



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL

WEDSON LUÍS DOS REIS DANTAS

**FATORES DETERMINANTES DA ACIDENTALIDADE RODOVIÁRIA NO
NORDESTE BRASILEIRO: ANÁLISE ESPACIAL E IMPLICAÇÕES PARA A
SEGURANÇA VIÁRIA**

Recife
2025

WEDSON LUÍS DOS REIS DANTAS

**FATORES DETERMINANTES DA ACIDENTALIDADE RODOVIÁRIA NO
NORDESTE BRASILEIRO: ANÁLISE ESPACIAL E IMPLICAÇÕES PARA A
SEGURANÇA VIÁRIA**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Civil da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de mestre em Engenharia Civil. Área de concentração: Transportes e Gestão das Infraestruturas Urbanas.

Orientador (a): Prof. Dr. Maurício Oliveira de Andrade

Recife

2025

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Dantas, Wedson Luís Dos Reis.

Fatores determinantes da acidentalidade rodoviária no Nordeste brasileiro: análise espacial e implicações para a segurança viária / Wedson Luís Dos Reis Dantas. - Recife, 2025. 102f.: il.

Dissertação (Mestrado) - Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, Programa de Pós-Graduação em Engenharia Civil, 2025.

Orientação: Maurício Oliveira de Andrade.

Inclui referências.

1. Sinistros rodoviários; 2. Análise espacial; 3. Segurança viária; 4. Regressão geoponderada; 5. Rodovias federais. I. Andrade, Maurício Oliveira de. II. Título.

UFPE-Biblioteca Central

WEDSON LUÍS DOS REIS DANTAS

**FATORES DETERMINANTES DA ACIDENTALIDADE RODOVIÁRIA
NO NORDESTE BRASILEIRO: ANÁLISE ESPACIAL E
IMPLICAÇÕES PARA A SEGURANÇA VIÁRIA**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Civil da Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, como requisito parcial para obtenção do título de Mestre em Engenharia Civil, Área de Transporte e Gestão das Infraestruturas Urbanas.

Aprovada em 24/02/2025

Orientador: Prof. Dr. Maurício Oliveira de Andrade – UFPE

BANCA EXAMINADORA

participação por videoconferência
Prof. Dr. Anísio Brasileiro de Freitas Dourado (examinador interno)
Universidade Federal de Pernambuco

participação por videoconferência
Prof.^a Dr.^a Márcia Rejane Oliveira Barros Carvalho Macedo (examinadora externa)
Universidade de Pernambuco

participação por videoconferência
Prof. Dr. Erison Rosa de Oliveira Barros (examinador externo)
Universidade Federal de Pernambuco

AGRADECIMENTOS

A Deus, por me conceder força, saúde e perseverança ao longo da caminhada, iluminando meu caminho e me permitindo chegar até aqui.

À minha noiva Josiedna Góis, por todo amor, paciência e apoio incondicional em todos os momentos, especialmente nos momentos mais difíceis enfrentados durante essa jornada. Seu incentivo foi essencial para que nunca desistisse.

À minha família, em especial aos meus pais Lucenildo e Rosimar e minha irmã Ana Beatriz, que sempre acreditaram em mim e me deram todo suporte necessário para que pudesse me dedicar aos estudos. Sem a base sólida que me proporcionaram, esta conquista não seria possível.

Ao meu orientador professor Dr. Maurício Oliveira de Andrade, por sua paciência, orientação e incentivo ao longo de todo o processo. Seus conhecimentos e dedicação foram fundamentais para a realização deste trabalho.

Aos professores, por todo o conhecimento partilhado e pela dedicação em transmitir o que sabem. Cada ensinamento contribuiu significativamente para minha formação acadêmica, profissional e enquanto ser humano.

Aos membros da banca, pela disposição em ler e avaliar minha pesquisa, pelas contribuições e sugestões enriquecedoras que ajudaram a aprimorar este trabalho.

Aos meus amigos, em especial João Pedro e Iranilza, que, de diversas formas, estiveram ao meu lado, compartilhando desafios, vitórias e momentos de descontração que tornaram essa caminhada mais leve.

Aos colegas de curso, com quem compartilhei essa jornada acadêmica. A troca de conhecimentos, as discussões e a parceria durante o mestrado tornaram essa experiência ainda mais enriquecedora.

Por fim, a todos que, de alguma forma, contribuíram para esta conquista, meu sincero agradecimento.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001

RESUMO

A segurança viária é uma preocupação global, devido aos elevados índices de acidentalidade rodoviária. A análise dos fatores relacionados a essas ocorrências, apresentam-se como fundamentais, para compreender sua relevância de maneira especializada. Portanto, este trabalho pretende analisar os fatores determinantes da acidentalidade rodoviária no Nordeste brasileiro, considerando variáveis socioeconômicas, demográficas e de infraestrutura viária, com ênfase na distribuição espacial dos acidentes. A metodologia inclui a coleta de dados secundários, aplicação de técnicas estatísticas, com a regressão por Mínimos Quadrados Ordinários (MQO), Índice de Moran e a análise espacial por meio da Regressão Geograficamente Ponderada (RGP). Os resultados indicam que os acidentes rodoviários apresentam padrões espaciais influenciados por fatores socioeconômicos e pela infraestrutura viária. A análise estatística indicou que áreas com maior densidade demográfica e maior fluxo de tráfego registram maior frequência de acidentes e influência dos fatores analisados, especialmente nas capitais e regiões metropolitanas, como Fortaleza, Natal, Recife, João Pessoa, Teresina e São Luís. A extensão e a densidade viária estão positivamente associadas ao aumento da acidentalidade. A análise espacial identificou áreas críticas que requerem atenção prioritária para a implementação de políticas de segurança viária. A RGP revelou que a relação entre os fatores explicativos e os acidentes não é homogênea, variando entre os municípios. Destaca-se a necessidade de medidas integradas para a mitigação dos acidentes rodoviários, incluindo planejamento viário, fiscalização e campanhas educativas. Conclui-se que a acidentalidade rodoviária no Nordeste brasileiro é influenciada por fatores socioeconômicos, demográficos e de infraestrutura viária, reforçando a importância de políticas baseadas em evidências e abordagens regionais e metodologias espaciais para a compreensão da acidentalidade viária.

Palavras-chave: sinistros rodoviários; análise espacial; segurança viária; regressão geponderada; rodovias federais.

ABSTRACT

Traffic safety is a global concern due to the high rates of crashes. Analyzing the factors related to these occurrences is essential to understanding their significance in a specialized manner. Therefore, this study aims to investigate the determining factors of crashes in Northeastern Brazil, considering socioeconomic, demographic, and road infrastructure variables, emphasizing the spatial distribution of accidents. The methodology includes collecting secondary data and applying statistical techniques, such as Ordinary Least Squares (OLS) regression, Moran's Index, and spatial analysis using Geographically Weighted Regression (GWR). The results indicate that crashes exhibit spatial patterns influenced by socioeconomic factors and road infrastructure. Statistical analysis shows that areas with higher population density and traffic flow record higher accident frequency, particularly in capital cities and metropolitan regions such as Fortaleza, Natal, Recife, João Pessoa, Teresina, and São Luís. Road extension and density are positively associated with increased accident rates. The spatial analysis identified critical areas that require priority attention for implementing traffic safety policies. The GWR analysis revealed that the relationship between explanatory factors and crashes is not homogeneous, varying across municipalities. This study highlights the need for integrated measures to mitigate crashes, including urban planning, enforcement, and educational campaigns. The findings conclude that road crash rates in Northeastern Brazil are influenced by socioeconomic, demographic, and infrastructure factors, reinforcing the importance of evidence-based policies, regional approaches, and spatial methodologies for a deeper understanding of road accident patterns.

Keywords: crashes; spatial analysis; road safety; geographically weighted regression; federal highways.

LISTA DE ILUSTRAÇÕES

Figura 1 - Fluxograma das etapas do método e atividades de pesquisa	24
Figura 2 - Total de acidentes rodoviários e óbitos por região do Brasil, dados de 2023	25
Figura 3 – Matriz de Correlação Linear de Pearson	45
Figura 4 – Índice de Moran Local para o modelo da variável dependente Acidentes por tráfego	71
Figura 5 – Índice de Moran Local para o modelo da variável dependente Vítimas Fatais	72
Figura 6 – Índice de Moran Local para o modelo da variável dependente Acidentes Totais	74
Figura 7 – Coeficiente de determinação (R^2) local para a variável dependente acidentes totais	78
Figura 8 – Coeficiente de determinação (R^2) local, variável dependente acidentes por quilômetro	79
Figura 9 – Coeficiente de determinação (R^2) local, variável dependente acidentes por milhão de veículos do tráfego	80
Figura 10 – Coeficiente de determinação (R^2) local para a variável dependente acidentes sem vítimas.	81
Figura 11 – Coeficiente de determinação (R^2) local, variável dependente acidentes c/ vítimas feridas	82
Figura 12 – Coeficiente de determinação (R^2) local, variável dependente acidentes com vítimas fatais	83
Figura 13 – Coeficiente de determinação (R^2) local, variável dependente acidentes urbanos	84
Figura 14 – Coeficiente de determinação (R^2) local, variável dependente acidentes rurais	85
Figura 15 – Estimativa (β) local, modelo de acidentes totais (ACID_TOTAIS), para o Intercepto	87
Figura 16 – Estimativa (β) local, modelo de acidentes totais (ACID_TOTAIS), para as variáveis extensão de vias e densidade viária	88
Figura 17 – Estimativa (β) local, modelo de acidentes totais (ACID_TOTAIS), para as variáveis VMDA e área municipal	89

Figura 18 – Estimativa (β) local, modelo de acidentes totais (ACID_TOTAIS), para as variáveis percentual da população ocupada e IDHM 90

Figura 19 – Estimativa (β) local, modelo de acidentes totais (ACID_TOTAIS), para as variáveis população e densidade demográfica 91

LISTA DE TABELAS

Tabela 1 – Bases de dados utilizadas para obtenção das variáveis explicativas	26
Tabela 2 – Modelos Estatísticos utilizados e principais particularidades	28
Tabela 3 – Descrição das variáveis dependentes e independentes	39
Tabela 4 – Estatísticas descritivas das variáveis dependentes e independentes	40
Tabela 5 – Outliers das variáveis dependentes e independentes.	43
Tabela 6 – Teste de Fator de Inflação da Variância (VIF) das variáveis explicativas	48
Tabela 7 – Fator de Inflação da Variância (VIF) das variáveis explicativas após retirada de variáveis	49
Tabela 8 – Principais métricas indicadoras de performance do modelo MQO.	50
Tabela 9 – Coeficientes das variáveis do modelo preliminar de MQO	51
Tabela 10 – Principais métricas indicadoras de performance do modelo MQO.	53
Tabela 11 – Métricas indicadoras de performance dos modelos de MQO para a variável ACID_TOTAL, modelo com 12 e com 8 variáveis explicativas.	55
Tabela 12 – Coeficientes das variáveis dos modelos de MQO: ACID_TOTAL, ACID_KM e ACID_MI_V	57
Tabela 13 – Coeficientes das variáveis dos modelos de MQO: SEM_VIT, VIT_FERIDA e VIT_FATAIS	58
Tabela 14 – Coeficientes das variáveis dos modelos de MQO: ACID_URB e ACID_RUR	59
Tabela 15 – Principais métricas indicadoras de performance do modelo final de MQO para a variável acidentes totais	64
Tabela 16 – Coeficientes das variáveis do modelo final de MQO	64
Tabela 17 – Resultados do teste de Shapiro-Wilk	65
Tabela 18 – Resultados do teste de autocorrelação dos resíduos	66
Tabela 19 – Resultados do teste de homocedasticidade	67
Tabela 20 – Resultados do índice de Moran para todos os modelos de MQO.	69
Tabela 21 – Principais métricas indicadoras de performance do modelo RGP.	76
Tabela 22 – Métricas de performance dos modelos de MQO e RGP para a variável ACID_TOTAL.	77
Tabela 23 – Estimativa de β para as variáveis do modelo ajustado de MQO e RGP, variável dependente acidentes totais (ACID_TOTAL)	86

SUMÁRIO

1 INTRODUÇÃO	12
1.1 PROBLEMA E HIPÓTESE DE PESQUISA	14
1.2 OBJETIVOS	14
1.2.1 Objetivo Geral	14
1.2.2 Objetivos Específicos	15
1.3 JUSTIFICATIVA	15
2 REFERENCIAL TEÓRICO	17
2.1 ESTUDOS EM SEGURANÇA VIÁRIA	17
2.2 INDICADORES SOCIOECONÔMICOS E DEMOGRÁFICOS E ACIDENTES RODOVIÁRIOS	18
2.2.1 Justiça Social no âmbito da acidentalidade	18
2.2.2 Relevância de indicadores socioeconômicos e acidentes	19
2.2.3 Demografia e acidentalidade	21
2.3 TRÁFEGO, INFRAESTRUTURA VIÁRIA E TERRITORIAL	21
2.4 CARACTERÍSTICAS DA FROTA E DOS CONDUTORES	22
3 METODOLOGIA	23
3.1 DEFINIÇÕES, COLETA E TRATAMENTO DE DADOS	24
3.2 ESTATÍSTICAS DESCRITIVAS E MATRIZ DE CORRELAÇÃO DE PEARSON	27
3.3 MODELAGEM ESTATÍSTICA: MÍNIMOS QUADRADOS ORDINÁRIOS (MQO) E REGRESSÃO GEOGRAFICAMENTE PONDERADA (RGP)	27
3.3.1 Mínimos Quadrados Ordinários (MQO)	28
3.3.2 Autocorrelação Espacial e Índice de Moran	31
3.3.3 Regressão Geograficamente Ponderada (RGP)	33
3.3.4 Limitações, contribuições e perspectivas da análise espacial	34
3.4 AVALIAÇÃO DOS RESULTADOS: RELAÇÃO ENTRE VARIÁVEIS SELECIONADAS E ACIDENTES RODOVIÁRIOS	36

4 RESULTADOS E DISCUSSÃO	38
4.1 ESTATÍSTICAS DESCRITIVAS E MATRIZ DE CORRELAÇÃO	38
4.1.1 Estatísticas descritivas das variáveis	38
4.1.2 Matriz de Correlação de Pearson	44
4.2 TESTE DE FATOR DE INFLAÇÃO DA VARIÂNCIA (VIF)	47
4.3 MÍNIMOS QUADRADOS ORDINÁRIOS (MQO)	49
4.3.1 Modelo preliminar de Mínimos Quadrados Ordinários para a variável acidentes totais (ACID_TOTAL)	50
4.3.2 Modelo ajustados de Mínimos Quadrados Ordinários	53
4.3.2.1 Indicadores de desempenho dos modelos ajustados	53
4.3.2.2 Coeficientes das variáveis relacionadas ao tráfego, infraestrutura viária e territorial	59
4.3.2.3 Coeficientes das variáveis relacionadas aos indicadores socioeconômicos e demográficos	61
4.3.2.4 Coeficientes das variáveis relacionadas a frota e condutores	62
4.3.3 Modelo final para a variável dependente acidentes totais (ACID_TOTAL)	64
4.3.4 Análise dos pressupostos do modelo linear	65
4.3.5 Análise geral sobre os modelos de MQO	67
4.4 ÍNDICE DE MORAN	68
4.4.1 Índice de Moran Global	68
4.4.2 Índice de Moran Local ou <i>LISA (Local Indicators of Spatial Association)</i>	70
4.5 REGRESSÃO GEOGRAFICAMENTE PONDERADA (RGP)	75
4.5.1 Indicadores de Desempenho da RGP	75
4.5.2 Coeficientes de determinação (R^2) locais do modelo de RGP	77
4.5.3 Distribuição espacial das estimativas de β locais	86
4.6 DISCUSSÃO GERAL DOS RESULTADOS	91
5 CONCLUSÃO	93
REFERÊNCIAS	97

1 INTRODUÇÃO

A segurança viária destaca-se como uma pauta de crescente relevância em todo o mundo. Nesse contexto, a Organização Mundial da Saúde (OMS, 2021), no âmbito da Década de Ação pela Segurança Viária, estabeleceu uma meta global de reduzir em 50% as mortes e lesões no trânsito até 2030. Essa meta está alinhada aos Objetivos de Desenvolvimento Sustentável (ODS), em especial ao quarto objetivo (Saúde e Bem-estar), que enfatiza a importância dessa temática. A relevância global da segurança viária evidencia a necessidade de estudos que busquem reduzir os riscos de acidentes, possibilitando a proposição de estratégias eficazes para a mitigação desses indicadores.

Nesse sentido, diversos trabalhos exploram fatores associados à accidentalidade, incluindo condições socioeconômicas (NANTULYA & REICH, 2003), interações entre pedestres e veículos (KONG & YANG, 2010), veículos motorizados e não motorizados (LEE, ABDEL-ATY & JIANG, 2015), ocorrências de acidentes com capotamento (CHEN *et al.*, 2016), risco de acidentes por modo de transporte (WANG, HUANG & ZENG, 2017), usuários vulneráveis das vias (CAI, ABDEL-ATY & LEE, 2017), densidade populacional e segurança viária (GUERRA, DONG & KONDO, 2019) e análise espacial relacionando acidentes com pedestres e o ambiente construído (ZAFRI & KHAN, 2022).

Dentre essas abordagens, destaca-se a análise espacial dos sinistros. Gomes (2018) investigou a influência da distribuição espacial dos registros, destacando a necessidade de compreender suas características por meio da delimitação de áreas e da consideração de variáveis descritivas, como fatores socioeconômicos e demográficos, características da rede viária e o grau de exposição. Esse estudo revelou a existência de autocorrelação espacial entre variáveis, indicando dependências mútuas. Além disso, concluiu que as variações nas estatísticas descritivas dependem da escala de análise, bem como do formato das variáveis e dos critérios de delimitação das áreas.

A análise de Quddus (2008) reforçou essa perspectiva ao explorar as relações entre sinistros de trânsito e fatores contribuintes em diferentes unidades de agregação espacial, como bairros em Londres. O estudo demonstrou que os sinistros estão relacionados a fatores socioeconômicos, infraestrutura viária e condições de tráfego

locais, e ressaltou a necessidade de análises em unidades maiores para minimizar efeitos de agregação espacial nos modelos.

Nos países de renda baixa e média, a análise espacial dos sinistros ganha ainda mais relevância. Mesa-Arango *et al.* (2018) destacou os desafios próprios desses países, evidenciando que o aumento da produtividade pode estar associado à redução de acidentes registrados. Contudo, a literatura apresenta divergências, como em Lee, Abdel-Aty & Jiang (2015), que associaram o aumento da produtividade e da atração de viagens ao crescimento das taxas de sinistros viários. Esses achados reforçam a importância de investigar as causas subjacentes a essas tendências e de compreender como os sinistros se relacionam com indicadores socioeconômicos.

Diversos trabalhos abordaram a acidentalidade com foco em características próprias dos sinistros rodoviários da região nordeste. Estes estudos abordam aspectos relevantes sobre a acidentalidade na região Nordeste. Nas vias rurais, destaca-se o risco de acidentes fatais em comparação com áreas urbanas, em que apontou 3,3 vezes mais chances de acidentes fatais (MACEDO *et al.*, 2022). Enquanto que nas zonas urbanas verificando o efeito das medidas da Década de Ação para a Segurança Viária, observa-se que apesar da redução no número de acidentes graves e óbitos, as intervenções em alguns casos apresentaram pouco ou nenhum impacto sobre parte dos acidentes registrados (SILVA, GOMES & CUNTO, 2025).

Além disso, outros estudos apontam para a influência das características da infraestrutura viária na acidentalidade. Fatores como extensão e densidade de vias, sinalização, comprimento de ciclovias e calçadas para pedestres demonstram que uma infraestrutura adequada pode reduzir significativamente os riscos de acidentes rodoviários (SIDDIQUI, ABDEL-ATY & CHOI, 2012).

Os riscos também estão associados às características do fluxo de tráfego e ao comportamento dos condutores. Velocidades elevadas e áreas com alto volume de tráfego contribuem para o aumento dos conflitos no trânsito, resultando em maiores índices de sinistros registrados (HUANG, WANG & PATON, 2018).

Portanto, torna-se essencial explorar as características socioeconômicas, demográficas, de infraestrutura viária, territoriais, relacionadas ao tráfego e aos condutores, sob o aspecto da distribuição espacial dos sinistros. Essa abordagem auxilia a propor estratégias eficazes de redução dos sinistros viários.

Desse modo, esta dissertação propõe uma metodologia inovadora e ainda pouco explorada no Brasil para o contexto dos acidentes rodoviários, ao adotar uma

análise em escala mais ampla, em nível regional e utilizando a unidade de agregação municipal. Essa abordagem destaca-se ao focar a região Nordeste do Brasil, uma área que carece de estudos abrangentes sobre a temática, visando preencher a lacuna existente na literatura ao compreender as disparidades regionais e os fatores que contribuem para as diferenças nos indicadores de acidentalidade.

1.1 PROBLEMA E HIPÓTESE DE PESQUISA

O presente trabalho baseia-se na seguinte pergunta de pesquisa: Em que medida as características socioeconômicas, demográficas e de infraestrutura viária influenciam a distribuição e a incidência de acidentes rodoviários em rodovias federais do Nordeste brasileiro?

Com base nesse questionamento, propõe-se a seguinte hipótese: A ocorrência de acidentes apresenta autocorrelação espacial, sendo influenciada por fatores socioeconômicos, demográficos e de infraestrutura viária. Regiões com maior vulnerabilidade socioeconômica tendem a registrar maior frequência e severidade de acidentes, resultado da interação entre variáveis como fatores demográficos, econômicos, educacionais, características dos condutores, infraestrutura viária, volume de tráfego e composição da frota de veículos.

1.2 OBJETIVOS

1.2.1 Objetivo Geral

Analisar a relação entre variáveis socioeconômicas, demográficas e de infraestrutura viária na distribuição espacial e nos índices de acidentes rodoviários nas rodovias federais do Nordeste brasileiro, visando identificar padrões e tendências que possam fornecer informações sobre a relação espacial entre essas variáveis e os acidentes.

1.2.2 Objetivos Específicos

- Examinar as relações entre características socioeconômicas e demográficas, identificando padrões estatísticos associados aos índices de acidentes rodoviários e verificando sua relevância no contexto da região Nordeste;
- Avaliar a influência dos fatores relacionados à infraestrutura viária e territorial, compreendendo de que forma a extensão e densidade viárias, bem como as características das vias, impactam na ocorrência de sinistros;
- Explorar a relação entre características do tráfego, como volume e composição da frota de veículos, e os acidentes rodoviários, considerando os níveis de exposição ao risco;
- Identificar padrões espaciais de concentração de acidentes, considerando frequência e gravidade, com o intuito de inferir os fatores locais e regionais determinantes e destacar as regiões que apresentam maiores desafios;
- Contribuir para a formulação de políticas públicas direcionadas à redução de acidentes rodoviários, apresentando evidências concretas sobre a importância de integrar fatores socioeconômicos regionais na implementação de medidas de segurança nas rodovias federais.

1.3 JUSTIFICATIVA

A segurança viária é um tema cada vez mais relevante em nível global, com impactos sobre a saúde pública, economia e qualidade de vida. De acordo com a Organização Mundial da Saúde (OMS, 2021), os acidentes de trânsito representam uma das principais causas de morte em diversos países, especialmente de renda baixa e média. As ocorrências em rodovias merecem maior atenção devido à complexidade das interações, que englobam trechos rurais e urbanos e apresentam riscos que devem ser identificados.

Portanto, torna-se relevante desenvolver estudos que permitam a proposição de estratégias para a redução dos índices nas rodovias. Pretende-se, por meio deste trabalho, contribuir para essa discussão adotando uma abordagem ampliada de análise espacial, que visa compreender os fatores que influenciam os sinistros viários na região Nordeste do Brasil.

Outro fator que reforça a relevância da pesquisa é a carência de estudos com metodologias que investiguem a distribuição dos sinistros e os fatores destacados em escalas maiores. Grande parte da literatura concentra-se em análises no âmbito de bairros, blocos de quarteirões, zonas de tráfego ou zonas censitárias de menores dimensões (GOMES, 2018).

Entretanto, outros autores destacam que podem ser usadas como unidades de agregação, áreas maiores e até mais complexas, como condados (delimitação local equivalente a municípios em países como os Estados Unidos, Canadá e Reino Unido), de modo a considerar as particularidades locais que influenciam a dinâmica dos acidentes rodoviários (CAI *et al.*, 2016). Este trabalho busca explorar por meio da agregação municipal, um nível de análise pouco estudado, pretendendo abranger a lacuna observada na literatura, permitindo uma compreensão mais ampla das variações territoriais e dos fatores a elas relacionadas.

A relevância do estudo é reforçada por suas contribuições observadas para a formulação de políticas públicas eficazes. Ao identificar os fatores que contribuem para as disparidades regionais, espera-se fornecer subsídios para a elaboração de estratégias adaptadas à realidade dos municípios nordestinos. O que pode incluir desde melhorias na infraestrutura viária até iniciativas educacionais e de fiscalização, visando a redução dos índices de accidentalidade.

Ademais, ao combinar análise espacial com dados da região Nordeste, este estudo contribui para o avanço do conhecimento científico na área de segurança viária. Ele abre caminho para novas pesquisas e abordagens, ao mesmo tempo em que promove visibilidade a uma região frequentemente marginalizada nos debates sobre infraestrutura e desenvolvimento, reforçando a importância de políticas públicas adaptadas às especificidades locais.

A presente dissertação se divide em cinco capítulos em que se introduz, explica, discute e se conclui a respeito do problema de pesquisa proposto e com base nos objetivos enunciados. Este capítulo primeiro apresenta a Introdução, contando com a apresentação da introdução ao tema, problema de pesquisa, objetivos e justificativa. O Capítulo 2 apresenta o referencial teórico, abordando os principais estudos relacionados à segurança viária, bem como os indicadores socioeconômicos e demográficos, tráfego, infraestrutura viária e territorial, características da frota e dos condutores, além das abordagens estatísticas empregadas, incluindo o método dos mínimos quadrados ordinários e a regressão geograficamente ponderada.

O Capítulo 3 descreve a metodologia adotada, detalhando as definições e processos de coleta de dados, estatísticas descritivas, matriz de correlação de Pearson, modelagem estatística e os critérios de avaliação dos resultados. No Capítulo 4, são apresentados e discutidos os resultados da pesquisa, incluindo as estatísticas descritivas, matriz de correlação, teste de fator de inflação da variância, análise espacial por meio do índice de Moran, bem como a modelagem estatística baseada nos mínimos quadrados ordinários e na regressão geograficamente ponderada. Por fim, o Capítulo 5 apresenta as conclusões do estudo, destacando as principais contribuições, limitações e sugestões para pesquisas futuras.

2 REFERENCIAL TEÓRICO

2.1 ESTUDOS EM SEGURANÇA VIÁRIA

Diversos estudos no âmbito da segurança viária buscam mensurar e compreender melhor a ocorrência dos acidentes rodoviários em localidades específicas, estados ou países. Esses trabalhos procuram analisar aspectos relevantes a respeito dessas ocorrências, considerando tanto fatores relacionados às vítimas, quanto elementos relacionados ao contexto dos acidentes, como a região, os dias da semana, o tipo do acidente, as características da pista, condições da via, entre outros (BARROSO JÚNIOR, BERTHO & VEIGA, 2016).

Inicialmente, os sinistros rodoviários têm sido predominantemente estudados sob a perspectiva da saúde pública. No Brasil, diversas pesquisas têm abordado essa temática, investigando, por exemplo, a morbidade hospitalar por acidentes de trânsito em hospitais públicos (DESLANDES & SILVA, 2000), analisando estimativas de lesões no trânsito (MALTA *et al.*, 2008), a influência da geometria da via e acidentes (KARLAFTIS & GOLIAS, 2002; MACEDO *et al.*, 2021) ou avaliando o impacto de variáveis demográficas e comportamentais sobre as ocorrências (LIMA & CRUZ JÚNIOR, 2016).

Paralelamente, outros estudos focam nos aspectos relacionados à estatística e o setor de transportes, com o intuito de compreender a influência dos acidentes sobre o sistema viário. Nessa perspectiva, os sinistros são vistos como uma externalidade a ser mitigada, de modo a promover maior segurança aos usuários do sistema de transporte. Os exemplos incluem análises sobre os riscos que envolvem acidentes

com pedestres (KONG & YANG, 2010), avaliando espacialmente os sinistros, para identificar modelos que capturem padrões espaciais relevantes (CAI *et al.*, 2016) ou verificando a relação entre acidentes e acessibilidade rodoviária, demografia e uso do solo (KIM, PANT & YAMASHITA, 2010).

Desse modo, analisar acidentes rodoviários engloba diversos fatores que merecem ser analisados de maneira conjunta. Logo, o impacto da infraestrutura viária e territorial, dos indicadores socioeconômicos e demográficos, do tamanho da frota e dos condutores envolvidos, merecem ser analisados, de modo a fornecer uma visão mais abrangente a respeito dos acidentes (SIDDIQUI, ABDEL-ATY & CHOI, 2012; CHIMBA, MUSINGUZI & KIDANDO, 2018; ZAFRI & KHAN, 2022).

2.2 INDICADORES SOCIOECONÔMICOS E DEMOGRÁFICOS E ACIDENTES RODOVIÁRIOS

2.2.1 Justiça Social no âmbito da acidentalidade

A justiça social apresenta uma relação com o setor de transportes marcada pelas desigualdades no acesso a recursos, infraestruturas e mobilidade sustentável. O trabalho de Lucas (2012) destaca o conceito de exclusão social no conceito dos transportes, que resulta na falta de acessibilidade a recursos essenciais como empregos, saúde e educação. Também reforça que a mobilidade afeta a equidade social, uma vez que as pessoas que sofrem com essa falta de acessibilidade apresentam barreiras que limitam sua participação social.

Vale destacar o que se conceitua como a justiça redistributiva, ou seja, a necessidade de políticas que garantam maior equidade na acessibilidade ao transporte. Apesar disso, a exclusão social ligada ao transporte é um desafio latente, para avançar, é necessário políticas integradas a outras políticas como por exemplo o planejamento urbano, habitação, saúde, educação e programas de bem-estar social (LUCAS, 2012)

Lucas *et al.* (2016) destaca sobre o fato que parte da população está exposta de maneira desproporcional às externalidades negativas do setor de transportes como por exemplo doenças crônicas relacionadas a poluição gerada pelo tráfego, mas

também ligada aos acidentes rodoviários que ocorrem e deixam vítimas feridas ou fatais.

Além disso, a pobreza de acessibilidade destaca-se como fator agravante para essa problemática. A pobreza de acessibilidade é caracterizada pela dificuldade de parcelas da população em alcançar bens e serviços. Para isso, torna-se fundamental, nesse âmbito, considerar a qualidade do transporte público e a segurança das experiências de viagem, em diferentes contextos. A presença de infraestrutura inadequada e sistemas de transporte ineficientes diminui as possibilidades em comunidades mais vulneráveis, enquanto aumentam os riscos associados à mobilidade em áreas periféricas e de baixa renda (LUCAS *et al.*, 2016).

Azetsop (2010) aborda essa discussão, enfatizando que a justiça social exige uma distribuição equitativa dos bens sociais, essenciais para a segurança viária, como estradas bem concebidas, presença policial efetiva, programas de educação para o trânsito e acesso a sistemas de saúde adequados. Porém, a realidade mostra que áreas mais pobres, frequentemente distantes dos grandes centros, têm menos acesso a esses recursos, aumentando os riscos potenciais.

As desigualdades observadas destacam uma falha sistêmica, que comprometem os direitos fundamentais das pessoas, ao respeito, à educação, à saúde e à segurança, pilares do bem-estar humano e fundamentais para a segurança viária (AZETSOP, 2010). Portanto, a segurança viária não engloba apenas uma análise técnica, mas também uma questão de justiça social. Políticas públicas que abordem essas disparidades estruturais, são fundamentais para garantir aos cidadãos, que independente de sua localização, tenham acesso a infraestruturas e sistemas de transporte seguros, equitativos e inclusivos.

2.2.2 Relevância de indicadores socioeconômicos e acidentes

Diversos estudos buscam compreender a associação entre variáveis socioeconômicas e acidentes rodoviários. O trabalho pioneiro de Nantulya e Reich (2003) abordou a questão dos acidentes rodoviários, comparando países de baixa e média renda. Esse estudo demonstrou, por meio de dados, a disparidade entre essas áreas, evidenciando o fardo desproporcional de morbidade e mortalidade em regiões economicamente vulneráveis, em comparação com países desenvolvidos. Isso ilustra

a desigualdade social e o impacto significativo atribuído a determinados países e grupos socioeconômicos.

Conforme estudo de Siddiqui, Abdel-Aty e Choi (2012), a renda domiciliar tende a ser negativamente associada a acidentes, o que demonstra que áreas com menor renda apresentam maior probabilidade de ocorrência de acidentes, sobretudo envolvendo pedestres. A presença de um fenômeno conhecido como gentrificação, que consiste na atração de novos investimentos imobiliários e de negócios em áreas anteriormente negligenciadas, também influencia a segurança viária.

De acordo com González, Laukaitou-Sideris e Chapple (2019), estudos conduzidos em Los Angeles mostraram que o desenvolvimento orientado ao transporte público nos entornos de estações de metrô aumentou os riscos de colisões com ciclistas e pedestres, devido ao aumento do fluxo nessas regiões.

Por outro lado, Li *et al.* (2022) destaca a relação entre vulnerabilidade social e fatores socioeconômicos com acidentes fatais. Comunidades de baixa renda em regiões densamente povoadas e com altos índices de vulnerabilidade social demonstram forte correlação com a ocorrência de acidentes. Tomando por base o conceito de justiça social, diversos aspectos dessas variáveis podem ser considerados. Liu, Das e Khan (2024) identificaram que áreas com maior concentração de trabalhadores de baixa renda, altos níveis de emprego e entropia domiciliar apresentam taxas elevadas de acidentes.

Noland e Laham (2018) abordaram a relação entre grupos de renda, etnias e raças nos Estados Unidos, constatando que os riscos de fatalidades entre esses grupos não apresentaram variações significativas. Contudo, grupos com melhores condições socioeconômicas que utilizam mais os veículos do que outros, estão sujeitos a maior exposição, o que pode implicar em um fardo desproporcional para parcelas vulneráveis no atual sistema de transporte.

Entretanto, as conclusões apontadas pela literatura não são completamente replicáveis, uma vez que as características socioeconômicas observadas em uma região podem não se repetir de maneira idêntica em outras. As particularidades de uma área específica nem sempre se aplicam a diferentes regiões ou países com composições socioeconômicas e demográficas distintas. Portanto, é essencial que novos estudos examinem as características específicas de cada localidade ou região.

2.2.3 Demografia e acidentalidade

O estudo pioneiro de LaScala, Gerber e Gruenewald (2000) investigou os acidentes envolvendo pedestres e sua relação com fatores demográficos. Os resultados desse trabalho mostraram que as taxas de lesões envolvendo pedestres estavam relacionadas à densidade populacional, à composição etária da população e a outros fatores, como fluxo de tráfego, níveis de desemprego, gênero e educação. Esses fatores são analisados de diferentes formas conforme o contexto local.

Chimba, Musinguzi e Kidando (2018) também analisaram a relação entre fatores socioeconômicos e demográficos, dando ênfase a acidentes com pedestres. Os resultados desses autores reforçam uma relação positiva entre densidade populacional, faixa etária entre 15 e 64 anos, bairros onde os moradores caminham para o trabalho e domicílios sem veículos, com a frequência de acidentes envolvendo pedestres.

No entanto, é importante notar que tais resultados não se repetem de maneira igual em diferentes regiões. Guerra, Dong e Kondo (2019) analisaram a relação entre densidade populacional e segurança de pedestres, apontando uma relação fraca entre esses fatores. O estudo indicou que variáveis como o traçado das vias e a presença de transporte público podem exercer uma influência maior que a densidade populacional, no que diz respeito à segurança viária.

2.3 TRÁFEGO, INFRAESTRUTURA VIÁRIA E TERRITORIAL

O fluxo de tráfego em uma via apresenta forte relação com os acidentes rodoviários. De acordo com Cai *et al.* (2016), o Volume Médio de Tráfego aumenta a propensão de registros. Assim, o aumento do volume de tráfego implica no aumento dos índices de acidentalidade. O tráfego é uma medida de exposição dos veículos, logo o aumento das ocorrências é coerente com o aumento da exposição.

A infraestrutura viária é fundamental para garantir a segurança aos usuários da via e evitar a ocorrência dos acidentes. De acordo com Chimba, Musinguzi e Kidando (2018), a falta de investimento em infraestrutura rodoviária está associada a um aumento de até 6,3% na taxa de fatalidades. Analisaram também que a presença adequada de canteiro central, acostamentos e calçadas contribui para a diminuição dos acidentes com pedestres. Esse resultado reforça a importância de investimentos

nessas infraestruturas, nos trechos rodoviários urbanos, conferindo mais segurança aos pedestres.

A presença de interseções é um fator relevante no que diz respeito ao traçado das vias. Os acidentes têm forte associação com a presença de interseções, principalmente envolvendo pedestres em zonas urbanas, devido à presença de um grande número de cruzamentos de pedestres nesses pontos e também à necessidade de manobras veiculares mais complexas. Esses fatores levam a uma maior exposição dos pedestres nas vias, aumentando a chance de colisões (SIDDIQUI, ABDEL-ATY & CHOI, 2012).

Outro fator importante da infraestrutura viária é sua extensão e densidade de vias. Mais quilômetros de rodovias resultam em mais interações e, portanto, maior risco de acidentes (LI et al., 2022). De acordo com Liu, Das e Khan (2024), a densidade de vias, que considera a extensão das vias pela área da unidade de agregação (em milhas por milha quadrada ou quilômetros por quilômetro quadrado), apresenta correlação significativa com os acidentes com danos à propriedade e com vítimas, corroborando os resultados de densidade de interseções, densidade de redes orientadas para pedestres e densidade de redes multimodais, o que se dá devido ao fato que redes mais densas, podem levar a cenários mais complexos, com o aumento das interações entre os usuários das vias. Melhorias são necessárias para buscar alternativas de infraestrutura que acomodem diferentes modos de transporte.

Desse modo, em comparando os acidentes que apresentam somente danos à propriedade e os registros com vítimas, a presença de maior densidade de infraestruturas para pedestres, contribuiu para menores índices de acidentes graves ou fatais (LIU, DAS & KHAN, 2024).

A área da unidade de agregação é um fator relevante, uma vez que apresenta relação com os acidentes rodoviários. A presença de áreas maiores tende a uma menor acidentalidade, uma vez que na maioria das vezes observa-se a dispersão desses registros ao longo de uma área extensa (WIER *et al.*, 2009).

2.4 CARACTERÍSTICAS DA FROTA E DOS CONDUTORES

A motorização, crescente em muitas regiões ao redor do mundo, tem um impacto direto na acidentalidade. A concentração de veículos motorizados eleva a complexidade das interações viárias, intensificando os conflitos entre os modos de

transporte e elevando os riscos aos pedestres e ciclistas nas vias (ZAFRI & KHAN, 2022).

Portanto, é evidente que a gestão adequada do volume de veículos, associada a intervenções no planejamento urbano e viário, é fundamental, para mitigar os efeitos da motorização na segurança viária. Levando em consideração que a motorização elevada contribui para acidentes graves ou letais em países em desenvolvimento (MESA-ARANGO *et al.*, 2018).

A presença de motocicletas na frota de uma determinada localidade é um fator relevante para a segurança viária. Os acidentes envolvendo motociclistas são bastante expressivos no contexto da acidentalidade. Quanto aos riscos de letalidade dos sinistros, observa-se que os ocupantes de motocicletas, dentre os veículos motorizados, apresentam risco elevado, maior que o de ocupantes de automóveis (BARROSO JÚNIOR, BERTHO & VEIGA, 2019).

No que diz respeito à idade dos condutores, muitos autores têm buscado encontrar uma relação entre os acidentes e esse fator. A ocorrência de acidentes com pedestres está relacionada positivamente com a idade dos envolvidos em sinistros, apresentando forte relação na faixa etária de 16 a 29 anos. Além disso, as dificuldades cognitivas em razão da idade também podem contribuir para tais ocorrências (LASCALA, GERBER & GRUENEWALD, 2000).

De acordo com Chimba, Musinguzi e Kidando (2018), o percentual de pessoas entre 15 e 64 anos está relacionado positivamente com os acidentes, possivelmente porque essa faixa engloba majoritariamente a população economicamente ativa, segundo a Organização Internacional do Trabalho (OIT), que precisa se deslocar com maior frequência e, conseqüentemente está mais exposta aos riscos nas vias.

3 METODOLOGIA

A proposta metodológica utilizada nesta dissertação para alcançar os objetivos geral e específicos foi realizada a partir das seguintes etapas sequenciais: i) Definições, Coleta e Tratamento de dados; ii) Estatísticas descritivas e Matriz de Correlação de Pearson; iii) Modelagem Estatística: Mínimos Quadrados Ordinários (MQO) e Regressão Geograficamente Ponderada (RGP); iv) Avaliação dos Resultados: Relação entre variáveis selecionadas e acidentes rodoviários. A Figura 1 ilustra o fluxograma das etapas da metodologia utilizada nessa pesquisa.

Figura 1 - Fluxograma das etapas do método e atividades de pesquisa



Fonte: O autor (2025).

3.1 DEFINIÇÕES, COLETA E TRATAMENTO DE DADOS

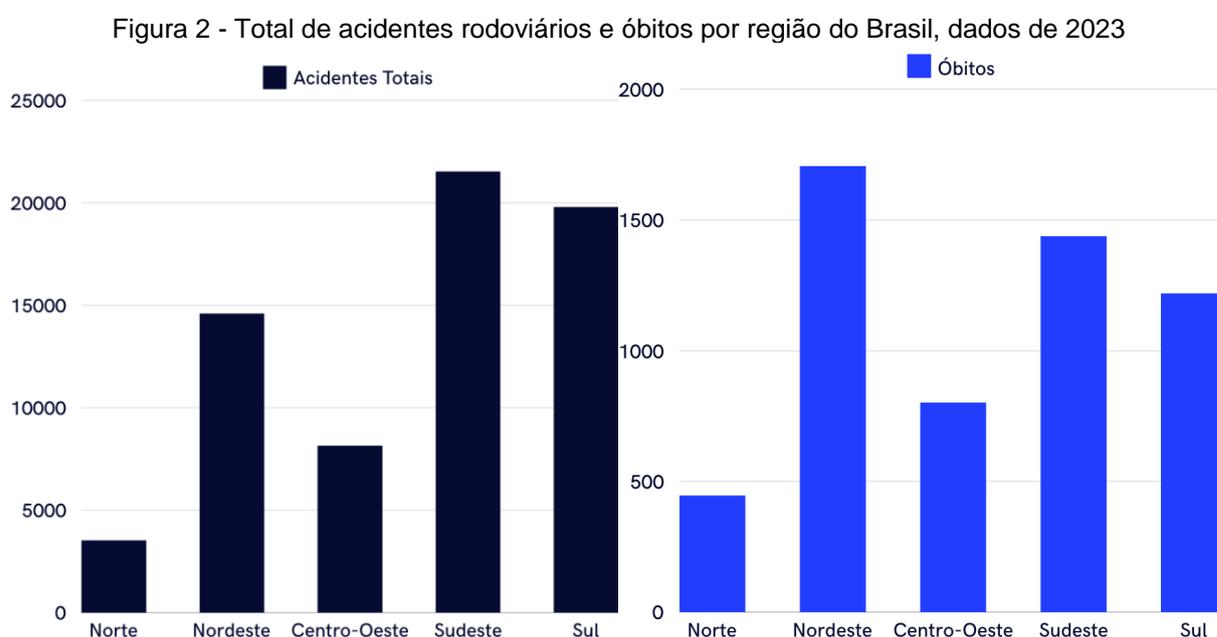
Nesta seção, são apresentadas as etapas adotadas para a definição, coleta e tratamento dos dados utilizados nesta pesquisa. O intuito foi selecionar as variáveis e os dados de interesse, considerando os mais adequados ao contexto dos modelos estatísticos analisados pela pesquisa.

Nas definições iniciais da pesquisa buscou-se compreender o recorte a ser utilizado na pesquisa, no qual optou-se pelo recorte transversal, concentrando-se em um único ano de referência, no caso 2022. O ano em questão apresenta um número de acidentes que indica estabilidade dos números em comparação com anos anteriores, até com os registros anteriores à pandemia de Covid-19, demonstrando se tratar de um ano típico, no que diz respeito aos índices de acidentes (CNT, 2023).

O recorte temporal do ano de 2022, também se deve ao fato de haver sido realizado o Censo demográfico, no mesmo ano, conferindo ao período a presença de

informações socioeconômicas atualizadas, a serem incorporadas às análises estatísticas.

A escolha do recorte espacial dos municípios região Nordeste do Brasil, como foco dessa pesquisa deve-se à importância dessa região para a temática, dado os elevados índices de acidentes rodoviários em toda sua malha, às características de sua infraestrutura rodoviária, sua heterogeneidade em zonas urbanas e rurais e principalmente devido ao elevado número de óbitos, maior em comparação com outras regiões, com malhas viárias maiores. A Figura 2 a seguir demonstra por meio dos dados de acidentes rodoviários da Polícia Rodoviária Federal – PRF (2023), os índices elevados dessa região comparada com as demais.



Fonte: O autor (2025), com dados da Polícia Rodoviária Federal (2023).

Os indicadores de acidentes e óbitos por região evidenciam que o Nordeste ocupa a terceira posição em número total de acidentes. Contudo, no que se refere ao total de óbitos por acidentes rodoviários, a região destacou-se como a que registrou o maior número de fatalidades, indicando um número elevado de ocorrências letais, destacando a gravidade dos registros. Esses dados reforçaram a importância de aprofundar os estudos sobre acidentes rodoviários na região Nordeste, levando em consideração as peculiaridades regionais e os fatores que contribuem para a alta mortalidade.

Os dados foram coletados a partir de diversas bases disponíveis, buscando adequar os dados que se pretende obter aos objetivos da pesquisa. As bases de dados foram selecionadas de acordo com sua relevância, atualidade e qualidade da informação a respeito dos acidentes rodoviários, indicadores socioeconômicos e geográficos e de tráfego. A Tabela 1 destaca as bases de dados utilizadas por grupo de variáveis selecionadas, de acordo com a natureza de cada variável.

Tabela 1 – Bases de dados utilizadas para obtenção das variáveis explicativas

Grupos de variáveis	Base de dados
Acidentes rodoviários	Polícia Rodoviária Federal (2023)
Infraestrutura viária e territorial	DNIT (2021) e IBGE (2023)
Indicadores socioeconômicos e demográficos	IBGE (2023)
Características relacionadas à frota e condutores	Ministério dos Transportes (2022) e IBGE (2023)

Fonte: O autor (2025).

Os dados utilizados foram integrados e incorporados, por meio do QGIS, em formato vetorial das shapefiles dos municípios da região nordeste. Esse trabalho fez uso desses dados vetoriais e descritivos ou tabulares (presentes nas tabelas de atributos dos modelos), com os dados de cada unidade de agregação para as variáveis dependentes e independentes utilizadas. Esse formato permitiu a adequada execução dos modelos lineares e espaciais.

O processo de seleção das variáveis, em que foram selecionados os parâmetros que seriam utilizados nas análises, envolveu uma análise minuciosa da natureza das variáveis disponíveis, bem como sua relevância no que diz respeito à literatura existente que trata de análises de acidentes e sócio economia. O critério para selecionar as variáveis visou maximizar o potencial explicativo dos modelos utilizados, buscando que os dados apresentem relação direta com as questões levantadas pela pesquisa.

Após a coleta, as bases de dados foram unificadas, considerando a unidade de agregação espacial utilizada, que neste caso foram os municípios. O processo de junção dos dados envolveu a padronização das informações, de modo a assegurar a compatibilidade entre as diferentes bases e permitindo a análise conjunta, garantindo também o adequado tratamento, de modo a garantir a qualidade e confiabilidade dos dados utilizados nos modelos estatísticos.

3.2 ESTATÍSTICAS DESCRITIVAS E MATRIZ DE CORRELAÇÃO DE PEARSON

Como etapa preliminar à modelagem estatística, foram realizadas análises para fornecer um entendimento inicial das características dos dados, com o intuito de identificar padrões, tendências e possíveis relações entre as variáveis. Além disso, essas análises permitiram atestar a qualidade dos dados, para identificar problemas não detectados nas etapas anteriores de tratamento. Para isso, foram geradas tabelas de análise descritiva de dados, que possibilitaram a visualização da distribuição das variáveis e de suas relações espaciais. Essas ferramentas permitiram identificar a presença de assimetrias, valores extremos e padrões de distribuição distintos.

O cálculo das estatísticas descritivas, inclui medidas de tendência central (mediana, média), medidas de posição (quartis, mínimo e máximo) e medidas de dispersão (coeficiente de variação e desvio padrão). Essas medidas fornecem informações detalhadas a respeito das características dos dados, contribuindo para identificar variáveis que poderiam necessitar de transformações adicionais antes de serem utilizadas na modelagem (ZAFRI & KHAN, 2022).

A matriz de correlação de Pearson foi elaborada para avaliar as relações lineares entre as variáveis. Essa análise possibilitou identificar variáveis mais fortemente associadas aos acidentes rodoviários e aquelas com alta correlação entre si, evidenciando possíveis problemas de multicolinearidade. As informações obtidas nessa etapa foram fundamentais para auxiliar na execução e interpretação dos modelos estatísticos, contribuindo para a validação e compreensão dos resultados.

3.3 MODELAGEM ESTATÍSTICA: MÍNIMOS QUADRADOS ORDINÁRIOS (MQO) E REGRESSÃO GEOGRAFICAMENTE PONDERADA (RGP)

A modelagem estatística realizada nesta pesquisa aplicou diferentes técnicas para analisar as relações entre as variáveis explicativas e os acidentes rodoviários, destacando a dimensão espacial desse fenômeno. Nesta seção detalham-se os modelos utilizados e a metodologia empregada, com o objetivo de garantir a assertividade das análises. A Tabela 2, apresenta as principais diferenças entre os dois modelos estatísticos empregados, ressaltando as abordagens complementares utilizadas.

Tabela 2 – Modelos Estatísticos utilizados e principais particularidades

Modelo	Descrição	Principais recursos
MQO (Mínimos Quadrados Ordinários)	Modelo Global de Regressão que minimiza a soma dos quadrados residuais	Assume relações constantes no espaço geográfico, espaço não-estacionário. Fornecer uma equação única para toda a regressão.
RGP (Regressão Geograficamente Ponderada)	Expande a Regressão Global a um modelo que permite que as relações entre as variáveis variem espacialmente	Utiliza-se de pesos espaciais, para conferir importância maior para dados mais próximos. Assume escala univariada, para todos os preditores.

Fonte: Adaptado de Shabrina, Buyuklieva e Ming (2021).

3.3.1 Mínimos Quadrados Ordinários (MQO)

A modelagem estatística de acidentes rodoviários é amplamente utilizada para avaliação de riscos, identificação de variáveis críticas e fatores mais relevantes no contexto da segurança viária. Nesse sentido, diversos trabalhos fazem uso de modelos de regressão linear por mínimos quadrados ordinários, para avaliar os efeitos das variáveis explicativas no contexto da acidentalidade (WIER *et al.*, 2009; HUANG, WANG & PATON, 2018; LI *et al.*, 2022).

Muitos trabalhos focados na frequência de acidentes, conhecidos como modelos de previsão de acidentes, basearam-se em regressão linear múltipla com uso de mínimos quadrados ordinários. No entanto, diversos estudos demonstraram as limitações desse modelo na previsão de acidentes devido às características dos dados de acidentes rodoviários (KARLAFTIS & TARKO, 1998)

A utilização de modelos de regressão linear simples assume variações espaciais constantes, o que nem sempre é verdadeiro em análises geográficas, que devem considerar os aspectos locais. Esses modelos precisam de validações dos pressupostos do modelo, por meio de testes de normalidade dos resíduos, homocedasticidade, presença de outliers e autocorrelação de resíduos. Apesar dessas limitações, a abordagem de regressão linear com o uso de mínimos quadrados ordinários vem sendo utilizada como alternativa preliminar para demonstrar as relações mais gerais entre as variáveis presentes nos modelos e atestar a superioridade de modelos espaciais para a modelagem de dados (SHABRINA,

BUYUKLIEVA & MING NG, 2021; LI *et al.*, 2022). Assim, o modelo de regressão linear e múltipla pode ser expresso pela Equação 1:

$$y = \beta_0 + \sum_{i=1}^n \beta_i x_i + \varepsilon \quad (1)$$

Onde:

- Y = vetor de variáveis dependentes (respostas).
- X = matriz de variáveis independentes (preditores).
- β = vetor de coeficientes a serem estimados.
- ε = vetor de erros aleatórios.

Antes da aplicação dos modelos, foi realizado o Teste do Fator de Inflação da Variância (*Variance Inflation Factor* – VIF), ferramenta essencial para avaliar a multicolinearidade entre os fatores explicativos utilizados no modelo. O VIF quantifica o quanto a variância do coeficiente de uma variável aumenta devido à correlação linear com outros presentes no modelo, identificando a redundância entre esses indicadores, valores de VIF superiores a 5 são normalmente considerados indicativos de multicolinearidade (O'Brien, 2007). A Equação X, demonstra como é definido o cálculo desse teste:

$$VIF = \frac{1}{1 - R_i^2}$$

Em que se utiliza o coeficiente de determinação para a i -ésima variável independente. Destaca-se também a importância em não considerar o valor isoladamente, mas analisando o contexto dos fatores envolvidos. Desse modo, para variáveis com altos valores de VIF, optou-se por revisá-las e, de acordo com suas características, removê-las, de modo a garantir maior estabilidade e precisão aos modelos.

O modelo de Mínimos Quadrados Ordinários (MQO) foi adotado como ponto de partida para a análise das relações lineares entre as variáveis explicativas e dependentes. Este modelo assume constantes os coeficientes em todo o espaço analisado, desconsiderando as relações espaciais que possam estar presentes nos dados (SHABRINA, BUYUKLIEVA & MING NG, 2021). Apesar de suas limitações, o modelo de MQO é amplamente utilizado devido à sua capacidade de fornecer uma visão global das relações entre os indicadores observados. Além disso, os resultados

do MQO servem como base para comparação com modelos mais complexos, como a Regressão Geograficamente Ponderada.

Também é válido destacar, que para avaliar a qualidade do modelo MQO, foram analisadas métricas como o coeficiente de determinação ajustado (R^2 ajustado) e a significância estatística dos coeficientes, por meio do p-valor.

A estimativa dos parâmetros, por meio dos coeficientes das variáveis explicativas permitem compreender o impacto relativo de cada variável independente na variável dependente (relacionadas aos indicadores dos acidentes rodoviários), permitindo a identificação dos fatores com maior influência e seu sentido (positivo ou negativo). O coeficiente β nos fornece essa medida e o impacto sobre a variável dependente, por exemplo, um coeficiente positivo com o valor de 0,002 para uma variável como VMDA, indica que, um aumento de mil veículos no volume médio diário de tráfego, geraria na variável dependente um aumento de duas unidades da variável dependente (por exemplo dois acidentes totais). Portanto, essa medida é fundamental para conhecer as relações entre essas variáveis explicativas e as variáveis dependentes.

Por meio dos resultados do modelo linear de MQO, é necessário avaliar o efeito das variáveis explicativas na variável dependente. Para isso, verificando os resultados é possível ajustar os modelos, por meio de retirada ou adaptações de variáveis, de modo a garantir resultados mais precisos e conseqüentemente com maior capacidade explicativa.

Além das métricas de desempenho dos modelos lineares é necessário investigar se o modelo atende aos pressupostos da regressão linear. Os pressupostos analisados são: a normalidade dos resíduos (por meio do teste de *Shapiro-Wilk*), o teste da homocedasticidade do modelo (teste de *White* padronizado e o teste de *Breusch-Pagan*) e a autocorrelação dos resíduos (teste de *Durbin-Watson*).

A formulação dos modelos envolveu também a escolha das variáveis resposta a serem utilizadas no modelo. A escolha das variáveis buscou abordar aspectos importantes sobre os acidentes, que são mais relevantes no que diz respeito à bibliografia na temática. Principalmente no que diz respeito: a frequência de acidentes (LI *et al.*, 2022); o nível de gravidade dos acidentes (LIU, DAS & KHAN, 2024); a localização, quanto aos trechos de rodovia localizados em zonas urbanas ou rurais e por taxa de acidentes por extensão viária ou veículos no tráfego, visando estabelecer relações entre os registros e esses indicadores (ZAFRI & KHAN, 2022).

As seguintes variáveis foram utilizadas, sendo relacionadas aos acidentes rodoviários: totais, por quilômetro de via, por milhão de veículos, sem vítimas, com vítimas feridas, com vítimas fatais, urbanos e rurais. Essa escolha ocorreu devido à importância em analisar a frequência dos acidentes, os índices relativos em comparação a extensão viária e ao tráfego, gravidade do acidente e localização (urbana e rural).

Essa abordagem visou incluir o máximo de informações relacionadas aos acidentes, a serem identificadas as relações das variáveis explicativas sobre os acidentes. Por outro lado, os fatores de acidentes envolvendo motociclistas e acidentes envolvendo ciclistas não puderam ser utilizadas como resposta pela alta presença de zeros, o que comprometeria o desempenho dos modelos. Para a utilização dessas variáveis seria preciso utilizar de um modelo de Poisson ou Binomial Negativo inflados de zeros, conforme ressaltado por Lord e Mannering (2010).

3.3.2 Autocorrelação Espacial e Índice de Moran

Para identificar a heterogeneidade espacial presente no modelo de MQO, é comum avaliar os resíduos, buscando identificar a autocorrelação espacial. O teste do Índice de Moran Global, verifica a influência dos padrões espaciais nos dados utilizados (HUANG, WANG & PATTON, 2018), fornecendo informações relevantes a respeito da região em análise. Atestada a autocorrelação espacial positiva, deve-se utilizar um modelo de regressão espacial.

De acordo com Anselin (1995), a Equação 2, expressa a fórmula do Índice de Moran Global:

$$I = \frac{n}{\sum \sum w_{ij}} \frac{\sum \sum w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum (y_i - \bar{y})^2} \quad (2)$$

Em que y_i é o valor da variável y na região i , y_j é o valor da variável y na região j , \bar{y} é a média de y , w_{ij} é o elemento ij da matriz de proximidade espacial e n é o número de observações.

Os resíduos do modelo são testados quanto à presença de autocorrelação espacial, utilizando-se o Índice de Moran Global. Desse modo, o índice foi aplicado para identificar se, globalmente, os dados de acidentes rodoviários apresentavam padrões espaciais. Essa métrica, proposta inicialmente por Moran (1948) indica que,

para valores positivos do índice, localidades próximas tendem a apresentar características semelhantes quanto aos acidentes, sugerindo agrupamento espacial, enquanto valores negativos indicam padrões de dispersão.

Os resultados obtidos, que apresentam significância estatística e valores positivos levam à conclusão de que a distribuição dos acidentes no espaço não é aleatória, reforçando a necessidade de utilização de modelos geograficamente ponderados.

Complementando a análise do Índice de Moran Global, foi utilizado o Índice de Moran Local, que permite identificar em que pontos os padrões espaciais, se destacam, seja em pontos específicos ou sub-regiões, com clusters das distribuições dos dados de acidentes ao longo do espaço geográfico.

Essa métrica fornece uma perspectiva diferente, pois destaca as dinâmicas locais e permite mapear as áreas críticas e características particulares que contribuem para esses agrupamentos ou dispersões nos diferentes modelos empregados (ANSELIN, 1995). O Índice de Moran Local, Local Indicators of Spatial Association – LISA, foi executado com 100 simulações e considerando dez vizinhos próximos, por meio do software QGIS 3.34.11.

Desse modo para a identificação dos agrupamentos espaciais nos dados, essa ferramenta é relevante, pois pode complementar o Índice de Moran Global, com o uso do Índice de Moran Local, que fornece as associações significativas, entre frequência de sinistros e variáveis, em nível local, de acordo com a gravidade do acidente rodoviário (LIU, DAS & KHAN, 2024).

O índice de Moran local é uma decomposição do Índice de Moran Global, possibilitando a construção de um mapa de dependência espacial local. Isso fornece mais detalhes a respeito das características locais, em comparação com o Índice de Moran Global. Esse indicador comumente é nomeado como *Local Indicators of Spatial Association - LISA* (Anselin, 1995). Expressa pela Equação 3:

$$I_i = \frac{\sum w_{ij} z_i z_j}{\sum_{i=1}^n z_i^2} \quad (3)$$

Em que I_i é o Índice de Moran Local, z_i é a diferença entre o atributo local i e a média de todos os atributos, z_j é a diferença entre o valor do atributo local j e a média de todos os atributos, w_{ij} são os pesos ou graus de conectividade atribuídos conforme a relação topológica entre as localidades i e j .

3.3.3 Regressão Geograficamente Ponderada (RGP)

O modelo de Regressão Geograficamente Ponderada (RGP), apresenta-se como uma abordagem alternativa ao método de MQO, permitindo investigar e modelar as variações espaciais complexas, oferecendo ferramentas importantes para a estimativa de parâmetros locais. Em contrapartida o RGP, considera a heterogeneidade espacial presente em cada unidade de agregação, permitindo que a estimativa de parâmetros varie localmente e propiciando resultados mais precisos a respeito das variáveis explicativas selecionadas (SHABRINA, BUYUKLIEVA & MING NG, 2021).

Conforme Brundson, Fotheringham e Charlton (1996), essa abordagem expande os modelos lineares tradicionais, permitindo que os coeficientes variem espacialmente. Estima-se os coeficientes locais para cada localização i , em cada variável independente, a Regressão Geograficamente Ponderada – RGP, pode ser definida por meio da Equação 4 a seguir:

$$y_i = \sum_{j=0}^{j=n} \beta_{ij} \times x_{ij} + \varepsilon_i \quad (4)$$

Em que y_i é a variável dependente para a localidade i x_{ij} são as variáveis independentes e ε_i é o erro, que pode ser independente e normalmente distribuído. Cada coeficiente β_{ij} nesse modelo é estimado para cada unidade espacial i , baseado nas amostras utilizadas. As observações são ponderadas de acordo com a sua proximidade com a localidade i , seguindo uma função de decaimento de distância.

Desse modo, a Regressão Geograficamente Ponderada (RGP) captura as variações espaciais nos coeficientes dos fatores explicativos. O modelo de RGP estima coeficientes locais, permitindo analisar como as relações entre os indicadores presentes nas diferentes áreas de agregação (CHANG CHIEN, CARVER & COMBER, 2020). Neste trabalho foram utilizados o nível de agregação municipal, com os municípios da região Nordeste do Brasil.

Esse modelo considera a distribuição espacial das observações, atribuindo pesos maiores às observações mais próximas. Portanto, são gerados mapas para ilustrar a variação espacial dos coeficientes, permitindo identificar regiões onde determinados fatores, tiveram maior impacto nos acidentes.

O desempenho dos modelos foi avaliado por meio de diversas métricas. No caso do MQO, foram analisados o R^2 ajustado e o Critério de Informação de Akaike

(AIC) e a significância dos coeficientes. O Critério de Informação de Akaike (AIC) auxilia na identificação do melhor ajuste para o modelo, para esse índice, os menores valores indicam melhor ajuste (OSHAN *et al.*, 2019).

Para a RGP, a avaliação incluiu a comparação entre o AIC dos modelos global e local, além da interpretação visual dos mapas gerados. Esse critério é utilizado para comparar os modelos e identificar as melhorias observadas por meio desse índice, demonstrando os ajustes e a eficácia dos modelos utilizados (CHARLTON, FOTHERINGHAM & BRUNSDON, 2009).

A presença da heterogeneidade espacial refere-se às diferenças de comportamento das variáveis, de acordo com a unidade espacial utilizada. Portanto, gera instabilidade aos coeficientes do modelo, violando a suposição da variância constante, assumida pelos modelos lineares (GOMES, 2018).

A heterogeneidade espacial pode ser atribuída a características da estrutura espacial, erros de medição nos dados ou má especificação do modelo, sendo as duas últimas geradoras de dependência espacial. Em contrapartida, o modelo geponderado considera de forma integrada tanto os efeitos da dependência quanto da heterogeneidade espacial (GOMES, 2018).

Sob o contexto da acidentalidade e indicadores socioeconômicos, esse modelo se mostra útil na identificação de *clusters* que permitem indicar a vulnerabilidade social e sua relação com os acidentes rodoviários. Destaca-se ao fornecer informações valiosas sobre a relação entre as desigualdades sociais e os indicadores de acidentes (LI *et al.*, 2022).

3.3.4 Limitações, contribuições e perspectivas da análise espacial

Esses resultados foram comparados para identificar as vantagens e limitações de cada modelo. Enquanto o MQO forneceu uma visão geral das relações, a RGP destacou variações locais importantes. A combinação de abordagens permitiu uma compreensão mais abrangente dos fatores que influenciam os acidentes rodoviários na Região Nordeste.

As análises espaciais de acidentes rodoviários enfrentam desafios que comprometem sua precisão e aplicabilidade. A dependência espacial e a heterogeneidade são frequentemente ignoradas em abordagens mais convencionais. À medida em que não se leva em consideração essas características pode levar a

ocorrência de estimativas enviesadas e a conclusões equivocadas sobre os acidentes. O modelo de RGP fornece um modelo estatístico mais ajustado, que os modelos lineares como o de MQO (ZHENG *et al.*, 2011).

No cenário internacional, modelos avançados, como a Regressão Geograficamente Ponderada de Multiescala (RGPM), apontam como muito eficientes para a análise espacial. Liu, Das e Khan (2024) apontam que essa metodologia supera as abordagens lineares, por considerar os aspectos locais dos fatores contribuintes, seja para variáveis socioeconômicas ou mesmo às características da infraestrutura viária. Esses modelos são eficazes na identificação de pontos críticos, que devem passar por intervenções e na proposição de políticas públicas de segurança viária.

Vale destacar, por outro lado, que a literatura identificou, nos modelos baseados em Poisson, a superdispersão dos dados. Esse fenômeno contribui para resultados enviesados. Logo, pontua-se que a Regressão Binomial Negativa Geograficamente Ponderada (RBNGP) é o modelo mais adequado, pois considera em sua equação, o parâmetro da superdispersão dos dados. O uso da regressão de Poisson não apresenta em geral estimativas imprecisas, porém é possível que esse método subestime a variância dos coeficientes, produzindo mais variáveis significativas (LORD & MANNERING, 2010; XU & HUANG, 2015). A forma geral de um modelo de RBNGP é a da Equação 5 abaixo:

$$y_j \sim NB \left[t_j \exp \left(\sum_k \beta_k (u_j, v_j) x_{jk} \right), \alpha (u_j, v_j) \right] \quad (5)$$

Em que (u_j, v_j) são as coordenadas de localização dos pontos dos dados j , para $j=1, \dots, n$, t_j é a variável offset, α é o parâmetro de superdispersão, β_k são os coeficientes relacionados às variáveis explicativas x_k , para $k=1, \dots, k$, e y_j é a j -ésima variável dependente.

O trabalho pioneiro no Brasil, de Gomes (2018) destaca em profundidade a temática, enfatiza a falta de estudos que explorem a aplicabilidade da metodologia de Regressão Binomial Negativa Geograficamente Ponderada (RBNGP) a escalas maiores. Segundo o autor, essa ampliação de escala é essencial para comprovar a robustez do modelo na análise espacial de acidentes, levando em consideração a superdispersão observada nos dados. Também ressalta a importância de aplicar essas ferramentas em diferentes contextos e sob outros aspectos, aumentando sua utilidade para modelos macroscópicos e destacando a necessidade de considerar o volume de tráfego, dado sua relevância para esse tipo de análise.

A escolha desses modelos foi pautada pela necessidade de capturar diferentes aspectos das relações entre as variáveis explicativas e a ocorrência de acidentes rodoviários. O MQO permite obter-se uma visão inicial e global das associações entre as variáveis, identificando fatores mais relevantes, porém a sua principal limitação está em não considerar os aspectos espaciais, O Índice de Moran demonstra a presença, em alguns modelos, de autocorrelação espacial, o que destaca a importância da RGP para capturar essas variações espaciais. O modelo de RGP permite capturar nuances locais que seriam mascaradas em uma abordagem global, revelando como as variáveis explicativas selecionadas influenciam os sinistros rodoviários de maneira diferente em diferentes regiões. A utilização desses modelos proporcionou uma análise robusta, integrando perspectivas globais e locais, levando à compreensão das dinâmicas espaciais dos acidentes rodoviários.

Além disso o autor também destaca que a presença de bases de dados incompletas e a limitação a modelos globais, apresentam-se como uma barreira significativa para a realização de análises mais robustas, sob a perspectiva de uma análise espacial.

Portanto, a partir das limitações e lacunas de pesquisa pontuadas por meio da literatura atualizada e de acordo com o estado da arte no tema. Esta dissertação busca implementar métodos avançados de análise espacial, de modo a propor uma análise em uma escala mais ampla, em nível regional, alinhando-se às melhores pesquisas internacionais na temática, adaptando a metodologia às especificidades locais, de modo a prover subsídios para a implementação de políticas públicas mais eficientes na segurança viária e contribuir com o avanço científico na temática no Brasil.

3.4 AVALIAÇÃO DOS RESULTADOS: RELAÇÃO ENTRE VARIÁVEIS SELECIONADAS E ACIDENTES RODOVIÁRIOS

Neste tópico, descrevem-se os procedimentos adotados para avaliar os resultados obtidos por meio da modelagem estatística. A análise concentra-se em compreender o impacto das variáveis independentes nos modelos observados e de que modo os padrões espaciais presentes nos dados influenciaram os modelos aplicados.

A avaliação dos coeficientes estimados pelos modelos permite observar a influência de cada variável explicativa na ocorrência de acidentes rodoviários. O foco

está em interpretar a magnitude e o sentido desses coeficientes, fornecendo subsídios para identificar os fatores mais relevantes e verificar sua relação estatística com a variável dependente. Dessa forma, as variáveis são avaliadas de acordo com seus grupos: Acidentes rodoviários; Infraestrutura viária e territorial; Indicadores socioeconômicos e demográficos; características relacionadas à frota e condutores. Essa maneira de sistematizar a avaliação dos resultados contribui para destacar a relevância de cada grupo de variáveis separadamente com relação aos indicadores de acidentes da Região Nordeste.

A validação desses resultados se dá por meio da comparação entre diferentes métricas estatísticas dos modelos como o Critério de Informação de Akaike (AIC) e indicadores de desempenho do modelo. Desse modo, é possível compreender se os coeficientes apresentam coerência e se as interpretações extraídas a respeito das variáveis estão alinhadas com a literatura e o modelo como um todo.

4 RESULTADOS E DISCUSSÃO

Esta seção apresenta os resultados das análises realizadas, destacando as principais relações entre as variáveis explicativas e os acidentes rodoviários na Região Nordeste do Brasil. Para proporcionar uma visão mais ampla, os resultados estão organizados de acordo com a seguinte estrutura, iniciando com a análise descritiva dos dados, seguida pela avaliação das correlações e da multicolinearidade entre as variáveis.

Na sequência, são apresentados os modelos de Mínimos Quadrados Ordinários (MQO), seguidos pela aplicação do Índice de Moran Global e Local, utilizado para identificar padrões espaciais significativos e investigar a autocorrelação espacial nos dados. Por fim, a Regressão Geograficamente Ponderada (RGP) é aplicada, possibilitando uma análise global das variáveis, assim como uma interpretação detalhada das dinâmicas espaciais.

4.1 ESTATÍSTICAS DESCRITIVAS E MATRIZ DE CORRELAÇÃO

4.1.1 Estatísticas descritivas das variáveis

As estatísticas descritivas contribuem para demonstrar a variabilidade dos dados, por meio das medidas de posição, tendência central e de dispersão. Sendo possível compreender de forma prática as características dos dados de acidentes, para cada variável, dependentes e independentes. A presença de outliers também foi analisada para aferir a variabilidade e as características dos dados.

A Tabela 3 apresenta a descrição de cada uma das variáveis dependentes e independentes.

Tabela 3 – Descrição das variáveis dependentes e independentes

Variáveis	Descrição
Acidentes rodoviários	
ACID_TOTAL	Total de Acidentes
ACID_KM	Acidentes por km de via
ACID_MI_V	Acidentes por milhão de veículos do tráfego
SEM_VIT	Acidentes sem vítimas
VIT_FERIDA	Acidentes com vítimas feridas
VIT_FATALIS	Acidentes com vítimas fatais
ACID_URB	Acidentes urbanos
ACID_RUR	Acidentes rurais
MOTO	Acidentes envolvendo motociclistas
CICLI	Acidentes envolvendo ciclistas
IDADE_MED	Idade média dos condutores
ANO_VEIC	Ano de fabricação do veículo
IDADE_VEIC	Idade do veículo
Infraestrutura viária e territorial	
EXT_VIAS	Extensão das vias
VMDA	Volume Médio Diário Anual
DENS_VIA	Densidade de vias
AREA_UN	Área da unidade territorial (km ²)
Indicadores socioeconômicos e demográficos	
PER_REND_C	Percentual da população com rendimento nominal mensal per capita de até 1/2 salário mínimo
P_POP_OCUP	População ocupada
SALARIO_M	Salário médio mensal dos trabalhadores formais (salários mínimos)
IDHM	Índice de Desenvolvimento Humano Municipal (IDHM)
PIB_P_CPT	PIB per capita (R\$)
TX_ESCOLAR	Taxa de escolarização de 6 a 14 anos de idade
POPULACAO	População no último censo
DENS_DEM	Densidade demográfica (hab/km ²)
Características relacionadas à frota e condutores	
PROP_AT_24	Proporção de condutores até 24 anos
PROP_25_64	Proporção de condutores entre 25 a 64 anos
PROP_65_M	Proporção de condutores a partir de 65 anos
PROP_MOTO	Proporção de motocicletas da frota
HAB_VEI	Habitantes/veículo

Fonte: O autor (2025).

A Tabela 4 apresenta as estatísticas descritivas das variáveis dependentes, contando com medidas de posição (quartis, mínimo e máximo), dispersão (desvio padrão e coeficiente de variação) e tendência central (média e mediana).

Tabela 4 – Estatísticas descritivas das variáveis dependentes e independentes

Variáveis	Mín.	Q1	Mediana	Média	Q3	Máx.	Des. Pad.	CV
Acidentes rodoviários								
ACID_TOTAL	1	4	8	21	19,75	436	43	2,057
ACID_KM	0,002	0,179	0,369	0,845	0,858	40	2,306	2,356
ACID_MI_V	0,292	2,959	5,922	11,404	12,406	204,621	17,547	1,539
SEM_VIT	0	0	1	4	4	86	8	2,218
VIT_FERIDA	0	2	5	15	13	337	33	2,218
VIT_FATAIS	0	0	1	2	3	23	3	1,250
ACID_URB	0	1	3	9	9	222	20	2,170
ACID_RUR	0	2	4	11	11	204	21	1,960
MOTO	0	0	0	1,22	1	57	3,88	3,185
CICLI	0	0	0	0,67	0	21	2,24	3,362
IDADE_MED	0	35,74	38,77	38,66	42	73	7,28	0,377
ANO_VEIC	1980	2010	2012	2011,75	2014	2022	3,46	0,002
IDADE_VEIC	1	8,441	9,839	10,25	11,5	42	3,46	0,338
Infraestrutura viária e territorial								
EXT_VIAS	1	13,76	22,33	29,94	36,45	305,80	27,75	0,927
VMDA	628	2495,1	3213,5	5310,68	5465,7	58333	6232,01	1,173
DENS_VIA	0	0,025	0,042	0,051	0,064	0,502	0,042	0,814
AREA_UN	26	262,1	535,1	1031,664	1199,1	15634,320	1568,763	1,521
Indicadores socioeconômicos e demográficos								
PER_REND_C	0	0,4682	0,502	0,498	0,531	0,613	0,050	0,100
P_POP_OCUP	0	0,086	0,114	0,134	0,162	0,701	0,073	0,547
SALARIO_M	1	1,6	1,8	1,860	2	6,90	0,392	0,211
IDHM	0	0,57	0,601	0,605	0,631	0,772	0,048	0,079
PIB_P_CPT	5732,2	10602	13253	18656,96	18876	210498,16	20186,90	1,082
TX_ESCOLAR	1	0,962	0,972	0,967	0,978	0,996	0,019	0,020
POPULACAO	1318	10584	19809	52894	38741	2428708	169653	3,207
DENS_DEM	1	16,88	35,45	144,760	81,44	7775,520	568,014	3,924
Características relacionadas à frota e condutores								
PROP_AT_24	0	0,242	0,321	0,329	0,417	1,000	0,164	0,498
PROP_25_64	0	0,533	0,632	0,626	0,711	1,000	0,160	0,256
PROP_65_M	0	0	0,031	0,046	0,066	0,500	0,062	1,359
PROP_MOTO	0	0	0,436	0,451	0,541	0,780	0,117	0,259
HAB_VEI	0	2,657	3,556	4,793	4,641	99,741	7,197	1,501

Fonte: O autor (2025).

Os resultados evidenciam a heterogeneidade dos dados e das variáveis a eles relacionados. Os acidentes totais (ACID_TOTAL) varia entre 1 a 436, com uma média de 21 e desvio padrão de 43, indicando que boa parte das localidades possuem números baixos de acidentes, mas também há municípios com alta concentração de sinistros rodoviários. Esse comportamento também se reproduz nas demais variáveis relacionadas à acidentes, seja sem vítimas (SEM_VIT, vítimas feridas (VIT_FERIDA) e fatais (VIT_FATAIS).

No que diz respeito às variáveis relacionadas à infraestrutura viária e territorial, as variáveis extensão das vias (EXT_VIAS), volume de tráfego (VMDA), densidade das vias (DENS_VIA) e área dos municípios (AREA_UN), apresentam alta variabilidade. As medidas de tendência central e de variabilidade, evidenciam a heterogeneidade das observações. As diferenças presentes entre os municípios estão relacionadas aos fluxos de tráfego, às regiões mais densamente povoadas ou com maior malha viária.

Esses fatores contribuem para a variabilidade registrada por essas variáveis e são um indicador da exposição de usuários das vias aos riscos de acidentes ou fatalidades. Considerando outro aspecto dessa exposição, Nantulya e Reich (2003) pontuam sua relação com a vulnerabilidade social observada.

Os indicadores socioeconômicos também destacam as desigualdades da Região Nordeste. O salário médio dos trabalhadores formais (SALARIO_M) apresenta alta variabilidade, com variação entre 1 a 6,9 salários mínimos por municípios e média de 1,86. O percentual da população com renda mensal per capita de até 1/2 salários mínimos (PER_REND_C), com média de 49,8%, indicador das desigualdades sociais presentes nesses municípios.

Desse modo, compreende-se que essas disparidades podem influenciar os padrões de accidentalidade, especialmente. A desigualdade social, discutida por Mesa-Arango *et al.* (2018), presente em países em desenvolvimento apresenta relação com os acidentes rodoviários. No que diz respeito a áreas com uma atividade econômica intensa, a tendência observada foi de redução dos índices.

Os indicadores da frota e dos condutores apresentam variáveis que revelam informações importantes sobre esses sinistros. A proporção de motocicletas na frota (PROP_MOTO), com média de 45,1%, evidencia a forte presença de motocicletas, em alguns municípios com valor máximo de 78%. A gravidade dos sinistros pode ser influenciada pela presença extensiva das motocicletas na composição da frota corroborando com o apontado por Barroso Júnior, Bertho e Veiga (2016). A idade média da frota de veículos (IDADE_VEIC), de 10,25 anos, reflete uma frota relativamente antiga, o que pode contribuir para falhas mecânicas ou a probabilidade dos acidentes.

As variáveis relacionadas às contagens dos acidentes rodoviários apresentam alta variabilidade. Verifica-se a presença de máximos muito elevados com relação ao Q3, principalmente nas variáveis: acidentes totais (ACID_TOTAL), acidentes por

quilômetro (ACID_KM) e acidentes urbanos (ACID_URB). Essa diferença elevada sugere a presença de *outliers* ou alta assimetria nos modelos.

Outras variáveis com escalas distintas, como por exemplo: o volume médio diário de tráfego (VMDA), população (POPULACAO), densidade demográfica (DENS_DEM), apresentam uma variabilidade muito grande. Além da presença de valores máximos elevados em comparação ao Q3, a presença de um coeficiente de variação elevado indica variação, influenciada pelas características próprias das variáveis.

Enquanto isso, algumas variáveis acabam sendo mais compactas, demonstrando uma menor variabilidade. O ano de fabricação dos veículos (ANO_VEIC), percentual da população de baixa renda (PER_REND_C), salário médio dos trabalhadores formais (SALARIO_M), taxa de escolarização (TX_ESCOLAR) e Índice de Desenvolvimento Humano (IDHM), apresentam os intervalos mais curtos, devido as características desses fatores, levando a uma menor variabilidade.

A Tabela 5 apresenta os *outliers* que compõe as variáveis dos modelos analisados, essa tabela conta com a contagem de *outliers*, o percentual com relação a totalidade dos dados, a média dos valores discrepantes para cada variável, a média das variáveis com e sem os *outliers*.

Tabela 5 – Outliers das variáveis dependentes e independentes.

Variáveis	Outliers	Outliers (%)	Média dos outliers	Méd. da var. com out.	Méd. da var. sem out.
<i>Acidentes rodoviários</i>					
ACID_TOTAL	65	9,94	116	20,9	10,4
ACID_KM	69	10,6	5,38	0,979	0,46
ACID_MI_V	64	9,79	50,6	11,4	7,15
SEM_VIT	45	6,88	26,6	3,59	1,89
VIT_FERIDA	69	10,6	85,7	15,1	6,77
VIT_FATAIS	32	4,89	10,9	2,17	1,72
ACID_URB	56	8,56	59,8	9,41	4,69
ACID_RUR	62	9,48	60,4	10,9	5,75
MOTO	71	10,9	8,48	1,22	0,333
CICLI	144	22	3,03	0,667	0
IDADE_MED	53	8,1	44,4	39,4	38,9
ANO_VEIC	42	6,42	2007	2012	2012
IDADE_VEIC	42	6,42	15,1	10,2	9,91
<i>Infraestrutura viária e territorial</i>					
EXT_VIAS	44	6,73	106	29,9	24,4
VMDA	60	9,17	21376	5311	3688
DENS_VIA	25	3,82	0,195	0,051	0,0453
AREA_UN	53	8,1	5025	1032	680
<i>Indicadores socioeconômicos e demográficos</i>					
PER_REND_C	10	1,53	0,351	0,498	0,5
P_POP_OCUP	35	5,35	0,353	0,134	0,122
SALARIO_M	26	3,98	3,08	1,86	1,81
IDHM	12	1,83	0,756	0,605	0,602
PIB_P_CPT	58	8,87	64799	18657	14167
TX_ESCOLAR	35	5,35	0,91	0,967	0,971
POPULACAO	70	10,7	301737	52894	23067
DENS_DEM	68	10,4	1009	145	44,5
<i>Características relacionadas à frota e condutores</i>					
PROP_AT_24	14	2,14	0,861	0,329	0,317
PROP_25_64	42	6,42	0,753	0,626	0,617
PROP_65_M	31	4,74	0,246	0,0458	0,0358
PROP_MOTO	0	0	NaN	0,451	0,451
HAB_VEI	45	6,88	21,4	4,79	3,57

Fonte: O autor (2025).

Algumas variáveis com maiores escalas, apresentaram *outliers* que aumentaram significativamente a média das variáveis, em comparação com a média das variáveis desconsiderando os valores discrepantes. Esse fator observa-se principalmente na variável de frequência de acidentes (ACID_TOTAL), volume de tráfego (VMDA), população total (POPULACAO), PIB per capita (PIB_P_CPT) e densidade demográfica (DENS_DEM). Essas variáveis têm como característica a

presença de altos valores, situadas principalmente nos grandes centros, que apresentam muitos registros, maior população e fluxo.

Em contrapartida, vale destacar os comportamentos de outras variáveis com relação a presença dos valores discrepantes. A única variável sem a presença de *outliers* foi a variável proporção de motocicletas. Outras variáveis, relacionadas aos veículos, às características socioeconômicas, a frota e condutores, não apresentaram muitos *outliers* e conseqüentemente menos variabilidade, com menor influência na média.

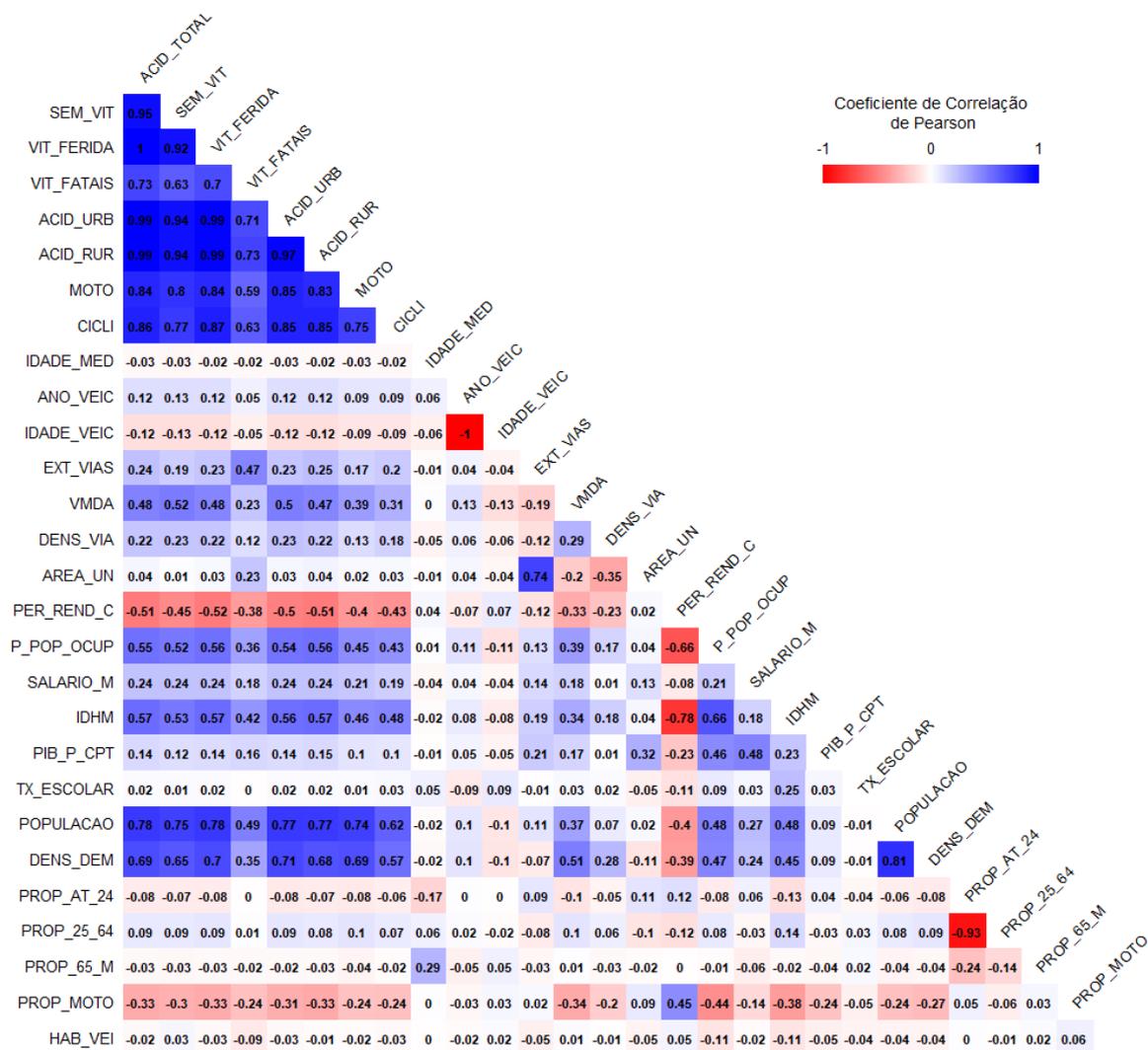
Portanto, essas observações reforçam a complexidade dos fatores que influenciam os sinistros viários. Desse modo, verifica-se a diferença de escala e a presença de variáveis presentes entre as variáveis do modelo, essa diferença conduz a reflexões a respeito da necessidade de adaptação nessas variáveis.

4.1.2 Matriz de Correlação de Pearson

No âmbito das análises a respeito das características dos dados e dos fatores utilizados no modelo, a matriz de correlação de Pearson auxilia na identificação das relações lineares entre os indicadores. De acordo com o grupo de variável e avaliando potenciais problemas de multicolinearidade que são verificados posteriormente pelo Teste de Fator de Inflação da Variância (VIF). Essa etapa permitiu compreender quantitativamente a magnitude e o sentido das associações, facilitando a seleção das variáveis a serem utilizadas nos modelos estatísticos. A Figura 3 apresenta os resultados obtidos na Matriz de Correlação Linear de Pearson, evidenciando os coeficientes obtidos e as relações entre as variáveis.

Para analisar os coeficientes de correlação das variáveis, é essencial observar as relações entre as variáveis explicativas e as variáveis de resposta, que auxiliam na identificação das relações entre estas e os acidentes rodoviários. Posteriormente, a análise dos coeficientes de correlação entre as variáveis explicativas poderá dar um indicativo a respeito da multicolinearidade entre elas.

Figura 3 – Matriz de Correlação Linear de Pearson



Fonte: O autor (2025).

O grupo de variáveis relacionadas aos acidentes rodoviários, dentre os quais as variáveis explicativas: idade média dos condutores (IDADE_MED), ano de fabricação do veículo (ANO_VEIC) e idade do veículo (IDADE_VEIC); apresentaram coeficientes de correlação baixos, com sentido positivo para ano do veículo e negativo para os demais, indicando pouca relação com as variáveis dependentes.

As variáveis de infraestrutura viária e territorial: extensão de vias (EXT_VIAS), volume de tráfego (VMDA), densidade de vias (DENS_VIA) e área do município (AREA_UN); apresentaram correlação moderada com as variáveis resposta, sempre positivas, o que indica que uma maior infraestrutura de vias, unidade territorial maior e maior fluxo de tráfego, estão associados a um aumento da ocorrência de acidentes, o que é coerente com Su, Sze e Bai (2021), que apontam que o maior fluxo de tráfego e infraestrutura viária contribuem para a ocorrência de mais acidentes.

Com relação aos indicadores socioeconômicos e demográficos, foram observados coeficientes variados. As variáveis demográficas, como população e densidade demográfica, apresentaram forte correlação ($r > 0,69$) com o total de acidentes. Esse resultado está alinhado com a evidência documentada na literatura, que indica que o aumento populacional está associado ao crescimento no número de registros de acidentes rodoviários (WANG, HUANG & ZENG, 2017).

Entre as variáveis socioeconômicas, o percentual da população ocupada, o salário médio dos trabalhadores formais, o Índice de Desenvolvimento Humano Municipal (IDHM) e o PIB per capita (PIB_P_CPT) apresentaram correlações positivas moderadas. Esses resultados sugerem que populações mais ativas economicamente tendem a apresentar maior frequência de acidentalidade, em conformidade com Kim, Pant e Yamashita (2010), também apontando que os grandes centros onde estão localizados mais empregos, tendem a concentrar mais acidentes que zonas residenciais.

A taxa de escolarização (TX_ESCOLAR) apresentou correlação positiva muito fraca. Por outro lado, o percentual da população com rendimento nominal mensal per capita de até 1/2 salário mínimo (PER_REND_C) exibiu uma correlação negativa, o que indicaria que regiões com maior concentração de pessoas de baixa renda, e conseqüentemente em situação econômica mais vulnerável, tenderiam a apresentar menos registros de acidentes, o que pode demonstrar-se como um contrassenso a literatura, merecendo ser melhor investigado por outros modelos.

Com relação à frota e condutores, a proporção de condutores por idade (PROP_AT_24, PROP_25_64 e PROP_65_M) e número de habitantes por veículo (HAB_VEIC) apresentaram correlação muito fraca, próxima de zero. A proporção de motociclistas (PROP_MOTO) com relação à totalidade de veículos da frota apresentou correlação negativa com as variáveis resposta, o que pode indicar uma relação inversa entre a proporção de motociclistas e a ocorrência de acidentes nas rodovias, o que pode ser um sinal contraditório em comparação com a literatura na temática, em trabalhos como Blazot *et al.* (2013).

No que diz respeito aos coeficientes de correlação entre as variáveis explicativas, algumas observações se destacam. Foram identificadas fortes correlações positivas, como entre extensão de vias (EXT_VIAS) e área municipal (AREA_UN), o que é coerente com as características territoriais, já que municípios maiores tendem a apresentar maior extensão viária. Outro exemplo é a correlação

entre percentual da população ocupada (P_POP_OCUP) e IDHM, indicando uma relação direta entre níveis de desenvolvimento econômico e humano.

Além disso, observou-se uma correlação elevada entre população (POPULACAO) e densidade demográfica (DENS_DEM) ($r = 0,81$), que sugere a possibilidade de multicolinearidade entre essas variáveis. Correlações positivas moderadas também foram verificadas entre as variáveis populacionais e algumas variáveis socioeconômicas e de infraestrutura viária, como percentual da população ocupada (P_POP_OCUP), salário médio dos trabalhadores formais (SALARIO_M), IDHM, VMDA e densidade de vias (DENS_VIA), reforçando a interdependência esperada entre esses indicadores.

Entre as correlações negativas, destacaram-se a relação perfeita e inversa entre idade do veículo (IDADE_VEIC) e ano de fabricação do veículo (ANO_VEIC) ($r = -1$), evidenciando que uma variável é diretamente consequência da outra. Da mesma forma, a proporção de condutores entre 25 e 64 anos (PROP_25_64) apresentou uma forte correlação negativa com a proporção de condutores até 24 anos (PROP_AT_24) ($r = -0,93$), o que reflete a interdependência entre essas variáveis, dada a natureza complementar das faixas etárias. Essas correlações negativas e positivas elevadas entre variáveis explicativas indicam possíveis problemas de multicolinearidade, que serão avaliados com maior rigor por meio do Teste de Fator de Inflação da Variância (VIF).

4.2 TESTE DE FATOR DE INFLAÇÃO DA VARIÂNCIA (VIF)

O Teste de Fator de Inflação da Variância foi aplicado para verificar a existência de multicolinearidade entre as variáveis explicativas utilizadas. A Tabela 6 demonstra os valores obtidos por meio desse teste, para todas as variáveis utilizadas.

Tabela 6 – Teste de Fator de Inflação da Variância (VIF) das variáveis explicativas

Variáveis Explicativas	VIF
IDADE_MED	1,124
ANO_VEIC	1,61E+09
IDADE_VEIC	1,61E+09
EXT_VIAS	2,716
VMDA	1,600
DENS_VIA	1,486
AREA_UN	3,008
PER_REND_C	3,112
P_POP_OCUP	2,825
SALARIO_M	1,508
IDHM	3,467
PIB_P_CPT	2,060
TX_ESCOLAR	1,145
POPULACAO	3,776
DENS_DEM	4,154
PROP_AT_24	1,61E+05
PROP_25_64	1,54E+05
PROP_65_M	2,33E+04
PROP_MOTO	1,390
HAB_VEI	1,027

Fonte: O autor (2025).

Para analisar os valores de VIF das variáveis explicativas, procurou-se levar em consideração o destacado por O'Brien (2007), que apontou para a importância de não utilizar técnicas mecânicas de corte, para retirada de variáveis, mas de fazer uma interpretação do VIF de maneira contextualizada, levando em conta outros fatores e compreendendo que há variáveis que podem ser transformadas, ou mesmo haver a retirada de somente parte das variáveis com altos valores, verificando posteriormente a presença de multicolinearidade.

Levando em consideração os elevados valores do índice para as variáveis ano de fabricação do veículo e idade do veículo, verificou-se que as duas variáveis, ratificando o observado na matriz de correlação, representam variáveis perfeitamente inversas, portanto, optou-se por retirar o ano de fabricação do veículo e manter a idade do veículo, partindo do pressuposto que com essa retirada eliminasse a multicolinearidade.

Na proporção da idade dos condutores, de acordo com as faixas etárias, até 24 anos (PROP_AT_24), entre 25 e 64 (PROP_25_64) e a partir de 65 anos (PROP_65_M), observou-se alta multicolinearidade, devido serem variáveis correlacionadas entre si, desse modo foi mantida a variável proporção de condutores

até 24 anos (PROP_AT_24), excluindo-se as outras faixas, para os modelos subsequentes. A Tabela 7 abaixo, destaca os valores do VIF após a retirada das três variáveis destacadas.

Tabela 7 – Fator de Inflação da Variância (VIF) das variáveis explicativas após retirada de variáveis

Variáveis Explicativas	VIF
IDADE_MED	1,060
IDADE_VEIC	1,055
EXT_VIAS	2,921
VMDA	1,654
DENS_VIA	1,649
AREA_UN	3,071
PER_REND_C	3,107
P_POP_OCUP	2,832
SALARIO_M	1,638
IDHM	3,403
PIB_P_CPT	1,934
TX_ESCOLAR	1,157
POPULACAO	3,639
DENS_DEM	3,978
PROP_AT_24	1,092
PROP_MOTO	1,429
HAB_VEI	1,023

Fonte: O autor (2025).

Desse modo, verificou-se os valores do VIF após a retirada das três variáveis, atestando que não havia mais multicolinearidade entre os fatores, todas permanecendo abaixo de 5, concluindo que as variáveis em questão poderiam ser utilizadas para o modelo de MQO.

4.3 MÍNIMOS QUADRADOS ORDINÁRIOS (MQO)

A modelagem estatística por meio de Mínimos Quadrados Ordinários foi fundamental para identificar as relações globais entre as variáveis explicativas e cada uma das variáveis dependentes nos modelos elaborados. Para esse modelo, foram utilizadas as variáveis destacadas na Tabela 7, como variáveis explicativas. As variáveis dependentes, foram relacionadas aos acidentes totais, as taxas de acidentes (por extensão viária e tráfego), a localização dos acidentes (urbanos e rurais) a gravidade dos acidentes (sem vítimas, com vítimas feridas e fatais).

Esta seção irá abordar os resultados dos indicadores de desempenho de cada modelo utilizado, a análise dos coeficientes obtidos e testes dos pressupostos do modelo linear. Logo, pretende-se identificar as variáveis que mais impactaram nos acidentes registrados nos municípios da região Nordeste do Brasil.

4.3.1 Modelo preliminar de Mínimos Quadrados Ordinários para a variável acidentes totais (ACID_TOTAL)

Inicialmente, utilizou-se de um modelo preliminar empregado para verificar a eficiência e a significância estatística para a variável dependente acidentes totais (ACID_TOTAL). O modelo apresentou as métricas de desempenho de acordo com a Tabela 8.

Tabela 8 – Principais métricas indicadoras de performance do modelo MQO.

Indicadores	ACID_TOTAL
R ²	0,7294
R ² (ajustado)	0,719
RSE	23,48 com 439 DF
p-valor	<2,2e-16
AIC	4201,35
MAE	12,30676
MSE	408,3077
R ² (teste)	0,7370998

Fonte: O autor (2025).

O valor do R², considera que: entre 0,00 a 0,30 o modelo explica pouco a variabilidade dos dados (fraco); 0,30 a 0,60 o modelo tem uma capacidade explicativa moderada; 0,60 a 0,90 o modelo explica bem os dados e 0,90 a 1,00 o modelo explica quase toda a variação (podendo indicar *overfitting* ou sobreajuste) (OSHAN *et al.*, 2019). O resultado obtido demonstrou um bom ajuste do modelo preliminar, o R² ajustado foi próximo do valor do R², demonstrando que as variáveis explicativas são em sua maioria relevantes e o modelo não foi penalizado pela presença variáveis redundantes.

O erro padrão da estimativa (RSE) é considerado baixo, dado a escala da variável dependente, demonstrando que o modelo apresenta boa capacidade preditiva. O valor da estatística F e do p-valor apontaram para afirmação de que o

modelo em questão demonstrou ser globalmente significativo. O valor do coeficiente AIC poderá ser comparado com os modelos subsequentes.

Para o modelo preliminar, verificou-se também os coeficientes obtidos para as variáveis em questão, contando com os valores dos coeficientes, erro, t-valor e p-valor para cada variável, conforme a Tabela 9.

Tabela 9 – Coeficientes das variáveis do modelo preliminar de MQO

Modelo Variáveis	ACID_TOTAL			
	Coeficiente	Erro	t-valor	p-valor
(Intercepto)	14,900	59,620	0,250	0,8027
EXT_VIAS	0,568	0,069	8,271	1,60E-15 ***
VMDA	0,002	0,000	7,097	5,17E-12 ***
IDADE_MED	-0,044	0,103	-0,427	0,6692
IDADE_VEIC	-0,151	0,310	-0,488	0,6257
PROP_AT_24	-0,995	6,962	-0,143	0,8865
PER_REND_C	-65,670	38,470	-1,707	0,0885 .
P_POP_OCUP	46,760	24,790	1,886	0,0600 .
POPULACAO	0,00013	0,00001	11,925	<2E-16 ***
DENS_DEM	-0,005	0,004	-1,322	0,1869
SALARIO_M	1,439	3,406	0,422	0,6729
IDHM	82,650	41,600	1,987	0,0476 *
PIB_P_CPT	-0,00011	0,00008	-1,344	0,1798
TX_ESCOLAR	-39,930	57,990	-0,689	0,4914
AREA_UN	-0,006	0,002	-3,954	8,95E-05 ***
PROP_MOTO	-9,446	11,260	-0,839	0,4019
DENS_VIA	54,030	33,910	1,593	0,1119
HAB_VEI	0,020	0,182	0,107	0,9145

Legenda: . significância ao nível 0,10

* significância ao nível de 0,05

** significância ao nível de 0,01

*** significância ao nível de 0,001

Fonte: O autor (2025).

As variáveis Extensão de vias e VMDA (Volume Médio Diário Anual), destacaram-se como altamente significativas, dado o p-valor obtido e coeficiente, demonstraram a associação positiva dessas variáveis, demonstrando que um aumento da infraestrutura viária está correlacionado ao crescimento no número de registros de acidentes, do mesmo modo, o VMDA demonstrou impacto positivo, refletindo um o aumento da acidentalidade em locais com maior fluxo.

A variável população (POPULACAO) também mostrou forte significância, com coeficiente positivo que, embora pequeno reflete a escala da população e sua relação proporcional com os acidentes rodoviários, indicando que quanto maior a população, maior a acidentalidade. A variável referente à área do município (AREA_UN),

apresentou relação inversa significativa com a variável dependente, o que sugere que municípios com maior área tendem a registrar menos acidentes, possivelmente devido a menor densidade viária, que reduz as interações entre os veículos. Esses resultados reforçam que fatores infraestruturais, territoriais e demográficos são relevantes na explicação da acidentalidade.

Com relação às variáveis socioeconômicas, o Índice de Desenvolvimento Humano Municipal (IDHM), apresentou associação positiva e significativa com a variável dependente, sugerindo que regiões com um maior IDHM, por apresentar maior mobilidade e atividade econômica, podem estar mais expostas a acidentes rodoviários.

Outras variáveis como percentual da população com rendimento nominal mensal per capita de até 1/2 salário mínimo (PER_REND_C) e percentual da população ocupada (P_POP_OCUP) apresentaram significância marginal, a 10%, os sinais dos coeficientes indicaram que quanto maior o percentual de famílias com baixa renda per-capita, menor os registros de acidentes, por outro lado, quanto maior o percentual da população ocupada, mais acidentes tenderiam a ser registrados. Essas relações, que podem aparentar serem contraditórias, podem ser devido ao fato de a concentração de empregos estar em regiões com maior acidentalidade, logo, outras regiões afastadas desses centros, com maiores percentuais de população de baixa renda, possuem menores índices.

Em contrapartida, outras variáveis não apresentaram significância estatística, sugerindo baixa contribuição para o modelo. Desse modo, foram retiradas para os modelos subsequentes as variáveis: idade média dos condutores (IDADE_MED); idade média dos veículos (IDADE_VEIC), salário médio dos trabalhadores formais (SALARIO_M), PIB per capita (PIB_P_CPT) e taxa de escolarização de 6 a 14 anos de idade (TX_ESCOLAR).

Por outro lado, algumas variáveis não significativas, como proporção de condutores até 24 anos de idade (PROP_AT_24), densidade demográfica (DENS_DEM), proporção de motociclistas da frota (PROP_MOTO), densidade de vias (DENS_VIA) e razão habitantes/veículo (HAB_VEI), foram mantidas nos modelos subsequentes, por possuírem relevância teórica na literatura como em Guerra, Dong e Kondo (2019), Zafri e Khan (2022) e Liu, Das e Khan (2024), podendo ter sido comprometida, no modelo, por redundância com outras variáveis. O resultado do modelo preliminar forneceu uma base para ajustes no modelo, com a exclusão de

variáveis não significativas e a validação de variáveis marginais. Assim, o modelo ajustado poderá capturar, de forma mais robusta, os fatores associados à acidentalidade rodoviária.

4.3.2 Modelo ajustados de Mínimos Quadrados Ordinários

4.3.2.1 Indicadores de desempenho dos modelos ajustados

Portanto, após a exclusão das variáveis destacadas anteriormente foram executados os modelos ajustados, com doze das dezessete variáveis explicativas presentes no modelo preliminar.

Portanto, foram executados oito modelos com as seguintes variáveis dependentes, respectivamente, acidentes totais (ACID_TOTAL), acidentes por km (ACID_KM), acidentes por milhão de veículos do tráfego (ACID_MI_V), acidentes sem vítimas (SEM_VIT), acidentes com vítimas feridas (VIT_FERIDA), acidentes com vítimas fatais (VIT_FATAIS), acidentes urbanos (ACID_URB), acidentes rurais (ACID_RUR). A Tabela 10 demonstra os principais resultados obtidos para os indicadores dos modelos de MQO.

Tabela 10 – Principais métricas indicadoras de performance do modelo MQO.

Indicadores	ACID_TOTAL	ACID_KM	ACID_MI_V	ACID_URB
R ²	0,7464	0,5597	0,5377	0,7422
R ² (ajustado)	0,7417	0,5514	0,529	0,7374
RSE	21,81 em 641 DF	1,66 em 641 DF	12,05 em 641 DF	10,47 em 641 DF
p-valor	<2,2e-16	<2,2e-16	<2,2e-16	<2,2e-16
AIC	5902,701	5902,701	5126,688	4943,186
MAE	12,223	0,599	6,484	5,992
MSE	399,529	6,475	77,798	106,697
R ² (teste)	0,7427	0,3444	0,3969	0,7297
Indicadores	ACID_RUR	SEM_VIT	VIT_FERIDA	VIT_FATAIS
R ²	0,7343	0,6936	0,7418	0,4832
R ² (ajustado)	0,7293	0,6878	0,737	0,4736
RSE	11,15 em 641 DF	4,45 em 641 DF	17,19 em 641 DF	1,967 em 641 DF
p-valor	<2,2e-16	<2,2e-16	<2,2e-16	<2,2e-16
AIC	5025,338	3823,983	5591,333	2755,383
MAE	6,378	2,590	9,288	1,301
MSE	92,551	20,341	232,669	3,343
R ² (teste)	0,7443	0,6235	0,7539	0,4588

Fonte: O autor (2025).

A significância estatística dos modelos foi confirmada, uma vez que todos apresentaram p-valores globais inferiores a $2,2 \times 10^{-16}$, indicando que, coletivamente, as variáveis explicativas incluídas contribuem de forma significativa para explicar a variabilidade das variáveis dependentes.

A análise dos indicadores de desempenho dos modelos de Mínimos Quadrados Ordinários (MQO) revelou variações na capacidade explicativa em função das diferentes variáveis dependentes analisadas. O modelo com maior poder explicativo é o ACID_TOTAL ($R^2 = 0,7464$ e R^2 ajustado = $0,7417$), enquanto VIT_FATAIS tem o pior desempenho ($R^2 = 0,4832$ e R^2 ajustado = $0,4736$). Isso indica que a acidentalidade total é mais previsível do que os acidentes fatais, que podem ter causas mais aleatórias ou não capturadas pelas variáveis do modelo.

Outros modelos com bons ajustes, caracterizados por elevados valores de R^2 e R^2 ajustado ($R^2 > 0,70$) foram os de ACID_URB, ACID_RUR e VIT_FERIDA. Esses resultados sugerem que as variáveis explicativas capturam, de forma eficiente, os fatores associados à ocorrência de acidentes totais, com vítimas feridas e em áreas urbanas e rurais. Também vale destacar o baixo desempenho dos modelos para ACID_KM e ACID_MI_V, com ajustes moderados ou baixos ($R^2 < 0,56$), indicando que as variáveis explicativas para esses modelos conseguiram explicar menos de 56% da variabilidade dos dados. A menor capacidade explicativa nesses casos indica que as variáveis independentes utilizadas não capturam adequadamente todos os fatores que influenciam a variabilidade dessas variáveis dependentes. Para aprimorar esses modelos, é necessário incorporar variáveis adicionais que capturem características específicas relacionadas aos fenômenos em estudo.

As métricas de erro padrão da estimativa (RSE), erro absoluto médio (MAE) e erro quadrático médio (MSE) corroboram as conclusões sobre os ajustes dos modelos. Os modelos de maior precisão foram os de ACID_KM (RSE = 1,66; MAE = 0,599; MSE = 6,475) e VIT_FATAIS (RSE = 1,967; MAE = 1,301; MSE = 3,343). Em contrapartida, os maiores erros foram registrados nos modelos de ACID_TOTAL (RSE = 21,81; MAE = 12,223; MSE = 399,529) e VIT_FERIDA (RSE = 17,19; MAE = 9,288; MSE = 232,669).

Esses resultados indicam que, embora os modelos de ACID_TOTAL e VIT_FERIDA apresentem bom desempenho em termos de R^2 , os altos valores de erro revelam a necessidade de ajustes para melhorar a precisão das previsões. Os erros observados podem estar relacionados às diferenças de escala entre as variáveis, com

a presença de muitos *outliers* em algumas variáveis. A utilização de modelos espaciais multiescalares poderiam diminuir os erros dos modelos lineares (SHABRINA, BUYUKLIEVA & MING NG, 2021).

Os modelos de ACID_KM e VIT_FATAIS, por outro lado, destacaram-se pelos baixos valores de erro, apesar de apresentarem menor capacidade explicativa em termos de R^2 . Isso sugere que, embora esses modelos sejam precisos em suas estimativas, a inclusão de variáveis adicionais ou transformações poderiam melhorar seu poder explicativo.

Para o modelo de Acidentes Totais (ACID_TOTAL) foi executado um modelo ajustado, retirando algumas variáveis que não tiveram significância estatística no primeiro modelo, de modo a verificar se haveria melhoria no modelo, com a expectativa de redução das métricas de erro. A Tabela 11 a seguir apresenta uma comparação entre o modelo anteriormente apresentado e esse modelo ajustado.

Tabela 11 – Métricas indicadoras de performance dos modelos de MQO para a variável ACID_TOTAL, modelo com 12 e com 8 variáveis explicativas.

Indicadores	Preliminar	Ajustado
R^2	0,7294	0,7464
R^2 (ajustado)	0,719	0,7417
RSE	23,48 com 439 DF	21,81 em 641 DF
p-valor	<2,2e-16	<2,2e-16
AIC	4201,35	5902,701
MAE	12,30676	12,223
MSE	408,3077	399,529
R^2 (teste)	0,7370998	0,7427

Fonte: O autor (2025).

Com base nas métricas gerais indicadoras da performance dos modelos, verificou-se com base no R^2 e R^2 ajustado, uma leve diminuição desse indicador, que sugeriu que a redução das variáveis não comprometeu sua capacidade explicativa de maneira significativa. O erro padrão da estimativa (SRE) também foi praticamente constante, sugerindo que a precisão dos modelos foi praticamente inalterada.

O Critério de Informação de Akaike (AIC), que ajuda a mensurar a eficiência do ajuste do modelo, em que os modelos que obtiverem os menores valores representam melhores ajuste. A pequena diminuição do valor de AIC no modelo ajustado, em comparação ao modelo completo, indica que o modelo ajustado é mais eficiente e apresenta melhor qualidade em seu ajuste.

A métrica do R^2 do modelo ajustado, que corresponde ao conjunto de teste, apresentou melhoria de 0,72 para 0,7465, indicando que a simplificação pode ter melhorado ligeiramente a capacidade explicativa do modelo no conjunto de teste. O modelo ajustado também apresentou uma leve melhoria do MAE e do MSE, indicando que a simplificação do modelo resultou em previsões ligeiramente mais precisas.

A análise dos coeficientes das variáveis explicativas permitirá inferir informações importantes sobre os dados de acidentes, no modelo de MQO. As Tabelas 12, 13 e 14 apresentam os coeficientes para os modelos de MQO, para cada variável dependente, esses modelos contam com as doze variáveis independentes selecionadas após os resultados do modelo preliminar.

Tabela 12 – Coeficientes das variáveis dos modelos de MQO: ACID_TOTAL, ACID_KM e ACID_MI_V

Modelo Variáveis	ACID_TOTAL			
	Coeficiente	Erro	t-valor	p-valor
(Intercept)	-26,010	30,270	-0,860	0,3904
EXT_VIAS	0,479	0,050	9,588	<2e-16 ***
VMDA	0,00135	0,00017	7,863	1,59E-14 ***
PROP_AT_24	-1,671	5,308	-0,315	0,753
PER_REND_C	-48,210	29,690	-1,624	0,1049
P_POP_OCUP	29,680	17,450	1,701	0,0895 .
POPULACAO	0,00013	0,00001	13,908	<2e-16 ***
DENS_DEM	0,005	0,003	1,714	0,0869 .
IDHM	68,100	31,380	2,170	3,04E-02 *
AREA_UN	-0,004	0,001	-4,141	3,92E-05 ***
PROP_MOTO	-7,638	8,524	-0,896	0,3706
DENS_VIA	63,420	24,810	2,556	0,0108 *
HAB_VEI	0,193	0,120	1,605	0,109
Modelo Variáveis	ACID_KM			
Coeficiente	Erro	t-valor	p-valor	
(Intercepto)	0,7707	2,311	0,334	0,738835
EXT_VIAS	-0,001478	0,00381	-0,388	6,98E-01
VMDA	9,50E-05	0,000013	7,258	1,14E-12 ***
PROP_AT_24	-0,09204	0,4053	-0,227	0,820403
PER_REND_C	-3,4	2,266	-1,500	0,134008
P_POP_OCUP	0,54	1,332	0,405	0,685363
POPULACAO	-3,05E-06	7,347E-07	-4,154	3,71E-05 ***
DENS_DEM	0,003058	0,0002318	13,189	<2e-16 ***
IDHM	3,023	2,396	1,262	0,207512
AREA_UN	-5,01E-05	0,0000674	-0,742	4,58E-01
PROP_MOTO	-0,6605	0,6508	-1,015	0,310525
DENS_VIA	-7,104	1,894	-3,750	0,000193 ***
HAB_VEI	0,0003654	0,009167	0,040	0,968213
Modelo Variáveis	ACID_MI_V			
Coeficiente	Erro	t-valor	p-valor	
(Intercept)	0,9683	16,72	0,058	0,953844
EXT_VIAS	0,2991	0,02758	10,848	<2e-16 ***
VMDA	-5,31E-04	0,0000947	-5,609	3,03E-08 ***
PROP_AT_24	-0,4291	2,933	-0,146	0,883731
PER_REND_C	-44,93	16,4	-2,739	0,006331 **
P_POP_OCUP	8,53E+00	9,642	0,885	0,376429
POPULACAO	4,98E-05	5,32E-06	9,358	<2e-16 ***
DENS_DEM	-5,74E-03	0,001678	-3,422	0,000662 ***
IDHM	46,87	17,34	2,703	0,007052 **
AREA_UN	-0,001254	0,0004882	-2,568	0,010455 *
PROP_MOTO	-10,21	4,71	-2,167	0,030594 *
DENS_VIA	26,1	13,71	1,903	0,057445 .
HAB_VEI	0,01845	0,06634	0,278	0,781024

Legenda: . significância ao nível 0,10; * significância ao nível de 0,05; ** significância ao nível de 0,01 *** significância ao nível de 0,001

Fonte: O autor (2025).

Tabela 13 – Coeficientes das variáveis dos modelos de MQO: SEM_VIT, VIT_FERIDA e VIT_FATAIS

Modelo Variáveis	SEM_VIT			
	Coeficiente	Erro	t-valor	p-valor
(Intercepto)	-15,02	6,177	-2,432	1,53E-02 *
EXT_VIAS	0,06949	0,01019	6,822	2,08E-11 ***
VM DA	3,41E-04	3,50E-05	9,751	<2e-16 ***
PROP_AT_24	-0,1726	1,083	-0,159	0,873437
PER_REND_C	4,018	6,059	0,663	5,07E-01
P_POP_OCUP	8,014	3,562	2,250	0,024773 *
POPULACAO	2,74E-05	1,96E-06	13,964	<2e-16 ***
DENS_DEM	-6,28E-04	6,20E-04	-1,013	0,311563
IDHM	16,27	6,405	2,540	0,011325 *
AREA_UN	-5,29E-04	1,80E-04	-2,934	0,003465 **
PROP_MOTO	-0,4482	1,74	-0,258	0,796775
DENS_VIA	16,59	5,064	3,275	0,001112 **
HAB_VEI	0,08529	0,0245	3,480	0,000535 ***
Modelo Variáveis	VIT_FERIDA			
Coeficiente	Erro	t-valor	p-valor	
(Intercepto)	-13,71	23,85	-0,575	0,56576
EXT_VIAS	0,3569	0,03934	9,073	<2e-16 ***
VM DA	9,53E-04	1,35E-04	7,052	4,59E-12 ***
PROP_AT_24	-1,784	4,184	-0,426	0,6699
PER_REND_C	-46,67	23,4	-1,995	4,65E-02 *
P_POP_OCUP	22,78	13,75	1,656	0,09816 .
POPULACAO	1,00E-04	7,59E-06	13,185	<2e-16 ***
DENS_DEM	0,00625	0,002393	2,611	0,00923 **
IDHM	50,23	24,74	2,031	0,04268 *
AREA_UN	-2,90E-03	6,97E-04	-4,164	3,55E-05 ***
PROP_MOTO	-5,613	6,719	-0,835	0,40381
DENS_VIA	42,78	19,56	2,187	0,02908 *
HAB_VEI	0,1236	0,09464	1,306	0,19189
Modelo Variáveis	VIT_FATAIS			
Coeficiente	Erro	t-valor	p-valor	
(Intercepto)	2,717	2,729	0,996	0,319769
EXT_VIAS	0,05215	0,0045	11,590	<2e-16 ***
VM DA	5,41E-05	1,55E-05	3,498	0,000502 ***
PROP_AT_24	0,2857	0,4786	0,597	0,550745
PER_REND_C	-5,556	2,677	-2,076	0,038297 *
P_POP_OCUP	-1,119	1,573	-0,711	0,477312
POPULACAO	0,00000641	8,676E-07	7,388	4,68E-13 ***
DENS_DEM	-4,16E-04	2,74E-04	-1,518	0,129389
IDHM	1,595	2,829	0,564	0,57305
AREA_UN	-2,30E-04	7,97E-05	-2,884	0,004056 **
PROP_MOTO	-1,578	0,7685	-2,053	0,040492 *
DENS_VIA	4,051	2,237	1,811	0,070621 .
HAB_VEI	-0,01623	0,01083	-1,500	0,134186

Legenda: . significância ao nível 0,10; * significância ao nível de 0,05; ** significância ao nível de 0,01 *** significância ao nível de 0,001

Fonte: O autor (2025).

Tabela 14 – Coeficientes das variáveis dos modelos de MQO: ACID_URB e ACID_RUR

Modelo	ACID_URB				
	Variáveis	Coeficiente	Erro	t-valor	p-valor
(Intercepto)		-14,15	14,53	-0,974	0,330453
EXT_VIAS		0,2222	0,02397	9,271	<2e-16 ***
VMDA		6,91E-04	8,23E-05	8,401	2,86E-16 ***
PROP_AT_24		-1,344	2,549	-0,527	0,598137
PER_REND_C		-20,1	14,26	-1,410	0,15911
P_POP_OCUP		10,96	8,38	1,308	0,191284
POPULACAO		5,73E-05	4,62E-06	12,396	<2e-16 ***
DENS_DEM		0,004983	0,001458	3,417	0,000673 ***
IDHM		31,62	15,07	2,098	0,036256 *
AREA_UN		-1,63E-03	4,24E-04	-3,845	0,000133 ***
PROP_MOTO		-1,571	4,093	-0,384	0,701326
DENS_VIA		27,7	11,92	2,325	0,020408 *
HAB_VEI		0,06701	0,05766	1,162	0,245568

Modelo	ACID_RUR				
	Variáveis	Coeficiente	Erro	t-valor	p-valor
(Intercepto)		-12,3	15,48	-0,795	0,4269
EXT_VIAS		2,45E-01	0,02552	9,602	<2E-16 ***
VMDA		6,40E-04	8,76E-05	7,307	8,16E-13 ***
PROP_AT_24		-0,3663	2,714	-0,135	0,8927
PER_REND_C		-23,81	15,18	-1,568	0,1173
P_POP_OCUP		1,77E+01	8,923	1,987	0,0474 *
POPULACAO		6,66E-05	4,92E-06	13,527	<2E-16 ***
DENS_DEM		0,002048	0,001553	1,319	0,1876
IDHM		3,44E+01	16,05	2,145	0,0323 *
AREA_UN		-0,001877	4,52E-04	-4,154	3,71E-05 ***
PROP_MOTO		-6,004	4,359	-1,377	0,1689
DENS_VIA		30,07	12,69	2,370	0,0181 *
HAB_VEI		0,1179	0,0614	1,920	0,0552 .

Legenda: . significância ao nível 0,10

* significância ao nível de 0,05

** significância ao nível de 0,01

*** significância ao nível de 0,001

Fonte: O autor (2025).

4.3.2.2 Coeficientes das variáveis relacionadas ao tráfego, infraestrutura viária e territorial

As variáveis referentes à infraestrutura viária e territorial, presentes nesses modelos, são: extensão de vias (EXT_VIAS), volume médio diário anual (VMDA), densidade de vias (DENS_VIA) e área da unidade territorial (AREA_UN). Esses fatores apresentaram no geral alta significância estatística para a maior parte dos modelos.

A extensão de vias (EXT_VIA) apresentou relação positiva e estatisticamente significativa em todos os modelos, exceto acidentes por quilômetro de via (ACID_KM), possivelmente devido à relação direta da variável dependente com a extensão das vias. Nos demais modelos a relação positiva indica que a medida que há mais extensão viária, conseqüentemente aumenta a probabilidade de ocorrência de acidentes, variável que pode estar relacionada a uma maior exposição aos riscos para os usuários, em diferentes contextos, de gravidade e localização (urbana e rural) (LI *et al.*, 2022).

A variável que reflete o tráfego, volume médio diário anual (VMDA), apresentou alta significância estatística em todos os modelos, apresentando sinal positivo, demonstrando sua relevância no contexto da acidentalidade. Esse resultado é consistente com o que pontua a literatura, uma vez que, quanto maior o tráfego, maior será o número de registros de acidentes observados (CAI *et al.*, 2016).

A área dos municípios estudados, apresentou coeficientes negativos e estatisticamente significativos em todos os modelos, exceto acidentes por quilômetro de via (ACID_KM). Os coeficientes negativos são um indicativo de que unidades territoriais maiores estão associadas a um menor número de acidentes (WIER *et al.*, 2009), esse resultado pode ser explicado pela tendência de dispersão do tráfego em áreas maiores, reduzindo os pontos críticos que podem gerar colisões. Vale ressaltar uma possível influência pelo fato dos municípios com maiores áreas estarem localizados em regiões com menores fluxos e população. O coeficiente negativo para os registros em áreas rurais reforça essa afirmação, demonstrando que áreas maiores, tendem a concentrar menos registros.

A densidade de vias demonstrou-se significativa estatisticamente e com coeficiente positivo, na maior parte dos modelos, especialmente em acidentes totais (ACID_TOTAL), sem vítimas (SEM_VIT), com vítimas feridas (VIT_FERIDA), urbanos (ACID_URB) e rurais (ACID_RUR). Esse resultado evidencia que áreas com maior densidade de vias, tendem a apresentar maiores índices de acidentes, devido a maior exposição que se registra nessas localidades (LIU, DAS & KHAN, 2024). Desse modo, a presença de mais infraestrutura viária sem um planejamento ou adequada gestão de segurança viária, pode levar a mais acidentes.

4.3.2.3 Coeficientes das variáveis relacionadas aos indicadores socioeconômicos e demográficos

As variáveis referentes aos indicadores socioeconômicos e demográficos presentes nesses modelos, o percentual da população com renda per-capita até um salário mínimo e meio (P_REND_C), percentual da população ocupada (P_POP_OCUP), Índice de Desenvolvimento Humano Municipal (IDHM), população (POPULACAO) e densidade demográfica (DENS_DEM), apresentaram resultados heterogêneos nos modelos analisados, sendo as variáveis população e IDHM as que apresentaram maior significância dentre as desse grupo.

O percentual da população com renda per-capita até um salário mínimo e meio, apresentou resultados heterogêneos nos diferentes modelos, mas demonstrou significância estatística em: acidentes por milhão de veículos do tráfego (ACID_MI_V), vítimas feridas (VIT_FERIDA) e vítimas fatais (VIT_FATAIS). Entretanto, o sinal do coeficiente foi negativo, sugerindo que os municípios com maior percentual de pessoas com baixa renda tendem a apresentarem menores taxas de acidentalidade pelo tráfego e registros graves ou fatais. Esse indicador pode estar relacionado com o fato de que em áreas com maior atividade econômica, que tendem a apresentar menos pessoas de baixa renda, apresentam mais registros de acidentes, enquanto áreas afastadas desses centros econômicos demonstram menos ocorrências (GONZÁLEZ, LAUKAITOU-SIDERIS & CHAPPLE, 2019).

Por outro lado, o percentual da população ocupada, apresentou significância estatística para dois modelos, de acidentes sem vítimas (SEM_VIT) e acidentes rurais (ACID_RUR), os coeficientes foram positivos, indicando que quanto maior o percentual de população em trabalhos formais, conseqüentemente maior o número de acidentes rurais e sem vítimas, o que também pode estar associado ao fato de que quanto maior o percentual de pessoas ocupadas, maior o fluxo de tráfego, aumentando a probabilidade de ocorrência de acidentes. Entretanto, a diferença, em comparação a outros modelos, pode se dever ao identificado por Mesa-Arango *et al.* (2018), que apontaram que o aumento da população economicamente ativa, reduziria os riscos de acidentes.

Com relação ao Índice de Desenvolvimento Humano Municipal (IDHM), este demonstrou significância estatística para maior parte dos modelos, exceto acidentes por quilômetro (ACID_KM) e com vítimas fatais (VIT_FATAIS). Nos modelos em que

foi significativo apresentou coeficientes positivos, podendo ser um indicativo de que regiões com maiores índices, possivelmente por possuírem maior densidade viária, mais veículos e maior fluxo, tendem a registrar maior número de acidentes.

Vale destacar, a limitação dos indicadores socioeconômicos, de acordo com Noland e Laham (2018), que destacam que a falta de informações a respeito das condições socioeconômicas específicas das vítimas nesses estudos, utilizar das informações referentes a uma unidade de agregação, pode ser uma solução, porém aumenta a chance de erros interpretativos.

A população do município também apresentou coeficiente positivo e alta significância estatística em todos os modelos, exceto em acidentes por quilômetro (ACID_KM), demonstrando que o aumento populacional sugere o crescimento no número de acidentes, esse resultado evidencia que o tamanho populacional é um dos principais fatores relacionados à acidentalidade, influenciados principalmente por ser um dos fatores contribuintes para a atração de viagens (LEE, ABDEL-ATY & JIANG, 2015). O sinal negativo do coeficiente, para o modelo da variável dependente acidente por quilômetro (ACID_KM), pode se dar devido ao fato do modelo não apresentar boa capacidade explicativa com $R^2 = 0,5597$.

A densidade demográfica só apresentou significância estatística para os modelos de acidentes por quilômetro (ACID_KM), com vítimas feridas (VIT_FERIDA), urbanos (ACID_URB) e por milhão de veículos do tráfego (ACID_MI_V). Os três primeiros apresentaram coeficientes positivos, indicando que áreas mais densamente povoadas estão associadas a uma maior probabilidade de acidentes e ferimentos, principalmente em zonas urbanas, esse resultado é consistente com o que é abordado pela literatura nesse tema (CHIMBA, MUSINGUZI & KIDANDO, 2018)

O sinal negativo do coeficiente da variável densidade demográfica, foi observado apenas no modelo da variável dependente acidentes por milhão de veículos (ACID_MI_V), o que também pode ser um indicativo da pouca capacidade explicativa do modelo, com $R^2=0,5377$. Os sinais conflitantes e a falta de significância em alguns modelos podem estar de acordo com o observado por Guerra, Dong e Kondo (2019), que constataram não haver tanta significância nessa variável, em comparação com outros fatores relacionados aos acidentes.

4.3.2.4 Coeficientes das variáveis relacionadas a frota e condutores

As variáveis referentes à composição da frota e aos condutores, que consistem em proporção de motocicletas na frota de veículos (PROP_MOTO) e o índice de habitantes por veículo (HAB_VEIC) e proporção de condutores até 24 anos (PROP_AT_24), não apresentaram significância estatística na maioria dos modelos, salvo exceções, indicando baixa relação com os acidentes na região Nordeste.

A proporção de motocicletas em relação ao total da frota de veículos, apresentou significância estatística em modelos específicos, como acidentes por milhão de veículos (ACID_MI_V) e com vítimas fatais (VIT_FATAIS). Em ambos os casos o coeficiente negativo sugere que regiões com maior proporção de motocicletas têm menos acidentes por milhão de veículos ou com vítimas fatais. Esse resultado vai de encontro a literatura, que de maneira clara associa as motocicletas a um risco viário elevado (BLAIZOT *et al.*, 2013). Esse resultado pode se dar, por algum fator não percebido pelo modelo, imprecisão, ou características regionais específicas. A falta de significância estatística observada na maioria dos modelos pode indicar que este não é um fator determinante de maneira isolada, mas condicionado a outros fatores, que podem ser melhor avaliados em outros trabalhos.

O índice de motorização, representado pela variável habitantes por veículos (HAB_VEIC), apresentou pouca significância estatística nos modelos analisados, relevante apenas no modelo de acidentes sem vítimas (SEM_VIT). O coeficiente positivo sugere que o aumento desse índice, implicando em mais pessoas por veículos registrados, apresenta tendência a mais ocorrências sem vítimas. Porém, a falta de significância nos demais modelos podem ser um indicativo que essa variável, de maneira isolada não é determinante para a ocorrência dos sinistros rodoviários, nesse contexto.

A proporção de condutores até 24 anos (PROP_AT_24), indicador da faixa etária dos condutores envolvidos nos sinistros viários, não apresentou significância estatística para os modelos. Apesar de não obter significância estatística, o coeficiente negativo, em boa parte dos modelos, pode ser um indicativo da pouca expressividade desse grupo em relação às demais faixas etárias, representando baixa influência sobre os acidentes. Esse resultado corrobora com outros trabalhos, que destacam a forte influência da faixa etária que engloba maior parte da população economicamente ativa, entre 15 a 65 anos (CHIMBA, MUSINGUZI & KIDANDO, 2018).

4.3.3 Modelo final para a variável dependente acidentes totais (ACID_TOTAL)

Após os resultados dos modelos ajustados, foi executado um modelo final, com 8 variáveis explicativas, para a variável dependentes acidentes totais. A Tabela 15 apresenta os indicadores da performance desse modelo, permitindo comparar com os resultados obtidos nos modelos anteriores.

Tabela 15 – Principais métricas indicadoras de performance do modelo final de MQO para a variável acidentes totais

Indicadores	ACID_TOTAL
R ²	0,7436
R ² (ajustado)	0,7404
RSE	21,86 com 645 DF
p-valor	<2,2e-16
AIC	5901,923
MAE	12,1721
MSE	393,6915
R ² (teste)	0,7465

Fonte: O autor (2025).

O modelo manteve os resultados praticamente iguais ao modelo ajustado, atestando que a retirada das variáveis não representou mudança significativa para a performance do modelo. A Tabela 16 apresenta os coeficientes obtidos para as variáveis explicativas utilizadas no modelo final.

Tabela 16 – Coeficientes das variáveis do modelo final de MQO

Modelo Variáveis	ACID_TOTAL			
	Coeficiente	Erro	t-valor	p-valor
(Intercepto)	-74,390	13,840	-5,375	1,07E-07 ***
EXT_VIAS	0,477	0,050	9,538	<2E-16 ***
VMDA	0,001	0,0002	8,251	8,83E-16 ***
P_POP_OCUP	39,700	16,430	2,416	0,01596 *
POPULACAO	0,00013	0,00001	13,952	<2E-16 ***
DENS_DEM	0,005	0,003	1,564	0,11835
IDHM	100,500	25,100	4,004	6,94E-05 ***
AREA_UN	-0,004	0,001	-4,250	2,45E-05 ***
DENS_VIA	69,290	24,690	2,807	0,00515 **

Legenda: * significância ao nível de 0,05

** significância ao nível de 0,01

*** significância ao nível de 0,001

Fonte: O autor (2025).

A análise dos coeficientes demonstrou que as variáveis extensão de vias (EXT_VIAS), volume médio diário anual (VMDA), população (POPULACAO), Índice de Desenvolvimento Humano Municipal (IDHM) e densidade de Vias (DENS_VIA)

apresentaram impactos positivos e estatisticamente significativos sobre a variável dependente. Destacando que o aumento dessas variáveis implica no aumento da totalidade dos acidentes observados.

Por outro lado, a variável área da unidade territorial – município (AREA_UN) apresentou coeficiente negativo altamente estatisticamente significativo, indicando que quanto maior a área do município, menor a acidentalidade observada, o que pode estar relacionado à menor concentração de acidentes em áreas maiores.

Em contrapartida, a densidade demográfica (DENS_DEM) não se mostrou significativa, sugerindo que não exerce influência expressiva no modelo, possivelmente influenciada pela presença da variável população total. Uma Análise de Componentes Principais pode apresentar-se como sugestões para trabalhos futuros, de modo a combinar variáveis correlacionadas sem perder informações valiosas aos modelos.

4.3.4 Análise dos pressupostos do modelo linear

Para investigar os pressupostos do modelo linear, analisou-se a normalidade dos resíduos, a autocorrelação dos resíduos e a homocedasticidade do modelo. Portanto, esses testes são fundamentais para verificar a conformidade dos modelos de MQO, atendendo aos pressupostos estatísticos, para garantir a validade das inferências realizadas. A Tabela 17 apresenta os resultados do teste do pressuposto de normalidade dos resíduos dos modelos ajustados e final.

Tabela 17 – Resultados do teste de Shapiro-Wilk

Teste de normalidade dos resíduos - Shapiro-Wilk			
Modelos	W	p-valor	Conclusão
ACID_TOTAL	0,67468	<2,20e-16	Os resíduos não seguem uma distribuição normal (violação da suposição de normalidade dos resíduos)
ACID_KM	0,31006	<2,20e-16	
ACID_MI_V	0,69575	<2,20e-16	
SEM_VIT	0,71495	<2,20e-16	
VIT_FERIDA	0,66662	<2,20e-16	
VIT_FATAIS	0,91639	<2,20e-16	
ACID_URB	0,69781	<2,20e-16	
ACID_RUR	0,69465	<2,20e-16	
ACID_TOTAL (final)	0,6775	<2,20e-16	

Fonte: O autor (2025).

O teste de normalidade de Shapiro-Wilk foi aplicado aos resíduos dos modelos, para verificar o padrão de normalidade na distribuição dos dados. Os resultados do valor estatístico W variam de 0,31006 (ACID_KM) e 0,91639 (VIT_FATAIS) e todos os p-valores são menores que $2,2 \times 10^{-16}$, indicando a rejeição da hipótese nula em todos os modelos. A rejeição da hipótese nula implica que os resíduos do modelo não seguem uma distribuição normal.

A Tabela 18 apresenta os resultados para o teste da autocorrelação dos resíduos de *Durbin-Watson*.

Tabela 18 – Resultados do teste de autocorrelação dos resíduos

Teste de autocorrelação dos resíduos – Durbin-Watson			
Modelos	DW	p-valor	Conclusão
ACID_TOTAL	2,013	0,5511	Os valores de DW \approx 2 indicam que não há autocorrelação dos resíduos
ACID_KM	2,0093	0,5323	
ACID_MI_V	2,0644	0,7844	
SEM_VIT	2,0163	0,5679	
VIT_FERIDA	2,0166	0,5692	
VIT_FATAIS	1,9767	0,3681	
ACID_URB	2,0453	0,7064	
ACID_RUR	1,9728	0,3492	
ACID_TOTAL (final)	2,0047	0,5182	

Fonte: O autor (2025).

A presença de autocorrelação dos resíduos foi analisada por meio do teste de Durbin-Watson, em que os valores da estatística DW aproximadamente igual a dois indicam ausência de autocorrelação. Nos modelos analisados os valores oscilaram entre 1,9728 a 2,0644, os p-valores foram relativamente altos, entre 0,3492 a 0,7844. Logo, os resultados indicam que não há evidências de autocorrelação significativa nos resíduos dos modelos, o que é um aspecto positivo, sugerindo que os erros não estão correlacionados uns com os outros ao longo das observações. Portanto, a ausência da autocorrelação dos resíduos atende a um dos pressupostos da regressão.

A Tabela 19 apresenta os resultados do teste de homocedasticidade de *White* e de *Breush-Pagan*

Tabela 19 – Resultados do teste de homocedasticidade

Teste de homocedasticidade – White BP <i>studentized</i> e Breush-Pagan					
Modelos	BP (st.)	p-valor	BP	p-valor	Conclusão
ACID_TOTAL	396,57	<2,2e-16	6615,6	<2,2e-16	Os resíduos têm heterocedasticidade (violação da suposição de homocedasticidade)
ACID_KM	56,228	1,089e-07	6486,6	<2,2e-16	
ACID_MI_V	172,09	<2,2e-16	3392,5	<2,2e-16	
SEM_VIT	314,39	<2,2e-16	5300,4	<2,2e-16	
VIT_FERIDA	391,76	<2,2e-16	6287,1	<2,2e-16	
VIT_FATAIS	211,75	<2,2e-16	915,34	<2,2e-16	
ACID_URB	385,74	<2,2e-16	5122,5	<2,2e-16	
ACID_RUR	404,2	<2,2e-16	6956,7	<2,2e-16	
ACID_TOTAL (final)	399,68	<2,2e-16	6549,4	<2,2e-16	

Fonte: O autor (2025).

A homocedasticidade dos modelos foi analisada por meio dos testes de *White studentized* e *Breuch-Pagan*, para esses dois testes a hipótese nula assume a variância constantes dos resíduos. Os valores dos testes estatísticos foram elevados para todos os modelos, indicando que a hipótese nula de homocedasticidade foi rejeitada. Esse resultado evidencia a presença de heterocedasticidade nos modelos, sugerindo que a variância dos erros não é constante e que os modelos podem estar sendo influenciados por observações maiores, como por exemplo variáveis com alta variabilidade no conjunto de dados e presença de *outliers*.

Desse modo, os resultados dos testes dos pressupostos demonstram que os modelos não atenderam a dois dos três pressupostos analisados. A influência dos *outliers* apresentados nas estatísticas descritivas contribui para os resultados. Desse modo a distribuição que não segue um padrão normalidade para os resíduos e a heterocedasticidade. Por outro lado, a ausência de autocorrelação dos resíduos aparece como contraponto a esses indicadores. Logo, a presença de valores discrepantes interfere nos resultados dos modelos e em sua significância estatística.

4.3.5 Análise geral sobre os modelos de MQO

Portanto, com base nos indicadores de performance dos modelos e dos coeficientes observados para as variáveis explicativas, verificou-se para os modelos ajustados uma boa capacidade explicativa para as variáveis. Destacaram-se principalmente os modelos de: acidentes totais (ACID_TOTAL), com vítimas feridas (VIT_FERIDA), urbanos (ACID_URB) e rurais (ACID_RUR).

Os modelos de acidentes sem vítimas (SEM_VIT) e com vítimas fatais (VIT_FATAL) apresentaram desempenho mediano, tanto nos indicadores de desempenho, como na análise de variáveis explicativas. Os modelos de acidentes por milhão de veículos (ACID_MI_V) e acidentes por quilômetro de via (ACID_KM), apresentaram desempenho inferior, com indicadores ruins e inconsistências na análise das variáveis explicativas.

Com relação aos grupos de variáveis, verificou-se um bom desempenho das variáveis referentes aos indicadores socioeconômicos e demográficos e ligados à infraestrutura viária e territorial, enquanto as variáveis ligadas à frota e condutores não apresentaram significância para os modelos. Esses resultados destacam a capacidade do modelo de capturar as relações entre as variáveis.

Os resultados das Tabelas 17, 18 e 19 indicam que os pressupostos da regressão linear foram violados, o que compromete a confiabilidade dos modelos lineares tradicionais (MQO). Os principais problemas identificados foram: Os resíduos não seguem uma distribuição normal e a heterocedasticidade nos resíduos. Verificou-se também indícios de que não há evidências autocorrelação nos resíduos dos modelos.

Para realizar uma avaliação preliminar do efeito espacial nos dados, precisamos identificar padrões espaciais na accidentalidade rodoviária e avaliar se a localização geográfica influencia os acidentes, pode ser utilizado: o Índice de Moran: Mede a autocorrelação espacial dos acidentes. Se o valor for alto, indica que os acidentes não estão distribuídos aleatoriamente, mas seguem um padrão espacial; mapa de clusters de acidentes (LISA - Indicadores Locais de Associação Espacial): Identifica áreas com alta ou baixa concentração de acidentes; Teste de dependência espacial dos resíduos da regressão, caso os resíduos do modelo linear apresentem correlação espacial, isso sugere que o Modelo de Regressão Geograficamente Ponderado (RGP) pode ser mais adequado, permitindo captar padrões espaciais para fornecer uma compreensão mais abrangente e detalhada sobre os fatores associados aos acidentes rodoviários.

4.4 ÍNDICE DE MORAN

4.4.1 Índice de Moran Global

A análise do Índice de Moran Global, foi realizada com o objetivo de verificar a autocorrelação espacial dos resíduos dos modelos de MQO. Essa etapa visa identificar padrões espaciais que podem indicar a necessidade de utilizar de modelos de regressão espacial. A Tabela 20 a seguir apresenta os resultados do Índice de Moran Global para os nove modelos utilizados, para as oito variáveis dependentes, com dois modelos da variável acidentes totais (ACID_TOTAL), sendo o primeiro modelo, com todas as variáveis inicialmente empregados e o outro, após os ajustes necessários e retirada dos fatores que não apresentaram significância estatística.

Tabela 20 – Resultados do índice de Moran para todos os modelos de MQO.

Indicadores	ACID_TOTAL	ACID_KM	ACID_MI_V
Índice de Moran	0,01634	-0,00840	0,04837
Índice Esperado	-0,00153	-0,00153	-0,00153
Variância	0,00026	0,00018	0,00026
z-score	1,110	-0,517	3,115
p-valor	0,134	0,698	9,18E-04
I Moran (Monte Carlo)	0,01634	-0,00840	0,04837
p-valor (Monte Carlo)	0,13860	0,65350	0,00990
Indicadores	ACID_URB	ACID_RUR	SEM_VIT
Índice de Moran	0,00845	0,01736	0,00383
Índice Esperado	-0,00153	-0,00153	-0,00153
Variância	0,00026	0,00026	0,00026
z-score	0,617	1,175	0,333
p-valor	0,269	0,120	0,369
I Moran (Monte Carlo)	0,00845	0,01736	0,00383
p-valor (Monte Carlo)	0,26730	0,14850	0,30690
Indicadores	VIT_FERIDA	VIT_FATAIS	ACID_TOTAL_2
Índice de Moran	0,01418	0,08735	0,02386
Índice Esperado	-0,00153	-0,00153	-0,00153
Variância	0,00026	0,00027	0,00026
z-score	0,975	5,414	1,577
p-valor	0,165	3,07E-08	0,057
I Moran (Monte Carlo)	0,01418	0,08735	0,02386
p-valor (Monte Carlo)	0,17820	0,00990	0,06931

Fonte: O autor (2025).

Os valores do Índice de Moran variaram entre os diferentes modelos analisados. Esse índice apresentou valores positivos para maior parte dos modelos, exceto o modelo de acidentes por quilômetro (ACID_KM). Por outro lado, com relação a significância estatística desses modelos, observou-se apenas no modelo de acidentes por milhão de veículos do tráfego (ACID_MI_V) e acidentes com vítimas fatais (VIT_FATAIS), a um nível de significância de 0,05. Considerando um nível de

significância de 0,10, é possível considerar o modelo ajustado de acidentes totais (ACID_TOTAL).

O modelo para a variável dependente acidentes com vítimas fatais (VIT_FATAIS) apresenta o maior Índice de Moran e p-valor significativo, indicando forte correlação a espacial entre os acidentes fatais. Modelos de acidentes por milhão de veículos do tráfego (ACID_MI_V) e acidentes rurais (ACID_RUR), que apresentaram tendência espacial, poderiam ser testados quanto a modelos híbridos, como Regressão Espacial por erros (SAR) ou modelos mistos.

Esses resultados indicam que os resíduos desses modelos não estão distribuídos de forma aleatória na região nordeste e que fatores espaciais não foram capturados pelo modelo de MQO. Destaca-se, portanto, a importância de incorporar abordagens espaciais, como a Regressão Geograficamente Ponderada (RGP), para compreender melhor os fatores que influenciam esses padrões.

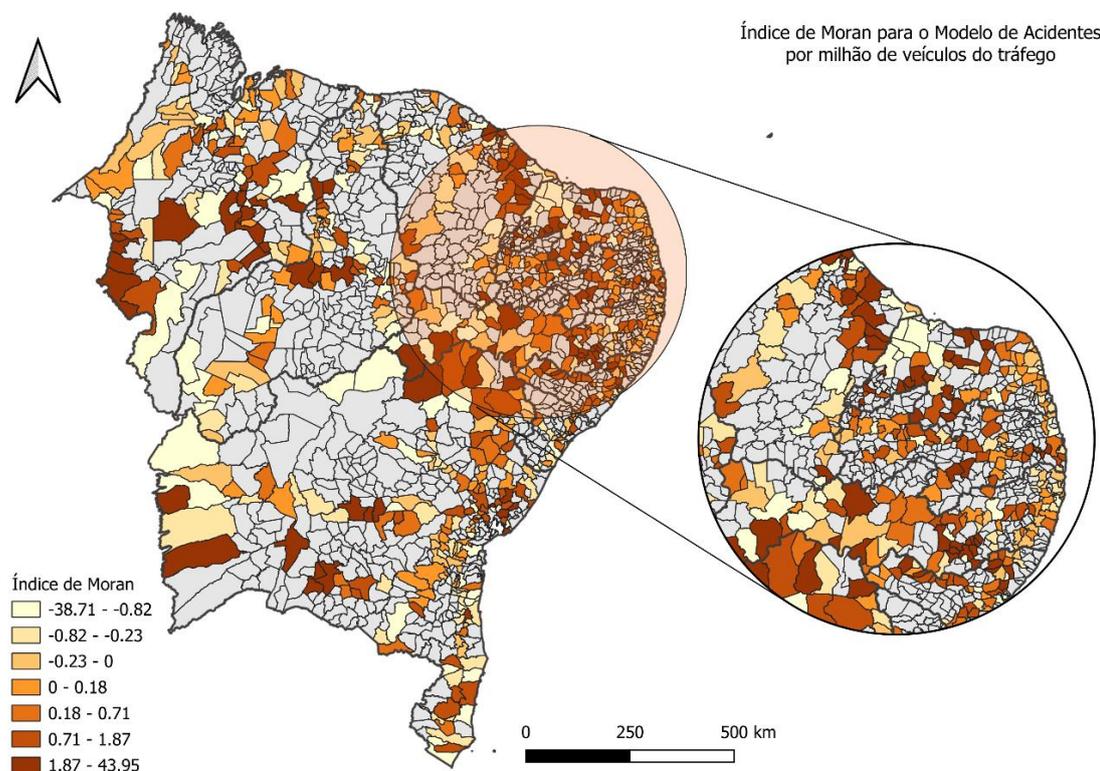
A ausência de autocorrelação espacial com significância estatística em outros modelos sugere que fatores explicativos tenham sido capazes de capturar a dinâmica espacial dessas variáveis, ou mesmo destacar que a distribuição espacial de acidentes menos graves é mais homogênea e menos influenciada por fatores locais específicos.

A análise desse Índice demonstrou que, embora nem todos os modelos tenham apresentado padrões espaciais significativos, há evidências robustas de agrupamento espacial, principalmente relacionado às vítimas fatais. Esses resultados reforçam a necessidade aprofundar essa análise, por meio da regressão geograficamente ponderada, enquanto modelo espacial, contribuindo para orientar políticas públicas mais eficazes e direcionadas.

4.4.2 Índice de Moran Local ou *LISA (Local Indicators of Spatial Association)*

Enquanto o Índice de Moran Global indica se há autocorrelação espacial no conjunto todo, o Índice de Moran Local identifica onde essa autocorrelação ocorre. De modo a identificar os *clusters locais* de agrupamentos dos acidentes por milhão de veículos do tráfego (ACID_MI_V) acidentes totais (ACID_TOTAL) e vítimas fatais (VIT_FATAIS), foi analisado o Índice de Moran Local. A Figura 4 apresenta o mapa do Índice de Moran Local para a variável dependente acidentes por milhão de veículos do tráfego.

Figura 4 – Índice de Moran Local para o modelo da variável dependente Acidentes por tráfego



Fonte: O autor (2025)

Os resultados do Índice de Moran local para o modelo da variável dependente acidentes por milhão de veículos do tráfego (ACID_MI_V), permitiram a identificação de padrões espaciais e agrupamentos em algumas regiões, evidenciando pontos com maiores índices, enquanto outras regiões apresentaram valores inferiores. Essas relações, para o modelo em questão, indicam a alta ou baixa incidência de acidentes relacionados ao tráfego de veículos, podendo estar mais presente em regiões com maior fluxo de veículos, devido a altas taxas de acidentes por tráfego em regiões geograficamente próximas.

Vale destacar, com relação aos indicadores obtidos, que o modelo de MQO, apesar de ter um Índice de Moran significativo, apresentou performance inferior ao esperado, com R^2 de 0,5377 e R^2 ajustado de 0,529, o que pode influenciar negativamente a precisão desse Índice de Moran Local.

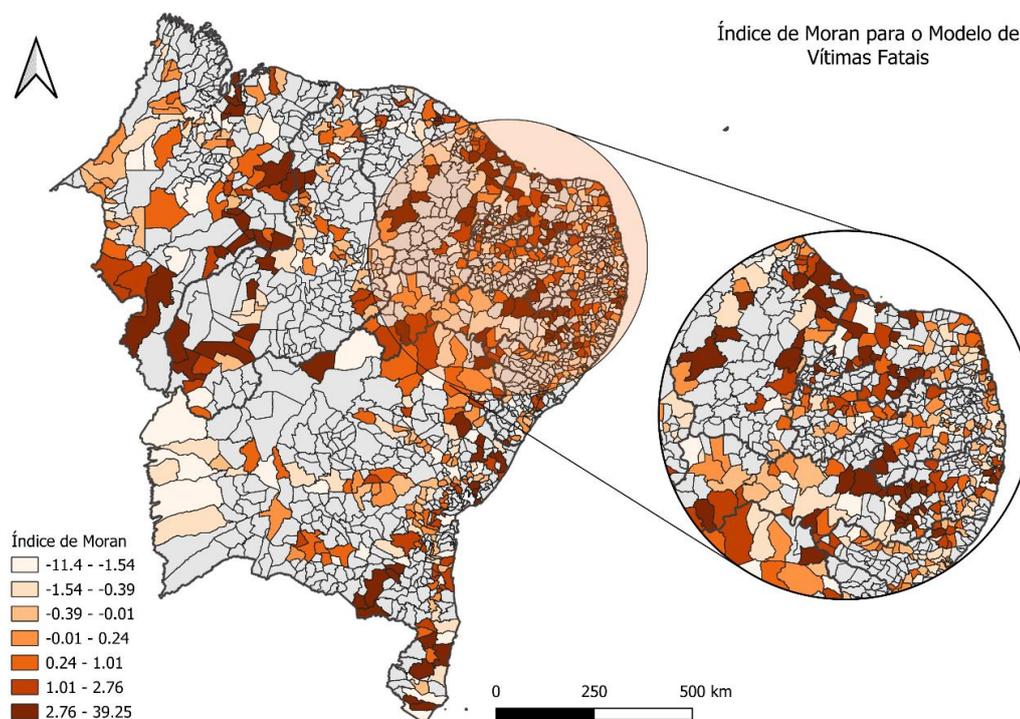
Apesar disso, observou-se alguns agrupamentos que podem ser evidenciados. No Estado do Ceará os maiores índices foram observados na Região Metropolitana de Fortaleza e municípios próximos, certamente influenciado pelo tráfego intenso e municípios da mesorregião do Jaguaribe, em trechos da BR-116. No Estado de

Pernambuco, um dos agrupamentos observados, foi na mesorregião do Agreste, na microrregião meridional, englobando Garanhuns, sugerindo uma interação da acidentalidade forte, nesse trecho das rodovias BR-423 e BR-424.

O Estado do Rio Grande do Norte apresentou agrupamentos mais significativos nas microrregiões do Seridó Potiguar e do Sertão do Apodi, região que conta com trechos das rodovias, BR-427 e BR-226. No Estado da Bahia, os agrupamentos com índices positivos mais relevantes estão entre municípios do vale do São Francisco, na divisa com Pernambuco e municípios da Região Metropolitana de Salvador. O sul do Maranhão também apresenta alguns agrupamentos, situados em trechos da BR-230 e a região metropolitana de São Luís, também apresentou indicadores relevantes.

Por outro lado, algumas capitais, não apresentaram indicadores significativos, como por exemplo, Natal, João Pessoa, Recife, Maceió e Aracaju. A ausência de índices mais elevados nessas capitais, pode evidenciar uma baixa capacidade explicativa do modelo, nesses pontos. A Figura 5 apresenta o Índice de Moran Local para a variável dependente acidentes com vítimas fatais.

Figura 5 – Índice de Moran Local para o modelo da variável dependente Vítimas Fatais



Fonte: O autor (2025)

O Índice de Moran Local, para o modelo de MQO da variável dependente acidentes com vítimas fatais (VIT_FATAL), também apresentou indicadores

interessantes, a respeito dos sinistros viários na região nordeste. Porém, vale destacar a baixa capacidade explicativa desse modelo, com R^2 de 0,4832 e de R^2 ajustado de 0,4736, o que pode contribuir para índices locais enviesados, com incoerências.

No Estado da Bahia, alguns índices positivos significativos, foram observados em municípios a norte de Salvador, Entre rios, Esplanada e Olindina, em trechos das rodovias BR-101 e BR-110, em trechos rurais dessas rodovias, podendo evidenciar a gravidade dos acidentes nesses locais, mais críticos, também nesse estado, destacaram-se os índices próximos a Vitória da Conquista e Porto Seguro, em trechos das rodovias, BR-116, BR-101 e BR-367.

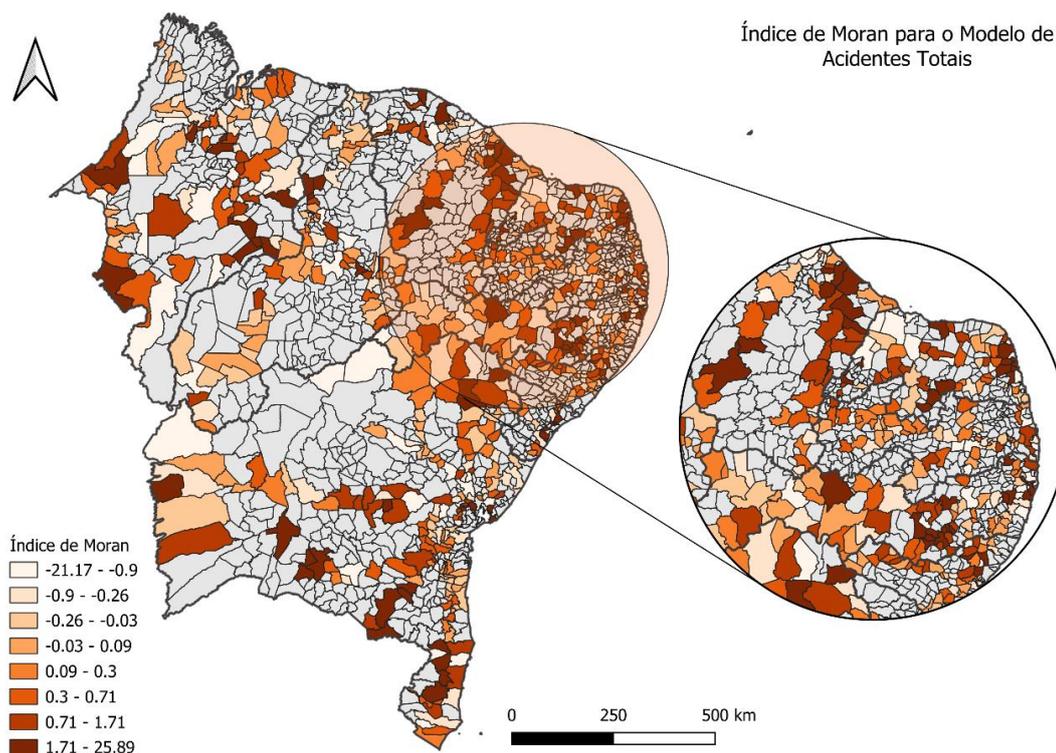
No Estado de Pernambuco, os indicadores positivos se concentraram na região do agreste, próximo a Garanhuns (BR-424), bem como em um conjunto de municípios cortados pela BR-232, no estado, de Caruaru a Custódia, evidenciando a gravidade dos acidentes com vítimas fatais nessas localidades, alguns municípios da Região Metropolitana do Recife (RMR), apresentaram fortes indicadores, mostrando a necessidade de intervenções nesses trechos, em alguns municípios, como Jaboatão dos Guararapes e São Lourenço da Mata.

O Estado do Rio Grande do Norte também apresentou indicadores fortes na região do Seridó e do Sertão do Apodi. Com relação ao Ceará, a mesorregião do Jaguaribe apresentou fortes indicadores, evidenciando a importância de maior monitoramento na região, visando diminuir os riscos de acidentes fatais na região.

Na região próxima ao município de Bom Jesus, no Piauí, foi identificado um forte agrupamento, com índices elevados, em região cortada pelas rodovias BR-135 e BR-235. Na região Metropolitana de São Luís e Teresina, divisa entre Maranhão e Piauí, apresentaram índices elevados, o que evidencia a forte influência da ocorrência de acidentes fatais, nessas capitais. O Maranhão também apresenta valores positivos relevantes, para a região próxima ao município de Colinas e São João dos Patos, cortados pelas rodovias BR-230 e BR-135.

Por outro lado, algumas regiões do sertão maranhense, do interior do Piauí, sertão do Pernambuco, interior de Alagoas e Sergipe, apresentaram clusters baixos, evidenciando a baixa dependência espacial em algumas localidades. Esses resultados reforçam a necessidade de intervenções específicas, voltados para trechos rurais das rodovias, ou em algumas capitais, como por exemplo São Luís, Teresina e Recife. A Figura 6 apresenta o Índice de Moran Local para a variável acidentes totais.

Figura 6 – Índice de Moran Local para o modelo da variável dependente Acidentes Totais



Fonte: O autor (2025)

No que diz respeito aos resultados do Índice de Moran Local para a variável dependente acidentes totais (ACID_TOTAL), foram obtidos agrupamentos relevantes, que podem ser analisados, sob a perspectiva da dependência espacial observada entre alguns municípios. Os resultados de performance do modelo de MQO, que apresentou R^2 de 0,7436 e R^2 ajustado de 0,7404, são satisfatórios, e podem contribuir para a identificação de agrupamentos relevantes para os dados utilizados.

Vale destacar, que para essa variável, indicadora da frequência dos acidentes rodoviários, em municípios de todas as Regiões Metropolitanas, foram obtidos Índices positivos, indicando a relevância dessas regiões, que concentram o tráfego e a acidentalidade, em muitos estados da região Nordeste.

Além disso, observou-se fortes indicadores no agreste pernambucano, próximo ao município de Garanhuns, no Seridó Potiguar (no Rio Grande do Norte), em municípios do Vale do São Francisco e em municípios próximos a Vitória da Conquista e Porto Seguro, na Bahia. A mesorregião do Jaguaribe, no Estado do Ceará, também apresentou fortes indicadores, evidenciando a importância de identificar ações importantes nesse ponto crítico presente em todas as três análises conduzidas.

Portanto, com base nas análises conduzidas pelo Índice de Moran, destaca-se a importância de compreender melhor a natureza desses eventos e as causas que envolvem essas ocorrências em diversos municípios. Destaca-se nesse sentido a importância de analisar e estudar individualmente o contexto das Regiões Metropolitanas, compreendendo sua relevância e complexidade no contexto da acidentalidade.

Torna-se válido evidenciar a importância de alguns trechos em específico, como os municípios em volta de Vitória da Conquista, na Bahia, Garanhuns, em Pernambuco, Currais Novos, no Rio Grande do Norte e Russas, no Ceará. Ao analisar por meio do Índice de Moran Local essas regiões, verificou-se uma forte presença de dependência espacial, entre os municípios próximos, destacando também a sua relevância sob o contexto da acidentalidade, podendo ser conduzidos outros estudos, evidenciando as características dessas regiões, que levam a um maior risco, dos acidentes rodoviários observados.

4.5 REGRESSÃO GEOGRAFICAMENTE PONDERADA (RGP)

4.5.1 Indicadores de Desempenho da RGP

Os resultados dos indicadores de desempenho e precisão do modelo de Regressão Geograficamente Ponderada (RGP) demonstram a eficácia do modelo como um todo, na identificação dos padrões espaciais. A Tabela 21 a seguir demonstra os principais indicadores dos modelos de RGP.

Tabela 21 – Principais métricas indicadoras de performance do modelo RGP.

Indicadores	ACID_TOTAL	ACID_KM	ACID_MI_V
Coef. de Correlação (r)	0,9626	0,9122	0,9257
R ² (médio)	0,8585	0,7701	0,7052
RSE	11,72 em 645 DF	1,025 em 645 DF	6,70 em 645 DF
RSE (%)	56,19%	104,78%	58,74%
Log-verossimilhança	4,9082	0,0364	3,7900
AIC	8,1837	17,9271	10,4200
Indicadores	ACID_URB	ACID_RUR	SEM_VIT
Coef. de Correlação (r)	0,9616	0,9564	0,9448
R ² (médio)	0,8541	0,8414	0,8201
RSE	5,655 em 645 DF	6,316 em 645 DF	2,628 em 645 DF
RSE (%)	60,09%	57,79%	73,25%
Log-verossimilhança	3,4512	3,6722	1,9190
AIC	11,0975	10,6556	14,1620
Indicadores	VIT_FERIDA	VIT_FATAIS	
Coef. de Correlação (r)	0,9611	0,8525	
R ² (médio)	0,8495	0,6069	
RSE	9,331 em 645 DF	1,432 em 645 DF	
RSE (%)	61,81%	66,12%	
Log-verossimilhança	4,4528	0,7050	
AIC	9,0943	16,5901	

Fonte: O autor (2025).

Os resultados do modelo de RGP demonstraram a eficiência do modelo, na identificação das previsões por meio de uma análise geográfica. O coeficiente de correlação elevado ($r > 0,85$), demonstra que os valores observados e previstos, estão com alta correlação entre si, demonstrando a eficiência do modelo.

Os coeficientes de determinação (R^2), apresentaram melhorias significativas em comparação com os coeficientes dos modelos de MQO, apresentados anteriormente nas Tabelas 10 e 11. Este resultado demonstra que os modelos de RGP explicam melhor a variabilidade dos dados de acidentes, que se distribuem geograficamente, para todas as variáveis dependentes dos modelos.

O erro padrão de estimativa (RSE) também apresentou redução dos indicadores para todos os modelos. Esse resultado demonstra a eficiência desse modelo, contribuindo para a redução do erro padrão de estimativa (RSE). O modelo de RGP da variável de acidentes totais apresentou melhor desempenho dentre os modelos destacados.

Para esse modelo também foi calculada a mediana do p-valor, indicando a significância estatística do modelo. O resultado obtido, com p-valor de 0,1239 demonstra uma diminuição da significância estatística do modelo para parte das

localidades, em comparação com o modelo de MQO, verifica-se uma melhoria acentuada de outros indicadores de performance do modelo. A Tabela 22 a seguir, compara os índices obtidos para os modelos de MQO e de RGP, para a variável de frequência dos acidentes (ACID_TOTAL).

Tabela 22 – Métricas de performance dos modelos de MQO e RGP para a variável ACID_TOTAL.

Indicadores	MQO	RGP
R ²	0,7436	0,8585
RSE	21,86 em 645 DF	11,72 em 645 DF
AIC	5901,923	8,184
p-valor	<2,2e-16	0,1239

Fonte: O autor (2025).

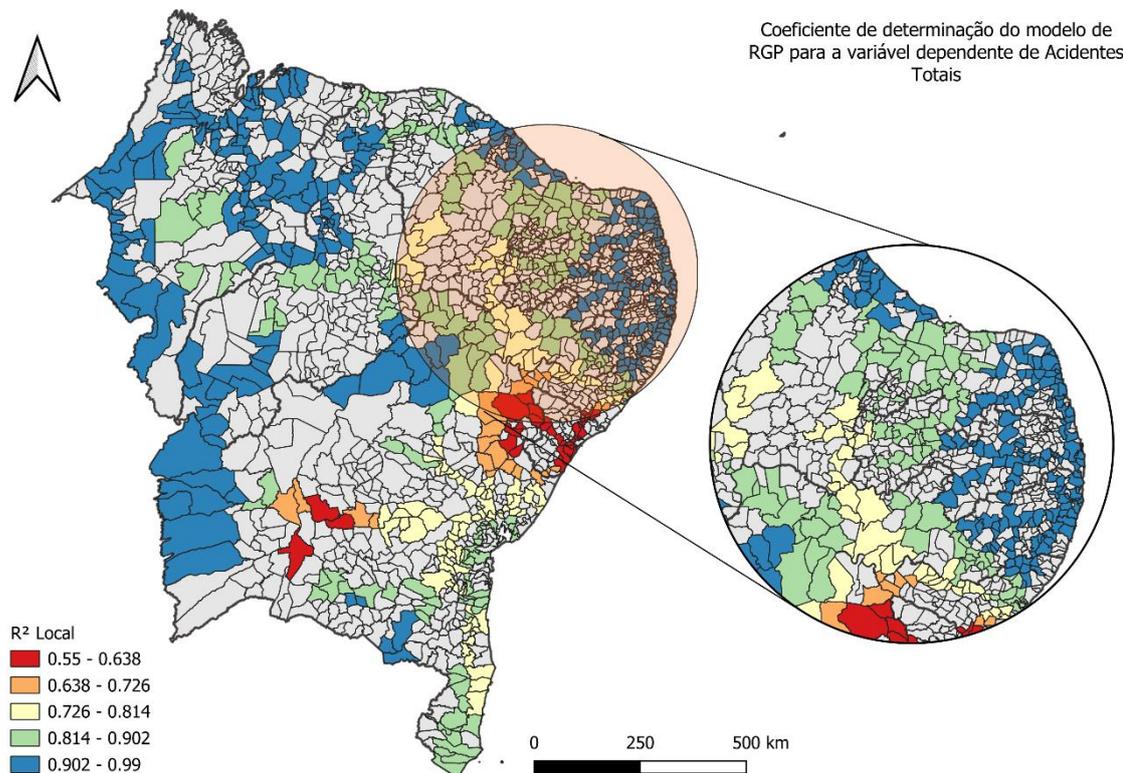
De modo comparativo, os valores dos indicadores apresentaram melhorias significativas. Verificou-se a redução do valor do Critério de Informação de Akaike (AIC), o que indica um melhor ajuste do modelo, em comparação com o modelo linear utilizado. Da mesma forma, a melhoria do coeficiente de determinação (R²) e a redução do erro padrão de estimativa (RSE), confirmaram essa melhoria do modelo geoponderado. Entretanto, o p-valor do RGP (0,1239) não é significativo ($p > 0.05$), o que indica que o modelo pode não estar capturando um padrão estatisticamente relevante.

4.5.2 Coeficientes de determinação (R²) locais do modelo de RGP

Os coeficientes de determinação (R²) locais obtidos pelo modelo de Regressão Geograficamente Ponderada (RGP) permitem avaliar a distribuição espacial da capacidade explicativa do modelo em diferentes municípios da região Nordeste do Brasil. Essa abordagem possibilita identificar as relações locais entre as variáveis independentes e dependentes, destacando padrões regionais de maior ou menor ajuste do modelo.

A Figura 7 apresenta os resultados dos coeficientes locais para a variável dependente acidentes totais (ACID_TOTAL).

Figura 7 – Coeficiente de determinação (R^2) local para a variável dependente acidentes totais



Legenda: O valor do coeficiente de determinação (R^2 local) representa o grau de ajuste do modelo estatístico em cada localidade.

Fonte: O autor (2025).

Os resultados indicam significativa heterogeneidade espacial, com altos valores ($R^2 > 0,814$) predominantemente em áreas metropolitanas e clusters localizados nos estados do Ceará, Paraíba, Pernambuco, Maranhão e Rio Grande do Norte. Esses municípios, representados por cores verdes e azuis no mapa, evidenciam alta capacidade explicativa do modelo.

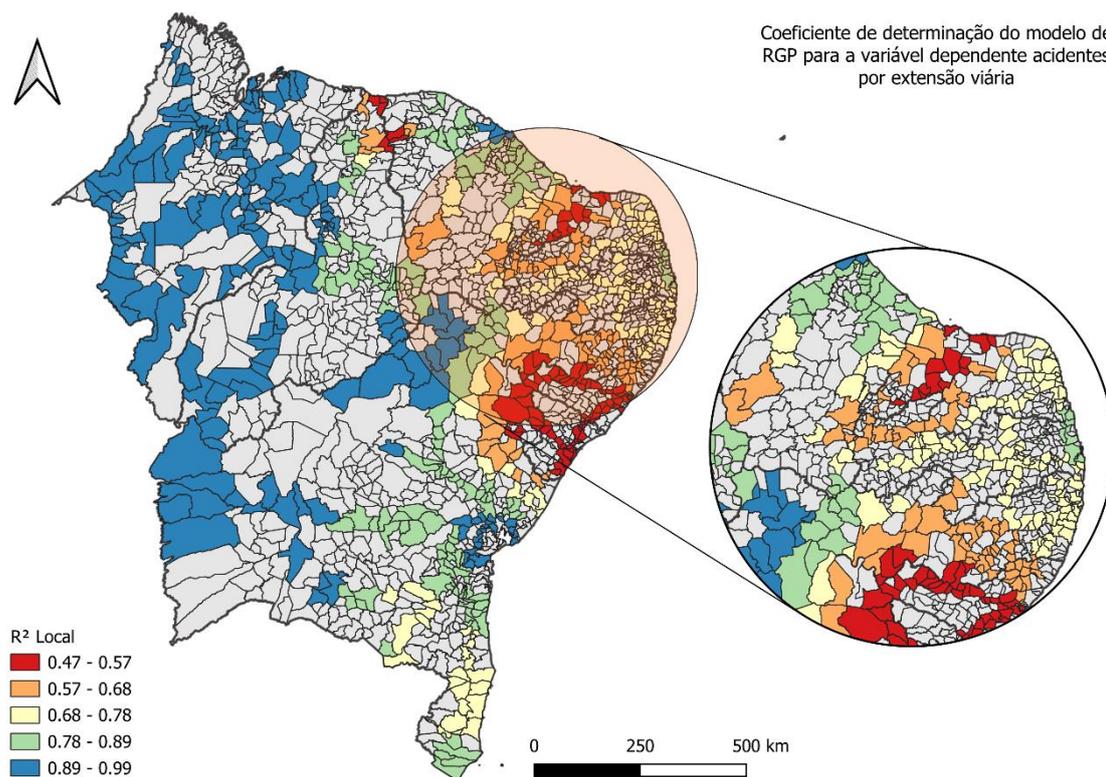
Agrupamentos de altos valores (R^2) em regiões específicas, como o Vale do Rio São Francisco (entre Bahia e Pernambuco) e o Agreste Pernambucano. Os agrupamentos observados entre municípios próximos reforçam a importância de considerar os efeitos espaciais, uma vez que fatores regionais influenciam municípios adjacentes. As regiões com melhor desempenho, indicam que as estratégias de mitigação podem ser direcionadas aos fatores identificados como mais relevantes.

Por outro lado, áreas com valores mais baixos ($R^2 < 0,638$), destacadas em vermelho, sugerem limitações na explicação dos acidentes devido à ausência de variáveis relevantes ou fatores locais não considerados. Por outro lado, áreas de

menor desempenho, como partes da Bahia, Sergipe e Alagoas, evidenciam a necessidade de estudos complementares para aprimorar a modelagem.

As Figuras 8 e 9 apresentam os coeficientes de determinação (R^2) locais para as variáveis dependentes relacionadas às taxas de acidentes, pela extensão viária (ACID_KM) e por milhão de veículos do tráfego anual (ACID_MI_V).

Figura 8 – Coeficiente de determinação (R^2) local, variável dependente acidentes por quilômetro



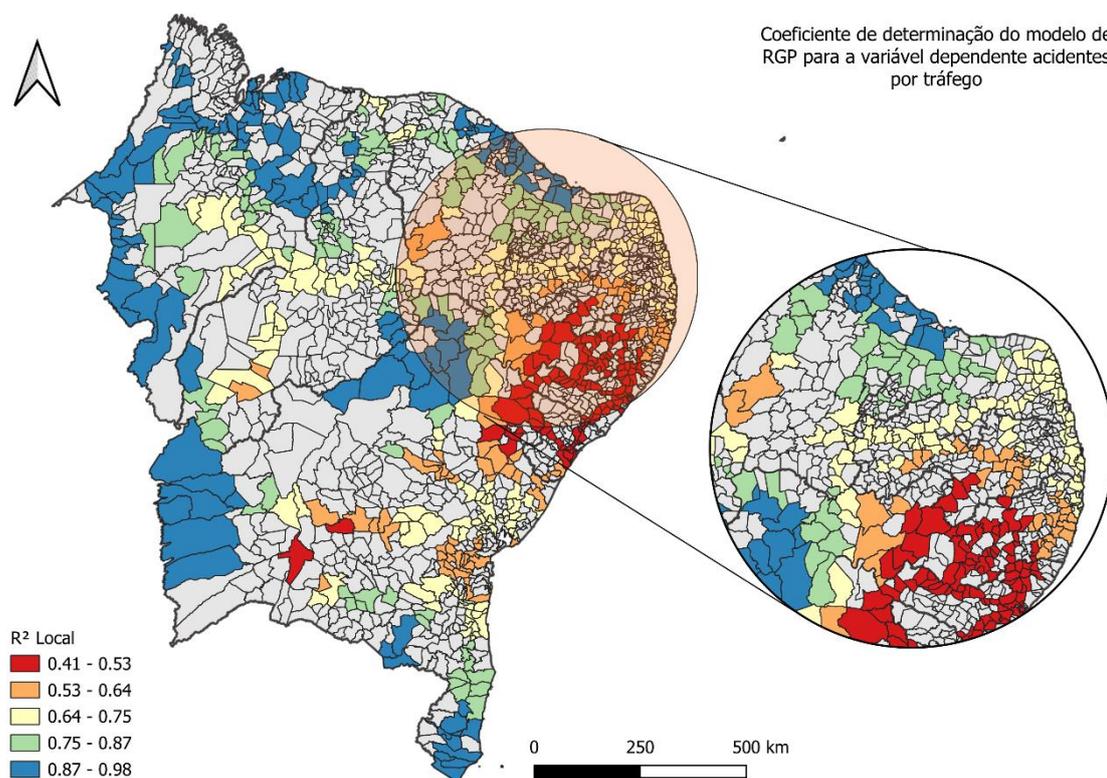
Legenda: O valor do coeficiente de determinação (R^2 local) representa o grau de ajuste do modelo estatístico em cada localidade.

Fonte: O autor (2025).

Observa-se que os valores de R^2 variam amplamente entre os municípios da região Nordeste. Vale destacar a prevalência de valores altos ($R^2 > 0,78$), em todo o estado do Maranhão, parte do vale do Rio São Francisco, no sul do Piauí e em municípios do interior da Bahia. No que diz respeito às regiões metropolitanas, os coeficientes mais altos ($R^2 > 0,78$) foram registrados em São Luís, Teresina, Salvador e João Pessoa. Esses resultados indicam que o modelo é capaz de explicar a variação dos acidentes em função da extensão viária, demonstrando uma relação forte entre a infraestrutura viária e a frequência de acidentes.

No entanto, algumas regiões de Sergipe, Alagoas, agreste de Pernambuco e sertão do Rio Grande do Norte, apresentam agrupamentos com valores baixos ($R^2 < 0,68$), representados pelas áreas em vermelho e laranja. A discrepância observada, pode ser atribuída à ausência de variáveis que forneçam subsídios explicativos para a variável dependente, essas variáveis podem estar relacionadas à geometria de via, qualidade de pavimento ou outros fatores ambientais, não considerados por esse modelo. O desempenho limitado para o modelo de acidentes por extensão viária é consistente com os resultados do modelo de MQO ($R^2 = 0,5597$), sugerindo a necessidade de considerar outros fatores que não incluídos pela modelagem.

Figura 9 – Coeficiente de determinação (R^2) local, variável dependente acidentes por milhão de veículos do tráfego



Legenda: O valor do coeficiente de determinação (R^2 local) representa o grau de ajuste do modelo estatístico em cada localidade.

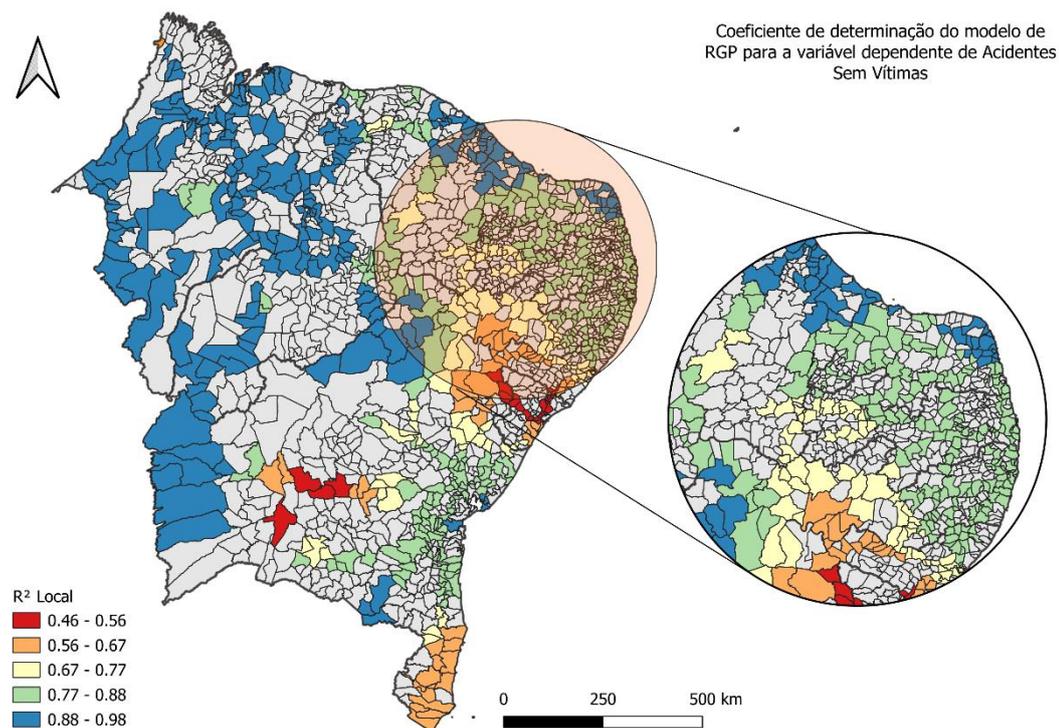
Fonte: O autor (2025).

O desempenho do modelo para essa variável acidentes por milhão de veículos do tráfego, apresentou algumas regiões com R^2 elevado ($R^2 > 0,75$). Os maiores coeficientes foram observados no Maranhão, vale do rio São Francisco, oeste da Bahia e as regiões metropolitanas de São Luís, Teresina e Fortaleza.

Entretanto, áreas de baixa densidade de tráfego, como o agreste de Pernambuco e partes de Alagoas e Sergipe, apresentaram coeficientes reduzidos ($R^2 < 0,53$). Esses resultados indicam que o modelo carece de variáveis capazes de capturar a complexidade das condições locais de tráfego. A baixa capacidade preditiva do modelo também reforça os resultados do modelo de MQO, com $R^2 = 0,5377$ e pode estar relacionada a fatores não considerados por meio do modelo utilizado.

As Figuras 10, 11 e 12 apresentam os coeficientes de determinação (R^2) locais para as variáveis dependentes relacionadas à gravidade dos sinistros viários, acidentes sem vítimas (SEM_VIT), com vítimas feridas (VIT_FERIDA) e com vítimas fatais (VIT_FATAIS).

Figura 10 – Coeficiente de determinação (R^2) local para a variável dependente acidentes sem vítimas.



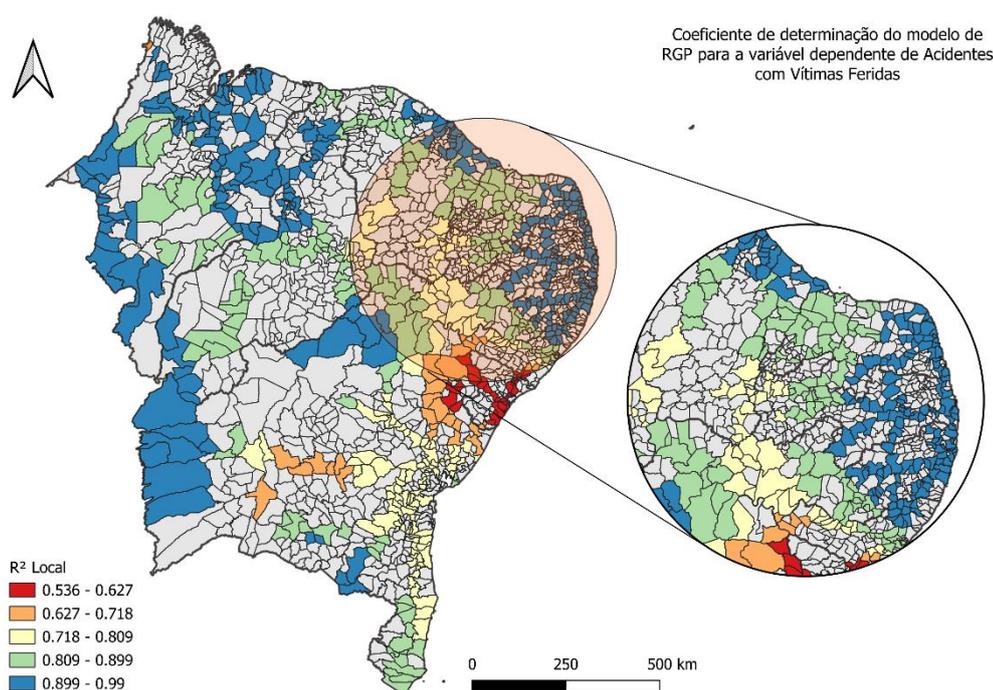
Legenda: O valor do coeficiente de determinação (R^2 local) representa o grau de ajuste do modelo estatístico em cada localidade.

Fonte: O autor (2025).

No que diz respeito aos coeficientes locais dos acidentes sem vítimas, predominam a presença de valores altos ($R^2 > 0,77$). A distribuição espacial segue um padrão similar ao de acidentes totais, com indicadores positivos nas regiões metropolitanas como Fortaleza, João Pessoa, Natal, Recife, São Luís e Teresina. Os

coeficientes elevados em áreas de maior densidade demográfica, podem se dar devido ao maior número de registros sem vítima nos grandes centros, em comparação a outras localidades, em que muitas vezes ocorre menor fiscalização desse tipo de ocorrência. Os valores inferiores ($R^2 < 0,53$) são observados em regiões do interior de Pernambuco, Bahia, Alagoas e Sergipe, demonstrando baixa capacidade explicativa nessas regiões.

Figura 11 – Coeficiente de determinação (R^2) local, variável dependente acidentes c/ vítimas feridas



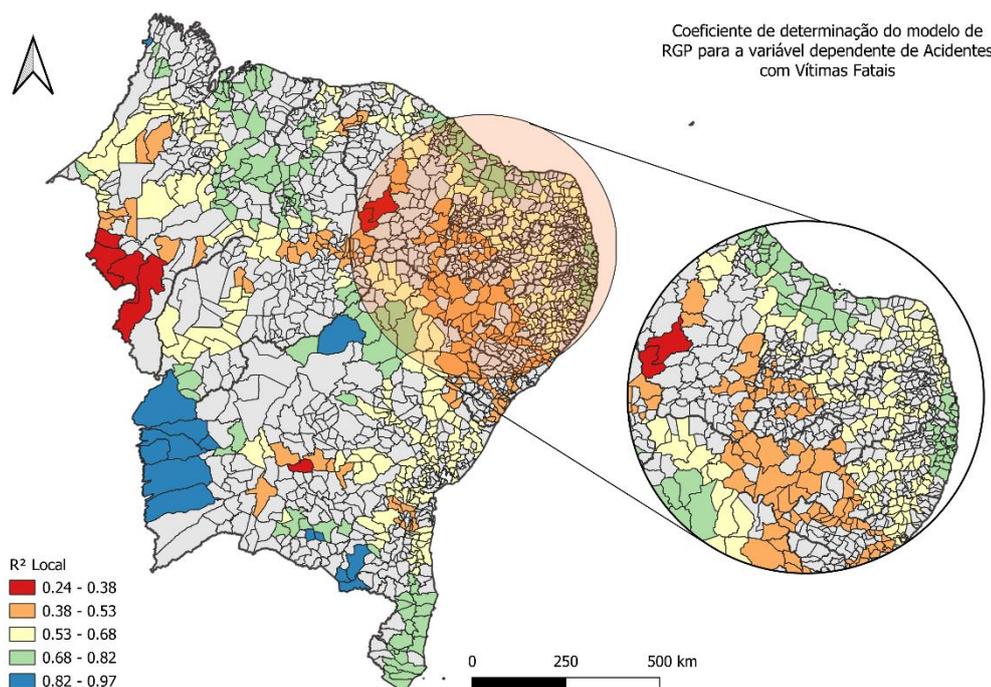
Legenda: O valor do coeficiente de determinação (R^2 local) representa o grau de ajuste do modelo estatístico em cada localidade.

Fonte: O autor (2025).

Para acidentes com vítimas feridas, os altos coeficientes ($R^2 > 0,718$) em regiões metropolitanas sugerem que os acidentes com vítimas feridas podem estar relacionados a fatores como densidade populacional, volume de tráfego e infraestrutura viária, também destacados pelo modelo de MQO. Algumas áreas do interior também apresentaram bons resultados, principalmente nos estados do Maranhão, Piauí, Ceará, Rio Grande do Norte, Paraíba e Pernambuco. Esses coeficientes evidenciam que o modelo captura fatores relacionados a esse tipo de acidente.

Os valores inferiores registrados ($R^2 < 0,718$) foram identificados principalmente em alguns municípios da Bahia, Alagoas e Sergipe, indicando que modelos focados nessas regiões devem incluir variáveis explicativas adicionais, para explicar a variabilidade dessa variável dependente.

Figura 12 – Coeficiente de determinação (R^2) local, variável dependente acidentes com vítimas fatais



Legenda: O valor do coeficiente de determinação (R^2 local) representa o grau de ajuste do modelo estatístico em cada localidade.

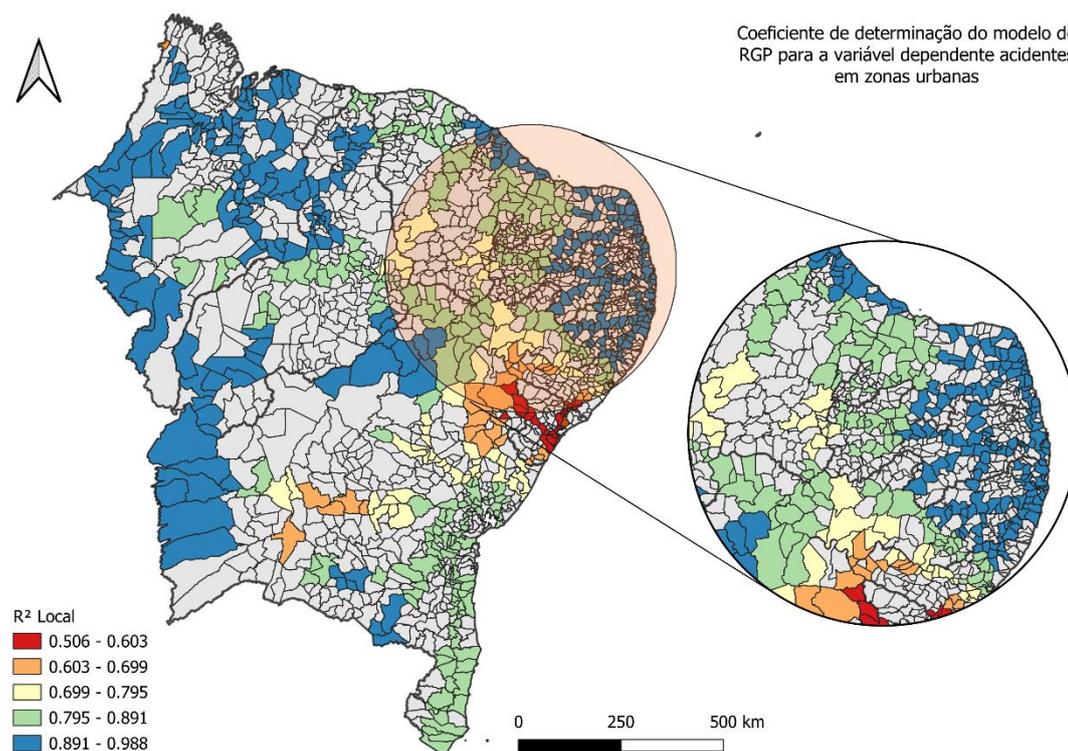
Fonte: O autor (2025).

Os coeficientes locais referentes a variável dependente de acidentes com vítimas fatais apresentaram maior variabilidade, com coeficientes medianos (R^2 entre 0,53 e 0,68). Os menores valores ($R^2 < 0,53$), refletindo uma maior dificuldade do modelo em explicar os acidentes com vítimas fatais, estiveram presentes em áreas do mais voltadas para os municípios do interior dos nove estados do Nordeste.

Nas regiões metropolitanas de São Luís, Teresina, Fortaleza, João Pessoa e Recife, os coeficientes são ligeiramente melhores ($R^2 > 0,68$). Isso indica que os fatores relacionados a estas capitais estão mais relacionados aos acidentes fatais, porém são inferiores aos observados em outras variáveis dependentes, demonstrando a baixa capacidade explicativa para essa variável.

As Figuras 13 e 14 apresentam os coeficientes de determinação (R^2) locais para as variáveis dependentes relacionadas à localização dos sinistros, acidentes urbanos (ACID_URB) e rurais (ACID_RUR).

Figura 13 – Coeficiente de determinação (R^2) local, variável dependente acidentes urbanos



Legenda: O valor do coeficiente de determinação (R^2 local) representa o grau de ajuste do modelo estatístico em cada localidade.

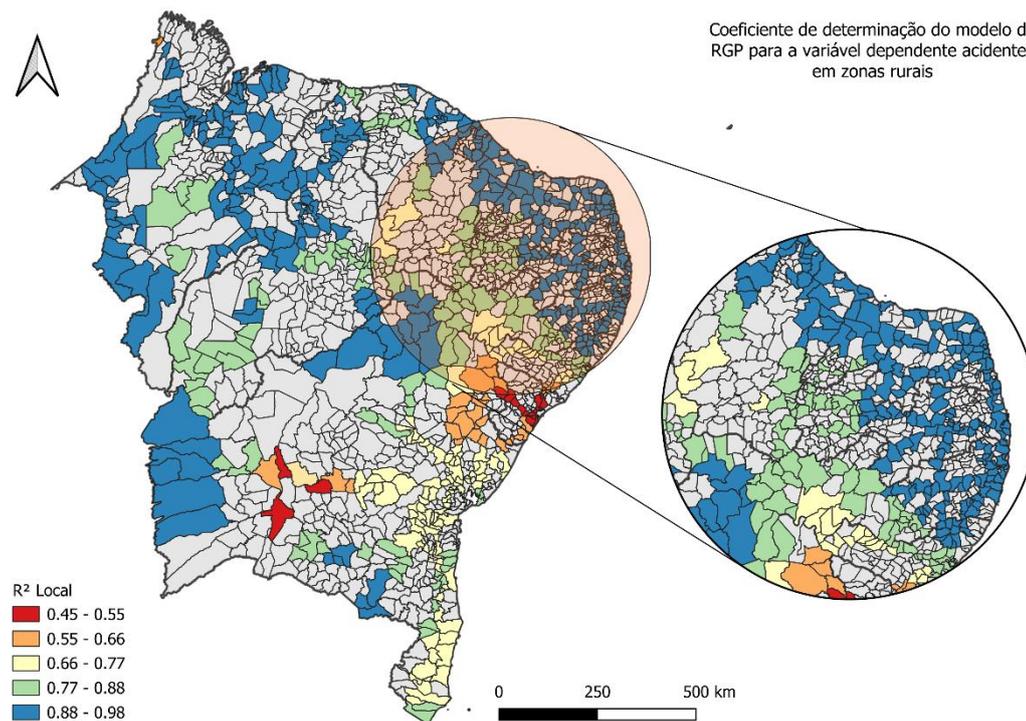
Fonte: O autor (2025).

Nessa figura, os valores de R^2 para acidentes urbanos obtiveram um desempenho geral satisfatório, com uma concentração de valores altos ($R^2 > 0,795$), principalmente nas capitais e regiões metropolitanas da região Nordeste. A tendência positiva também predomina em boa parte dos estados do Nordeste, com exceção do estado de Sergipe. Esse resultado demonstra uma maior homogeneidade das características dos sinistros em áreas urbanas e da capacidade do modelo em capturar os fatores explicativos relevantes, como extensão de vias, tráfego, características populacionais, IDHM e área municipal.

Os valores mais baixos ($R^2 < 0,699$), foram registrados principalmente no estado de Sergipe, parte de Alagoas e no estado da Bahia. Para essas regiões, demonstra-se a necessidade de considerar outras variáveis não consideradas ou

mesmo a presença de inconsistências nos dados de acidentes, principalmente nos registros do estado de Sergipe.

Figura 14 – Coeficiente de determinação (R^2) local, variável dependente acidentes rurais



Legenda: O valor do coeficiente de determinação (R^2 local) representa o grau de ajuste do modelo estatístico em cada localidade.

Fonte: O autor (2025).

Os coeficientes locais apresentam valores altos ($R^2 > 0,77$) que predominam nos estados do Maranhão, Piauí, Ceará, Rio Grande do Norte, Paraíba e Pernambuco. Esses agrupamentos demonstraram eficácia na explicação da variável dependente, relacionada aos acidentes em trechos rodoviários rurais. O resultado, tomando por base as variáveis explicativas mais significativas estatisticamente, relacionadas a extensão viária, volume de tráfego, população, densidade demográfica e área do município.

Os indicadores baixos ($R^2 < 0,66$) estiveram presentes em Sergipe, parte de Alagoas e alguns municípios da Bahia. Os resultados sugerem uma baixa capacidade explicativa nessas regiões, o que pode estar relacionado com a heterogeneidade das

condições viárias em regiões rurais, corroborando a falta de variáveis explicativas que expliquem de forma suficiente os acidentes nessas regiões.

Logo, a partir da análise dos coeficientes de determinação dos coeficientes de determinação (R^2) médios e dos coeficientes locais dos modelos, verificou-se que o melhor desempenho observado, foi identificado no modelo de acidentes totais (ACID_TOTAL). A partir daí, foi possível analisar os resultados das estimativas de β locais das variáveis independentes para esse modelo.

4.5.3 Distribuição espacial das estimativas de β locais

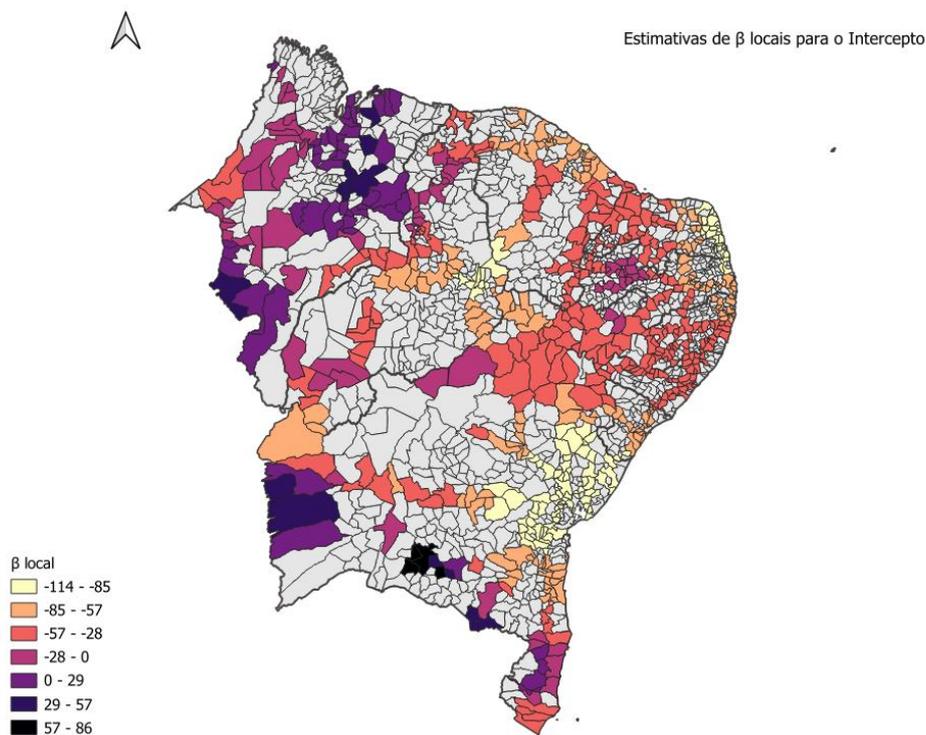
O modelo da variável que descreve a frequência de acidentes (ACID_TOTAL), foi analisado quanto a estimativa β de seus parâmetros. O modelo de RGP foi mensurado por meio da mediana dos indicadores locais. A Tabela 23 apresentou um comparativo para a estimativa β entre os dois modelos de MQO e RGP.

Tabela 23 – Estimativa de β para as variáveis do modelo ajustado de MQO e RGP, variável dependente acidentes totais (ACID_TOTAL)

Modelo Variáveis	MQO Estimativa	RGP Estimativa (mediana)
(Intercepto)	-74,390	-47,267
EXT_VIAS	0,477	0,501
VM DA	0,001	0,002
P_POP_OCUP	39,700	14,789
POPULACAO	0,00013	0,000314
DENS_DEM	0,005	-0,028
IDHM	100,500	61,617
AREA_UN	-0,004	-0,00615
DENS_VIA	69,290	74,521

Fonte: O autor (2025).

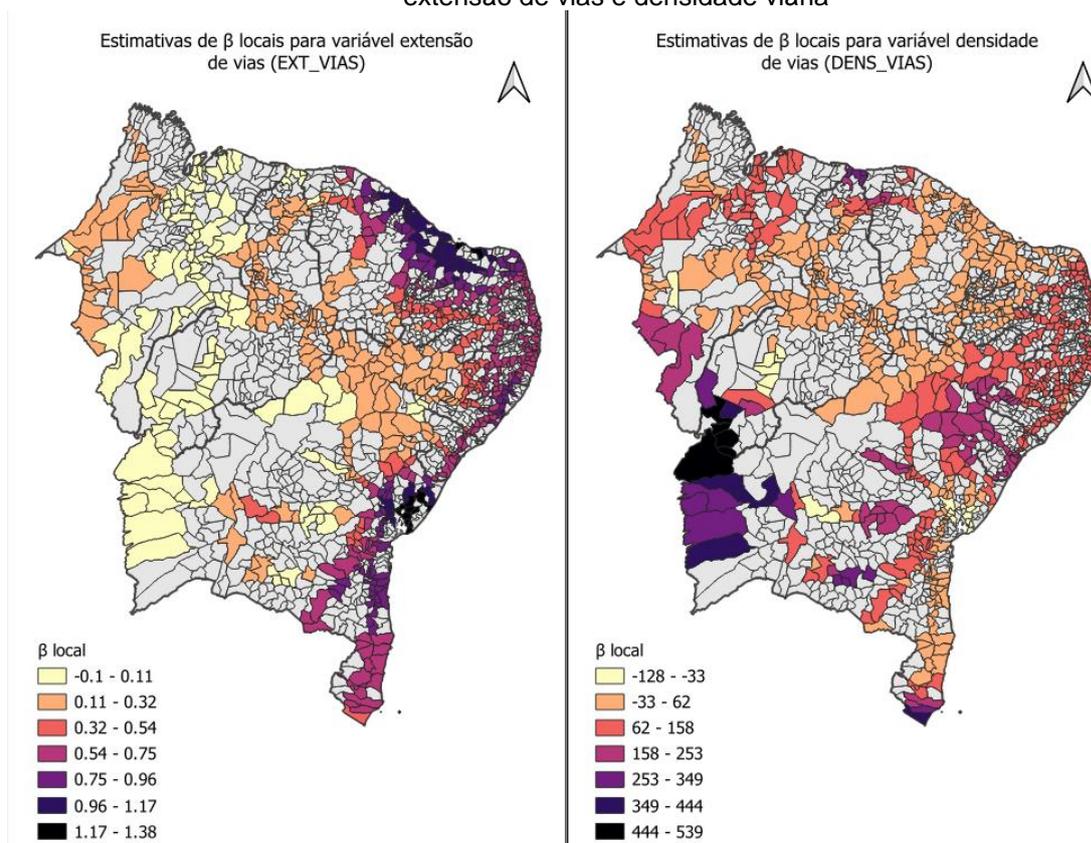
Por meio dessa tabela, é possível verificar que os coeficientes mantiveram sinais similares para os dois modelos, exceto na variável densidade demográfica, apesar de apresentar um valor próximo de zero nas duas estimativas. Esse resultado confirma as tendências observadas no modelo linear. As figuras 15, 16, 17, 18 e 19 a seguir apresentam as estimativas locais para o intercepto e todas as variáveis explicativas.

Figura 15 – Estimativa (β) local, modelo de acidentes totais (ACID_TOTAIS), para o Intercepto

Fonte: O autor (2025).

A análise das estimativas do intercepto do modelo traz informações relevantes sobre a distribuição espacial da frequência de acidentes. As regiões em que o intercepto apresenta valores positivos elevados indicam que, mesmo na ausência dos efeitos das variáveis explicativas, há uma tendência de maior incidência de acidentes. Por outro lado, interceptos negativos indicam que a frequência de acidentes tende a ser menor quando os fatores (variáveis explicativas) não são considerados, podendo indicar que as áreas com menor estimativa demonstram maior influência das variáveis explicativas.

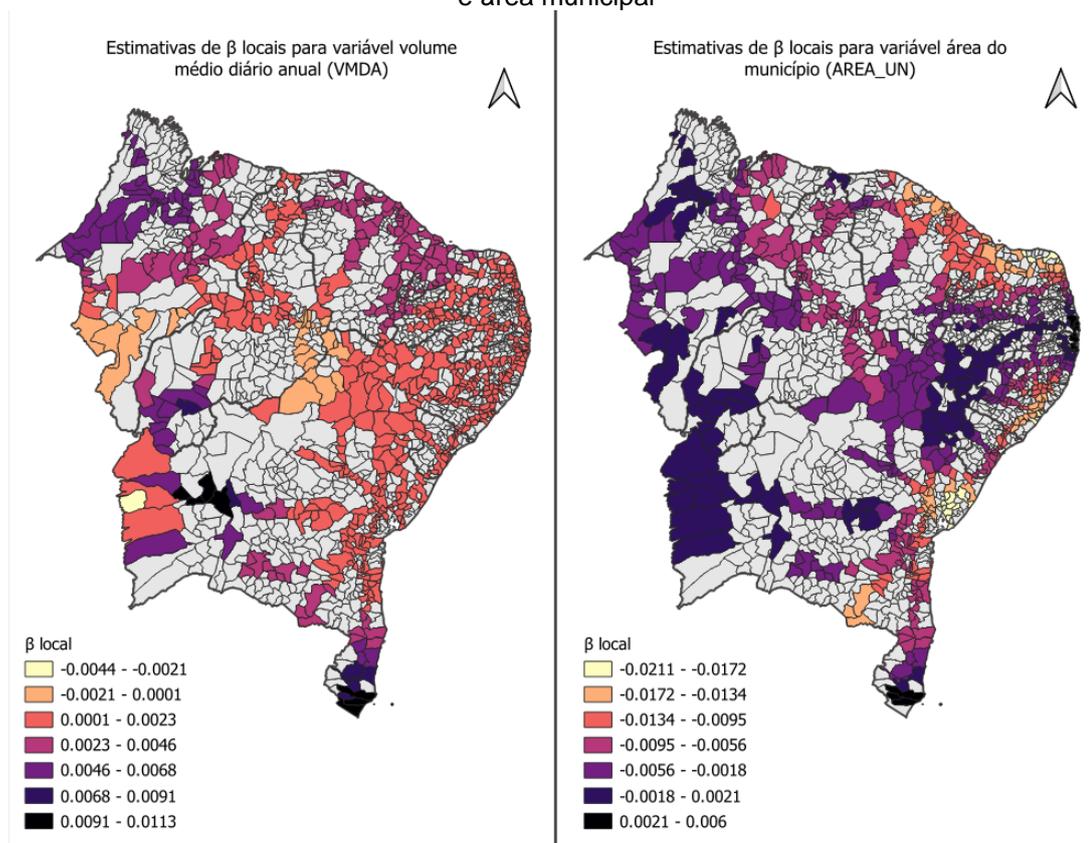
Figura 16 – Estimativa (β) local, modelo de acidentes totais (ACID_TOTAIS), para as variáveis extensão de vias e densidade viária



Fonte: O autor (2025).

Os resultados da GWR indicam que a extensão de vias (EXT_VIAS) e a densidade de vias (DENS_VIAS) apresentam padrões espaciais diferentes com relação à frequência de acidentes, nos dois casos, verificou-se a prevalência de estimativas positivas. No caso da extensão de vias, observa-se que os coeficientes locais são mais elevados nas regiões litorâneas e metropolitanas, sugerindo que a maior malha viária está fortemente associada ao aumento da ocorrência de acidentes. Por outro lado, áreas do interior apresentam coeficientes mais baixos, indicando que a extensão viária pode ter menor influência na frequência de acidentes nessas regiões. A densidade de vias, apresentou uma variação mais acentuada nos coeficientes, com valores elevados em certas regiões do interior, mas estimativas positivas em boa parte da região. Esse comportamento pode estar relacionado à concentração da malha viária em áreas urbanizadas, onde há maior fluxo de veículos.

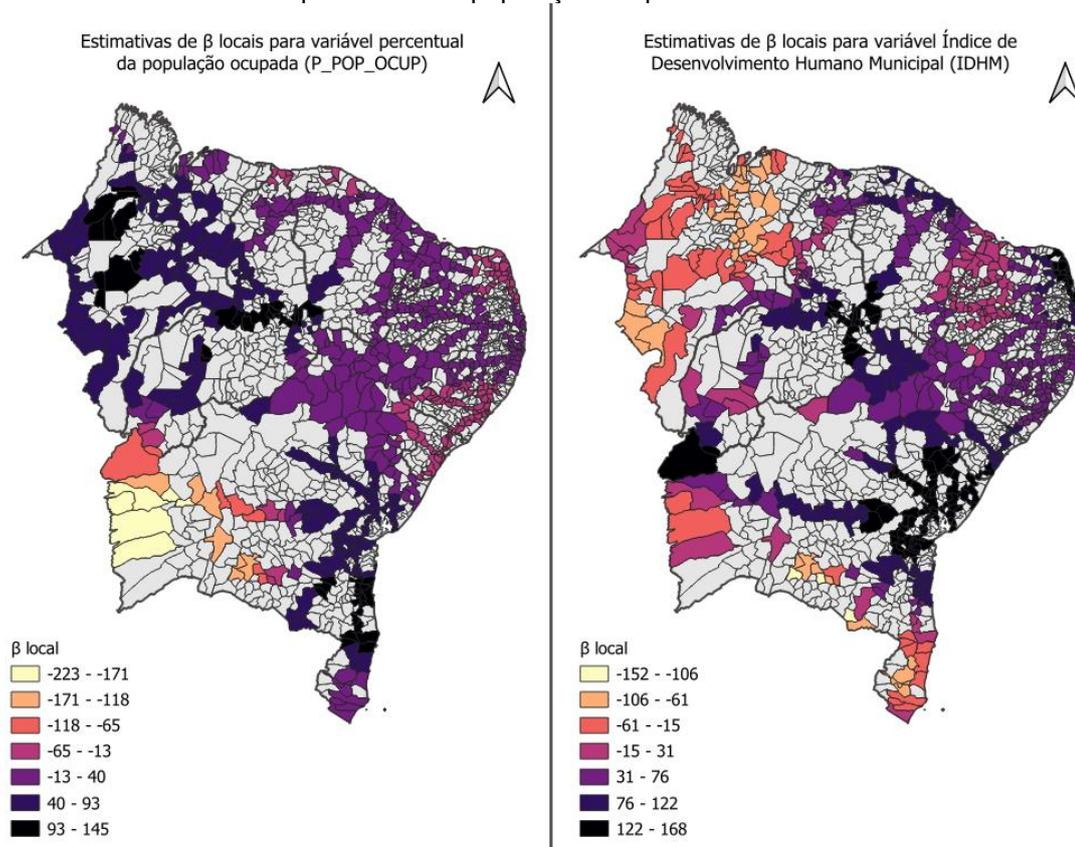
Figura 17 – Estimativa (β) local, modelo de acidentes totais (ACID_TOTAIS), para as variáveis VMDA e área municipal



Fonte: O autor (2025).

As estimativas β locais para o volume médio diário anual (VMDA) e a área do município (AREA_UN) também evidencia relações espaciais relevantes. Os coeficientes para o VMDA são positivos em boa parte dos municípios concentrados em áreas de grande fluxo viário, mas também em áreas do interior, sugerindo que o volume de tráfego tem um efeito positivo para a ocorrência de acidentes, uma vez que o aumento do fluxo de veículos eleva a probabilidade de sinistros rodoviários. Em contrapartida, a variável área do município apresenta coeficientes beta predominantemente negativos, mas próximos de zero, indicando que municípios com maior extensão territorial tendem a ter menor probabilidade de ocorrência de acidentes. Esse resultado pode estar relacionado ao fato de que, em municípios de maior área, a distribuição da malha viária e a dispersão podem reduzir a densidade de tráfego em determinados trechos, diminuindo o risco de acidentes.

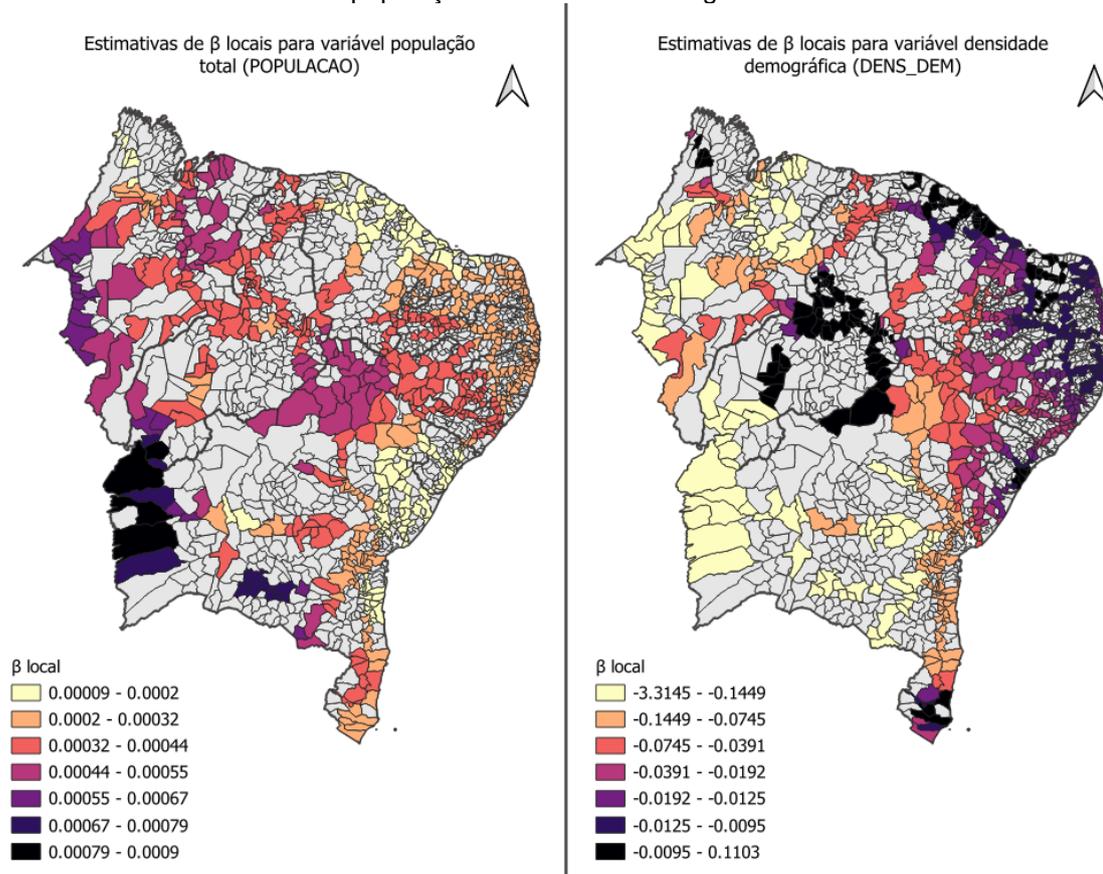
Figura 18 – Estimativa (β) local, modelo de acidentes totais (ACID_TOTAIS), para as variáveis percentual da população ocupada e IDHM



Fonte: O autor (2025).

Fatores socioeconômicos também demonstram relação direta e predominantemente positiva na ocorrência de acidentes. Conforme evidenciado pelas estimativas locais de percentual da população ocupada (P_POP_OCUP) e Índice de Desenvolvimento Humano Municipal (IDHM), os valores predominam para essas variáveis em boa parte da região nordeste, com concentração em regiões litorâneas e regiões metropolitanas, o que reforça a influência positiva dessas variáveis sobre os acidentes rodoviários. O aumento de população ocupada e de IDHM sugere o aumento da ocorrência de acidentes em muitos municípios da região Nordeste.

Figura 19 – Estimativa (β) local, modelo de acidentes totais (ACID_TOTAIS), para as variáveis população e densidade demográfica



Fonte: O autor (2025).

A relação entre a ocorrência de acidentes e a população total e a densidade demográfica apresentaram características distintas. Para a população, todos os coeficientes foram maiores que zero, indicando relação positiva sobre a accidentalidade e reforçando que quanto maior o aumento populacional, maior a probabilidade de ocorrência de sinistros. A densidade demográfica, com valores predominantemente negativos, próximos de zero, apresentou parâmetros mais fortes em muitas regiões metropolitanas e em suas proximidades, destacando a influência das regiões metropolitanas sobre os acidentes rodoviários.

4.6 DISCUSSÃO GERAL DOS RESULTADOS

Os modelos de MQO forneceram inferências relevantes a respeito dos dados e constituíram de dados importantes que permitiram uma visão mais ampla a respeito das interações presentes nas variáveis, sendo atestadas por meio da literatura no tema. Apesar disso, apresentou dificuldades no que diz respeito a sua capacidade

explicativa, para algumas variáveis, a presença de outliers influenciou negativamente o modelo, com diferença elevada de escalas, levando a ausência da distribuição normal dos resíduos e destacada heterocesticidade. O modelo apresentou resultados interessantes, mas não foi o mais adequado para explicar de maneira mais assertiva o conjunto de dados, que apresentam relações espaciais que foram exploradas.

O Índice de Moran Global e Local, permitiu identificar a presença da dependência espacial entre municípios próximos, revelando que as ocorrências de acidentes são influenciadas por fatores regionais. Com base nos resíduos do modelo de MQO, o índice local identificou algumas localidades que merecem atenção quanto a acidentalidade, a mesorregião do Jaguaribe e a região metropolitana de Fortaleza (Ceará), o Agreste pernambucano, o Seridó potiguar, o Vale do Rio São Francisco (entre Bahia, Pernambuco e municípios próximos a Vitória da Conquista e Porto Seguro).

Os resultados da RGP demonstraram melhorias ao considerar os padrões espaciais nos sinistros rodoviários. O modelo da variável dependente que descrevia a frequência total de acidentes apresentou os melhores resultados, de acordo com os indicadores da performance dos modelos. Apesar disso, os resultados de outros modelos focados em aspectos específicos dos acidentes (gravidade, localização, extensão viária e tráfego) também obtiveram resultados interessantes.

No que diz respeito aos coeficientes das variáveis explicativas do modelo de RGP, as tendências do modelo linear se mantiveram para boa parte das variáveis. A análise das estimativas de β locais permitiu compreender melhor os padrões espaciais, destacando a influência com sentido positivo dos elevados fluxos de veículos, população, extensão viária, população ocupada e IDHM.

5 CONCLUSÃO

Essa dissertação teve como objetivo compreender a relação entre as variáveis socioeconômicas, demográficas e a infraestrutura viária na distribuição espacial e nos índices de acidentes rodoviários ocorridos no Nordeste brasileiro. A hipótese levantada pontuava que a ocorrência de acidentes é influenciada por fatores socioeconômicos, demográficos e de infraestrutura viária. Com base na análise dos resultados obtidos, é possível afirmar que os objetivos estabelecidos foram atingidos e que a hipótese levantada foi confirmada, permitindo compreender melhor as relações espaciais e a dinâmica dos acidentes rodoviários na região Nordeste.

A investigação das características socioeconômicas e demográficas revelou padrões estatisticamente significativos. As análises identificaram, principalmente, uma forte relação entre os acidentes e variáveis como o Índice de Desenvolvimento Humano Municipal (IDHM), o percentual da população com baixa renda, densidade demográfica e tamanho populacional. Esses resultados evidenciam a tendência de crescimento dos índices de acidentes conforme a população aumenta, com os grandes centros concentrando maior parte dos registros de sinistros.

O modelo espacial também permitiu identificar a influência do percentual da população ocupada e o Índice de Desenvolvimento Humano Municipal, contribuindo para aumentar a probabilidade de ocorrência de acidentes na região. A distribuição espacial para essas variáveis não demonstrou um padrão específico bem definido, entretanto as regiões metropolitanas e muitas áreas do interior nordestino apresentaram maiores estimativas nesses parâmetros.

Logo, no que se refere aos indicadores socioeconômicos, uma das principais limitações identificadas foi a ausência de um índice unificado de vulnerabilidade social. Essa lacuna impossibilitou uma análise mais aprofundada sobre a relação direta entre vulnerabilidade social e a incidência de acidentes rodoviários, limitando a consideração de outros aspectos relevantes não contemplados por este estudo. Assim, recomenda-se que pesquisas futuras incluam variáveis que representem melhor esses fatores, permitindo uma avaliação com mais informações a respeito dessas relações.

Os fatores relacionados à infraestrutura viária e territorial evidenciaram a influência da exposição dos usuários da via nas ocorrências. As variáveis referentes à área municipal, à extensão e à densidade viária apresentaram alta significância

estatística. Observou-se que quanto maior a área do municipal, menor a ocorrência de acidentes, possivelmente influenciada pela dispersão dos registros em territórios menos populosos e com menor fluxo viário. Por outro lado, a extensão e densidade viária, demonstraram uma influência positiva na frequência de acidentes, pois a presença de maior infraestrutura, por exemplo em regiões metropolitanas, quando ausente de planejamento e gestão adequados, pode levar a um aumento dos acidentes rodoviários.

A exploração da relação entre as características do tráfego e os acidentes rodoviários, reforçou a importância do fluxo veicular na dinâmica dos sinistros. A alta significância estatística e a presença de coeficientes positivos em todos os modelos demonstram que o fluxo de veículos atua como uma variável de exposição, de modo que um aumento no volume de tráfego está associado a uma maior probabilidade de ocorrência de acidentes rodoviários. Outros modelos podem se aprofundar na relação entre essa variável e os sinistros rodoviários, compreendendo melhor as concentrações espaciais que mais influenciam essas ocorrências.

No que se refere aos fatores relacionados à frota e aos condutores, variáveis como a proporção de motocicletas na frota, índices de motorização e proporção de condutores até 24 anos, não apresentaram significância estatística nos modelos. Isso indica que essas variáveis não contribuíram de maneira substancial para explicar a ocorrência dos sinistros rodoviários dentro do escopo desta dissertação. Verifica-se como uma limitação do estudo, sugerindo que essas variáveis poderiam ser abordadas por pesquisas futuras sob perspectivas diferentes e com metodologias mais eficientes para compreender melhor tais relações, uma vez que os modelos lineares e geponderados não forneceram subsídios que permitissem conclusões mais relevantes.

No âmbito dos modelos de Mínimos Quadrados Ordinários (MQO), observou-se a dificuldade em analisar as variáveis referentes às taxas de acidentes, por tráfego e por extensão viária. Sugere-se a implementação de modelos que abordem de forma mais precisa as taxas de acidentes em relação à extensão viária e volume de tráfego, aprimorando a capacidade explicativa de tais modelos.

A identificação de padrões espaciais nos acidentes, por meio da aplicação de ferramentas estatísticas, permitiu compreender melhor a dinâmica espacial dos sinistros viários. O Índice de Moran e o modelo de Regressão Geograficamente Ponderada (RGP) foram fundamentais para detectar a autocorrelação espacial em

alguns modelos, permitiram a identificação de áreas críticas em que as variáveis explicativas influenciavam mais diretamente a variável dependente, o que permitiu extrair mais informações a respeito da distribuição espacial dos acidentes rodoviários.

A determinação das áreas críticas, foi fundamental para identificar as regiões que devem ser o foco para a implementação de ações mitigadoras. As regiões metropolitanas demonstraram relação muito forte com os sinistros, dentre elas é possível destacar: Fortaleza, João Pessoa, Natal, Recife, São Luís e Teresina. Essas regiões, devido à sua relevância socioeconômica, tamanho populacional, elevado fluxo de tráfego, necessitam de monitoramento constante e a adoção de medidas mitigadoras para redução dos acidentes.

Além das regiões metropolitanas, algumas áreas do interior também apresentaram altos agrupamentos de acidentes para boa parte dos modelos, indicando a necessidade de intervenções. Entre elas, destacam-se o Vale do Rio São Francisco (entre Pernambuco e Bahia), o Agreste pernambucano, e os estados do Maranhão, Piauí e Rio Grande do Norte.

Em contrapartida, algumas áreas que merecem atenção devido aos agrupamentos de menores valores, estados como Sergipe e parte de Alagoas, que apresentaram resultados inferiores. Possivelmente, esse resultado pode ser devido a falhas na qualidade dos dados gerando modelos que não conseguem classificar os padrões espaciais nessas regiões. Portanto, sugere a necessidade de estudos mais detalhados, com utilização de modelos aprimorados, para capturar os fatores não identificados nessas regiões.

Outro ponto relevante para futuras pesquisas é a análise mais aprofundada da acidentalidade envolvendo motociclistas e ciclistas. Os modelos utilizados neste trabalho não foram suficientemente eficazes para capturar essa dinâmica, indicando a necessidade de metodologias alternativas para explorar esses fatores.

A dinâmica espacial observada permitiu inferir aspectos importantes sobre os acidentes na região Nordeste, fornecendo evidências para a formulação de políticas públicas voltadas à segurança viária. Os resultados indicam que medidas integradas, envolvendo infraestrutura viária bem planejada, educação e fiscalização viária mais rigorosa, são fundamentais para reduzir os índices de acidentes nas rodovias federais do Nordeste brasileiro. Essas iniciativas devem ser direcionadas, prioritariamente, para as áreas críticas, garantindo maior eficácia nas estratégias de prevenção e controle da acidentalidade rodoviária.

Dessa forma, esta dissertação contribui significativamente para o entendimento dos fatores que influenciam a ocorrência de acidentes rodoviários no Nordeste brasileiro, ressaltando a importância da integração entre variáveis socioeconômicas, demográficas e de infraestrutura viária na formulação de alternativas eficazes para mitigação das ocorrências. A identificação de padrões espaciais e áreas críticas reforçam a necessidade de políticas públicas direcionadas e baseadas em evidências. Assim, espera-se que os achados neste estudo sirvam como referência para futuras pesquisas e para o desenvolvimento de ações concretas que visem à redução dos sinistros rodoviários, promovendo maior segurança para as pessoas.

REFERÊNCIAS

ANSELIN, L. Local Indicators of Spatial Association—LISA. **Geographical Analysis**, v. 27, n. 2, p. 93–115, abr. 1995.

AZETSOP, J. Social Justice Approach to Road Safety in Kenya: Addressing the Uneven Distribution of Road Traffic Injuries and Deaths across Population Groups. **Public Health Ethics**, v. 3, n. 2, p. 115–127, 1 jul. 2010.

BARROSO JUNIOR, G. T.; BERTHO, A. C. S.; VEIGA, A. D. C. A letalidade dos acidentes de trânsito nas rodovias federais brasileiras. **Revista Brasileira de Estudos de População**, v. 36, p. 1–22, 16 jul. 2019.

BLAIZOT, S. et al. Injury incidence rates of cyclists compared to pedestrians, car occupants and powered two-wheeler riders, using a medical registry and mobility data, Rhône County, France. **Accident Analysis & Prevention**, v. 58, p. 35–45, set. 2013.

BRUNSDON, C.; FOTHERINGHAM, A. S.; CHARLTON, M. E. Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. **Geographical Analysis**, v. 28, n. 4, p. 281–298, out. 1996.

CAI, Q. et al. Macro-level pedestrian and bicycle crash analysis: Incorporating spatial spillover effects in dual state count models. **Accident Analysis & Prevention**, v. 93, p. 14–22, ago. 2016.

CAI, Q.; ABDEL-ATY, M.; LEE, J. Macro-level vulnerable road users crash analysis: A Bayesian joint modeling approach of frequency and proportion. **Accident Analysis & Prevention**, v. 107, p. 11–19, out. 2017.

CHANG CHIEN, Y.-M.; CARVER, S.; COMBER, A. Using geographically weighted models to explore how crowdsourced landscape perceptions relate to landscape physical characteristics. **Landscape and Urban Planning**, v. 203, p. 103904, nov. 2020.

CHARLTON, Martin; FOTHERINGHAM, Stewart; BRUNSDON, Chris. Geographically weighted regression. **White paper. National Centre for Geocomputation. National University of Ireland Maynooth**, v. 2, 2009.

CHEN, C. et al. Investigating driver injury severity patterns in rollover crashes using support vector machine models. **Accident Analysis & Prevention**, v. 90, p. 128–139, maio 2016.

CHIMBA, D.; MUSINGUZI, A.; KIDANDO, E. Associating pedestrian crashes with demographic and socioeconomic factors. **Case Studies on Transport Policy**, v. 6, n. 1, p. 11–16, mar. 2018.

CONFEDERAÇÃO NACIONAL DO TRANSPORTE - CNT. **Painel CNT de consultas dinâmicas dos acidentes rodoviários 2022**. 2023. Disponível em: <https://www.cnt.org.br/painel-acidente>. Acesso em: 03 out. 2024.

DESLANDES, S. F.; SILVA, C. M. F. P. D. Análise da morbidade hospitalar por acidentes de trânsito em hospitais públicos do Rio de Janeiro, RJ, Brasil. **Revista de Saúde Pública**, v. 34, n. 4, p. 367–372, ago. 2000.

DEPARTAMENTO NACIONAL DE INFRAESTRUTURA DE TRANSPORTES - DNIT. Volume Médio Diário Anual (VMDA) - 2021. Departamento Nacional de Infraestrutura de Transportes, 2021. Disponível em: <https://servicos.dnit.gov.br/vgeo/>. Acesso em: 03 out. 2024.

GOMES, Marcos José Timbó Lima. **Proposta metodológica de avaliação dos efeitos espaciais de dependência e heterogeneidade dos acidentes de trânsito em diferentes configurações de áreas com a incorporação da superdispersão**. 2018. 111 f. Tese (Doutorado em Engenharia de Transportes) – Universidade Federal do Ceará, Fortaleza, 2018.

GONZÁLEZ, S. R.; LOUKAITOU-SIDERIS, A.; CHAPPLE, K. Transit neighborhoods, commercial gentrification, and traffic crashes: Exploring the linkages in Los Angeles and the Bay Area. **Journal of Transport Geography**, v. 77, p. 79–89, maio 2019.

GUERRA, E.; DONG, X.; KONDO, M. Do Denser Neighborhoods Have Safer Streets? Population Density and Traffic Safety in the Philadelphia Region. **Journal of Planning Education and Research**, p. 0739456X19845043, 5 maio 2019.

HUANG, Y.; WANG, X.; PATTON, D. Examining spatial relationships between crashes and the built environment: A geographically weighted regression approach. **Journal of Transport Geography**, v. 69, p. 221–233, maio 2018.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. **Cidades e Estados - Panorama**. Instituto Brasileiro de Geografia e Estatística, 2023. Disponível em: <https://cidades.ibge.gov.br/brasil/al/panorama>. Acesso em: 03 out. 2024.

KARLAFTIS, M. G.; GOLIAS, I. Effects of road geometry and traffic volumes on rural roadway accident rates. **Accident Analysis & Prevention**, v. 34, n. 3, p. 357–365, maio 2002.

KARLAFTIS, M. G.; TARKO, A. P. Heterogeneity considerations in accident modeling. **Accident Analysis & Prevention**, v. 30, n. 4, p. 425–433, jul. 1998.

KIM, K.; PANT, P.; YAMASHITA, E. Accidents and Accessibility: Measuring Influences of Demographic and Land Use Variables in Honolulu, Hawaii. **Transportation Research Record: Journal of the Transportation Research Board**, v. 2147, n. 1, p. 9–17, jan. 2010.

KONG, C.; YANG, J. Logistic regression analysis of pedestrian casualty risk in passenger vehicle collisions in China. **Accident Analysis & Prevention**, v. 42, n. 4, p. 987–993, jul. 2010.

LASCALA, E. A.; GERBER, D.; GRUENEWALD, P. J. Demographic and environmental correlates of pedestrian injury collisions: a spatial analysis. **Accident Analysis & Prevention**, v. 32, n. 5, p. 651–658, set. 2000.

LEE, J.; ABDEL-ATY, M.; JIANG, X. Multivariate crash modeling for motor vehicle and non-motorized modes at the macroscopic level. **Accident Analysis & Prevention**, v. 78, p. 146–154, maio 2015.

LI, X. et al. Do underserved and socially vulnerable communities observe more crashes? A spatial examination of social vulnerability and crash risks in Texas. **Accident Analysis & Prevention**, v. 173, p. 106721, ago. 2022.

LIMA, L. C. de; CRUZ JUNIOR, V. da S. Estudo dos acidentes de trânsito no Brasil à luz da Pesquisa Nacional de Saúde 2013. In: VII CONGRESO DE LA ASOCIACIÓN LATINOAMERICANA DE POBLACIÓN e XX ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 17-22 out. 2016, Foz do Iguaçu. **Anais [...]**. Foz do Iguaçu: ABEP, 2016.

LIU, J.; DAS, S.; KHAN, M. N. Decoding the impacts of contributory factors and addressing social disparities in crash frequency analysis. **Accident Analysis & Prevention**, v. 194, p. 107375, jan. 2024.

LORD, D.; MANNERING, F. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. **Transportation Research Part A: Policy and Practice**, v. 44, n. 5, p. 291–305, jun. 2010.

LUCAS, K. Transport and social exclusion: Where are we now? **Transport Policy**, v. 20, p. 105–113, mar. 2012.

LUCAS, K. *et al.* Transport poverty and its adverse social consequences. **Proceedings of the Institution of Civil Engineers - Transport**, v. 169, n. 6, p. 353–365, dez. 2016.

MACEDO, M. R. O. B. C. ; MAIA, M. L. A. ; RABBANI, E. R. K. ; LIMA NETO, O. ; **ANDRADE, M. O.** . Traffic accident prediction model for rural highways in Pernambuco. **CASE STUDIES ON TRANSPORT POLICY**, v. 9, p. 2000, 2021.

MALTA, D. C. *et al.* Análise das ocorrências das lesões no trânsito e fatores relacionados segundo resultados da Pesquisa Nacional por Amostra de Domicílios (PNAD) Brasil, 2008. **Ciência & Saúde Coletiva**, v. 16, n. 9, p. 3679–3687, set. 2011.

MESA-ARANGO, R. et al. Influence of Socioeconomic Conditions on Crash Injury Severity for an Urban Area in a Developing Country. **Transportation Research Record: Journal of the Transportation Research Board**, v. 2672, n. 31, p. 41–53, dez. 2018.

MINISTÉRIO DOS TRANSPORTES. **Estatísticas da frota de veículos por município**. Secretaria Nacional de Trânsito (Senatran), 2022. Disponível em: <https://www.gov.br/transportes/pt-br/assuntos/transito/conteudo-Senatran/estatisticas-frota-de-veiculos-senatran>. Acesso em: 03 out. 2024.

MORAN, P.A.P. **The interpretation of statistical maps**. J. R. Stat. Soc. Ser. B 10, 2, 1948. Disponível em: <https://doi.org/10.1111/j.2517-6161.1948.tb00012.x>.

NANTULYA, V. M.; REICH, M. R. Equity dimensions of road traffic injuries in low- and middle-income countries. **Injury Control and Safety Promotion**, v. 10, n. 1–2, p. 13–20, abr. 2003.

NOLAND, R. B.; LAHAM, M. L. Are low income and minority households more likely to die from traffic-related crashes? **Accident Analysis & Prevention**, v. 120, p. 233–238, nov. 2018.

O'BRIEN, R. M. A Caution Regarding Rules of Thumb for Variance Inflation Factors. **Quality & Quantity**, v. 41, n. 5, p. 673–690, 11 set. 2007.

ORGANIZAÇÃO MUNDIAL DA SAÚDE – OMS. **Global Plan for the Decade of Action for Road Safety 2021-2030**. Genebra. Organização Mundial da Saúde, 2021. Disponível em: <<https://www.who.int/pt/publications/m/item/global-plan-for-the-decade-of-action-for-road-safety-2021-2030>>. Acesso em: 29 de outubro de 2024.

OSHAN, T. et al. mgwr: A Python Implementation of Multiscale Geographically Weighted Regression for Investigating Process Spatial Heterogeneity and Scale. **ISPRS International Journal of Geo-Information**, v. 8, n. 6, p. 269, 8 jun. 2019.

POLÍCIA RODOVIÁRIA FEDERAL - PRF. **Dados abertos de acidentes de trânsito - 2022**. Polícia Rodoviária Federal, 2022. Disponível em: <https://www.gov.br/prf/pt-br/aceso-a-informacao/dados-abertos/dados-abertos-da-prf>. Acesso em: 03 out. 2024.

QUDDUS, M. A. Modelling area-wide count outcomes with spatial correlation and heterogeneity: An analysis of London crash data. **Accident Analysis & Prevention**, v. 40, n. 4, p. 1486–1497, jul. 2008.

SHABRINA, Z.; BUYUKLIEVA, B.; NG, M. K. M. Short-term rental platform in the urban tourism context: a geographically weighted regression (GWR) and a multiscale GWR (MGWR) approaches. **Geographical Analysis**, v. 53, p. 686–707, 2021. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1111/gean.12259>. Acesso em: 20 de dez. de 2024.

SIDDIQUI, C.; ABDEL-ATY, M.; CHOI, K. Macroscopic spatial analysis of pedestrian and bicycle crashes. **Accident Analysis & Prevention**, v. 45, p. 382–391, mar. 2012.

SILVA, M. N., GOMES, M. J. T. L., & CUNTO, F. J. C. Investigating temporal trends in risk factors related to injury severity of crashes with pedestrians in urban areas. **Traffic Injury Prevention**, p. 1–7, fev. 2025.

SU, J.; SZE, N. N.; BAI, L. A joint probability model for pedestrian crashes at macroscopic level: Roles of environment, traffic, and population characteristics. **Accident Analysis & Prevention**, v. 150, p. 105898, fev. 2021.

WANG, J.; HUANG, H.; ZENG, Q. The effect of zonal factors in estimating crash risks by transportation modes: Motor vehicle, bicycle and pedestrian. **Accident Analysis & Prevention**, v. 98, p. 223–231, jan. 2017.

WIER, M. et al. An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. **Accident Analysis & Prevention**, v. 41, n. 1, p. 137–145, jan. 2009.

XU, P.; HUANG, H. Modeling crash spatial heterogeneity: random parameter versus geographically weighting. **Accident Analysis e Prevention**, v. 75, p. 16-25, 2015.

ZAFRI, N. M.; KHAN, A. A spatial regression modeling framework for examining relationships between the built environment and pedestrian crash occurrences at macroscopic level: A study in a developing country context. **Geography and Sustainability**, v. 3, n. 4, p. 312–324, dez. 2022.

ZHENG, L. et al. All Accidents are Not Equal: Using Geographically Weighted Regressions Models to Assess and Forecast Accident Impacts. In: 3rd INTERNATIONAL CONFERENCE ON ROAD SAFETY AND SIMULATION, 14-16 set. 2011, Indianapolis, USA. **Proccedings [...]**. 2011.