

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

1. Introdução	2
1.1 Contextualização	2
1.2 Objetivos	4
2. Coleta dos dados	5
2.1 Dados OpenDataSUS	5
2.2 Dados do Instituto Brasileiro de Geografia e Estatística	7
3. Processamento	8
3.1 Dados do Instituto Brasileiro de Geografia e Estatística	8
3.2 OpenDataSUS	9
3.3 Transformação	9
3.3.1 Notificações	11
3.3.2 Idade	12
3.3.3 Raça	13
3.3.4 Sexo	14
3.3.5 Comorbidades	15
3.3.6 Recursos do tratamento	16
3.3.7 Evolução	18
4. Análise Exploratória	20
4.1 Número absoluto de mortes de covid-19 no Brasil	20
4.2 Taxa de mortalidade da covid-19 no Brasil	22
4.3 Taxa de letalidade da covid-19 no Brasil	24
5. Inferência Estatística	27
5.1 Perfil do brasileiro com maior e menor risco de morte	27
6. Machine learning	31
6.1 Árvore de Desisão de Classificação	31
6.2 Arvore de Decisão e Seleção de Atributos	31
6.3 Árvore de Decisão de Classificação otimizado	32
6.4 Naive Bayse	33
Apêndice A - Taxa de Mortalidade Estados e Capitais	35
Anexo A - Ficha de Notificação	36
Anexo B - Dicionário de dados	38

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, que pode ser acessada no link: https://drive.google.com/drive/folders/1l52SO-RMHYfs2ci_2pBAZJM

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",
"SUPPORT_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_TOMOGRAFIA_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_CLASSIFICACAO_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 no link: [https://www.ibge.gov.br/estatisticas/downloads-estatisticas.h](https://www.ibge.gov.br/estatisticas/downloads-estatisticas.html) no diretório: Estimativa_de_Populacao/Estimativas_2020, em formato xls, conforme figura 1 abaixo:

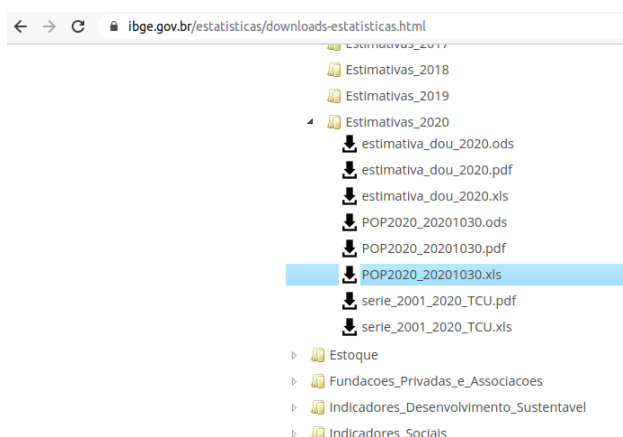


Figura 1: Arquivo IBGE

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:

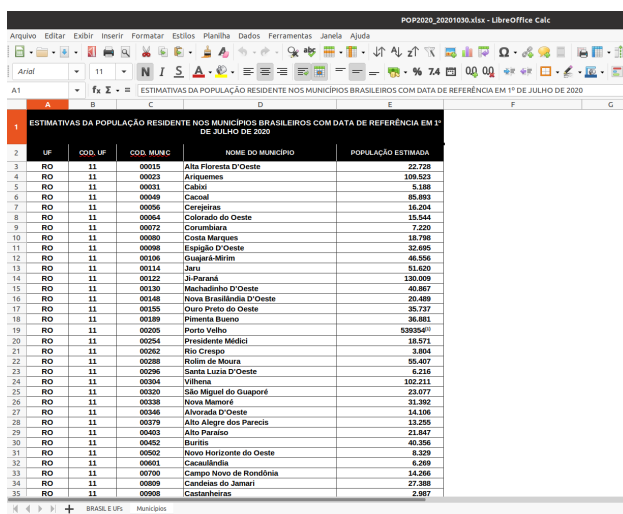
```
arquivoIBGE=read.xlsx("dados/POP2020_20201030.xlsx", sheet = "Municípios",
startRow = 3, colNames =FALSE, rowNames = FALSE)
```

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.

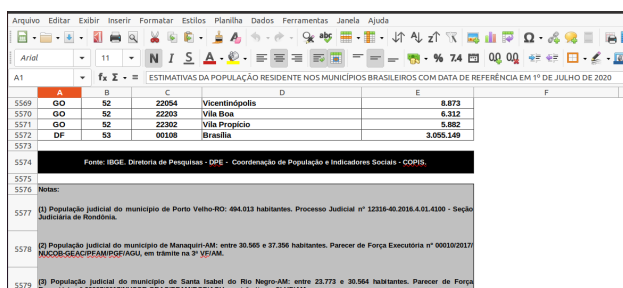
3.1 Dados do Instituto Brasileiro de Geografia e Estatística

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	Ariquemes	106.523
RO	11	00031	Cabral	5.188
RO	11	00048	Canoinhas	88.853
RO	11	00056	Cerejeiras	16.204
RO	11	00064	Colorado do Oeste	15.544
RO	11	00072	Corumbiara	7.235
RO	11	00080	Costa Marques	18.798
RO	11	00098	Engelão D'Oeste	32.695
RO	11	00106	Guajará-Mirim	46.555
RO	11	00114	Jaru	51.620
RO	11	00122	Ji-Paraná	126.009
RO	11	00130	Machadinho D'Oeste	40.867
RO	11	00148	Novo Brasilândia D'Oeste	20.489
RO	11	00156	Osório Neto do Oeste	26.727
RO	11	00189	Pimenta Bueno	36.881
RO	11	00205	Porto Velho	539241
RO	11	00254	Presidente Médici	18.571
RO	11	00262	Rio Crespo	3.804
RO	11	00288	Rolim de Moura	56.407
RO	11	00296	Santa Luzia D'Oeste	6.216
RO	11	00304	Vilhena	102.211
RO	11	00320	São Miguel do Guaiporé	23.677
RO	11	00338	Nova Mamoré	31.392
RO	11	00346	Ahorada D'Oeste	14.106
RO	11	00378	Alto Alegre dos Parecis	13.256
RO	11	00403	Alto Paraíso	21.847
RO	11	00452	Buritis	40.356
RO	11	00502	Novo Horizonte do Oeste	6.529
RO	11	00601	Cacaulândia	6.269
RO	11	00700	Campo Novo de Rondônia	14.266
RO	11	00800	Candonga do Jari	27.388
RO	11	00908	Castanheiras	2.587

Figura 2: Arquivo IBGE



UF	COD_UF	COD_MUNIC	NOME DO MUNICÍPIO	POPULAÇÃO ESTIMADA
GO	52	22054	Vicentinópolis	8.873
GO	52	22203	Vila Boa	6.312
GO	52	22302	Vila Propício	5.882
DF	53	00108	Brasília	3.095.149

Fonte: IBGE, Diretoria de Pesquisas - DPE - Coordenação de População e Indicadores Sociais - COPIS.

Nota:

(1) População judicial do município de Porto Velho-RO: 494.013 habitantes. Processo Judicial nº 12218-40.2018.4.01.4100 - Seção Judiciária de Rondônia.

(2) População judicial do município de Manauzin-AM: entre 30.565 e 27.356 habitantes. Parecer de Força Executória nº 00013/2017/MACOB-GEAC/PFAM/PQFAGUL em trâmite na 3ª VJ/AM.

(3) População judicial do município de Santa Isabel do Rio Negro-AM: entre 23.773 e 30.564 habitantes. Parecer de Força Executória nº 00007/2017/MACOB-GEAC/PFAM/PQFAGUL em trâmite na 3ª VJ/AM.

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 com fluxo do pré-processamento dos dados: [4](#)

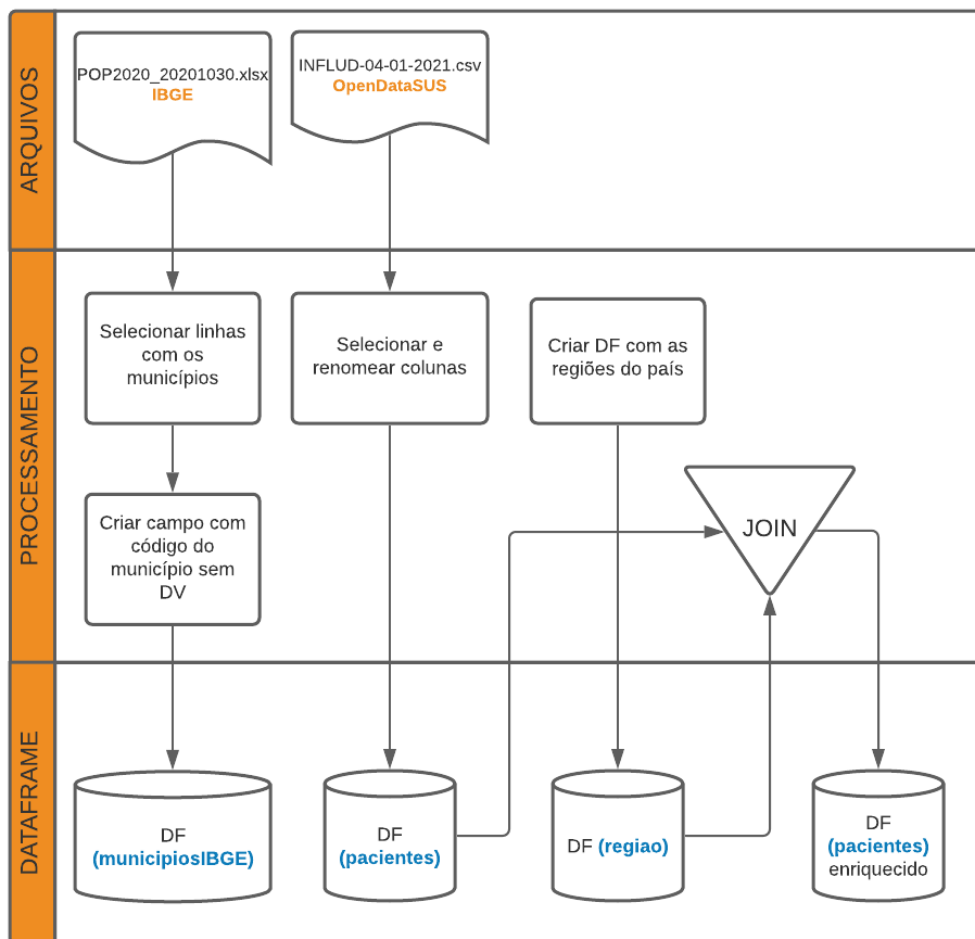


Figura 4: Pré-processamento dos dados

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.

```
#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.

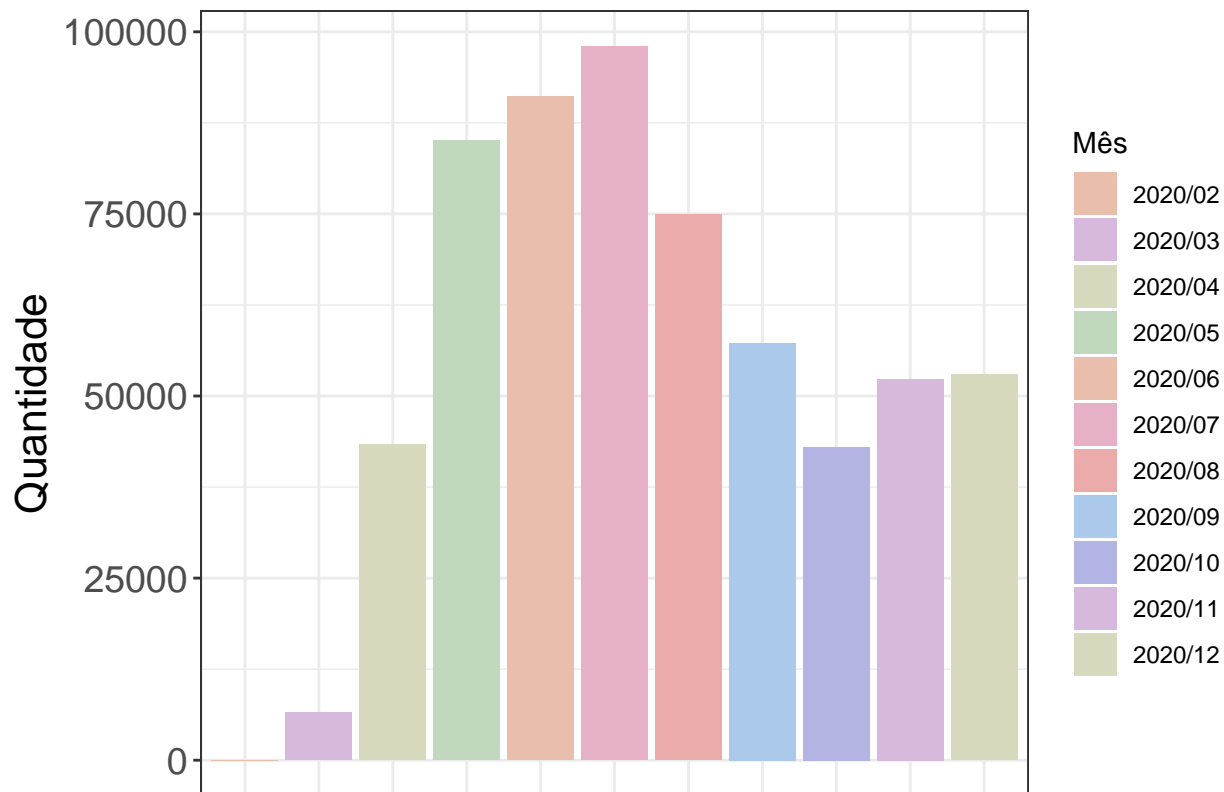


Gráfico 1: Notificações de Covid-19 por Mês

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, sudeste e sul;
- aproximadamente 75

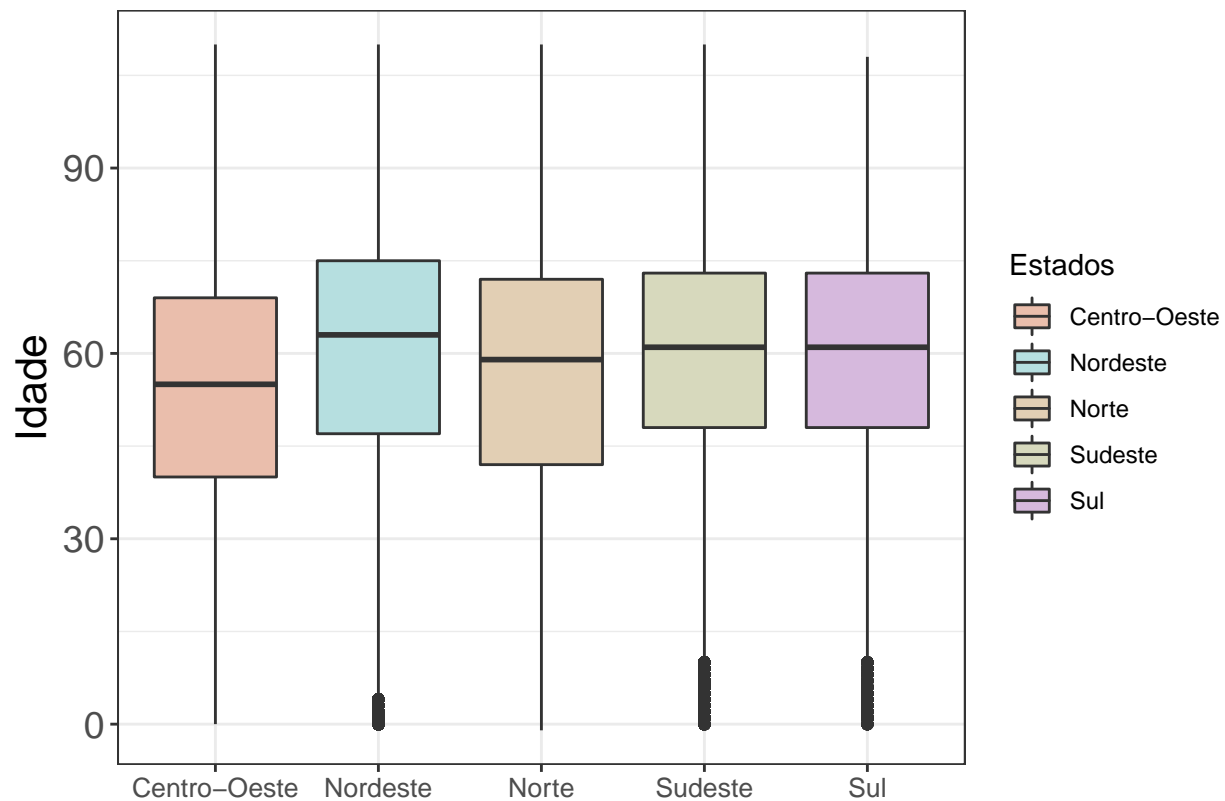


Gráfico 2: Boxplot de Idades dos casos graves de covid-19 por região

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 indígina foi a que registrou menos casos;

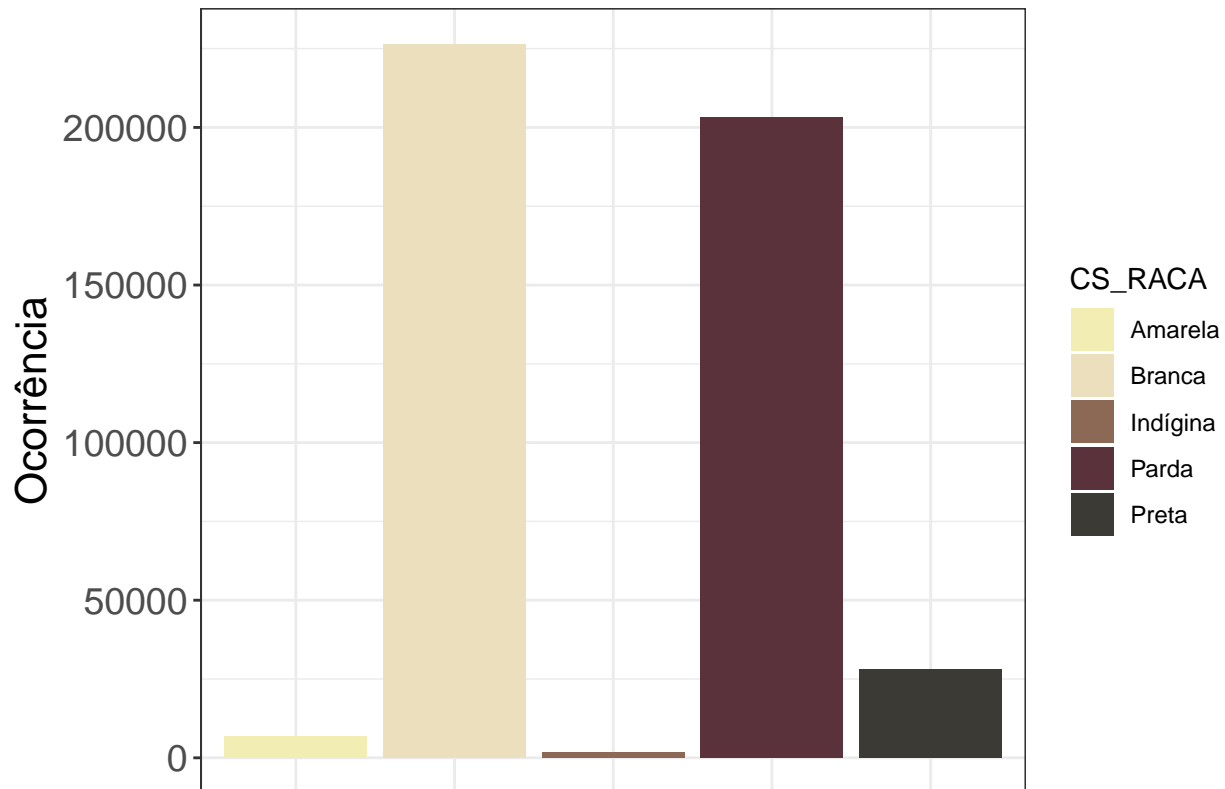


Gráfico 3: Quantitativo dos casos graves de covid-19 por raça

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 apresntam 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$CS_SEX0 = factor(pacientes$CS_SEX0)
```

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

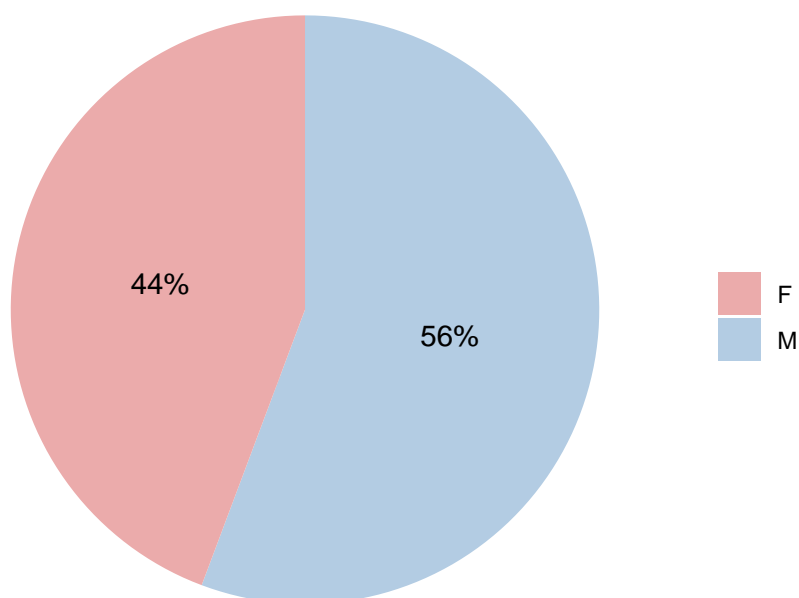


Gráfico 4: Percentual dos casos graves de covid-19 por sexo

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
#Alterando domínios
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
#Alterando domínios
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.

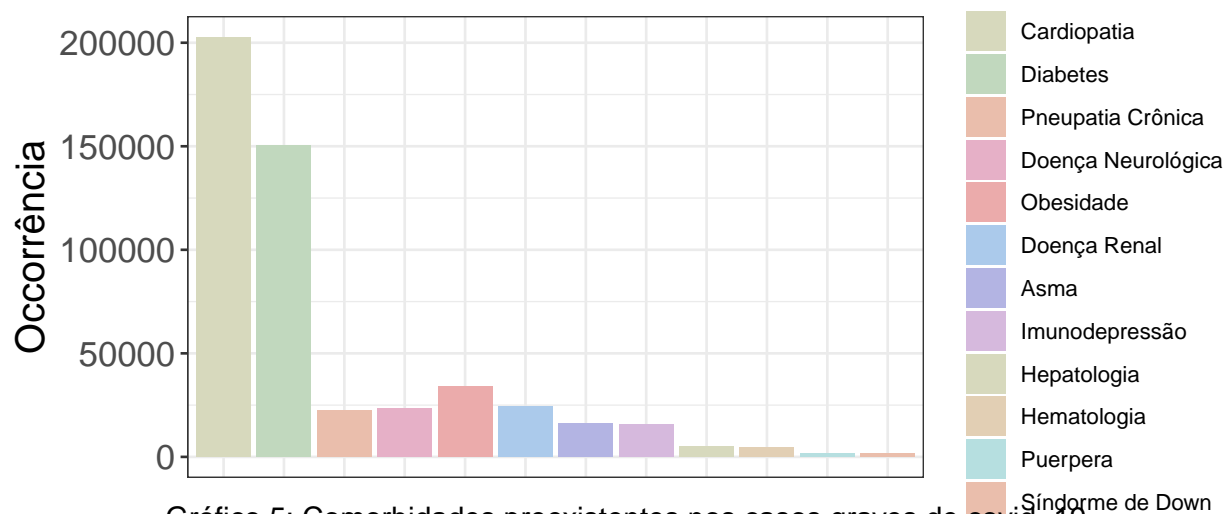


Gráfico 5: Comorbidades preexistentes nos casos graves de covid-19

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:

```
#Usou Suporte Ventilatório Invasivo / Não Invasivo
#Foi Hospitalizado
#Alterando domínios
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 conclui-se que aproximadamente metade dos pacientes hospitalizados precisaram de oxigênio não invasivo.

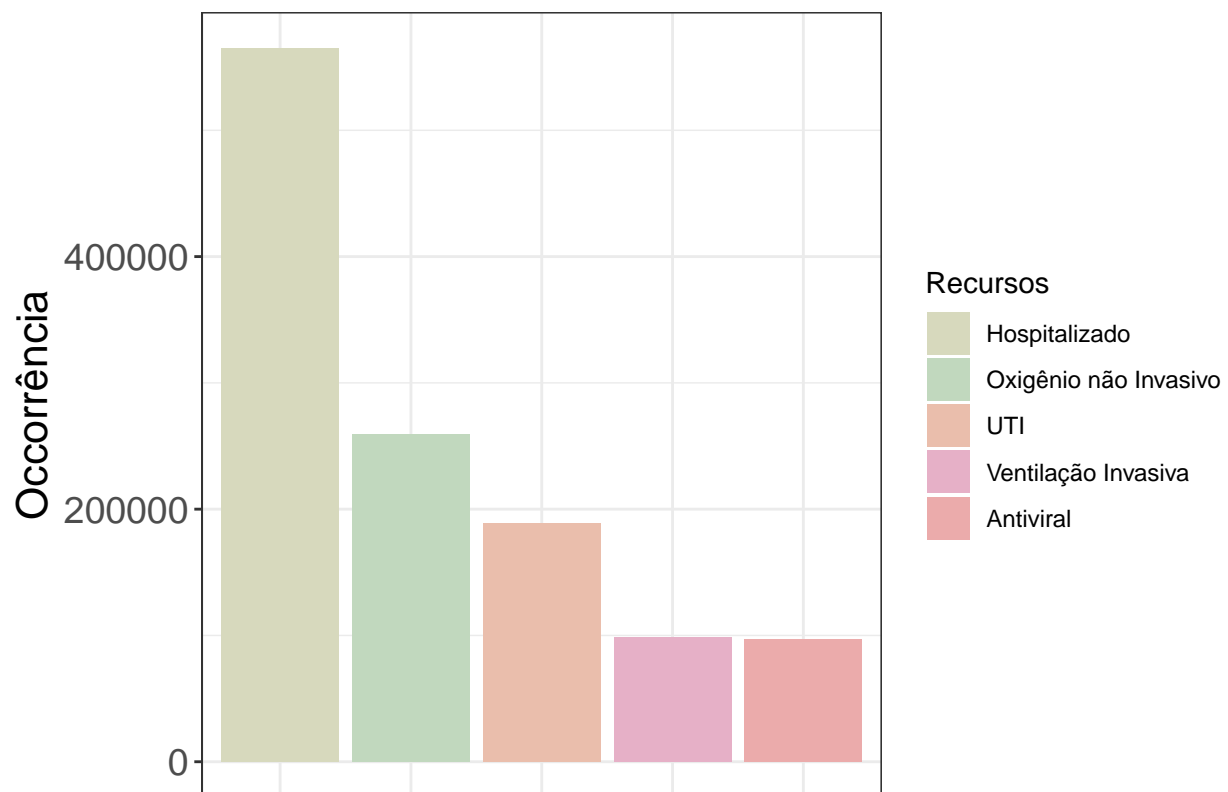


Gráfico 6: Recursos utilizados nos casos graves de covid-19

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), os quais foram removidos na elaboração dos gráficos e modelos de machine learning.

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$CS_EVOLUCAO=factor(pacientes$CS_EVOLUCAO)
```

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

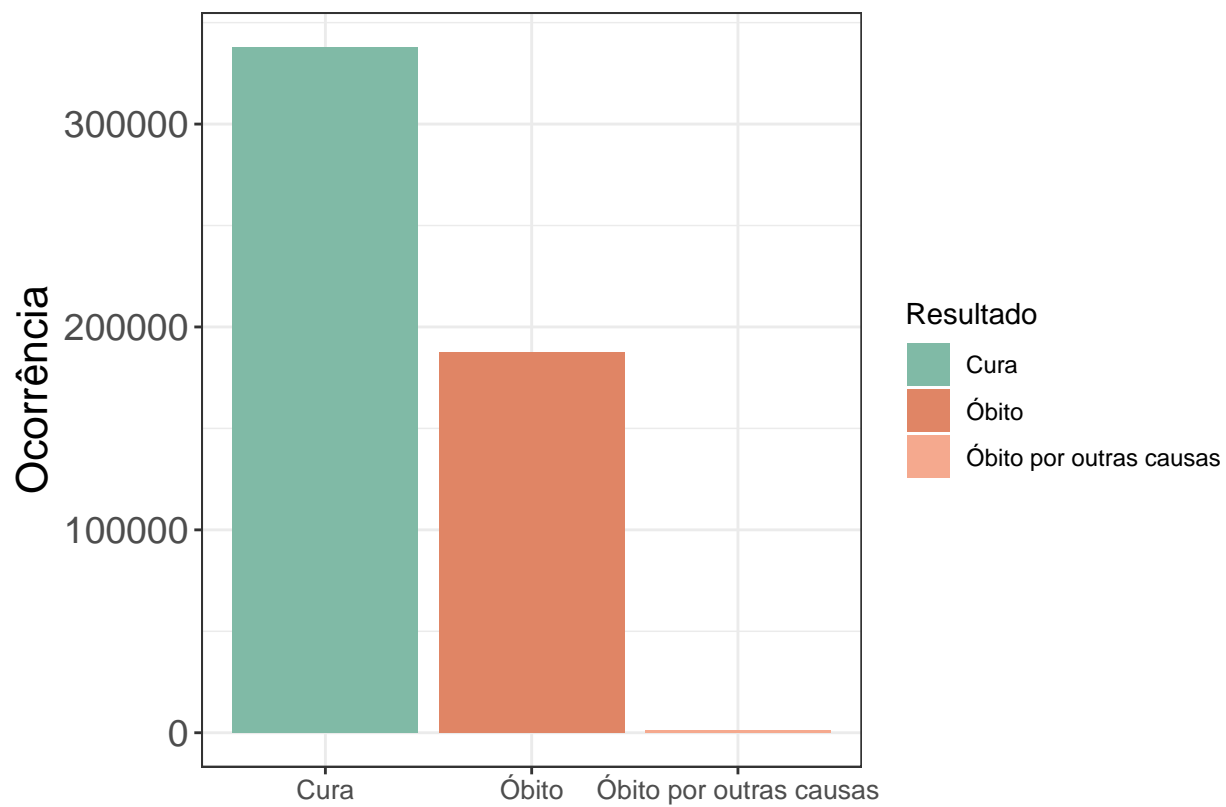


Gráfico 7: Evolução dos casos graves de covid-19

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:

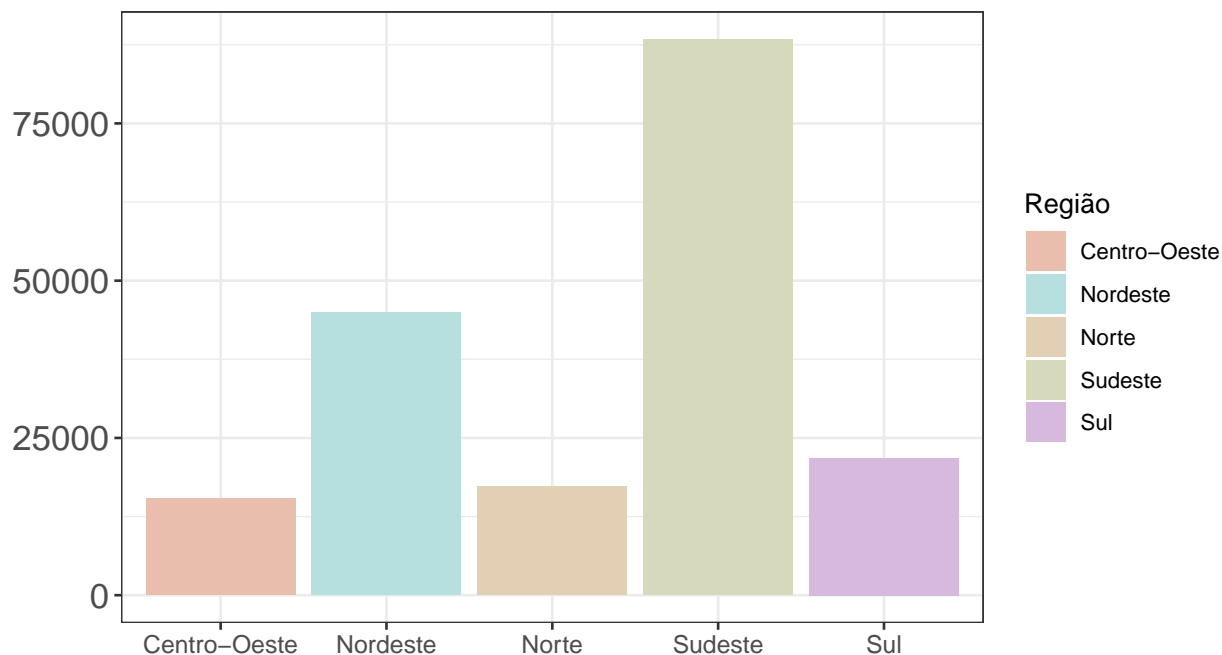


Gráfico 8: Mortes de covid-19 por região

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:

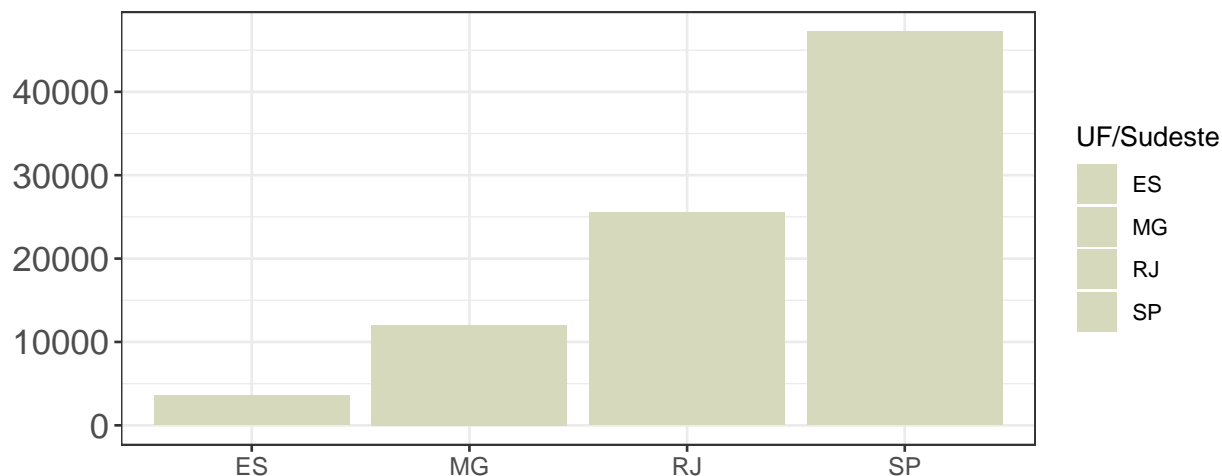


Gráfico 9: Mortes por covid-19 na região Sudeste

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.

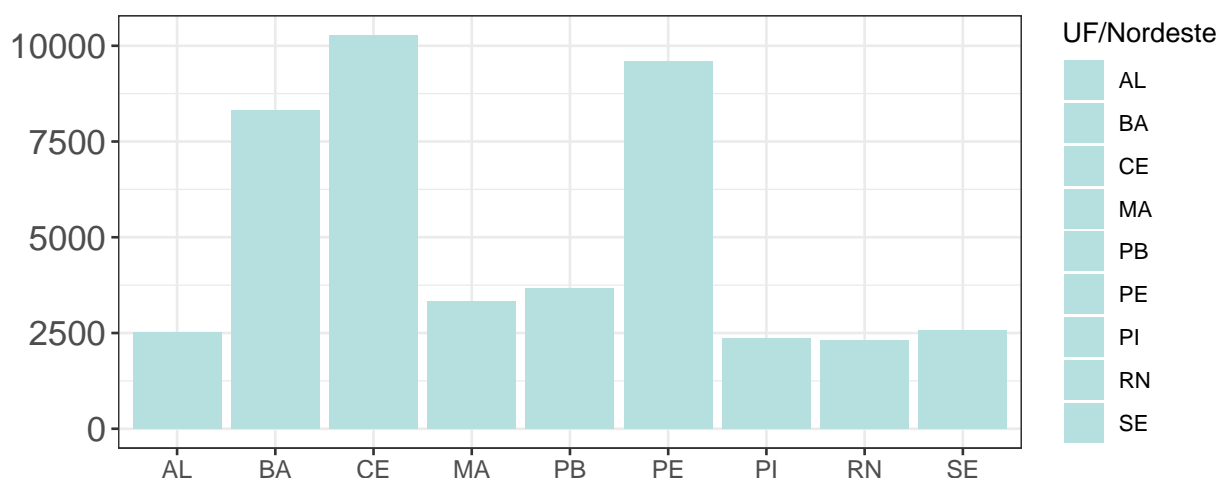


Gráfico 10: Mortes por covid-19 na região Nordeste

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.

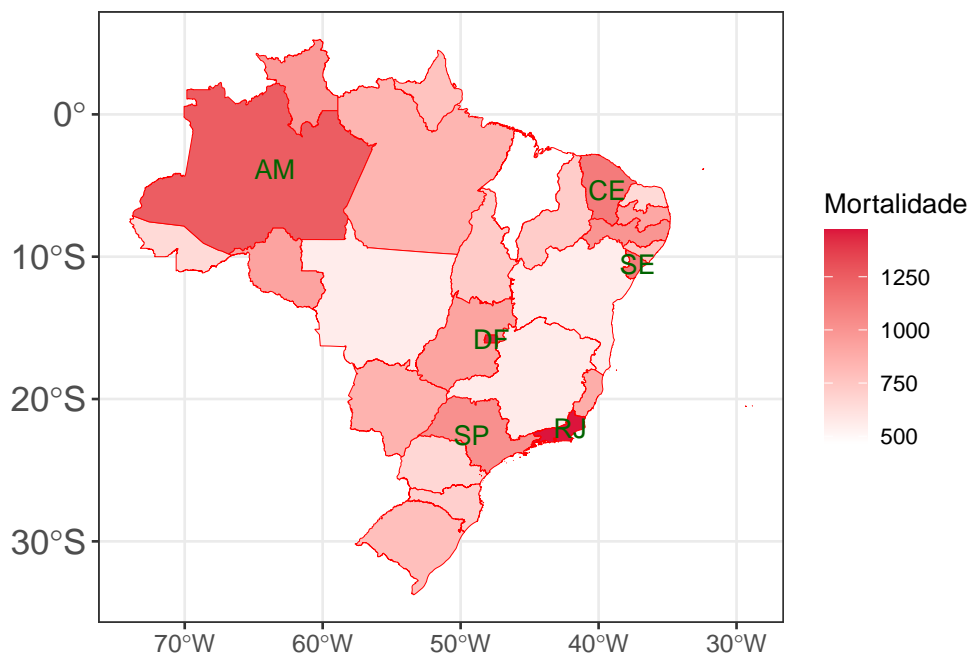


Gráfico 11: Mapa de calor das taxas de mortalidade por milhão de habitantes

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 19ª (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, pondendo-se, ainda, visualizar mais detalhes das taxas de mortalidade por covid-19 nas capitais no Apêndice A, deste trabalho.

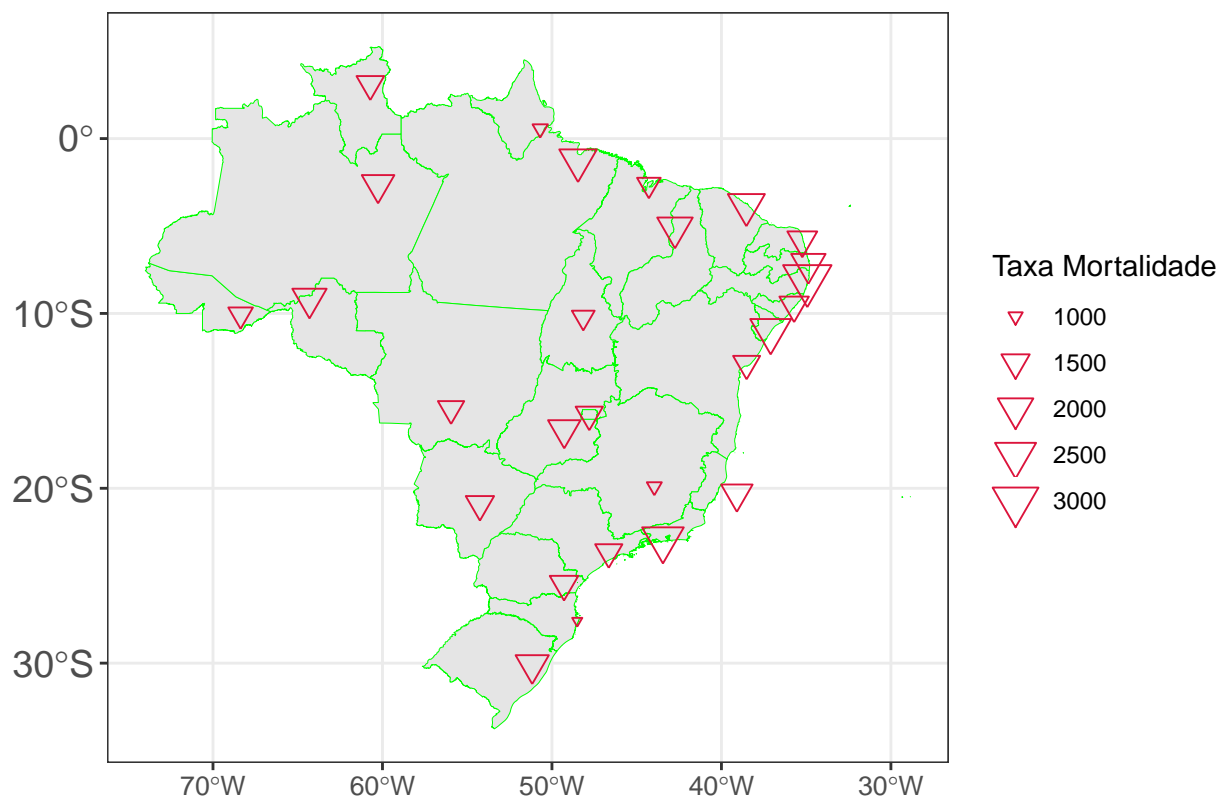


Gráfico 12: Taxa de mortalidade por covid-19 nas capitais

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 35% (trinta e cinco por cento), com nível de significância de 95% (noventa e cinco por cento). Para este teste, utilizou-se uma amostra com 493.035 (quatrocentos e noventa e três mil e trinta e cinco) pacientes hospitalizados. Esta taxa coincide com a 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 = 0.88042, df = 1, p-value = 0.3481
## alternative hypothesis: true p is not equal to 0.35
## 95 percent confidence interval:
##  0.3493066 0.3519725
## sample estimates:
##      p
```

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^a (vigésima) semana.

Nas semanas seguintes, a região Norte apresentou queda acentuada de sua taxa, chegando ao seu ponto mais baixo na 43^a (quadragésima terceira) semana. Nesta ocasião, chegou a

apresentar a menor média do país. Contudo, à partir da semana 44^a (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, à partir da 48^a (quadragésima oitava) semana, passou a apresentar uma subida acentuada de sua taxa.

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

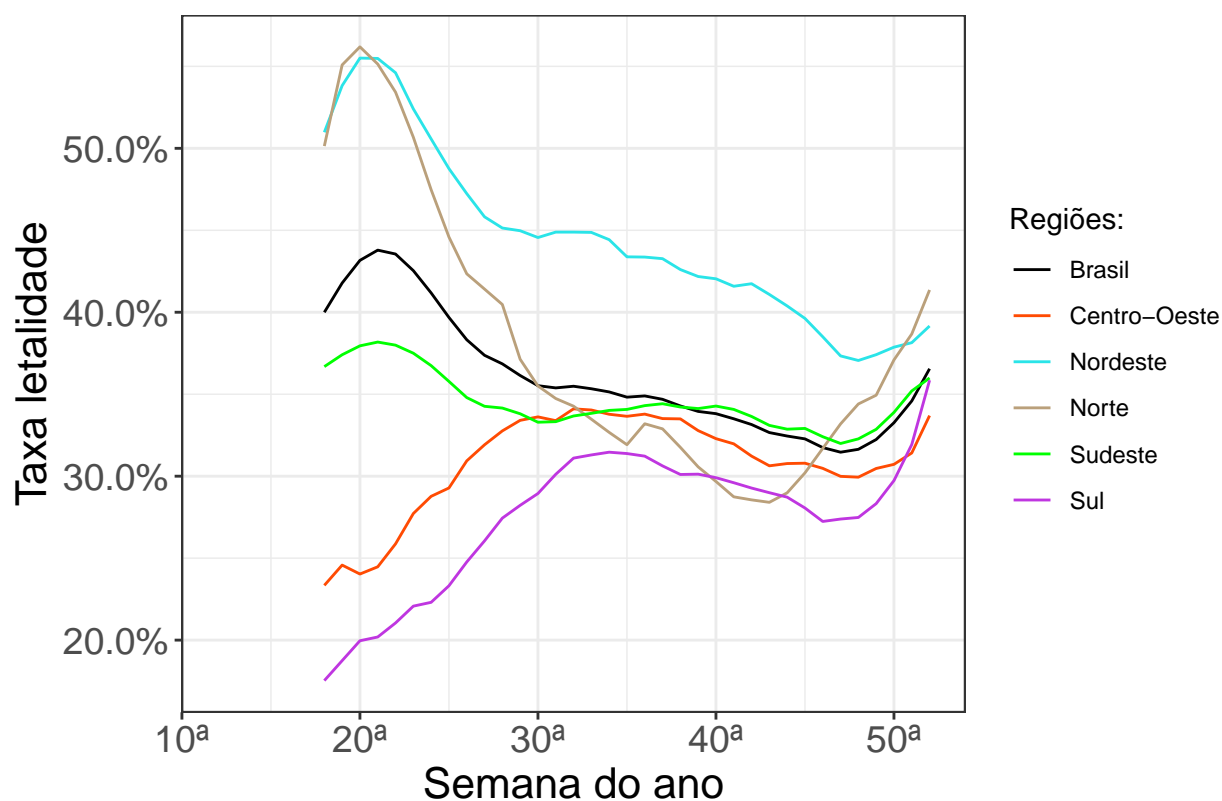


Gráfico 13: Média móvel semanal da taxa de letalidade

5. Inferência Estatística

No capítulo anterior, foi analisado os diferentes impactos da covid-19 a parti da regiões, estados e município. Neste capítulo, o foco será identificar quais as características do paciente que impactam a taxa de letalidade da covid-19, no Brasil.

5.1 Perfil do brasileiro com maior e menor risco de morte

Para reduzir a variação dos valores as idades foram agrupadas em faixas. Neste momento, foram adotados os mesmos intervalos utilizados pelo IBGE.

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

Primeira Execução

```
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" )
```

Na primeira execução o modelo apresentou valores de baixa significância.

Para a faixa etária, nos intervalos de 5 a 9 anos e 10 a 14 anos o modelo apresentou baixa significância, o modelo, também não exibiu a faixa de 0 a 4 anos. Assim, esses intervalos foram reagrupadas em um novo intervalo de 0 a 14 anos.

No que se refere à raça do paciente, o modelo apresentou baixa significância para raça branca e não exibiu a raça amarela. Então, optou-se por agrupar as raças amarela e branca em um grupo misto. Conforme tabela [5](#) da página [28](#).

	β^1	Std. Error	z value	Pr(> z)
β^0	-3.1647	0.0743	-42.61	0.0000
Homem	0.2465	0.0075	32.99	0.0000
05 a 09 anos	-0.1450	0.1098	-1.32	0.1865
10 a 14 anos	0.1796	0.1155	1.56	0.1199
15 a 19 anos	0.4714	0.0943	5.00	0.0000
20 a 24 anos	0.4353	0.0809	5.38	0.0000
25 a 29 anos	0.4097	0.0751	5.46	0.0000
30 a 34 anos	0.5109	0.0712	7.18	0.0000
35 a 39 anos	0.6436	0.0691	9.31	0.0000
40 a 44 anos	0.8465	0.0683	12.40	0.0000
45 a 49 anos	1.0456	0.0678	15.43	0.0000
50 a 54 anos	1.2462	0.0673	18.51	0.0000
55 a 59 anos	1.5130	0.0670	22.58	0.0000
60 a 64 anos	1.8458	0.0668	27.62	0.0000
65 a 69 anos	2.1292	0.0668	31.89	0.0000
70 a 74 anos	2.3683	0.0668	35.45	0.0000
75 a 79 anos	2.5917	0.0670	38.71	0.0000
80 anos ou mais	2.9861	0.0666	44.85	0.0000
Branca	0.0225	0.0316	0.71	0.4764
Indígena	0.3425	0.0667	5.13	0.0000
Parda	0.1790	0.0316	5.66	0.0000
Preta	0.3760	0.0343	10.95	0.0000
Tem comorbidades	0.5298	0.0085	62.17	0.0000
Nordeste	0.4173	0.0158	26.37	0.0000
Norte	0.3576	0.0181	19.79	0.0000
Sudeste	0.0532	0.0144	3.70	0.0002
Sul	-0.1697	0.0169	-10.04	0.0000

Tabela 5: Modelo com valores de baixa significância presentes

Segunda Execução

```
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" )
```

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 (β^1 : 2.9888), já apresenta doenças preexistentes (β^1 : 0.5299), reside na região nordeste (β^1 : 0.4171), é de cor preta (β^1 : 0.3543) e homem (β^1 : 0.2465), com probabilidade de óbito de 80%.

Já o perfil com menor risco de morte no Brasil: reside na região sul (β^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%.

Os valores de β^0 e β^1 do modelo otimizado podem ser verificados na tabela: [6](#) da página [30](#).

	β^1	Std. Error	z value	Pr(> z)
β^0	-3.1459	0.0488	-64.53	0.0000
Homem	0.2465	0.0075	32.98	0.0000
15 a 19 anos	0.4742	0.0817	5.80	0.0000
20 a 24 anos	0.4380	0.0659	6.65	0.0000
25 a 29 anos	0.4124	0.0585	7.05	0.0000
30 a 34 anos	0.5136	0.0534	9.62	0.0000
35 a 39 anos	0.6463	0.0507	12.76	0.0000
40 a 44 anos	0.8491	0.0494	17.17	0.0000
45 a 49 anos	1.0483	0.0487	21.51	0.0000
50 a 54 anos	1.2489	0.0481	25.95	0.0000
55 a 59 anos	1.5157	0.0477	31.79	0.0000
60 a 64 anos	1.8485	0.0474	38.96	0.0000
65 a 69 anos	2.1319	0.0473	45.03	0.0000
70 a 74 anos	2.3709	0.0474	50.03	0.0000
75 a 79 anos	2.5944	0.0476	54.48	0.0000
80 anos ou mais	2.9888	0.0471	63.48	0.0000
Indígena	0.3206	0.0595	5.39	0.0000
Parda	0.1574	0.0090	17.55	0.0000
Preta	0.3543	0.0156	22.64	0.0000
Tem comorbidade	0.5299	0.0085	62.18	0.0000
Nordeste	0.4171	0.0158	26.36	0.0000
Norte	0.3576	0.0181	19.79	0.0000
Sudeste	0.0535	0.0144	3.72	0.0002
Sul	-0.1689	0.0169	-10.01	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] 0.689072
```

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)
```

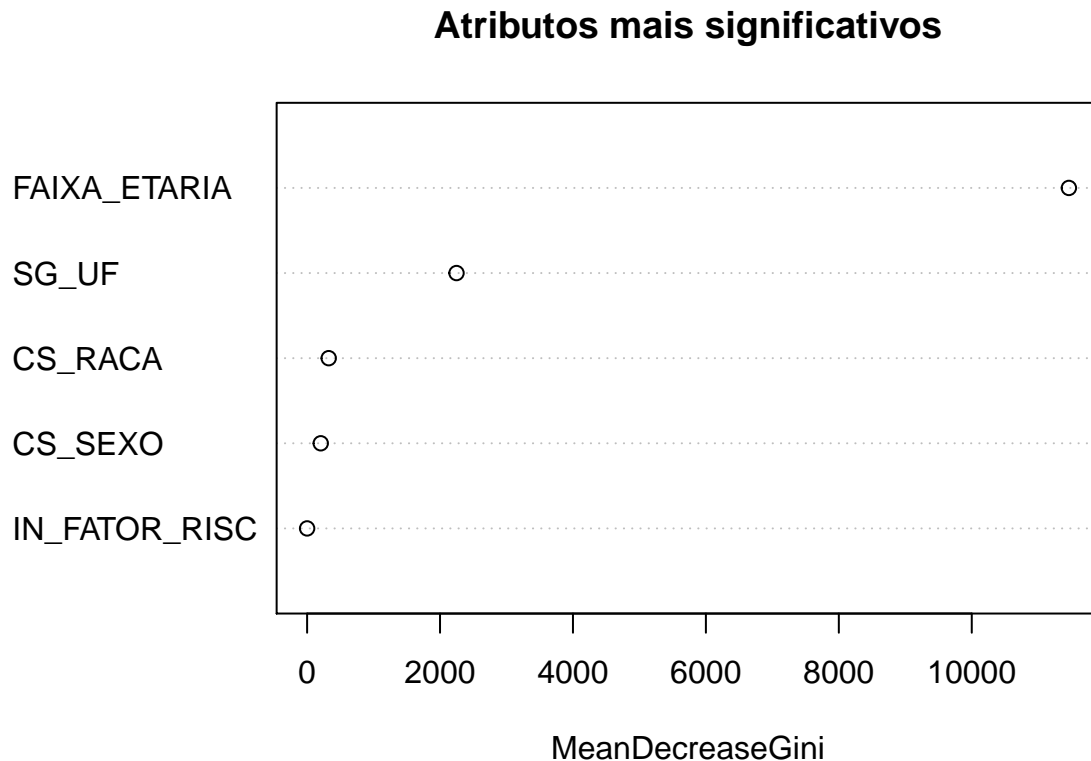
```
teste = predict(selecaoAtributos, dadosTeste)  
previsao=cbind(teste, dadosTeste)  
previsao["PREVISAO"]=previsao$teste
```

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

```
## [1] 0.6975923
```

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] 0.7234923
```

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] 0.7469846
```

6.4 Naive Bayse

Com os mesmo 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] 0.715159
```

\section

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

Nº



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

CASO DE SÍNDROME RESPIRATÓRIA AGUDA GRAVE (SRAG-HOSPITALIZADO): Indivíduo com *SG que apresente: dispneia/desconforto respiratório OU pressão persistente no tórax OU saturação de O₂ 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 1 ^{os} sintomas
3	UF:	4	Município:
			Código (IBGE):
5	Unidade de Saúde:		Código (CNES):
	6	CPF do cidadão:	
	7	Nome:	8
	9	Data de nascimento:	10
			(Ou) Idade: 1-Dia 2-Mês 3-Ano
	11	Gestante:	
			1-1 ^o Trimestre 2-2 ^o Trimestre 3-3 ^o Trimestre
	12	Raça/Cor:	
		1-Branca 2-Preta 3-Amarela 4-Parda 5-Indígena 9-Ignorado	
	13	Se indígena, qual etnia?	
	14	Escolaridade:	
		0-Sem escolaridade/Analfabeto 1-Fundamental 1 ^o ciclo (1 ^a a 5 ^a série) 2-Fundamental 2 ^o ciclo (6 ^a a 9 ^a série) 3-Médio (1 ^o ao 3 ^o 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:
	25	Zona:	26
		1-Urbana 2-Rural 3-Periurbana 9-Ignorado	País: (se residente fora do Brasil)
	27	Paciente tem histórico de viagem internacional até 14 dias antes do início dos sintomas?	
		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?	
		1-Sim 2-Não 9-Ignorado	
	33	Trata-se de caso nosocomial (infecção adquirida no hospital)?	
		1-Sim 2-Não 9-Ignorado	
	34	Paciente trabalha ou tem contato direto com aves, suínos, ou outro animal?	
		1-Sim 2-Não	
	35	Sinais e Sintomas:	
		1-Sim 2-Não 9-Ignorado	
		Febre Tosse Dor de Garganta Dispneia	
		Desconforto Respiratório Saturação O ₂ <95% Diarreia Vômito Dor abdominal Fadiga	
		Perda do olfato Perda do paladar Outros	
	36	Possui fatores de risco/comorbidades?	
		1-Sim 2-Não 9-Ignorado	
		Se sim, qual(is)? (Marcar X)	
		Puérpera (até 45 dias do parto) Doença Cardiovascular Crônica Doença Hematológica Crônica	
		Síndrome de Down Doença Hepática Crônica Asma	
		Diabetes mellitus Doença Neurológica Crônica Outra Pneumopatia Crônica	
		Imunodeficiência/Imunodepressão Doença Renal Crônica Obesidade, IMC	
		Outros	
	37	Recebeu vacina contra Gripe na última campanha?	38
		1-Sim 2-Não 9-Ignorado	Data da vacinação:
		Se < 6 meses: a mãe recebeu a vacina? 1-Sim 2-Não 9-Ignorado	
		a mãe amamenta a criança? 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 1 ^a dose: (1 ^a dose para crianças vacinadas pela primeira vez)	
		Data da 2 ^a dose: (2 ^a dose para crianças vacinadas pela primeira vez)	

Dados de Atendimento	39	Usou antiviral para gripe? <input type="checkbox"/> <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado	40	Qual antiviral? <input type="checkbox"/> <input type="checkbox"/> 1-Osetamivir 2-Zanamivir 3-Outro, especifique: _____	41	Data início do tratamento ____/____/____
	42	Houve internação? <input type="checkbox"/> <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"/> <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"/> <input type="checkbox"/> 1-Sim, invasivo 2-Sim, não invasivo 3-Não 9-Ignorado	51	Raio X de Tórax: <input type="checkbox"/> <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"/> <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"/> <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado	56	Data da coleta: _____	57	Tipo de amostra: <input type="checkbox"/> <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"/> <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"/> <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"/> <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado Se sim, qual influenza? <input type="checkbox"/> <input type="checkbox"/> 1-Influenza A 2-Influenza B Positivo para outros vírus? <input type="checkbox"/> <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"/> <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"/> <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado Se sim, qual influenza? <input type="checkbox"/> <input type="checkbox"/> 1-Influenza A 2-Influenza B Influenza A, qual subtipo? <input type="checkbox"/> <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"/> <input type="checkbox"/> 1-Victoria 2-Yamagata 3-Não realizado 4-Inconclusivo 5-Outro, especifique: _____ Positivo para outros vírus? <input type="checkbox"/> <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"/> <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"/> <input type="checkbox"/> 1-Teste rápido 2-Elisa 3- Quimiluminescência 4- Outro, qual? _____ Resultado do Teste Sorológico para SARS-CoV-2: <input type="checkbox"/> <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"/> <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"/> <input type="checkbox"/> 1- Laboratorial 2- Clínico Epidemiológico 3- Clínico 4- Clínico-Imagem	
74		Evolução do Caso: <input type="checkbox"/> <input type="checkbox"/> 1-Cura 2-Óbito 3-Óbito por outras Causas 9-Ignorado	75	Data da alta ou óbito: _____	76	Data do Encerramento: _____
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	Campo Interno 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	Date DD/MM/AAAA		Data de preenchimento da ficha de notificação.	Campo Obrigatório 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	Campo Interno Calculado a partir da data dos Primeiros Sintomas (SS)	SEM_NOT

				<ul style="list-style-type: none">▪ <u>Unidade Sentinela</u> – o campo é preenchido automaticamente pelo sistema com a Regional do Município onde está localizada a unidade de notificação.▪ <u>Municipal</u> – o campo é preenchido automaticamente pelo sistema com a regional do município do usuário.				
5-Unidade de Saúde Código (CNES)	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.	<p>Campo Obrigatório</p> <p>Preenchendo o nome da unidade, o código é preenchido automaticamente, e vice-versa;</p> <p>Se usuário que está digitando a ficha for de nível:</p> <ul style="list-style-type: none">▪ <u>Unidade Sentinela</u> - o campo é preenchido automaticamente pelo sistema.▪ <u>Municipal</u> – abre tabela a penas com as unidades sentinelas do município.▪ <u>Estadual ou Federal</u> – abre tabela com as unidades sentinelas do município selecionado o campo 4.	ID_UNIDADE OU CO_UNI_NOT			
6-CPF do paciente	Varchar2(15)		Número do Cadastro de Pessoa Física (CPF) do paciente notificado	<p>Campo Essencial</p> <p>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.</p>	NU_CPF			
7-Nome	Varchar2(70)		Nome completo do paciente (sem abreviações)	<p>Campo Obrigatório</p>	NM_PACIENT			
8-Sexo	Varchar2 (1)	1-Masculino 2-Feminino 9-Ignorado	Sexo do paciente.	<p>Campo Obrigatório</p>	CS_SEXO			
9-Data de nascimento	Date DD/MM/AAAA		Data de nascimento do paciente.	<p>Campo Essencial</p> <p>Data deve ser <= a data dos primeiros sintomas.</p>	DT_NASC			
10-(ou) Idade	Varchar2(3)		Idade informada pelo paciente quando não se sabe a data de nascimento. Na falta desse dado é registrada a idade	<p>Campo Obrigatório</p> <p>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>.</p> <p>Idade deve ser <= 150.</p>	NU_IDADE_N			

(ou) Tipo/Idade	Varchar2(1)		aparente.	Campo Obrigatório	TP_IDADE
	1-Dia 2-Mês 3-Ano			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> . 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. 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. 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.	
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.	Campo Obrigatório Se selecionado categoria 2-Feminino no campo Sexo. 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. 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.	CS_GESTANT
12-Raça/Cor	Varchar2(2)	1-Branca 2-Preta 3-Amarela 4-Parada 5-Indígena 9-Ignorado	Cor ou raça declarada pelo paciente: Branca; Preta; Amarela; Parada (pessoa que se declarou mulata, cabocla, cafuza, mameluca ou mestiça de preto com pessoa de outra cor ou raça); e, Indígena.	Campo Obrigatório	CS_RACA
13-Se indígena, qual etnia?	Varchar2(4)	Tabela do SIA/SI com código e nomes das etnias indígenas.	Nome e código da etnia do paciente,	Campo Essencial	CS_ETINIA

				quando indígena.	Habilitado se 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 (1º a 3º 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.	Campo Essencial Preenchido automaticamente 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”.	CS_ESCOL_N
15- Ocupação		Varchar2(6)	Tabela com código da Ocupação da Classificação Brasileira de Ocupações (CBO).	Ocupação profissional do paciente	Campo Essencial	PAC_COCBO ou PAC_DSCBO
16-Nome da mãe		Varchar2(70)		Nome completo da mãe do paciente (sem abreviações).	Campo Essencial	NM_MAE_PAC
17-CEP		Varchar2(8)		CEP de residência do paciente.	Campo Essencial Validado a partir da tabela de CEP dos Correios.	NU_CEP
18-UF		Varchar2(2)	Tabela com código e siglas das UF padronizados pelo IBGE.	Unidade Federativa de residência do paciente.	Campo Obrigatório Se campo 25-País for Brasil. Se preenchido o campo CEP, a UF é preenchida automaticamente pelo sistema e desabilitada para edição.	SG_UF
Regional de Saúde de Residência Código (IBGE)		Varchar2 (6)	Tabela 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.	Campo Interno Preenchido o nome da regional de saúde de residência, o código é preenchido automaticamente, e vice-versa;	ID_RG_RES OU CO_RG_RES
19-Município Código (IBGE)		Varchar2(6)	Tabela com código e nome dos Municípios padronizados pelo IBGE.	Município de residência do paciente.	Campo Obrigatório 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.	ID_MN_RES OU CO_MUN_RES

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

Início dos sintomas?		9- Ignorado	País onde foi realizada a viagem		
28- Se sim: Qual país?	Varchar2(3)	Tabela com código e nome dos Países.		Campo Obrigatório Habilitado se campo 27- Paciente tem histórico de viagem, selecionado categoria 1-Sim em Paciente tem histórico de viagem internacional até 14 dias antes do início dos sintomas?	PAIS_VGM OU CO_PS_VGM
29- Em qual local?	Varchar2(30)		Local (cidade, estado, província e outros) onde foi realizada a viagem	Campo Essencial Habilitado se selecionado categoria 1-Sim em campo 27- Paciente tem histórico de viagem internacional até 14 dias antes do início dos sintomas?	LO_PS_VGM
30- Data da viagem	Date DD/MM/AAAA		Data em que foi realizada a viagem	Campo Obrigatório Se preenchido campo 27- Paciente tem histórico de viagem internacional até 14 dias antes do início dos sintomas? (campo 26), deve ser informada a data de viagem	DT_VGM
31- Data de retorno	Date DD/MM/AAAA		Data em que retornou de viagem	Campo Obrigatório Se preenchido campo 27- Paciente tem histórico de viagem internacional até 14 dias antes do início dos sintomas? (campo 26), deve ser informada a data de retorno	DT_RT_VGM
32- É caso proveniente de surto de SG que evoluiu para SRAG?	Varchar2 (1)	1-Sim 2-Não 9-Ignorado	Caso é proveniente de surto de SG?	Campo Essencial	SURTO_SG
33- Trata-se de caso nosocomial (infecção adquirida no hospital)?	Varchar2 (1)	1-Sim 2-Não 9-Ignorado	Caso de SRAG com infecção adquirida após internação.	Campo Essencial Quando o campo 33 for igual a 1, é permitido digitar data de início dos sintomas posterior a data de internação.	NOSOCOMIAL
34- Paciente trabalha ou tem contato direto com aves, suínos, ou outro animal?	Varchar2(1)	1-Sim 2-Não 9-Ignorado	Caso com contato direto com aves ou suínos.	Campo Essencial	AVE_SUINO
Paciente trabalha ou tem contato direto com aves, suínos/Outro animal (especificar)	Varchar2(60)		Informar o animal que o paciente teve contato se selecionado a opção 3.	Campo Essencial Habilitado de campo 34- Contato com outro animal = 3 (Outro).	OUT_ANIM
35- Sinais e Sintomas/Febre	Varchar2 (1)	1-Sim 2-Não 9-Ignorado	Paciente apresentou febre?	Campo Essencial	FEBRE
35- Sinais e Sintomas/Tosse	Varchar2 (1)	1-Sim 2-Não	Paciente apresentou tosse?	Campo Essencial	TOSSE

35-Sinais e Sintomas/Dor de Garganta	Varchar2(1)	9-Ignorado	Paciente apresentou dor de garganta?	Campo Essencial		GARGANTA
35-Sinais e Sintomas/Dispneia	Varchar2(1)	1-Sim 2-Não 9-Ignorado	Paciente apresentou dispneia?	Campo Essencial		DISPNEIA
35-Sinais e Sintomas/Desconforto Respiratório	Varchar2(1)	1-Sim 2-Não 9-Ignorado	Paciente apresentou desconforto respiratório?	Campo Essencial		DESC_RESP
35-Sinais e Sintomas/Saturação O ₂ < 95%	Varchar2(1)	1-Sim 2-Não 9-Ignorado	Paciente apresentou saturação O ₂ < 95%?	Campo Essencial		SATURACAO
35-Sinais e Sintomas/Diarreia	Varchar2(1)	1-Sim 2-Não 9-Ignorado	Paciente apresentou diarreia?	Campo Essencial		DIARREIA
35-Sinais e Sintomas/Vômito	Varchar2(1)	1-Sim 2-Não 9-Ignorado	Paciente apresentou vômito?	Campo Essencial		VOMITO
35-Sinais e Sintomas/Dor abdominal	Varchar2(1)	1-Sim 2-Não 9-Ignorado	Paciente apresentou dor abdominal?	Campo Essencial		DOR_ABD
35-Sinais e Sintomas/Fadiga	Varchar2(1)	1-Sim 2-Não 9-Ignorado	Paciente apresentou fadiga?	Campo Essencial		FADIGA
35-Sinais e Sintomas/Perda do Olfato	Varchar2(1)	1-Sim 2-Não 9-Ignorado	Paciente apresentou perda do olfato?	Campo Essencial		PERD_OLFT
35-Sinais e Sintomas/Perda do Paladar	Varchar2(1)	1-Sim 2-Não 9-Ignorado	Paciente apresentou perda do paladar?	Campo Essencial		PERD_PALA
35-Sinais e Sintomas/Outros	Varchar2(1)	1-Sim 2-Não 9-Ignorado	Paciente apresentou outro(s) sintoma(s)?	Campo Essencial		OUTRO_SIN
35-Sinais e Sintomas/Outros (Descrição)	Varchar2(30)		Listar outros sinais e sintomas	Campo Essencial		OUTRO_DES

				apresentados pelo paciente.	Habilitado e 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	Campo Essencial	FATOR_RISC
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)?	Campo Essencial Habilitado e selecionado no campo 8- Sexo Feminino.	PUERPERA
36-Fatores de risco/ Doença Cardiovascular Crônica	Varchar2 (1)	1-Sim 2-Não 9-Ignorado		Paciente possui Doença Cardiovascular Crônica ?	Campo Essencial	CARDIOPATI
36-Fatores de risco/ Doença Hematológica Crônica	Varchar2 (1)	1-Sim 2-Não 9-Ignorado		Paciente possui Doença Hematológica Crônica ?	Campo Essencial	HEMATOLOGI
36-Fatores de risco/ Síndrome de Down	Varchar2 (1)	1-Sim 2-Não 9-Ignorado		Paciente possui Síndrome de Down?	Campo Essencial	SIND_DOWN
36-Fatores de risco/ Doença Hepática Crônica	Varchar2 (1)	1-Sim 2-Não 9-Ignorado		Paciente possui Doença Hepática Crônica ?	Campo Essencial	HEPATICA
36-Fatores de risco/ Asma	Varchar2 (1)	1-Sim 2-Não 9-Ignorado		Paciente possui Asma ?	Campo Essencial	ASMA
36-Fatores de risco/ Diabetes mellitus	Varchar2 (1)	1-Sim 2-Não 9-Ignorado		Paciente possui Diabetes mellitus?	Campo Essencial	DIABETES
36-Fatores de risco/ Doença Neurológica Crônica	Varchar2 (1)	1-Sim 2-Não 9-Ignorado		Paciente possui Doença Neurológica?	Campo Essencial	NEUROLOGIC
36-Fatores de risco/ Outra Pneumopatia Crônica	Varchar2 (1)	1-Sim 2-Não 9-Ignorado		Paciente possui outra pneumopatia crônica?	Campo Essencial	PNEUMOPATI
36-Fatores de risco/ Imunodeficiência	Varchar2 (1)	1-Sim		Paciente possui	Campo Essencial	IMUNODEPRE

ou Imunodepressão		2-Não 9-Ignorado	Imunodeficiência ou Imunodepressão (diminuição da função do sistema imunológico)?		
36-Fatores de risco/ Doença Renal Crônica	Varchar2(1)	1-Sim 2-Não 9-Ignorado	Paciente possui Doença Renal Crônica?	Campo Essencial	RENAL
36-Fatores de risco/ Obesidade	Varchar2(1)	1-Sim 2-Não 9-Ignorado	Paciente possui obesidade?	Campo Essencial	OBESIDADE
36-Fatores de risco/ Obesidade (Descrição IMC)	Varchar2(3)		Valor do IMC (Índice de Massa Corporal) do paciente calculado pelo profissional de saúde.	Campo Essencial Habilitado se selecionado categoria 1-Sim em Fatores de risco/Obesidade.	OBES_IMC
36-Fatores de risco/ Outros	Varchar2(1)	1-Sim 2-Não 9-Ignorado	Paciente possui outro(s) fator(es) de risco?	Campo Essencial	OUT_MORBI
36-Fatores de risco/ Outros (Descrição)	Varchar2(30)		Listar outro(s) fator(es) de risco do paciente.	Campo Essencial Habilitado se selecionado categoria 1-Sim em Fatores de risco/Outros.	MORB_DESC
37-Recebeu vacina contra Gripe na última campanha?	Varchar2(1)	1-Sim 2-Não 9-Ignorado	Informar se o paciente foi vacinado contra gripe na última campanha, após verificar a documentação / cade meta. Caso o paciente não tenha a cade meta, direcionar a pergunta para ele ou responsável e preencher o campo com o código correspondente a	Campo Essencial	VACINA

				para tratamento da doença?			
40- Qual antiviral?	Varchar2 (1)	2- Não 9- Ignorado	Qual antiviral utilizado?	Campo Essencial Habilitado se 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.	Campo Essencial Habilitado se 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.	Campo Essencial Habilitado se 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?	Campo Essencial 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.	Campo Obrigatório 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 padronizados pelo IBGE.	Unidade Federativa de internação do paciente.	Campo Essencial Habilitado se 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 padronizados pelo IBGE.	Regional de Saúde onde está localizado o Município de internação do paciente.	Campo Interno 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.	Campo Essencial Habilitado se campo 42- Houve internação? for igual a 1		ID_MN_INTE OU CO_MU_INTE	

46- Unidade de Saúde de Internação/ Código CNES	Varchar2(20)	Tabela com códigos CNES e nomes das Unidades de Saúde.	Unidade Sentinela que realizou a internação do paciente.	Campo Essencial Habilitado se campo 42- Houve internação? for igual a 1	ID_UN_INTE OU CO_UN_INTE
47- Internado em UTI?	Varchar2(1)	1-Sim 2-Não 9-Ignorado	O paciente foi internado em UTI?	Campo Essencial	UTI
48- Data da entrada na UTI	Date DD/MM/AAAA		Data de entrada do paciente na unidade de Terapia intensiva (UTI).	Campo Essencial Habilitado se campo 47- Internado em UTI? for igual a 1. Data deve ser maior ou igual a 2- Data de 1ºs sintomas da SRAG e menor ou igual a data da digitação (atual).	DT_ENTUTI
49- Data da saída da UTI	Date DD/MM/AAAA		Data em que o paciente saiu da Unidade de Terapia intensiva (UTI).	Campo Essencial Habilitado se campo 47- Internado em UTI? for igual a 1. Data deve ser maior ou igual a 48- Data da entrada na UTI e menor ou igual a data da digitação (atual).	DT_SAIDUTI
50- Uso de suporte ventilatório?	Varchar2(1)	1-Sim, invasivo 2-Sim, não invasivo 3-Não 9-Ignorado	O paciente fez uso de suporte ventilatório?	Campo Essencial	SUPORT_VEN
51- Raio X de Tórax	Varchar2(1)	1-Normal 2-Infiltrado intersticial 3-Consolidação 4-Misto 5-Outro 6-Não realizado 9-Ignorado	Informar resultado de Raio X de Tórax.	Campo Essencial	RAIOX_RES
Raio X de Tórax/ Outro (especificar)	Varchar2(30)		Informar o resultado do RX de tórax e selecionado a opção 5-Outro.	Campo Essencial Habilitado se campo 51- Raio X de Tórax = 5 (Outro).	RAIOX_OUT
52- Data do Raio X	Date DD/MM/AAAA		Se realizou RX de Tórax, especificar a data do exame.	Campo Essencial Habilitado se campo 51- Raio X de Tórax for igual a 1, 2, 3, 4 ou 5.	DT_RAIOX
53- Aspecto Tomografia	Number(3)	1-Típico COVID-19 2- Indeterminado COVID-19	Informar o resultado da	Campo Essencial	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 selecionado a opção 5-Outro	Campo Essencial Habilitado 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.	Campo Essencial Habilitado 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?	Campo Essencial	AMOSTRA
56- Data da Coleta	Data DD/MM/AAAA			Data da coleta da amostra para realização do teste diagnóstico.	Campo Essencial Habilitado 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 orofaríngea 2- Lavado Bronco-alveolar 3- Tecido post-mortem 4- Outra, qual? 5- LCR 9- Ignorado		Tipo da amostra clínica coletada para o teste diagnóstico.	Campo Essencial Habilitado 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.	Campo Essencial Campo habilitado e selecionado categoria 4-Outra, qual em Tipo de amostra.	OUT_AMOST
58- Nº da Requisição do GAL				Número da requisição de exames gerado	Campo Essencial	REQUISI_GAL

59- Tipo do Teste antígeno	Number(3)		1-Imunofluorescência (IF) 2- Teste rápido antígeno	pe lo sistema GAL.		Campo Essencial		TP_TES_AN
60- Data do resultado teste Antígeno	Data DD/MM/AAAA			Data do resultado do teste a antígeno.		Campo Essencial Data deve ser maior ou igual a 56- Data da Coleta		DT_RES_AN
61- Resultado do Teste Antígeno	Varchar2(1)		1-positivo 2-Negativo 3- Inconclusivo 4-Não realizado 5-Aguardando resultado 9-Ignorado	Resultado do Teste Antígeno		Campo Essencial Este campo virá marcado com 5-Aguardando Resultado e estará habilitado se o campo 55- Coletou amostra? = 1		RES_AN
62-Laboratório que realizou o Teste antígeno	Varchar2(70)		Nomes dos laboratórios cadastrados no sistema	Laboratório responsável pela liberação do resultado do teste antígeno.		Campo Essencial Habilitado se selecionado categoria 1-Positivo, 2-Negativo ou 3-Inconclusivo preenchendo o nome do laboratório, o código é preenchido automaticamente, ou vice-versa		LAB_AN
62-Laboratório que realizou o Teste antígeno	Varchar2(7)		Tabela com códigos CNES			Campo Essencial		CO_LAB_AN
63-Agente etiológico – Teste Antígeno. Positivo para Influenza?	Varchar2(1)		1-Sim 2-Não 9-Ignorado	Resultado do Teste Antígeno que foi positivo para Influenza		Campo Essencial		POS_AN_FLU
63-Agente etiológico – Teste Antígeno. Se sim, qual Influenza?	Varchar2(1)		1-Influenza A 2-Influenza B	Resultado do Teste Antígeno, para o tipo de Influenza.		Campo Essencial Habilitado se campo 63-Agente etiológico – Teste Antígeno: Positivo para Influenza? = 1.		TP_FLU_AN
63-Agente etiológico – Teste Antígeno. Positivo para outros vírus?	Varchar2 (1)		1-Sim 2-Não 9-Ignorado	Resultado do Teste Antígeno, que foi positivo para outro vírus respiratório.		Campo Essencial		POS_AN_OUT
63-Agente etiológico – Teste Antígeno. SARS-CoV-2	Varchar2(1)		1-marcado pelo usuário Vazio - não marcado	Resultado do Teste Antígeno, para SARS-CoV-2.		Campo Essencial Habilitado se campo 63-Agente etiológico, Teste Antígeno. Positivo para outros vírus? = 1.		AN_SARS2
63-Agente etiológico – Teste Antígeno.	Varchar2(1)		1-marcado pelo usuário Vazio - não marcado	Resultado do Teste Antígeno, para		Campo Essencial		AN_VSR

VSR			VSR.	Habilitado se campo 63-Agente etiológico, Teste Antigenico. Positivo para outros virus? = 1.	
63-Agente etiológico – Teste Antigenico. Parainfluenza 1	Varchar2(1)	1-marca do pelo usuário Vazio - não marcado	Resultado do Teste Antigenico, para Parainfluenza 1.	Campo Essencial Habilitado se campo 63-Agente etiológico, Teste Antigenico. Positivo para outros virus? = 1.	AN_PARA1
63-Agente etiológico – Teste Antigenico. Parainfluenza 2	Varchar2(1)	1-marca do pelo usuário Vazio - não marcado	Resultado do Teste Antigenico, para Parainfluenza 2.	Campo Essencial Habilitado se campo 63-Agente etiológico, Teste Antigenico. Positivo para outros virus? = 1.	AN_PARA2
63-Agente etiológico – Teste Antigenico. Parainfluenza 3	Varchar2(1)	1-marca do pelo usuário Vazio - não marcado	Resultado do Teste Antigenico, para Parainfluenza 3.	Campo Essencial Habilitado se campo 63-Agente etiológico, Teste Antigenico. Positivo para outros virus? = 1.	AN_PARA3
63-Agente etiológico – Teste Antigenico. Adenovirus	Varchar2(1)	1-marca do pelo usuário Vazio - não marcado	Resultado do Teste Antigenico, Adenovirus.	Campo Essencial Habilitado se campo 63-Agente etiológico, Teste Antigenico. Positivo para outros virus? = 1.	AN_ADENO
63- Agente etiológico – Teste Antigenico. Outro virus respiratório	Varchar2(1)	1-marca do pelo usuário Vazio - não marcado	Resultado do Teste Antigenico. Outro virus respiratório.	Campo Essencial Habilitado se campo 63-Agente etiológico, Teste Antigenico. Positivo para outros virus? = 1.	AN_OUTRO
63- Agente etiológico – Teste Antigenico. Outro virus respiratório (Descrição)	Varchar2(30)		Nome do outro virus respiratório identificado pelo Teste Antigenico.	Campo Essencial Habilitado se campo 63-Agente etiológico, Teste Antigenico. Positivo para outros virus? = 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.	Campo Essencial 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	Date DD/MM/AAAA		Data do Resultado RT-PCR/outro método por Biologia Molecular	Campo Essencial 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	Campo Essencial	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.	Campo Essencial Habilitado 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) pd m09 2-Influenza A (H3N2) 3-Influenza A não subtipado 4-Influenza A não subtipável 5-In conclusivo 6-Outro, especifique:	Subtipo para Influenza A.	Campo Essencial Habilitado 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.	Campo Essencial Habilitado 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-In conclusivo 5-Outro, especifique:	Linhagem para Influenza B.	Campo Essencial Habilitado se campo 66-Agente etiológico – RT_PCR/outr método por Biologia Molecular: Se sim, qual Influenza? = 2.	PCR_FLUBU
66- Agente etiológico – RT-PCR/outr método por Biologia Molecular: Se Influenza B, qual linhagem? Outro, especifique:	Varchar2 (30)		Outra linhagem para Influenza B.	Campo Essencial Habilitado se 66- Agente etiológico – RT-PCR/outr método por Biologia Molecular: Se Influenza B, qual linhagem? = 5.	FLUBU_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	Campo Essencial	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).	Campo Essencial Habilitado 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).	Campo Essencial Habilitado se campo 66- Agente etiológico – RT-PCR/outr método por	PCR_VSR

						Biologia Molecular: Positivo para outros vírus? = 1	
66- Agente etiológico – RT-PCR/ou método por Biologia Molecular: Parainfluenza 1	Varchar2 (1)	1-marca do pelo usuário Vazio - não marcado	Resultado diagnóstico do RT- PCR para Parainfluenza 1.	Campo Essencial Habilitado se campo 66- Agente etiológico – RT-PCR/ou método por Biologia Molecular: Positivo para outros vírus? = 1	PCR_PARA1		
66- Agente etiológico – RT-PCR/ou método por Biologia Molecular: Parainfluenza 2	Varchar2 (1)	1-marca do pelo usuário Vazio - não marcado	Resultado diagnóstico do RT- PCR para Parainfluenza 2.	Campo Essencial Habilitado se campo 66- Agente etiológico – RT-PCR/ou método por Biologia Molecular: Positivo para outros vírus? = 1	PCR_PARA2		
66- Agente etiológico – RT-PCR/ou método por Biologia Molecular: Parainfluenza 3	Varchar2 (1)	1-marca do pelo usuário Vazio - não marcado	Resultado diagnóstico do RT- PCR para Parainfluenza 3.	Campo Essencial Habilitado se campo 66- Agente etiológico – RT-PCR/ou método por Biologia Molecular: Positivo para outros vírus? = 1	PCR_PARA3		
66- Agente etiológico – RT-PCR/ou método por Biologia Molecular: Parainfluenza 4	Varchar2 (1)	1-marca do pelo usuário Vazio - não marcado	Resultado diagnóstico do RT- PCR para Parainfluenza 4.	Campo Essencial Habilitado se campo 66- Agente etiológico – RT-PCR/ou método por Biologia Molecular: Positivo para outros vírus? = 1	PCR_PARA4		
66- Agente etiológico – RT-PCR/ou método por Biologia Molecular: Adenovírus	Varchar2 (1)	1-marca do pelo usuário Vazio - não marcado	Resultado diagnóstico do RT- PCR para Adenovírus.	Campo Essencial Habilitado se campo 66- Agente etiológico – RT-PCR/ou método por Biologia Molecular: Positivo para outros vírus? = 1	PCR_ADENO		
66- Agente etiológico – RT-PCR/ou método por Biologia Molecular: Metapneumovírus	Varchar2 (1)	1-marca do pelo usuário Vazio - não marcado	Resultado diagnóstico do RT- PCR para Metapneumovírus.	Campo Essencial Habilitado se campo 66- Agente etiológico – RT-PCR/ou método por Biologia Molecular: Positivo para outros vírus? = 1	PCR_METAP		
66- Agente etiológico – RT-PCR/ou método por Biologia Molecular: Bocavírus	Varchar2 (1)	1-marca do pelo usuário Vazio - não marcado	Resultado diagnóstico do RT- PCR para Bocavírus.	Campo Essencial Habilitado se campo 66- Agente etiológico – RT-PCR/ou método por Biologia Molecular: Positivo para outros vírus? = 1	PCR_BOCA		
66- Agente etiológico – RT-PCR/ou método por Biologia Molecular: Rinovírus	Varchar2 (1)	1-marca do pelo usuário Vazio - não marcado	Resultado diagnóstico do RT- PCR para Rinovírus.	Campo Essencial Habilitado se campo 66- Agente etiológico – RT-PCR/ou método por Biologia Molecular: Positivo para outros vírus? = 1	PCR_RINO		

							Biologia Molecular: Positivo para outros vírus? = 1	
66- Agente etiológico – RT-PCR/ou outro método por Biologia Molecular: Outro vírus respiratório, especifique:	Varchar2 (1)	1-marca do pelo usuário Vazio - não marcado	Resultado diagnóstico do RT- PCR para Outro vírus respiratório.	Campo Essencial		PCR_OUTRO		
66- Agente etiológico – RT-PCR/ou outro método por Biologia Molecular: Outro vírus respiratório (Descrição)	Varchar2 (30)		Nome do outro vírus respiratório identificado pelo RT-PCR.	Campo Essencial		DS_PCR_OUT		
67-Laboratório que realizou RT- PCR/ou método por Biologia Molecular Código (CNES)	Varchar2 (7)	Tabela com códigos CNES e nomes dos laboratórios cadastrados no sistema.	Laboratório responsável pela liberação do resultado do teste diagnóstico (RT- PCR) da amostra do paciente.	Campo Essencial		LAB_PCR OU CO_LAB_PCR		
68- Tipo de Amostra Sorológica para SARS-Cov-2	Number(3)	1- Sangue/plasma/soro 2-Outra, qual? 9-Ignorado	Tipo de amostra sorológica que foi coletada.	Campo Essencial		TP_AM_SOR		
Tipo de Amostra Sorológica para SARS- Cov-2/Outra, qual?			Descrição do tipo da amostra clínica, caso diferente das listadas na categoria um (1) do campo.	Campo Essencial		SOR_OUT		
69- Data da coleta	Data DD/MM/AAAA		Data da coleta do material para diagnóstico por Sorologia.	Campo Essencial		DT_CO_SOR		
70- Tipo de Sorologia para SARS-Cov-2	Number(3)	1-Teste rápido 2-Elisa 3- Quimioluminescência 4-Outro, qual	Tipo do Teste Sorológico que foi realizado	Campo Essencial		TP_SOR		

70- Tipo de Sorologia para SARS-Cov-2	Varchar2(100)		Descrição do tipo de Teste Sorológico	Campo Essencial Campo habilitado se selecionado categoria 4-Outro, qual?	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	Campo Essencial	RES_SOR
70- Resultado do Teste Sorológico para SARS-Cov-2:	Varchar2(1)	IgG	Resultado da Sorologia para SARS-Cov-2	Campo Essencial	RES_IGG
70- Resultado do Teste Sorológico para SARS-Cov-2:	Varchar2(1)	IgM	Resultado da Sorologia para SARS-Cov-2	Campo Essencial	RES_IGM
70- Resultado do Teste Sorológico para SARS-Cov-2:	Varchar2(1)	IgA	Resultado da Sorologia para SARS-Cov-2	Campo Essencial	RES_IGA
71- Data do Resultado	Date DD/MM/AAAA		Data do Resultado do Teste Sorológico	Campo Essencial 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.	Campo Obrigatório	CLASSI_FIN
72-Classificação final do caso 3-SRAG por outra causa , qual:	Varchar2(30)		Descrição de qual outro agente etiológico foi identificado	Campo Obrigatório Se campo 72-Classificação final do caso = 3.	CLASSI_OUT

73—Critério de Encerramento	Varchar2(1)	1. Laboratorial 2. Clínico Epidemiológico 3. Clínico 4. Clínico Imagem	Indicar qual o critério de confirmação.	Habilitado se campo 72- Classificação final do caso = 3. Campo Essencial	CRITERIO
74—Evolução do caso	Varchar2(1)	1-Cura 2-Óbito 3-Óbito por outras causas 9-Ignorado	Evolução do caso	Campo Essencial	EVOLUCAO
75—Data da alta ou óbito	Date DD/MM/AAAA		Data da alta ou óbito	Campo Essencial Data da alta ou do óbito deve ser > ou = a data dos primeiros sintomas e <= a data da digitação (atual). Habilitado se campo 74- Evolução do caso = 1 ou 2.	DT_EVOLUCA
76-Data do Encerramento	Date DD/MM/AAAA		Data do encerramento do caso.	Campo Obrigatório 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).	DT_ENCERRA
77- Número D.O			Número da Declaração de Óbito	Campo Essencial Habilitado se o campo 74- Evolução do caso = 2	NU_DO
78-Observações	Varchar2(999)		Outras observações sobre o paciente consideradas pertinentes.	Campo Opcional	OBSERVA
79-Profissional de Saúde Responsável	Varchar2(60)		Nome completo do profissional de saúde (sem abreviações) responsável pela notificação.	Campo Essencial	NOME_PROF
80-Registro Conselho/Matrícula	Varchar2(15)		Número do conselho ou matrícula do	Campo Essencial	REG_PROF

				profissional de saúde responsável pela notificação (Ex: CRM/RJ 1234)			
Data da digitação		Date DD/MM/AAAA		Data de inclusão do registro no sistema.	Campo Interno 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. Não é atualizada se houver alterações posteriores de dados.		DT_DIGITA