

**UNIVERSIDADE FEDERAL DE MINAS GERAIS**

**Programa de Pós-Graduação em Estatística**

Hiago Andrade Duarte

**INVESTIGAÇÃO DA RELAÇÃO ENTRE A COR DO VEÍCULO E O RISCO DE  
FURTO E ROUBO:**

Uma análise para o estado de São Paulo de 2016 a 2021

Belo Horizonte

2023

Hiago Andrade Duarte

**INVESTIGAÇÃO DA RELAÇÃO ENTRE A COR DO VEÍCULO E O RISCO DE  
FURTO E ROUBO:**

Uma análise para o estado de São Paulo de 2016 a 2021

**Versão final**

Monografia de especialização apresentada  
à Universidade Federal de Minas Gerais,  
como requisito parcial à obtenção do título  
de Especialista em Estatística, com Ênfase  
em Indústria e Mercado.

Orientadora: Edna Afonso Reis

Belo Horizonte

2023

2023, Hiago Andrade Duarte.  
Todos os direitos reservados.

Duarte, Hiago Andrade

D812i      Investigação da relação entre a cor do veículo e o risco de furto e roubo: uma análise para o estado de São Paulo de 2016 a 2021 [manuscrito]/ Hiago Andrade Duarte—2023.  
1 recurso online (34 f. il, color).

Orientadora . Edna Afonso Reis.

Monografia (especialização) - Universidade Federal de Minas Gerais, Instituto de Ciências Exatas, Departamento de Estatística

Referências: 33 – 34.

1. Estatística. 2. Veículos – Cor. 3. Veículos – Cor – Furto. I. Reis, Edna Afonso. II. Universidade Federal de Minas Gerais. I. Instituto de Ciências Exatas, Departamento de Estatística. III. Título.

CDU 519.2 (043)

Ficha catalográfica elaborada pela bibliotecária Belkiz Inez Rezende Costa  
CRB 6/1510 Universidade Federal de Minas Gerais – ICEX



**Universidade Federal de Minas Gerais**  
**Instituto de Ciências Exatas**  
**Departamento de Estatística**  
**Programa de Pós-Graduação / Especialização**  
Av. Pres. Antônio Carlos, 6627 - Pampulha  
31270-901 – Belo Horizonte – MG

E-mail: [pgest@ufmg.br](mailto:pgest@ufmg.br)  
Tel: 3409-5923 – FAX: 3409-5924

## **ATA DO 290ª. TRABALHO DE FIM DE CURSO DE ESPECIALIZAÇÃO EM ESTATÍSTICA DE HIAGO ANDRADE DUARTE.**

Aos dezoito dias do mês de maio de 2023, às 17:00 horas, com utilização de recursos de videoconferência a distância, reuniram-se os professores abaixo relacionados, formando a Comissão Examinadora homologada pela Comissão do Curso de Especialização em Estatística, para julgar a apresentação do trabalho de fim de curso do aluno **Hiago Andrade Duarte**, intitulado: “Investigação da relação entre a cor do veículo e o risco de furto e roubo: uma análise para o estado de São Paulo de 2016 a 2021”, como requisito para obtenção do Grau de Especialista em Estatística. Abrindo a sessão, o Presidente da Comissão, Professora Edna Afonso Reis – Orientadora, após dar conhecimento aos presentes do teor das normas regulamentares, passou a palavra ao candidato para apresentação de seu trabalho. Seguiu-se a arguição pelos examinadores com a respectiva defesa do candidato. Após a defesa, os membros da banca examinadora reuniram-se sem a presença do candidato e do público, para julgamento e expedição do resultado final. Foi atribuída a seguinte indicação: o candidato foi considerado Aprovado condicional às modificações sugeridas pela banca examinadora no prazo de 30 dias a partir da data de hoje por unanimidade. O resultado final foi comunicado publicamente ao candidato pela Presidente da Comissão. Nada mais havendo a tratar, a Presidente encerrou a reunião e lavrou a presente Ata, que será assinada por todos os membros participantes da banca examinadora. Belo Horizonte, 18 de maio de 2023.

Prof.<sup>a</sup> Edna Afonso Reis (Orientadora)  
Departamento de Estatística / ICEx / UFMG

Jussiane Nader Gonçalves Assinado de forma digital por Jussiane Nader Gonçalves  
Dados: 2023.05.23 09:59:20 -03'00'

Prof.<sup>a</sup> Jussiane Nader Gonçalves  
Departamento de Estatística / ICEx / UFMG

Prof.<sup>a</sup> Thais Rotsen Corrêa  
Departamento de Estatística / ICEx / UFMG



**Universidade Federal de Minas Gerais**  
**Instituto de Ciências Exatas**  
**Departamento de Estatística**  
**Programa de Pós-Graduação / Especialização**  
Av. Pres. Antônio Carlos, 6627 - Pampulha  
31270-901 – Belo Horizonte – MG

**E-mail: pgest@ufmg.br**  
Tel: 3409-5923 – FAX: 3409-5924

## **DECLARAÇÃO DE CUMPRIMENTO DE REQUISITOS PARA CONCLUSÃO DO CURSO DE ESPECIALIZAÇÃO EM ESTATÍSTICA.**

Declaro para os devidos fins que Hiago Andrade Duarte, número de registro 2021677049, cumpriu todos os requisitos necessários para conclusão do curso de Especialização em Estatística e que entregou para sua orientadora, a professora Edna Afonso Reis, o trabalho, que aprovou a versão final. O trabalho foi apresentado no dia 18 de maio de 2023 com o título “Investigação da relação entre a cor do veículo e o risco de furto e roubo: uma análise para o estado de São Paulo de 2016 a 2021”.

Belo Horizonte, 06 de junho de 2023

Roberto da Costa  
Quinino:8087129172  
0

Assinado de forma digital por  
Roberto da Costa  
Quinino:80871291720  
Dados: 2023.06.06 15:14:20 -03'00'

**Prof. Roberto da Costa Quinino**  
**Coordenador do curso de**  
**Especialização em Estatística**  
**Departamento de Estatística / UFMG**

## RESUMO

O presente estudo teve como objetivo investigar se há uma relação entre a cor do veículo e o risco de ocorrer um evento de furto ou roubo. Os dados utilizados foram obtidos da Secretaria de Segurança Pública do Estado de São Paulo, contendo os boletins de ocorrência de furtos ou roubos de veículos entre 2016 e 2021. Como não há dados individualizados de veículos que não foram roubados/furtados, foi aplicado o método de exposição induzida. Para aplicação do método neste trabalho, a condição de luminosidade no momento do roubo ou furto do veículo foi o fator escolhido para definir os dois grupos de comparação: um grupo denominado “propenso a cores”, formado pelas ocorrências nas quais a luminosidade natural era suficiente e, desse modo, a cor do veículo poderia ter contribuído para a ocorrência de um furto ou roubo; e outro grupo, denominado “exposição induzida”, formado pelos registros de furtos e roubos que ocorreram em condições de baixa luz natural, menos propícias à influência da coloração do automóvel em razão da menor visibilidade. As distribuições Binomial Negativa e Poisson foram utilizadas para modelar o número de roubos e furtos por meio de modelos lineares generalizados. Os testes de aderência aplicados mostraram que apenas o modelo binomial negativo se ajustou aos dados, que apresentaram um comportamento de sobredispersão. As taxas de roubo/furto (divisão do número de ocorrências pelo tamanho da frota) para os veículos das cores preta, branca, prata e vermelho foram maiores que aquelas das cores azul, verde, marrom e amarela. Entretanto, os resultados do modelo binomial negativo mostraram que, estatisticamente, não há uma cor de automóvel mais ou menos propícia a sofrer um roubo ou furto quando comparada a veículos de coloração branca, cor utilizada como de linha de base comparativa do estudo.

Palavras-chaves: Cor do veículo, Risco, Furto, Roubo, Exposição induzida.

## ABSTRACT

The present study aimed to investigate whether there is a relationship between the color of the vehicle and the risk of a theft or robbery event. The data used were obtained from the Department of Public Security of the State of São Paulo, containing the reports of thefts or thefts of vehicles between 2016 and 2021. As there is no individualized data of vehicles that were not stolen/stolen, the induced exposure method was applied. For the application of the method in this work, the luminosity condition at the time of theft or theft of the vehicle was the factor chosen to define the two comparison groups: a group called "prone to colors", formed by the occurrences in which the natural luminosity was sufficient and, thus, the color of the vehicle could have contributed to the occurrence of a theft or robbery; and another group, called "induced exposure", formed by the records of thefts and robberies that occurred in conditions of low natural light, less conducive to the influence of the color of the automobile due to the lower visibility. The Binomial Negative and Poisson distributions were used to model the number of robberies and thefts using generalized linear models. The adhesion tests applied showed that only the negative binomial model adjusted to the data, which presented an overdispersion behavior. The rates of theft/theft (division of the number of occurrences by the size of the fleet) for the vehicles of the colors black, white, silver and red were higher than those of the colors blue, green, brown and yellow. However, the results of the negative binomial model showed that, statistically, there is no car color more or less conducive to suffering a robbery or theft when compared to white vehicles, a color used as a comparative baseline of the study.

Keywords: Vehicle color, Risk, Theft, Theft, Induced exposure.

## LISTA DE TABELAS

<b>Tabela 1</b> – Informações relativas aos automóveis ou à ocorrência dos registros SPP-SP	15
<b>Tabela 2</b> – Frota automotiva média do estado de São Paulo	16
<b>Tabela 3</b> – Descrição das informações obtidas nos boletins de ocorrência	25
<b>Tabela 4</b> – Resultados dos critérios de seleção de modelos	29
<b>Tabela 5</b> – Estimativas e Intervalos segundo cor do automóvel em relação à cor branca	29
<b>Tabela 6</b> – Estimativas e Intervalos estratificado por período da semana	30
<b>Tabela 7</b> – Estimativas e Intervalos estratificado por ano de fabricação do automóvel	31
<b>Tabela 8</b> – Estimativas e Intervalos estratificado por período do dia	31



## LISTA DE GRÁFICOS

<b>Gráfico 1</b> – Taxa de furtos e roubos segundo cor do automóvel	26
<b>Gráfico 2</b> – Taxa de furtos e roubos por ano ocorrência e cor do automóvel	27
<b>Gráfico 3</b> – Percentual de furtos e roubos por cor do veículo segundo ano de fabricação	27
<b>Gráfico 4</b> – Percentual de furtos e roubos por cor do veículo segundo período da semana	28
<b>Gráfico 5</b> – Percentual de furtos e roubos por cor do veículo segundo período do dia	29

## SUMÁRIO

1. INTRODUÇÃO.....	11
1.1. Objetivos.....	11
2. FUNDAMENTAÇÃO TEÓRICA.....	12
3. METODOLOGIA.....	14
3.1. Tipo de Pesquisa .....	14
3.2. Coleta de Dados .....	15
3.3.1. Método de Exposição Induzido .....	17
3.3.2. Modelagem Estatística de Contagens .....	18
3.3.2.1. Modelo de Regressão Poisson .....	19
3.3.2.2. Modelo de Regressão Binomial Negativo (BN) .....	20
3.3.3. Critérios de Seleção de Modelos .....	21
3.3.4. Formulação de Modelos para Contagens de Roubo e Furto .....	22
4. RESULTADOS .....	24
4.1. Análises Descritivas .....	24
4.2. Modelagem do Risco de Roubo e Furto .....	29
5. CONCLUSÕES E CONSIDERAÇÕES FINAIS.....	32
REFERÊNCIAS BIBLIOGRÁFICAS.....	34

## 1. INTRODUÇÃO

Para a maioria das pessoas, a cor desempenha um papel importante quando se trata de comprar um veículo, roupas ou outros itens, mas muitas vezes ela é escolhida apenas por preferência, normalmente não sendo considerada como um fator de segurança. Ao longo do tempo, foram identificados diversos fatores de risco relativos à segurança contra furtos e roubos de veículos, sendo eles, por exemplo: o local em que o veículo foi estacionado, a existência de rota de fuga, objetos dentro do automóvel, utilização de equipamentos de segurança, etc. No entanto, um aspecto não muito estudado é se a cor de um veículo pode contribuir nas ocorrências de furtos e roubos.

Dados disponibilizados pelo Ministério da Infraestrutura (MI, 2022) informam que, em julho de 2022, a frota de automóveis no Brasil correspondeu a 59 milhões de veículos, sendo que, desse contingente, 33% estavam registrados no estado de São Paulo. Com relação aos casos de furtos e roubos, foram registradas três mil ocorrências em São Paulo no mês de julho de 2022, segundo informações da Secretaria de Segurança Pública (SSP/SP, 2022). Isso significa que, em média, quatro carros foram roubados por hora no estado. Dessa forma, é possível notar que São Paulo possui grande peso no cenário automotivo nacional e carece que cidadãos e operadoras de seguro efetuem medidas para que possam contornar os riscos existentes relacionados a roubos e furtos de automóveis.

Atualmente, não parece haver uma preocupação de compradores de veículos e seguradoras quanto a uma possível contribuição da cor do veículo para o risco de roubo ou furto. Possíveis perguntas sobre esse tema podem interessar consumidores e entidades supervisionadas a fim de mitigá-lo, como por exemplo: se a cor interfere no risco de o veículo ser roubado ou furtado, qual cor de veículo pode refletir mais segurança e se os prêmios de seguro deveriam depender disso. Nesse sentido, este trabalho se concentra em investigar se há uma relação entre a coloração automotiva e o risco de roubo ou furto.

### 1.1. Objetivos

O objetivo geral deste trabalho consistiu em investigar a possível relação entre a cor do veículo e o risco de roubo e furto no estado de São Paulo, no período de 2016 a 2021. Já em relação aos objetivos específicos, pretendeu-se analisar a ocorrência de roubo ou furto em função do ano de fabricação do automóvel, período do dia e dia da semana.

## 2. FUNDAMENTAÇÃO TEÓRICA

Poucas pesquisas conclusivas foram realizadas associando e identificando a cor do automóvel como fator de risco. A maioria delas foram realizadas na tentativa de estabelecer se há correlação entre a cor e o risco de acidentes. Os pesquisadores tentaram descobrir se existem veículos com uma cor específica que são mais propensos a acidentes. Embora os estudos sejam conclusivos, algumas de suas conclusões se contradizem. Pesquisadores como Nathan (1969), Furness et al. (2003), Newstead e D'Elia (2007) e Owusu-Ansah (2010) conduziram estudos para entender a relação entre cor e visibilidade.

O interesse em descobrir relações referentes a cor do veículo como fator de risco começou em 1969. Nathan (1969) foi um dos primeiros estudiosos a tentar estabelecer se alguns veículos são mais seguros do que outros por causa de sua cor e sua conspicuidade. Em seu estudo, observou que a coloração de um veículo pode ter um efeito relativo a problemas de visibilidade para o tipo de colisão traseira. Citando uma pesquisa realizada na Universidade da Califórnia em Los Angeles, Nathan (1969) observou que a cor afeta a capacidade de se perceber se um veículo está se aproximando. Essa pesquisa descobriu que os veículos de cor azul e amarelo parecem mais próximos, enquanto os veículos de cor cinza pareciam estar mais distantes do que sua localização real. Nathan (1969) concluiu que, de acordo com o estudo desempenhado, a cor mais segura seria aquela cor altamente visível sob diferentes condições climáticas e de iluminação. No entanto, observou a partir de uma pesquisa realizada na Suécia, onde os pesquisadores analisaram 31.000 acidentes de veículos, que os veículos pretos, que representavam apenas 4,4% dos veículos pesquisados, são os mais inseguros, uma vez que 22,5% dos acidentes de veículos envolveram veículos pretos. Por outro lado, embora a porcentagem dos veículos cor de rosa não tenha sido indicada, os veículos cor de rosa foram considerados os mais seguros, uma vez que estavam envolvidos em apenas 2,4% dos acidentes (Nathan, 1969).

Furness et al. (2003) realizou um estudo em Auckland, Nova Zelândia para analisar o efeito da cor do veículo sobre o risco de um acidente mortal por meio do modelo retrospectivo de pesquisa de caso-controle. O grupo de casos era composto por condutores de veículos vítimas de acidentes onde um ou mais ocupantes do veículo foram gravemente feridos ou submetidos a ferimentos. Por outro lado, o grupo controle era composto por condutores de veículos escolhidos aleatoriamente em horários aleatórios. A pesquisa utilizou análises multivariadas ajustadas por diversos fatores, como sexo do motorista, idade, escolaridade, consumo de álcool,

idade do veículo, velocidade do veículo e uso do cinto de segurança. Eles descobriram que os veículos de cor prata eram os mais seguros.

Além disso, Newstead e D'Elia (2007) realizaram um estudo sobre a relação entre a cor do veículo e o risco de acidente usando abordagens estatísticas de análise de dados e um método de exposição induzido. A partir de suas pesquisas, determinaram uma relação estatisticamente significativa entre a cor do veículo e o risco de colisão. Vários veículos estavam ligados ao maior risco de colisão em comparação com veículos brancos. Essas cores, de acordo com Newstead e D'Elia (2007), têm um índice de visibilidade geralmente menor e incluem preto, azul, verde, cinza, prata e vermelho. Seu estudo descobriu que nenhuma cor era estatisticamente significativamente mais segura que o branco, embora algumas cores não pudessem ser diferenciadas do branco estatisticamente em termos de seu risco relativo de colisão. Além disso, Newstead e D'Elia (2007) descobriram que havia relação entre a cor do veículo e o risco de acidente de forma mais significativa, cerca de 10%, quando a luz do dia estava presente. Por fim, observaram que a relação entre a cor do veículo e o acidente de carro poderia ser influenciada por fatores ambientais (Newstead e D'Elia, 2007).

Owusu-Ansah (2010) em seu estudo de tese realizou uma pesquisa para investigar a relação entre a cor de um veículo e o risco de acidentes. Ele utilizou o procedimento de exposição induzido semelhante ao utilizado no estudo de Newstead e D'Elia (2007), mas Owusu-Ansah usou tanto os modelos Poisson quanto o binômio negativo (BN) para encaixar os dados, enquanto Newstead e D'Elia (2007) usaram apenas o modelo Poisson. Owusu-Ansah (2010) usou dados de acidentes do estado de Nevada para o período de 2003 a 2008. Seu estudo descobriu que o modelo BN se encaixava melhor nos dados do que o modelo Poisson devido à grande dispersão nos dados e concluiu que nenhuma cor do veículo era estatisticamente mais segura ou arriscada do que a coloração branca.

Por fim, Alanazi (2016) realizou estudo com objetivo semelhante a Owusu-Ansah (2010). Em seu trabalho, Alanazi (2016) utilizou dados de acidentes de trânsito que ocorreram no estado de Ohio de 2011 a 2015. No estudo, também inferiu que o modelo BN se encaixava melhor nos dados do que o modelo Poisson e, estatisticamente, além de variações aleatórias, não foi encontrada cor do veículo mais segura ou arriscada que a branca, a cor utilizada como cor de linha de base do modelo.

Tendo analisado a literatura sobre cor do veículo e risco de acidente, pode-se notar que, embora vários pesquisadores tenham tentado esclarecer o mistério entre a cor do veículo e o risco de acidentes, alguns dos achados da literatura se contradizem. Por exemplo, o estudo de Newstead e D'Elia (2007) observou que os veículos de cor prata não são seguros, mas o estudo

da Furness et al. (2003) descobriu que os veículos prateados são mais seguros que os veículos com outras cores. Por fim, Owusu-Ansah (2010) e Alanazi (2016) observaram que nenhuma cor do veículo é mais segura do que outras em termos de risco de acidente.

Mesmo os trabalhos relatados sendo investigações referentes ao risco de acidentes, é possível abstrair deles inspiração e orientação na investigação quanto ao risco de furto e roubo, uma vez que possuem recursos de dados semelhantes, baseados em contagem. Esse tipo de dado resume-se a eventos que ocorrem em uma determinada taxa, nos quais as observações podem somente tomar valores inteiros não negativos. Nesse sentido, esse trabalho se propõe a investigar a relação entre a cor do veículo e o risco de furto e roubo, usando abordagens estatísticas e análise de dados.

### **3. METODOLOGIA**

#### **3.1. Tipo de Pesquisa**

Quanto aos objetivos, esta pesquisa pode ser classificada como descritiva, uma vez que descreve características relacionadas ao risco de furto ou roubo automotivo, assim com classifica a coloração dos automóveis como “maior risco”, “menor risco” ou “não significativo”, com base nas características apresentadas. De acordo com Gil (2002), a pesquisa descritiva tem como principal objetivo a descrição das características de determinada população ou fenômeno, ou então, o estabelecimento de relações entre as variáveis. Ela mede, avalia ou coleta dados sobre diversos aspectos, dimensões ou componentes do fenômeno a ser pesquisado.

Quanto aos procedimentos, pode ser caracterizada como uma pesquisa documental, uma vez que se utiliza de boletins de ocorrência divulgados pela Secretaria de Segurança Pública de São Paulo, e número de veículos em circulação divulgados pelo Departamento Estadual de Trânsito de São Paulo para a realização das análises. Segundo Gil (2002), a pesquisa documental baseia-se em materiais que ainda não receberam um tratamento analítico ou que podem ser reelaborados de acordo com os objetivos da pesquisa.

Quanto à abordagem, trata-se de um estudo quantitativo pois, emprega quantificação a partir do momento que os dados abordam eventos que ocorrem em uma determinada taxa, permitindo utilização de ferramentas para consolidação de variáveis de forma a gerar uma métrica única, capaz de representar a complexidade da análise de desempenho e efeito de

diversas variáveis. De acordo com Richardson (1985), a pesquisa quantitativa caracteriza-se pelo emprego de quantificação tanto nas modalidades de coleta de informações, quanto no tratamento delas através de técnicas estatísticas, desde as mais simples como percentual, média, desvio-padrão, às mais complexas, como coeficiente de correlação, análise de regressão etc.

### 3.2. Coleta de Dados

Esta pesquisa foi desenvolvida a partir de duas bases de dados do estado de São Paulo, uma proveniente da Secretaria de Segurança Pública (SSP/SP) e outra do Departamento Estadual de Trânsito (DETRAM/SP), compreendidas entre os anos de 2016 e 2021.

A Secretaria de Segurança Pública de São Paulo disponibiliza, em seu banco de dados, registros de boletins de ocorrência de furto ou roubo de veículos. Foram mantidas apenas as ocorrências de veículos caracterizados com o tipo “automóvel”, sendo retiradas, por exemplo, ocorrências do tipo “motocicleta” e “caminhão”. Além disso, todos os casos com registros ausentes ou incompletos para variáveis que continham o estado da ocorrência e a cor do automóvel também foram removidos. Como resultado, os dados consistiam em 721.231 registros de roubos ou furtos de automóveis.

A Tabela 1 lista as informações retiradas dos boletins de ocorrência sobre os automóveis envolvidos e características da ocorrência. Devido ao número relativamente grande de variações de coloração, foram definidas oito categorias com cores mais frequentes e uma categoria agrupando as cores menos frequentes. Além disso, veículos classificados com as cores “bege” e “dourado” foram reclassificados para “amarelo”; e aqueles classificados como “cinza” foram reclassificados para “prata”.

Tabela 1- Informações relativas aos automóveis ou à ocorrência dos registros SPP-SP

<b>Variável</b>	<b>Categorias</b>
Ano	2016
	2017
	2018
	2019
	2020
	2021
Cor do automóvel	Amarelo (inclui bege e dourado)
	Azul
	Branco
	Marrom

	Prata (inclui cinza)
	Preta
	Verde
	Vermelho
	Demais cores
Período do dia	A Noite
	À Tarde
	De Madrugada
	Em Hora Incerta
	Pela Manhã
Dia da semana	Segunda-feira
	Terça-feira
	Quarta-feira
	Quinta-feira
	Sexta-feira
	Sábado
	Domingo
Tipo de BO	Roubo
	Furto
Ano de fabricação do automóvel	Anterior a 2013
	2013
	2014
	2015
	2016
	2017
	2018
	2019
	2020
	2021

Fonte: SPP-SP (2022)

Os dados do DETRAM/SP fornecem informações quantitativas da frota veicular em circulação no estado de São Paulo agrupada por ano, mês, cor e tipo (Tabela 2). Em média, 18,6 milhões de automóveis compõem a frota veicular do estado por ano. Com relação as cores, 38% da frota é composta por automóveis de coloração prata, 16,8% de cor preta e 15,5% branca, isso implica que quase três quartos dos veículos paulistas pertencem a uma dessas três cores.

Tabela 2 - Frota automotiva do estado de São Paulo

Variável	Parâmetro	Frota	Percentual
Ano	2016	17.680.290	-
	2017	18.035.686	-
	2018	18.533.509	-
	2019	19.054.875	-
	2020	19.326.890	-
	2021	19.285.143	-



	<b>Média</b>	<b>18.652.732</b>	<b>-</b>
Cor	Amarelo	957.828	5,1%
	Azul	1.416.866	7,6%
	Branco	2.892.749	15,5%
	Marrom	236.506	1,3%
	Prata	7.085.077	38,0%
	Preta	3.139.143	16,8%
	Verde	1.028.719	5,5%
	Vermelho	1.839.269	9,9%
	Demais cores	56.370	0,3%
	Não Informado	205	0,01%
	<b>Total</b>	<b>18.652.732</b>	<b>100%</b>

Fonte: DETRAM.SP (2022)

### 3.3. Análise Estatística

Inicialmente são apresentadas as distribuições de frequência, tanto absolutas quanto relativas, das variáveis relacionadas ao boletim de ocorrência de roubo ou furto, conforme listado na Tabela 1. Em seguida, realiza-se uma análise descritiva para investigar a influência da cor do automóvel na ocorrência de roubo e furto. Essa análise é conduzida por meio do cálculo da taxa de roubo e furto para as variáveis cor do veículo e ano, obtida pela divisão do número de roubos e furtos de automóveis de determinada cor pelo tamanho da frota veicular correspondente. As taxas de roubo e furto são comparadas entre as diferentes cores, tanto nos dados gerais como dentro de grupos de estratificação definidos pelas demais variáveis mencionadas na Tabela 1.

#### 3.3.1. Método de Exposição Induzido

Uma das principais fraquezas identificadas nos desenhos do estudo de caso-controle realizado por Furness et al. (2003) foi a limitação na gama de fatores de confusão mensuráveis que poderiam ser controlados na análise. A falta de medição e combinação dos principais fatores que influenciam o risco de acidente no projeto de caso-controle pode resultar em um viés significativo na associação de medidas entre a cor do veículo e o risco de acidente.

O método de exposição induzida tem sido utilizado por pesquisadores como uma abordagem para contornar situações semelhantes à descrita anteriormente. Essa abordagem oferece um potencial significativo para avaliar de forma mais definitiva a relação entre a cor do veículo e a variável de interesse, permitindo um controle intrínseco sobre os fatores de

confusão, inclusive aqueles que geralmente não são mensuráveis. Burton et al. (2004) utilizaram o método para avaliar mudanças no risco de acidente associado aos freios ABS no veículo analisado. Owusu-Ansah (2010) e Alanazi (2016) também utilizaram o método para estudar a relação entre a cor do veículo e seu risco de acidente; esses dois estudos desempenharam um papel motivador para o estudo atual.

Para implementar o método de exposição induzido, é necessário classificar os dados em dois grupos. O primeiro é o grupo no qual as ocorrências de furtos e roubos estão mais propensas à coloração do veículo; já o segundo grupo, denominado exposição induzido, é composto pelas ocorrências menos propensas ao fator cor. A definição desses grupos é descrita a seguir.

- i. **Grupo de ocorrências propensas a cores:** Este grupo constitui diferentes tipos de ocorrências onde a visibilidade do veículo em termos de cor pode contribuir para a ocorrência de um furto ou roubo. Nesse sentido, são furtos e roubos de vários veículos, mas apenas aqueles que ocorreram no período de maior luminosidade natural (pela manhã e a tarde) e, portanto, a maior conspicuidade pode ter desempenhado um papel na chance de um furto ou roubo de acordo com a cor do automóvel.
- ii. **Grupo de exposição induzida:** Este grupo envolve roubos e furtos de veículos que ocorreram no período naturalmente menos luminoso (à noite e de madrugada). Por haver menor conspicuidade, ocorrências podem ter ocorrido por outras razões que não envolveram a visibilidade do veículo.

### 3.3.2. Modelagem Estatística de Contagens

As distribuições de probabilidade Poisson e Binomial Negativa são as mais utilizadas para modelar dados de contagem de algum evento durante determinado intervalo de tempo. Um Modelo Linear Generalizado (MLG) é usado nestes casos para verificar fatores que podem afetar a frequência de ocorrências do evento de interesse.

Pesquisas que serviram como fundamentação teórica para este trabalho utilizaram o modelo de regressão de Poisson em seus estudos para identificar a significância da variável cor na ocorrência de acidentes de trânsito, como por exemplo, o trabalho de Newstead e D'Elia

(2007). No entanto, as pesquisas conduzidas por Owusu-Ansah (2010) e Alanazi (2016) apontaram que as conclusões derivadas do trabalho de Newstead e D'Elia (2007) não foram satisfatórias, devido aos dados não apresentarem uma relação de variância igual à média, requisito exigido pelo modelo de regressão de Poisson.

Este critério bastante restrito geralmente não é satisfeito por dados do mundo real. Frequentemente, a variância é maior do que a média, uma propriedade chamada sobredispersão, e às vezes a variância é menor que a média, chamada de subdispersão. Em tais casos é necessário usar um modelo de regressão que não fará a suposição de equidispersão, ou seja, não assume que variância possui o mesmo valor da média.

O modelo de regressão Binomial Negativo (BN) é reconhecido como um modelo eficaz para lidar com a sobredispersão. Devido a essa capacidade, estudos mais recentes têm preferido sua aplicação como uma abordagem promissora. Trabalhos que também estudavam a relação da cor com acidentes de trânsito, utilizaram tanto o modelo de regressão Binomial Negativo quanto o modelo Poisson (Owusu-Ansah, 2010 e Alanazi, 2016), constatando que o modelo de distribuição de regressão BN oferece um ajuste melhor quando comparado ao modelo Poisson.

A distribuição binomial negativa é obtida aplicando o teorema do binômio com um expoente negativo. Esse modelo é comumente chamado de modelo Poisson-Gama, pois pode ser considerado uma combinação das distribuições Poisson e Gama. Ao introduzir um efeito aleatório na média da distribuição Poisson, o parâmetro da Poisson é assumido como sendo distribuído segundo a distribuição Gama (Agresti, 2018).

Serão apresentadas e discutidas as formulações matemáticas dos modelos de regressão Poisson e Binomial Negativo. A ênfase na apresentação dos modelos será em aprofundar suas características, permitindo uma compreensão mais ampla e aprofundada dos mesmos, bem como o entendimento dos motivos à sua utilização.

### 3.3.2.1. Modelo de Regressão Poisson

A regressão de Poisson é um modelo linear aleatório onde a variável de resposta é assumida como tendo distribuição Poisson. A probabilidade de uma variável de contagem  $Y$  pode ser descrita pela Equação 1 :

$$P_r\{Y = k\} = \frac{e^{-\mu} \mu^k}{k!}, \quad (1)$$

sendo  $\mu > 0$ ;  $\mu = E(Y) = Var(Y)$ ;  $k = 0,1,2, \dots$

A relação entre a média e a variância é determinante na modelagem de dados utilizando o modelo de regressão Poisson. Conforme apontado por Larget (2007), quando a média é igual à variância, quaisquer parâmetros que afetem a média também afetarão a variância. No caso do modelo Poisson, as variáveis explicativas influenciam o valor esperado da variável resposta, e como a média deve ser positiva, enquanto a combinação linear das variáveis explicativas pode assumir qualquer valor, é necessária a utilização de uma função de ligação. A função de ligação comumente adotada é o logaritmo natural, como destacado por Rodriguez (2007).

Considere uma amostra de  $n$  observações,  $y_1, y_2 \dots y_n$  que podem ser consideradas como concepções de variáveis aleatórias independentes seguindo uma distribuição Poisson, com  $Y_i \sim P(\mu_i)$ . Se a média  $\mu_i$ , que também é igual a variância, depende de um vetor de variáveis explicativas  $x_i$ , é possível desenvolver um modelo linear simples da forma mostrada na Equação 2:

$$\mu_i = x_i' \beta \quad (2)$$

No entanto, esse modelo possui uma limitação, pois o estimador linear, à direita, pode tomar qualquer valor real, enquanto a média Poisson, à esquerda, que representa a contagem esperada, não pode assumir valores negativos. Para resolver esse problema, recorreremos à transformação logarítmica. Realizamos a transformação logarítmica  $\eta_i = \log(\mu_i)$  e assumimos que a média transformada segue um modelo linear, expresso como  $\eta_i = \beta_0 + \beta_1 x_1 + \dots \beta_k x_k$ . Isso pode ser representado na Equação 3:

$$\log(\mu) = \beta_0 + \beta_1 x_1 + \dots \beta_k x_k \quad (3)$$

Dessa forma, o uso da função logarítmica na média transformada permite a modelagem de forma linear, evitando valores negativos e garantindo que a relação entre as variáveis explicativas e a média da distribuição Poisson seja apropriada.

### 3.3.2.2. Modelo de Regressão Binomial Negativo (BN)

Se o fenómeno de sobredispersão for identificado nos dados, a regressão Binomial Negativa é normalmente usada ao invés de Poisson. O modelo BN também é conhecido como modelo Poisson-Gama sendo semelhante ao modelo Poisson, no qual a variável dependente  $Y$  é modelada como uma mistura de variáveis Poisson com uma média  $\mu$  assumida para seguir uma distribuição Gama.

O modelo Poisson-Gamma é uma combinação de duas distribuições e é usado para explicar a sobredispersão que geralmente é observada na utilização de conjuntos de dados de contagem (Lord et al., 2005). É importante destacar que a distribuição Binomial Negativa é adequada, em especial, para dados heterogêneos, nos quais a taxa de ocorrência dos eventos pode variar. Ela é utilizada para modelar situações em que ocorre uma contagem de eventos até que um determinado número de sucessos seja alcançado. A média e a variância de uma distribuição Binomial Negativa podem ser apresentadas como mostrado na Equação 4:

$$E(Y) = \mu, \quad Var(Y) = \mu + k\mu^2 \quad (4)$$

onde  $k =$  *parâmetro de dispersão não negativo*.

### 3.3.3. Critérios de Seleção de Modelos

Um dos problemas da subestimação dos erros padrão é a atribuição ao teste de significância, nos testes de hipóteses, de mais relevância do que realmente retrata. Nesse sentido, um método de deteção da sobredispersão nos dados é o uso da Equação 5:

$$k = \frac{D}{DF} \quad (5)$$

onde  $k =$  *parâmetro de dispersão não negativo*;  $D =$  *deviance*;  $DF =$  *grau de liberdade*.

Se  $k$  é maior que um, há indícios que os dados estão sobredispersos, o que significa que a variância é maior do que a média; por outro lado, se  $k$  for menor que um, isso significa

subdispersão nos dados. Caso seja observado qualquer um desses casos, não podemos submeter os dados ao modelo Poisson visto na Equação 1.

O modelo Binomial Negativo é o modelo mais indicado caso seja observado sobredispersão ou subdispersão em dados de contagem. A qualidade do ajuste entre os dados observados e os valores estimados de uma distribuição Binomial Negativa ou uma distribuição de Poisson é geralmente determinada usando a estatística log-verossimilhança  $G^2$  como mostrado na Equação 6:

$$G^2 = 2 \sum \text{observado} \times \frac{\log(\text{observado})}{\text{esperado}} \quad (6)$$

### 3.3.4. Formulação de Modelos para Contagens de Roubo e Furto

O modelo utilizado neste trabalho é baseado nos procedimentos desenvolvidos por Newstead e D'Elia (2007) e estendido por Alanazi (2016). A cor branca é definida como linha de base com a qual as demais cores são comparadas em termos de risco de furto ou roubo. O teste de hipóteses é definido da seguinte forma:

- I. Hipótese nula ( $H_0$ ): não há diferença no risco de ocorrência de furtos ou roubos atribuíveis à cor do veículo em relação à cor branca;
- II. Hipótese alternativa ( $H_1$ ): há diferença no risco de furto ou roubo atribuível à cor do veículo em relação à cor branca.

O modelo linear logarítmico seguindo uma distribuição Poisson ou Binomial Negativa é expresso como mostrado na Equação 7:

$$\text{Ln}(\eta_{jc}) = \alpha + \beta_j + \gamma_{jc} + \zeta_{jcl} + \delta_{jca} + \varepsilon \quad (7)$$

onde:

- $\eta$  = contagem de ocorrências de roubo ou furto;
- $\alpha, \beta, \gamma, \zeta$  e  $\delta$  = coeficientes (efeitos) a serem estimados;
- $\varepsilon$  = erro aleatório;

- $j$  = índice para o grupo comparação (*propenso à cor* = 0 ou *exposição induzida* = 1 );  
 $c$  = índice para a cor do veículo;  
 $l$  = índice para o dia da semana da ocorrência do roubo ou furto;  
 $a$  = índice para o ano de fabricação do automóvel.

A Equação 7 pode ser modificada para proporcionar um efeito médio ao longo do dia da semana da ocorrência e do ano de fabricação. De acordo com Newstead & D'Elia (2007), o parâmetro de interação com quatro fatores pode então ser reduzido a interações de três ou dois fatores, considerando, três ou dois parâmetros apenas da seguinte forma:

- Risco geral de furto ou roubo,  $\gamma_{jcla}$  torna-se um termo de interação com dois fatores,  $\gamma_{jc}$
- Risco de furto ou roubo pelo período da semana da ocorrência,  $\gamma_{jcla}$  torna-se um termo de interação com três fatores,  $\gamma_{jcl}$
- Risco de furto e roubo pelo ano de fabricação do veículo,  $\gamma_{jcla}$  torna-se um termo de interação com três fatores,  $\gamma_{jca}$

O grupo de furto e roubo propenso a cores pode ser estimado referenciando a cor branca como linha de base ( $c = 0$ ) e tomando o grupo de exposição induzida ( $j = 1$ ) também como linha de base. Portanto, a razão de risco relativa  $\rho_i$ , representando o risco de furto ou roubo de um veículo atribuível à sua cor, pode ser expressa como mostrado na Equação 8:

$$\rho_i = e^{\gamma_{jc}} \quad (8)$$

Em seguida, foram calculados os intervalos de confiança de 95%, ( $\alpha = 0,05$ ), para a razão de risco relativa de cada cor. Esses intervalos foram estimados utilizando o método demonstrado pela Equação 9:

$$\ln(\rho_i) = \pm Z_{\frac{\alpha}{2}}(SE) \quad (9)$$

onde:

$Z_{\frac{\alpha}{2}}$  = valor da distribuição normal padrão com nível de confiança  $(1-\alpha)$

$SE$  = erro padrão estimado.

A Equação 9 é aplicada sob a suposição de que a distribuição da interação parâmetros é normalmente distribuída, com média zero e variância unitária, o logaritmo natural deve ser calculado para atingir a normalidade (Agresti, 2007). Assim, espera-se que: uma cor com razão da taxa de risco de furto ou roubo menor que um, apresente um risco de roubo ou furto menor do que a cor branca; uma cor com razão da taxa de risco de maior que um, espera-se que tenha maior risco de furto ou roubo do que a cor branca.

A seguir, serão apresentados testes de ajuste dos dados para determinar, entre os modelos Poisson e Binomial Negativo, qual se encaixa melhor, o modelo que apresentar melhor ajuste será utilizado no teste de hipóteses. Além disso, serão apresentados os resultados dos riscos estimados de furtos e roubos por cor, juntamente com os intervalos de confiança e o valor  $p$ , permitindo verificar, ao compararmos com a cor branca, se os resultados são estatisticamente significativos e, sustentando ou não, a hipótese de que há uma relação entre a cor do veículo e o risco de furto e roubo, que é o escopo deste trabalho.

Os softwares Microsoft Excel®, Alteryx® e Python® (pacote *Statsmodels*) foram utilizados como ferramentas para extração, transformação, carregamento dos dados e aplicação dos modelos GLM propostos.

## 4. RESULTADOS

### 4.1. Análises Descritivas

As informações na Tabela 3 revelam que 59,4% das ocorrências com automóveis no estado de São Paulo entre 2016 e 2021 selecionadas no estudo foram associadas a casos de roubos e 40,6% a furtos.

Houve uma queda no número de ocorrências de 2016 a 2021, com destaque para os dois últimos anos, para a qual uma das explicações pode estar associada ao isolamento social imposto pela pandemia de Covid 19 (Tabela 3). Nesse período, o governo de São Paulo promoveu políticas que restringiam o contato social para desestimular a transmissão viral, dessa forma havia menos veículos em circulação expostos ao risco de sofrer um furto ou um roubo.



Durante o período analisado, verificou-se que a maioria das ocorrências de roubo ou furto ocorreram durante a noite, representando 34,6% do total, enquanto o menor número de casos aconteceu durante a madrugada, com apenas 14,7%. Em relação aos dias da semana, constatou-se que a maioria dos casos foram entre segunda e quinta-feira, abrangendo 61,3% do total, enquanto o restante dos casos, 38,7%, aconteceram durante os finais de semana (sexta, sábado e domingo).

Quanto ao ano de fabricação do automóvel, encontrou-se que 68% dos registros de roubo ou furto foram relacionados a automóveis fabricados anteriormente a 2013 e que o número de registros é menor para automóveis mais recentes (Tabela 3), podendo essa queda estar relacionada ao menor número de automóveis mais novos em circulação.

Ocorrências envolvendo veículos de cor prata, preto ou branco são as mais frequentes, representando 41%, 20% e 17% dos casos, respectivamente (Tabela 3). Como mostrado na Tabela 2, essas três cores possuem uma grande relevância na frota automotiva do estado; logo, não é de surpreender ver que as cores mais populares também sejam a maioria dos veículos envolvidos em roubos ou furtos.

Tabela 3 – Descrição das informações obtidas nos boletins de ocorrência por motivos de furto ou roubo no estado de São Paulo segundo variáveis selecionadas no estudo (n=721.231)

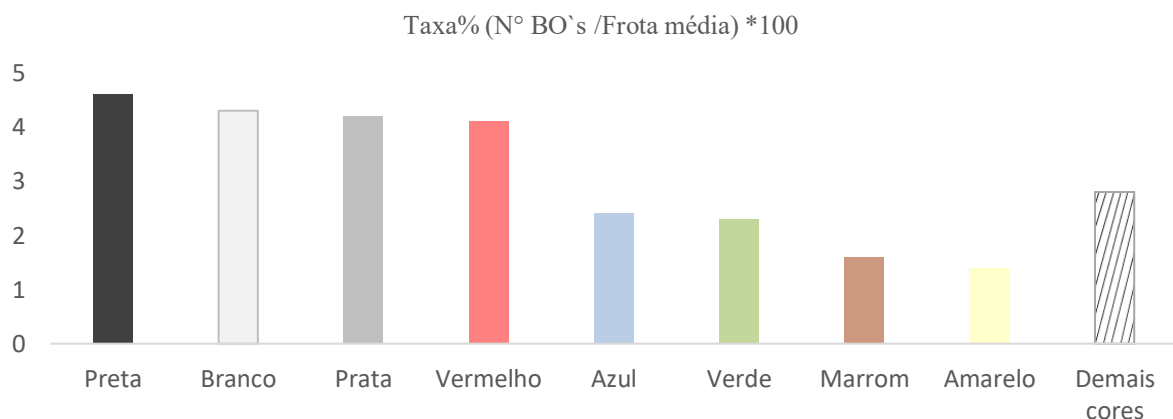
<b>Variável</b>	<b>Categorias</b>	<b>Frequência</b>	<b>Percentual</b>
Tipo de BO	Roubo	428.282	59,4
	Furto	292.949	40,6
Ano de registro da ocorrência	2016	161.590	22,4
	2017	146.585	20,3
	2018	130.929	18,2
	2019	118.188	16,4
	2020	80.132	11,1
	2021	83.807	11,6
Período do dia	À noite	249.629	34,6
	Pela manhã	178.210	24,7
	À tarde	154.476	21,4
	De madrugada	105.812	14,7
	Em hora incerta	33.104	4,6
Período da semana	Segunda à quinta	442.185	61,3
	Sexta à domingo	279.046	38,7
Ano Fabricação	Anterior a 2013	454.318	63,0
	2013	59.930	8,3
	2014	56.078	7,8
	2015	41.852	5,8
	2016	31.333	4,3
	2017	30.184	4,2
	2018	23.133	3,2

	2019	16.658	2,3
	2020	5.733	0,8
	2021	2.012	0,3
Cor	Prata	300.141	41,6
	Preta	144.048	20
	Branco	124.784	17,3
	Vermelho	74.633	10,3
	Azul	34.524	4,8
	Verde	24.061	3,3
	Amarelo	13.534	1,9
	Marrom	3.716	0,5
	Demais cores	1.790	0,2

Fonte: SSP/SP (2022)

O Gráfico 1 apresenta, separadamente para cada cor de automóvel, a *taxa de furto e roubo*, calculada pela divisão entre o número de boletins de ocorrência e a frota veicular média em cada cor. As taxas de roubo/furto para os veículos das cores preta, branca, prata e vermelho foram maiores que aquelas das cores azul, verde, marrom e amarela. Observa-se que, mesmo a cor prata tendo os maiores números absolutos de boletins de ocorrência e frota veicular, é apenas a terceira com a maior taxa dentre as cores classificadas. Veículos que possuem coloração vermelha apresentaram taxas de furto e roubo compatíveis com as cores mais comuns de automóveis (branco, preto e prata).

Gráfico 1- Taxa de furtos e roubos segundo cor do automóvel

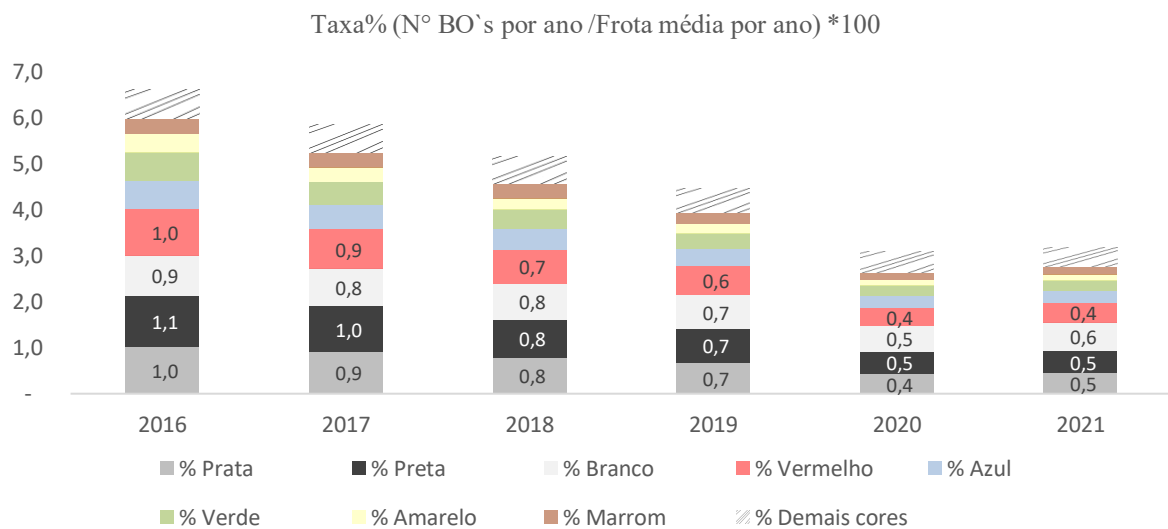


Fonte: Elaboração Própria

O Gráfico 2 ilustra a evolução da taxa de furto e roubo ao longo do tempo, considerando também a cor dos veículos. A análise dessa taxa também foi obtida através da divisão do número

de boletins de ocorrência registrados pela frota veicular média, porém correspondente a cada ano. Ao examinarmos o gráfico, é possível constatar uma tendência decrescente das taxas ao longo dos anos, com uma redução significativa e já esperada no número de ocorrências nos anos de 2020 e 2021. No que diz respeito à variação das taxas, é interessante notar que os veículos da cor branca apresentaram a menor queda ao longo do tempo, por outro lado, os veículos de cor preta tiveram a maior redução na taxa, indicando menores ocorrências desse tipo de crime em relação a essa cor do automóvel ao longo do tempo.

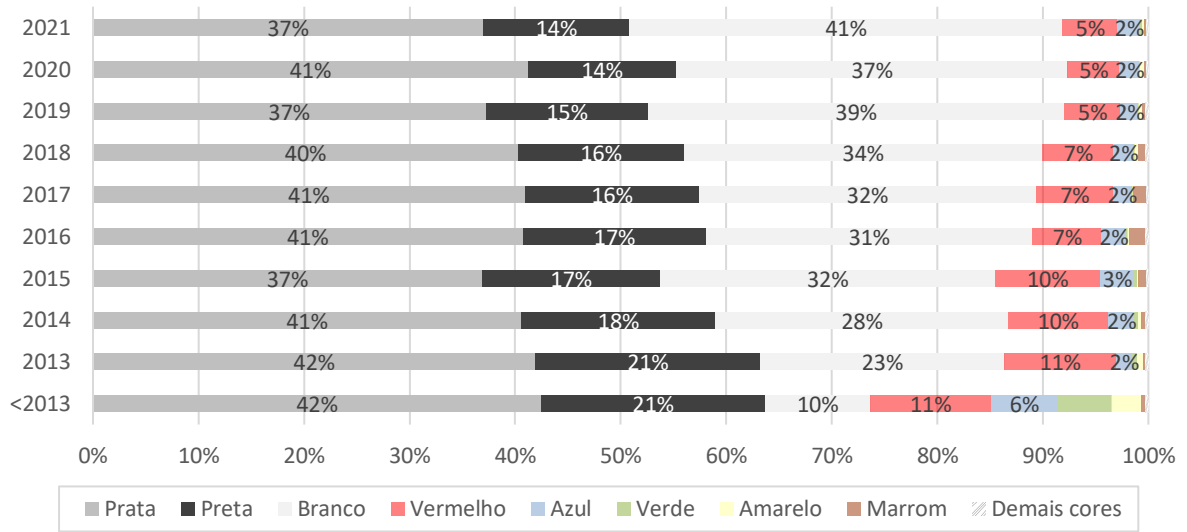
Gráfico 2 – Taxa de furtos e roubos por ano ocorrência e cor do automóvel



Fonte: Elaboração Própria

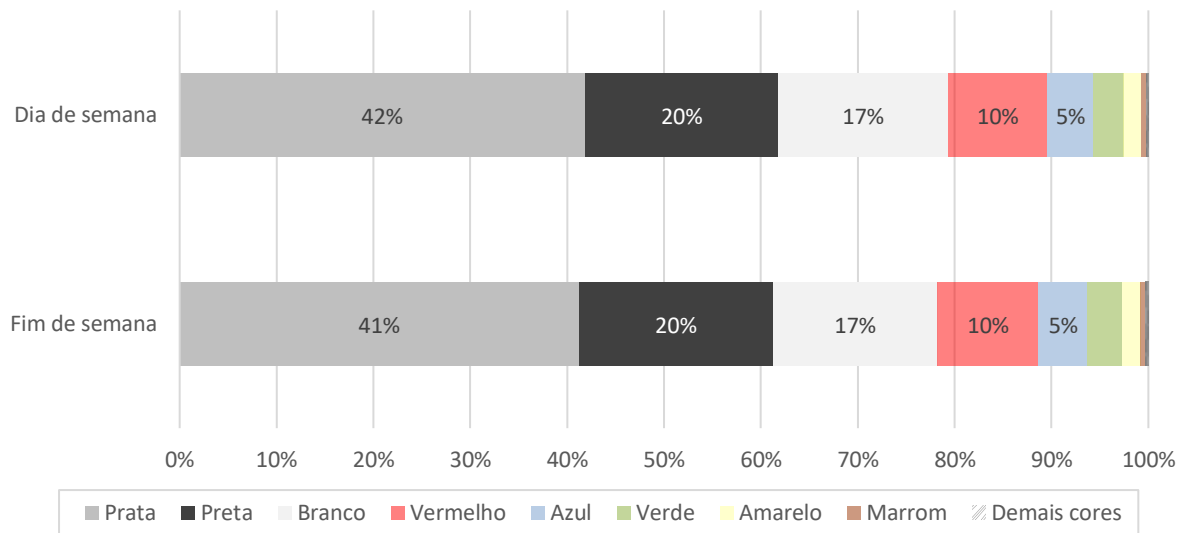
Os gráficos 3 a 5 apresentam a distribuição das ocorrências, em termos absolutos, de roubo e furto segundo a cor do veículo, estratificada pelo ano da ocorrência e da fabricação do veículo, período da semana e do dia. A distribuição percentual das ocorrências entre as cores dos veículos aparenta ser influenciada apenas pelo período do dia. Durante a noite e a madrugada, quando a luminosidade é menor, observa-se uma distribuição de cores diferente em comparação aos períodos de maior luminosidade, ou seja, pela manhã e tarde.

Gráfico 3 – Percentual de furtos e roubos por cor do veículo segundo ano de fabricação



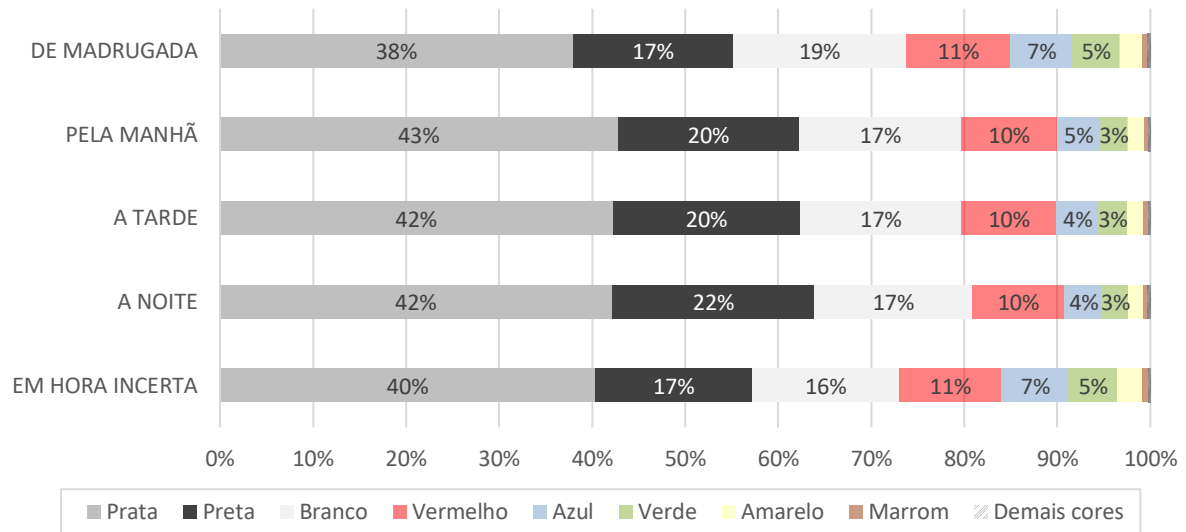
Fonte: Elaboração Própria

Gráfico 4 - Percentual de furtos e roubos por cor do veículo segundo período da semana



Fonte: Elaboração Própria

Gráfico 5 – Percentual de furtos e roubos por cor do veículo segundo período do dia



Fonte: Elaboração Própria

#### 4.2. Modelagem do Risco de Roubo e Furto

A Tabela 4 apresenta os resultados dos ajustes dos modelos Poisson e Binomial Negativo. Observa-se que ambos valores  $D/GL$  e  $X^2/GL$  são muito maiores que um no modelo Poisson, indicando sobredispersão nos dados, ou seja, a variância não é igual à média, como definido no modelo de Poisson. No entanto, no modelo Binomial Negativo, essas medidas estão mais próximas de um, indicando que esse modelo se ajusta melhor aos dados quando comparado ao modelo Poisson. Nesse sentido, uma vez que, dentre os dois modelos considerados, somente o Binomial Negativo mostrou-se adequado para modelagem dos dados, apenas seus resultados são apresentados a seguir.

Tabela 4 - Resultados dos critérios de seleção de modelos.

Parâmetro	Modelo Poisson	Modelo Binomial Negativo
Deviance (D)	14.207	2.579
Qui-Quadrado de Pearson ( $X^2$ )	14.709	2.629
Grau de Liberdade (GL)	2.182	2.182
$D/GL$	6,511	1,182
$X^2/GL$	6,741	1,205

Fonte: Elaboração Própria

A Tabela 5 mostra as razões de risco relativas para um modelo Binomial Negativo de cada cor dos veículos em comparação com as ocorrências de veículos brancos. Há de se destacar que todas cores têm risco de roubo ou furto relativos à cor branca próximo de 1, sendo todos não estatisticamente significativos. Como resultado, não se rejeita a hipótese nula de que não há diferença no risco de ocorrência em furtos ou roubos atribuíveis à cor do veículo em relação ao veículo branco.

**Tabela 5 – Estimativas pontuais e Intervalo de Confiança de 95% do Risco Relativo de roubo ou furto estimados no modelo Binomial Negativo, segundo cor do automóvel em relação à cor branca.**

Cor	Risco Relativo	Intervalo de Confiança de 95%		Valor p
		Limite Inferior	Limite Superior	
Amarelo	1,08	0,8108	1,4275	0,4733
Azul	1,02	0,7911	1,3127	0,6995
Marrom	1,09	0,7521	1,5768	0,8544
Prata	0,98	0,7697	1,2564	0,1729
Preta	1,10	0,8575	1,4219	0,2198
Verde	1,19	0,9101	1,5659	0,3326
Vermelho	1,17	0,9125	1,4912	0,4416
Demais Cores	1,55	0,7328	3,2595	0,3428

Fonte: Elaboração Própria

As razões relativas de risco de ocorrências para o modelo estendido por outros fatores de análise, incluindo o período da semana da ocorrência do roubo ou furto, ano de fabricação do automóvel e período do dia da ocorrência são demonstrados nas Tabelas 6, 7 e 8, sendo que nenhum dos resultados foi estatisticamente significativo a 5%, embora os valores-p tenham sido relativamente baixos para as cores amarelo e vermelho, com riscos menores de roubo ou furto em relação aos brancos, em carros fabricados em períodos mais recentes.

**Tabela 6 – Estimativas pontuais e Intervalo de Confiança de 95% do Risco Relativo de roubo ou furto estimados no modelo Binomial Negativo, segundo cor do automóvel, estratificado por período da semana.**

Período da Semana	Cor	Risco Relativo	Intervalo de Confiança de 95%		P-Valor
			Limite Inferior	Limite Superior	
Dia de semana	Amarelo	1,08	0,8839	1,3314	0,3591
	Azul	1,02	0,8450	1,2384	0,6632
	Marrom	1,27	0,9950	1,6397	0,3683
	Prata	0,98	0,8127	1,1850	0,2609
	Preta	0,92	0,7557	1,1048	0,3973

	Verde	1,11	0,9102	1,3521	0,5527
	Vermelho	1,08	0,8944	1,3037	0,5462
	Demais Cores	1,06	0,7122	1,5820	0,7459
Fim de Semana	Amarelo	1,01	0,8186	1,2509	0,6857
	Azul	1,03	0,8429	1,2468	0,6525
	Marrom	1,28	0,9825	1,6683	0,3874
	Prata	1,01	0,8305	1,2217	0,1960
	Preta	0,94	0,7733	1,1407	0,2376
	Verde	1,13	0,9228	1,3873	0,4467
	Vermelho	1,05	0,8685	1,2770	0,5479
	Demais Cores	0,97	0,6359	1,4998	0,5804

Fonte: Elaboração Própria

**Tabela 7 – Estimativas pontuais e Intervalo de Confiança de 95% do Risco Relativo de roubo ou furto estimados no modelo Binomial Negativo, segundo cor do automóvel, estratificado por ano de fabricação do automóvel.**

Ano Fabricação	Cor	Risco Relativo	Intervalo de Confiança de 95%		Valor p
			Limite Inferior	Limite Superior	
<2013	Amarelo	0,81	0,6489	1,0199	0,1862
	Azul	0,96	0,7794	1,1746	0,4485
	Marrom	1,17	0,8844	1,5573	0,7481
	Prata	0,86	0,7060	1,0534	0,1267
	Preta	0,83	0,6739	1,0164	0,2174
	Verde	0,94	0,7517	1,1655	0,2664
	Vermelho	0,91	0,7387	1,1022	0,1349
	Demais Cores	1,42	0,8296	1,7121	0,6241
2013-2016	<b>Amarelo</b>	<b>0,76</b>	<b>0,5957</b>	<b>0,9456</b>	<b>0,0663</b>
	Azul	0,91	0,7413	1,1250	0,2548
	Marrom	1,16	0,8568	1,5611	0,8362
	Prata	0,87	0,7141	1,0718	0,1600
	Preta	0,79	0,6401	0,9708	0,2997
	Verde	0,91	0,7292	1,1374	0,1825
	<b>Vermelho</b>	<b>0,87</b>	<b>0,7129</b>	<b>1,0699</b>	<b>0,0731</b>
	Demais Cores	1,10	0,6234	1,9312	0,8365
2017-2021	<b>Amarelo</b>	<b>0,97</b>	<b>0,7718</b>	<b>1,2143</b>	<b>0,0869</b>
	Azul	1,05	0,8594	1,2912	0,5074
	Marrom	1,06	0,7727	1,4307	0,6468
	Prata	1,05	0,8609	1,2832	0,1652
	Preta	0,94	0,7692	1,1503	0,2329
	Verde	1,02	0,8236	1,2647	0,7061
	Vermelho	0,70	0,4281	1,1363	0,4141
	Demais Cores	1,15	0,9419	1,4042	0,8685

Fonte: Elaboração Própria

**Tabela 8 – Estimativas pontuais e Intervalo de Confiança de 95% do Risco Relativo de roubo ou furto estimados no modelo Binomial Negativo, segundo cor do automóvel, estratificado por período do dia.**

Período do Dia	Cor	Risco Relativo	Intervalo de Confiança de 95%		P-Valor
			Limite Inferior	Limite Superior	
À noite	Amarelo	1,03	0,8254	1,2581	0,3435

	Azul	0,96	0,8082	1,1845	0,3363
	Marrom	1,22	0,9261	1,5548	0,1524
	Prata	0,93	0,7725	1,1485	0,2478
	Preta	0,88	0,6958	1,0663	0,3695
	Verde	1,06	0,8691	1,2567	0,1336
	Vermelho	1,04	0,8545	1,2016	0,1194
	Demais Cores	0,99	0,6783	1,4722	0,7026
Pela manhã	Amarelo	0,94	0,7906	1,1809	0,1401
	Azul	0,98	0,8152	1,1839	0,4135
	Marrom	1,24	0,9228	1,5676	0,3693
	Prata	0,96	0,7758	1,1404	0,1860
	Preta	0,88	0,7313	1,1007	0,2265
	Verde	1,05	0,8521	1,3002	0,4243
	Vermelho	0,98	0,8381	1,2343	0,5221
	Demais Cores	0,92	0,6081	1,4143	0,5361
À tarde	Amarelo	0,94	0,7792	1,1875	0,6420
	Azul	0,95	0,7786	1,2019	0,6282
	Marrom	1,19	0,9428	1,589	0,3635
	Prata	0,97	0,8039	1,1694	0,1826
	Preta	0,87	0,7205	1,0644	0,2220
	Verde	1,08	0,8647	1,3003	0,427
	Vermelho	0,99	0,8247	1,1767	0,5201
	Demais Cores	0,94	0,607	1,4172	0,5425
De madrugada	Amarelo	0,96	0,7675	1,1563	0,6482
	Azul	0,96	0,8045	1,2048	0,6328
	Marrom	1,21	0,9339	1,5773	0,366
	Prata	0,96	0,7671	1,162	0,1826
	Preta	0,9	0,7402	1,0678	0,2231
	Verde	1,05	0,8902	1,2868	0,4203
	Vermelho	1,01	0,8339	1,1802	0,5255
	Demais Cores	0,94	0,6167	1,4487	0,5427

Fonte: Elaboração Própria

## 5. CONCLUSÕES E CONSIDERAÇÕES FINAIS

Este estudo investigou a relação entre a cor do veículo e o risco de ocorrência de furto ou roubo com base nos registros do estado de São Paulo no período de 2016 a 2021. Além disso, foram analisados fatores como o ano de fabricação do veículo, o dia da semana e o período do dia para avaliar se eles contribuem para variações no risco de acordo com a cor do veículo.



Com base nos dados utilizados neste estudo, não foi encontrada uma associação estatisticamente significativa entre a cor do veículo e a probabilidade de roubo ou furto, quando comparada a veículos de cor branca. No entanto, é importante ressaltar que esses resultados são baseados em dados de apenas um estado brasileiro e que esta área de estudo não foi amplamente explorada academicamente. Portanto, recomenda-se que mais pesquisas sejam conduzidas, abrangendo diferentes regiões geográficas e considerando outros parâmetros de interação, a fim de comparar resultados e obter conclusões mais abrangentes.

Sugere-se também a adoção de modelos adicionais que considerem interações entre as variáveis estudadas, além de uma reagrupação das variáveis que não possuem uma distribuição semelhante. Dessa forma, espera-se que futuras pesquisas possam servir como referência para aqueles interessados no assunto, contribuindo para ampliar o interesse nessa área não apenas no setor de seguros e segurança, mas em diversos setores da economia.

## REFERÊNCIAS BIBLIOGRÁFICAS

AGRESTI, Alan. **An introduction to categorical data analysis**. John Wiley & Sons, 2018.

ALANAZI, Fayez Khalaf. **Investigating the Effect of Vehicle Color on Crash Risk**. 2016. Tese de Doutorado. University of Dayton.

Burton, D., Delaney, A., Newstead, S., Logan, D. & Fildes, B. (2004) **Evaluation of Antilock Braking Systems Effectiveness**, Research report 04/01, Royal Automobile Club of Victoria, Melbourne, Australia.

DETRAM.SP – Departamento Estadual de trânsito do estado de São Paulo. **Frota por carroceria, categoria, cor, espécie e tipo**. Disponível em <https://detran.sp.gov.br/wps/portal/portaldetran/detran/estatisticastransito/3a410653-0dd2-45df-a324-bdfc4711d988>. Acesso em: 31/08/2022.

FURNESS, S. et al. **Car colour and risk of car crash injury: population based case control study**. *Bmj*, v. 327, n. 7429, p. 1455-1456, 2003.

GIL, Antonio Carlos et al. **Como elaborar projetos de pesquisa**. São Paulo: Atlas, 2002.

GOV.SP. CORONAVIRUS – **Normas legais de autoria do Executivo paulista elaboradas para dar andamento às ações de combate e prevenção à pandemia do novo coronavírus**. Disponível em <https://www.saopaulo.sp.gov.br/coronavirus/legislacao/>. Acesso em: 20/09/2022.

LARGET, Bret. Poisson Regression. **University of Wisconsin—Madison**. <http://www.stat.wisc.edu/courses/St572-larget/spring2007/handouts24-2.pdf> retrieved on, v. 23, n. 09, p. 2011, 2007.

LORD, Dominique; WASHINGTON, Simon P.; IVAN, John N. Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory. **Accident Analysis & Prevention**, v. 37, n. 1, p. 35-46, 2005.

MI – Ministério da Infraestrutura. **Frota de Veículos Automotores**. Disponível em <https://dados.gov.br/dataset/frota-de-veiculos>. Acesso em: 31/08/2022.

NATHAN, Robert A. **What's the safest color for a motor vehicle?** Traffic Safety (Chicago), 1969.

NEWSTEAD, Stuart; D'ELIA, Angelo. **An investigation into the relationship between vehicle colour and crash risk.** *Prevention*, v. 17, n. 1, p. 47-56, 2007.

OWUSU-ANSAH, **Stephen Osei.** **Investigation of the relationship between vehicle color and safety.** 2010. Tese de Doutorado. University of Dayton.

RICHARDSON, Roberto Jarry et al. **Pesquisa social: métodos e técnicas.** São Paulo: atlas, 1985.

RODRÍGUEZ, Germán. Poisson models for count data. **Lecture notes on generalized linear models**, 2007.

SSP – Secretaria de Segurança Pública do estado de São Paulo. **Furto de Veículos.** Disponível em <http://www.ssp.sp.gov.br/transparenciassp/>. Acesso em: 31/08/2022.