

## PERFIL DA CRIMINALIDADE EM MINAS GERAIS: UMA APLICAÇÃO DE ANÁLISE DE AGRUPAMENTO

Talita Egevardt de Castro<sup>1</sup>  
Iara Maira da Silveira<sup>2</sup>  
Viviane Silva Lírio<sup>3</sup>  
Daniel Arruda Coronel<sup>4</sup>

**Resumo:** Este trabalho tem por objetivo identificar o perfil de criminalidade nos municípios do Estado de Minas Gerais, com base nos seguintes indicadores: taxa de crimes violentos, taxa de crimes violentos contra o patrimônio, taxa de crimes violentos contra a pessoa e taxa de homicídios. Para a análise dos dados foram aplicadas técnicas de estatística multivariada entre elas a Análise Fatorial e Análise de *Cluster*. A base de dados utilizada foi do Índice Mineiro de Responsabilidade Social, referente ao ano de 2010. Os resultados indicaram que os municípios que apresentam os maiores índices de criminalidade estão localizados na Região Metropolitana de Belo Horizonte e no Vale do Rio Doce.

**Palavras-chaves:** criminalidade, análise de *cluster*, Minas Gerais.

**Abstract:** This paper aims to identify the profile of crime in the cities of Minas Gerais, based on the following indicators: violent crime rate, violent crime rate against property, violent crime rate against persons and homicide rate. Were applied multivariate statistical techniques including the Factor Analysis and Cluster Analysis. The database used was the Mining Social Responsibility Index for the year 2010. Results indicated that municipalities have the highest crime rates are located in the Metropolitan Region of Belo Horizonte and Vale do Rio Doce.

**Keywords:** criminality, cluster analysis, Minas Gerais.

**Área temática:** Economia

---

<sup>1</sup> Mestranda em Economia Aplicada pela UFV.

<sup>2</sup> Mestranda em Economia Aplicada pela UFV.

<sup>3</sup> Professora Associada do Programa de Pós-Graduação em Economia Aplicada da UFV.

<sup>4</sup> Professor Adjunto do Programa de Pós-Graduação em Administração e Diretor da Editora da UFSM.

## 1. Introdução

O ônus da criminalidade para a sociedade é alto, pois além de reduzir o bem-estar social devido ao aumento dos gastos públicos e privados com segurança, reduz a qualidade de vida, diminui as atividades turísticas e os investimentos em uma determinada região.

De acordo com o IBGE (2010) o Estado de Minas Gerais apresenta uma população de 19.597.330, com um total de 853 municípios. Segundo Brasil (2010) a distribuição espacial da Criminalidade Violenta no Estado mantém o padrão constatado em anos anteriores, ou seja, as maiores taxas de crimes violentos estão nos maiores municípios, com destaque para a Região Metropolitana de Belo Horizonte<sup>5</sup>, o Triângulo Mineiro, a região Noroeste e municípios como Pirapora, Montes Claros e Governador Valadares.

Quanto à distribuição espacial do crime violento contra a pessoa, essa também manteve o padrão observado anteriormente em que a maior incidência desse tipo de crime está na região Nordeste, com destaque para a Região Metropolitana de Belo Horizonte, Triângulo Mineiro e a região Noroeste (BRASIL, 2010).

As maiores taxas do crime violento contra o patrimônio estão nos maiores municípios mineiros, seguindo o mesmo padrão já mencionado para o crime violento, com destaque para a Região Metropolitana de Belo Horizonte, o Triângulo Mineiro e a região Noroeste.

Diante disso o presente trabalho tem por objetivo identificar o perfil de criminalidade dos municípios do Estado de Minas Gerais, com base em indicadores de criminalidade divulgados pelo Índice Mineiro de Responsabilidade Social.

O presente trabalho está estruturado em três seções, além desta introdução. Na segunda seção, apresentam-se os procedimentos metodológicos, na terceira, os resultados obtidos são analisados e discutidos e, concluindo, são delineadas algumas considerações finais sobre o tema discutido.

## 2. Metodologia

Um modelo de análise fatorial, de acordo com Mingoti (2005), é dado, genericamente, em forma matricial por:

$$X_i = a_{ij}F_j + \varepsilon_i \quad (1)$$

em que:

$X_i = (X_1, X_2, \dots, X_p)^t$  é um vetor transposto de variáveis aleatórias observáveis;

$a_{ij}$  = é uma matriz (p x m) de coeficientes fixos denominados cargas fatoriais, os quais descrevem o relacionamento linear de  $X_i$  e  $F_j$ ;

$F_j = (F_1, F_2, \dots, F_p)^t$  é um vetor transposto (m < p) de variáveis latentes que descrevem os elementos não observáveis da amostra; e

---

<sup>5</sup> Municípios da Região Metropolitana de Belo Horizonte: Baldim, Belo Horizonte, Betim, Brumadinho, Caeté, Capim Branco, Confins, Contagem, Esmeraldas, Florestal, Ibirité, Igarapé, Itaguara, Itatiaiuçu, Jaboticatubas, Nova União, Juatuba, Lagoa Santa, Mário Campos, Mateus Leme, Matozinhos, Nova Lima, Pedro Leopoldo, Raposos, Ribeirão das Neves, Rio Acima, Rio Manso, Sabará, Santa Luzia, São Joaquim de Bicas, São José da Lapa, Sarzedo, Taquaraçu de Minas, Vespasiano.

$\varepsilon_i = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p)^t$  é um vetor transposto dos erros aleatórios, correspondentes aos erros de medição e à variação de  $X_i$  que não é explicada pelos fatores comuns  $F_j$ .

O método utilizado no presente trabalho para a realização da análise fatorial foi o método de componentes principais. O procedimento é que de posse da matriz R, toma-se como fatores comuns os m componentes principais dessa matriz. Com o objetivo de verificar se a análise fatorial utilizada ajusta-se aos dados do modelo, utilizaram-se os testes de Esfericidade de Bartlett e o Critério de Kaiser-Meyer-Olkin (KMO). O primeiro fornece a probabilidade estatística de que a matriz de correlação tenha correlações significantes entre pelo menos algumas das variáveis e compara a matriz de correlação populacional com a matriz identidade. Para que os dados sejam adequados a essa análise, o resultado desse teste deve ser a rejeição da hipótese nula de igualdade das matrizes. O outro teste, o de KMO, verifica a adequação dos dados a partir da criação de um índice que varia de 0 a 1, que compara as correlações simples e parciais entre as variáveis, sendo que valores superiores a 0,5 demonstram que os dados são adequados à análise fatorial (MINGOTI, 2005; HAIR et al., 2009).

Outro método utilizado nesse trabalho consistiu na Análise de Agrupamentos também conhecida como Análise de *Cluster*. Essa técnica consiste em dividir os elementos da amostra em grupos cujos elementos do mesmo grupo sejam semelhantes entre si, dadas às características analisadas, e os elementos de outros grupos heterogêneos em relação às características, (MINGOTI, 2005).

A análise de *cluster* é uma técnica multivariada de classificação que objetiva reduzir a dimensionalidade dos dados. Agrupa um conjunto de dados em subconjuntos, utilizando um critério fixado que pode variar ligeiramente em virtude do método de agrupamento utilizado.

Utilizou-se no presente trabalho os municípios que foram agregados por meio do método não-hierárquico (*k-means*), que dispõe os itens em  $k$  grupos, onde os grupos de itens baseiam-se na sua proximidade em relação a um centróide. Johnson e Wichern (1992) expõem o processo em três etapas: i) Divisão aleatória dos itens em  $k$  grupos iniciais; ii) Disposição de cada item ao grupo onde o centróide é mais próximo. Feito isso, recalcula-se o centróide para os novos grupos formados; iii) Repetição do passo anterior até que não ocorra mudança em duas iterações consecutivas dos grupos.

A determinação do número de grupos depende do conhecimento do problema ou de uma avaliação do conjunto de dados. No primeiro caso, a escolha tem um maior grau de subjetividade. No segundo caso, pode-se escolher  $k$  de modo que se maximize a variabilidade entre os grupos em relação à variabilidade intragrupos. Optou-se pela utilização de 12 agrupamentos, dada a diversidade e a quantidade de municípios existentes no estado de Minas Gerais.

Feito o agrupamento procedeu-se com o cálculo do índice de criminalidade. A classificação dos municípios foi feita com base nos resultados da Análise de *Cluster*. Os índices dos grupos foram calculados a partir dos escores fatoriais, ou seja, dos valores dos fatores para cada um dos 12 grupos. De acordo com Santos et. al. (2011) o procedimento para a construção do índice de criminalidade foi feito por meio do cálculo da média dos fatores ponderada pela proporção de explicação da variância total associada a cada um deles, dado pela expressão (2):

$$IC_i = \frac{\sum_{j=1}^2 w_j \times FP_{ij}}{\sum_{j=1}^2 w_j} \quad (2)$$

Onde  $i = 1, 2, \dots, 12$  e  $j = 1, 2$ .

$IC_i$  - é o Índice de Criminalidade do  $i$ -ésimo município;

$w_j$  - proporção da variância explicada pelo j-ésimo fator; e

$FP_{ij}$  - valor do j-ésimo escore fatorial padronizado associado ao i-ésimo município.

Ainda de acordo com Santos et. al. (2011) “o escore fatorial foi padronizado com a finalidade de se obter valores positivos dos escores originais e permitir a hierarquização dos municípios, uma vez que o IC deve variar entre zero a 100”. A expressão utilizada foi a seguinte:

$$FP_{ij} = \left( \frac{F_{ij} - F_{ij}^{\min}}{F_{ij}^{\max} - F_{ij}^{\min}} \right) \times 100$$

(3)

Em que:

$F_{ij}^{\min}$  - é o menor valor do j-ésimo escore fatorial observado entre todos os municípios;

$F_{ij}^{\max}$  - é o maior valor do j-ésimo escore fatorial observado entre todos os municípios.

Dados os valores do IC, quanto maior os valores do IC maior o índice de criminalidade nos municípios pertencentes ao grupo.

## 2.1 Fonte de dados

Os dados utilizados no presente trabalho referem-se ao Índice Mineiro de Responsabilidade Social (IMRS) para o ano de 2010, dos 853 municípios do Estado de Minas Gerais disponibilizados pela Fundação João Pinheiro.

Foram analisadas as seguintes variáveis: taxa de crimes violentos contra pessoa, taxa de crimes contra o patrimônio, taxa de crimes violentos e taxa de homicídios. A Figura 1 apresenta a descrição das variáveis conforme o IMRS.

Variável	Nome da variável	Descrição
Tx_cv	Taxa de crimes violentos	Medida pela razão entre o número de ocorrências registradas de crimes violentos (homicídio, homicídio tentado, estupro, roubo e roubo a mão armada) e a população do município (por cem mil habitantes).
Tx_cvcpt	Taxa de crimes violentos contra o patrimônio	Razão entre o número de ocorrências registradas de crime contra o patrimônio (roubo e roubo a mão armada) e a população do município (por cem mil habitantes).
Tx_cvcps	Taxa de crimes violentos contra a pessoa	Razão entre o número de ocorrências registradas contra a pessoa (homicídio, homicídio tentado e estupro) e a população do município (por cem mil habitantes)
Tx_h	Taxa de homicídios	Razão entre o número de ocorrências registradas de homicídio e a população do município (por cem mil habitantes).

**Figura 1** – Descrição das variáveis utilizadas para analisar o perfil da criminalidade em Minas Gerais.

Fonte: Fundação João Pinheiro.

### 3. Análise e discussão dos resultados

Inicialmente realizou-se uma análise fatorial das variáveis para verificar a possibilidade de redução dessas variáveis para a Análise de *Cluster*.

O teste de Bartlett indicou significância para o teste ao nível de 1%, demonstrando que as variáveis não são correlacionadas, fato que, conseqüentemente, demonstra a adequabilidade para o uso da técnica de análise fatorial. O outro procedimento realizado, o teste de KMO apresentou valor de 0,7, o qual é maior que valor crítico e adequado ao emprego da análise fatorial, que, conforme classificação de Hair *et al.* (2009), corresponde a 0,5. Portanto, ambos os testes demonstram a possibilidade de realização da análise fatorial. Sendo assim, os testes estatísticos realizados permitem concluir que o conjunto de dados que compõem a amostra utilizada são adequados ao emprego da análise fatorial.

Tabela 1 – Testes de Bartlett e KMO.

Teste	Chi2
Teste de esfericidade de Bartlett	15000
KMO	0,7025

Fonte: Resultados da pesquisa.

A Tabela 2 apresenta os resultados obtidos com a estimação da análise fatorial. Pode-se observar que apenas dois fatores obtiveram raiz característica maior do que a unidade, sendo que o fator 1 explica 65% e o fator 2 explica 27%. Os dois fatores em conjunto explicam 92% da variância total das variáveis utilizadas.

**Tabela 2** – Valores das raízes características, variância total explicada e variância acumulada pelos fatores identificados na análise fatorial.

Fator	Raiz Característica	Variância explicada pelo fator (%)	Variância acumulada (%)
1	2.60166	0,6504	0,6504
2	1.08505	0,2713	0,9217
3	0.31329	0,0783	1,0000
4	0.00000	0,0000	1,0000

Fonte: Resultados da pesquisa.

Tendo em vista a existência de diferentes critérios para estabelecer o número de fatores principais que devem ser extraídos, de acordo com Freitas *et. al.* (2007), optou-se pela inclusão apenas dos componentes que conseguem sintetizar uma variância acumulada em torno de 70%. Sendo assim, utilizam-se os dois primeiros fatores que, conjuntamente, explicam 92,1% da variância total das variáveis analisadas, conforme se pode verificar na última coluna da Tabela 2.

Na Tabela 3 estão apresentadas as cargas fatoriais, ou os coeficientes de correlação entre os fatores e cada um dos indicadores e as respectivas comunalidades, após a rotação ortogonal. Para facilitar a visualização das correlações mais expressivas (acima de 0,7) entre a carga fatorial (F) e o indicador, essas estão destacadas em negrito.

**Tabela 3** – Cargas fatoriais após a rotação ortogonal e comunalidades, obtidas na análise fatorial das variáveis.

Variáveis	Cargas Fatoriais		Comunalidades
	F1	F2	
Tx_cv	<b>0,9194</b>	0,3901	0,99
Tx_cvcpt	<b>0,9908</b>	0,0641	0,98
Tx_cvcps	0,2793	<b>0,8789</b>	0,85
Tx_h	0,0951	<b>0,9187</b>	0,85

Fonte: Resultados da pesquisa.

Verifica-se, portanto, que o Fator 1 está correlacionado com a taxa de crimes violentos e a taxa de crimes violentos contra o patrimônio. O Fator 2 está correlacionado a taxa de crimes violentos contra a pessoa e taxa de homicídios. Assim, o Fator 1 foi denominado de crimes contra o patrimônio e o Fator 2 de crimes contra a pessoa.

### 3.1 Análise de Agrupamento e Índice de Criminalidade

A análise de *Cluster* ou Agrupamento tem por objetivo identificar grupos cujos elementos sejam homogêneos entre si, porém heterogêneos em relação aos demais grupos. O método utilizado foi o método não hierárquico *kmeans*, que dispõe os itens em *k* grupos, onde os grupos de itens baseiam-se na sua proximidade em relação a um centróide.

Com relação a análise de agrupamentos foram selecionados 12 grupos representativos dos 853 municípios de Minas Gerais utilizando o método de *kmeans*, como mencionado na metodologia.

A partir dos grupos formados foram calculados os escores fatoriais padronizados e em seguida o índice de criminalidade, baseados nos escores padronizados. O índice de criminalidade é apresentado na Tabela 4 e observa-se que os grupos que apresentam maiores índices de criminalidade são os grupos 4 e 8 (valores em negrito) com índice de aproximadamente 55% e 53%, respectivamente.

Convém destacar que no grupo 4 o Fator 1 é mais latente, ou seja, há maior evidência de crimes contra o patrimônio, este grupo compreende os municípios da Região Metropolitana de Belo Horizonte. No grupo 8 observa-se que o Fator 2 é o mais latente, ou seja, há maior evidência de crimes contra a pessoa, este grupo compreende os municípios da mesorregião do Vale do Rio Doce e da Região Metropolitana de Belo Horizonte.

Os índices mais baixos são os dos grupos 5, 7 e 12, que compreendem as regiões Sul/Sudeste, Sul/Sudoeste, Norte de Minas e Zona da Mata. Nessas regiões, os crimes relacionados ao Fator 2 são mais latentes, ou seja, há uma maior incidência de crimes contra o patrimônio.

**Tabela 4** – Escores fatoriais e Índice de criminalidade para os doze grupos de municípios homogêneos de Minas Gerais.

Grupos	Escores Fatoriais		Índice de criminalidade (%)
	F <sub>1</sub>	F <sub>2</sub>	
1	2,28911	0,28980	30,45
2	0,71593	-0,64631	10,52
3	-0,58338	0,60549	15,44
<b>4</b>	<b>7,21103</b>	<b>-0,32605</b>	<b>55,36</b>
5	-0,53692	-0,12857	7,99
6	0,04579	-0,15376	11,44
7	-0,07385	-0,83403	3,49
<b>8</b>	<b>-0,14674</b>	<b>3,90102</b>	<b>53,03</b>
9	1,07023	0,44426	24,30
10	0,13870	0,76136	21,70
11	-0,50475	1,65880	27,07
12	-0,62100	-0,77150	0,66

Fonte: Resultados da pesquisa.

Peixoto *et. al.* (2004) em estudo feito para analisar a criminalidade na região metropolitana de Belo Horizonte, encontraram que os crimes contra a pessoa então relacionados positivamente com um indicador que mede o grau de urbanização da região. E, além disso, verificaram que a taxa de homicídio pode estar relacionada à taxa de homicídio da vizinhança, indicando que a criminalidade possui um processo de difusão espacial.

Esses dois resultados encontrados pelos autores corroboram os resultados encontrados no presente estudo, indicando que as regiões onde há maior urbanização há também maior incidência de criminalidade principalmente no tocante aos homicídios. Já nas regiões menos urbanizadas há maior incidência de crimes contra o patrimônio.

As regiões em que o Fator 1 (crimes contra o patrimônio) é mais latente são as seguintes: Central Mineira, Jequitinhonha, Metropolitana de Belo Horizonte, Norte de Minas, Oeste de Minas, Triângulo Mineiro/Alto Paranaíba, Vale do Mucuri, Vale do Rio Doce, Zona da Mata, Campo das Vertentes, Noroeste de Minas, Sul/Sudoeste de Minas e Noroeste de Minas.

As regiões em que o Fator 2 (crimes contra a pessoa) é mais latente são: Zona da Mata, Central Mineira, Jequitinhonha, Metropolitana de Belo Horizonte, Norte de Minas, Oeste de Minas, Sul/Sudoeste de Minas, Triângulo Mineiro/Alto Paranaíba, Vale do Mucuri, Vale do Rio Doce, Campo das Vertentes, Noroeste de Minas, Norte de Minas e Oeste de Minas.

#### 4. Conclusões

Os resultados indicaram que os métodos de análise estatística multivariada servem adequadamente para identificar o perfil de criminalidade dos municípios do Estado de Minas Gerais. A Análise Fatorial mostrou que os 4 indicadores utilizados estão correlacionados com 2 fatores, os quais foram denominados de crimes contra o patrimônio e crimes contra a pessoa.

O cálculo do Índice de Criminalidade mostrou que os municípios que apresentam os maiores índices compreendem a Região Metropolitana de Belo Horizonte e do Vale do Rio Doce, enquanto que os municípios que apresentam os menores índices

estão nas regiões Sul/Sudeste, Sul/Sudoeste, Norte de Minas e Zona da Mata. Assim, verifica-se, que na Região Metropolitana de Belo Horizonte há predominância de crimes violentos e crimes violentos contra o patrimônio.

O presente estudo ficou limitado a um curto período de tempo, não podendo ser levantadas maiores inferências para os próximos anos. Outro ponto a ser destacado e que é pertinente para futuras pesquisas a incorporação de variáveis de cunho social com o objetivo de mensurar a relação com o crime.

## Referências

BRASIL. Fundação João Pinheiro. **Anuário de informações criminais de Minas Gerais**. Belo Horizonte, MG, 2010.

FREITAS, C. A.; PAZ, M. V.; NICOLA, D. S. Analisando a modernização da agropecuária gaúcha: uma aplicação de análise fatorial e *cluster*. **Análise Econômica**, ano 25, n° 47, mar/2007. Porto Alegre, Faculdade de Ciências Econômicas, UFRGS, 2007.

HAIR, F.J.; BLACK, W. C.; BABIN, B.; ANDERSON, R. E.; TATHAN, R. L. **Análise multivariada de dados**. 6ª ed. Porto Alegre: Bookman, 2009.

IBGE - Instituto Brasileiro de Geografia e Estatística. **Censo demográfico**: resultados preliminares, 2010.

JOHNSON, R. A., WICHERN, D. W. **Applied multivariate statistical analysis**. 3 ed. New Jersey: Prentice Hall, 1992. p. 356-383.

MINGOTI, S. A. **Análise de dados através métodos de estatística multivariada**: uma abordagem aplicada. Belo Horizonte: Editora UFMG, 2005.

PEIXOTO, B.; MORO, S.; ANDRADE, M. Criminalidade na Região Metropolitana de Belo Horizonte: uma análise espacial. In: *XI Seminário de Economia Mineira*, 2004. Diamantina, 2004.

SANTOS, M. A. S.; SANTANA, A. C.; RAIOL, L. C. B. Índice de modernização da pecuária leiteira no estado de Rondônia: determinantes e hierarquização. **Perspectiva Econômica**, v.7, n.2, p. 93-106, 2011.