

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE INFORMÁTICA  
CURSO DE ENGENHARIA DE COMPUTAÇÃO

FABRÍCIO SZCZESNY DE ALMEIDA

**Desenvolvimento de uma ferramenta para  
análise e estudo, do reflexo da vacinação  
sobre os dados da pandemia da COVID-19**

Monografia apresentada como requisito parcial  
para a obtenção do grau de Bacharel em  
Engenharia da Computação

Orientador: Prof. Dr.  
José Rodrigo Furlaneto Azambuja

Porto Alegre  
2022

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Carlos André Bulhões Mendes

Vice-Reitora: Prof.<sup>a</sup> Patricia Helena Lucas Pranke

Pró-Reitora de Graduação: Prof.<sup>a</sup> Cíntia Inês Boll

Diretora do Instituto de Informática: Prof.<sup>a</sup>. Carla Maria Dal Sasso Freitas

Diretora da Escola de Engenharia: Prof.<sup>a</sup>. Carla Schwengber Ten Caten

Coordenador do Curso de Engenharia de Computação: Prof. Walter Fetter Lages

Bibliotecária-chefe do Instituto de Informática: Beatriz Regina Bastos Haro

Bibliotecária-chefe da Escola de Engenharia: Rosane Beatriz Allegretti Borges

*“Sometimes life will hit a brick on the head.*

*Do not lose faith”*

— STEVE JOBS

## **AGRADECIMENTOS**

Agradeço primeiramente a Deus por estar comigo durante toda minha caminhada, me dando saúde, força e fé de que tudo daria certo.

Aos meus pais Oneida R. S. de Almeida e Luciano de Almeida que nunca mediram esforços para que eu tivesse uma educação de qualidade e conseguisse chegar onde cheguei.

A instituição universitária FURG e seus professores, onde iniciei minha jornada acadêmica e descobri minha real vocação profissional, a instituição universitária UFRGS e seus professores, onde segui minha carreira acadêmica e cresci muito profissionalmente e pessoalmente.

Ao meu orientador José Rodrigo F. Azambuja, por todo apoio e dedicação comigo ao longo do desenvolvimento deste trabalho.

A minha esposa, Andreza de L. C. Silva, que esteve ao meu lado nos últimos anos desse ciclo e sempre me motivou e me deu carinho para seguir.

Aos meus tios Edson C. do Valle e Silvana de Almeida por serem os maiores motivadores do meu ingresso na universidade, e terem sido exemplos a serem seguidos.

A todos meus familiares e amigos, que direta e indiretamente tiveram participações ao longo de toda minha jornada acadêmica, e muitas vezes compreenderam minhas ausências com eles, em virtude de envolvimento relacionados a universidade.

## RESUMO

Dados de proliferação de doenças são utilizados para analisar a evolução da contaminação em uma determinada área geográfica, e traçar políticas de combate e tratamento. Atualmente, estamos vivendo, a nível global, um grande período pandêmico da doença COVID-19, e prever surtos com exatidão, traçar os melhores planos de prevenção e combate ao vírus causador da doença, são procedimentos difíceis devido a grande variedade de fatores geográficos e populacionais envolvidos. Visando esse cenário, este trabalho tem como objetivo contextualizar e implementar uma aplicação de visualização e consulta a dados divulgados pelo governo brasileiro, sobre a pandemia da COVID-19, aplicando técnicas, visuais e estatísticas, de análise de dados para auxiliar na definição de políticas de priorização da vacinação, dentre os grupos definidos pelo governo com base em idade, comorbidades e atuação no mercado de trabalho.

**Palavras-chave:** Ciência de dados. Análise de dados. Pandemia. Doenças. Surtos. Combate. Vacinação. COVID-19.

## **Development of a tool for analysis and study of the vaccination reflex about COVID-19 pandemic data**

### **ABSTRACT**

Disease proliferation data are used to analyze the evolution of contamination in a given geographic area, and to outline policies for combating and treating it. We are currently experiencing, globally, a major pandemic period of the COVID-19 disease, and accurately predicting outbreaks, drawing up the best plans to prevent and combat the virus that causes the disease, are difficult procedures due to the wide variety of geographic and populations involved. Aiming at this scenario, this work aims to contextualize and implement an application to visualize and consult data released by the Brazilian government on the COVID-19 pandemic, applying techniques, visuals and statistics, of data analysis to assist in the definition of policies. prioritization of vaccination, among the groups defined by the government based on age, comorbidities and performance in the labor market.

**Keywords:** Data science, Data analysis, Pandemic, Illnesses, Outbreaks, Combat, Vaccination, COVID-19.

## **LISTA DE ABREVIATURAS E SIGLAS**

UFRGS	Universidade Federal do Rio Grande do Sul
FURG	Universidade Federal do Rio Grande
SGBD	Sistema de gerenciamento de banco de dados
MVC	Model View Controller
HTML	HyperText Markup Language
CSS	Cascading Style Sheets
UML	Linguagem de Modelagem Unificada
MSB	Ministério da saúde brasileiro
OMS	Organização Mundial da Saúde
RF	Requisito funcional
RFN	Requisito não funcional

## LISTA DE FIGURAS

Figura 1.1	Casos diários de COVID-19 no Brasil de março/2020 a Abril/2022. ....	12
Figura 1.2	Mortes diárias por COVID-19 no Brasil de março/2020 a Abril/2022. ....	13
Figura 2.1	Interface de dados do Brasil - Pátria Vacinada. ....	15
Figura 2.2	Visão gráfica dos dados pelo Coronavírus Brasil. ....	16
Figura 2.3	Visão numérica de dados pelo Coronavírus Brasil. ....	16
Figura 2.4	Interface do Dados COVID-19 em Porto Alegre. ....	17
Figura 2.5	Visão geral da ferramenta COVID-19 Analysis Tools. ....	18
Figura 2.6	Visão geral do painel Monitora COVID-19. ....	19
Figura 2.7	Visão gráfica fornecida pelo Monitora COVID-19. ....	19
Figura 3.1	Filtro de segmentação. ....	20
Figura 3.2	Visão de dados consultados. ....	22
Figura 3.3	Estatísticas grupo de vacinação. ....	22
Figura 3.4	Diagrama UML de componentes da aplicação e ilustração do MVC. ....	23
Figura 3.5	Diagrama UML do SGBD. ....	26
Figura 4.1	Página inicial da aplicação. ....	32
Figura 4.2	Tela de filtragem e seleção para consulta de dados. ....	32
Figura 4.3	Resultado da consulta de dados. ....	33
Figura 4.4	Tela de filtragem e seleção para consulta sobre reflexo da vacinação. ....	34
Figura 4.5	Resultado da consulta sobre reflexo da vacinação. ....	34
Figura B.1	Página inicial da aplicação. ....	52
Figura B.2	Página de segmentação da consulta. ....	53
Figura B.3	Página de segmentação da consulta. ....	53
Figura B.4	Página de segmentação da consulta a vacinas. ....	54
Figura B.5	Página de segmentação da consulta de vacinação. ....	54



## LISTA DE TABELAS

Tabela 3.1	Requisitos funcionais.....	21
Tabela 3.2	Requisitos não funcionais.....	22
Tabela 3.3	Dicionário de dados da base de casos e óbitos.....	27
Tabela 3.4	Dicionário de dados da base de vacinação. ....	28
Tabela 3.5	Dicionário de dados da base de ocupação hospitalar.....	29
Tabela 4.1	Comparação de funcionalidades oferecidos. ....	35
Tabela 4.2	Comparação de abrangência geográfica e origem dos dados. ....	36
Tabela 4.3	Comparação de dados utilizados. ....	36
Tabela 4.4	Comparação de filtros oferecidos. ....	36

## SUMÁRIO

<b>1 INTRODUÇÃO</b>	<b>11</b>
1.1 Contextualização	11
1.2 Motivação	12
1.3 Objetivo	14
<b>2 ESTUDO DE FERRAMENTAS EXISTENTES</b>	<b>15</b>
2.1 Brasil - Pátria Vacinada	15
2.2 Coronavírus Brasil	15
2.3 Dados COVID-19 em Porto Alegre	17
2.4 Websensors	17
2.5 COVID-19 Analysis Tools	18
2.6 Monitora COVID-19	18
<b>3 APRESENTAÇÃO DA SOLUÇÃO</b>	<b>20</b>
<b>3.1 Levantamento de requisitos</b>	<b>20</b>
3.1.1 Requisitos funcionais	20
3.1.2 Requisitos não funcionais	22
<b>3.2 Arquitetura</b>	<b>23</b>
3.2.1 Controller (Back-end)	24
3.2.2 View (Front-end)	25
3.2.3 Model (SGBD)	25
<b>3.3 Dados utilizados</b>	<b>25</b>
3.3.1 Casos e óbitos	27
3.3.2 Vacinação	27
3.3.3 Ocupação hospitalar	29
<b>4 RESULTADOS</b>	<b>31</b>
<b>4.1 Funcionalidades</b>	<b>31</b>
4.1.1 Página inicial	31
4.1.2 Consulta de dados	32
4.1.3 Consulta sobre reflexo da vacinação	33
<b>4.2 Comparativos</b>	<b>35</b>
<b>5 CONCLUSÃO</b>	<b>37</b>
<b>REFERÊNCIAS</b>	<b>38</b>
<b>APÊNDICEA PROCESSAMENTO DE DADOS</b>	<b>39</b>
A.1 Casos e óbitos	39
A.2 Ocupação hospitalar	44
A.3 Vacinação	48
<b>APÊNDICEB GUIA DE UTILIZAÇÃO</b>	<b>52</b>
B.1 Acessando a ferramenta	52
B.2 Consultando dados	52
B.3 Consultando reflexo da vacinação	54

# 1 INTRODUÇÃO

## 1.1 Contextualização

Desde seus primórdios, a humanidade sofre com surgimento de novas doenças, fazendo com que a ela se adaptasse e desenvolvesse formas de prevenção e combate a essas doenças para garantir a continuidade de sua existência. Essas doenças, originadas de vírus, bactérias ou fungos, normalmente apresentam características de surtos sazonais e regionais, causando epidemias, mas em situações mais severas acabam tomando proporções mundiais caracterizando pandemias.

Em dezembro de 2019 (GRUBER, 2020) surgiram alguns casos de pessoas com sintomas gripais severos na cidade de Wuhan, na China, e dentro de alguns dias pesquisadores constataram que essas pessoas haviam sido contaminadas por um novo vírus, altamente contagioso e que agredia severamente o sistema respiratório da pessoa infectada. Devido a muitas semelhanças do vírus que infectou esse grupo de pessoas, na cidade de Wuhan, com o vírus SARS-CoV (comum entre espécies de morcegos e pangolins), pesquisadores acreditam que ele foi resultado da mutação do SARS-CoV e o classificaram como SARS-CoV-2, conhecido popularmente como o novo Coronavírus ou COVID-19.

Assim que identificados os primeiros casos e a gravidade da doença que havia surgido, as autoridades locais de Wuhan, tomaram todas as medidas sanitárias cabíveis para que ocorresse apenas casos locais e a doença não tomasse maiores proporções, mas devido ao grande potencial de transmissão do vírus e o grande movimento de pessoas entre países que há hoje em dia, reflexo da globalização, em meados de janeiro de 2020 foram identificados os primeiros casos da doença na Europa e em pequeno espaço de tempo nos demais continentes, os sintomas severos, aliados às difíceis políticas sanitárias de prevenção a transmissão e a ausência de um tratamento eficaz, tanto de prevenção, como de cura para contaminados fez a doença tomar proporções mundialmente alarmantes, sendo assim anunciada pela OMS a pandemia da COVID-19, em 11 de março de 2020.

No Brasil o primeiro caso da COVID-19 foi identificado em 26 de fevereiro de 2020 na cidade de São Paulo (MSB, 2020b), onde o infectado era um homem que havia retornado de viagem da Europa. Assim que identificados os primeiros casos, os governos regionais passaram a tomar medidas sanitárias como a obrigatoriedade do uso de máscaras, higienização de mãos e redução de aglomerações, afim de reduzir o contágio, enquanto mundialmente pesquisadores de todas as áreas buscavam identificar soluções

para o quadro de caos na saúde que havia se instaurado.

## 1.2 Motivação

Desde a identificação dos primeiros casos da doença e início da pandemia da COVID-19, os pesquisadores da área da saúde buscaram constantemente a descoberta de meios eficazes de combate a disseminação do vírus, tratamento de prevenção e sintomas severos em infectados. Mundialmente, uma das maiores frentes de pesquisa foi sobre o desenvolvimento de vacinas que realizassem a imunização da população, pesquisa que devido o grande tempo de desenvolvimento e testes necessários, aliados à falta de incentivo governamental em diferentes países acabou por gerar resultados satisfatórios e eficazes apenas ao final do ano de 2020 (BBC, 2021), fazendo com que a pandemia da COVID-19 já tivesse atingido altíssimos números de óbitos e pessoas com sequelas da doença. No Brasil, em especial, tivemos números jamais vistos antes em uma crise sanitária (figuras 1.2 e ).

Figura 1.1: Casos diários de COVID-19 no Brasil de março/2020 a Abril/2022.

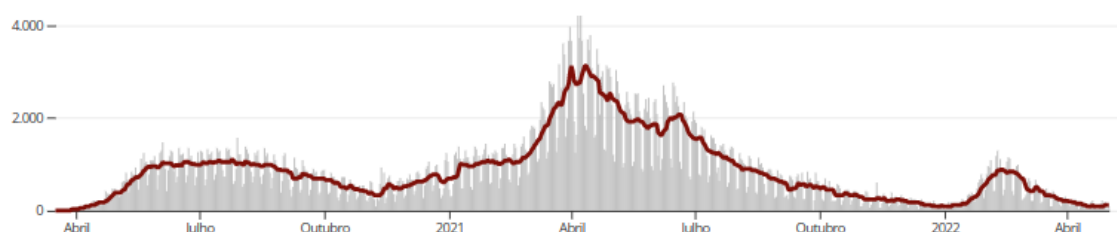


Fonte: G1 via Consórcio de veículos de imprensa.

Após o início da vacinação no Brasil, em 17 de janeiro de 2021 (FIOCRUZ, 2022), os pesquisadores da área passaram também a ter como objetivo de pesquisa a definição de grupos de pessoas que teriam prioridade na vacinação, utilizando critérios como a idade, área de atuação no mercado de trabalho, histórico de doenças crônicas, entre outros, afim de maximizar o efeito na população como um todo, e causar a maior redução possível

no número de casos graves e óbitos causados pela COVID-19. Sem amostras de dados que ilustrasse o impacto de imunizações anteriores sob a população e baseados apenas em estudos de comprovação da eficácia do imunizante, ainda em fase de testes, e em dados de campanhas de vacinação da influenza, aplicada anualmente, definiu-se uma política de vacinação baseada no agrupamento e priorização da população conforme comorbidades, faixa etária, e área de atuação no mercado de trabalho.

Figura 1.2: Mortes diárias por COVID-19 no Brasil de março/2020 a Abril/2022.



Fonte: G1 via Consórcio de veículos de imprensa.

A centralização de dados relacionados a pandemia da COVID-19, em diversos meios de fácil e livre acesso, proporcionada pela globalização tecnológica vivenciada nos últimos anos, alidada a inúmeras pesquisas na área da ciência de dados e ao interesse de entidades governamentais em mapear locais de maior crescimento no número de novos casos e óbitos, para aplicar medidas de prevenção ao avanço da COVID-19, tivemos o desenvolvimento de uma série de soluções que visavam prover uma melhor visualização e interpretação dos dados disponibilizados. Estas soluções por muitas vezes possuíam uma pequena abrangência regional e limitações na natureza de dados utilizada e filtros disponibilizados, como será apresentado no capítulo 2.

Conforme especialistas, (BRITO, 2020) soluções utilizando a ciência de dados, por mais simples que fossem, foram de grande importância para identificar picos de casos, localidades de maior incidência de novos casos e localidades onde a doença possuía maior letalidade, auxiliando as entidades governamentais em tomadas de decisões sobre medidas preventivas e expansões hospitalares, além de também serem utilizadas por ferramentas de mídia para informativo a população.

### 1.3 Objetivo

Baseado no cenário pandêmico, vivido nos anos de 2020 e 2021, a facilidade e alta disponibilidade ao acesso de dados da pandemia da COVID-19, aliado a importância da análise dos mesmos para o combate da doença, este trabalho tem como objetivo:

- Realizar uma análise das ferramentas atualmente disponíveis para consulta de dados da pandemia da COVID-19.
- Identificar fontes de dados confiáveis para casos, óbitos, ocupação hospitalar e vacinação.
- Processar e filtrar os dados obtidos.
- Projetar e construir um banco de dados que relacione os dados obtidos.
- Propor, projetar e desenvolver uma plataforma para visualização dos dados obtidos e armazenados na base de dados criada, permitindo ao seu usuário realizar análises correlacionadas.

Onde visamos agregar os seguintes valores a comunidade em geral:

- Facilitar o acesso a informações referente a dados da COVID-19.
- Facilitar a interpretação dos dados.
- Auxiliar entidades governamentais a analisar o reflexo da vacinação sobre a população e traçar planos mais eficientes a cada ano.
- Possibilitar segmentações, tanto geográficas, como temporais, bem específicas.
- Disponibilizar uma base de dados que centralize todas informações referentes a métricas da COVID-19.

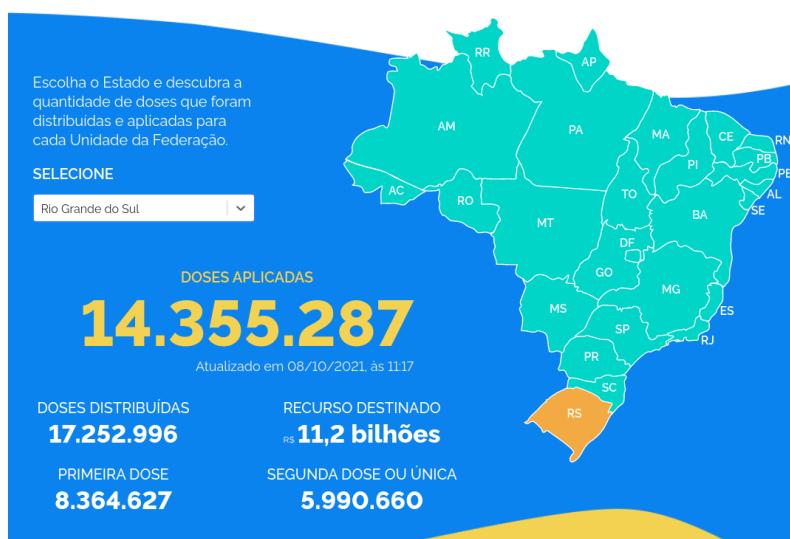
## 2 ESTUDO DE FERRAMENTAS EXISTENTES

A seguir será apresentada a relação das principais soluções da ciência de dados sobre os dados da COVID-19 atualmente disponíveis no Brasil e suas funcionalidades:

### 2.1 Brasil - Pátria Vacinada

O Brasil - Pátria Vacinada (MSB, 2020a) é uma ferramenta informativa, desenvolvida pelo governo federal para ser o portal de informações do ministério da saúde a população brasileira, sobre o estado da vacinação no país contra a COVID-19. Ele foi disponibilizado em <<https://www.gov.br/saude/pt-br/vacinacao>> e utiliza dados oficiais do ministério da saúde do Brasil para fornecer métricas sobre o número de vacinados em todas as modalidades de dose da vacina. Ele possibilita que o usuário visualize os dados de vacinação contra a COVID-19 no Brasil de forma gráfica, possibilitando o filtro por estado da federação (figura 2.1).

Figura 2.1: Interface de dados do Brasil - Pátria Vacinada.



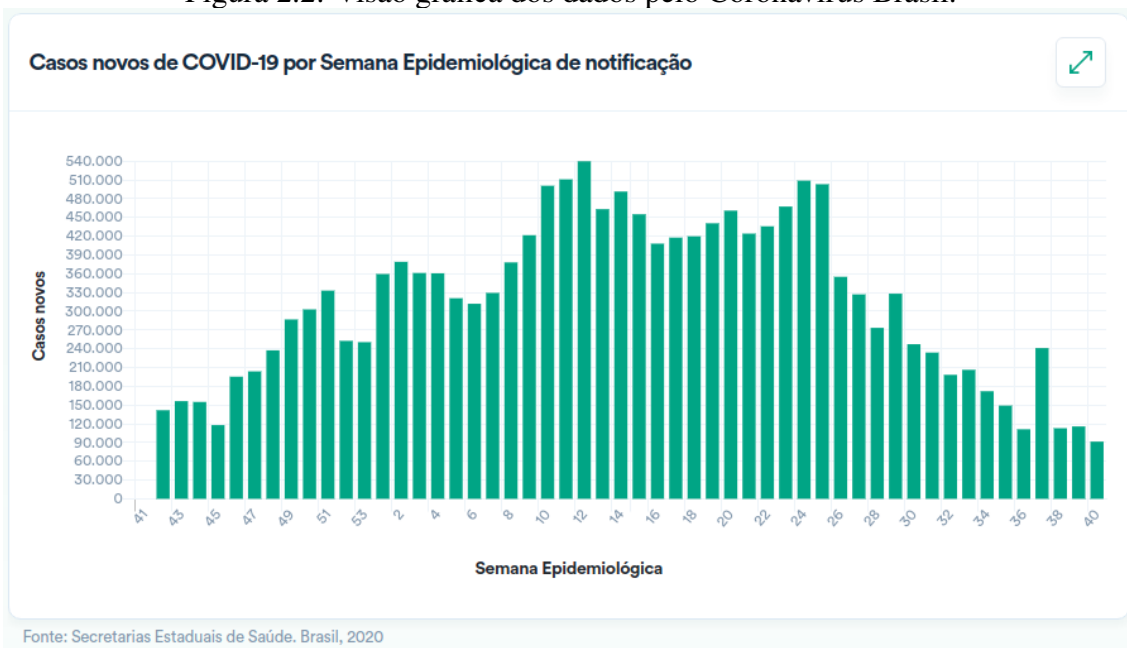
Fonte: Brasil - Pátria Vacinada.

### 2.2 Coronavírus Brasil

O Coronavírus Brasil (MSB, 2020a) é uma ferramenta de visualização de dados, desenvolvida pelo governo federal brasileiro, para ser o portal informativo do ministério da saúde, de métricas da pandemia da COVID-19, à população. Ele foi disponibilizado

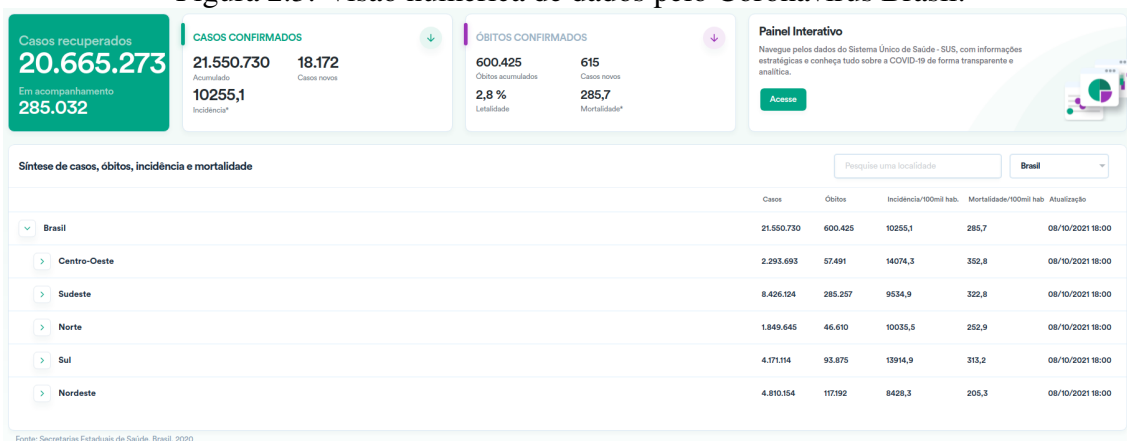
pelo governo brasileiro em <<https://covid.saude.gov.br>> e utiliza dados oficiais do ministério da saúde do Brasil para fornecer métricas diárias e acumuladas sobre casos, óbitos e recuperados, a nível nacional, possibilitando ao usuário a visualização dos dados gráfica e textual (figuras 2.3 e 2.2).

Figura 2.2: Visão gráfica dos dados pelo Coronavírus Brasil.



Fonte: Coronavírus Brasil.

Figura 2.3: Visão numérica de dados pelo Coronavírus Brasil.

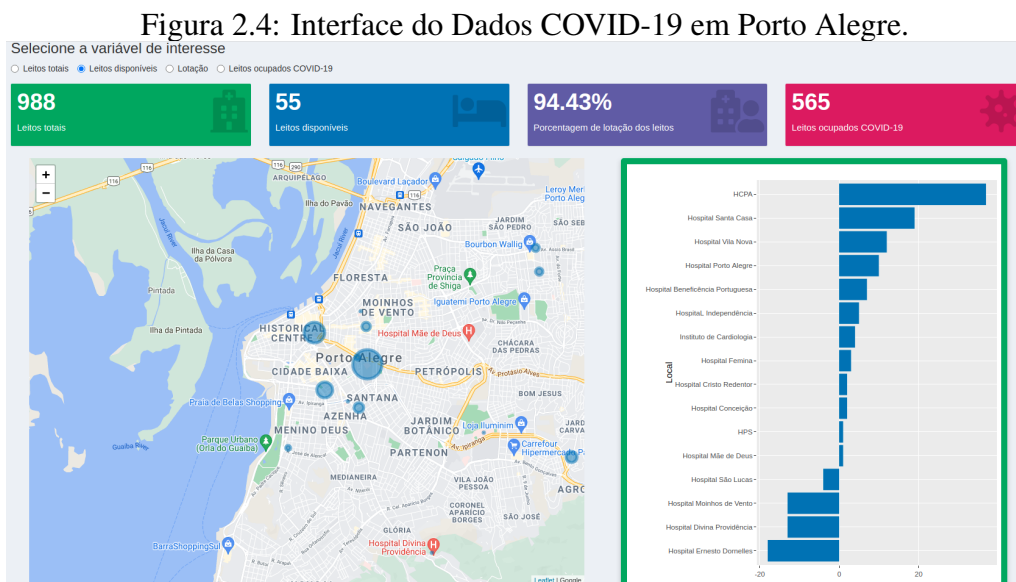


Fonte: Coronavírus Brasil.



### 2.3 Dados COVID-19 em Porto Alegre

O Dados COVID-19 em Porto Alegre (UFRGS, 2020) é uma ferramenta de visualização de dados da pandemia da COVID-19, no âmbito municipal da cidade de Porto Alegre/RS, disponível em <[https://mhbarbian.shinyapps.io/covid19\\_poa](https://mhbarbian.shinyapps.io/covid19_poa)>, foi desenvolvida por um grupo de estudantes e professores do Departamento de estatística da Universidade Federal do Rio Grande do Sul (UFRGS). O Dados COVID-19 em Porto Alegre utiliza dados fornecidos pela prefeitura de Porto Alegre/RS, fornecendo ao seu usuário informações de casos e óbitos por bairro da cidade e ocupação hospitalar em cada hospital da cidade. (figura 2.4).



Fonte: Dados COVID-19 em Porto Alegre.

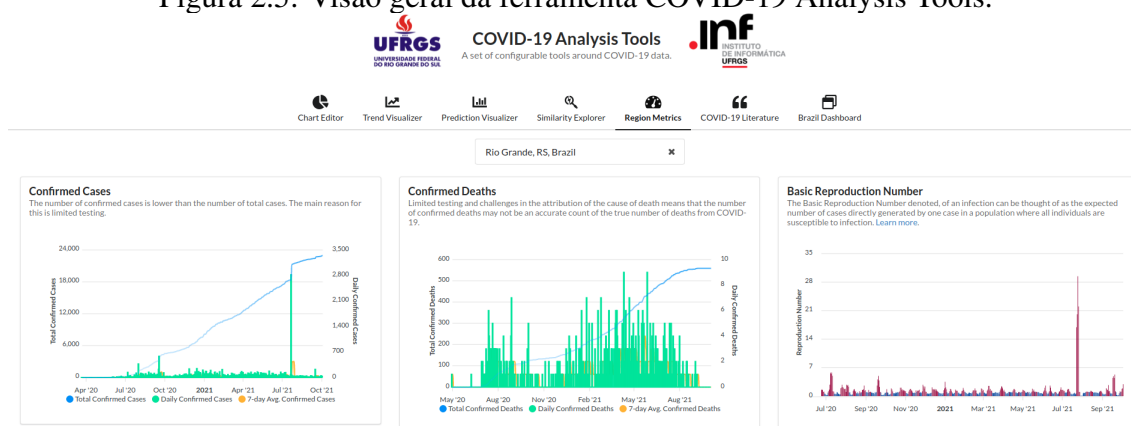
### 2.4 Websensors

O Websensors (USP, 2020b) é uma ferramenta desenvolvida por pesquisadores do Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo (USP), disponível em <<http://websensors.net.br/projects/covid19/>>, realiza a extração de dados de notícias e textos a respeito da pandemia da COVID-19, e a partir de técnicas de mineração de dados e análise de padrões, realiza o pré processamento dos dados e os aplica em uma rede neural treinada com dados de curvas de contágio de outros países, para obter uma estimada predição do contágio no Brasil.

## 2.5 COVID-19 Analysis Tools

A COVID-19 Analysis Tools (UFRGS-UCHILE, 2020) é uma ferramenta de visualização de dados da pandemia da COVID-19 a âmbito mundial, disponível em <<https://covid19.ufrgs.dev/tools/metrics/Brazil.regions.RS:Rio%20Grande>>, foi desenvolvida por um grupo de pesquisa universitário do Instituto de Informática da Universidade Federal do Rio Grande do Sul (UFRGS) (figura 2.5), que provê a visualização gráfica de dados de casos e óbitos acumulados e novos, também realiza a predição da evolução dos dados trabalhados com base em registros passados de localidades que possuem comportamento similar na evolução de métricas da pandemia.

Figura 2.5: Visão geral da ferramenta COVID-19 Analysis Tools.



Fonte: COVID-19 Analysis Tools.

## 2.6 Monitora COVID-19

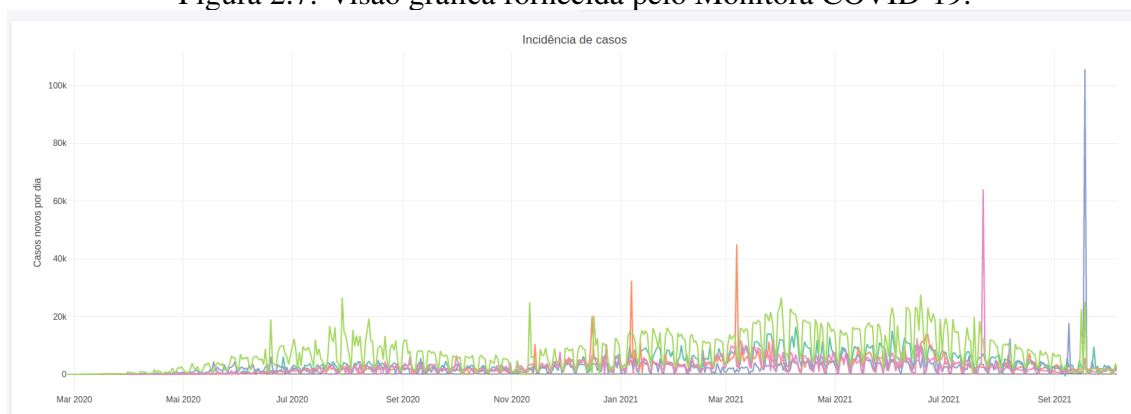
O Monitora COVID-19 (USP, 2020a) é uma ferramenta de visualização de dados da pandemia da COVID-19 (figura 2.6), desenvolvido pela Fundação Oswaldo Cruz (FIOCRUZ) e disponível em <<https://bigdata-covid19a.icict.fiocruz.br>>. Utiliza dados do ministério da saúde do Brasil e de secretarias estaduais de saúde para fornecer ao seu usuário métricas de casos, óbitos e vacinados, tanto acumulados, quanto novos. Apresenta uma serie de filtros para segmentação da pesquisa, mas sem realizar correlação entre dados de casos/óbitos e vacinação (figura 2.7).

Figura 2.6: Visão geral do painel Monitora COVID-19.



Fonte: Monitora COVID-19.

Figura 2.7: Visão gráfica fornecida pelo Monitora COVID-19.



Fonte: Monitora COVID-19.

Além das ferramentas citadas acima, uma série de outras aplicações semelhantes foram desenvolvidas, por entidades de diversas origens, para fornecer informações focadas e direcionadas em sua região, como municípios e estados. Estas aplicações não foram citadas acima pelo caráter semelhante as demais já referenciadas.

### 3 APRESENTAÇÃO DA SOLUÇÃO

Analisando as ferramentas referenciadas no capítulo 2, e disponíveis para livre acesso, podemos verificar a carência em soluções que possibilitem uma análise do impacto da vacinação, sobre a incidência de novos casos, óbitos e internações, podendo vir a auxiliar na definição de estratégias de vacinação em futuras campanhas de imunização.

Com o intuito de sanar esta carência, auxiliando na visualização de dados relacionados a COVID-19, definição e priorização de grupos de vacinação, a serão apresentados os requisitos da plataforma Consulta COVID Brasil, uma plataforma de consulta, visualização e estudo de dados relacionados a indicativos da pandemia da COVID-19 no Brasil, correlacionando dados de diferentes naturezas disponibilizados pelo governo federal brasileiro através de técnicas de ciência de dados.

#### 3.1 Levantamento de requisitos

Os requisitos da aplicação desenvolvida serão apresentados divididos conforme suas classificações, funcionais e não funcionais, tendo o autor na posição de cliente da solução, definindo os requisitos da aplicação a partir de sua visão crítica das ferramentas analisadas no capítulo 2 e da forma como as mídias digitais e os governos vinham disponibilizando a informação.

##### 3.1.1 Requisitos funcionais

Figura 3.1: Filtro de segmentação.

A imagem mostra a interface de usuário para o filtro de segmentação da plataforma 'Consulta Covid Brasil'. No topo, há uma barra verde com dois botões: 'Consultar dados' (destacado) e 'Consultar vacinas'. À direita da barra, o texto 'Consulta Covid Brasil' é exibido. Abaixo, o título 'Consultar dados' é centralizado. O formulário contém os seguintes elementos:

- Um menu suspenso para 'Estado' com o valor 'Todos' selecionado.
- Um menu suspenso para 'Município' com o valor 'Todos' selecionado.
- Dois campos de entrada para o 'Período', com o primeiro contendo '10/10/2020' e o segundo contendo '10/10/2021'.
- Um botão verde 'Buscar' localizado abaixo dos campos de data.

Fonte: O autor.

Tabela 3.1: Requisitos funcionais.

Identificador do requisito	Descrição do requisito
RF-1	Deve possuir uma interface que permita a consulta e dados de casos, óbitos, ocupação hospitalar e vacinação, tanto em valores acumulados como novos diários.
RF-2	Deve disponibilizar o resultado da consulta RF-1 de forma gráfica, em conjunto com a data e o número de ocorrências no pico da métrica amostrada (figura 3.2).
RF-3	Deve possuir uma interface que permita realizar o relacionamento de dados de casos, óbitos e ocupação hospitalar com dados de vacinação.
RF-4	Deve realizar o relacionamento dos dados de casos, óbitos e ocupação hospitalar com dados de vacinação, identificando os grupos de vacinação que quando vacinados tiveram maior reflexo na diminuição das seguintes métricas: Novos casos, novos óbitos, internações e todas as métricas anteriores.
RF-5	Deve disponibilizar a informação da consulta RF-3 de forma gráfica, relacionando as incidências da métrica analisada com o número de vacinados. E também informar o número total de vacinados por dose, média de idade e distribuição percentual de gênero dos vacinados (figura 3.3).
RF-6	Deve possuir informação de todos estados e cidades brasileiras.
RF-7	Deve possibilitar filtros de estado, cidade e período para as consultas RF-2 e RF-3 (figura 3.1), possibilitando uma segmentação a nível regional e municipal da informação.
RF-8	Dever ser acessível via aplicativos de navegação na internet.

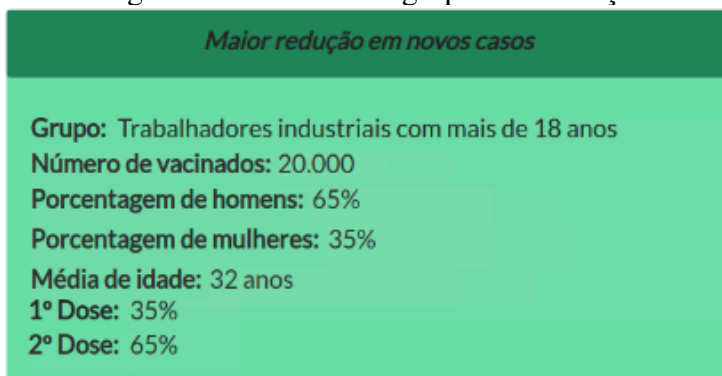
Fonte: O autor.

Figura 3.2: Visão de dados consultados.



Fonte: O autor.

Figura 3.3: Estatísticas grupo de vacinação.



Fonte: O autor.

### 3.1.2 Requisitos não funcionais

Tabela 3.2: Requisitos não funcionais.

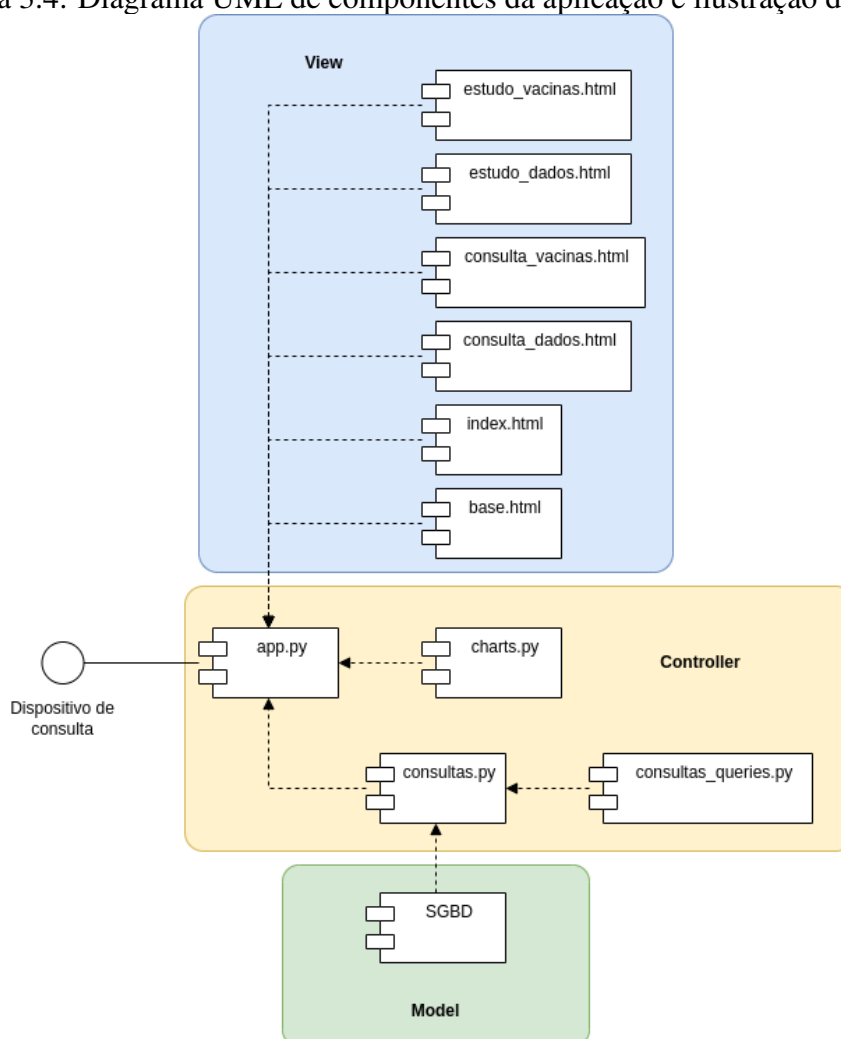
Identificador do requisito	Descrição do requisito
RNF-1	Ser simples e intuitivo.
RNF-2	Possuir baixa latência nas pesquisas.
RNF-3	Ser acessível de diferentes tipos de dispositivos, como por exemplo: Computadores e celulares.
RNF-4	Executar em servidor Linux.

Fonte: O autor.

### 3.2 Arquitetura

A fim de oferecer suporte de acesso a plataforma desenvolvida em diferentes tipos de dispositivos e sistemas operacionais, optou-se pela utilização do modelo de arquitetura MVC, onde teremos uma aplicação centralizada em um servidor, juntamente com os dados consultados, módulos de responsabilidades distintas e bem definidas e a compatibilidade com os diferentes tipos de dispositivos (figura 3.4).

Figura 3.4: Diagrama UML de componentes da aplicação e ilustração do MVC.



Fonte: O autor.

Para implementação da arquitetura MVC, utilizou-se o framework Flask devido a sua capacidade de implementar um servidor de requisições de baixa latência, de forma simples e eficiente e permitir a adoção do SGBD que melhor se encaixe com a demanda da aplicação. A seguir detalharemos a implementação de cada um dos módulos que compõem o MVC da solução:

### 3.2.1 Controller (Back-end)

Para o desenvolvimento do módulo Controller do MVC implementado pelo framework Flask, utilizou-se a linguagem de programação Python, na versão 3.8, em conjunto com uma série de bibliotecas disponíveis para auxílio na manipulação dos dados, são elas:

- **Pandas:** Biblioteca utilizada para manipulação dos dados consultados na base de dados. Oferece uma série de funções que facilitam a manipulação de manipulação de grandes conjuntos de dados, fornecendo rapidez e robustez nas consultas e filtragens.
- **Json:** Biblioteca utilizada para formatar os dados na comunicação entre os módulos controller e View. Esta biblioteca é de grande importância na aplicação devido o framework Flask utilizar o padrão Jinja (JINJA, 2022) para comunicação entre os módulos controller e view, exigindo assim a formatação dos mesmos no forma JSON.
- **Plotly:** Biblioteca especializada na criação de gráficos que podem ser renderizados diretamente no módulo view, utilizando a linguagem de programação Python, através de dados previamente formatados pela biblioteca Pandas. Oferece uma série de funcionalidades que são de grande utilidade na aplicação desenvolvida:
  - Análise de dados em determinado ponto do gráfico correlacionando todos os dados representados pelo mesmo.
  - Análise localizada de dos dados, selecionando um determinado período para zoom.
  - Possibilidade de ocultar métricas do gráfico.
  - Download do gráfico em formato de imagem.
  - Recursos gráficos de expansão para tela cheia.
- **Sqlalchemy:** Biblioteca utilizada para conexão do módulo controller com o SGB escolhido. Optou-se pela utilização da biblioteca Sqlalchemy, mediante tantas outras mais da mesma natureza existentes para a linguagem de programação Python, devido sua compatibilidade com diferentes tipos de SGBDs relacionais e o alto grau de abstração para estabelecer a conexão com o SGBD e realizar consultas no mesmo.



A escolha pela utilização do framework Flask, em conjunto com a linguagem de programação Python se justifica pela sua grande facilidade de manipular grandes volumes de dados, além da familiaridade do desenvolvedor com a tecnologia.

### **3.2.2 View (Front-end)**

Para o desenvolvimento do módulo view do MVC implementado pelo framework Flask, utilizou-se as linguagens de programação HTML, JavaScript e CSS, afim de desenvolver uma interface multiplataforma acessível via navegadores de internet, simples, intuitiva e de baixo recurso computacional do dispositivo utilizado para acesso. Também utilizou-se os frameworks Bootstrap, para obtermos elementos HTML de fins específicos, como formulários e cabeçalhos com efeitos e estilos gráficos pré definidos, e JQuery para identificar e tratar eventos provenientes de elementos HTML da interface.

### **3.2.3 Model (SGBD)**

Para o desenvolvimento do módulo Model do MVC implementado pelo framework Flask, utilizou-se o SGBD relacional PostgreSQL, na versão 13. A escolha pelo PostgreSQL como SGBD para a aplicação desenvolvida justifica-se pelos seguintes aspectos:

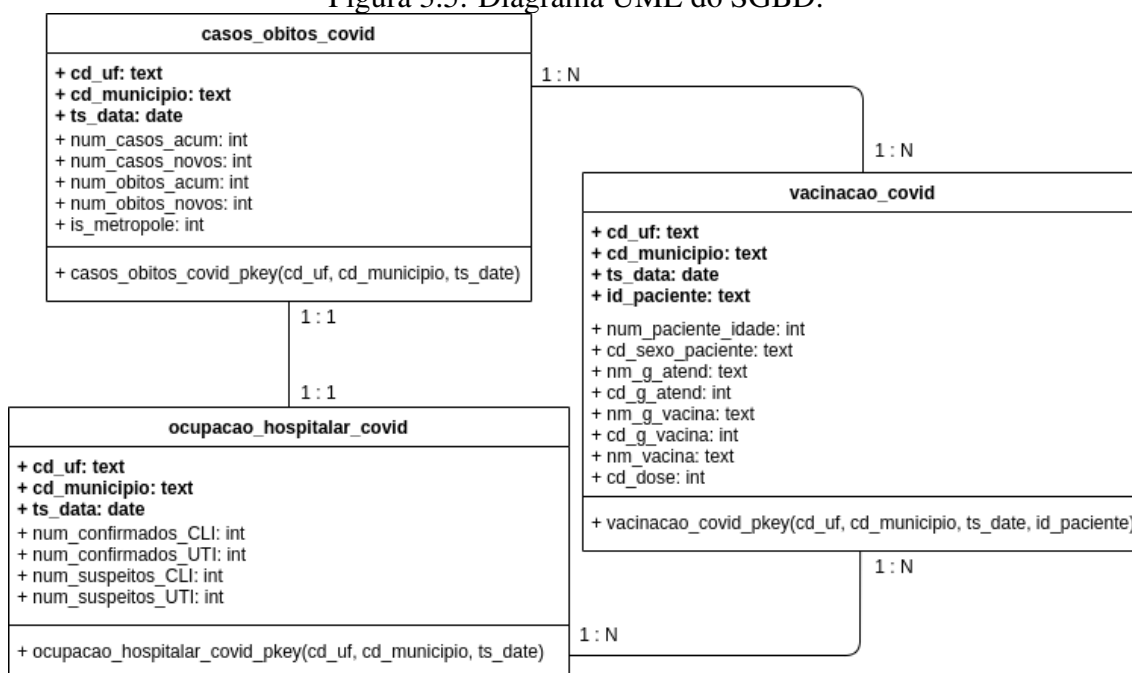
- Suporte a grandes volumes de consultas.
- Suporte a consultas complexas.
- Totalmente gratuito e de código aberto.
- Utilização simples e eficiente, entregando respostas com baixa latência.
- Consultas simples de serem codificadas.

## **3.3 Dados utilizados**

Em virtude de não termos uma base de dados pronta e de acesso aberto disponível, que relacione dados de casos, óbitos, ocupação hospitalar e vacinação relacionados a COVID-19, uma das etapas de maior importância do desenvolvimento da aplicação foi mapear e tratar dados dessa natureza para projeto e criação de uma única base de dados que fornecesse essas informações.

Para criação da base de dados da aplicação foram utilizadas três bases de dados distintas de origem do ministério da saúde do Brasil, de disponibilidade gratuita e aberta, onde cada uma fornece dados de uma natureza para a aplicação e projetado um banco de dados otimizado para correlação de dados e consultas, visto que a natureza da aplicação desenvolvida ser uma plataforma de consulta e visualização de dados, sem efetuar escritas de dados (figura 3.5).

Figura 3.5: Diagrama UML do SGBD.



Fonte: O autor.

Para o correlacionamento dos dados utilizou-se as chaves *cd\_uf*, *cd\_município* e *ts\_date*, presentes em todas as bases de dados utilizadas, possibilitando assim a agregação de informações via operações do tipo *JOIN* utilizando estas colunas e a fácil obtenção de dados de casos, óbitos, ocupação hospitalar e vacinação para qualquer município brasileiro em qualquer data cuja base ofereça suporte, bem como a realização de filtros temporais em consultas a mesma.

É importante salientar que todas as bases de dados utilizadas, após obtidas por vias manuais de suas fontes, necessitaram de tratamentos para eliminação de dados ruidosos, que apresentavam tipagem distinta ou valor totalmente fora do esperado, seleção das colunas de relevância para a aplicação e normalização dos nomes das mesmas. Para estes procedimentos foram utilizados os algoritmos disponíveis nos apêndices A.1, A.2, A.3.

### 3.3.1 Casos e óbitos

Base de dados disponibilizada pelo ministério da saúde na plataforma Coronavírus Brasil, contem dados nacionais, estaduais e municipais, atualizados diariamente. Sua disponibilização ocorre em planilhas no formato CSV, sendo fornecido um arquivo por semestre.

Esta base de dados está modelada de forma que cada linha do arquivo se refere aos números diários de uma cidade específica, seguindo o dicionário de dados para as colunas conforme a tabela 3.3, onde também é sinalizada quais as colunas que foram utilizadas.

Tabela 3.3: Dicionário de dados da base de casos e óbitos.

Nome da coluna	Descrição	Será utilizado
regiao	Região geográfica do Brasil	Não
estado	UF do estado brasileiro	Sim
município	Nome do município	Sim
coduf	Código do estado no IBGE	Não
codmun	Código do município no IBGE	Não
data	Data do registro	Sim
populacaoTCU2019	População conforme senso 2019	Sim
casosAcumulado	Número de casos acumulados	Sim
casosNovos	Número de novos casos no dia	Sim
obitosAcumulado	Número de óbitos acumulados	Sim
obitosNovos	Número de novos óbitos no dia	Sim
interior/metropolitana	Indicador se a cidade é metropole ou interior	Sim

Fonte: O autor.

Para o tratamento de seleção, normalização e carga dos dados desta natureza no banco de dados foi utilizado o algoritmo disponível no apêndice A.1.

### 3.3.2 Vacinação

Base de dados disponibilizada pelo ministério da saúde na plataforma Open Data SUS, contem dados nacionais, estaduais e municipais atualizados diariamente. Sua disponibilização ocorre em uma planilha no formato CSV, sendo fornecido um arquivo com os registros de todo o período de vacinação para cada estado brasileiro.

Tabela 3.4: Dicionário de dados da base de vacinação.

<b>Nome da coluna</b>	<b>Descrição</b>	<b>Será utilizado</b>
paciente_id	Id do paciente	Sim
paciente_idade	Idade do paciente	Sim
paciente_datanascimento	Data de nascimento do paciente	Não
paciente_enumsexobiologico	Sexo do paciente	Sim
paciente_racacor_codigo	Cor de pele do paciente	Não
paciente_racacor_valor	Id de cor de pele do paciente	Não
paciente_endereco_coibgemunicípio	Endereço do paciente	Não
paciente_endereco_copais	Identificador de estrangeiro	Não
paciente_endereco_nmmunicípio	município do paciente	Não
paciente_endereco_nmpais	Pais do paciente	Não
paciente_endereco_uf	Estado do paciente	Não
paciente_endereco_cep	CEP do paciente	Não
paciente_nacionalidade	Nacionalidade do paciente	Não
estabelecimento_valor	Código do local de vacinação	Não
estabelecimento_razaosocial	Razão social do local de vacinação	Não
estabelecimento_nofantasia	Nome fantasia do local de vacinação	Não
estabelecimento_município_codigo	Código IBGE município	Não
estabelecimento_município_nome	Nome município	Sim
estabelecimento_uf	Estado município	Sim
vacina_grupoatendimento_codigo	Código do grupo de atendimento	Sim
vacina_grupoatendimento_nome	Nome do grupo de atendimento	Sim
vacina_categoria_codigo	Código do grupo de vacinação	Sim
vacina_categoria_nome	Nome do grupo de vacinação	Sim
vacina_lote	Lote da vacina	Não
vacina_fabricante_nome	Fabricante da vacina	Sim
vacina_fabricante_referencia	Código do fabricante da vacina	Sim
vacina_dataaplicacao	Data da aplicação	Sim
vacina_descricao_dose	Identificador de 1º ou 2º dose	Sim
vacina_codigo	Código da vacina	Sim
vacina_nome	Nome da vacina	Sim
sistema_origem	Sistema de origem	Não
data_importacao_rnds	Data da importação	Não

Esta base de dados está modelada de forma que cada linha do arquivo se refere a vacinação de um paciente, seguindo o dicionário de dados para as colunas conforme a tabela 3.4, onde também é sinalizada quais as colunas que foram utilizadas na ferramenta desenvolvida.

Para o tratamento de seleção, normalização e carga dos dados desta natureza no banco de dados foi utilizado o algoritmo disponível no apêndice A.3.

### 3.3.3 Ocupação hospitalar

Base de dados disponibilizada pelo ministério da saúde na plataforma Open Data SUS, contem dados nacionais, estaduais e municipais atualizados diariamente. Sua disponibilização ocorre em uma planilha no formato CSV, sendo fornecido um arquivo com os registros do período todo de pandemia.

Esta base de dados está modelada de forma que cada linha do arquivo se refere aos números diários de uma cidade específica, seguindo o dicionário de dados para as colunas conforme a tabela 3.5, onde também é sinalizada quais as colunas que foram utilizadas na ferramenta desenvolvida.

Tabela 3.5: Dicionário de dados da base de ocupação hospitalar.

Nome da coluna	Descrição	Será utilizado
dataNotificacao	Data de notificação	Sim
ocupacaoSuspeitoCli	Número suspeitos CLI	Sim
ocupacaoSuspeitoUti	Número suspeitos UTI	Sim
ocupacaoConfirmadoCli	Número confirmados CLI	Sim
ocupacaoConfirmadoUti	Número confirmados UTI	Sim
saidaSuspeitaObitos	Saída por óbito suspeito	Não
saidaSuspeitaAltas	Saída por alta suspeito	Não
saidaConfirmadaObitos	Saída por óbito confirmado	Não
saidaConfirmadaAltas	Saída por alta confirmado	Não
estadoNotificacao	UF do estado de notificação	Sim
municípioNotificacao	Município de notificação	Sim
estado	UF estado de nesciência	Não
município	Município estado de nesciência	Não

Fonte: O autor.

Para o tratamento de seleção, normalização e carga dos dados desta natureza no banco de dados foi utilizado o algoritmo disponível no apêndice A.2.

Realizando uma análise prévia desta base de dados, observa-se que a característica da centralização dos dados desta natureza em grandes metrópoles regionais, fato que demonstra a busca por tratamento em centros de maiores referências e até mesmo a busca por tratamento em cidades vizinhas, por moradores de cidades que não possuem leito hospitalar.

## 4 RESULTADOS

Visando disponibilizar livre acesso à aplicação desenvolvida, utilizou-se o serviço de hospedagem para aplicações *Heroku*, nele a aplicação desenvolvida encontra-se acessível por meio do endereço <<http://consultacovidbrasil.herokuapp.com/>> e pode ser utilizada seguindo as orientações disponíveis no apêndice B.

A versão disponibilizada possui dados de casos, óbitos e ocupação hospitalar disponíveis para todos municípios brasileiros, entre os períodos de fevereiro de 2020 a abril de 2022, e dados de vacinação limitados os estados do Acre, Roraima e Amapá, tal limitação ocorre devido o volume de dados total de vacinação para todos estados brasileiros totaliza mais de 1Tb de dados, tornando inviável o armazenamento deste volume de dados em um serviço de hospedagem de baixo custo.

A seguir será apresentado cada uma das funcionalidades da aplicação desenvolvida.

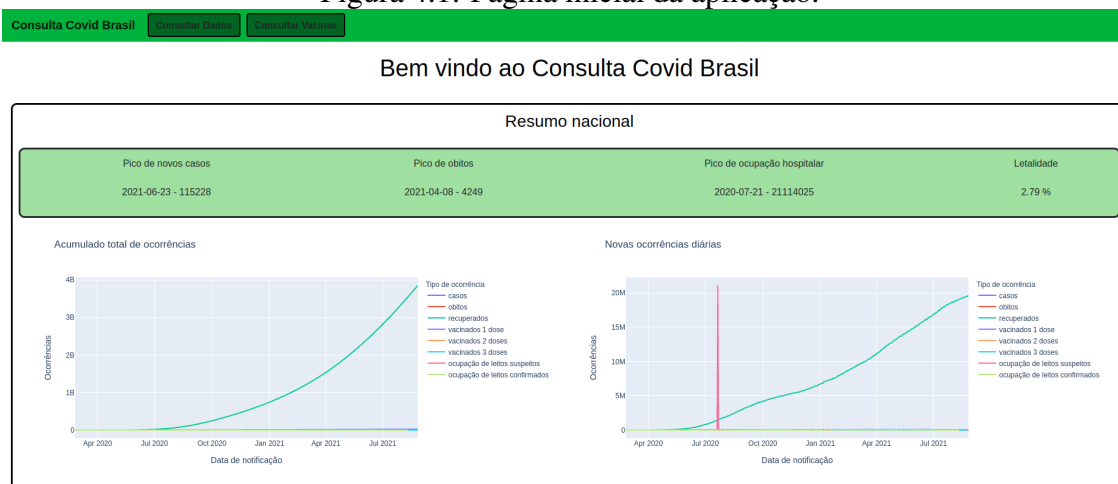
### 4.1 Funcionalidades

#### 4.1.1 Página inicial

Ao acessar a plataforma Consulta Covid Brasil o usuário terá acesso a página inicial da plataforma (figura 4.1), onde, para âmbito nacional e agrupado por regiões geográficas, são disponibilizadas métricas gerais da pandemia como:

- Data e número de ocorrências do pico de novos casos.
- Data e número de ocorrências de pico de óbitos.
- Data e número de ocorrências do pico de ocupação hospitalar.
- Letalidade média (razão entre óbitos e casos acumulados).
- Gráfico de ocorrências acumuladas e diárias para:
  - Casos
  - Óbitos
  - Recuperados
  - Vacinados em 1, 2 e 3 dose
  - Ocupação hospitalar por confirmados e suspeitos

Figura 4.1: Página inicial da aplicação.



Fonte: O autor.

#### 4.1.2 Consulta de dados

Figura 4.2: Tela de filtragem e seleção para consulta de dados.

Fonte: O autor.

Ao acessar, no menu da página inicial, a funcionalidade *Consultar Dados* o usuário terá acesso a uma interface que oferece uma gama de filtros (figura 4.2) tanto geográficos, possibilitando filtrar a consulta desde o nível nacional até o nível municipal, como filtros temporais, possibilitando indicar um determinado espaço de tempo no qual se deseja realizar a consulta.

Após definir os filtros desejados e realizar a consulta, será realizada uma consulta simples de relacionamento e filtragem das diferentes bases de dados citadas na seção 3.3 e disponibilizando ao usuário as métricas listadas a seguir e ilustradas na figura 4.3.

- Parâmetros de pesquisa.
- Data e número de ocorrências do pico de novos casos.



- Data e número de ocorrências de pico de óbitos.
- Data e número de ocorrências do pico de ocupação hospitalar.
- Letalidade média (razão entre óbitos e casos acumulados).
- Gráfico de ocorrências acumuladas e diárias para:
  - Casos
  - Óbitos
  - Recuperados
  - Vacinados em 1, 2 e 3 dose
  - Ocupação hospitalar por confirmados e suspeitos

Figura 4.3: Resultado da consulta de dados.



Fonte: O autor.

#### 4.1.3 Consulta sobre reflexo da vacinação

Ao acessar, no menu da página inicial, a funcionalidade *Consultar Vacinas* o usuário terá acesso a uma interface que oferece uma gama de filtros (figura 4.4) tanto geográficos, possibilitando filtrar a consulta desde o nível nacional até o nível municipal, como filtros temporais, possibilitando indicar um determinado espaço de tempo no qual se deseja realizar a consulta.

Figura 4.4: Tela de filtragem e seleção para consulta sobre reflexo da vacinação.

Fonte: O autor.

Figura 4.5: Resultado da consulta sobre reflexo da vacinação.



Fonte: O autor.

Após definir os filtros desejados e realizar a consulta, utilizando a eurística de identificar o grupo vacinal de maior redução na métrica analisada, observando uma janela temporal de 30 dias anterior a cada dia do período analisado, onde cada pessoa contabilizada está vacinada a mais de 15 dias (BUTANTAN, 2021), será realizada uma consulta de relacionamento e filtragem das diferentes bases de dados citadas na seção 3.3 e disponibilizando ao usuário, para cada um dos indicadores: Gerais, casos, óbitos e internações as métricas listadas a seguir e ilustradas na figura 4.5.

- Parâmetros de pesquisa.
- Grupo vacinal de maior redução em métricas.
- Total de vacinados.
- Percentual masculino e feminino. (distribuição por sexo).

- Média de idade dos vacinados.
- Número de vacinados em cada dose.
- Gráfico de ocorrências acumuladas para:
  - Casos
  - Óbitos
  - Vacinados em 1, 2 e 3 dose pertencentes ao grupo sinalizado
  - Ocupação hospitalar por confirmados e suspeitos

## 4.2 Comparativos

Após finalizada a implementação da plataforma proposta e detalhada nos capítulos, 3 e 4 deste documento, podemos realizar a comparação da mesma com as soluções já existentes e citadas no capítulo 2, em quatro características principais:

- Abrangência e origem dos dados (tabela 4.2).
- Completude de dados (tabela 4.3).
- Capacidade de segmentação (tabela 4.4).
- Funcionalidades finais (tabela 4.1).

Tabela 4.1: Comparação de funcionalidades oferecidos.

Ferramenta	Análise e identificação do reflexo da vacinação	Formato de disponibilidade da informação
Consulta Covid Brasil	Sim	Texto e gráficos
Coronavírus Brasil	Não	Texto e gráficos
Brasil - Pátria Vacinada	Não	Texto e gráficos
COVID-19 em Porto Alegre	Não	Texto e gráficos
Websensors	Não	Texto e gráficos
COVID-19 Analysis Tools	Não	Texto e gráficos
Monitora COVID-19	Não	Texto e gráficos

Fonte: O autor.

Tabela 4.2: Comparação de abrangência geográfica e origem dos dados.

Ferramenta	Abrangência	Origem dos dados
Consulta Covid Brasil	Nacional	MSB
Coronavírus Brasil	Nacional	MSB
Brasil - Pátria Vacinada	Nacional	MSB
COVID-19 em Porto Alegre	Municipal	Prefeitura
Websensors	Nacional	MSB
COVID-19 Analysis Tools	Mundial	Notícias
Monitora COVID-19	Nacional	MSB

Fonte: O autor.

Tabela 4.3: Comparação de dados utilizados.

Ferramenta	Casos	Óbitos	Ocupação Hospitalar	Vacinação
Consulta Covid Brasil	Sim	Sim	Sim	Sim
Coronavírus Brasil	Sim	Sim	Sim	Não
Brasil - Pátria Vacinada	Não	Não	Não	Sim
COVID-19 em Porto Alegre	Sim	Sim	Sim	Não
Websensors	Sim	Sim	Não	Não
COVID-19 Analysis Tools	Sim	Sim	Não	Não
Monitora COVID-19	Sim	Sim	Sim	Sim

Fonte: O autor.

Tabela 4.4: Comparação de filtros oferecidos.

Ferramenta	Filtro por período	Menor seguimentação geográfica
Consulta Covid Brasil	Sim	Municipal
Coronavírus Brasil	Não	Estadual
Brasil - Pátria Vacinada	Não	Estadual
COVID-19 em Porto Alegre	Não	Municipal
Websensors	Sim	Estadual
COVID-19 Analysis Tools	Sim	Estadual
Monitora COVID-19	Sim	Estadual

Fonte: O autor.

## 5 CONCLUSÃO

Ao longo deste trabalho se identificou uma limitação existente em plataformas de acesso a informações e dados da pandemia da COVID-19, onde analisou-se uma série de soluções existentes e mapeou-se pontos fortes e deficiências de cada uma delas, visando propor e construir uma plataforma de consulta a informações referentes a pandemia da COVID-19, que apresenta soluções e preenche as lacunas de funcionalidade das demais soluções.

Com base na análise do resultado obtido através das comparações citadas na seção 4.2, realizada entre as diferentes ferramentas da mesma natureza, podemos concluir que a plataforma originada deste trabalho atendeu os objetivos propostos, apresentando soluções para todos pontos de carência das ferramentas atualmente disponíveis e possui grande potencial de auxiliar profissionais da saúde e órgãos governamentais, a traçar planos de vacinação cada vez mais eficazes, conforme a agregação de novos dados a cada ano.

Além disso, é possível identificar que a plataforma resultante deste trabalho, ainda possui grande potencial e espaço de crescimento, evolução e refinamento de suas funcionalidades, podendo dar origem a novos trabalhos de pesquisa e extensão, orientados a demandas específicas.

## REFERÊNCIAS

- BBC. Vacinas contra a covid-19. In: <https://www.bbc.com/portuguese/brasil-57124616>. [S.l.]: BBC New Brasil, 2021.
- BRITO, A. M. S. de. Ciênciadados:ferramentasparacombate à covid-19. In: <https://www.nucleodoconhecimento.com.br/tecnologia/ciencia-de-dados>. [S.l.]: Núcleo do conhecimento, 2020.
- BUTANTAN. Quanto tempo demora para adquirir imunidade após tomar a vacina contra a covid-19? In: <https://butantan.gov.br/noticias/quanto-tempo-demora-para-adquirir-imunidade-apos-tomar-a-vacina-contra-a-covid-19>: :text=Em%20geral[S.l.]: BUTANTAN, 2021.
- FIOCRUZ, P. Brasil celebra um ano da vacina contra a covid-19. In: <https://portal.fiocruz.br/noticia/brasil-celebra-um-ano-da-vacina-contra-covid-19>. [S.l.]: Portal FIOCRUZ, 2022.
- GRUBER, A. Covid-19: o que se sabe sobre a origem da doença. In: <https://jornal.usp.br/artigos/covid2-o-que-se-sabe-sobre-a-origem-da-doenca>. [S.l.]: Jornal da USP, 2020.
- JINJA. Jinja documentation. In: <https://jinja.palletsprojects.com/en/3.1.x/>. [S.l.]: JINJA, 2022.
- MSB. Coronavirus brasil. In: <https://covid.saude.gov.br>. [S.l.]: Governo Brasileiro - Ministerio da Saúde, 2020.
- MSB. Primeiro caso de covid-19 no brasil. In: <https://www.gov.br/saude/pt-br/assuntos/noticias/2020/julho/primeiro-caso-de-covid-19-no-brasil-permanecendo-o-de-26-de-fevereiro>. [S.l.]: Governo Brasileiro - Ministerio da Saúde, 2020.
- UFRGS. Dados covid-19 em porto alegre. In: [https://mhbarbian.shinyapps.io/covid19\\_poa](https://mhbarbian.shinyapps.io/covid19_poa). [S.l.]: Departamento de estatística, 2020.
- UFRGS-UCHILE. Covid-19 analysis tools. In: <https://covid19.ufrgs.dev/tools/metrics/Brazil.regions.RS:Rio%20Grande>. [S.l.]: Instituto de informatica, 2020.
- USP. Monitoracovid-19. In: <https://bigdata-covid19.iciet.fiocruz.br>. [S.l.]: Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT), 2020.
- USP. Websensors. In: <http://websensors.net.br/projects/covid19>. [S.l.]: Instituto de Ciências Matemáticas e de Computação, 2020.

## ApêndiceA PROCESSAMENTO DE DADOS

### A.1 Casos e óbitos

O algoritmo a seguir, codificado na linguagem de programação Python, foi utilizado para realizar a manipulação dos dados brutos de casos e óbitos e carregá-los no banco de dados utilizado.

```
import pandas as pd
import psycopg2
from unidecode import unidecode

# Declaracao de variaveis de parametro -- ALTERAR
n_base_casos_covid_primeiro_semestre_2020 = 'file_path'
n_base_casos_covid_segundo_semestre_2020 = 'file_path'
n_base_casos_covid_primeiro_semestre_2021 = 'file_path'
n_base_casos_covid_segundo_semestre_2021 = 'file_path'
n_base_casos_covid_primeiro_semestre_2022 = 'file_path'
arg_ip = 'ip_db'
arg_name = 'db_name'
arg_user = 'db_user'
arg_pass = 'db_pass'

ufs = [('rio grande do sul', 'rs'),
       ('santa catarina', 'sc'),
       ('parana', 'pr'),
       ('sao paulo', 'sp'),
       ('rio de janeiro', 'rj'),
       ('minas gerais', 'mg'),
       ('mato grosso do sul', 'ms'),
       ('mato grosso', 'mt'),
       ('goias', 'go'),
       ('espirito santo', 'es'),
       ('bahia', 'ba'),
       ('alagoas', 'al'),
       ('tocantins', 'to'),
       ('rondonia', 'ro'),
```

```
('acres', 'ac'),  
( 'amazonas', 'am'),  
( 'roraima', 'rr'),  
( 'para', 'pa'),  
( 'amapa', 'ap'),  
( 'maranhao', 'ma'),  
( 'sergipe', 'se'),  
( 'piaui', 'pi'),  
( 'ceara', 'ce'),  
( 'pernambuco', 'pe'),  
( 'rio grande do norte', 'rn'),  
( 'paraiba', 'pb')]
```

```
base_covid_p_sem_2020 = pd.read_csv(  
    n_base_casos_covid_primeiro_semestre_2020,  
    error_bad_lines=False, sep=';')  
base_covid_s_sem_2020 = pd.read_csv(  
    n_base_casos_covid_segundo_semestre_2020,  
    error_bad_lines=False, sep=';')  
base_covid_p_sem_2021 = pd.read_csv(  
    n_base_casos_covid_primeiro_semestre_2021,  
    error_bad_lines=False, sep=';')  
base_covid_s_sem_2021 = pd.read_csv(  
    n_base_casos_covid_segundo_semestre_2021,  
    error_bad_lines=False, sep=';')  
base_covid_9_sem_2022 = pd.read_csv(  
    n_base_casos_covid_primeiro_semestre_2022,  
    error_bad_lines=False, sep=';')  
  
base_covid = pd.concat([base_covid_p_sem_2020,  
                        base_covid_s_sem_2020,  
                        base_covid_p_sem_2021,  
                        base_covid_s_sem_2021,  
                        base_covid_9_sem_2022],  
                        ignore_index=True)
```



```

base_covid_br = base_covid.loc[base_covid['regiao'] == 'Brasil'] [
    ['data', 'semanaEpi',
     'populacaoTCU2019', 'casosAcumulado',
     'casosNovos', 'obitosAcumulado',
     'obitosNovos', 'Recuperadosnovos',
     'emAcompanhamentoNovos']]

base_covid_br.columns = ['dt_notificacao', 'num_semana',
                          'num_popul_censo', 'num_casos_acum',
                          'num_casos_novos', 'num_obitos_acum',
                          'num_obitos_novos', 'num_recup_novos',
                          'num_em_acomp_novos']

base_covid_br_2 = base_covid_br.fillna(-1)

conn = psycopg2.connect(
    f"host={arg_ip} dbname={arg_name} user={arg_user} password={arg_pass}")

cur = conn.cursor()
for index, row in base_covid_br_2.iterrows():
    cur.execute(
        f"""INSERT INTO covid.casos_covid_br VALUES (
            TO_DATE('{row[0]}', 'YYYY-MM-DD'),
            {row[1]},{row[2]},{row[3]},{row[4]},
            {row[5]},{row[6]},{row[7]},{row[8]}
        ) ON CONFLICT ON CONSTRAINT casos_covid_br_pk DO NOTHING"""
    )
conn.commit()

base_covid_estados = base_covid.loc[
    (base_covid['regiao'] != 'Brasil') &
    (base_covid['município'].isnull() == True)] [
    ['estado', 'coduf', 'data', 'semanaEpi',
     'populacaoTCU2019', 'casosAcumulado',
     'casosNovos', 'obitosAcumulado', 'obitosNovos',
     'Recuperadosnovos', 'emAcompanhamentoNovos']]

```

```

base_covid_estados.columns = ['ds_uf', 'cd_uf', 'dt_notificacao',
                              'num_semana', 'num_popul_censo',
                              'num_casos_acum', 'num_casos_novos',
                              'num_obitos_acum', 'num_obitos_novos',
                              'num_recup_novos', 'num_em_acomp_novos']

base_covid_estados['ds_uf'] = base_covid_estados[
    'ds_uf'].str.lower()
base_covid_estados['ds_uf'] = base_covid_estados[
    'ds_uf'].apply(lambda x: unicode(x))

for old_uf, new_uf in ufs:
    base_covid_estados['ds_uf'] = base_covid_estados[
        'ds_uf'].str.replace(old_uf, new_uf)

base_covid_estados = base_covid_estados[
    base_covid_estados['ds_uf'].notna()]
base_covid_estados = base_covid_estados[
    base_covid_estados['dt_notificacao'].notna()]
base_covid_estados_2 = base_covid_estados.fillna(-1)

conn = psycopg2.connect(
    f"host={arg_ip} dbname={arg_name} user={arg_user} password={arg_pass}")

cur = conn.cursor()
for index, row in base_covid_estados_2.iterrows():
    cur.execute(
        f"""INSERT INTO covid.casos_covid_estados VALUES (
            '{row[0]}', {row[1]}, TO_DATE('{row[2]}', 'YYYY-MM-DD'),
            {row[3]}, {row[4]}, {row[5]}, {row[6]},
            {row[7]}, {row[8]}, {row[9]}, {row[10]}
        ) ON CONFLICT ON CONSTRAINT casos_covid_estados_pk DO NOTHING"""
    )
conn.commit()

```

```

base_covid_municipios = base_covid.loc[
    (base_covid['regiao'] != 'Brasil') &
    (base_covid['municipio'].isnull() == False)][
    ['estado', 'coduf', 'municipio', 'codmun',
     'data', 'semanaEpi', 'populacaoTCU2019',
     'casosAcumulado', 'casosNovos', 'obitosAcumulado',
     'obitosNovos', 'Recuperadosnovos',
     'emAcompanhamentoNovos', 'interior/metropolitana']]

base_covid_municipios.columns = ['ds_uf', 'cd_uf', 'ds_municipio',
                                  'cd_municipio', 'dt_notificacao',
                                  'num_semana', 'num_popul_censo',
                                  'num_casos_acum', 'num_casos_novos',
                                  'num_obitos_acum', 'num_obitos_novos',
                                  'num_recup_novos', 'num_em_acomp_novos',
                                  'is_metropole']

base_covid_municipios['ds_uf'] = base_covid_municipios[
    'ds_uf'].str.lower()
base_covid_municipios['ds_uf'] = base_covid_municipios[
    'ds_uf'].apply(lambda x: unidecode(x))

base_covid_municipios[
    'ds_municipio'] = base_covid_municipios['ds_municipio'].str.lower()
base_covid_municipios[
    'ds_municipio'] = base_covid_municipios[
    'ds_municipio'].apply(lambda x: unidecode(x))
base_covid_municipios[
    'ds_municipio'] = base_covid_municipios[
    'ds_municipio'].str.replace("\\'", "'")

for old_uf, new_uf in ufs:
    base_covid_municipios[
        'ds_uf'] = base_covid_municipios[
        'ds_uf'].str.replace(old_uf, new_uf)

```

```

base_covid_municipios = base_covid_municipios[
    base_covid_municipios['ds_municipio'].notna()]
base_covid_municipios = base_covid_municipios[
    base_covid_municipios['ds_uf'].notna()]
base_covid_municipios = base_covid_municipios[
    base_covid_municipios['dt_notificacao'].notna()]
base_covid_municipios_2 = base_covid_municipios.fillna(-1)

conn = psycopg2.connect(
    f"host={arg_ip} dbname={arg_name} user={arg_user} password={arg_pass}")

cur = conn.cursor()
for index, row in base_covid_municipios_2.iterrows():
    cur.execute(
        f"""INSERT INTO covid.casos_covid_municipios VALUES (
            '{row[0]}', '{row[2]}', TO_DATE('{row[4]}', 'YYYY-MM-DD'),
            {row[5]}, {row[6]}, {row[7]}, {row[8]}, {row[9]},
            {row[10]}, {row[11]}, {row[12]}, {row[13]}
        ) ON CONFLICT ON CONSTRAINT casos_covid_municipios_pk DO NOTHING"""
    )
conn.commit()

```

## A.2 Ocupação hospitalar

O algoritmo a seguir, codificado na linguagem de programação Python, foi utilizado para realizar a manipulação dos dados brutos de ocupação hospitalar e carregá-los no banco de dados utilizado.

```

import pandas as pd
import psycopg2
from unidecode import unidecode

# Declaracao de variaveis de parametro - ALTERAR
n_base_ocupacao_2020 = 'file_path'
n_base_ocupacao_2021 = 'file_path'
n_base_ocupacao_2022 = 'file_path'

```



```

base_2022 = pd.read_csv (n_base_ocupacao_2022,
    error_bad_lines=False, sep=',')
base_ocupacao = pd.concat (
    [base_2020,
    base_2021,
    base_2022],
    ignore_index=True)

base_ocupacao_municipios = base_ocupacao[
    ['dataNotificacao', 'ocupacaoSuspeitoCli',
    'ocupacaoSuspeitoUti', 'ocupacaoConfirmadoCli',
    'ocupacaoConfirmadoUti', 'saidaSuspeitaObitos',
    'saidaSuspeitaAltas', 'saidaConfirmadaObitos',
    'saidaConfirmadaAltas', 'municipioNotificacao',
    'estadoNotificacao']]

base_ocupacao_municipios.columns = ['dt_notificacao',
    'num_ocup_susp_cti', 'num_ocup_susp_uti',
    'num_ocup_confir_cti', 'num_ocup_confir_uti',
    'saida_susp_obito', 'saida_susp_alta',
    'saida_confir_obito', 'saida_confir_alta',
    'ds_municipio', 'ds_uf']

base_ocupacao_municipios[
    'ds_municipio'] = base_ocupacao_municipios[
    'ds_municipio'].str.lower()
base_ocupacao_municipios[
    'ds_municipio'] = base_ocupacao_municipios[
    'ds_municipio'].apply(lambda x : unidecode(x))
base_ocupacao_municipios[
    'ds_municipio'] = base_ocupacao_municipios[
    'ds_municipio'].str.replace("\\'", "")
base_ocupacao_municipios[
    'ds_uf'] = base_ocupacao_municipios['ds_uf'].str.lower()
base_ocupacao_municipios[
    'ds_uf'] = base_ocupacao_municipios[

```

```

    'ds_uf'].apply(lambda x : unicode(x))
base_ocupacao_municipios = base_ocupacao_municipios[
    base_ocupacao_municipios['dt_notificacao'].notna()]
base_ocupacao_municipios[
    'dt_notificacao'] = base_ocupacao_municipios[
    'dt_notificacao'].apply(lambda x : str(x).split('T')[0])
for old_uf, new_uf in ufs:
    base_ocupacao_municipios[
        'ds_uf'] = base_ocupacao_municipios[
            'ds_uf'].str.replace(old_uf, new_uf)

base_ocupacao_municipios = base_ocupacao_municipios[
    base_ocupacao_municipios['ds_municipio'].notna()]
base_ocupacao_municipios = base_ocupacao_municipios[
    base_ocupacao_municipios['ds_uf'].notna()]
base_ocupacao_municipios = base_ocupacao_municipios[
    base_ocupacao_municipios['dt_notificacao'].notna()]
base_ocupacao_municipios_2 = base_ocupacao_municipios.fillna(-1)

conn = psycopg2.connect(
    f"host={arg_ip} dbname={arg_name} user={arg_user} password={arg_pass}")

cur = conn.cursor()
for index, row in base_ocupacao_municipios_2.iterrows():
    cur.execute(
        f"""INSERT INTO covid.ocupacao VALUES (
            TO_DATE('{row[0]}', 'YYYY-MM-DD'),{row[1]},{row[2]},
            {row[3]},{row[4]},{row[5]},{row[6]},
            {row[7]},{row[8]},{row[9]},{row[10]}
        ) ON CONFLICT ON CONSTRAINT ocupacao_pk DO NOTHING"""
    )
conn.commit()

```

### A.3 Vacinação

O algoritmo a seguir, codificado na linguagem de programação Python, foi utilizado para realizar a manipulação dos dados brutos de vacinação e carregá-los no banco de dados utilizado.

```
import pandas as pd
import psycopg2
from unidecode import unidecode
import os

path = '/home/fsalmeida/Pictures/vacinacao/amapa'
arg_ip = 'localhost'
arg_name = 'covid'
arg_user = 'postgres'
arg_pass = 'admin'

ufs = [('rio grande do sul', 'rs'),
       ('santa catarina', 'sc'),
       ('parana', 'pr'),
       ('sao paulo', 'sp'),
       ('rio de janeiro', 'rj'),
       ('minas gerais', 'mg'),
       ('mato grosso do sul', 'ms'),
       ('mato grosso', 'mt'),
       ('goias', 'go'),
       ('espirito santo', 'es'),
       ('bahia', 'ba'),
       ('alagoas', 'al'),
       ('tocantins', 'to'),
       ('rondonia', 'ro'),
       ('acre', 'ac'),
       ('amazonas', 'am'),
       ('roraima', 'rr'),
       ('para', 'pa'),
       ('amapa', 'ap'),
       ('maranhao', 'ma'),
```



```

        ('sergipe', 'se'),
        ('piaui', 'pi'),
        ('ceara', 'ce'),
        ('pernambuco', 'pe'),
        ('rio grande do norte', 'rn'),
        ('paraiba', 'pb')]

files = os.listdir(path)

base_vacinacao = pd.DataFrame(
    columns=['num_idade', 'cd_sexo', 'ds_raca',
            'ds_municipio', 'ds_uf', 'cd_grupo',
            'cd_categoria', 'dt_aplicacao',
            'num_dose', 'cd_vacina'])

for file in files:
    n_base_vacinacao = f"{path}/{file}"
    base_vacinacao_body = pd.read_csv (
        n_base_vacinacao, error_bad_lines=False,
        header=None, skiprows=1, sep=';')
    base_vacinacao_body = base_vacinacao_body.loc[
        :, [2, 4, 6, 18, 19, 20, 22, 27, 28, 29]]
    base_vacinacao_body = pd.DataFrame(
        base_vacinacao_body.values, columns=[
            'num_idade', 'cd_sexo', 'ds_raca',
            'ds_municipio', 'ds_uf', 'cd_grupo',
            'cd_categoria', 'dt_aplicacao',
            'num_dose', 'cd_vacina'])
    base_vacinacao_body = base_vacinacao_body[
        base_vacinacao_body['num_idade'].notna()]
    base_vacinacao_body = base_vacinacao_body[
        base_vacinacao_body['cd_sexo'].notna()]
    base_vacinacao_body = base_vacinacao_body[
        base_vacinacao_body['ds_raca'].notna()]
    base_vacinacao_body = base_vacinacao_body[
        base_vacinacao_body['ds_municipio'].notna()]

```

```

base_vacinacao_body = base_vacinacao_body[
    base_vacinacao_body['ds_uf'].notna()]
base_vacinacao_body = base_vacinacao_body[
    base_vacinacao_body['cd_grupo'].notna()]
base_vacinacao_body = base_vacinacao_body[
    base_vacinacao_body['cd_categoria'].notna()]
base_vacinacao_body = base_vacinacao_body[
    base_vacinacao_body['dt_aplicacao'].notna()]
base_vacinacao_body[
    'ds_uf'] = base_vacinacao_body['ds_uf'].str.lower()
base_vacinacao_body[
    'ds_uf'] = base_vacinacao_body[
    'ds_uf'].apply(lambda x : unidecode(x))
base_vacinacao_body[
    'ds_municipio'] = base_vacinacao_body[
    'ds_municipio'].str.lower()
base_vacinacao_body['ds_municipio'] = base_vacinacao_body[
    'ds_municipio'].apply(lambda x : unidecode(x))
base_vacinacao_body = base_vacinacao_body[
    base_vacinacao_body['num_dose'].notna()]
base_vacinacao_body = base_vacinacao_body[
    base_vacinacao_body['cd_vacina'].notna()]
for old_uf, new_uf in ufs:
    base_vacinacao_body['ds_uf'] = base_vacinacao_body[
        'ds_uf'].str.replace(old_uf, new_uf)
base_vacinacao_body['num_dose'] = base_vacinacao_body[
    'num_dose'].apply(lambda x : str(x).split('a')[0])
base_vacinacao = pd.concat(
    [base_vacinacao, base_vacinacao_body], ignore_index=True)

conn = psycopg2.connect(
    f"host={arg_ip} dbname={arg_name} user={arg_user} password={arg_pass}")

cur = conn.cursor()
for index, row in base_vacinacao.iterrows():
    cur.execute(

```

```
f"""INSERT INTO covid.vacinacao VALUES (  
    {row[0]}, '{row[1]}', '{row[2]}', '{row[3]}',  
    '{row[4]}', {row[5]}, {row[6]},  
    TO_DATE('{row[7]}', 'YYYY-MM-DD'), '{row[8]}',  
    {row[9]}  
) """  
  
)  
conn.commit()
```

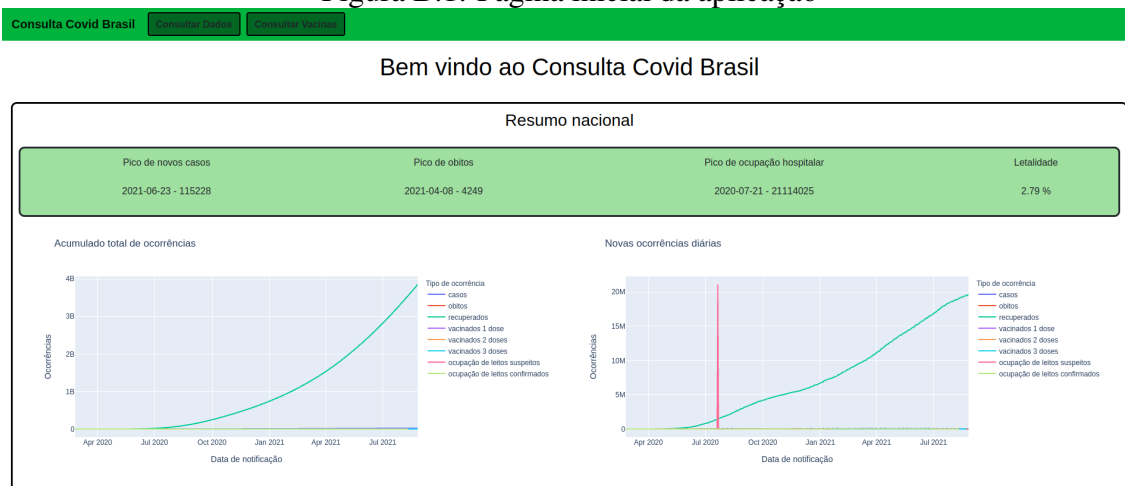
## ApêndiceB GUIA DE UTILIZAÇÃO

A seguir será apresentado os instruções para acesso e utilização da plataforma desenvolvida:

### B.1 Acessando a ferramenta

- *Passo 1:* Abra o navegador de internet de sua preferência.
- *Passo 2:* Se estiver acessando a plataforma em uma execução local acesse <localhost: 5000>, caso esteja acessando a instância remota disponibilizada, acesse <http://consultacovidbrasil.herokuapp.com/>.
- *Passo 3:* Após o sistema carregar a pagina inicial (figura B.1) você terá acesso a informações de resumo nacional e regional de métricas gerais da pandemia:

Figura B.1: Página inicial da aplicação



### B.2 Consultando dados

- *Passo 1:* Acesse a tela de consulta de dados por meio do botão *Consultar Dados* presente no cabeçalho da pagina.
- *Passo 2:* O sistema carregará a pagina de filtro (figura B.2):

Figura B.2: Página de segmentação da consulta

Fonte: O autor

- *Passo 3:* Selecione a região geográfica e período desejado para consulta, lembrando que caso os campos sejam deixados em branco são consultados dados para todos os valores disponíveis.
- *Passo 4:* Clique no botão *Buscar*.
- *Passo 5:* O sistema carregará a interface de resultados mostrando as métricas referentes a segmentação realizada (figura B.3).

Figura B.3: Página de segmentação da consulta



Fonte: O autor

- *Passo 6:* Para retornar a página inicial basta clicar em *Consulta Covid Brasil* no cabeçalho da aplicação.

### B.3 Consultando reflexo da vacinação

- *Passo 1:* Acesse a tela de consulta de dados por meio do botão *Consultar Vacinas* presente no cabeçalho da pagina.
- *Passo 2:* O sistema carregará a pagina de filtro (figura B.4):

Figura B.4: Página de segmentação da consulta a vacinas

Fonte: O autor

- *Passo 3:* Selecione a região geográfica e período desejado para consulta, lembrando que caso os campos sejam deixados em branco são consultados dados para todos os valores disponíveis.
- *Passo 4:* Clique no botão *Buscar*.
- *Passo 5:* O sistema carregará a interface de resultados mostrando as métricas referentes a segmentação realizada (figura B.5).

Figura B.5: Página de segmentação da consulta de vacinação



Fonte: O autor