

**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, PROVENIENTES DO SARS-CoV-2, NO BRASIL,  
EM 2020, PARA SUBSIDIAR A TOMADA DE DECISÃO  
E O PLANEJAMENTO NA ÁREA DE SAÚDE**

**Belo Horizonte**  
**Fevereiro de 2021**

**Fábio Roberto Barros Padilha**

**ANÁLISE DE DADOS NACIONAIS SOBRE A EVOLUÇÃO DAS SÍNDROMES  
RESPIRATÓRIAS AGUDAS GRAVES, PROVENIENTES DO SARS-CoV-2, NO BRASIL,  
EM 2020, PARA SUBSIDIAR A TOMADA DE DECISÃO  
E O PLANEJAMENTO NA ÁREA DE SAÚDE**

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  
Fevereiro de 2021**

# Sumário

<b>1. Introdução</b>	<b>2</b>
1.1 Contextualização . . . . .	3
1.2 Objetivo . . . . .	4
<b>2. Coleta dos dados</b>	<b>5</b>
2.1 Dados OpenDataSUS . . . . .	5
2.1 Dados do Instituto Brasileiro de Geografia e Estatística . . . . .	7
<b>3. Processamento</b>	<b>8</b>
3.1 Dados do Instituto Brasileiro de Geografia e Estatística . . . . .	8
3.2 Dados do Sistema Único de Saúde . . . . .	9
3.3 Transformação . . . . .	9
3.3.1 Notificações . . . . .	10
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 . . . . .	17
<b>4. Análise Exploratória</b>	<b>18</b>
4.1 Número absoluto de mortes de covid-19 no Brasil . . . . .	18
4.2 Taxa de mortalidade da covid-19 no Brasil . . . . .	21
4.3 Taxa de letalidade da covid-19 no Brasil . . . . .	22
<b>5. Inferência Estatística</b>	<b>25</b>
5.1 Perfil do brasileiro com maior e menor risco de morte . . . . .	25
<b>6. Machine learning</b>	<b>29</b>
6.1 Árvore de Decisão de Classificação . . . . .	29
6.2 Árvore de Decisão e Seleção de Atributos . . . . .	29
6.3 Árvore de Decisão de Classificação otimizado . . . . .	30
6.4 Naive Bayes . . . . .	31
<b>Referência bibliográfica</b>	<b>32</b>
<b>Apêndice A - Taxa de Mortalidade Estados e Capitais</b>	<b>33</b>
<b>Anexo A - Ficha de Notificação</b>	<b>34</b>
<b>Anexo B - Dicionário de dados</b>	<b>36</b>

# 1. Introdução

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

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

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

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

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

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

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) com ênfase da 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. De modo, a subsidiar a implementação de políticas de saúde pública. A exemplo, a 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, o 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 e de um plano de vacinação nacional.

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 corona vírus.

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)[6]

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 da página 3.

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 Objetivo

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 médico 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 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 dos números de óbitos e contenção do novo coronavírus. Desta forma, têm-se os seguintes objetivos específicos:

- identificar as regiões brasileiras com maiores número de óbitos;
- identificar as taxas de mortalidade e de letalidade do: país, estados e capitais;
- encontrar os perfis dos pacientes com maiores riscos de morte.

## 2. Coleta dos dados

### 2.1 Dados OpenDataSUS

Dentre os dados utilizados neste trabalho, destacam-se os dados fornecidos pelo Ministério da Saúde por meio da Secretaria de Vigilância em Saúde no portal OpenDataSUS. Sobre os dados coletados informa-se no portal OpenDataSUS:

O Ministério da Saúde (MS), por meio da Secretaria de Vigilância em Saúde (SVS), desenvolve a vigilância da Síndrome Respiratória Aguda Grave (SRAG) no Brasil, desde a pandemia de Influenza A(H1N1)pdm09 em 2009. A partir disso, esta vigilância foi implantada na rede de Influenza e outros vírus respiratórios, que anteriormente atuava apenas com a vigilância sentinela de Síndrome Gripal (SG). Recentemente (2020), a vigilância da COVID-19, a infecção humana causada pelo novo Coronavírus, foi incorporada na rede de vigilância da Influenza e outros vírus respiratórios.

A Secretaria de Vigilância disponibiliza o arquivo para [download em formato CSV](#). Esse arquivo é atualizado semanalmente, no momento do download o arquivo disponibilizado continha dados até 04/01/2021.

Como a Secretaria de Vigilância em Saúde atualiza o arquivo CSV semanalmente, para garantir o acesso ao arquivo exato, analisado neste trabalho, fez-se necessário manter uma cópia do arquivo no Google drive. O arquivo pode ser acessado no link: [https://drive.google.com/drive/folders/1l52SO-RMHYfs2ci\\_2pBAZJNw2hA5i2me?usp=sharing](https://drive.google.com/drive/folders/1l52SO-RMHYfs2ci_2pBAZJNw2hA5i2me?usp=sharing)

O arquivo contém 1.103.183 (um milhão cento e três mil e cento e oitenta e três) observações e 154 variáveis. 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 na seção anexo.

```
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, bem como renomeadas a fim de facilitar a identificação nas seções seguintes.

```
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_PCRROUT", "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")]
```

```
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", "INDIARREIA",
"IN_VOMITO", "IN_DOR_ABDOMINAL", "IN_FADIGA", "IN_PERDA_OLFATO", "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_SUORTE_VENTILATORIO",
"CS_RAIOX_RESULTADO", "CS_TOMOGRFIA_RESULTADO", "IN_COLETOU_AMOSTRA",
"CS_TESTE_ANTIGENO", "CS_RESULTADO_ANTIGENO", "IN_POSITIVO_INFLUENZA",
"CS_TIPO_INFLUENZA", "IN_POSITIVO_OUT_VIRUS", "IN_POSITIVO_SARS2",
"IN_POSITIVO_VSR", "IN_POSITIVO_PARAINFLUENZA1", "IN_POSITIVO_PARAINFLUENZA2",
"IN_POSITIVO_PARAINFLUENZA3", "IN_POSITIVO_ADENOVIRUS", "CS_PCR_RESULTADO",
"IN_POSITIVO_PCR_INFLUENZA", "CS_TIPO_INFLUENZA_PCR", "CS_PCR_INFLUENZA_A_SUBTIPO",
"CS_PCR_FLU_B_LINHAGEM", "IN_PCR_POSITIVO_OUT_VIRUS", "IN_PCR_POSITIVO_SARS2",
"IN_PCR_POSITIVO_VSR", "IN_PCR_POSITIVO_PARAINFLUENZA1",
"IN_PCR_POSITIVO_PARAINFLUENZA2", "IN_PCR_POSITIVO_PARAINFLUENZA3",
"IN_PCR_POSITIVO_PARAINFLUENZA4", "IN_PCR_POSITIVO_ADENOVIRUS",
"IN_PCR_POSITIVO_METAPNEUMOVIRUS", "IN_PCR_POSITIVO_BOCAVIRUS",
"IN_PCR_POSITIVO_RINOVIRUS", "CS_TIPO_AMOSTRA_SOROLOGICA_SARS_COV2",
"CS_TESTE_SOROLOGICO", "IN_SOROLOGICO_POSITIVO_IGG", "IN_SOROLOGICO_POSITIVO_IGM",
"IN_RESULTADO_POSITIVO_IGA", "CS_CLASSICACAO_FINAL", "CS_CRITERIO_CONFIRMACAO",
"CS_EVOLUCAO", "DT_EVOLUCAO")
```



## 2.1 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.html>, no diretório: Estimativa\_de\_Populacao/Estimativas\_2020, em formato xls. Conforme figura 1 da página 7. Após o download, o arquivo foi convertido para o formato xlsx, no LibreOffice, para facilitar a importação para o 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, dado que a população por estado ou por região pode ser facilmente obtida a partir do somatório populacional dos municípios.

A importação iniciou-se, a partir da terceira linha do arquivo, a fim de desconsiderar as linhas referentes ao cabeçalho.

```
arquivoIBGE=read.xlsx("dados/POP2020_20201030.xlsx", sheet = "Municípios",  
startRow = 3, colNames =FALSE, rowNames = FALSE)  
#Download de conjuntos de dados espaciais oficiais do Brasil.  
municipios <- read_municipality(year=2019)
```

```
## Using year 2019
```

```
## Loading data for the whole country. This might take a few minutes.
```

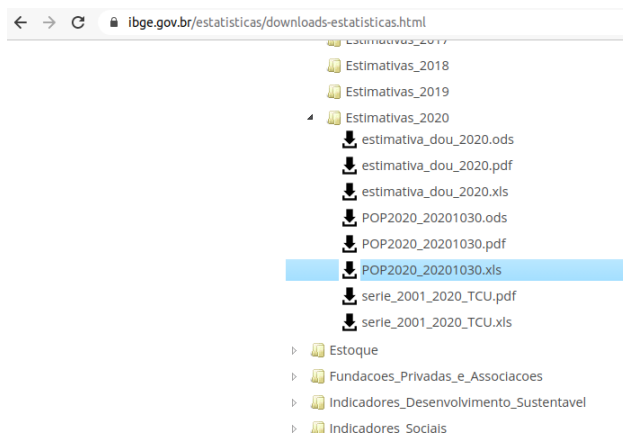


Figura 1: Arquivo IBGE

### 3. Processamento

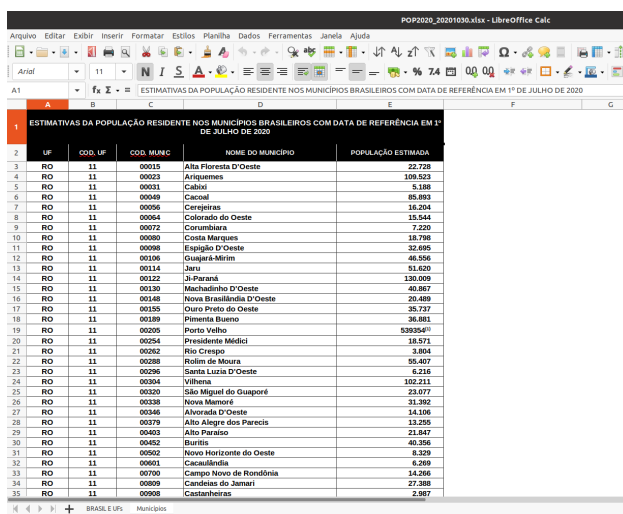
Neste capítulo, contempla-se a maioria das transformações as quais foram submetidos os arquivos CSV disponibilizado no portal OpenDataSUS e também o arquivo xls disponibilizado no portal do IBGE. Também foi removido as linhas dos data frames resultante da importação dos arquivos que não são objeto dessa análise.

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

A planilha disponibilizada pelo IBGE contém cabeçalho e rodapé com dados que não são relevantes para este trabalho. Os dados populacionais necessários estão presentes da linha 3 até a linha 5572. Veja figura 2 e a figura 3 na página 8.

A importação dos dados foi iniciada a partir da linha três do arquivo, descartando o cabeçalho. Para descartar o rodapé, o data frame resultante da importação foi selecionado somente até a linha 5570. Por fim, foi criando 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 a 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.



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	00040	Caculé	85.853
RO	11	00056	Cerejeiras	16.204
RO	11	00064	Colorado do Oeste	15.544
RO	11	00072	Corumbeira	7.220
RO	11	00080	Costa Marques	18.798
RO	11	00090	Epigênia D'Oeste	32.695
RO	11	00106	Guaçu-Mirim	46.556
RO	11	00114	Jaru	51.620
RO	11	00122	Ji-Paraná	126.009
RO	11	00130	Machadinho D'Oeste	40.867
RO	11	00148	Nova 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	539.547
RO	11	00241	Presidente Médici	18.571
RO	11	00262	Rio Crespo	3.804
RO	11	00288	Rolim de Moura	95.407
RO	11	00296	Santa Luzia D'Oeste	6.216
RO	11	00304	Vilhena	102.211
RO	11	00320	São Miguel do Guaaporé	23.677
RO	11	00338	Nova Memoré	21.362
RO	11	00346	Alvorada D'Oeste	14.106
RO	11	00378	Alto Alegre dos Parecis	12.256
RO	11	00403	Alto Paraíso	21.847
RO	11	00452	Buritis	40.356
RO	11	00502	Novo Horizonte do Oeste	6.329
RO	11	00601	Cacaulândia	6.269
RO	11	00700	Campo Novo de Rondônia	14.266
RO	11	00809	Candonga do Jamari	27.388
RO	11	00908	Castanheiras	2.987

Figura 2: Arquivo IBGE

A1	B	C	D	E	F
5569	GO	52	22054	Vicentópolis	8.873
5570	GO	52	22203	Vila Boa	6.312
5571	GO	52	22302	Vila Príncipe	5.862
5572	DF	53	00108	Brasília	3.095.149
5573					
5574	Fonte: IBGE, Diretoria de Pesquisas - DPE - Coordenação de População e Indicadores Sociais - COPIS.				
5575	Notas:				
5576					
5577	(1) População judicial do município de Porto Velho-RO: 494.013 habitantes. Processo Judicial nº 12336-40.2016.4.01.4200 - Seção Judiciária de Rondônia.				
5578	(2) População judicial do município de Manauçu-AM: entre 30.565 e 27.356 habitantes. Parecer de Força Executória nº 0001302017/MACOB-GEAC/PPFAMPQF/AGU, em trâmite na 3ª VCIAM.				
5579	(3) População judicial do município de Santa Isabel do Rio Negro-AM: entre 23.773 e 30.564 habitantes. Parecer de Força Executória nº 000072017/MACOB-GEAC/PPFAMPQF/AGU, em trâmite na 3ª VCIAM.				

Figura 3: Rodapé arquivo IBGE

Nas seções seguintes serão criados gráficos sumarizados pelas regiões do país. Entretanto, a planilha com os dados populacionais do IBGE e o CSV disponibilizado no portal OpenDataSUS, não consta essa informação. Veja figura 2 na página 8.

Assim, fez-se necessário 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.

```
#Essas regiões são utilizadas nos gráficos.
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 Dados do Sistema Único de Saúde

A transformação dos dados do Sistema Único do Saúde - SUS 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. Na segunda etapa outras transformações aconteceram com o propósito de melhor ajustar os dados para as execuções dos algoritmos de machine learning.

Os campos tratados na seção 3 (três) foram selecionados a partir das colunas utilizadas nos algoritmos e gráficos das seções: 4, 5 e 6.

## 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. Veja figura 4 na página 10.

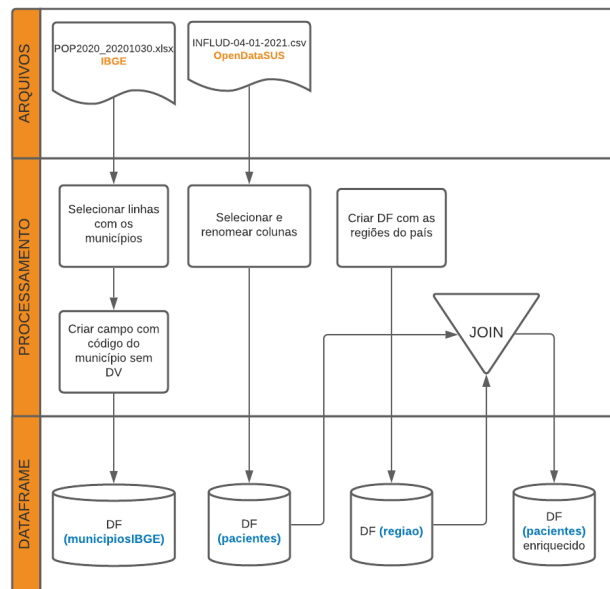


Figura 4: Coleta dos dados

### 3.3.1 Notificações

O campo DT\_NOTIFICACAO contém a data do preenchimento da ficha de notificação, esse campo foi ajustado para um campo do tipo data, formato dd/mm/aaaa. Depois de converter o campo, os registros com data de notificação superior a 31/12/2020 foram removidos do data frame, visto que o escopo desse trabalho é o ano de 2020.

Nessa seção também foram removidos do data frame as notificações que não foram classificadas como casos de covid-19.

Não há valores omissos nessa coluna.

```
#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"),]

pacientes=pacientes[complete.cases(pacientes$CS_CLASSICACAO_FINAL) &
pacientes$CS_CLASSICACAO_FINAL==5,]
```

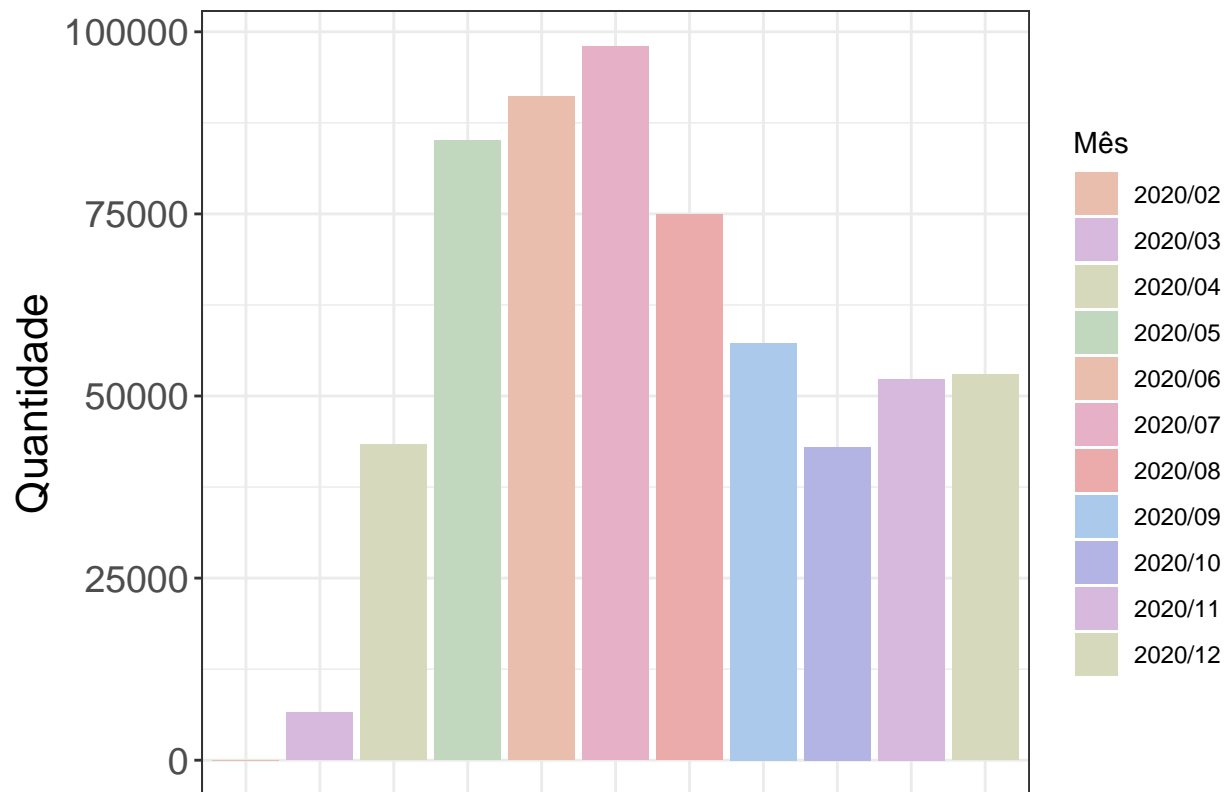


Gráfico 1: Notificações de Covid-19 por Mês

No Brasil, em 2020 os primeiros casos de covid-19 foram registrados em fevereiro e teve o maior número de casos em julho.

### 3.3.2 Idade

O processamento da coluna idade, foi atribuído a média da idade nacional, 59 (cinquenta e nove) anos nas ocorrências com valores superiores a 110 anos. Essa transformação teve como objetivo minimizar os outliers e os prováveis erros de digitação.

Não há valores omissos nessa coluna.

```
#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))
```

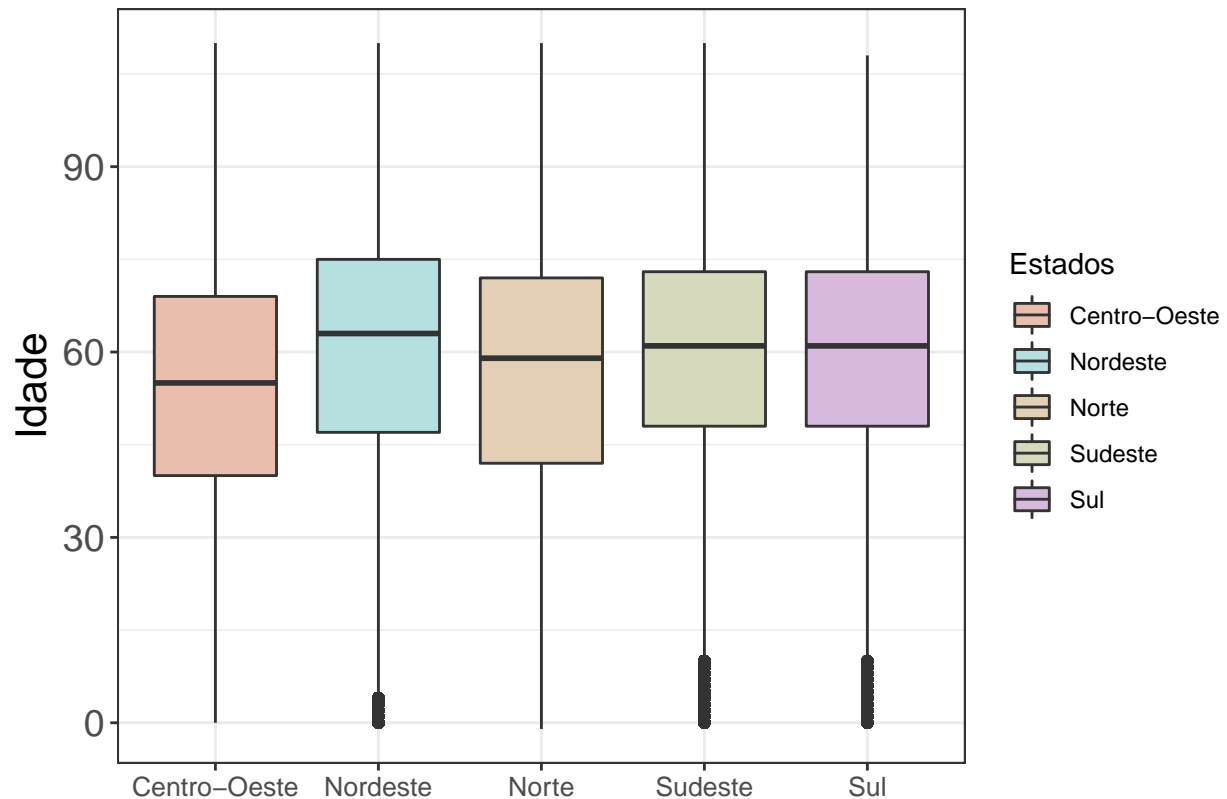


Gráfico 2: Idades por Região

### 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.

O percentual de valores omissos na coluna CS\_RACA é de 22% (vinte e dois por cento). Esses valores foram removidos.

```
#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)
```

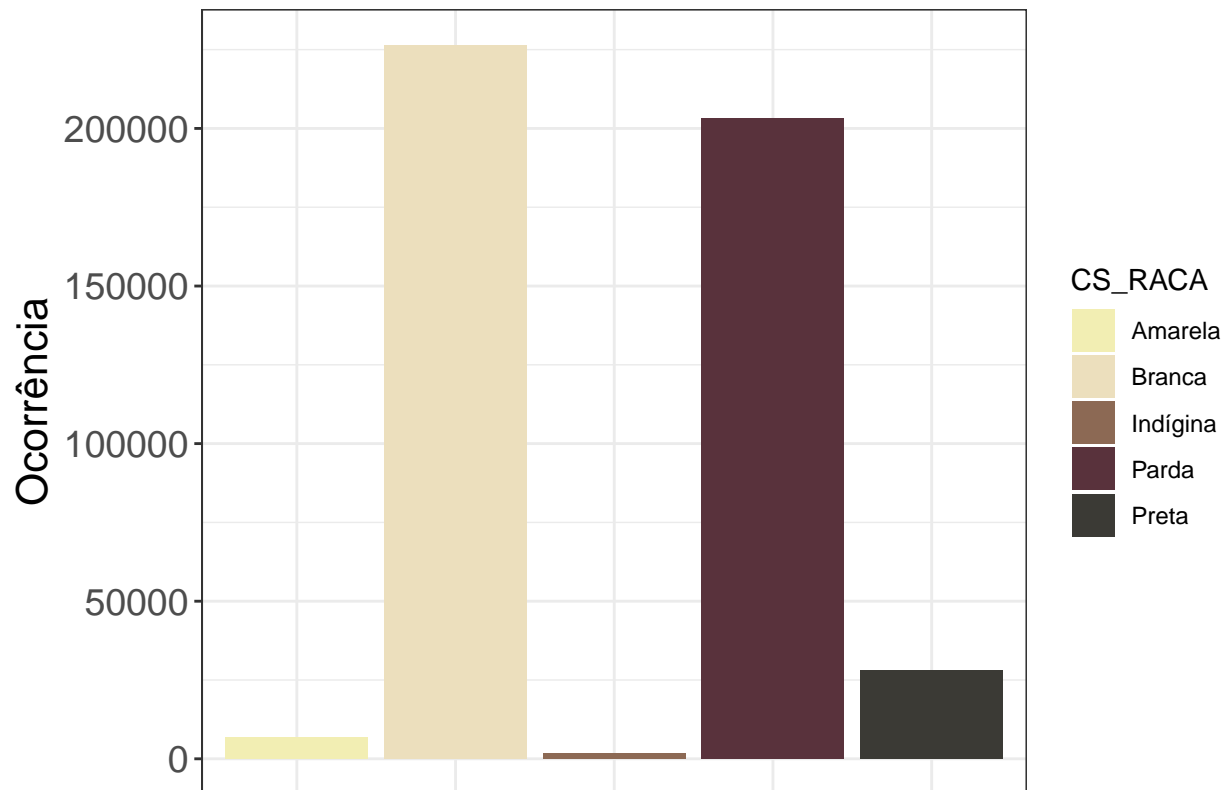


Gráfico 3: Raça

### 3.3.4 Sexo

O percentual de valores omissões na coluna sexo é de 0.01% (zero virgula zero um por cento). Esses valores omissos foram tratados a partir da coluna gestação. Atribuiu-se sexo feminino, quando o coluna gestacional informa existência de gestação. E quando não, atribuiu-se sexo masculino, dado que a presença do sexo masculino é maior e a incidência de valores omissos é muito pequena.

```
pacientes[is.na(pacientes$CS_SEX0) & pacientes$CS_GESTANTE>4, ]$CS_SEX0="M"  
pacientes[is.na(pacientes$CS_SEX0) & pacientes$CS_GESTANTE<=4, ]$CS_SEX0="F"  
pacientes$CS_SEX0 = factor(pacientes$CS_SEX0)
```

#### Sexo

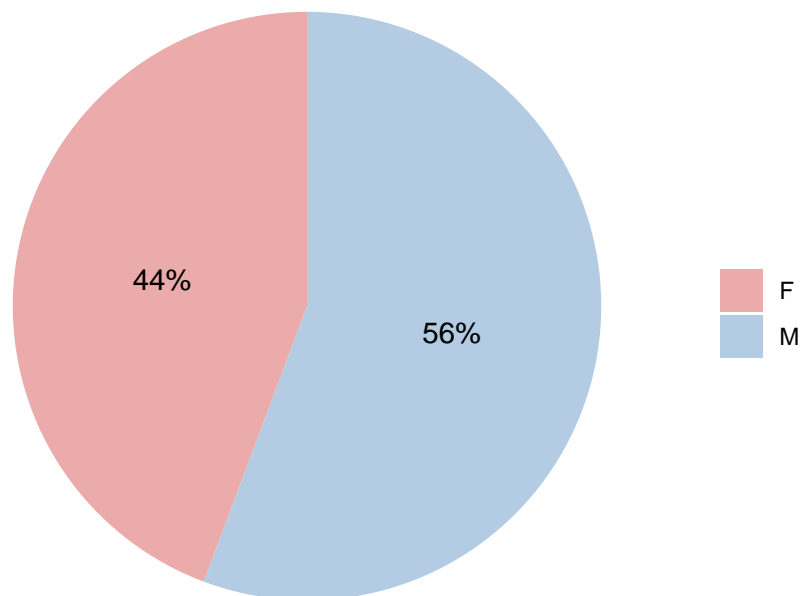


Gráfico 4: Sexo



### 3.3.5 Comorbidades

Nas colunas referente as comorbidades do paciente, o processamento foi converter os valores numéricos para as strings correspondentes, conforme descrito no dicionário de dados, disponível na seção anexo.

Não há valores omissos na coluna IN\_FATOR\_RISC, essa coluna informa se o paciente tem alguma comorbidade.

As comorbidades mais presentes nos casos de covid-19, são as cardiopatias e diabetes.

```
#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...
```

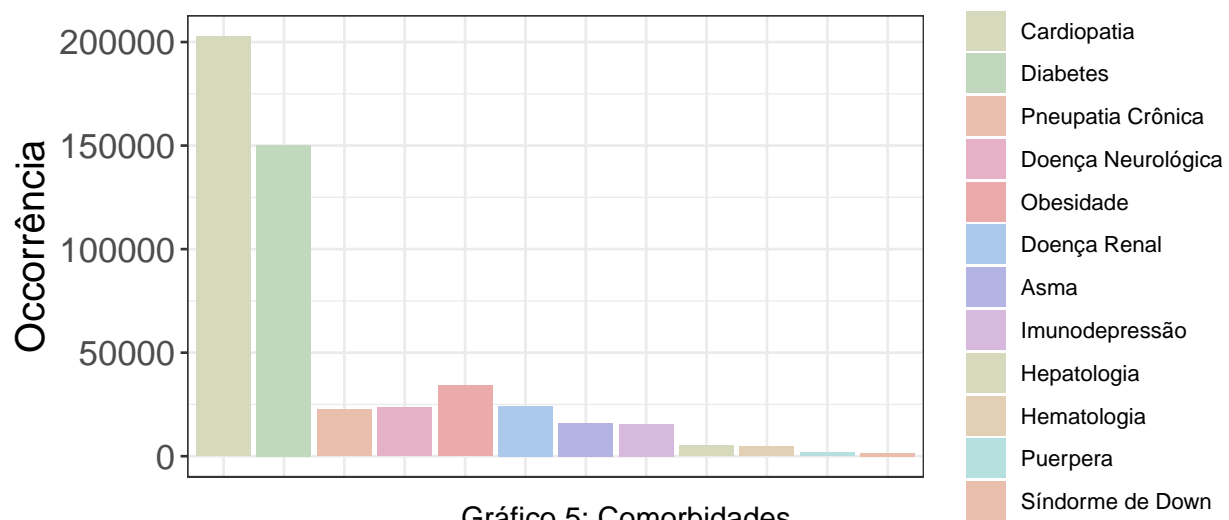


Gráfico 5: Comorbidades

### 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.

O percentual de valores omissões na coluna IN\_HOSPITALIZADO é de 2.61% (dois virgula sessenta e um por cento). Os valores omissos referentes a tratamento foram removidos.

```
#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)
```

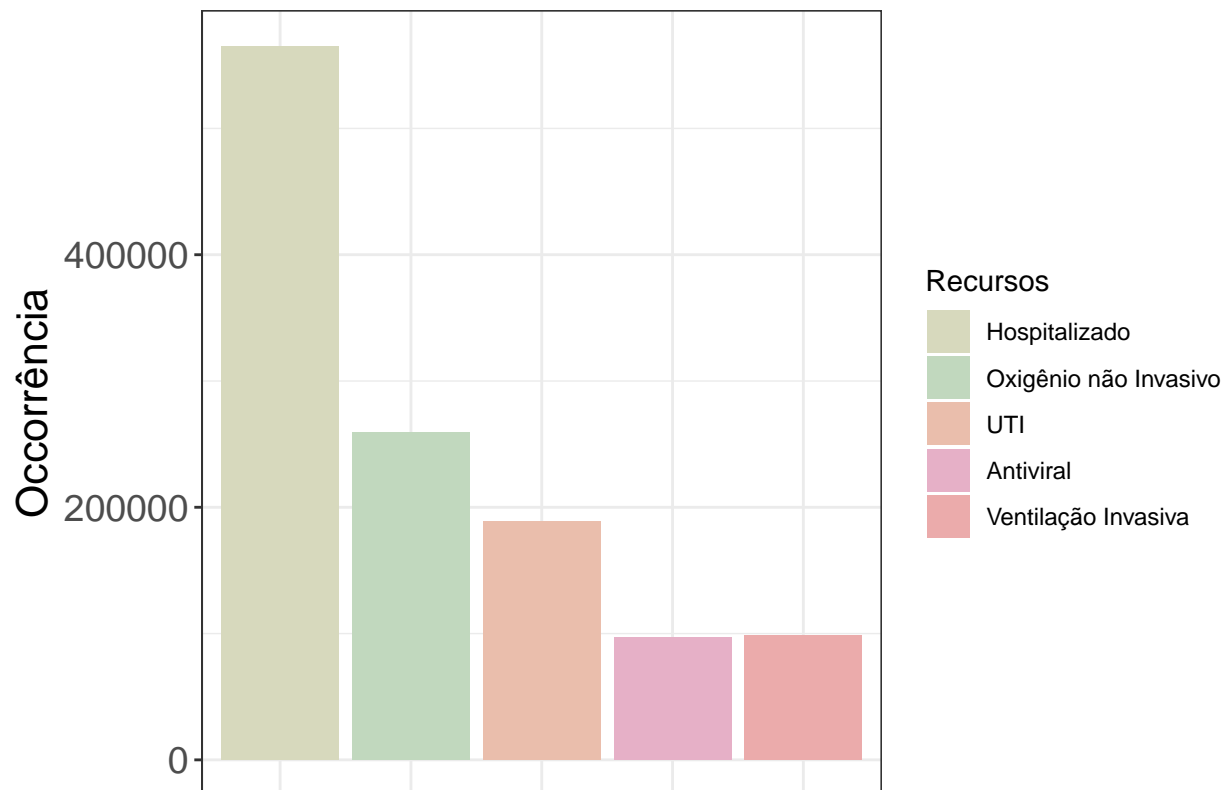


Gráfico 6: Recursos

### 3.3.7 Evolução

Para os valores numéricos presentes na coluna CS\_EVOLUCAO foram atribuídos strings mais significativas para elaboração do gráfico.

O percentual de valores omissos nesta coluna é de 12.31%. Os valores omissos referentes a evolução foram removidos.

```
#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)
```

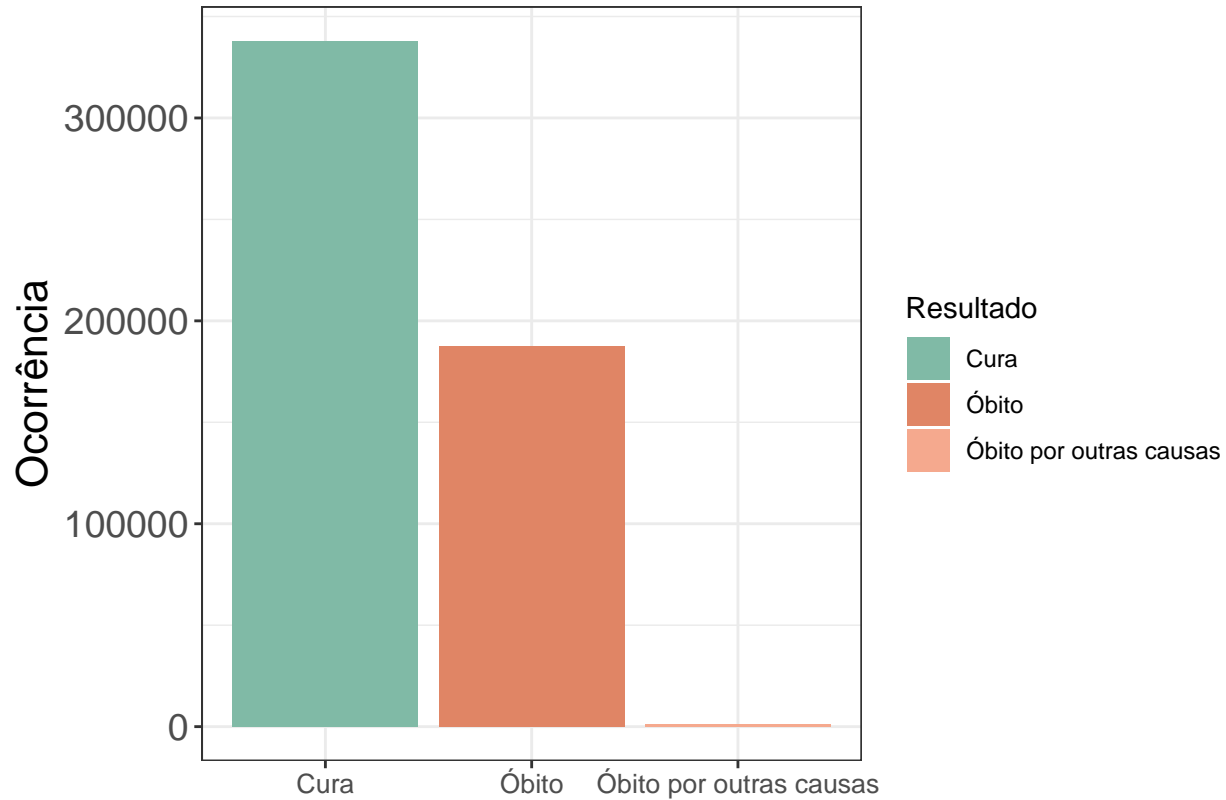


Gráfico 7: Evolução / Desfecho

## 4. Análise Exploratória

Nesse capítulo é feita uma análise detalhada da situação do Brasil em relação a covid-19. Para isto, foram contabilizados os números de mortes absolutos, bem como as 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.

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

É importante destacar que as taxas foram calculadas exclusivamente a partir dos casos graves de covid-19. As taxas gerais, dos casos leves até casos mais graves tem menor valor.

### 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.

A este respeito, o gráfico 1 apresenta a distribuição das mortes por covid-19 por região, conforme pode ser visto abaixo:

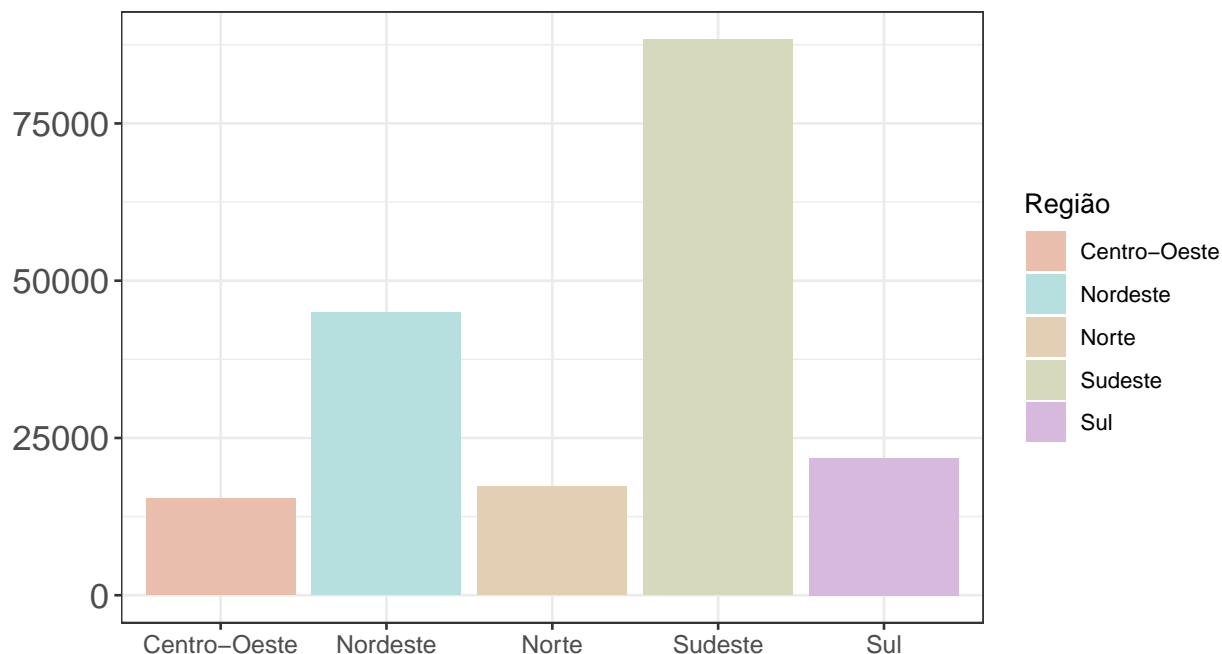


Gráfico 8: Mortes por Região

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 2 apresentado abaixo:

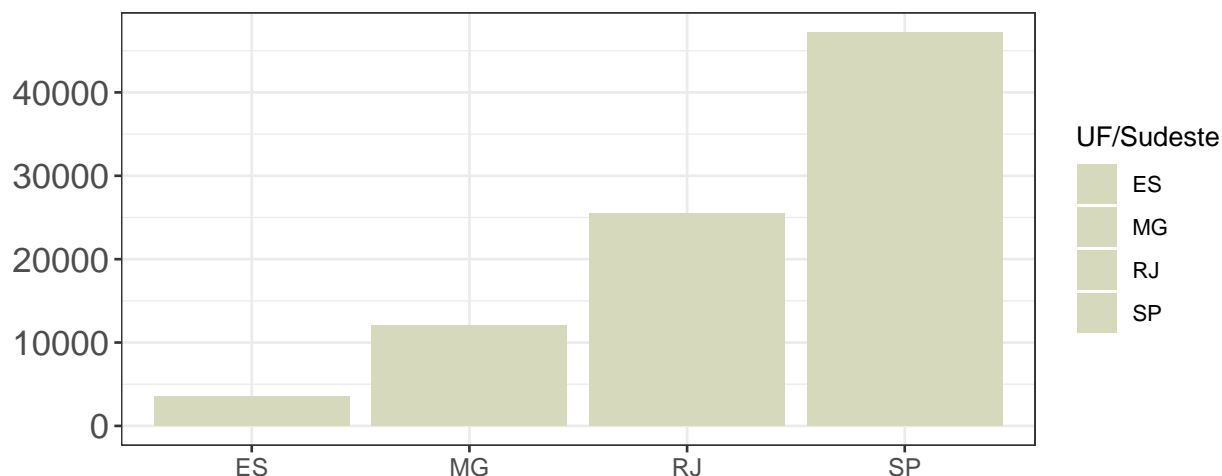


Gráfico 9: Mortes na Região Sudeste

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 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, tem quase 60% das mortes do Nordeste. Sobre este assunto, o gráfico 3, exibido abaixo, apresenta a distribuição de mortes por covid-19 na região.

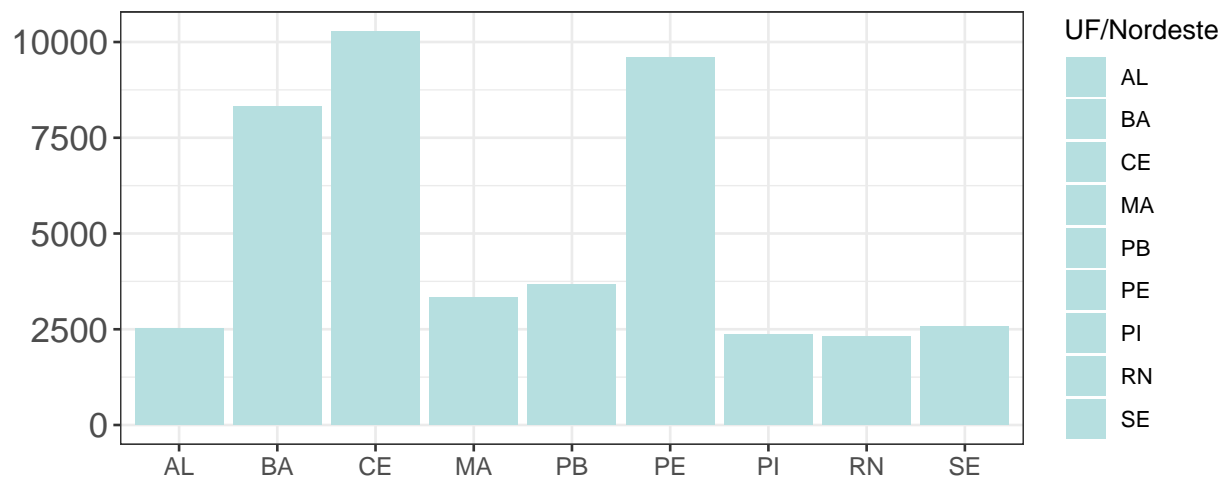


Gráfico 10: Mortes na Região Nordeste

## 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, e 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 do Distrito Federal. Veja 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

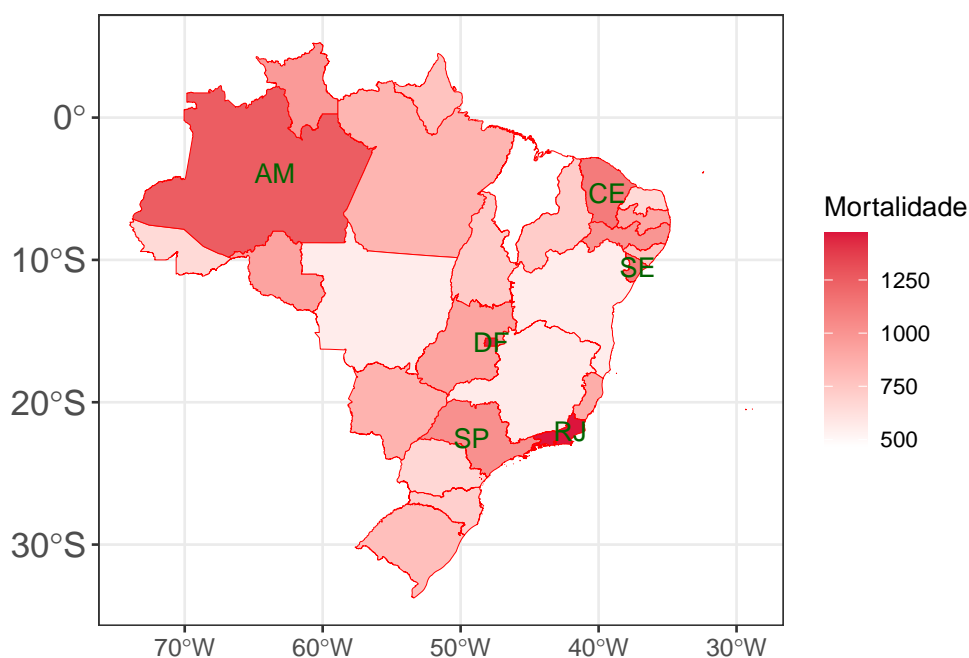


Gráfico 11: Taxa de Mortalidade por Estado

Nas capitais do país, a taxa de mortalidade são ainda maiores que as estaduais. A média da taxa de mortalidade nas capitais é de 1.689 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 óbitos, 96% (noventa e seis por cento) superior a média nacional das capitais. As capitais das regiões norte e nordeste ocupam as primeiras posições no rank das capitais com maiores taxas de mortalidade. Já o município de São Paulo tem taxa de mortalidade inferior a média nacional e ocupa a 19<sup>a</sup> (décima nona) posição do rank, com 1.436 mortes por milhão de habitantes.

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

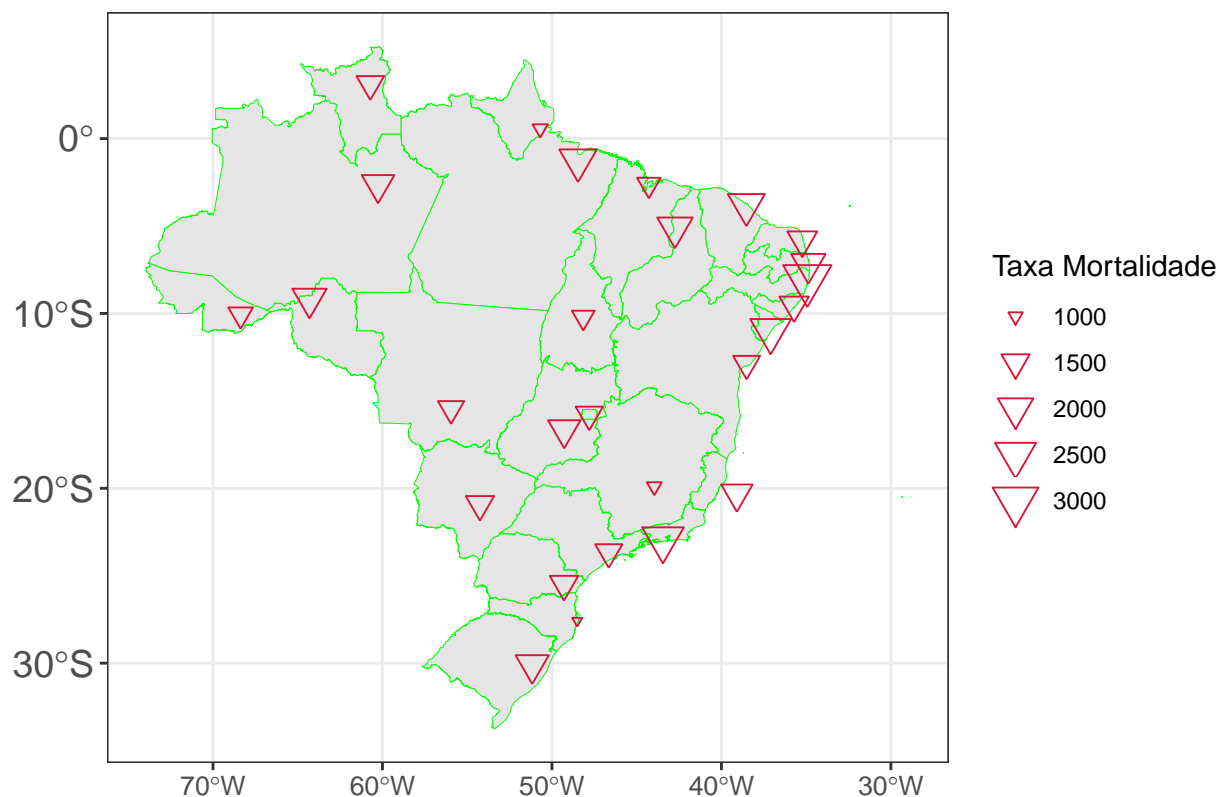


Gráfico 12: Taxa de Mortalidade Capitais

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

O teste de hipótese realizado a partir de uma amostra com 493.035 (quatrocentos e noventa e três mil e trinta e cinco) pacientes hospitalizados é de 35% (trinta e cinco por cento) com um nível de significância de 95%. A taxa equipara-se ao estudo realizado no Reino Unido, no qual compara-se a taxa de letalidade da covid-19 em pacientes hospitalizados ao vírus ebola.

```
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
## 0.3506384
```



De acordo com James Gallagher, Repórter de ciência e saúde (2020)[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.



Figura 5: Matéria BBC NEWS

As regiões Norte e Nordeste, no início da pandemia, por volta da 18<sup>a</sup> (decima oitava) semana do ano de 2020, tinham as taxas de letalidade mais altas do país, com 45% (quarenta e cinco por cento) e 36% (trinta e seis por cento), respectivamente. Pouco tempo depois, na 21<sup>a</sup> (vigésima primeira) semana do ano, a região Norte, chegará ao pior momento de 2020, momento em que a covid-19 atingiu a taxa de letalidade de 49% na região, essa foi a mais alta taxa de letalidade de todo país, em 2020.

As semanas seguintes, na região Norte, foram de queda acentuada da taxa de letalidade. A Região Norte, na 24<sup>a</sup> semana do ano, tinha a taxa de letalidade inferior a região Nordeste, esta com 46% e a região norte, com 45%. No decorrer do ano, a taxa de letalidade da região Norte continuou a cair. Então o Nordeste passou a liderar com as mais altas taxas de letalidade por quase todo ano de 2020.

Nas últimas semanas do ano a taxa de letalidade da região Norte voltou a subir. Já nas últimas semana do ano a região Norte ultrapassou a região Nordeste.

A Região sul que havia permanecido com as menores taxas de letalidade de covid-19 do país, em 2020, em dezembro deve uma subida acentuada, tendo as taxas de letalidade equiparáveis com as taxas do Nordeste. O Nordeste foi a região que por mais tempo, manteve as mais altas taxas de letalidade do país, em 2020.

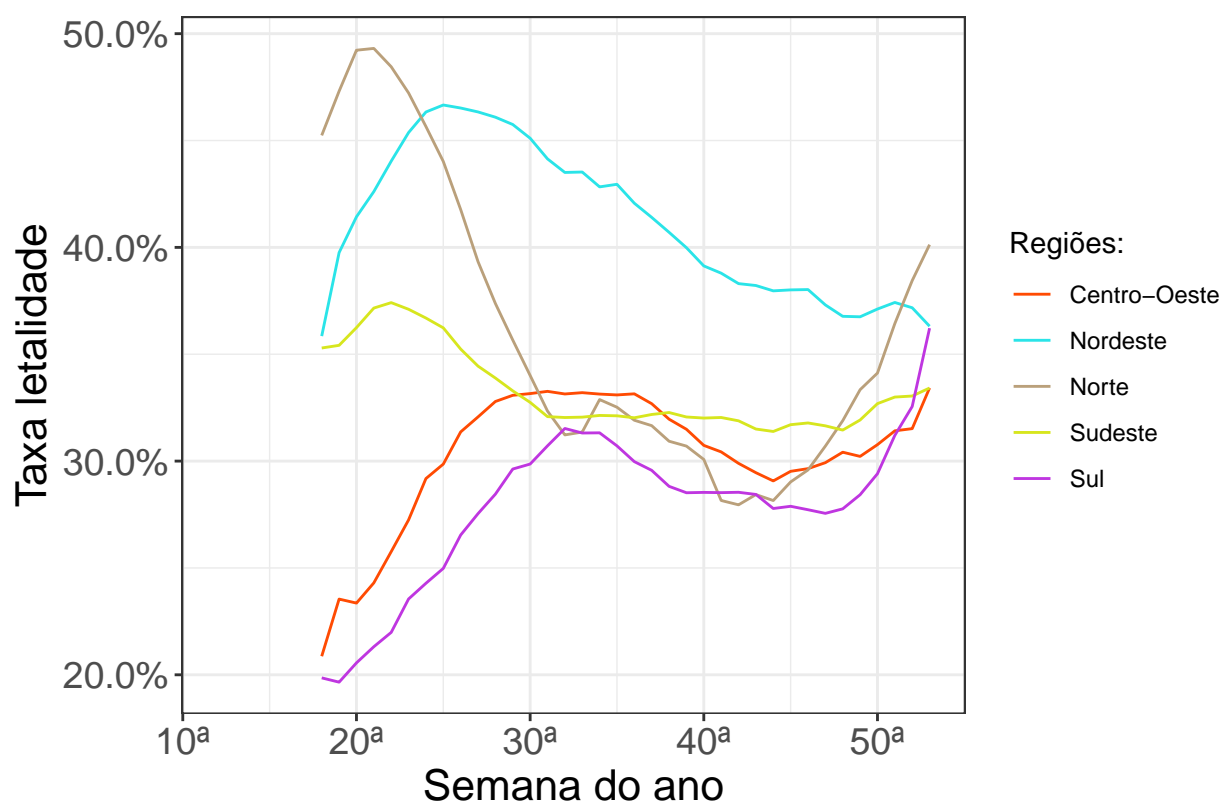


Gráfico 6: Média Móvel Taxa de Letalidade

## 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 [26](#).

	$\beta^1$	Std. Error	z value	Pr(> z )
$\beta^0$	-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 ( $\beta^1$  : 2.9888), já apresenta doenças preexistentes ( $\beta^1$  : 0.5299), reside na região nordeste ( $\beta^1$  : 0.4171), é de cor preta ( $\beta^1$  : 0.3543) e homem ( $\beta^1$  : 0.2465), com probabilidade de óbito de 80%.

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

Os valores de  $\beta^0$  e  $\beta^1$  do modelo otimizado podem ser verificados na tabela: [6](#) da página [28](#).

	$\beta^1$	Std. Error	z value	Pr(> z )
$\beta^0$	-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 Decisão 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] 0.6928662
```

### 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)
```

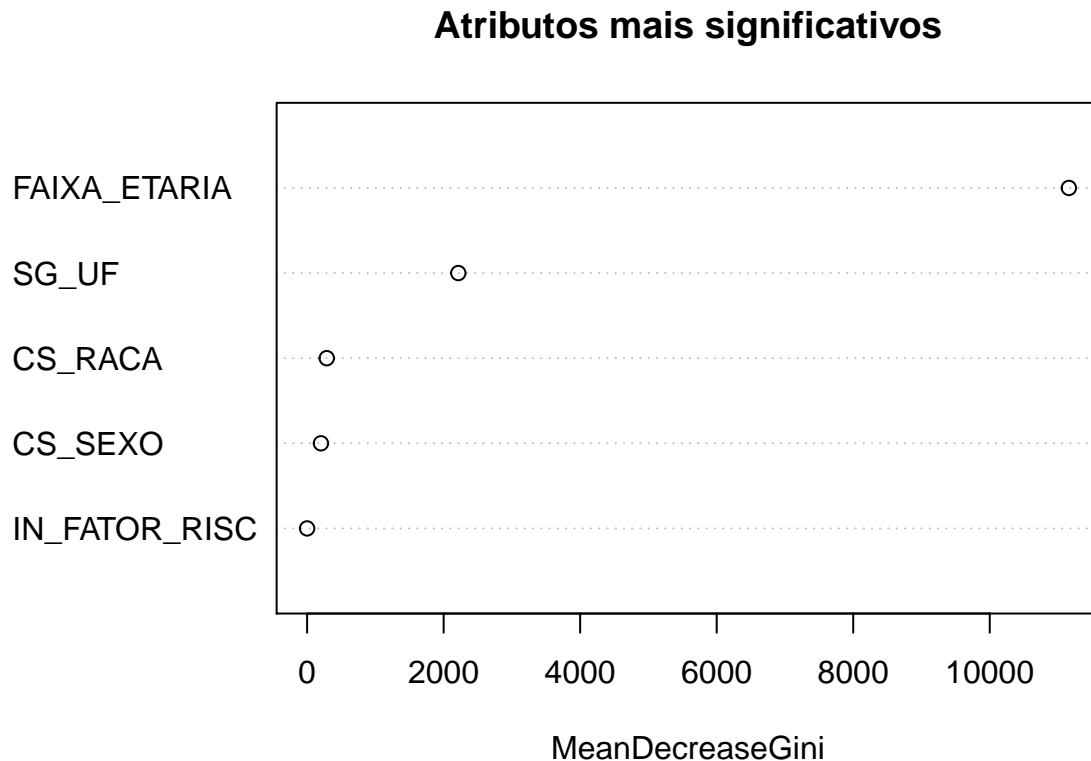
```
teste = predict(selecaoAtributos, dadosTeste)  
previsao=cbind(teste , dadosTeste)  
previsao["PREVISAO"]=previsao$teste
```

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

```
## [1] 0.7039405
```

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] 0.7188396
```



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] 0.7454913
```

## 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] 0.7165047
```

# Referência bibliográfica

## 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] 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	17366189	1470	Rio de Janeiro	6747815	2556
2	DF	3055149	1439	Brasília	3055149	1439
3	AM	4207714	1254	Manaus	2219580	1774
4	CE	9187103	1119	Fortaleza	2686612	2126
5	SE	2318822	1110	Aracaju	664908	2472
6	SP	46289333	1020	São Paulo	12325232	1436
7	PE	9616621	997	Recife	1653461	3313
8	RR	631181	969	Boa Vista	419652	1458
9	RO	1796460	927	Porto Velho	539354	1902
10	GO	7113540	924	Goiânia	1536097	1764
11	PB	4039277	906	João Pessoa	817511	1913
12	ES	4064052	879	Vitória	365855	1697
13	MS	2809394	849	Campo Grande	906092	1498
14	PA	8690745	839	Belém	1499641	2175
15	RS	11422973	788	Porto Alegre	1488252	1820
16	AP	861773	770	Macapá	512902	1015
17	AL	3351543	751	Maceió	1025360	1566
18	TO	1590248	728	Palmas	306296	1227
19	PI	3281480	723	Teresina	868075	1999
20	SC	7252502	706	Florianópolis	508826	972
21	PR	11516840	666	Curitiba	1948626	1508
22	RN	3534165	655	Natal	890480	1593
23	AC	894470	655	Rio Branco	413418	1306
24	MG	21292666	565	Belo Horizonte	2521564	1009
25	MT	3526220	560	Cuiabá	618124	1389
26	BA	14930634	557	Salvador	2886698	1426
27	MA	7114598	467	São Luís	1108975	1264

Tabela 7: Taxa Mortalidade Estados e Capitais

# Anexo A - Ficha de Notificação

Nº	
 <div> <div>MINISTÉRIO DA SAÚDE</div> <div>SECRETARIA DE VIGILÂNCIA EM SAÚDE</div> </div> <div> <div>SIVEP Gripe - SISTEMA DE INFORMAÇÃO DE VIGILÂNCIA EPIDEMIOLÓGICA DA GRIPE</div> <div>FICHA DE REGISTRO INDIVIDUAL - CASOS DE SÍNDROME RESPIRATÓRIA AGUDA GRAVE HOSPITALIZADO - 27/07/2020</div> </div>	
<b>CASO DE SÍNDROME RESPIRATÓRIA AGUDA GRAVE (SRAG-HOSPITALIZADO):</b> Indivíduo com *SG que apresente: dispneia/desconforto respiratório OU pressão persistente no tórax OU saturação de O <sub>2</sub> menor que 95% em ar ambiente OU coloração azulada dos lábios ou rosto. (*SG: Indivíduo com quadro respiratório agudo, caracterizado por pelo menos dois (2) dos seguintes sinais e sintomas: febre (mesmo que referida), calafrios, dor de garganta, dor de cabeça, tosse, coriza, distúrbios olfativos ou gustativos). Para efeito de notificação no Sivep-Gripe, devem ser considerados os casos de SRAG hospitalizados ou os óbitos por SRAG independente de hospitalização.	
1	Data do preenchimento da ficha de notificação:
2	Data de 1ºs sintomas
3	UF: _____
4	Município: _____
	Código (IBGE): _____
5	Unidade de Saúde: _____
	Código (CNES): _____
<div> <div>6</div> <div>CPF do cidadão: _____</div> </div>	
<div> <div>7</div> <div>Nome: _____</div> </div>	
<div> <div>8</div> <div>Sexo: <input type="checkbox"/> 1-Masc. 2-Fem. 9-Ign</div> </div>	
<div> <div>9</div> <div>Data de nascimento: _____</div> </div>	
<div> <div>10</div> <div>(Ou) Idade: _____</div> </div>	
<div> <div>11</div> <div>Gestante: <input type="checkbox"/></div> </div>	
<div> <div>12</div> <div>Raça/Cor: <input type="checkbox"/> 1-Branca 2-Preta 3-Amarela 4-Parda 5-Indígena 9-Ignorado</div> </div>	
<div> <div>13</div> <div>Se indígena, qual etnia? _____</div> </div>	
<div> <div>14</div> <div>Escolaridade: <input type="checkbox"/> 0-Sem escolaridade/Analfabeto 1-Fundamental 1º ciclo (1ª a 5ª série) 2-Fundamental 2º ciclo (6ª a 9ª série) 3-Médio (1ª ao 3º ano) 4-Superior 5-Não se aplica 9-Ignorado</div> </div>	
<div> <div>15</div> <div>Ocupação: _____</div> </div>	
<div> <div>16</div> <div>Nome da mãe: _____</div> </div>	
<div> <div>17</div> <div>CEP: _____</div> </div>	
<div> <div>18</div> <div>UF: _____</div> </div>	
<div> <div>19</div> <div>Município: _____</div> </div>	
<div> <div>20</div> <div>Bairro: _____</div> </div>	
<div> <div>21</div> <div>Logradouro (Rua, Avenida, etc.): _____</div> </div>	
<div> <div>22</div> <div>Nº: _____</div> </div>	
<div> <div>23</div> <div>Complemento (apto, casa, etc.): _____</div> </div>	
<div> <div>24</div> <div>(DDD) Telefone: _____</div> </div>	
<div> <div>25</div> <div>Zona: <input type="checkbox"/> 1-Urbana 2-Rural 3-Periurbana 9-Ignorado</div> </div>	
<div> <div>26</div> <div>País: (se residente fora do Brasil) _____</div> </div>	
<div> <div>27</div> <div>Paciente tem histórico de viagem internacional até 14 dias antes do início dos sintomas? <input type="checkbox"/> 1-Sim 2-Não 9-Ign</div> </div>	
<div> <div>28</div> <div>Se sim: Qual país? _____</div> </div>	
<div> <div>29</div> <div>Em qual local? _____</div> </div>	
<div> <div>30</div> <div>Data da viagem: _____</div> </div>	
<div> <div>31</div> <div>Data do retorno: _____</div> </div>	
<div> <div>32</div> <div>É caso proveniente de surto de SG que evoluiu para SRAG? <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado</div> </div>	
<div> <div>33</div> <div>Trata-se de caso nosocomial (infecção adquirida no hospital)? <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado</div> </div>	
<div> <div>34</div> <div>Paciente trabalha ou tem contato direto com aves, suínos, ou outro animal? <input type="checkbox"/> 1-Sim 2-Não 3-Outro, qual _____ 9-Ignorado</div> </div>	
<div> <div>35</div> <div>Sinais e Sintomas: 1-Sim 2-Não 9-Ignorado <input type="checkbox"/> Febre <input type="checkbox"/> Tosse <input type="checkbox"/> Dor de Garganta <input type="checkbox"/> Dispneia <input type="checkbox"/> Desconforto Respiratório <input type="checkbox"/> Saturação O<sub>2</sub>&lt;95% <input type="checkbox"/> Diarreia <input type="checkbox"/> Vômito <input type="checkbox"/> Dor abdominal <input type="checkbox"/> Fadiga <input type="checkbox"/> Perda do olfato <input type="checkbox"/> Perda do paladar <input type="checkbox"/> Outros _____</div> </div>	
<div> <div>36</div> <div>Possui fatores de risco/comorbidades? <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado Se sim, qual(is)? (Marcar X)</div> </div>	
<div> <div>37</div> <div>Recebeu vacina contra Gripe na última campanha? <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado</div> </div>	
<div> <div>38</div> <div>Data da vacinação: _____</div> </div>	
<div> <div>39</div> <div>Se &lt; 6 meses: a mãe recebeu a vacina? <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado Se sim, data: _____ a mãe amamenta a criança? <input type="checkbox"/> 1-Sim 2-Não 9-Ignorado</div> </div>	
<div> <div>40</div> <div>Se &gt;= 6 meses e &lt;= 8 anos: Data da dose única 1/1: _____ (dose única para crianças vacinadas em campanhas de anos anteriores) Data da 1ª dose: _____ (1ª dose para crianças vacinadas pela primeira vez) Data da 2ª dose: _____ (2ª dose para crianças vacinadas pela primeira vez)</div> </div>	

Dados de Atendimento	39	Usou antiviral para gripe?  __  1-Sim 2-Não 9-Ignorado	40	Qual antiviral?  __  1-Oseltamivir 2-Zanamivir 3-Outro, especifique: _____	41	Data início do tratamento ____/____/____	
	42	Houve internação?  __  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?  __  1-Sim 2-Não 9-Ignorado	48	Data da entrada na UTI: _____	49	Data da saída da UTI: _____	
	50	Uso de suporte ventilatório:  __  1-Sim, invasivo 2-Sim, não invasivo 3-Não 9-Ignorado	51	Raio X de Tórax:  __  1-Normal 2-Infiltrado intersticial 3-Consolidação 4-Misto 5-Outro: _____ 6-Não realizado 9-Ignorado	52	Data do Raio X: _____	
	53	Aspecto Tomografia  __  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  __  1-Sim 2-Não 9-Ignorado	56	Data da coleta: _____	57	Tipo de amostra:  __  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:  __  1-Imunofluorescência (IF) 2-Teste rápido antígeno		
	60	Data do resultado da pesquisa de antígenos: _____		61	Resultado da Teste antígeno:  __  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?  __  1-Sim 2-Não 9-Ignorado Se sim, qual influenza?  __  1-Influenza A 2-Influenza B Positivo para outros vírus?  __  1-Sim 2-Não 9-Ignorado Se outros vírus respiratórios qual(is)? (marcar X)  __  SARS-CoV-2  __  Vírus Sincial Respiratório  __  Parainfluenza 1  __  Parainfluenza 2  __  Parainfluenza 3  __  Adenovírus  __  Outro vírus respiratório, especifique: _____					
	64	Resultado da RT-PCR/outro método por Biologia Molecular:  __  1-Detectável 2-Não Detectável 3-Inconclusivo 4-Não realizado 5-Aguardando resultado 9-Ignorado		65	Data do resultado RT-PCR/outro método por Biologia Molecular: _____		
	66	Agente Etiológico - RT-PCR/outro método por Biologia Molecular: Positivo para Influenza?  __  1-Sim 2-Não 9-Ignorado Se sim, qual influenza?  __  1-Influenza A 2-Influenza B Influenza A, qual subtipo?  __  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?  __  1-Victoria 2-Yamagata 3-Não realizado 4-Inconclusivo 5-Outro, especifique: _____ Positivo para outros vírus?  __  1-Sim 2-Não 9-Ignorado Se outros vírus respiratórios, qual(is)? (marcar X)  __  SARS-CoV-2  __  Vírus Sincial Respiratório  __  Parainfluenza 1  __  Parainfluenza 2  __  Parainfluenza 3  __  Parainfluenza 4  __  Adenovírus  __  Metapneumovírus  __  Bocavírus  __  Rinovírus  __  Outro vírus respiratório, especifique: _____					
	67	Laboratório que realizou RT-PCR/outro método por Biologia Molecular: _____		Código (CNES):  _ _ _ _ _ _ _ _			
	68	Tipo de amostra sorológica para SARS-Cov-2:  __  1-Sangue/plasma/soro 2-Outra, qual? _____ 9-Ignorado		69	Data da coleta: _____		
	70	Tipo de Sorologia para SARS-Cov-2:  __  1-Teste rápido 2-Elisa 3-Quimiluminescência 4-Outro, qual? _____ Resultado do Teste Sorológico para SARS-Cov-2:  __  IgG  __  IgM  __  IgA 1-Positivo 2-Negativo 3-Inconclusivo 4-Não realizado 5-Aguardando resultado 9-Ignorado		71	Data do resultado: _____		
	Conclusão	72	Classificação final do caso:  __  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:  __  1-Laboratorial 2-Clinico Epidemiológico 3-Clinico 4-Clinico-Imagem	
74		Evolução do Caso:  __  1-Cura 2-Óbito 3-Óbito por outras Causas 9-Ignorado		75	Data da alta ou óbito: _____		
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	Varchar(12)		Número do registro	<b>Campo Interno</b> Número sequencial gerado automaticamente pelo sistema.  Utilizar o padrão: 320120000123  Dígito 1: caracteriza o tipo da ficha (1=SG, 2=SRAG-UTI e 3=SRAG Hospitalizado).  Dígitos 2 a 12: número sequencial gerado automaticamente pelo sistema.	NU_NOTIFIC
1-Data do preenchimento da ficha de notificação	Date DD/MM/AAAA		Data de preenchimento da ficha de notificação.	<b>Campo Obrigatório</b> Data deve ser <= a data da digitação.	DT_NOTIFIC
Semana Epidemiológica do preenchimento da ficha de notificação	Varchar(2(6))		Semana Epidemiológica do preenchimento da ficha de	<b>Campo Interno</b> Calculado a partir da data dos Primeiros Sintomas. (SS)	SEM_NOT



				<ul style="list-style-type: none"><li>▪ <b>Unidade Sentinela</b> – o campo é preenchido automaticamente pelo sistema com a Regional do Município onde está localizada a unidade de notificação.</li><li>▪ <b>Municipal</b> – o campo é preenchido automaticamente pelo sistema com a regional do município do usuário.</li></ul>	ID_UNIDADE OU CO_UNI_NOT
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.	<b>Campo Obrigatório</b>  Preenchendo o nome da unidade, o código é preenchido automaticamente, e vice-versa;  Se o usuário que está digitando a ficha for de nível: <ul style="list-style-type: none"><li>▪ <b>Unidade Sentinela</b> – o campo é preenchido automaticamente pelo sistema.</li><li>▪ <b>Municipal</b> – abre tabela apenas com as unidades sentinelas do município.</li><li>▪ <b>Esta dual ou Federal</b> – abre tabela com as unidades sentinelas do município selecionado o campo 4.</li></ul>	
6-CPF do paciente	Varchar2(15)		Número do Cadastro de Pessoa Física (CPF) do paciente notificado	<b>Campo Essencial</b>  Quando preenchido o número do CPF o sistema deverá preencher o Nome, Sexo, Data de Nascimento, Idade, Raça/Cor e o Nome da mãe do paciente.	NU_CPF
7-Nome	Varchar2(70)		Nome completo do paciente (sem abreviações)	<b>Campo Obrigatório</b>	NM_PACIENT
8-Sexo	Varchar2 (1)	1-Masculino 2-Feminino 9-Ignorado	Sexo do paciente.	<b>Campo Obrigatório</b>	CS_SEXO
9-Data de nascimento	Date DD/MM/AAAA		Data de nascimento do paciente.	<b>Campo Essencial</b>  Data deve ser <= a data dos primeiros sintomas.	DT_NASC
10-(ou) Idade	Varchar2(3)		Idade informada pelo paciente quando não se sabe a data de nascimento.	<b>Campo Obrigatório</b>  Se digitada a data de nascimento, a Idade é calculada e preenchida automaticamente pelo sistema: considerando o intervalo entre a data de nascimento e a <u>data dos primeiros sintomas</u> .  Idade deve ser <= 150.	NU_IDADE_N



(ou) Tipo/Idade	Varchar2(1)		aparente.	<b>Campo Obrigatório</b> 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.	TP_IDADE
11-Gestante	Varchar2(1)	1-1º Trimestre 2-2º Trimestre 3-3º Trimestre 4-Idade Gestacional Ignorada 5-Não 6-Não se aplica 9-Ignorado	Idade gestacional da paciente.	<b>Campo Obrigatório</b> Se selecionado categoria 2-Feminino no campo Sexo.  Se selecionado sexo igual a Masculino ou a idade for menor ou igual a 9 anos o campo é preenchido automaticamente com 6-Não se aplica.  Se selecionado sexo igual a Feminino 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-Parda 5-Indígena 9-Ignorado	Cor ou raça declarada pelo paciente: Branca; Preta; Amarela; Parda (pessoa que se declarou mulata, cabocla, cafuza, mestiça de preto com pessoa de outra cor ou raça); e, Indígena.	<b>Campo Obrigatório</b>	CS_RACA
13-Se indígena, qual etnia?	Varchar2(4)	Tabela do SIASI com código e nomes das etnias indígenas.	Nome e código da etnia do paciente,	<b>Campo Essencial</b>	CS_ETNIA

				quando indígena.	Habilitado se campo 11-Raça/Cor for igual a 5-Indígena.	
14-Escolaridade	Varchar2(1)	0-Sem escolaridade/ Analábeto 1-Fundamental 1º ciclo (1ª a 5ª série) 2-Fundamental 2º ciclo (6ª a 9ª série) 3-Médio (1º ao 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.	<b>Campo Essencial</b>  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	<b>Campo Essencial</b>	PAC_COCBO ou PAC_DSCBO	
16- Nome da mãe	Varchar2(70)		Nome completo da mãe do paciente (sem abreviações).	<b>Campo Essencial</b>	NM_MAE_PAC	
17- CEP	Varchar2(8)		CEP de residência do paciente.	<b>Campo Essencial</b> 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.	<b>Campo Obrigatório</b>  Se campo 25-País for Brasil.  Se preenchido o campo CEP, a UF é preenchida automaticamente pelo sistema e desabilitada para edição.	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.	<b>Campo Interno</b>  Preenchendo o nome da regional de saúde e de residência, o código é preenchido automaticamente, e vice-versa;	ID_RG_RESI OU CO_RG_RESI	
19- Município Código (IBGE)	Varchar2(6)	Tabela com código e nome dos Municípios padronizados pelo IBGE.	Município de residência do paciente.	<b>Campo Obrigatório</b>  Se campo 25-País for Brasil.  Se preenchido o campo CEP, o Município e seu respectivo código IBGE são preenchidos automaticamente pelo sistema e desabilitados para edição.	ID_MN_RESI 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.		
<b>20-Bairro</b>		Varchar2(72)	Tabela com código e nome dos Bairros padronizados pelos Correios.	Bairro de residência do paciente.	<b>Campo Essencial</b>  Se preenchido o campo CEP, o Bairro é preenchido automaticamente pelo sistema.	<b>NM_BAIRRO</b>	
<b>21-Logradouro (Rua, Avenida, etc.)</b>		Varchar2(50)	Tabela com código e nome dos Logradouros padronizados pelos Correios.	Logradouro (rua, avenida, quadra, travessa, etc.) do endereço de residência do paciente.	<b>Campo Essencial</b>  Se preenchido o campo CEP, o logradouro é preenchido automaticamente pelo sistema.	<b>NM_LOGRADO</b>	
<b>22-Nº</b>		Varchar2(8)		Nº do logradouro (nº da casa ou do edifício).	<b>Campo Essencial</b>	<b>NU_NUMERO</b>	
<b>23-Complemento (apto, casa, etc.)</b>		Varchar2(15)		Complemento do logradouro (bloco, apto, casa etc.).	<b>Campo Essencial</b>	<b>NM_COMPLEM</b>	
<b>24-(DDD) Telefone</b>		Varchar2(4) Varchar2(10)		Código DDD e número de telefone para contato do paciente.	<b>Campo Essencial</b>	<b>NU_DDD_TEL OU NU_TELEFON</b>	
<b>25-Zona</b>		Varchar2(1)	1-Urbana 2-Rural 3-Periurbana 9-Ignorado	Zona geográfica do endereço de residência do paciente.	<b>Campo Essencial</b>	<b>CS_ZONA</b>	
<b>26-País (se residente fora do Brasil)</b>		Varchar2(3)	Tabela com código e nome dos Países.	País de residência do paciente.	<b>Campo Obrigatório</b>  Se preenchido CEP, ou for selecionada uma UF, o campo País é preenchido automaticamente pelo sistema e desabilitado para edição.  Se selecionado País diferente de Brasil, os campos 17 a 23 são desabilitados.	<b>ID_PAIS OU CO_PAIS</b>	
<b>27- Paciente tem histórico de viagem internacional até 14 dias antes do</b>		Varchar2 (1)	1-Sim 2-Não		<b>Campo Obrigatório</b>	<b>HISTO_VGM</b>	

Início dos sintomas?	9-Ignorado						
28- Se sim: Qual país?	Varchar2(3)	Tabela com código e nome dos Países.	País onde foi realizada a viagem	<b>Campo Obrigatório</b> Habilitado se campo 27- Paciente tem histórico de viagem, selecionado categoria 1-Sim e Paciente tem histórico de viagem internacional até 14 dias antes do início dos sintomas?	<b>PAIS_VGM OU CO_PS_VGM</b>		
29- Em qual local?	Varchar2(30)		Local (cidade, estado, província e outros) onde foi realizada a viagem	<b>Campo Essencial</b> Habilitado se selecionado categoria 1-Sim em campo 27- Paciente tem histórico de viagem internacional até 14 dias antes do início dos sintomas?	<b>LO_PS_VGM</b>		
30- Data da viagem	Date DD/MM/AAAA		Data em que foi realizada a viagem	<b>Campo Obrigatório</b> 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	<b>DT_VGM</b>		
31- Data de retorno	Date DD/MM/AAAA		Data em que retornou de viagem	<b>Campo Obrigatório</b> 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	<b>DT_RT_VGM</b>		
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?	<b>Campo Essencial</b>	<b>SURTO_SG</b>		
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.	<b>Campo Essencial</b> Quando o campo 33 for igual a 1, é permitido digitar data de início dos sintomas posterior a data de internação.	<b>NOSOCOMIAL</b>		
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.	<b>Campo Essencial</b>	<b>AVE_SUINO</b>		
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.	<b>Campo Essencial</b> Habilitado de campo 34- Contato com outro animal = 3 (Outro).	<b>OUT_ANIM</b>		
35- Sinais e Sintomas/Febre	Varchar2(1) 1-Sim 2-Não 9-Ignorado		Paciente apresentou febre?	<b>Campo Essencial</b>	<b>FEBRE</b>		
35- Sinais e Sintomas/Tosse	Varchar2(1) 1-Sim 2-Não		Paciente apresentou tosse?	<b>Campo Essencial</b>	<b>TOSSE</b>		

35-Sinais e Sintomas/Dor de Garganta	Varchar2(1)	9-Ignorado 1-Sim 2-Não 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 <sub>2</sub> < 95%	Varchar2(1)	1-Sim 2-Não 9-Ignorado	Paciente apresentou saturação O <sub>2</sub> < 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 se selecionado categoria 1-Síntomas e Sinais e Outros.	FATOR_RISC
36-Fatores de risco	Varchar2(1)	1-Sim 2-Não 9-Ignorado		Paciente apresenta algum fator de risco	<b>Campo Essencial</b>	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)?	<b>Campo Essencial</b> Habilitado se 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?	<b>Campo Essencial</b>	CARDIOPATI
36-Fatores de risco/ Doença Hematológica Crônica	Varchar2(1)	1-Sim 2-Não 9-Ignorado		Paciente possui Doença Hematológica Crônica?	<b>Campo Essencial</b>	HEMATOLOGI
36-Fatores de risco/ Síndrome de Down	Varchar2(1)	1-Sim 2-Não 9-Ignorado		Paciente possui Síndrome de Down?	<b>Campo Essencial</b>	SIND_DOWN
36-Fatores de risco/ Doença Hepática Crônica	Varchar2(1)	1-Sim 2-Não 9-Ignorado		Paciente possui Doença Hepática Crônica?	<b>Campo Essencial</b>	HEPATICA
36-Fatores de risco/ Asma	Varchar2(1)	1-Sim 2-Não 9-Ignorado		Paciente possui Asma?	<b>Campo Essencial</b>	ASMA
36-Fatores de risco/ Diabetes mellitus	Varchar2(1)	1-Sim 2-Não 9-Ignorado		Paciente possui Diabetes mellitus?	<b>Campo Essencial</b>	DIABETES
36-Fatores de risco/ Doença Neurológica Crônica	Varchar2(1)	1-Sim 2-Não 9-Ignorado		Paciente possui Doença Neurológica?	<b>Campo Essencial</b>	NEUROLOGIC
36-Fatores de risco/ Outra Pneumopatia Crônica	Varchar2(1)	1-Sim 2-Não 9-Ignorado		Paciente possui outra pneumopatia crônica?	<b>Campo Essencial</b>	PNEUMOPATI
36-Fatores de risco/ Imunodeficiência	Varchar2(1)	1-Sim		Paciente possui	<b>Campo Essencial</b>	IMUNODEPRE

ou imunodepressão			2-Não 9-Ignorado		Imunodeficiência ou imunopressão (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 / caderneta.  Caso o paciente não tenha a caderneta, direcionar a pergunta para ele ou responsável e preencher o campo com o código correspondente a	Campo Essencial	VACINA

38-Data da vacinação	Date DD/MM/AAAA		resposta.	<b>Campo Essencial</b> Habilitado se campo 31-Recebeu vacina contra Gripe for igual a 1. Data deve ser <= a data da digitação (data atual).	DT_UT_DOSE
Se < 6 meses: a mãe recebeu a vacina?	Varchar2(1) 1-Sim 2-Não 9-Ignorado		Se paciente < 6 meses, a mãe recebeu vacina?	<b>Campo Essencial</b> Habilitar campo Se a idade do caso for < 6 meses.	MAE_VAC
Se sim, data	Date DD/MM/AAAA		Se a mãe recebeu vacina, qual a data?	<b>Campo Essencial</b> Habilitado se campo Se < 6 meses: a mãe recebeu a vacina for igual a 1. Data deve ser <= a data da digitação (data atual).	DT_VAC_MAE
Se < 6 meses: a mãe amamenta a criança?	Varchar2(1) 1-Sim 2-Não 9-Ignorado		Se paciente < 6 meses, a mãe amamenta a criança?	<b>Campo Essencial</b> Habilitar campo se Se a idade do caso for < 6 meses.	M_AMAMENTA
Se >= 6 meses e <= 8 anos: Data da dose única 1/1	Date DD/MM/AAAA		Se >= 6 meses e <= 8 anos, data da dose única para crianças vacinadas em campanhas de anos anteriores	<b>Campo Essencial</b> Habilitar campo Se a idade do caso for >= 6 meses e <= 8 anos	DT_DOSEUNI
Se >= 6 meses e <= 8 anos: Data da 1ª dose	Date DD/MM/AAAA		Se >= 6 meses e <= 8 anos, data da 1ª dose para crianças vacinadas pela primeira vez	<b>Campo Essencial</b> Habilitar campo Se a idade do caso for >= 6 meses e <= 8 anos	DT_1_DOSE
Se >= 6 meses e <= 8 anos: Data da 2ª dose	Date DD/MM/AAAA		Se >= 6 meses e <= 8 anos data da 2ª dose para crianças vacinadas pela primeira vez	<b>Campo Essencial</b> Habilitar campo Se a idade do caso for >= 6 meses e <= 8 anos	DT_2_DOSE
39-Usou antiviral para gripe?	Varchar2(1) 1-Sim		Fez uso de antiviral	<b>Campo Essencial</b>	ANTIVIRAL



			2-Não 9-Ignorado	para tratamento da doença?			
40-Qual antiviral?	Varchar2(1)		1-Oseltamivir 2-Zanamivir 3-Outro	Qual antiviral utilizado?	<b>Campo Essencial</b> 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.	<b>Campo Essencial</b> 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.	<b>Campo Essencial</b> Habilitado se campo 39-Uso antiviral para gripe? for igual a 1. Data deve ser <= a data da digitação (data atual).	DT_ANTIVIR	
42-Houve internação?	Varchar2(1)		1-Sim 2-Não 9-Ignorado	O paciente foi internado?	<b>Campo Essencial</b> Caso o campo não seja igual a 1 – Sim o sistema emitirá um aviso indicando que não atende a definição de caso.	HOSPITAL	
43-Data da internação por SRAG	Date DD/MM/AAAA			Data em que o paciente foi hospitalizado.	<b>Campo Obrigatório</b> Data deve ser maior ou igual a 2- Data de 1ºs sintomas e menor ou igual a data da digitação (atual).	DT_INTERNA	
44-UF de internação	Varchar2(2)		Tabela com código e siglas das UF padronizadas pelo IBGE.	Unidade Federativa de internação do paciente.	<b>Campo Essencial</b> 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.	<b>Campo Interno</b> Preenchendo o nome da regional de saúde de internação, o código é preenchido automaticamente, e vice-versa.	ID_RG_INTE OU CO_RG_INTE	
45-Município de Internação/ Código (IBGE)	Varchar2(20)		Tabela com código e nomes dos Municípios padronizados pelo IBGE.	Município onde está localizado a Unidade de Saúde onde o paciente internou.	<b>Campo Essencial</b> 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	Varchar(20)	Tabela com códigos CNES e nomes das Unidades de Saúde.	Unidade Sentinela que realizou a internação do paciente.	<b>Campo Essencial</b> Habilitado se campo 42- Houve internação? for igual a 1	ID_UN_INTE OU CO_UN_INTE
47- Internado em UTI?	Varchar(1)	1- Sim 2- Não 9- Ignorado	O paciente foi internado em UTI?	<b>Campo Essencial</b>	UTI
48- Data da entrada na UTI	Date DD/MM/AAAA		Data de entrada do paciente na unidade de Terapia intensiva (UTI).	<b>Campo Essencial</b> 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).	<b>Campo Essencial</b> 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?	Varchar(1)	1- Sim, invasivo 2- Sim, não invasivo 3- Não 9- Ignorado	O paciente fez uso de suporte ventilatório?	<b>Campo Essencial</b>	SUPORT_VEN
51- Raio X de Tórax	Varchar(1)	1- Normal 2- Infiltrado intersticial 3- Consolidação 4- Misto 5- Outro 6- Não realizado 9- Ignorado	Informar resultado de Raio X de Tórax.	<b>Campo Essencial</b>	RAIOX_RES
Raio X de Tórax/ Outro (especificar)	Varchar(30)		Informar o resultado de RX de tórax e se selecionado a opção 5- Outro.	<b>Campo Essencial</b> Habilitado de 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.	<b>Campo Essencial</b> 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	<b>Campo Essencial</b>	TOMO_RES

		3-Atípico COVID-19 4-Negativo para Pneumonia 5-Outro 6-Não realizado 9-Ignorado	tomografia.		
Aspecto Tomografia/Outro (especificar)	Varchar2(100)		Informar o resultado da tomografia selecionando a opção 5-Outro	<b>Campo Essencial</b> 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.	<b>Campo Essencial</b> 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?	<b>Campo Essencial</b>	AMOSTRA
56- Data da Coleta	Data DD/MM/AAAA		Data da coleta da amostra para realização do teste diagnóstico.	<b>Campo Essencial</b> 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 de Nasofaringe 2-Lavado Bronco-alveolar 3-Tecido <i>post-mortem</i> 4-Outra, qual? 5-LCR 9-Ignorado	Tipo da amostra clínica coletada para o teste diagnóstico.	<b>Campo Essencial</b> Habilitado de campo 55- Coletou amostra? = 1.	TP_AMOSTRA
Tipo de amostra/Outra	Varchar2(30)		Descrição tipo da amostra clínica, caso diferente das listadas nas categorias do campo.	<b>Campo Essencial</b> 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	<b>Campo Essencial</b>	REQUIL_GAL

59- Tipo do Teste antígeno	Number(3)			pelo sistema GAL	<b>Campo Essencial</b>		TP_TES_AN
60- Data do resultado teste Antígeno	Data DD/MM/AAAA	1-Imunofluorescência (IF) 2-Teste rápido antígeno		Tipo do teste antígeno que foi realizado.	<b>Campo Essencial</b>		DT_RES_AN
61- Resultado do Teste Antígeno	Varchar2(1)	1-positivo 2-Negativo 3- Inconclusivo 4-Não realizado 5-Aguardando resultado 9-Ignorado		Data do resultado do teste antígeno.	<b>Campo Essencial</b> Data deve ser maior ou igual a 56- Data da Coleta		RES_AN
62-Laboratório que realizou o Teste antígeno	Varchar2(70)	Nomes dos Laboratórios cadastrados no sistema		Resultado do Teste Antígeno	<b>Campo Essencial</b> Este campo virá marcado com 5-Aguardando Resultado e estará habilitado se o campo 55-Coletou amostra? = 1		LAB_AN
62-Laboratório que realizou o Teste antígeno	Varchar2(7)	Tabela com códigos CNES		Laboratório responsável pela liberação do resultado do teste antígeno.	<b>Campo Essencial</b> Habilitado se selecionado categoria 1-Positivo, 2-Negativo ou 3-Inconclusivo Preenchendo o nome do Laboratório, o código é preenchido automaticamente, ou vice-versa		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	<b>Campo Essencial</b>		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.	<b>Campo Essencial</b> 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.	<b>Campo Essencial</b>		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.	<b>Campo Essencial</b> 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	<b>Campo Essencial</b>		AN_VSR

VSR			VSR.	Habilitado se campo 63-Agente etiológico, Teste Antígeno. Positivo para outros vírus? = 1.	
63-Agente etiológico – Teste Antígeno. Parainfluenza 1	Teste	Varchar2(1)	Resultado do Teste Antígeno, para Parainfluenza 1. 1-marcado pelo usuário Vazio - não marcado	<b>Campo Essencial</b> Habilitado se campo 63-Agente etiológico, Teste Antígeno. Positivo para outros vírus? = 1.	AN_PARA1
63-Agente etiológico – Teste Antígeno. Parainfluenza 2		Varchar2(1)	1-marcado pelo usuário Vazio - não marcado	<b>Campo Essencial</b> Habilitado se campo 63-Agente etiológico, Teste Antígeno Positivo para outros vírus? = 1.	AN_PARA2
63-Agente etiológico – Teste Antígeno. Parainfluenza 3	Teste	Varchar2(1)	1-marcado pelo usuário Vazio - não marcado	<b>Campo Essencial</b> Habilitado se campo 63-Agente etiológico, Teste Antígeno. Positivo para outros vírus? = 1.	AN_PARA3
63-Agente etiológico – Teste Antígeno. Adenovírus	Teste	Varchar2(1)	1-marcado pelo usuário Vazio - não marcado	<b>Campo Essencial</b> Habilitado se campo 63-Agente etiológico, Teste Antígeno. Positivo para outros vírus? = 1.	AN_ADENO
63- Agente etiológico – Teste Antígeno. Outro vírus respiratório	Teste	Varchar2(1)	1-marcado pelo usuário Vazio - não marcado	<b>Campo Essencial</b> Habilitado se campo 63-Agente etiológico, Teste Antígeno. Positivo para outros vírus? = 1.	AN_OUTRO
63- Agente etiológico – Teste Antígeno. Outro vírus respiratório (Descrição)	Teste	Varchar2(30)		<b>Campo Essencial</b> Habilitado se campo 63-Agente etiológico, Teste Antígeno. Positivo para outros vírus? = 1.	DS_AN_OUT
64-Resultado da RT-PCR/outro método por Biologia Molecular		Varchar2(1)	1-Detectável 2-Não Detectável 3-Inconclusivo 4-Não Realizado 5-Aguardando Resultado 9-Ignorado	<b>Campo Essencial</b> Este campo virá marcado com 5-Aguardando Resultado e estará habilitado se o campo 55-Coletou amostra? = 1.	PCR_RESUL
65-Data do Resultado RT-PCR/outro método por Biologia Molecular		Date DD/MM/AAAA		<b>Campo Essencial</b> Campo habilitado e selecionado ca 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/outra método por Biologia Molecular: Positivo para Influenza?	Varchar2(1)	1-Sim 2-Não 9-Ignorado	Resultado da RT-PCR foi positivo para Influenza	<b>Campo Essencial</b>	POS_PCRFLU
66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Se sim, qual Influenza?	Varchar2(1)	1-Influenza A 2-Influenza B	Resultado diagnóstico do RT-PCR para o tipo de Influenza.	<b>Campo Essencial</b> Habilitado se campo 66-Agente etiológico – RT-PCR/outra método por Biologia Molecular: Positivo para Influenza? = 1.	TP_FLU_PCR
66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Se Influenza A, qual subtipo?	Varchar2(1)	1-Influenza A(H1N1)pdm09 2-Influenza A(H3N2) 3-Influenza A não subtipado 4-Influenza A não subtipável 5-Inconclusivo 6-Outro, especifique:	Subtipo para Influenza A.	<b>Campo Essencial</b> Habilitado se campo 66-Agente etiológico – RT-PCR/outra método por Biologia Molecular: Se sim, qual Influenza? = 1.	PCR_FLUASU
66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Se Influenza A, qual subtipo? Outro, especifique:	Varchar2(30)		Outro subtipo para Influenza A.	<b>Campo Essencial</b> Habilitado se Agente etiológico – RT-PCR/outra método por Biologia Molecular: Se Influenza A, qual subtipo? = 6.	FLUASU_OUT
66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Se Influenza B, qual linhagem?	Varchar2(1)	1-Victoria 2-Yamagata 3-Não realizado 4-Inconclusivo 5-Outro, especifique:	Linhagem para Influenza B.	<b>Campo Essencial</b> Habilitado se campo 66-Agente etiológico – RT-PCR/outra método por Biologia Molecular: Se sim, qual Influenza? = 2.	PCR_FLUBLI
66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Se Influenza B, qual linhagem? Outro, especifique:	Varchar2(30)		Outra linhagem para Influenza B.	<b>Campo Essencial</b> Habilitado se 66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Se Influenza B, qual linhagem? = 5.	FLUBLI_OUT
66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Positivo para outros vírus?	Varchar2(1)	1-Sim 2-Não 9-Ignorado	Resultado da RT-PCR foi positivo para outro vírus respiratório	<b>Campo Essencial</b>	POS_PCROUT
66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: SARS-CoV-2	Varchar2(1)	1-marcado pelo usuário Vazio - não marcado	Resultado diagnóstico do RT-PCR para (SARS-CoV-2).	<b>Campo Essencial</b> Habilitado se campo 66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Positivo para outros vírus? = 1.	PCR_SARS2
66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: VSR	Varchar2(1)	1-marcado pelo usuário Vazio - não marcado	Resultado diagnóstico do RT-PCR para (VSR).	<b>Campo Essencial</b> Habilitado se campo 66- Agente etiológico – RT-PCR/outra método por	PCR_VSR

66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Parainfluenza 1	Varchar2(1)	1-marcado pelo usuário Vazio - não marcado	Resultado diagnóstico do RT-PCR para Parainfluenza 1.	Biologia Molecular: Positivo para outros vírus? = 1  Campo Essencial Habilitado se campo 66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Positivo para outros vírus? = 1	PCR_PARA1
66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Parainfluenza 2	Varchar2(1)	1-marcado pelo usuário Vazio - não marcado	Resultado diagnóstico do RT-PCR para Parainfluenza 2.	Campo Essencial Habilitado se campo 66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Positivo para outros vírus? = 1	PCR_PARA2
66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Parainfluenza 3	Varchar2(1)	1-marcado pelo usuário Vazio - não marcado	Resultado diagnóstico do RT-PCR para Parainfluenza 3.	Campo Essencial Habilitado se campo 66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Positivo para outros vírus? = 1	PCR_PARA3
66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Parainfluenza 4	Varchar2(1)	1-marcado pelo usuário Vazio - não marcado	Resultado diagnóstico do RT-PCR para Parainfluenza 4.	Campo Essencial Habilitado se campo 66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Positivo para outros vírus? = 1	PCR_PARA4
66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Adenovírus	Varchar2(1)	1-marcado pelo usuário Vazio - não marcado	Resultado diagnóstico do RT-PCR para Adenovírus.	Campo Essencial Habilitado se campo 66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Positivo para outros vírus? = 1	PCR_ADENO
66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Metapneumovírus	Varchar2(1)	1-marcado pelo usuário Vazio - não marcado	Resultado diagnóstico do RT-PCR para Metapneumovírus.	Campo Essencial Habilitado se campo 66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Positivo para outros vírus? = 1	PCR_METAP
66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Bocavírus	Varchar2(1)	1-marcado pelo usuário Vazio - não marcado	Resultado diagnóstico do RT-PCR para Bocavírus.	Campo Essencial Habilitado se campo 66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Positivo para outros vírus? = 1	PCR_BOCA
66- Agente etiológico – RT-PCR/outra método por Biologia Molecular: Rinovírus	Varchar2(1)	1-marcado pelo usuário Vazio - não marcado	Resultado diagnóstico do RT-PCR para Rinovírus.	Campo Essencial Habilitado se campo 66- Agente etiológico – RT-PCR/outra 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-marcado pelo usuário Vazio - não marcado	Resultado diagnóstico do RT- PCR para Outro vírus respiratório.	<b>Campo Essencial</b>  Habilitado se campo 66- Agente etiológico – RT-PCR/ou outro método por Biologia Molecular: Positivo para outros vírus? = 1	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.	<b>Campo Essencial</b>  Habilitado se 66- Agente etiológico – RT-PCR/ou outro método por Biologia Molecular: Outro vírus respiratório, especifique:	DS_PCR_OUT	
67- Laboratório que realizou RT- PCR/ou outro 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.	<b>Campo Essencial</b>  Habilitado e selecionado categoria 1- Detectável, 2- Não Detectável ou 3- Inconclusivo em 66- Resultado da RT-PCR/ou outro método por Biologia Molecular.  Preenchendo o nome do laboratório, o código é preenchido automaticamente, ou vice-versa.	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.	<b>Campo Essencial</b>	TP_AM_SOR	
Tipo de Amostra Sorológica para SARS- Cov-2/Outra, qual?			Descrição tipo da amostra clínica, caso o diferente das listadas na categoria um (1) do campo.	<b>Campo Essencial</b>  Campo habilitado se selecionado categoria 2- Outra, qual? em Tipo de Amostra Sorológica.	SOR_OUT	
69- Data da coleta	Data DD/MM/AAAA		Data da coleta do material para diagnóstico por Sorologia.	<b>Campo Essencial</b>  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_CO_SOR	
70- Tipo de Sorologia para SARS-Cov-2	Number(3)	1- Teste rápido 2- Elisa 3- Qui miluminescência 4- Outro, qual	Tipo do Teste Sorológico que foi realizado	<b>Campo Essencial</b>	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? em Tipo de Sorologia.	OUT_SOR
70- Tipo de Sorologia para SARS-Cov-2/Outro, qual?	Varchar2(100)		Outro tipo de amostra Sorológica		SOR_OUT
70- Resultado do Teste Sorológico para SARS-Cov-2:	Varchar2(1)	1-Positivo 2-Negativo 3-Inconclusivo 4-Não realizado 5-Aguarda resultado 9 Ignorado	Resultado da Sorologia para SARS-Cov-2	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

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

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