

**ANÁLISE TEMPORAL E ESPACIAL DOS CASOS DE COVID-19 NAS REGIÕES GEOGRÁFICAS
IMEDIATAS DO BRASIL**

**TEMPORAL AND SPATIAL ANALYSIS OF COVID-19 CASES IN IMMEDIATE GEOGRAPHIC
REGIONS OF BRAZIL**

Pedro Loures Alzamora

Departamento de Ciência da computação, Universidade Federal de Minas Gerais
ploures.alzamora@gmail.com

Daniel Victor Ferreira

Departamento de Ciência da computação, Universidade Federal de Minas Gerais
vicdaniel.ferreira@gmail.com

Isadora Cristina de Matos Rodrigues

Departamento de Ciência da computação, Universidade Federal de Minas Gerais
isadorac2008@hotmail.com

Julia Fonseca de Sena

Departamento de Ciência da computação, Universidade Federal de Minas Gerais
jufs128@gmail.com

Marcelo Ganem

Departamento de Ciência da computação, Universidade Federal de Minas Gerais
masganem@gmail.com

Marcelo Sartori Locatelli

Departamento de Ciência da computação, Universidade Federal de Minas Gerais
marcelosartlocatelli@gmail.com

Thiago Henrique Moreira Santos

Departamento de Ciência da computação, Universidade Federal de Minas Gerais
thiago.santos@dcc.ufmg.br

Evandro Landulfo Teixeira Paradela Cunha

Faculdade de Letras, Universidade Federal de Minas Gerais
evandrocunha@dcc.ufmg.br

Janaina Teodoro Guiginski

Centro de Desenvolvimento e Planejamento Regional, Universidade Federal de Minas Gerais
janainaguiginski@gmail.com

Ramon Adrian Salinas Franco

Centro das Ciências Exatas e das Tecnologias, Universidade Federal do Oeste da Bahia
ramon.franco@ufob.edu.br

Tereza Bernardes

Universidade de Minas Gerais
tbernardesfaria@gmail.com

Ana Paula Couto da Silva

Departamento de Ciência da computação, Universidade Federal de Minas Gerais
ana.coutosilva@dcc.ufmg.br

Wagner Meira Junior

Departamento de Ciência da computação, Universidade Federal de Minas Gerais
meira@dcc.ufmg.br

RESUMO

O geoprocessamento de dados e as análises espaciais são ferramentas importantes para o estudo de fenômenos como a disseminação de doenças pelo território e ao longo do tempo. O objetivo deste estudo é investigar, utilizando a Análise Exploratória de Dados Espaciais (AEDE), as alterações nos padrões de distribuição geográfica da Covid-19 no Brasil em dois períodos distintos da pandemia: (i) entre abril e agosto de 2020; e (ii) entre novembro de 2020 e março de 2021. Para tanto, as estatísticas I de Moran e LISA foram aplicadas aos dados referentes a três indicadores epidemiológicos: casos acumulados, novos casos e letalidade da doença. Os resultados encontrados e as visualizações propostas apresentam uma perspectiva ampla sobre a variação nos casos de Covid-19 nas regiões brasileiras e colaboram para um melhor entendimento sobre as dinâmicas epidemiológicas no Brasil no primeiro ano da pandemia de Covid-19.

Palavras-chave: Coronavírus. Pandemia. Indicadores epidemiológicos. I de Moran. LISA. Geografia da saúde.

ABSTRACT

Data geoprocessing and spatial analysis are important tools for the study of phenomena such as the spread of diseases across the territory and over time. The aim of this study is to analyze, using Exploratory Spatial Data Analysis (ESDA), changes in the geographic distribution patterns of Covid-19 in Brazil in two distinct periods of the pandemic: (i) between April and August 2020; and (ii) between November 2020 and March 2021. For this purpose, Moran's I and LISA were applied to data referring to three epidemiological indicators: accumulated cases, new cases and lethality of the disease. The results found and the visualizations proposed present a broad perspective on the variation of Covid-19 cases in Brazilian regions and contribute to a better understanding of the epidemiological dynamics in Brazil in the first year of the Covid-19 pandemic.

Keywords: Coronavirus. Pandemic. Epidemiological indicators. Moran's I. LISA. Health geography.

INTRODUÇÃO

O geoprocessamento de dados e as análises espaciais podem auxiliar na compreensão de fenômenos complexos, heterogêneos e autocorrelacionados, como é o caso da disseminação de doenças ao longo do tempo e do território (ALBUQUERQUE e RIBEIRO, 2020; GUIMARÃES et al., 2020). Diversos estudos brasileiros utilizaram a análise espacial para investigar a pandemia de Covid-19. Cardoso et al. (2020) fizeram o mapeamento dos casos confirmados até abril de 2020 por bairros no município de São Gonçalo (RJ). Também nesse estado, Cavalcante e Abreu (2020) utilizaram ferramentas de análise espacial para descrever a dispersão de casos, óbitos e vários indicadores na cidade do Rio de Janeiro até o início de abril de 2020. Para o Ceará, Pedrosa e Albuquerque (2020) chamaram a atenção para a concentração dos casos e recursos hospitalares no estado e indícios preocupantes de interiorização da doença logo nas primeiras semanas após o registro do primeiro caso, em meados de março de 2020. Abarcando todo o Brasil, Lima et al. (2021) investigaram os casos e mortes por Covid-19 nas microrregiões brasileiras, considerando características da população, como distribuição por sexo e idade, e disponibilidade de serviços de saúde, como quantidade de médicos e leitos de hospital. Guimarães et al. (2020) também consideraram todo o território nacional ao defender o uso de mapas e ferramentas próprias da Geografia para analisar a disseminação da Covid-19.

Nesse contexto, o presente artigo tem o objetivo de investigar a existência de padrões espaciais significativos traçados pelos casos de Covid-19 e óbitos ocorridos no Brasil por meio da Análise Exploratória de Dados Espaciais (AEDE). A AEDE compreende um conjunto de técnicas estatísticas para análise de informações geográficas com o propósito central de identificar padrões de associação espacial nos dados, impondo a menor estrutura possível, ou seja, deixando que os dados falem por si mesmos (ALMEIDA, 2004). Para tanto, foram calculadas as estatísticas I de Moran e LISA (*Local Indicators of Spatial Association*) para três indicadores epidemiológicos referentes à Covid-19 no Brasil: (i) casos acumulados por 100 mil habitantes; (ii) novos casos semanais por 100 mil habitantes; e (iii) letalidade da doença (razão entre óbitos por Covid-19 e casos confirmados). A unidade básica de análise é formada pelas regiões geográficas imediatas do Brasil e a investigação compreende o período de abril de 2020 a março de 2021 – isto é, aproximadamente o primeiro ano da pandemia.

O propósito central deste artigo é avaliar a evolução da pandemia do ponto de vista espacial e temporal simultaneamente, oferecendo uma perspectiva atualizada, ampla e abrangente no tempo e no espaço da Covid-19. Espera-se que, dentre as contribuições do presente artigo, a consideração das mudanças na distribuição geográfica da Covid-19 ao longo do tempo e em todo o território nacional seja das mais importantes.

O artigo está em quatro partes, incluindo esta introdução. A seguir, são apresentados a fonte de dados e os métodos utilizados. Na terceira parte, são apresentados os principais resultados encontrados. Por fim, na quarta parte são traçadas as considerações finais a título de conclusão.

DADOS E MÉTODOS

Fonte de dados e indicadores epidemiológicos

Os dados utilizados nas análises foram coletados a partir do Projeto Brasil.io². Os dados englobam o total de casos e óbitos diários de Covid-19 ocorridos em todos os municípios brasileiros, compilados pelas 27 Secretarias Estaduais de Saúde. Tal nível de detalhamento territorial e temporal confere grande versatilidade aos dados, que podem, assim, ser agrupados em diferentes períodos e níveis geográficos. A partir destes dados e da projeção bayesiana da população estimada por Freire et al. (2019) para os municípios de cada região imediata em 2020, foram calculados três indicadores do avanço da pandemia de Covid-19 (MOREIRA et al., 2021):

- total de casos acumulados por 100.000 habitantes;
- novos casos semanais por 100.000 habitantes;
- letalidade da Covid-19, definida como a razão entre o total de óbitos e o total de casos confirmados da doença.

O cálculo dos indicadores em razão da população (isto é, indicadores por 100.000 habitantes) permite a comparação entre regiões com diferentes contingentes populacionais. A Tabela 1 apresenta um trecho da base de dados utilizada, ilustrando os dados disponíveis e os três indicadores calculados. De modo a garantir maior estabilidade nos dados, os valores diários foram agregados em valores semanais e os municípios foram agregados em regiões geográficas imediatas. Assim, os três indicadores foram calculados por semana epidemiológica e para cada região geográfica imediata do Brasil. As regiões geográficas imediatas foram selecionadas como unidades de análise por terem uma precisão espacial maior do que o nível estadual e, ao mesmo tempo, uma flutuação de dados menor do que o nível municipal (IBGE, 2017).

Com o objetivo de verificar a existência de padrões espaciais significativos para os valores de cada indicador epidemiológico, foram aplicadas técnicas de AEDE, que permitem classificar as regiões

² <https://brasil.io/home/>

Análise temporal e espacial dos casos de Covid-19 nas regiões geográficas imediatas do Brasil

Pedro Loures Alzamora
Daniel Victor Ferreira
Isadora Cristina de Matos Rodrigues
Julia Fonseca de Sena
Marcelo Ganem
Marcelo Sartori Locatelli
Thiago Henrique Moreira Santos
Evandro L. Teixeira Paradela Cunha
Janaina Teodoro Guiginski
Ramon Adrian Salinas Franco
Tereza Bernardes
Ana Paula Couto da Silva
Wagner Meira Junior

geográficas imediatas em grupos homogêneos segundo a variável de interesse. A seguir, são apresentadas, em linhas gerais, as estatísticas de autocorrelação espacial utilizadas. Para as estimativas de autocorrelação global, foi empregada a estatística I de Moran global e, para detectar padrões locais de autocorrelação, utilizou-se a estatística LISA (I de Moran local).

Tabela 1 – Trecho exemplificativo da base de dados contendo dados processados e tratados dos casos acumulados por 100.000 habitantes, novos casos por 100.000 habitantes e letalidade da Covid-19. Brasil, 2020-2021.

Região	UF	Região imediata	Código IBGE	Semana Epidemiológica	Data início da semana	População estimada 2020	Total casos acumulados	Total óbitos acumulados	Novos casos na semana	Casos acum./ 100.000 hab.	Novos casos/ 100.000 hab.	Letalidade = óbitos/casos * 100
Norte	PA	Abaetetuba	150003	16	12/4/2020	439624	6	0	4	1,36	0,91	0,00
Sudeste	MG	Abaeté	310070	16	12/4/2020	39168	1	0	1	2,55	2,55	0,00
Nordeste	CE	Acaraú	230017	16	12/4/2020	216119	4	1	3	1,85	1,39	25,00
Sudeste	SP	Adamantina - Lucélia	350019	16	12/4/2020	134556	1	0	1	0,74	0,74	0,00
Nordeste	PE	Afogados da Ingazeira	260014	16	12/4/2020	167955	2	1	2	1,19	1,19	50,00
(...)												
Norte	PA	Abaetetuba	150003	48	22/11/2020	439624	8582	232	82	1952,12	18,65	2,70
Sudeste	MG	Abaeté	310070	48	22/11/2020	39168	187	3	21	477,43	53,61	1,60
Nordeste	CE	Acaraú	230017	48	22/11/2020	216119	8351	150	108	3864,08	49,97	1,80
Sudeste	SP	Adamantina - Lucélia	350019	48	22/11/2020	134556	1611	44	127	1197,27	94,38	2,73
Nordeste	PE	Afogados da Ingazeira	260014	48	22/11/2020	187106	5502	94	333	2940,58	177,97	1,71
(...)												

Fonte: Elaboração própria a partir dos dados das Secretarias Estaduais de Saúde e consolidados pelo Projeto Brasil.io.

I de Moran global

A estatística I de Moran fornece um único valor como medida da associação espacial de todos os polígonos, representando, nesse caso, a correlação linear do indicador epidemiológico entre as regiões. Quando positivo e estatisticamente significativo, indicado pelo p-valor do índice, o coeficiente I de Moran global indica a presença de correlação espacial positiva, ou seja, regiões com valores altos para determinado indicador epidemiológico tendem a se agrupar no espaço, assim como regiões com valores baixos tendem a formar agrupamentos com regiões semelhantes (ALMEIDA, 2004; REY et al., 2020). Seguindo a fórmula proposta por Moran (1948), o coeficiente de autocorrelação I de Moran global pode ser calculado por:

$$I = \frac{n}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} z_i z_j}{\sum_i z_i^2}$$

sendo n o número de observações (unidades espaciais), z_i e z_j os valores padronizados do indicador epidemiológico no par de regiões imediatas i e j , respectivamente, e w_{ij} o peso espacial para o par de regiões i e j na matriz de pesos espaciais W . A matriz de pesos espaciais W foi definida com base na ideia de vizinhança do tipo *queen* (rainha), que considera como vizinhas as regiões que possuem uma fronteira comum ou um nó (vértice) comum. Se duas regiões são vizinhas, W assume valor 1; caso contrário, assume valor 0.

Estatística LISA (I de Moran local)

As estatísticas LISA (*Local Indicators of Spatial Association*) foram originalmente propostas por Anselin (1995) e fornecem um valor para cada região, permitindo identificar agrupamentos espaciais significativos. Para identificar esses agrupamentos, utiliza-se o diagrama de dispersão de Moran e o mapa de *clusters*. O mapa de *clusters* combina a significância das medidas de associação local – o mapa de significância – com as informações do diagrama de dispersão de Moran. O diagrama de dispersão informa a medida global de associação linear – I de Moran (representado pela reta de regressão) – e o tipo de associação linear para cada observação. Existem quatro tipos possíveis de associação espacial linear: alto-alto, baixo-baixo, alto-baixo e baixo-alto.

Desse modo, a estatística I de Moran local permite identificar padrões locais de associação linear que sejam estatisticamente significativos. O cálculo é semelhante ao do I de Moran global, mas, neste caso, é aplicado para cada uma das observações. Ou seja, cada unidade geográfica recebe um valor semelhante ao I de Moran global que descreve o seu grau de similaridade ou dissimilaridade em relação às localidades vizinhas. Esse valor é denominado I de Moran local e sua representação formal pode ser descrita por:

$$I_i = \frac{z_i}{m_2} \sum_j w_{ij} z_j ; m_2 = \frac{\sum_i z_i^2}{n}$$

em que m_2 é a variância da distribuição de valores, $z_i = y_i - \bar{y}$, w_{ij} é o peso espacial para o par observado e n é o número de observações.

Método de análise e apresentação dos resultados

O período de análise divide a pandemia em dois intervalos de tempo. O primeiro intervalo compreende as semanas epidemiológicas entre 12 de abril de 2020 e 08 de agosto de 2020 (semanas epidemiológicas 16/2020 a 32/2020) e será referido no texto como primeira fase da pandemia. O segundo intervalo de tempo, que corresponde à segunda fase da pandemia, compreende o período entre 22 de novembro de 2020 e 20 de março de 2021 (semanas epidemiológicas 48/2020 a 11/2021). Essa divisão foi necessária devido à ausência de dados desagregados para os municípios do estado do Mato Grosso entre os meses de agosto e novembro de 2020, período que foi desconsiderado na análise.

As estatísticas de autocorrelação espacial de cada indicador epidemiológico foram calculadas para cada uma das 32 semanas epidemiológicas consideradas na análise. Foi utilizada a biblioteca de código aberto para análise espacial PySal, disponibilizada por Rey e Anselin (2010) e Rey et al. (2020) em linguagem Python. Os códigos e roteiro de análise foram desenvolvidos no ambiente Jupyter Notebook, disponibilizados publicamente como repositório na plataforma Github³. No total, foram gerados 96 mapas de *clusters* LISA – 32 mapas para cada um dos três indicadores epidemiológicos. Tradicionalmente, as análises exploratórias de dados espaciais não trabalham com um volume tão grande de mapas. Em geral, as comparações empregam indicadores de caráter cumulativo (casos totais, por exemplo) e analisam poucos pontos no tempo (GUIMARÃES et al.,

³[CDA-EPCWeb/Analise-Espacial \(github.com\)](https://github.com/CDA-EPCWeb/Analise-Espacial)

2020; SZWARCOWALD et al., 2000). Neste trabalho, assim como Hass (2021), optamos por utilizar indicadores de caráter instantâneo (novos casos por semana, por exemplo), tornando necessária a criação de uma nova metodologia de sumarização dos dados.

Nesta metodologia, os 96 mapas originais foram sumarizados em seis mapas, que representam a situação dos novos casos, casos acumulados e letalidade na primeira e na segunda fases da pandemia. Ou seja, os 32 mapas de *clusters* LISA gerados para cada indicador epidemiológico foram convertidos em dois mapas, cada um representando uma fase da pandemia. Cada mapa expressa o tipo de *cluster* predominante no período considerado. Os *clusters* são apresentados nas cores azul e vermelho. A cor azul representa os *clusters* do tipo baixo-baixo, a cor vermelha representa os *clusters* do tipo alto-alto, a cor verde representa os *clusters* baixo-alto e a cor amarelo os *clusters* alto-baixo. A intensidade da cor representa a consistência com que cada região imediata esteve incluída em um determinado tipo de *cluster*. Por exemplo, se uma determinada região imediata foi classificada mais vezes como pertencente a um *cluster* do tipo alto-alto do que a outro tipo de *cluster*, esta região é apresentada em vermelho; quanto mais intensa a cor, maior o número de semanas em que esta região integrou aquele *cluster*.

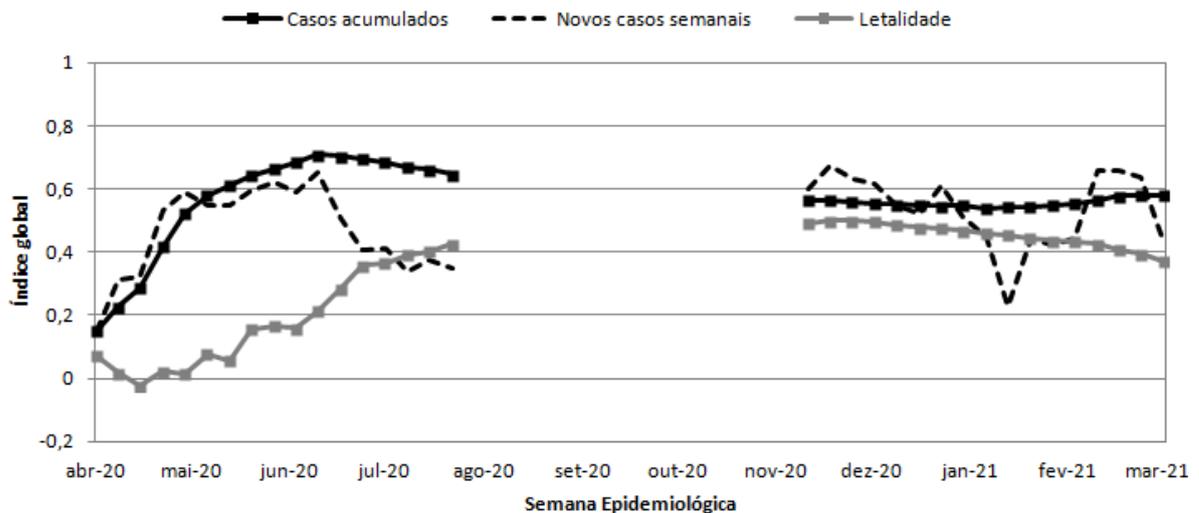
Com o intuito de facilitar a compreensão e não extrapolar o escopo do trabalho, não foram analisados os *clusters* do tipo alto-baixo e baixo-alto. Esses *clusters*, embora relevantes para uma compreensão mais profunda da distribuição regional dos indicadores, destacam as regiões que destoam significativamente de seus vizinhos em termos epidemiológicos, o que foge ao objetivo deste artigo de observar regiões positivamente correlacionadas com seus vizinhos. Essa estratégia permitiu que a informação de todas as semanas epidemiológicas fosse condensada em uma única visualização, ressaltando inclusive as situações epidemiológicas mais representativas de cada fase nas regiões geográficas imediatas.

RESULTADOS E DISCUSSÃO

Para todas as semanas, os coeficientes de correlação espacial I de Moran foram positivos e estatisticamente significativos, indicando que o total de casos, os novos casos semanais e a letalidade da Covid-19 não estão aleatoriamente distribuídos no território. Ou seja, regiões vizinhas tendem a apresentar valores similares para um mesmo indicador. A Figura 1 apresenta o valor do índice de Moran global, calculado a cada semana epidemiológica, para os três indicadores considerados. Para todos os indicadores, os valores da estatística I de Moran são positivos e mostram tendência de crescimento nas primeiras semanas e estabilidade no segundo período observado, com exceção do I de Moran para os novos casos semanais, que demonstra maior oscilação ao longo das semanas. Isso ocorre porque os indicadores de casos acumulados e de letalidade são calculados a partir de dados acumulados, proporcionando a esses indicadores uma tendência de estabilização com o passar do tempo. Nas primeiras semanas da pandemia, os casos acumulados e a letalidade estavam sujeitos a grandes mudanças. Entretanto, à medida que o tempo avança, a distribuição espacial dos indicadores tende a mudar cada vez menos. Como e por que cada um dos indicadores se comporta individualmente será abordado nas próximas seções. Aqui, para entender a estabilização do I de Moran global, basta saber que, à medida que os casos se acumulam, torna-se mais difícil alterar suas configurações espaciais. Por isso, observamos menor variação nos valores da estatística I de Moran global com o passar do tempo.

Análise temporal e espacial dos casos de Covid-19 nas regiões geográficas imediatas do Brasil

Figura 1 – I de Moran global a cada semana epidemiológica para os indicadores: casos acumulados por 100.000 habitantes, novos casos por 100.000 habitantes e letalidade da Covid-19. Brasil, 2020-2021.



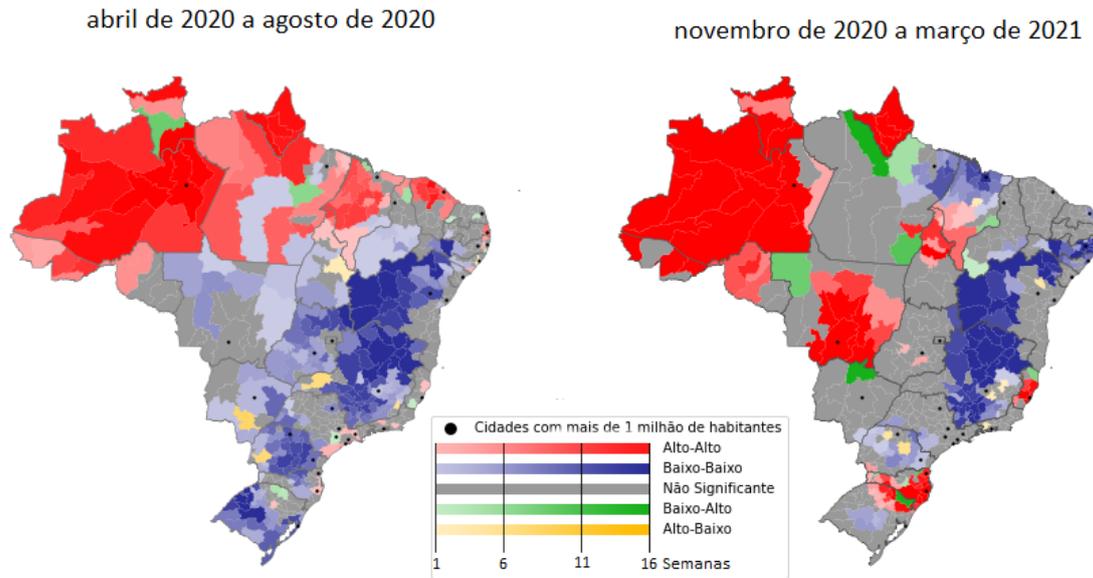
Fonte: Elaboração própria a partir dos dados das Secretarias Estaduais de Saúde e consolidados pelo Projeto Brasil.io.

Foram produzidos dois mapas para cada um dos indicadores epidemiológicos (isto é: total de casos por 100.000 habitantes, novos casos semanais por 100.000 habitantes e letalidade). As regiões imediatas predominantemente classificadas como *clusters* do tipo alto-alto – regiões com valores altos para o indicador epidemiológico em análise, rodeadas por vizinhos com valores altos – estão representadas em vermelho. As regiões predominantemente classificadas como baixo-baixo – regiões com baixos valores rodeadas por vizinhos com baixos valores para um mesmo indicador epidemiológico – estão apresentadas na cor azul. Em cinza, estão representadas as regiões sem autocorrelação espacial significativa ou com predominância de *cluster* do tipo alto-baixo ou baixo-alto. A intensidade das cores indica o número de vezes que a região integrou aquele agrupamento.

Casos acumulados por 100.000 habitantes

Os mapas de *clusters* para o indicador de casos acumulados por 100.000 habitantes estão dispostos na Figura 2. Antes de avaliar os mapas, é importante considerar que este indicador é calculado cumulativamente com base no total de casos dividido pela população e multiplicado por 100.000. Desse modo, o total de casos de uma semana é sempre condicionado e superior ao valor da semana anterior. Logo, espera-se que o número de ilhas com autocorrelação espacial significativa diminua com o tempo. Isso porque, à medida que o indicador se aproxima de 100.000 (seu próprio denominador), as variações semanais dos casos acumulados tornam-se cada vez menores. Ou seja, o total de casos torna-se cada vez mais condicionado pelo seu próprio valor na semana anterior e cada vez menos pelos valores das regiões vizinhas.

Figura 2 – Mapas de *clusters* para casos acumulados por 100.000 habitantes. Brasil, 2020-2021.



Fonte: Elaboração própria a partir dos dados das Secretarias Estaduais de Saúde e consolidados pelo Projeto Brasil.io.

A comparação entre os mapas da primeira e da segunda fase da pandemia, apresentados na Figura 2, indica que houve redução do tamanho dos agrupamentos com autocorrelação espacial significativa ao longo do tempo. Entretanto, algumas regiões chamam a atenção pela persistência de seus agrupamentos, tanto baixo-baixo quanto alto-alto. Um exemplo de persistência no agrupamento do tipo baixo-baixo é a região da bacia do Rio São Francisco. Possível explicação pode ser encontrada no fato de que a região é formada por cidades de pequeno e médio porte, pouco integradas à malha rodoviária federal, com menor tráfego de pessoas e, conseqüentemente, baixa taxa de transmissão do coronavírus. No sul de Minas Gerais também aparecem *clusters* do tipo baixo-baixo, predominantes nas duas fases da pandemia.

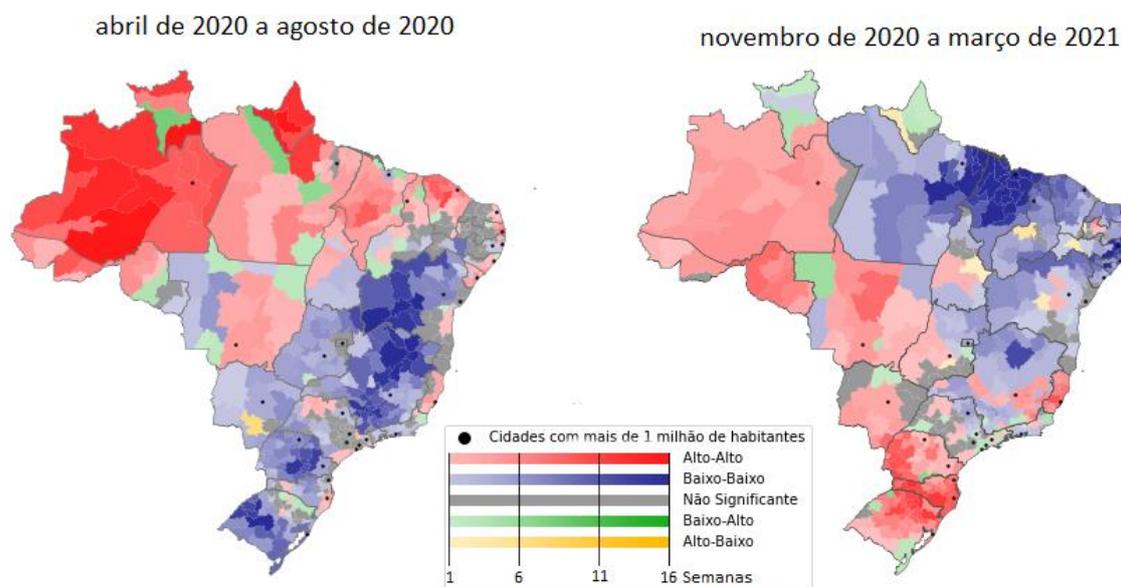
Em contrapartida, na maioria dos estados da região Norte (Acre, Amapá, Amazonas, Rondônia e Roraima) observa-se predominância de *clusters* do tipo alto-alto. O Norte aparece como uma região de alto risco, com estrutura etária jovem e em piores condições socioeconômicas do que as demais (LIMA et al., 2021). Além disso, o surgimento de uma nova cepa mais transmissível do coronavírus em Manaus pode ter contribuído para o agravamento da situação nesta região (FREITAS et al., 2021; NAVECA et al., 2021).

Na primeira fase da pandemia, as regiões geográficas imediatas do Mato Grosso e de Santa Catarina eram áreas com baixa concentração de casos, mas tornaram-se regiões de alta concentração no segundo período. O surgimento de *clusters* do tipo alto-alto também é observado no norte do Tocantins e no Espírito Santo, ponto que será abordado na próxima subseção, a partir dos mapas de novos casos por semana. Por sua vez, nas regiões próximas a Belém (PA) e São Luís (MA), observou-se o surgimento de *clusters* do tipo baixo-baixo em substituição aos *clusters* alto-alto. Como possível explicação, destacamos que esses estados estabeleceram medidas de isolamento mais rígidas no começo da pandemia (SCHAFER et al., 2020).

Novos casos semanais por 100.000 habitantes

A Figura 3 apresenta os mapas de *clusters* para o indicador de novos casos semanais de Covid-19 nas duas fases da pandemia. Assim como o indicador de casos acumulados, o indicador de novos casos também é relativo à população. Porém, enquanto os casos acumulados em uma semana, por definição, estão associados aos casos registrados nas semanas anteriores, os novos casos não são cumulativos ao longo do tempo. Os novos casos semanais podem aumentar ou diminuir de uma semana para a outra sem estarem, necessariamente, relacionados com os valores registrados na semana anterior.

Figura 3 – Mapa de *clusters* para novos casos semanais por 100.000 habitantes. Brasil, 2020-2021.



Fonte: Elaboração própria a partir dos dados das Secretarias Estaduais de Saúde e consolidados pelo Projeto Brasil.io.

Na primeira fase da pandemia, pode-se observar que o mapa de novos casos, apresentado na Figura 3, não difere muito do mapa de casos acumulados mostrado anteriormente na Figura 2. Isso ocorre porque o indicador de casos acumulados pode ser interpretado como um valor acumulado de novos casos. Assim, ambos os indicadores tendem a ser mais semelhantes no início do período analisado. Enquanto o total de casos informa sobre o efeito acumulado da doença no decorrer da pandemia, os novos casos descrevem o impacto imediato dela. Na segunda fase, porém, é visível a neutralização de alguns e a intensificação de outros *clusters* de novos casos.

Entre a primeira e a segunda fases da pandemia, observa-se uma mudança no eixo geográfico do indicador de novos casos semanais. Na primeira fase, a pandemia afligiu muito mais e com muito mais força o eixo Norte-Nordeste (Leste-Oeste), regiões com população mais jovem e em pior situação socioeconômica, em relação ao restante do país. Na segunda fase, o Nordeste passou a ser uma região menos afetada, com predominância de *clusters* do tipo baixo-baixo para o indicador de novos casos.

Segundo Schaefer et al. (2020), sete estados brasileiros do Norte e Nordeste foram mais rápidos na implementação de intervenções para enfrentar o coronavírus e decretaram medidas de fechamento das escolas e a suspensão de eventos públicos antes mesmo da confirmação do primeiro caso de Covid-19 no estado. Estas medidas parecem ter surtido efeitos positivos no médio prazo para a ocorrência de novos casos da doença no Amapá, Pará, Maranhão e Piauí, estados que deixaram de apresentar predominância de *clusters* alto-alto na segunda fase da pandemia. Em Roraima e Tocantins, ainda aparecem *clusters* alto-alto, mas surgiram novos *clusters* baixo-baixo. Uma diferença encontrada entre esses dois estados e os quatro citados anteriormente é que Roraima e Tocantins, embora tenham decretado medidas antecipadas de isolamento, não possuíam um plano formalizado de flexibilização das medidas de distanciamento social até o final de julho de 2020 (MORAES et al., 2020). A única exceção é Mato Grosso que, apesar de também ter antecipado medidas de isolamento, apresentou aumento dos *clusters* do tipo alto-alto na segunda fase da pandemia.

Acre, Amazonas e Rondônia, na região Norte, e os estados das regiões Centro-Oeste e Sul passaram a ser mais afetados na segunda fase, com agrupamentos significativos de localidades com alta incidência de casos por semana, formando um eixo mais inclinado, na direção Norte-Sul. Na segunda fase da pandemia, o estado de Minas Gerais chama a atenção por formar um caminho de ilhas alto-alto que corta o estado em várias direções. Essas ilhas se situam às margens da BR-381, que liga São Paulo ao Espírito Santo, passando por Belo Horizonte e Governador Valadares; e da BR-262, que passa por Uberaba e Belo Horizonte e leva a Vitória, no Espírito Santo. O caminho identificado também conta com a influência da BR-356 e da BR-354, que ligam o Distrito Federal e Goiás ao estado do Rio de Janeiro. Além disso, a região conta também com ferrovias, que conectam Uberlândia, Belo Horizonte e Vitória.

Em perspectiva mais ampla, pode-se notar um caminho de *clusters* alto-alto de novos casos que parte do Mato Grosso, passa pela porção inferior do estado de Goiás e pelo caminho já citado em Minas Gerais, chegando ao Espírito Santo. Em outras palavras, existe um conjunto de *clusters* com alta incidência de casos que se sobrepõe às regiões imediatas pertencentes ao Corredor Sudeste (Eixo Vitória) de exportação da soja e milho. Esta é uma evidência importante de que o sistema de transportes – e consequente circulação de pessoas – atua como um vetor de transmissão da Covid-19 e de interiorização da doença, tal qual antecipado por Silva et al. (2019), que chegaram a um entendimento semelhante ao analisar os primeiros sinais da pandemia no estado da Bahia em abril de 2020. A partir desse resultado, pode-se levantar a hipótese de que o agronegócio de exportação (mercado da soja, por exemplo) colaborou para a interiorização da doença, possivelmente por envolver transporte rodoviário, modalidade que agrega mais pessoas do que outros meios de transporte. Corroborando esta hipótese a inclusão dos caminhoneiros e demais trabalhadores do transporte entre os grupos prioritários da vacinação em janeiro de 2021 pelo Ministério da Saúde (CNT, 2021). Devido à grande mobilidade dos trabalhadores do setor, além de risco mais alto de contrair a infecção e disseminá-la ao longo do trajeto, em várias regiões do Brasil, os caminhoneiros passaram a ter prioridade no recebimento da vacina de dose única da Janssen (ESTADÃO, 2021; EXAME, 2021).

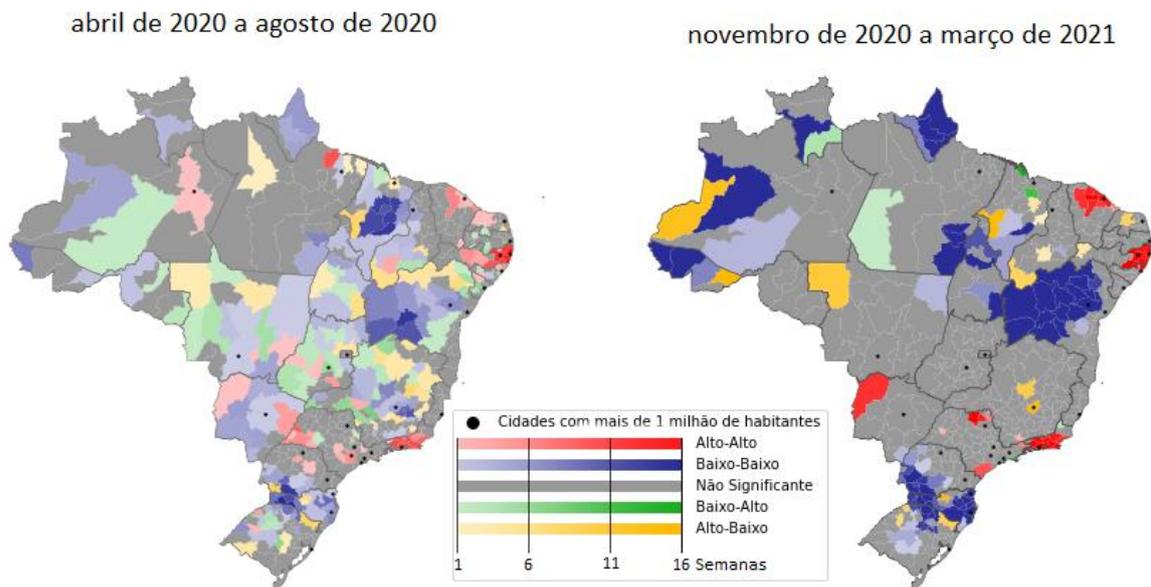
Letalidade

Os mapas de *clusters* para a letalidade estão dispostos na Figura 4. O indicador de letalidade é calculado como a razão entre os óbitos acumulados e os casos acumulados. Assim como o indicador de casos acumulados, a letalidade também é calculada de maneira cumulativa. Porém, enquanto o indicador de casos acumulados sempre aumenta e tende à estabilidade à medida que se aproxima de 100.000 (caso em que quase toda a população terá sido infectada), a letalidade pode aumentar ou diminuir e tende a se estabilizar em diferentes níveis, dependendo da região.

O aumento ou decréscimo da letalidade reflete a relação entre o número de óbitos e casos registrados e qualquer mudança em um ou em outro altera seu valor. Como este indicador tende a se

estabilizar com o passar do tempo, a análise dos *clusters* na segunda fase da pandemia é mais interessante.

Figura 4 – Mapa de *clusters* para a letalidade da Covid-19. Brasil, 2020-2021.



Fonte: Elaboração própria a partir dos dados das Secretarias Estaduais de Saúde e consolidados pelo Projeto Brasil.io.

Ao comparar as duas fases da pandemia, quase não se registra o surgimento de novos *clusters* no segundo período, sendo mais comum a filtragem de regiões e a intensificação das cores representadas no mapa. Observa-se também que não há casos de inversão de predominância, isto é, nenhuma região que possuía uma classificação predominantemente alta se tornou uma região de predominância baixa, e vice-versa.

Uma observação particular está no fato de que as regiões que formaram *clusters* do tipo alto-alto para letalidade – Pantanal no Mato Grosso do Sul, Ribeirão Preto e litoral sul de São Paulo, Rio de Janeiro e litoral do Ceará – não apresentaram agrupamentos significativos para o indicador de casos acumulados. Uma exceção é observada no litoral de Pernambuco, que apresentou *cluster* baixo-baixo para total de casos e *cluster* alto-alto para letalidade, o que pode estar relacionado a um baixo nível de testagem de casos suspeitos nesta região. Segundo dados da última edição da PNAD Covid-19, de novembro de 2020, Pernambuco foi o segundo estado brasileiro que menos aplicou testes para detecção da doença em sua população. Enquanto 13,5% da população brasileira havia realizado algum teste diagnóstico de Covid-19, apenas 9,3% da população de Pernambuco havia sido testada (IBGE, 2020).

Algumas regiões com baixa letalidade (Figura 4) pertencem a agrupamentos com valores altos para casos acumulados (Figura 2). Esse é o caso do Acre, oeste do Amazonas, Amapá, norte do Tocantins e praticamente todo o estado de Santa Catarina. Isso indica que o aumento no total de casos tem sido, em média, proporcionalmente superior ao aumento do número de óbitos. Ou seja,

esses agrupamentos com baixa letalidade são também agrupamentos com grande quantidade de casos acumulados.

E, finalmente, é interessante destacar o interior da Bahia, que apresenta *clusters* do tipo baixo-baixo tanto para os indicadores de casos, novos ou acumulados, quanto para a letalidade nas duas fases da pandemia.

CONCLUSÃO

Neste trabalho, é apresentada uma análise espacial e temporal dos casos de Covid-19 e sua letalidade nas regiões geográficas imediatas do Brasil. O maior desafio encontrado foi a necessidade de lidar com uma grande quantidade de dados brutos, o que demandou a utilização de técnicas computacionais de manipulação e geoprocessamento dos dados e ferramentas de Análise Exploratória de Dados Espaciais (AEDE), de modo a condensar as informações e extrair conhecimentos sobre a evolução da pandemia. Além disso, os dados foram normalizados em relação à população residente em cada localidade, pois os números brutos, sem qualquer ajuste populacional, não permitem comparações regionais e poderiam levar a conclusões equivocadas sobre a situação sanitária.

Ao investigar dados sobre dois períodos distintos da pandemia no país (abril a agosto de 2020 e novembro de 2020 a março de 2021), observou-se a autocorrelação espacial global positiva para os casos de Covid-19 e existência de *clusters* espaciais significativos. Os resultados mostraram a persistência de agrupamentos com baixos níveis de casos acumulados de Covid-19 por 100.000 habitantes na bacia do Rio São Francisco e sul de Minas Gerais. Pará e Maranhão apresentaram evolução positiva, com o aparecimento de *clusters* baixo-baixo na segunda fase da pandemia. Por outro lado, observou-se a predominância de agrupamentos com altos níveis de casos acumulados em alguns estados da região Norte e o surgimento de agrupamentos desse tipo no Mato Grosso e em Santa Catarina. A análise do indicador de novos casos semanais de Covid-19 por 100.000 habitantes expõe uma mudança importante no eixo geográfico de incidência de casos ao longo do tempo. Na primeira fase havia uma concentração de *clusters* alto-alto nas regiões Nordeste e Norte (eixo Leste-Oeste), situação que é alterada visivelmente no período seguinte. Na região Nordeste, passaram a predominar *clusters* baixo-baixo e, por outro lado, a presença de *clusters* alto-alto intensificou-se nos estados das regiões Centro-Oeste e Sul. Como resultado, o eixo geográfico de maior incidência de casos de Covid-19 é observado no sentido Norte-Sul. Outro ponto de destaque está na identificação de um conjunto de *clusters* significativos no Corredor Sudeste de exportação (Eixo Vitória), que forma um caminho de regiões com altos níveis de novos casos semanais de Covid-19 a partir de novembro de 2020.

O estudo aqui apresentado possui algumas limitações relevantes. Uma delas é o fato de que os mapas apresentados condensam a informação sobre o tipo de *cluster* predominante em cada fase, mas não mostram se houve maior predominância no início ou no final de cada período, ou seja, se há tendência de aumento ou redução ao longo do tempo. Outra limitação é que o mapa de *clusters* para a letalidade da Covid-19 (Figura 4) fornece informações parciais, dificultando a comparação entre regiões, já que pressupõe que o nível de testagem de casos suspeitos é semelhante em todas as regiões. Uma forma de contornar essa dificuldade seria incluir mapas para mortalidade (número de óbitos por 100.000 habitantes) de modo a complementar a análise sobre a letalidade. Outros indicadores epidemiológicos não incluídos aqui têm o potencial de enriquecer a análise e merecem maior investigação em trabalhos futuros. Esse é o caso do número relativo de mortes, por semana e acumulado, fatores de crescimento de casos e óbitos, além de indicadores mais refinados que controlem o nível de testagem das regiões. Outro ponto importante a ser explorado está na ampliação do tempo de análise, de modo a incluir o segundo ano da pandemia, que conta com mudanças importantes no cenário geral, como a crescente flexibilização das medidas de isolamento social, e a vacinação, que, embora em alguns momentos tenha sido insuficiente e com ritmo inconstante, pode

ter tido efeitos sobre os casos e a letalidade da Covid-19. Além disso, trabalhos futuros também poderão abordar a questão da autocorrelação temporal dos indicadores, principalmente de casos acumulados e letalidade, por usarem os valores acumulados de casos e óbitos.

De maneira geral, conclui-se que a metodologia de análise espacial dos casos de Covid-19 nas regiões geográficas imediatas do Brasil por meio de técnicas de geoprocessamento de dados e análise espacial temporal simultânea proposta no presente artigo apresenta uma visão abrangente do aumento dos casos de Covid-19 nas regiões brasileiras, contribuindo para um melhor entendimento das dinâmicas epidemiológicas no país no primeiro ano da pandemia em questão.

REFERÊNCIAS

- ALBUQUERQUE, M. V.; RIBEIRO, L. H. L. Desigualdade, situação geográfica e sentidos da ação na pandemia da COVID-19 no Brasil. **Cadernos de Saúde Pública**, v. 36, n. 12, 2020. <https://doi.org/10.1590/0102-311x00208720>
- ALMEIDA, E. S. **Curso de Econometria Espacial Aplicada**. Piracicaba, 2004.
- ANSELIN, L. Local indicators of spatial association—LISA. **Geographical Analysis**, v. 27, n. 2, p. 93–115, 1995. <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>
- CARDOSO, P. V.; SEABRA, V. S.; BASTOS, I. B.; COSTA, E.C. P. A importância da análise espacial para tomada de decisão: um olhar sobre a pandemia de COVID-19. **Revista Tamoios**, v. 16, n. 1, p. 125–137, 2020. <https://doi.org/10.12957/tamoios.2020.50440>
- CAVALCANTE, J. R.; ABREU, A. J. L. COVID-19 no município do Rio de Janeiro: análise espacial da ocorrência dos primeiros casos e óbitos confirmados. **Epidemiologia e Serviços de Saúde**, v. 29, n. 3, e2020204, 2020. <https://doi.org/10.5123/S1679-49742020000300007>
- CNT - Confederação Nacional do Transporte. Acompanhe o ritmo de vacinação dos trabalhadores do transporte pelo país. **Agência CNT Transporte Atual**, Brasília, 25 maio 2021. Disponível em: <https://cnt.org.br/agencia-cnt/acompanhe-o-ritmo-de-vacinacao-dos-trabalhadores-do-transporte-pelo-pais>. Acesso em: 24 ago. 2021.
- ESTADÃO. Caminhoneiros podem receber dose única da vacina contra covid-19. **Estadão**, São Paulo, 29 jun. 2021. Disponível em: <https://estradao.estadao.com.br/caminhoes/caminhoneiros-podem-receber-dose-unica-da-vacina-contracovid-19/>. Acesso em: 24 ago. 2021.
- EXAME. Governos aproveitam Janssen para vacinar caminhoneiros e moradores de rua. **Exame**, 28 jun. 2021. Disponível em: <https://exame.com/brasil/governos-aproveitam-janssen-para-vacinar-caminhoneiros-e-moradores-de-rua/>. Acesso em: 24 ago. 2021.
- FREIRE, F. H. M. A.; GONZAGA, M. R.; QUEIROZ, B. L. Projeção populacional municipal com estimadores bayesianos, Brasil 2010-2030. **Seguridade Social Municipal**. Projeto Brasil, v. 3, 2019.
- FREITAS, A. R. R.; GIOVANETTI, M.; ALCANTARA, L. C. J. Emerging variants of SARS-CoV-2 and its public health implications. **InterAmerican Journal of Medicine and Health**, v. 4, 2021.
- GUIMARÃES, R. B.; CATÃO, R. C.; MARTINUCI, O. S.; PUGLIESI, E. A.; MATSUMOTO, P. S. S. O raciocínio geográfico e as chaves de leitura da Covid-19 no território brasileiro. **Estudos Avançados**, v. 34, n. 99, p. 119–140, 2020. <https://doi.org/10.1590/s0103-4014.2020.3499.008>
- HASS F. S; ARSANJANI J. J., 2021. **The Geography of the Covid-19 Pandemic: A Data-Driven Approach to Exploring Geographical Driving Forces** *International Journal of Environmental Research and Public Health* 18, 6 (2021). <https://doi.org/10.3390/ijerph18062803>
- IBGE - Instituto Brasileiro de Geografia e Estatística. **Divisão regional do Brasil em regiões geográficas imediatas e regiões geográficas intermediárias**. Rio de Janeiro: IBGE, 2017.

IBGE - Instituto Brasileiro de Geografia e Estatística. 28,6 milhões de pessoas fizeram algum teste para a Covid-19 até novembro. Agência IBGE Notícias, Rio de Janeiro, 23 dez. 2020. Disponível em: <https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/29771-28-6-milhoes-de-pessoas-fizeram-algum-teste-para-a-covid-19-ate-novembro>. Acesso em: 13 jul. 2021.

LIMA, E. E. C.; GAYAWAN, E.; BAPTISTA, E. A.; QUEIROZ, B. L. Spatial pattern of COVID-19 deaths and infections in small areas of Brazil. **PLOS One**, v. 16, n. 2, e0246808, 2021. <https://doi.org/10.1371/journal.pone.0246808>

MOREIRA, Pedro et al. Covid Data Analytics: Repositório de Dados Provenientes de Múltiplas Fontes sobre a Pandemia de COVID-19 no Brasil. In: DATASET SHOWCASE WORKSHOP (DSW), 3. , 2021, Rio de Janeiro. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2021. p. 107-116. <https://doi.org/10.5753/dsw.2021.17419>

MORAES, R. F.; SILVA, L. L. S.; TOSCANO, C. M. **Covid-19 e medidas de distanciamento social no Brasil**: análise comparativa dos planos estaduais de flexibilização. Brasília: Ipea, 2020. (Nota Técnica, n. 25).

MORAN, P. A. P. The Interpretation of Statistical Maps. **Journal of the Royal Statistical Society. Series B (Methodological)**, v. 10, n. 2, 1948, p. 243–251. <https://doi.org/10.1111/j.2517-6161.1948.tb00012.x>

NAVECA, F. et al. SARS-CoV-2 reinfection by the new Variant of Concern (VOC) P. 1 in Amazonas, Brazil. **virological.org**, 2021. Disponível em: <https://virological.org/t/spike-e484k-mutation-in-the-first-sars-cov-2-reinfection-case-confirmed-in-brazil-2020/584>. Acesso em: 14 jul. 2021.

PEDROSA, N. L.; ALBUQUERQUE, N. L. S. Análise Espacial dos Casos de COVID-19 e leitos de terapia intensiva no estado do Ceará, Brasil. **Ciência & Saúde Coletiva**, v. 25, p. 2461–2468, 2020. <https://doi.org/10.1590/1413-81232020256.1.10952020>

PROJETO BRASIL.IO. **ESPECIAL COVID-19—Dados por Município**. Boletins informativos e casos do coronavírus por município por dia. Disponível em: https://brasil.io/dataset/covid19/caso_full/. Acesso em: 01 maio 2021.

REY, S. J.; ANSELIN, L. PySAL: A Python library of spatial analytical methods. In: **Handbook of applied spatial analysis**. Springer: Berlin, Heidelberg, 2010. p. 175–193. https://doi.org/10.1007/978-3-642-03647-7_11

REY, S. J.; ARRIBAS-BEL, D.; WOLF L. J. Part II - Spatial Data Analysis. In: REY, S. J.; ARRIBAS-BEL, D.; WOLF L. J. **Geographic Data Science with PySAL and the PyData Stack**. 2020. Disponível em: https://geographicdata.science/book/intro_part_ii.html. Acesso em: 16 jul. 2021.

SCHAEFER, B. M.; RESENDE, R. C., EPITÁCIO, S. D. S. F.; ALEIXO, M. T.. Ações governamentais contra o novo coronavírus: evidências dos estados brasileiros. **Revista de Administração Pública**, v. 54, p. 1429–1445, 2020. <https://doi.org/10.1590/0034-761220200503>

SILVA, R. J.; SILVA, K. B.; MATTOS, J. B.. Análise espacial sobre a dispersão da covid-19 no Estado da Bahia. Technical Note. **SciELO Preprints**, 2020. <https://doi.org/10.1590/SciELOPreprints.39>

SZWARCWALD, C. L. et. al. 2000. A disseminação da epidemia da AIDS no Brasil, no período de 1987-1996: uma análise espacial. **Cadernos de Saúde Pública** 16 (2000), S07–S19. <https://doi.org/10.1590/S0102-311X2000000700002>