

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ**

**AMILTON LUCIANO GARCIA DA SILVA**

**ANÁLISE ESPACIAL DA TAXA DE CRESCIMENTO DO NÚMERO DE CASOS DE  
COVID-19 NOS BAIRROS DE TOLEDO-PR**

**TOLEDO**

**2021**

**AMILTON LUCIANO GARCIA DA SILVA**

**ANÁLISE ESPACIAL DA TAXA DE CRESCIMENTO DO NÚMERO DE CASOS DE  
COVID-19 NOS BAIRROS DE TOLEDO-PR**

**SPATIAL ANALYSIS OF THE GROWTH RATE OF THE NUMBER OF CASES OF  
COVID-19 IN THE NEIGHBORHOOD OF TOLEDO-PR**

Trabalho de conclusão de curso de graduação  
apresentado como requisito para obtenção do título de  
Licenciado em Matemática da Universidade  
Tecnológica Federal do Paraná (UTFPR).

Orientador: Prof. Dr. Gustavo Henrique Dalposso.

**TOLEDO**

**2021**



[4.0 Internacional](https://creativecommons.org/licenses/by-sa/4.0/)

Esta licença permite remixe, adaptação e criação a partir do trabalho, mesmo para fins comerciais, desde que sejam atribuídos créditos ao(s) autor(es) e que licenciem as novas criações sob termos idênticos. Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

**AMILTON LUCIANO GARCIA DA SILVA**

**ANÁLISE ESPACIAL DA TAXA DE CRESCIMENTO DO NÚMERO DE CASOS DE  
COVID-19 NOS BAIRROS DE TOLEDO-PR**

Trabalho de Conclusão de Curso de Graduação  
apresentado como requisito para obtenção do título de  
Licenciado em Matemática da Universidade  
Tecnológica Federal do Paraná (UTFPR).

Data de aprovação: 09/dezembro/2021

---

Daniela Trentin Nava  
Doutorado  
Universidade Tecnológica Federal do Paraná

---

Gustavo Henrique Dalposso  
Doutorado  
Universidade Tecnológica Federal do Paraná

---

Márcio Paulo de Oliveira  
Doutorado  
Universidade Tecnológica Federal do Paraná

**TOLEDO  
2021**

Dedico este trabalho aos meus pais, Lúcia Garcia e Milton Ramiro, pelo apoio incondicional em todos os momentos difíceis da minha trajetória acadêmica.

Também, à tia Circe Bonotto, que com bons conselhos sempre me incentivou, e à meu avô Aristides Garcia (*in memoriam*).

## AGRADECIMENTOS

Agradeço ao meu orientador Prof. Dr. Gustavo Henrique Dalposso, pela sabedoria com que me guiou nesta trajetória, e rendeu uma grande amizade.

Gostaria de deixar registrado também, o meu reconhecimento à minha família, especialmente aos meus pais Milton Ramiro e Lucia Garcia, e à tia Circe Bonotto, pois acredito que sem o apoio deles seria muito difícil vencer esse desafio.

Ao meu parceiro Alex Dias Mariano pela companhia excepcional e pela parceria durante todos esses anos de convivência.

À professora Vera Quadros que me ensinou a desvendar os mistérios da escrita.

Às professoras Simone Fermino e Mariza Zanini e ao professor Émerson Brugnerotto, que me mostraram a beleza dos números.

Ao corpo docente do curso, especialmente à professora Dra. Daniela Trentin Nava e ao professor Dr. Márcio Paulo de Oliveira.

À Deus por nunca me desamparar.

Ouse coisas grandiosas.

## RESUMO

A estatística espacial de áreas fornece metodologias que são de grande importância para estudos envolvendo dados relacionados à saúde, pois permite determinar a existência de padrões espaciais, identificando regiões que apresentam características semelhantes, o que permite a alocação de recursos financeiros e humanos na realização de tratamentos localizados, minimizando os impactos causados por doenças. Neste sentido, este trabalho utilizou as metodologias da estatística espacial de áreas para investigar a taxa de crescimento do número de casos de Covid-19 nos bairros de Toledo. Foram analisados dezesseis períodos compreendidos entre os dias 25/06/2020 a 20/10/2020. Elaborou-se mapas temáticos da taxa de crescimento, investigou-se a existência de autocorrelação espacial entre os bairros e gerou-se mapas identificando regiões em que a autocorrelação foi significativa. Concluiu-se que a estatística espacial de áreas é uma boa alternativa para estudar a taxa de contaminação por Covid-19, pois fornece de forma clara e precisa informações coerentes com a realidade, o que auxilia as instâncias decisórias na tomada de decisões.

Palavras-chave: estatística espacial de áreas; autocorrelação espacial; índice de Moran; LISA.

## **ABSTRACT**

The spatial statistics of areas provides methodologies that are of great importance for studies involving data related to health, as it allows to determine the existence of spatial patterns, identifying regions that have similar characteristics, which allows the allocation of financial and human resources in carrying out treatments located, minimizing the impacts caused by diseases. In this sense, this work used the methodologies of spatial statistics of areas to investigate the growth rate of the number of Covid-19 cases in Toledo neighborhoods. Sixteen periods between 06/25/2020 at 10/20/2020 were analyzed. Thematic maps of the growth rate were elaborated, the existence of spatial autocorrelation between neighborhoods was investigated, and maps were generated identifying regions where the autocorrelation was significant. It was concluded that the spatial statistics of areas is a good alternative to study the Covid-19 contamination rate, as it provides clear and precise information consistent with reality, which helps decision-makers in decision-making.

Keywords: spatial statistics of areas; spatial autocorrelation; Moran index; LISA.



## LISTA DE ILUSTRAÇÕES

<b>Figura 1 - Mapa dos 22 bairros do município de Toledo-PR</b>	<b>19</b>
<b>Figura 2 - Representação dos tipos de adjacência entre as áreas. (A) Adjacência <i>Queen</i> “rainha”, (B) Adjacência <i>Rock</i> “torre” e (C) Adjacência <i>Bishop</i> “bispo”</b>	<b>22</b>
<b>Figura 3 - Matriz de proximidade espacial, normalizada pelas linhas</b>	<b>23</b>
<b>Figura 4 - Construção do gráfico de espalhamento de Moran</b>	<b>25</b>
<b>Figura 5 - Legenda do mapa do gráfico de espalhamento de Moran</b>	<b>26</b>
<b>Figura 6 - Evolução temporal da Taxa de contaminação por Covid-19 nos bairros de Toledo com amplitude de classe igual a 40</b>	<b>31</b>
<b>Figura 7 - Evolução temporal da Taxa de contaminação por Covid-19 nos bairros de Toledo com amplitude de classe por quintis</b>	<b>32</b>
<b>Figura 8 - Diagrama de espalhamento de Moran (Moran scatterplot) da Taxa de contaminação por Covid-19 nos bairros de Toledo</b>	<b>35</b>
<b>Figura 9 - Mapa de significância LISA da Taxa de contaminação por Covid-19 nos bairros de Toledo. Apenas valores com significância acima de 95% estão mostrados</b>	<b>37</b>
<b>Figura 10 - LISA <i>cluster map</i> da Taxa de contaminação por Covid-19 nos bairros de Toledo</b>	<b>39</b>

## LISTA DE TABELAS

<b>Tabela 1 - Estatísticas descritivas da Taxa de contaminação por Covid-19</b>	<b>29</b>
<b>Tabela 2 - Índice I de Moran da Taxa de contaminação por Covid-19</b>	<b>34</b>

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>11</b>
<b>1.1</b>	<b>Objetivo Geral</b>	<b>12</b>
<b>1.2</b>	<b>Objetivos Específicos</b>	<b>12</b>
<b>2</b>	<b>REVISÃO BIBLIOGRÁFICA</b>	<b>13</b>
<b>2.1</b>	<b>Coronavírus (Covid-19)</b>	<b>13</b>
<b>2.2</b>	<b>Estatística Espacial</b>	<b>15</b>
2.2.1	Histórico	15
2.2.2	Tipos de análise espacial	16
<u>2.2.2.1</u>	<u>Eventos ou padrões pontuais</u>	<u>16</u>
<u>2.2.2.2</u>	<u>Geoestatística</u>	<u>17</u>
<u>2.2.2.3</u>	<u>Estatística espacial de áreas</u>	<u>17</u>
<b>3</b>	<b>MATERIAL E MÉTODOS</b>	<b>19</b>
<b>3.1</b>	<b>Material</b>	<b>19</b>
<b>3.2</b>	<b>Métodos</b>	<b>20</b>
3.2.1	Análise exploratória	20
3.2.2	Análise exploratória espacial	21
3.2.3	Matriz de proximidade espacial	21
3.2.4	Índice de autocorrelação global I de Moran	23
3.2.5	Gráfico de espalhamento de Moran	24
3.2.6	Índice de Autocorrelação Local – LISA	26
3.2.7	Mapa de significância LISA	27
3.2.8	Mapa de <i>cluster</i> LISA	27
3.2.9	Recursos computacionais	28
<b>4</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>29</b>
<b>4.1</b>	<b>Análise exploratória</b>	<b>29</b>
<b>4.2</b>	<b>Análise exploratória espacial</b>	<b>30</b>
<b>4.3</b>	<b>Índice de Moran</b>	<b>33</b>
<b>4.4</b>	<b>Moran <i>scatterplot</i></b>	<b>34</b>
<b>4.5</b>	<b>Índice LISA</b>	<b>36</b>
<b>5</b>	<b>CONCLUSÕES</b>	<b>41</b>
	<b>REFERÊNCIAS</b>	<b>42</b>

## 1 INTRODUÇÃO

A análise espacial de dados teve como precursor o médico inglês John Snow, que foi o primeiro a incorporar a localização espacial em uma investigação que ocorreu em 1854, em um surto de cólera em Londres. O doutor Snow localizou os casos de cólera, e fazendo o uso direto dos mapas da cidade, concluiu que a enfermidade estava relacionada com o endereço das residências e o consumo de água insalubre dos poços de água existentes naquelas ruas, conseguindo acabar com a epidemia (BRASIL, 2007).

Com o avanço tecnológico, o esforço em traduzir a estatística para uma linguagem mais acessível resultou no desenvolvimento de softwares que são conhecidos como Sistemas de Informações Geográficas ou SIGs, que viabilizam a criação, a manipulação e extração de informações de bases de dados geográficos. A utilização desses métodos têm proporcionado aplicações em diversas áreas, em especial ao setor da saúde (BRASIL, 2007; DRUCK et al., 2004).

Com a pandemia do vírus SARS-CoV-2 (*Severe Acute Respiratory Syndrome Coronavirus 2*), responsável por ocasionar infecções agudas e graves no trato respiratório, a análise espacial de dados tem dado grande suporte a gestores de saúde pública. Permitindo a visualização espacial dos problemas e a associação de outras variáveis espaciais que podem agravar o quadro. Além disso, possibilita a avaliação de tendências futuras, auxiliando as instâncias decisórias, com o propósito de reduzir o número de mortes e o colapso do sistema de saúde (Brasil, 2007; NOBRE et al., 2014).

Neste contexto, a estatística espacial possibilita determinar a existência de padrões espaciais, considerando a localização geográfica do fenômeno em estudo. As técnicas empregadas na estatística espacial difere das técnicas utilizadas na estatística clássica, por levar em consideração a localização das informações no processo de coleta, descrição ou análise dos dados (SEFFRIN, 2017).

A análise espacial considera basicamente três tipos de dados: (i) eventos ou padrões pontuais que são fenômenos expressos por ocorrências identificadas como pontos localizados no espaço, como por exemplo, localização de crimes e espécies vegetais; (ii) superfícies contínuas estimadas com base em um conjunto de amostras de campo podendo estar regularmente ou irregularmente distribuídas, usualmente esse tipo de dado é resultante de levantamento de recursos naturais; (iii) estatística

espacial de áreas - técnicas utilizadas quando não se dispõe da localização exata dos eventos, mas de um valor por área delimitada por polígonos, que podem representar cidades, bairros, talhões ou setores censitários ( DRUCK et al., 2004).

Assim, para aplicação das metodologias da estatística espacial de áreas no estudo da taxa de contaminação por Covid-19, escolheu-se os bairros do município de Toledo, que conta com uma população estimada em 150 mil habitantes, sendo o 12º município paranaense mais populoso e também considerada a capital do agronegócio do Paraná (IBGE, 2021).

### **1.1 Objetivo geral**

O objetivo geral do trabalho é utilizar a metodologia da estatística espacial de áreas para investigar a taxa de crescimento do número de casos de COVID-19 nos bairros de Toledo-PR.

### **1.1 Objetivos específicos**

Elaborar mapas temáticos da distribuição da taxa de crescimento do número de casos de COVID-19 nos bairros de Toledo-PR;

Investigar a existência de autocorrelação espacial da taxa de crescimento do número de casos de COVID-19 nos bairros de Toledo-PR;

Utilizar os *softwares* R para os cálculos, GeoDa e QGIS para o geoprocessamento.

## 2 REVISÃO BIBLIOGRÁFICA

### 2.1 Coronavírus (Covid-19)

Desde os primeiros relatos de um surto de pneumonia de origem desconhecida em Wuhan, província de Hubei, China, no final de 2019, os temores se ascenderam. Mais tarde, descobriu-se que as infecções eram causadas por um novo tipo de coronavírus, o SARS-CoV-2. A sua rápida proliferação a nível mundial fez com que no dia 11 de março de 2020, o diretor geral da Organização Mundial da Saúde (OMS), Tedros Adhanom, declarasse a infecção por Covid-19 como pandemia. Até este dia já havia mais de 118 mil casos em 114 países e mais de 4 mil óbitos (BBC, 2020).

O primeiro caso confirmado no Brasil foi em 26 de fevereiro de 2020, um homem de 61 anos que voltou de uma viagem à Itália (PINHEIRO; RUPRECHT, 2020). Em Toledo, o primeiro caso confirmado foi em 4 de abril de 2020, um homem de 55 anos (HANSEN, 2020).

Os coronavírus infectam humanos e animais e são responsáveis por ocasionar infecções agudas e graves no trato respiratório. O SARS-CoV-2 (*Severe Acute Respiratory Syndrome*) é o sétimo coronavírus conhecido a infectar humanos (ANDERSEN et al., 2020).

Os coronavírus são vírus de RNA de sentido positivo, ou seja, as sequências abertas de leitura (ORFs) que codificam as proteínas virais estão presentes na mesma orientação do genoma. Os vírus RNA são os únicos organismos que possuem RNA como genoma, precisando se adaptar a condições impostas pelas células hospedeiras para poder se multiplicar. As células eucariontes não sintetizam RNA a partir de moldes RNA, etapa fundamental para a replicação do genoma do vírus. Entretanto, a seleção natural solucionou este impasse, pois o genoma de um vírus RNA codifica a sua própria enzima replicativa (FLORES, 2007).

Os coronavírus têm composição proteica simples, com um conjunto básico de quatro espécies de proteína, a proteína nucleocapsídica (N), a proteína *spike* (S), a proteína de membrana (SM), e a glicoproteína de membrana (M) e uma glicoproteína de membrana adicional (HE) (ROTTIER, 1995).

O SARS-CoV-2 tem uma característica notável, pois pode se ligar ao ACE2 humano com alta afinidade, porém as simulações computacionais preveem que a interação não é ideal. Assim, esta ligação de alta afinidade da proteína *spike* ao ACE2

humano está associado resultado de seleção natural. Esta é uma forte evidência de que o vírus não é produto de manipulação intencional, pois se a manipulação genética tivesse sido realizada, um dos vários sistemas de genética reversa disponíveis para Betacoronavírus possivelmente teria sido utilizado. Deste modo, segundo (ANDERSEN et al., 2020) duas hipóteses podem explicar de forma plausível a origem do vírus:

1. Seleção natural em um animal hospedeiro antes da transferência zoonótica - como muitos dos primeiros casos são associados ao mercado de frutos do mar em Huanan em Wuhan, é possível que uma fonte animal estivesse no local, e que essa fonte muito possivelmente seriam os morcegos, dada a semelhança do SARS-CoV-2 com o SARS-CoV de morcego.

Os pangolins malaios, mamíferos ameaçados de extinção, vendidos ilegalmente no mercado de frutos do mar, também contém coronavírus semelhantes ao SARS-CoV-2.

2. Seleção natural em humanos após transferência zoonótica - é possível que um precursor do SARS-CoV-2 saltou para humanos, adquirindo características de adaptação durante a transmissão não detectada em humanos. Quando adquiridas, essas adaptações permitiram que a pandemia decolasse. Todos os genomas do SARS-Co-2 sequenciados até agora, mostram que as características genômicas derivam de um ancestral comum, que também as possuía.

Entretanto, vale destacar também que não deve-se descartar a possibilidade de transmissão da SARS-CoV-2 aos humanos por meio de fuga de laboratório, como já ocorrido em 2003 com o SARS-CoV (LIM et al., 2004).

A transmissão entre humanos se dá por gotículas, em que o vírus é carregado em pequenas aerossóis, originárias da boca e do nariz de pessoas infectadas ao falar, tossir ou espirrar (SOARES et al., 2021). A infecção pode ocorrer também quando uma pessoa toca superfícies ou objetos contaminados e imediatamente em seguida toca seus olhos, nariz ou boca. O tempo estimado do período de incubação do SARS-CoV-2 é de 14 dias, com mediana de 4 a 6 dias, no entanto há relatos de períodos de incubação de até 24 dias (BAI et al., 2020).

## 2.2 Estatística espacial

A estatística espacial é o ramo da estatística que inclui em sua análise a localização espacial dos fenômenos, ou seja, o objetivo é incorporar o espaço a análise que se deseja fazer (DRUCK et al., 2004).

Deste modo, a estatística espacial surge em complementação às técnicas estatísticas comuns, pois um dos pressupostos da estatística é a independência entre as observações. No entanto, quando observamos o espaço, esse pressuposto é pouco realista, pois de acordo com a primeira lei da geografia de (Tobler, 1979), “*todas as coisas são parecidas, mas coisas mais próximas se parecem mais que coisas mais distantes*” (BRASIL, 2007).

Assim, a maior parte das ocorrências, sejam elas sociais ou naturais, se relacionam, e isso se enfraquece com a distância. E é justamente esta a especificidade da análise espacial: técnicas que possibilitam modelar e gerenciar fenômenos que a distribuição é afetada pela localização geográfica (BRASIL, 2007).

### 2.2.1 Histórico

A estatística espacial teve como precursor o médico inglês John Snow, que foi o primeiro a fazer uso explícito de mapas em uma análise durante um surto de cólera em Londres, em 1854. O doutor Snow produziu diversos mapas a fim de avaliar sua hipótese sobre a transmissão da cólera, ele localizou as residências de cada caso de cólera, a idade, a profissão, dentre outras variáveis de interesse, incluindo a localização dos locais de captação de água. Assim, ele concluiu que as mortes por cólera eram principalmente em decorrência do consumo de água insalubre e se concentravam em torno do poço de água da *Broad Street*, atual *Broadwick Street* (BRASIL, 2007).

Um dos mapas mais conhecidos de John Snow mostra a concentração de pontos que representam as mortes por cólera, em que a medida que se afasta do centro do mapa, diminui a concentração de pontos, demonstrando um padrão de distribuição espacial (BRASIL, 2007).

O desenvolvimento acelerado da tecnologia possibilitou traduzir a estatística para uma linguagem mais acessível, com a criação de softwares capazes de viabilizar a criação, a manipulação de grandes volumes de dados e a extração de informações de bases de dados geográficos. O resultado tem sido o desenvolvimento de Sistemas



de Informações Geográficas (SIGs). Os SIGs se destacam como ferramentas de geoprocessamento, principalmente em análises que envolvem fatores ambientais e epidemiológicos, pois permitem interpretar a distribuição dos eventos (LONGLEY et al., 2013; ROJAS; BARCELLOS; PEITER, 1999).

O primeiro modelo de SIG foi desenvolvido no Canadá. O governo canadense tinha interesse em quantificar as riquezas naturais da nação, e isso foi facilitado pelo Sistema de Informação Geográfica do Canadá (*Canada Geographic Information System - CGIS*). Já na década de 1970, o governo norte americano, através da Universidade de Harvard, aprimorou o *software CGIS* com o objetivo de realizar o censo demográfico da década, resultando no SIG *Odyssey* (LONGLEY et al., 2013).

A história moderna dos SIGs tem seu início a partir da década de 1980 quando o preço dos computadores suficientemente potentes caiu. Hoje diversos softwares estão disponíveis para facilitar as tarefas que envolvem a análise de dados espaciais (LONGLEY et al., 2013).

## 2.2.2 Tipos de análise espacial

Na análise espacial de dados, são considerados principalmente três tipos de dados georreferenciados. Sua diferenciação está relacionada com sua natureza estocástica. Desse modo, diferentes metodologias estatísticas são empregadas em cada um.

Para este trabalho, apenas a análise de dados de área será discutida de forma mais detalhada. Porém, na sequência será apresentado uma breve introdução a respeito dos três tipos de dados georreferenciados.

### 2.2.2.1 Eventos ou padrões pontuais

São fenômenos expressos como ocorrências identificadas como pontos localizados no espaço, denominados processos pontuais ou eventos. Ou seja, o objeto de interesse é a própria localização espacial do fenômeno. O objetivo primário nesse tipo de análise é determinar se os eventos ou pontos observados exibem algum padrão sistemático, em contraste a uma distribuição aleatória. Assim, o interesse principal consiste nas coordenadas geográficas que associa a localização exata dos acontecimentos como por exemplo, localização de crimes, ocorrência de doenças e

espécies vegetais (DRUCK, et al., 2004; MELO; CARRENO; PADILHA-FELTRIN, 2012).

Na prática, o foco é detectar a existência de padrões de aglomerados espaciais (*cluster*), através da constatação de um número acima do esperado de casos excessivamente próximos. Isso indicará a existência de uma distribuição espacial diferente da aleatória, merecendo ser objeto de maior análise (DRUCK, et al., 2004; MELO; CARRENO; PADILHA-FELTRIN, 2012).

#### 2.2.2.2 Geoestatística

Tendo sua metodologia utilizada na Agricultura de Precisão, Lavra e Prospecção Mineira, Climatologia, Engenharia Florestal, Epidemiologia, Geologia do Petróleo, entre outros, a Geoestatística ou Teoria das Variáveis Regionalizadas teve seu desenvolvimento a partir dos trabalhos investigativos de Daniel G. Krige (KRIGE, 1951), engenheiro de minas que trabalhou nas minas de ouro do Rand, na África do Sul (YAMAMOTO; LANDIM, 2013).

A Teoria das Variáveis Regionalizadas tem dois objetivos: teoricamente descrever a correlação espacial; na prática, minimizar o erro da estimação de dados em locais não amostrados (YAMAMOTO; LANDIM, 2013).

Esses métodos têm o objetivo de fornecer um conjunto de técnicas necessárias para entender a aparente aleatoriedade dos dados, os quais, apresentam uma possível estruturação espacial, estabelecendo, assim, uma função de correlação espacial (YAMAMOTO; LANDIM, 2013).

Esse tipo de dado é geralmente resultante de recursos naturais, tendo seus resultados usualmente expressos por mapas geológicos, topográficos, ecológicos, fitogeográficos e pedológicos (DRUCK et al., 2004).

#### 2.2.2.3 Estatística espacial de áreas

Neste tipo de análise, se dispõe apenas de um valor por área, que pode representar talhões, bairros, municípios, estados entre outros. O objetivo desse tipo de análise é identificar autocorrelações espaciais em nível global e local, e a forma de apresentação desses resultados é feita utilizando mapas temáticos com o padrão espacial do fenômeno (DALPOSSO, 2010; DRUCK et al., 2004).

Assim, a estatística espacial de áreas fornece metodologias que possibilitam elaborar análises robustas de dados, permitindo realizar diversas leituras do problema e subsidiar o seu gerenciamento no espaço e no tempo.

Segundo (BRASIL, 2007; LIMA-COSTA; BARRETO, 2003) as principais aplicações da estatística espacial de áreas na saúde são:

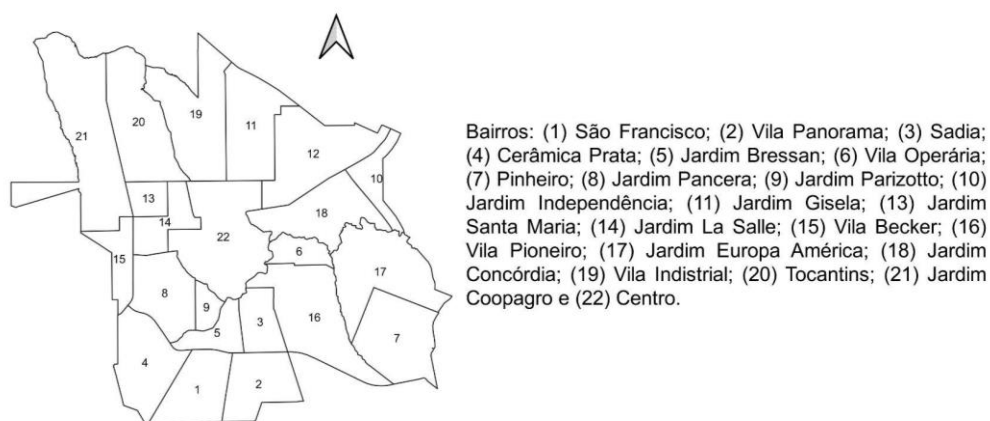
1. Mapeamento de doenças - consiste na descrição do processo de distribuição espacial com objetivo de avaliar a variação geográfica na sua ocorrência, identificando locais de risco e orientando a alocação de recursos;
2. Estudos ecológicos - visa estudar a relação entre incidência de doenças e exposição entre grupos de indivíduos, para verificar a possível existência de associação entre elas;
3. Monitoramento de problemas ambientais - estima a distribuição espacial de fatores ambientais relevantes para a saúde como monitoramento de potenciais fontes ambientais de risco à saúde.

### 3 MATERIAL E MÉTODOS

#### 3.1 Material

A área de estudo do trabalho abrange os 22 bairros do município de Toledo - Paraná. A análise do conjunto de dados referentes aos números de casos reagentes positivos para o vírus SARS-CoV-2, será efetuada com o propósito de ilustrar a aplicação das técnicas da estatística espacial de áreas. Para a criação do banco de dados espaciais, será utilizado um *shapefile* dos bairros do município representados na Figura 1.

**Figura 1 - Mapa dos 22 bairros do município de Toledo-PR.**



**Fonte: Autoria própria com software QGIS**

Os dados foram fornecidos pela Secretaria de Saúde do município de Toledo - Paraná do ano de 2020. Tendo em vista que o trabalho aborda apenas os dados agrupados por áreas, não se obteve nenhuma identificação pessoal, o que, de acordo com a Resolução nº 510/2016, do Conselho Nacional de Saúde, isenta a necessidade prévia de aprovação pelos órgãos competentes em pesquisas com seres humanos (BRASIL, 2016).

Desse modo, primeiramente é feita uma análise exploratória espacial com mapas (legenda feita por amplitude ou quartis) para depois investigar se existe autocorrelação espacial.

Para o cálculo da taxa de crescimento do número de casos será utilizada a Equação 1 (RAMOS; SANCHEZ, 2000).

$$T = \frac{(V_p - V_a)}{V_a} \cdot 100 \quad (1)$$

em que:

$T$  : taxa de crescimento percentual

$V_p$ : valor presente

$V_a$ : valor anterior

## 3.2 Métodos

### 3.2.1 Análise exploratória

A análise exploratória possibilita um estudo detalhado dos dados. Nessa abordagem, a finalidade é obter dos dados a maior quantidade possível de informações (MEDRI, 2011). Assim, calcula-se mínimo, máximo, quartis, assimetria, desvio padrão, média e mediana.

A forma mais simples e intuitiva de análise exploratória é a visualização de valores extremos no mapa. Vale ressaltar que, o uso de diferentes pontos de corte da variável induz a visualização de diferentes aspectos. Isso implica na busca de uma descrição dos eventos apresentados, tentando gerar hipóteses ou mesmo modelos mais apropriados. Nesse caso usa-se basicamente três métodos de corte de variável: intervalos iguais, percentis e desvios-padrão (BRASIL, 2007; DALPOSSO, 2010; DRUCK et al., 2004).

No caso de intervalos iguais, em que os valores máximo e mínimo são divididos pelo número de classes de interesse, se ocorre uma distribuição muito concentrada de um lado para a variável, este corte deixa apenas um número muito pequeno de áreas nas classes da cauda mais alongada da distribuição, como resultado, a maior parte das áreas é alocada a uma ou duas cores. O uso de percentis para definição de classes destina a alocação dos polígonos em quantidades iguais pelas cores, porém isso pode mascarar diferenças significativas em valores extremos e dificultar a identificação de áreas críticas. Já o uso de desvios-padrão, no qual a distribuição da variável é apresentada em gradações de cores diferentes para valores acima e abaixo da média, faz a suposição da normalidade da distribuição da variável (DRUCK et al., 2004).

Vale destacar que, na análise exploratória é importante experimentar diferentes tipos de corte da variável na visualização dos mapas. Esses procedimentos são adaptações de técnicas estatísticas consagradas, como histograma, quartis,

desvio-padrão e pontos discrepantes, porém, agora esses recursos são representados em mapas temáticos, indicando as localizações (DALPOSSO, 2010; DRUCK et al., 2004).

### 3.2.2 Análise exploratória espacial

A análise exploratória espacial permite descrever e visualizar as distribuições espaciais tanto em nível global quanto em nível local, descobrir padrões de associação espacial (*clusters*), sugerir instabilidades espaciais (não estacionariedade) e identificar situações atípicas (*outliers*) (NUNES, 2013; SEFFRIN, 2017).

Esses procedimentos são um conjunto de métodos genéricos de análise exploratória e a visualização dos dados é por intermédio de mapas temáticos com o padrão espacial do fenômeno. Desta forma, essas técnicas permitem descrever a distribuição das variáveis em estudo, identificar distribuições atípicas não só em relação ao tipo de distribuição, mas também em relação aos vizinhos, investigando a existência de padrões na distribuição espacial do fenômeno (DRUCK et al., 2004; NUNES, 2013).

Assim, a análise espacial é composta por um conjunto de mecanismos que busca identificar a estrutura de correlação espacial que melhor descreva os dados, tendo como parâmetro, a estimativa da magnitude da autocorrelação espacial entre as áreas (NUNES, 2013).

Deste modo, geralmente o processo de modelagem é precedido de uma fase de análise exploratória associada à apresentação visual dos dados em forma de gráficos e/ou de mapas temáticos com o padrão espacial do fenômeno, buscando então, a identificação de padrões de autocorrelação espacial. O modelo inferencial de autocorrelação espacial objetiva quantificar a dependência espacial entre os valores das amostras (DRUCK et al., 2004; NUNES, 2013).

Dessa forma, serão elaborados mapas das taxas de crescimento considerando intervalos de mesma amplitude e quintis.

### 3.2.3 Matriz de proximidade espacial

Encontrada na literatura como matriz de vizinhança, matriz de conectividade ou matriz de proximidade espacial, ela indica a relação espacial de cada área com as demais. Ou seja, a matriz de proximidade espacial é uma ferramenta bastante útil para

descrever o arranjo espacial dos objetos (BRASIL, 2007; DALPOSSO, 2010; NEVES et al., 2000).

Dado um conjunto de  $n$  áreas  $\{A_1, A_2, \dots, A_n\}$ , construímos a matriz  $W$ , quadrada,  $n \times n$  em que cada um dos elementos  $W_{ij}$  representa uma medida de proximidade entre a área  $A_i$  e a área  $A_j$ , sendo  $n$  o número total de áreas. Essa medida de proximidade pode ser calculada a partir de um dos seguintes critérios:

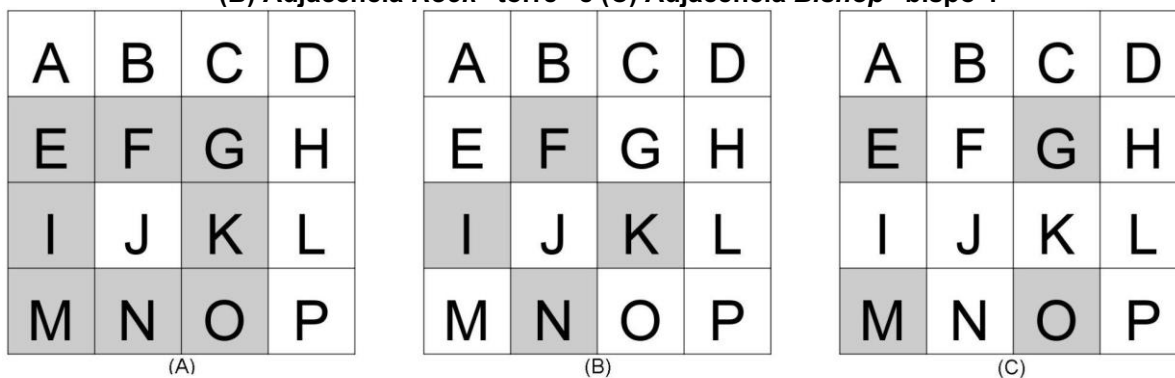
i)  $W_{ij} = 1$ , se o centróide de  $A_i$  está a uma determinada distância de  $A_j$ ; caso contrário  $W_{ij} = 0$ ; para  $i \neq j = 1, \dots, n$ .

ii)  $W_{ij} = 1$ , se  $A_i$  compartilha um lado comum com  $A_j$ ; caso contrário  $W_{ij} = 0$ ; para  $i \neq j = 1, \dots, n$ .

iii)  $W_{ij} = \frac{l_{ij}}{l_i}$ , onde  $l_{ij}$  é o comprimento da fronteira entre  $A_i$  e  $A_j$ ; e  $l_i$  é o perímetro de  $A_i$ ; para  $i \neq j = 1, \dots, n$ .

Ao utilizar matrizes de proximidade espacial que considerem a adjacência, é importante convencionar as formas de vizinhança. Para isso, os critérios são baseados em movimentos de peças do jogo de xadrez, como rainha, torre e bispo.

**Figura 2 - Representação dos tipos de adjacência entre áreas. (A) Adjacência Queen “rainha”, (B) Adjacência Rock “torre” e (C) Adjacência Bishop “bispo”.**



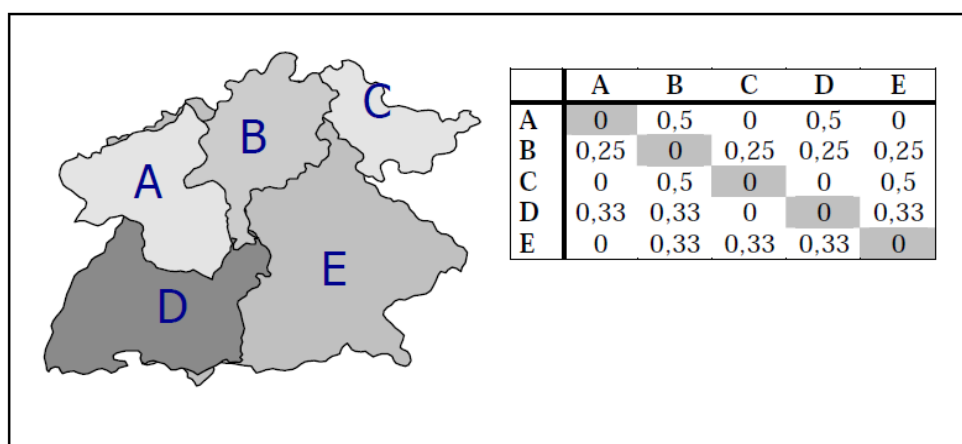
Fonte: Autoria própria

A Figura 2 exemplifica os critérios de vizinhança para identificar os vizinhos da área J. Na Figura 2 (A) o critério adotado é o rainha “Queen”, ou seja, todas as áreas que tem interseção não nula com a área J serão vizinhas da área J. Na Figura 2 (B) é apresentado o critério torre “Rock”, em que os vizinhos serão apenas aqueles que tiverem um lado em comum. Na Figura 2 (C) o critério adotado é o bispo “Bishop”, ou seja, apenas as áreas que se localizam nas diagonais são considerados vizinhos.

O conceito de matriz de proximidade espacial pode ser generalizada para vizinhos dos vizinhos, e isso denota a ordem da matriz. Por exemplo, considerando o critério de vizinhança *Rock* e uma matriz de ordem 2, os vizinhos B e L também seriam considerados vizinhos de J.

Como a matriz de proximidade é utilizada em cálculos de indicadores na fase de análise exploratória, é muito útil normalizar suas linhas, para que a soma dos pesos de cada linha seja igual a 1.

**Figura 3 - Matriz de proximidade espacial, normalizada pelas linhas.**



Fonte: DRUCK et al., 2004

A Figura 3 mostra um exemplo de matriz de proximidade espacial em que os valores dos elementos da matriz refletem o critério de adjacência Queen e foram normalizados.

Assim, para este trabalho será considerada a matriz *Queen* de primeira ordem, em que são considerados vizinhos todos os polígonos possuem interseção não nula com o polígono desejado.

### 3.2.4 Índice de autocorrelação global I de Moran

O índice de Moran fornece uma medida geral da associação espacial existente no conjunto dos dados, ou seja, é uma medida cujo resultado indica se os dados se distribuem ou não de forma aleatória no espaço. Seu valor varia de -1 a 1. Caso esta autocorrelação espacial seja positiva, então tem-se que, se uma determinada região apresentar valores elevados ou (baixos) para uma variável qualquer, então seu entorno deve apresentar valores elevados ou (baixos), o que



caracteriza uma espécie de *cluster* no espaço. Caso essa estatística seja negativa, então pode-se esperar que se uma região apresentar valores elevados ou (baixos) para uma variável, deverá estar cercada por regiões cujos valores para esta mesma variável devem ser baixos ou (elevados) (DALPOSSO, 2010; NEVES et al, 2000).

Segundo DALPOSSO (2010, apud ANSELIN 2005) o índice I de Moran é dado por :

$$I = \frac{Z^t \cdot W_z}{Z^t \cdot Z} \quad (2)$$

em que:

Z: vetor dos desvios

$Z^t$ : vetor dos desvios transposto

$W_z$ : vetor das médias ponderadas

### 3.2.5 Gráfico de espalhamento de Moran

O gráfico de espalhamento de Moran é uma das formas de interpretar o índice I de Moran, sendo então uma forma adicional de visualizar a dependência espacial.

Construído com base nos valores normalizados (valores de atributos subtraídos de sua média e dividido pelo desvio padrão). A ideia é comparar os valores normalizados do atributo numa área com a média dos seus vizinhos, construindo um gráfico bidimensional de Z (valores normalizados) por  $W_z$  (média dos vizinhos, também normalizadas), com o objetivo de identificar pontos com associação espacial positiva ou não (DRUCK et al., 2004).

O gráfico de espalhamento de Moran apresentado na Figura 4 é dividido em quadrantes. Esses quadrantes correspondem a quatro padrões de associação espacial entre as regiões e seus vizinhos (DALPOSSO, 2010).

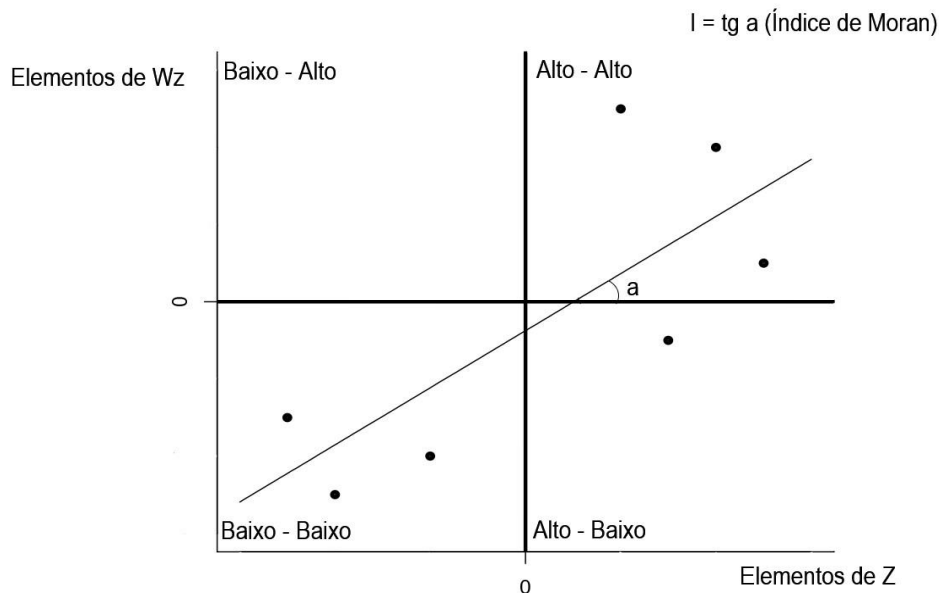
- O primeiro quadrante alto-alto (AA) mostra as regiões que apresentam altos valores para a variável em análise cercadas por regiões que apresentam valores semelhantes.
- O segundo quadrante baixo-baixo (BB) mostra as regiões com valores baixos cercados por vizinhos que apresentam baixos valores.

As regiões que estão localizadas no primeiro e terceiro quadrante apresentam autocorrelação espacial positiva, ou seja, essas regiões formam *clusters* de valores similares.

- O terceiro quadrante alto-baixo (AB) é formado pelas regiões com altos valores para as variáveis em análise cercadas por regiões com baixos valores.
- O quarto quadrante baixo-alto (BA) mostra as regiões com valores baixos cercados por vizinhos que apresentam valores altos.

As regiões apresentadas no segundo e quarto quadrante podem ser vistas como regiões que não seguem o mesmo processo de dependência espacial das demais observações, ou seja, apresentam autocorrelação espacial negativa, indicando regiões em que seu comportamento é diferente de seus vizinhos. Notoriamente, o gráfico reflete a estrutura espacial nas duas escalas de análise: vizinhança e tendência (DALPOSSO, 2010; DRUCK et al., 2004).

**Figura 4 - Construção do gráfico de espalhamento de Moran.**



**Fonte: Autoria própria**

O índice  $I$  de Moran é equivalente ao coeficiente de regressão linear que indica a inclinação da reta de regressão ( $a$ ) de  $W_z$  em  $Z$ . O gráfico de espalhamento de Moran também pode ser apresentado na forma de um mapa temático bidimensional, em que

cada polígono é apresentado segundo seu quadrante no gráfico de espalhamento de Moran. Este mapa é conhecido como *Box Map* e sua legenda é constituída segundo a Figura 5 (DRUCK et al., 2004).

**Figura 5 - Legenda do mapa do gráfico de espalhamento de Moran.**

	Alto - Alto	AA
	Baixo - Baixo	BB
	Baixo - Alto	BA
	Alto - Baixo	AB

**Fonte: Autoria própria**

### 3.2.6 Índice de Autocorrelação Local - LISA

O índice de autocorrelação global I de Moran fornece um único valor como medida de associação espacial para todo o conjunto de dados, o que é útil na caracterização da região de estudo como um todo. Porém, um único valor de autocorrelação espacial pode “mascarar” situações locais. Assim, outra abordagem para a análise da estrutura espacial é a função LISA (*Local Indicator of Spatial Association*), ou Índice de Autocorrelação Local. Quando usados em conjunto com o Índice de Autocorrelação Global I de Moran, eles refinam nosso conhecimento sobre os processos que dão origem à dependência espacial, pois nos permitem encontrar “bolsões” de dependência espacial que não são evidenciados pelos índices globais (BRASIL, 2007; KAMPEL; CÂMARA; QUINTANILHA, 2000).

Neste sentido, os indicadores locais produzem um valor específico para cada objeto, permitindo assim, a identificação de agrupamentos de objetos com valores de atributos semelhantes (*clusters*) e objetos anômalos (*outliers*) (DALPOSSO, 2010).

Segundo (DALPOSSO, 2010 apud ANSELIN 1995), um LISA deve atender dois objetivos:

- I. permitir a identificação de padrões de associação espacial significativos;
- II. ser uma decomposição do índice global de associação espacial.

O LISA utilizado neste trabalho é o índice local de Moran. Uma das formas de representação deste índice, em função dos elementos básicos, é de acordo com a Equação 3:

$$I_i = \frac{(z_i \cdot Wz_i)}{\sigma^2} \quad (3)$$

em que:

$I_i$ : índice local para o objeto  $i$ ;  $i = 1, \dots, n$ ;

$z_i$ : valor do desvio do objeto  $i$ ;  $i = 1, \dots, n$ ;

$Wz_i$ : valor médio dos desvios dos objetos vizinhos de  $i$ ;

$\sigma^2$ : variância da distribuição dos valores dos desvios.

### 3.2.7 Mapa de significância LISA

A significância estatística do uso do índice de Moran local é computada de forma similar ao caso do índice global. Para cada área, calcula-se o índice local, e depois permuta-se aleatoriamente o valor das demais áreas, até obter uma pseudo-distribuição para a qual possamos computar os parâmetros de significância. Uma vez determinada a significância estatística do índice local de Moran, é útil gerar um mapa indicando as regiões que apresentam correlação local significativamente diferente do resto dos dados. Essas regiões podem ser vistas como “bolsões” de não-estacionariedade, pois são áreas com dinâmica espacial própria e que merecem análise detalhada (DRUCK et al., 2004).

### 3.2.8 Mapa de *cluster* LISA

Ao combinar as informações do gráfico de espalhamento de Moran e a informação do mapa de significância das medidas de associação local, obtêm-se o mapa de *cluster* LISA. Desta forma, as regiões detectadas como significativas no mapa de significância LISA são classificadas conforme sua localização no gráfico de espalhamento de Moran. Assim, este será dividido em *clusters* do tipo alto-alto (AA), alto-baixo (AB), baixo-baixo (BB), baixo-alto (BA) e em regiões em cinza mais claro, que significa que tal região não se encontra em nenhum agrupamento porque o valor da variável para essa região não é diferente, em termos estatísticos, da média de

todas as regiões, ou seja, ela não apresentou significância (DALPOSSO, 2010; MARCELLO et al., 2019 apud ALMEIDA, 2012).

### 3.2.9 Recursos computacionais

Neste trabalho foram utilizados para o geoprocessamento os *softwares* GeoDa e QGIS e para os cálculos o *software* R.

## 4 RESULTADOS E DISCUSSÃO

### 4.1 Análise Exploratória

A Tabela 1 apresenta as estatísticas descritivas da Taxa de contaminação por Covid-19 nos bairros de Toledo, calculada através da Equação (1). Vale destacar que três bairros do município são complexos industriais, sendo Sadia (3), Cerâmica Prata (4) e Jardim Independência (10). Deste modo, o número de casos contabilizados nesses locais foi zero. Consequentemente não apresentando Taxa de crescimento.

**Tabela 1 - Estatísticas descritivas da Taxa de contaminação por Covid-19**

	Mínimo	$Q_1^*$	Mediana	Média	$Q_3^{**}$	Desvio Padrão	Assimetria	Máximo
Taxa 1	0	41,36	53,57	52,07	69,51	34,00	-0,223	103,23
Taxa 2	0	19,89	30,31	40,32	44,86	41,40	2,670	200
Taxa 3	0	21,73	28,99	39,18	38,54	44,79	2,592	200
Taxa 4	0	8,27	12,36	13,55	18,06	9,01	0,570	34,37
Taxa 5	0	4,31	6,42	6,03	8,21	4,36	0,320	16,66
Taxa 6	0	6,75	10,41	11,34	16,60	8,52	0,295	30,16
Taxa 7	0	1,09	2,79	3,79	4,13	5,66	3,414	27,27
Taxa 8	0	9,30	13,28	16,76	24,66	14,04	1,314	61,02
Taxa 9	0	2,04	5,57	6,41	8,93	6,62	1,548	27,27
Taxa 10	0	3,57	6,40	7,92	10,97	6,50	1,075	24,49
Taxa 11	0	4,81	8,58	9,24	11,38	8,58	2,283	40,98
Taxa 12	0	5,06	8,72	10,16	15,51	6,68	0,051	20,33
Taxa 13	0	5,24	10,71	10,71	15,17	7,11	0,154	25,86
Taxa 14	0	5,32	8,14	7,80	10,53	5,52	0,450	22,22
Taxa 15	0	0,19	2,22	2,19	3,52	1,76	0,190	5,96
Taxa 16	0	2,18	3,33	3,58	4,86	2,68	0,966	11,53

Valores dos dados em porcentagem. \*1º Quartil. \*\*3º Quartil.

Taxa 1: 25/06 a 03/07; Taxa 2: 03/07 a 09/07; Taxa 3: 09/07 a 20/07; Taxa 4: 20/07 a 27/07; Taxa 5: 27/07 a 04/08; Taxa 6: 04/08 a 11/08; Taxa 7: 11/08 a 17/08; Taxa 8: 17/08 a 25/08; Taxa 9: 25/08 a 31/08; Taxa 10: 31/08 a 09/09; Taxa 11: 09/09 a 14/09; Taxa 12: 14/09 a 21/09; Taxa 13: 21/09 a 29/09; Taxa 14: 29/09 a 06/10; Taxa 15: 06/10 a 12/10; Taxa 16: 12/10 a 20/10.

Observa-se que no primeiro período (Taxa 1: 25/06 a 03/07) a Taxa máxima de variação foi de 103,23%, em que os casos passaram de 31 para 63 no bairro Centro (22). Já no segundo período (Taxa 2: 03/07 a 09/07) analisado a Taxa máxima de variação foi de 200%, em que os casos passaram de 2 para 6 no bairro Vila Becker (15). A última Taxa máxima de variação expressiva foi do terceiro período, também com 200%, em que os casos passaram de 3 para 9 no bairro Jardim Parizotto (9).

Nota-se também, que no penúltimo período analisado (Taxa 15: 06/10 a 12/10) a Taxa máxima de variação foi de apenas 5,96%, em que os casos passaram de 160 para 165 no bairro Jardim Bressan (5).

Em relação ao primeiro quartil, a partir do terceiro período observa-se que a Taxa de contaminação por Covid-19 não ultrapassa os dois dígitos, com destaque para a Taxa 15 (06/10 a 12/10) com valor de 0,19%.

#### 4.2 Análise Exploratória Espacial

Aplicamos as técnicas descritas neste trabalho ao conjunto de dados utilizados no estudo, como forma de ilustrar a sua utilização prática, segundo os critérios de proximidade *Queen*, em que são considerados vizinhos todas as áreas que possuem intersecção não nula com a área em questão.

A Figura 6 apresenta a evolução temporal da Taxa de contaminação por Covid-19 nos bairros de Toledo. Os intervalos foram definidos agrupando os valores de todas as Taxas para obter o máximo (200) o mínimo (0) e a amplitude total (200) das Taxas. Posteriormente, dividiu-se a amplitude total pelo número de classes desejado (5) para obter o intervalo de classe.

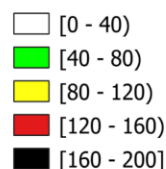
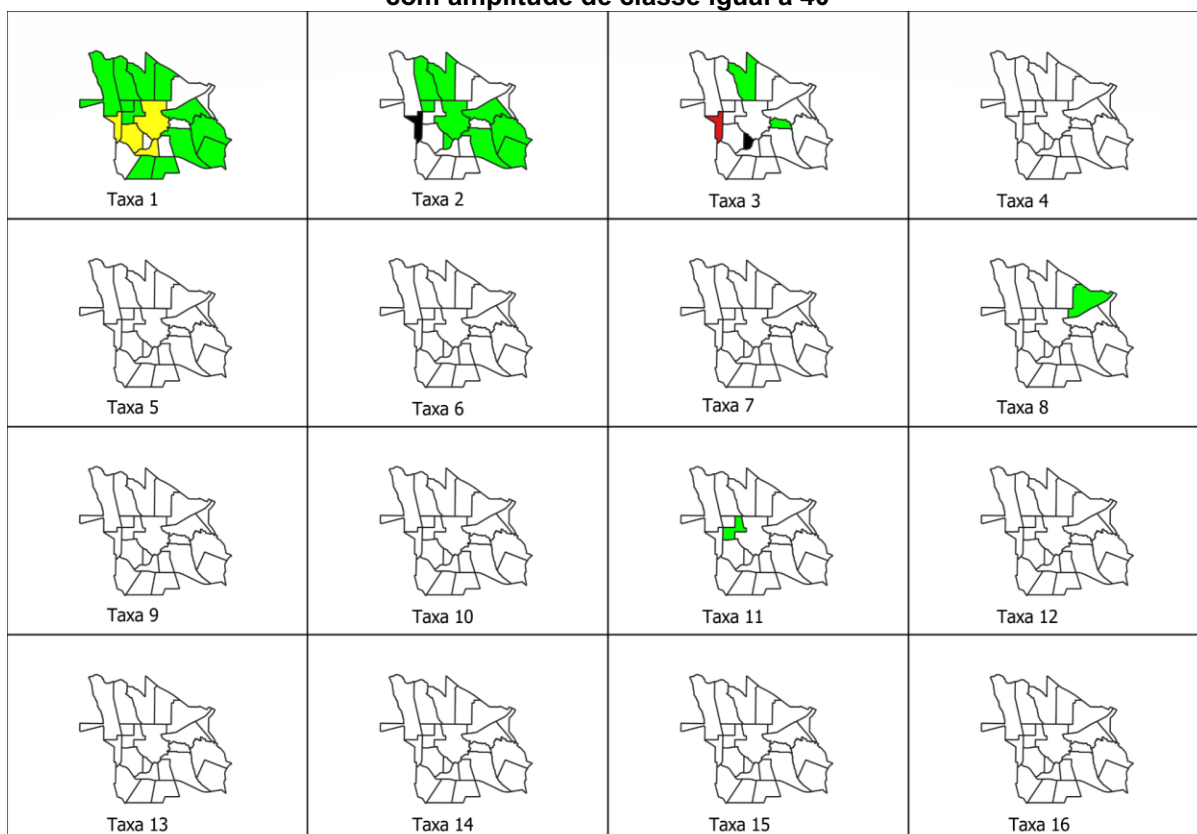
Observa-se que no primeiro período analisado (Taxa 1: 25/06 a 03/07) os bairros Centro (22), Jardim Pancera (8), Jardim Parizotto (9) e Vila Becker (16) apresentaram Taxa de contaminação entre 80% a 120%. Já no segundo período analisado, o bairro Vila Becker (16) apresentou Taxa de contaminação entre 160% a 200%, trocando de patamar no terceiro período, ficando entre 120% a 160%. No terceiro período o destaque fica para o Bairro Jardim Parizotto (9) que apresentou Taxa de contaminação entre 160% a 200%.

Nos três primeiros períodos analisados (Taxa 1 a Taxa 3: 25/06 a 20/07), o bairro Vila Industrial (19) permaneceu com Taxa de contaminação no intervalo entre 40% a 80%.

No oitavo período analisado (Taxa 8: 17/08 a 25/08), o bairro Jardim Porto Alegre (12) apresentou Taxa de contaminação entre 40% a 80%, enquanto que no restante dos bairros a Taxa de contaminação ficou entre 0% a 40%. Igualmente acontece na Taxa 11 (09/09 a 14/09) em que apenas o bairro La Salle (14) apresenta Taxa de contaminação entre 40% a 80%, e o restante do município com Taxa de contaminação entre 0% a 40%.

No entanto, a Figura 6 não apresenta informações robustas para uma análise mais detalhada da situação.

**Figura 6 - Evolução temporal da Taxa de contaminação por Covid-19 nos bairros de Toledo com amplitude de classe igual a 40**



**Fonte: Autoria própria com software QGIS**

Deste modo, para visualizar os agrupamentos espaciais existentes confeccionou-se o mapa de agrupamento por quintis para cada período analisado. Assim, com o uso do *software* R, utilizou-se a função *quantile* para a obtenção dos intervalos, que são apresentados na Figura 7.

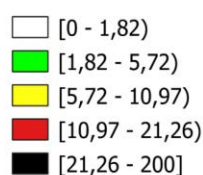
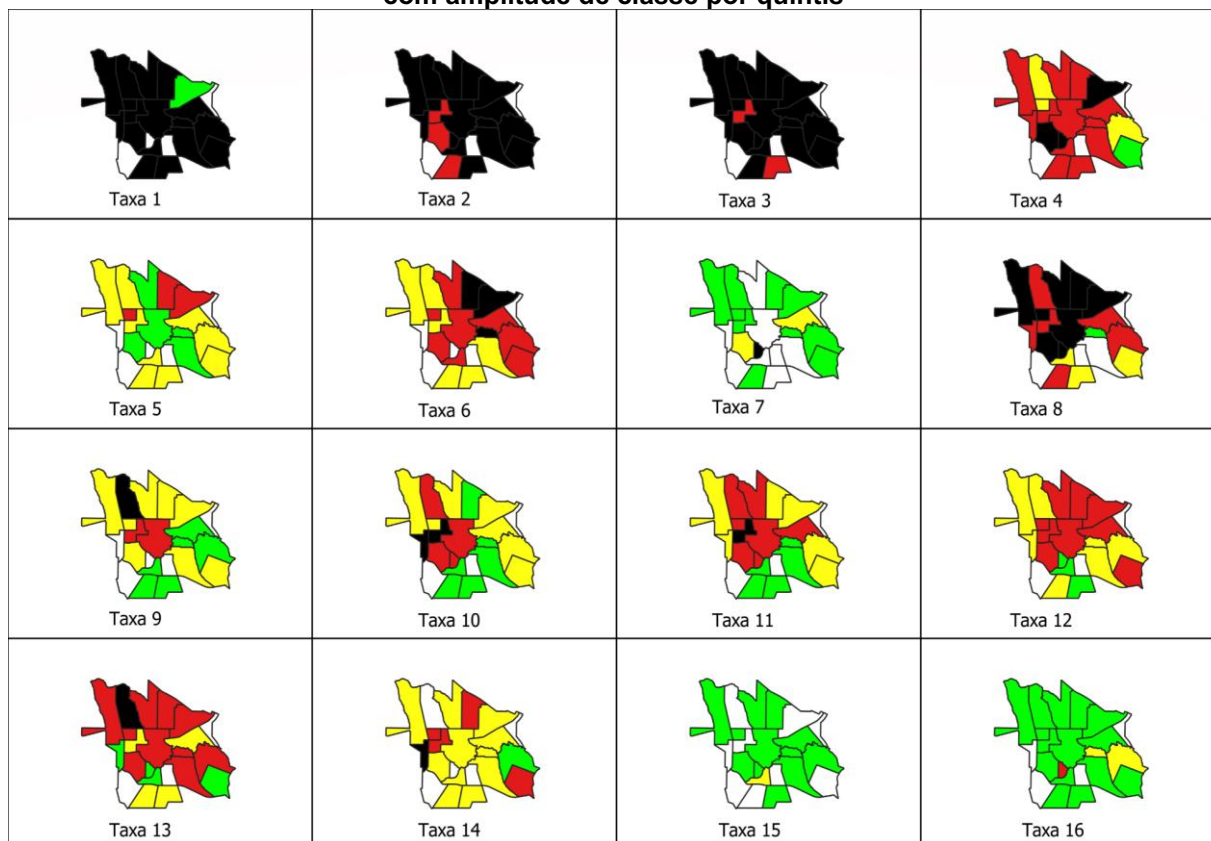
É possível perceber que os três primeiros períodos foram os mais críticos (Taxa 1 a Taxa 3: 25/06 a 20/07), em que praticamente todo o município de Toledo apresentou Taxa de contaminação entre 21,26% a 200%.

Observa-se também que na Taxa 7 (11/08 a 17/08) foi o período em que houve nove bairros com Taxa mínima de contaminação variando de 0% a 1,82%. Porém a situação piorou no período seguinte (17/08 a 25/08), em que os bairros



Centro (22) e Vila Industrial (19) passaram da Taxa mínima de contaminação entre 0% a 1,82% para a Taxa máxima de contaminação entre 21,26% a 200%.

**Figura 7 - Evolução temporal da Taxa de contaminação por Covid-19 nos bairros de Toledo com amplitude de classe por quintis**



**Fonte: Autoria própria com *software* QGIS**

Já os bairros Jardim Coopagro (21), Jardim Gisela (11) e Jardim Porto Alegre (12) passaram de 1,82% a 5,72% para 21,26% a 200%.

Ainda com relação à Taxa 7 (11/08 a 17/08), nota-se que o bairro Jardim Parizotto (9) foi o único que apresentou Taxa máxima de contaminação de 21,26% a 200%, mantendo-se no mesmo patamar no período subsequente.

Pode-se notar também que nas Taxas 11,12 e 13 (09/09 a 29/09) existem aglomerados de bairros que mantêm a Taxa de contaminação constante de um

período para o outro, sendo Centro (22), Jardim Pancera (8) e Vila Industrial (19), em que a Taxa de contaminação permaneceu entre 10,97% a 21,26%.

#### 4.3 Índice I de Moran

A Tabela 2 apresenta os Índices I de Moran da Taxa de contaminação por Covid-19 com seus respectivos níveis descritivos (p-valor), referente à hipótese de independência espacial, pois é importante estabelecer sua validade estatística, ou seja, a partir desses valores pode-se estimar a significância deste índice.

Deste modo, foram geradas 999 permutações dos valores de atributos associados às regiões; cada permutação produz um novo arranjo espacial, em que os valores são redistribuídos entre as áreas.

Assim, como apenas um dos arranjos corresponde à situação observada, pode-se construir uma distribuição empírica de I. Se o valor do índice I medido originalmente corresponder a um “extremo” da distribuição simulada, então trata-se de um valor com significância estatística.

No caso da Tabela 2 o Índice Global de Moran foi calculado para os 16 períodos analisados.

Observa-se que para as Taxas 1, 2 e 3 (25/06 a 20/07) o índice aponta uma autocorrelação espacial inversa, ou seja, são bairros que possuem alta Taxa de contaminação e estão cercados por bairros que possuem baixa Taxa de contaminação e o contrário também ocorre, bairros com Taxa de contaminação baixa cercado por bairros que possuem alta Taxa de contaminação.

No entanto, a Taxa 1 (25/06 a 03/07) apresenta um nível de significância de 0,193, não sendo significativo, o mesmo ocorre com a Taxa 3 (09/07 a 20/07), o que não descarta a possibilidade de autocorrelação espacial, no entanto sendo uma autocorrelação espacial muito fraca.

Já a Taxa 2 (03/07 a 09/07) apresenta um alto índice de autocorrelação espacial inversa em que uma área que apresenta valores baixos para uma determinada variável e está cercada por vizinhos que apresentam valores altos para esta mesma variável, e o contrário também é válido, com um nível de significância de 97,3%, evidenciando a presença de aglomerados de valores semelhantes.

Nota-se também que a Taxa 7 (11/08 a 17/08) apresenta um valor muito próximo de 0 para o índice de Moran, com nível de significância<sup>1</sup> de 0,235 o que indica independência espacial.

**Tabela 2 - Índice I de Moran da Taxa de contaminação por Covid-19**

	<b>Índice I de Moran</b>	<b>p-valor</b>
Taxa 1	-0,173	0,193
Taxa 2	-0,230	0,027
Taxa 3	-0,047	0,417
Taxa 4	0,045	0,251
Taxa 5	0,066	0,217
Taxa 6	0,173	0,066
Taxa 7	0,004	0,235
Taxa 8	0,149	0,077
Taxa 9	0,186	0,055
Taxa 10	0,213	0,027
Taxa 11	0,121	0,061
Taxa 12	0,088	0,166
Taxa 13	0,185	0,057
Taxa 14	-0,032	0,459
Taxa 15	-0,045	0,464
Taxa 16	-0,053	0,478

Diferentemente, a Taxa 10 (31/08 a 09/09) apresenta um índice de 0,213, significativo ao nível de 97,3%, indicando autocorrelação espacial positiva, ou seja, existem bairros que apresentam Taxa alta de contaminação cercados por bairros com alta Taxa de contaminação, e o contrário também ocorre, bairros com baixa Taxa de contaminação cercados por bairros com baixa Taxa de contaminação, esses aglomerados espaciais podem ser vistos na Figura 7.

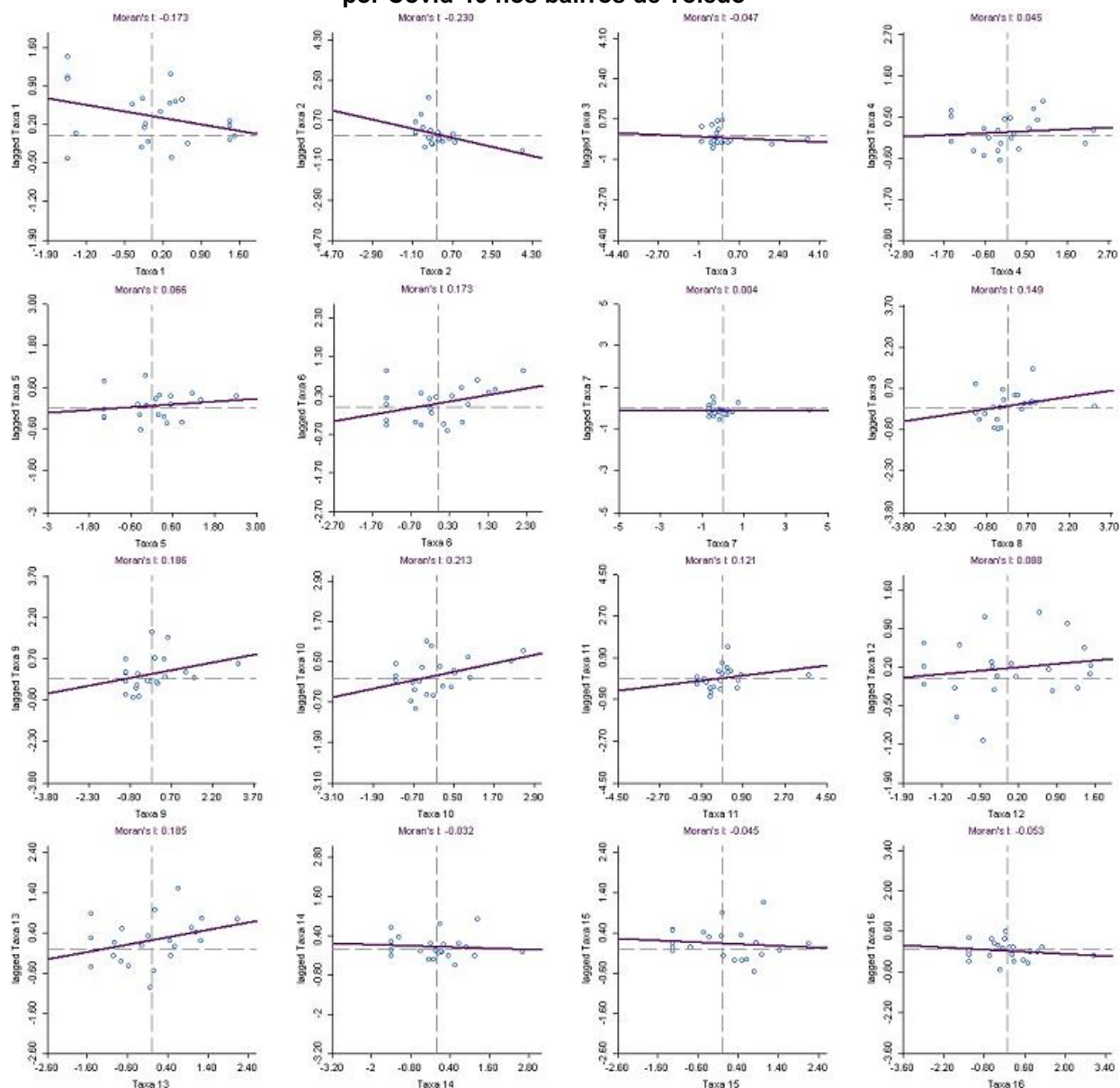
#### 4.4 Moran Scatterplot

A fim de uma melhor representação dos dados apresentados na Tabela 2, o Gráfico de Espalhamento de Moran é uma maneira adicional de visualizar a dependência espacial. Deste modo, compara-se os valores normalizados do atributo numa área com a média dos seus vizinhos. Assim, o eixo da abscissa é referente aos valores normalizados e o eixo da ordenada é referente à média dos vizinhos.

<sup>1</sup> Recomenda-se a leitura de (FERREIRA; PATINO, 2015) para maior entendimento sobre p-valor. Disponível em: <<https://www.scielo.br/j/jbpneu/a/SWk5XsCsXTW7GBZq8n7mVMJ/?lang=en>>

A Figura 8 apresenta os diagramas de espalhamento de Moran (Moran scatterplot) para os 16 períodos analisados (25/06 a 20/10), cada ponto representa um bairro do município de Toledo - PR.

**Figura 8 - Diagrama de espalhamento de Moran (Moran scatterplot) da Taxa de contaminação por Covid-19 nos bairros de Toledo**



Fonte: Autoria própria com software GeoDa

Considerando a Taxa 2 (03/07 a 09/07), nota-se que 14 dos 22 bairros do município (63,6%) estão nos quadrantes BA e AB, o que evidencia a autocorrelação espacial inversa, com nível significativo de 0,027.

Observando o período que compreende as Taxas 2, 3, 4, 5 e 6 (03/07 a 11/08), é evidente o período de transição, em que a maioria dos bairros passa de uma autocorrelação espacial inversa para uma autocorrelação espacial direta em que bairros com alta Taxa de contaminação estão cercados por bairros com alta Taxa de contaminação e o inverso também ocorre, bairros com baixa Taxa de contaminação estão cercados por bairros com baixa Taxa de contaminação. Para a Taxa 6 (04/08 a 11/08) são 13 dos 22 bairros nesta situação (59,09%).

Já na Taxa 7(11/08 a 17/08), em que o índice de Moran é muito próximo de zero, a reta de regressão tende ao eixo da abscissa, indicando independência espacial, ou seja, a distribuição dos dados não aponta uma causalidade, estando assim distribuídos de forma aleatória.

Nota-se que os períodos compreendidos entre as Taxas 8, 9, 10, 11, 12 e 13 (17/08 a 29/09) são períodos que apresentam autocorrelação espacial direta, ou seja, são bairros que apresentam alta Taxa de contaminação cercados por bairros que também apresentam alta Taxa de contaminação, ocorrendo também o contrário, bairros com baixa Taxa de contaminação cercados por bairros que também apresentam baixa Taxa de contaminação.

No entanto, da Taxa 13 (21/09 a 29/09) para a Taxa 14 (29/09 a 06/10) há uma transição de autocorrelação espacial direta na Taxa 13 para autocorrelação espacial inversa na Taxa 14, mantendo seu decréscimo até o final do período de análise.

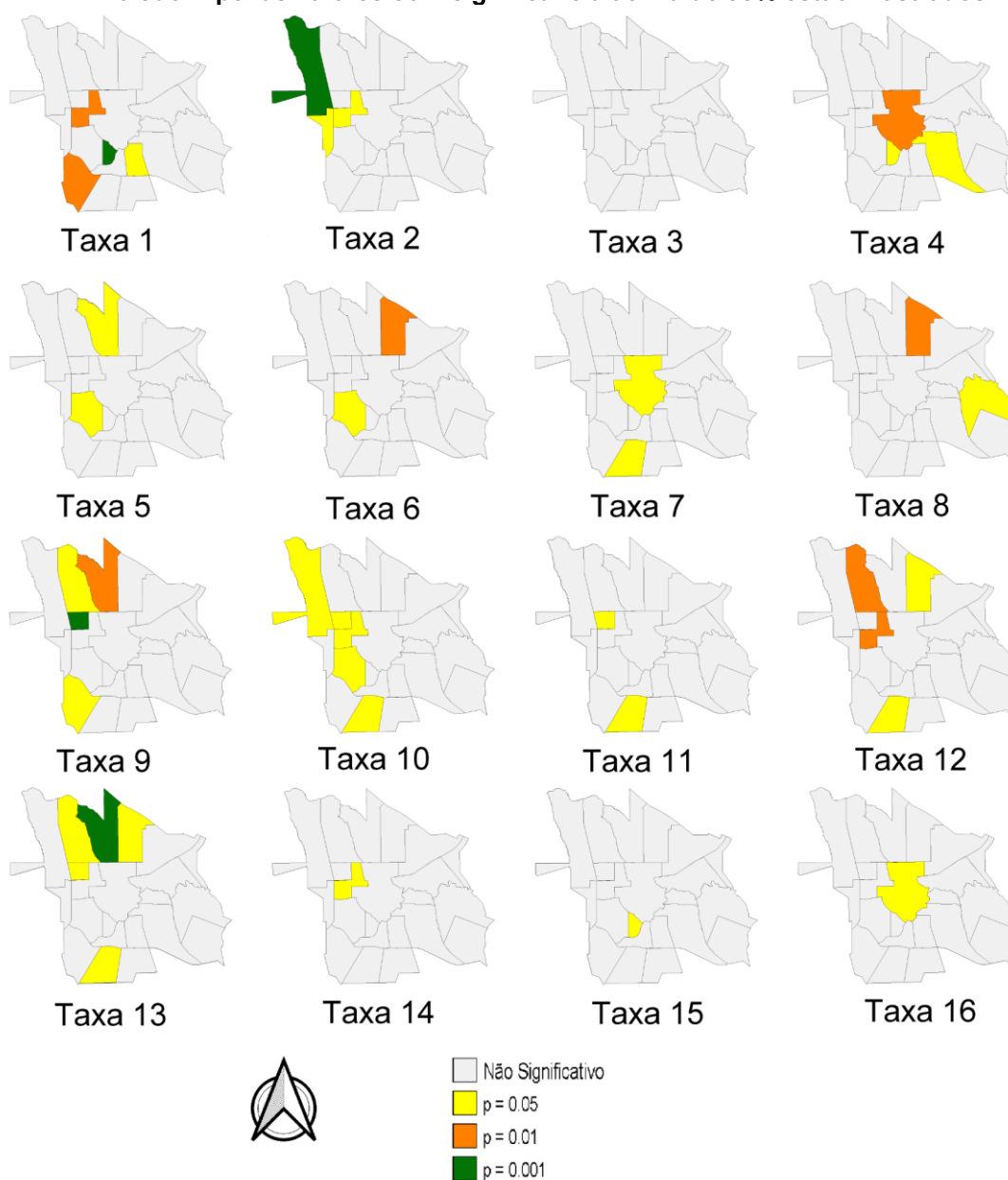
#### **4.5 Índice LISA**

O Índice de autocorrelação espacial global I de Moran fornece um único valor de associação espacial para todo o conjunto de dados. No entanto, é útil utilizar indicadores de associação espacial que possam ser associados às diferentes localizações da variável em estudo. Assim, os indicadores locais produzem um valor específico para cada área, desta forma, permitindo a identificação de agrupamentos (DRUCK et al., 2004).

A significância estatística para este índice é computada de forma similar ao índice global. Para cada área, calcula-se o índice local, e depois permuta-se aleatoriamente o valor das demais áreas, até obter uma pseudo-distribuição para a qual pode-se computar os parâmetros de significância.

Uma vez determinada a significância estatística do índice LISA, gerou-se um mapa (Figura 9) indicando as regiões que apresentam correlação local significativamente diferente do resto dos dados, o que permite a identificação de agrupamentos de objetos com valores de atributos semelhantes (*clusters*) e objetos anômalos (*outliers*).

**Figura 9 - Mapa de significância LISA da Taxa de contaminação por Covid-19 nos bairros de Toledo. Apenas valores com significância acima de 95% estão mostrados**



Fonte: Autoria própria com software GeoDa

A legenda do mapa da Figura 9 é formada por quatro classes, sendo que a primeira indica os bairros em que o índice local não é significativo, a segunda indica os bairros em que o índice local é significativo ao nível de 5% de probabilidade, a terceira indica os bairros em que o índice local é significativo ao nível de 1% de probabilidade, e a última classe indica os bairros em que o índice local é significativo ao nível de 0,1% de probabilidade.

Observando a Figura 9 é possível identificar o comportamento da Taxa de contaminação por Covid-19 nos bairros do município de Toledo. Na Taxa 3 (09/07 a 20/07) é notório que não há correlação espacial significativa entre os bairros. Na Taxa 9 é possível identificar um agrupamento de bairros na região norte. Este agrupamento indica os bairros com correlação espacial significativa. Eles são compostos pelos bairros Tocantins (20), Vila Industrial (19) e Jardim Santa Maria (13), sendo que o bairro Tocantins (20) apresenta significância de 95% o bairro Vila Industrial (19) apresenta nível de significância de 99% e o Jardim Santa Maria (13) apresenta nível de significância de 99,9%.

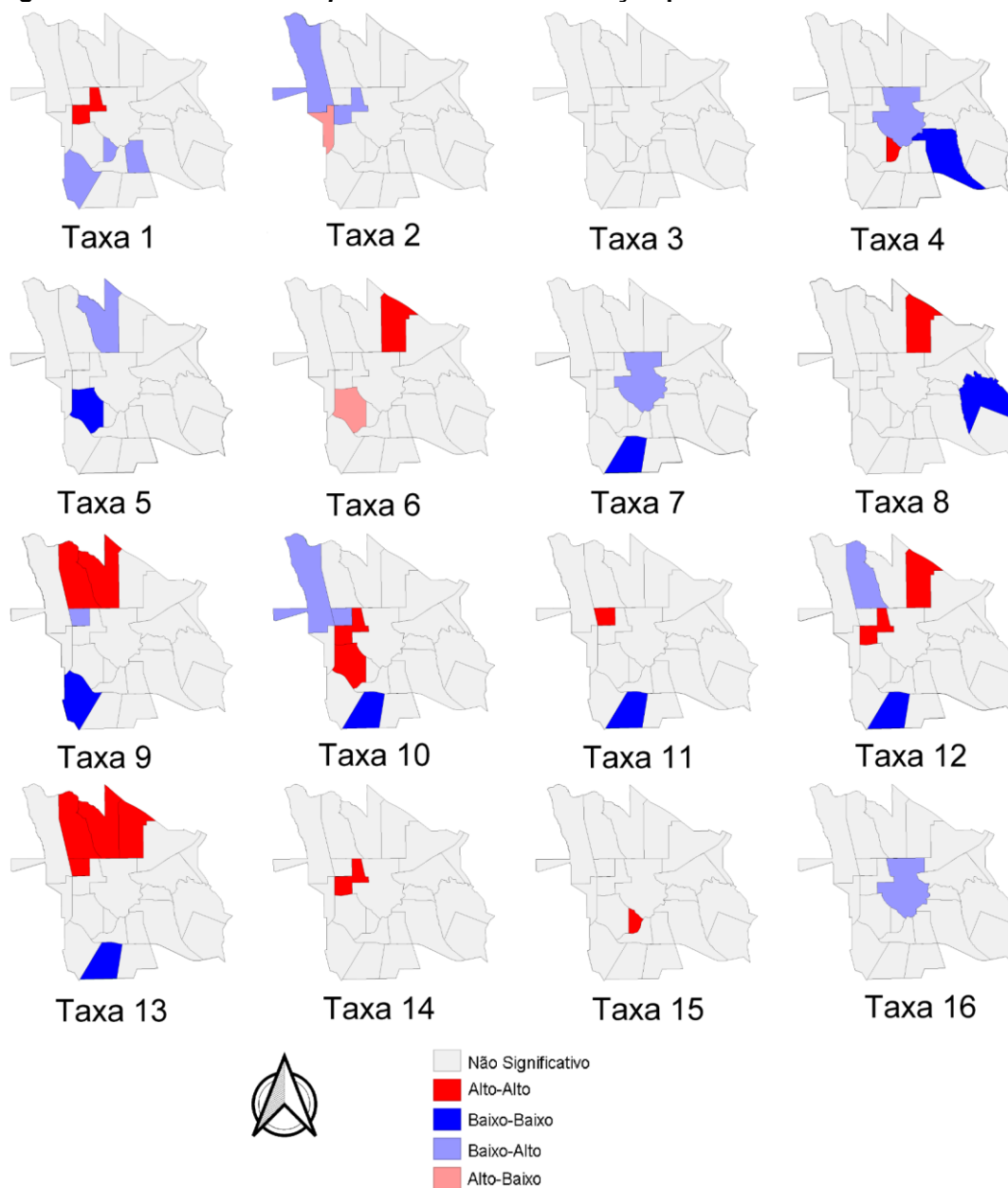
Já na Taxa 10 (31/08 a 09/09) destaca-se a presença de um agrupamento de bairros vizinhos, compostos pelos bairros Jardim Coopagro (21), Jardim Santa Maria (13), Jardim La Salle (14) e Jardim Pancera (8), todos com nível de significância de 95%. No entanto, observa-se que este agrupamento não se mantém no período seguinte, ficando apenas o bairro Santa Maria (13) na Taxa 11 (09/09 a 14/09) com correlação espacial significativa de 95%.

Destaca-se também o bairro São Francisco (1) que nas Taxas 10, 11, 12 e 13 (31/08 a 29/09) permaneceu com correlação espacial significativa de 95%.

Na Taxa 13 (21/09 a 29/09) é possível notar outro agrupamento de bairros no norte do município, sendo Jardim Santa Maria (13), Tocantins (20), Vila Industrial (19) e Jardim Gisela (11), sendo que o bairro Vila Industrial (19) apresenta significância de 99,9% e os demais bairros citados apresentam significância de 95% juntamente com outro bairro no sul do município o bairro São Francisco (1).

A partir dos dos bairros identificados como significantes gerou-se o mapa conhecido como LISA *cluster map* (Figura 10) que permite identificar qual tipo de correlação existe entre eles. O LISA *cluster map* é classificado conforme a legenda do gráfico de espalhamento de Moran (DRUCK et al., 2004).

**Figura 10 - LISA cluster map da Taxa de contaminação por Covid-19 nos bairros de Toledo**



**Fonte: Autoria própria com software Geoda**

Observa-se que na Taxa 3 (09/07 a 20/07) foi um período que não apresentou correlação espacial significativa.

Na Taxa 5 (27/07 a 04/08) o bairro Jardim Parizotto (8) descreve correlação espacial direta significativa, pois apresenta baixa Taxa de contaminação e está cercado por bairros que também apresentam baixa Taxa de contaminação. No entanto, no período seguinte, Taxa 6 (04/08 a 11/08), o mesmo bairro apresentou correlação espacial inversa significativa, pois passou de baixa Taxa de contaminação



para alta Taxa de contaminação. Porém, seus vizinhos continuaram apresentando baixa Taxa de contaminação.

Assim, este bairro merece mais atenção das autoridades sanitárias, e o contrário também é válido, em que os bairros vizinhos a este recebendo mais atenção das autoridades sanitárias evite que isso se espalhe na vizinhança, passando de baixa Taxa de contaminação para alta Taxa de contaminação.

Na Taxa 9 (25/08 a 31/08) verifica-se que o aglomerado de bairros que apresentam correlação espacial significativa é formado por dois bairros que possuem alta Taxa de contaminação e estão cercados por bairros que também possuem alta Taxa de contaminação. O primeiro é o bairro Tocantins (20) e Vila Industrial (19).

Na Taxa 10 (31/08 a 09/09) nota-se uma área de transição entre a região norte e a região centro-oeste, em que os bairros do norte Jardim Coopagro (21) e Jardim Santa Maria (13) são classificados como Baixo-Alto (BA) e os bairros centro-oeste Jardim La Salle (14) e Jardim Pancera (8) são classificados como Alto-Alto (AA). Deste modo, os bairros nortenhos merecem mais atenção das autoridades sanitárias para evitar que toda a região se transforme em um grande aglomerado Alto-Alto (AA).

Observando as Taxas 10, 11, 12 e 13 (31/08 a 29/09) verifica-se que o bairro São Francisco (1) apresentou correlação espacial inversa significativa, pois apresenta baixa Taxa de contaminação em todos os períodos, sendo acompanhado por seus vizinhos.

Na Taxa 13 (21/09 a 29/09) nota-se um aglomerado de bairros na região norte do município que apresentam alta Taxa de contaminação e são acompanhados por seus vizinhos que também apresentam alta Taxa de contaminação, sendo o bairro Tocantins (20), Jardim Santa Maria (13), Vila Industrial (19) e Jardim Gisela (11).

## 5 CONCLUSÕES

Os resultados mostram que a estatística espacial de áreas é uma boa alternativa para estudar a taxa de contaminação por Covid-19, pois oferece de forma clara e precisa informações coerentes com a realidade, permitindo identificar regiões com características semelhantes.

Assim, por exemplo, ao identificar bairros com correlação alto-baixo (AB) é possível orientar medidas aos residentes para evitar, digamos, um progresso para o *status* alto-alto (AA) com seus vizinhos.

Isso serve de informação robusta para as autoridades sanitárias do município, auxiliando na tomada de decisões e na alocação de recursos humanos e financeiros, permitindo que as autoridades possam tomar providências localizadas. Minimizando os impactos causados pela doença.

## REFERÊNCIAS

ANDERSEN, K. G. et al. The proximal origin of SARS-CoV-2. **Nature Medicine**, v.26, n.4, p. 450-452, 2020.

BAI, Y. et al. Presumed asymptomatic carrier transmission of COVID-19. **JAMA**, v.323, n.14, p. 1406-1407, 2020.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. **Introdução à Estatística Espacial Para a Saúde Pública**. v. 3. Brasília: Ministério da Saúde, 2007.

BRASIL. **Resolução nº 510, de 07 de abril de 2016**. Diário Oficial da União, Brasília, DF, 07 abr. 2016. Seção 1, p. 44-46. Disponível em: <<http://conselho.saude.gov.br/resolucoes/2016/Reso510.pd>>. Acesso em: 17 mar. 2021.

Coronavírus: OMS declara pandemia. **BBC News Brasil**, São Paulo, 11 mar. 2020. Notícias. Disponível em: <<https://www.bbc.com/portuguese/geral-51842518>>. Acesso em: 04 abr. 2021.

DALPOSSO, G. H. **Estatística Espacial Aplicada à Agricultura de Precisão**. 2010. 66 f. Dissertação (Mestrado em Engenharia Agrícola) - Universidade Estadual do Oeste do Paraná, Cascavel, 2010.

DRUCK, S. et al. (Org). **Análise espacial de dados geográficos**. Planaltina, DF: EMBRAPA Cerrados, 2004.

FLORES, E. F. (Org.). **Virologia Veterinária**. Santa Maria: Ed. UFSM, 2007.

HANSEN, R. **Primeiro caso de coronavírus é confirmado em Toledo**. Prefeitura de Toledo, Toledo, 07 abr. 2020. Painel. Disponível em: <<https://www.toledo.pr.gov.br/noticia/primeiro-caso-de-coronavirus-e-confirmado-em-toledo>>. Acesso em: 04 abr. 2021.

Instituto Brasileiro de Geografia e Estatística. Cidades e Estados - Toledo. **IBGE**. Disponível em: <<https://www.ibge.gov.br/cidades-e-estados/pr/toledo.html>>> Acesso em: 12 dez. 2021.

KAMPEL, S. A.; CÂMARA, G.; QUINTANILHA, J. A. **Análise exploratória das relações espaciais do desflorestamento na Amazônia legal brasileira**. ResearchGate, GISBrasil 2000, Salvador, jan. 2000.

LIM, P. L. et al. Laboratory-acquired severe acute respiratory syndrome. **The New England Journal of Medicine**, v.350, n.17, p.1740-1745, 2004.

LIMA-COSTA, M. F.; BARRETO, S. M. Tipos de estudos epidemiológicos: conceitos básicos e aplicações na área do envelhecimento. **Epidemiologia e Serviços de Saúde**, Brasília, v.12, n. 4, dez. 2003.

LONGLEY, P. A. et al. **Sistemas e ciência da informação geográfica**. 3. ed. Porto Alegre: Bookman, 2013.

- MARCELLO, T. et al. Análise espacial da taxa de mortalidade infantil no Paraná. **Brazilian Journal of Development**, Curitiba, v. 5, n. 10, p. 18862-18876, out. 2019.
- MEDRI, W. **Análise exploratória de dados**. 2011. Centro de Ciências Exatas / Departamento de Estatística - Curso de Especialização "Lato Sensu" em Estatística - Universidade Estadual de Londrina, Londrina, 2011.
- MELO, J. D.; CARRENO, E. M.; PADRILHA-FELTRIN, A. Análise espacial de eventos pontuais na alocação de transformadores da rede elétrica secundária de distribuição. In: CONGRESSO BRASILEIRO DE AUTOMÁTICA, 21., 2012, Campina Grande. **Anais [...]** Campina Grande: CBA, 2012.
- NEVES, M. C. et al. Análise exploratória espacial de dados sócio-econômicos de São Paulo. Campinas: Embrapa Meio Ambiente, **Instituto Nacional de Pesquisas Espaciais**; 2000.
- NOBRE, A. F. S. et al. Primeira detecção de coronavírus humano associado à infecção respiratória aguda na Região Norte do Brasil. **Revista Pan-Amazônica de Saúde**, v. 5, n. 2, p. 37-41, jun. 2014.
- NUNES, F. G. Análise exploratória espacial de indicadores de desenvolvimento socioambiental das regiões de planejamento do norte e nordeste goiano. **Revista Ateliê Geográfico**, Goiânia, v. 7, n. 1, p. 237-259, jan./abr. 2013.
- PINHEIRO, C.; RUPRECHT, T. Coronavírus: primeiro caso é confirmado no Brasil. O que fazer agora?. **Veja**, São Paulo, 26 fev. 2020. Saúde. Disponível em: <<https://saude.abril.com.br/medicina/coronavirus-primeiro-caso-brasil/>>. Acesso em: 04 abr. 2021.
- RAMOS, C. S.; SANCHEZ, M. C. Estudo metodológico de classificação de dados para cartografia temática. **Geografia**, Rio Claro, v. 25(2), p. 23-52, ago. 2000.
- ROJAS, L. I.; BARCELLOS, C.; PEITER, P. Utilização de mapas no campo da epidemiologia no Brasil: reflexões sobre trabalhos apresentados no IV congresso brasileiro de epidemiologia. **Informe Epidemiológico do SUS**. Brasília, v.8, n.2, jun. 1999.
- ROTTIER, P. J. M. The Coronavirus Membrane Glycoprotein. **The Viruses: The Coronaviridae**, p. 115-139, Springer, Boston, MA, 1995.
- SEFFRIN, R. **Análise exploratória de dados espaciais aplicada a produtividade de milho no estado do Paraná**. 2017. 96 f. Dissertação (Mestrado em Tecnologias Computacionais para o Agronegócio) - Universidade Tecnológica Federal do Paraná, Medianeira, 2017.
- SOARES, A. K. T. A importância da paramentação e desparamentação seguras em infecções por aerossol, com foco à Covid-19: uma revisão da literatura. **Revista Eletrônica Acervo Saúde**, v. 13, n. 6, p. e7786, jun. 2021.
- YAMAMOTO, J. K.; LANDIM, P. M. B. **Geoestatística: conceitos e aplicações**. São Paulo: Oficina de Textos, 2013.