



**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO PIAUÍ**  
**CAMPUS FLORIANO**  
**CURSO TECNOLOGIA EM ANÁLISE E DESENVOLVIMENTO DE SISTEMAS**

**IAGO BEZERRA DA SILVA**

**AVALIAÇÃO DAS ESTRATÉGIAS DE COMBATE À COVID-19 UTILIZANDO**  
**TÉCNICAS DE ANÁLISE DE DADOS: UM ESTUDO DE CASO SOBRE A**  
**REALIDADE BRASILEIRA**

**FLORIANO**  
**2023**

IAGO BEZERRA DA SILVA

AVALIAÇÃO DAS ESTRATÉGIAS DE COMBATE À COVID-19 UTILIZANDO TÉCNICAS  
DE ANÁLISE DE DADOS: UM ESTUDO DE CASO SOBRE A REALIDADE BRASILEIRA

Trabalho de Conclusão de Curso (artigo científico)  
apresentado como exigência parcial para obtenção do  
diploma do Curso de Tecnologia em Análise e  
desenvolvimento de sistemas do Instituto Federal de  
Educação, Ciência e Tecnologia do Piauí, Campus  
Floriano.

Orientador: Prof. Dr. Willamys Rangel Nunes de Sousa.

FLORIANO  
2023

# **AVALIANDO AS ESTRATÉGIAS DE COMBATE À COVID-19 UTILIZANDO TÉCNICAS DE ANÁLISE DE DADOS: UM ESTUDO DE CASO SOBRE A REALIDADE BRASILEIRA**

Iago Bezerra da Silva<sup>1</sup>  
Willamys Rangel Nunes de Sousa<sup>2</sup>

## **RESUMO**

A pandemia do COVID-19 assolou diversos países devido à falta de métodos eficazes para combater a pandemia. Durante esse período, não houve desenvolvimento de melhores estratégias para combater essa epidemia. Este trabalho teve como objetivo analisar as estratégias e ações adotadas pelo governo brasileiro no combate à COVID-19 e avaliar o impacto dessas medidas. Para alcançar esse objetivo, foram utilizados diversos tipos de dados, incluindo informações demográficas, econômicas e de saúde, bem como o Índice de Restrições. Apesar das limitações enfrentadas, esta pesquisa contribuiu significativamente para o entendimento do desempenho das cidades brasileiras durante a pandemia. Dessa forma, é possível identificar as cidades que conseguiram obter uma melhor resposta à pandemia e as estratégias adotadas por elas.

**Palavras-chave:** COVID-19; Pandemia; Aprendizado de máquina; Stringency index.

## **ABSTRACT**

The COVID-19 pandemic has devastated several countries due to the lack of effective methods to combat the pandemic. During the pandemic, there was no development of better strategies to combat COVID-19, the objective of this work was to analyze the strategies and actions adopted by the Brazilian government to combat COVID-19 and evaluate the impact of these measures. To achieve this objective, several types of data were used, including demographic, economic and health information, as well as the Restrictions Index. Despite the limitations faced, this research contributed significantly to understanding the performance of Brazilian cities during the pandemic. In this way, it is possible to identify the cities that managed to obtain a better response to the pandemic and the strategies adopted by them.

**Keywords:** COVID-19; Pandemic; Machine learning; Stringency index.

Data de aprovação: 16/10/2023.

---

<sup>1</sup> Currículo sucinto do(a) autor(a). Instituição de ensino. E-mail

<sup>2</sup> Currículo sucinto do(a) professor(a) orientador(a). Instituição de ensino. E-mail

## 1 INTRODUÇÃO

A pandemia de COVID-19 tem se revelado como um dos desafios mais significativos do século XXI, afetando de forma profunda a saúde pública, a economia e a sociedade em escala global (OMS, 2020). Desde os primeiros casos identificados na cidade de *Wuhan*, na China, em dezembro de 2019, a disseminação do vírus se acelerou rapidamente, resultando em mais de 2 milhões de casos e 120 mil óbitos em todo o mundo em poucos meses (OMS, 2020).

Esses números alarmantes demonstram a velocidade e o impacto avassalador que o vírus teve sobre a saúde e o bem-estar das pessoas em todo o planeta, levando os países a adotarem medidas drásticas para conter sua propagação e mitigar os efeitos da pandemia.

Contudo, a resposta à disseminação do vírus foi marcada pela falta de padronização ao combate à pandemia (CEPI, 2021). Com o objetivo de estabelecer diretrizes uniformes e identificar as melhores estratégias de combate à COVID-19, surgiu o *Stringency Index* (HAAS *et al.*, 2020).

Este índice busca analisar e classificar as medidas adotadas por cada país, permitindo que outras nações possam replicar as estratégias mais eficazes no enfrentamento da pandemia. Ao promover o compartilhamento de conhecimento e experiências, o *Stringency Index* desempenha um papel fundamental na busca por soluções mais assertivas e na busca por uma resposta global coordenada diante dessa crise de saúde pública (HAAS *et al.*, 2020).

No contexto brasileiro, a gravidade da situação também teve um impacto significativo no país. Durante a pandemia, o Brasil enfrentou várias ondas de contágio, com picos de casos e óbitos ocorrendo em momentos distintos em diferentes regiões (MACIEL; CASTRO-SILVA; FARIAS, 2020).

Essa variação temporal e geográfica das infecções trouxe desafios adicionais para as autoridades de saúde e para o sistema de saúde como um todo, que precisavam se adaptar e direcionar recursos de acordo com as demandas emergentes. A disseminação acelerada do vírus, combinada com a diversidade socioeconômica e geográfica do país, contribuiu para a complexidade da resposta à pandemia no Brasil, exigindo medidas abrangentes e adaptadas a cada realidade local (COSSA, *et al.*, 2021).

A análise de dados tem se mostrado fundamental na tomada de decisões embasadas em evidências, fornecendo informações valiosas para a gestão da crise sanitária e a importância de compreender e analisar as estratégias que foram utilizadas para o combate à COVID-19 adotadas no contexto brasileiro, utilizando técnicas de análise de dados e o *Stringency Index* como medida de avaliação.

Assim, este trabalho de pesquisa tem como objetivo realizar uma avaliação das estratégias adotadas no combate à COVID-19 no Brasil, utilizando técnicas de análise de dados, como uma ferramenta para compreender a realidade brasileira e buscar soluções que foram efetivas durante a pandemia.

Para tanto, foi realizado um estudo de caso, focado na realidade brasileira, a fim de examinar a eficácia das medidas implementadas pelo governo e pelos órgãos de saúde, bem como seus impactos na contenção da propagação do vírus e na saúde pública.

## 2 FUNDAMENTAÇÃO TEÓRICA

### 2.1 COVID-19

De acordo com a Organização Mundial da Saúde (OMS 2020), a Covid-19 é uma doença infecciosa causada pelo coronavírus (SARS-CoV-2), que foi identificada pela primeira vez na cidade de *Wuhan*, na China, em dezembro de 2019. A doença espalhou-se rapidamente pelo mundo, tornando-se uma pandemia em março de 2020. Esta doença afeta principalmente o sistema respiratório, causando vários sintomas que variam de leves a graves, como tosse seca, febre, fadiga, falta de ar e dores musculares.

A Covid-19 é uma doença altamente contagiosa e seu principal modo de transmissão é por meio de gotículas respiratórias expelidas por uma pessoa infectada ao falar, tossir ou espirrar, podendo ser inaladas por outra pessoa de forma direta ou indireta. Essa forma de transmissão torna-se um grande desafio na prevenção da doença.

Por essa razão, a Organização Mundial da Saúde (OMS 2020) recomendou diversas medidas preventivas para evitar a disseminação do vírus, como o distanciamento social, a utilização de máscaras faciais e a higiene regular das mãos. Essas medidas são de extrema importância para reduzir o risco de contágio e proteger a saúde pública.

### 2.2 PANDEMIA GLOBAL

Desde o início dos primeiros surtos de Covid-19, em 31 de dezembro de 2019, a Organização Mundial da Saúde (OMS) foi alertada sobre vários casos de pneumonia, na cidade de *Wuhan*, na província de *Hubei* (OMS, 2019). Essa pneumonia foi identificada como uma nova variação do vírus SARS-CoV-2, que ainda não havia sido identificado no corpo humano (OMS, 2020).

Infelizmente, esta nova cepa do vírus pode causar uma série de sintomas, como febre, falta de ar, tosse, espirros, coriza, dor de garganta, diarreia e cansaço. Devido à sua alta taxa de transmissão, o vírus se espalhou rapidamente e acabou se tornando uma pandemia mundial.

Em 30 de janeiro de 2020, a OMS declarou que o surto do novo coronavírus constitui uma Emergência de Saúde Pública de Importância Internacional (ESPII) - o nível mais alto de alerta da Organização. Esta decisão foi tomada para melhorar a cooperação e coordenação de como combater essa crise mundial entre os países. Essa crise não afetou apenas o sistema de saúde em escala global, mas também teve grandes impactos sociais, econômicos, políticos e culturais.

A estimativa da OMS sobre o número de infectados e mortos não apenas afeta diretamente a capacidade dos sistemas de saúde, mas também tem um impacto significativo em nossas vidas diárias. A exposição contínua da população ao vírus pode agravar ainda mais os impactos relacionados ao nosso modo de vida (OMS, 2020).

Com o objetivo de controlar a disseminação do vírus, a OMS propôs várias recomendações, sendo o *lockdown* das cidades uma das medidas mais marcantes, além do uso de máscaras e a higienização adequada das mãos (OMS, 2020).

Embora essas medidas tenham sido eficazes no combate ao vírus, elas também tiveram impacto na sustentabilidade econômica do sistema financeiro dos países e na saúde mental das pessoas que se encontram em situação de confinamento (SILVA, SANTOS, 2020).

Diante deste cenário caótico, muitos foram os desafios enfrentados pelos países que buscavam controlar rapidamente a propagação do vírus em suas regiões. No entanto, ficou evidente que não houve uma padronização na conduta global, com alguns países adotando medidas mais rigorosas do que outros desde o início da pandemia.

## 2.3 PANDEMIA NO BRASIL

No Brasil, a pandemia de COVID-19 foi considerada uma das maiores do mundo, classificando o país como o terceiro com a mais alta taxa de letalidade absoluta, ficando atrás apenas dos Estados Unidos e do Reino Unido (IMPERIAL COLLEGE LONDON, 2021). O primeiro caso confirmado de COVID-19 no Brasil ocorreu em 26 de fevereiro de 2020, envolvendo um idoso que havia retornado da Itália. Em menos de dois meses, o país já contabilizava 135.106 casos confirmados e 9.416 óbitos (MACIEL, CASTRO - SILVA, FARIAS, 2020).

Diante do aumento no número de casos e óbitos nas diferentes Unidades da Federação (UF) do Brasil, uma das principais medidas adotadas para conter a propagação do vírus foi a implementação do distanciamento social e o *lockdown* em cidades. Apesar de todas as medidas de contenção contra a disseminação do coronavírus, o Brasil registrou um total de 37 milhões de casos, sendo a região nordeste a mais afetada, representando 31,2% do total de casos no país (MINISTÉRIO DA SAÚDE, 2020).

## 2.4 VACINAÇÃO

Durante a pandemia do COVID-19, depositou-se uma esperança promissora e ansiosa nas vacinas como solução crucial para combater a disseminação da doença, que já havia se estabelecido em todo o mundo (GAO *et al.*, 2020).

Diante do crescente número de óbitos, houve uma grande preocupação com a produção de vacinas, levando os cientistas a se empenharem em sua criação. Surpreendentemente, os primeiros imunizantes foram desenvolvidos em tempo recorde, sendo disponibilizados para a população a partir de meados de 2020.

Enquanto, tradicionalmente, levaria mais de uma década para a aprovação de uma vacina, aquela destinada a combater a COVID-19 alcançou os braços das pessoas em questão de meses. A primeira pessoa a ser vacinada ocorreu em dezembro de 2020, no Reino Unido.

No Brasil, à medida que a pandemia se desenrolava, a primeira pessoa só viria ser vacinada em meados de fevereiro de 2021, onde começou a ser aplicada no país, após obter aprovação emergencial da Agência Nacional de Vigilância Sanitária (ANVISA) (OPAS, 2022).

## 2.5 APRENDIZADO DE MÁQUINA

De acordo com SAS (2022), a inteligência artificial (IA) pode ser definida como a ciência de reproduzir as habilidades humanas. Nesse sentido, o aprendizado de máquina (*Machine Learning*) é uma vertente específica da IA, que baseia-se em modelos matemáticos e estatísticos que aprendem através de exemplos para realizar previsões ou tomar decisões, onde pode ser utilizados em saúde, previsão de doenças, identificação de padrões em exames médicos, em dados de finanças, detecção de fraudes, previsão de tendências de mercado, análise de risco de crédito, além de indústria e em marketing digital. O desempenho do modelo de aprendizado de

máquina é aprimorado de acordo com a experiência adquirida, o que o torna cada vez mais preciso e eficiente em sua tomada de decisão. De acordo com Alpaydin (2010), o Aprendizado de Máquina é uma subárea da Inteligência Artificial que se concentra no desenvolvimento de algoritmos e técnicas que permitem que os computadores aprendam a partir de dados, sem a necessidade de serem explicitamente programados.

Essa abordagem tem uma ampla variedade de aplicações em áreas como a classificação de imagens, análise de sentimentos em mídias sociais, reconhecimento de voz e detecção de fraudes financeiras.

Segundo Ludermir (2021), existem dois tipos de classificação, uma abordagem supervisionada e não supervisionada, onde a aprendizagem supervisionada envolve treinar um modelo para fazer previsões ou tomar decisões com base em um conjunto de dados de entrada e saída conhecidos, sendo amplamente utilizado em detecção de anomalias, detecção de fraudes, entre outros. Em contrapartida, a classificação não supervisionada, como o algoritmo K-Means, é um processo que pode revelar padrões que não foram inicialmente utilizados em técnicas de agrupamento. O K-Means é um exemplo de algoritmo de agrupamento não supervisionado que tem como objetivo dividir um conjunto de dados em grupos ou clusters, onde os pontos dentro de um mesmo cluster são mais similares entre si do que com os pontos de outros clusters. Esse método é útil em diversas aplicações, como segmentação de clientes, análise de agrupamento geográfico e muito mais.

## 2.6 STRINGENCY INDEX

O Índice de Restrições Governamentais, também conhecido como *Stringency Index* (HAAS *et al.*, 2020), é uma medida utilizada para avaliar o grau de restrições impostas pelos governos durante a pandemia de COVID-19.

Esta métrica leva em consideração uma variedade de indicadores, como o fechamento de escolas, restrições de viagens, proibições de eventos em massa e o uso obrigatório de máscaras faciais. Ela foi amplamente utilizada para comparar e monitorar as medidas adotadas pelos países ao redor do mundo, para conter a disseminação do vírus. Essas informações podem ser obtidas no banco de dados da Universidade de Oxford, que compila os dados e calcula o *Stringency Index* para cada país (HAAS *et al.*, 2020).

Durante a pandemia, alguns estudos utilizaram esse método com o propósito de avaliar o desempenho de seus países em relação a outros durante esse período. Um desses estudos foi conduzido por Barberia *et al.* (2021), que empregaram o mesmo índice para avaliar como a rigidez das políticas impactava a mobilidade da população.

Outra pesquisa que trata desse tema é a de Darapaneni (2023), que busca identificar as estratégias mais eficazes para combater a Covid-19. Este trabalho utilizou técnicas de aprendizado de máquina para analisar o efeito de diversas medidas governamentais em países de natureza semelhante.

## 2.7 TRABALHOS RELACIONADOS

O trabalho de Darapaneni (2023), intitulado *Machine Learning Approach For Clustering Of Countries To Identify The Best Strategies To Combat Covid-19*, apresenta uma abordagem de aprendizagem de máquina para a clusterização de países com características semelhantes,

buscando avaliar o impacto de diferentes medidas governamentais no contexto da pandemia da Covid-19.

Esta pesquisa utiliza dados históricos da Covid-19, além de considerar parâmetros socioeconômicos e outros fatores relevantes, com o objetivo de identificar estratégias potenciais a serem adotadas em cada fase da pandemia, no qual este trabalho tem como base a sua metodologia.

Já no artigo de Chen *et al.* (2021) *Exploratory Data Analysis on the Usage of COVID-19 Vaccine* apresenta uma análise detalhada do panorama da vacinação contra a COVID-19 em diferentes países.

Este trabalho aborda questões essenciais, como a limpeza de dados, a análise exploratória e a visualização do conjunto de dados, tendo em vista proporcionar uma compreensão mais clara da situação global no enfrentamento da epidemia. O objetivo principal deste estudo foi destacar a importância crucial da vacinação como ferramenta fundamental no combate à COVID-19. Além disso, o trabalho enfatiza a necessidade contínua de monitorar e analisar os dados, o que contribui para uma compreensão mais abrangente da situação global da epidemia.

E por fim, Silva *et al.* (2020), analisa as medidas de distanciamento social adotadas pelos estados brasileiros durante a pandemia da COVID-19. Ele inclui uma análise epidemiológica por estado, com informações sobre o tipo de medida adotada e o momento de sua implementação.

### 3 METODOLOGIA

Este estudo adotou uma abordagem quantitativa, que trabalha com variáveis expressas como dados numéricos e utiliza recursos e técnicas estatísticas para classificação e análise, incluindo porcentagem, média, desvio padrão, entre outros (FONTELLES, 2009).

A metodologia adotada neste trabalho foi inspirada no estudo realizado por Darapaneni (2023). Para descrever a construção da base de dados, foi apresentado um processo detalhado na Figura 1, que ilustra todas as etapas envolvidas.

Figura 1 - Construção da Base de Dados



Fonte: elaborado pelo autor (2023).



Na construção da base de dados, foram utilizados três conjuntos de dados principais: dados demográficos, dados relativos às medidas e respostas governamentais e dados epidemiológicos sobre a pandemia de COVID-19 no Brasil. Estes dados foram adquiridos através do Instituto Brasileiro de Geografia e Estatística (IBGE), DataSus, Ministério da Saúde e *Our World ind Data* (BRASIL, 2023a; BRASIL, 2023b; DATA, 2023; IBGE, 2023)

A seleção das cidades para análise de dados foi baseada no Produto Interno Bruto (PIB), utilizando uma base de dados do IBGE. Foram escolhidas as 5 (cinco) primeiras cidades com o maior PIB de cada unidade federativa, resultando em um total de 131 registros. Essa seleção representa as principais cidades de cada Estado brasileiro.

Logo após a primeira seleção de cidades, foi realizada uma seleção inicial de atributos para refinar a base de dados, incluindo apenas aqueles que seriam mantidos na base final. No Quadro 1, é possível verificar os atributos que foram retidos, juntamente com seus *metadados* e a base original à qual eles pertenciam.

Quadro 1 - Quadro de Metadados

| Base de Dados Original | Nome           | Metadados                                         |
|------------------------|----------------|---------------------------------------------------|
| IBGE                   | UF             | Sigla da Unidade Federativa                       |
| IBGE                   | COD_UF         | Código da Unidade Federativa                      |
| IBGE                   | COD_MUNIC      | Código do Município da Cidade                     |
| IBGE                   | NOME_MUNICÍPIO | Nome da Cidade                                    |
| IBGE                   | População      | População de cada cidade                          |
| IBGE                   | COD_REGIÃO     | O código em qual região a cidade fica localizada  |
| IBGE                   | NOME_REGIÃO    | O nome que a região fica localizado a cidade      |
| DATASUS                | SEMIÁRIDO      | Este atributo fica em uma área de clima semiárido |
| IBGE                   | PIB_BRUTO      | O produto interno bruto da cidade                 |
| Ministério da Saúde    | QTD_CNES       | Quantidade de estabelecimentos de saúde existem   |
| Fusão de atributos     | COD_FULL       | Esta coluna é uma junção de duas outras colunas.  |

Continua...

Quadro 1 - Quadro de Metadados (continuação)

| Base de Dados Original | Nome             | Metadados                                                                                                                                                                                              |
|------------------------|------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Our In Data            | STRINGENCY_INDEX | Esse atributo está relacionado a medidas e resposta governamentais. Que é a média 4 fatores principais: resposta governamentais, índice de contenção de saúde, suporte econômico, e políticas públicas |
| Fusão de atributos     | COD_FULL         | Esta coluna é uma junção de duas outras colunas.                                                                                                                                                       |

Fonte: elaborado pelo autor (2023)

Durante o processo de seleção de atributos, foram necessárias algumas alterações e inclusões. A primeira modificação foi feita no atributo "*SEMIÁRIDO*", cujos registros, originalmente, possuíam os valores "*sim*" ou "*não*". Optou-se por transformá-los em valores binários, sendo *0* representando "*não*" e *1* representando "*sim*". Além disso, um novo atributo denominado "*COD\_FULL*" foi adicionado, para representar o código completo do IBGE para cada registro da base de dados.

Após o pré-processamento dos dados, foi iniciada a etapa de análise, utilizando o algoritmo não supervisionado *K-means* (COSSA, *et al.*, 2021).

Esta etapa da metodologia permitiu a construção de *clusters*<sup>3</sup> e a realização de uma análise dos dados. Ademais, foi realizada uma análise detalhada de cada agrupamento gerado pelo algoritmo. Nesta fase, foram investigadas as características e os padrões presentes em cada *cluster*, identificando possíveis correlações e *insights* relevantes para o estudo.

Após a análise inicial, dados relacionados ao COVID-19 foram incorporados, onde esses dados tiveram um recorte temporal do ano 2021, como podemos observar no Quadro 2, onde o propósito era de avaliar se os *clusters* previamente definidos se manteriam em seus agrupamentos originais ou passariam por alterações significativas.

<sup>3</sup> Definição de Clusters 10.1590/S0104-12902015000100003

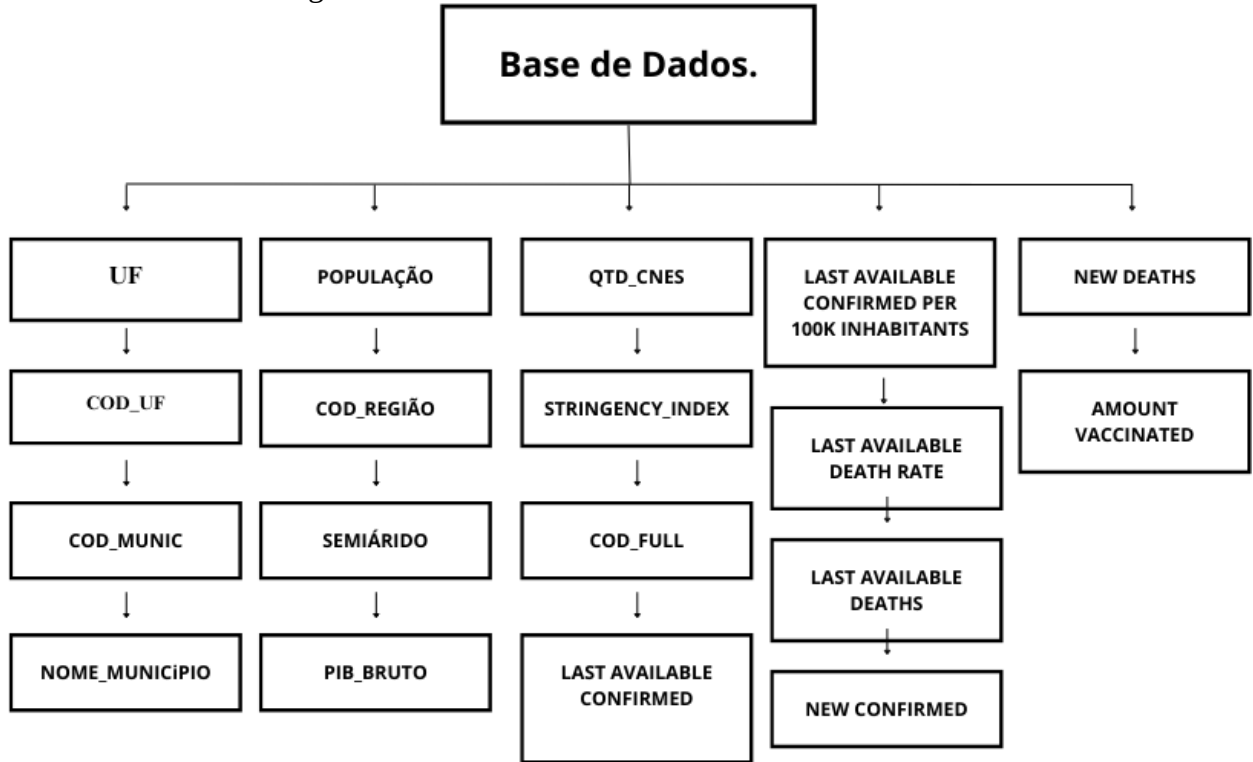
Quadro 2 - Quadro de Metadados após inserção de dados do COVID-19

| Base de Dados Original | Nome                                          | Metadados                                                                                                                                                                                   |
|------------------------|-----------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| DATASUS                | Epidemiological Week                          | Unidade de medida que marca a semana em curso na pandemia e permite determinar como está ocorrendo a propagação da doença durante um determinado período de tempo, que costuma ser 07 dias. |
| DATASUS                | Last Available Confirmed                      | Representa o número total mais recente de casos confirmados de uma doença, como uma infecção viral, em uma determinada área.                                                                |
| DATASUS                | Last Available Confirmed per 100k Inhabitants | Indica o número de casos confirmados da doença por 100.000 habitantes, o que ajuda a normalizar os dados para áreas com diferentes tamanhos de população.                                   |
| DATASUS                | Last Available Death Rate                     | Refere-se à taxa de mortalidade mais recente, que é a proporção de mortes em relação ao número total de casos confirmados.                                                                  |
| DATASUS                | Last Available Deaths                         | Representa o número mais recente de mortes causadas pela COVID-19.                                                                                                                          |
| DATASUS                | New Confirmed                                 | Indica o número de novos casos confirmados relatados em um único dia.                                                                                                                       |
| DATASUS                | New Deaths                                    | Indica o número de novas mortes relatadas em um único dia.                                                                                                                                  |
| DATASUS                | Amount Vaccinated                             | Representa o número total de pessoas vacinadas pelo COVID-19.                                                                                                                               |

Fonte: elaborado pelo autor (2023)

O objetivo era discernir se esses agrupamentos se manteriam estáveis ou se ocorreria uma reconfiguração. Esse procedimento teve como meta investigar as causas subjacentes à movimentação de determinadas cidades para fora de seus clusters originais durante a pandemia, buscando compreender os fatores que influenciaram essas mudanças. Abaixo, na Figura 2, estão sendo apresentados os atributos da base de dados final.

Figura 2 - Amostra das colunas da base de dados final.



Fonte: elaborado pelo autor (2023)

#### 4 ANÁLISE E RESULTADOS

Antes de serem adicionados os dados do COVID-19, foi realizada a primeira clusterização, para determinar a qual *cluster* cada cidade pertenceria.

Os resultados mostraram que 55 cidades foram alocadas no *cluster 0*, outras 58 cidades no *cluster 1* e 18 cidades no *cluster 2*, conforme apresentado na Tabela 2.

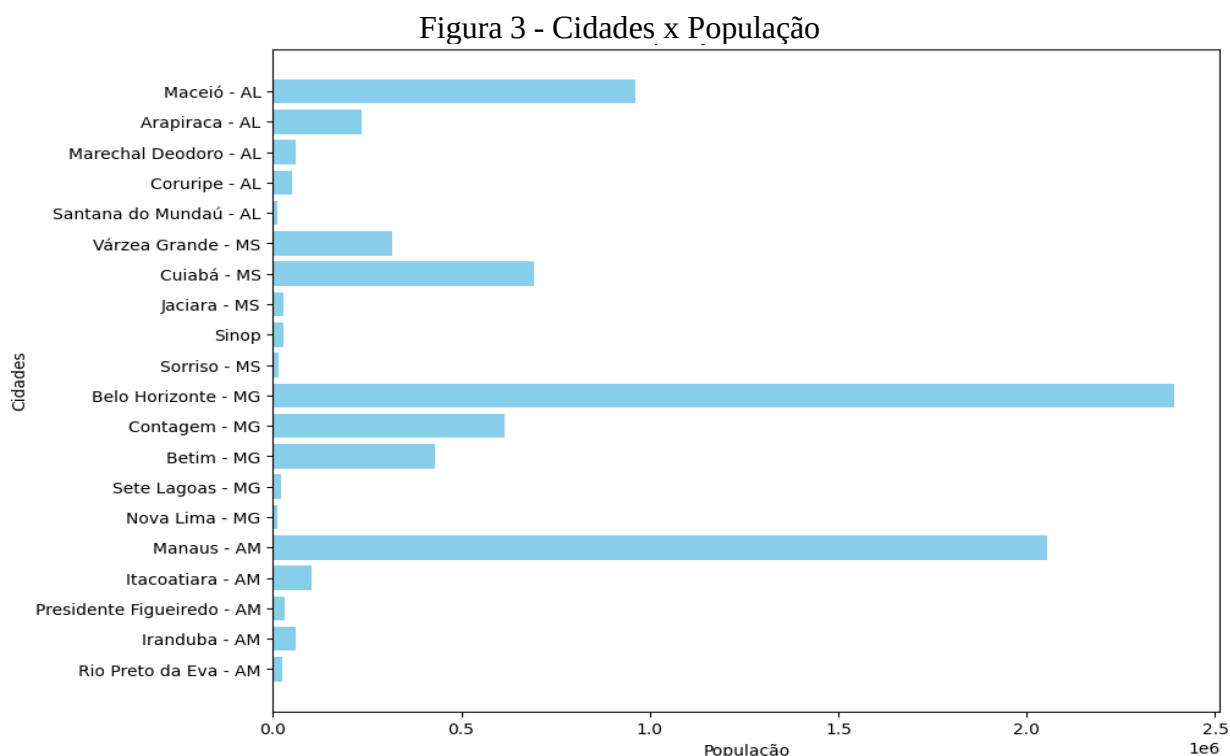
Quadro 3 - Amostra de Cluster pré - Covid.

| Cluster   | Cidade                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|-----------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Cluster 0 | Maceió - AL, Arapiraca - AL, Marechal Deodoro - AL, Coruripe - AL, Santana do Mundaú - AL, Brasília - DF, Salvador - BA, Camaçari - BA, Feira de Santana - BA, São Francisco do Conde - BA, Lauro de Freitas - BA, Fortaleza - CE, Maracanaú - CE, Caucaia - CE, Juazeiro do Norte - CE, Sobral - CE, Belo Horizonte - MG, Contagem - MG, Sete Lagoas - MG, Nova Lima - MG, Manaus - AM, Itacoatiara - AM, Presidente Figueiredo - AM, Iranduba - AM, Rio Preto da Eva - AM, Goiânia - GO, Luziânia - GO, Aparecida de Goiânia - GO, São Simão - GO, Senador Canedo - GO, Belém - PA, São Luís - MA, Imperatriz - MA, Timon - MA, Açailândia - MA, Estreito - MA, Campo Grande - MS, Dourados - MS, Ponta Porã - MS, Maracaju - MS, Cuiabá - MT, Várzea Grande - MT, Jaciara - MT, Sinop - MT, Sorriso - MT, Vitória - ES, Serra - ES, Vila Velha - ES, Cariacica - ES, Viana - ES.                                                                               |
| Cluster 1 | Santarém - PA, Castanhal - PA, Marituba - PA, Ananindeua - PA, Coruripe - MS,                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|           | Porto Alegre - RS, São Leopoldo - RS, Canoas - RS, Gravataí - RS, Rio Grande - RS, São Paulo - SP, Osasco - SP, Guarulhos - SP, Campinas - SP, Barueri - SP, Riode Janeiro - RJ, Duque de Caxias - RJ, Niterói - RJ, Marica - RJ, São Gonçalo - RJ, Curitiba - PR, São José dos Pinhais - PR, Londrina - PR, Maringá - PR, Araucária - PR, Recife - PE, Jaboatão dos Guararapes - PE, Ipojuca - PE, Cabo de Santo Agostinho - PE, Petrolina - PE, Uruçuí - PI, Aracaju - SE, Canindé de São Francisco - SE, Nossa Senhora do Socorro - SE, Estância - SE, Itabaiana - SE, Itabaiana - SE, Joinville - SC, Itajaí - SC, Florianópolis - SC, Blumenau - SC, Chapecó - SC, Porto Velho - RO, Ji - Paraná - RO, Vilhena - RO, Ariquemes - RO, Cacoal - RO, Boa Vista - RR, Rorainópolis - RR, Bonfim - RR, Caracaraí - RR, Iracema - RR, Porto Nacional - TO, Gurupi - TO, Rio Branco - AC, Cruzeiro do Sul - AC, Sena Madureira - AC, Tarauacá - AC, Brasiléia - AC. |
| Cluster 2 | Teresina - PI, Floriano - PI, Picos - PI, Parnaíba - PI, Palmas - TO, Peixe - TO, Paraíso do Tocantins - TO, Natal - RN, Extremoz - RN, Parnamirim - RN, Goianinha - RN, São Gonçalo do Amarante - RN, João Pessoa - PB, Campina Grande - PB, Cabedelo - PB, Santa Rita - PB, Alhandra - PB.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |

Fonte: elaborado pelo autor (2023)

Ao realizar uma análise das cidades, antes da aplicação dos dados de COVID-19, elas foram inicialmente agrupadas em três clusters distintos. No *cluster 0*, foi possível observar a presença de algumas cidades que compartilham características específicas, como a sua proximidade geográfica, que é o caso das cidades do Nordeste (AL, BA, CE, MA), Centro-Oeste

(DF, GO, MS, MT), Sudeste(ES, MG) e Norte (AM, PA). Além disso, algumas dessas cidades têm uma população semelhante, como podemos ver na Figura 3.

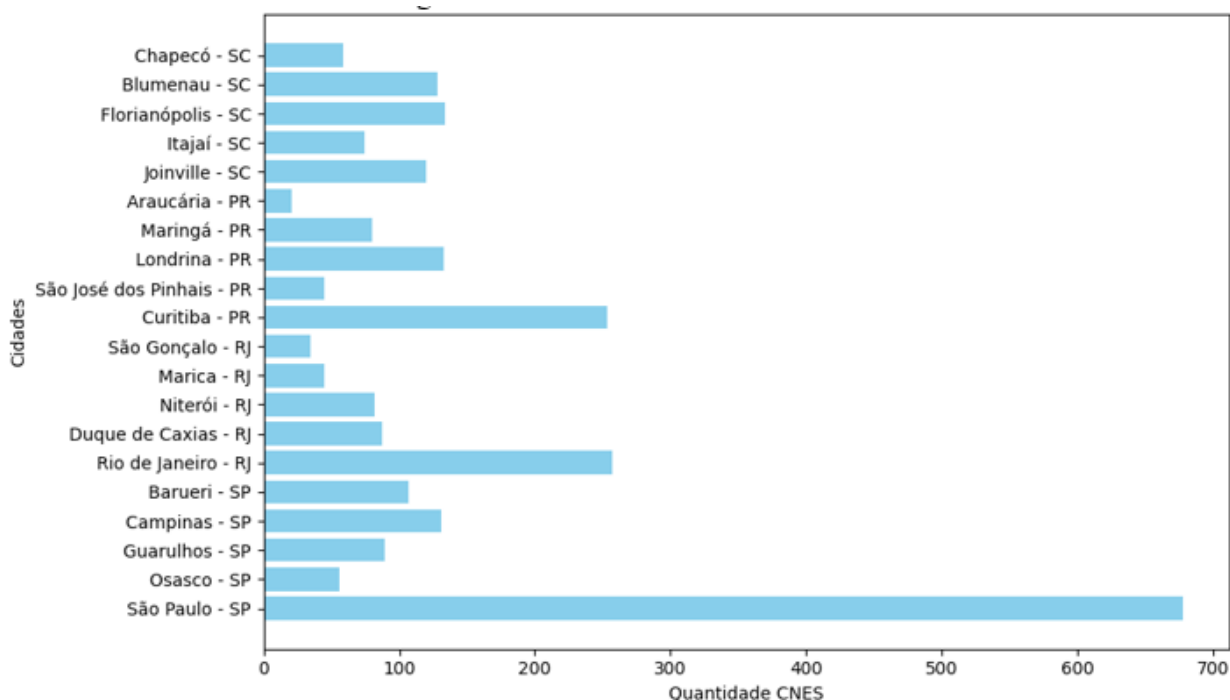


Fonte: elaborado pelo autor (2023).

Na Figura 3, destacam-se cidades oriundas de distintas regiões do Brasil, a saber: Maceió -AL, Cuiabá - MS, Belo Horizonte - MG e Manaus - AM. Essas cidades estão distribuídas em diversas regiões geográficas do país, abrangendo o Nordeste, Centro-Oeste, Sudeste e Norte. Essas cidades demonstram níveis populacionais relativamente próximos. Isso nos permite perceber que as cidades agrupadas no *cluster 0* compartilham semelhanças notáveis em seus níveis populacionais.

Por sua vez, o *Cluster 1* apresenta um agrupamento de cidades que compartilham a característica de estarem localizadas em diferentes estados do Brasil. No entanto, a lista inclui diversas cidades de várias regiões do país, abrangendo o Norte, Nordeste, Sul e Sudeste. Uma semelhança notável é identificada na infraestrutura dessas cidades, especialmente em relação à quantidade de Cadastro Nacional de Estabelecimentos de Saúde (CNES), que se aproxima de valores semelhantes. Esse aspecto pode ser observado na Figura 4.

Figura 4 - Cidades x CNES



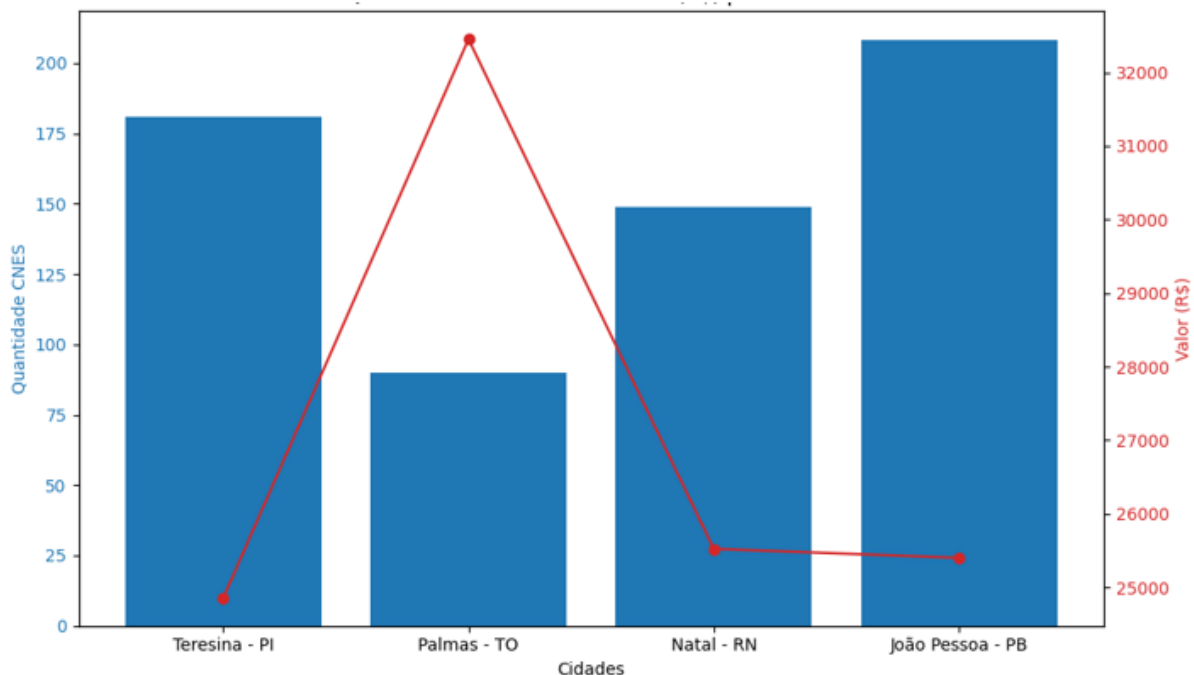
Fonte: elaborado pelo autor (2023).

Conforme ilustrado na Figura 4, percebe-se que diversas cidades de distintas regiões do país, compõem o *cluster 1*. Essas cidades apresentam quantidades de CNES semelhantes, com um destaque especial para as capitais dos estados. Estas, por sua vez, sobressaem-se das demais cidades de seus respectivos estados, caracterizando-se pela grande quantidade de CNES. Esse padrão é notável nas cidades de São Paulo - SP, Rio de Janeiro - RJ e Curitiba - PR, as quais registram as mais expressivas quantidades de CNES entre todas as cidades analisadas na Figura 4.

Ainda mais interessante é a observação de que as cidades remanescentes demonstram uma uniformidade significativa em suas quantidades de CNES, com valores que se mantêm abaixo do limite de 150 CNES. Esse fenômeno destaca a tendência geral no *cluster 1*, onde a distribuição das estruturas de saúde, indica que, embora possam existir variações dentro deste *cluster*, é comum a grande maioria das cidades terem no máximo 150 estabelecimentos de saúde.

Por fim, no *Cluster 2*, foram agrupadas cidades localizadas apenas nas regiões Nordeste e Norte do Brasil. Observa-se que essas cidades possuem uma notável dependência de suas respectivas capitais, uma vez que tanto a quantidade de CNES quanto o PIB são significativamente maiores nas capitais. Esse cenário é ilustrado na Figura 5, apresentada abaixo.

Figura 5- Quantidade de CNES e PIB das Capitais



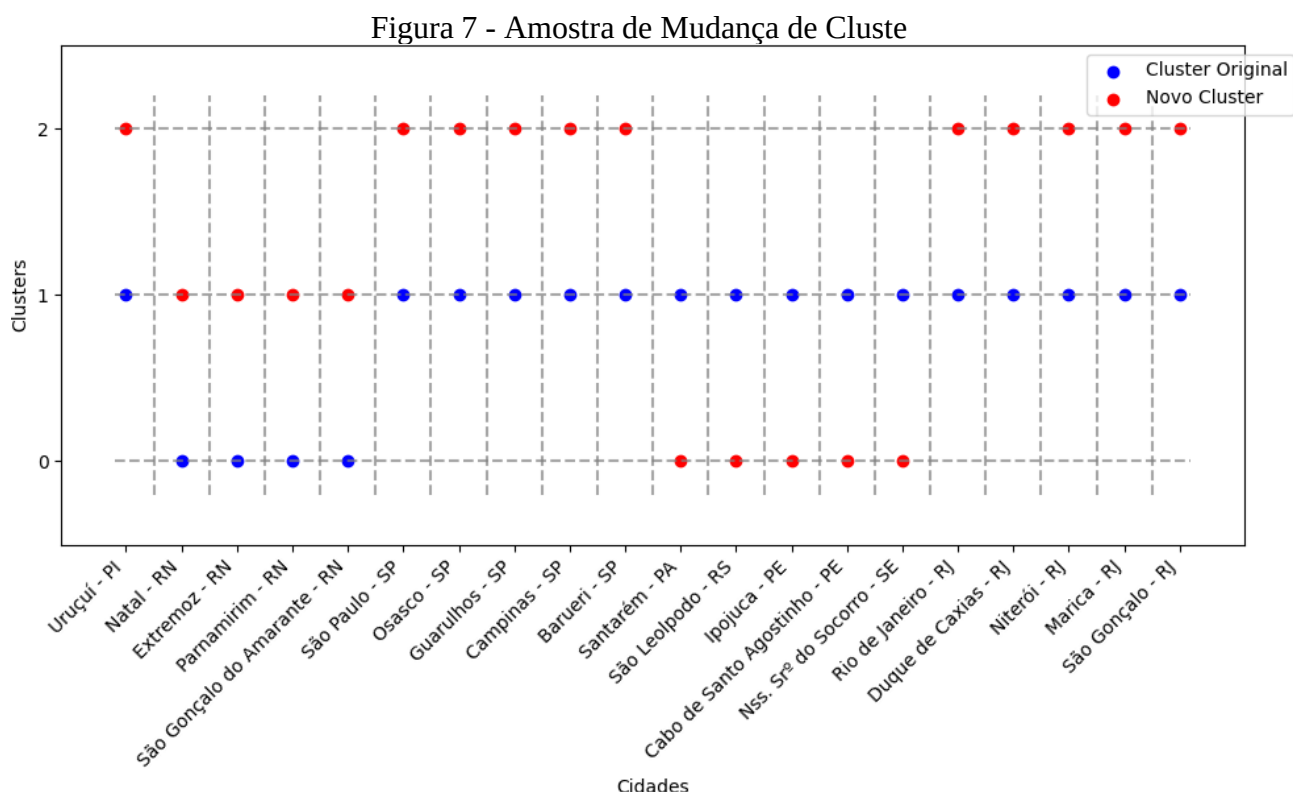
Fonte: elaborado pelo autor (2023).

Através do Figura 5, pode-se verificar a quantidade de CNES e os valores do PIB em reais, para as cidades de Teresina - PI, Palmas - TO, Natal - RN e João Pessoa - PB. A barra azul representaa quantidade de CNES e a linha vermelha representa o valor do PIB em reais.

Conforme evidenciado na Figura 5, nota-se que as cidades em questão são, de fato, as capitais de seus respectivos Estados. É interessante destacar que, para além desse aspecto, o Produto Interno Bruto (PIB) destas cidades compartilha uma notável similaridade. Esse padrão é ainda mais significativo ao considerar que o PIB destas cidades é o mais elevado dentro de seus estados, o que realça a sua influência econômica e papel central na região. E a suas infraestruturas apresentam umasimilaridade considerável, o que podemos evidenciar pela quase equivalência em suas quantidades deCNES. Esses valores variam em uma faixa muito estreita, com diferenças mínimas entre eles.

Após a aplicação dos dados COVID-19 na base de dados, foram observadas alterações nos *clusters*, resultando na realocação de algumas cidades em relação aos seus *clusters* anteriores, conforme apresentado na Figura 6.



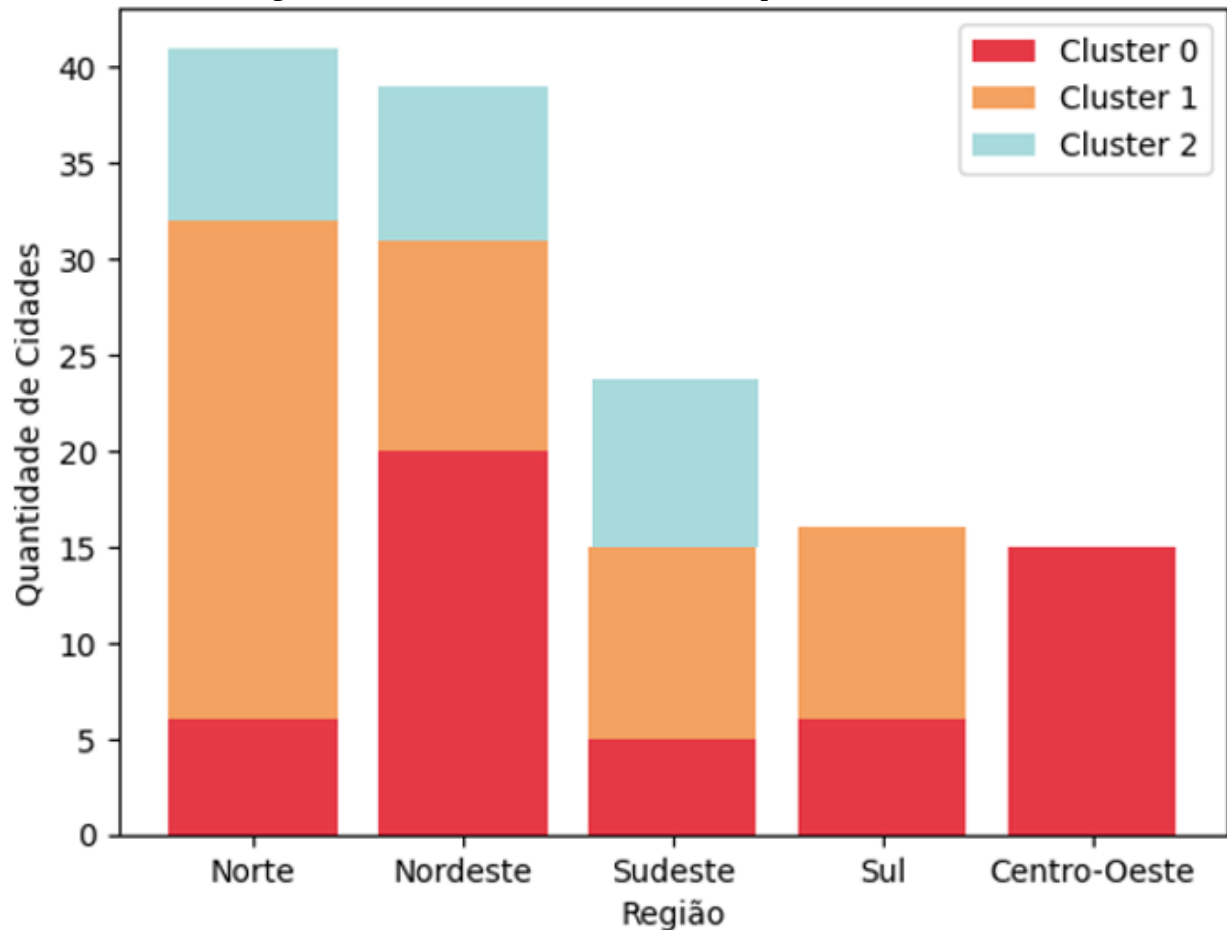


Fonte: elaborado pelo autor (2023).

Na Figura 7, examinamos as cidades que experimentaram mudanças nos seus agrupamentos (*clusters*). Os pontos azuis representam os *clusters* originais, enquanto os pontos vermelhos indicamos *clusters* novos, resultantes da inclusão dos dados de COVID-19. É possível notar que algumas cidades migraram para diferentes *clusters* após a análise, evidenciando as mudanças nos agrupamentos. A Figura 7 ilustra claramente quais cidades tiveram suas posições alteradas nos clusters. Essa alteração pode ser atribuída não apenas a fatores como a quantidade de CNES, mas também ao *Stringency Index*, que pode ter influenciado diretamente essa mudança ao avaliar o índice de restrição e controle de vacinação.

Já na Figura 8, é possível observar um gráfico que ilustra a quantidade de cidades em cada *cluster*, com as cidades agrupadas por região geográfica.

Figura 8 - Amostra de Cidades dividido por Cluster

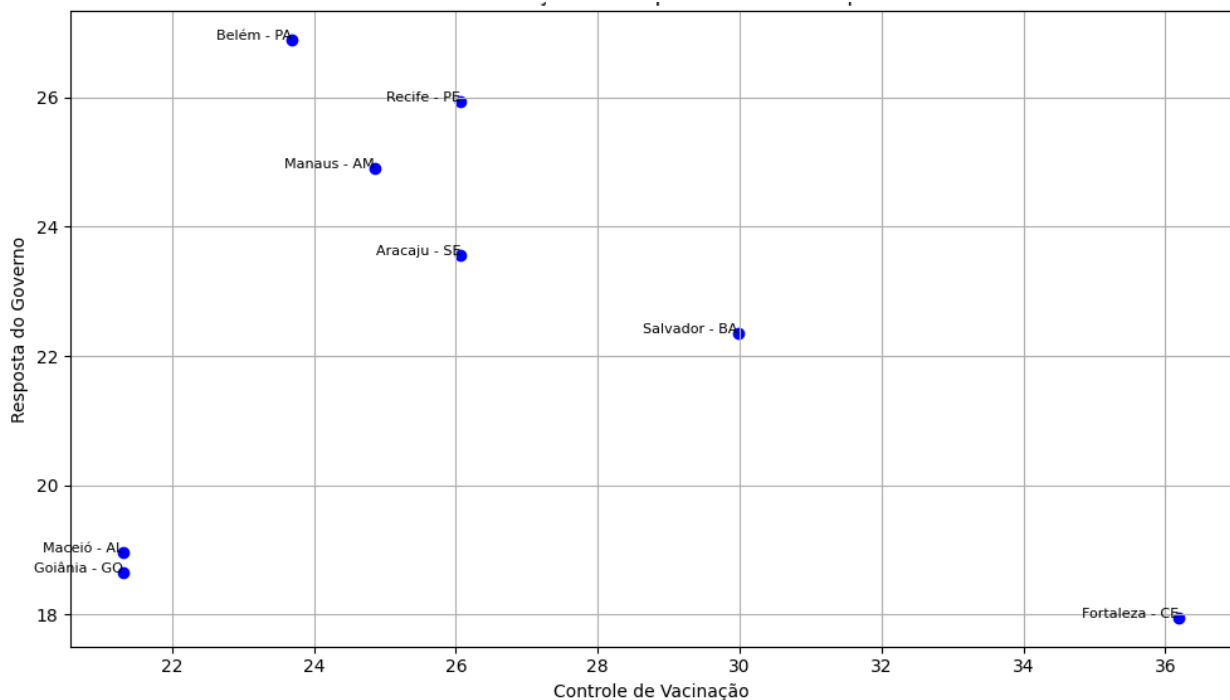


Fonte: elaborado pelo autor (2023).

Na Figura 8, é possível visualizar de maneira clara a dispersão das cidades em suas respectivas regiões geográficas.

Após algumas análises no *Cluster 0*, observa-se uma mistura de cidades provenientes de todas as regiões do país, onde temos uma grande maioria das cidades do Nordeste. Além disso, as cidades possuem pouca influência sobre as políticas públicas e o controle de vacinações em geral segundo o *stringency index*, o que parece contribuir para sua tendência de agrupamento, conforme pode-se observar na Figura 9.

Figura 9- Controle de Vacinação vs. Resposta do Governo por Cidade



Fonte: elaborado pelo autor (2023)

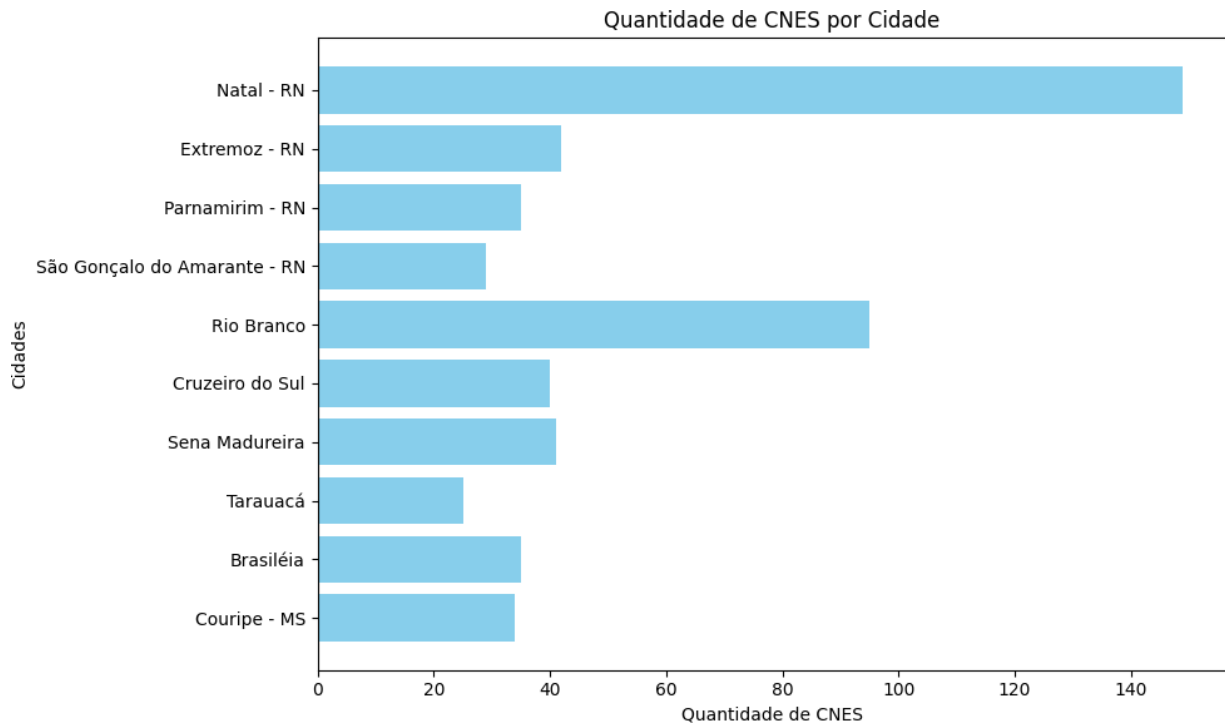
Conforme a evolução da pandemia se desenrolava, ao analisar as respostas governamentais direcionadas a várias cidades, como ilustrado na Figura 9, torna-se evidente que a maior parte das medidas adotadas pelo governo estava voltada para a implementação de restrições. Como podemos observar na Figura 9, as cidades que apresentam maiores valores no eixo Y, como é o caso de Belém

- Pará, demonstram uma resposta governamental significativa. No entanto, é importante ressaltar que, apesar dessas medidas robustas, o controle da vacinação nessas cidades permanece baixo, mesmo considerando a disponibilidade de vacinas, como evidenciado pelo gráfico.

De modo geral, as restrições aplicadas pelas cidades não se concentraram predominantemente no distanciamento social, mas sim em outras medidas restritivas, como o fechamento de escolas e o cancelamento de eventos, entre outras abordagens. Por outro lado, é interessante notar que algumas cidades exibem uma tendência oposta, como é o caso de Fortaleza - CE, que investe mais no controle da vacinação. Isso é evidenciado no eixo X, onde cidades com um maior controle de vacinação tendem a adotar menos medidas restritivas adicionais por parte do governo.

Já no *Cluster 1*, foram encontradas cidades que compartilham afinidades tanto em relação ao Produto Interno Bruto (PIB) quanto aos dados demográficos. Essas cidades também estão geograficamente próximas.

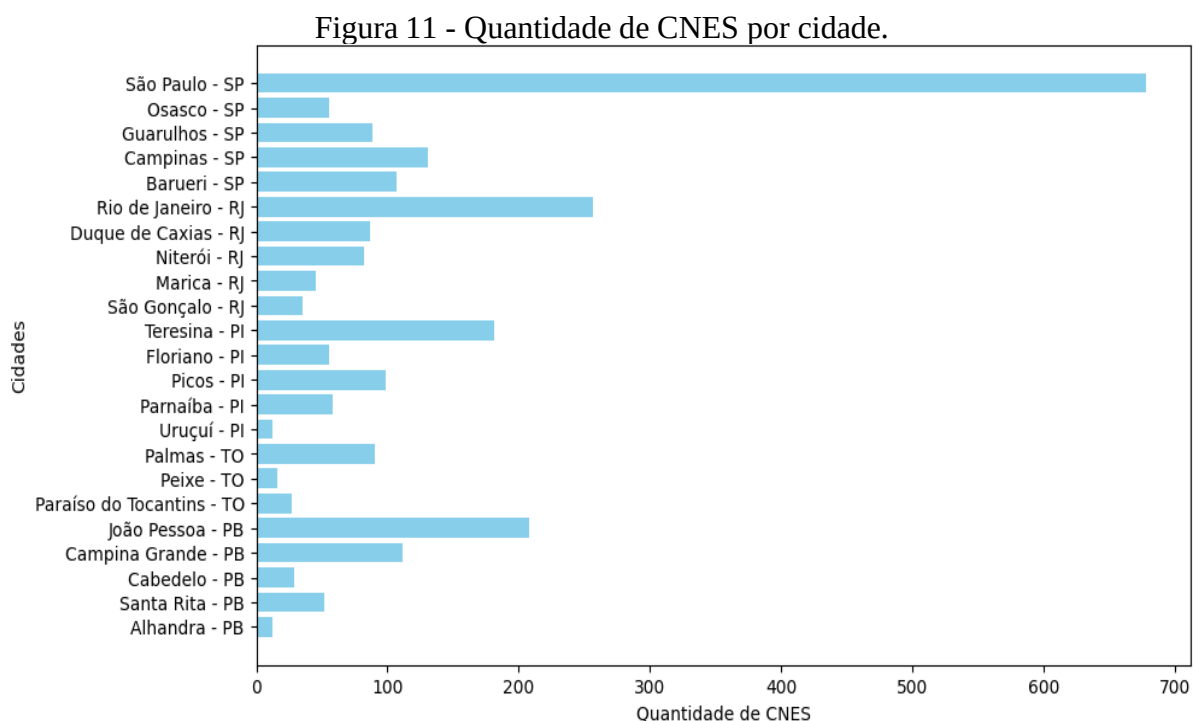
Figura 10 - Quantidade de CNES por cidade.



Fonte: elaborado pelo autor (2023)

Na Figura 10, podemos observar que a maioria das cidades apresenta uma quantidade de CNES inferior a 60 estabelecimentos de saúde. Esse cenário suscita uma preocupação considerável durante uma pandemia, uma vez que a escassez de leitos e recursos para atendimento médico se torna um fator de extrema importância.

Uma observação inicial que desperta interesse no *cluster 2* é a quantidade de CNES em cada cidade, conforme pode ser visualizado na Figura 11.

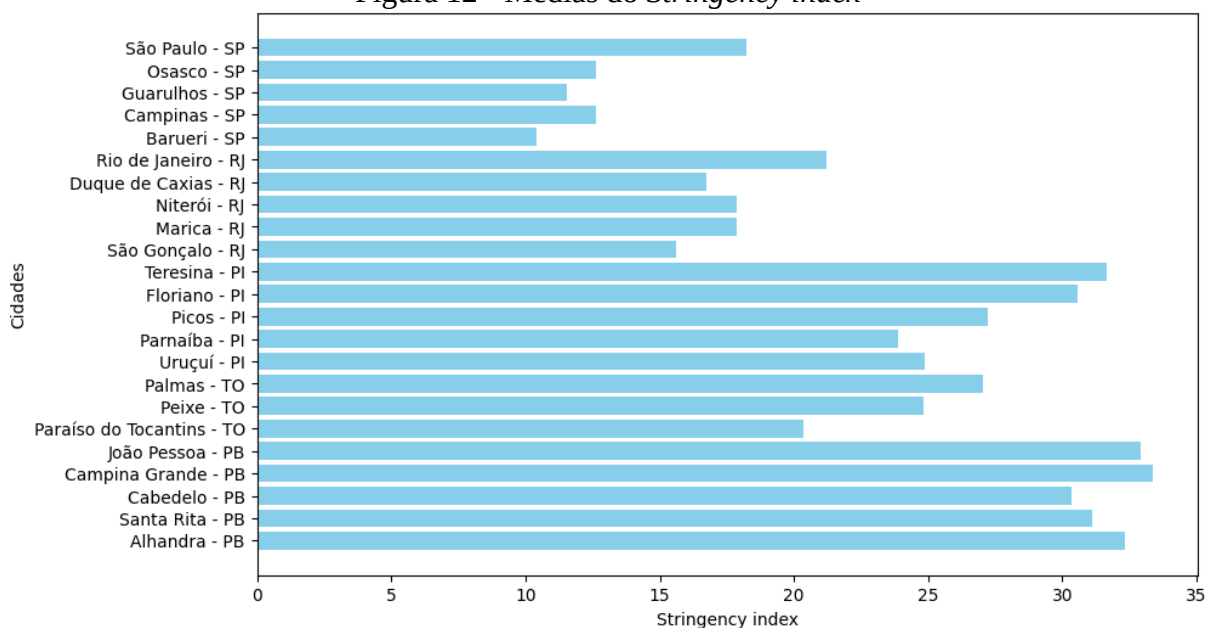


Fonte: elaborado pelo autor (2023)

Na Figura 11, é possível observar que as cidades realocadas para o *cluster 2* exibem quantidades de CNES equivalentes, com algumas variações notáveis.

No entanto, é importante destacar que as capitais, que também estão presentes neste *cluster*, tendem a apresentar quantidades de CNES mais elevadas. Por outro lado, a mudança de *cluster* não se deve apenas a esses números. O *stringency index*, que avalia o rigor das medidas de restrição, também desempenha um papel. Dentro do *cluster 2*, essas cidades destacam-se por possuírem os maiores índices de *stringency index*, dentre as 131 cidades analisadas. De acordo com esse método, essas cidades demonstraram um desempenho acima das demais cidades durante a pandemia, conforme apresentado na Figura 12.

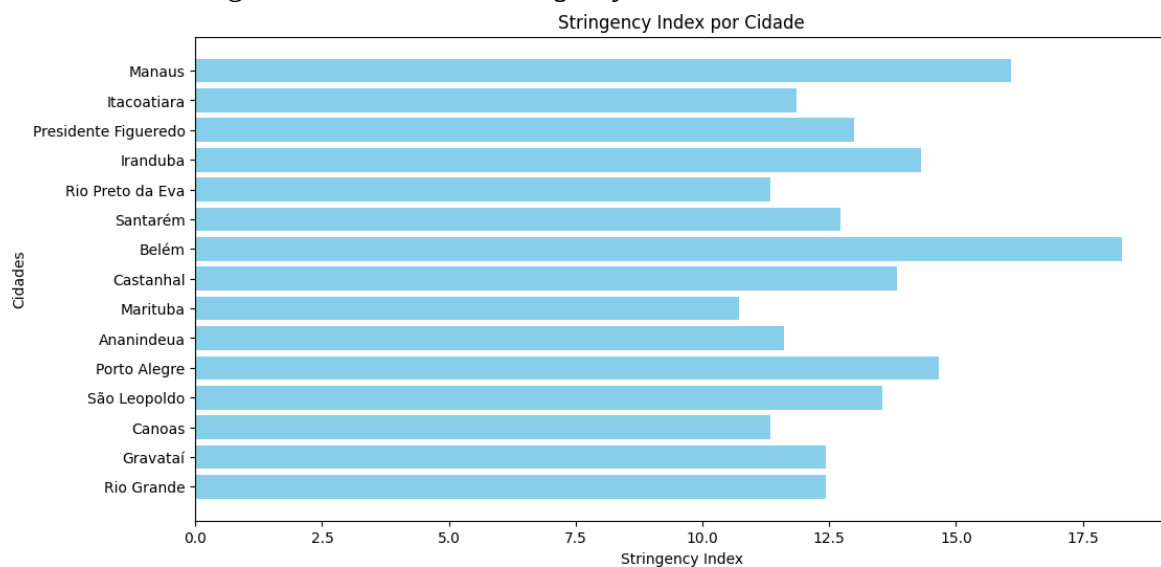
Figura 12 - Médias do *Stringency index*



Fonte: elaborado pelo autor (2023)

Como podemos observar nos dados do *Cluster 2*, a maioria das cidades apresentou um desempenho semelhante durante o período da pandemia, conforme ilustrado na Figura 12, baseada no *Stringency Index*. Essas cidades do *Cluster 2* exibem médias superiores em comparação aos demais *clusters*, como podemos ver na Figura, o que sugere uma abordagem mais eficaz no enfrentamento da pandemia.

Figura 13 - Médias do *Stringency index* dos demais Clusters



Fonte: elaborado pelo autor (2023)

Conforme observado nas demais cidades dos clusters apresentada no Figura 13, algumas cidades apresentam médias significativamente mais baixas no Stringency Index do que o Cluster 2. Isso indica que as cidades pertencentes a este último cluster tiveram um desempenho notavelmente melhor durante a pandemia em comparação com as demais.

Ademais, é importante notar que, dentro desse *Cluster 2*, as cidades não apenas possuem quantidades de CNES equivalentes à sua densidade populacional, mas também demonstram uma combinação de fatores que influenciam seu desempenho. Respostas governamentais ágeis, índices de contenção da propagação da doença, que frequentemente incluem campanhas de vacinação, suporte econômico adequado e implementação de políticas públicas eficazes, tudo isso parece contribuir para o sucesso dessas cidades em lidar com os desafios da pandemia.

#### 4 CONCLUSÃO

No decorrer deste estudo, foi investigado de forma abrangente a dinâmica das cidades brasileiras durante o período da pandemia. Utilizando técnicas de análise de *clusters*, onde foi identificado agrupamentos distintos entre as cidades, o que nos permitiu compreender suas similaridades e diferenças em termos de indicadores econômicos, demográficos e de saúde.

Ao explorar os resultados, notamos a formação de *clusters* que não apenas refletiam aspectos econômicos e geográficos, mas também revelavam a influência de fatores regionais, como infraestrutura de saúde e políticas públicas. Especificamente, o *Cluster 1* reuniu cidades de São Paulo, Rio de Janeiro, além de algumas cidades do Nordeste, indicando afinidades em termos de PIB, localização e indicadores de saúde. O *Cluster 0*, por outro lado, agrupou cidades de diversas regiões com menor presença de estabelecimentos de saúde e menor influência nas políticas de vacinação.

A aplicação dos dados de Covid-19 teve influência sobre os dados e revelou mudanças nos agrupamentos originais, com algumas cidades sendo realocadas para o *Cluster 2*. Nesse novo agrupamento, observamos cidades que, apesar de possuírem quantidades equivalentes de CNES, compartilhavam desempenhos notavelmente consistentes durante a pandemia. O *Stringency Index* destacou-se como um indicador importante, sugerindo que cidades com rigorosas medidas de contenção e estratégias de resposta governamental mais eficazes tendiam a se destacar no *Cluster 2*.

Quanto às limitações deste estudo, é essencial destacar os desafios que enfrentamos no acesso aos dados. Não foi possível obter uma base de dados completa, e encontramos dificuldades no acesso às fontes de dados disponíveis.

No entanto, a contribuição deste trabalho é fundamental para a compreensão do desempenho de diferentes cidades brasileiras durante o período da pandemia. Como resultado, conseguimos identificar as cidades que se destacaram positivamente durante esse período. Para pesquisas futuras, é importante analisar essas mesmas cidades em futuras ondas de Covid-19, a fim de avaliar se elas conseguiram manter o mesmo padrão de desempenho.

## REFERÊNCIAS

ALPAYDIN, E. **Introduction to machine learning**. 2º ed. Cambridge: MIT Press, 2010.

BARBERIA, L. *et al.* The effect of state-level social distancing policy stringency on mobility in the states of Brazil. **Revista de Administração Pública**, [S.L.]. 2021; vol. 55(1): 27-49. FapUNIFESP (SciELO). <http://dx.doi.org/10.1590/0034-761220200549>. Disponível em: <http://dx.doi.org/10.1590/0034-761220200549>. Acesso 06/08/2023.

BRASIL. **Painel Coronavírus**: SRAG Casos por região. Publicado 2020. Disponível em: [www.covid.saude.gov.br](http://www.covid.saude.gov.br). Acesso em 25/04/2023.

BRASIL. **COVID-19**. S/ d. Disponível em: [https://brasil.io/dataset/covid19/caso\\_full/](https://brasil.io/dataset/covid19/caso_full/). Acesso em 01/07/2023a.

BRASIL. **Dados abertos Ministério da Saúde**. S/ d. Disponível em: <https://dados.gov.br/dados/organizacoes/visualizar/ministerio-da-saude>. Acesso 01/07/2023b.

CHEN, L. *et al.* Exploratory Data Analysis on the Usage of COVID-19 Vaccine. 2021 6Th International Symposium On Computer And Information Processing Technology (Iscipt). **IEEE**. 2021; vol. 2(1): 1-5. Disponível em: <http://dx.doi.org/10.1109/iscipt53667.2021.00107>. Acesso em 10/06/2023.

COSSA, G. S. *et al.* Measures to fight the COVID-19 pandemic and the impact of health systems: a comparative analysis between brazil, italy, and the usa. **O Mundo da Saúde**. 2021; vol. 45: 379-389. Disponível em: <http://dx.doi.org/10.15343/0104-7809.202145379389>. Acesso em 08/07/2023.

CEPI. **New Vaccines For A Safer World**. S/ d. Disponível em: <https://cepi.net>. Acesso em 11/07/2023.

DARAPANENI, N. *et al.* **Machine learning approach for clustering of countries to identify the best strategies to combat Covid-19**. IEEE. 2021: 1-7. Disponível em: <https://ieeexplore.ieee.org/document/123456789>. Acesso em: 03/03/2023.

DATA, O. I. COVID-1: **Stringency index**. S/ d. Disponível em: <https://ourworldindata.org/covid-stringency-index>. Acesso em 01/07/2023.

FONTELLES, M. *et al.* Metodologia da pesquisa científica: diretrizes para a elaboração de um protocolo de pesquisa. **Revista paraense de medicina**. 2009; vol. 23(3): 1-8.

GAO, Q. *et al.* Rapid development of an inactivated vaccine for SARS- CoV-2. **Bui Rxiv**. 2020: 1-2. Disponível em: <https://www.biorxiv.org/content/10.1101/2020.04.17.046375v1>. Acesso em



27/05/2023.

HAAS, A. *et al.* Worldwide Data on Government Responses to COVID-19. **Nature Human Behaviour**. 2020; vol. 4: 879-888. Disponível em: <https://www.nature.com/articles/s41562-020-01009-0>. Acesso em 25/05/2023.

IBGE. **Catálogo**. S/d. Disponível em: <https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=2101637>. Acesso em 01/07/2023.

LUDERMIR, T. B. Inteligência Artificial e Aprendizado de Máquina: estado atual e tendências. **FapUNIFESP**. 2021; vol. 35(101): 85-94. 2021. Disponível em: <http://dx.doi.org/10.1590/s0103-4014.2021.35101.007>. Acesso em 27/05/2023.

MACIEL, J. A. C.; CASTRO-SILVA, I. I.; FARIAS, M. R. Análise inicial da correlação espacial entre a incidência de COVID-19 e o desenvolvimento humano nos municípios do estado do Ceará no Brasil. **Revista Brasileira de Epidemiologia**. 2020; vol. 23(1):2-3. Disponível em: <http://dx.doi.org/10.1590/1980-549720200057>. Acesso em 22/05/2023.

OMS. **Covid-19**. S/ d. Disponível em: <https://www.who.int/pt/emergencies/disease-outbreak-news/item/2020-DON-27-january-2020-novel-coronavirus-China>. Acesso em 12/02/2023.

OPAS. **Vacinas contra COVID-19**. S/d. Disponível em: <https://www.paho.org/pt/vacinas-contracovid-19>. Acesso em 28/05/2023.

SAS. **Aprendizado de máquina**. Disponível em: [https://www.sas.com/pt\\_br/insights/analytics/machine-learning.html#:~:text=O%20aprendizado%20de%20máquina%20\(em,o%20mínimo%20de%20inter venção%20humana](https://www.sas.com/pt_br/insights/analytics/machine-learning.html#:~:text=O%20aprendizado%20de%20máquina%20(em,o%20mínimo%20de%20inter%20venção%20humana). Acesso em 12/02/2023.

SILVA, L. A.; SANTOS, C. M. Impactos da Pandemia do Coronavírus (COVID- 19) na Economia e na Saúde Mental da População. **Revista de Gestão em Sistemas de Saúde**. 2020; vol. 9(1): 3796-3812. Disponível em: <https://doi.org/10.5585/rgss.v9i1.3796>. Acesso em 12/02/2023.

SILVA, L. L. S. *et al.* Medidas de distanciamento social para o enfrentamento da COVID-19 no Brasil: caracterização e análise epidemiológica por estado. **Cadernos de Saúde Pública FapUNIFESP**. 2020; vol. 36(9): 1-15. Disponível em: <http://dx.doi.org/10.1590/0102-311x00185020>. Acesso em 06/07/2023.

WHO. **Coronavirus disease (COVID-19)**. S/d. Disponível em: [https://www.who.int/health-topics/coronavirus#tab=tab\\_1](https://www.who.int/health-topics/coronavirus#tab=tab_1). Acesso em 06/07/2023.