



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO TECNOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TRANSPORTES E
GESTÃO TERRITORIAL

Loany Steffane Rezende Falcão

**Determinantes da Gravidade dos Sinistros de Trânsito em Rodovias Federais
Utilizando Regressão Logística Binária**

Florianópolis
2024

Loany Steffane Rezende Falcão

**Determinantes da Gravidade dos Sinistros de Trânsito em Rodovias Federais
Utilizando Regressão Logística Binária**

Dissertação submetida ao Programa de Pós-Graduação em Engenharia de Transportes e Gestão Territorial da Universidade Federal de Santa Catarina como requisito parcial para a obtenção do título de Mestra em Sistemas de Transportes.

Orientador(a): Prof. Fernando Seabra, Dr.

Florianópolis

2024

Falcão, Loany Steffane Rezende

Determinantes da Gravidade dos Sinistros de Trânsito em Rodovias Federais Utilizando Regressão Logística Binária / Loany Steffane Rezende Falcão ; orientador, Fernando Seabra, 2024.

115 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia de Transportes e Gestão Territorial, Florianópolis, 2024.

Inclui referências.

1. Engenharia de Transportes e Gestão Territorial. 2. segurança viária. 3. regressão logística binária. 4. severidade. 5. sistemas de transportes. I. Seabra, Fernando. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia de Transportes e Gestão Territorial. III. Título.

Loany Steffane Rezende Falcão

**DETERMINANTES DA GRAVIDADE DOS SINISTROS DE TRÂNSITO EM
RODOVIAS FEDERAIS UTILIZANDO REGRESSÃO LOGÍSTICA BINÁRIA**

O presente trabalho em nível de mestrado foi avaliado e aprovado por banca
examinadora composta pelos seguintes membros:

Prof. Arnaldo Debatin Neto, Dr.
Universidade Federal de Santa Catarina

Prof. Eduardo Lobo, Dr.
Universidade Federal de Santa Catarina

Prof. Flávio Jose Craveiro Cunto, Dr.
Universidade Federal do Ceará

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi
julgado adequado para obtenção do título de mestre em Mestre em Engenharia de
Transportes e Gestão

Prof. Dr.(a) Rogério Cid Bastos
Coordenador(a) do Programa

Prof. Dr. Fernando Seabra
Orientador

Florianópolis, 11 de setembro de 2024.

AGRADECIMENTOS

À minha mãe Eva, pai Fábio (in memoriam) e irmã Nayara que me apoiaram em todas as escolhas e acreditaram no meu potencial desde o início.

Ao Ícaro, meu parceiro, pelo incentivo constante e pelo apoio incondicional, mesmo nos momentos em que minha presença não era frequente. Sua compreensão e paciência foram fundamentais para que eu pudesse buscar meus sonhos.

Ao meu orientador, prof. Fernando Seabra minha imensa gratidão. Tive a sorte de ser orientada de um professor que me deu todo o suporte possível, com paciência e disponibilidade durante todo curso.

Às minhas amigas Carol, Letícia, Thálita e Sara, que, mesmo à distância, torcem e vibram por cada pequena conquista. Vocês são meu porto seguro e minha fonte inesgotável de inspiração e força.

Aos colegas e amigos da Arteris, que me acolheram com generosidade e de quem tanto aprendi nesses anos, meu sincero agradecimento por todo o conhecimento compartilhado e pelo apoio.

RESUMO

Reconhecendo a necessidade de reduzir a gravidade dos sinistros de trânsito, especialmente considerando que no Brasil esses sinistros são a oitava maior causa de mortes, este trabalho busca identificar os principais fatores que levam à ocorrência de vítimas graves ou fatais nas rodovias federais brasileiras, focando nos estados de Minas Gerais, São Paulo, Paraná e Santa Catarina. Para isso, realizou-se uma análise sistemática da literatura, que fundamentou a escolha do modelo de regressão logística binária para analisar a gravidade dos sinistros, considerando exclusivamente vítimas moderadas, graves e fatais em veículos automotores. As variáveis analisadas incluíram características das vítimas, dos veículos, das vias e das condições climáticas. Adicionalmente, uma variável de agrupamento utilizando o método DBSCAN foi incluída para identificar áreas críticas de sinistros. As análises revelaram que a gravidade dos sinistros tende a aumentar quando o condutor é do gênero feminino, em condições de pista molhada, em áreas urbanas, em regiões classificadas como pontos críticos e em curvas acentuadas com pista molhada. Em contrapartida, a gravidade dos sinistros diminui durante período noturno, inverno, em declives, quando há envolvimento de veículos pesados ou veículos de passageiros e com o aumento do volume de tráfego. Além disso, sinistros envolvendo mais de um veículo e um deles é veículo leve ou pesado também mostram uma menor gravidade.

Palavras-chave: sistemas de transportes, segurança viária; regressão logística binária, severidade.

ABSTRACT

Recognizing the urgent need to reduce the severity of traffic accidents, especially considering that in Brazil these accidents are the eighth leading cause of death, this study aims to identify the main factors that lead to severe or fatal injuries on federal highways in Brazil, focusing on the states of Minas Gerais, São Paulo, Paraná, and Santa Catarina. To achieve this, a systematic literature review was conducted, which supported the selection of the binary logistic regression model to analyze the severity of accidents, focusing exclusively on moderate, severe, and fatal victims in motor vehicles. The variables analyzed included characteristics of the victims, vehicles, roads, and weather conditions. Additionally, a clustering variable using the DBSCAN method was included to identify critical areas for accidents. The analysis revealed that the severity of accidents tends to increase when the driver is female, under wet road conditions, in urban areas, in regions classified as critical areas, and on sharp curves with wet roads. Conversely, the severity of accidents decreases during the night, in winter, on downhill slopes, when heavy vehicles or cars are involved, and with increased traffic volume. Furthermore, accidents involving more than one vehicle, where one is a car or heavy vehicle, also show reduced severity.

Keywords: transport system, road safety, binomial logistic regression, severity.

LISTA DE FIGURAS

Figura 1 - Componentes de custo dos sinistros nas rodovias federais (2014)	20
Figura 2 - Frota de veículos e número de fatalidades de 2008 a 2022.....	27
Figura 3 – Países de acordo com as leis de velocidade vigentes em 2017..	31
Figura 4 - Distribuição da Malha Rodoviária Brasileira	33
Figura 5 - Curvas ROC dos diferentes tipos de discriminação.	55
Figura 6 - Agrupamento com algoritmo <i>k-means</i> (a) $K = 2$; (b) $K = 3$; (c) $K = 4$	61
Figura 7 - Representação dos passos para o agrupamento do algoritmo hierárquico.....	63
Figura 8 - As células em azul são chamadas de populares.....	64
Figura 9 - Resultados do uso do agrupamento DBSCAN	65
Figura 10 - Etapas para a definição do grupo.....	66
Figura 11 - Fluxograma metodológico	69
Figura 12 - Mapa das rodovias analisadas	70
Figura 13 – Tráfego da Fernão Dias na praça de pedágio 6, sentido norte em 2022	73
Figura 14 - Influência de cada praça de pedágio.....	73
Figura 15 – Soma de UPS por ano por tipo de sinistro.....	77
Figura 16 - Gravidade das Vítimas por Porcentagem do total	78
Figura 17 - Porcentagem de vítimas por Concessionária	79
Figura 18 - Mapa com a geolocalização dos 15 grupos com maior UPS	82
Figura 19 - 15 Grupos com maior número de UPS da Litoral Sul.....	83
Figura 20 - 15 Grupos com maior número de UPS da Régis Bittencourt	84
Figura 21 - 15 Grupos com maior número de UPS da Fernão Dias	85

LISTA DE TABELAS

Tabela 1 - Dissertações similares PPGTG	16
Tabela 2 - Tipos de classificação de sinistros de trânsito.....	19
Tabela 3 - Distribuição das fatalidades no trânsito por Estado.....	26
Tabela 4 - Palavras-chave utilizadas na revisão sistemática da literatura....	35
Tabela 5 - Trabalhos acadêmicos selecionados na RSL.....	37
Tabela 6 – Demais informações relevantes sobre modelos utilizados nos trabalhos acadêmicos selecionados na RSL.....	40
Tabela 7 - Pesos atribuídos aos sinistros	57
Tabela 8 - Medidas mitigadoras de sinistros de trânsito.....	59
Tabela 9 - Número de registros com base nos filtros aplicados	79
Tabela 10 - Ranking com os 15 grupo com maior número de UPS.....	81
Tabela 11 - Variáveis presentes em cada sinistro	86
Tabela 12 – Variáveis Utilizadas em Cada Modelo	92
Tabela 13 - Medida de qualidade	93
Tabela 14 - Variáveis Independentes e Parâmetros.....	95
Tabela 15 - Variáveis Independentes e Parâmetros do Modelo 1	113
Tabela 16 - Variáveis Independentes e Parâmetros do Modelo 2	115

LISTA DE ABREVIATURAS E SIGLAS

ANTT	Agência Nacional de Transportes Terrestres
ABNT	Associação Brasileira de Normas Técnicas
CBT	Código Brasileiro de Trânsito
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
BIC	Critério de Informação Bayesiano
AIC	Critério de Informação de <i>Akaike</i>
DPVAT	Danos Pessoais Causados por Veículos Automotores de Vias Terrestres
DASV	Década de Ações para a Segurança Viária
DATASUS	Departamento de Informação e Informática do Sistema Único de Saúde
FHA	<i>Federal Highway Administration</i>
ODS	Objetivos do Desenvolvimento Sustentável
OR	<i>Odds ratio</i>
ONU	Organização das Nações Unidas
PNCV	Plano Nacional de Controle Eletrônico de Velocidade
PNATRANS	Plano Nacional de Redução de Mortes e Lesões no Trânsito
PRF	Polícia Rodoviária Federal
PIB	Produto Interno Bruto
ROC	<i>Receiver Operator Characteristic</i>
RENAEST	Registro Nacional de Sinistros e Estatísticas de Trânsito
RENAVAN	Registro Nacional de Veículos Automotores
RSL	Revisão Sistemática da Literatura
SENATRAN	Secretaria Nacional De Trânsito
SIM	Sistema de Informação de Mortalidade
SNV	Sistema Nacional de Viação
UPS	Unidade Padrão De Severidade
UFSC	Universidade Federal de Santa Catarina

SUMÁRIO

1	INTRODUÇÃO	14
1.1	ADERÊNCIA AO PPGTG	15
1.2	OBJETIVOS.....	17
1.2.1	Objetivo Geral	17
1.2.2	Objetivos Específicos	17
2	FUNDAMENTAÇÃO TEÓRICA	19
2.1	CUSTO ECONÔMICO DOS SINISTROS RODOVIÁRIOS.....	19
2.2	DÉCADA DE AÇÕES PARA A SEGURANÇA VIÁRIA	21
2.2.1	Sistema Seguro	24
2.2.2	Panorama Mundial	25
2.2.3	Panorama Brasileiro	26
2.2.4	Fatores Contribuintes Para Sinistros Rodoviários	28
2.2.4.1	<i>Direção Agressiva</i>	30
2.2.4.2	<i>Controle de Velocidade</i>	30
2.3	RODOVIAS CONCEDIDAS NO BRASIL	32
2.3.1	Infraestrutura Brasileira	32
2.4	REVISÃO SISTEMÁTICA DA LITERATURA.....	34
2.4.1	Planejamento da Revisão	34
2.4.2	Execução da Revisão	36
2.4.3	Relatório	44
2.5	MODELO DE REGRESSÃO LOGÍSTICA BINÁRIA.....	45
2.5.1	Ajuste do modelo	46
2.5.2	Métodos de Avaliação do Modelo Logístico	47
2.5.2.1	<i>Teste de Razão de Verossimilhança</i>	47
2.5.2.2	<i>Critério de Informação de Akaike</i>	48
2.5.2.3	<i>Critério de Informação Bayesiano</i>	49
2.5.2.4	<i>Estatística Deviance</i>	50
2.5.2.5	<i>Pseudo R²</i>	50
2.5.2.5.1	<i>R² de Cox e Snell</i>	50
2.5.2.5.2	<i>R² de Nagelkerke</i>	51
2.5.2.6	<i>Teste de Wald</i>	52
2.5.2.7	<i>Resíduos</i>	52

2.5.2.7.1	Resíduo de Pearson	53
2.5.2.7.2	Resíduos padronizados:	53
2.5.2.8	<i>Curva Receiver Operator Characteristic</i>	54
2.5.3	Razão de Chances (<i>Odds Ratio</i>)	55
2.6	UNIDADE PADRÃO DE SEVERIDADE	56
2.7	TÉCNICAS DE AGRUPAMENTO APLICADAS A SINISTROS RODOVIÁRIOS	58
2.7.1	Ponto Crítico	58
2.7.2	Método de particionamento (<i>k-means</i>)	60
2.7.3	Algoritmo Hierárquico	62
2.7.4	Método baseado em grade (CLIFE)	63
2.7.5	Método baseado em densidade (DBSCAN)	64
2.8	LINGUAGEM DE PROGRAMAÇÃO	67
3	MATERIAIS E MÉTODOS	69
1	70
1	70
3.1	DEFINIÇÃO DA ÁREA DE ESTUDO	70
3.1.1	Litoral Sul	71
3.1.2	Régis Bittencourt	71
3.1.3	Fernão Dias	72
3.2	TRATAMENTO PRELIMINAR DOS DADOS	72
3.2.1	Dados Incluídos	72
3.2.1.1	<i>Volume Mensal Médio</i>	72
3.2.1.2	<i>Informações geográficas</i>	74
3.2.1.3	<i>Iluminação</i>	74
3.2.1.4	<i>Presença de vias marginais</i>	75
3.2.1.5	<i>Análise temporal</i>	75
3.2.1.6	<i>Unidade Padrão de Severidade</i>	75
3.2.2	Filtros Utilizados	76
3.2.2.1	<i>Sinistros com vítimas</i>	76
3.2.2.2	<i>Sinistros com vítimas moderadas, graves ou fatais</i>	78
3.3	ANÁLISE DESCRITIVA	79
3.3.1	Modelo DBSCAN	80
3.3.2	Dados do modelo	80

3.3.2.1	<i>Litoral Sul</i>	83
3.3.2.2	<i>Régis Bittencourt</i>	84
3.3.2.3	<i>Fernão Dias</i>	85
3.4	MODELO LOGIT.....	86
3.4.1	Variáveis Analisadas	86
3.4.2	Método de Seleção	91
3.4.3	Modelos Propostos	92
3.4.4	Testes de Validação	93
4	RESULTADOS	95
5	CONSIDERAÇÕES FINAIS	101
	REFERÊNCIAS	104
	APÊNDICE	113

1 INTRODUÇÃO

A *World Health Organization* WHO (2018), relata que os sinistros de trânsito são uma das principais causas de mortalidade em todo o mundo, afetando especialmente pessoas entre cinco e 29 anos. Anualmente, ocorrem cerca de 1,35 milhão de mortes e 50 milhões de feridos devido a esses incidentes, com países de média e baixa renda apresentando taxas de mortalidade duas vezes maiores do que países de alta renda.

No contexto brasileiro, dados da Polícia Rodoviária Federal (PRF) e do Registro Nacional de Sinistros e Estatísticas de Trânsito (RENAEST) revelam que, em 2021, ocorreram 28 mil mortes decorrentes de sinistros de trânsito. Com base nos estudos desenvolvidos por um grupo de pesquisadores denominado *Diseases and Injuries Collaborators* (2020), abrangendo o período de 2009 a 2019, verificou-se que os sinistros de trânsito no Brasil permaneceram como a oitava principal causa de morte no país ao longo do período. Isso destaca a persistência desse problema e a necessidade de análise e solução.

De acordo com Carvalho (2020) os sinistros de trânsito no Brasil representaram um custo aproximado de 3% do Produto Interno Bruto (PIB) em 2015. Esse valor abrange uma variedade de despesas, incluindo cuidados com saúde, perda de produtividade devido a lesões ou mortes, danos à infraestrutura e veículos envolvidos, além de custos operacionais como remoção de veículos acidentados.

Diante da importância e urgência na redução do número de sinistros, a Assembleia Geral das Nações Unidas (ONU) declarou a segunda Década de Ação para a Segurança no Trânsito 2021-2030, estabelecendo a meta de reduzir em pelo menos 50% as mortes e lesões no trânsito até 2030 por meio da Resolução 74/299 de 31 de agosto de 2020 da Assembleia Geral da ONU.

No Brasil, desde o lançamento dessa meta pela ONU, foram estabelecidos planos para redução de sinistros, incluindo o Plano Nacional de Redução de Mortes e Lesões no Trânsito (PNATRANS), instituído pela Lei nº 13.614, de 11 de janeiro de 2018, e pela Resolução do CONTRAN nº 870, de 13 de setembro de 2021. Este plano define seis pilares para a redução de sinistros: gestão da segurança no trânsito, vias seguras, segurança veicular, educação para o trânsito, atendimento às vítimas e normatização e fiscalização.

Dentre as principais abordagens adotadas pela Resolução, podemos destacar as abordagens do Sistema Seguro integrado à metodologia do Visão Zero, que

definem que sinistros são inevitáveis, porém nenhuma morte no trânsito é aceitável. Para a adoção de intervenções com vistas à redução de sinistros é necessário compreender as razões que levam à sua ocorrência, podendo-se relacionar o número de sinistros e características locais.

Com o intuito de analisar a ocorrência de sinistros no Brasil, o IPEA realiza a compilação dos dados e define frentes de análise dos sinistros. Em 2015, o Instituto publicou uma pesquisa com foco nos estados brasileiros, destacando aqueles com maior número de sinistros. Dessa forma, fica mais visível para o governo entender os estados que possuem maior necessidade de melhorias (CARVALHO, 2016).

Segundo Carvalho (2016) os estados com as maiores taxas de sinistros por quilômetro rodado são Santa Catarina, Rio de Janeiro, São Paulo, Minas Gerais e Paraná. Esses estados também se destacam pelo alto número de mortes relacionadas a sinistros. Nesse contexto, este trabalho irá considerar quatro, dos cinco estados com as maiores taxas de sinistros.

Por fim, entendendo que sinistros envolvendo pedestres e ciclistas diferem significativamente daqueles que envolvem veículos automotores, devido à maior vulnerabilidade desses usuários e ao seu perfil distinto de utilização da rodovia, o presente trabalho focará exclusivamente na análise de veículos automotores, excluindo pedestres e ciclistas de todas as análises.

1.1 ADERÊNCIA AO PPGTG

Esta dissertação busca estimar os determinantes da gravidade dos sinistros envolvendo veículos automotores em rodovias federais brasileiras. Como tal, está inserida na área de concentração Sistemas de Transportes e na linha de pesquisa Planejamento de Sistemas de Transportes. No PPGTG, esta linha visa adequar as necessidades de transporte de uma região ao seu desenvolvimento de acordo com suas características estruturais. Isto significa implantar novos sistemas ou melhorar os existentes.

No histórico do PPGTG, foram encontradas cinco dissertações que guardam afinidade com o tema deste trabalho. Na Tabela 1, a seguir, estão destacados os trabalhos considerados de contexto mais próximo ao desta dissertação.

Tabela 1 - Dissertações similares PPGTG

Autor	ANBT	Observações
Martins, Fernanda S. (2023)	Martins, Fernanda S. Procedimento metodológico para a localização de pontos de parada e descanso em rodovias Dissertação (mestrado) - Universidade Federal de Santa Catarina, Florianópolis, 2023.	O estudo focou na otimização da localização de áreas de descanso para caminhoneiros em rodovias nacionais, utilizando técnicas de Pesquisa Operacional e Sistemas de Informação Geográfica, com o objetivo de melhorar a segurança viária, o conforto dos motoristas e a eficiência no transporte de cargas.
Sobrosa, Camila Q. (2023)	Sobrosa, Camila Q. Contribuição metodológica para alocação de áreas de escape para veículos pesados em rodovias brasileiras. Dissertação (mestrado) - Universidade Federal de Santa Catarina, Florianópolis, 2023.	O estudo foi desenvolvido para criar uma metodologia que identifica pontos críticos e otimiza a alocação de áreas de escape para veículos pesados em rodovias, visando reduzir o número de sinistros e melhorar a segurança viária, com foco em um estudo de caso na BR-116.
Kaesemodel, Luciano (2021)	Kaesemodel, Luciano. Método georreferenciado de identificação de trechos concentradores de acidentes para suporte à tomada de decisão em relação à fiscalização dos limites de velocidade em rodovias. Dissertação (mestrado) - Universidade Federal de Santa Catarina, Florianópolis, 2021.	Este estudo propôs um método georreferenciado para identificar trechos concentradores de sinistros, visando auxiliar a seleção de locais prioritários para instalação de fiscalização eletrônica de velocidade, com base em técnicas de clustering e densidade de sinistros, e foi validado com dados da malha rodoviária federal de Santa Catarina
Pianezer, Tiago A. (2019)	Pianezer, Tiago A. Análise comparativa de metodologias para identificação de segmentos críticos em rodovias rurais. Dissertação (mestrado) - Universidade Federal de Santa Catarina, Florianópolis, 2019.	Este trabalho aborda a segurança viária e tem como objetivo estudar diferentes metodologias para identificar locais críticos de sinistros em rodovias rurais, aplicando três delas em um estudo de caso na BR-116 em Santa Catarina, e analisando suas correlações para entender a eficácia na identificação de segmentos críticos
Thomazi, Kenia R. (2018)	Thomazi, Kenia R. Elaboração de modelos de previsão de acidentes de trânsito em rodovias rurais de pista simples do estado de Santa Catarina com base no Highway Safety Manual (HSM). Dissertação (mestrado) - Universidade Federal de Santa Catarina, Florianópolis, 2018.	O estudo desenvolveu modelos de previsão de sinistros de trânsito em rodovias rurais de pista simples em Santa Catarina, utilizando a metodologia do Highway Safety Manual (HSM) e validando sua aplicabilidade local, com o objetivo de aprimorar a segurança viária e reduzir sinistros.

Pode-se perceber que há trabalhos que possuem contexto de aplicação semelhante ao da presente dissertação, bem como objetivos relacionados à melhoria da segurança viária em rodovias brasileiras. Destacam-se os trabalhos de Martins (2023) e Sobrosa (2023), que focaram na otimização de áreas de descanso e áreas de escape para veículos pesados, contribuindo para a redução de sinistros e aumento da segurança.

Por outro lado, os trabalhos de Thomazi (2018) e Kaesemodel (2021) abordaram a previsão de sinistros de trânsito e a identificação de trechos críticos em rodovias utilizando metodologias de análise de dados e técnicas georreferenciadas, o

que está alinhado com o uso da regressão logística binária e do método DBSCAN neste estudo para identificar áreas críticas de sinistros e variáveis associadas à gravidade dos sinistros. Essas abordagens metodológicas compartilham o objetivo comum de aprimorar a segurança viária por meio da identificação de padrões e pontos críticos em rodovias.

Por fim, o trabalho de Pianezzer (2019) também apresenta relação com a presente dissertação, ao comparar metodologias de identificação de segmentos críticos em rodovias. Embora com diferentes técnicas, o foco em rodovias federais e a busca por correlacionar fatores de risco nos sinistros reforça a importância de metodologias robustas e comparáveis.

Observando-se o histórico de trabalhos do PPGTG, nota-se que a presente dissertação traz como contribuição específica a análise das principais características associadas à gravidade dos sinistros. Nesse sentido, esta dissertação pode vir a ser subsídio para novas pesquisas em áreas relacionadas à segurança viária, servindo como base para o desenvolvimento de campanhas educacionais, bem como melhorias viárias direcionadas à redução da gravidade dos sinistros de trânsito. Além disso, a aplicação de regressão logística binária e a identificação de áreas críticas por meio do método DBSCAN oferece um novo panorama metodológico que pode ser expandido e aprimorado em estudos futuros.

1.2 OBJETIVOS

1.2.1 Objetivo Geral

Estimar os determinantes da gravidade dos sinistros envolvendo veículos automotores em rodovias federais nos estados de Minas Gerais, São Paulo, Paraná e Santa Catarina.

1.2.2 Objetivos Específicos

Afim de atingir o objetivo geral deste trabalho, foram definidos os seguintes objetivos específicos:

- I. Considerar as principais teorias, estudos e descobertas científicas que elucidam os fatores causadores de sinistros de trânsito e as variáveis mais frequentemente associadas a esses eventos;
- II. Propor modelo de clusterização a fim de determinar as principais áreas críticas para a ocorrência de sinistros graves ou fatais;
- III. Propor modelo de regressão logística binário e analisar os fatores que influenciam a severidade dos sinistros de trânsito, identificando aqueles que são preponderantes para o aumento da gravidade desses eventos,

2 FUNDAMENTAÇÃO TEÓRICA

Conforme a Associação Brasileira de Normas Técnicas (ABNT, 2020), sinistros ou sinistros de trânsito são:

Todo evento que resulte em dano ao veículo ou a sua carga e/ou em lesões a pessoas e/ou animais, e que possa trazer dano material ou prejuízos ao trânsito, à via ou ao meio ambiente, em que pelo menos uma das partes está em movimento nas vias terrestres ou em áreas abertas ao público.

As classificações dos tipos de sinistro de trânsito são demonstradas pela Tabela 1:

Tabela 2 - Tipos de classificação de sinistros de trânsito

Tipos de sinistro de trânsito	Descrição
Sem vítima	Aquele que não resulte em vítima, mas que traga dano material ou prejuízos ao trânsito, à via ou ao meio ambiente
Com vítima não fatal	Aquele que não resulte em vítima fatal
Com vítima fatal	Aquele que resulte em vítima morrendo imediatamente ou em até 30 dias, como resultados do sinistro ou suas complicações.

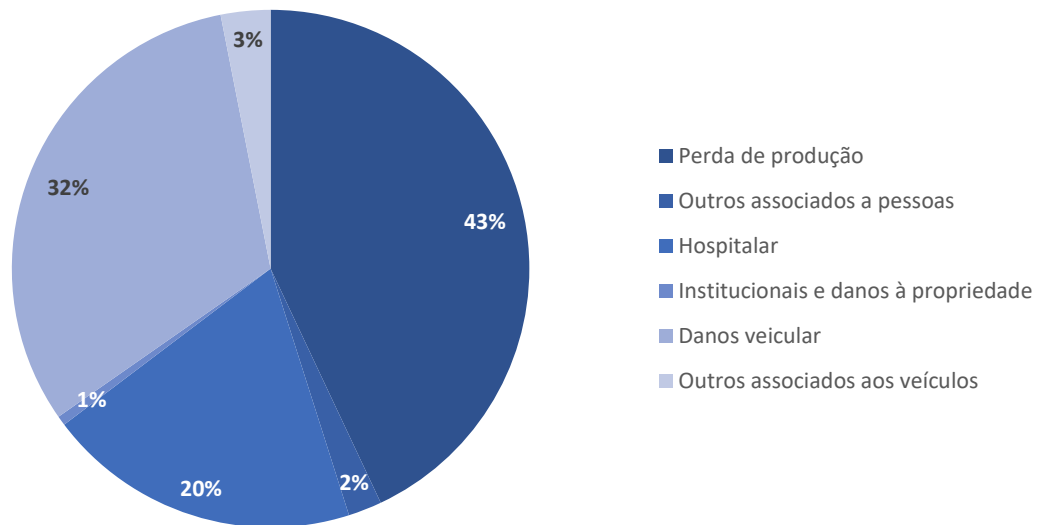
Fonte: ABNT (2020)

2.1 CUSTO ECONÔMICO DOS SINISTROS RODOVIÁRIOS

Segundo Carvalho (2020), um aspecto negativo de se analisar quanto aos sinistros viários, está relacionado ao alto custo envolvido, gerando grande impacto na economia do país. Dentro dos custos associados aos sinistros, podemos citar o atendimento das vítimas, danos ao veículo, danos patrimoniais, custos institucionais, entre outros.

Nas rodovias federais em 2014, os custos dos sinistros para a sociedade foram em torno de R\$12,8 bilhões, sendo que 62% dos custos são relacionados às vítimas e 37,4% relacionados ao veículo, como perda de carga, danos ao veículo e remoção Carvalho (2020). A Figura 1 ilustra as divisões dos custos associados aos sinistros.

Figura 1 - Componentes de custo dos sinistros nas rodovias federais (2014)



Fonte: Carvalho (2020)

Dessa forma, entende-se que os sinistros com vítimas e fatais apresentam os maiores custos, pois além dos investimentos financeiros em recuperação dos danos materiais, há aqueles vinculados ao tratamento das vítimas. Conforme destacado por Carvalho (2020) os sinistros fatais compõem 5% do total de ocorrências, em contra partida representam 35% dos gastos totais relacionados à sinistros.

O impacto causado devido a perdas de produção justifica o número expressivo de indenizações registradas pelo seguro Danos Pessoais Causados por Veículos Automotores de Vias Terrestres (DPVAT). Em 2020, foram registradas 310.710 indenizações relacionadas a sinistros de trânsito no Brasil, sendo 67% decorrentes de invalidez permanente, 22% de despesas médicas e 11% por mortes (SEGURADORA LÍDER, 2020)

Wijnen e Stipdonk (2016) conduziram um estudo abrangendo 17 países, dos quais 10 possuíam PIB per capita elevado, e 7 possuíam PIB per capita médio e baixo. Eles descobriram que no primeiro grupo, o custo médio de sinistros atinge 2,7% do PIB, enquanto no segundo grupo atinge 2,2%.

Ao considerar o Brasil como um país com PIB per capita médio, conforme os demais países analisados, 2,2% do PIB representaria um gasto de 54 bilhões de reais no ano de 2014. Ao compreender os altos custos associados aos sinistros, podemos avaliar não apenas as perdas sociais, mas também as perdas econômicas

decorrentes. Dessa forma, é fundamental entender os fatores que geram tais ocorrências.

2.2 DÉCADA DE AÇÕES PARA A SEGURANÇA VIÁRIA

A ONU lançou, na Assembleia Geral, a Década de Ações para a Segurança Viária (DASV), por meio da Resolução 64/255, de 02 de março de 2010. A principal proposta da Resolução foi a redução em 50% das fatalidades no trânsito entre os anos de 2011 e 2020, a partir de planos de ações que cada país participante deveria implementar no seu território.

Porém, antes da definição da DASV, em 2009 foi lançado o Relatório Global do Estado da Segurança Viária: Tempo de Ação, o qual definiu a importância de tratar pontos como: (i) a segurança dos usuários vulneráveis (pedestres, ciclistas e motociclistas), os quais representaram metade das fatalidades no trânsito; (ii) melhoria de legislações e estudos para a redução de sinistros, como a implantação do sistema “Visão Zero” da Suíça, priorização e investimento no transporte não motorizado no Chile, dentre outros; (iii) melhoria na consolidação de dados relacionados à segurança viária por um órgão confiável, visto que apenas 58% dos países reportaram possuir um plano nacional de segurança viária (WHO, 2009).

Após a publicação do relatório houve a Primeira Conferência Ministerial Global de Segurança Viária, em Moscou, na Rússia. A Conferência possuiu grande importância por ser o primeiro encontro global para se discutir segurança viária. Sendo a principal declaração da conferência dada pela Assembleia Geral da ONU ao definir o período entre 2011-2020 como a DASV (PAVARINO FILHO, 2016).

Dessa forma, a Organização Mundial da Saúde lançou em 2011 um guia intitulado Plano Global, como um documento de referência para dar apoio à implementação da Década de Ação 2011-2020. O Plano define a importância de uma agência reguladora capaz de propor planos e/ou estratégias mensuráveis com foco em segurança viária (WHO, 2011).

As principais intervenções propostas no Plano Global são:

- Projetos de uso e ocupação do solo com incorporação de recursos de segurança viária;
- Projetos de malha viária mais seguras;

- Exigir auditorias independentes de segurança viária para novos projetos de construção;
- Implementar melhorias à segurança dos veículos;
- Promover o transporte público;
- Gestão eficaz da velocidade pela polícia e por uso de vias calmas (traffic-calming);
- Definição e cumprimento de leis de segurança do usuário como o uso de cintos de segurança, capacetes e sistemas de retenção para crianças (cadeirinha);
- Definição e aplicação de limites de concentração de álcool no sangue para motoristas;
- Melhorar o atendimento pós-sinistro para vítimas de sinistros de trânsito.

O Plano segue ainda a abordagem “Sistema Seguro”, na qual entende-se que o motorista está sujeito a cometer erros/falhas, e o sistema é responsável por minimizar os danos causados considerando a vulnerabilidade do corpo humano. Parte da aceitação do erro humano envolve a compreensão de que os sinistros de trânsito não podem ser completamente evitados; no entanto, o sistema deve ser projetado para evitar vítimas graves e fatais (WHO, 2011).

Por fim, o Plano define cinco pilares de atuação:

- Pilar 1 - Gestão de Segurança no Trânsito: Aderir e/ou implementar integralmente os instrumentos jurídicos da ONU, incentivar a criação de parcerias multissetoriais e a designação de agências líderes com capacidade para desenvolver e liderar a entrega de estratégias, planos e metas nacionais de segurança no trânsito;
- Pilar 2 - Infraestrutura Viária e Mobilidade: Trata sobre a melhoria da qualidade e segurança da malha viária, a partir de planejamentos, projetos, construções e exploração das vias em termos de segurança;
- Pilar 3 - Segurança Veicular: uso de tecnologia padrão universal para melhorar a segurança dos veículos, informações ao consumidor e incentivos à adoção de novas tecnologias;

- Pilar 4 - Usuários Mais Seguros: Desenvolvimento de programas que melhorem o comportamento dos usuários no trânsito. Através da implantação e cobrança de leis, combinada com a conscientização/educação para o uso de dispositivos de segurança e cumprimento de leis quanto à velocidade máxima permitida e não uso de álcool;
- Pilar 5 – Resposta aos Sinistros: Aumentar a capacidade de resposta às emergências, além de melhorar a capacidade dos sistemas de saúde e reabilitação de longo prazo para as vítimas de sinistro.

Em 2015, a comunidade internacional reunida na sede das Nações Unidas em Nova York (EUA), aprovou a Agenda 2030, delineando 17 Objetivos do Desenvolvimento Sustentável (ODS) e um total de 169 metas, com objetivo de guiar a economia mundial, bem-estar e diplomacia. O Objetivo Global 3.6, que visa assegurar uma vida saudável e promover o bem-estar para todos, em todas as idades, incluiu além da redução em 50% das fatalidades prevista pela DASV, a redução em 50% das vítimas decorrentes de sinistros de trânsito até 2020 (WHO, 2015).

De maneira geral, a maioria dos países não alcançaram a Meta 2011-2020 proposta, sendo os países da União Europeia aqueles com as maiores taxas de redução de fatalidades no período. O Brasil apresentou redução de 24% considerando os dados do Departamento de Informação e Informática do Sistema Único de Saúde (DATASUS) do período (MINISTÉRIO DA SAÚDE, 2023).

Entendendo que segurança viária ainda é um problema a ser resolvido após o não cumprimento da Meta proposta entre 2011 e 2020, a Assembleia Geral da ONU aprovou a Resolução 74/299, intitulada “Melhorando a segurança rodoviária global”, definindo a Segunda Década Ação pela Segurança Viária, com proposta de redução em 50% o número de mortes e lesões até 2030 (WHO, 2021).

Assim como na década anterior, foi lançado o Plano Global, com as diretrizes para os países alcançarem a meta proposta. O Plano continua com os cinco pilares, com pequenas alterações nas ações propostas dentro de cada cenário. O Plano também prevê formas para a implantação das melhorias a partir de financiamento, implantação de leis, gestão da velocidade, desenvolvimento de profissionais de segurança viária, garantir a expectativa de gênero no planejamento de transportes,

adaptação de tecnologias do Sistema Seguro, e foco em países de baixa e média renda (WHO, 2021).

2.2.1 Sistema Seguro

O Sistema Seguro foi lançado em 1992 Koornstra et al. (1992) na Holanda, local onde foi bem aceito, e a partir de 1995 um programa de implementação foi iniciado. Um estudo de avaliação cobrindo o período de 1998 a 2007 revelou uma redução de 30% no número de fatalidades. Os benefícios dos investimentos foram quatro vezes maiores do que os custos Wegman et al. (2023).

Mais duas atualizações foram publicadas posteriormente (WEGMAN; AARTS, 2006; SWOV, 2018). Os princípios nessas três versões diferentes são muito semelhantes. O foco das duas primeiras versões está apenas no design do Sistema Seguro, a terceira edição acrescenta dois princípios organizacionais: alocação eficaz de responsabilidade e aprendizado e inovação no sistema de tráfego

A ideia principal de uma abordagem de Sistema Seguro é criar um ambiente para os usuários das estradas que elimine ou reduza drasticamente o risco de sinistros e, caso a eliminação do risco de sinistros não seja (ainda) realista e os sinistros ainda ocorram, reduzir os riscos de ferimentos graves (WEGMAN et al., 2023).

Um Sistema Seguro envolve uma abordagem pró-ativa baseada na compreensão dos riscos inerentes ao transporte rodoviário que devem ser evitados antes que os sinistros comecem a ocorrer. Finalmente, existe uma responsabilidade compartilhada entre aqueles que usam as estradas e aqueles que projetam, constroem e gerenciam os componentes do sistema de transporte (estradas, veículos, tecnologias, cuidados pós-acidente).

Na terceira edição de Segurança Sustentável, cinco princípios são essenciais: três princípios de design (1, 2 e 3) e dois princípios de organização (4 e 5).

1. Funcionalidade das Estradas: Estradas devem ajustar a velocidade e o design para garantir segurança, especialmente entre veículos e usuários vulneráveis.
2. (Bio)mecânica: O design da estrada deve minimizar diferenças de velocidade, direção e tamanho dos veículos para reduzir o risco de sinistros, focando especialmente na segurança de ciclistas.

3. Psicologia: O tráfego deve ser projetado para ser claro e intuitivo, com sinais e ambiente que facilitem a compreensão e adaptação dos usuários, principalmente os mais velhos.
4. Atribuindo Responsabilidades: As regras de trânsito devem ser claras e indulgentes com erros humanos, prevenindo sinistros graves.
5. Aprendizado e Inovação: O sistema de trânsito deve se adaptar continuamente, aplicando inovações baseadas em dados para reduzir sinistros e melhorar a segurança.

O Sistema Seguro busca tornar o uso das estradas o mais seguro possível, levando em conta as necessidades e capacidades dos usuários atuais e futuros. Reconhece que certos modos de transporte, como veículos de duas rodas, são mais perigosos, e que alguns grupos, como crianças e idosos, são mais vulneráveis a lesões. O objetivo é garantir segurança máxima para todos Wegman et al. (2023).

Um componente crítico dessa visão é a redução das velocidades de impacto, já que menores velocidades reduzem a gravidade dos sinistros, especialmente em áreas com usuários vulneráveis. Esse princípio é central para evitar ferimentos graves e salvar vidas Wegman et al. (2023). Para atingir essa segurança máxima, propõe-se um Sistema Seguro com três etapas:

- a) Eliminação: Evitar que situações perigosas ocorram.
- b) Minimização: Reduzir o número de situações perigosas e desencorajar o uso de modos de transporte mais arriscados.
- c) Mitigação: Minimizar as consequências dos riscos quando eles ocorrem, adotando medidas adequadas.

A terceira edição também destaca a importância da "dimensão humana", tanto para os usuários das estradas quanto para os profissionais que projetam e gerenciam o sistema de trânsito. As falhas em qualquer elemento do sistema devem ser compensadas por outros, assegurando que ele seja o mais seguro possível, especialmente através de medidas como a redução de velocidades.

2.2.2 Panorama Mundial

O índice de fatalidades a cada 100 mil habitantes é o mais utilizado para análises comparativas entre os diferentes países do mundo. Em 2016, o índice médio mundial foi de 18,2, na Europa, os países de alta renda apresentaram taxas de 5,1 e

países média renda 14,4, número três vezes maior. O mesmo ocorre nos demais países do mundo, quanto maior a renda, menor a taxa (WHO, 2018).

Planos bem desenhados, políticas mais estruturadas, melhoria na infraestrutura das vias e veículos mais seguros são os principais pontos destacados pelo *Global Status Report on Road Safety*, para a redução de fatalidades entre jovens em países com maior renda (WHO, 2018).

2.2.3 Panorama Brasileiro

De acordo com os dados abertos da PRF e RENAEST, os Estados de São Paulo, Paraná, Minas Gerais, Bahia, e Rio Grande do Sul apresentam maiores números de fatalidades no trânsito, respectivamente. Sendo as regiões Sul e Sudeste com maior número de fatalidades no trânsito, conforme pode ser demonstrado pela Tabela 3.

Tabela 3 - Distribuição das fatalidades no trânsito por Estado

Estado	Part. (%)		Estado	Part. (%)		Estado	Part. (%)	
	Fatalidades	Frota		Fatalidades	Frota		Fatalidades	Frota
São Paulo	18.1%	28.3%	Rio de Janeiro	3.1%	6.6%	Mato Grosso do Sul	0.8%	1.6%
Paraná	9.7%	7.5%	Paraíba	2.6%	1.3%	Tocantins	0.7%	0.7%
Minas Gerais	9.5%	11.1%	Mato Grosso	2.6%	2.2%	Roraima	0.6%	0.2%
Bahia	7.5%	4.2%	Alagoas	2.5%	0.9%	Amazonas	0.5%	0.9%
Rio Grande do Sul	6.9%	7.0%	Ceará	2.3%	3.2%	Acre	0.3%	0.3%
Pernambuco	5.6%	3.0%	Piauí	2.3%	1.2%	Sergipe	0.2%	0.8%
Santa Catarina	4.8%	5.2%	Rio Grande do Norte	2.1%	1.3%	Amapá	0.1%	0.2%
Pará	4.7%	2.1%	Rondônia	1.9%	1.0%			
Goiás	4.4%	3.9%	Espírito Santo	1.1%	1.9%	Total Geral	100%	100%
Maranhão	4.0%	1.7%	Distrito Federal	1.0%	1.7%			

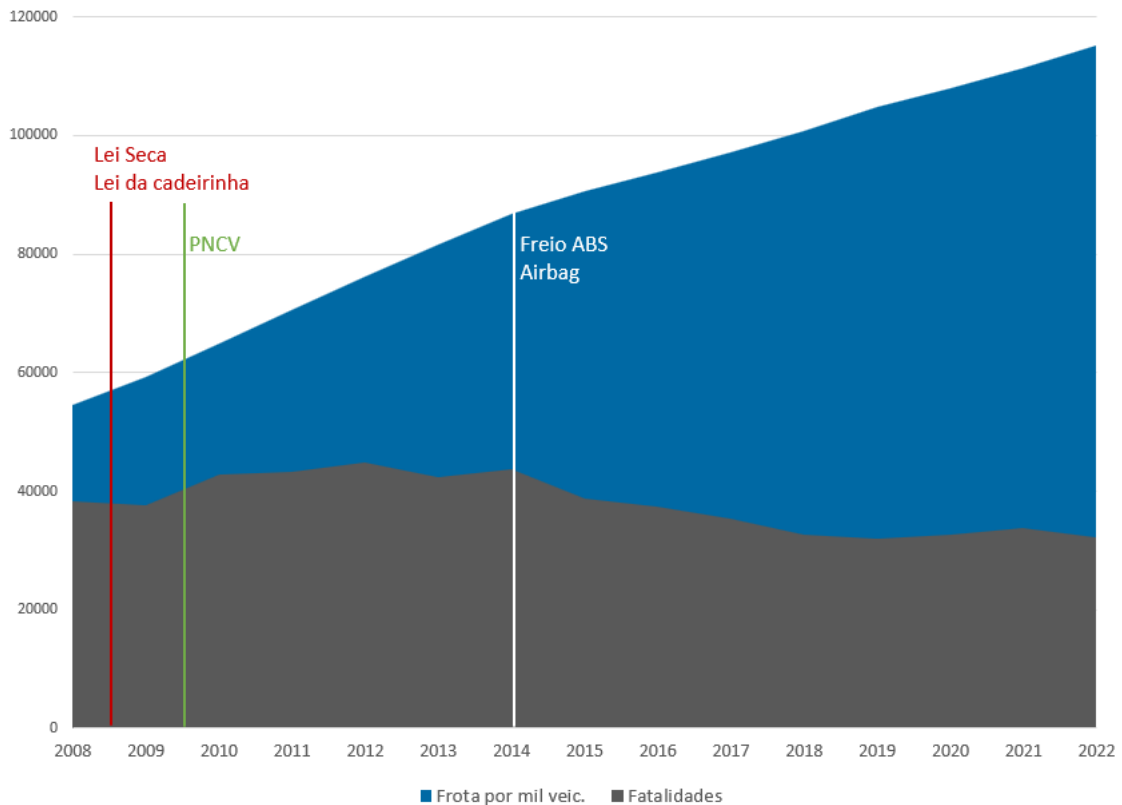
Fonte: Autora (2024)

Também é possível analisar a participação de cada Estado na frota de veículos, com base nos dados do Registro Nacional de Veículos Automotores (RENAVAN). De forma geral, vale ressaltar que os Estados com maior número de fatalidades no trânsito possuem as maiores frotas de veículos, como mostrado na Tabela 3. Dessa forma, entende-se que existe relação entre o aumento do número de veículos e o aumento no número de sinistros, devido à maior exposição do motorista (MÂNICA, 2007).

Porém, de acordo com o Sistema de Informação de Mortalidade (SIM) do DATASUS, nos últimos 10 anos houve redução média de 3% no número de

fatalidades. Enquanto a frota de veículos apresentou um aumento médio de 4%, de acordo com dados da Secretaria Nacional De Trânsito (SENATRAN). Tais informações são apresentadas na Figura 2.

Figura 2 - Frota de veículos e número de fatalidades de 2008 a 2022



Fonte: Elaboração própria com dados do Sistema de Informação de Mortalidade e Secretaria Nacional de Trânsito (2023).

Tal crescimento da frota de veículos está atrelado à expansão dos centros urbanos, tornando necessário o uso de meios de transporte motorizados individuais para a mobilidade. No comparativo dos últimos 10 anos, além do automóvel, que apresentou aumento de 50%, a motocicleta também se destaca pelo crescimento em torno de 60%.

Alguns pontos que podem ser definidos como marcos para a redução das fatalidades no decorrer dos anos são: (i) Lei Seca e a Lei da cadeirinha, ambas em vigor em 2008; (ii) o Plano Nacional de Controle Eletrônico de Velocidade, criado em 2009; (iii) Obrigatoriedade da instalação de freio ABS e *airbag* em todos os carros produzidos a partir de 2014.

Porém, em alguns anos é notável o aumento de fatalidades comparado com o ano anterior, como 2010 que apresentou aumento de 14%, e mais recentemente 2021 com aumento de 3% comparado com 2020. Dentre outros fatores associados a esse aumento podemos citar a ingestão de bebida alcoólica, excesso de velocidade e não uso de dispositivo de segurança. Pesquisas feitas por Damacena et al. (2016) mostraram que de 1840 entrevistados que se envolveram em sinistros de trânsito, 12% dirigiram sob efeito de bebida alcoólica nos 12 meses anteriores à pesquisa. E vale destacar, que dentro dessa porcentagem, 54% dos envolvidos eram motociclistas.

Quanto ao uso de dispositivos de segurança, a pesquisa de Malta et al. (2016) apresenta 80% dos adultos, com idade acima de 18 anos, utilizam o cinto de segurança no banco da frente, porém a porcentagem cai para 50% dos entrevistados quando os mesmos estão no banco de trás. O capacete é utilizado por 80% dos condutores e passageiros de motocicletas. Por fim, o estudo mostrou que as menores taxas de uso de equipamentos de segurança estão, principalmente, nas regiões rurais.

Nesse cenário, a segurança viária torna-se a medida de desempenho do sistema viário, definindo o quanto o sistema é seguro em relação a ocorrência de sinistros e as consequências relacionadas (CARDOSO, 2006).

2.2.4 Fatores Contribuintes Para Sinistros Rodoviários

Os eventos antes e durante o sinistro mostram os fatores que contribuíram para sua gravidade e como as soluções de engenharia ou mudanças tecnológicas podem reduzir a ocorrência ou severidade do sinistro (AASHTO, 2009). Porém, a definição do conjunto de causas do sinistro é um dos maiores desafios na melhoria da segurança viária, visto que, de forma geral, envolve a interação de diferentes fatores na sua ocorrência.

De acordo com o Ministério dos Transportes (2002), os principais fatores que podem contribuir na ocorrência de sinistros são divididos em três categorias: (i) fator humano, associado ao comportamento e tempo de reação do motorista; (ii) fator via e meio ambiente, relacionado a questões de projeto, manutenção da via e da vegetação próxima ao trecho, e (iii) fator veículo, associado a eventuais problemas que possam ocorrer decorrente de falhas no veículo.

Dessa forma, a classificação é o resultado da análise entre a relação direta do risco de ocorrência do sinistro em um determinado local e as características geométricas, ambientais, condições de trafegabilidade e do veículo que circula na via, assim como a reação do condutor a tais eventos (MINISTÉRIO DOS TRANSPORTES, 2002).

Segundo Gold (1998) existe mais um fator que pode ter ligação direta aos sinistros, o fator institucional/social. Tal fator estaria relacionado a questões de regulamentação e fiscalização, como a obrigatoriedade do uso de cadeirinha para crianças, uso de cinto de segurança, não ingestão de bebida alcoólica, determinação de velocidade máxima permitida, dentre outros fatores. Porém, mesmo esse fator sendo considerado influente na redução de sinistros não é, habitualmente, considerado nos estudos.

Elvik (2006) apresenta quatro leis com o objetivo de justificar a ocorrência de um sinistro de trânsito: (i) A primeira é a lei universal do aprendizado, que define que quanto mais experiência o motorista adquire, melhor se torna a sua capacidade em detectar e controlar riscos para a ocorrência de sinistros, (ii) a segunda refere-se a eventos raros, dessa forma entende-se que quanto mais raro um evento é de ser encontrado, mais difícil se torna a previsão de sua ocorrência e menores são as chances de evitar que o sinistro ocorra, (iii) a terceira está relacionada à complexidade, ou seja, a união de diversos eventos relevantes, podendo gerar dificuldades cognitivas, (iv) a quarta está associada à capacidade cognitiva, ocorre devido a reduzida necessidade que o motorista tem de tomar decisões de curto prazo e concentrar a atenção na via.

As quatro leis são definidas como leis da causalidade de sinistros e são hipóteses empiricamente testáveis. As leis não definem que existem apenas esses fatores que influenciam a ocorrência de sinistros, podendo ser somados aos fatores anteriormente descritos (ELVIK, 2006).

De acordo com a Organização Mundial da Saúde (2018), os sinistros podem estar atrelados a diversos fatores, como: rápido crescimento dos centros urbanos, padrões de segurança precários (tanto a segurança dos veículos quanto da via), falta de fiscalização, distração ou cansaço ao dirigir, dirigir sob o efeito do uso de drogas lícitas ou ilícitas, excesso de velocidade e falta do uso de elementos de segurança (cadeirinha para crianças, cinto de segurança, capacete, dentre outros).

A partir do exposto, entende-se a importância da previsão de sinistros de tráfego. Os modelos de segurança viária vêm sendo utilizados para auxiliar na compreensão e mitigação das ocorrências observadas nos trechos estudados.

2.2.4.1 *Direção Agressiva*

A condução perigosa é caracterizada por uma ampla gama de decisões que colocam em risco a segurança dos condutores e de outros usuários da estrada. O *Fatality Analysis Reporting Systems* indica que nos Estados Unidos cerca de 53% dos sinistros fatais são devidos à condução agressiva (AAA FOUNDATION FOR TRAFFIC SAFETY, 2016). Além disso, uma pesquisa de motoristas realizada pela Fundação para Segurança no Trânsito da *American Automobile Association* em 2011 indicou que cerca de 90% dos motoristas viam a direção perigosa como uma ameaça muito séria ou algo sério à sua própria segurança (AAA FOUNDATION FOR TRAFFIC SAFETY, 2016).

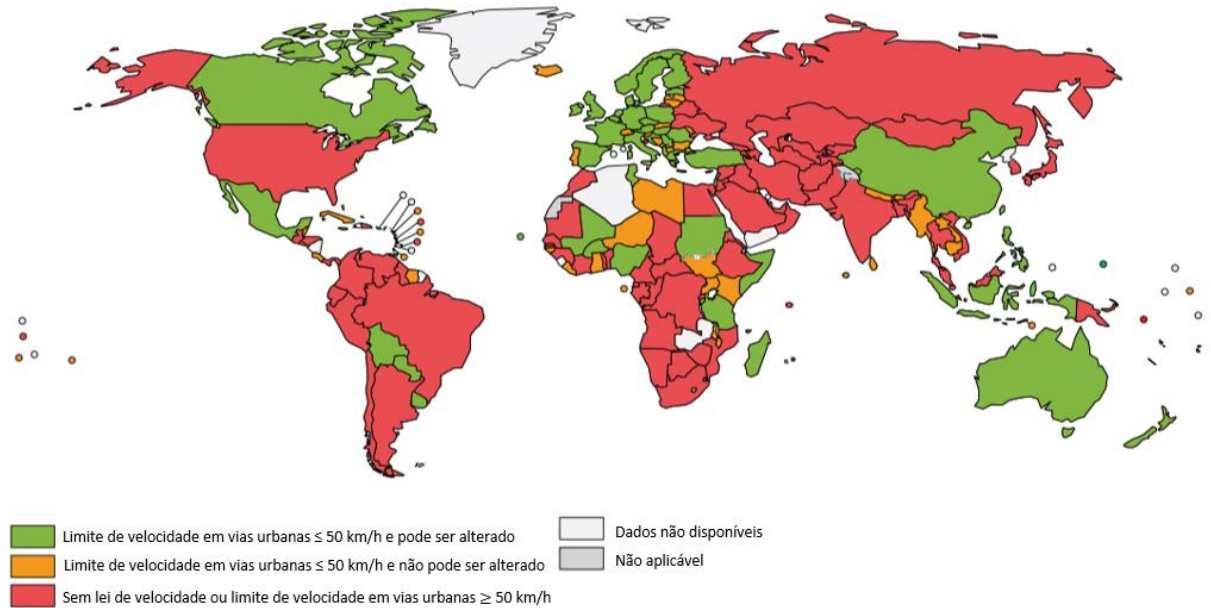
No Brasil não há definição oficial quanto à classificação de direção perigosa ou agressiva, consta apenas a definição de manobra perigosa no Art. 175 do Código Brasileiro de Trânsito (CBT) que é caracterizada pela arrancada brusca, derrapagem ou frenagem com deslizamento ou arrastamento de pneus, sendo a infração classificada como gravíssima.

2.2.4.2 *Controle de Velocidade*

A velocidade apresenta grande importância nas análises de segurança viária devido ao grande impacto desta na gravidade do sinistro. A alta velocidade necessita de maior tempo de frenagem até a parada fazendo com que a diferença entre a distância de frenagem de um veículo e outro mais rápido seja consideravelmente maior. Tal fator contribui para maior gravidade de sinistros em geral (ELVIK ET AL., 2015).

Em suas diretrizes, a OMS propõe a adoção de 50km/h como velocidade limite em áreas urbanas WHO (2018). A Figura 3 ilustra a velocidade vigente nos países com base na proposta da OMS.

Figura 3 – Países de acordo com as leis de velocidade vigentes em 2017



Fonte: WHO, 2018.

Entendendo a gravidade do tema, o Brasil possui Plano Nacional de Controle Eletrônico de Velocidade (PNCV), desenvolvido em 2016 pelo Ministério da Infraestrutura. O Plano possui metodologias, procedimentos e critérios para a implantação de dispositivos para a redução de velocidade. Porém não há diretrizes claras de fiscalização do sistema, ficando a cargo das autoridades locais.

Por esse motivo, no Brasil existem definições sobre o limite de velocidade de acordo com o tipo de via. De acordo com o Art. 61 do Cap. III do CTB em vias nas quais não há sinalização com indicação de velocidade, a velocidade máxima permitida será de:

1. Nas vias urbanas:
 - a. 80 km/h, nas vias de trânsito rápido;
 - b. 60 km/h, nas vias arteriais;
 - c. 40 km/h, nas vias coletoras;
 - d. 30 km/h, nas vias locais;
2. Nas vias rurais:
 - a. nas rodovias de pista dupla:
 - i. 110 km/h para automóveis, camionetas, caminhonetes e motocicletas;
 - ii. 90 km/h para os demais veículos;

- b. nas rodovias de pista simples:
 - i. 100 km/h (cem quilômetros por hora) para automóveis, camionetas, caminhonetes e motocicletas;
 - ii. 90 km/h (noventa quilômetros por hora) para os demais veículos;
- c. nas estradas: 60 km/h (sessenta quilômetros por hora).

Qualquer motorista que não seguir o limite de velocidade definido na via pode estar sujeito a multa. Porém, devido ao alto custo de fiscalização, apenas alguns pontos são monitorados. A monitoração das velocidades dos veículos pode ser feita por meio de controladores de velocidade, que geram multas quando o motorista ultrapassa o limite de velocidade.

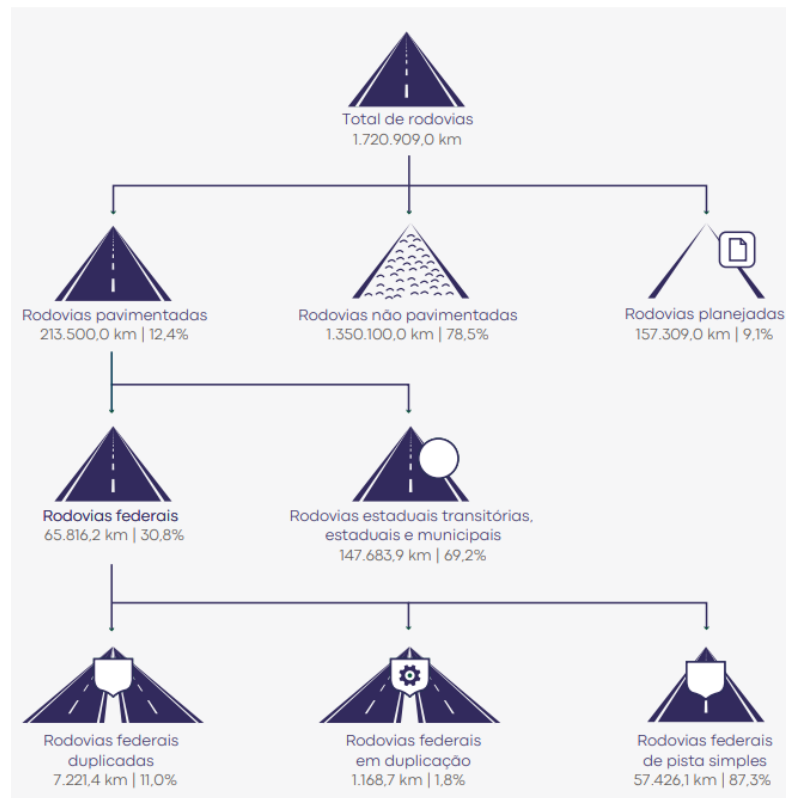
Contudo, a efetividade dessas medidas depende também da qualidade da infraestrutura rodoviária. É nesse contexto que as rodovias concedidas desempenham um papel crucial, pois as concessões podem trazer investimentos significativos em infraestrutura e manutenção.

2.3 RODOVIAS CONCEDIDAS NO BRASIL

2.3.1 Infraestrutura Brasileira

A rede rodoviária brasileira, conforme o Sistema Nacional de Viação (SNV), abrange aproximadamente 1,7 milhão de quilômetros de extensão, dos quais apenas 12,4% são pavimentados. A distribuição da malha rodoviária, de acordo com o tipo de superfície e jurisdição, pode ser visualizada na Figura 4.

Figura 4 - Distribuição da Malha Rodoviária Brasileira



Fonte: CNT, SEST SENAT e ITL (2023)

Ao analisarmos a malha rodoviária em relação à extensão territorial, observamos que o Brasil possui apenas 25 quilômetros de vias pavimentadas para cada mil quilômetros quadrados de território. Em comparação com países de dimensões semelhantes, como China e Estados Unidos, essa disparidade é evidente. A China conta com 447 quilômetros e os EUA com 437 quilômetros de vias pavimentadas para cada mil quilômetros quadrados de território, destacando-se ainda pela diversidade de modos de transporte disponíveis (CNT ET AL., 2023).

Além da escassez de vias pavimentadas, o sistema rodoviário brasileiro se destaca como o principal meio de transporte do país, movimentando 65% das cargas e 95% dos passageiros. No entanto, um problema evidente surge ao analisar a qualidade do pavimento dessas vias. De fato, cerca de 67% da extensão total da rede rodoviária apresenta condições de conservação classificadas como Regular, Ruim ou Péssima (CNT ET AL., 2023).

De acordo com Schwab (2019), o Brasil é o décimo país na categoria “tamanho de mercado”. Por outro lado, a pontuação para a malha viária varia de 1 (pior cenário) a 7 (melhor cenário), o Brasil recebeu a nota 3, ocupando a 116ª posição

de 141 países na classificação. Essa situação ressalta a necessidade de investimentos e melhorias na infraestrutura rodoviária para garantir a eficiência e segurança do transporte no Brasil.

2.4 REVISÃO SISTEMÁTICA DA LITERATURA

Este estudo utiliza a metodologia de Revisão Sistemática da Literatura (RSL) descrita em Kitchenham e Charters (2007). Esta metodologia descreve um processo rigoroso de revisão da literatura para um tópico de interesse. O processo garante que a revisão identifique, avalie e interprete todos os artigos de pesquisa relevantes de maneira reproduzível.

A revisão da literatura consiste em três fases principais: planejamento, execução e relatório. Durante o planejamento é definido um protocolo para o estudo, que inclui o objetivo do estudo, as questões de pesquisa a serem respondidas, as fontes de busca dos artigos, as palavras de busca, critérios de inclusão e exclusão, itens a serem extraídos dos estudos primários e métodos de análise a serem utilizados.

Durante a execução, as palavras de busca são aplicadas conforme especificado no protocolo, os critérios de inclusão e exclusão são aplicados para identificar os estudos primários, e todos os itens de dados são extraídos desses artigos.

Por último, durante o relatório, os dados extraídos são analisados e interpretados para responder às questões de investigação e para obter informações úteis do estudo.

2.4.1 Planejamento da Revisão

O objetivo deste estudo de revisão é identificar o impacto na gravidade de sinistro de trânsito considerando os fatores ambientais, viários e humanos. Para orientar a pesquisa, foram formuladas seis perguntas complementares:

- a) Quais fatores ambientais são mais frequentemente associados à gravidade dos sinistros de trânsito?
- b) Quais fatores humanos têm maior impacto na gravidade dos sinistros de trânsito?

- c) Quais fatores viários apresentam maior impacto na gravidade de um sinistro de trânsito?
- d) Como os diferentes tipos de vias influenciam a gravidade dos sinistros de trânsito?
- e) Quais estratégias de mitigação têm sido eficazes na redução da gravidade dos sinistros de trânsito?
- f) Existe variação na relação entre fatores ambientais, humanos e viários e a gravidade dos sinistros de trânsito entre diferentes regiões do mundo?

Para responder a tais perguntas foram realizadas buscas no Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Este portal, vinculado ao Ministério da Educação, oferece acesso unificado a periódicos e bases de dados, tanto de acesso aberto quanto restrito. Para acessar esse conteúdo, utilizamos os recursos disponíveis na Universidade Federal de Santa Catarina (UFSC).

Devido à sua relevância na área de estudo e à eficácia de suas ferramentas de busca, selecionamos três bases de dados específicas: *Scopus*, *ScienceDirect*, *Web of Science* e *IEEE Xplore*.

Para a definição dos termos de busca utilizou-se palavras-chave com o intuito de responder as questões de pesquisa propostas. Foram utilizadas palavras-chave e seus sinônimos em inglês e português, conforme apresentado na Tabela 4, para garantir que a ferramenta de busca identificasse publicações em ambas as línguas.

Tabela 4 - Palavras-chave utilizadas na revisão sistemática da literatura

Palavras-chave em inglês	Palavras-chave em português
Traffic accident, road, severity, victim, fatality, logistic regression, collision	Sinistros rodoviários, rodovia, severidade, vítima, fatalidade, regressão logística, colisão

Fonte: Autora (2024)

Foram estabelecidos critérios rigorosos para a inclusão e exclusão de estudos na análise, com base nos títulos, resumos e palavras-chave. Os artigos foram inicialmente filtrados pelos títulos e resumos e, posteriormente, submetidos à leitura

integral e avaliação da qualidade. Os seguintes critérios de inclusão foram aplicados aos títulos, resumos e palavras-chave:

Critérios de Inclusão (CI):

CI1: Estudos publicados no período entre 2000 e 2023.

CI2: Trabalhos relacionados à área de Engenharia e Ciência da Computação;

CI3: Aderência pelo título;

CI4: Aderência pelo conteúdo.

Os seguintes critérios indicaram a exclusão de um estudo.

Critérios de Exclusão (CE):

CE1: Obras de conferências ou editoriais;

CE2: Estudos não acessíveis na íntegra;

CE3: Duplicatas de estudos;

CE4: Estudos que envolvem ciclistas e pedestres.

2.4.2 Execução da Revisão

As bases de dados mencionadas anteriormente foram verificadas e os critérios de inclusão e exclusão foram aplicados de forma sistemática para garantir a seleção adequada e consistente dos estudos a serem incluídos na análise. Com base nesses critérios, foram selecionados 21 artigos, listados na Tabela 5, onde é possível verificar o título, autores, ano de publicação e número de citações. Na Tabela 6, são apresentados o modelo de regressão utilizado, assim como as variáveis dependentes e independentes consideradas.

Tabela 5 - Trabalhos acadêmicos selecionados na RSL

Número	Nome	Autores	País	Modelo	Variável Dependente
1	Influence of Road Safety Barriers on the Severity of Motorcyclist Injuries in Horizontal Curves	Ferko et al. (2022)	Croácia	Regressão Logística Binária	1: Vítima Grave ou Fatal 0: Caso contrário
2	The effect of weather on the severity of multi-vehicle crashes a case study of Iran	Yazdani; Nassiri (2021)	Irã	Regressão Logística Ordenada	1: Sinistro apenas com danos materiais 2: Sinistros com feridos 3: cidentes com vítimas fatais
3	Modeling severities of motorcycle crashes using random parameters	Farid; Ksaibati (2021)	EUA	1: Regressão Logística Binária; 2: Regressão Logística Binária Mista	1: Vítima Fatal 0: Caso contrário
4	Analyzing injury severity of motorcycle at-fault crashes using machine learning techniques, decision tree and logistic regression models	Rezapour et al. (2020)	EUA	Regressão logística binária + Arvore de Classificação	1: Vítima Fatal ou Vítima com lesão 0: Sinistro com danos materiais
5	Examination of the severity of two-lane highway traffic barrier crashes using the mixed logit model	Rezapour; Wulff; Ksaibati, (2019)	EUA	Regressão Logística Binária Mista	1: Com Vítima Fatal 0: Sem Vítima
6	Identifying Fatality Risk Factors for the Commercial Vehicle Driver Population	Islam; Ozkul, (2019)	EUA	Regressão Logística Binária	1: Vítima Fatal 0: Caso contrário
7	Factors associated with crash severities in built-up areas along rural highways of Nevada: A case study of 11 towns	Shrestha; Shrestha (2017)	EUA	Regressão Logística Binária	1: Vítima com Lesão 0: Caso Contrário
8	Factors influential in making an injury severity difference to older drivers involved in fixed object–passenger car crashes	Dissanayake; Lu, (2002)	EUA	Regressão Logística Binária	Modelo 1: 1-Sinistro com Vítima; 0-Sinistro sem Vítima Modelo 2: 1-Sinistro com Vítima Moderada/Grave/Fatal; 0-

Número	Nome	Autores	País	Modelo	Variável Dependente
					Sinistro com Víctima Leve Modelo 3: 1-Sinistro com Víctima Grave/Fatal; 0-Sinistro com Víctima Moderada Modelo 4: 1-Sinistro com Víctima Fatal; 0-Sinistro com Víctima Grave Modelo 5: 1-Sinistro com Víctima Fatal; 0-Sinistro com Víctima Grave/Moderada/Leve Modelo 6: 1-Sinistro com Víctima Grave; 0-Sinistro com Víctima Moderada/Leve Modelo 7: 1-Sinistro com Víctima Moderada; 0-Sinistro com Víctima Leve
9	Application of classification algorithms for analysis of road safety risk factor dependencies	Kwon et al. (2015)	EUA	1: Modelo de Naive Bayes; 2: Árvore de Decisão; 3: Regressão Logística Binária.	1: Víctima Fatal ou Ferida 0: Sem Víctima
10	Survival hazards of road environment factors between motor-vehicles and motorcycles	Li et al., (2009)	Taiwan	1: Regressão Logística Binária; 2: Regressão de Risco Proporcional de Cox	1: Víctima Fatal 0: Caso contrário
11	Analyzing crash injury severity for a mountainous freeway incorporating real-time traffic and weather data	Yu; Abdel-Aty (2014)	EUA	1: Modelo Logístico de Parâmetro Fixo; 2: Modelo Logístico de Parâmetro Aleatório 3: Máquina de Vetores de Suporte	1: Sinistro Grave 0: Caso Contrário
12	Analysis of driver injury severity in truck-involved accidents using a non-parametric classification tree model	Chang; Chien (2013)	Taiwan	Modelo Não Paramétrico de Árvore de Classificação e Regressão (CART)	1: Víctima Fatal 2: Víctima Ferida 3: Víctima Ilesa

Número	Nome	Autores	País	Modelo	Variável Dependente
13	Traffic accidents involving fatigue driving and their extent of casualties	Zhang et al. (2016)	China	Regressão Logística Binária	1: Sinistro com Vítima 0: Sem Vítima
14	Factors affecting hospital admission and recovery stay duration of in-patient motor victims in Spain	Santolino et al. (2012)	Espanha	Regressão Tobit Baseado na Distribuição Generalizada	Número de dias hospitalizado
15	Incorporating real-time traffic and weather data to explore road accident likelihood and severity in urban arterials	Theofilatos, (2017)	Grecia	Regressão Logística Bayesiana	Número de sinistros Gravidade de sinistros
16	Bivariate ordered-response probit model of driver's and passenger's injury severities in collisions with fixed objects	YAMAMOTO; SHANKAR (2004)	EUA	Modelo Probit Bivariado de Resposta Ordenada	1: Sinistro com Danos Materiais; 2: Sinistro com Vítima Leve; 3: Sinistro com Vítima Moderada; 4: Sinistro com Vítima Grave; 5: Sinistro com Vítima Fatal
17	Factors propelling fatalities during road crashes: A detailed investigation and modelling of historical crash data with field studies	MOHANTY et al. (2022)	India	1: Regressão Logística Binária 2: Árvore de Classificação e Regressão (CART); 3: <i>K-Means</i>	1: Sinistro Fatal 0: Caso Contrário
18	Toll road crash severity using mixed logit model incorporating heterogeneous mean structures	DAS et al. (2024)	EUA	Modelo Logit Misto	Modelo 1: 1: Sinistro com Vítimas Leves ou Moderadas 0: Sinistro com Danos materiais vs Ferimentos leves Modelo 2: 1: Vítimas Fatais ou Graves 0: Vítimas Moderadas
19	Exploring factors affecting injury severity of crashes in freeway tunnels	MA et al. (2016)	China	Modelo Logit Ordenado Generalizado	1: Vítima Fatal 2: Vítima Ferida 3: Vítima Ilesa

Número	Nome	Autores	País	Modelo	Variável Dependente
20	An investigation of influential factors of downgrade truck crashes: A logistic regression approach	MOOMEN; REZAPOUR; KSAIBATI (2019)	EUA	Regressão Logística Binária	1: Sinistro Envolvendo Veículo Pesado 0: Sinistro Envolvendo Outros Veículos
21	Exploring factors affecting the injury severity of freeway tunnel crashes: A random parameters approach with heterogeneity in means and variances	PERVEZ; LEE; HUANG (2022)	China	Modelo de parâmetros aleatórios com heterogeneidade em médias e variâncias	1: Vítima Grave/Fatal 0: Caso Contrário

Fonte: Autora (2024)

Tabela 6 – Demais informações relevantes sobre modelos utilizados nos trabalhos acadêmicos selecionados na RSL

Número	Variável independente	Objetivo	Filtros
1	Característica do sinistro: Velocidade inadequada, fora da estrada, veículo único; Característica da via: Curva; Característica temporal: durante a noite.	Análise da efetividade de barreiras para prevenção de fatalidades em sinistros envolvendo motocicletas	Sinistro envolvendo apenas um veículo e apenas motocicletas que saíram de pista ao se acidentarem.
2	Característica da via: equipamento rodoviário (guarda-corpos, sistemas de iluminação e três faixas), tráfego de veículos; Característica ambiental: Temperatura abaixo de -7°C, temperatura acima de 19°C, precipitação superior de 5mm/d; Característica temporal: dia útil, outono, inverno.	Principais fatores ambientais que podem aumentar o risco de gravidade de um sinistro	Sem filtro
3	Característica do condutor: idade, gênero, uso de álcool; Característica do sinistro: tipos de sinistro; Característica da via: duas pistas, condição da pista, velocidade permitida, perfil da pista; Característica temporal: dia da semana, período do dia.	Risco de sinistro grave envolvendo apenas uma motocicleta ou vários veículos e motocicleta	Sinistros envolvendo motocicleta
4	Característica do condutor: gênero, idade, uso de álcool; Característica do sinistro: ponto de impacto, a manobra do veículo, o tráfego, velocidade, número de veículos envolvidos; Característica da via: curvas verticais e horizontais, tipo de local, limite de velocidade; Característica ambiental: condições meteorológicas, condições da estrada e condições de iluminação; Característica temporal: dia da semana, período do dia.	Analisar qual o modelo que se comporta melhor na análise da gravidade de sinistros envolvendo motocicleta	Sinistros envolvendo motocicleta

5	<p>Característica do condutor: gênero, idade;</p> <p>Característica do sinistro: houve ação inadequada do motorista;</p> <p>Característica da via: condição da pista, traçado da pista, limite de velocidade;</p> <p>Característica do veículo: tipo de veículo;</p> <p>Característica ambiental: horário do dia.</p>	Analisar a performance das barreiras	Sinistros onde ocorre colisão com barreiras laterais
6	<p>Característica do condutor: status de carta de condução;</p> <p>Característica do sinistro: número de vítimas, tipo de sinistro, número de veículos envolvidos, velocidade do veículo;</p> <p>Característica da via: classe da rodovia, número de faixas, existência de barreira, condição da pista;</p> <p>Característica do veículo: peso, tipo de veículo;</p> <p>Característica ambiental: condição climática;</p> <p>Característica temporal: período do dia, dia da semana, período do ano.</p>	Analisar se o sinistro fatal envolvendo caminhão está relacionado com a idade. Foram gerados 4 modelos, sendo cada um com uma faixa de idade	Sinistro envolvendo caminhão
7	<p>Característica do condutor: fadiga, uso de álcool;</p> <p>Característica do sinistro: tipo de veículo, causa do sinistro, número de veículos envolvidos;</p> <p>Característica ambiental: condições meteorológicas, condições de iluminação;</p> <p>Característica temporal: período do ano, dia da semana, período do dia.</p>	Determinar os fatores associados à gravidade dos sinistros ocorridos em áreas urbanas ao longo das rodovias rurais de Nevada, EUA	Sem filtro
8	<p>Característica do condutor: gênero, uso de álcool, motorista ejetado, culpado pelo sinistro, município de residência;</p> <p>Característica do sinistro: velocidade do veículo, tipo de sinistro;</p> <p>Característica da via: tipo de local, traçado da pista, tipo de estrada;</p> <p>Característica do veículo: problema mecânico</p> <p>Característica ambiental: condições meteorológicas, condições de iluminação;</p>	Identificar fatores que influenciam a gravidade das lesões em motoristas mais velhos em sinistros com objetos fixos	Colisão com objeto fixo envolvendo apenas um veículo, onde o motorista possui idade igual ou superior a 65 anos
9	<p>Característica do condutor: uso do celular, movimento antes da colisão, uso do álcool, etnia;</p> <p>Característica da via: preservação da pista, intersecção, tipo de local;</p> <p>Característica do veículo: marca do veículo;</p> <p>Característica ambiental: condições de iluminação;</p> <p>Característica temporal: dia da semana, mês do ano, ano;</p>	Comparar o modelo de Naive Bayes, árvore de decisão e regressão logística binária	Sem filtro

10	<p>Característica do condutor: idade, gênero; Característica do sinistro: tipo de sinistro; Característica da via: tipo de local; Característica ambiental: condições de iluminação, condições de visibilidade; Característica temporal: ano.</p>	Investigar os efeitos das variáveis na incidência de sinistros e no tempo até a morte.	Não foram considerados óbitos sem número de identificação. Indivíduos envolvidos em múltiplas colisões não foram considerados. O consumo de álcool, que estava associado a lesões mais graves em sinistros, não foi considerado porque os dados relevantes não estavam disponíveis.
11	<p>Característica da via: Inclinação do segmento, velocidade; Característica ambiental: condições climáticas e temperatura.</p>	Comparação entre o 1 Modelo Logístico de parâmetro fixo, Modelo Logístico de parâmetro aleatório e Máquina de Vetores de Suporte.	Sem filtro
12	<p>Característica do condutor: gênero, uso do cinto, uso de álcool; Característica do sinistro: tipo de sinistro, causa do sinistro; Característica do veículo: tipo do veículo.</p>	Estabelecer a relação entre a gravidade das vítimas e as variáveis independentes	Sinistros envolvendo caminhões
13	<p>Característica do condutor: gênero, idade, experiência, ocupação, status da carta de condução; Característica do sinistro: tipo de sinistro, causa do sinistro; Característica da via: tipo de estrada, número de faixas, equipamento rodoviário, condição da pista; Característica do veículo: tipo do veículo, condição do veículo; Característica ambiental: condições climáticas, iluminação; Característica temporal: dia de semana, estação do ano, ano.</p>	Estabelecer a relação entre a gravidade das vítimas e as variáveis independentes	Sem filtro
14	<p>Característica do condutor: gênero, idade, região lesionada; Característica da vítima: gênero, idade, região lesionada.</p>	Analisa as Característica das lesões que explicam a hospitalização de vítimas de sinistros de trânsito ocorridos na Espanha	Apenas vítimas hospitalizadas e que se recuperaram ou possuem lesões estáveis
15	<p>Característica do sinistro: gravidade do sinistro, tipo de sinistro; Característica da via: traçado da pista, equipamento rodoviário, condição do pavimento; Característica ambiental: condição do tempo;</p>	Analisar o impacto de variáveis como tráfego e informações meteorológicas no número e gravidade de sinistros	Sem filtro
16	<p>Característica do condutor: gênero, idade, fadiga, uso de álcool; Característica do sinistro: causa do sinistro, velocidade, número de passageiros; Característica da via: intersecção, perfil da pista, tipo de local, equipamento rodoviário;</p>	Analisar o impacto das variáveis na gravidade do condutor e passageiro mais gravemente ferido	Sinistros envolvendo apenas um veículo que colidiu com objeto fixo

	<p>Característica do veículo: tipo do veículo, condição do veículo.</p> <p>Característica ambiental: condição do tempo, iluminação;</p> <p>Característica temporal: estação do ano, ano;</p>		
17	<p>Característica da via: existência de cruzamento, presença de acostamento, tipo de local, sinalização, velocidade permitida;</p> <p>Característica do veículo: tipo de veículo;</p> <p>Característica ambiental: condições climáticas;</p> <p>Característica temporal: dia da semana, estação do ano, período do dia;</p>	Desenvolver um modelo de previsão de fatalidades em sinistros e propor um índice de gravidade de sinistros com <i>k-means</i>	Sem filtro
18	<p>Característica do condutor: gênero, idade, houve infração;</p> <p>Característica do sinistro: tipo de sinistro, houve sinistro anterior;</p> <p>Característica da via: velocidade permitida, iluminação, condição da pista;</p> <p>Característica do veículo: tipo de veículo, número de ocupantes;</p> <p>Característica temporal: período do dia.</p>	Compreender os fatores que afetam a gravidade dos sinistros em rodovias pedagiadas.	Sem filtros
19	<p>Característica do sinistro: tipo de sinistro, número de veículos envolvidos, tipo de veículo envolvido, velocidade</p> <p>Característica da via: comprimento do túnel, tipo de local, traçado da pista, tráfego;</p> <p>Característica ambiental: condições climáticas;</p> <p>Característica temporal: período do dia, dia da semana, estação do ano.</p>	Analisar principais variáveis que impactam a gravidade de um sinistro que ocorre em túnel	Sinistros que ocorrem em tuneis.
20	<p>Característica do condutor: idade, gênero;</p> <p>Característica do sinistro: velocidade, número de veículos envolvidos</p> <p>Característica da via: velocidade permitida, condição da pista;</p> <p>Característica ambiental: condições climáticas;</p> <p>Característica temporal: dia da semana, período do dia.</p>	Análise de fatores significativos que impactam as colisões de caminhões em descidas	Sinistros envolvendo caminhões em serra
21	<p>Característica do condutor: gênero, idade, experiência, infrações;</p> <p>Característica do sinistro: tipo de sinistro;</p> <p>Característica da via: tipo de local, traçado da pista;</p> <p>Característica do veículo: tipo de veículo</p> <p>Característica ambiental: condições climáticas;</p> <p>Característica temporal: período do dia, dia da semana, estação do ano.</p>	Analisar principais variáveis que impactam a gravidade de um sinistro que ocorre em túnel	Sinistros que ocorrem em tuneis.

Fonte: Autora (2024)

Após a análise dos trabalhos acadêmicos selecionados na RSL, foi possível identificar os principais filtros aplicados na análise de sinistros, bem como as principais variáveis dependentes e independentes utilizadas.

2.4.3 Relatório

Dos 21 artigos analisados, 15 apresentaram estudos relacionados à gravidade dos sinistros, com pelo menos um modelo utilizando vítima fatal ou grave como variável dependente. Desses, 4 artigos (3, 4, 9, 17) realizaram testes para determinar o melhor modelo para prever a gravidade dos sinistros. Entre os modelos testados, a regressão logística binária foi um dos métodos avaliados, e em 2 desses 4 artigos, o modelo de regressão logística binária demonstrou o melhor desempenho na previsão da gravidade dos sinistros.

A análise permite separar as variáveis independentes em subgrupos, como característica do condutor, do sinistro, da via, do veículo, ambientais e temporais.

Característica do condutor: 71% dos artigos apresentaram este grupo de variáveis. Abrangem tanto característica físicas, como idade e gênero do condutor, quanto característica culturais, como etnia e local de residência, além de fatores psicológicos, como o uso de álcool e outras situações.

Característica do sinistro: 76% dos artigos apresentaram este grupo de variáveis. Estão relacionadas ao sinistro em si, como o tipo de sinistro, a velocidade no momento da ocorrência e o número de veículos envolvidos.

Característica da via: 86% dos artigos apresentaram este grupo de variáveis, sendo o mais presente nos artigos estudados. Referem-se à infraestrutura da via, como o número de faixas, presença de curvas ou declives, tipo de uso do solo, além de dados sobre a capacidade viária, como tráfego e velocidade.

Característica do veículo: 38% dos artigos apresentaram este grupo de variáveis, sendo o menos presente nos artigos avaliados. Geralmente, referem-se ao tipo do veículo e, em alguns casos, à marca do veículo, peso e condição do mesmo.

Característica ambiental: 76% dos artigos apresentaram este grupo de variáveis. Incluem informações meteorológicas, como temperatura e condições climáticas (chuva, neve), além de dados sobre iluminação e visibilidade.

Característica temporais: 71% dos artigos apresentaram este grupo de variáveis. Referem-se ao período no tempo, como hora do dia, dia da semana, estação do ano e feriados.

Os estudos descritos desempenharam um papel fundamental no embasamento teórico desta pesquisa, uma vez que a classificação de variáveis dependentes em vítimas moderadas, graves e fatais é amplamente discutida na literatura. As variáveis dependentes utilizadas em trabalhos anteriores também foram de extrema importância, proporcionando um direcionamento mais claro sobre quais variáveis poderiam ser incluídas no modelo.

Dessa forma, identificou-se a possibilidade de empregar um modelo logístico binário com o objetivo de compreender as principais variáveis que influenciam a gravidade dos sinistros no Brasil, país não contemplado nas bases de dados analisadas.

2.5 MODELO DE REGRESSÃO LOGÍSTICA BINÁRIA

Ao longo dos anos, uma série de estudos tem aplicado diversas técnicas metodológicas para explorar a relação entre a gravidade dos sinistros e seus fatores contribuintes, incluindo o uso de modelos de regressão logística (WAHAB; JIANG, 2019). Destaca-se que entre os modelos de regressão, o modelo de regressão logística binária é amplamente reconhecido e utilizado para essa finalidade (KONONEN; FLANNAGAN; WANG, 2011; KWON; RHEE; YOON, 2015; OÑA; MUJALLI; CALVO, 2011)

A regressão logística binária é uma técnica estatística usada para modelar a relação entre uma variável dependente Y dicotômica (ou seja, que possui apenas duas categorias, como "sim" ou "não", "1" ou "0"). Sendo as variáveis independentes dicotômicas, nominais, ordinais ou quantitativas (HOSMER; LEMESHOW; STURDIVANT, 2013).

Considerando sinistros rodoviários como a base da análise, a variável independente seria a ocorrência de um sinistro com vítima e as variáveis independentes seriam os fatores associados ao sinistro, como idade do condutor, condição do tempo, traçado da pista e outras características que podem ser entendidas como relevantes para o modelo.

Conforme apontado por Moomen, Rezapour e Ksaibati (2019), na estrutura do modelo de regressão logística binária, a probabilidade (P_i) de encontrar um sinistro com vítima (i) é definida como segue.

$$E[Y|X] = \pi(x) = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}} \quad (1)$$

Em que:

Y : é a variável dependente;

x_1, \dots, x_p : são as variáveis independentes, ou seja, os fatores que contribuem para o sinistro com vítima ocorrer;

β_1, \dots, β_p : os coeficientes estimados a partir do conjunto de dados por meio da máxima verossimilhança.

Uma das propriedades da regressão logística é a possibilidade de linearização por meio da transformação logit, que é a função logarítmica da razão de chances. Ao aplicar a transformação logit, a regressão logística é transformada em uma regressão linear, o que simplifica a interpretação dos coeficientes das variáveis independentes (Tay; Rifaat; Chin, 2008b). Logo, a função logística pode ser expressa por:

$$\text{logit}(P_i) = \ln\left(\frac{P_i}{1 - P_i}\right) = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n \quad (2)$$

Onde:

P_i : probabilidade condicional de Y ser igual a 1 dado X

$\text{logit}(P_i)$: o logit da probabilidade de P

β_0 : o intercepto da regressão

β_1, \dots, β_n : coeficientes associados a cada uma das variáveis independentes

x_1, \dots, x_p

2.5.1 Ajuste do modelo

O método de máxima verossimilhança é amplamente empregado para ajustar o modelo de regressão logística binária, com o objetivo de estimar os coeficientes ($\beta_0 + \beta_1 + \dots + \beta_n$) com base nos dados observados. Esta abordagem estatística

busca encontrar os valores dos parâmetros que maximizam a probabilidade de reproduzir o padrão de observação dos dados amostrados, utilizando a função de verossimilhança. As estimativas dos parâmetros, obtidas por meio deste método, são consideradas as melhores estimativas dentro do contexto do modelo, uma vez que maximizam a probabilidade de reproduzir o padrão observado nos dados amostrados (HOSMER; LEMESHOW; STURDIVANT, 2013).

A máxima verossimilhança (\hat{l}), compara os resultados ($y=1$ se o sinistro ocorreu e $y=0$ caso contrário) com as probabilidades previstas de acordo com a equação a seguir.

$$\hat{l} = \sum_{i=1}^n y_i \ln(P_i) + (1 - y_i) \ln(1 - P_i) \quad (3)$$

Onde \hat{l} é a log-verossimilhança, y_i é a variável resposta para a i -ésima observação (1 se o evento ocorreu e 0 caso contrário), e P_i é a probabilidade prevista pelo modelo de que a i -ésima observação tenha $y_i = 1$. O logaritmo natural é aplicado a P_i e $1 - P_i$ para garantir que a função de verossimilhança seja maximizada.

2.5.2 Métodos de Avaliação do Modelo Logístico

O presente capítulo utilizará como principal referência o livro dos autores Hosmer, Lemeshow e Sturdivant (2013) reconhecido por sua importância e influência na área de estatística, especialmente em regressão logística. O livro define várias medidas de qualidade de ajuste que podem ser incluídas na análise, tais como o teste de razão de verossimilhança, AIC, BIC, *deviance*, *pseudo R²* de *Nagelkerke*, teste de *Wald* e a curva ROC. Estas medidas são fundamentais para avaliar a precisão e a robustez dos modelos estatísticos, proporcionando uma visão abrangente do desempenho e da adequação do modelo aos dados observados.

2.5.2.1 Teste de Razão de Verossimilhança

O teste de razão de verossimilhança é realizado para avaliar se os parâmetros de um modelo de regressão logística influenciam significativamente a variável dependente. Este teste compara a log-verossimilhança do modelo saturado e modelo

reduzido. A diferença na log-verossimilhança entre esses modelos é usada para determinar se os parâmetros adicionais no modelo completo melhoram significativamente o ajuste aos dados (HOSMER; LEMESHOW; STURDIVANT, 2013).

$$D = -2 \ln \left[\frac{\text{Função da máxima verossimilhança do modelo reduzido}}{\text{Função da máxima verossimilhança do modelo saturado}} \right] \quad (4)$$

Onde o modelo saturado é aquele que contém todas as variáveis e suas interações, enquanto o modelo reduzido contém apenas as variáveis desejadas para o estudo. Logo o teste pode ser representado por:

$$D = -2 \sum_{i=1}^n \left[y_i \ln \left(\frac{\hat{\pi}_i}{y_i} \right) + (1 - y_i) \ln \left(\frac{1 - \hat{\pi}_i}{1 - y_i} \right) \right] \quad (5)$$

O teste possui as seguintes hipóteses:

$$\begin{cases} H_0: \beta_1 = \dots = \beta_p = 0 \\ H_1: \text{Existe algum } \beta \neq 0 \end{cases}$$

Ou seja, na hipótese nula (H_0) nenhuma das variáveis independentes tem efeito significativo na variável dependente. Enquanto que na hipótese alternativa (H_1), pelo menos uma das variáveis independentes tem efeito significativo na variável dependente.

Portanto, ao testar a hipótese nula, todos os coeficientes do modelo são igualados a zero, enquanto a hipótese alternativa é representada pelo modelo saturado, que mantém todos os seus coeficientes. Ao rejeitar a hipótese nula, concluímos que existe pelo menos uma variável independente que tem um efeito significativo na variável dependente

2.5.2.2 Critério de Informação de Akaike

Para determinar o melhor modelo a ser avaliado, é importante analisar os resíduos gerados. Dessa forma, Akaike (1987) apresenta o Critério de Informação de Akaike (AIC). O AIC é calculado visando minimizar a informação de *Kullback-Leibler* (Kullback e Leibler (1951)), uma medida da discrepância entre a distribuição real e o modelo proposto. Essa métrica é derivada da soma do logaritmo máximo da função

de verossimilhança com a penalidade associada ao número de parâmetros do modelo. Em essência, quanto menos informações são perdidas no modelo proposto, melhor sua qualidade, refletindo-se em valores menores de AIC.

$$AIC = -2\hat{l} + 2k \quad (6)$$

Onde \hat{l} é o máximo valor da função de log-verossimilhança para o modelo estimado e k é o número de parâmetros incluídos no modelo (variância).

De maneira geral, o AIC é uma ferramenta utilizada para comparar e selecionar modelos estatísticos. Ele é útil quando se deseja testar vários modelos e determinar qual se ajusta melhor a um conjunto de dados. Pois o método avalia a qualidade relativa dos modelos, sendo preferível escolher o modelo com o menor valor de AIC.

2.5.2.3 *Critério de Informação Bayesiano*

O Critério de Informação Bayesiano (BIC) recebe esse nome devido ao argumento Bayesiano utilizado por Schwarz (1978) para sua formulação. Ele é um critério de seleção de modelos que considera a probabilidade a posteriori. O BIC é derivado sob a hipótese assintótica de que os dados observados são gerados por modelos probabilísticos pertencentes à Família Exponencial.

Assim como o AIC, valores menores de BIC indicam um melhor ajuste do modelo. Além disso, o BIC também penaliza modelos com muitas variáveis, mas de uma maneira mais rigorosa. Essa abordagem torna o BIC particularmente útil para selecionar modelos em conjuntos de dados grandes ou complexos.

$$BIC = -2\hat{l} + k \ln(n) \quad (7)$$

Onde \hat{l} é o máximo valor da função de máxima verossimilhança para o modelo estimado e k é o número de parâmetros incluídos no modelo (variância) e n o tamanho da amostra.

2.5.2.4 Estatística Deviance

O *deviance*, ou desvio, em um modelo logístico corresponde à soma dos quadrados dos resíduos, similar ao modelo de regressão linear. Esta definição é estabelecida por Hosmer et al. (2013) como:

$$Dev = 2Ln \left[\frac{L_{red}}{L_{sat}} \right] \quad (8)$$

Onde L_{red} e L_{sat} representam, respectivamente, a verossimilhança do modelo reduzido e a verossimilhança do modelo saturado. A estatística *deviance* segue uma distribuição X^2 com $n-q$ graus de liberdade, sendo um indicativo de melhor ajuste quanto menor for o seu valor.

2.5.2.5 Pseudo R^2

O pseudo R^2 é uma abordagem alternativa para avaliar a eficácia do modelo, semelhante ao conceito de R^2 na regressão linear. Enquanto na regressão linear, o R^2 mede o impacto linear das variáveis independentes sobre a variável dependente, na regressão logística, calcular o R^2 não é viável devido à natureza probabilística da variável dependente (HOSMER et al., 2013)

Por isso, na regressão logística, é comum recorrer ao pseudo R^2 , derivado dos testes de Cox; Snell (1989) e Nagelkerke (1991). No entanto, é importante ressaltar que esses testes são parâmetros e devem ser interpretados com prudência e sensibilidade pelo pesquisador.

Em termos de interpretação, o valor do pseudo R^2 varia de 0 a 1, indicando a porcentagem dos dados explicada pelo modelo. Quanto maior o valor do pseudo R^2 , mais explicativo é o modelo e melhor ele se ajusta aos dados observados na amostra.

2.5.2.5.1 R^2 de Cox e Snell

Primeiramente tem-se \hat{l} (modelo ajustado) definido como a probabilidade condicional da variável de respostas dadas às variáveis preditoras. Se houver n observações no conjunto de dados, então \hat{l} é o produto das n probabilidades,

enquanto \hat{l}_0 representa o modelo inicial. Portanto, ao tomar a raiz n-ésima do produto $\hat{l} - \hat{l}_0$ obtém-se uma estimativa da probabilidade de cada valor (COX; SNELL, 1989). Dessa forma, o pseudo R^2 de Cox e Snell é obtido de uma transformação estatística usada para verificar a convergência de uma regressão logística, representado através da equação a seguir:

$$R_{CS}^2 = 1 - e^{\frac{2\hat{l}_0 - 2\hat{l}}{n}} \quad (9)$$

Onde:

R_{CS}^2 : é a função do pseudo R^2 de Cox e Snell;

\hat{l} é o modelo ajustado;

\hat{l}_0 é o modelo inicial; e n representa cada observação da amostra.

Note-se que em seu máximo nunca terá valor igual a 1 devido ao fato de que esta estatística não o atinge, mesmo quando o ajustamento é perfeito.

2.5.2.5.2 R^2 de Nagelkerke

Nagelkerke propôs uma modificação no pseudo R^2 de Cox e Snell para torná-lo mais informativo. Ele ajustou o pseudo R^2 para variar de 0 a 1, onde 1 indica um ajuste perfeito do modelo aos dados. Essa adaptação permite uma interpretação mais direta dos valores intermédios, tornando pseudo R^2 de Nagelkerke uma medida mais útil para entender quando o comportamento da variável dependente é explicado pelas variáveis independentes no modelo. Sendo assim, o compara a máxima verossimilhança do modelo desenvolvido (\hat{l}) e do modelo nulo (\hat{l}_0) (NAGELKERKE, 1991)

$$R_N^2 = \frac{1 - e^{\frac{2\hat{l}_0 - 2\hat{l}}{n}}}{1 - e^{\frac{2\hat{l}_0}{n}}} = \frac{R_{CS}^2}{1 - e^{\frac{2\hat{l}_0}{n}}} \quad (10)$$

Onde:

R_N^2 : é a função do pseudo R^2 de Nagelkerke;

R_{CS}^2 : é a função do pseudo R^2 de Cox e Snell;

\hat{l}_0 : é o modelo;

n: representa cada observação da amostra.

2.5.2.6 *Teste de Wald*

De acordo com Hosmer et al. (2013), além de ajustar o modelo aos dados, é importante avaliar a contribuição individual dos preditores. Na regressão linear, o teste t é comumente utilizado para verificar a qualidade do ajuste. Já na regressão logística, emprega-se a estatística de Wald, que indica se o coeficiente β_i de cada preditor é significativamente diferente de zero. Se o resultado for significativo, então o preditor contribui de forma significativa para a previsão de Y.

A equação do teste de Wald é calculada pela divisão entre a estimativa do máximo da verossimilhança do parâmetro β_i pelo estimador do erro padrão $SE(\beta_i)$, conforme apresentado na equação a seguir.

$$W_i = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)} \quad (11)$$

Essa estatística de teste segue uma distribuição normal padrão. Dessa forma, se o valor-p associado à estatística de Wald for menor que um certo nível de significância predefinido (geralmente 0,05), rejeitamos a hipótese nula e concluímos que o coeficiente é estatisticamente significativo. Caso contrário, não rejeitamos a hipótese nula e concluímos que o coeficiente não é estatisticamente diferente de zero (HOSMER et al., 2013).

2.5.2.7 *Resíduos*

De acordo com Turkman e Silva (2000) a análise de resíduos desempenha um papel fundamental na avaliação da qualidade de ajuste de um modelo estatístico. Os resíduos permitem não apenas verificar se o modelo escolhido, incluindo a distribuição e a função de ligação, é apropriado, mas ajuda a identificar observações que não são bem explicadas pelo modelo.

Para cada observação individual (representada pela i-ésima observação), o resíduo é definido como a diferença entre o valor observado e o valor estimado pelo

modelo. É importante que os resíduos estejam livres de *outliers*, que podem distorcer as análises.

Outliers podem ser identificados observando os resíduos padronizados, os quais devem estar dentro dos intervalos de $[-3, 3]$. Se *outliers* forem identificados, é recomendado revisar o modelo sem essas observações para determinar sua influência no modelo e decidir se devem ser removidos.

A obtenção dos resíduos padronizados é realizada com base nos resíduos de Pearson, conforme apresentado no próximo item.

2.5.2.7.1 Resíduo de Pearson

O resíduo de Pearson é calculado pela seguinte expressão:

$$r_{ip} = \frac{y_i - \hat{\mu}_i}{\sqrt{V(\hat{\mu}_i)}}, i = 1, 2, \dots, n \quad (12)$$

O resíduo r_{ip} representa a contribuição da i -ésima observação para o cálculo da estatística de Pearson generalizada, conforme expresso abaixo:

$$X^2 = \sum_{i=1}^n \frac{(y_i - \hat{\mu}_i)^2}{V(\hat{\mu}_i)} \quad (13)$$

Onde $V(\hat{\mu}_i)$ é a função de variância estimada para a distribuição do modelo em estudo.

2.5.2.7.2 Resíduos padronizados:

A partir dos resíduos de Pearson, podemos derivar os resíduos padronizados, que são resíduos ajustados pelo respectivo desvio padrão. Essa padronização nos permite comparar a magnitude dos resíduos entre diferentes observações e facilita a identificação de valores extremos ou *outliers*. Dessa forma, o resíduo de Pearson padronizado é dado por:

$$r_{iE} = \frac{y_i - \hat{\mu}_i}{\sqrt{V(\hat{\mu}_i)(1 - h_{ii})}} \quad (14)$$

Onde h_{ii} representa os valores da diagonal da matriz de projeção $H = W^{\frac{1}{2}}X(X'WX)^{-1}X'W^{1/2}$.

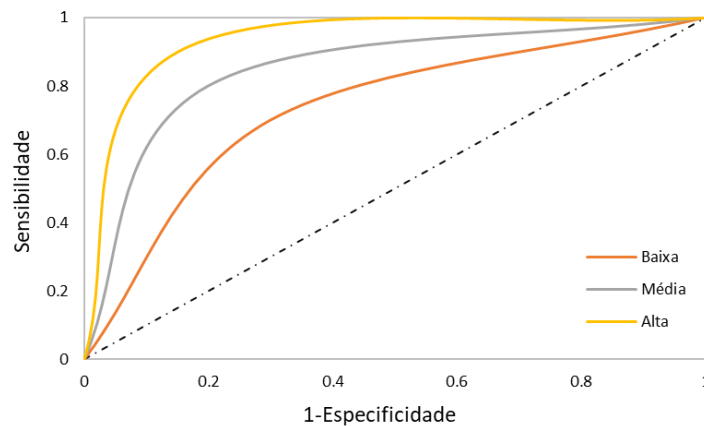
2.5.2.8 Curva Receiver Operator Characteristic

Conforme apresentado por Oña, Mujalli e Calvo (2011) a curva *Receiver Operator Characteristic* (ROC) é outro indicador que pode ser analisado. A curva ROC representa a taxa de verdadeiros positivos (Sensibilidade) versus a taxa de falsos positivos (1 - Especificidade) em um teste estatístico. A área sob a curva, também conhecida como *Area Under the Curve* (AUC), é particularmente útil para descrever o desempenho geral do teste, com um valor máximo de 1,00 indicando um teste perfeito e uma área ROC de 0,50 indicando um teste sem valor.

A sensibilidade é definida como a proporção de positivos corretamente identificados, enquanto a especificidade é a proporção de negativos corretamente identificados. A curva ROC é utilizada para determinar o ponto de corte ideal na curva, de modo a maximizar tanto a sensibilidade quanto a especificidade, minimizando as taxas de falsos positivos e falsos negativos (WAN et al., 2017).

Para Hosmer, Lemeshow e Sturdivant (2013) a área sob a curva ROC, que varia de 0,5 a 1, fornece uma medida da capacidade de discriminar entre indivíduos que experimentam o resultado de interesse e aqueles que não o experimentam. Quanto maior for o valor dessa área, melhor será a capacidade preditiva do modelo ajustado.

Figura 5 - Curvas ROC dos diferentes tipos de discriminação.



Fonte: Adaptado de Braga (2000)

Hosmer, Lemeshow e Sturdivant (2013) apresentam uma regra interpretativa da AUC, relacionada ao poder preditivo do modelo ajustado, da maneira que segue:

- Se $AUC = 0,5$, não indica discriminação;
- Se $0,7 \leq AUC < 0,8$, indica uma discriminação aceitável;
- Se $0,8 \leq AUC < 0,9$, indica uma discriminação excelente;
- Se $AUC \geq 0,9$, indica discriminação excepcional.

2.5.3 Razão de Chances (*Odds Ratio*)

De acordo com Hosmer et. al (2013), além das estatísticas de ajuste do modelo, os parâmetros são interpretados por meio de suas razões de chances, também chamada de *odds ratio* (OR). A razão de chances é uma medida que expressa a probabilidade de um evento ocorrer, dado o efeito de um parâmetro específico, em comparação com a probabilidade de o evento não ocorrer, quando esse parâmetro não está presente, mantendo todos os outros fatores constantes

Considerando por exemplo, a razão de chances entre duas categorias A e B com probabilidades P_a e P_b , respectivamente, a razão de chances indica o quão provável um evento é em uma categoria em relação ao mesmo evento na outra categoria. O OR é calculado conforme a equação abaixo:

$$OR = \frac{\text{odds}(A)}{\text{odds}(B)} = \frac{\frac{P_a}{1 - P_a}}{\frac{P_b}{1 - P_b}} = \frac{P_a(1 - P_b)}{P_b(1 - P_a)} \quad (15)$$

A interpretação do OR pode ser feita de duas maneiras:

Em termos absolutos: $OR = 1$ indica que o evento tem a mesma probabilidade de ocorrer em ambos os grupos. Valores de OR maiores que 1 indicam que o evento é mais provável no grupo A em comparação com o grupo B, enquanto valores de OR menores que 1 indicam o contrário (Tay; Rifaat; Chin, 2008b; Zhang et al., 2016).

Em termos percentuais, calculamos $OR_p = (OR - 1) \times 100\%$. Uma $OR_p > 0\%$ sugere um aumento na probabilidade de o evento ocorrer no grupo A em relação ao grupo B, enquanto valores de $OR_p < 0\%$ indicam o oposto. Quando $OR_p = 0\%$, o evento tem a mesma probabilidade de ocorrer em ambos os grupos.

No modelo de regressão logística, a razão de chances também pode ser interpretada como a contribuição de cada variável para o modelo. Em outras palavras, é uma maneira de quantificar a variação na probabilidade de ocorrência do evento quando $X_i = 1$ em relação a $X_i = 0$.

Para a criação do modelo, além dos testes estatísticos, é fundamental incluir variáveis baseadas em cálculos previamente estudados. Uma dessas variáveis é a unidade padrão de severidade, que será analisada a seguir para entender como ela contribui para a modelagem, influenciando potencialmente a probabilidade dos eventos que estamos investigando.

2.6 UNIDADE PADRÃO DE SEVERIDADE

Além da frequência de sinistros, é possível estabelecer pesos diferentes de acordo com a gravidade dos sinistros para a definição da criticidade dos pontos analisados (AASHTO, 2009). Dessa forma, surge a Unidade Padrão de Severidade (UPS), criada pelo DENATRAN em 1987, na qual sinistros com vítimas e fatais possuem maiores pesos por colocarem mais pressão sobre a infraestrutura.

O Ministério dos Transportes (2020), entendendo a necessidade de tratar os sinistros envolvendo pedestres de forma separada aos sinistros com vítima ferida, aperfeiçoa o UPS, com a inclusão de um peso, como mostrado na Tabela 7, o qual apresenta o peso de cada sinistro de acordo com essa severidade.

Tabela 7 - Pesos atribuídos aos sinistros

Gravidade do sinistro de trânsito	Peso atribuído
Sinistro com somente danos materiais	1
Sinistro com vítima não fatal	4
Sinistro com vítima não fatal envolvendo pedestre	6
Sinistro com vítima fatal	13

Fonte: Ministério dos Transportes (2002)

O valor do segmento analisado é então expresso pela soma dos produtos do número de sinistros de acordo com a severidade do mesmo. Dessa forma, o cálculo do valor da severidade no trecho é feito através da equação:

$$UPS = (1 \times N_{ad}) + (4 \times N_{av}) + (6 \times N_{ap}) + (13 \times N_{af}) \quad (16)$$

Onde:

UPS: Unidade Padrão de Severidade;

N_{ad} : Número de sinistros com danos materiais;

N_{av} : Número de sinistros com vítimas não fatais;

N_{ap} : Número de sinistros com vítimas não fatais envolvendo pedestre;

N_{af} : Número de sinistros com vítimas fatais.

Dentre as dificuldades na análise de sinistros de acordo com a gravidade, a principal problemática está na forma de classificação da severidade do sinistro. Um sinistro pode gerar diversas vítimas, com grau de ferimentos diferentes. Dessa forma, uma das abordagens utilizadas é a classificação de acordo com a vítima mais grave. Assim, sinistros com óbitos, sempre serão sinistros com vítimas fatais (LORD; MANNERING, 2010).

Além do UPS, outra variável que pode ser incluída no modelo refere-se ao agrupamento dos sinistros de acordo com características pré-definidas, como geolocalização, tipo de veículo, idade do condutor, dentre outras. A análise baseada na geolocalização visa identificar os locais com maior incidência de sinistros, permitindo definir regiões de atenção prioritária para a implementação de medidas preventivas.

2.7 TÉCNICAS DE AGRUPAMENTO APLICADAS A SINISTROS RODOVIÁRIOS

Existem diversas formas de análise de sinistros, sendo duas abordagens principais: a análise descritiva e a análise preditiva.

A análise descritiva é utilizada para identificar grupos ou padrões nos dados, com base em semelhanças entre esses objetos decorrentes de interações entre variáveis independentes. Por outro lado, a análise preditiva busca prever eventos ou comportamentos futuros, mapeando um conjunto de valores de entrada para um valor de saída (COMI; POLIMENI; BALSAMO, 2022).

Dentre os modelos para análises descritivas, destaca-se a análise de *grupo* ou agrupamentos, um método de mineração de dados que permite segmentar os elementos dos dados em grupos com base na homogeneidade ou heterogeneidade dos dados (FRALEY; RAFTERY, 2002).

Antes de adentrarmos na descrição dos métodos de agrupamento, é fundamental entender o conceito de pontos críticos em relação aos sinistros rodoviários. Compreender essa definição é crucial para contextualizar a análise e aplicação dos métodos de agrupamento na identificação dos riscos associados a esses locais.

2.7.1 Ponto Crítico

Existem muitas definições de pontos críticos, dentre elas, Hakket e Mahalel (1978) entendem que pontos críticos não são necessariamente os locais com muitos sinistros, mas são aqueles em que o número real de sinistros é significativamente maior do que o número esperado. Enquanto Mahalel et al. (1982) entendem que melhorias viárias possuem limitações na redução de sinistros e define pontos críticos como locais nos quais há uma redução máxima esperada de sinistros, mesmo após intervenções viárias.

De acordo com Cheng e Washington (2005) um local pode apresentar grande número de sinistros devido a dois fatores: (i) problemas de segurança viária, podendo ser relacionada à via ou a alta exposição no local; ou (ii) uma flutuação aleatória durante o período de contagem de sinistros. Sendo assim, é necessário identificar

locais inseguros com baixo número de sinistros, chamados de falso negativo, e locais seguros com alto índice de sinistros, chamados de falsos positivos.

Dessa forma, a simples contagem de sinistros não indica qual dos cenários prevalece no local, sendo necessário utilizar de métodos estatísticos para a definição dos pontos críticos. E, a partir de identificação correta de tais pontos, pode-se definir investimentos assertivos para melhoria da segurança viária.

Para entender o problema de segurança viária, o primeiro passo é mapear os locais das ocorrências de sinistros, sendo possível definir, na maioria dos casos, os locais concentradores, ou pontos críticos. O desenvolvimento de índices ou indicadores de acidentalidade desses locais pode ser realizado de duas maneiras: de acordo com a frequência que ocorrem ou de acordo com a severidade (SATRIA; CASTRO, 2016).

A principal diferença entre as duas formas de análise deve-se ao fato da frequência, diferentemente da gravidade dos sinistros, ser um estudo de segurança objetiva, por ser uma medida independente do observador (AASHTO, 2009).

A partir dos estudos desenvolvidos no decorrer dos últimos anos, entende-se o uso do termo frequência média esperada de sinistros para descrever a estimativa da frequência média de sinistros a longo prazo, visto que um indicador a curto prazo não define a tendência do problema (AASHTO, 2009).

A partir da análise dos pontos críticos são definidas formas de reduzir a ocorrência de sinistros, como demonstrado pelo *Federal Highway Administration* (FHA, 2018) na Tabela 8

Tabela 8 - Medidas mitigadoras de sinistros de trânsito

Categoria	Medida
Pedestres/Ciclistas	<ul style="list-style-type: none"> - Construção de ciclovias e faixas de pedestres; - Melhoria da visibilidade e sinalização na faixa de pedestre; - Canteiros centrais ou ilhas de refúgio para pedestres; - Semáforos bom botoeira em vias rápidas para pedestres; - Alterar a via com foco nos ciclistas; - Melhoria das calçadas
Controle de Velocidade	<ul style="list-style-type: none"> - Limite de velocidade apropriado para todos os usuários da via; - Câmeras de segurança de velocidade; - Limite de velocidade variável
Geometria da Via	<ul style="list-style-type: none"> - Delineado compatível com a curva horizontal; - Sinalização horizontal e vertical correta;

Categoria	Medida
Intersecções	<ul style="list-style-type: none"> - Linhas de bordo largas; - Barreiras centrais adequadas; - <i>Rumble Strips</i> para vias com sentido duplo; - Inclusão da tecnologia <i>Safety Edge</i>. <ul style="list-style-type: none"> - Semáforo com borda refletiva; - Faixas de conversão a esquerda em vias com dois sentidos; - Melhoria do design e sinalização de acessos às vias; - Rotatórias; - Faixas exclusivas de conversão à esquerda e à direita; - Melhoria de sinalização e demais medidas de baixo custo; - Melhoria no tempo de amarelo
Demais Melhorias	<ul style="list-style-type: none"> - Planos para melhoria da segurança viária; - Melhoria da aderência do pavimento; - Melhoria da iluminação; - Auditorias de segurança rodoviária.

Fonte: Federal Highway Administration (2018)

A definição de pontos críticos pode variar dependendo do contexto, sendo uma das abordagens mais comuns na literatura a utilização de técnicas de agrupamento. Essas técnicas permitem identificar áreas onde a incidência de sinistros é significativamente alta ou onde ocorrem eventos com consequências graves, fornecendo uma abordagem sistemática para delinear esses locais de risco.

De acordo com o Thang e Kim (2011), existem muitos algoritmos de agrupamento na literatura, podendo ser categorizados em:

1. Método de particionamento (*k-means*);
2. Método hierárquico;
3. Método baseado em densidade (DBSCAN);
4. Método baseado em grade (CLIFE).

2.7.2 Método de particionamento (*k-means*)

O agrupamento é uma técnica empregada para agrupar informações com base em seu grau de similaridade, organizando os dados em um número definido de grupos. Essa abordagem é amplamente utilizada em diversas áreas do conhecimento.

O algoritmo *k-means* é uma heurística de agrupamento não hierárquico, que possui como premissa a divisão dos dados em *k* grupos com algum nível de homogeneidade. Em seguida é realizada a divisão de todos os itens pelos *k* grupos preestabelecidos e a melhor partição dos *n* itens será aquela que otimiza o critério

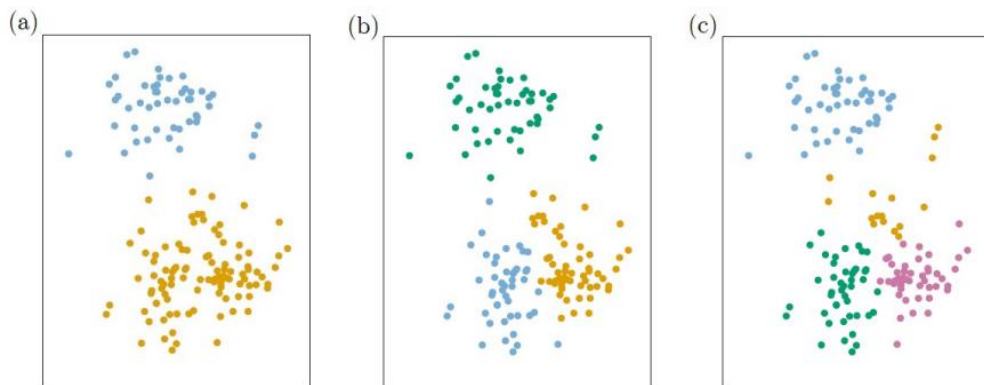
escolhido. Para realizar essa divisão o modelo mede a proximidade dos grupos usando a distância euclidiana entre os centroides deles (JAMES et al., 2013).

De forma simplificada, as etapas para a execução do algoritmo são:

- g) Define-se o número de agrupamento (k) que serão gerados a partir dos n itens;
- h) O algoritmo seleciona aleatoriamente k itens do conjunto de dados, para servir como centroides iniciais;
- i) Em seguida agrupa-se os objetos restantes aos grupos cujos centroides se encontram mais próximos;
- j) Atualiza-se o cálculo dos centroides dos grupos até que não ocorra variação significativa na distância mínima de cada objeto da base de dados a cada um dos centroides dos k grupos.

Para a definição das medidas de similaridade, a distância euclidiana é uma das métricas mais utilizadas para análises de grupo. O método euclidiano define que uma pequena distância entre dois objetos implica em uma forte semelhança, enquanto uma grande distância implica em baixa semelhança (ERMAN; ARLITT; MAHANTI, 2006).

Figura 6 - Agrupamento com algoritmo *k-means* (a) $K = 2$; (b) $K = 3$; (c) $K = 4$



Fonte: (JAMES et al., 2013)

Dessa forma, busca-se criar grupos com elementos mais semelhantes entre si quanto possível, enquanto que as semelhanças entre os elementos de grupos distintos sejam as menores possíveis. Tais divisões são criadas a fim de criar grupos com itens que possuam a maior semelhança.

De acordo com Linden (2009), o algoritmo possui grande simplicidade comparado com outros modelos, além da velocidade de execução ser extremamente alta, sendo necessário poucas interações para alcançar a estabilidade. Porém, o modelo apresenta duas principais críticas:

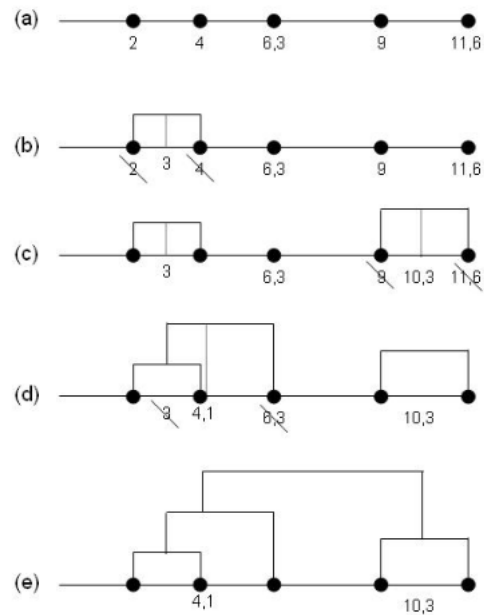
- a) a aleatoriedade na escolha dos centroides, o que pode gerar uma má separação dos conjuntos;
- b) a escolha do número de grupos é feita pelo usuário, fazendo com que um número grande pode quebrar um grupo em dois, enquanto um número pequeno pode englobar dois grupos naturais.

Além da seleção dos grupos iniciais que irão influenciar as agregações e o resultado final, é importante atentar para a escolha das variáveis. James et al. (2013) definem que variáveis que assumem o mesmo valor para todos os itens são pouco discriminatórias e a inclusão apenas polui a amostra. O mesmo ocorre para variáveis com muita discriminação e pouco relevantes para a solução do problema, devendo-se excluir tais variáveis do modelo.

2.7.3 Algoritmo Hierárquico

Os algoritmos hierárquicos criam, a partir de relacionamentos, uma hierarquia entre seus elementos, conforme mostrado na Figura 7. Existem duas formas de desenvolver o algoritmo: o aglomerativo, que cria conjuntos a partir de elementos segregados, e o divisivo, que parte de um conjunto maior e particiona o mesmo em conjuntos menores (LINDEN, 2009).

Figura 7 - Representação dos passos para o agrupamento do algoritmo hierárquico.



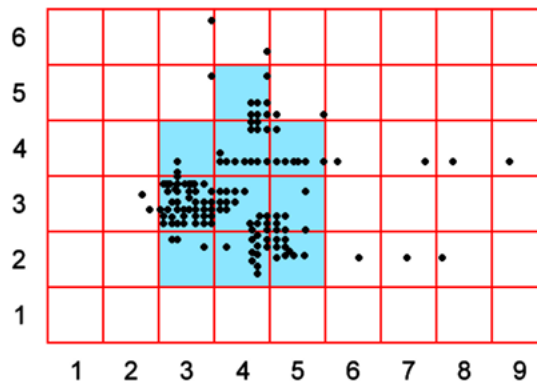
Fonte: Linden, 2009.

2.7.4 Método baseado em grade (CLIFE)

As abordagens de agrupamento baseadas em grade definem agrupamentos baseados em estruturas de célula em grade e na densidade. O “CLIFE” inicialmente particiona a base de dados em subespaços, sendo a densidade baseada no número de objetos em cada célula e cada item passa a ser representado pela célula que ocupa na grade (TSAI; YEN, 2008).

Dentre as vantagens do uso do algoritmo, a principal se dá pela melhora no desempenho, com a redução no tempo de agrupamento devido ao número de interações, visto que são baseadas no número de células de grade. Embora o CLIQUE tenha um tempo de agrupamento rápido, seus limites de agrupamento são horizontais ou verticais, devido à natureza da grade retangular (TSAI; YEN, 2008).

Figura 8 - As células em azul são chamadas de populares



Fonte: Tsai e Yen (2008).

2.7.5 Método baseado em densidade (DBSCAN)

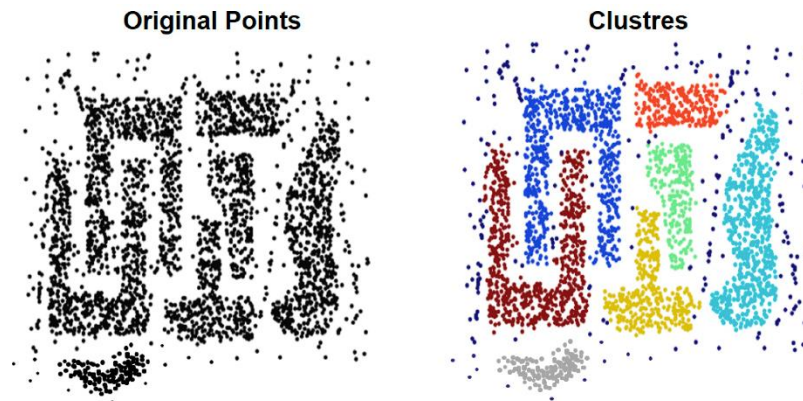
Métodos baseados em densidade, como o DBSCAN, assumem que os elementos pertencentes a um determinado grupo seguem uma mesma distribuição, modelando os grupos como regiões densas do conjunto de dados divididas por áreas de regiões esparsas. Esses métodos têm a vantagem de identificar grupos de formatos arbitrários, ao contrário dos métodos baseados em particionamento, como o *k-means*, que apresentam melhores resultados com grupos de formato circular. Enquanto o *k-means* gera grupos esféricos, o DBSCAN é capaz de gerar grupos com formas mais complexas (ERMAN; ARLITT; MAHANTI, 2006).

O algoritmo possui dois parâmetros de entrada: *épsilon* e *minpts*. *Épsilon* representa a semelhança entre duas conexões de rede e *minpts* é o número mínimo de pontos dentro do grupo. Sendo assim, o *épsilon* juntamente com o *minpts* definem o número de pontos semelhantes ao centroide analisado (THANG; KIM, 2011).

Diferentemente dos demais modelos, o DBSCAN gera o que chamamos de ruídos, ou seja, pontos que não pertencem a qualquer grupo (ERMAN; ARLITT; MAHANTI, 2006, THANG; KIM, 2011).

A Figura 9 apresenta pontos aleatórios na primeira imagem e a aplicação do algoritmo DBSCAN na segunda imagem, onde os pontos azul-escuros são os ruídos e não pertencem a nenhum grupo.

Figura 9 - Resultados do uso do agrupamento DBSCAN



Fonte: Xue et al. (2018)

As definições utilizadas no DBSCAN são:

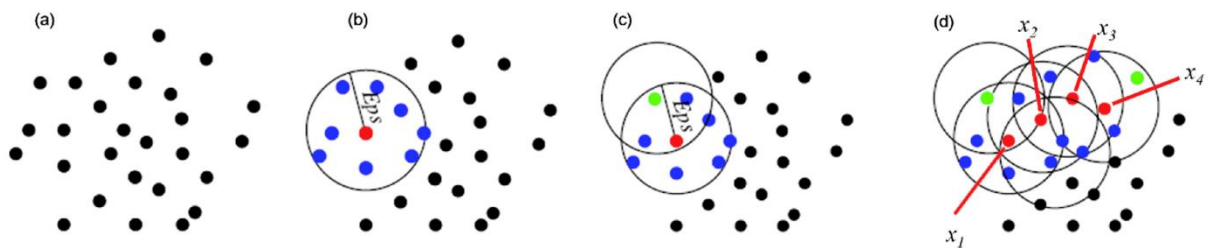
1. $N_{eps}(p)$: é definido como o agrupamento de pontos q que se encontram dentro de uma distância eps em relação ao ponto p . É frequentemente referido como o agrupamento de vizinhança de p .
2. Diretamente alcançável por densidade (DDR): um ponto p é considerado DDR a partir de um ponto q se $p \in N_{eps}(q)$ e $|N_{eps}(q)| \geq minPoints$.
3. Alcançável por densidade (DR): um ponto p é considerado alcançável por densidade a partir de um ponto q , se houver uma sequência de pontos $\{p_1, \dots, p_n\}$ onde $p_1 = p$ e $p_n = q$, tal que p_{i+1} é DDR a partir de p_i .
4. Conectado por densidade (DC): Um ponto p está conectado por densidade a um ponto q se existe um ponto o tal que p e q são DR a partir de o .
5. Ponto central (*Core point*): um ponto p é categorizado como ponto central se o número de pontos em sua vizinhança $|N_{eps}(p)| \geq minPoints$.
6. Ponto de borda (*Border point*): um ponto p é categorizado como ponto de borda se o número de pontos em sua vizinhança $|N_{eps}(p)| < minPoints$ e p for DDR a partir de um ponto central.

7. Ruído (*Noise*): Um ponto p é categorizado como ruído se o número de pontos em sua vizinhança $|N_{eps}(o)| < minPoints$ e p não for DDR a partir de nenhum ponto central.

A figura apresenta quatro imagens que apresentam como é realizada a construção do grupo:

- Pontos a serem analisados;
- O ponto vermelho é o ponto central;
- O ponto verde é o ponto de borda;
- Todos os pontos vermelhos são DR.

Figura 10 - Etapas para a definição do grupo



Fonte: Xue et al. (2018)

De acordo com (SZENASI; CSIBA, 2014; XUE et al., 2018), para a construção do algoritmo, é necessário seguir alguns passos:

Primeiro, para cada ponto p não visitado, o algoritmo encontra o conjunto de seus vizinhos $N_{eps}(p)$ dentro de uma determinada distância. Se esse conjunto de vizinhos for maior que um valor mínimo predefinido ($minPoints$), um novo grupo é criado e o ponto atual, juntamente com seus vizinhos, são atribuídos a esse grupo. Em seguida, os pontos não visitados do novo grupo são expandidos em um processo semelhante. Este processo continua até que todos os elementos tenham sido visitados.

Na expansão de um grupo, a expansão começa a partir de um ponto p , encontrando todos os elementos que estão conectados por densidade (DC) a p . A função recebe como entrada o ponto p , seu conjunto de vizinhos ($NeighborPts$), o identificador do grupo a ser expandido (C), e os parâmetros eps e $minPoints$. Para cada ponto p' no conjunto de vizinhos de p , se p' ainda não foi visitado, sua vizinhança é recuperada e adicionada ao conjunto de vizinhos ($NeighborPts$) que serão

verificados. Depois de verificar p' , se ele não pertencer a nenhum grupo, ele é adicionado ao grupo atual. Este processo continua até que não haja mais pontos para verificar no conjunto *NeighborPts*.

As análises de agrupamentos surgem como uma ferramenta eficaz na identificação de pontos críticos em dados relacionados a sinistros. Ao detectar padrões de agrupamento em conjuntos de dados, essas análises podem destacar áreas geográficas ou condições específicas que apresentam uma concentração significativa de ocorrências de sinistros.

Porém, além de identificar os pontos críticos de sinistros, é essencial compreender os fatores associados. Para isso, modelos de regressão podem ser ferramentas valiosas. Enquanto as análises de agrupamentos revelam áreas de maior concentração de sinistros, modelos de regressão permitem explorar os fatores que influenciam a probabilidade ou gravidade desses eventos.

2.8 LINGUAGEM DE PROGRAMAÇÃO

Com o objetivo de utilizar uma ferramenta estatística robusta, foram consideradas diversas opções de implementação, e concluiu-se que a melhor escolha seria uma linguagem de programação de código aberto. Essa linguagem deveria demandar menos esforço de programação, mas ainda assim possibilitar o desenvolvimento de um produto de alta qualidade. Diante disso, Python foi selecionada como a linguagem de programação ideal.

Python é uma linguagem de programação de alto nível que se destaca pela vasta quantidade de bibliotecas disponíveis, como as que oferecem suporte para matemática simbólica e integração com sistemas e criação de gráficos interativos. Essas bibliotecas tornam o ambiente de desenvolvimento extremamente versátil, permitindo a construção de modelos estatísticos avançados que podem resolver problemas complexos com eficiência (MCKINNEY, 2017).

Desde seu lançamento em 1991, Python evoluiu para se tornar uma das linguagens de programação mais populares. Sua natureza de código aberto permitiu uma ampla colaboração da comunidade de computação científica, o que contribuiu para seu desenvolvimento contínuo e crescente adoção em aplicações industriais e acadêmicas (MCKINNEY, 2017).

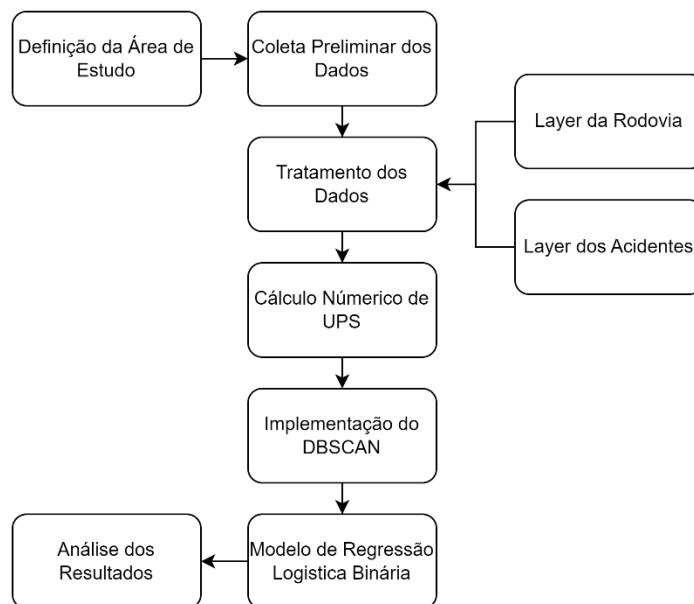
Python é uma linguagem poderosa para análise de dados e modelos estatísticos, especialmente em regressão logística binomial, utilizando bibliotecas como *pandas*, *numpy*, *statsmodels* e *scikit-learn*. *Pandas* e *numpy* são essenciais para o tratamento e manipulação de dados, fornecendo estruturas eficientes para organizar e realizar cálculos numéricos (MCKINNEY, 2017). *Statsmodels* é utilizada para realizar análises estatísticas detalhadas e ajustar modelos de regressão logística, oferecendo métricas de diagnóstico e testes de hipótese. Por fim, *scikit-learn* é empregada para aprendizado de máquina, facilitando a implementação prática de regressão logística, previsão e avaliação de desempenho (Scikit-learn, 2024). Essas bibliotecas juntas fornecem uma plataforma robusta para desenvolver, avaliar e otimizar modelos de regressão logística binomial.

3 MATERIAIS E MÉTODOS

Neste capítulo, são apresentados os materiais e métodos utilizados na pesquisa. Inicialmente, foram obtidos dados de sinistros rodoviários, que passaram por um tratamento extensivo. Utilizou-se o algoritmo DBSCAN, conforme apresentado no item 2.8.5, com o objetivo de identificar grupos de sinistros de trânsito com vítimas moderadas, graves ou fatais.

Após a identificação dos grupos, estes foram utilizados como uma variável explicativa no modelo logístico binário desenvolvido para avaliar os fatores que aumentam a gravidade dos sinistros de trânsito. O modelo logístico considerou, além dos grupos identificados, variáveis relacionadas à característica do condutor, do sinistro, da via, do veículo, ambientais e temporais, como apresentado no item 2.5.3. O fluxograma metodológico é apresentado na Figura 11 e as etapas são descritas a seguir:

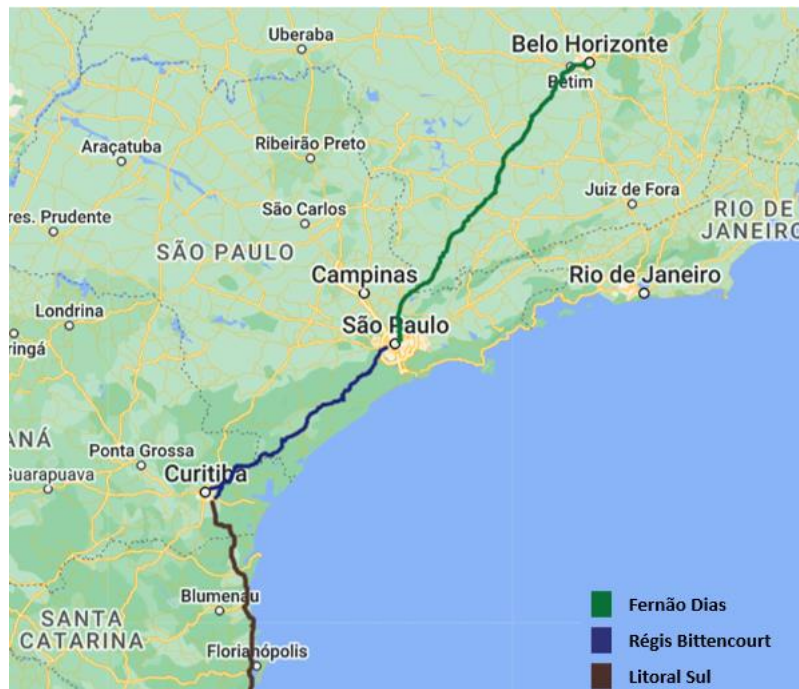
Figura 11 - Fluxograma metodológico



3.1 DEFINIÇÃO DA ÁREA DE ESTUDO

O escopo deste estudo abrangeu as rodovias federais concedidas em 2008, atualmente sob a administração da Arteris. Especificamente, foram analisadas três concessionárias do grupo: Fernão Dias, Régis Bittencourt e Litoral Sul, conforme representado na Figura 12. Essas rodovias totalizam mais de mil quilômetros de extensão e foram objeto de estudo detalhado no contexto da ocorrência de sinistros rodoviários e suas características associadas.

Figura 12 - Mapa das rodovias analisadas



Fonte: Autor, 2024

Todos os quatro estados estão no *ranking* dos seis Estados brasileiros com maior frota de veículos do Brasil, sendo São Paulo, Minas Gerais e Paraná os primeiros da lista, respectivamente (SECRETARIA DE TRÂNSITO NACIONAL, 2022). O tráfego de veículos está relacionado ao eixo comercial existente na região, que também abrange os estados com maior PIB nacional (IBGE, 2020).

As três concessionárias apresentam três rodovias e pelo menos uma Serra no trecho, aumentando o risco de sinistros no segmento devido às curvas e inclinação acentuada.

3.1.1 Litoral Sul

A Concessionária Litoral Sul abrange uma extensão de 357 quilômetros, conectando as cidades de São José dos Pinhais, na região metropolitana de Curitiba, e Palhoça, na região metropolitana de Florianópolis.

Esta rodovia é reconhecida pelo seu intenso fluxo de veículos, sendo a de maior tráfego entre as três estudadas. Isso se deve principalmente à sua função vital na ligação entre a capital do estado do Paraná e importantes cidades de Santa Catarina, como Joinville, Itajaí e Florianópolis.

Por se situar ao longo do litoral e atravessar áreas urbanas significativas, a Litoral Sul registra um elevado volume de veículos leves, especialmente devido às viagens pendulares e de lazer. O tráfego de veículos leves pode representar até 70% do tráfego total nessa rodovia, consolidando-a como a concessionária com o maior fluxo absoluto de veículos entre as três analisadas.

3.1.2 Régis Bittencourt

A Rodovia Régis Bittencourt, com seus 315 quilômetros de extensão, atravessa 17 municípios desde Taboão da Serra, na região metropolitana de São Paulo, até Curitiba, no estado do Paraná.

Destaca-se pela presença de três serras ao longo do percurso, o que torna o traçado da via sinuoso e propenso a sinistros. Com aproximadamente 80% do tráfego composto por veículos pesados, a rodovia desempenha um papel crucial na conexão entre grandes centros urbanos, como São Paulo e Curitiba, e os portos marítimos de Santos e Paranaguá.

Em termos de segurança viária, um dos principais desafios enfrentados pela concessionária é a alta incidência de tombamentos de caminhões. Esse problema é atribuído tanto à quantidade de curvas acentuadas nas serras quanto ao intenso fluxo de veículos pesados.

3.1.3 Fernão Dias

A Rodovia Fernão Dias, com seus 562 quilômetros de extensão, percorre um trajeto que abrange 33 municípios, desde a cidade de Belo Horizonte, em Minas Gerais, até Guarulhos, na região metropolitana de São Paulo. Trata-se de um corredor comercial, que conecta duas importantes metrópoles do país, São Paulo e Belo Horizonte. Cerca de 60% do tráfego nessa rodovia é composto por veículos pesados.

3.2 TRATAMENTO PRELIMINAR DOS DADOS

Os dados referentes aos sinistros foram fornecidos pela Arteris, empresa encarregada da concessão dos trechos rodoviários em questão. O início das operações das três concessionárias ocorreu em 2008, o que implica que os dados desse ano não estão completos. Portanto, para fins de análise, o ano de 2008 foi excluído do conjunto de dados.

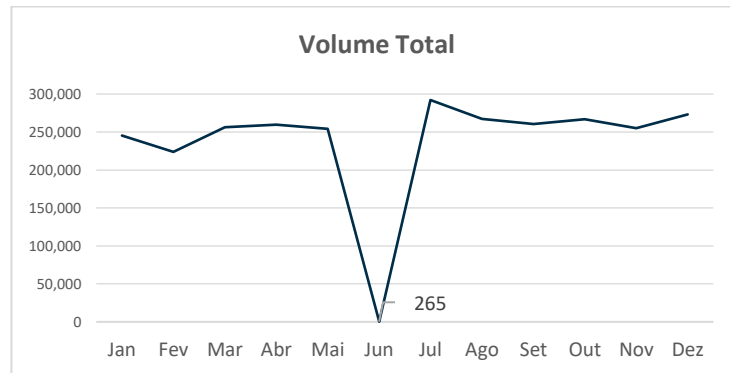
3.2.1 Dados Incluídos

3.2.1.1 Volume Mensal Médio

Além dos dados de sinistros, foram incluídos os dados de tráfego de veículos no trecho, com base nos dados abertos da Agência Nacional de Transportes Terrestres (ANTT). Como os dados da ANTT não incluem o ano de 2009, esse ano também foi desconsiderado para o estudo. Portanto, o período de análise compreendeu os anos de 2010 a 2022.

Inicialmente, foi necessário tratar os meses que apresentaram dados zerados ou com números significativamente inferiores aos meses anteriores, conforme ilustrado na Figura 13. Para isso, foi realizada uma regressão linear nos pontos, visando corrigir essas inconsistências nos dados.

Figura 13 – Tráfego da Fernão Dias na praça de pedágio 6, sentido norte em 2022



Fonte: Autora (2024)

Para calcular o volume de tráfego no ponto específico onde ocorreu cada sinistro, foi adotada uma abordagem que considera a localização das praças de pedágio como pontos de referência. A área de influência de cada praça de pedágio foi definida como a metade do trecho entre a própria praça e a próxima, conforme apresentado na Figura 14, onde cada cor corresponde à abrangência de uma praça de pedágio, assumindo que o tráfego se distribui de maneira uniforme ao longo desse intervalo.

Figura 14 - Influência de cada praça de pedágio



Fonte: Autora (2024)

Na figura, a praça 3 está localizada no quilômetro 79,3 da BR101/SC com sua área de abrangência estendendo-se do quilômetro 40,3 ao 118,3 da mesma rodovia. Essa delimitação se justifica pelo fato de que o quilômetro 40,3 está localizado entre o quilômetro 1,35 (Praça 2) e o quilômetro 79,3 (Praça 3), enquanto o quilômetro 118,3 se situa entre o quilômetro 79,3 (Praça 3) e o quilômetro 157,4 (Praça 4).

Quanto às praças de pedágio localizadas no início ou no final do trecho, como a Praça 1 na Figura 14, a área de influência foi considerada até o início do trecho concedido, uma vez que não há outra praça subsequente. Essa abordagem visa capturar o volume de tráfego em cada ponto da rodovia, levando em conta a importância das praças de pedágio como pontos de controle e cobrança.

Para integrar os dados das praças de pedágio com os registros de sinistros, foram considerados o sentido da via, o mês e o ano em que cada sinistro ocorreu. Essa abordagem permitiu relacionar os dados específicos de cada sinistro com as informações correspondentes das praças de pedágio mais próximas, levando em conta o fluxo de veículos em cada direção da rodovia e em cada período de tempo.

3.2.1.2 *Informações geográficas*

Para complementar a análise dos sinistros, foram incluídas informações sobre a localização geográfica dos eventos, distinguindo entre regiões urbanas e rurais. Esse processo foi realizado por meio de análises de imagens de satélite, que permitiram identificar as características do ambiente onde ocorreram os sinistros.

No caso das regiões urbanas, foi realizada uma verificação adicional para determinar se a área em questão correspondia à uma cidade com mais de 200 mil habitantes. Isso é importante devido às particularidades do tráfego e da infraestrutura viária em áreas urbanas mais densamente povoadas, que podem influenciar os padrões de sinistros.

3.2.1.3 *Iluminação*

Adicionalmente, foi realizada uma análise específica para os sinistros que ocorreram durante a noite, a fim de determinar se a região estava iluminada. Essa

informação foi obtida por meio dos dados de iluminação noturna disponibilizados pela ANTT, permitindo uma avaliação mais precisa do ambiente de condução durante a noite e sua relação com a ocorrência de sinistros.

3.2.1.4 *Presença de vias marginais*

Também foram incluídas informações sobre a presença de vias marginais nos trechos onde ocorreram os sinistros. Esses dados foram obtidos a partir das informações disponibilizadas pela ANTT, fornecendo (dados, informações, reflexões) *insights* adicionais sobre a configuração das rodovias e sua influência na segurança viária. Essa análise detalhada das características do ambiente rodoviário contribuiu para uma compreensão mais abrangente dos padrões de sinistros e pode orientar a implementação de medidas preventivas e de segurança.

3.2.1.5 *Análise temporal*

Para avaliar o impacto de diversas medidas de segurança implementadas antes de 2015, como a obrigatoriedade do uso de cadeirinhas para crianças, a criação da Lei Seca, a implementação do PNCV, e a exigência da instalação de freios ABS e *airbags*, foi criada uma nova variável dicotômica 'Anterior a 2015'. Esta variável classifica os sinistros em duas categorias: aqueles que ocorreram antes ou durante o ano de 2014, codificados como 0 (falsa), e aqueles que ocorreram após 2014, codificados como 1 (verdadeira)

A escolha do ano de 2014 como ponto de corte foi deliberada, uma vez que todas as medidas de segurança mencionadas foram implementadas antes ou durante desse ano. Dessa forma, a variável dicotômica permite uma análise comparativa entre os períodos pré e pós-implementação das medidas, possibilitando uma avaliação do impacto dessas políticas públicas na redução de sinistros de trânsito.

3.2.1.6 *Unidade Padrão de Severidade*

Essa variável foi incluída com base no cálculo apresentado anteriormente. No entanto, como este trabalho foca apenas em sinistros com vítimas, o peso 1 não foi

considerado, assim como o peso 6, que se aplica a sinistros envolvendo pedestres. Dessa forma, foram incluídos apenas os pesos 4 e 13, correspondentes à vítima mais grave no sinistro, para refletir com precisão a severidade dos sinistros analisados.

Essa variável será utilizada como critério no modelo de agrupamento para definir o tamanho mínimo dos agrupamentos. Além disso, a soma do UPS dentro de cada cluster determinará o peso do agrupamento, que será utilizado como parâmetro, juntamente com a geolocalização.

3.2.2 Filtros Utilizados

Entendendo a necessidade de abordagens distintas para sinistros envolvendo pedestres e ciclistas, devido à sua maior vulnerabilidade, esses casos foram excluídos da amostra de análise. Da mesma forma, sinistros ocorridos em vias marginais foram removidos da análise, uma vez que essas vias frequentemente apresentam características físicas e padrões de tráfego diferentes das vias principais.

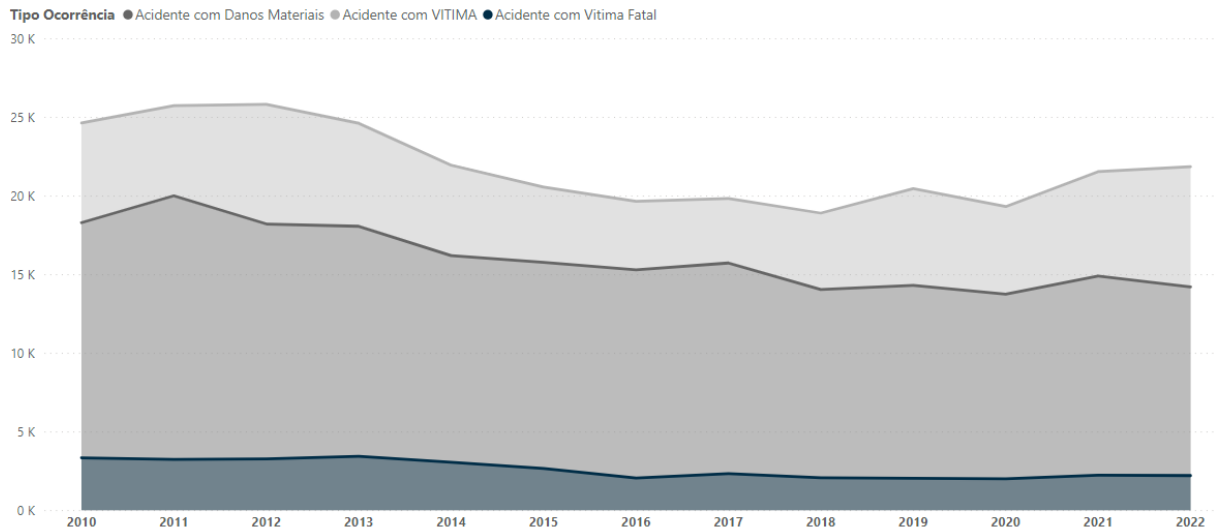
3.2.2.1 Sinistros com vítimas

Após essa exclusão, dos 338.843 sinistros iniciais, restaram 282.545 sinistros e 542.891 vítimas para a análise preliminar. Esse refinamento da amostra permitirá uma investigação mais específica e direcionada sobre os padrões e fatores associados aos sinistros nas rodovias estudadas, contribuindo para uma compreensão mais aprofundada dos fenômenos envolvidos.

No período entre 2010 e 2022, 68% das ocorrências foram sinistros com danos materiais, enquanto 30% envolveram vítimas e 2% foram classificados como sinistros com vítimas fatais. No entanto, ao considerarmos o valor de UPS, conforme descrito na seção 2.6, a distribuição desses sinistros se altera significativamente. Nesse contexto, os sinistros com vítimas passam a representar 58% do peso total, enquanto os sinistros com vítimas fatais correspondem a 10% do peso total.

A Figura 15 oferece uma representação visual dessa distribuição ao longo dos anos, destacando a proporção de sinistros com diferentes níveis de gravidade com base no valor de UPS.

Figura 15 – Soma de UPS por ano por tipo de sinistro



Fonte: Autora (2024)

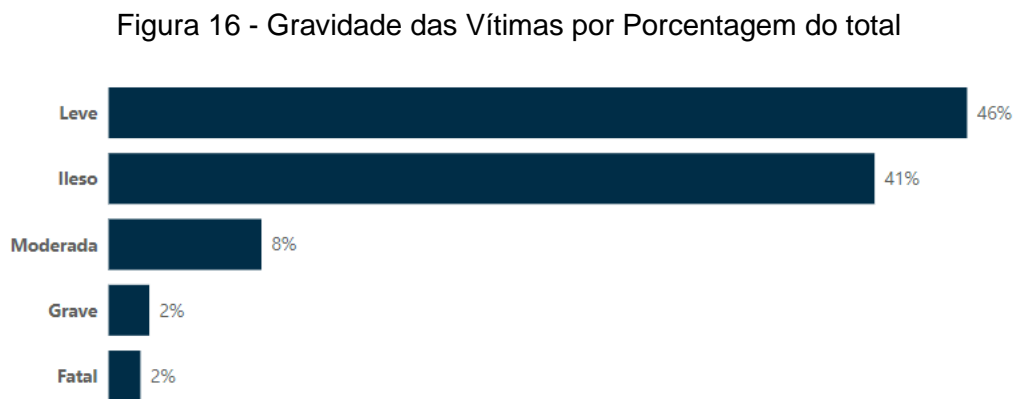
Ao considerar o objetivo principal deste estudo, que é a redução de vítimas fatais e graves em sinistros rodoviários, percebeu-se a necessidade de concentrar a análise nos casos que mais impactam diretamente na segurança viária. Portanto, decidiu-se focar exclusivamente nos sinistros que envolvem vítimas.

Essa decisão foi tomada levando em consideração que os sinistros sem vítimas representam uma parcela significativa dos dados totais, e sua inclusão poderia distorcer a análise e dificultar a identificação de padrões e tendências específicas relacionadas à ocorrência de lesões graves e fatais. Dessa forma, ao se concentrar nos sinistros com vítimas, é possível direcionar melhor os esforços para medidas preventivas e intervenções que visam mitigar os impactos mais severos dos sinistros de trânsito nas rodovias estudadas.

É importante ressaltar que um único sinistro pode envolver múltiplas vítimas, e, portanto, sinistros com vítimas podem abranger todas as classificações possíveis. Conforme observado por Kwon, Rhee e Yoon (2015), as vítimas são categorizadas em diferentes níveis de gravidade, incluindo vítima ileso, vítima leve, vítima moderada, vítima grave e vítima fatal. Assim, mesmo em um sinistro com vítima fatal, é possível que haja também vítimas classificadas como ilesas.

3.2.2.2 Sinistros com vítimas moderadas, graves ou fatais

Assim como observado nas ocorrências de sinistros com danos materiais, há uma discrepância significativa na contagem total quando comparada com outros tipos de sinistros. Da mesma forma, ao analisar a gravidade das vítimas, é notável a diferença no número de vítimas leves e ilesas em comparação com as demais vítimas, como ilustrado na Figura 16.



Fonte: Autora (2024)

Devido à discrepância observada na contagem de vítimas leves e ilesas em comparação com as demais vítimas em nossos dados de análise de sinistros, optamos por focar nossa atenção nas vítimas com gravidade moderada em diante. Essa decisão foi tomada para garantir uma análise mais precisa e significativa das diferentes categorias de gravidade das vítimas.

Assim, para simplificar nossa abordagem e concentrar-nos nas categorias mais relevantes de gravidade, agrupamos as vítimas em duas categorias distintas: moderadas e graves ou fatais.

É importante ressaltar que, após aplicarmos esse filtro, o número inicial de 200 mil vítimas foi reduzido para 21 mil, compreendendo apenas as vítimas moderadas, graves ou fatais. Em termos numéricos, a Fernão Dias representa 39% do total de vítimas analisadas, seguida pela Litoral Sul com 36% e pela Régis Bittencourt com 25%.

Dessa forma, a Tabela 9 apresenta o número de sinistros de acordo com os filtros aplicados.

Tabela 9 - Número de registros com base nos filtros aplicados

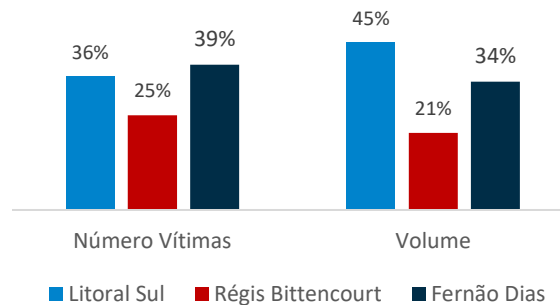
Tipo de Filtro Aplicado	Número de Sinistros	Número Vítimas
Todos Os Sinistros de 2009- 2022	331.403	696.057
Sinistros de 2010-2022	309.027	645.770
Excluindo Sinistros com Pedestres e Ciclistas	307.517	642.384
Excluindo Sinistros em Marginais	287.512	604.571
Excluindo Sinistros com Danos Materiais	78.402	196.227
Apenas Vítimas Moderadas, Graves ou Fatais	17.093	21.360

Fonte: Autora (2024)

3.3 ANÁLISE DESCRITIVA

A distribuição dos dados de número de vítimas e tráfego viário é apresentada na Figura 17. Observa-se que a Litoral Sul tem o maior volume de tráfego, representando 45% do total. Em contrapartida, a Fernão Dias é a concessionária com o maior número de vítimas, correspondendo a 39% dos casos.

Figura 17 - Porcentagem de vítimas por Concessionária



Fonte: Autora (2024)

Em relação às demais variáveis analisadas:

- Horário do sinistro: 54% dos sinistros com vítimas ocorrem durante o dia.
- Tipo de veículo envolvido:
 - 30% envolvem motocicletas.
 - 59% envolvem veículos leves.
 - 39% envolvem veículos pesados.

- Menos de 1% envolvem veículos de passageiros.
- Localização do sinistro: 18% dos sinistros ocorrem em regiões que possuem vias marginais.
- Perfil das vítimas:
 - 32% das vítimas são graves ou fatais;
 - 65% das vítimas são condutores;
 - 75% das vítimas são homens.
- Condições da pista: 22% dos sinistros ocorrem em pista molhada.

3.3.1 Modelo DBSCAN

O modelo DBSCAN foi escolhido devido à natureza dos dados, que envolvem sinistros distribuídos ao longo da malha viária, sofrendo influência de fatores como curvatura, altitude e outras infraestruturas da via. A análise por densidade é a abordagem mais adequada para representar o agrupamento dos grupos nesse contexto.

A implementação do modelo foi realizada em *Python*, onde o valor mínimo de pontos (minpts) foi definido como cinco. Para determinar a densidade (épsilon), optou-se por um valor de 100 metros. Um critério adotado foi estabelecer um valor mínimo de UPS para cada grupo, a fim de evitar a formação de grupos com extensões pequenas. Portanto, grupos com menos de 20 UPS foram descartados do modelo.

Considerando as particularidades de cada rodovia e concessionária, foram gerados um total de 18 modelos. Isso ocorre porque há três concessionárias, cada uma operando em três rodovias distintas, com dois sentidos de tráfego para cada uma delas. Essa abordagem permite uma análise mais detalhada e específica de cada contexto rodoviário.

3.3.2 Dados do modelo

Assim, foram identificados 810 grupos, cobrindo uma extensão de 93 quilômetros de via, o que representa aproximadamente 4% da extensão total da rodovia e abrangem cerca de 29% do total de vítimas fatais e graves registradas. Vale destacar que alguns grupos apresentaram uma extensão total de 0 metros, pois

apenas sinistros ocorridos exatamente no mesmo ponto foram considerados nesses grupos. Um exemplo é o grupo 5 da Tabela 10, que abrange o trecho do km 525.5.

Dentre os 15 grupos com o maior número de UPS, oito estão localizados na Fernão Dias, quatro na Régis Bittencourt e três na Litoral Sul, conforme detalhado na Tabela 10. Essa distribuição destaca áreas críticas ao longo dessas rodovias, onde uma concentração significativa de sinistros graves e fatais foi observada.

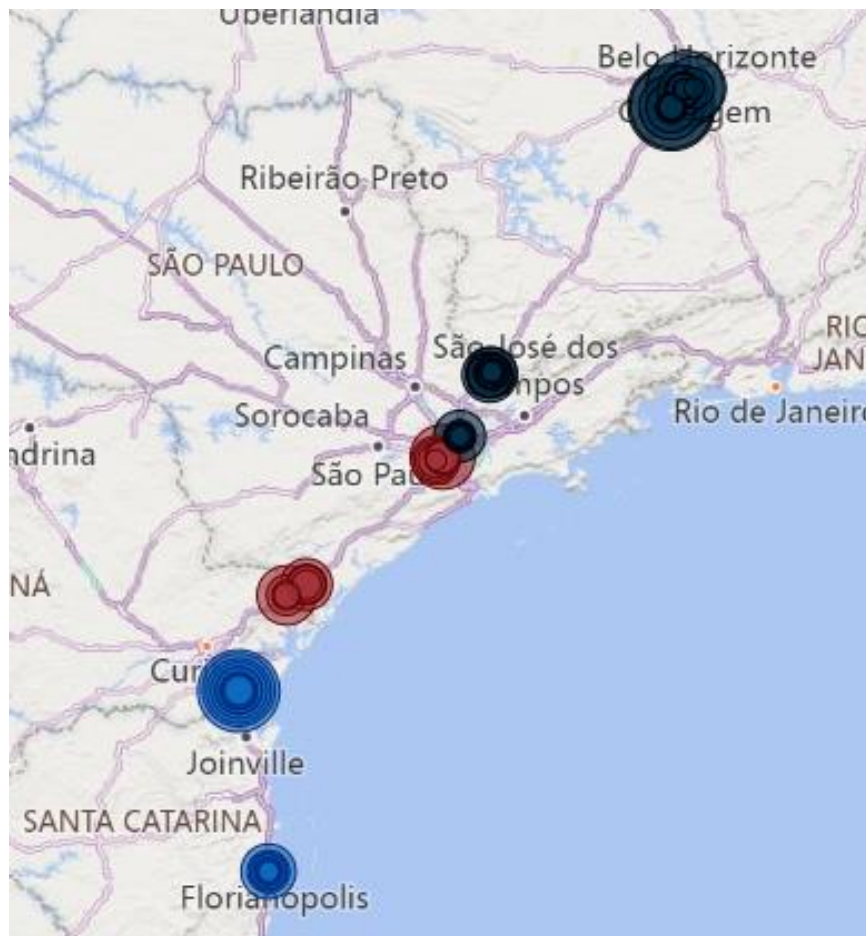
Tabela 10 - Ranking com os 15 grupo com maior número de UPS

Ranking	Concessionaria	Rodovia/ Sentido	Km Inicial	Km Final	Extensão (metros)	Número Acidentes	Soma UPS
1	Litoral Sul	BR376/ Sul	667.90	668.80	900	70	433
2	Fernão Dias	MG-BR381/ Sul	524.90	525.00	100	33	258
3	Litoral Sul	BR376/ Sul	667.50	667.60	100	35	230
4	Régis Bittencourt	SP-BR116/ Norte	511.50	511.90	400	34	226
5	Fernão Dias	MG-BR381/ Sul	525.50	525.50	0	28	202
5	Fernão Dias	MG-BR381/ Sul	525.70	526.20	500	28	202
7	Litoral Sul	BR101/ Sul	205.90	206.70	800	39	201
8	Fernão Dias	MG-BR381/ Norte	493.90	494.00	100	31	151
9	Régis Bittencourt	SP-BR116/ Norte	544.90	545.20	300	26	149
10	Fernão Dias	MG-BR381/ Norte	934.00	934.40	400	30	138
11	Fernão Dias	MG-BR381/ Sul	929.50	930.00	500	26	131
12	Régis Bittencourt	SP-BR116/ Norte	269.80	270.00	200	27	126
13	Fernão Dias	MG-BR381/ Sul	482.80	483.00	200	20	125
14	Régis Bittencourt	SP-BR116/ Sul	278.70	279.00	300	22	124
15	Fernão Dias	SP-BR381/ Sul	78.70	79.10	400	21	120

Fonte: Dados brutos da Concessionária, elaborado pela autora (2024)

Os grupos apresentados na Tabela 10 podem ser observados na Figura 18.

Figura 18 - Mapa com a geolocalização dos 15 grupos com maior UPS



Fonte: Autora (2024)

Destaca-se que, dos três grupos da Litoral Sul, dois deles estão concentrados na descida da Serra Dona Francisca (1, 3), enquanto o último encontra-se na área da Grande Florianópolis (7). Quanto à Régis Bittencourt, o grupo com o maior número de ocorrências UPS está localizado na Serra do Azeite (4), seguido pelo segundo na Serra da Bica da Onça (9), além de dois grupos situados na região da Grande São Paulo (12, 14).

Já na Fernão Dias, os três primeiros estão localizados na serra Igarapé (2, 5, 6), três na região da Grande Belo Horizonte (8, 13, 15) e dois na região de Extrema/SP (10, 11). Esses dados ressaltam áreas específicas ao longo dessas rodovias que apresentam uma concentração significativa de sinistros graves, possibilitando uma análise mais precisa e direcionada para a implementação de medidas preventivas e de segurança viária.

3.3.2.1 Litoral Sul

A Figura 19 ilustra os 15 grupos identificados na Litoral Sul com os maiores valores de UPS. É importante ressaltar que 13 grupos estão concentrados na região da Serra Dona Francisca e da Grande Florianópolis. Em conjunto, esses 15 grupos abrangem aproximadamente 0.6% da extensão total da malha viária da região, porém representam cerca de 16% do total de vítimas fatais registradas pela concessionária.

Figura 19 - 15 Grupos com maior número de UPS da Litoral Sul



Ranking	Rodovia/ Sentido	Km Inicial	Km Final	Extensão (metros)	Número Acidentes	Soma UPS
1	BR376/ Sul	667.90	668.80	900	70	433
2	BR376/ Sul	667.50	667.60	100	35	230
3	BR101/ Sul	205.90	206.70	800	39	201
4	BR376/ Sul	673.90	674.00	100	16	118
5	BR376/ Sul	665.90	666.40	500	19	112
5	BR376/ Sul	666.70	667.10	400	19	112
7	BR376/ Norte	615.00	615.50	500	21	111
8	BR101/ Norte	202.90	203.00	100	16	100
9	BR101/ Sul	205.00	205.00	0	20	98
10	BR376/ Sul	665.50	665.60	100	12	84
11	BR101/ Sul	207.10	207.40	300	16	82
12	BR376/ Sul	667.70	667.80	100	11	80
13	BR101/ Norte	118.00	118.00	0	15	78
14	BR101/ Norte	211.90	212.00	100	12	75
15	BR101/ Norte	210.90	211.00	100	16	73

Fonte: Dados brutos da Concessionária, elaborado pela autora (2024)

Na BR376/PR, dos oito grupos identificados, sete encontram-se no sentido sul, em curvas acentuadas (grupos 1, 2, 4, 5, 6, 10, 12), enquanto um está localizado mais próximo de Curitiba (grupo 7). Nessa região, a predominância de sinistros é atribuída ao intenso tráfego de veículos pesados, que, especialmente em áreas montanhosas, enfrentam maior probabilidade de ocorrência de sinistros.

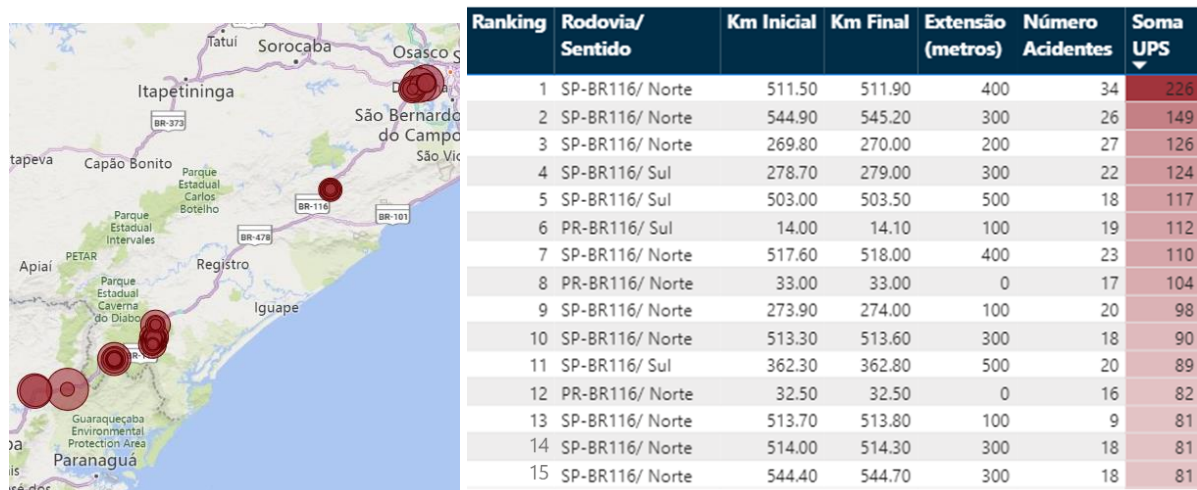
Dos sete grupos identificados na BR101/SC, seis estão situados na região da Grande Florianópolis (grupos 3, 8, 9, 11, 14, 15), enquanto um está localizado na área de Itajaí (grupo 13). Essas regiões são conhecidas por apresentarem um tráfego intenso de veículos, tanto devido às viagens diárias pendulares e ao transporte de cargas durante a semana, quanto às viagens recreativas nos fins de semana.

3.3.2.2 Régis Bittencourt

A Figura 20 oferece uma visualização da distribuição dos 15 grupos com o maior número de UPS ao longo da Régis Bittencourt. É importante ressaltar que, mesmo representando apenas 4% da malha viária total, esses grupos correspondem à uma parcela significativa das fatalidades, totalizando 17%.

Ao examinar a distribuição geográfica desses grupos, observamos que nove deles estão concentrados em áreas serranas, sendo seis situados na Serra do Azeite (grupos 1, 5, 7, 10, 13, 14), dois na Serra da Bica da Onça (grupos 2 e 13) e um na região da Serra do Cafezal (grupo 11). Além disso, três estão localizados na região da Grande São Paulo (grupos 3, 4, 9) e três na área de Campina Grande do Sul (grupos 6, 8 e 12).

Figura 20 - 15 Grupos com maior número de UPS da Régis Bittencourt



Fonte: Dados brutos da Concessionária, elaborado pela autora (2024)

É relevante ressaltar que nas áreas situadas em serras, as condições da via são desafiadoras devido ao número de curvas e declives acentuados. Esses trechos, combinados ao tráfego composto principalmente por veículos pesados, criam um ambiente propenso à ocorrência de sinistros. Além disso, os grupos em Campina Grande do Sul estão posicionados em uma região com curvas acentuadas e leve acive, devido à proximidade da Serra do Mar.

No que se refere aos pontos próximos à Grande São Paulo, é importante destacar que, além do intenso fluxo de veículos característico, o grupo 3 apresenta

uma particularidade adicional. Localizado no fim do trecho de concessão, torna-se mais desafiador implementar medidas de controle de fluxo e fiscalização

3.3.2.3 Fernão Dias

A Figura 21 oferece uma visualização dos 15 grupos com maior número de UPS, abrangendo 0,3% da malha viária e representando 6% das fatalidades e vítimas graves da Concessionária.

Quatro grupos estão localizados na região da Grande Belo Horizonte (grupos 4, 7, 8 e 11), enquanto três estão situados na região da Grande São Paulo (grupos 9, 14 e 15). Além disso, observa-se uma presença marcante de grupos em regiões caracterizadas por condições viárias desafiadoras, incluindo curvas acentuadas e declives, como a Serra de Igarapé (grupos 1, 2, 3 e 10), na região de Extrema/MG (grupos 5 e 6) e Itatiaiuçu/MG (grupo 12).

Figura 21 - 15 Grupos com maior número de UPS da Fernão Dias



Ranking	Rodovia/Sentido	Km Inicial	Km Final	Extensão (metros)	Número Acidentes	Soma UPS
1	MG-BR381/ Sul	524.90	525.00	100	33	258
2	MG-BR381/ Sul	525.50	525.50	0	28	202
2	MG-BR381/ Sul	525.70	526.20	500	28	202
4	MG-BR381/ Norte	493.90	494.00	100	31	151
5	MG-BR381/ Norte	934.00	934.40	400	30	138
6	MG-BR381/ Sul	929.50	930.00	500	26	131
7	MG-BR381/ Sul	482.80	483.00	200	20	125
8	SP-BR381/ Sul	78.70	79.10	400	21	120
9	MG-BR381/ Sul	496.50	497.00	500	21	120
10	MG-BR381/ Norte	531.90	532.00	100	17	113
11	MG-BR381/ Sul	493.90	494.00	100	19	112
12	MG-BR381/ Norte	540.00	540.30	300	18	108
13	MG-BR381/ Norte	688.00	688.10	100	11	98
14	SP-BR381/ Norte	74.90	75.00	100	24	96
15	SP-BR381/ Norte	67.80	68.00	200	17	95

Fonte: Dados brutos da Concessionária, elaborado pela autora (2024)

A análise da rodovia Fernão Dias revela padrões semelhantes às demais analisadas, com uma alta densidade de tráfego próximo aos grandes centros urbanos e em regiões serranas. Nessas áreas, o elevado fluxo de veículos pesados contribui significativamente para o aumento do número de sinistros. No entanto, há uma exceção notável no grupo 13, onde a presença de acessos de retorno em desnível pode desempenhar um papel na ocorrência de sinistros, mesmo em áreas que não são tradicionalmente consideradas de alto risco.

3.4 MODELO LOGIT

O tipo de gravidade do sinistro rodoviário como uma variável dependente binária no modelo, foi recodificado em dois níveis: 0 para vítimas moderadas e 1 para vítimas graves ou fatais. No modelo proposto, o logit foi o logaritmo natural da probabilidade de a variável resposta Y ser uma vítima grave ou fatal ($Y=1$) versus uma vítima moderada ($Y=0$).

3.4.1 Variáveis Analisadas

Conforme a Tabela 11, cada sinistro possui 21 variáveis, sendo que estação do ano foi incluída com base na data do sinistro, período do dia baseado na hora da ocorrência e “área crítica” refere-se ao modelo DBSCAN, que determina se um sinistro está presente ou não em um agrupamento ou área crítica identificado pelo algoritmo.

Tabela 11 - Variáveis presentes em cada sinistro

Variável	Descrição da Variável	Fatal/Grave		Moderada	
		N	%	N	%
Característica da Vítima					
Sexo	Feminino	1368	6%	3879	18%
	Masculino	5509	26%	10604	50%
Faixa Etária	1-17 anos	404	2%	821	4%
	18-24 anos	1061	5%	2578	12%
	25-34 anos	1687	8%	3982	19%
	35-49 anos	1770	8%	3947	18%
	Maior que 50 anos	1148	5%	2430	11%
	Não Informado	807	4%	725	3%
Posição da Vítima	Condutor	4633	22%	9233	43%
	Passageiro	2244	11%	5250	25%
Característica do Sinistro					
Tipo Sinistro	Colisão Traseira	2107	10%	4070	19%

Variável	Descrição da Variável	Fatal/Grave		Moderada	
		N	%	N	%
	Queda de Moto	590	3%	1769	8%
	Colisão Lateral/Transversal	673	3%	1299	6%
	Capotamento	1016	5%	2569	12%
	Choque	717	3%	1589	7%
	Saida de Pista	306	1%	901	4%
	Tombamento	711	3%	1298	6%
	Colisão Frontal	393	2%	317	1%
	Atropelamento - Animal	35	0%	59	0%
	Outros	329	2%	612	3%
Número de Veículos	Um	3618	17%	8224	39%
	Dois ou mais	3259	15%	6259	29%
Área Crítica	Sim	2419	11%	5691	27%
	Não	4458	21%	8792	41%
Característica da Via					
Condição Pista	Seca	5541	26%	11131	52%
	Molhada	1336	6%	3352	16%
Perfil Pista	Em Nivel	4113	23%	9187	51%
	Declive/Aclive	1709	9%	3181	17%
Traçado Pista	Reta	4207	20%	8696	41%
	Curva Suave	1450	7%	3164	15%
	Curva Acentuada	1220	6%	2623	12%
Tipo Local	Rural	5020	24%	10094	47%
	Urbano	1857	9%	4389	21%
Iluminação Noturna	Presente	581	6%	1381	14%
	Não presente	2829	29%	4957	51%
Marginal	Presente	1162	5%	2762	13%
	Não presente	5715	27%	11721	55%
Tráfego	Variável Continua	-	-	-	-

Característica do Veículo

Variável	Descrição da Variável	Fatal/Grave		Moderada	
		N	%	N	%
Motocicleta	Não Presente	4844	23%	9936	47%
	Presente	2033	10%	4547	21%
Veículo Leve	Não Presente	3058	14%	5564	26%
	Presente	3819	18%	8919	42%
Veículo Pesado	Não Presente	3502	16%	9590	45%
	Presente	3375	16%	4893	23%
Veículo Passageiro	Não Presente	6317	30%	13636	64%
	Presente	560	3%	847	4%
Idade do veículo	0-5 anos	1736	8%	3223	15%
	6-10 anos	1019	5%	2380	11%
	11-20 anos	718	3%	1941	9%
	Maior que 20 anos	305	1%	438	2%
	Não informado	2976	14%	6624	31%
Característica Temporal					
Estação do ano	Verão	1854	9%	4074	19%
	Primavera	1704	8%	3813	18%
	Outono	1646	8%	3362	16%
	Inverno	1673	8%	3234	15%
Período do dia	Dia	3467	16%	8145	38%
	Noite	3410	16%	6338	30%
Anterior a 2015	Sim	3529	17%	7593	36%
	Não	3348	16%	6890	32%

Fonte: Autora (2024)

A equação considerando todas as variáveis é apresentada a seguir:

$$\begin{aligned}
\text{logit}(P_i) = & \beta_0 + \beta_1 * \text{Sexo}_i + (\beta_2 * \text{Faixa_Etaria}_{18-24,i} \\
& + \beta_3 * \text{Faixa_Etaria}_{25-34,i} + \beta_4 * \text{Faixa_Etaria}_{35-49,i} \\
& + \beta_5 * \text{Faixa_Etaria}_{50+,i} + \beta_6 * \text{Faixa_Etaria}_{naoinfo,i}) \\
& + \beta_7 * \text{Posicao_Vitima}_i \\
& + (\beta_8 * \text{Tipo_Acidente}_{colisao_traseira,i} \\
& + \beta_9 * \text{Tipo_Acidente}_{colisao_lateral,i} \\
& + \beta_{10} * \text{Tipo_Acidente}_{capotamento,i} \\
& + \beta_{11} * \text{Tipo_Acidente}_{choque,i} \\
& + \beta_{12} * \text{Tipo_Acidente}_{saida_pista,i} \\
& + \beta_{13} * \text{Tipo_Acidente}_{tombamento,i} \\
& + \beta_{14} * \text{Tipo_Acidente}_{colisao_frontal,i} \\
& + \beta_{15} * \text{Tipo_Acidente}_{atropelamento_animal,i} \\
& + \beta_{16} * \text{Tipo_Acidente}_{outros,i}) + \beta_{17} * \text{Num_Veiculos}_i \\
& + \beta_{18} * \text{Area_Critica}_i + \beta_{19} * \text{Condicao_Pista}_i \\
& + \beta_{20} \text{Perfil_Pista}_i + (\beta_{21} \text{Tracado_Pista}_{curva_suave,i} \\
& + \beta_{22} \text{Tracado_Pista}_{curva_acentuada,i}) + \beta_{23} * \text{Tipo_Local}_i \\
& + \beta_{24} * \text{Iluminacao}_i + \beta_{25} * \text{Marginal}_i + \beta_{26} * \text{Trafego}_i \\
& + \beta_{27} * \text{Motocicleta}_i + \beta_{28} * \text{Veiculo_Leve}_i \\
& + \beta_{29} * \text{Veiculo_Pesado}_i + \beta_{30} * \text{Veiculo_Passageiro}_i \\
& + (\beta_{31} * \text{Idade_Veiculo}_{6-10,i} + \beta_{32} * \text{Idade_Veiculo}_{11-20,i} \\
& + \beta_{33} * \text{Idade_Veiculo}_{20+,i} + \beta_{34} * \text{Idade_Veiculo}_{naoinfo,i}) \\
& + (\beta_{35} * \text{Estacao}_{primavera,i} + \beta_{36} * \text{Estacao}_{outono,i} \\
& + \beta_{37} * \text{Estacao}_{inverno,i}) + \beta_{38} * \text{Periodo_Dia}_i \\
& + \beta_{39} * \text{Anterior_2015}_i
\end{aligned}$$

Onde:

P_i : é a probabilidade de ocorrência de uma vítima grave ou fatal ocorrer para a observação i ;

β_0 : é o intercepto do modelo;

β_1 : Coeficiente para a variável Sexo, considerando a categoria de referência “masculino”;

$\beta_2, \beta_3, \beta_4, \beta_5, \beta_6$: Coeficientes para a variável dicotômica “Faixa Etária”, considerando a categoria de referência “1-17 anos”;

β_7 : Coeficiente para a variável dicotômica “Posição da Vítima”, considerando a categoria de referência “condutor”;

$\beta_8, \beta_9, \beta_{10}, \beta_{11}, \beta_{12}, \beta_{13}, \beta_{14}, \beta_{15}, \beta_{16}$: Coeficientes para a variável categórica “Tipo de Sinistro”, considerando a categoria de referência “Queda de Moto”;

β_{17} : Coeficiente para a variável dicotômica “Número de Veículos”, considerando a categoria de referência “Um”;

β_{18} : Coeficiente para a variável dicotômica “Área Crítica”, considerando a categoria de referência “Sim”;

β_{19} : Coeficiente para a variável dicotômica “Condição da Pista”, considerando a categoria de referência “Seca”;

β_{20} : Coeficiente para a variável dicotômica “Perfil da Pista”, considerando a categoria de referência “Em Nível”;

β_{21}, β_{22} : Coeficiente para a variável categórica “Traçado da Pista”, considerando a categoria de referência “Reta”;

β_{23} : Coeficiente para a variável dicotômica “Tipo da Local”, considerando a categoria de referência “Rural”;

β_{24} : Coeficiente para a variável dicotômica “Iluminação”, considerando a categoria de referência “Presente”;

β_{25} : Coeficiente para a variável dicotômica “Marginal”, considerando a categoria de referência “Presente”;

β_{26} : Coeficiente para a variável contínua “Trafego”;

β_{27} : Coeficiente para a variável dicotômica “Motocicleta”, considerando a categoria de referência “Presente”;

β_{28} : Coeficiente para a variável dicotômica “Veículos Leves”, considerando a categoria de referência “Presente”;

β_{29} : Coeficiente para a variável dicotômica “Veículos Pesados”, considerando a categoria de referência “Presente”;

β_{30} : Coeficiente para a variável dicotômica “Veículos de Passageiros”, considerando a categoria de referência “Presente”;

$\beta_{31}, \beta_{32}, \beta_{33}, \beta_{34}$: Coeficiente para a variável categórica “Idade do Veículo”, considerando a categoria de referência “0-5 anos”;

$\beta_{35}, \beta_{36}, \beta_{37}$: Coeficiente para a variável categórica “Estação”, considerando a categoria de referência “Verão”;

β_{38} : Coeficiente para a variável dicotômica “Período do Dia”, considerando a categoria de referência “Dia”;

β_{39} : Coeficiente para a variável dicotômica “Anterior a 2015”, considerando a categoria de referência “Sim”;

3.4.2 Método de Seleção

No modelo de regressão logística é necessário definir como será realizada a seleção das variáveis, para que apenas variáveis preditoras permaneçam no modelo. Para isso existem algumas técnicas, sendo a seleção *Backward* e a seleção *Forward* as mais utilizadas e conhecidas.

No método *Forward* o modelo inicia sem variáveis e, posteriormente são adicionadas todas as possíveis variáveis preditoras, iniciando com aquela que apresenta o menor valor-p. O processo continua até que todas as variáveis preditoras selecionadas tenham sido testadas e aquelas que não atingirem um limite pré-definido de valor-p (geralmente p acima de 0,05) sejam excluídas do modelo (HOSMER et al., 2013)

No atual estudo foi utilizado o método *Backward*, o qual inicia com um modelo completo que inclui todas as variáveis consideradas. Em seguida, as variáveis menos significativas são removidas, uma de cada vez, começando com aquela que tem o

maior valor-p. Esse processo continua até que reste apenas as variáveis mais significativas para o modelo, ou até que nenhuma variável seja deixada no modelo (HOSMER et al., 2013).

3.4.3 Modelos Propostos

Três modelos foram selecionados com base em critérios específicos. O modelo 1 foi escolhido por ser o mais abrangente e atender ao pré-requisito de que os resíduos estejam dentro do intervalo aceitável de -3 a 3. No modelo 2, foram removidas as variáveis com valores de p maiores que 0,1, indicando menor significância estatística.

Por fim, no modelo 3, foram excluídas as variáveis cujos resultados não faziam sentido lógico no contexto do estudo, como a redução da gravidade em sinistros envolvendo motocicletas, o que contradiz a literatura sobre o comportamento desse tipo de variável. O Modelo 1 incorporou um total de 24 variáveis, enquanto o Modelo 2 utilizou 17 variáveis e o Modelo 3, 16 variáveis, como demonstrado na Tabela 12.

Tabela 12 – Variáveis Utilizadas em Cada Modelo

Nome variável	Modelo 1	Modelo 2	Modelo 3
Sexo da Vítima	✓	✓	✓
Faixa Etária da Vítima	✓	✗	✗
Posição da Vítima	✓	✗	✗
Tipo de Sinistro	✓	✗	✗
Período do dia	✓	✓	✓
Estação do ano	✓	✓	✓
Anterior a 2015	✓	✗	✗
Ano do Veículo da Vítima	✓	✗	✗
Condição da Pista	✓	✓	✓
Perfil da Pista	✓	✓	✓
Traçado da Pista	✓	✓	✓
Tipo de Local	✓	✓	✓
Houve Motocicleta Envolvida?	✓	✓	✗
Houve Veículo de Passeio Envolvido?	✓	✓	✓

Nome variável	Modelo 1	Modelo 2	Modelo 3
Houve Veículo Pesado Envolvido?	✓	✓	✓
Houve Veículo de Passageiros Envolvido?	✓	✓	✓
Presença de Marginal	✓	✗	✗
Iluminação Noturna	✓	✗	✗
Tráfego Mensal Absoluto	✓	✓	✓
Número de Veículos Envolvidos	✓	✓	✓
Sinistro Presente em Agrupamento?	✓	✓	✓
Traçado + Condição da Pista	✓	✓	✓
Número de Veículos + Veículo Pesado	✓	✓	✓
Número de Veículos + Veículo Leve	✓	✓	✓

Fonte: Autora (2024)

3.4.4 Testes de Validação

Por fim, considerando todas as informações apresentadas sobre os três modelos construídos, os testes de razão de verossimilhança produziram valores de $p < 2,2e^{-16}$ tanto para a comparação entre o Modelo 1 e o Modelo 2, quanto entre o Modelo 2 e o Modelo 3. Esses valores de p são todos menores que 0,05, indicando que os modelos são estatisticamente diferentes entre si.

Por fim, ao considerar todas as informações apresentadas sobre os dois modelos construídos, o teste de razão de verossimilhança apresentou valor de p igual a 0.00786, ou seja, menor que 0.05. O resultado indica que os modelos são estatisticamente diferentes entre si.

As demais medidas de qualidade de cada modelo são apresentadas na Tabela 13.

Tabela 13 - Medida de qualidade

Modelo	Resíduos	Graus de Liberdade	AIC	BIC	Deviance
Modelo 1	[-2,81; 1,86]	21.291	25.789	26.044	25.717

Modelo 2	[-2,72; 2,32]	21.322	25.561	25.869	25.491
Modelo 3	[-2,72; 2,31]	21.324	25.527	25.840	25.448

Fonte: Autora (2024)

O modelo 3 apresenta o menor AIC, BIC, e *deviance*, sugerindo que oferece o melhor equilíbrio entre ajuste e complexidade do modelo. O Modelo 2 também apresenta um bom desempenho, especialmente em termos de amplitude dos resíduos. O Modelo 1, embora ainda dentro de um intervalo aceitável e valores mais altos de AIC, BIC e *deviance*.

Esses resultados indicam que o Modelo 3 é o mais robusto e eficiente na previsão da gravidade dos sinistros, considerando as variáveis incluídas no modelo. A seguir apresentam-se os resultados.

4 RESULTADOS

Estão representados na

Tabela 14 os resultados da regressão logística binomial: o coeficiente estimado, estatística t, o valor p, os valores das OR, os intervalos de confiança. O asterisco representa a variável utilizada como referência e os dados em vermelho simbolizam que aquela categoria apresentou um valor de p menor que 0.05, sendo significativo para o modelo.

Tabela 14 - Variáveis Independentes e Parâmetros

	Estimate	t-statistic	p-value		OR	Limite Inferior	Limite Superior
Gênero							
Feminino	0.297	0.038	5.00E-15	***	1.35	1.25	1.45
* Masculino							
Período do dia							
* Dia							
Noite	-0.171	0.031	2.75E-08	***	0.84	0.79	0.90
Estação							
* Verão							
Primavera	0.017	0.042	6.88E-01		1.02	0.94	1.10
Outono	-0.078	0.043	7.02E-02	.	0.92	0.85	1.01
Inverno	-0.116	0.043	7.24E-03	**	0.89	0.82	0.97
Condição da Pista							
* Dia							
Molhada	0.141	0.052	6.40E-03	**	1.15	1.04	1.28
Perfil da Pista							
* Em Nível							
Declive	-0.151	0.039	1.04E-04	***	0.86	0.80	0.93
Active	-0.020	0.045	6.53E-01		0.98	0.90	1.07
Traçado da Pista							
* Reta							

	Estimate	t-statistic	p-value		OR	Limite Inferior	Limite Superior
Curva Suave	0.075	0.043	8.37E-02	.	1.08	0.99	1.17
Curva Acentuada	0.032	0.052	5.29E-01		1.03	0.93	1.14
Tipo de Local							
* Rural							
Urbano	0.098	0.038	9.96E-03	**	1.10	1.02	1.19
Veículo Leve							
* Não presente							
Presente	0.196	0.043	5.86E-06	***	1.22	1.12	1.33
Veículo Pesado							
* Não presente							
Presente	-0.213	0.067	1.51E-03	**	0.81	0.71	0.92
Veículo de Passageiros							
* Não presente							
Presente	-0.163	0.062	8.90E-03	**	0.85	0.75	0.96
Volume							
Variável contínua	-0.179	0.0404	9.42E-06	***	0.84	0.77	0.90
Número de Veículos							
* Um							
Dois ou mais	-0.025	0.061	6.84E-01		0.98	0.87	1.10
Área Crítica							
* Não							
Sim	0.100	0.033	2.60E-03	**	1.10	1.04	1.18
Condição da Pista vs Traçado da Pista							
* Pista Molhada e Reta							
Pista Molhada e Curva Suave	-0.054	0.094	5.67E-01		0.95	0.79	1.14
Pista Molhada e Curva Acentuada	0.221	0.095	1.93E-02	*	1.25	1.04	1.50
Número de Veículos e Veículo Pesado							
* Com Veículo Pesado e um Veículo Envolvido							

	Estimate	t-statistic	p-value		OR	Limite Inferior	Limite Superior
Com Veículo Pesado e Dois ou Mais Veículos Envolvidos	-0.475	0.076	4.76E-10	***	0.62	0.54	0.72
Número de Veículos e Veículo Leve							
* Com Veículo Leve e um Veículo Envolvido							
Com Veículo Leve e Dois ou Mais Veículos Envolvidos	-0.226	0.085	8E-03	**	0.80	0.67	0.94

Fonte: Autora (2024)

Das 24 variáveis analisadas, 16 variáveis foram significativas para o modelo. Entre estas, 13 das 21 categorias foram consideradas significativas. A razão de chances para uma vítima do sexo feminino foi de 1,35, o que denota que mulheres têm 35% mais chance de ser uma vítima fatal ou grave em sinistros, quando comparados aos homens, como apresentado nos trabalhos de Dissanayake e Lu (2002), Martin e Lenguerrand (2008), Majdzadeh et al. (2008), Mergia et al. (2013), Ye et al. (2021) e Ryan et al. (2022). Essa diferença pode estar relacionada à morfologia corporal e à força física.

De acordo com Evans (2001) há diferenças fundamentais dependentes de gênero entre motoristas do sexo feminino e masculino, de modo que um motorista do sexo feminino tem mais probabilidade de morrer ao sofrer impactos físicos semelhantes e potencialmente letais.

Para a condição de pista molhada, a razão de chances foi de 1.15. Isso sugere que as condições de pista molhada estão associadas a um aumento de 15% na chance de uma vítima ser fatal ou grave em sinistros, em comparação com as condições de pista seca. Risco também verificado nos trabalhos de Yan et al. (2005^a), Rifaat e Tay (2009), Islam e Ozkul (2019) e Meng et al. (2020).

A variável combinada de Condição da Pista e Perfil da Pista indica que sinistros em curvas acentuadas e com pista molhada resultam em um aumento de 25% na probabilidade de o sinistro ser grave ou fatal. Isso ocorre devido à dificuldade de estabilizar o veículo em condições de menor atrito. Estudos de Dissanayake e Lu (2002b), Farid e Ksaibati (2021), Mohanty et al. (2022) e Zhu et al. (2019) demonstram que sinistros em curvas já estão associados a um aumento no número de sinistros.

Portanto, é compreensível que a combinação de pista molhada, outro fator que contribui para o aumento da gravidade, reforça ainda mais esse impacto (YAMAMOTO; SHANKAR, 2004; YAZDANI; NASSIRI, 2021).

As variáveis "veículo pesado" e "veículo de passageiros" apresentaram redução na razão de chances de 19% e 15%, respectivamente, na gravidade dos sinistros. Segundo Yan, Radwan e Abdel-Aty (2005), veículos grandes, por serem mais lentos e mais visíveis no trânsito, reduzem o risco de fatalidade quando envolvidos em sinistros. Rahimi et al. (2020) destacam as restrições de velocidade mais rigorosas para veículos pesados, o que contribui para uma maior atenção dos motoristas à estrada e, conseqüentemente, reduz a gravidade dos sinistros.

Esse efeito é ainda mais destacado na interação entre a variável "número de veículos e veículo pesado", onde um sinistro envolvendo um veículo pesado e mais de um outro veículo pode reduzir o risco de ferimentos graves ou fatais em 38%. Isso sugere que a maior visibilidade e a menor velocidade dos veículos pesados contribuem para uma menor gravidade nos sinistros em que estão envolvidos.

A razão de chances para a variável volume de tráfego é de 0.84. Isso indica que um aumento de uma unidade no logaritmo do volume está associado a uma redução de cerca de 16% nas chances de uma vítima ser classificada como grave ou fatal. Sendo uma unidade de 'Volume', que está em escala logarítmica, igual a 217.117 veículos. Portanto, podemos interpretar que um aumento de 217.117 veículos absolutos mensalmente está correlacionado a uma redução de aproximadamente 16% nas chances de uma vítima ser classificada como grave ou fatal, considerando todas as outras variáveis constantes.

Tais resultados diferem dos dados apresentados por Yu e Abdel-Aty (2014) e Theofilatos (2017), que mostram que o aumento do tráfego aumenta a gravidade dos sinistros. No entanto, de acordo com Holdridge, Shankar e Ulfarsson (2005), Martin e Lenguerrand (2008) e Meng et al. (2020) à medida que o limite de velocidade aumenta, a propensão a lesões também aumenta.

Isso pode estar em linha com os resultados encontrados no estudo, especialmente ao considerar a relação entre o aumento do tráfego e a saturação da rede viária, que, em alguns casos, pode levar à redução da velocidade dos veículos em vias congestionadas.

Para entender o impacto do aumento de 217.117 veículos no tráfego, analisamos os dados de 2022. Na rodovia Fernão Dias, um aumento de uma unidade

no volume logaritmo na praça de pedágio mais movimentada, na Grande Belo Horizonte, implica em um aumento de 21% no tráfego atual. Para a praça menos movimentada, esse aumento chega a 83%. Na rodovia Litoral Sul, o mesmo incremento representa um acréscimo de 27% na praça de pedágio mais movimentada e 49% na praça menos movimentada. Na rodovia Régis Bittencourt, observa-se um aumento de 45% na praça mais movimentada e 68% na praça menos movimentada.

Considerando que nos últimos 10 anos houve um aumento médio de 15% no tráfego nas três concessionárias, é improvável que a redução de 15% nas vítimas graves e fatais não seja alcançada nos próximos 10 anos apenas com o aumento do tráfego. Isso sugere que outros fatores ou medidas além do aumento do tráfego precisarão ser considerados e implementados para reduzir as fatalidades.

Algumas variáveis analisadas apresentaram resultados diferentes daqueles comumente encontrados na literatura. Por exemplo, o modelo indicou que sinistros ocorridos durante a noite tendem a diminuir a gravidade em 16% em comparação com os sinistros que ocorrem durante o dia. Possivelmente devido à maior cautela dos motoristas no período noturno. Porém estudos como Shrestha e Shrestha (2017), Lin et al. (2019), Useche et al. (2022) e Jashami et al. (2024) encontraram que sinistros em período noturno aumentam a gravidade do sinistro.

Considerando que o inverno é uma estação chuvosa, seria esperado um aumento na gravidade dos sinistros devido às condições da pista. No entanto, no modelo apresentado, essa variável mostrou uma redução em 11% na gravidade dos sinistros, o que pode ser comparado aos achados de Khattak e Knapp (2001), que observaram uma redução similar na gravidade dos sinistros durante o inverno em regiões com neve. No entanto, a condição de neve, que influencia a redução da gravidade, não se aplica ao contexto brasileiro.

Além disso, estudos de Gill e Goldacre (2009) e Sun et al. (2022) atribuíram o aumento da gravidade dos sinistros durante o inverno à redução do período de luz do dia. Porém, é importante destacar que, no contexto brasileiro, essa redução de luminosidade no inverno é mínima, com uma variação de, no máximo, duas horas, ao contrário dos países estudados, onde a diferença na duração da luz do dia entre o inverno e o verão pode chegar a oito horas. Essa diferença torna o impacto da redução da luminosidade menos significativo no Brasil.

Para o tipo de local "Urbano", a razão de chances foi de 1.10. Isso sugere que sinistros em áreas urbanas estão associados a um aumento de 10% na chance de

uma vítima ser fatal ou grave, em comparação com áreas rurais. Isso pode ser atribuído ao maior número de acessos e à maior propensão à poluição visual, fator que podem distrair o motorista e aumentar a gravidade dos sinistros.

Esse resultado contrasta com os achados de Li et al. (2009), que observaram uma redução na gravidade dos sinistros em regiões urbanas. No entanto, essa redução foi amplamente atribuída à diminuição da velocidade em áreas com semáforos, característica ausente nas rodovias analisadas neste estudo.

Além disso, este trabalho introduziu variáveis com algumas modificações, bem como novas variáveis, que podem enriquecer a literatura existente. Enquanto Apronti et al. (2019) e Yan et al. (2005b) destacam na frequência na ocorrência de sinistros em regiões de subida ou descida, principalmente em áreas montanhosas, eles não abordam o impacto na gravidade dos sinistros nessas situações. No presente estudo, os locais em declive apresentam uma redução de 14% na gravidade dos sinistros. Isso pode ser explicado pelas diversas melhorias implementadas em regiões montanhosas, como iluminação e construção de duas áreas de escape. Em locais onde a média anual de vítimas fatais ou graves era cinco vezes maior antes das melhorias.

O envolvimento de veículos leves resultou em um aumento de 22% na gravidade dos sinistros. No entanto, ao considerar o número de veículos envolvidos, observa-se que a gravidade tende a diminuir em 20% quando mais de um veículo está envolvido no sinistro. Isso sugere que a maior gravidade dos sinistros com veículos leves está possivelmente associada a situações em que o veículo está sozinho na rodovia, o que pode estar relacionado a um aumento da velocidade e à falta de atenção à via. Nesses casos, o veículo pode perder o controle e colidir com objetos ou estruturas, resultando em sinistros mais graves.

A variável "áreas críticas," que avalia se o sinistro ocorreu em uma área crítica definida com base em grupos, revela que sinistros nessas áreas aumenta o risco de fatalidade em 10%. Isso pode ser atribuído ao fato de que muitos dos pontos críticos estão localizados em regiões urbanas, que no modelo também apresentou aumento da gravidade de sinistros.

5 CONSIDERAÇÕES FINAIS

Nos últimos anos, a busca pela redução da gravidade dos sinistros de trânsito tem se intensificado, visando diminuir o número de fatalidades. No entanto, no Brasil, ainda são poucos os estudos que analisam especificamente as rodovias brasileiras. Este trabalho buscou identificar os principais fatores que influenciam a gravidade dos sinistros.

Sendo uma área de estudo recente, muitos ajustes são necessários para o desenvolvimento de um modelo adequado ao cenário brasileiro. Isso se deve às características únicas das rodovias e aos comportamentos específicos dos motoristas, que diferem daqueles observados em outros países.

Diante desse contexto, este trabalho propôs um modelo de regressão logística para analisar sinistros com vítimas moderadas, graves e fatais nos estados de Santa Catarina, Paraná, São Paulo e Minas Gerais, entre os anos de 2010 e 2022 nas rodovias administradas pelas concessionárias pela Fernão Dias, Régis Bittencourt e Litoral Sul. A análise foi baseada em dados sobre infraestrutura viária, características das vítimas, dos veículos envolvidos e do ambiente.

Com base nos resultados apresentados nesse trabalho, foi possível identificar diversas áreas onde intervenções específicas podem contribuir para a redução da gravidade dos sinistros nas rodovias analisadas. Estes resultados fornecem um auxílio para a implementação de melhorias que abrangem desde a segurança viária até campanhas de conscientização.

Os dados indicam que as mulheres são fisicamente mais vulneráveis do que os homens em sinistros de trânsito, o que ressalta a importância de adaptações específicas nos equipamentos de segurança veicular para atender às diferenças de gênero. Os cintos de segurança, airbags e outros dispositivos de proteção podem, por exemplo considerar essas diferenças, entendendo que certos modelos de carros possuem uma maior proporção de usuárias mulheres, tais adaptações podem ter um impacto significativo na redução de fatalidades e lesões graves.

Além disso, a combinação de curvas acentuadas e pistas molhadas foi identificada como um fator de risco significativo, aumentando a gravidade dos sinistros. Para mitigar esse risco, é essencial que esses pontos sejam priorizados nos projetos de manutenção e melhoria das rodovias. Melhorias na sinalização e a instalação de dispositivos de drenagem eficientes são medidas essenciais para evitar

a formação de pistas molhadas. O uso de materiais de alta aderência pode ajudar a estabilizar os veículos em condições adversas, contribuindo para a segurança dos motoristas.

A análise também indicou uma redução na gravidade dos sinistros com o aumento do volume de tráfego, possivelmente devido ao efeito do volume sobre a capacidade, que tende a diminuir a velocidade dos veículos em trechos de saturação da rede. Considerando essa relação, a implementação de sistemas de gerenciamento de tráfego que ajustem dinamicamente os limites de velocidade com base nas condições do tráfego surge como uma medida promissora e defendida no Sistema Seguro.

Além disso, a ONU já destacou que o Brasil carece de controle adequado sobre a velocidade no trânsito. Essa deficiência é especialmente relevante para áreas urbanas e regiões identificadas neste estudo como críticas, ambas apresentando um aumento na gravidade dos acidentes. A falta de controle mais rígido da velocidade nesses trechos permanece um fator crítico, reforçando a urgência de medidas que busquem mitigar tais problemas.

Os desafios propostos para trabalhos futuros incluem: análise da gravidade dos sinistros por tipo de veículo, uso de outros modelos estatísticos, como parâmetros aleatórios ou modelos mistos, análise detalhada dos clusters de sinistros, estudo de sinistros noturnos em rodovias federais, e impacto do volume de veículos no nível de serviço, velocidade e gravidade dos sinistros.

Por fim, a conscientização pública desempenha um papel crucial na segurança rodoviária. Com base nos resultados, é possível desenvolver campanhas educativas direcionadas que abordem os principais riscos identificados, como dirigir em condições de pista molhada, durante o dia em áreas urbanas, e em curvas acentuadas, que exigem maior atenção. Essas campanhas devem promover comportamentos seguros e alertar os motoristas sobre os riscos associados às diferentes condições rodoviárias, contribuindo para um ambiente mais seguro para todos.

REFERÊNCIAS

- AAA FOUNDATION FOR TRAFFIC SAFETY. **Prevalence of Self-Reported Aggressive Driving Behavior: United States**. Washington, DC, 2016.
- AASHTO. **Highway Safety Manual**. 2009.
- AKAIKE, H. Factor analysis and AIC. **Psychometrika**, v. 52, p. 317–331, 1987.
- APRONTI, D. T.; SAHA, P.; MOOMEN, M.; KSAIBATI, K. Truck safety evaluation on Wyoming mountain passes. **Accident Analysis & Prevention**, v. 122, p. 342–349, 2019.
- BRAGA, A. C. DA S. **Curvas ROC: aspectos funcionais e aplicações.**, 2000. Tese (Doutorado em Engenharia de Produção e Sistemas), Braga: Universidade do Minho.
- CARVALHO, C. H. R. **MORTES POR ACIDENTES DE TRANSPORTE TERRESTRE NO BRASIL: ANÁLISE DOS SISTEMAS DE INFORMAÇÃO DO MINISTÉRIO DA SAÚDE**. Rio de Janeiro, 2016.
- CARVALHO, C. H. R. **CUSTOS DOS ACIDENTES DE TRÂNSITO NO BRASIL: ESTIMATIVA SIMPLIFICADA COM BASE NA ATUALIZAÇÃO DAS PESQUISAS DO IPEA SOBRE CUSTOS DE ACIDENTES NOS AGLOMERADOS URBANOS E RODOVIAS**. Brasília/DF, 2020.
- CHANG, L.-Y.; CHIEN, J.-T. Analysis of driver injury severity in truck-involved accidents using a non-parametric classification tree model. **Safety Science**, v. 51, n. 1, p. 17–22, 2013.
- CHENG, W.; WASHINGTON, S. P. Experimental evaluation of hotspot identification methods. **Accident Analysis and Prevention**, v. 37, n. 5, p. 870–881, 2005.
- CNT; SEST SENAT; ITL. **Pesquisa CNT de Rodovias 2023**. Brasília, 2023.
- COMI, A.; POLIMENI, A.; BALSAMO, C. Road Accident Analysis with Data Mining Approach: Evidence from Rome. *Transportation Research Procedia*. **Anais...** . v. 62, p.798–805, 2022. Elsevier B.V.
- COX, D. R.; SNELL, E. J. **Analysis of Binary Data**. 2º ed. Londres: Routledge, 1989.
- DAMACENA, G. N.; MALTA, D. C.; BOCCOLINI, C. S.; et al. Consumo abusivo de álcool e envolvimento em acidentes de trânsito na população brasileira,

2013. **Ciencia e Saude Coletiva**, v. 21, n. 12, p. 3777–3786, 2016. Associacao Brasileira de Pos - Graduacao em Saude Coletiva.

DAS, S.; JAFARI, M.; HOSSAIN, A.; CHAKRABORTY, R.; MIMI, M. S. Toll road crash severity using mixed logit model incorporating heterogeneous mean structures. **Transportmetrica A: Transport Science**, 2024.

DEPARTAMENTO NACIONAL DE INFRAESTRUTURA DE TRANSPORTES. **INSTRUÇÃO NORMATIVA Nº 27/2020/DNIT SEDE, DE 30 DE JULHO DE 2020**. Brasil, 2020.

DISEASES AND INJURIES COLLABORATORS. Global burden of 369 diseases and injuries in 204 countries and territories, 1990 - 2019: a systematic analysis for the Global Burden of Disease Study 2019. **Lancet**, v. 396, n. 10258, p. 1204–1222, 2020.

DISSANAYAKE, S.; LU, J. J. Factors influential in making an injury severity difference to older drivers involved in fixed object–passenger car crashes. **Accident Analysis & Prevention**, v. 34, n. 5, p. 609–618, 2002.

ELVIK, R. Laws of accident causation. **Accident Analysis and Prevention**, v. 38, n. 4, p. 742–747, 2006.

ELVIK, R.; HØYE, A.; VAA, T.; SØRENSEN, M. **O Manual de Medidas de Segurança Viária**. Madrid, 2015.

ERMAN, J.; ARLITT, M.; MAHANTI, A. **Traffic Classification Using Clustering Algorithms**. Calgary, Canadá, 2006.

EVANS, L. Female compared with male fatality risk from similar physical impacts. **Journal of Traffic Medicine**, p. 10–19, 2001.

FARID, A.; KSAIBATI, K. Modeling severities of motorcycle crashes using random parameters. **Journal of Traffic and Transportation Engineering (English Edition)**, v. 8, n. 2, p. 225–236, 2021.

FERKO, M.; BABIĆ, DARIO; BABIĆ, DARKO; et al. Influence of Road Safety Barriers on the Severity of Motorcyclist Injuries in Horizontal Curves. **Sustainability**, v. 14, n. 22, 2022.

FRALEY, C.; RAFTERY, A. E. Model-based clustering, discriminant analysis, and density estimation. **Journal of the American Statistical Association**, v. 97, n. 458, 2002.

GILL, M.; GOLDACRE, M. J. Seasonal variation in hospital admission for road traffic injuries in England: analysis of hospital statistics. **Injury Prevention**, v. 15, n. 6, p. 374–378, 2009.

GOLD, P. A. **Segurança no Trânsito: Aplicações de Engenharia para Reduzir Acidentes**. New York: Banco Interamericano de Desenvolvimento, 1998.

HAKKERT, A. S.; MAHALEL, D. ESTIMATING THE NUMBER OF ACCIDENTS AT INTERSECTIONS FROM A KNOWLEDGE OF THE TRAFFIC FLOWS ON THE APPROACHES. **Accident Analysis & Prevention**, v. 10, n. 1, p. 69–79, 1978. Pergamon Press.

HOLDRIDGE, J. M.; SHANKAR, V. N.; ULFARSSON, G. F. The crash severity impacts of fixed roadside objects. **Journal of Safety Research**, v. 36, n. 2, p. 139–147, 2005.

HOSMER, D. W.; LEMESHOW, S.; STURDIVANT, R. X. **Applied Logistic Regression**. 3^o ed. New Jersey: Wiley, 2013.

ISLAM, M.; OZKUL, S. Identifying Fatality Risk Factors for the Commercial Vehicle Driver Population. **Transportation Research Record: Journal of the Transportation Research Board**, v. 2673, n. 9, p. 297–310, 2019.

JAMES, G.; WITTEN, D.; HASTIE, T.; TIBSHIRANI, R. **An Introduction to Statistical Learning: With Applications in R**. Springer, 2013.

JASHAMI, H.; ANDERSON, J. C.; MOHAMMED, H. A.; COBB, D. P.; HURWITZ, D. S. Contributing factors to right-turn crash severity at signalized intersections: An application of econometric modeling. **International Journal of Transportation Science and Technology**, v. 13, p. 243–257, 2024.

KHATTAK, A.; KNAPP, K. Interstate Highway Crash Injuries During Winter Snow and Nonsnow Events. **Transportation Research Record**, p. 30–36, 2001.

KITCHENHAM, B.; CHARTERS, S. **Guidelines for performing Systematic Literature Reviews in Software Engineering**. Reviews in Software Engineering, 2007. Reviews in Software Engineering.

KONONEN, D. W.; FLANNAGAN, C. A. C.; WANG, S. C. Identification and validation of a logistic regression model for predicting serious injuries associated with motor vehicle crashes. **Accident Analysis & Prevention**, v. 43, n. 1, p. 112–122, 2011.

KOORNSTRA, M. J.; MATHIJSSSEN, M. P. M.; MULDER, J. A. G.; ROSZBACH, R.; WEGMAN, F. C. M. **Naar een duurzaam veilig wegverkeer:**

Nationale Verkeersveiligheidsverkenning voor de jaren 1990/2010.

Leidschendam, Netherlands, 1992.

KULLBACK, S.; LEIBLER, R. A. On Information and Sufficiency. **The Annals of Mathematical Statistics**, v. 22, n. 1, p. 79–86, 1951.

KWON, O. H.; RHEE, W.; YOON, Y. Application of classification algorithms for analysis of road safety risk factor dependencies. **Accident Analysis & Prevention**, v. 75, p. 1–15, 2015.

LI, M.-D.; DOONG, J.-L.; HUANG, W.-S.; LAI, C.-H.; JENG, M.-C. Survival hazards of road environment factors between motor-vehicles and motorcycles. **Accident Analysis & Prevention**, v. 41, n. 5, p. 938–947, 2009.

LIN, P.-S.; GUO, R.; BIALKOWSKA-JELINSKA, E.; KOURTELLIS, A.; ZHANG, Y. Development of countermeasures to effectively improve pedestrian safety in low-income areas. **Journal of Traffic and Transportation Engineering (English Edition)**, v. 6, n. 2, p. 162–174, 2019.

LINDEN, R. Técnicas de Agrupamento. **Revista de Sistemas de Informação da FSMA**, p. 18–36, 2009. Disponível em: <<http://www.fsma.edu.br/si/sistemas.html>>.

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, 2010. Elsevier Ltd.

MA, Z.; CHIEN, S. I.-J.; DONG, C.; HU, D.; XU, T. Exploring factors affecting injury severity of crashes in freeway tunnels. **Tunnelling and Underground Space Technology**, v. 59, p. 100–104, 2016.

MAHALEL, D.; HAKKERT, A. S.; PRASHKER, J. N. A SYSTEM FOR THE ALLOCATION OF SAFETY RESOURCES ON A ROAD NETWORK. **Accident Analysis & Prevention**, v. 14, n. 1, p. 45–56, 1982.

MAJDZADEH, R.; KHALAGI, K.; NARAGHI, K.; MOTEVALIAN, A.; ESHRAGHIAN, M. R. Determinants of traffic injuries in drivers and motorcyclists involved in an accident. **Accident Analysis and Prevention**, v. 40, n. 1, p. 17–23, 2008.

MALTA, D. C.; ANDRADE, S. S. C. DE A.; GOMES, N.; et al. Lesões no trânsito e uso de equipamento de proteção na população Brasileira, segundo estudo de base populacional. **Ciencia e Saude Coletiva**, v. 21, n. 2, p. 399–409, 2016. Associação Brasileira de Pós - Graduação em Saúde Coletiva.

MÂNICA, A. G. **MODELO DE PREVISÃO DE ACIDENTES RODOVIÁRIOS ENVOLVENDO MOTOCICLETAS**, 2007. Mestrado, Porto Alegre: Universidade Federal do Rio Grande do Sul.

MARTIN, J.-L.; LENGUERRAND, E. A population based estimation of the driver protection provided by passenger cars: France 1996–2005. **Accident Analysis & Prevention**, v. 40, n. 6, p. 1811–1821, 2008.

MENG, F.; XU, P.; SONG, C.; et al. Influential Factors Associated with Consecutive Crash Severity: A Two-Level Logistic Modeling Approach. **International Journal of Environmental Research and Public Health**, v. 17, n. 15, 2020.

MERGIA, W. Y.; EUSTACE, D.; CHIMBA, D.; QUMSIYEH, M. Exploring factors contributing to injury severity at freeway merging and diverging locations in Ohio. **Accident Analysis & Prevention**, v. 55, p. 202–210, 2013.

MINISTÉRIO DOS TRANSPORTES. Programa para Tratamento de Segmentos Críticos. Disponível em: <<https://www.gov.br/dnit/pt-br/1a-semana-do-planejamento/10ProgramadetratamentodesegmentoscriticosSS.pdf>>. Acesso em: 23/10/2022.

MOHANTY, M.; PANDA, R.; GANDUPALLI, S. R.; ARYA, R. R.; LENKA, S. K. Factors propelling fatalities during road crashes: A detailed investigation and modelling of historical crash data with field studies. **Heliyon**, v. 8, n. 11, p. e11531, 2022.

MOOMEN, M.; REZAPOUR, M.; KSAIBATI, K. An investigation of influential factors of downgrade truck crashes: A logistic regression approach. **Journal of Traffic and Transportation Engineering (English Edition)**, v. 6, n. 2, p. 185–195, 2019.

NAGELKERKE, N. J. D. A note on a general definition of the coefficient of determination. **Biometrika**, v. 78, n. 3, p. 691–692, 1991.

OÑA, J. DE; MUJALLI, R. O.; CALVO, F. J. Analysis of traffic accident injury severity on Spanish rural highways using Bayesian networks. **Accident Analysis & Prevention**, v. 43, n. 1, p. 402–411, 2011.

PAVARINO FILHO, R. V. As declarações de Moscou e Brasília sobre a segurança no trânsito – um paralelo entre dois momentos no tema da saúde. **Ciencia e Saude Coletiva**, v. 21, n. 12, p. 3649–3660, 2016. Associação Brasileira de Pós - Graduação em Saúde Coletiva.

PERVEZ, A.; LEE, J.; HUANG, H. Exploring factors affecting the injury severity of freeway tunnel crashes: A random parameters approach with heterogeneity in means and variances. **Accident Analysis & Prevention**, v. 178, p. 106835, 2022.

RAHIMI, E.; SHAMSHIRIPOUR, A.; SAMIMI, A.; MOHAMMADIAN, A. (KOUROS). Investigating the injury severity of single-vehicle truck crashes in a developing country. **Accident Analysis & Prevention**, v. 137, 2020.

REZAPOUR, M.; MEHRARA MOLAN, A.; KSAIBATI, K. Analyzing injury severity of motorcycle at-fault crashes using machine learning techniques, decision tree and logistic regression models. **International Journal of Transportation Science and Technology**, v. 9, n. 2, p. 89–99, 2020.

REZAPOUR, M.; WULFF, S. S.; KSAIBATI, K. Examination of the severity of two-lane highway traffic barrier crashes using the mixed logit model. **Journal of Safety Research**, v. 70, p. 223–232, 2019.

RIFAAT, S. M.; TAY, R. Effects of Street Patterns on Injury Risks in Two-Vehicle Crashes. **Transportation Research Record: Journal of the Transportation Research Board**, v. 2102, n. 1, p. 61–67, 2009.

RYAN, A.; TAINTER, F.; FITZPATRICK, C.; et al. The impact of sex on motor vehicle crash injury outcomes. **Journal of Transportation Safety and Security**, v. 14, n. 5, p. 818–842, 2022. Taylor and Francis Ltd.

SANTOLINO, M.; BOLANCÉ, C.; ALCANIZ, M. Factors affecting hospital admission and recovery stay duration of in-patient motor victims in Spain. **Accident Analysis & Prevention**, v. 49, p. 512–519, 2012.

SATRIA, R.; CASTRO, M. GIS Tools for Analyzing Accidents and Road Design: A Review. *Transportation Research Procedia*. **Anais...** . v. 18, p.242–247, 2016. Elsevier B.V.

SCHWAB, K. **The Global Competitiveness Report**. Suíça, 2019.

SCHWARZ, G. Estimating the Dimension of a Model. **The Annals of Statistics**, v. 6, n. 2, p. 461–464, 1978.

SEGURADORA LÍDER. **Relatório Anual 2020**. 2020.

SHRESTHA, P. P.; SHRESTHA, K. J. Factors associated with crash severities in built-up areas along rural highways of Nevada: A case study of 11 towns. **Journal of Traffic and Transportation Engineering (English Edition)**, v. 4, n. 1, p. 96–102, 2017.

SUN, Z.; XING, Y.; WANG, J.; et al. Exploring injury severity of vulnerable road user involved crashes across seasons: A hybrid method integrating random parameter logit model and Bayesian network. **Safety Science**, v. 150, 2022.

SWOV. **Principles for design and organization of a casualty-free road traffic system**. 2018.

SZENASI, S.; CSIBA, P. CLUSTERING ALGORITHM IN ORDER TO FIND ACCIDENT BLACK SPOTS IDENTIFIED BY GPS COORDIANTES. . p.17–26, 2014. Bulgaria: Geoconference on Informatics, Geoinformatics and Remote Sensing.

TAY, R.; RIFAAT, S. M.; CHIN, H. C. A logistic model of the effects of roadway, environmental, vehicle, crash and driver characteristics on hit-and-run crashes. **Accident Analysis & Prevention**, v. 40, n. 4, p. 1330–1336, 2008.

THANG, T. M.; KIM, J. The anomaly detection by using DBSCAN clustering with multiple parameters. 2011 International Conference on Information Science and Applications, ICISA 2011. **Anais...** , 2011.

THEOFILATOS, A. Incorporating real-time traffic and weather data to explore road accident likelihood and severity in urban arterials. **Journal of Safety Research**, v. 61, p. 9–21, 2017.

TSAI, C.-F.; YEN, C.-C. Unsupervised Anomaly Detection Using HDG-Clustering Algorithm. **ICONIP 2007 Parte II**, p. 356–365, 2008.

TURKMAN, A. A.; SILVA, G. L. Modelos Lineares Generalizados - da Teoria à Prática. VIII Congresso Anual da Sociedade Portuguesa de Estatística. **Anais...** , 2000. Lisboa.

USECHE, S. A.; CENDALES, B.; ALONSO, F.; MONTORO, L. Multidimensional prediction of work traffic crashes among Spanish professional drivers in cargo and passenger transportation. **International Journal of Occupational Safety and Ergonomics**, v. 28, n. 1, p. 20–27, 2022.

WAHAB, L.; JIANG, H. A comparative study on machine learning based algorithms for prediction of motorcycle crash severity. **PLOS ONE**, v. 14, n. 4, p. e0214966, 2019.

WAN, P.; WU, C.; LIN, Y.; MA, X. On-road experimental study on driving anger identification model based on physiological features by ROC curve analysis. **IET Intelligent Transport Systems**, v. 11, n. 5, p. 290–298, 2017.

WEGMAN, F.; AARTS, L. **Advancing sustainable safety: National road safety outlook for 2005–2006**. Leidschendam, The Netherlands, 2006.

WEGMAN, F.; AARTS, L.; VAN DER KNAAP, P. Sustainable Safety: A Short History of a Safe System Approach in the Netherlands. **The Vision Zero Handbook**. p.307–336, 2023. Cham: Springer International Publishing.

WHO, W. H. O. **Global status report on road safety: time for action.** Geneva: World Health Organization, 2009.

WHO, W. H. O. **Global Plan for the Decade of Action for Road Safety 2011-2020.** Geneva, 2011.

WHO, W. H. O. **Transformando Nosso Mundo: A Agenda 2030 para o Desenvolvimento Sustentável.** 2015.

WHO, W. H. O. **Global Status Report on Road Safety 2018.** 2018.

WHO, W. H. O. **PLANO GLOBAL DÉCADA DE AÇÃO PELA SEGURANÇA NO TRÂNSITO 2021 - 2030.** 2021.

WIJNEN, W.; STIPDONK, H. Social costs of road crashes: An international analysis. **Accident Analysis & Prevention**, v. 94, p. 97–106, 2016.

XUE, Q.; WANG, Y.; ZHAI, H.; CHANG, X. Automatic Identification of Fractures Using a Density-Based Clustering Algorithm with Time-Spatial Constraints. **Energies**, v. 11, n. 3, 2018.

YAMAMOTO, T.; SHANKAR, V. N. Bivariate ordered-response probit model of driver's and passenger's injury severities in collisions with fixed objects. **Accident Analysis & Prevention**, v. 36, n. 5, p. 869–876, 2004.

YAN, X.; RADWAN, E.; ABDEL-ATY, M. Characteristics of rear-end accidents at signalized intersections using multiple logistic regression model. **Accident Analysis & Prevention**, v. 37, n. 6, p. 983–995, 2005a.

YAN, X.; RADWAN, E.; ABDEL-ATY, M. Characteristics of rear-end accidents at signalized intersections using multiple logistic regression model. **Accident Analysis & Prevention**, v. 37, n. 6, p. 983–995, 2005b.

YAZDANI, M.; NASSIRI, H. The effect of weather on the severity of multi-vehicle crashes: A case study of Iran. **Proceedings of the Institution of Civil Engineers: Transport**, v. 174, n. 5, p. 333–342, 2021. ICE Publishing.

YE, F.; CHENG, W.; WANG, C.; LIU, H.; BAI, J. Investigating the severity of expressway crash based on the random parameter logit model accounting for unobserved heterogeneity. **Advances in Mechanical Engineering**, v. 13, n. 12, 2021.

YU, R.; ABDEL-ATY, M. Analyzing crash injury severity for a mountainous freeway incorporating real-time traffic and weather data. **Safety Science**, v. 63, p. 50–56, 2014.

ZHANG, G.; YAU, K. K. W.; ZHANG, X.; LI, Y. Traffic accidents involving fatigue driving and their extent of casualties. **Accident Analysis & Prevention**, v. 87, p. 34–42, 2016.

ZHU, L.; LU, L.; ZHANG, W.; ZHAO, Y.; SONG, M. Analysis of Accident Severity for Curved Roadways Based on Bayesian Networks. **Sustainability**, v. 11, n. 8, 2019.

APÊNDICE

Tabela 15 - Variáveis Independentes e Parâmetros do Modelo 1

	Estimate	t-statistic	p-value		OR	Limite Inferior	Limite Superior
Gênero							
Feminino	0.154	0.051	2.48E-03	**	1.17	1.06	1.29
* Masculino							
Faixa Etária							
1-17 anos	-0.228	0.090	1.92E-03	**	0.80	0.69	0.92
* 18-24 anos							
25-34 anos	-0.034	0.056	7.93E-01		0.97	0.75	1.25
35-49 anos	0.033	0.057	8.15E-01		1.03	0.78	1.37
Maior que 50 anos	-0.032	0.063	8.16E-01		0.97	0.74	1.27
Não Informado	-0.020	0.077	9.24E-01		0.98	0.66	1.47
Posição da Vítima							
* Condutor							
Passageiro	0.350	0.050	2.22E-12	***	1.42	1.29	1.57
Tipo de Sinistro							
* Queda de Moto							
Atropelamento - Animal	-0.186	0.277	3.52E-01		0.83	0.56	1.23
Capotamento	-0.464	0.103	6.51E-06	***	0.63	0.51	0.77
Choque	-0.232	0.096	2.46E-01		0.79	0.54	1.17
Colisão Frontal	-0.418	0.129	3.66E-02		0.66	0.44	0.97
Colisão Lateral/Transversal	-0.187	0.098	3.50E-01		0.83	0.56	1.23
Colisão Traseira	-0.276	0.089	1.84E-03	**	0.76	0.64	0.90
Outros	-0.252	0.118	2.08E-01		0.78	0.53	1.15
Saida de Pista	-0.312	0.115	1.37E-03	**	0.73	0.60	0.89
Tombamento	-0.037	0.105	8.53E-01		0.96	0.65	1.43
Período do dia							
* Dia							
Noite	-0.072	0.036	4.90E-02	*	0.93	0.87	1.00
Estação							
* Verão							
Primavera	-0.048	0.049	3.31E-01		0.95	0.87	1.05
Outono	-0.175	0.051	5.80E-04	***	0.84	0.76	0.93
Inverno	-0.115	0.051	2.44E-02	*	0.89	0.81	0.99
Anterior a 2015							
Sim	0.033	0.037	3.83E-01		1.03	0.96	1.11
* Não							
Ano do Veículo da Vítima							
* 0-5 anos							
6-10 anos	0.270	0.766	1.76E-01		1.31	0.89	1.94
11-20 anos	0.139	0.696	4.86E-01		1.15	0.78	1.70
Maior que 20 anos	0.059	0.296	7.67E-01		1.06	0.72	1.57

	Estimate	t-statistic	p-value		OR	Limite Inferior	Limite Superior
Não informado	-0.206	0.414	3.02E-01		0.81	0.55	1.20
Condição da Pista							
* Dia							
Molhada	0.079	0.061	1.96E-01		1.08	0.96	1.22
Perfil da Pista							
* Em Nível							
Declive	-0.081	0.046	7.98E-02	.	0.92	0.84	1.01
Aclive	-0.008	0.053	8.78E-01		0.99	0.90	1.10
Traçado da Pista							
* Reta							
Curva Suave	0.117	0.051	2.28E-02	*	1.12	1.02	1.24
Curva Acentuada	0.008	0.061	8.94E-01		1.01	0.89	1.14
Tipo de Local							
* Rural							
Urbano	0.046	0.923	3.56E-01		1.05	0.95	1.15
Motocicleta							
* Não presente							
Presente	-0.278	0.060	4.12E-06	***	0.76	0.67	0.85
Veículo Leve							
* Não presente							
Presente	0.172	0.089	5.45E-02	.	1.19	1.00	1.41
Veículo Pesado							
* Não presente							
Presente	-0.207	0.090	2.17E-02	*	0.81	0.68	0.97
Veículo de Passageiros							
* Não presente							
Presente	0.461	0.081	1.43E-08	***	1.59	1.35	1.86
Presença de Marginal							
* Não presente							
Presente	0.048	0.055	3.80E-01		1.05	0.94	1.17
Iluminação Noturna							
* Não presente							
Presente	-0.044	0.077	2.14E-01		0.04	0.89	1.03
Não se aplica	-0.057	0.108	3.92E-02		0.04	0.88	1.01
Volume							
Variável contínua	-0.310	0.048	1.18E-10	***	0.73	0.67	0.81
Número de Veículos							
* Um							
Dois ou mais	0.067	0.096	4.88E-01		1.07	0.89	1.29
Área Crítica							
* Não							
Sim	-0.047	0.039	2.23E-01		0.95	0.88	1.03
Condição da Pista vs Traçado da Pista							
* Pista Molhada e Reta							

	Estimate	t-statistic	p-value		OR	Limite Inferior	Limite Superior
Pista Molhada e Curva Suave	-0.148	0.110	1.79E-01		0.86	0.70	1.07
Pista Molhada e Curva Acentuada	0.119	0.111	2.84E-01		1.13	0.91	1.40
Número de Veículos e Veículo Pesado							
* Com Veículo Pesado e um Veículo Envolvido							
Com Veículo Pesado e Dois ou Mais Veículos Envolvidos	-0.337	0.101	8.23E-04	***	0.71	0.59	0.87
Número de Veículos e Veículo Leve							
* Com Veículo Leve e um Veículo Envolvido							
Com Veículo Leve e Dois ou Mais Veículos Envolvidos	-0.036	0.100	7.21E-01		0.96	0.79	1.17

Tabela 16 - Variáveis Independentes e Parâmetros do Modelo 2

	Estimate	t-statistic	p-value		OR	Limite Inferior	Limite Superior
Gênero							
* Feminino	0.286	0.038	6.34E-14	***	1.33	1.24	1.44
* Masculino							
Período do dia							
* Dia							
Noite	-0.186	0.031	1.94E-09	***	0.83	0.78	0.88
Estação							
* Verão							
Primavera	0.020	0.042	6.33E-01		1.02	0.94	1.11
Outono	-0.082	0.043	5.79E-02	.	0.92	0.85	1.00
Inverno	-0.114	0.043	8.32E-03	**	0.89	0.82	0.97
Condição da Pista							
* Dia							
Molhada	0.131	0.052	1.17E-02	*	1.14	1.03	1.26
Perfil da Pista							
* Em Nível							
Declive	-0.155	0.039	7.05E-05	***	0.86	0.79	0.92
Active	-0.026	0.045	5.57E-01		0.97	0.89	1.06
Traçado da Pista							
* Reta							
Curva Suave	0.073	0.043	9.17E-02	.	1.08	0.99	1.17
Curva Acentuada	0.025	0.052	6.26E-01		1.03	0.93	1.13
Tipo de Local							
* Rural							
Urbano	0.133	0.039	5.71E-04	***	1.14	1.06	1.23
Motocicleta							
* Não presente							
Presente	-0.246	0.051	1.79E-06	***	0.78	0.71	0.87
Veículo Leve							
* Não presente							

	Estimate	t-statistic	p-value		OR	Limite Inferior	Limite Superior
Presente	0.267	0.077	4.96E-04	***	1.31	1.12	1.52
Veículo Pesado							
* Não presente							
Presente	-0.221	0.077	4.40E-03	**	0.80	0.69	0.93
Veículo de Passageiros							
* Não presente							
Presente	-0.203	0.064	1.46E-03	**	0.82	0.72	0.93
Volume							
Variável contínua	-0.165	0.041	4.80E-05	***	0.85	0.78	0.92
Número de Veículos							
* Um							
Dois ou mais	0.174	0.083	3.64E-02	*	1.19	1.01	1.40
Área Crítica							
* Não							
Sim	0.104	0.033	1.70E-03	**	1.11	1.04	1.19
Condição da Pista vs Traçado da Pista							
* Pista Molhada e Reta							
Pista Molhada e Curva Suave	-0.056	0.094	5.53E-01		0.95	0.79	1.14
Pista Molhada e Curva Acentuada	0.222	0.095	1.94E-02	*	1.25	1.04	1.50
Número de Veículos e Veículo Pesado							
* Com Veículo Pesado e um Veículo Envolvido							
Com Veículo Pesado e Dois ou Mais Veículos Envolvidos	-0.596	0.086	5.06E-12	***	0.55	0.47	0.65
Número de Veículos e Veículo Leve							
* Com Veículo Leve e um Veículo Envolvido							
Com Veículo Leve e Dois ou Mais Veículos Envolvidos	-0.226	0.085	8.32E-03	**	0.80	0.67	0.94