

ANÁLISE DOS DADOS DE VIOLÊNCIA E CRIMINALIDADE DO RIO GRANDE DO NORTE

Jesaías Kendson Carvalho Nunes (UFERSA)

Ana Cláudia Souza Vidal de Negreiros (UFERSA)

Miriam Karla Rocha (UFRN)

Resumo

A preocupação dos órgãos responsáveis pela segurança pública é compreender os motivos que contribuem para a grande incidência de crimes no estado do Rio Grande do Norte. Diante disso, este estudo propõe uma análise dos principais fatores relacionados aos índices de criminalidade, utilizando os dados fornecidos pela Polícia Civil do RN referentes ao ano de 2021. Por meio da preparação e clusterização desses dados, é possível a identificação de padrões e tendências relevantes. Os resultados obtidos evidenciam, por exemplo, que cidades com populações semelhantes apresentam níveis comparáveis de criminalidade. Compreender esses fatores essenciais que influenciam a ocorrência dos crimes é fundamental para orientar a formulação de políticas e estratégias eficazes de combate à criminalidade.

Palavras-chave: Índices de Criminalidade, Clusterização, Identificação de Padrões.

1. Introdução

De acordo com a Organização Mundial da Saúde (OMS, 2018), o Brasil é nono país mais violento do mundo, com uma taxa de 31,1 homicídios a cada 100 mil pessoas, perdendo apenas para Venezuela, com 56,8, no continente sul-americano. Em números totais, cerca de 1,2 milhão de pessoas perderam sua vida em crimes violentos no Brasil entre 1991 e 2017 (UNODC, 2019). Segundo Minayo (2016), é observado que, em regiões mais pobres das cidades há mais violência e as autoridades policiais cometem abusos, criando situações preconceituosas devido a raça e condição financeira. Segundo o autor, há um perfil do praticante da violência no Brasil, como: “faixa etária entre 14 e 29 anos, de cor preta ou parda, residentes em áreas mais vulneráveis, com escolaridade de, no máximo, ensino fundamental, com renda baixa e participantes de programas sociais e que vivem num ambiente familiar marcado por ausências dos pais, maus tratos ou agressões de todas as naturezas. Mas isso não significa que os moradores de certas áreas sejam considerados criminosos.”

Contudo, viver em um lugar tranquilo e sossegado é o desejo de qualquer indivíduo, a segurança pública é um dos fatores fundamentais para toda e qualquer sociedade e, tendo em vista dados atuais de segurança, o Rio Grande do Norte não é um estado que proporciona a segurança adequada a seus habitantes. Atualmente, vive-se em uma realidade que afeta a vida e segurança de todos, as drogas, por exemplo, terminam sendo um dos pivôs para o aumento dos homicídios e roubos (LIMA; PAIVA, 2023).

Nesse sentido, de acordo com Machado (2019), o Rio Grande do Norte apresenta a maior taxa de violência do país, com 62,8 mortes violentas por 100 mil habitantes em 2017, o que representa o maior índice entre os estados brasileiros. Além disso, a taxa de homicídios nesse estado aumentou em 320% entre 2006 e o ano anterior à pesquisa. Vale destacar que os jovens potiguaros, na faixa etária de 15 a 29 anos, são os mais afetados, com uma taxa de 152 mortes por 100 mil habitantes, representando um aumento de 482% desde 2006. Em comparação, em São Paulo, esse índice é significativamente menor, com 18 mortes por 100 mil habitantes. Esses dados foram obtidos do Sistema Único de Saúde (SUS) e compilados pelo Atlas da Violência, uma publicação anual do Instituto de Pesquisa Econômica Aplicada (Ipea) e do Fórum Brasileiro de Segurança Pública.

Diante do exposto, a análise de dados de segurança pode proporcionar informações importantes para que possam ser levadas em conta na tomada de decisão de estados e municípios quanto a ações que visem a redução dos níveis de criminalidade. Ao utilizar as técnicas de análises de dados que serão apresentados no decorrer deste trabalho, é possível detectar padrões comportamentais dos crimes, como também as cidades e regiões com um maior número de crimes. Esses dados concretos podem indicar um melhor destino dos recursos públicos, possibilitando uma melhor segurança para todos os potiguaros. Como também o acesso aos dados de segurança pública com mais facilidade e clareza ajuda a melhorar a confiança da população nas autoridades.

Vale ressaltar que com o passar do tempo, o volume de informações disponíveis para análise está cada vez maior. Essa expansão traz desafios significativos, uma vez que a quantidade de dados dificulta a análise e compreensão. Esse aumento no volume de dados apresenta um desafio importante: como transformar essa grande quantidade de informações em interpretações úteis e relevantes? A incapacidade de realizar análises eficientes, pode resultar em escolhas erradas, como a perda de oportunidades significativas e ineficácia nas operações.

Os dados revelam informações que ainda não trazem retornos relevantes, por ainda não terem sido tratados. A análise serve para isso, ao tratar são extraídas informações importantes que serão usadas em diversas situações. Destaca-se a necessidade crescente de ferramentas que possibilitem uma eficaz coleta, organização e interpretação desses dados, em áreas como transporte, saúde, serviço público, entre outros (BASTOS, 2023).

A análise de dados é o processo de técnicas estatísticas e lógicas usado como métrica para obter informações diante de determinados dados. Seu foco é conseguir informações importantes a partir dos mesmos. Com isso, é possível tomar decisões concretas em uma solução de um problema, para alcançar o objetivo. Como o volume de dados cada vez maior, as análises necessárias passam por algumas dificuldades (MONTENEGRO, 2023).

Portanto, se torna imprescindível para a tomada de decisões importantes. Sem uma compreensão dos padrões presentes nos dados, as decisões se baseiam em suposições inadequadas, levando a resultados irrelevantes. Como também a necessidade de ter estratégias e interpretações que possam tratar os desafios específicos dos grandes volumes de dados, como a carência de um processamento mais rápido e garantias em termos de qualidade e segurança deles (TAYRANE, 2022).

O objetivo deste trabalho é realizar um estudo dos principais fatores que contribuem para o aumento da criminalidade no Rio Grande do Norte. Através da análise desses fatores, é possível identificar possíveis ações que possam ser implementadas para reduzir a incidência desses crimes no estado. O reconhecimento desses elementos pode ser fundamental para o desenvolvimento de estratégias eficazes. Além disso, a identificação das origens dos crimes permitirá direcionar melhorias específicas para áreas vulneráveis. A segurança de todos os cidadãos é de extrema importância e, portanto, evitar a adoção de medidas que coloquem em risco a vida das pessoas.

2. Referencial Teórico

Para uma compreensão mais completa deste trabalho, é fundamental explorar alguns tópicos teóricos que fundamentam o seu processo de criação. Nesse sentido, é importante abordar três temas específicos durante o referencial teórico: a análise de dados, criminalidade e a clusterização, que serão abordados nas seções 2.1, 2.2 e 2.3, respectivamente.

2.1 Análise de dados

Dados são informações registradas ou resultantes de alguma medição, que servem como matéria prima. A disponibilidade desses dados cria oportunidades importantes para a obtenção de informações significativas. Eles podem ser adquiridos por meio da percepção sensorial, como a observação direta, ou por meio da execução de procedimentos de medição ou através de consultas em algum banco de dados (UNICAMP, 2020).

Por sua vez, uma operação de medição é um método que mostra a relação entre a propriedade de interesse e um conjunto de classificações. Os sistemas de medição incluem unidades padrão para diversas grandezas e procedimentos para determinar valores de acordo com essas informações pretendidas (SILVA, 2020).

Assim, a análise de dados é indispensável em uma época de informações em grande escala e digitais. É essencial para os insights, além da tomada de decisões baseadas em evidências e no avanço do conhecimento em diferentes áreas. Para realizar a análise de dados é necessário algumas técnicas e métodos para interpretar e englobar informações em determinados bancos de dados, que podem ser grandes e complexos (BULIAN e ALENCAR, 2017).

Fundamentada em princípios estatísticos e matemáticos existem diversas técnicas de análise de dados, desde análises estatísticas básicas até métodos mais robustos. Essa análise usa a tecnologia da informação, modelos matemáticos e estatísticos, bem como técnicas de visualização, a fim de extrair informações valiosas de conjuntos de dados brutos (MAKERZINE, 2023).

Além disso, é dividida em algumas etapas importantes, como a definição do objeto de análise, a escolha o tipo de dados que serão analisados, definição de como será a mensuração dos dados, realização da coleta de dados, a análise dos dados e por fim sua interpretação. Para a coleta de dados deve existir uma estratégia para coletar e agregar os dados que o ajudarão nessa análise. Após coletar os dados, eles não estarão prontos para serem analisados. Logo, os dados devem ser limpos. Durante esse processo de limpeza, devem ser removidos os erros, informações que podem estar duplicadas e discrepâncias, como também dados irrelevantes para a análise (MONTENEGRO, 2023).

Após a “limpeza” dos dados, a análise pode começar. Há alguns métodos, como: Análise descritiva; análise diagnóstica; análise preditiva; análise prescritiva. A análise descritiva concentra-se em descrever eventos que já aconteceram com base em fatos e resultados conhecidos, a análise diagnóstica explora as razões relacionadas a esses eventos, identificando padrões e relações nos dados. A análise preditiva, embora não permita prever a certeza do futuro, busca tendências antecipadas com base em dados históricos, enquanto a análise prescritiva fornece recomendações estratégicas com base nas estratégias, orientando esforços para alcançar resultados maximizados. Cada uma mostra a sua importância diante de uma tomada de decisão, conforme demonstrado na Figura 1 (MONTENEGRO, 2023).

Figura 1: Tipos de análise



Fonte: TAYRANE (2022)

2.2 Criminalidade

Segundo Alvarenga (2017), a criminalidade não é apenas a contagem dos crimes ocorridos em um determinado tempo e local. É um fenômeno social, que vem de uma junção de fatores das bases da vida social em um contexto específico. Surgindo da interação de diversos elementos, a criminalidade reflete

uma série de condições que comprometem gravemente a ordem social e as condições existenciais em uma comunidade ou região em particular

As autoridades tem o compromisso de buscar métodos e estratégias para instruir, prevenir e reprimir toda e qualquer forma de violência em qualquer região do país. Mas, devido a diversos problemas, tanto legislativas quanto na efetivação por parte das autoridades, a sociedade continua a enfrentar os desafios impostos por comportamentos violentos realizados por menores infratores e por indivíduos que realizam assaltos, sequestros e homicídios. As leis que deveriam resguardar os cidadãos praticamente não existem, ao passo que as falhas presentes nas legislações vigentes acabam, infelizmente, incentivando a prática criminosa (LUCENA,2022).

2.3 Clusterização

A clusterização surgiu como uma abordagem fundamental em análise de dados multivariados. Utilizando métodos numéricos e considerando as variáveis de cada caso, seu propósito é agrupar automaticamente os dados similares em conjuntos distintos, conhecidos como clusters ou agrupamentos (CASSIANO, 2014).

O pesquisador busca ordenar e agrupar observações, é recomendado realizar uma análise de agrupamentos. Nessa parte, o foco é na organização em clusters para o referido estudo. Pode-se optar por determinar a quantidade ideal de clusters formados ou especificar o número desejado de grupos com base em critérios específicos. Posteriormente, é essencial testar como as observações são organizadas e alocadas dentro dos grupos determinados. Além disso, é importante conduzir análises utilizando softwares apropriados (FAVERO, 2017).

Figura 2: Análise de agrupamentos



Fonte: FAVERO (2017)

A ideia principal é garantir que os elementos dentro de cada grupo sejam semelhantes entre eles, como também que sejam diferentes dos outros *clusters*. Basicamente, o processo de clusterização busca homogeneidade dentro de cada grupo e a diversidade entre os *clusters*. Um dos pontos positivos da utilização da clusterização é que, ao reunir dados semelhantes, torna-

se viável descrever de forma mais precisa e econômica as singularidades de cada grupo. Essa abordagem proporciona uma compreensão mais ampla do conjunto de dados de modo geral, permitindo o desenvolvimento de métodos de classificação para novos dados e a identificação de correlações os dados que poderiam passar despercebidas sem o uso dessas técnicas. (CASSIANO,2014).

3. Banco de dados

Os indicadores de violência e criminalidade são essenciais para contribuir na análise e trazer melhorias para a segurança de todos. Mas, o registro de fatores contribuintes requer uma disponibilidade de dados que são imprescindíveis para o método a ser utilizado. A enumeração de fatores contribui e deve ser completa, como também, o mais acessível para facilitar o trabalho de registro e análise destas informações. A seleção do método correto para desenvolvimento de uma pesquisa depende do objetivo e posteriormente dos objetivos que o pesquisador deseja atingir.

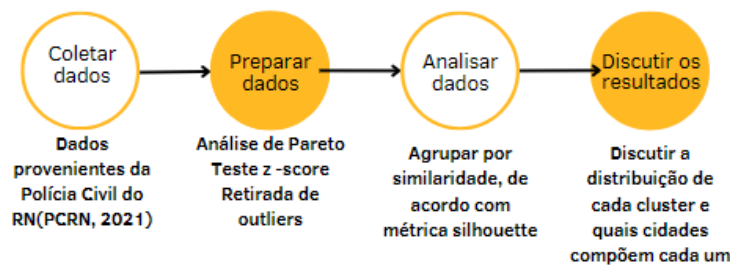
O banco de dados utilizado neste estudo é proveniente da base de dados da Polícia Civil do Rio Grande do Norte (PCRN,2021). O banco de dados é composto por 78734 observações de criminalidades no estado do Rio Grande do Norte. Estes dados constituem cerca de quarenta variáveis, que foram averiguados e colocados em intervalos corretos de numeração. Este banco de dados conta com informações referentes a 2021 (janeiro a dezembro).

4. Procedimentos metodológicos

Esse estudo é caracterizado como uma pesquisa aplicada, que de acordo com Gil (2019) consiste na busca por dados que se tornam em aplicações práticas, com o objetivo de resolver problemas relacionados ao estudo. O estudo se enquadra como quantitativo, assim, utilizando um banco de dados da Polícia Civil do Rio Grande do Norte (PCRN, 2021), o método irá seguir uma sequência bem estabelecida que envolve a coleta de dados, preparação dos dados, análise dos dados e, por fim, a discussão e análise dos resultados. Quanto ao objetivo, descritivo, mostra os fatores que determinam a ocorrência do fenômeno e os procedimentos será uma pesquisa bibliográfica.

Os dados foram coletados, logo após foram processados em variáveis padronizadas, onde as variáveis irão resultar em uma média igual a 0 e um desvio padrão igual a 1, utilizando a técnica z-score, que mede quantos desvios padrão acima ou abaixo da média um ponto está (VILELA JUNIOR, 2020). Logo em sequência, utilizando o software JASP, foi aplicada a clusterização, que segundo Hair (2005) é o agrupamento de dados por similaridade, tornando uma única análise, o que antes era obrigatório ser feito por diversas vezes e por fim, a discussão de como cada cluster foi distribuído (ver Figura 3).

Figura 3 - Procedimento da pesquisa



Fonte: autoria própria (2024)

Baseado no método hierárquico, assim será identificado o número de grupos que serão agrupados por similaridade e a consequente análise de como eles serão distribuídos em cada cluster

4.1 Coleta de dados

O banco de dados utilizado foi o relatório estatístico de crimes no Rio Grande do Norte, com média de 140 cidades e 40 variáveis de crimes, do ano de 2021. Esses dados servem para divulgar as estatísticas dos crimes no referido estado para o governo estadual ter o conhecimento e averiguar possíveis ações para mitigar tais feitos. Esse conjunto de dados pode ser encontrado no site oficial da polícia civil do RN.

Conforme será explicado nas próximas seções, os dados foram preparados utilizando a análise de Pareto e do z-score, e os crimes definidos para o presente artigo foram ameaça, estelionato, furto, roubo e roubo com aumento de pena se há concurso de duas ou mais pessoas, conforme Quadro 1, que foi construído de acordo as informações da Polícia Civil do Distrito Federal (PCDF,2024).

Quadro 1 – Descrição das variáveis

Natureza das observações	Descrição
Ameaça	Ocorre quando o indivíduo afirma que irá causar um mal injusto ou grave à outra pessoa (Art. 147, do Código Penal)
Estelionato	Ocorre quando o indivíduo tenta ter vantagem ilícita para si próprio ou para outra pessoa em prejuízo da vítima, induzindo-a ou mantendo-a em erro, mediante qualquer meio fraudulento (art. 171, Código Penal)
Furto	Ocorre quando uma ou mais pessoas subtraem (retiram) coisa alheia (objetos, documentos valores) de outra pessoa sem empregar violência física ou grave ameaça.
Roubo	Ocorre quando o autor empregou grave ameaça (utilização de arma de fogo ou outro instrumento) para subtração do(s) objeto(s)

Roubo com aumento de pena se há concurso de duas ou mais pessoas	Ocorre quando há mais de dois ou três autores, que empregaram grave ameaça (utilização de arma de fogo ou outro instrumento) para subtração do(s) objeto(s)
--	---

Fonte: autoria própria (2024)

4.2 Preparação e análise dos dados

Após a coleta, foram processados e padronizados para evitar inconsistências, mas ainda não estavam preparados para serem analisados, devido ao seu grande volume, então foi utilizada a diagrama de Pareto e a retirada dos *outliers*, que são valores muito discrepantes dos demais.

Para ajudar no processamento, adotou-se a técnica de Pareto, conforme Slack (2018), em qualquer processo que busca uma melhoria, é essencial diferenciar entre o que é realmente importante e o que é menos relevante. O diagrama de Pareto, tem por objetivo identificar as questões menos importantes em contraposição às muitas questões imprescindíveis. É relativamente simples, envolvendo a classificação de cada observação de acordo com a ordem de importância dos problemas. A análise de Pareto foca na observação de que um número relativamente pequeno de causas é responsável pela maioria dos efeitos.

Logo em seguida, os *outliers* foram identificados como as cidades com maior população (Grande Natal e Mossoró) e retirados devido ao seu alto número de valor máximo e mínimo conforme a Tabela 1.

Tabela 1 – Padronização com a técnica z-score

	Ameaça	Estelionato	Furto	Roubo	Roubo com aumento de pena se há concurso de duas ou mais pessoas
Média	89,0209	165,2949196	160,4	139,08	27,89094472
Desvio padrão	403,247	934,9508571	905,72	841,89	146,5398552
Mínimo	0	0	0	0	0
Máximo	3538	8289	8165	7447	1138

Fonte: autoria própria (2024)

Recomendadas por Rousseeuw et al (2017), os dados que estiverem fora desses limites serão determinados como *outliers* e serão retirados para evitar incoerência na formação dos *clusters* (ver Tabela 2).

Tabela 2 - Padronização com a técnica z-score

	Ameaça	Estelionato	Furto	Roubo	Roubo com aumento de pena se há concurso de duas ou mais pessoas
Média	27,7862	31,38187741	36,011	15,712	3,465309573
Desvio padrão	40,9922	56,33273611	66,771	39,377	8,052348936
Mínimo	0	0	0	0	0
Máximo	201	354	412	251	51

Fonte: autoria própria (2024)

4.3 Etapa de discussão dos resultados

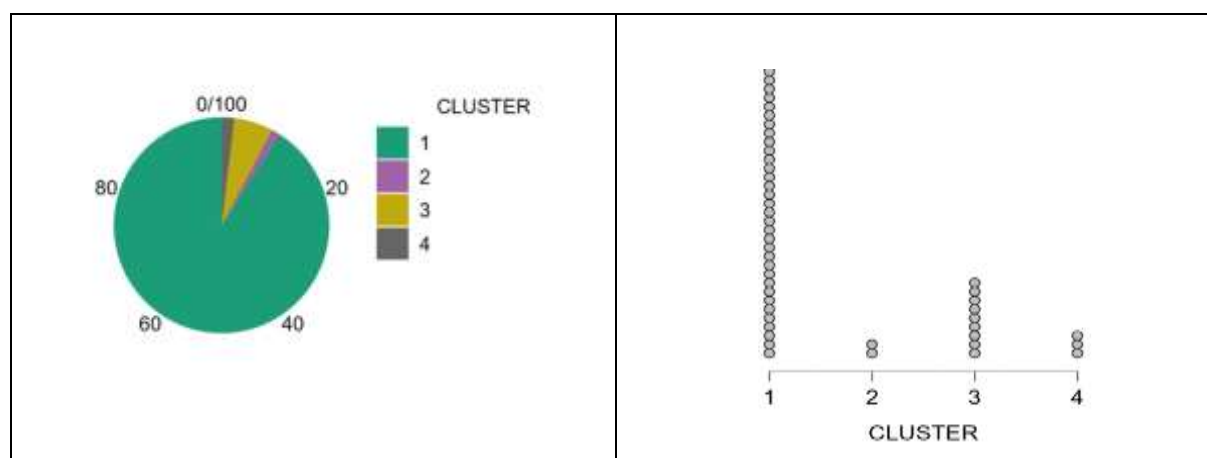
Ao final, os dados obtidos durante o desenvolvimento deste trabalho foram discutidos e analisados, e foram observadas informações relevantes sobre a situação observada e o problema real identificado, os quais estão apresentados na seção 5 deste documento. Durante essa análise, foi utilizada a técnica de clusterização para examinar os resultados obtidos, que consistem nas descrições dos agrupamentos identificados. Além disso, foi conduzida uma investigação sobre a distribuição de cada *cluster*, que foi feita através da métrica *silhouette*, com o intuito de identificar possíveis estratégias de adaptação e ações que poderiam ser implementadas para reduzir o número de crimes no estado do Rio Grande do Norte. Após a abordagem desses tópicos, foi viável realizar uma análise abrangente e visualizar os resultados da pesquisa, os quais estão detalhados na seção 5.

5. Resultados

Após todo o procedimento os dados estão prontos para serem analisados. A análise é composta por verificações, que foram feitas através software JASP. O banco de dados foi reduzido e apto para realizar a clusterização, que começa com o a quantidade de agrupamentos que é definida pela técnica *silhouette*, nessa medição, a análise varia entre -1 e 1, ou seja, quanto mais rente de +1, mostra que a similaridade dos dados analisados é maior em seu *cluster* do que com relação aos outros *clusters*, confirmando que a clusterização seria melhor nessa análise (Rousseeuw, 1987). O *silhouette* obtido foi 0,720 e foram formados quatro *clusters*, que são similares em cidades que possuem a mesma quantidade de habitantes e sua consequente quantidade de crimes.

Conforme a Figura 4, é possível observar que o “*cluster 1*” prevalece devido a maioria das cidades terem uma população pequena e como consequência a mesma quantidade média de crimes.

Figura 4 - Gráfico de pizza e diagrama de pontos: divisão dos clusters

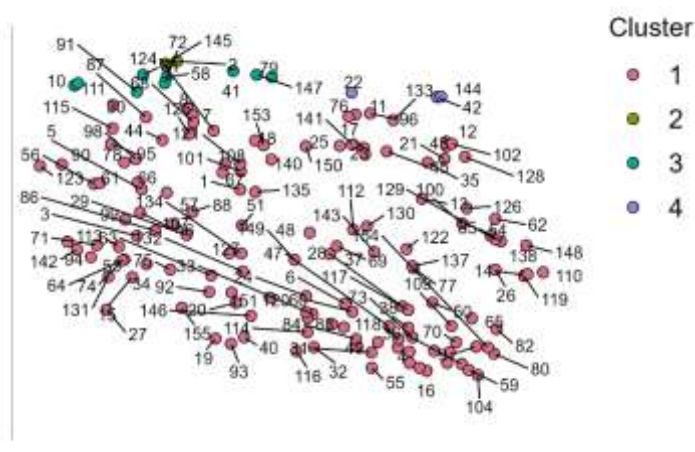


Fonte: autoria própria (2024)

Para uma melhor visualização dos quatro *clusters* formados, foi utilizado o *plot*, que é uma ferramenta do software utilizado, onde é identificado graficamente onde os *clusters* estão localizados após a distribuição (ver figura 5).

Após passar por vários testes, os *clusters* estão definidos e agora é necessário a interpretação de cada um. As cidades que compõem cada *cluster* são as que possuem a mesma quantidade com relação a população, em valores aproximados de acordo com o IBGE(2022), alguns exemplos são: a “observação 10” e “111” no “*cluster 3*”, que são Apodi e Santa Cruz, respectivamente, com 36 mil e 37 mil habitantes, aproximadamente, como também a “56” que corresponde a Jucurutu com 18 mil habitantes e a “61” que é Jardim de Piranhas com 14 mil, é observado que nesse *cluster* a população não é a mesma, mas continua com valores aproximados. Por fim, no “*cluster 1*”, é visualizado que a “observação 2” corresponde a Assú com 56 mil habitantes e a “81” a Nova Cruz com 34 mil habitantes. Portanto, fica comprovado que as cidades que possuem populações semelhantes ou muito próximas, possuem a mesma quantidade de crimes.

Figura 5: *Clusters*



Fonte: autoria própria (2024)

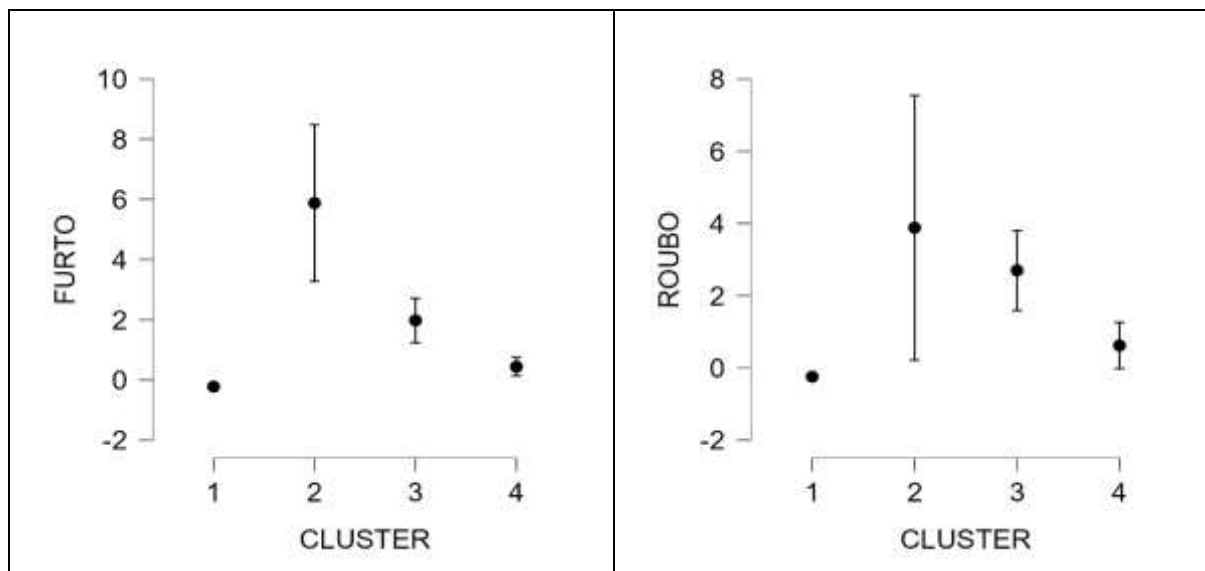
5.1 Gráficos de intervalo

Os gráficos de intervalos (ver figuras 6, 7, e 8) são essenciais para a visualização da distribuição e análise dos *clusters*, para explicar o comportamento de cada variável. Possuem linhas verticais que representam os intervalos de variação para cada *cluster* em comparação aos demais dados. Cada linha possui por um ponto central que indica a média dos valores nesses *clusters* (IBM,2024).

Na “variável furto”, é observado que cada um está em um intervalo diferente, onde no “*cluster 1*” pode-se ver as cidades de Messias Targino e Triunfo Potiguar, que estão no mesmo intervalo, já no “*cluster 2*” foram observados as cidades de Açu e Tibau Do Sul, também no mesmo intervalo.

Na “variável roubo”, o “*cluster 2*” tem como participantes as duas últimas cidades citadas, com uma variação maior, devido a certa diferença entre as populações das mesmas, mas ainda se equiparam devido ao grande fluxo de turistas da cidade de Tibau do Sul, que é reconhecida internacionalmente por promover turismo sustentável com a praia de pipa (MTUR,2023) (ver Figura 6).

Figura 6 – Gráficos de intervalo

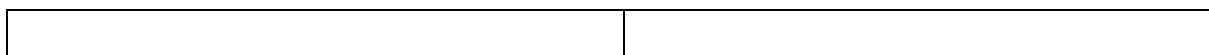


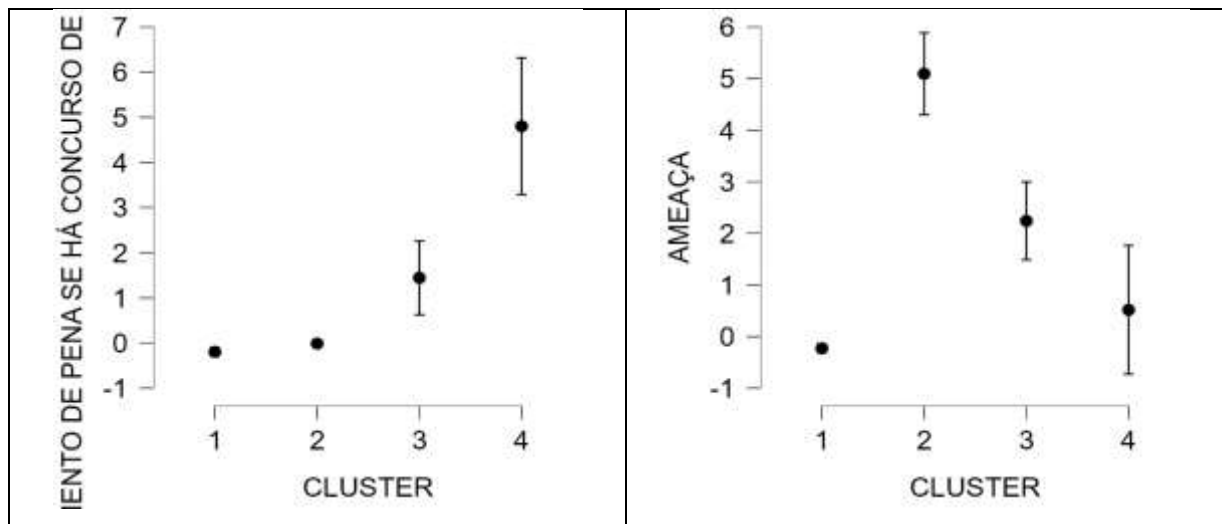
Fonte: autoria própria (2024)

Na análise seguinte, os “*clusters 1,2,3 e 4*” que correspondem ao roubo com aumento de pena se há concurso de duas ou mais pessoas, os dois primeiros *clusters* tiveram uma dispersão menor, o que significa que os intervalos são mais estreitos, indicando que os referidos *clusters* são mais semelhantes. As cidades identificadas foram Apodi, Goianinha, João Câmara, Macau, Monte Alegre, Nova Cruz, Pau Dos Ferros, Santa Cruz e Touros, que de acordo com o último censo possuem o número de habitantes muito próximo entre cada uma delas (IBGE,2022).

Na ameaça, o “*cluster 2*” teve um intervalo maior, que corresponde também as cidades de Açu e Tibau do Sul, já justificados anteriormente. Já o “*cluster 3*” são as cidades de Touros e Macau, semelhantes com relação a sua população de acordo com o último censo com 33 mil e 32 mil habitantes, respectivamente. No “*cluster 4*” estão as cidades de Governador Dix-Sept Rosado e Tibau. É importante ressaltar que ambos estão em intervalos diferentes e mais largos (IBGE,2022) (ver Figura 7).

Figura 7 - Gráficos de intervalo

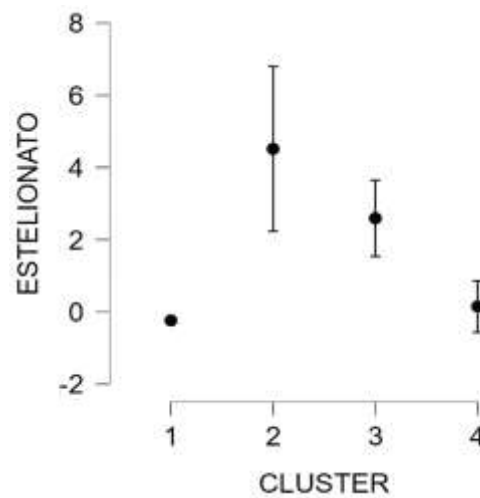




Fonte: autoria própria (2024)

Na Figura 8, do estelionato, o “*cluster 1*” é bem compacto, já o *cluster* dois prevaleceu novamente e teve um intervalo maior, assim como o “3” e o “4”, sendo em proporção menor. Foram identificadas as cidades de Canguaretama e Tibau, da mesma forma da primeira análise, há uma diferença entre a população, mas que foi equiparada novamente devido ao grande fluxo turístico de Tibau (ver Figura 7)

Figura 8 - Gráfico de intervalo



Fonte: autoria própria (2024)

6. Considerações finais

Os dados sobre criminalidade desempenham um papel imprescindíveis na identificação dos fundamentos que são importantes e na compreensão das estatísticas disponíveis. Os registros convencionais fornecem percepções sobre onde e quando os crimes ocorreram. Diante da complexidade e da grande quantidade de variáveis disponíveis, é observado que a preparação dos dados, aliada à utilização da métrica de avaliação de clusters conhecida como silhouette, facilitou a identificação de padrões mais claros na separação dos crimes. Nesse contexto, foram definidos padrões orientados por variáveis como ameaça, estelionato, roubo, furto e roubo com agravante de concurso de duas ou mais pessoas, proporcionando uma análise primordial.

Compreender os principais fatores associados à ocorrência dos crimes é crucial para o desenvolvimento de planos e estratégias eficazes de redução da criminalidade. Por meio das análises realizadas, foi possível obter uma compreensão mais aprofundada de como os crimes se manifestam em diferentes regiões do estado. Além disso, foram identificados padrões distintos de comportamento das variáveis em resposta a variações em cada um deles.

Os métodos de clusterização aplicados demonstraram resultados satisfatórios e importantes. Vale ressaltar a importância de empregar outros tipos de métodos na análise de bancos de dados, uma vez que esses métodos podem oferecer opções complementares e impulsionadoras para uma compreensão mais completa dos fenômenos estudados.

Referências

ALVARENGA, Dílio Procópio Drummond de. **Crime e criminalidade**. [S.L.]: Revista Jus Vigilantibus, 2017.

Disponível em: <https://www.jusbrasil.com.br/artigos/crime-e-criminalidade/448338731#:~:text=Criminalidade%2C%20por%20sua%20vez%2C%20n%C3%A3o,tempo%20e%20em%20certo%20lugar> .. Acesso em: 25 abr. 2024.

BASTOS, Athena. **Análise de dados: uma ferramenta para criar melhores estratégias de negócio**. São

Paulo: Alura, 2023. Disponível em: <https://www.alura.com.br/empresas/artigos/analise-de-dados>. Acesso em: 25 mar. 2024

BULIAN, Braian de Souza; ALENCAR, Cícero Aparecido. **A Importância do Big Data para a Compreensão dos Hábitos de Compra dos Consumidores, Aumentando as Vantagens Competitivas no Comércio**

Varejista. Revista Científica Multidisciplinar Núcleo do Conhecimento. Edição 08. Ano 02, Vol. 03. pp 18-41, novembro de 2017. Disponível em: <https://www.nucleodoconhecimento.com.br/administracao/big-data-comercio-varejista> Acesso em: 25 fev. 2024.

CASSIANO, K. **Análise de Séries Temporais Usando Análise Espectral Singular (SSA) e Clusterização de Suas Componentes Baseada em Densidade**. 2014. Tese (Doutorado em Engenharia Elétrica) – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro – PUC-Rio, Disponível em:

https://www.maxwell.vrac.puc-rio.br/24787/24787_5.PDF . Acesso em: 30 de março de 2024.

Como aplicar a análise de dados na tomada de decisão? Rio de Janeiro: Ipsense, 2020. Disponível em:

<https://www.ipsense.com.br/tecnologia/como-aplicar-a-analise-de-dados-na-tomada-de-decisao/> . Acesso em: 20 set. 2023

Escritório das Nações Unidas sobre Drogas e Crime, UNODC. **Brasil tem segunda maior taxa de homicídios da América do Sul, diz relatório da ONU**. Brasília, 2019. Disponível em: <https://www.unodc.org/lpo-brazil/pt/frontpage/2019/07/brasil-tem-segunda-maior-taxa-de-homicidios-da-amrica-do-sul--diz-relatorio-da-onu.html>.

Acesso em: 26 mar. 2024.

FÁVERO, Luiz Paulo; BELFIORE, Patrícia. **Manual de Análise de Dados - Estatística e Modelagem Multivariada com Excel®, SPSS e Stata**. Rio de Janeiro: Gen Ltc, 2017.

GIL, Antonio Carlos. Métodos e Técnicas de Pesquisa Social. 7. ed. São Paulo: Atlas, 2019.

GUANIS DE BARROS VILELA JUNIOR. ESTATÍSTICA: **Teste Z (ou Escore Padronizado)**. Campinas: Centro de Pesquisas Avançadas em Qualidade de Vida, 2020. Disponível em: https://www.cpaqv.org/estatistica/teste_z.pdf. Acesso em: 05 mar. 2024.

HAIR, Joseph F.; BLACK, William C.; TATHAM, Ronald L.; ANDERSON, Rolph E.; TATHAM, Ronald L.. **Análise Multivariada de Dados**. 5. ed. São Paulo: Bookman, 2005.

IBM. **Tipos de gráficos de indicadores de alcance**. Disponível em: <https://www.ibm.com/docs/pt-br/cognos-analytics/12.0.0?topic=types-range-indicator-charts>. Acesso em: 05 mar. 2024.

Instituto Brasileiro de Geografia e Estatística, IBGE. **IBGE Cidades Rio Grande do Norte Panorama**. Brasília, 2022. Disponível em: <https://cidades.ibge.gov.br/brasil/rn/panorama>. Acesso em: 25 mar. 2024.

LIMA, Renato Sérgio de; PAIVA, Luiz Fábio Silva. **Por que a crise no Rio Grande do Norte é um retrato do brasil**. [S.L.]: Revista Piauí, 2023. Disponível em: <https://piaui.folha.uol.com.br/por-que-crise-no-rio-grande-do-norte-e-um-retrato-do-brasil/> . Acesso em: 20 mar. 2024.

LUCENA, Jorge. **A violência e a criminalidade**. [S.L.]: Brasil Escola, 2022. Disponível em: <https://meuartigo.brasilecola.uol.com.br/atualidades/a-violencia-criminalidade.htm> . Acesso em: 25 abr. 2024.

MACHADO, Leandro. **Por que o Rio Grande do Norte é o pior lugar para ser jovem no Brasil**. Natal: Bbc, 2019. Disponível em: <https://www.bbc.com/portuguese/brasil-49187027>. Acesso em: 25 mar. 2024.

MIGUEL, P. A. C.; **Metodologia de pesquisa em engenharia de produção e gestão de operações**. 3. ed. Rio de Janeiro: Elsevier Ltd, 2010.

MINAYO, Maria Cecília de Souza. **Saúde e Segurança Pública: Desafios em territórios marcados pela violência**. Rio de Janeiro: Filhos do Vento, 2016.

MONTENEGRO, Bruna. **Análise de dados: metodologia, tipos e técnicas mais usadas**. [s.l.]: Ebac, 2023. Disponível em: <https://ebaonline.com.br/blog/analise-de-dados-metodologia-tecnicas-tipos>. Acesso em: 22 set. 2023.

MTUR. **Tibau do Sul é reconhecida internacionalmente por promover turismo sustentável**. São Paulo: Brasil Travel News, 2021. Disponível em: <https://brasiltravelnews.com.br/noticias/tibau-do-sul-e-reconhecida-internacionalmente-por-promover-turismo-sustentavel/>. Acesso em: 24 mar. 2024.

Polícia Civil do Distrito Federal. **Delegacia Eletrônica**. Disponível em: <https://www.pcdf.df.gov.br/servicos/delegacia-eletronica>. Acesso em: 06 abr. 2024

Polícia civil do Rio Grande do norte. **Estatísticas da Polícia Civil**. Natal. Disponível em: <http://www.policiacivil.rn.gov.br/Conteudo.asp?TRAN=ITEM&TARG=5517&ACT=&PAGE=0&PARM=&LB=L=Dados+Estat%EDsticos>. Acesso em: 10 fev. 2024

Redação Tribuna do Norte. **Criminalidade cresce em Natal**. Natal, 2006. Disponível em:
<https://tribunadonorte.com.br/natal/criminalidade-cresce-em-natal/>. Acesso em: 26 mar. 2024.

ROUSSEEUW, Peter J. **Silhouettes: a graphical aid to the interpretation and validation of cluster analysis**. 1987, J Comput Appl Math 20:53–65

ROUSSEEUW, Peter J.; HUBERT, Mia. **Anomaly detection by robust statistics**. WIREs Data Mining and Knowledge Discovery. Lovaina, Belgium, p. 3-4, nov. 2017.

SILVA, Paulo Roque Martins. **Conceitos Técnicos da Metrologia**. [S.L.]: Instituto de Metrologia, Qualidade e Tecnologia – Inmetro, 2020. Disponível em:
<https://repositorio.enap.gov.br/bitstream/1/6687/5/M%C3%B3dulo%20%20Conceitos%20T%C3%A9cnicos%20em%20Metrologia.pdf>. Acesso em: 25 fev. 2024.

SLACK, N.; BRANDON-JONES, A.; JOHNSTON R. **Administração da Produção**. 8ª Edição. Editora Atlas, 2018.

TAYRANE. **Metodologia de análise de dados: um guia completo sobre o tema**. São Paulo: Ploomes, 2022. Disponível em: <https://blog.ploomes.com/analise-de-dados/> . Acesso em: 05 fev. 2024

TERRA, Rodrigo. **O que é Análise de dados?** SI: Makerzine, 2023. Disponível em:
<https://www.makerzine.com.br/programacao/o-que-e-analise-de-dados/> . Acesso em: 21 set. 2023

Universidade Estadual de Campinas, Unicamp, 2020. **O que são dados?**, Disponível em:
<https://www.ime.unicamp.br/~hildete/dados.pdf> , Acesso em: 22 set. 2023