <chr> "false", "false", "false", "false", "~
## $ Meio.de.Diagnostico <chr> "SDO", "HISTOLOGIA DO TUMOR PRIMÁRIO"~
## $ Extensão <chr> "SEM INFORMAÇÃO", "SEM INFORMAÇÃO", "~
## $ Lateralidade <chr> "", "", "", "", "", "", "", "", "", "", "~
## $ Estadiamento <chr> "", "", "", "", "", "", "", "", "", "", "~
## $ TNM <chr> "", "", "", "", "", "", "", "", "", "", "~
## $ Status.Vital <chr> "", "", "", "", "", "", "", "", "", "", "~
## $ Tipo.do.Óbito <chr> "CÂNCER", "", "CÂNCER", "", "CÂNCER",~
## $ Data.do.Óbito <chr> "14/07/2000", "", "24/01/2001", "", "~
## $ Data.de.Último.Contato <chr> "", "", "", "", "", "", "", "", "", "", "~
## $ Data.de.Diagnostico <chr> "14/07/2000", "01/01/2000", "14/11/20~
## $ Metástase.à.distância <chr> "", "", "", "", "", "", "", "", "", "", "~
## $ falhas <chr> "1", "0", "1", "0", "1", "1", "1", "1~

```

```

# =====
#   posiveis fatores
# =====
#
# # fatores
#
#   sexo
#   grau de instrução
#   Estado civil
#   nome.profissão
#   raca.cor
#   nacionalidade
#   cidade.endereço
#   Descrição.da.Doença.Infantil
#   Código.da.Doença.Infantil
#   Indicador.de.Caso.Raro
#   Meio.de.Diagnostico
#   Status.Vital
#   Metástase.à.distância

```

```
pulmao %>%
  dplyr::select(where(is.numeric)) %>%
  glimpse()
```

```
## Rows: 4,427
## Columns: 4
## $ Código.do.Paciente <int> 474827, 474595, 475047, 475238, 475304, 475344, 4~
## $ Idade <int> 67, 49, 69, 65, 47, 72, 74, 63, 59, 49, 69, 77, 5~
## $ Código.Profissão <int> 10, 0, 142, 0, 0, 0, 0, 0, 0, 630, 0, 0, 0, ~
## $ Código.da.Morfologia <int> 80003, 82303, 81403, 80103, 80103, 80703, 80703, ~
```

```
# obs
# =====
# Código.do.Paciente que está sendo identificada como uma variável número na verdade
# é o do tipo texto ,ou seja não é um valor quantitativo
# vamos converter essa variável
pulmao$Código.do.Paciente<-as.character(pulmao$Código.do.Paciente)
```

```
# código da profissão tbm é apenas um código
```

```
pulmao$Código.Profissão<-as.character(pulmao$Código.Profissão)
```

```
pulmao %>%
  dplyr::select(where(is.numeric)) %>%
  glimpse()
```

```
## Rows: 4,427
## Columns: 2
## $ Idade <int> 67, 49, 69, 65, 47, 72, 74, 63, 59, 49, 69, 77, 5~
## $ Código.da.Morfologia <int> 80003, 82303, 81403, 80103, 80103, 80703, 80703, ~
```

```
# fazendo o mesmo para código da morfologia
```

```
pulmao$Código.da.Morfologia<-as.character(pulmao$Código.da.Morfologia)
```

```
pulmao %>%
  dplyr::select(where(is.numeric)) %>%
  glimpse()
```

```
## Rows: 4,427
## Columns: 1
## $ Idade <int> 67, 49, 69, 65, 47, 72, 74, 63, 59, 49, 69, 77, 57, 75, 81, 72, ~
```

```
#pulmao %>% select(Idade) %>% table()
```

## Convertendo os possíveis fatores

```

# sexo
# grau de instrução
# Estado civil
# nome.profissão
# raca.cor
# nacionalidade
# cidade.endereço
# Descrição.da.Doenca.Infantil
# Código.da.Doenca.Infantil
# Indicador.de.Caso.Raro
# Meio.de.Diagnostico
# Status.Vital
# Metástase.à.distância

pulmao$Sexo<- as.factor(pulmao$Sexo)

glimpse(pulmao)

```

```

## Rows: 4,427
## Columns: 39
## $ Código.do.Paciente      <chr> "474827", "474595", "475047", "475238~
## $ Nome.do.RCBP            <chr> "RCBP BELO HORIZONTE", "RCBP BELO HOR~
## $ Sexo                    <fct> MASCULINO, MASCULINO, MASCULINO, MASC~
## $ Data.de.Nascimento      <chr> "11/05/1933", "01/12/1950", "16/05/19~
## $ Idade                   <int> 67, 49, 69, 65, 47, 72, 74, 63, 59, 4~
## $ Raca.Cor                 <chr> "SEM INFORMAÇÃO", "SEM INFORMAÇÃO", "~
## $ Nacionalidade           <chr> "Não informado", "Não informado", "Nã~
## $ Naturalidade.Estado     <chr> "Não informado", "Não informado", "Nã~
## $ Naturalidade            <chr> "", "", "", "", "", "", "", "", "~
## $ Grau.de.Instrução       <chr> "SEM INFORMAÇÃO", "SEM INFORMAÇÃO", "~
## $ Estado.Civil            <chr> "SEM INFORMAÇÃO", "SEM INFORMAÇÃO", "~
## $ Código.Profissão        <chr> "10", "0", "142", "0", "0", "0", "0",~
## $ Nome.Profissão          <chr> "MEMBROS DAS FORÇAS ARMADAS\t", "SEM ~
## $ Estado.Endereço         <chr> "MINAS GERAIS", "MINAS GERAIS", "MINA~
## $ Cidade.Endereço        <chr> "BELO HORIZONTE", "BELO HORIZONTE", "~
## $ Descrição.da.Topografia <chr> "PULMAO, SOE", "PULMAO, SOE", "PULMAO~
## $ Código.da.Topografia    <chr> "C349", "C349", "C349", "C349", "C349~
## $ Descrição.da.Morfologia <chr> "NEOPLASMA MALIGNO", "CARCINOMA SOLID~
## $ Código.da.Morfologia    <chr> "80003", "82303", "81403", "80103", "~
## $ Descrição.da.Doenca    <chr> "BRONQUIOS OU PULMOES, SOE", "BRONQUI~
## $ Código.da.Doenca        <chr> "C349", "C349", "C349", "C349", "C349~
## $ Descrição.da.Doenca.Infantil <chr> "", "", "", "", "", "", "", "", "~
## $ Código.da.Doenca.Infantil <chr> "", "", "", "", "", "", "", "", "~
## $ Descrição.da.Doenca.Adulto.Jovem <chr> "", "", "", "", "", "", "", "", "~
## $ Código.da.Doenca.Adulto.Jovem <chr> "", "", "", "", "", "", "", "", "~
## $ Indicador.de.Caso.Raro  <chr> "false", "false", "false", "false", "~
## $ Meio.de.Diagnostico     <chr> "SD0", "HISTOLOGIA DO TUMOR PRIMÁRIO~
## $ Extensão                <chr> "SEM INFORMAÇÃO", "SEM INFORMAÇÃO", "~
## $ Lateralidade            <chr> "", "", "", "", "", "", "", "", "~
## $ Estadiamento           <chr> "", "", "", "", "", "", "", "", "~
## $ TNM                     <chr> "", "", "", "", "", "", "", "", "~
## $ Status.Vital            <chr> "", "", "", "", "", "", "", "", "~

```

```
## $ Tipo.do.Obito      <chr> "CÂNCER", "", "CÂNCER", "", "CÂNCER", ~
## $ Data.do.Óbito     <chr> "14/07/2000", "", "24/01/2001", "", "~
## $ Data.de.Último.Contato <chr> "", "", "", "", "", "", "", "", "", "", "~
## $ Data.de.Diagnostico <chr> "14/07/2000", "01/01/2000", "14/11/20~
## $ Metástase.à.distância <chr> "", "", "", "", "", "", "", "", "", "", "~
## $ X                 <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ falhas            <chr> "1", "0", "1", "0", "1", "1", "1", "1", "1~
```

```
#c(pulmao$Sexo,pulmao$Nacionalidade)<-lapply(c(pulmao$Sexo,pulmao$Nacionalidade), as.factor)
```

```
pulmao$Nacionalidade<- as.factor(pulmao$Nacionalidade)
pulmao$Raca.Cor<- as.factor(pulmao$Raca.Cor)
pulmao$Naturalidade<- as.factor(pulmao$Naturalidade)

pulmao$Grau.de.Instrução<- as.factor(pulmao$Grau.de.Instrução)

pulmao$Estado.Civil<- as.factor(pulmao$Estado.Civil)

pulmao %>% select(
  Descrição.da.Morfologia
) %>% table()
```

```
## Descrição.da.Morfologia
##          ADENOCARCINOMA BRONQUIOLO-ALVEOLAR
##                                     30
##          ADENOCARCINOMA BRONQUIOLO-ALVEOLAR, SOE
##                                     79
##          ADENOCARCINOMA DE CELULAS CLARAS, SOE
##                                     11
##          ADENOCARCINOMA MUCINOSO
##                                     16
##          ADENOCARCINOMA PAPILAR, SOE
##                                     10
##          ADENOCARCINOMA PRODUTOR DE MUCINA
##                                     17
##          ADENOCARCINOMA TRABECULAR
##                                     1
##          ADENOCARCINOMA TUBULAR
##                                     1
##          ADENOCARCINOMA, SOE
##                                     837
##          CARCINOMA ADENOESCAMOSO
##                                     3
##          CARCINOMA ADENOIDE CISTICO
##                                     4
##          CARCINOMA ANAPLASICO, SOE
##                                     3
##          CARCINOMA BRONQUIOLO ALVEOLAR MUCINOSO
##                                     15
##          CARCINOMA BRONQUIOLO ALVEOLAR NAO MUCINOSO
##                                     10
##          CARCINOMA DE CEL. ESC. GRANDES, NAO CERATINIZADO
##                                     3
##          CARCINOMA DE CEL. ESC. PEQUENAS, NAO CERATINIZADO
```

##		1
##	CARCINOMA DE CEL. ESCAMOSAS CERATINIZADO, SOE	
##		3
##	CARCINOMA DE CELULAS ACINOSAS	
##		2
##	CARCINOMA DE CELULAS EM ANEL DE SINETE	
##		4
##	CARCINOMA DE CELULAS ESCAMOSAS, SOE	
##		154
##	CARCINOMA DE CELULAS GRANDES, SOE	
##		38
##	CARCINOMA DE CELULAS NAO PEQUENAS	
##		103
##	CARCINOMA DE CELULAS PEQUENAS E DE CELULAS GRANDES	
##		1
##	CARCINOMA DE CELULAS PEQUENAS, SOE	
##		28
##	CARCINOMA EPITELIAL-MIOEPITELIAL	
##		1
##	CARCINOMA ESC. DE CEL. GDES., NAO CERAT., SOE	
##		2
##	CARCINOMA ESC. DE CEL. PEQUENAS, NAO CERATINIZADO	
##		2
##	CARCINOMA ESCAMOCELULAR CERATINIZADO, SOE	
##		2
##	CARCINOMA ESCAMOCELULAR IN SITU, SOE	
##		3
##	CARCINOMA ESCAMOCELULAR, SOE	
##		327
##	CARCINOMA IN SITU DE CELULAS ESCAMOSAS, SOE	
##		2
##	CARCINOMA IN SITU, SOE	
##		1
##	CARCINOMA INDIFERENCIADO, SOE	
##		18
##	CARCINOMA MUCOEPIDERMIDE	
##		4
##	CARCINOMA NEUROENDOCRINO	
##		15
##	CARCINOMA NEUROENDOCRINO DE GRANDES CELULAS	
##		2
##	CARCINOMA NEUROENDOCRINO, SOE	
##		35
##	CARCINOMA OAT CELL	
##		227
##	CARCINOMA PAPILAR, SOE	
##		3
##	CARCINOMA PSEUDOSSARCOMATOSO	
##		4
##	CARCINOMA SOLIDO, SOE	
##		4
##	CARCINOMA, SOE	
##		379
##	HISTIOCITOSE DE CEL. DE LANGERHANS, DISSEMINADA	



```

##                                     1
##                               HISTIOCITOSE MALIGNA
##                                     1
##       LINFOMA DE HODGKIN, CELULARIDADE MISTA, SOE
##                                     1
##       LINFOMA DE ZONA MARGINAL TIPO CELULAS B, SOE
##                                     2
##       LINFOMA MALIGNO DE LINFOCITOS PEQUENOS, SOE
##                                     1
##       LINFOMA MALIGNO, CELS. GRANDES B, DIFUSO, SOE
##                                     2
##               LINFOMA MALIGNO, DIFUSO, SOE
##                                     1
##               LINFOMA MALIGNO, SOE
##                                     3
##       LINFOMA MALIGNO, TIPO NAO HODGKIN, SOE
##                                     2
##               MESOTELIOMA MALIGNO
##                                     1
##               NEOPLASIA MALIGNA
##                               1461
##               NEOPLASMA MALIGNO
##                               493
##               PLASMOCITOMA, SOE
##                                     1
##               RABDOMIOSSARCOMA, SOE
##                                     1
##               SARCOMA DE KAPOSI
##                                     1
##               SARCOMA FUSOCELULAR
##                                     2
##               SARCOMA, SOE
##                                     2
## TU MALIGNO DE CEL. SEMELHANTES A ENTEROCROMAFINICA
##                                     1
##               TUMOR CARCINOIDE ATIPICO
##                                     1
##               TUMOR CARCINOIDE, SOE
##                               39
##               TUMOR MALIGNO DE CELULAS PEQUENAS
##                                     5

```

```

pulmao$Descrição.da.Morfologia<- as.factor(pulmao$Descrição.da.Morfologia)
pulmao %>% select(
  Descrição.da.Doenca
) %>% table()

```

```

## Descrição.da.Doenca
##       BRONQUIOS E PULMOES
##                               6
##       BRONQUIOS OU PULMOES, SOE
##                               4404
##       CELULARIDADE MISTA
##                               1

```

```
## DOENCA DE LETTERER-SIWE
## 1
## GRANDES CELULAS (DIFUSO)
## 4
## HISTIOCILOSE MALIGNA
## 1
## LINFOMA NAO-HODGKIN DIFUSO, SOE
## 1
## LINFOMA NAO-HODGKIN, SOE
## 5
## MESOTELIOMA DE OUTRAS LOCALIZACOES
## 1
## PEQUENAS CELULAS (DIFUSO)
## 1
## PLASMOCITOMA EXTRAMEDULAR
## 1
## SARCOMA DE KAPOS DE OUTRAS LOCALIZACOES
## 1
```

```
pulmao$Descrição.da.Doenca<- as.factor(pulmao$Descrição.da.Doenca)
pulmao$Descrição.da.Doenca.Adulto.Jovem<- as.factor(pulmao$Descrição.da.Doenca.Adulto.Jovem)
pulmao$Descrição.da.Doenca.Infantil<- as.factor(pulmao$Descrição.da.Doenca.Infantil)
pulmao$Indicador.de.Caso.Raro<- as.factor(pulmao$Indicador.de.Caso.Raro)
pulmao$Meio.de.Diagnostico<- as.factor(pulmao$Meio.de.Diagnostico)
pulmao$Extensão<- as.factor(pulmao$Extensão)
pulmao$Tipo.do.Obito<- as.factor(pulmao$Tipo.do.Obito)
pulmao$Nome.Profissão<- as.factor(pulmao$Nome.Profissão)

glimpse(pulmao)
```

```
## Rows: 4,427
## Columns: 39
## $ Código.do.Paciente      <chr> "474827", "474595", "475047", "475238~
## $ Nome.do.RCBP            <chr> "RCBP BELO HORIZONTE", "RCBP BELO HOR~
## $ Sexo                     <fct> MASCULINO, MASCULINO, MASCULINO, MASC~
## $ Data.de.Nascimento       <chr> "11/05/1933", "01/12/1950", "16/05/19~
## $ Idade                    <int> 67, 49, 69, 65, 47, 72, 74, 63, 59, 4~
## $ Raca.Cor                 <fct> SEM INFORMAÇÃO, SEM INFORMAÇÃO, BRANC~
## $ Nacionalidade            <fct> Não informado, Não informado, Não inf~
## $ Naturalidade.Estado      <chr> "Não informado", "Não informado", "Nã~
## $ Naturalidade             <fct> , , , , , , , , , , , , , , , ~
## $ Grau.de.Instrução        <fct> SEM INFORMAÇÃO, SEM INFORMAÇÃO, SEM I~
## $ Estado.Civil             <fct> SEM INFORMAÇÃO, SEM INFORMAÇÃO, CASAD~
## $ Código.Profissão         <chr> "10", "0", "142", "0", "0", "0", "0", ~
## $ Nome.Profissão           <fct> "MEMBROS DAS FORÇAS ARMADAS\t", "SEM ~
## $ Estado.Endereço          <chr> "MINAS GERAIS", "MINAS GERAIS", "MINA~
## $ Cidade.Endereço          <chr> "BELO HORIZONTE", "BELO HORIZONTE", "~
## $ Descrição.da.Topografia  <chr> "PULMAO, SOE", "PULMAO, SOE", "PULMAO~
## $ Código.da.Topografia     <chr> "C349", "C349", "C349", "C349", "C349~
## $ Descrição.da.Morfologia  <fct> "NEOPLASMA MALIGNO", "CARCINOMA SOLID~
## $ Código.da.Morfologia     <chr> "80003", "82303", "81403", "80103", "~
## $ Descrição.da.Doenca      <fct> "BRONQUIOS OU PULMOES, SOE", "BRONQUI~
## $ Código.da.Doenca         <chr> "C349", "C349", "C349", "C349", "C349~
## $ Descrição.da.Doenca.Infantil <fct> , , , , , , , , , , , , , , , ~
```

```
## $ Código.da.Doenca.Infantil      <chr> "", "", "", "", "", "", "", "", "", ""
## $ Descrição.da.Doenca.Adulto.Jovem <fct> "", "", "", "", "", "", "", "", "", ""
## $ Código.da.Doenca.Adulto.Jovem   <chr> "", "", "", "", "", "", "", "", "", ""
## $ Indicador.de.Caso.Raro           <fct> false, false, false, false, false, fa~
## $ Meio.de.Diagnostico              <fct> SDO, HISTOLOGIA DO TUMOR PRIMÁRIO, HI~
## $ Extensão                         <fct> SEM INFORMAÇÃO, SEM INFORMAÇÃO, SEM I~
## $ Lateralidade                     <chr> "", "", "", "", "", "", "", "", "", ""
## $ Estadiamento                     <chr> "", "", "", "", "", "", "", "", "", ""
## $ TNM                              <chr> "", "", "", "", "", "", "", "", "", ""
## $ Status.Vital                     <chr> "", "", "", "", "", "", "", "", "", ""
## $ Tipo.do.Obito                    <fct> CÂNCER, , CÂNCER, , CÂNCER, CÂNCER, C~
## $ Data.do.Úbito                    <chr> "14/07/2000", "", "24/01/2001", "", ""
## $ Data.de.Último.Contato            <chr> "", "", "", "", "", "", "", "", "", ""
## $ Data.de.Diagnostico               <chr> "14/07/2000", "01/01/2000", "14/11/20~
## $ Metástase.à.distância             <chr> "", "", "", "", "", "", "", "", "", ""
## $ X                                 <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ falhas                            <chr> "1", "0", "1", "0", "1", "1", "1", "1"
```

## Convertendo as variáveis de datas

```
pulmao$Data.de.Nascimento<- as.Date(pulmao$Data.de.Nascimento)

pulmao$Data.de.Último.Contato<- as.Date(pulmao$Data.de.Último.Contato)

pulmao$Data.de.Diagnostico<- as.Date(pulmao$Data.de.Diagnostico)

pulmao$Data.do.Úbito<- as.Date(pulmao$Data.do.Úbito)

table(pulmao$Data.de.Último.Contato)
```

```
##
## 0001-01-20 0001-02-20 0001-03-20 0001-04-20 0001-05-20 0001-06-20 0001-07-20
##          7          4          5          9          5          5          3
## 0001-08-20 0001-09-20 0001-10-20 0001-11-20 0001-12-20 0002-01-20 0002-02-20
##          7          7          4          8          2          5          4
## 0002-03-20 0002-04-20 0002-05-20 0002-06-20 0002-07-20 0002-08-20 0002-09-20
##          7          3          4          5          7          6          4
## 0002-10-20 0002-11-20 0002-12-20 0003-01-20 0003-02-20 0003-03-20 0003-04-20
##          9          2          5          3          10         4          10
## 0003-05-20 0003-06-20 0003-07-20 0003-08-20 0003-09-20 0003-10-20 0003-11-20
##          5          10         4          3          8          7          4
## 0003-12-20 0004-01-20 0004-02-20 0004-03-20 0004-04-20 0004-05-20 0004-06-20
##          5          5          10         4          3          8          4
## 0004-07-20 0004-08-20 0004-09-20 0004-10-20 0004-11-20 0004-12-20 0005-01-20
##          6          6          6          6          4          6          3
## 0005-02-20 0005-03-20 0005-04-20 0005-05-20 0005-06-20 0005-07-20 0005-08-20
##          4          4          5          8          4          7          2
## 0005-09-20 0005-10-20 0005-11-20 0005-12-20 0006-01-20 0006-02-20 0006-03-20
##          4          4          2          7          6          6          10
## 0006-04-20 0006-05-20 0006-06-20 0006-07-20 0006-08-20 0006-09-20 0006-10-20
##          9          3          5          4          3          4          4
```

##	0006-11-20	0006-12-20	0007-01-20	0007-02-20	0007-03-20	0007-04-20	0007-05-20
##	4	7	7	6	5	4	6
##	0007-06-20	0007-07-20	0007-08-20	0007-09-20	0007-10-20	0007-11-20	0007-12-20
##	3	3	9	3	11	3	7
##	0008-01-20	0008-02-20	0008-03-20	0008-04-20	0008-05-20	0008-06-20	0008-07-20
##	6	6	3	8	8	8	6
##	0008-08-20	0008-09-20	0008-10-20	0008-11-20	0008-12-20	0009-01-20	0009-02-20
##	3	3	1	6	5	4	8
##	0009-03-20	0009-04-20	0009-05-20	0009-06-20	0009-07-20	0009-08-20	0009-09-20
##	5	6	8	4	7	7	5
##	0009-10-20	0009-11-20	0009-12-20	0010-01-20	0010-02-20	0010-03-20	0010-04-20
##	11	4	6	8	6	2	5
##	0010-05-20	0010-06-20	0010-07-20	0010-08-20	0010-09-20	0010-10-20	0010-11-20
##	8	8	9	8	2	4	7
##	0010-12-20	0011-01-20	0011-02-20	0011-03-20	0011-04-20	0011-05-20	0011-06-20
##	3	6	3	6	8	5	4
##	0011-07-20	0011-08-20	0011-09-20	0011-10-20	0011-11-20	0011-12-20	0012-01-20
##	6	6	8	6	2	1	3
##	0012-02-20	0012-03-20	0012-04-20	0012-05-20	0012-06-20	0012-07-20	0012-08-20
##	13	5	8	10	2	9	2
##	0012-09-20	0012-10-20	0012-11-20	0012-12-20	0013-01-20	0013-02-20	0013-03-20
##	8	7	1	6	7	4	4
##	0013-04-20	0013-05-20	0013-06-20	0013-07-20	0013-08-20	0013-09-20	0013-10-20
##	2	9	2	6	5	7	9
##	0013-11-20	0013-12-20	0014-01-20	0014-02-20	0014-03-20	0014-04-20	0014-05-20
##	6	6	2	5	5	5	4
##	0014-06-20	0014-07-20	0014-08-20	0014-09-20	0014-10-20	0014-11-20	0014-12-20
##	7	7	6	8	2	6	3
##	0015-01-20	0015-02-20	0015-03-20	0015-04-20	0015-05-20	0015-06-20	0015-07-20
##	6	1	3	8	7	4	7
##	0015-08-20	0015-09-20	0015-10-20	0015-11-20	0015-12-20	0016-01-20	0016-02-20
##	3	7	6	1	10	6	5
##	0016-03-20	0016-04-20	0016-05-20	0016-06-20	0016-07-20	0016-08-20	0016-09-20
##	5	5	8	9	5	7	7
##	0016-10-20	0016-11-20	0016-12-20	0017-01-20	0017-02-20	0017-03-20	0017-04-20
##	2	6	6	6	5	6	7
##	0017-05-20	0017-06-20	0017-07-20	0017-08-20	0017-09-20	0017-10-20	0017-11-20
##	9	7	7	6	5	5	8
##	0017-12-20	0018-01-20	0018-02-20	0018-03-20	0018-04-20	0018-05-20	0018-06-20
##	2	4	6	3	6	6	2
##	0018-07-20	0018-08-20	0018-09-20	0018-10-20	0018-11-20	0018-12-20	0019-01-20
##	5	9	4	10	7	10	10
##	0019-02-20	0019-03-20	0019-04-20	0019-05-20	0019-06-20	0019-07-20	0019-08-20
##	3	4	6	5	3	3	3
##	0019-09-20	0019-10-20	0019-11-20	0019-12-20	0020-01-20	0020-02-20	0020-03-20
##	2	3	6	4	4	5	11
##	0020-05-20	0020-06-20	0020-07-20	0020-08-20	0020-09-20	0020-10-20	0020-11-20
##	6	5	9	5	2	2	5
##	0020-12-20	0021-01-20	0021-02-20	0021-03-20	0021-04-20	0021-05-20	0021-06-20
##	4	6	2	5	4	4	8
##	0021-07-20	0021-08-20	0021-09-20	0021-10-20	0021-11-20	0021-12-20	0022-01-20
##	4	4	7	6	3	5	5
##	0022-02-20	0022-03-20	0022-04-20	0022-05-20	0022-06-20	0022-07-20	0022-08-20
##	5	3	3	3	8	8	4

```

## 0022-09-20 0022-10-20 0022-11-20 0022-12-20 0023-01-20 0023-02-20 0023-03-20
##          9          3          5          8          3          11          5
## 0023-04-20 0023-05-20 0023-06-20 0023-07-20 0023-08-20 0023-09-20 0023-10-20
##          1          6          1          1          5          8          8
## 0023-11-20 0023-12-20 0024-01-20 0024-02-20 0024-03-20 0024-04-20 0024-05-20
##          2          6          6          5          5          2          4
## 0024-06-20 0024-07-20 0024-08-20 0024-09-20 0024-10-20 0024-11-20 0024-12-20
##          5          8          6          3          6          5          5
## 0025-01-20 0025-02-20 0025-03-20 0025-04-20 0025-05-20 0025-06-20 0025-07-20
##          9          4          4          7          5          5          8
## 0025-08-20 0025-09-20 0025-10-20 0025-11-20 0025-12-20 0026-01-20 0026-02-20
##          7          5          7          4          2          9          3
## 0026-03-20 0026-04-20 0026-05-20 0026-06-20 0026-07-20 0026-08-20 0026-09-20
##          9          10         1          4          9          8          3
## 0026-10-20 0026-11-20 0026-12-20 0027-01-20 0027-02-20 0027-03-20 0027-04-20
##          3          1          7          5          9          4          3
## 0027-05-20 0027-06-20 0027-07-20 0027-08-20 0027-09-20 0027-10-20 0027-11-20
##          1          9          5          3          2          2          9
## 0027-12-20 0028-01-20 0028-02-20 0028-03-20 0028-04-20 0028-05-20 0028-06-20
##          1          6          4          5          1          4          8
## 0028-07-20 0028-08-20 0028-09-20 0028-10-20 0028-11-20 0028-12-20 0029-01-20
##          4          2          8          4          5          6          5
## 0029-02-20 0029-03-20 0029-04-20 0029-05-20 0029-06-20 0029-07-20 0029-08-20
##          1          8          3          7          9          5          4
## 0029-09-20 0029-10-20 0029-11-20 0029-12-20 0030-01-20 0030-03-20 0030-04-20
##          5          7          6          4          9          8          7
## 0030-05-20 0030-06-20 0030-07-20 0030-08-20 0030-09-20 0030-10-20 0030-11-20
##          4          8          4          7          5          5          6
## 0030-12-20 0031-01-20 0031-03-20 0031-05-20 0031-07-20 0031-08-20 0031-10-20
##          7          6          9          5          4          3          7
## 0031-12-20
##          7

```

```
glimpse(dados)
```

```

## Rows: 160,743
## Columns: 38
## $ Código.do.Paciente      <int> 475186, 474801, 474803, 474809, 47481~
## $ Nome.do.RCBP            <chr> "RCBP BELO HORIZONTE", "RCBP BELO HOR~
## $ Sexo                    <chr> "FEMININO", "MASCULINO", "MASCULINO", ~
## $ Data.de.Nascimento      <chr> "28/12/1950", "10/02/1919", "14/11/19~
## $ Idade                   <int> 49, 81, 60, 79, 81, 78, 86, 44, 51, 7~
## $ Raca.Cor                <chr> "SEM INFORMAÇÃO", "SEM INFORMAÇÃO", "~
## $ Nacionalidade           <chr> "Não informado", "Não informado", "Nã~
## $ Naturalidade.Estado     <chr> "Não informado", "Não informado", "Nã~
## $ Naturalidade            <chr> "", "", "", "", "", "Não Informado", ~
## $ Grau.de.Instrução       <chr> "SEM INFORMAÇÃO", "SEM INFORMAÇÃO", "~
## $ Estado.Civil            <chr> "SEM INFORMAÇÃO", "SEM INFORMAÇÃO", "~
## $ Código.Profissão        <int> 0, 723, 0, 0, 621, 0, 0, 0, 0, 10, ~
## $ Nome.Profissão          <chr> "SEM INFORMAÇÃO\t", "TRABALHADORES DE~
## $ Estado.Endereço         <chr> "MINAS GERAIS", "MINAS GERAIS", "MINA~
## $ Cidade.Endereço        <chr> "BELO HORIZONTE", "BELO HORIZONTE", "~
## $ Descrição.da.Topografia <chr> "UTERO, SOE", "BEXIGA, SOE", "ESTOMAG~
## $ Código.da.Topografia    <chr> "C559", "C679", "C169", "C250", "C169~

```

```
## $ Descrição.da.Morfologia      <chr> "NEOPLASMA MALIGNO", "NEOPLASMA MALIG~
## $ Código.da.Morfologia         <int> 80003, 80003, 80003, 80003, 80003, 80~
## $ Descrição.da.Doenca         <chr> "UTERO, SOE", "BEXIGA, SOE", "ESTOMAG~
## $ Código.da.Doenca            <chr> "C55", "C679", "C169", "C250", "C169"~
## $ Descrição.da.Doenca.Infantil <chr> "", "", "", "", "", "", "", "", "", ""~
## $ Código.da.Doenca.Infantil    <chr> "", "", "", "", "", "", "", "", "", ""~
## $ Descrição.da.Doenca.Adulto.Jovem <chr> "", "", "", "", "", "", "", "", "", ""~
## $ Código.da.Doenca.Adulto.Jovem <chr> "", "", "", "", "", "", "", "", "", ""~
## $ Indicador.de.Caso.Raro       <chr> "false", "false", "false", "false", ""~
## $ Meio.de.Diagnostico         <chr> "SDO", "SDO", "SDO", "SDO", "SEM INFO~
## $ Extensão                    <chr> "SEM INFORMAÇÃO", "SEM INFORMAÇÃO", ""~
## $ Lateralidade                <chr> "", "", "", "", "", "", "", "", "", ""~
## $ Estadiamento               <chr> "", "", "", "", "", "", "", "", "", ""~
## $ TNM                         <chr> "", "", "", "", "", "", "", "", "", ""~
## $ Status.Vital                <chr> "", "", "", "", "", "", "", "", "", ""~
## $ Tipo.do.Óbito               <chr> "CÂNCER", "CÂNCER", "CÂNCER", "CÂNCER~
## $ Data.do.Óbito               <chr> "11/02/2000", "22/04/2000", "12/04/20~
## $ Data.de.Último.Contato      <chr> "", "", "", "", "", "", "", "", "", ""~
## $ Data.de.Diagnostico         <chr> "11/02/2000", "22/04/2000", "12/04/20~
## $ Metástase.à.distância       <chr> "", "", "", "", "", "", "", "", "", ""~
## $ X                           <lg1> NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
```

```
library(lubridate)
# t<-pulmao %>%
#   dplyr::select(where(is.Date))

pulmao <-pulmao %>% arrange(Data.do.Óbito)
pulmao$teste<-floor_date(mdy(pulmao$Data.do.Óbito),"month")
```

```
## Warning: All formats failed to parse. No formats found.
```

```
# pulmao %>% filter(Data.do.Óbito>Data.de.Último.Contato,1,0 )

pulmao$tempo <- pulmao$Data.do.Óbito

pulmao$mes<-( month(as.Date( pulmao$Data.de.Diagnostico))-month( as.Date(pulmao$Data.do.Óbito)) )
```