

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS  
NÚCLEO DE EDUCAÇÃO À DISTÂNCIA  
Pós-graduação Lato Sensu em Ciência de Dados e Big Data

**Fábio Roberto Barros Padilha**

ANÁLISE DE DADOS NACIONAIS SOBRE A EVOLUÇÃO  
DAS SÍNDROMES RESPIRATÓRIAS AGUDAS GRAVES  
POR SARS-CoV-2, NO ANO DE 2020

**Belo Horizonte**  
**2021**

**Fábio Roberto Barros Padilha**

**ANÁLISE DE DADOS NACIONAIS SOBRE A EVOLUÇÃO  
DAS SÍNDROMES RESPIRATÓRIAS AGUDAS GRAVES  
POR SARS-CoV-2, NO ANO DE 2020**

Trabalho de Conclusão de Curso apresentado ao Curso de Especialização em Ciência de Dados e Big Data como requisito parcial à obtenção do título de especialista.

**Belo Horizonte  
2021**

# Sumário

<b>1. Introdução</b>	<b>2</b>
1.1 Contextualização . . . . .	2
1.2 Objetivos . . . . .	4
<b>2. Coleta dos dados</b>	<b>5</b>
2.1 Dados OpenDataSUS . . . . .	5
2.2 Dados do Instituto Brasileiro de Geografia e Estatística . . . . .	7
<b>3. Processamento</b>	<b>8</b>
3.1 Dados do Instituto Brasileiro de Geografia e Estatística . . . . .	8
3.2 OpenDataSUS . . . . .	9
3.3 Transformação . . . . .	9
3.3.1 Notificações . . . . .	10
3.3.2 Idade . . . . .	11
3.3.3 Raça . . . . .	12
3.3.4 Sexo . . . . .	13
3.3.5 Comorbidades . . . . .	14
3.3.6 Recursos do tratamento . . . . .	15
3.3.7 Evolução . . . . .	16
<b>4. Análise Exploratória</b>	<b>18</b>
4.1 Número absoluto de mortes de covid-19 no Brasil . . . . .	18
4.2 Taxa de mortalidade da covid-19 no Brasil . . . . .	20
4.3 Taxa de letalidade da covid-19 no Brasil . . . . .	22
<b>5. Inferência Estatística</b>	<b>25</b>
5.1 Perfil do brasileiro com maior e menor risco de morte . . . . .	25
<b>6. Machine learning</b>	<b>28</b>
6.1 Árvore de Decisão de Classificação . . . . .	28
6.2 Árvore de Decisão e Seleção de Atributos . . . . .	28
6.3 Árvore de Decisão de Classificação otimizado . . . . .	29
6.4 Naive Bayes . . . . .	30
<b>Referências Bibliográficas</b>	<b>31</b>
<b>Apêndice A - Taxa de Mortalidade Estados e Capitais</b>	<b>32</b>
<b>Anexo A - Ficha de Notificação</b>	<b>33</b>
<b>Anexo B - Dicionário de dados</b>	<b>35</b>

# 1. Introdução

Nos últimos anos, o aumento de dados disponíveis digitalmente e o aparecimento de novas tecnologias têm favorecido a produção de conhecimento útil para sociedade.

De acordo com Gilson Medeiros de Oliveira Junior (2010):

O uso de técnicas de inteligência artificial, mais especificamente algoritmos de aprendizagem para classificação de padrões, tem crescido bastante nos últimos anos pelo fato de auxiliarem na resolução de problemas complexos do mundo real que podem envolver um elevado número de variáveis [...]. (JUNIOR, 2010, p.1)[4]

Esses algoritmos de aprendizagem estão relacionados ao aprendizado de máquina (machine learning), que é uma parte de Inteligência Artificial (IA) onde um algoritmo computacional cria, a partir dos dados disponíveis, modelos de aprendizado para a resolução de problemas (MEHTA, 2017).[5]

Dentre as técnicas de aprendizagem de máquina, pode-se citar o algoritmo de Naive Bayes, importante classificador probabilístico, que é, frequentemente, utilizado em processamento de diagnósticos médicos (cálculo de probabilidades em cima de diagnóstico de doenças), segundo Pedro César Tebaldi Gomes (2019) [2].

Outros importantes algoritmos de aprendizagem são as máquinas de vetores suporte - SVMs (Support Vector Machines), que “[...] têm a capacidade de resolver problemas de classificação e regressão, adquirindo com o aprendizado na etapa de treinamento a capacidade de generalização” (JUNIOR, 2010, p.5). [4]

Por fim, pode-se destacar a técnica de mineração de dados de árvore de decisão na aprendizagem de máquina. De acordo com Hosokawa (2011) [3], a árvore de decisão é um modelo preditivo, representado por um gráfico em forma de árvore, contendo as decisões a serem tomadas e seus possíveis riscos, custo e prejuízos.

Neste sentido, este trabalho utilizou as técnicas citadas acima, dentre outras, para análise de dados do Sistema Único de Saúde (SUS) relacionados às Síndromes Respiratórias Agudas Graves (SARS) relacionadas à covid-19, em 2020. De forma a identificar as regiões do Brasil mais impactadas pela covid-19, bem como os perfis com maior risco de morte, contribuindo com informações importantes para implementação de políticas de saúde pública, a exemplo da elaboração de um plano nacional de vacinação.

## 1.1 Contextualização

O ano de 2020 foi marcado por isolamento, medo e insegurança, causados pela pandemia do novo coronavírus SARS-CoV-2 e, com ela, pelo aumento considerável dos casos

de óbitos por covid-19.

Durante todo o curso da pandemia, um dos principais problemas que as autoridades e os profissionais de saúde enfrentaram foi a falta de informação: de recursos médicos, como respiradores, e de um plano adequado para sua distribuição.

Como o novo coronavírus é a principal causa para o aumento dos casos de mortes de Síndrome Respiratória Aguda Grave (SARS), identificar as regiões do Brasil mais afetadas, bem como ser capaz de prever os perfis com maior risco de morte, será de grande ajuda para as autoridades na elaboração de políticas públicas na área da saúde.

Dado o atual volume de informações relacionados ao SARS-CoV-2 (Covid-19) e às Síndromes Respiratórias Agudas Graves disponíveis, é possível se utilizar da tecnologia, como a inteligência artificial, para realizar previsões que auxiliem as autoridades da área de saúde e os médicos na elaboração de planos no combate ao novo coronavírus.

Sobre o uso das informações geradas com base nestas análises preditivas, Roberto Moreira Santos (2017) afirma que:

A aplicação desses resultados provenientes das análises é de suma importância para vários setores: como o setor da saúde, que consegue mapear os riscos de certas doenças que ocorrem em determinados grupos de pacientes, ou a resposta aos melhores medicamentos [...] (SANTOS, 2017, p.10)[7]

Neste contexto, “A analítica preditiva pode ser apoiada por técnicas da computação como aprendizagem de máquina (machine learning) e a mineração de dados para aumentar seu potencial de precisão” (WAYNE, 2007 apud SANTOS, 2017, p.11).

Dada a situação de calamidade pública provocada pelo novo Coronavírus SARS-CoV-2, que levou o Brasil a ultrapassar a marca de 180 mil mortes por covid-19, em 2020, e a chegada da vacina, buscou-se utilizar, neste trabalho, técnicas de mineração de dados e machine learning, com apoio do RStudio para o desenvolvimento de um projeto de ciência de dados com propósito de contribuir com a contenção da doença no Brasil. As bibliotecas utilizadas na elaboração do projeto estão descritas na Tabela 1 apresentada abaixo:

ggplot2	gridExtra	geobr	geobr	dplyr	latexpdf
stringr	openxlsx	rpart	e1071	randomForest	scales

Tabela 1: Pacotes R

Assim, espera-se que as previsões apresentadas com base no uso dessas tecnologias gerem conhecimentos úteis que auxiliem as autoridades públicas na implementação de políticas públicas na área da saúde e capacitem melhor o país na contenção da doença.

## 1.2 Objetivos

A pandemia do SARS-CoV-2 está impondo ao Brasil e ao mundo o maior desafio sanitário dos últimos 100 anos, tendo sido a Gripe Espanhola, ocorrida no ano de 1918, a última pandemia registrada de grande impacto mundial. Neste contexto, há uma atuação ativa por parte das autoridades de saúde na implementação de políticas públicas para conter o avanço da doença e o crescente número de óbitos.

Embora o Brasil, tenha um dos melhores sistemas de saúde da América Latina, a capacidade de atendimento em todo país é bastante desigual. Nas regiões mais pobres, onde a capacidade de atendimento de saúde é menor, têm-se observado a existência de situações críticas, como a falta de médicos, de recursos e de leitos hospitalares.

Neste contexto, este trabalho tem como objetivo geral entender, de forma regionalizada, a evolução dos casos graves da doença, a fim de contribuir com as autoridades da área de saúde na construção de políticas públicas voltadas para redução do número de óbitos e para contenção do novo coronavírus.

É importante destacar que, para este trabalho, são considerados casos graves de covid-19 aqueles existentes no arquivo CSV do OpenDataSUS.

Por fim, têm-se os seguintes objetivos específicos que irão ajudar na consecução do objetivo geral deste trabalho de conclusão de curso:

- identificar as regiões brasileiras com maior número de óbitos por covid-19;
- identificar a taxa de mortalidade por covid-19 dos estados e capitais;
- realizar teste de hipótese da taxa de letalidade do Brasil;
- apresentar a média móvel da taxa de letalidade de covid-19 do Brasil e regiões, ao longo do ano de 2020;
- encontrar os perfis dos pacientes com maiores e menores riscos de morte por covid-19;
- utilizar técnicas de machine learning para identificar o percentual de acerto do risco de morte por covid-19, a partir dos atributos selecionados na pesquisa.

## 2. Coleta dos dados

Nesse trabalho foram utilizada 2 (duas) fontes de dados, uma disponibilizada pela Secretaria de Vigilância em Saúde no portal OpenDataSUS e a outra disponibilizada no portal do Instituto Brasileiro de Geografia e Estatística (IBGE), cujo detalhamento será apresentado nas seções seguintes.

### 2.1 Dados OpenDataSUS

Dentre os dados utilizados neste trabalho, destacam-se os dados fornecidos pela Secretaria de Vigilância em Saúde, que está vinculada ao Ministério da Saúde, no portal OpenDataSUS. Sobre os dados coletados, o portal OpenDataSUS informa que:[6]

Esta página tem como finalidade disponibilizar o legado dos bancos de dados (BD) epidemiológicos de SRAG, da rede de vigilância da Influenza e outros vírus respiratórios, desde o início da sua implantação (2009) até os dias atuais (2021), com a incorporação da vigilância da COVID-19. Atualmente, o sistema oficial para o registro dos casos e óbitos de SRAG é o Sistema de Informação da Vigilância Epidemiológica da Gripe (SIVEP-Gripe).

O arquivo utilizado neste trabalho foi disponibilizado para [download em formato CSV](#). É importante observar que este arquivo é atualizado semanalmente e que, no momento do download, continha dados até 04/01/2021.

Como a Secretaria de Vigilância em Saúde atualiza os dados semanalmente, para garantir o acesso ao arquivo analisado neste trabalho, fez-se necessário manter uma cópia no [Google drive](#).

O arquivo contém 1.103.183 (um milhão cento e três mil e cento e oitenta e três) linhas e 154 colunas. Cada linha do arquivo corresponde ao registro de um paciente com a síndrome respiratória aguda grave. As 154 colunas são campos coletados a partir da [ficha de notificação do paciente](#), além de informações geradas pelo sistema. A ficha de notificação e o dicionário de dados estão disponíveis, respectivamente, nos anexos A e B. O trecho de código para importação do arquivo segue descrito abaixo:

```
arquivo = read.csv("dados/INFLUD-04-01-2021.csv", header = T, sep = ";",  
na.strings = "", stringsAsFactors = T)
```

Após a importação dos dados, as colunas objeto da análise foram selecionadas e renomeadas, a fim de facilitar a identificação nas seções seguintes. Os trechos de código, para selecionar e renomear as colunas seguem descritos abaixo:

A) Selecionar:

```
pacientes = arquivo[ , c("DT_NOTIFIC", "SEM_NOT", "DT_SIN_PRI", "SEM_PRI", "SG_UF_NOT",  
"ID_MUNICIP", "CO_MUN_NOT", "ID_UNIDADE", "CO_UNI_NOT",  
"CS_SEXO", "NU_IDADE_N", "CS_GESTANT", "CS_RACA", "SURTO_SG", "NOSOCOMIAL", "FEBRE",  
"TOSSE", "GARGANTA", "DISPNEIA", "DESC_RESP", "SATURACAO", "DIARREIA", "VOMITO",  
"DOR_ABD", "FADIGA", "PERD_OLFT", "PERD_PALA", "FATOR_RISC", "PUERPERA",  
"CARDIOPATI", "HEMATOLOGI", "SIND_DOWN", "HEPATICA", "ASMA", "DIABETES",  
"NEUROLOGIC", "PNEUMOPATI", "IMUNODEPRE", "RENAL", "OBESIDADE", "OBES_IMC", "VACINA",  
"ANTIVIRAL", "HOSPITAL", "DT_INTERNA", "UTI", "DT_ENTUTI", "DT_SAIDUTI",  
"SUPORT_VEN", "RAIOX_RES", "TOMO_RES", "AMOSTRA", "TP_TES_AN", "RES_AN",  
"POS_AN_FLU", "TP_FLU_AN", "POS_AN_OUT", "AN_SARS2", "AN_VSR",  
"AN_PARA1", "AN_PARA2", "AN_PARA3", "AN_ADENO", "PCR_RESUL", "POS_PCRFLU",  
"TP_FLU_PCR", "PCR_FLUASU", "PCR_FLUBLI", "POS_PCROUT", "PCR_SARS2", "PCR_VSR",  
"PCR_PARA1", "PCR_PARA2", "PCR_PARA3", "PCR_PARA4", "PCR_ADENO", "PCR_METAP",  
"PCR_BOCA", "PCR_RINO", "TP_AM_SOR", "TP_SOR", "RES_IGG", "RES_IGM", "RES_IGA",  
"CLASSI_FIN", "CRITERIO", "EVOLUCAO", "DT_EVOLUCA")] ]
```

B) Renomear:

```
colnames(pacientes) = c("DT_NOTIFICACAO", "SEMANA_NOTIFICACAO", "DT_PRIMEIROS_SINTOMAS",  
"SEMANA_PRIMEIROS_SINTOMAS", "SG_UF", "MUNICIPIO",  
"COD_MUNICIPIO", "HOSPITAL", "COD_HOSPITAL", "CS_SEXO", "NU_IDADE", "CS_GESTANTE",  
"CS_RACA", "IN_SURTO_SINDROME_GRIPAL", "IN_NOSOCOMIAL", "IN_FEBRE", "IN_TOSSE",  
"IN_GARGANTA", "IN_DISPNEIA", "IN_DESCONFORTO_RESP", "IN_SATURACAO", "IN_DIARREIA",  
"IN_VOMITO", "IN_DOR_ABDOMINAL", "IN_FADIGA", "IN_PERDA_OLFTATO", "IN_PERDA_PALADAR",  
"IN_FATOR_RISC", "IN_PUERPERA", "IN_CARDIOPATI", "IN_HEMATOLOGI", "IN_SINDROME_DOWN",  
"IN_DOENCA_HEPATICA", "IN_ASMA", "IN_DIABETES", "IN_DOENCA_NEUROLOGICA",  
"IN_PNEUMOPATIA_CRONICA", "IN_IMUNODEPRESSAO", "IN_DOENCA_RENAL", "IN_OBESIDADE",  
"NU_OBESIDADE_IMC", "IN_RECEBEU_VACINA_GRIPE", "IN_USOU_ANTIVIRAL", "IN_HOSPITALIZADO",  
"DT_INTERNACAO", "IN_UTI", "DT_ENTRADA_UTI", "DT_SAIDA_UTI", "CS_SUPORTE_VENTILATORIO",  
"CS_RAIOX_RESULTADO", "CS_TOMOGRFIA_RESULTADO", "IN_COLETOU_AMOSTRA",  
"CS_TESTE_ANTIGENO", "CS_RESULTADO_ANTIGENO", "IN_POSITIVO_INFLUENZA",  
"CS_TIPO_INFLUENZA", "IN_POSITIVO_OUT_VIRUS", "IN_POSITIVO_SARS2",  
"IN_POSITIVO_VSR", "IN_POSITIVO_PARAINFLUENZA1", "IN_POSITIVO_PARAINFLUENZA2",  
"IN_POSITIVO_PARAINFLUENZA3", "IN_POSITIVO_ADENOVIRUS", "CS_PCR_RESULTADO",  
"IN_POSITIVO_PCR_INFLUENZA", "CS_TIPO_INFLUENZA_PCR", "CS_PCR_INFLUENZA_A_SUBTIPO",  
"CS_PCR_FLU_B_LINHAGEM", "IN_PCR_POSITIVO_OUT_VIRUS", "IN_PCR_POSITIVO_SARS2",  
"IN_PCR_POSITIVO_VSR", "IN_PCR_POSITIVO_PARAINFLUENZA1",  
"IN_PCR_POSITIVO_PARAINFLUENZA2", "IN_PCR_POSITIVO_PARAINFLUENZA3",  
"IN_PCR_POSITIVO_PARAINFLUENZA4", "IN_PCR_POSITIVO_ADENOVIRUS",  
"IN_PCR_POSITIVO_METAPNEUMOVIRUS", "IN_PCR_POSITIVO_BOCAVIRUS",  
"IN_PCR_POSITIVO_RINOVIRUS", "CS_TIPO_AMOSTRA_SOROLOGICA_SARS_COV2",  
"CS_TESTE_SOROLOGICO", "IN_SOROLOGICO_POSITIVO_IGG", "IN_SOROLOGICO_POSITIVO_IGM",  
"IN_RESULTADO_POSITIVO_IGA", "CS_CLASSICACAO_FINAL", "CS_CRITERIO_CONFIRMACAO",  
"CS_EVOLUCAO", "DT_EVOLUCAO")
```



## 2.2 Dados do Instituto Brasileiro de Geografia e Estatística

Fez-se necessário enriquecer este trabalho com dados do Instituto Brasileiro de Geografia e Estatística - IBGE, tanto para o cálculo das taxas de mortalidades, bem como para elaboração dos mapas.

O IBGE disponibiliza o arquivo neste [link](#), no diretório: `\textbf{\ Estima-tiva_de_Populacao/Estimativas_2020}`, em formato xls, conforme Figura 1 abaixo:

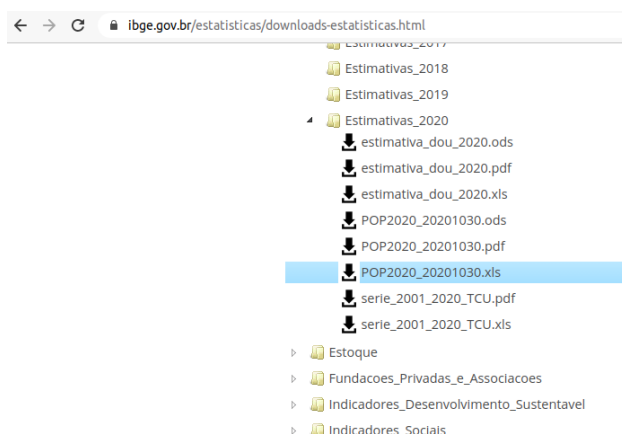


Figura 1: Arquivo IBGE

Após o download, o arquivo foi convertido para o formato xlsx, no LibreOffice, para facilitar a importação para RStudio. O arquivo contém duas abas, uma com a estimativa populacional por unidade federativa e outra aba com a estimativa populacional por municípios.

Para este trabalho, importamos apenas os dados da aba municípios, uma vez que a população por estado ou por região pode ser facilmente obtida a partir do somatório populacional dos municípios.

A importação dos dados foi iniciada a partir da terceira linha do arquivo, a fim de desconsiderar as linhas referentes ao cabeçalho, conforme trecho de código a seguir:

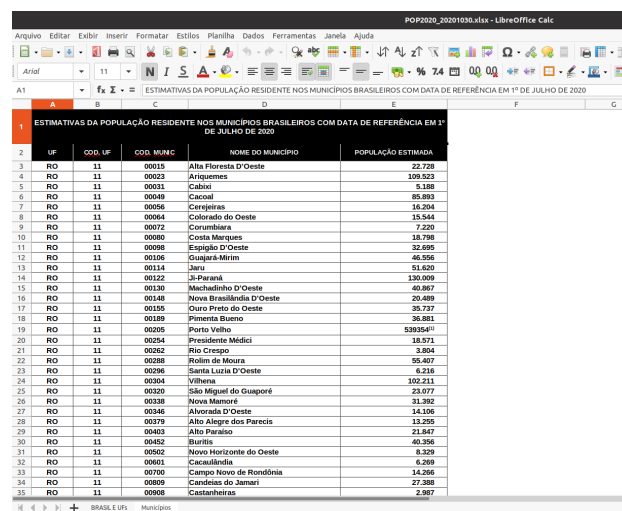
```
arquivoIBGE=read.xlsx("dados/POP2020_20201030.xlsx", sheet = "Municípios",  
startRow = 3, colNames =FALSE, rowNames = FALSE)
```

### 3. Processamento

Este capítulo contempla as transformações realizadas nos data frames resultantes das importações dos arquivos CSV (disponibilizado no portal OpenDataSUS) e XLS (disponibilizado no portal do IBGE). Estes campos tratados são os principais atributos utilizados nos algoritmos e gráficos constantes nos capítulos seguintes.

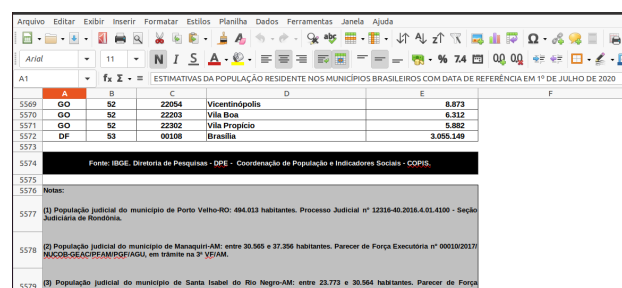
#### 3.1 Dados do Instituto Brasileiro de Geografia e Estatística

Os dados populacionais do IBGE, necessários para realização deste trabalho, estão presentes da linha 3 até a linha 5572, de acordo com as Figuras 2 e 3:



UF	COD. UF	COD. MUNIC	NOME DO MUNICÍPIO	POPULAÇÃO ESTIMADA
RO	11	00015	Alta Floresta D'Oeste	22.728
RO	11	00023	Araruama	109.623
RO	11	00031	Araruama	5.188
RO	11	00049	Caracol	85.893
RO	11	00056	Cariacás	16.254
RO	11	00064	Colorado do Oeste	15.544
RO	11	00072	Conceição	7.220
RO	11	00080	Costa Marques	18.786
RO	11	00088	Espigão D'Oeste	32.695
RO	11	00106	Guajará-Mirim	46.556
RO	11	00114	Itaú	51.620
RO	11	00122	J. Paraná	130.009
RO	11	00130	Machadinho D'Oeste	40.867
RO	11	00148	Nova Brasília D'Oeste	20.489
RO	11	00155	Ouro Preto do Oeste	35.737
RO	11	00169	Pimenta Bueno	36.881
RO	11	00205	Porto Velho	539354 <sup>1</sup>
RO	11	00254	Presidente Médici	18.571
RO	11	00262	Rio Crespo	3.854
RO	11	00288	Rolim de Moura	55.407
RO	11	00296	Santa Luzia D'Oeste	6.216
RO	11	00304	Vilhena	102.111
RO	11	00320	São Miguel do Guaporé	23.077
RO	11	00338	Nova Mamoré	31.362
RO	11	00346	Alvarado D'Oeste	14.156
RO	11	00379	Alto Alegre dos Parecis	13.255
RO	11	00403	Alto Páris	21.847
RO	11	00452	Buritis	40.356
RO	11	00502	Novo Horizonte do Oeste	8.329
RO	11	00611	Cavaliândia	6.269
RO	11	00700	Campo Novo de Rondônia	14.266
RO	11	00809	Candonga do Jari	27.388
RO	11	00908	Castanheira	2.387

Figura 2: Cabeçalho arquivo IBGE



Fonte: IBGE, Diretoria de Pesquisas - DPE - Coordenação de População e Indicadores Sociais - COPIS.				
Notas:				
(1) População judicial do município de Porto Velho-RO: 484.513 habitantes. Processo Judicial nº 12319-40.2018.4.01.4100 - Seção Judiciária de Rondônia.				
(2) População judicial do município de Manaus-AM: entre 30.565 e 37.356 habitantes. Parecer de Força Executória nº 00018/2023-NUCOP-GEAC-PPAM/PPF/AGU, em trâmite na 2ª VFIAM.				
(3) População judicial do município de Santa Isabel do Rio Negro-AM: entre 23.773 e 30.564 habitantes. Parecer de Força Executória nº 00007/2023-NUCOP-GEAC-PPAM/PPF/AGU, em trâmite na 2ª VFIAM.				

Figura 3: Rodapé arquivo IBGE

A importação dos dados foi iniciada a partir da linha três do arquivo, descartando o cabeçalho. Para descartar o rodapé do data frame resultante da importação, foi selecionado somente até a linha 5.570 (cinco mil quinhentos e setenta). Por fim, foi criado um campo extra no data frame para armazenar o código do município sem o último dígito.

A necessidade de criar uma coluna extra para o município sem o dígito verificador ocorreu devido à ausência do dígito verificar no arquivo CSV disponibilizado no portal OpenDataSUS. O campo criado com o código do município sem o dígito verificador foi utilizado em JOINS nas seções seguintes, onde serão criados gráficos pelas diversas localidades do país.

É importante destacar que a informação das regiões do país (norte, nordeste, centro-oeste, sudeste e sul) não estavam presentes nos arquivos apresentados. Tornando-se necessária a criação de uma matriz com os estados e suas respectivas regiões. Essa matriz permitiu o enriquecimento dos dados e a produção dos gráficos sumarizados, conforme trecho de código abaixo:

```
regiao = matrix(data = NA, nrow = 27, ncol = 2, byrow = TRUE)
regiao[1,] = c("AC", "Norte")
regiao[2,] = c("AP", "Norte")
regiao[3,] = c("AM", "Norte")
regiao[4,] = c("PA", "Norte")
regiao[5,] = c("RO", "Norte")
regiao[6,] = c("RR", "Norte")
regiao[7,] = c("TO", "Norte")
regiao[8,] = c("AL", "Nordeste")
regiao[9,] = c("BA", "Nordeste")
#Continua para todos os 27 estados...
```

```
pacientes = pacientes %>% left_join(regiao, c("SG_UF" = "SG_UF"))
```

## 3.2 OpenDataSUS

A transformação dos dados do SUS, disponibilizados no portal OpenDataSUS, foi realizada em duas etapas. Na primeira etapa, o processamento teve como finalidade favorecer a elaboração de gráficos no ggplot2 a fim de apresentar a distribuição dos dados. Já na segunda etapa, outras transformações aconteceram com o propósito de ajustar os dados para as execuções dos algoritmos de machine learning.

## 3.3 Transformação

A partir dessa seção será feita a limpeza e transformação dos dados, de maneira mais significativa. Nas seções anteriores, o processamento dos dados se restringiu a importar os arquivos objeto da análise e enriquecer um dos data frame resultante da importação com as regiões do país. Para melhor visualizar o que já foi realizado até o momento, segue Figura 4 com fluxo do pré-processamento dos dados:

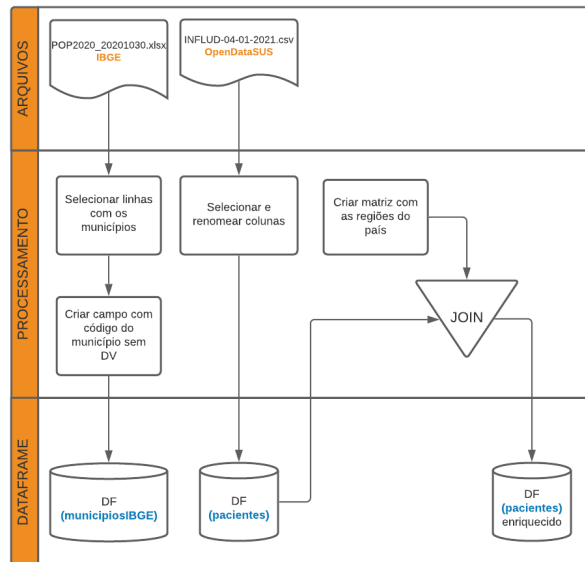


Figura 4: Pré-processamento dos dados

Nas subseções abaixo, será realizado o detalhamento das principais colunas do arquivo CSV disponibilizado no portal OpenDataSUS. Ressalta-se que não foi necessário detalhar a transformação do arquivo disponibilizado pelo IBGE, uma vez que o mesmo não precisou de transformações significativas, além das já mencionadas no item 3.1.

É importante enfatizar, ainda, que foi necessário elaborar gráficos preliminares nessas subseções, a fim de validar o processamento. No entanto, os gráficos mais elaborados e que trazem informações mais importantes para este trabalho estão dispostos no capítulo 4.

### 3.3.1 Notificações

O campo DT\_NOTIFICACAO, que contém a data do preenchimento da ficha de notificação, foi ajustado para o formato dd/mm/aaaa. Após a conversão do campo, foi necessária a realização de algumas limpezas nos dados, como a remoção dos registros com data de notificação superior a 31/12/2020 do data frame, visto que o escopo desse trabalho é o ano de 2020, bem como a retirada das notificações que não foram classificadas como casos de covid-19. Conforme trecho de código abaixo:

```
#Redefinir tipo de dados da coluna
pacientes$DT_NOTIFICACAO = as.Date(pacientes$DT_NOTIFICACAO, format="%d/%m/%Y")
pacientes["MES_NOTIFICACAO"] = format(pacientes$DT_NOTIFICACAO, "%Y/%m")
#Removendo ocorrencia com datas de notificação superior a 31/12/2020
pacientes=pacientes[pacientes$DT_NOTIFICACAO<as.Date("01/01/2021", format="%d/%m/%Y"),]
```

É importante destacar que não existem valores omissos na coluna DT\_NOTIFICACAO. A partir deste processamento foi possível a criação do Gráfico 1 (Notificações de

Covid-19 por Mês) e verificar que no Brasil, em 2020, os primeiros casos de covid-19 foram registrados em fevereiro e que, no mês de julho, ocorreu o maior número de casos.

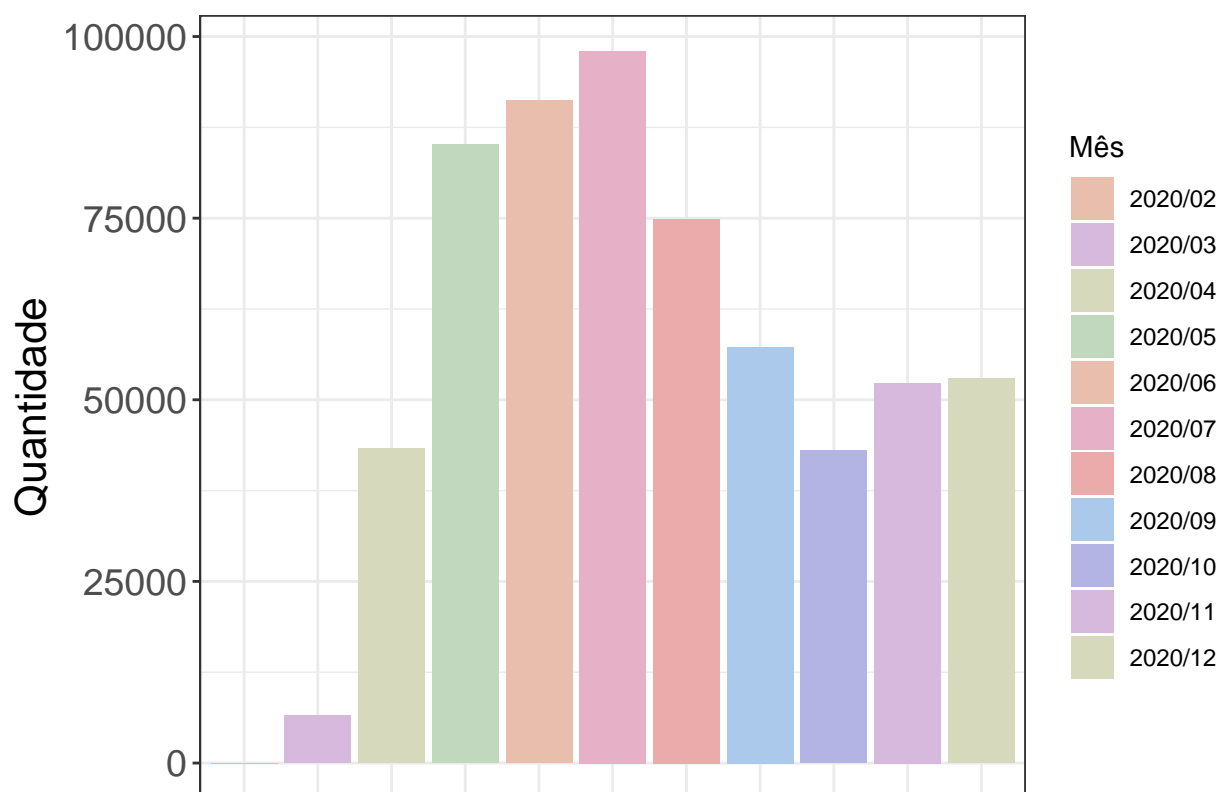


Gráfico 1: Notificações de Covid-19 por Mês

### 3.3.2 Idade

No processamento da coluna `NU_IDADE`, atribuiu-se a média da idade nacional dos casos de covid-19, que é de 59 (cinquenta e nove) anos, nas ocorrências com valores superiores a 110 (cento e dez) anos. Essa transformação teve como objetivo minimizar os valores outliers e os prováveis erros de digitação. Conforme trecho de código abaixo:

```
#Idades acima de 110 anos atribui a idade média.
pacientes$NU_IDADE=as(pacientes$NU_IDADE, "integer");
pacientes[pacientes$NU_IDADE>110,]$NU_IDADE=as.integer(mean(pacientes$NU_IDADE, na.rm=T))
```

Igualmente ao ocorrido no item anterior indentificou-se que não há valores omissos na coluna `NU_IDADE`.

Com base nesse processamento foi elaborado o Gráfico 2 (Boxplot de Idades dos casos graves de covid-19 por região), no qual fica evidenciado que:

- a região centro-oeste e a região nordeste apresentam respectivamente a menor e a maior média nacional de idade dos casos graves de covid-19;
- existem outlier abaixo de 15 anos nas regiões nordeste, suldeste e sul;

- aproximadamente 75% dos casos graves de covid-19 possuem idade acima dos 40 anos.

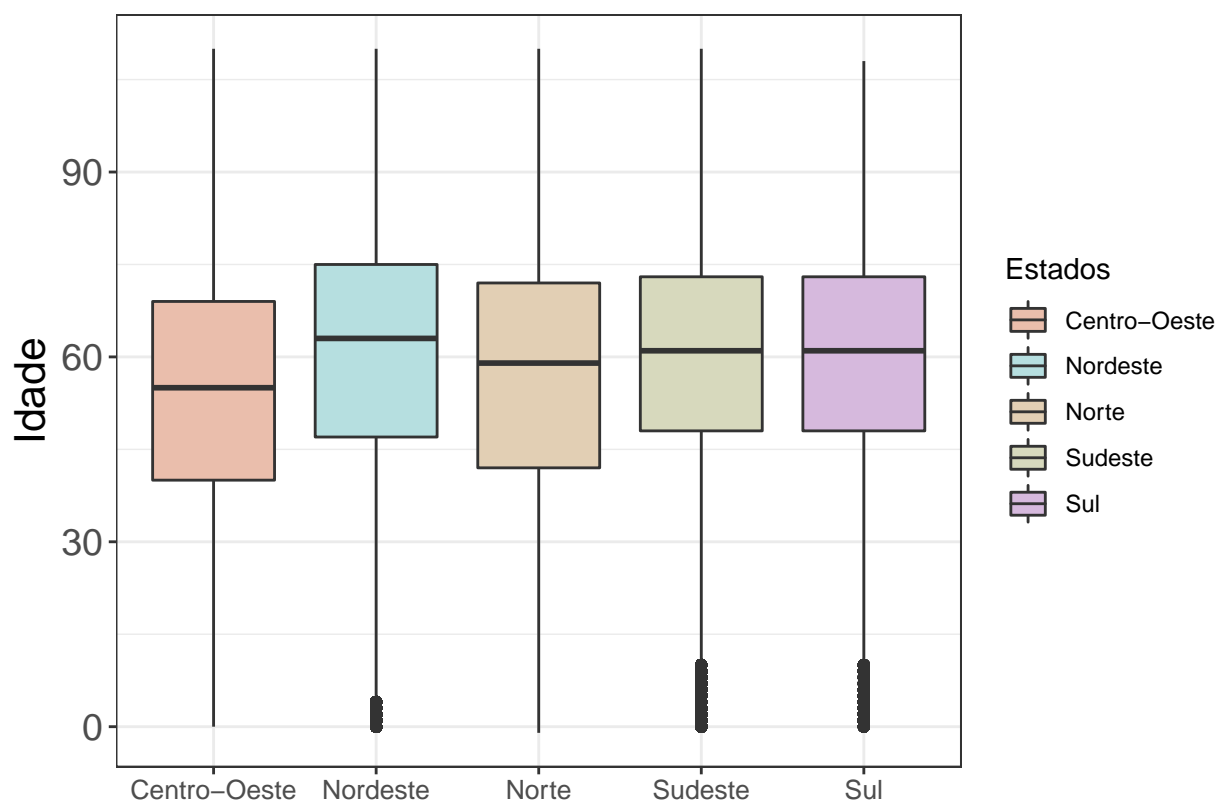


Gráfico 2: Boxplot de Idades dos casos graves de covid-19 por região

### 3.3.3 Raça

Na coluna CS\_RACA, os valores numéricos foram substituídos pelas strings correspondentes, conforme descrito no dicionário de dados disponível na seção anexo. Essas alterações são verificáveis no trecho de código a seguir:

```
#Transformando domínios da coluna raça em valores mais significativos para o gráfico
pacientes[ complete.cases(pacientes$CS_RACA) & pacientes$CS_RACA==9,]$CS_RACA = NA
pacientes[ complete.cases(pacientes$CS_RACA) & pacientes$CS_RACA==1,]$CS_RACA = "Branca"
pacientes[ complete.cases(pacientes$CS_RACA) & pacientes$CS_RACA==2,]$CS_RACA = "Preta"
pacientes[ complete.cases(pacientes$CS_RACA) & pacientes$CS_RACA==3,]$CS_RACA = "Amarela"
pacientes[ complete.cases(pacientes$CS_RACA) & pacientes$CS_RACA==4,]$CS_RACA = "Parda"
pacientes[ complete.cases(pacientes$CS_RACA) & pacientes$CS_RACA==5,]$CS_RACA = "Indígena"
pacientes$CS_RACA = factor(pacientes$CS_RACA)
```

É importante destacar que, na coluna CS\_RACA, o percentual de valores é de 22% (vinte e dois por cento). Esses valores foram removidos nas elaborações de gráficos e modelos de machine learning.

Este processamento possibilitou a construção do Gráfico 3 (Quantitativo dos casos graves de covid-19 por raça), revelando que:

- a raça com maior representatividade de casos graves de covid-19 é a branca seguida

pela parda;

- a raça indígena foi a que registrou menos casos.

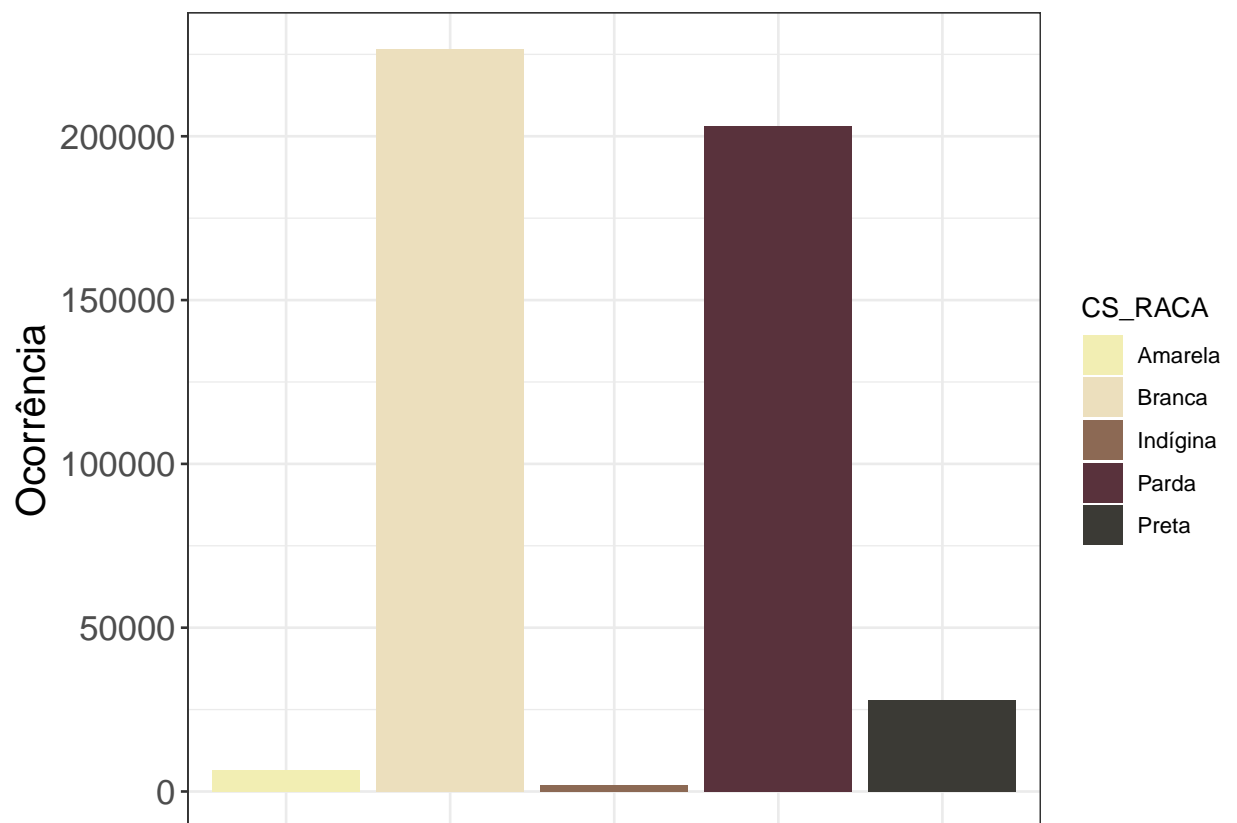


Gráfico 3: Quantitativo dos casos graves de covid-19 por raça

### 3.3.4 Sexo

O percentual de valores omissões na coluna CS\_SEXO é de 0.01% (zero vírgula zero um por cento). Esses valores foram tratados a partir da coluna CS\_GESTANTE, atribuiu-se sexo feminino, quando esta coluna informa existência de gestação. Em contrapartida, pra os valores omissões de sexo que não apresem indicação de gestação, atribuiu-se sexo masculino, dado que a presença do sexo masculino é maior (vide Gráfico 4 abaixo) e a incidência de valores omissos é muito pequena.

Abaixo segue o trecho de código responsável por este processamento:

```
pacientes[is.na(pacientes$CS_SEXO) & pacientes$CS_GESTANTE>4, ]$CS_SEXO="M"  
pacientes[is.na(pacientes$CS_SEXO) & pacientes$CS_GESTANTE<=4, ]$CS_SEXO="F"  
pacientes$CS_SEXO = factor(pacientes$CS_SEXO)
```

Com base neste processamento foi gerado o Gráfico 4 (Percentual dos casos graves de covid-19 por sexo), que destaca a maior incidência de casos graves de covid-19 no sexo masculino.

## Sexo

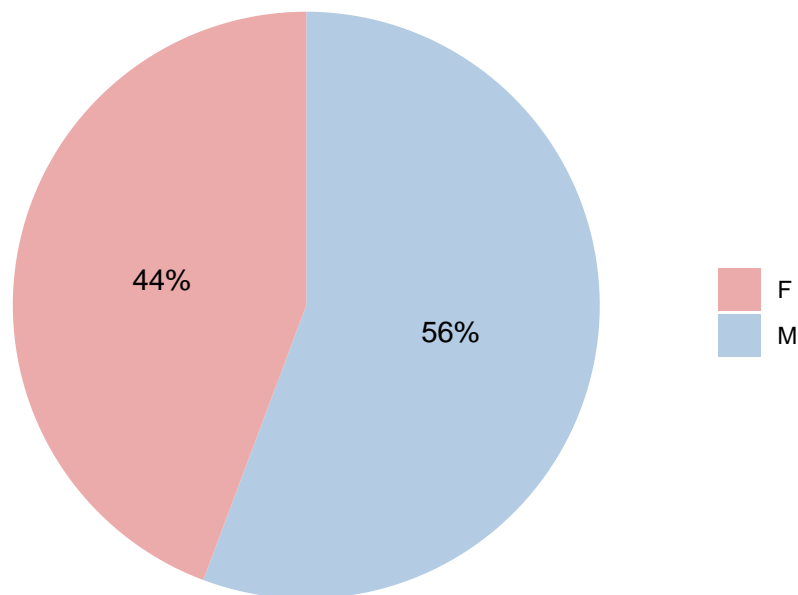


Gráfico 4: Percentual dos casos graves de covid-19 por sexo

### 3.3.5 Comorbidades

Nas colunas referente às comorbidades do paciente, foram convertidos os valores numéricos para as strings correspondentes, conforme descrito no dicionário de dados, disponível no Anexo B, deste trabalho.

A coluna `IN_FATOR_RISC`, informa a existência de comorbidade no paciente. Nela não foram identificados valores omissos.

No trecho de código abaixo verifica-se parte do processamento realizado nas colunas referente a comorbidades.

```
#Puerpera?
pacientes[complete.cases(pacientes$IN_PUERPERA) &
  pacientes$IN_PUERPERA==9,]$IN_PUERPERA=NA
pacientes[complete.cases(pacientes$IN_PUERPERA) &
  pacientes$IN_PUERPERA==1,]$IN_PUERPERA="Sim"
pacientes[complete.cases(pacientes$IN_PUERPERA) &
  pacientes$IN_PUERPERA==2,]$IN_PUERPERA="Não"
pacientes$IN_PUERPERA=factor(pacientes$IN_PUERPERA)
#Cardiopatia?
pacientes[complete.cases(pacientes$IN_CARDIOPATI) &
  pacientes$IN_CARDIOPATI==9,]$IN_CARDIOPATI=NA
```



```
pacientes[complete.cases(pacientes$IN_CARDIOPATI) &
  pacientes$IN_CARDIOPATI==1, ]$IN_CARDIOPATI="Sim"
pacientes[complete.cases(pacientes$IN_CARDIOPATI) &
  pacientes$IN_CARDIOPATI==2, ]$IN_CARDIOPATI="Não"
pacientes$IN_CARDIOPATI=factor(pacientes$IN_CARDIOPATI)
#Continua para todos os campos do grupo comorbidades...
```

Com base no processamento acima foi possível gerar o Gráfico 5 (Comorbidades preexistentes nos casos graves de covid-19) no qual verifica-se a maior incidência de comorbidades referente a cardiopatias e diabetes.

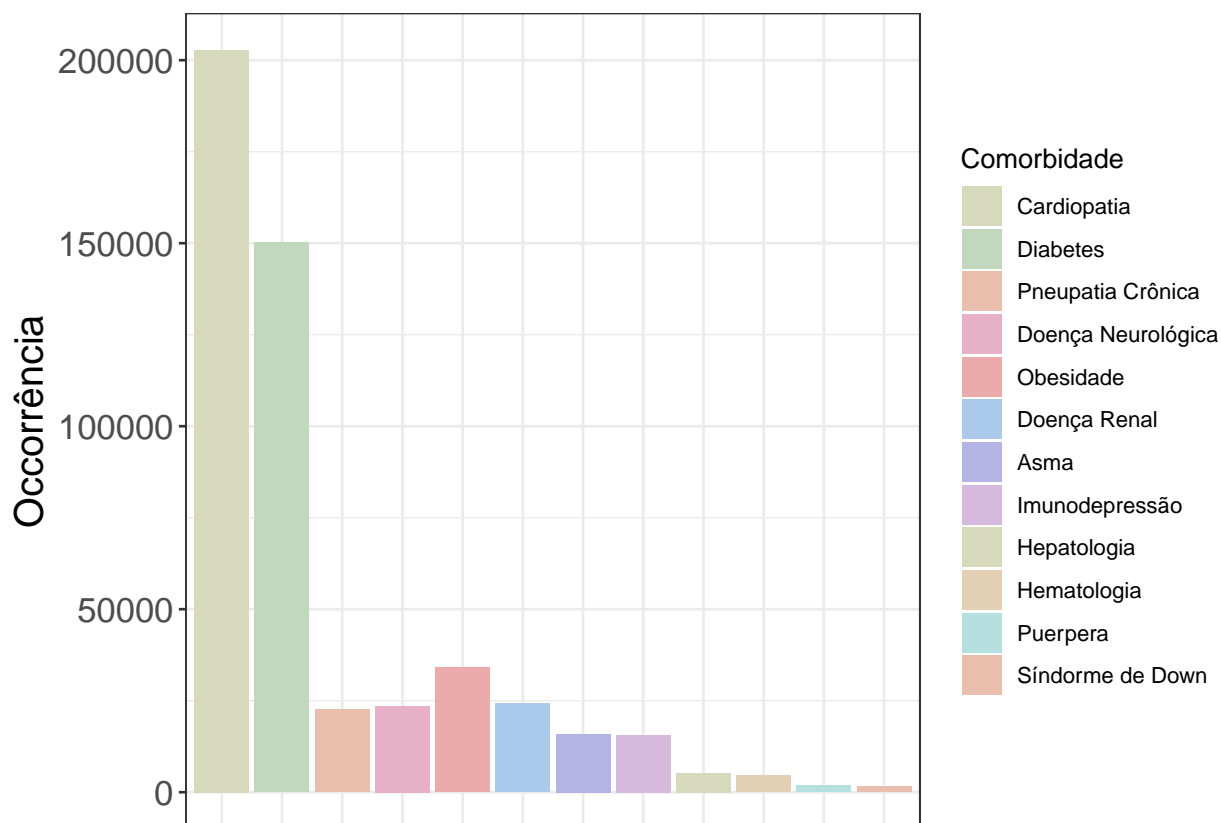


Gráfico 5: Comorbidades preexistentes nos casos graves de covid-19

### 3.3.6 Recursos do tratamento

Nesta seção constam os recursos que foram utilizados pelo paciente durante o tratamento. Os recursos são: se o paciente precisou de oxigênio, se precisou de suporte de ventilação pulmonar, se foi pra UTI, se tomou algum antiviral ou se foi hospitalizado.

Ressalta-se que a coluna mais utilizada deste grupo foi a IN\_HOSPITALIZADO, cujo percentual de valores omissos é de 2.61% (dois virgula sessenta e um por cento), estes valores foram removidos na geração dos gráficos e modelos de machine learning.

Nas colunas referente a recursos do tratamento do paciente, foram convertidos os valores numéricos para as strings correspondentes, conforme descrito no dicionário de dados,

disponível no Anexo B. O trecho de código abaixo reflete este processamento:

```
#Foi Hospitalizado?
pacientes[complete.cases(pacientes$IN_HOSPITALIZADO) &
  pacientes$IN_HOSPITALIZADO==1, ]$IN_HOSPITALIZADO="Sim"
pacientes[complete.cases(pacientes$IN_HOSPITALIZADO) &
  pacientes$IN_HOSPITALIZADO==2, ]$IN_HOSPITALIZADO="Não"
pacientes$IN_HOSPITALIZADO=factor(pacientes$IN_HOSPITALIZADO)
#Continua para os demais campos...
```

Esse processamento permitiu a elaboração do Gráfico 6 (Recursos utilizados nos casos graves de covid-19), do qual concluí-se que aproximadamente metade dos pacientes hospitalizados precisaram de oxigênio não invasivo.

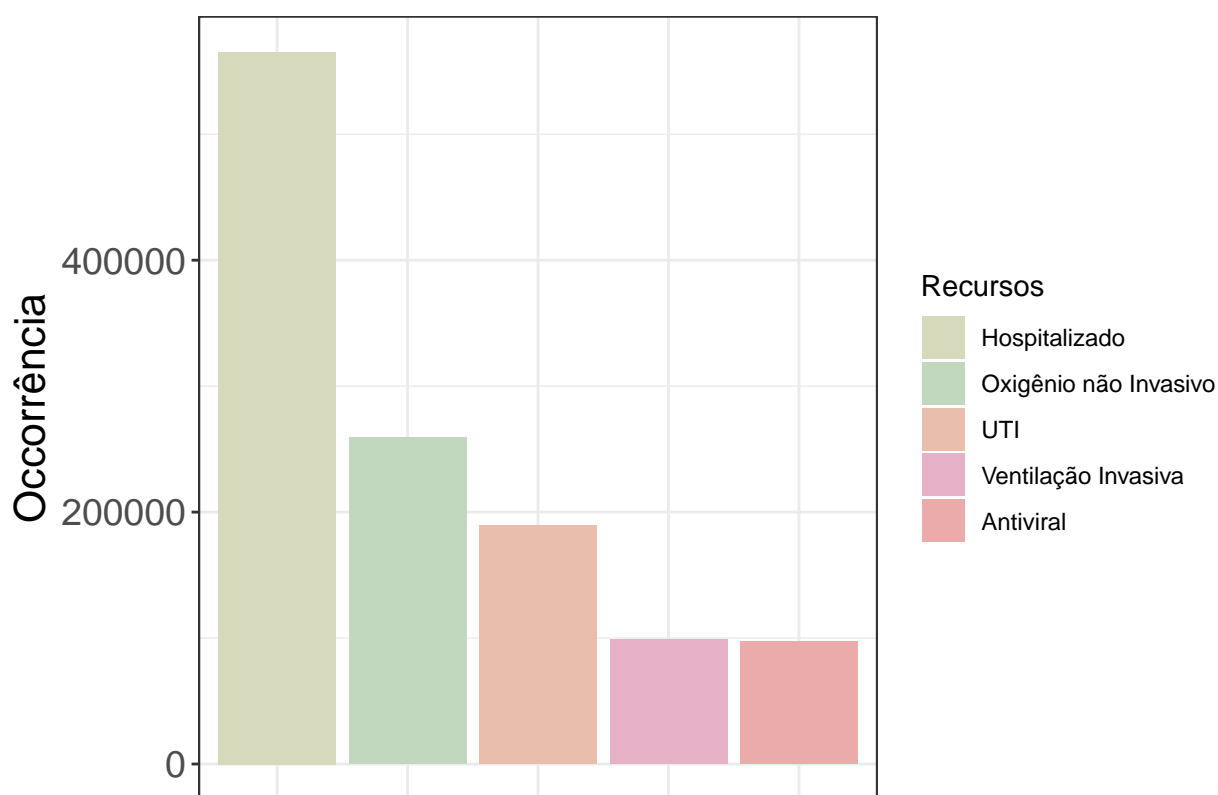


Gráfico 6: Recursos utilizados nos casos graves de covid-19

### 3.3.7 Evolução

Na coluna CS\_EVOLUCAO, os valores numéricos foram convertidos para as strings correspondentes, conforme descrito no dicionário de dados, disponível no Anexo B, deste trabalho.

O percentual de valores omissos nesta coluna é de 12,31% (doze virgula trinta e um por cento), para esses valores foram atribuído a valor “Em tratamento”.

O trecho de código abaixo reflete o processamento desta coluna:

```
#Data Evolução / Desfecho
pacientes$DT_EVOLUCAO = as.Date(pacientes$DT_EVOLUCAO, format="%d/%m/%Y")
pacientes["SEMANA_EVOLUCAO"] = format(pacientes$DT_EVOLUCAO, "%U");
pacientes$SEMANA_EVOLUCAO=as.integer(pacientes$SEMANA_EVOLUCAO)
#Resultado Evolução / Desfecho
pacientes[complete.cases(pacientes$CS_EVOLUCAO) &
  pacientes$CS_EVOLUCAO==1, ]$CS_EVOLUCAO="Cura"
pacientes[complete.cases(pacientes$CS_EVOLUCAO) &
  pacientes$CS_EVOLUCAO==2, ]$CS_EVOLUCAO="Óbito"
pacientes[complete.cases(pacientes$CS_EVOLUCAO) &
  pacientes$CS_EVOLUCAO==3, ]$CS_EVOLUCAO="Óbito por outras causas"
pacientes[is.na(pacientes$CS_EVOLUCAO), ]$CS_EVOLUCAO="Em tratamento"
pacientes[complete.cases(pacientes$CS_EVOLUCAO) & pacientes$CS_EVOLUCAO==9,]$CS_EVOLUCAO="Em tratamento"
pacientes$CS_EVOLUCAO=factor(pacientes$CS_EVOLUCAO)
```

A partir desse processamento foi elaborado o Gráfico 7 (Evolução dos casos graves de covid-19), do qual concluí-se que, aproximadamente, 1/3 (um terço) dos pacientes graves de covid-19 evoluem para óbito.

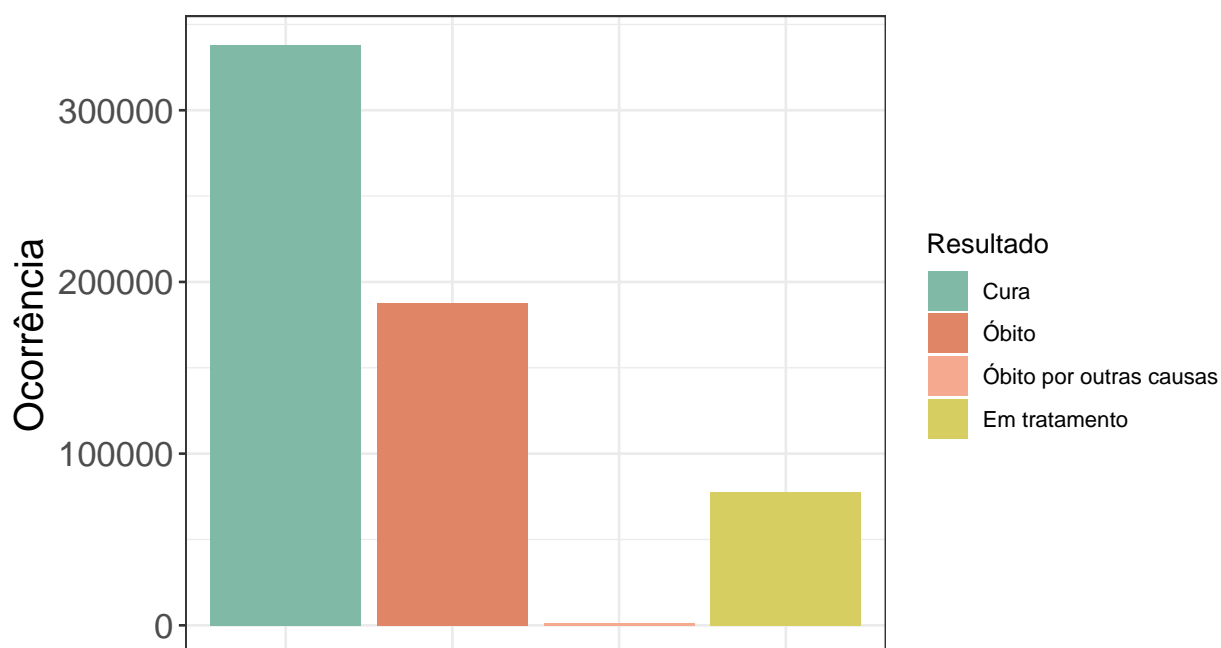


Gráfico 7: Evolução dos casos graves de covid-19

## 4. Análise Exploratória

Neste capítulo, serão detalhadas as taxas de mortalidade e letalidade, bem como o número de mortes absolutos decorrentes de covid-19 no Brasil, em 2020. É importante observar que essa análise utilizou a técnica de drill down, detalhando as informações até o nível de capitais.

Com intuito de facilitar as análises presente neste capítulo, torna-se necessário nivelar alguns conceitos, tais como: taxas de mortalidade e letalidade.

A taxa de mortalidade é a razão entre o número total de óbitos e o total da população vezes uma constante utilizada como unidade de medida, neste trabalho será adotado 1.000.000 (um milhão) de habitantes.

Já a taxa de letalidade é a razão entre o número total de óbitos e o número total de casos graves.

Essas taxas foram calculadas exclusivamente a partir dos casos graves de covid-19. Deve-se lembrar que, para este trabalho, considera-se casos graves de covid-19 aqueles presente no arquivo CSV do OpenDataSUS.

### 4.1 Número absoluto de mortes de covid-19 no Brasil

No ano de 2020, o Brasil teve 187.727 (cento e oitenta e sete mil e setecentos e vinte e sete) mortes por covid-19. Neste sentido, o Sudeste apresentou quase a metade dos casos de mortes por covid-19 do país, com 88.362 (oitenta e oito mil e trezentos e sessenta e duas) mortes. As demais regiões registraram juntas um número de mortes um pouco superior, totalizando 99.365 (noventa e nove mil e trezentos e sessenta e cinco) mortes, o que representa 52,93% (cinquenta e dois vírgula noventa e três por cento) do total de óbitos registrados no país por covid-19.

A este respeito, o Gráfico 8 (Mortes de covid-19 por região) apresenta a distribuição das mortes decorrente de covid-19, de forma regionalizada, conforme pode ser visto abaixo:

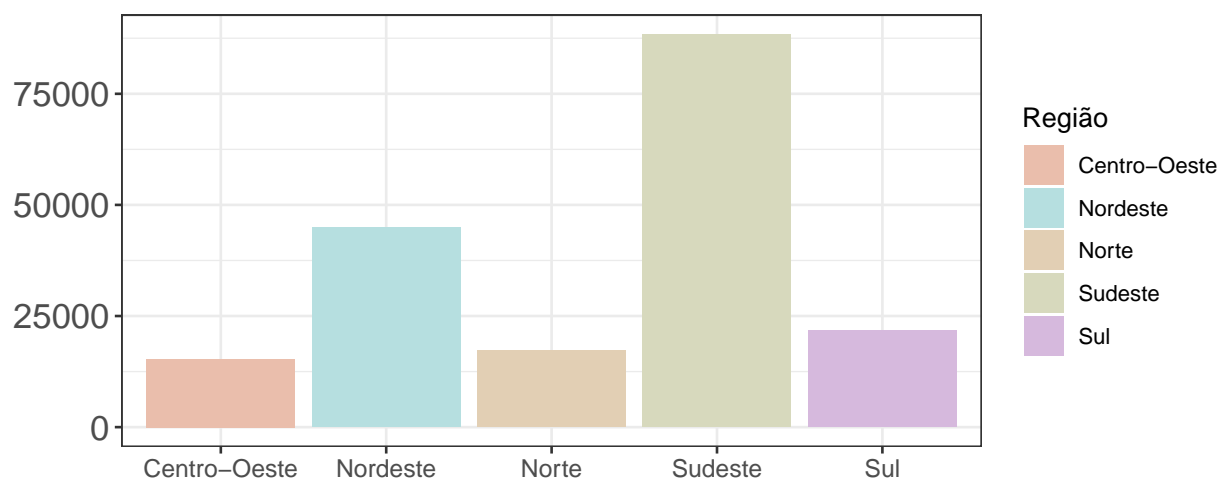


Gráfico 8: Mortes de covid-19 por região

Após a região sudeste, o Nordeste aparece como a segunda região com maior número de mortes do Brasil, totalizando 44.965 (quarenta e quatro mil e novecentos e sessenta e cinco) mortes, ou seja, 23,95% (vinte e três vírgula noventa e cinco por cento) do total de mortes por covid-19 do país.

No Sudeste, o Estado de São Paulo foi responsável por 53,45% (cinquenta e três vírgula quarenta e cinco por cento) das mortes, registrando 47.226 (quarenta e sete mil duzentos e vinte e seis) óbitos, ou seja, mais da metade das mortes na região, como pode ser visto no gráfico 9 (Mortes por covid-19 na região Sudeste) apresentado abaixo:

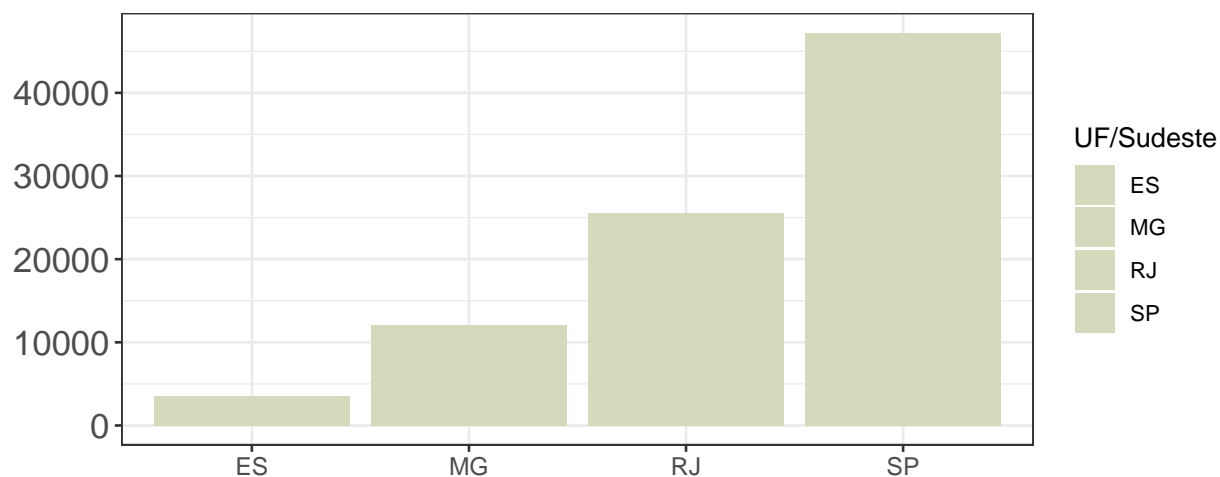


Gráfico 9: Mortes por covid-19 na região Sudeste

É importante observarmos que o Estado de São Paulo apresenta mais mortes do que a região nordeste inteira, que é a segunda maior região com casos de morte por covid-19 no país.

No que se refere ao Nordeste, o Estado do Ceará apresenta, aproximadamente, 1/4 (um quarto) do número de óbitos, no total de 10.281 (dez mil duzentos e oitenta e um) óbitos, seguido por Pernambuco e Bahia, com 9.597 (nove mil quinhentos e noventa e sete) e 8.321

(oito mil trezentos e vinte e um) óbitos, respectivamente. Esses três estados, juntos, têm quase 60% (sessenta por cento) das mortes por covid-19 no Nordeste. Sobre este assunto, o Gráfico 10 (Mortes por covid-19 na região Nordeste), exibido abaixo, apresenta a distribuição de mortes por covid-19 na região.

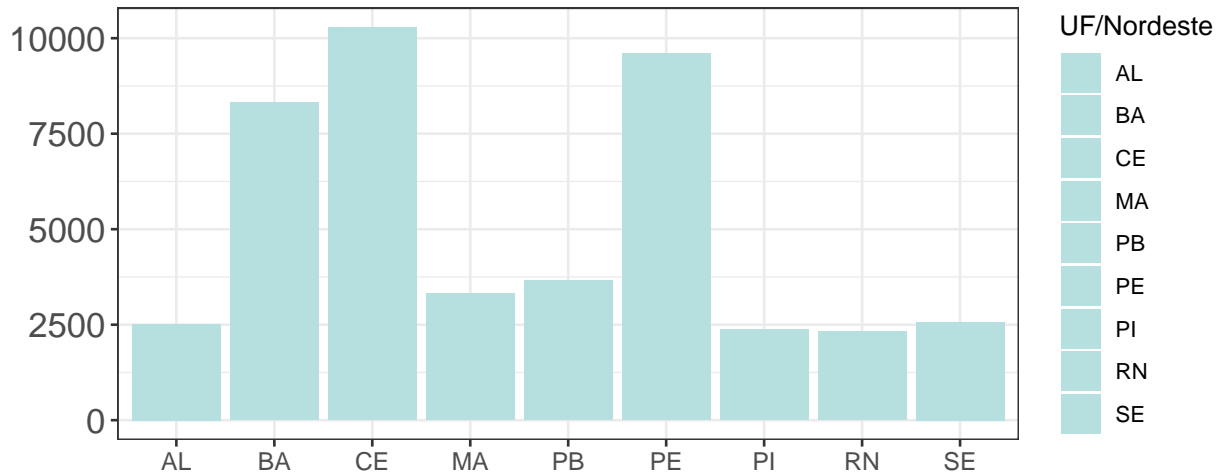


Gráfico 10: Mortes por covid-19 na região Nordeste

## 4.2 Taxa de mortalidade da covid-19 no Brasil

Apesar do estado de São Paulo liderar o números de mortes absoluto por covid-19 no Brasil, em relação ao tamanho da população, a situação é diferente. Verifica-se que 5 (cinco) estados do país ultrapassam São Paulo em mortes por milhão de habitantes. O Rio de Janeiro é o estado com a mais alta taxa de mortalidade do país, seguido pelo Distrito Federal. Para desmostrar segue a Tabela 2}.

RJ	DF	AM	CE	SE	SP
1.470	1.439	1.254	1.119	1.110	1.020

Tabela 2: Estados com maiores taxas de mortalidade por milhão de habitantes

Apesar da Tabela 2} apresentar apenas os 6 (seis) estados com as maiores taxas de mortalidade, pode-se visualizar no Gráfico 11 o mapa de calor das taxas de mortalidade por estado. De forma complementar, tem-se, ainda, a tabela completa com as taxas de mortalidade por covid-19 de todos os estados disponível no Apêndice A.

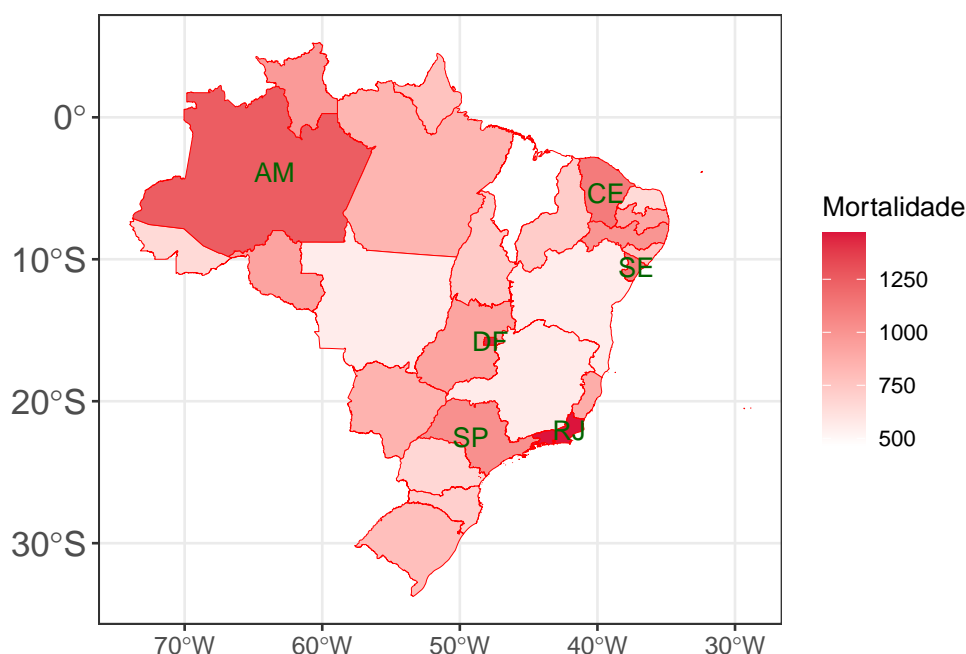


Gráfico 11: Mapa de calor das taxas de mortalidade por milhão de habitantes

No que se refere as capitais brasileiras, as taxas de mortalidade são ainda maiores que as estaduais. A média da taxa de mortalidade nas capitais é de 1.689 (um mil seiscentos e oitenta e nove) óbitos por milhão de habitantes, superior ao Estado do Rio de Janeiro, Estado com maior taxa de mortalidade do país (veja Tabela 2}).

Recife é a capital com maior número de mortes por milhão de habitantes, com 3.313 (três mil trezentos e treze) óbitos por covid-19, o que representa uma taxa 96% (noventa e seis por cento) superior à média nacional das capitais.

Das 7 (sete) maiores taxas de mortalidade por covid-19 nas capitais as regiões norte e nordeste ocupam 6 posições no ranking, com exceção da cidade do Rio de Janeiro. Já o município de São Paulo tem taxa de mortalidade inferior a média nacional e ocupa a 19<sup>a</sup> (décima nona) posição do ranking, com 1.436 mortes por milhão de habitantes, esse ranking consta na Tabela 3} abaixo.

Recife	Rio de Janeiro	Aracaju	Belém	Fortaleza	Teresina	João Pessoa
3313	2556	2472	2175	2126	1999	1913

Tabela 3: Capitais com maiores taxas de mortalidade

As taxas de mortalidade das capitais por covid-19 estão representadas em forma geométrica de triângulo invertido no Gráfico 12 (Taxa de mortalidade por covid-19 nas capitais), sendo o maior triângulo, a maior taxa de mortalidade, podendo-se, ainda, visualizar mais detalhes das taxas de mortalidade por covid-19 nas capitais no Apêndice A, deste trabalho.

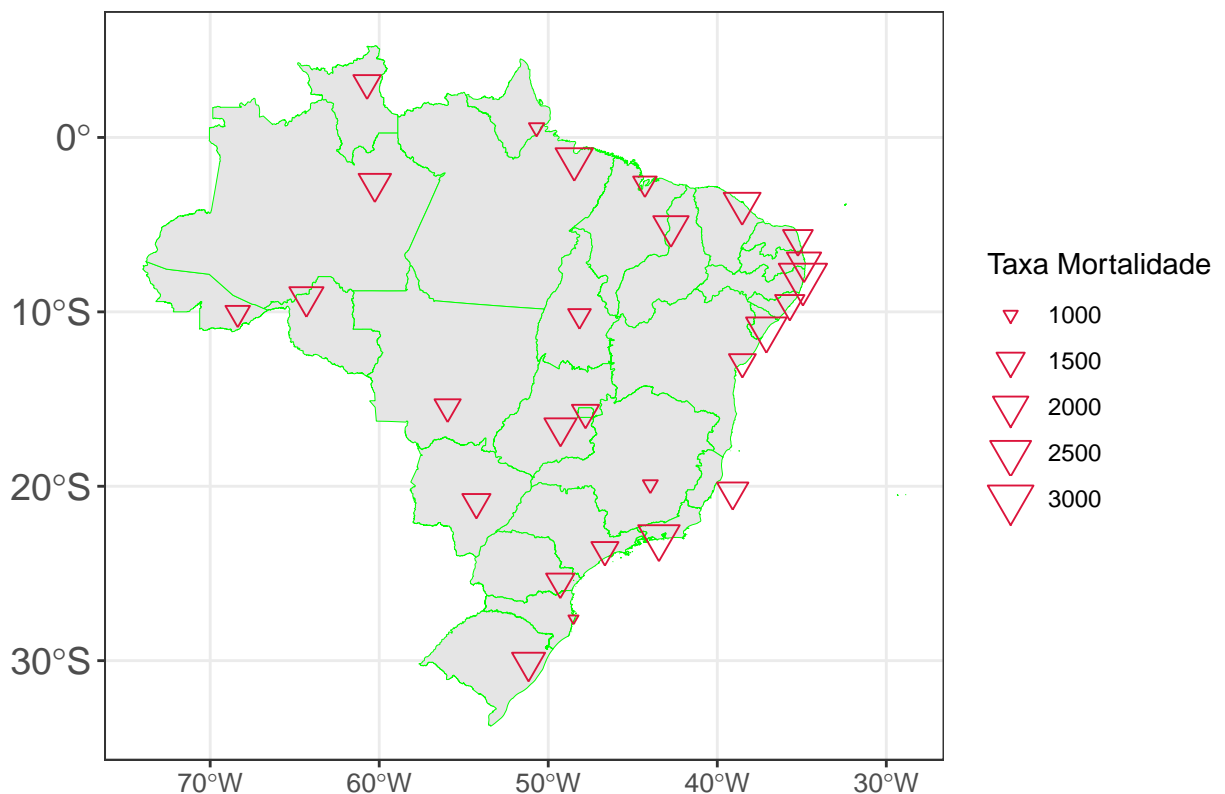


Gráfico 12: Taxa de mortalidade por covid-19 nas capitais

### 4.3 Taxa de letalidade da covid-19 no Brasil

A taxa de letalidade de covid-19 no Brasil, com verificação a partir de teste de hipótese, é de 30% (trinta por cento), com nível de significância de 95% (noventa e cinco por cento). Para este teste, utilizou-se uma amostra com 564,093 (quinhentos e sessenta e quatro mil e noventa e três) pacientes hospitalizados. Esta taxa é menor do que a taxa identificada no estudo realizado no Reino Unido, no qual comparou-se a taxa de letalidade da covid-19 com a do vírus ebola. A seguir, verifica-se o trecho de código para realização do teste de hipótese:

```
qtdObitos=dim(pacientesCovidTxLetalidade[pacientesCovidTxLetalidade$CS_EVOLUCAO=="Óbito",,])[1]
totalPacientes=dim(pacientesCovidTxLetalidade[,,])[1]
prop.test(qtdObitos, totalPacientes, 0.35, conf.level = 0.95)

##
## 1-sample proportions test with continuity correction
##
## data:  qtdObitos out of totalPacientes, null probability 0.35
## X-squared = 4698.4, df = 1, p-value < 0.00000000000000022
## alternative hypothesis: true p is not equal to 0.35
## 95 percent confidence interval:
##  0.3052663 0.3076743
## sample estimates:
##      p
## 0.306469
```



A Figura 5} mostra a manchete da matéria da BBC NEWS, que faz referência ao estudo Britânico citado acima.



Figura 5: Matéria BBC NEWS

De acordo com James Gallagher (2020), repórter de ciência e saúde[1]:

Pessoas internadas em hospitais por conta do novo coronavírus têm a mesma probabilidade de morrer do que aquelas hospitalizadas pelo vírus do ebola, segundo nova pesquisa conduzida no Reino Unido.

No que se refere à média móvel das taxas de letalidade por covid-19 do Brasil e regiões, Norte e Nordeste apresentaram, no início da pandemia, as maiores médias do país, ambas acima de 50% (cinquenta por cento), atingindo pico na 20<sup>a</sup> (vigésima) semana.

Nas semanas seguintes, a região Norte apresentou queda acentuada de sua taxa, chegando ao seu ponto mais baixo na 43<sup>a</sup> (quadragésima terceira) semana. Nesta ocasião, chegou a apresentar a menor média do país. Contudo, à partir da semana 44<sup>a</sup> (quadragésima quarta) semana voltou a subir, mantendo-se em ascensão até o término do ano de 2020, ocasião na qual ultrapassou a média das demais regiões.

Quanto ao Nordeste, este se manteve com média acima das apresentadas no Brasil e demais regiões, praticamente, durante todo ano de 2020, com exceção do mês de dezembro quando foi ultrapassada pela Região Norte.

É importante destacar que a Região Sul apresentou, durante quase todo ano de 2020, as menores médias do país. No entanto, no fim do ano, obteve média móvel superior a média nacional.

Estes resultados podem ser observados no Gráfico 13 (Média móvel semanal da taxa de letalidade), que está apresentado abaixo:

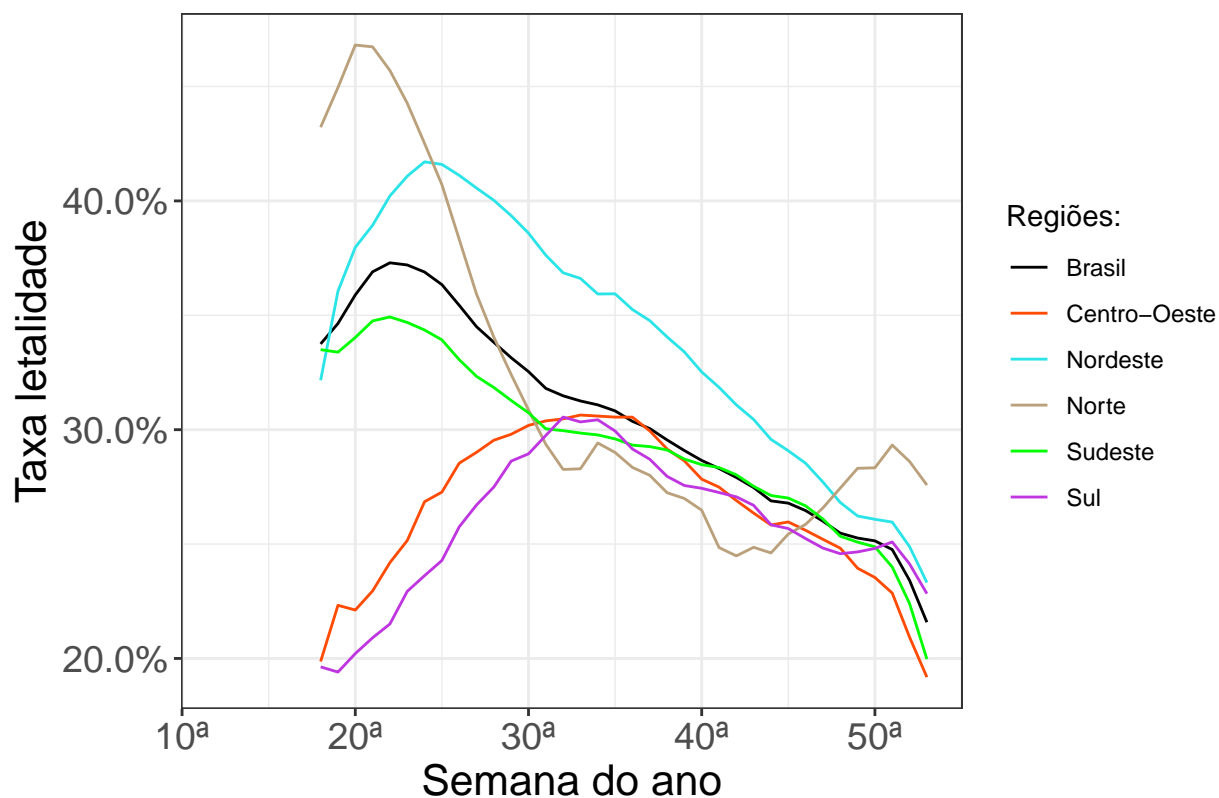


Gráfico 13: Média móvel semanal da taxa de letalidade

## 5. Inferência Estatística

No capítulo anterior, foram analisados os número absolutos de óbitos por covid-19, a média móvel da taxa de letalidade nacional por região, bem como as taxas de mortalidade de covid-19 nas regiões, estados e capitais brasileiras.

Neste capítulo, o foco será identificar o perfil do paciente hospitalizado com maior e menor risco de morte por covid-19, no Brasil.

### 5.1 Perfil do brasileiro com maior e menor risco de morte

Dos atributos disponíveis no data frame, foram selecionadas as seguintes colunas para construção dos perfis de pacientes hospitalizados com maior risco de morte: sexo, idade, existência de comorbidades, raça e região.

Na tentativa de reduzir os número de perfis e torná-los mais representativos para este modelo, foram adotados alguns tratamentos que serão detalhados nos próximos parágrafos.

No que se refere à idade, os valores foram agrupados em intervalos, adotando-se os mesmos intervalos utilizados na pirâmide etária adotada pelo Instituto Brasileiro de Geografia e Estatística (IBGE), conforme Tabela 4} abaixo:

0 a 4 anos	20 a 24 anos	40 a 44 anos	60 a 64 anos	80 anos ou mais
5 a 9 anos	25 a 29 anos	45 a 49 anos	65 a 69 anos	
10 a 14 anos	30 a 34 anos	50 a 54 anos	70 a 74 anos	
15 a 19 anos	35 a 39 anos	55 a 59 anos	75 a 79 anos	

Tabela 4: Faixa Etária

No que se refere à comorbidade, as colunas que detalham os tipos de comorbidades foram ignoradas para construção deste modelo, sendo adotada apenas a coluna fator de risco, que indica a existência de doença preexistente ou não.

As execuções necessárias para a otimização deste modelo foram realizadas em duas etapas:

A) Primeira execução:

Na primeira execução o modelo de regressão logística apresentou valores de baixa significância para:

- faixa etária: nos intervalos de 5 a 9 anos, 10 a 14 anos, bem como não exibiu a faixa de 0 a 4 anos;
- raça do paciente: na cor branca, bem como não exibiu a cor amarela.

O código utilizado na execução do modelo de regressão logística pode ser verificado no trecho de código abaixo, bem como o resultado da execução na Tabela 5}.

```
modeloRegressaoLogistica = glm(formula = IN_OBITO_COVID ~ CS_SEXO +
FAIXA_ETARIA + CS_RACA + IN_FATOR_RISCO + REGIAO,
data = dadosRegressaoLogistica, family = binomial(link = "logit"))
modeloRegressaoLogistica.table <- xtable(modeloRegressaoLogistica)
print(modeloRegressaoLogistica.table, type = "LaTeX" )
```

	$\beta^1$	Std. Error	z value	Pr(> z )
$\beta^0$	-3.1647	0.0743	-42.61	0.0000
Homem	0.2391	0.0071	33.49	0.0000
05 a 09 anos	-0.1248	0.1091	-1.14	0.2526
10 a 14 anos	0.1606	0.1144	1.40	0.1603
15 a 19 anos	0.4831	0.0934	5.17	0.0000
20 a 24 anos	0.4469	0.0802	5.57	0.0000
25 a 29 anos	0.4227	0.0745	5.68	0.0000
30 a 34 anos	0.5345	0.0706	7.57	0.0000
35 a 39 anos	0.6648	0.0686	9.69	0.0000
40 a 44 anos	0.8634	0.0677	12.75	0.0000
45 a 49 anos	1.0554	0.0672	15.70	0.0000
50 a 54 anos	1.2525	0.0668	18.76	0.0000
55 a 59 anos	1.5067	0.0665	22.67	0.0000
60 a 64 anos	1.8241	0.0663	27.52	0.0000
65 a 69 anos	2.0901	0.0662	31.57	0.0000
70 a 74 anos	2.3196	0.0662	35.03	0.0000
75 a 79 anos	2.5236	0.0663	38.04	0.0000
80 anos ou mais	2.9123	0.0660	44.13	0.0000
Branca	0.0187	0.0302	0.62	0.5363
Indígena	0.3564	0.0641	5.56	0.0000
Parda	0.1592	0.0302	5.28	0.0000
Preta	0.3490	0.0328	10.65	0.0000
Fator de Risco	0.5169	0.0082	62.95	0.0000
Nordeste	0.3166	0.0150	21.12	0.0000
Norte	0.3953	0.0173	22.87	0.0000
Sudeste	0.0907	0.0138	6.60	0.0000
Sul	-0.1248	0.0162	-7.69	0.0000

Tabela 5: Modelo com valores de baixa significância presentes

B) Segunda execução:

Após o ajuste do modelo e a reexecução em uma amostra com 383.266 (trezentos e oitenta e três mil e duzentos e sessenta e seis) pacientes. Conclui-se que o perfil com maior risco de morte tem: idade de 80 (oitenta) anos ou mais ( $\beta^1$  : 2.9888), já apresenta doenças

preexistentes ( $\beta^1$  : 0.5299), reside na região nordeste ( $\beta^1$  : 0.4171), é de cor preta ( $\beta^1$  : 0.3543) e homem ( $\beta^1$  : 0.2465), com probabilidade de óbito de 80%.

Já o perfil com menor risco de morte no Brasil: reside na região sul ( $\beta^1$  : -0.1689), tem idade de 0 (zero) a 14 (quatorze) anos, não tem doenças preexistentes, é de cor branca ou amarela e do sexo feminino, com probabilidade de óbito por covid-19 de 3,5%. Abaixo segue o trecho de código responsável pela segunda execução do modelo.

```
modeloRegressaoLogistica = glm(formula = IN_OBITO_COVID ~ CS_SEXO +
FAIXA_ETARIA + CS_RACA + IN_FATOR_RISC + REGIAO,
data = dadosRegressaoLogistica, family = binomial(link = "logit"))
modeloRegressaoLogistica.table <- xtable(modeloRegressaoLogistica)
print(modeloRegressaoLogistica.table, type = "LaTeX" )
```

Os valores de  $\beta^0$  e  $\beta^1$  resultado da reexecução do modelo ajustado podem ser verificados na Tabela: 6}.

	$\beta^1$	Std. Error	z value	Pr(> z )
$\beta^0$	-3.1459	0.0488	-64.53	0.0000
Homem	0.2390	0.0071	33.48	0.0000
15 a 19 anos	0.4834	0.0809	5.98	0.0000
20 a 24 anos	0.4472	0.0653	6.85	0.0000
25 a 29 anos	0.4230	0.0580	7.30	0.0000
30 a 34 anos	0.5348	0.0530	10.10	0.0000
35 a 39 anos	0.6651	0.0502	13.24	0.0000
40 a 44 anos	0.8637	0.0490	17.62	0.0000
45 a 49 anos	1.0557	0.0483	21.85	0.0000
50 a 54 anos	1.2528	0.0477	26.26	0.0000
55 a 59 anos	1.5069	0.0473	31.89	0.0000
60 a 64 anos	1.8243	0.0470	38.81	0.0000
65 a 69 anos	2.0904	0.0469	44.58	0.0000
70 a 74 anos	2.3199	0.0469	49.45	0.0000
75 a 79 anos	2.5238	0.0471	53.59	0.0000
80 anos ou mais	2.9126	0.0466	62.50	0.0000
Indígena	0.3382	0.0572	5.91	0.0000
Parda	0.1413	0.0086	16.48	0.0000
Preta	0.3310	0.0149	22.21	0.0000
Fator de Risco	0.5170	0.0082	62.96	0.0000
Nordeste	0.3165	0.0150	21.11	0.0000
Norte	0.3953	0.0173	22.86	0.0000
Sudeste	0.0910	0.0137	6.62	0.0000
Sul	-0.1241	0.0162	-7.67	0.0000

Tabela 6: Modelo de regressão logística otimizado

## 6. Machine learning

Neste capítulo, utiliza-se técnicas de machine learning para identificar outros atributos com impacto no risco de óbito dos pacientes, além dos atributos já descritos nas seções anteriores novos atributos serão incorporados aos modelos. Os algoritmos, também serão diversificados.

A eficiência dos atributos selecionados, bem como o algoritmo utilizado, será mensurada a partir da taxa de acerto do algoritmo.

Os dados utilizados nos treinamentos dos modelos, bem como os dados utilizados nos testes, serão extraídos a partir de uma amostra aleatória com 1000 (um mil) registros. Sendo 70% (setenta por cento) da amostra destinada para o treinamento e 30%(trinta por cento) destinada para os testes.

### 6.1 Árvore de Desisão de Classificação

Na primeira execução, o algoritmo utilizado será o rpart e os atributos serão os mesmos utilizados na seção de inferência estatística. Com essa configuração a taxa de acerto do modelo foi 69% (sessenta e nove por cento).

```
arvoreClassificacao = rpart(formula = CS_EVOLUCAO ~ CS_SEXO +  
FAIXA_ETARIA + CS_RACA + IN_FATOR_RISC + REGIAO,  
data = dadosMachineLearningObito, method = "class")
```

```
teste = predict(arvoreClassificacao, newdata = dadosTeste)  
previsao=cbind(teste, dadosTeste)  
previsao["PREVISAO"]=ifelse(previsao$CURA>.5,"CURA", "OBITO")
```

```
confusao=table(previsao$CS_EVOLUCAO, previsao$PREVISAO)  
taxaAcerto = (confusao[1]+confusao[4])/sum(confusao)  
taxaAcerto  
## [1] 0.6418863
```

### 6.2 Arvore de Decisão e Seleção de Atributos

Com método randomForest e mantendo praticamente os mesmo atributos da execução anterior, com exceção do atributo REGIAO, substituído pelo atributo SG\_UF a taxa de acerto referente ao risco de morte do paciente, apresenta uma pequena melhora. A após a mudança, a taxa de acerto passou para 70% (setenta por cento).

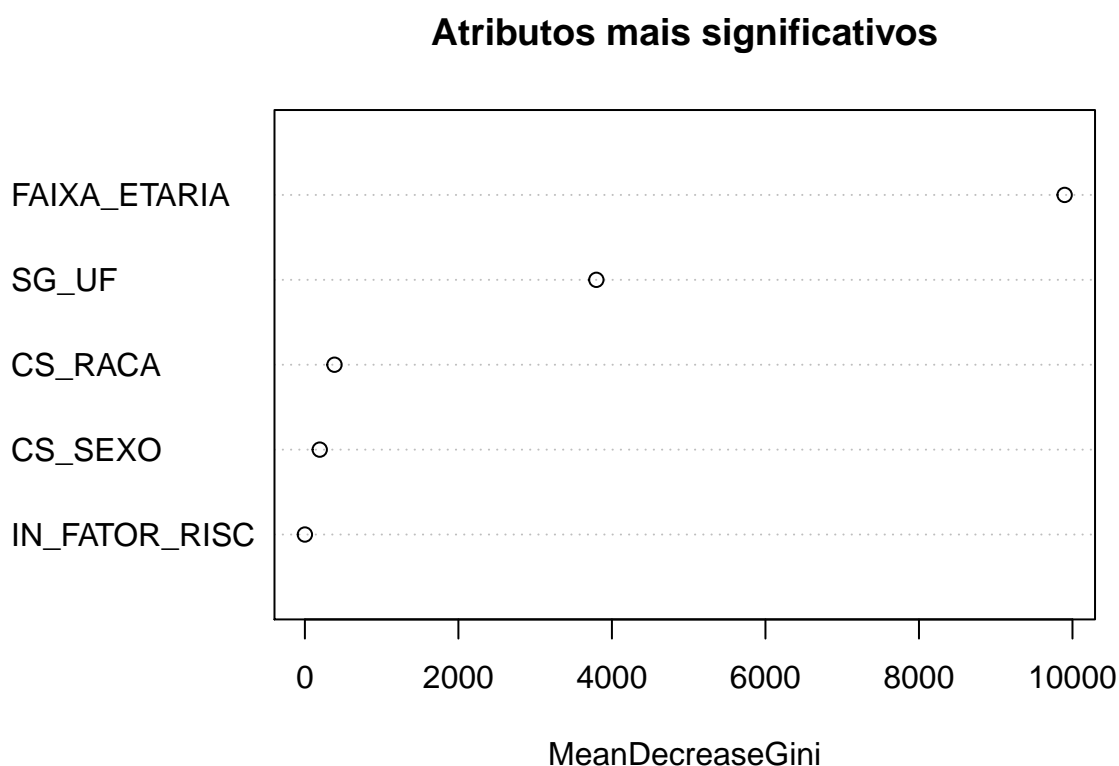
```
selecaoAtributos = randomForest(CS_EVOLUCAO ~ CS_SEXO +  
FAIXA_ETARIA + CS_RACA + IN_FATOR_RISC + SG_UF, data = dadosTreino)
```

```
teste = predict(selecaoAtributos, dadosTeste)  
previsao=cbind(teste , dadosTeste)  
previsao["PREVISAO"]=previsao$teste
```

```
confusao=table(previsao$CS_EVOLUCAO, previsao$PREVISAO)
taxaAcerto = (confusao[1]+confusao[4])/sum(confusao)
taxaAcerto
## [1] 0.6605315
```

O gráfico abaixo apresenta os atributos mais significativos na classificação de risco de morte do paciente. A FAIXA\_ETARIA do paciente é o atributo com maior peso na classificação do registro em relação ao risco de morte. Nas próximas execuções detalharemos mais os atributos FAIXA\_ETARIA e SG\_UF.

```
varImpPlot(selecaoAtributos, main="Atributos mais significativos")
```



### 6.3 Árvore de Decisão de Classificação otimizado

Utilizando novamente a função rpart agora com atributo NU\_IDADE e MUNICIPIO, a taxa de acerto passou a ser de 71% (setenta e um por cento)

```
arvoreClassificacao = rpart(formula = CS_EVOLUCAO ~ NU_IDADE + MUNICIPIO,
                             data = dadosMachineLearningObito, method = "class")
```

```
teste = predict(arvoreClassificacao, newdata = dadosTeste)
previsao=cbind(teste, dadosTeste)
previsao["PREVISAO"]=ifelse(previsao$CURA>.5,"CURA", "OBITO")
```

```
confusao=table(previsao$CS_EVOLUCAO, previsao$PREVISAO)
taxaAcerto = (confusao[1]+confusao[4])/sum(confusao)
taxaAcerto
```

```
## [1] 0.6782066
```

Quando adicionado o hospital de atendimento do paciente como atributo do modelo, a taxa de acerto passa a ser de 74% (setenta e quatro por cento).

```
arvoreClassificacao = rpart(formula = CS_EVOLUCAO ~ NU_IDADE + MUNICIPIO + HOSPITAL,  
                             data = dadosMachineLearningObito, method = "class")
```

```
teste = predict(arvoreClassificacao, newdata = dadosTeste)  
previsao=cbind(teste, dadosTeste)  
previsao["PREVISAO"]=ifelse(previsao$CURA>.5,"CURA", "OBITO")
```

```
confusao=table(previsao$CS_EVOLUCAO, previsao$PREVISAO)  
taxaAcerto = (confusao[1]+confusao[4])/sum(confusao)  
taxaAcerto
```

```
## [1] 0.7006119
```

## 6.4 Naive Bayse

Com os mesmos atributos da seção anterior e o método naiveBayes do pacote e1071 a taxa de acerto foi menor, 71% (setenta e um por cento)

```
modeloNaiveBayes = naiveBayes(formula = CS_EVOLUCAO ~ NU_IDADE + MUNICIPIO + HOSPITAL,  
                                data = dadosTreino)
```

```
teste = predict(modeloNaiveBayes, dadosTeste)  
previsao=cbind(teste , dadosTeste)  
previsao["PREVISAO"]=previsao$teste
```

```
confusao=table(previsao$CS_EVOLUCAO, previsao$PREVISAO)  
taxaAcerto = (confusao[1]+confusao[4])/sum(confusao)  
taxaAcerto
```

```
## [1] 0.6838349
```



# Referências Bibliográficas

## Referências

- [1] Gallagher, James. *Coronavírus é 'tão mortal quanto ebola' em casos de internação, diz estudo britânico*. DATA GEEKS, 29 abr. 2020. Disponível em: <https://www.bbc.com/portuguese/geral-52481818> - ( Acessado em: 14 jan. 2021 )
- [2] GOMES, Pedro César Tebaldi. *Classificação com Naive Bayes*. DATA GEEKS, 22 fev. 2019. Disponível em: <https://www.datageeks.com.br/naive-bayes/> - ( Acessado em: 14 jan. 2021 )
- [3] Hosokawa, Eric Ossamu. *Técnica de Árvore de Decisão em Mineração de Dados. 2011. Trabalho de Conclusão de Curso (Tecnólogo em Processamento de Dados)* - Faculdade de Tecnologia de São Paulo, São Paulo, 2011.
- [4] JUNIOR, Gilson Medeiros de Oliveira. *Máquina de Vetores Suporte: estudo e análise de parâmetros para otimização de resultado. 2010. Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação)* - Universidade Federal de Pernambuco, Recife, 2010.
- [5] MEHTA, R. *Big Data Analytics with Java*. Birmingham: Packt Publishing Ltd, 2017.
- [6] Portal OpenDataSUS, *SRAG 2021 - Banco de Dados de Síndrome Respiratória*. OpenDataSUS, 10 jan. 2021. Disponível em: <https://opendatasus.saude.gov.br/dataset/bd-srag-2021>
- [7] SANTOS, Roberto Moreira. *A importância do uso de ferramentas de analítica preditiva gratuitas para as empresas. 2017. Trabalho de Conclusão de Curso (Bacharelado em Sistemas de Informação)* - Universidade Federal de Pernambuco, Recife, 2017.

## Apêndice A - Taxa de Mortalidade Estados e Capitais

	UF	População	Mortalidade	Capital	População	Mortalidade
1	RJ	17.366.189	1.470	Rio de Janeiro	6.747.815	2.556
2	DF	3.055.149	1.439	Brasília	3.055.149	1.439
3	AM	4.2077.14	1.254	Manaus	2.219.580	1.774
4	CE	9.187.103	1.119	Fortaleza	2.686.612	2.126
5	SE	2.318.822	1.110	Aracaju	664.908	2.472
6	SP	46.289.333	1.020	São Paulo	12.325.232	1.436
7	PE	9.616.621	997	Recife	1.653.461	3.313
8	RR	631.181	969	Boa Vista	419.652	1.458
9	RO	1.796.460	927	Porto Velho	539.354	1.902
10	GO	7.113.540	924	Goiânia	1.536.097	1.764
11	PB	4.039.277	906	João Pessoa	817.511	1.913
12	ES	4.064.052	879	Vitória	365.855	1.697
13	MS	2.809.394	849	Campo Grande	906.092	1.498
14	PA	8.690.745	839	Belém	1.499.641	2175
15	RS	11.422.973	788	Porto Alegre	1.488.252	1.820
16	AP	861.773	770	Macapá	512.902	1.015
17	AL	3.351.543	751	Maceió	1.025.360	1.566
18	TO	1.590.248	728	Palmas	306.296	1.227
19	PI	3.281.480	723	Teresina	868.075	1.999
20	SC	7.252.502	706	Florianópolis	508.826	972
21	PR	11.516.840	666	Curitiba	1.948.626	1.508
22	RN	3.534.165	655	Natal	890.480	1.593
23	AC	894.470	655	Rio Branco	413.418	1.306
24	MG	21.292.666	565	Belo Horizonte	2.521.564	1.009
25	MT	3.526.220	560	Cuiabá	618.124	1.389
26	BA	14.930.634	557	Salvador	2.886.698	1.426
27	MA	7.5114.598	467	São Luís	1.108.975	1.264

Tabela 7: Taxa Mortalidade Estados e Capitais

# Anexo A - Ficha de Notificação



MINISTÉRIO DA SAÚDE  
SECRETARIA DE VIGILÂNCIA EM SAÚDE

SIVEP Gripe - SISTEMA DE INFORMAÇÃO DE VIGILÂNCIA EPIDEMIOLÓGICA DA GRIPE  
FICHA DE REGISTRO INDIVIDUAL - CASOS DE SÍNDROME RESPIRATÓRIA AGUDA GRAVE HOSPITALIZADO - 27/07/2020

Nº

**CASO DE SÍNDROME RESPIRATÓRIA AGUDA GRAVE (SRAG-HOSPITALIZADO):** Indivíduo com \*SG que apresente: dispneia/desconforto respiratório OU pressão persistente no tórax OU saturação de O<sub>2</sub> menor que 95% em ar ambiente OU coloração azulada dos lábios ou rosto. (\*SG: Indivíduo com quadro respiratório agudo, caracterizado por pelo menos dois (2) dos seguintes sinais e sintomas: febre (mesmo que referida), calafrios, dor de garganta, dor de cabeça, tosse, coriza, distúrbios olfativos ou gustativos).  
Para efeito de notificação no Sivep-Gripe, devem ser considerados os casos de SRAG hospitalizados ou os óbitos por SRAG independente de hospitalização.

1	Data do preenchimento da ficha de notificação:		2	Data de 1ºs sintomas		
3	UF:	4	Município:	Código (IBGE):		
5	Unidade de Saúde:			Código (CNES):		
Dados do Paciente	6	CPF do cidadão: _____				
	7	Nome: _____			8	Sexo: <input type="checkbox"/> 1-Masc. 2-Fem. 9-Ign
	9	Data de nascimento:	10	(Ou) Idade: _____	11	Gestante: <input type="checkbox"/>
				1-Dia 2-Mês 3-Ano _____		
	12	Raça/Cor: <input type="checkbox"/> 1-Branca 2-Preta 3-Amarela 4-Parda 5-Indígena 9-Ignorado			1-1º Trimestre 2-2º Trimestre 3-3º Trimestre	
	13	Se indígena, qual etnia? _____			4-Idade Gestacional Ignorada 5-Não se aplica 9-Ignorado	
Dados de residência	14	Escolaridade: <input type="checkbox"/> 0-Sem escolaridade/Analfabeto 1-Fundamental 1º cido (1ª a 5ª série) 2-Fundamental 2º cido (6ª a 9ª série) 3-Médio (1º ao 3º ano) 4-Superior 5-Não se aplica 9-Ignorado				
	15	Ocupação: _____	16	Nome da mãe: _____		
	17	CEP: _____ - _____				
	18	UF: _____	19	Município: _____	Código (IBGE): _____	
	20	Bairro: _____	21	Logradouro (Rua, Avenida, etc.): _____	22	Nº: _____
	23	Complemento (apto, casa, etc.): _____		24	(DDD) Telefone: _____	
Dados Clínicos e Epidemiológicos	25	Zona: <input type="checkbox"/> 1-Urbana 2-Rural 3-Periurbana 9-Ignorado		26	País: (se residente fora do Brasil) _____	
	27	Paciente tem histórico de viagem internacional até 14 dias antes do início dos sintomas? <input type="checkbox"/> 1-Sim 2-Não 9-Ign				
	28	Se sim: Qual país? _____		29	Em qual local? _____	
	30	Data da viagem: _____		31	Data do retorno: _____	
	32	É caso proveniente de surto de SG que evoluiu para SRAG? <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado				
	33	Trata-se de caso nosocomial (infecção adquirida no hospital)? <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado				
	34	Paciente trabalha ou tem contato direto com aves, suínos, ou outro animal? <input type="checkbox"/> 1-Sim 2-Não 3-Outro, qual _____ 9-Ignorado				
	35	Sinais e Sintomas: 1-Sim 2-Não 9-Ignorado <input type="checkbox"/> Febre <input type="checkbox"/> Tosse <input type="checkbox"/> Dor de Garganta <input type="checkbox"/> Dispneia <input type="checkbox"/> Desconforto Respiratório <input type="checkbox"/> Saturação O <sub>2</sub> <95% <input type="checkbox"/> Diarreia <input type="checkbox"/> Vômito <input type="checkbox"/> Dor abdominal <input type="checkbox"/> Fadiga <input type="checkbox"/> Perda do olfato <input type="checkbox"/> Perda do paladar <input type="checkbox"/> Outros _____				
	36	Possui fatores de risco/comorbidades? <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado Se sim, qual(is)? (Marcar X) <input type="checkbox"/> Puérpera (até 45 dias do parto) <input type="checkbox"/> Doença Cardiovascular Crônica <input type="checkbox"/> Doença Hematológica Crônica <input type="checkbox"/> Síndrome de Down <input type="checkbox"/> Doença Hepática Crônica <input type="checkbox"/> Asma <input type="checkbox"/> Diabetes mellitus <input type="checkbox"/> Doença Neurológica Crônica <input type="checkbox"/> Outra Pneumopatia Crônica <input type="checkbox"/> Imunodeficiência/Imunodepressão <input type="checkbox"/> Doença Renal Crônica <input type="checkbox"/> Obesidade, IMC _____ <input type="checkbox"/> Outros _____				
	37	Recebeu vacina contra Gripe na última campanha? <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado			38	Data da vacinação: _____
	Se < 6 meses: a mãe recebeu a vacina? <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado Se sim, data: _____ a mãe amamenta a criança? <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado					
	Se >= 6 meses e <= 8 anos: Data da dose única 1/1: _____ (dose única para crianças vacinadas em campanhas de anos anteriores) Data da 1ª dose: _____ (1ª dose para crianças vacinadas pela primeira vez) Data da 2ª dose: _____ (2ª dose para crianças vacinadas pela primeira vez)					

Dados de Atendimento	39	Usou antiviral para gripe? <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado	40	Qual antiviral? <input type="checkbox"/> 1-Oseltamivir 2-Zanamivir 3-Outro, especifique: _____	41	Data início do tratamento _____	
	42	Houve internação? <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado	43	Data da internação por SRAG: _____	44	UF de internação: _____	
	45	Município de internação: _____	Código (IBGE): _____				
	46	Unidade de Saúde de internação: _____	Código (CNES): _____				
	47	Internado em UTI? <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado	48	Data da entrada na UTI: _____	49	Data da saída da UTI: _____	
	50	Uso de suporte ventilatório: <input type="checkbox"/> 1-Sim, invasivo 2-Sim, não invasivo 3-Não 9-Ignorado	51	Raio X de Tórax: <input type="checkbox"/> 1-Normal 2-Infiltrado intersticial 3-Consolidação 4-Misto 5-Outro: _____ 6-Não realizado 9-Ignorado	52	Data do Raio X: _____	
	53	Aspecto Tomografia: <input type="checkbox"/> 1-Típico COVID-19 2-Indeterminado COVID-19 3- Atípico COVID-19 4-Negativo para Pneumonia 5-Outro 6-Não realizado 9-Ignorado	54	Data da tomografia: _____			
	55	Coletou amostra <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado	56	Data da coleta: _____	57	Tipo de amostra: <input type="checkbox"/> 1-Secreção de Naso-orofaringe 2-Lavado Bronco-alveolar 3-Tecido <i>post-mortem</i> 4-Outra, qual? _____ 5-LCR 9-Ignorado	
Dados Laboratoriais	58	Nº Requisição do GAL: _____		59	Tipo do teste para pesquisa de antígenos virais: <input type="checkbox"/> 1-Imunofluorescência (IF) 2- Teste rápido antígeno		
	60	Data do resultado da pesquisa de antígenos: _____		61	Resultado da Teste antígeno: <input type="checkbox"/> 1-positivo 2-Negativo 3- Inconclusivo 4-Não realizado 5-Aguardando resultado 9-Ignorado		
	62	Laboratório que realizou o Teste antígeno: _____ Código (CNES): _____					
	63	Agente Etiológico - Teste antígeno: Positivo para Influenza? <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado Se sim, qual influenza? <input type="checkbox"/> 1-Influenza A 2-Influenza B Positivo para outros vírus? <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado Se outros vírus respiratórios qual(is)? (marcar X) <input type="checkbox"/> SARS-CoV-2 <input type="checkbox"/> Vírus Sincial Respiratório <input type="checkbox"/> Parainfluenza 1 <input type="checkbox"/> Parainfluenza 2 <input type="checkbox"/> Parainfluenza 3 <input type="checkbox"/> Adenovírus <input type="checkbox"/> Outro vírus respiratório, especifique: _____					
	64	Resultado da RT- PCR/outra método por Biologia Molecular: <input type="checkbox"/> 1-Detectável 2-Não Detectável 3-Inconclusivo 4-Não realizado 5-Aguardando resultado 9-Ignorado		65	Data do resultado RT-PCR/outra método por Biologia Molecular: _____		
	66	Agente Etiológico - RT-PCR/outra método por Biologia Molecular: Positivo para Influenza? <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado Se sim, qual influenza? <input type="checkbox"/> 1-Influenza A 2-Influenza B Influenza A, qual subtipo? <input type="checkbox"/> 1-Influenza A(H1N1)pdm09 2-Influenza A/H3N2 3-Influenza A não subtipado 4-Influenza A não subtipável 5-Inconclusivo 6-Outro, especifique: _____ Influenza B, qual linhagem? <input type="checkbox"/> 1-Victoria 2-Yamagata 3-Não realizado 4-Inconclusivo 5-Outro, especifique: _____ Positivo para outros vírus? <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado Se outros vírus respiratórios, qual(is)? (marcar X) <input type="checkbox"/> SARS-CoV-2 <input type="checkbox"/> Vírus Sincial Respiratório <input type="checkbox"/> Parainfluenza 1 <input type="checkbox"/> Parainfluenza 2 <input type="checkbox"/> Parainfluenza 3 <input type="checkbox"/> Parainfluenza 4 <input type="checkbox"/> Adenovírus <input type="checkbox"/> Metapneumovírus <input type="checkbox"/> Bocavírus <input type="checkbox"/> Rinovírus <input type="checkbox"/> Outro vírus respiratório, especifique: _____					
	67	Laboratório que realizou RT-PCR/outra método por Biologia Molecular: _____ Código (CNES): _____					
	68	Tipo de amostra sorológica para SARS-Cov-2: <input type="checkbox"/> 1- Sangue/plasma/soro 2-Outra, qual? _____ 9-Ignorado			69	Data da coleta: _____	
	70	Tipo de Sorologia para SARS-Cov-2: <input type="checkbox"/> 1-Teste rápido 2-Elisa 3- Quimiluminescência 4- Outro, qual? _____ Resultado do Teste Sorológico para SARS-CoV-2: <input type="checkbox"/> IgG <input type="checkbox"/> IgM <input type="checkbox"/> IgA 1-Positivo 2-Negativo 3- Inconclusivo 4-Não realizado 5-Aguarda resultado 9 Ignorado			71	Data do resultado: _____	
	Conclusão	72	Classificação final do caso: <input type="checkbox"/> 1-SRAG por influenza 2-SRAG por outro vírus respiratório 3-SRAG por outro agente etiológico, qual _____ 4-SRAG não especificado 5-SRAG por COVID-19			73	Critério de Encerramento: <input type="checkbox"/> 1- Laboratorial 2- Clínico Epidemiológico 3- Clínico 4- Clínico-Imagem
74		Evolução do Caso: <input type="checkbox"/> 1-Cura 2-Óbito 3-Óbito por outras Causas 9-Ignorado		75	Data da alta ou óbito: _____		76
77 Número D.O: _____ - _____							
78 OBSERVAÇÕES: _____							
79 Profissional de Saúde Responsável: _____					80 Registro Conselho/Matrícula: _____		

# Anexo B - Dicionário de dados

## Dicionário de Dados

### FICHA DE REGISTRO INDIVIDUAL – CASOS DE SÍNDROME RESPIRATÓRIA AGUDA GRAVE HOSPITALIZADOS

Este documento tem como finalidade descrever as variáveis exportadas para o banco de dados em DBF.

CAMPO OBRIGATÓRIO é aquele cuja ausência de dado impossibilita a inclusão do registro no sistema.  
CAMPO ESSENCIAL é aquele que, apesar de não ser obrigatório, registra dado necessário à investigação do caso ou ao cálculo de indicador epidemiológico ou operacional.  
CAMPO INTERNO é aquele que apesar de não constar na ficha e não aparecer no display da tela, é preenchido automaticamente pelo sistema.  
CAMPO OPCIONAL é aquele que só deve ser preenchido caso seja necessário, aparece no display da tela e consta no banco de dados.

Nome do campo	Tipo	Categoria	Descrição	Características	DBF
Nº	Varchar2(12)		Número do registro	<b>Campo Interno</b> Número sequencial gerado automaticamente pelo sistema. Utilizar o padrão: 320120000123 Dígito 1: caracteriza o tipo da ficha (1=SG, 2=SRAG-UTI e 3-SRAG Hospitalizado). Dígitos 2 a 12: número sequencial gerado automaticamente pelo sistema.	NU_NOTIFIC
1-Data do preenchimento da ficha de notificação	Data DD/MM/AAAA		Data de preenchimento da ficha de notificação.	<b>Campo Obrigatório</b> Data deve ser <= a data da digitação.	DT_NOTIFIC
Semana Epidemiológica do preenchimento da ficha de notificação	Varchar2(6)		Semana Epidemiológica do preenchimento da ficha de	<b>Campo Interno</b> Calculado a partir da data dos Primeiros Sintomas. (SS)	SEM_NOT

				notificação.					
<b>2-Data de 1ºs sintomas</b>	DD/MM/AAAA	DD/MM/AAAA		Data de 1ºs sintomas do caso.	<b>Campo Obrigatório</b> Data deve ser <= a data da digitação e data do preenchimento da ficha de notificação	<b>DT_SIN_PRI</b>			
<b>Semana Epidemiológica dos Primeiros Sintomas</b>	Varchar2(6)			Semana Epidemiológica do início dos sintomas.	<b>Campo Interno</b> Calculado a partir da data dos Primeiros Sintomas. (SS)	<b>SEM_PRI</b>			
<b>3-UF</b>	Varchar2(2)	Tabela com código e siglas das UF padronizados pelo IBGE.		Unidade Federativa onde está localizada a Unidade Sentinela que realizou a notificação.	<b>Campo Obrigatório</b> Se usuário que está digitando a ficha for de nível: <ul style="list-style-type: none"> <li>Unidade Sentinela - o campo é preenchido automaticamente pelo sistema com a UF, município e unidade onde está cadastrado o usuário.</li> <li>Municipal - o campo é preenchido automaticamente pelo sistema com a UF e município onde está cadastrado o usuário.</li> <li>Estadual - o campo é preenchido automaticamente pelo sistema com a UF do usuário.</li> <li>Federal - abre tabela com todas as UF que possuem unidades sentinelas cadastradas no sistema.</li> </ul>	<b>SG_UF_NOT</b>			
<b>4-Município Código (IBGE)</b>	Varchar2(6)	Tabela com código e nomes dos Municípios padronizados pelo IBGE.		Município onde está localizada a Unidade Sentinela que realizou a notificação.	<b>Campo Obrigatório</b> Preenchendo o nome do município de notificação, o código é preenchido automaticamente, e vice-versa; Se usuário que está digitando a ficha for de nível: <ul style="list-style-type: none"> <li>Unidade Sentinela - o campo é preenchido automaticamente pelo sistema com o Município onde está localizada a unidade de notificação.</li> <li>Municipal - o campo é preenchido automaticamente pelo sistema com o município do usuário.</li> <li>Estadual ou Federal - abre tabela com todos os municípios da UF selecionada no campo 3 que possuem unidades sentinelas cadastradas no sistema.</li> </ul>	<b>ID_MUNICIP OU CO_MUN_NOT</b>			
<b>Regional de Saúde de Notificação Código (IBGE)</b>	Varchar2(6)	Tabela com código e nomes das Regionais de Saúde dos municípios de notificação padronizados pelo IBGE.		Regional de Saúde onde está localizado o Município realizou a notificação.	<b>Campo Interno</b> Preenchendo o nome da regional de saúde de notificação, o código é preenchido automaticamente, e vice-versa; Se usuário que está digitando a ficha for de nível:	<b>ID_REGIONA OU CO_REGIONA</b>			

					<ul style="list-style-type: none"> <li>▪ <u>Unidade Sentinela</u> – o campo é preenchido automaticamente pelo sistema com a Regional do Município onde está localizada a unidade de notificação.</li> <li>▪ <u>Municipal</u> – o campo é preenchido automaticamente pelo sistema com a regional do município do usuário.</li> </ul>	
<b>5-Unidade de Saúde Código (CNES)</b>		Varchar2(7)	Tabela com códigos CNES e nomes das Unidades Sentinelas cadastradas no sistema.	Unidade Sentinela que realizou o atendimento, coleta de amostra e registro do caso.	<b>Campo Obrigatório</b> Preenchendo o nome da unidade, o código é preenchido automaticamente, e vice-versa;  Se usuário que está digitando a ficha for de nível: <ul style="list-style-type: none"> <li>▪ <u>Unidade Sentinela</u> - o campo é preenchido automaticamente pelo sistema.</li> <li>▪ <u>Municipal</u> – abre tabela apenas com as unidades sentinelas do município.</li> <li>▪ <u>Estadual ou Federal</u> – abre tabela com as unidades sentinelas do município selecionado o campo 4.</li> </ul>	ID_UNIDADE OU CO_UNI_NOT
<b>6-CPF do paciente</b>		Varchar2(15)		Número do Cadastro de Pessoa Física (CPF) do paciente notificado	<b>Campo Essencial</b> Quando preenchido o número do CPF o sistema deverá preencher o Nome, Sexo, Data de Nascimento, Idade, Raça/Cor e o Nome da mãe do paciente.	NU_CPF
<b>7-Nome</b>		Varchar2(70)		Nome completo do paciente (sem abreviações)	<b>Campo Obrigatório</b>	NM_PACIENT
<b>8-Sexo</b>		Varchar2(1)	1-Masculino 2-Feminino 9-Ignorado	Sexo do paciente.	<b>Campo Obrigatório</b>	CS_SEXO
<b>9-Data de nascimento</b>		Date DD/MM/AAAA		Data de nascimento do paciente.	<b>Campo Essencial</b> Data deve ser <= a data dos primeiros sintomas.	DT_NASC
<b>10-(ou) Idade</b>		Varchar2(3)		Idade informada pelo paciente quando não se sabe a data de nascimento.	<b>Campo Obrigatório</b> Se digitado a data de nascimento, a idade é calculada e preenchida automaticamente pelo sistema: considerando o intervalo entre a data de nascimento e a <u>data dos primeiros sintomas</u> .  Idade deve ser <= 150.	NU_IDADE_N

(ou) Tipo/Idade	Varchar2(1)	1-Dia 2-Mês 3-Ano	aparente.	<p><b>Campo Obrigatório</b></p> <p>Se digitado a data de nascimento, o campo Idade/Tipo é calculado e preenchido automaticamente pelo sistema: considerando o intervalo entre a data de nascimento e a <u>data dos primeiros sintomas</u>.</p> <p>Se a diferença for de 0 a 30 dias, o sistema grava em Idade = (nº dias) e em Tipo = 1-Dia. Por exemplo: se Data de nascimento = 05/12/2012 e Data dos 1ºs sintomas = 11/12/2012, então Idade = 6 e Tipo = 1-Dia.</p> <p>Se a diferença for de 1 a 11 meses, o sistema grava em Idade = (nº meses) e em Tipo = 2-Mês. Por exemplo: se Data de nascimento = 05/10/2012 e Data dos 1ºs sintomas = 11/12/2012, então Idade = 2 e Tipo = 2-Mês.</p> <p>Se a diferença for maior ou igual a 12 meses, o sistema grava em Idade = (nº anos) e em Tipo = 3-Ano. Por exemplo: se Data de nascimento = 05/10/2011 e Data dos 1ºs sintomas = 11/12/2012, então Idade = 1 e Tipo = 3-Ano.</p>	TP_IDADE
11-Gestante	Varchar2(1)	1-1º Trimestre 2-2º Trimestre 3-3º Trimestre 4-Idade Gestacional Ignorada 5-Não 6-Não se aplica 9-Ignorado	Idade gestacional da paciente.	<p><b>Campo Obrigatório</b></p> <p>Se selecionado categoria 2-Feminino no campo Sexo.</p> <p>Se selecionado sexo igual a <u>Masculino</u> ou a <u>idade for menor ou igual a 9 anos</u> o campo é preenchido automaticamente com 6-Não se aplica.</p> <p>Se selecionado sexo igual a <u>Feminino</u> e idade for maior que 9 anos, o campo não pode ser preenchido com 6-Não se aplica.</p>	CS_GESTANT
12-Raça/Cor	Varchar2(2)	1-Branca 2-Preta 3-Amarela 4-Parda 5-Indígena 9-Ignorado	Cor ou raça declarada pelo paciente: Branca; Preta; Amarela; Parda (pessoa que se declarou mulata, cabocla, cafuz, mameluca ou mestiça de preto com pessoa de outra cor ou raça); e, Indígena.	<p><b>Campo Obrigatório</b></p>	CS_RACA
13-Se indígena, qual etnia?	Varchar2(4)	Tabulado SIASI com código e nomes das etnias indígenas.	Nome e código da etnia do paciente,	<p><b>Campo Essencial</b></p>	CS_ETINIA



				quando indígena.	Ha bilitados e campo 11-Raça/Cor for igual a 5-Indígena.	
14-Escolaridade	Varchar2(1)	0-Sem escolaridade/ Analífabeto 1-Fundamental 1º ciclo (1ª a 5ª série) 2-Fundamental 2º ciclo (6ª a 9ª série) 3-Médio (1ª a o3ª ano) 4-Superior 5-Não se aplica 9-Ignorado	Nível de escolaridade do paciente.  Para os níveis fundamental e médio deve ser considerada a última série ou ano concluído.	<b>Campo Essencial</b>  Preenchido a utomaticamente com a categoria “não se aplica” quando idade for menor que 7 anos  Quando idade for maior que 7 anos, o campo não pode ser preenchido com “não se aplica”.	<b>CS_ESCOL_N</b>	
15- Ocupação	Varchar2(6)	Ta bela com código da Ocupação da Classificação Brasileira de Ocupações (CBO).	Ocupação profissional do paciente	<b>Campo Essencial</b>	<b>PAC_COCBO ou PAC_DSCBO</b>	
16- Nome da mãe	Varchar2(70)		Nome completo da mãe do paciente (sem abreviações).	<b>Campo Essencial</b>	<b>NM_MAE_PAC</b>	
17- CEP	Varchar2(8)		CEP de residência do paciente.	<b>Campo Essencial</b>  Validado a partir da tabela de CEP dos Correios.	<b>NU_CEP</b>	
18- UF	Varchar2(2)	Ta bela com código e siglas das UF padronizados pelo IBGE.	Unidade Federativa de residência do paciente.	<b>Campo Obrigatório</b>  Se campo 25-País for Brasil.  Se preenchido o campo CEP, a UF é preenchida automaticamente pelo sistema e desabilitada para edição.	<b>SG_UF</b>	
Regional de Saúde de Residência Código (IBGE)	Varchar2 (6)	Ta bela com código e nomes das Regionais de Saúde dos municípios de residência padronizados pelo IBGE.	Regional de Saúde onde está localizado o Município de residência do paciente.	<b>Campo Interno</b>  Preenchendo o nome da regional de saúde de residência, o código é preenchido automaticamente, e vice-versa;	<b>ID_RG_RESIOU CO_RG_RESI</b>	
19- Município Código (IBGE)	Varchar2(6)	Ta bela com código e nome dos Municípios padronizados pelo IBGE.	Município de residência do paciente.	<b>Campo Obrigatório</b>  Se campo 25-País for Brasil.  Se preenchido o campo CEP, o Município e seu respectivo código IBGE são preenchidos automaticamente pelo sistema e desabilitados para edição.	<b>ID_MN_RESIOU CO_MUN_RES</b>	

					Se o CEP não for preenchido, o campo é habilitado depois de selecionada uma UF no campo 17. Nesse caso, o sistema abre a tela com os municípios da UF.  Preenchendo o nome do município, o código é preenchido automaticamente, ou vice-versa.	
<b>20-Bairro</b>	Varchar2(72)	Tabela com código e nome dos Bairros padronizados pelos Correios.	Bairro de residência do paciente.	<b>Campo Essencial</b>  Se preenchido o campo CEP, o Bairro é preenchido automaticamente pelo sistema.	<b>NM_BAIRRO</b>	
<b>21-Logradouro (Rua, Avenida, etc.)</b>	Varchar2(50)	Tabela com código e nome dos logradouros padronizados pelos Correios.	Logradouro (rua, avenida, quadra, travessa, etc.) do endereço de residência do paciente.	<b>Campo Essencial</b>  Se preenchido o campo CEP, o logradouro é preenchido automaticamente pelo sistema.	<b>NM_LOGRADO</b>	
<b>22-Nº</b>	Varchar2(8)		Nº do logradouro (nº da casa ou do edifício).	<b>Campo Essencial</b>	<b>NU_NUMERO</b>	
<b>23-Complemento (apto, casa, etc.)</b>	Varchar2(15)		Complemento do logradouro (bloco, apto., casa, etc.).	<b>Campo Essencial</b>	<b>NM_COMPLEM</b>	
<b>24-(DDD) Telefone</b>	Varchar2(4) Varchar2(10)		Código DDD e número de telefone para contato do paciente.	<b>Campo Essencial</b>	<b>NU_DDD_TEL OU NU_TELEFON</b>	
<b>25-Zona</b>	Varchar2(1)	1-Urbana 2-Rural 3-Periurbana 9-Ignorado	Zona geográfica do endereço de residência do paciente.	<b>Campo Essencial</b>	<b>CS_ZONA</b>	
<b>26-País (se residente fora do Brasil)</b>	Varchar2(3)	Tabela com código e nome dos Países.	País de residência do paciente.	<b>Campo Obrigatório</b>  Se preenchido CEP, ou for selecionada uma UF, o campo País é preenchido automaticamente pelo sistema e desabilitado para edição.  Se selecionado País diferente de Brasil, os campos 17 a 23 são desabilitados.	<b>ID_PAIS OU CO_PAIS</b>	
<b>27- Paciente tem histórico de viagem internacional até 14 dias antes do</b>	Varchar2(1)	1-Sim 2-Não		<b>Campo Obrigatório</b>	<b>HISTO_VGM</b>	





					apresentados pelo paciente.	Ha bilitado se selecionado categoria 1-Sim em Sinais e Sintomas/Outros.	
36-Fatores de risco	Varchar2 (1)	1-Sim 2-Não 9-Ignorado			Paciente apresenta algum fator de risco	<b>Campo Essencial</b>	<b>FATOR_RISC</b>
36-Fatores de risco/ Puérpera	Varchar2 (1)	1-Sim 2-Não 9-Ignorado			Paciente é puérpera ou parturiente (mulher que pariu recentemente – até 45 dias do parto)?	<b>Campo Essencial</b> Ha bilitados e selecionado no campo 8- Sexo Feminino.	<b>PUERPERA</b>
36-Fatores de risco/ Doença Cardiovascular Crônica	Varchar2 (1)	1-Sim 2-Não 9-Ignorado			Paciente possui Doença Cardiovascular Crônica ?	<b>Campo Essencial</b>	<b>CARDIOPATI</b>
36-Fatores de risco/ Doença Hematológica Crônica	Varchar2 (1)	1-Sim 2-Não 9-Ignorado			Paciente possui Doença Hematológica Crônica ?	<b>Campo Essencial</b>	<b>HEMATOLOGI</b>
36-Fatores de risco/ Síndrome de Down	Varchar2 (1)	1-Sim 2-Não 9-Ignorado			Paciente possui Síndrome de Down?	<b>Campo Essencial</b>	<b>SIND_DOWN</b>
36-Fatores de risco/ Doença Hepática Crônica	Varchar2 (1)	1-Sim 2-Não 9-Ignorado			Paciente possui Doença Hepática Crônica ?	<b>Campo Essencial</b>	<b>HEPATICA</b>
36-Fatores de risco/ Asma	Varchar2 (1)	1-Sim 2-Não 9-Ignorado			Paciente possui Asma ?	<b>Campo Essencial</b>	<b>ASMA</b>
36-Fatores de risco/ Diabetes mellitus	Varchar2 (1)	1-Sim 2-Não 9-Ignorado			Paciente possui Diabetes <i>mellitus</i> ?	<b>Campo Essencial</b>	<b>DIABETES</b>
36-Fatores de risco/ Doença Neurológica Crônica	Varchar2 (1)	1-Sim 2-Não 9-Ignorado			Paciente possui Doença Neurológica?	<b>Campo Essencial</b>	<b>NEUROLOGIC</b>
36-Fatores de risco/ Outra Pneumopatia Crônica	Varchar2 (1)	1-Sim 2-Não 9-Ignorado			Paciente possui outra pneumopatia crônica?	<b>Campo Essencial</b>	<b>PNEUMOPATI</b>
36-Fatores de risco/ Imunodeficiência	Varchar2 (1)	1-Sim			Paciente possui	<b>Campo Essencial</b>	<b>IMUNODEPRE</b>

ou Imunodepressão		2-Não 9-Ignorado	Imunodeficiência ou Imunodepressão (diminuição da função do sistema imunológico)?			
36-Fatores de risco/ Doença Renal Crônica	Varcha r2 (1)	1-Sim 2-Não 9-Ignorado	Paciente possui Doença Renal Crônica?	<b>Campo Essencial</b>		RENAL
36-Fatores de risco/ Obesidade	Varcha r2 (1)	1-Sim 2-Não 9-Ignorado	Paciente possui obesidade?	<b>Campo Essencial</b>		OBESIDADE
36-Fatores de risco/ Obesidade (Descrição IMC)	Varcha r2(3)		Valor do IMC (Índice de Massa Corporal) do paciente calculado pelo profissional de saúde.	<b>Campo Essencial</b> Ha bilitado se selecionado categoria 1-Sim em Fatores de risco/Obesidade.		OBES_IMC
36-Fatores de risco/ Outros	Varcha r2(1)	1-Sim 2-Não 9-Ignorado	Paciente possui outro(s) fator(es) de risco?	<b>Campo Essencial</b>		OUT_MORBI
36-Fatores de risco/ Outros (Descrição)	Varcha r2(30)		Listar outro(s) fator(es) de risco do paciente.	<b>Campo Essencial</b> Ha bilitado se selecionado categoria 1-Sim em Fatores de risco/Outros.		MORB_DESC
37-Recebeu vacina contra Gripe na última campanha?	Varcha r2 (1)	1-Sim 2-Não 9-Ignorado	Informar se o paciente foi vacinado contra gripe na última campanha, após verificar a documentação / caderneta.  Caso o paciente não tenha a caderneta, direcionar a pergunta para ele ou responsável e preencher o campo com o código correspondente a	<b>Campo Essencial</b>		VACINA



			2-Não 9-Ignorado	para tratamento da doença?			
40-Qual antiviral?	Varchar2(1)		1- Oseltamivir 2- Zanamivir 3- Outro	Qual antiviral utilizado?	<b>Campo Essencial</b> Ha bilittado e campo 39- Qual antiviral? for igual a 1.	TP_ANTIVIR	
Qual antiviral /Outro, especifique	Varchar2(30)			Se o antiviral utilizado não foi Oseltamivir ou Zanamivir, informar qual antiviral foi utilizado.	<b>Campo Essencial</b> Ha bilittado e campo 40- Qual antiviral? for igual a 3.	OUT_ANTIV	
41-Data do início do tratamento	Date DD/MM/AAAA			Data em que foi iniciado o tratamento com o antiviral.	<b>Campo Essencial</b> Ha bilittado e campo 39-Usou antiviral para gripe? for igual a 1. Data deve ser <= a data da digitação (data atual).	DT_ANTIVIR	
42-Houve internação?	Varchar2(1)		1-Sim 2-Não 9-Ignorado	O paciente foi internado?	<b>Campo Essencial</b> Caso o campo não seja igual a 1 – Sim o sistema emitirá um aviso indicando que não atende a definição de caso.	HOSPITAL	
43-Data da internação por SRAG	Date DD/MM/AAAA			Data em que o paciente foi hospitalizado.	<b>Campo Obrigatório</b> Data deve ser maior ou igual a 2- Data de 1ºs sintomas e menor ou igual a data da digitação (atual).	DT_INTERNA	
44-UF de internação	Varchar2(2)		Tabela com código e siglas das UF padronizadas pelo IBGE.	Unidade Federativa de Internação do paciente.	<b>Campo Essencial</b> Ha bilittado e campo 42-Houve internação? for igual a 1	SG_UF_INTE	
Regional de Saúde de Internação Código (IBGE)	Varchar2(6)		Tabela com código e nomes das Regionais de Saúde dos municípios de internação padronizadas pelo IBGE.	Regional de Saúde onde está localizado o Município de internação do paciente.	<b>Campo Interno</b> Preenchendo o nome da regional de saúde de internação, o código é preenchido automaticamente, e vice-versa.	ID_RG_INTE OU CO_RG_INTE	
45-Município de internação/ Código (IBGE)	Varchar2(20)		Tabela com código e nomes dos Municípios padronizados pelo IBGE.	Município onde está localizado a Unidade de Saúde onde o paciente internou.	<b>Campo Essencial</b> Ha bilittado e campo 42-Houve internação? for igual a 1	ID_MN_INTE OU CO_MU_INTE	



<b>46- Unidade de Saúde de internação/ Código CNES</b>	Varchar(20)	Tabela com códigos CNES e nomes das Unidades de Saúde.	Unidade Sentinela que realizou a internação do paciente.	<b>Campo Essencial</b> Ha bilatado e campo <b>42-Houve internação?</b> for igual a 1	ID_UN_INTE OU CO_UN_INTE
<b>47- Internado em UTI?</b>	Varchar(1)	1-Sim 2-Não 9-Ignorado	O paciente foi internado em UTI?	<b>Campo Essencial</b>	UTI
<b>48- Data da entrada na UTI</b>	Data DD/MM/AAAA		Data de entrada do paciente na unidade de Terapia intensiva (UTI).	<b>Campo Essencial</b> Ha bilatado e campo <b>47-Internado em UTI?</b> for igual a 1. Data deve ser maior ou igual a <b>2-Data de 1ºs sintomas da SRAG</b> e menor ou igual a data da digitação (atual).	DT_ENT_UTI
<b>49- Data da saída da UTI</b>	Data DD/MM/AAAA		Data em que o paciente saiu da Unidade de Terapia intensiva (UTI).	<b>Campo Essencial</b> Ha bilatado e campo <b>47-Internado em UTI?</b> for igual a 1. Data deve ser maior ou igual a <b>48-Data da entrada na UTI</b> e menor ou igual a data da digitação (atual).	DT_SAID_UTI
<b>50- Uso de suporte ventilatório?</b>	Varchar(1)	1-Sim, invasivo 2-Sim, não invasivo 3-Não 9-Ignorado	O paciente fez uso de suporte ventilatório?	<b>Campo Essencial</b>	SUPORT_VEN
<b>51- Raio X de Tórax</b>	Varchar(1)	1-Normal 2-Infiltrado intersticial 3-Consolidação 4-Misto 5-Outro 6-Não realizado 9-Ignorado	Informar resultado de Raio X de Tórax.	<b>Campo Essencial</b>	RAIOX_RES
<b>Raio X de Tórax/ Outro (especificar)</b>	Varchar(30)		Informar o resultado do RX de tórax e selecionado a opção 5-Outro.	<b>Campo Essencial</b> Ha bilatado e campo <b>51- Raio X de Tórax</b> = 5 (Outro).	RAIOX_OUT
<b>52- Data do Raio X</b>	Data DD/MM/AAAA		Se realizou RX de Tórax, especificar a data do exame.	<b>Campo Essencial</b> Ha bilatado e campo <b>51- Raio X de Tórax</b> for igual a 1, 2, 3, 4 ou 5.	DT_RAIOX
<b>53- Aspecto Tomografia</b>	Number(3)	1-Típico COVID-19 2-Indeterminado COVID-19	Informar o resultado da	<b>Campo Essencial</b>	TOMO_RES

		3- Atípico COVID-19 4- Negativo para Pneumonia 5- Outro 6- Não realizado 9- Ignorado	tomografia.		
Aspecto Tomografia/Outro (especificar)	Varchar2(100)		Informar o resultado da tomografia selecionada opção 5-Outro	<b>Campo Essencial</b> Ha bilhettato de campo 53- Aspecto Tomografia = 5 (Outro)	TOMO_OUT
54- Data da Tomografia	Data DD/MM/AAAA		Se realizou tomografia, especificar a data do exame.	<b>Campo Essencial</b> Ha bilhettato de campo 53- Aspecto Tomografia for igual a 1, 2, 3, 4 ou 5.	DT_TOMO
55- Coletou amostra?	Varchar2(1)	1- Sim 2- Não 9- Ignorado	Foi realizado coleta de amostra para realização de teste diagnóstico?	<b>Campo Essencial</b>	AMOSTRA
56- Data da Coleta	Data DD/MM/AAAA		Data da coleta da amostra para realização do teste diagnóstico.	<b>Campo Essencial</b> Ha bilhettato de campo 55- Coletou amostra? = 1. Data deve ser maior ou igual a 2- Data de 1ºs sintomas e menor ou igual a data da digitação (atual).	DT_COLETA
57- Tipo de amostra	Varchar2(30)	1- Secreção de Naso-oro-farínge 2- Lava do Broco-alveolar 3- Tecido <i>post-mortem</i> 4- Outra, qual? 5- LCR 9- Ignorado	Tipo da amostra clínica coletada para o teste diagnóstico.	<b>Campo Essencial</b> Ha bilhettato de campo 55- Coletou amostra? = 1.	TP_AMOSTRA
Tipo de amostra/Outra	Varchar2(30)		Descrição do tipo da amostra clínica, caso diferente das listadas nas categorias do campo.	<b>Campo Essencial</b> Campo habilitados e selecionado categoria 4-Outra, qual em Tipo de amostra.	OUT_AMOST
58- Nº da Requisição do GAL			Número da requisição de exames gerado	<b>Campo Essencial</b>	REQUIL_GAL



VSR			VSR.	Ha bilittados e campo 63-Agente etiológico, Teste Antigénico. Positivo para outros vírus? = 1.	
63-Agente etiológico – Teste Antigénico. Parainfluenza 1	Varchar2(1)	1-marca do pelo usuário Vazio - não marcado	Resultado do Teste Antigénico, para Parainfluenza 1.	<b>Campo Essencial</b> Ha bilittados e campo 63-Agente etiológico, Teste Antigénico. Positivo para outros vírus? = 1.	AN_PARA1
63-Agente etiológico – Teste Antigénico. Parainfluenza 2	Varchar2(1)	1-marca do pelo usuário Vazio - não marcado	Resultado do Teste Antigénico, para Parainfluenza 2.	<b>Campo Essencial</b> Ha bilittados e campo 63-Agente etiológico, Teste Antigénico. Positivo para outros vírus? = 1.	AN_PARA2
63-Agente etiológico – Teste Antigénico. Parainfluenza 3	Varchar2(1)	1-marca do pelo usuário Vazio - não marcado	Resultado do Teste Antigénico, para Parainfluenza 3.	<b>Campo Essencial</b> Ha bilittados e campo 63-Agente etiológico, Teste Antigénico. Positivo para outros vírus? = 1.	AN_PARA3
63-Agente etiológico – Teste Antigénico. Adenovirus	Varchar2(1)	1-marca do pelo usuário Vazio - não marcado	Resultado do Teste Antigénico, Adenovirus.	<b>Campo Essencial</b> Ha bilittados e campo 63-Agente etiológico, Teste Antigénico. Positivo para outros vírus? = 1.	AN_ADENO
63- Agente etiológico – Teste Antigénico. Outro virus respiratório	Varchar2(1)	1-marca do pelo usuário Vazio - não marcado	Resultado do Teste Antigénico, Outro virus respiratório.	<b>Campo Essencial</b> Ha bilittados e campo 63-Agente etiológico, Teste Antigénico. Positivo para outros vírus? = 1.	AN_OUTRO
63- Agente etiológico – Teste Antigénico. Outro virus respiratório (Descrição)	Varchar2(30)		Nome do outro virus respiratório identificado pelo Teste Antigénico.	<b>Campo Essencial</b> Ha bilittados e campo 63-Agente etiológico, Teste Antigénico. Positivo para outros vírus? = 1.	DS_AN_OUT
64-Resultado da RT-PCR/outro método por Biologia Molecular	Varchar2(1)	1-Detectável 2-Não Detectável 3-Inconclusivo 4-Não Realizado 5-Aguardando Resultado 9-Ignorado	Resultado do teste de RT-PCR/outro método por Biologia Molecular.	<b>Campo Essencial</b> Este campo virá marcado com 5-Aguardando Resultado e estará habilitado se o campo 55-Coletou amostra? = 1.	PCR_RESUL
65-Data do Resultado RT-PCR/outro método por Biologia Molecular	Data DD/MM/AAAA		Data do Resultado RT-PCR/outro método por Biologia Molecular	<b>Campo Essencial</b> Campo habilitado se selecionado categoria 1-Detectável, 2-Não Detectável ou 3-Inconclusivo em Resultado da RT-PCR/outro método por Biologia Molecular.  Data deve ser >= a data da coleta- campo 56.	DT_PCR

66- Agente etiológico – RT-PCR/outr método por Biologia Molecular: Positivo para Influenza?	Varchar2 (1)	1-Sim 2-Não 9-Ignorado	Resultado da RT- PCR foi positivo para Influenza	<b>Campo Essencial</b>	POS_PCRFLU
66- Agente etiológico – RT-PCR/outr método por Biologia Molecular: Se sim, qual Influenza?	Varchar2(1)	1-Influenza A 2-Influenza B	Resultado diagnóstico do RT- PCR para o tipo de Influenza.	<b>Campo Essencial</b> Ha bilittado se campo 66-Agente etiológico – RT_PCR/outr método por Biologia Molecular: Positivo para Influenza? = 1.	TP_FLU_PCR
66- Agente etiológico – RT-PCR/outr método por Biologia Molecular: Se Influenza A, qual subtipo?	Varchar2(1)	1-Influenza A(H1N1)pdm09 2-Influenza A (H3N2) 3-Influenza A não subtipado 4-Influenza A não subtipável 5-Inconclusivo 6-Outro, especifique:	Subtipo para Influenza A.	<b>Campo Essencial</b> Ha bilittado se campo 66-Agente etiológico – RT_PCR/outr método por Biologia Molecular: Se sim, qual Influenza? = 1.	PCR_FLUASU
66- Agente etiológico – RT-PCR/outr método por Biologia Molecular: Se Influenza A, qual subtipo? Outro, especifique:	Varchar2 (30)		Outro subtipo para Influenza A.	<b>Campo Essencial</b> Ha bilittado se Agente etiológico – RT-PCR/outr método por Biologia Molecular: Se Influenza A, qual subtipo? = 6.	FLUASU_OUT
66- Agente etiológico – RT-PCR/outr método por Biologia Molecular: Se Influenza B, qual linhagem?	Varchar2(1)	1-Victoria 2-Yamagata 3-Não realizado 4-Inconclusivo 5-Outro, especifique:	Linhagem para Influenza B.	<b>Campo Essencial</b> Ha bilittado se campo 66-Agente etiológico – RT_PCR/outr método por Biologia Molecular: Se sim, qual Influenza? = 2.	PCR_FLUBLI
66- Agente etiológico – RT-PCR/outr método por Biologia Molecular: Se Influenza B, qual linhagem? Outro, especifique:	Varchar2 (30)		Outra linhagem para Influenza B.	<b>Campo Essencial</b> Ha bilittado se 66- Agente etiológico – RT-PCR/outr método por Biologia Molecular: Se Influenza B, qual linhagem? = 5.	FLUBLI_OUT
66- Agente etiológico – RT-PCR/outr método por Biologia Molecular: Positivo para outros vírus?	Varchar2 (1)	1-Sim 2-Não 9-Ignorado	Resultado da RT- PCR foi positivo para outro vírus respiratório	<b>Campo Essencial</b>	POS_PCROUT
66- Agente etiológico – RT-PCR/outr método por Biologia Molecular: SARS-CoV-2	Varchar2 (1)	1-marca do pelo usuário Vazio - não o marcado	Resultado diagnóstico do RT- PCR para (SARS-CoV- 2).	<b>Campo Essencial</b> Ha bilittado se campo 66- Agente etiológico – RT-PCR/outr método por Biologia Molecular: Positivo para outros vírus? = 1.	PCR_SARS2
66- Agente etiológico – RT-PCR/outr método por Biologia Molecular: VSR	Varchar2 (1)	1-marca do pelo usuário Vazio - não o marcado	Resultado diagnóstico do RT- PCR para (VSR).	<b>Campo Essencial</b> Ha bilittado se campo 66- Agente etiológico – RT-PCR/outr método por VSR	PCR_VSR



						<b>Biologia Molecular:</b> Positivo para outros vírus? = 1	
<b>66- Agente etiológico – RT-PCR/outror método por Biologia Molecular:</b> Outro vírus respiratório, especifique:	Varchar2 (1)	1-marca do pelo usuário Vazio - não marcado		Resultado diagnóstico do RT-PCR para Outro vírus respiratório.	<b>Campo Essencial</b> Ha bilitado se campo <b>66- Agente etiológico – RT-PCR/outror método por Biologia Molecular:</b> Positivo para outros vírus? = 1	<b>PCR_OUTRO</b>	
<b>66- Agente etiológico – RT-PCR/outror método por Biologia Molecular:</b> Outro vírus respiratório (Descrição)	Varchar2 (30)			Nome do outro vírus res piratório identificado pelo RT-PCR.	<b>Campo Essencial</b> Ha bilitado se <b>66- Agente etiológico – RT-PCR/outror método por Biologia Molecular:</b> Outro vírus respiratório, especifique:	<b>DS_PCR_OUT</b>	
<b>67- Laboratório que realizou RT-PCR/outror método por Biologia Molecular</b> Código (CNES)	Varchar2 (7)	Tabela com códigos CNES e nomes dos Laboratórios cadastrados no sistema.		La boratório res ponsável pela liberação do resultado do teste diagnóstico (RT-PCR) da amostra do paciente.	<b>Campo Essencial</b> Ha bilitados e selecionado categoria 1- Detectável, 2- Não Detectável ou 3- Inconclusivo em <b>66- Resultado da RT-PCR/outror método por Biologia Molecular.</b>  Preenchendo o nome do La boratório, o código é preenchido automaticamente, ou vice-versa.	<b>LAB_PCR OU CO_LAB_PCR</b>	
<b>68- Tipo de Amostra Sorológica para SARS-Cov-2</b>	Number(3)	1- Sangue/plasma/soro 2- Outra, qual? 9- Ignorado		Tipo de amostra sorológica que foi coletada.	<b>Campo Essencial</b>	<b>TP_AM_SOR</b>	
<b>Tipo de Amostra Sorológica para SARS-Cov-2/Outra, qual?</b>				Descrição do tipo da amostra clínica, caso diferente das listadas na categoria um (1) do campo.	<b>Campo Essencial</b> Campo o habilitado e selecionado categoria 2- Outra, qual? em Tipo de Amostra Sorológica.	<b>SOR_OUT</b>	
<b>69- Data da coleta</b>	Data DD/MM/AAAA			Da ta d a coleta do material para diagnóstico por Sorologia.	<b>Campo Essencial</b> Ha bilitado de campo <b>55- Coletou amostra? = 1.</b>  Da ta d e ve s er maior ou igual a <b>2- Data de 1ºs sintomas</b> e menor ou igual a data da digitação (atual).	<b>DT_CO_SOR</b>	
<b>70- Tipo de Sorologia para SARS-Cov-2</b>	Number(3)	1- Teste rápido 2- Elisa 3- Quimioluminescência 4- Outro, qual		Tipo do Teste Sorológico que foi realizado	<b>Campo Essencial</b>	<b>TP_SOR</b>	

70- Tipo de Sorologia para SARS-Cov-2	Varchar2(100)		Descrição do tipo de Teste Sorológico	<b>Campo Essencial</b> Campo habilitado se selecionado categoria 4-Outro, qual? em Tipo de Sorologia.	OUT_SOR
70- Tipo de Sorologia para SARS-Cov-2/Outro, qual?	Varchar2(100)		Outro tipo de amostra Sorológica		SOR_OUT
70- Resultado do Teste Sorológico para SARS-Cov-2:	Varchar2(1)	1-Positivo 2-Negativo 3-Inconclusivo 4-Não realizado 5-Aguarda resultado 9-Ignorado	Resultado da Sorologia para SARS-CoV-2	<b>Campo Essencial</b>	RES_SOR
70- Resultado do Teste Sorológico para SARS-Cov-2:	Varchar2(1)	IgG	Resultado da Sorologia para SARS-CoV-2	<b>Campo Essencial</b>	RES_IGG
70- Resultado do Teste Sorológico para SARS-Cov-2:	Varchar2(1)	IgM	Resultado da Sorologia para SARS-CoV-2	<b>Campo Essencial</b>	RES_IGM
70- Resultado do Teste Sorológico para SARS-Cov-2:	Varchar2(1)	IgA	Resultado da Sorologia para SARS-CoV-2	<b>Campo Essencial</b>	RES_IGA
71- Data do Resultado	Data DD/MM/AAAA		Data do Resultado do Teste Sorológico	<b>Campo Essencial</b> Data deve ser maior ou igual a 69- Data da Coleta	DT_RES
72-Classificação final do caso	Varchar2(1)	1-SRAG por influenza 2-SRAG por outro vírus respiratório 3-SRAG por outro agente etiológico, qual: 4-SRAG não especificado 5-SRAG por COVID-19	Diagnóstico final do caso.  Se tiver resultados divergentes entre as metodologias laboratoriais, priorizar o resultado do RT-PCR.	<b>Campo Obrigatório</b>	CLASSI_FIN
72-Classificação final do caso 3-SRAG por outra causa, qual:	Varchar2(30)		Descrição de qual outro agente etiológico foi identificado	<b>Campo Obrigatório</b> Se campo 72-Classificação final do caso = 3.	CLASSI_OUT



					Ha bilittados e campo 72- Classificação final do caso = 3.	
<b>73- Critério de Encerramento</b>	Varcha r2(1)	1. Laboratorial 2. Clínico Epidemiológico 3. Clínico 4. Clínico Imagem	Indicar qual o critério de confirmação.	<b>Campo Essencial</b>		<b>CRITERIO</b>
<b>74- Evolução do caso</b>	Varcha r2(1)	1- Cura 2- Óbito 3- Óbito por outras causas 9- Ignorado	Evolução do caso	<b>Campo Essencial</b>		<b>EVOLUCAO</b>
<b>75- Data da alta ou óbito</b>	Data DD/MM/AAAA		Data da alta ou óbito	<b>Campo Essencial</b> Data da alta ou do óbito deve ser > ou = a data dos primeiros sintomas e <= a data da digitação (atual). Ha bilittados e campo 74- <b>Evolução do caso = 1 ou 2.</b>		<b>DT_EVOLUCA</b>
<b>76- Data do Encerramento</b>	Data DD/MM/AAAA		Data do encerramento do caso.	<b>Campo Obrigatório</b> Se o campo 72- Classificação final do caso estiver preenchido. Data do encerramento deve ser > ou = a data do preenchimento. Data do encerramento deve ser < ou = a data da digitação (atual).		<b>DT_ENCERRA</b>
<b>77- Número D.O</b>			Número da Declaração de Óbito	<b>Campo Essencial</b> Ha bilittados e o campo 74- Evolução do caso = 2		<b>NU_DO</b>
<b>78- Observações</b>	Varcha r2(999)		Outras observações sobre o paciente consideradas pertinentes.	<b>Campo Opcional</b>		<b>OBSERVA</b>
<b>79- Profissional de Saúde Responsável</b>	Varcha r2(60)		Nome completo do profissional de saúde (sem abreviações) responsável pela notificação.	<b>Campo Essencial</b>		<b>NOME_PROF</b>
<b>80- Registro Conselho/Matrícula</b>	Varcha r2(15)		Número do conselho ou matrícula do	<b>Campo Essencial</b>		<b>REG_PROF</b>

				profissional de saúde responsável pela notificação (Ex: CRM/RJ 1234)		
Data da digitação		Data DD/MM/AAAA		Data de inclusão do registro no sistema.	<b>Campo Interno</b> Preenchido automaticamente pelo sistema com a data da digitação da ficha. Não é a data de preenchimento da ficha manualmente e sim a data em que é digitado no sistema. <b>Não é atualizada se houver alterações posteriores de dados.</b>	DT_DIGITA