



DADOS DE CRIMINALIDADE EM RESENDE/RJ: TRATAMENTO E ANÁLISE COM FOCO EM PADRÕES E TENDÊNCIAS

CRIME DATA IN RESENDE/RJ: TREATMENT AND ANALYSIS WITH A FOCUS ON PATTERNS AND TRENDS

Priscilla Maria Pires dos SANTOS
Universidade do Estado do Rio de Janeiro (FAT/UERJ)
Email: priscilla.pires@fat.uerj.br
ORCID: <http://orcid.org/0000-0001-6531-149X>

Ana Clara Castillo de Assis CALAZANS
Universidade do Estado do Rio de Janeiro (FAT/UERJ)
Email: ana.calazans@discentes.fat.uerj.br
ORCID: <http://orcid.org/0009-0002-6801-5420>

RESUMO

Este trabalho analisou a criminalidade na cidade de Resende no estado do Rio de Janeiro, com foco nos crimes patrimoniais a partir de dados históricos coletados pelo Instituto de Segurança deste estado, de 2003 a 2024. O objetivo é explorar ferramentas para tratamento de dados, utilizando códigos em linguagem Python, para identificar padrões em ocorrências de roubos e furtos de celulares e veículos, bem como apreensão, posse e tráfico de drogas. Basicamente, veículos automotores, celulares e drogas são elementos centrais na vida urbana, gerando grande impacto na vida cotidiana. Discutimos como eventos sazonais, fatores econômicos e mudanças na legislação podem influenciar as ocorrências de tais delitos. A metodologia incluiu a análise exploratória dos dados por meio de visualização gráfica e uma experimentação da aplicação do algoritmo Prophet para prever tendências. Embora o Prophet tenha captado razoavelmente bem tendências gerais, ele se mostrou limitado ao prever picos e variações abruptas nos dados, características comuns em dados de criminalidade. A conclusão aponta para a necessidade de uma abordagem híbrida que integre análises exploratórias e métodos preditivos avançados e complementares para apoiar políticas públicas de segurança mais eficazes e informadas.

Palavras-chave: Análise de Dados. Crimes patrimoniais. Segurança Pública.

ABSTRACT

This work analyzes crime in the city of Resende, in the state of Rio de Janeiro, focusing on property crimes based on historical data collected by the Security Institute of this state in the period from 2003 to 2024. The aim is to use codes in the Python programming language to explore data processing tools to identify patterns in robberies and thefts of cell phones and vehicles, as well as in the seizure, possession and trafficking of drugs. Fundamentally, motor vehicles, cell phones and drugs are central elements of urban life and have a major impact on daily life. We discuss how seasonal events, economic factors and changes in legislation can influence the occurrence of such crimes. The methodology involved exploratory data analysis through graphical visualization and experimentation with the application of the Prophet algorithm to predict trends. Although Prophet captured general trends quite well, it proved limited in predicting spikes and abrupt fluctuations in the data, which are common in crime data. The conclusion points to the need for a hybrid approach that integrates exploratory analysis and advanced and complementary predictive methods to support more effective and informed public safety policies.

Keywords: Data analytics. Property crime. Public safety.

INTRODUÇÃO

A Lei nº 3.329/1999 criou o Instituto de Segurança Pública do Estado do Rio de Janeiro (ISP/RJ). Em 2024, o estado do Rio de Janeiro ocupou a terceira posição em crimes violentos no Brasil, com média de 21,10 ocorrências por 100 mil habitantes (Globo, 2024a). Os registros de ocorrências da Polícia Civil do Rio de Janeiro estão disponíveis em bases on line (<https://www.ispdados.rj.gov.br/>).

O tratamento de grande volume de dados representa um desafio conhecido como “Big Data”. A Ciência de Dados e Inteligência Artificial (IA) ajudam a identificar padrões e anomalias, o que acelera investigações científicas (Miranda, 2019; Rezende, 2003). O município de Resende, com cerca de 130 mil habitantes (Ibge, 2024), tem registrado, nos últimos anos, um número relevante de operações policiais focadas no combate ao tráfico de drogas (Globo, 2024b), com uso inclusive de veículo blindado conhecido como “caveirão”.

Os dados do ISP, organizados mensalmente, formam uma série temporal que pode ser analisada quanto a tendências, sazonalidades, ciclos e resíduos (Harvey E Peters, 1990). A análise de séries temporais na Ciência de Dados permite fazer previsões a partir de dados históricos e apoiar políticas públicas de segurança.

Desenvolvido por uma estudante de graduação da Faculdade de Tecnologia (FAT), Campus da Universidade do Estado do Rio de Janeiro em Resende, este estudo utiliza dados do ISP/RJ para identificar tendências e padrões em alguns tipos de crimes na cidade, aplicando códigos em Python e na etapa final o algoritmo Prophet (Meta Inc., 2017). Este trabalho organiza-se em: metodologia na seção 2, análise exploratória e resultados nas seções 3 e 4, e conclusões na seção 5. A partir da seção 3, discutimos os principais aprendizados e limitações encontradas.

METODOLOGIA

Os dados foram coletados do site oficial do ISP/RJ, com foco no município de Resende/RJ, o qual possui registros mensais de janeiro de 2003 a setembro de 2024. O objetivo foi analisar crimes contra o patrimônio, nesta primeira parte, enquanto que uma segunda pesquisa abordará crimes contra a pessoa, incluindo violência contra a mulher, dada sua relevância social.

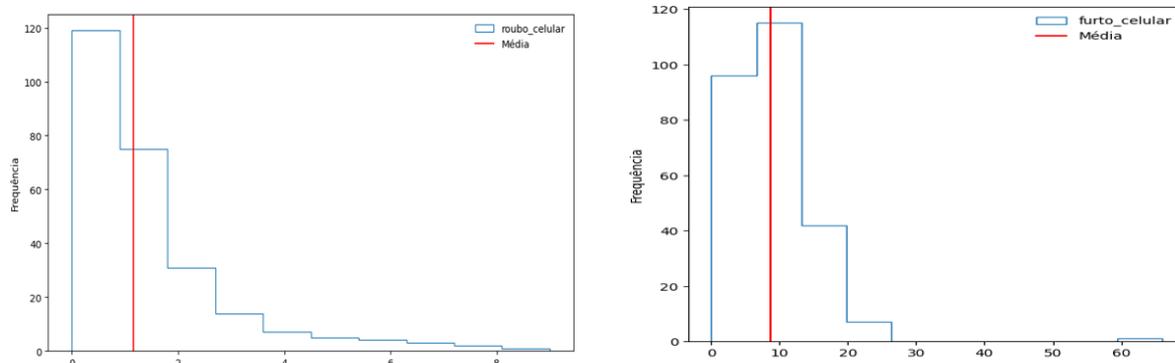
A análise exploratória dos dados, realizada no Google Colaboraty com bibliotecas como Pandas, apresentou 261 linhas e 59 colunas, todas completas. Isso facilitou o processo, evitando problemas comuns na Ciência de Dados relacionados a dados ausentes. Com o “dataset” verificado, avançamos para as análises exploratórias, cujos resultados serão discutidos na seção³.

ESTUDO DESCRITIVO DE ROUBOS E FURTOS DE CELULARES

Das 59 colunas do banco de dados, selecionamos para análise os crimes de roubo e furto de veículos e celulares, além de posse, apreensão e tráfico de drogas. Concentramos nesses delitos devido as altas incidências em zonas urbanas, impacto social e econômico para as vítimas, além da presença de tecnologias de rastreamento e monitoramento, que auxiliam na recuperação e prevenção. Todas essas características oferecem indicadores úteis para ações preventivas. Vale esclarecer que conforme o

Código Penal Brasileiro, roubo (Art. 157) envolve ameaça ou violência, enquanto furto (Art. 155) ocorre sem violência contra a vítima.

Figura 1: à esquerda histograma da distribuição de ocorrências de roubo de celulares e à direita histograma da distribuição de ocorrências de furto de celulares, evidenciando a concentração de registros.

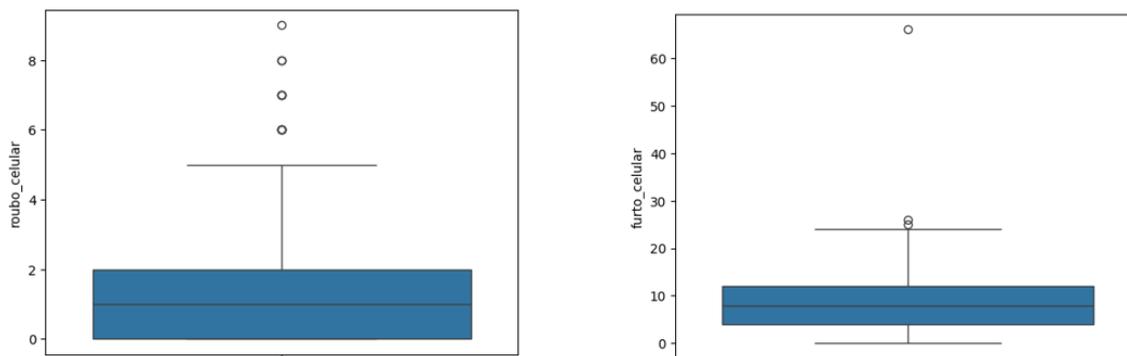


Fonte: elaborado pelos autores.

No gráfico de roubos (à esquerda), a maioria das ocorrências concentra-se em 1 a 2 casos mensais, com alguns valores altos esporádicos classificados como “outliers”. Isto nos indica uma distribuição assimétrica (Aggarwal, 2016). O número máximo de casos registrados em um único mês foi de 9 ocorrências. Em 119 meses (aproximadamente 45% do período analisado), não houve registros deste tipo de crime, o que pode indicar também uma subnotificação (o fato não foi comunicado ou contabilizado pela polícia) e não necessariamente a ausência do delito. A média mensal foi de pouco mais de 1 roubo. Já os furtos de celulares (à direita) mostram uma distribuição menos assimétrica, com a maioria dos meses registrando até 20 casos e pico de 66 em outubro de 2023. Apenas 7 meses não tiveram registros de furtos, e a média mensal foi de 8 casos. Assim, nota-se que, na última década, furtos superaram roubos de celulares na cidade.

O boxplot é útil para detectar “outliers” com base no intervalo interquartil (IQR), definido como a diferença entre o 3º (Q3) e o 1º quartil (Q1). Os “outliers” são valores fora do intervalo $Q1 - 1,5 \times IQR$ a $Q3 + 1,5 \times IQR$. O boxplot mostra cinco pontos principais: valor mínimo, Q1, mediana, Q3 e valor máximo (excluindo “outliers”).

Figura 2: à esquerda boxplot das ocorrências de roubo de celular e à direita boxplot das ocorrências de furto de celular, de 01/2003 até 09/2024.

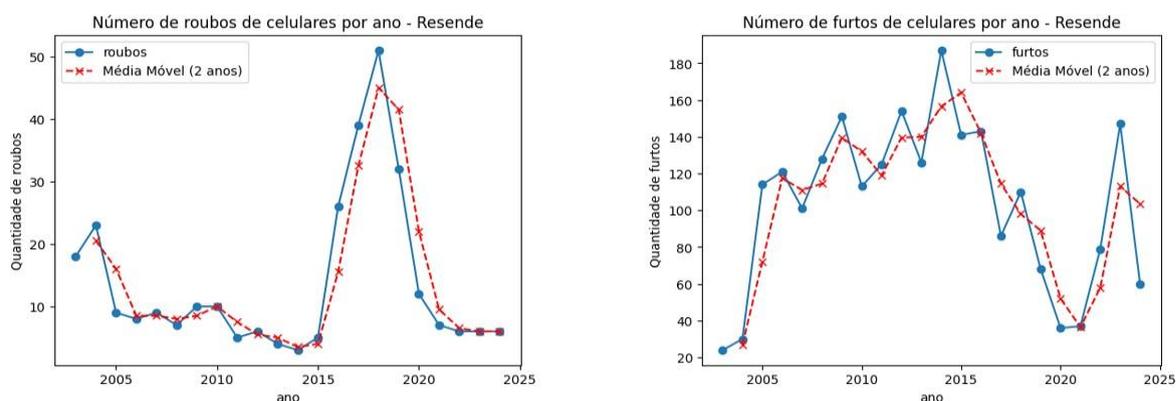


Fonte: elaborado pelos autores.

A distribuição dos dados (Figura 2) destaca a tendência central e a dispersão. Para roubos de celulares, o limite superior é 5, com “outliers” representando 3,8% da amostra. Uma maior variação no último quartil deste gráfico sugere picos em certos períodos. No gráfico dos furtos de celulares, o valor 66 é um “outlier” evidente. A média e mediana são cerca de 8 casos, com 75% dos meses registrando até 12 ocorrências. Apenas 2,4% dos dados de furtos são “outliers”, o que indica que esses valores são exceções na cidade.

A Figura 3 mostra o gráfico de linhas dos dados de roubos e furtos de celulares por ano, evidenciando os picos de “outliers” identificados no boxplot. A Média Móvel Simples (MMS) suaviza os dados, dando destaque para tendências e movimentos de longo prazo. Com uma janela de dois anos, a MMS permite observar crescimento ou declínio ao comparar períodos anuais consecutivos.

Figura 3: evolução anual da quantidade de roubos e furtos de celulares na cidade de Resende/RJ juntamente com as respectivas Médias Móveis simples.



Fonte: elaborado pelos autores.

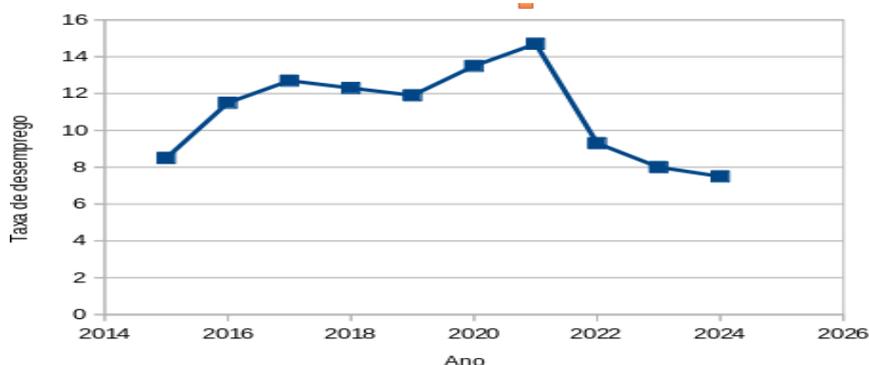
Por meio da Figura 3, vemos que o número significativo de furtos antes de 2015, se deve, provavelmente, a maior acessibilidade e popularização dos “smartphones”, a um controle menos rigoroso da revenda, menor adoção na época de medidas de bloqueios (tipo IMEI), isto é, os celulares eram facilmente reutilizáveis ou revendidos, enquanto os roubos variaram numa faixa 10 vezes menor. No entanto, de 2015 a 2020, vemos um aumento progressivo nos roubos e queda acentuada após 2020, coincidindo com o início da pandemia da COVID-19. Em relação aos furtos, a análise da distribuição anual, revela um movimento decrescente de 2015 a 2020, embora oscilatório, seguido por uma estabilidade em 2020 e 2021, e um crescimento gradual no período pós-pandêmico (após 2021). A MMS acompanhou de forma satisfatória a evolução dos registros.

A seguir, discutiremos alguns fatores que podem explicar o comportamento antes e depois do período pandêmico.

Creemos que o aumento de roubos, alguns anos antes de 2020, foi provavelmente influenciado por crises econômicas, desemprego e um menor controle sobre a revenda de aparelhos de terceiros, uma vez que “smartphones” são alvos atrativos financeiramente. A queda gradual após o período pandêmico (2020-2021) sugere que o sistema de bloqueio por IMEI, implementado pela ANATEL desde 2016, e melhorias na segurança pública na cidade tiveram impacto positivo.

A Figura 4 mostra a taxa de desemprego anual do Brasil desde 2014. Durante a crise econômica de 2015-2017, o desemprego subiu, agravado pela recessão e prejuízos de grandes empresas (Globo, 2017). Durante o auge da pandemia da COVID-19 (2020-2021) o desemprego aumentou novamente, mas entre 2022 e 2024 se observa uma queda gradual, aproximando-se dos níveis pré-pandemia.

Figura 4: evolução anual da taxa de desemprego no Brasil.

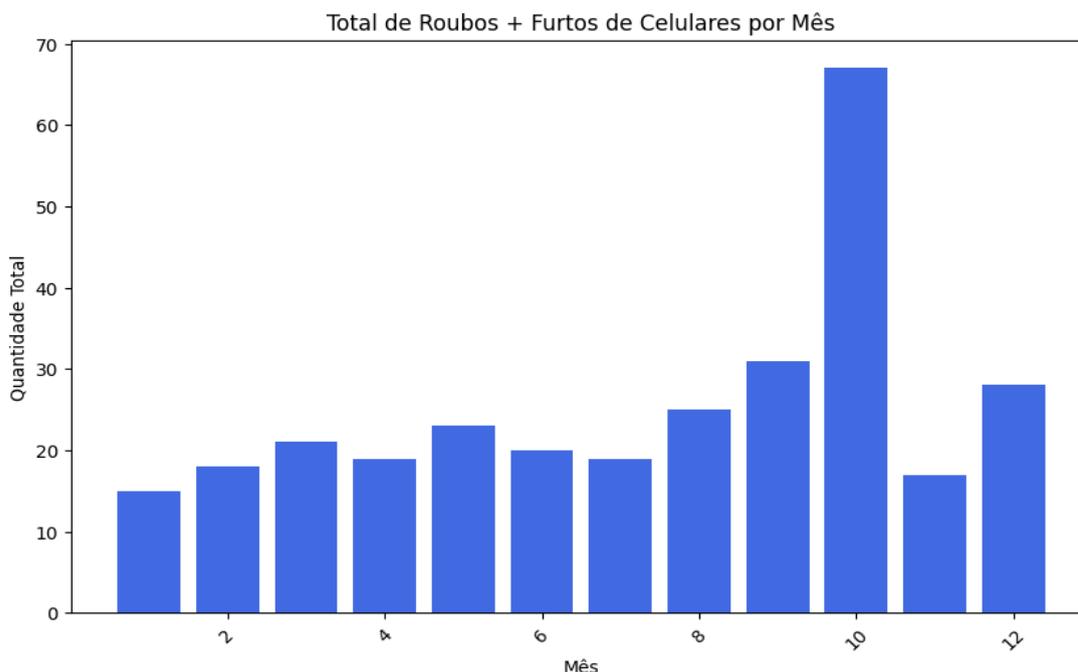


Fontes: Trading Economics (2024) e IBGE (2024), elaborada pelos autores.

Além do custo de reposição do celular para a vítima, há o risco de uma perda financeira ainda maior, caso os criminosos acessem contas bancárias ou informações pessoais no dispositivo. A Figura 3 indicou que a crise econômica de 2015 e anos subsequentes elevou a agressividade dos crimes, convertendo furtos em roubos de modo a garantir o sucesso da ação. Após 2021, com a retomada econômica e o aumento da circulação de pessoas nas ruas, os furtos cresceram enquanto os roubos diminuíram. Este fato sugere que furtos se tornaram uma forma de “ganho” menos arriscada em relação aos roubos, no geral, violentos.

A Figura 5 apresenta o total mensal de roubos e furtos de celulares. Esta comparação permite à segurança pública planejar estratégias ao observar aumentos ou reduções sazonais, possivelmente ligados a eventos locais. Por exemplo, setembro teve um leve aumento em relação a agosto, e outubro registrou um aumento expressivo em relação aos demais meses, o que sugere maior necessidade de policiamento preventivo neste período de máximas.

Figura 5: variação mensal do total de roubos e furtos (somados) de celulares na cidade de Resende/RJ.



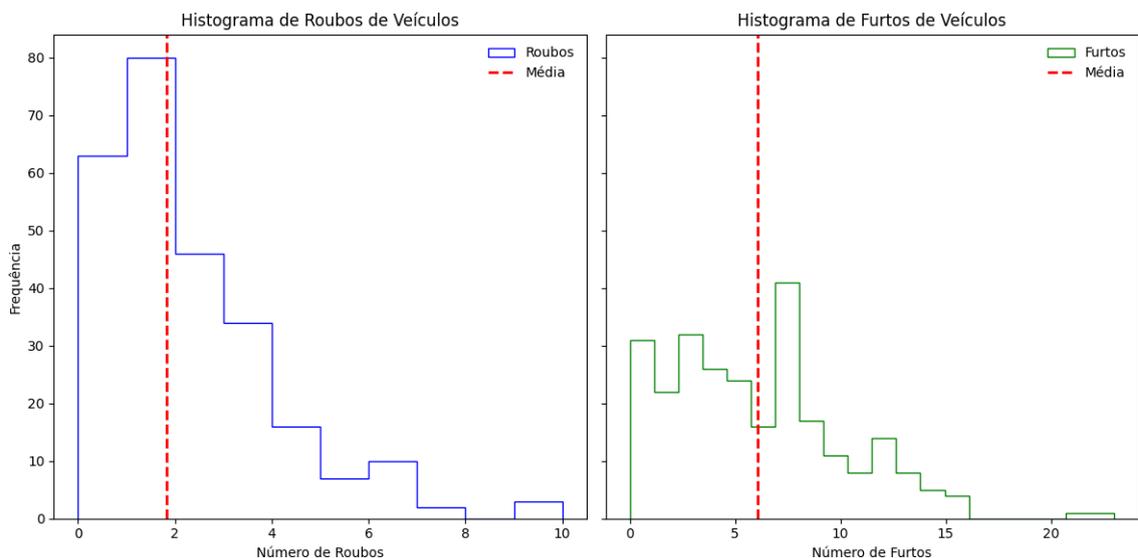
Fonte: Elaborado pelos autores.

Vale ressaltar que a Exposição Agropecuária, Industrial e Comercial (Exapicor), principal evento anual de Resende, ocorre entre setembro e outubro, com duração de cerca de 10 dias, incluindo sempre o dia 29 de setembro. O evento promove shows de grandes artistas nacionais de forma gratuita, feiras e atrai um grande público. Esse aumento de furtos e roubos de celulares verificado, destaca a necessidade de reforço no policiamento preventivo nessa época.

ESTUDO DESCRITIVO DE ROUBOS E FURTOS DE VEÍCULOS

Nesta seção, analisamos os roubos e furtos de veículos. A Figura 6 apresenta uma comparação das frequências mensais desses crimes por meio de histogramas. O gráfico à esquerda na Figura 6 mostra uma distribuição assimétrica para roubos de veículos, concentrada em 0 e 2 casos, com uma média de 2 e raros registros acima de 4. Tal fato sugere que roubos foram incomuns no período. Já o gráfico à direita da mesma figura mostra maior variabilidade para furtos, com frequências mensais até acima de 10 e uma média de 5. Tal fato sugere que furtos de veículos são mais frequentes na cidade. A comparação mostra que roubos são menos comuns e mais concentrados, enquanto os furtos são mais frequentes e espalhados mensalmente.

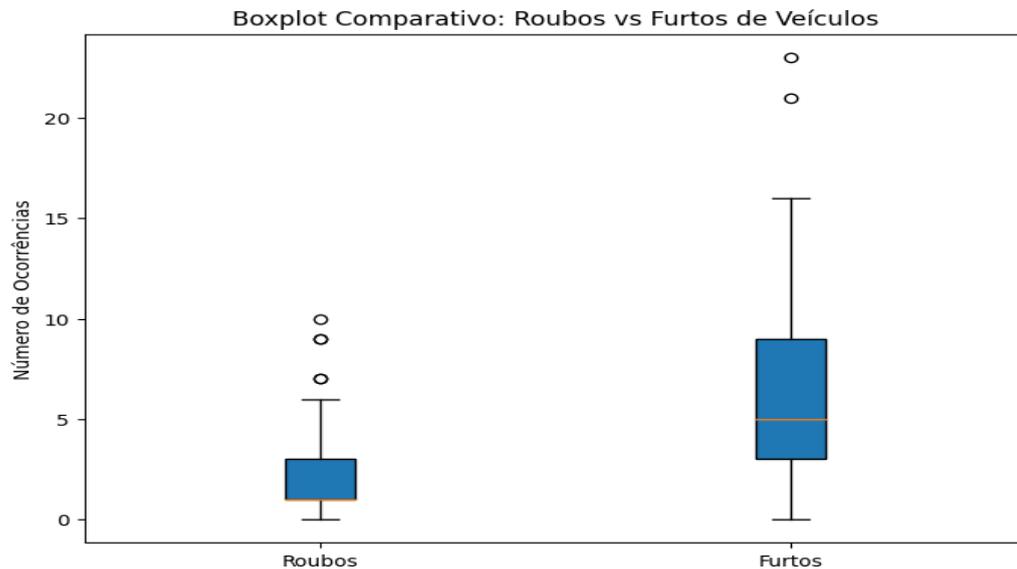
Figura 6: histogramas comparativos da frequência mensal de roubos e furtos de veículos, com destaque as variabilidades.



Fonte: elaborado pelos autores.

Na sequência, tem-se o gráfico “boxplot” relativo à Figura 6 para visualização dos quartis, mediana e pontos fora do padrão.

Figura 7: boxplot comparativo da distribuição do número de roubos e furtos de veículos.

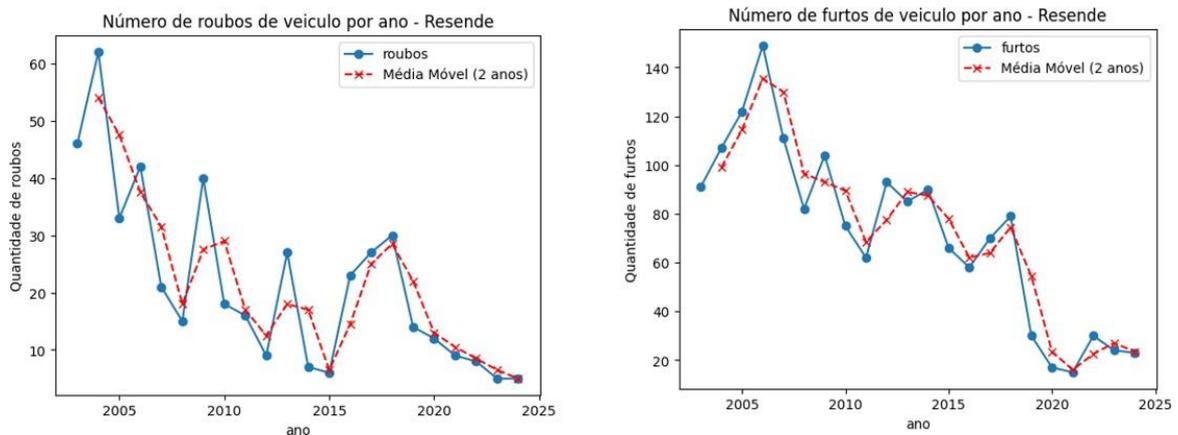


Fonte: Elaborado pelos autores.

Observa-se na Figura 7 que a mediana para ocorrências de roubos (linha dentro da caixa) é baixa, próxima de 1, indicando que a maioria dos casos de roubo não ultrapassa o valor de 5, com poucas variações. Em contraste, para furtos, a mediana é mais elevada, com os valores mínimos concentrando-se mais próximos da mediana (Q2 - 50%) do que os valores máximos. Esse comportamento sugere uma variação maior no último quartil, o que indica que, em alguns meses, o número de furtos foi significativamente mais alto que o usual.

Na Figura 8 é apresentada a quantidade anual de crimes contra a propriedade de veículos, acompanhada pela MMS, a qual suaviza as flutuações anuais e destaca as tendências ao longo do período total.

Figura 8: à esquerda se tem o número de roubos de veículos por ano, e à direita o número de furtos de veículos por ano, na cidade de Resende/RJ.



Fonte: Elaborado pelos autores.

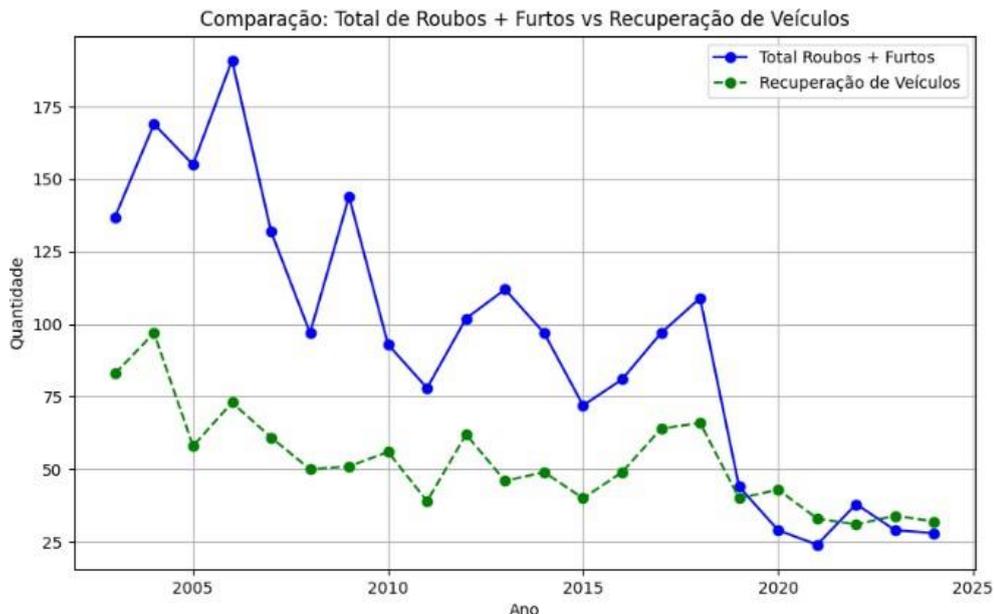
Nota-se um decréscimo em ambos os tipos de crimes contra veículos após 2005. O número de roubos caiu de 60 casos para menos de 5 em 2024, enquanto os furtos, que atingiram um pico de aproximadamente 140 ocorrências, reduziram para menos de 20 a partir de 2020. Embora a tendência geral seja de queda, observam-se oscilações anuais, com pequenos aumentos e quedas, especialmente no período entre 2005 e 2015. Destaca-se que de 2015 a 2018 ambos os crimes aumentaram anualmente. Após 2020, o número de roubos voltou a cair, possivelmente refletindo políticas de segurança mais eficazes ou menor exposição de veículos ao roubo. Em contraste, os furtos subiram levemente após 2020 e se estabilizaram. Conclui-se que a tendência geral é positiva, com uma redução dos crimes contra veículos na cidade.

A redução de casos pode ser atribuída ao aprimoramento das tecnologias de monitoramento urbano, à aplicação de legislações como a Lei do Desmonte de 2014 e ao possível desvio do foco criminoso para atividades como o tráfico de drogas, discutido posteriormente. A Figura 9 apresenta o total de roubos e furtos de veículos em comparação ao número de recuperação. A recuperação de veículos é um aspecto crucial nesta análise, pois indica a eficácia das políticas de segurança pública e seu impacto econômico.

Na Figura 9, a linha azul exibe um pico inicial em torno de 2005, seguido por uma queda gradual nas ocorrências ao longo dos anos, com uma acentuada redução a partir de 2018. A linha verde, que representa a recuperação de veículos, mostra um leve declínio ao longo do tempo, mas de forma menos pronunciada, mantendo-se relativamente

estável, mesmo com variações anuais. Observa-se uma discrepância entre os crimes e a recuperação no início da série histórica (2003 a 2010), sugerindo uma eficiência limitada de recuperação neste período. A partir de 2020, as linhas azul e verde se aproximam, refletindo uma queda significativa nos crimes, o que possibilitou uma recuperação proporcionalmente melhor.

Figura 9: evolução anual dos roubos e furtos de veículos somados (linha azul) em comparação com a recuperação de veículos (linha verde) entre aproximadamente 2003 e 2024.



Fonte: Elaborado pelos autores.

Citamos, entre os fatores que provavelmente impulsionaram a recuperação de veículos, o aprimoramento do rastreamento via GPS, o aumento do uso de câmeras de vigilância e a Lei do Desmonte de 2014, que dificulta o comércio ilegal de peças. No entanto, a linha verde mostrou uma estabilidade nas recuperações, enquanto as ocorrências caíram mais rapidamente, o que indica que há espaço para melhoria da eficiência das operações.

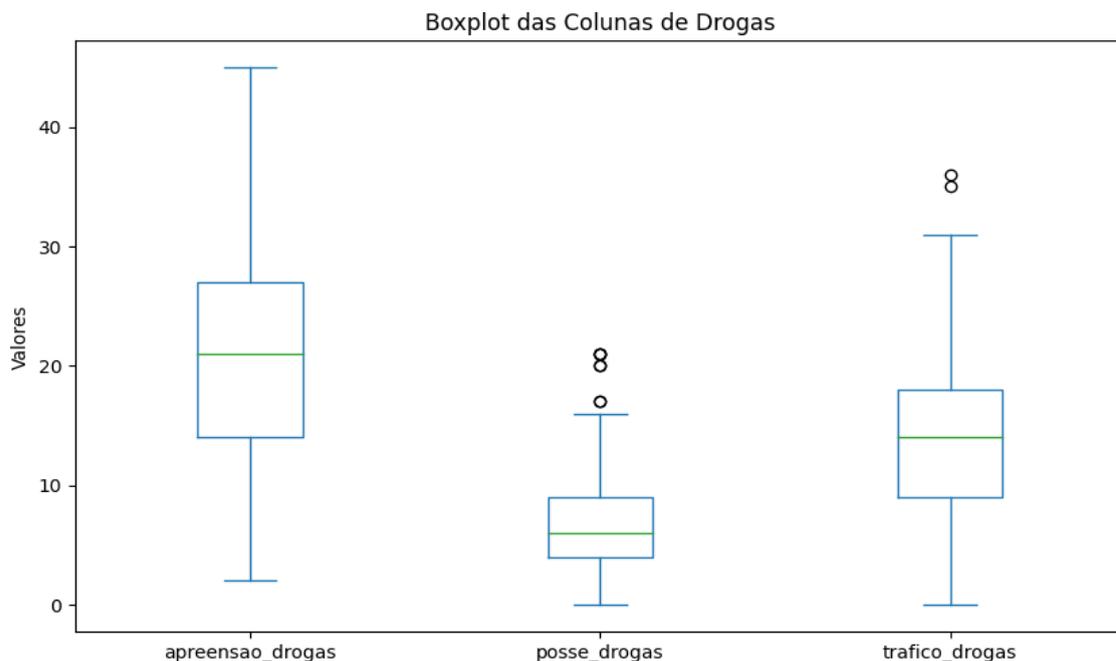
A partir deste ponto, a análise se volta para crimes relacionados a entorpecentes, de forma a entendermos sua dinâmica e possíveis padrões.

ANÁLISE DESCRITIVA DOS CRIMES ENVOLVENDO ENTORPECENTES

A Lei nº 11.343/2006 (Lei de Drogas) define as infrações: o Art. 28 considera a posse para o consumo pessoal uma infração administrativa, enquanto que o Art. 33 caracteriza tráfico como produção, venda ou transporte não autorizado legalmente. A apreensão de drogas se refere ao confisco de substâncias ilegais pelas autoridades policiais.

Na Figura 10 apresentamos gráficos do tipo boxplot para comparação das variáveis apreensão, tráfico e posse de drogas.

Figura 10: boxplots comparativos de apreensão, tráfico e posse de drogas.



Fonte: Elaborado pelos autores.

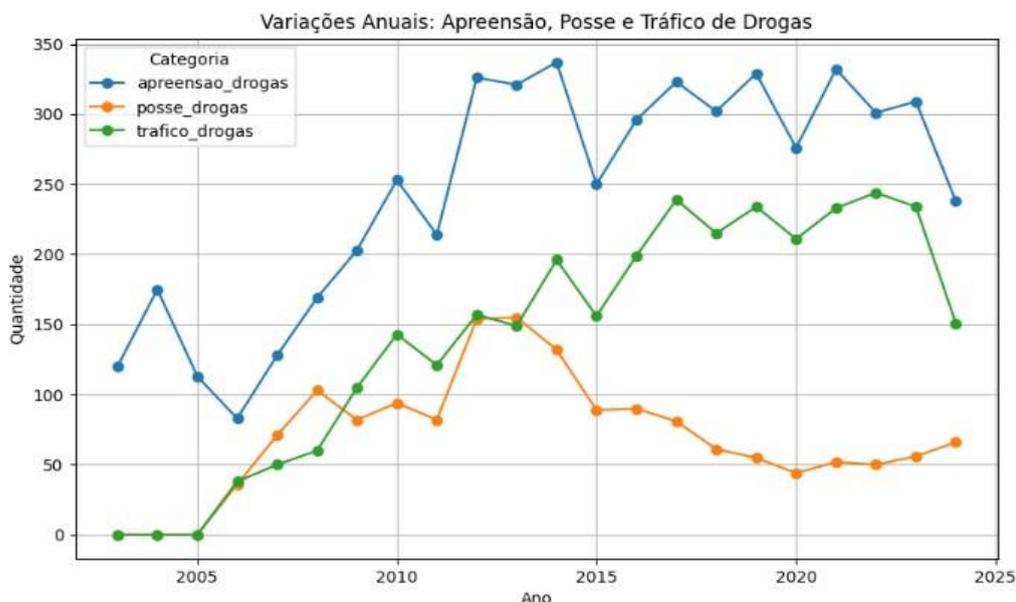
A seguir, detalharemos a análise da Figura 10 para cada variável. Em relação à apreensão de drogas, observa-se uma ampla dispersão na distribuição, variando de 0 a cerca de 40 casos mensais. A mediana está em torno de 20, o que indica que, em pelos menos metade dos meses entre 2003 e 2024, foram registradas cerca de 20 apreensões. A amplitude interquartil (área entre o Q1 e o Q3), mostra uma variabilidade significativa, sem valores atípicos identificados.

Para os casos de posse de drogas, a dispersão é menor, com uma mediana de aproximadamente 8 registros mensais. Alguns valores fora do padrão (“outliers”) foram registrados em determinados meses dos anos de 2012, 2013, 2014 e 2015. A amplitude interquartil reduzida sugere que a maioria dos dados se concentra em torno da mediana.

Em relação ao tráfico de drogas, a distribuição se concentra entre 10 e 20 casos mensais, com uma mediana próxima de 14. A amplitude interquartil é moderada, indicando alguma variabilidade. Foram identificados valores atípicos acima de 30, sugerindo picos eventuais. A dispersão é semelhante à observada para apreensão de drogas, mas em uma faixa inferior. Destacam-se picos nos meses de dezembro de 2021 e junho de 2023.

Em resumo, a posse de drogas mostrou uma distribuição mais concentrada, enquanto a apreensão de drogas e o tráfico apresentaram maior dispersão, embora em diferentes intervalos de valores. A seguir, são apresentados gráficos temporais para as três variáveis.

Figura 11: evoluções anuais de apreensão, posse e tráfico de drogas na cidade de Resende/RJ.



Fonte: Elaborado pelos autores.

O gráfico da Figura 11 apresenta a evolução anual de três categorias relacionadas aos registros de drogas: apreensão (linha azul), posse (linha laranja) e tráfico (linha verde). A linha azul mostra um aumento expressivo no número de apreensões de 2006 até 2014, seguido por um leve declínio após 2015 e uma retomada no crescimento entre 2016 e 2020. Este período de variações frequentes pode refletir ajustes nas estratégias de combate. A partir de 2020, observa-se uma tendência de queda, seguida por uma retomada e nova queda recente, possivelmente indicando mudanças nas políticas de repressão. A linha verde, que representa o tráfico, revela um crescimento constante, com

oscilações até 2023, quando ocorre uma recente queda. No entanto, o número de casos registrados de tráfico permanece elevado em comparação com os anos iniciais da série. Já a linha laranja, que representa posse, apresenta um crescimento do início até 2013, seguido por um declínio contínuo, com oscilações a partir de 2014, estabilizando-se em um patamar mais baixo.

Visualmente, nota-se uma forte correlação entre as curvas de apreensão e tráfico: quanto maior a repressão ao tráfico, maior a apreensão de drogas. A Figura 13 também sugere uma possível mudança no foco das autoridades para o combate ao tráfico e uma abordagem diferenciada nos casos de posse para consumo pessoal.

ASSOCIAÇÕES ENTRE CRIMES PATRIMONIAIS E DROGAS

Esta seção integra dados estatísticos que possibilitam a análise das associações entre diferentes tipos de delitos, ajudando a identificar padrões de atividade criminosa, de forma a visualizar, quando existente, a lógica que conecta crimes distintos.

Três métodos são amplamente utilizados para avaliar a correlação entre variáveis: o Método de Correlação de Pearson (Zou et al., 2003), o Método de Spearman (Mukaka, 2012) e o Método de Kendall (Origuela, 2018). Os métodos de correlação podem fornecer informações de forma a subsidiar políticas públicas integradas e identificar padrões de crimes que ocorrem de forma simultânea ou relacionada. O método mais adequado depende da natureza dos dados e do comportamento das variáveis.

Avaliamos a correlação entre crimes patrimoniais e crimes de drogas, considerando características como linearidade, presença de pontos fora da curva e distribuição dos dados. Optamos pelo Método de Spearman, mais adequado para relações não lineares e robusto a “outliers”, oferecendo à pesquisa precisão em cenários em que uma variável não responde diretamente à outra. Ele também lida melhor com meses atípicos, comuns em estudos criminais, fornecendo uma visão mais fiel das relações analisadas.

A Figura 12 apresenta um mapa de calor da correlação de Spearman entre crimes patrimoniais e de entorpecentes em Resende/RJ. Cores intensas indicam correlações fortes, enquanto cores neutras refletem correlações fracas ou inexistentes. Este mapa ajuda na identificação de padrões e relações entre os delitos, oferecendo uma visão aprofundada das dinâmicas criminais.

Figura 12: mapa de calor da correlação de Spearman entre crimes patrimoniais e crimes relacionados a drogas na cidade de Resende/RJ.



Fonte: Elaborado pelos autores.

Em relação à apreensão, observa-se uma correlação negativa com crimes relacionados a veículos e praticamente nula com celulares, sugerindo que as apreensões de drogas têm pouca ou nenhuma conexão direta com os crimes patrimoniais aqui explorados, especialmente os de celulares. Em relação à posse, vê-se que apresenta leves associações positivas com furtos de veículos e celulares, e correlações quase nulas com roubos, o que sugere uma possível relação entre o consumo de drogas e a prática de crimes menores, como furtos. Já em relação ao tráfico, a correlação de -0.46 com furtos de veículos indica uma associação negativa moderada. Isso sugere que os dois crimes não ocorrem de forma paralela. Já a correlação com furtos de celulares é ainda menor, de -0.29. Em resumo, isso indica que em áreas com maior incidência de tráfico há uma tendência de redução dos furtos. Em relação aos roubos de veículos e celulares, as correlações são ambas de intensidade fraca, sendo com veículos inversa e com celulares direta. A correlação levemente positiva com roubos de celulares sugere que, possivelmente, o tráfico de drogas pode ocorrer em áreas onde há uma leve propensão para roubos de celulares, mas como dissemos a relação é muito fraca para indicar o padrão de forma clara.

A próxima seção explora o uso do Facebook Prophet, uma ferramenta de código aberto para previsão de séries temporais, ideal para dados com tendências não lineares ou dados faltantes, de maneira a se obter projeções robustas.

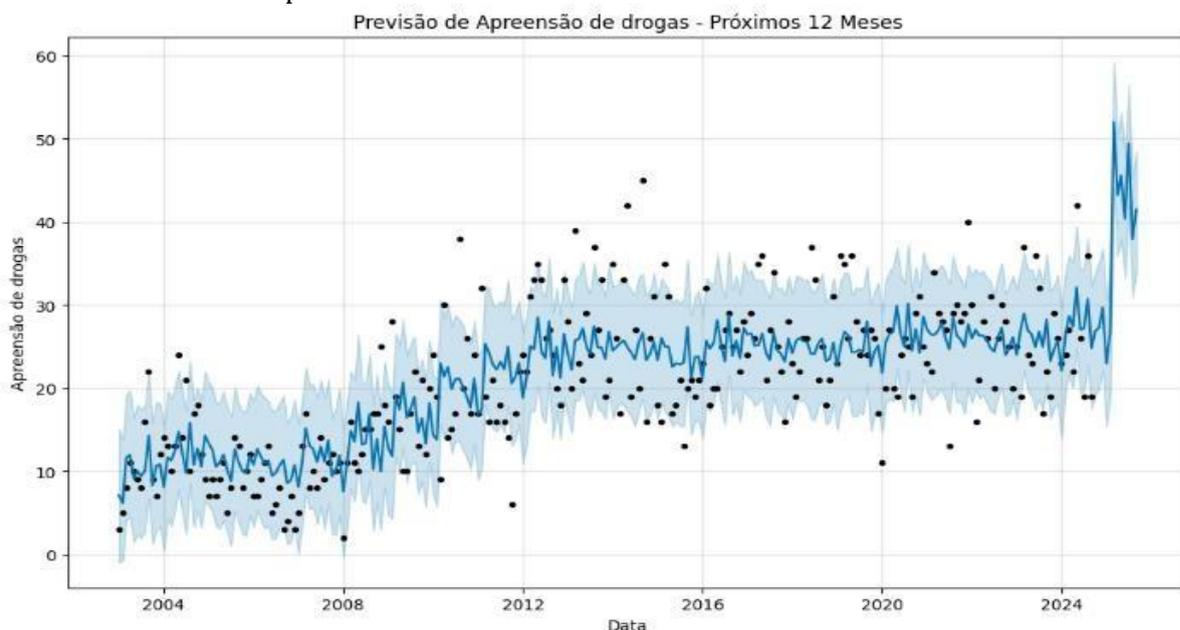
APLICAÇÃO DO PROPHET

O Prophet opera em quatro etapas principais: preparação dos dados: as colunas de datas e valores são nomeadas com ds e y para definir a sequência temporal; treinamento do modelo: ajuste do modelo aos dados históricos para capturar sazonalidade e tendências; previsões: o modelo projeta a evolução futura dos dados; avaliação dos resultados: comparações entre dados reais e validação do modelo por meio de métricas, como Erro Absoluto Médio (MAE), ou MAPE – Erro absoluto médio percentual.

Utilizando o Prophet, criamos um modelo para prever a apreensão de drogas nos próximos 12 meses (a partir do último dado tabelado). Este modelo experimental simplifica a análise e facilita a compreensão dos resultados, oferecendo percepções que poderão ser aplicadas a outras variáveis no futuro.

Nas Figuras 13 e 14, são apresentados os resultados utilizando o Prophet. Incluímos no modelo feriados nacionais e a Exapicor, este último devido a sua relevância anteriormente mencionada.

Figura 13: previsão da apreensão de drogas para os próximos 12 meses, gerada pelo modelo Prophet.



Fonte: Elaborado pelos autores.

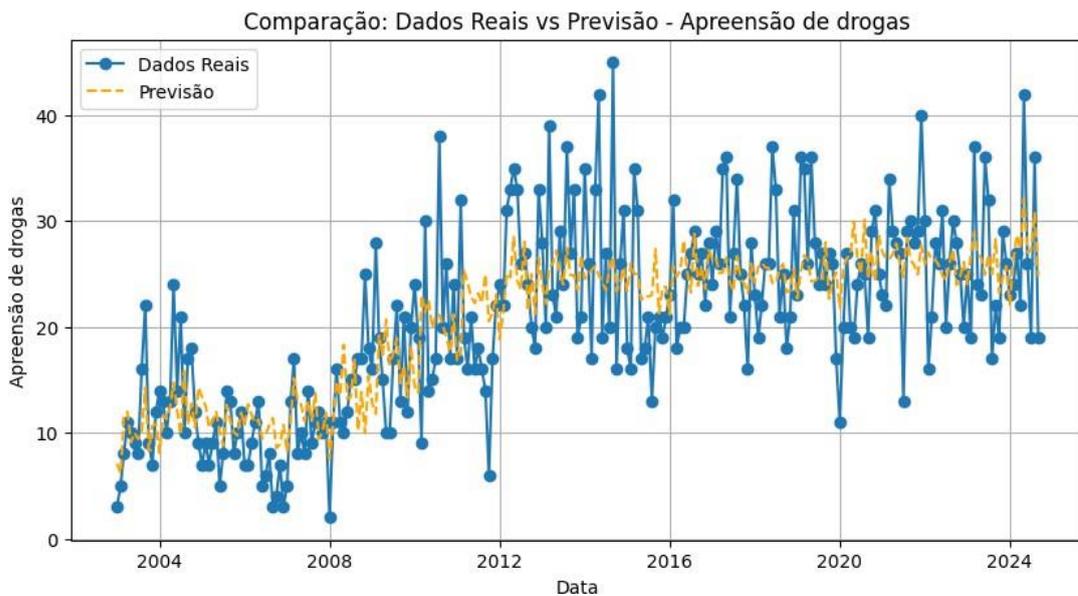
De forma geral, o modelo capturou bem a tendência de aumento ao longo do tempo (linha azul). O intervalo de confiança (sombra azul) é mais estreito no início e se amplia levemente ao longo do tempo na tentativa de acompanhar os pontos reais (pretos). Logo, reflete maior incerteza nas previsões de longo prazo, o que é comum em séries temporais com alta variabilidade, como as de dados sociais e criminalidade. O modelo apresentou um MAPE (Erro Percentual Absoluto Médio) de 28%, indicando uma precisão razoável, mas que poderia ser aprimorada, por exemplo, reduzindo o horizonte de previsão para 3 ou 6 meses.

De fato, a Figura 13 mostra que o Prophet conseguiu prever a tendência de aumento e estabilização nas apreensões de drogas. Por exemplo, entre 2003 e 2015, observa-se um aumento progressivo nas apreensões, indicado tanto pelos dados reais como pela previsão. No entanto, diversos picos nos dados reais não foram capturados com precisão pelo modelo, indicando que a previsão é suavizada e não refletiu as variações bruscas.

A presença de pontos reais acima e abaixo do intervalo de confiança reforça a variabilidade dos dados sociais. Ainda assim, a maioria dos pontos reais está dentro da faixa de confiança, uma vez que o modelo conseguiu, em grande parte, prever a faixa geral de variação, embora com limitações. O modelo apontou um aumento significativo das apreensões, influenciado pelos últimos dados da série histórica. Assim, faz-se necessário ajustes no modelo para evitar projeções baseadas apenas nos dados recentes.

A Figura 14 apresenta uma comparação dos dados reais de apreensão de drogas (linha azul sólida) com as previsões do Prophet (linha tracejada laranja). Embora o modelo capture a tendência geral, picos e flutuações dos dados reais não foram previstos com precisão.

Figura 14: comparação de dados reais contra a previsão fornecida pela Prophet em relação a variável apreensão de drogas.



Fonte: Elaborado pelos autores.

Todos os delitos analisados nesta pesquisa são altamente dependentes de operações policiais, denúncias e fatores aleatórios. Portanto, é esperado que os dados apresentem picos e variações sem um padrão claro, o que torna a previsão extremamente complexa. Logo, o Prophet se mostrou limitado, pois, embora seja eficaz para capturar tendências e sazonalidades gerais, ele não previu de forma satisfatória picos extremos ou variações aleatórias.

Após várias iterações, aumentamos o parâmetro `Fourier_order`, que define os harmônicos de Fourier na modelagem sazonal, e testamos prazos de previsões mais curtos (6 e 3 meses). Embora o MAPE inicial de 28% mostre espaço para melhorias, o modelo captou a tendência geral. Regressores adicionais, como dados de operações policiais, e técnicas complementares são necessários para lidar melhor com os valores extremos. Ao ajustar o `changepoint_prior_scale` do Prophet de 0.1 para 0.5, reduziu-se o MAPE para 26%, aumentando a sensibilidade do algoritmo a mudanças bruscas.

Portanto, optamos por não aplicar o algoritmo do Prophet às demais variáveis, uma vez que os resultados obtidos por meios de vários testes foram pouco precisos e não representam fielmente as flutuações complexas e características dos delitos analisados.

Em conclusão, acreditamos que o ideal seria utilizar modelos de aprendizado de máquina, como árvores de decisão ou redes neurais, de forma complementar ao Prophet para capturar picos fora do padrão - embora isso esteja além do escopo deste trabalho.

Além disso, acreditamos que previsões diárias ou semanais seriam mais eficazes do que previsões mensais (com apenas um dado por mês) para capturar variações mais detalhadas.

CONSIDERAÇÕES FINAIS

Este estudo abordou a criminalidade em Resende/RJ por meio de uma análise aprofundada de dados históricos, destacando crimes contra o patrimônio e atividades ligadas a drogas. A partir de um estudo detalhado de roubos e furtos de celulares e veículos, bem como dos registros de apreensão, posse e tráfico de drogas, identificamos padrões relevantes que revelam o impacto de eventos sazonais e econômicos, além da existência de legislação específicas no comportamento da criminalidade. O levantamento ofereceu uma visão ampla sobre a dinâmica dos crimes analisadas.

O uso do Prophet representou uma etapa adicional na busca por previsões de tendências, mostrando-se útil para captar movimentos gerais, mas limitado na precisão dos picos e variações aleatórias que caracterizam os dados de criminalidade. O MAPE de 28% sugere que, embora o modelo capture bem as tendências, há oportunidades de aprimoramento com a inclusão de variáveis externas e previsões de curto prazo.

Concluimos que, para entender a complexidade da criminalidade e apoiar políticas de segurança pública de forma eficaz, é necessário um olhar abrangente que combine análise exploratória com métodos preditivos avançados. Mostrou-se a importância de integrar dados históricos, contexto local e tecnologias de aprendizado de máquina, apontando para uma abordagem híbrida em pesquisas futuras. Desta maneira, será possível aprimorar a precisão e oferecer percepções profundas para a segurança de Resende e outras cidades ou regiões.

REFERÊNCIAS

AGGARWAL, C. C. **Outlier Analysis**. 2ed. Cham: Springer, 2016.

GLOBO. **Monitor da Violência**. Disponível em: <https://g1.globo.com/monitor-da-violencia/>. Acesso em: 30 out. 2024a.

GLOBO. **Polícia Militar faz operação contra o tráfico de drogas em Resende**. G1, Rio de Janeiro, 17 jan. 2024. Disponível em: <https://g1.globo.com/rj/sul-do-rio-costa-verde/noticia/2024/01/17/policia-militar-faz-operacao-contr-o-traffic-o-de-drogas-em-resende.ghtml>. Acesso em: 30 out. 2024b.

DADOS DE CRIMINALIDADE EM RESENDE/RJ: TRATAMENTO E ANÁLISE COM FOCO EM PADRÕES E TENDÊNCIAS. Priscilla Maria Pires dos SANTOS; Ana Clara Castillo de Assis CALAZANS. JNT Facit Business and Technology Journal. QUALIS B1. ISSN: 2526-4281 - FLUXO CONTÍNUO. 2024 - MÊS DE NOVEMBRO - Ed. 56. VOL. 01. Págs. 336-355-. <http://revistas.faculdefacit.edu.br>. E-mail: jnt@faculdefacit.edu.br.

GOOGLE. **Colaboratory: Collaborative data science**. 2024. Disponível em: <https://colab.research.google.com/>. Acesso em: 30 out. 2024.

HARVEY, A. C.; PETERS, S. Estimation procedures for structural time series models. **Journal of Forecasting**, v. 9, n.2, p. 89-108, 1990.

IBGE. **Instituto Brasileiro de Geografia e Estatística**. Disponível em: <https://www.ibge.gov.br>. Acesso em: 30 out. 2024.

META INC. **Prophet: Forecasting at scale**. 2017. Disponível em: <https://facebook.github.io/prophet/>. Acesso em: 30 out. 2024.

MIRANDA, G. H. B.; FILHO, J. A. F. **Introdução à Ciência de Dados: Conceitos e Aplicações**. Rio de Janeiro: Elsevier, 2019.

MUKAKA, M. M. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. **Malawi Medical Journal**, v. 24, n. 3, p. 69-71, 2012. Disponível em: <https://www.ajol.info/index.php/mmj/article/view/81576>. Acesso em: 29 out. 2024.

ORIGUELA, L. A. **Weighted networks from Pearson, Spearman, and Kendall correlations to characterize the influence of events on the Brazilian stock market structure**. 2018. Dissertação (Mestrado em Administração de Organizações) – Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto. Disponível em: <https://pdfs.semanticscholar.org>. Acesso em: 29 out. 2024.

REZENDE, S. O. **Sistemas Inteligentes: Fundamentos e Aplicações de Inteligência Artificial**. Rio de Janeiro: LTC, 2003.

TRADING ECONOMICS. Brazil **Unemployment Rate**. Disponível em: <https://pt.tradingeconomics.com/brazil/unemployment-rate>. Acesso em: 30 out. 2024.

ZOU, K. H.; TUNCALI, K.; SILVERMAN, S. G. Correlation and simple linear regression. **Radiology**, v. 227, n. 3, p. 617-622, 2003. DOI: 10.1148/radiol.2273011499.