

VIOLÊNCIA NO RECÔNCAVO BAIANO: UMA ANÁLISE MULTIVARIADA

*Luis Henrique Costa Ferreira**

RESUMO: Esta pesquisa, de caráter exploratório, teve como objetivo estudar as inter-relações entre as variáveis, taxa de matrícula no ensino médio, taxa de matrícula no ensino fundamental, densidade demográfica, rendimento nominal mensal, população ocupada, Índice GINI, PIB *per capita* e taxa de homicídio, representando as dimensões Educação, Economia, Trabalho e Renda, Desigualdade Social e Violência, com o emprego da análise multivariada. O objeto do estudo foram os municípios que constituem o Território de Identidade do Recôncavo Baiano. Como resultados observou-se que as variações entre os municípios não são influenciadas, de modo expressivo, pelas taxas de homicídios. Elas são explicadas pelas dimensões Trabalho e Renda e Educação. A dimensão Violência recebe fortes influências das taxas de homicídios, dos PIB(s) e das taxas de matrículas no ensino médio. Sendo que a correlação positiva com a última leva a despertar um sinal de alerta sobre o significado dela como indicador de indutor de inclusão social e demonstra que no caso analisado a simples matrícula no ensino não significa que esteja ocorrendo uma política de educação inclusiva..

Palavras-chave: Violência. Educação. Bahia. Análise multivariada. Homicídio.

* Mestre em Segurança Pública, Justiça e Cidadania. Possui graduação em Engenharia Civil pela Universidade Católica do Salvador (1986), graduação em Bacharelado Em Direito pela Universidade Católica do Salvador (1994) e graduação em Licenciatura Em Matemática pela Universidade Católica do Salvador (1986). Atualmente é Delegado de Polícia Civil Estado da Bahia.

1. INTRODUÇÃO

É do conhecimento comum que o planejamento e o acompanhamento de políticas públicas, na maioria das vezes, irão depender do monitoramento e da interpretação de muitas variáveis dos mais diversos matizes. As políticas públicas para o controle da violência não escapam dessa realidade. A violência em um território, ou em uma comunidade, se desenvolve em um contexto multivariado e para que se possa tentar controlá-la é imprescindível estudar como essas variáveis se associam, interagem umas com as outras e quais os respectivos impactos delas, individualmente e em conjunto, sobre o território estudado. Santos e Gomes (2006) observaram que a alocação de recursos para conter a violência não deve levar em consideração apenas as taxas de criminalidade, mas também outras variáveis que incorporem as diferenças entre os territórios ou entre as comunidades. Os valores das relações sociais devem ser vivenciados nos campos da cultura, da educação, da política e da economia, pois constituem a base dos direitos humanos (SILVA, 2010, p. 8). Silva (2006, p 16) admite a existência de uma cultura de violência que se espalha por todo o horizonte social.

Esta pesquisa, de caráter exploratório, teve como objetivo estudar as inter-relações entre as variáveis, **taxa de matrícula no ensino médio, taxa de matrícula no ensino fundamental, densidade demográfica, rendimento nominal mensal, população ocupada, Índice GINI, PIB per capita** e a **taxa de homicídio**, representando as dimensões **Educação, Economia, Trabalho e Renda, Desigualdade Social e Violência**, com o emprego da Estatística Multivariada. E, para tanto, fez uso da análise de agrupamento e da análise de componentes principais (ACP). O objeto de estudo foram os 19 municípios que constituem o Território de Identidade do Recôncavo Baiano, a opção por um Território de Identidade é uma tentativa de nivelar possíveis influências de alguma variação cultural. Os dados trabalhados foram

colhidos na Secretaria do Planejamento do Estado da Bahia, no IBGE e no Mapa da Violência.

Como resultado ficou demonstrado que as variações entre os municípios do Território de Identidade não são influenciadas de maneira expressiva pelas taxas de homicídios, que representam a dimensão **Violência**. Elas são melhores explicadas pelas dimensões Trabalho e Renda e Educação. Sendo que o componente **Violência** recebe fortes influências das **taxas de homicídios**, dos **PIB(s)** e das **taxas de matrículas no ensino médio**. Surgiu que estas últimas estão correlacionadas de maneira positiva com a Violência, o que leva a despertar um sinal de alerta sobre o significado da variável **taxas de matrículas no ensino médio** como indutor de inclusão social e demonstra que, no caso estudado, a matrícula não significa que esteja ocorrendo uma educação inclusiva.

Além desta introdução, este trabalho contém os seguintes tópicos: Referencial teórico, Desenvolvimento da Pesquisa e Conclusões.

2. REFERENCIAL TEÓRICO

Hair Jr *et al.* (2009) explica que análise de agrupamento é um conjunto de técnicas multivariadas cuja finalidade é agregar objetos com base nas características que eles possuem. A análise de agrupamento classifica objetos de modo que cada objeto é semelhante aos outros no agrupamento com base em um leque de características escolhidas. Ela tem sido chamada de análise Q, construção de tipologia, análise de classificação, análise de *cluster* ou taxonomia numérica e tem por objetivo definir a estrutura dos dados colocando as observações mais parecidas em grupos. Os agrupamentos resultantes de objetos devem então exibir elevada homogeneidade interna e elevada heterogeneidade externa. Silvestre (2007) a cita como uma das técnicas cuja finalidade é a produzir agrupamento de

dados. Para isso utiliza de três questões: medição da similaridade, formação dos agrupamentos e quantidade de agrupamentos. A análise de agrupamento é uma técnica estatística exploratória e deve ser interpretada como uma ferramenta para se atingir uma meta já definida (HAIR JR et al., 2009). A seguir uma síntese da doutrina sobre o método:

O agrupamento de dados, também chamado de *clustering*, é a tarefa de particionar um conjunto de dados em subconjuntos de itens semelhantes. Itens também se chamam instâncias, observações, entidades, objetos de dados ou itens de dados. (...) A força de uma relação entre dois ou mais itens pode ser quantificada como uma medida de similaridade: uma função matemática que calcula as semelhanças entre dois itens de dados. Os resultados desse cálculo, chamado de valores de similaridade, basicamente comparam um item específico a todos os outros no conjunto de dados. Esses outros itens serão mais ou menos similares em comparação ao item específico. (...) Cada grupo tem um item que melhor o representa, e esse item é chamado de representante do agrupamento. (...) O agrupamento de dados revela a estrutura nos dados extraindo agrupamentos naturais de um conjunto de dados. Portanto, descobrir agrupamentos é uma etapa essencial para a formulação de ideias e hipóteses sobre a estrutura de seus dados e produzir insights para melhor entendê-los (BARI, CHAOUCHI, e JUNG, 2019, p. 118-120).

Uma das características dos métodos de *cluster* é a inexistência de uma técnica que determine o melhor número de grupos a ser utilizado. Assim, o número de grupos a serem formados depende muitas vezes do conhecimento empírico do pesquisador e dos objetivos a serem alcançados sobre a realidade em estudo (Instituto Paranaense de Desenvolvimento Econômico e Social, 1995, p. 14).

Santos e Gomes (2006, p. 263) explicam que a técnica de análise de agrupamento é utilizada para identificar e classificar unidades (variáveis ou objetos) em grupos distintos, de acordo com determinadas características, a partir de indicadores de semelhança e que a análise de agrupamento utiliza o conceito de distância

entre as unidades de classificação. Os autores observam que existem diversos métodos para mensuração dessa distância, sendo o mais utilizado a distância euclidiana (SANTOS e GOMES, 2006).

Frei (2006, p. 55) explica que “uma das maiores dificuldades da Análise de Agrupamentos é a determinação do número de grupos. Além disso, agrupamentos diferenciados surgem quando usamos diferentes algoritmos”. O autor continua observando que não existe um procedimento padrão para resolver esta situação (FREI, 2006).

Jabine e Claude (2007) citam o uso de análise de agrupamento K-, análise de fatores, análise discriminativa gradual e metodologia em árvore para classificação e regressão em pesquisa sobre direitos humanos. Santos e Gomes (2006), em estudo para a identificação de regiões homogêneas no oferecimento de serviços de saúde pública em Minas Gerais utilizaram técnicas de análise fatorial e de análise de agrupamento. O resultado final concluiu pela identificação de dois grupos homogêneos e pela necessidade de políticas públicas diferenciadas para cada grupo (SANTOS e GOMES, 2006).

No estado de Minas Gerais a alocação de recursos do SUS não deve levar em consideração apenas a população, mas sim outras variáveis que incorporem as diferenças de cada grupo (SANTOS e GOMES, 2006, p. 268).

Em artigo no qual é proposta uma sistemática para o estudo e interpretação da estabilidade dos métodos de análise de agrupamento, através de vários algoritmos de agrupamento em dados de vegetação, Albuquerque e outros (2006) concluíram que “(...) em princípio, qualquer algoritmo de agrupamento estudado está estabilizado e existem, de fato, grupos entre os indivíduos verificados” (ALBUQUERQUE et al., 2006, p. 263).

Camilo, Carvalho e Xavier, no XXXII Encontro ANPAD (2008), apresentaram estudo sobre o comportamento de 18 setores da economia

nacional durante o período 2003 a 2006, utilizando métricas de desempenho econômico, financeiro e de estrutura, para análise de performance. Foram utilizados Análise de Componentes Principais e a Análise de *Cluster*. Três *Clusters* foram definidos: inferior 1, intermediário 2 e superior 3. Como conclusão, foi observado que o setor de Mineração manteve-se posicionado no grupo superior durante o período sob análise, o setor Atacado firmou posição no *Cluster 2*, o setor Autoindústria marcou grande volatilidade, enquanto determinados setores apresentam mutações constantes outros permanecem no mesmo *Cluster* e outros acompanham as movimentações conjuntamente.

No trabalho apresentado no Seminário sobre a Economia Mineira (2014), (Perfil da criminalidade em Minas Gerais: uma aplicação de análise de agrupamento) Castro et al., identificou o perfil de criminalidade nos municípios do Estado de Minas Gerais utilizando Análise de *Cluster* e Análise fatorial. As seguintes variáveis foram analisadas: taxa de crimes violentos contra pessoa, taxa de crimes contra o patrimônio, taxa de crimes violentos e taxa de homicídios. Os 853 municípios mineiros foram classificados em 12 grupos. O Instituto Paranaense de Desenvolvimento Econômico e Social (1995, p. 2), explica que “os município paranaenses, no que concerne às condições sócio-econômicas, possuem características diferenciadas decorrentes de uma série de fatores de natureza física, econômica e social. Essa heterogeneidade contribui de maneira acentuada na ocupação das cidades e organização física do espaço urbano”. “Diante disso, a melhoria das condições de vida da população requer um tratamento diferenciado, dadas as particularidades de cada município e as semelhanças entre alguns deles” (Instituto Paranaense de Desenvolvimento Econômico e Social, p. 2).

O foco da análise de agrupamento é a comparação de objetos com base na variável estatística, não na estimação da variável estatística em si. Por isso a definição da variável estatística é um passo crítico na análise.

Os objetivos da análise de agrupamentos não podem ser separados da seleção de variáveis usadas para caracterizar os objetos a serem agrupados. Seja o objetivo exploratório ou confirmatório, o pesquisador efetivamente restringe os possíveis resultados pelas variáveis selecionadas para uso. Os agrupamentos obtidos refletem a estrutura inerente dos dados e são definidos apenas pelas variáveis. Assim, a seleção das variáveis a serem incluídas na variável estatística de agrupamento deve ser feita em relação a considerações teóricas e conceituais, bem como práticas (HAIR JR et al. 2009, p. 438).

Vale observar que Hair Jr et al. (2009) aconselham a padronização das variáveis que serão utilizadas na análise de agrupamento e para tanto sugerem a conversão de cada variável em escores padrões (escore *Z*). A finalidade do procedimento é eliminar o viés introduzido pelas diferenças nas escalas empregadas. Os métodos para a combinação dos elementos estão classificados em hierárquicos e não-hierárquicos, sendo que aqueles são separados em divisivos ou aglomerativos.

Obtida uma classificação aceitável o passo seguinte segue a orientação de Pereira (2004, p. 83) que ensina que o esforço de síntese de “informações para produção do conhecimento envolve processamento de redução de dimensionalidade”. Neste trabalho, para atingir a redução da dimensionalidade, buscou-se auxílio nas técnicas exploratórias de análise fatorial, as quais são úteis quando as variáveis trabalhadas apresentam, entre si, coeficiente de correlação relativamente elevado e se deseja estabelecer novas variáveis que capturem o comportamento do conjunto das variáveis originais (FÁVERO e BELFIORE, 2017).

Fávero e Belfiore (2017) explicam que a análise fatorial é uma técnica multivariada exploratória que procura identificar uma quantidade relativamente pequena de fatores que representam o comportamento conjunto de variáveis originais interdependentes e que pode ser usada para a redução da dimensão dos dados. Os autores citam que dentre os

métodos para determinação de fatores o mais utilizado é o conhecido como componentes principais.

O que a análise de componentes principais faz é identificar primeiro a função que melhor discrimina os objetos e, em seguida derivar a função ortogonal (perpendicular) dessa primeira função. No novo sistema de coordenadas, a distinção entre objetos se fará por suas medidas originais nos novos eixos e as unidades de medida serão abstratas, nem X nem Y das medidas originais (PEREIRA, 2004, p. 118).

A análise de Componentes Principais (ACP) envolve um procedimento matemático que transforma um número de variáveis possivelmente correlacionadas em um número menor de variáveis não correlacionadas denominadas componentes principais, que encontra, entre outras, aplicações em estudos de desenvolvimento regional e diagnósticos socioeconômicos (LOESCH e HOELTGEBAUM, 2012).

Hair Jr *et al.* (2009, p. 94) explicam que o “propósito geral de técnicas de análise fatorial é encontrar um modo de condensar (resumir) a informação contida em diversas variáveis originais em um conjunto menor de novas dimensões compostas ou variáveis estatísticas (fatores) com uma perda mínima de informação”. Os autores reconhecem a existência de dois os métodos básicos de análise fatorial disponíveis para o pesquisador, a análise de fatores e a análise de componentes (HAIR JR *et al.*, 2009, p. 99).

Fávero e Belfiore (2017) explicam que a análise de componentes principais apresenta quatro objetivos principais: 1) redução estrutural; 2) verificação da validade de constructos; 3) elaboração de *rankings* e 4) extração de fatores ortogonais para posterior uso em técnicas multivariadas confirmatórias. Hair *et al.* (2009, p. 94) cita que “as técnicas de análise fatorial podem satisfazer um entre dois objetivos: (1) identificação da estrutura por meio do resumo de dados ou (2) redução de dados”. A análise de componentes principais busca subsidiar diagnósticos, ela não se aplica

quando a intenção é a produção de previsões.

Grings, Baço e Mello (2014) utilizaram o método de componentes principais em análise socioeconômica da criminalidade no estado do Paraná com resultados satisfatórios e observaram que, no Brasil, em razão da escassez de dados, a maioria dos estudos sobre o tema é baseada nas taxas de homicídios disponibilizadas no Sistema de Informações sobre Mortalidade, existindo assim lacunas que demandam pesquisas específicas. Os autores observaram que quanto maior o nível de renda da população maior o número de crimes contra o patrimônio; que a melhoria na educação não contribuiu para uma redução da criminalidade; quanto maiores o índice GINI e os indicadores de desigualdade maior a incidência de crimes contra o patrimônio.

Fernandez e Lobo (2005) utilizaram análise fatorial em estudo sobre a criminalidade na Região Metropolitana de Salvador e chegaram a conclusão que o nível de educação, a eficiência da polícia, a concentração de renda, a densidade demográfica, o grau de urbanização, assim como as rendas do município e do governo municipal são importantes elementos que explicam os índices de atividade criminosa naquela região. Os autores observaram que expansões dos níveis de educação, das rendas do município e do governo municipal, do grau de urbanização e da eficiência da polícia, assim como reduções da concentração de renda, contribuem significativamente para reduzir a atividade criminosa.

Grings, Baço e Mello (2014), citando Araújo Júnior e Fajnzylber, comentam a criminalidade nas microrregiões mineiras e observam que maiores níveis educacionais implicam menores taxas de crime contra pessoa e maiores taxas de crime contra a propriedade, e que a desigualdade de renda encontra-se associada a maiores taxas de homicídios e menores taxas de roubos de veículos, sendo que a educação produz um resultado, ambíguo podendo reduzir ou aumentar a criminalidade, conforme o contexto social. Grings, Baço e

Mello (2014) citam que, na Grande São Paulo, o aumento dos índices de desemprego e de concentração de renda, a redução do rendimento médio do trabalhador e a queda das eficiências da polícia e da justiça explicam o crescimento da atividade criminosa naquela região.

O Governo da Bahia (BAHIA), com o objetivo de identificar prioridades temáticas definidas a partir da realidade local, passou a reconhecer a existência de 27 Territórios de Identidade, constituídos a partir da especificidade de cada região.

Território é conceituado como um espaço físico, geograficamente definido, geralmente contínuo, caracterizado por critérios multidimensionais, tais como o ambiente, a economia, a sociedade, a cultura, a política e as instituições, e uma população com grupos sociais relativamente distintos, que se relacionam interna e externamente por meio de processos específicos, onde se pode distinguir um ou mais elementos que indicam identidade, coesão social, cultural e territorial (BAHIA).

Os Territórios de Identidade foram compostos com metodologia que teve como base o sentimento de pertencimento, “onde as comunidades, através de suas representações, foram convidadas a opinar” (BAHIA).

O Território de Identidade do Recôncavo é composto pelos seguintes municípios: Cabaceiras do Paraguaçu, Cachoeira, Castro Alves, Conceição do Almeida, Cruz das Almas, Dom Macedo Costa, Governador Mangabeira, Maragogipe, Muniz Ferreira, Muritiba, Nazaré, Salinas da Margarida, Santo Amaro, Santo Antônio de Jesus, São Felipe, São Félix, Sapeaçu, Saubara, Varzedo (BAHIA).

Valade (VALADE; BOUDAN, 1995) cita ter tido contato com mais de 163 definições para cultura e muitas acepções diferentes para a expressão. Ele a explica a expressão cultura como um conceito ambíguo, para por fim esclarecer que ela abarca o conjunto de elementos de uma vida social organizada. Laraia (2009) cita três famílias de teorias que buscam entender o sentido da expressão: as teorias idealistas de

culturas; as teorias que enxergam a cultura como sistemas simbólicos e as teorias para as quais a cultura é um sistema adaptativo. Para este trabalho é irrelevante a teoria sobre cultura, fundamental é que a mesma interpretação sobre a expressão seja aplicada de maneira uniforme a todos os municípios.

Silva (2010) explica que a dinâmica da vida, da convivência entre os homens, deve ter como base o respeito pelo outro e como orientação a solidariedade para com o outro, a satisfação das necessidades da vida em geral e a cooperação com o outro e a preservação daquilo que é comum à humanidade. E segundo o autor, não há como se negar a existência de uma cultura de violência (SILVA, p. 16). A expressão cultura da violência também é utilizada por Costa (2005).

Os valores dessas relações sociais devem ser vivenciados no campo da cultura, da educação, da política, da economia, pois constituem a base dos direitos humanos (SILVA, 2010, p. 8).

O indivíduo não existe por si mesmo, mas pela convivência social, pelas diferentes dimensões do desdobramento da vida coletiva. A violência tem suas raízes nessas relações de convivência e não na natureza do indivíduo. A violência é uma relação humana; ela não existiria se a vida humana fosse algo individual (*idem*, p. 7).

Vale observar que os dados sobre educação se referem ao que os educadores denominam de educação formal, ou escolar, em contraste com a educação não formal, decorrente das vivências práticas (GOHN, 2013). Sobre o tema é interessante observar a referência que Gohn (2013) faz a Montesquieu quando este reconhece três áreas de educação: a proveniente dos pais (não formal), a escolar (formal) e a advinda da experiência (não formal).

Viapiana (2006) fala das dificuldades em se buscar correlações entre indicadores sociais e crimes. Não são raros os resultados conflitantes em estudos diversos. O autor ressalva para as situações nas quais as análises das variáveis socioeconômicas são mais abrangentes,

envolvendo indicadores de pobreza, taxas de renda, desigualdade e desorganização social. Nessas condições surgem correlações positivas com alguns tipos de crimes. Ele cita como principais conclusões: a) Correlações positivas: municípios de renda mais alta tendem a apresentar menores taxas de homicídios; maior taxa de pobreza maior criminalidade; maior desorganização familiar maior taxa de homicídio e o efeito “contágio”; b) correlações negativas: maior desigualdade de renda menor taxa de homicídio. Viapiana (2006, p. 73) cita que o estudo do homicídio e observar “o papel de variáveis que descrevem mais diretamente o estilo de vida e as características individuais das vítimas de homicídios”, os resultados serão: a) A renda média, a escolaridade e a pobreza afetam as probabilidades de vitimização; b) as condições econômicas e sociais do indivíduo influem na probabilidade de vitimização; c) o álcool aumenta a probabilidade de vitimização e d) os não-brancos sofrem maiores riscos de sofrer com crimes não-economicamente motivados. Os crimes apresentam maiores incidências nos territórios onde “imperam as piores condições socioeconômicas” (VIAPIANA, 2006, p.74).

3. DESENVOLVIMENTO DA PESQUISA

Este trabalho, de caráter exploratório, teve como objetivo estudar as inter-relações entre as variáveis, taxa de matrícula no ensino médio, taxa de matrícula no ensino fundamental, densidade demográfica, rendimento nominal mensal, população ocupada, Índice GINI, PIB *per capita* e taxa de homicídio, representando as dimensões Educação, Economia, Trabalho e Renda, Desigualdade Social e Violência, com o emprego da análise multivariada. Para tanto fez uso da análise de agrupamento e da análise de componentes principais (ACP).

O objeto da pesquisa, selecionado por conveniência, foi o Território de

Identidade do Recôncavo Baiano. Ele é composto pelos seguintes municípios: Cabaceiras do Paraguaçu, Cachoeira, Castro Alves, Conceição do Almeida, Cruz das Almas, Dom Macedo Costa, Governador Mangabeira, Maragogipe, Muniz Ferreira, Muritiba, Nazaré, Salinas da Margarida, Santo Amaro, Santo Antônio de Jesus, São Felipe, São Félix, Sapeaçu, Saubara, Varzedo (BAHIA). A seleção da amostra com base nos Territórios de Identidade buscou utilizar os estudos prévios desenvolvidos pela Secretária do Planejamento (SEPLAN) que identificaram identidade, coesão social, cultural e territorial entre os municípios (BAHIA). Ou seja, espera-se que em termos culturais os itens sejam semelhantes.

Os dados utilizados foram coletados das seguintes fontes (além da SEPLAN), o Instituto Brasileiro de Geografia e Estatística – IBGE e o Mapa da Violência (WAISELFISZ, 2016). Na Tabela 1 – Variáveis estão os dados selecionados, as respectivas fontes e as dimensões associadas a eles. Os números entre os parênteses representam o ano em que o dado foi coletado de acordo com a fonte. Na Tabela 2, temos as siglas associadas as variáveis para fim de referência.

Tabela 1 – Variáveis

VARIÁVEL	FONTE	DIMENSÃO
Taxa de matrícula no ensino médio (2018)	IBGE	EDUCAÇÃO
Taxa de matrícula no ensino fundamental (2018)	IBGE	EDUCAÇÃO
Densidade demográfica (2010)	IBGE	ECONOMIA
Rendimento nominal mensal de até 1/2 SM (2010)	IBGE	TRABALHO & RENDA
População ocupada (2017)	IBGE	TRABALHO & RENDA
Índice GINI (2003)	IBGE	DESIGUALDADE SOCIAL
PIB per capita R\$ (2016)	IBGE	ECONOMIA
Taxa de homicídio (2010)	MAPA DA VIOLÊNCIA	VIOLÊNCIA
Fonte: O Autor		

Tabela 2 – Variáveis e siglas

VARIÁVEL	SIGLA	DIMENSÃO
Taxa de matrícula no ensino médio (2018)	X1	EDUCAÇÃO
Taxa de matrícula no ensino fundamental (2018)	X2	EDUCAÇÃO
Densidade demográfica (2010)	X3	ECONOMIA
Rendimento nominal mensal de até 1/2 SM (2010)	X4	TRABALHO & RENDA
População ocupada (2017)	X5	TRABALHO & RENDA
Índice GINI (2003)	X6	DESIGUALDADE SOCIAL
PIB per capita R\$ (2016)	X7	ECONOMIA
Taxa de homicídio (2010)	X8	VIOLÊNCIA
Fonte: O Autor		

Os dados em valores absolutos, Matrículas no ensino fundamental, Matrículas no ensino médio e Homicídios, foram transformados em taxas. Os dois primeiros após serem divididos pela

população no último censo ¹ e o último pelo mesmo processo mais a multiplicação por 100.000 (cem mil). As Tabelas 3, 3B, 4 e 4B apresentam os dados originais e os dados transformados em taxas.

Tabela 3A – Dados originais

Município	População censo (2010)	Taxas de matrículas no ensino médio (2018)	Taxas de matrículas no ensino fundamental (2018)	Densidade demográfica (2010)
CABACEIRAS DO PARAGUAÇU	17.327	696	2.900	76,66
CACHOEIRA	32.026	1.358	5.126	81,03
CASTRO ALVES	25.408	941	3.780	35,7
CONCEIÇÃO DO ALMEIDA	17.899	421	1.822	61,7
CRUZ DAS ALMAS	58.606	2.590	8.563	402,12
CONCEIÇÃO DO JACUIPE	30.123	1.429	5.312	256,3
DOM MACEDO COSTA	3.874	127	630	45,7
GOVERNADOR MANGABEIRA	19.818	1.212	3.225	186,4
MARAGOGIPE	42.815	1.409	5.660	97,27
MUNIZ FERREIRA	7.317	250	879	66,45
MURITIBA	28.899	1.062	3.692	323,58
NAZARÉ	27.274	1.470	3.876	107,47
SALINAS DA MARGARIDA	13.456	855	2.299	89,81
SANTO AMARO	57.800	2.832	7.844	117,26
SANTO ANTÔNIO DE JESUS	90.985	3.957	14.080	348,14
SÃO FELIPE	20.305	890	2.824	98,57
SÃO FELIX	14.098	373	1.409	142,11
SAPEAÇU	16.585	961	2.615	141,5
SAUBARA	11.201	427	1.737	68,51
VARZEDO	9.109	299	1.258	40,16

Fonte: O Autor.

Tabela 3B - Dados originais

Município	Rendimento nominal mensal de até 1/2 SM (2010)	População ocupada (2017)	Índice GINI (2003)	PIB per capita R\$ (2016)	Taxa de homicídios (2010)
CABACEIRAS DO PARAGUAÇU	0,568	0,054	0,36	R\$ 6.859,34	3
CACHOEIRA	0,486	0,112	0,46	R\$ 14.941,95	4
CASTRO ALVES	0,512	0,156	0,43	R\$ 13.329,84	4
CONCEIÇÃO DO ALMEIDA	0,517	0,102	0,41	R\$ 8.467,14	2
CRUZ DAS ALMAS	0,414	0,203	0,44	R\$ 14.897,85	10
CONCEIÇÃO DO JACUIPE	0,408	0,244	0,42	R\$ 60.168,45	10
DOM MACEDO COSTA	0,504	0,075	0,36	R\$ 7.521,17	0
GOVERNADOR MANGABEIRA	0,516	0,131	0,37	R\$ 12.741,64	2
MARAGOGIPE	0,548	0,047	0,45	R\$ 8.727,02	5
MUNIZ FERREIRA	0,509	0,064	0,41	R\$ 6.348,61	1
MURITIBA	0,464	0,084	0,4	R\$ 9.419,42	2
NAZARÉ	0,481	0,099	0,43	R\$ 8.377,23	13
SALINAS DA MARGARIDA	0,531	0,072	0,37	R\$ 10.759,51	2
SANTO AMARO	0,488	0,118	0,44	R\$ 12.211,38	22
SANTO ANTÔNIO DE JESUS	0,389	0,23	0,49	R\$ 20.442,10	45
SÃO FELIPE	0,548	0,089	0,4	R\$ 8.318,56	0
SÃO FELIX	0,515	0,09	0,43	R\$ 11.203,09	3
SAPEAÇU	0,506	0,087	0,4	R\$ 10.095,24	4
SAUBARA	0,513	0,081	0,38	R\$ 9.870,74	1
VARZEDO	0,51	0,059	0,35	R\$ 10.022,68	0

Fonte: O Autor.

¹ Censo do ano de 2010.

Tabela 4A – Dados transformados em taxas

Município	População censo (2010)	Taxas de matrículas no ensino médio (2018)	Taxas de matrículas no ensino fundamental (2018)	Densidade demográfica (2010)
CABACEIRAS DO PARAGUAÇU	17.327	4,0%	16,7%	76,66
CACHOEIRA	32.026	4,2%	16,0%	81,03
CASTRO ALVES	25.408	3,7%	14,9%	35,70
CONCEIÇÃO DO ALMEIDA	17.889	2,4%	10,2%	61,70
CRUZ DAS ALMAS	58.606	4,4%	14,6%	402,12
CONCEIÇÃO DO JACUÍPE	30.123	4,7%	17,6%	256,30
DOM MACEDO COSTA	3.874	3,3%	16,3%	45,70
GOVERNADOR MANGABEIRA	19.818	6,1%	16,3%	186,40
MARAGOGIPE	42.815	3,3%	13,2%	97,27
MUNIZ FERREIRA	7.317	3,4%	12,0%	66,45
MURITIBA	28.899	3,7%	12,8%	323,58
NAZARÉ	27.274	5,4%	14,2%	107,47
SALINAS DA MARGARIDA	13.456	6,4%	17,1%	89,81
SANTO AMARO	57.800	4,9%	13,6%	117,26
SANTO ANTÔNIO DE JESUS	90.985	4,3%	15,5%	348,14
SÃO FELIPE	20.305	4,4%	13,9%	98,57
SÃO FÉLIX	14.098	2,6%	10,0%	142,11
SAPEAÇU	16.585	5,8%	15,8%	141,50
SAUBARA	11.201	3,8%	15,5%	68,51
VARZEDO	9.109	3,3%	13,8%	40,16
MÉDIA	27.245,750	0,042	0,145	139,322
DESVIO PADRÃO	21.219,503	0,011	0,021	108,318

FONTE: O AUTOR

Tabela 4B – Dados transformados em taxas

Município	Rendimento nominal mensal de até 1/2 SM (2010)	População ocupada (2017)	Índice GINI (2003)	PIB per capita R\$ (2010)	Taxa de homicídios (2010)
CABACEIRAS DO PARAGUAÇU	56,8%	5,4%	0,36	R\$ 6.859,34	17,3
CACHOEIRA	48,6%	11,2%	0,46	R\$ 14.941,95	12,5
CASTRO ALVES	51,2%	15,6%	0,43	R\$ 13.329,84	15,7
CONCEIÇÃO DO ALMEIDA	51,7%	10,2%	0,41	R\$ 8.467,14	11,2
CRUZ DAS ALMAS	41,4%	20,3%	0,44	R\$ 14.897,85	17,1
CONCEIÇÃO DO JACUÍPE	40,8%	24,4%	0,42	R\$ 60.168,45	33,2
DOM MACEDO COSTA	50,4%	7,5%	0,36	R\$ 7.521,17	0,0
GOVERNADOR MANGABEIRA	51,6%	13,1%	0,37	R\$ 12.741,64	10,1
MARAGOGIPE	54,8%	4,7%	0,45	R\$ 8.727,02	11,7
MUNIZ FERREIRA	50,9%	6,4%	0,41	R\$ 6.348,61	13,7
MURITIBA	46,4%	8,4%	0,40	R\$ 9.419,42	6,9
NAZARÉ	48,1%	9,9%	0,43	R\$ 8.377,23	47,7
SALINAS DA MARGARIDA	53,1%	7,2%	0,37	R\$ 10.759,51	14,9
SANTO AMARO	48,8%	11,8%	0,44	R\$ 12.211,38	38,1
SANTO ANTÔNIO DE JESUS	38,9%	23,0%	0,49	R\$ 20.442,10	49,5
SÃO FELIPE	54,8%	8,9%	0,40	R\$ 8.318,56	0,0
SÃO FÉLIX	51,5%	9,0%	0,43	R\$ 11.203,09	21,3
SAPEAÇU	50,6%	8,7%	0,40	R\$ 10.095,24	24,1
SAUBARA	51,3%	8,1%	0,38	R\$ 9.870,74	8,9
VARZEDO	51,0%	5,9%	0,35	R\$ 10.022,68	0,0
MÉDIA	0,496	0,110	0,410	13.236,15	17,686
DESVIO PADRÃO	0,047	0,057	0,037	11.538,32	14,402

FONTE: O AUTOR

Apesar dos softwares estatísticos estarem preparados para trabalhar com dados brutos, para ter certeza da eliminação de qualquer viés induzido pelas características estruturais das variáveis, optou-se pela padronização dos dados com a

conversão de cada variável em escore padrão (escore Z), com média zero e desvio padrão igual a um. Na Tabela 5 – Dados Padronizados (ESCORE Z) - estão os escores padrões de cada uma das observações.

Tabela 5 – Dados padronizados (escore z)

Município	Taxa de matrículas no ensino médio (2018)	Taxa de matrículas no ensino fundamental (2018)	Densidade demográfica (2010)	Rendimento nominal mensal de até 1/2 SM (2010)	População ocupada (2017)	Índice GINI (2003)	PIB per capita R\$ (2016)	Taxa de homicídios (2010)
CABACEIRAS DO PARAGUAÇU	-0,17	1,06	-0,58	1,53	-0,98	-1,34	-0,55	-0,03
CACHOEIRA	0,03	0,71	-0,54	-0,22	0,04	1,34	0,15	-0,36
CASTRO ALVES	-0,46	0,18	-0,96	0,34	0,81	0,53	0,01	-0,13
CONCEIÇÃO DO ALMEIDA	-1,69	-2,04	-0,72	0,44	-0,14	0,00	-0,41	-0,45
CRUZ DAS ALMAS	0,19	0,05	2,43	-1,76	1,64	0,80	0,14	-0,04
CONCEIÇÃO DO JACUIPE	0,49	1,48	1,08	-1,89	2,36	0,27	4,07	1,08
DOM MACEDO COSTA	-0,85	0,83	-0,86	0,16	-0,61	-1,34	-0,50	-1,23
GOVERNADOR MANGABEIRA	1,74	0,84	0,43	0,42	0,37	-1,07	-0,04	-0,53
MARAGOGIPE	-0,84	-0,60	-0,39	1,11	-1,11	1,07	-0,39	-0,42
MUNIZ FERREIRA	-0,72	-1,17	-0,67	0,27	-0,81	0,00	-0,60	-0,28
MURITIBA	-0,49	-0,81	1,70	-0,69	-0,46	-0,27	-0,33	-0,75
NAZARÉ	1,08	-0,13	-0,29	-0,33	-0,19	0,53	-0,42	2,08
SALINAS DA MARGARIDA	1,96	1,22	-0,46	0,74	-0,67	-1,07	-0,21	-0,20
SANTO AMARO	0,63	-0,44	-0,20	-0,18	0,14	0,80	-0,09	1,41
SANTO ANTÔNIO DE JESUS	0,13	0,46	1,93	-2,30	2,12	2,14	0,62	2,21
SÃO FELIPE	0,16	-0,28	-0,38	1,11	-0,37	-0,27	-0,43	-1,23
SÃO FÉLIX	-1,43	-2,13	0,03	0,40	-0,35	0,53	-0,18	0,25
SAPEAÇU	1,45	0,60	0,02	0,21	-0,40	-0,27	-0,27	0,45
SAUBARA	-0,36	0,48	-0,65	0,36	-0,51	-0,80	-0,29	-0,61
VARZEDO	-0,84	-0,32	-0,92	0,29	-0,90	-1,60	-0,28	-1,23

FONTE: O AUTOR

Com os dados transformados em taxas e padronizados o passo seguinte foi a análise de agrupamento. Para isso foi utilizado o *software* MINITAB 19. Por se tratarem de variáveis qualitativas decidiu-se por trabalhar com o método usando distância euclidiana e ligação completa. Primeiro as observações foram agrupadas sem incluir a variável Taxa de homicídios

(2010) com a finalidade de construir um parâmetro para comparação. Em seguida foram agrupadas incluindo a variável Taxa de homicídios (2010).

Na Tabela 6 estão os passos do agrupamento. A opção foi por agrupar as observações em 5 (cinco) grupos. Ensaio preliminares indicaram que uma quantidade maior de grupos pulveriza a amostra.

Tabela 6 – Amalgamação

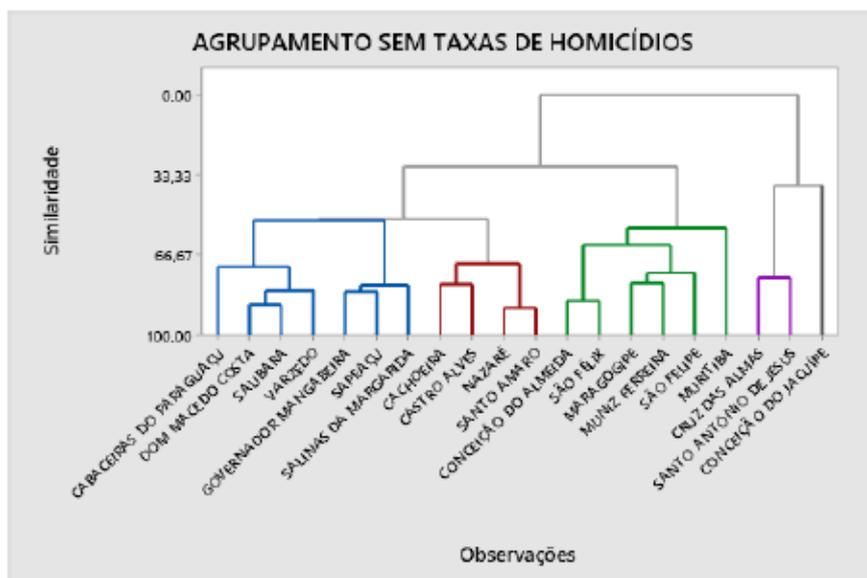
Passos de Amalgamação

	Número de Passo	Número de agrupados	Nível de similaridade	Nível de distância	Agrupados reunidos	Novo agrupado	Número de obs. no novo agrupado
1	19	89,1385	0,78484	12	14	12	2
2	18	87,7453	0,88552	7	19	7	2
3	17	86,0312	1,00938	4	17	4	2
4	16	82,1845	1,28734	8	18	8	2
5	15	81,7375	1,31964	7	20	7	3
6	14	79,4590	1,48428	8	13	8	3
7	13	79,1381	1,50747	2	3	2	2
8	12	78,5992	1,54641	9	10	9	2
9	11	76,2368	1,71711	5	15	5	2
10	10	74,3559	1,85303	9	16	9	3
11	9	71,7441	2,04176	1	7	1	4
12	8	70,5160	2,13049	2	12	2	4
13	7	62,6249	2,70070	4	9	4	5
14	6	55,6669	3,20348	4	11	4	6
15	5	52,4109	3,43876	1	8	1	7
16	4	51,7359	3,48754	1	2	1	11
17	3	37,8703	4,48946	5	6	5	3
18	2	29,8690	5,06762	1	4	1	17
19	1	0,0000	7,22594	1	5	1	20

Fonte: O Autor.

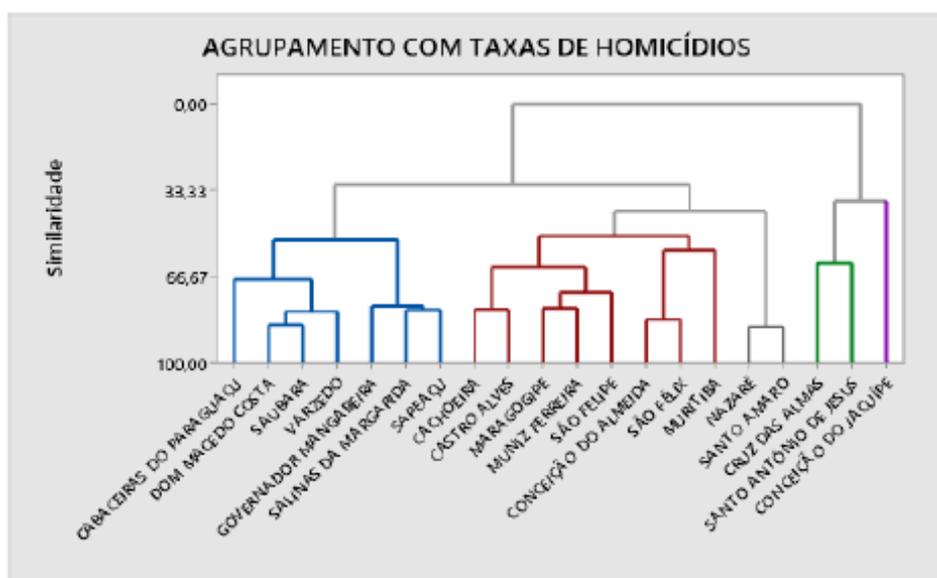
Na construção do agrupamento incluindo a variável Taxa de homicídios (2010) foram mantidos os 5 (cinco) grupos. Os Dendogramas 1 e 2 exibem os agrupamentos dos municípios sem a

inclusão da taxa de homicídio e com a inclusão desta variável, respectivamente. A Tabela 7 apresenta os municípios classificados em função da taxa de homicídio.



Fonte: O Autor.

Dendograma 1



Fonte: O Autor.

Dendograma 2

Tabela 7

TAXAS DE HOMICÍDIOS - CLASSIFICAÇÃO		
CLASSIFICAÇÃO	MUNICÍPIO	TAXA DE HOMICÍDIO
1	SANTO ANTÔNIO DE JESUS	49,459
2	NAZARÉ	47,664
3	SANTO AMARO	38,062
4	CONCEIÇÃO DO JACUIPE	33,197
5	SAPEAÇU	24,118
6	SÃO FÉLIX	21,280
7	CABACEIRAS DO PARAGUAÇU	17,314
8	CRUZ DAS ALMAS	17,063
9	CASTRO ALVES	15,743
10	SALINAS DA MARGARIDA	14,863
11	MUNIZ FERREIRA	13,667
12	CACHOEIRA	12,490
13	MARAGOGIPE	11,678
14	CONCEIÇÃO DO ALMEIDA	11,180
15	GOVERNADOR MANGABEIRA	10,092
16	SAUBARA	8,928
17	MURITIBA	6,921
18	DOM MACEDO COSTA	0,000
19	SÃO FELIPE	0,000
20	VARZEDO	0,000
FONTE: O AUTOR		

Sem a inclusão da taxa de homicídios ocorreram os seguintes agrupamentos: a) Cabaceiras do Paraguaçu, Dom Macedo Costa, Saubara, Varzedo, Governador Mangabeira, Sapeaçu e Salinas da Margarida formam o Grupo 1, com similaridade dentro do grupo de 52,41; b) Cachoeira, Castro Alves, Nazaré e Santo Amaro, formam o Grupo 2, com similaridade de 79,14; c) Conceição do Almeida, São Félix, Maragogipe, Muniz Ferreira, São Felipe e Muritiba, formam o Grupo 3, a similaridade é de 55,68; d) Cruz das Almas e Santo de Jesus, formam o Grupo 4, no grupo a similaridade é de 76,24 e e) Conceição do Jacuípe forma o Grupo 5. Entre os Grupos 4 e 5 a similaridade é de 37,88. Entre os demais grupos a similaridade é de 29,88. Entre os Grupos 4, 5 e os demais a similaridade é próxima de zero.

Com a inclusão da taxa de homicídios formaram-se os seguintes agrupamentos: a) Cabaceiras do Paraguaçu, Dom Macedo Costa, Saubara, Varzedo, Governador mangabeira, Sapeaçu e Salinas da Margarida, formam o Grupo 1, com similaridade dentro do grupo de 52,49; b) Cachoeira, Castro Alves, Marogogipe, Muniz Ferreira, São Felipe, Conceição do Almeida,

São Félix e Muritiba, formam o Grupo 2, com similaridade de 50,94; c) Nazaré e Santo Amaro, formam o Grupo 5, com similaridade de 86,07; d) Cruz das Almas e Santo Antonio de Jesus, formam o Grupo 3, com similaridade de 61,69 e e) Conceição do Jacuípe forma o Grupo 4. Entre o Grupo 4 e o Grupo 5 a similaridade é de 37,36. Entre os demais grupos a similaridade é de 31,13. A similaridade entre os Grupos 4, 5 e os demais é zero.

As interpretações das manutenções nos mesmos grupos dos pares os municípios de Nazaré e Santo Amaro, Santo Antonio de Jesus e Cruz das Almas, surge que o primeiro par (Nazaré e Santo Amaro) possui valores próximos para todas as variáveis. Já Santo Antonio de Jesus e Cruz das Almas se distinguem dos demais, de modo significativo, pelas densidades demográficas altas. Destaca-se a situação envolvendo Cruz das Almas e Santo Antonio de Jesus. O segundo com uma taxa de homicídios de 49,5 mortes/100.000 habitantes e o primeiro com uma taxa de homicídios de 17,1 mortes/100.000 habitantes, apesar disso os dois municípios possuem mais similaridades entre si do que com todos os demais.

Concluída a análise de agrupamento das observações, o passo seguinte foi executar a análise para identificar no conjunto de variáveis os componentes principais, para tanto mais uma vez se fez uso do *software* MINITAB. A Tabela 8 exhibe os cálculos das correlações Pearson para as variáveis, a indicação é de há correlações entre os dados.

Tabela 8 - Correlações**Correlações**

	Taxas de matrículas no ensino m	Taxas de matrículas no ensino f	Densidade demográfica (2010)	Rendimento nominal mensal de at	População ocupada (2017)
Taxas de matrículas no ensino f	0,654				
Densidade demográfica (2010)	0,213	0,095			
Rendimento nominal mensal de at	-0,127	-0,168	-0,790		
População ocupada (2017)	0,213	0,293	0,683	-0,850	
Índice GINI (2003)	-0,077	-0,221	0,420	-0,543	0,538
PIB per capita R\$ (2016)	0,177	0,398	0,406	-0,616	0,741
Taxa de homicídios (2010)	0,356	0,105	0,372	-0,553	0,538
	Índice GINI (2003)	PIB per capita R\$ (2016)			
Taxas de matrículas no ensino f					
Densidade demográfica (2010)					
Rendimento nominal mensal de at					
População ocupada (2017)					
Índice GINI (2003)					
PIB per capita R\$ (2016)	0,240				
Taxa de homicídios (2010)	0,614	0,370			

Fonte: O Autor.

Nas Tabelas 9 e 10 estão, respectivamente, a Matriz de Correlação e os Autovetores. O gráfico Biplot 1 expressa a

influência de cada variável nos dois primeiros componentes principais.

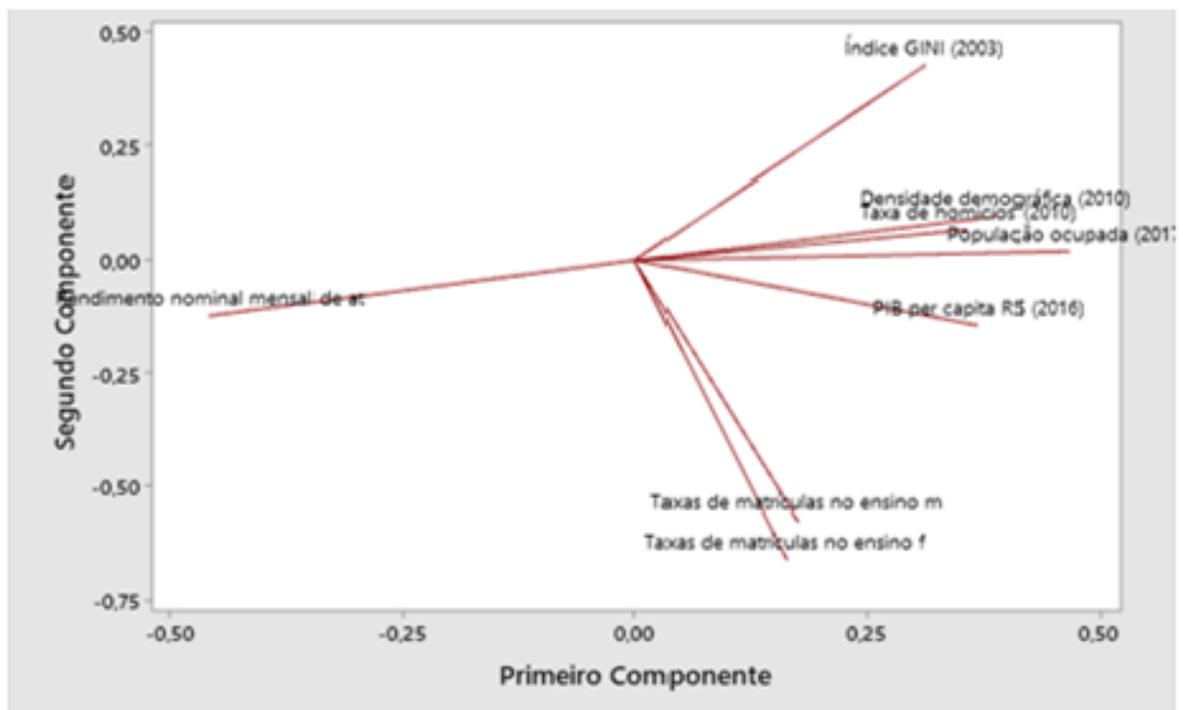
Tabela 9**Autoanálise (Autovalores e Autovetores) da Matriz de Correlação**

Autovalor	3,9483	1,7365	0,9202	0,6344	0,3080	0,2158	0,1388	0,0980
Proporção	0,494	0,217	0,115	0,079	0,038	0,027	0,017	0,012
Acumulado	0,494	0,711	0,826	0,905	0,943	0,970	0,988	1,000

Fonte: O Autor.**Tabela 10****Autovetores**

Variável	CP1	CP2	CP3	CP4	CP5	CP6	CP7
Taxas de matrículas no ensino m	0,175	-0,576	-0,458	0,262	0,101	0,467	-0,236
Taxas de matrículas no ensino f	0,163	-0,659	0,071	-0,115	-0,504	-0,411	0,284
Densidade demográfica (2010)	0,390	0,097	0,164	0,708	0,087	0,123	0,431
Rendimento nominal mensal de at	-0,458	-0,123	-0,184	-0,178	-0,055	0,444	0,174
População ocupada (2017)	0,467	0,020	0,204	-0,088	-0,140	0,054	-0,697
Índice GINI (2003)	0,313	0,428	-0,368	-0,168	-0,642	0,294	0,189
PIB per capita R\$ (2016)	0,370	-0,144	0,427	-0,528	0,293	0,420	0,320
Taxa de homicídios (2010)	0,359	0,068	-0,604	-0,267	0,456	-0,368	0,151
Variável	CP8						
Taxas de matrículas no ensino m	-0,273						
Taxas de matrículas no ensino f	0,132						
Densidade demográfica (2010)	0,320						
Rendimento nominal mensal de at	0,692						
População ocupada (2017)	0,473						
Índice GINI (2003)	-0,144						
PIB per capita R\$ (2016)	-0,130						
Taxa de homicídios (2010)	0,254						

Fonte: O Autor.



Fonte: O autor

Biplot 1

A interpretação do gráfico biplot 1 leva a concluir a existência de razoável associação positiva entre as variáveis Densidade demográfica, Taxa de Homicídios e População Ocupada, pois os vetores estão próximos e segundo para o mesmo quadrante. A variável Rendimento nominal mensal se associa de maneira negativa com as três variáveis citadas. As variáveis Taxa de matrículas no ensino médio e Taxa de matrículas no ensino fundamental também estão associadas. Comparando as variáveis com CP1 e CP2 percebe-se que Densidade demográfica, Taxa de homicídios e População ocupada se associam de maneira positiva com os estes componentes. Rendimento nominal mensal possui uma associação negativa com CP1 e CP2 e as variáveis Taxa de matrículas no ensino médio e Taxa de matrículas no ensino fundamental possuem associações positiva com CP1 e negativa com CP2. A variável Rendimento nominal mensal possui

forte influência com relação a CP1 e as variáveis Taxa de matrícula no ensino médio e Taxa de matrícula no ensino fundamental exercem forte influência em CP2.

Quando as informações das Tabelas 9 e 10 são interpretadas surge que os dois primeiros componentes principais (CP1 e CP2) explicam 71,1% da variação nos dados e os três primeiros componentes principais explicam 82,6% da variação nos dados (apesar de CP3 ter um autovalor menor que 1²). Decidiu-se, então, utilizar esses três componentes principais para compor o modelo descrito pelas equações 1, 2 e 3 a seguir expostas:

² Pelo critério de Kaiser devem ser considerados somente os componentes principais com os autovalores maiores que 1. Outro modelo é o Critério de Pearson (ou regra dos 80%) utilizado quando se trabalha com a matriz de covariância, nele devem considerar-se tantas componentes principais quantas as necessárias para que a percentagem de variância por elas explicada seja superior a 80%. Por fim, com a mesma finalidade pode ser a interpretação do gráfico scree plot.

Equação 1 – TRABALHO E RENDA

$$CP1 = 0,175x_1 + 0,163x_2 + 0,390x_3 - 0,458x_4 + 0,467x_5 + 0,313x_6 + 0,370x_7 + 0,359x_8$$

Equação 2 - EDUCAÇÃO

$$CP2 = -0,576x_1 - 0,659x_2 + 0,097x_3 - 0,123x_4 + 0,020x_5 + 0,428x_6 - 0,144x_7 + 0,068x_8$$

Equação 3 - VIOLÊNCIA

$$CP3 = -0,458x_1 + 0,071x_2 + 0,164x_3 - 0,184x_4 + 0,204x_5 - 0,368x_6 + 0,427x_7 - 0,604x_8$$

Onde os coeficientes foram obtidos na Tabela 10 (autovetores) e as variáveis x_i , conforme a Tabela 2, correspondem a:

x_1	= Taxa de matrículas no ensino médio
x_2	= Taxa de matrículas no ensino fundamental
x_3	= Densidade demográfica
x_4	= Rendimento nominal mensal
x_5	= População ocupada
x_6	= Índice GINI
x_7	= PIB per capita
x_8	= Taxa de homicídios

As variáveis Rendimento nominal mensal e População ocupada se destacam em CP1, a primeira com associação negativa e a segunda com associação positiva. CP1 pode ser denominado de componente **Trabalho e Renda** e representante dessa dimensão. Em CP2, identificamos as variáveis Taxa de matrículas no ensino fundamental e Taxas de matrículas no ensino fundamental, ambas se associam com CP2 de maneira negativa. CP2 pode ser denominado como componente **Educação**. CP3 está maior associado, com sinal negativo, a variável Taxa de homicídios, pode, assim, representar a dimensão **Violência**, menor o valor de CP3 indica uma maior violência.

Além da variável Taxa de homicídios o componente CP3 (**Violência**) é influenciado pela variável Taxa de matrículas no ensino médio, pela variável **PIB per capita** e pela variável **Índice GINI**. As variáveis **Taxa de matrícula no ensino médio** e **Índice GINI** se associam ao CP3 com o mesmo sinal da **Taxa de homicídios**. A quantidade de homicídios cresce junto a quantidade de matrículas e a redução da desigualdade social. A variável PIB per capita se associa com sinal inverso com a Taxa de homicídios. A variável Taxa de homicídios volta a ter uma associação significativa quando é analisado o quinto componente principal (CP5), responsável por 3,8% da variação do conjunto de dados. As variáveis que mais se associam com CP5 são Taxas de matrícula no ensino fundamental (-0,504), Índice GINI (-0,642) e Taxa de homicídios (0,456). As duas primeiras se relacionam com sinal negativo para com o quinto componente principal. Para a terceira variável (Taxa de homicídios) a correlação é positiva. Na Tabela 11, a seguir, temos os escores dos municípios calculados para os três primeiros componentes principais.

Tabela 11

MUNICÍPIO	ESCORES		
	TRABALHO & RENDA CP1	EDUCAÇÃO CP2	VIOLÊNCIA CP3
CABACEIRAS DO PARAGUAÇU	-1,8782	-1,3552	-0,1515
CACHOEIRA	0,3741	0,0154	-0,2133
CASTRO ALVES	-0,0762	0,2474	0,0592
CONCEIÇÃO DO ALMEIDA	-1,4897	2,2183	0,4999
CRUZ DAS ALMAS	2,8518	0,6579	0,7654
CONCEIÇÃO DO JACUIPE	4,6941	-1,2712	1,8768
DOM MACEDO COSTA	-1,7539	-0,7611	1,1727
GOVERNADOR MANGABEIRA	0,0523	-2,0447	0,0259
MARAGOGIPE	-1,3797	1,1689	-0,4609
MUNIZ FERREIRA	-1,4023	1,1408	-0,1638
MURITIBA	0,0761	0,9392	0,8874
NAZARÉ	0,8721	-0,0933	-2,1639
SALINAS DA MARGARIDA	-0,7721	-2,5223	-0,7374
SANTO AMARO	0,8348	0,3817	-1,4787
SANTO ANTÔNIO DE JESUS	4,5837	1,1088	-0,7076
SÃO FELIPE	-1,5245	-0,2253	0,2251
SÃO FÉLIX	-0,7405	2,4393	-0,0612
SAPEAÇU	0,0522	-1,3055	-1,0248
SAUBARA	-1,2183	-0,5664	0,4607
VARZEDO	-2,1558	-0,1725	1,1901

Fonte: O Autor

4. CONCLUSÕES

A aplicação da análise de agrupamento ao conjunto classificou os municípios pelas similaridades em cinco grupos conforme conveniência da investigação. As diferenças entre os grupos formados sem a inclusão da Taxa de homicídios e com a inclusão da Taxa de homicídios não foram impactantes. As mudanças mais significativas foram nos Grupos 2 e 3. Os Grupos 2 e 3 foram reunidos no Grupo 2 e os municípios de Nazaré e Santo Amaro foram extraídos para formar um novo Grupo 3. Os municípios de Cruz das Almas, Santo Antonio de Jesus e Conceição do Jacuípe foram considerados muito distantes dos demais, independente da inclusão da variável Taxa de homicídios na análise. Aqui vale lembrar que todos os municípios pertencem ao Território de Identidade do Recôncavo Baiano, logo

integrantes de um espaço físico, geograficamente definido, contínuo, caracterizado por critérios multidimensionais, como o ambiente, a economia, a sociedade, a cultura, a política e as instituições, e uma população com grupos sociais relativamente distintos e onde se pode distinguir um ou mais elementos que indicam identidade. Ou seja, pelo menos em termos culturais se espera uma identidade entre eles.

Com os dados que foram reunidos concluiu-se que a variação no conjunto de municípios pode ser explicada pelo componente **Trabalho e Renda**, o qual sofre maior influência das variáveis **Rendimento mensal** de até 1/2 SM³ e **População ocupada**. O componente

³ Rendimento mensal de até 1/2 SM = rendimento médio mensal de até metade do salário mínimo.

Educação completa a variação. Nele se destacam as variáveis Taxa de matrículas no ensino médio e Taxa de matrículas no ensino fundamental. Maior quantidade de alunos matriculados está associada a menor desigualdade social e menor quantidade pessoas com rendimento de até meio salário mínimo.

O componente **Violência** é explicado pelas variáveis **Taxa de homicídios, Taxa de matrículas no ensino médio e PIB per capita**. Quanto menor o componente maior a violência. As variáveis Taxas de matrículas no ensino médio e Taxas de homicídios atuam no mesmo sentido. A influência da Taxa de Matrículas no ensino fundamental é desprezível. O resultado parece conflitar o estudo de Fernandez e Lobo (2005) sobre a criminalidade na Região Metropolitana de Salvador onde indicaram que o nível de educação contribui significativamente para reduzir a atividade criminosa. Os dados parecem concordar com Grings, Baço e Mello (2014) quando citam o efeito ambíguo da educação sobre a violência encontrado por Araújo Júnior e Fajnzylber. Na tentativa para explicar os resultados encontrados surgiram as seguintes questões e hipóteses:

a) A correlação positiva entre matrículas na educação formal e as taxas de homicídios nos faz questionar o indicador de educação utilizado, uma vez que parece conflitar com a literatura estudada sobre o tema. Algumas questões surgem: As taxas de matrículas estão realmente expressando educação formal? Ou elas representam apenas estatísticas sem impactos práticos? As escolas estão educando? As famílias estão matriculando as crianças buscando educação formal, ou os interesses são outros?

b) Hoje temos o Bolsa Família, um Programa de transferência de renda em cujos condicionantes para habilitar uma família encontramos: “Ao ingressar no programa, os usuários devem estar atentos às chamadas condicionalidades do Bolsa Família, que são compromissos assumidos pelos beneficiários e pelo poder público para a superação da pobreza. Na área da

educação, Crianças e adolescentes com idades entre 6 e 15 anos devem ter, no mínimo, 85% de presença nas aulas. Para jovens de 16 a 17 anos, a frequência mínima exigida é de 75%” (BRASIL, 2017). Os números de matriculados espelham a busca pela educação formal ou a busca pelo Bolsa Família? A escola esta atingindo sua finalidade como *trampolim* da mobilidade social e da inclusão social?

c) Uma hipótese a ser discutida é o que Caldas (2019) cita como infrequência dos alunos matriculados em artigo publicado no Bahia Noticias. Ou seja, apesar de matriculado o aluno não frequenta a escola. Ele apenas aumenta o indicador de matrículas.

d) Sob outro ponto de vista, Odon (2018, p. 52) cita Fajnzylber, Lederman e Loayza para considerar que “existe um efeito de atraso de políticas educacionais sobre a taxa de crime, isto é, a redução do crime como efeito de investimentos em educação não se materializa enquanto os jovens estão sendo educados, mas, majoritariamente, quando eles se tornam adultos (efeito inercial). Além disso, há o efeito indireto da educação sobre a desigualdade social”.

e) O PIB⁴ per capita, em tese, expressa a riqueza do território, no caso dos municípios. Neste estudo o impacto do PIB sobre a violência é negativo. Um maior PIB per capita pode indicar uma maior circulação de riqueza e maior oportunidade de trabalho e construção de um projeto de vida, reduzindo a tentação do crime.

Quando se considera o conjunto de variáveis estudado, representando as dimensões **Educação, Economia, Trabalho e Renda, Desigualdade Social e Violência**, ocorre que as variações entre os municípios do Território de Identidade do Recôncavo não são influenciadas de maneira expressiva pelas taxas de homicídios. Elas são explicadas pelas dimensões **Trabalho e Renda e Educação**. O componente **Violência** recebe fortes influências das taxas

⁴ Produto interno bruto

de homicídios, dos PIB(s) e das taxas de matrículas no ensino médio. As taxas de matrículas no ensino médio estão correlacionadas de maneira positiva com a **Violência** o que leva a despertar um sinal de alerta sobre o significado da variável como indutor de inclusão social e demonstra que no caso estudado a matrícula no ensino não significa que esteja ocorrendo educação inclusiva.

Um fator que deve ser considerado quando da utilização dos dados e da interpretação dos resultados é a temporalidade dos dados. Pois foram colhidos em fontes diversas e nem sempre em um mesmo momento.

REFERÊNCIAS

ALBUQUERQUE, Márcio Augusto de; *et al.* Estabilidade em análise de agrupamento: estudo de caso em ciência florestal. **Sociedade de Investigações Florestais**, Viçosa, v. 30, n. 2, p.257-265, 2006. Disponível em: <<http://www.scielo.br/pdf/rarv/v30n2/a13v30n2>>. Acesso em: 21 jul. 2019.

BAHIA. SEPLAN. SEPLAN. **Territórios de Identidade**. Disponível em: <<http://www.seplan.ba.gov.br/modules/conteudo/conteudo.php?conteudo=17>>. Acesso em: 3 ago. 2019.

BARI, Anasse; CHAOUCHI, Mohamed; JUNG, Tommy. **Análise preditiva para leigos**. 2. ed. Rio de Janeiro: Alta Books, 2019.

BRASIL. IBGE. Ministério da Economia. **Municípios: Municípios de Bahia**. 2017. Disponível em: <<https://cidades.ibge.gov.br/>>. Acesso em: 3 ago. 2019.

_____. Secretaria Especial do Desenvolvimento Social. Ministério da Cidadania. **Bolsa Família: saiba quais são as regras para participar do programa**. 2017. Disponível em: <<http://mds.gov.br/area-de-imprensa/noticias/2017/setembro/bolsa-familia-saiba-quais-sao-as-regras-para-participar-do-programa>>. Acesso em: 25 ago. 2019.

CALDAS, Matheus. Secretaria de Educação recorre ao MP-BA para combater infrequência escolar em Salvador. **Bahia Notícias**. Salvador, p. 0-0. 17 set. 2019. Disponível em: <<https://www.bahianoticias.com.br/noticia/239638-secretaria-de-educacao-recorre-ao-mp-ba-para-combater-infrequencia-escolar-em-salvador.html>>. Acesso em: 17 set. 2019.

COSTA, Ivone Freire. **Polícia e sociedade: gestão de segurança pública, violência e controle social**. Salvador: Edufba, 2005.

FÁVERO, Luiz Paulo; BELFIORE, Patrícia. **Manual de análise de dados**. Rio de Janeiro: Elsevier, 2017.

FERNANDEZ, José Carrera; LOBO, Luiz Fernando. A criminalidade na região metropolitana de Salvador. **Análise Econômica**, Porto Alegre, v. 23, n. 44, p.3-37, 2005. Disponível em: <<https://seer.ufrgs.br/AnaliseEconomica/article/view/10816/6424>>. Acesso em: 10 jan. 2019

FREI, Fernando. **Introdução à análise de agrupamentos: teoria e prática**. São Paulo: Unesp, 2006.

GOHN, Maria da Glória. **Educação não formal e o educador social: atuação no desenvolvimento de projetos sociais**. São Paulo: Cortez, 2013.

GRINGS, Maicon; BAÇO, Fernanda Bezerra; MELLO, Gilmar Ribeiro de. **Análise socioeconômica da criminalidade no estado do Paraná**. In: III Congresso Nacional de Pesquisa em Ciências Sociais Aplicadas – III CONAPE, 3., 2014, Fernando Beltrão. Anais... . Francisco Beltrão: UNIOESTE, 2014. p. 1 - 21. Disponível em: <http://cac.php.unioeste.br/eventos/conape/anais/iii_conape/Arquivos/Artigos/Artigoscompletos/CIENCIASECONOMICAS/5.pdf>. Acesso em: 3 jan. 2019.

HAIR JUNIOR, Joseph F *et al.* **Análise multivariada de dados**. 6. ed. Porto Alegre: Bookman, 2009.

JABINE, Thomas B.; CLAUDE, Richard P. **Direitos Humanos e Estatística**: O arquivo posto a nu. São Paulo: Editora da Universidade de São Paulo, 2007. 552 p. Gilson César Cardoso de Souza.

LARAIA, Roque de Barros. **Cultura**: um conceito antropológico. Rio de Janeiro: Jorge Zahar, 2009.

LOESCH, Claudio; HOELTGEBAUM, Marianne. **Métodos estatísticos multivariados**. São Paulo, Saraiva, 2012.

ODON, Tiago Ivo. **Segurança pública e análise econômica do crime**: O desenho de uma estratégia para a redução da criminalidade no Brasil. *Ril Brasília, Brasília*, v. 55, n. 218, p.33-61, jun. 2018. Disponível em: <https://www12.senado.leg.br/ril/edicoes/55/218/ril_v55_n218_p33.pdf>. Acesso em: 1 ago. 2019.

OLIVEN, Ruben George. **Violência e Cultura no Brasil**. Rio de Janeiro: Centro Edelstein de Pesquisas Sociais, 2010.

PARANÁ. IPARDES - Instituto Paranaense de Desenvolvimento Econômico e Social. Governo do Estado do Paraná. **Agrupamento dos municípios paranaenses segundo variáveis sócio-econômicas**: uma análise estatística. Curitiba: IPARDES, 1995. 60 p. Disponível em: <http://www.ipardes.pr.gov.br/biblioteca/docs/agrupamento_municip_pr_12_95.pdf>. Acesso em: 21 jul. 2019.

PEREIRA, Júlio Cesar Rodrigues. **Análise de dados qualitativos**: Estratégias metodológicas para as ciências da saúde, humanas e sociais. 3. ed. São Paulo: Edusp, 2004.

SANTOS, Cristiane Márcia dos; GOMES, Adriano Provezano. **Identificação de regiões homogêneas no oferecimento de serviços de saúde pública em Minas Gerais**. *Bahia Análise e Dados, Salvador*, v. 16, p.257-268, set. 2006. Disponível em: <http://www.sei.ba.gov.br/index.php?option=com_content&view=article&id=981&Itemid=284>. Acesso em: 21 jul. 2019.

SILVA, Enio Waldir da. **Sociologia da violência**. Ijuí: Unijaí, 2010.

SILVA, Raphael Rossi; BENIN, Giovani. **Análises Biplot**: conceitos, interpretações e aplicações. *Ciência Rural, Santa Maria, Santa Maria*, v. 42, n. 8, p.1404-1412, ago. 2012. Disponível em: <<http://www.scielo.br/pdf/cr/v42n8/a23912cr5918.pdf>>. Acesso em: 1 jun. 2019.

SILVESTRE, António Luís. **Análise de dados e estatística descritiva**. Lisboa: Escolar Editora, 2007. 352 p.

VALADE Bernard; BOUDON, Raymond. **Tratado de sociologia**. Rio de Janeiro, Jorge Zahar, 1995.

VIAPIANA, Luiz Tadeu. **Economia do crime**: Uma explicação para a formação do criminoso. Porto Alegre: Age, 2006.

WASELFISZ, Julio Jacobo. **Mapa da Violência**. 2016. Disponível em: <https://www.mapadaviolencia.org.br/tabelas2012/taxas_mapas_2000_2010_estadual.xls>. Acesso em: 3 ago. 2019.

VIOLENCE IN THE BAHIA'S CONCAVE: A MULTIVARIATED ANALYSIS

ABSTRACT: This survey was designed as an exploratory study to analysis of the relationship between enrolment rate in secondary education, primary enrollment rate, population density, nominal efficiency, busy population, GINI index, per capita gross domestic product (GDP), homicide rate, represents the dimensions, Education, Economy, Employment and Income, Social Inequality e Violence, using a multivariate statistics. The municipalities in which at Bahia's Concave territory were objects of study for this research. We note that the variation between cities are not influenced by homicide rate and can be explained for the dimensions Employment and Income and Education. The dimension Violence receives strong influences from homicide rate, per capita gross domestic product (GDP) and enrolment rate in secondary education. The positive correlation between dimension Violence and enrolment rate in secondary education shows that analysis case school enrollment does not mean that inclusive educational policy.

Keywords: Violence. Education. Bahia. Multivariate statistics. Homicide.

Recebido em 22 de setembro de 2019

Aprovado em 10 de novembro de 2019