

## A UTILIZAÇÃO DA ANÁLISE DE CORRESPONDÊNCIA SIMPLES EM UM ESTUDO SOBRE OCORRÊNCIAS CRIMINAIS NAS SUB-REGIÕES DO VALE DO PARAÍBA E LITORAL NORTE DO ESTADO DE SÃO PAULO

### Autores

Denise Ferreira Laurito Nascimento<sup>1</sup>

Thales Domingos Machado<sup>2</sup>

### Resumo

A Análise de Correspondência (AC) é uma técnica que trabalha com dados de variáveis categóricas qualitativas e analisa as associações entre elas, cujos procedimentos permitem aos pesquisadores levarem em consideração todos os tipos de variáveis, de modo que é desnecessário recorrer aos usos incorretos de técnicas de ponderação arbitrária. O objetivo deste estudo é apresentar a estatística AC como método alternativo de estudo de variáveis categóricas: ocorrências criminais em cidades do Vale do Paraíba e Litoral Norte do Estado de São Paulo. A ferramenta SPSS (versão 20) foi essencial para determinar o teste Chi-Quadrado, as dimensões principais das variáveis e os gráficos bidimensionais biplot. Os resultados alcançados foram positivos, uma vez que a análise de correspondência simples conseguiu organizar os dados disponibilizados publicamente e apresentar uma análise de significância e determinar as principais categorias que contribuíram para as dimensões principais.

**Palavras-chave:** Análise de Correspondência; Segurança Pública; Vale do Paraíba; Litoral Norte; Variáveis Categóricas.

### *THE USE OF SIMPLE CORRESPONDENCE ANALYSIS IN A STUDY ON CRIMINAL OCCURRENCES IN THE SUB-REGIONS OF THE PARAÍBA VALLEY AND THE NORTHERN COAST OF THE STATE OF SÃO PAULO*

### *Abstract*

*Correspondence Analysis (CA) is a technique that works with data from qualitative categorical variables and analyzes the associations between them. The correspondence analysis procedures allow researchers to take into account all variable types, and it is unnecessary to resort to incorrect uses of arbitrary weighting techniques. The objective of this study is to present the AC statistic as an alternative method of study of categorical variables: criminal occurrences in cities of Vale do Paraíba and the North Coast of the State of São Paulo. The SPSS tool (version 20) was essential to determine the Chi-square test, the main dimensions of the variables and the two-dimensional biplot plots. The results achieved were positive, since the simple correspondence analysis managed to organize the publicly available data and present a significance analysis and determine the main categories that contributed to the Main dimensions.*

**Keywords:** *Correspondence Analysis; Public Security. Vale do Paraíba; North Coast; Categorical Variables.*

<sup>1</sup> doutorado em Engenharia de Materiais – USP – Lorena (SP) e docente na FATEC Prof. Waldomiro May – Email: denise.nascimento@fatec.sp.gov.br

<sup>2</sup> Graduando no Curso Superior de Tecnologia em Gestão da Produção Industrial das Faculdades de Tecnologia do Estado de São Paulo – FATEC Prof. Waldomiro May – Email: thales.d.machado@gmail.com

## INTRODUÇÃO

A análise de correspondência (AC) é um método exploratório que abrange diversas áreas do conhecimento como ciências sociais, engenharia, psicologia, entre outras. Nas ciências sociais e comportamentais, a AC possui fortes raízes, pois, é frequentemente utilizada para determinar e entender as relações e comportamentos com base em duas ou mais variáveis categóricas, que classificam os objetos em escalas qualitativas e podem ser descritas como nominais ou ordinárias (KHANGAR; KAMALJA, 2017).

De forma generalizada, as pesquisas categorizam suas observações, anotando a frequência de observações realizadas para cada variável utilizada. O aumento dos estudos dos dados categóricos ficou evidente a partir de 1960 e seu principal objetivo é estudar as associações entre as variáveis de um determinado experimento. Muitas teorias foram propostas e a AC se destacou, sendo utilizada em muitos campos da ciência. Karl Pearson (1857-1936), matemático britânico, desenvolveu ferramentas que denominou de correlação, com base no trabalho Galton (1822-1911), contribuindo para as áreas de psicologia, antropologia, medicina e sociologia. O estudioso também desenvolveu os testes de regressão e Chi-Quadrado (SCHNEIDER, 2005). Essas ferramentas serviram de base para a construção da técnica AC (FÁVERO; BELFIORE, 2015; KHANGAR; KAMALJA, 2017).

Na década de 1960, Jean-Paul Benzécri, matemático e linguista francês, contribuiu significativamente com estudos de casos utilizando a AC. Jan de Leeuw e Chikio Hayashi, alguns anos mais tarde, também deram contribuição importante para o campo de estudos da AC. Mas foi, em 1984, que M. J. Greenacre publicou seus estudos, intitulados *Theory and Applications of Correspondence Analysis*, e espalhou as técnicas da AC pelo mundo (FÁVERO; BELFIORE, 2015).

Quando há a necessidade de estudar variáveis que possuem o comportamento categórico e identificar a associação entre variáveis e categorias, a técnica exploratória análise de correspondência simples (ANACOR) pode ser utilizada. Ela poderá ser empregada apenas quando houver duas variáveis categóricas a serem estudadas e restringindo o estudo a técnicas bivariadas (FÁVERO; BELFIORE, 2015).

Entendendo as principais aplicabilidades da ANACOR, esta pesquisa utilizou dados públicos fornecidos pela Secretaria da Segurança Pública (SSP). Desde o ano de

1995, a SSP do estado São Paulo é a primeira a começar a disseminar os principais dados estatísticos por Estado, relacionando área, municípios e crime de todas as unidades policiais, entre outros. Os dados colaboram para o estado entender e acompanhar tendências e evoluções das taxas criminais (SSP; 2011).

Assim, o foco desta pesquisa foi coletar os dados das ocorrências criminais no período de 2018 das cidades do Vale do Paraíba e Litoral Norte disponibilizados pela SSP do Estado de São Paulo e utilizar a ferramenta bivariada ANACOR para relacioná-los estatisticamente utilizando o software SPSS (versão 20).

A questão de pesquisa que orienta o estudo é analisar quais são as principais categorias que contribuíram para as dimensões das principais variáveis categóricas das cidades do Vale do Paraíba e Litoral Norte do estado de São Paulo?

Assim, o objetivo deste estudo é apresentar a estatística ANACOR como método alternativo de estudo de variáveis categóricas: ocorrências criminais e cidades do Vale do Paraíba e Litoral Norte do Estado de São Paulo.

Os objetivos específicos são: apresentar o método de análise gráfica para pesquisas de dados qualitativos; mostrar outra possibilidade de estudar a área de segurança pública com as ferramentas estatísticas tratadas nessas pesquisas; contribuir academicamente com estudos relacionados à Análise de Correspondência Simples no campo da Segurança Pública; identificar as associações fortes entre as variáveis categóricas das cidades do Vale do Paraíba e do Litoral Norte (que foram distribuídas em sub-regiões) e a variável categórica ocorrências criminais no período de 2018.

## **2 FUNDAMENTAÇÃO TEÓRICA**

### **2.1 Análise de correspondência simples**

A ANACOR utiliza procedimentos bivariados para estudar duas variáveis categóricas, categorias e intensidade com base em uma tabela de contingência ou tabela de classificação cruzada. A matemática completa sobre o desenvolvimento da análise de correspondência simples pode ser encontrada em Fávero e Belfiore (2015, p. 179-191).

A partir de um conjunto de dados com duas variáveis categóricas I e J, é possível construir uma tabela de contingência, mostrando as frequências absolutas das categorias

de ambas as variáveis. Sendo as células  $ij$  em uma quantidade  $(= 1, \dots, = 1, \dots)$  de observações, o valor total do banco de dados, dado por  $N$ , pode ser calculado.

Um dos principais vértices de estudo da ANACOR é a associação estatística das variáveis categóricas, podendo utilizar o teste 2.

Para ANACOR, o teste 2 é dada pela Eq. (1), assumindo  $(- 1) (- 1)$  graus de liberdade.

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{[n_{ij} - \left( \frac{\sum_i n_{ij} \cdot \sum_j n_{ij}}{N} \right)]^2}{\left( \frac{\sum_i n_{ij} \cdot \sum_j n_{ij}}{N} \right)}$$

De acordo com um grau de liberdade e nível de significância estipulado, se o teste 2 for superior ao seu valor crítico, há uma relação estatística aleatória entre as duas variáveis categóricas, em que as hipóteses  $0 =$  ambas variáveis categóricas relacionadas estatisticamente de forma aleatória e  $1 =$  ambas variáveis categóricas não são relacionadas estatisticamente de forma aleatória.

De acordo com aumento do tamanho das amostras, a análise 2 poderá ser afetada negativamente, assim Beh (2004 apud FÁVERO; BELFIORE, 2015) propôs o uso da Inércia Principal Total na tabela de contingência, melhorando a identificações de pontos importantes e auxiliando nas construções de biplots.

Para determinar um biplot ou mapa simétrico, em que as linhas e colunas das categorias das variáveis possuem escalas semelhantes, ou geralmente chamando de normalização simétrica, caso seja necessário dar prioridades para principal linha e principal coluna, é preciso desconsiderar a normalização simétrica em um gráfico biplot.

Determinando as coordenadas de cada categoria, o biplot poderá ser construído com base em  $m$  dimensões, no caso de um estudo de análise de correspondência múltipla. Como a ANACOR é um estudo bivariado, que mostra apenas as duas principais dimensões, então,  $m = 2$ .

## 2.2 BILOT

A capacidade humana consegue inspecionar com facilidade um objeto plotado em planos bidimensionais e tridimensionais. Quando um objeto possui mais de três



A seguir, apresenta-se a estatística de criminalidade da SSP do estado de São Paulo.

### **2.3 Secretaria da Segurança Pública – estatística de criminalidade**

A SSP, mediante a Lei Estadual nº 9.155/95, adotou transparência na disponibilização para o público dos dados estatísticos de criminalidade no Diário Oficial do Estado, no período trimestral. Frequentemente, as pesquisas realizadas sobre a criminalidade no Brasil são realizadas com base nos dados estatísticos colhidos da base de dados da SSP do Estado de São Paulo porque possuem credibilidade e são dados públicos (KAHN, 2005).

Foi criado também o Sistema Estadual de Coleta de Estatísticas Criminais, concretizado no setor policial, mediante a Res SSP-160-08/05/2011. Esse sistema visa a unificação, racionalização e a elevação da qualidade dos dados estatísticos coletados pelas unidades policiais para serem dispostas a serviço da Administração Superior da Secretaria da Segurança Pública. Mediante a situação, os dados possuem muita precisão, objetividade e transparência (KAHN, 2005).

Para melhorar o sistema de coleta de dados, foi criada a Intranet da Polícia Civil, onde as unidades do Estado de São Paulo preenchem mensalmente todos os dados direto no Sistema. O objetivo do sistema é padronizar as informações disponíveis. No ano de 2001, todos os departamentos da polícia civil do Estado de São Paulo passaram por treinamentos realizados por uma equipe do Departamento de Telemática, atual Departamento de Inteligência da Polícia Civil (DIPOL) e do Departamento de Análise e Planejamento (KAHN, 2005).

Tendo plena consciência do direito à informação pública e liberdade da empresa, a Coordenadoria de Análise e Planejamento (CAP) da SSP criou algumas recomendações ao utilizar e interpretar os dados estatísticos sobre criminalidade (KAHN, 2005).

Em todo país, as estatísticas oficiais de criminalidade são empregadas para delinear as circunstâncias da segurança pública, possuindo validades e confiabilidade, ou seja, são uma amostra de ocorrências criminais. Os dados não retratam o universo de todos os crimes em um determinado local. Por exemplo, a elevação dos dados de criminalidade não é diretamente proporcional à degradação na segurança pública, mas é

o oposto. O aumento dos casos de criminalidade que são registrados é diretamente interligado às variáveis positivas que se referem à credibilidade e performance policial (KAHN, 2005).

Outras interpretações imprecisas são as sazonalidades, ou seja, aumento da criminalidade em períodos de verão (pessoas ficam nas ruas por mais tempo), férias, volta às aulas, finais de semana, feriados, entre outros (KAHN, 2005).

### **3 METODOLOGIA**

Em uma pesquisa científica, existem muitos tipos de metodologias a serem definidas e estudadas, por exemplo, definir-se o estudo como de caráter exploratório, diagnóstico ou teste de hipóteses, entre outros. A pesquisa deste trabalho está sendo embasada nas classificações de pesquisa descritiva, fundamental, qualitativa e quantitativa e, por fim, empírica. (KOTHARI, 2004).

a) Pesquisa descritiva: é a abordagem de um tópico existente em um campo específico de estudo;

b) Pesquisa Fundamental: busca estudar generalizações, fenômenos naturais ou matemáticos;

c) Pesquisas Quantitativa e Qualitativa: a pesquisa quantitativa é fomentada pela medição da quantidade, portanto, sua aplicação pode ser mensurada em termos de números. A pesquisa qualitativa trata dos estudos dos fenômenos de qualidade ou tipo, importante para o estudo das ciências comportamentais;

d) Pesquisas Empíricas: direcionam o trabalho para um caráter observável, ajudando o(s) pesquisador(es) a chegar a resultados finais sobre um determinado tema, com base em análises de experimentos ou observações do estudo. Com base nos tipos de classificações metodológicas abordados, os pesquisadores foram capazes de construir alguns procedimentos importantes durante o desenvolvimento desta pesquisa, com base em métodos e técnicas de Kothari (2004), conforme demonstrado no Quadro 1. Por meio da análise desse quadro, observa-se, de forma generalizada, os métodos e técnicas desta pesquisa. A seguir, serão detalhados de forma crítica os métodos da pesquisa.

#### **Quadro 1: Métodos e técnicas da pesquisa.**

Tipo	Métodos	Técnicas
Questionamento e Pesquisa Bibliográfica	(1) Formulação do questionamento. (2) Busca por artigos históricos e recentes sobre o tema abordado. (3) Análise.	(1) Buscar se há outros artigos com o mesmo tema já proposto, formular os objetivos, limitação do trabalho. (2) Buscar fontes na <i>internet</i> em periódicos renomados para determinar os principais autores que irão colaborar com referências teórica para a pesquisa; (3) compilar o conteúdo de forma crítica, de modo a extrair apenas informações pertinentes e importantes para construção das estatísticas e fundamentação teórica do trabalho, bem como as referências.
Hipóteses	(4) Formular as Hipóteses (5) Coletar dados. (6) Analisar os dados.	(4) Foram determinadas as hipóteses, ação que limita a pesquisa, focando em apenas um ponto. (5) Todos os dados foram coletados de dados públicos, disponíveis para qualquer cidadão. (6) Foi escolhido um <i>software</i> específico para analisar os dados com base nas hipóteses determinadas anteriormente. Caso necessário, voltar a etapa (5).
Resultados e discussões	(7) Resultados. (8) Discussões. (9) Considerações Finais (10) Referências	(7) Apresentar o experimento por meio de gráficos e resoluções matemáticas, de modo a expor os principais resultados. (8) Realizar uma discussão imparcial sobre os eventos obtidos. (9) Analisar se há possibilidades de continuar o trabalho, caso haja novos pesquisadores interessados no assunto tratado.

Fonte: Autores.

Os dados estatísticos do Estado de São Paulo, coletados na SSP, foram filtrados para o período de 2018, fazendo o somatório de todas as ocorrências dos meses (Janeiros à Dezembro). Esses dados levam em consideração todas as delegacias das cidades envolvidas. O Quadro mostra quais foram as ocorrências observadas nas cidades do Vale do Paraíba e Litoral Norte e suas respectivas siglas.

**Quadro 2:** Ocorrências Registradas no Ano

OCORRÊNCIAS	SIGLAS
TOTAL DE ROUBO – OUTROS	(TRO)
ROUBO - OUTROS	(RO)
ROUBO DE VEÍCULO	(RV)
FURTO - OUTROS	(FO)
FURTO DE VEÍCULO	(FV)

Fonte: (KAHN, 2005).

No Quadro 2 foram descritas as ocorrências que serão analisadas. O Total de Roubos – Outros é o somatório dos roubos de cargas e roubo a banco. Roubos – Outros é a subtração de bens alheios com violência ou ameaças graves. Em Roubo de Veículos, todos os tipos de veículos motorizados foram levados em consideração. Neste caso, foram somados apenas os veículos e não peças ou acessórios. Furto – Outros contempla a subtração de bens alheios móveis. Furto de Veículos é o mesmo caso para Roubo de Veículos, considerando que no furto não ocorre violência, enquanto o ato de roubar gera ameaças à vítima e violência, por exemplo, assalto à mão armada.

As cidades avaliadas nesta pesquisa foram: São José dos Campos (SJC), Jacareí (JAC), Caçapava (CA), Paraibuna (PB), Santa Branca (SB), Jambuí (JB), Monteiro Lobato (ML), Igaratá (ITA), Taubaté (TBE), Pindamonhangaba (PMG), Campos do Jordão (CJ), Tremembé (TRE), São Bento do Sapucaí (SBS), São Luiz do Paraitinga (SLP), Santo Antônio do Pinhal (SAP), Natividade da Serra (NS), Lagoinha (LGA), Redenção da Serra (RS), Guaratinguetá (GTA), Lorena (LNA), Aparecida (APA), Cachoeira Paulista (CPA), Potim (PTM), Cunha (CHA), Piquete (PQE), Roseira (RSA), Canas (CNS), Cruzeiro (CRZ), Queluz (QLZ), Bananal (BNL), Lavrinhas (LVS), Silveiras (SLS), São José do Barreiro (SJB), Areias (ARS), Arapeí (ARS), Caraguatatuba (CTA), Ubatuba (UBA), São Sebastião (SS) e Ilhabela (IBA).

Após definidas as cidades, utilizou-se os conceitos de divisão sub-regional, de acordo com a Empresa Paulista de Planejamento Metropolitana (EMPLASA), como mostrado na Figura 2.

**Figura 2:** Sub-regiões do Vale do Paraíba e Litoral Norte



As Figura 3-7 mostram as tabelas de contingência com frequências absolutas observadas para ocorrências e sub-regiões 1-5, respectivamente.

**Figura 3:** Tabela de contingências para as ocorrências e sub-regiões 1

Ocorrências Policiais	Cidade								
	SJC	JAC	CA	PAB	SB	JB	ML	ITA	Active Margin
TRO	2844	854	428	20	10	4	4	23	4187
RO	2822	774	402	19	9	4	4	19	4053
RV	707	312	138	7	4	1	0	12	1161
FO	5688	1953	934	177	135	74	36	144	9141
FV	2193	606	239	6	23	1	1	7	3076
Active Margin	14254	4499	2141	229	181	84	45	205	21638

Fonte: Adaptado de (SSP, 2018)

**Figura 4:** Tabela de contingências para as ocorrências e sub-regiões 2

Ocorrências	Cidades										
	TBE	PMG	CJ	TRE	SBS	SLP	SAP	NS	LGA	RS	Active Margin
TRO	1138	563	62	93	1	13	3	1	3	3	1880
RO	1130	549	62	93	1	13	2	1	3	2	1856
RV	179	93	1	24	0	1	0	0	0	0	298
FO	2978	1198	414	263	71	115	42	64	36	42	5223
FV	578	174	7	52	5	5	4	7	2	0	834
Active Margin	6003	2577	546	525	78	147	51	73	44	47	10091

Fonte: Adaptado de (SSP, 2018)

**Figura 5:** Tabela de contingências para as ocorrências e sub-regiões 3

Ocorrências	Cidades									
	GTA	LNA	APA	CPA	PTM	CHA	POE	RSA	CNS	Active Margin
TRO	330	314	97	46	24	29	13	19	14	686
RO	324	312	96	46	24	29	12	17	11	871
RV	57	19	14	2	7	5	1	5	0	110
FO	917	744	439	246	80	154	79	99	53	2811
FV	68	43	27	29	4	13	2	5	3	194
Active Margin	1696	1432	673	369	139	230	107	145	81	4872

Fonte: Adaptado de (SSP, 2018)

**Figura 6:** Tabela de contingências para as ocorrências e sub-regiões 4

Ocorrências	Cidades								
	CRZ	OLZ	BNA	LVS	SLS	SJB	ARS	ARI	Active Margin
TRO	124	16	1	18	9	2	0	2	172
RO	124	14	1	15	9	2	0	2	167
RV	6	3	0	2	0	0	0	0	11
FO	767	131	33	59	48	26	18	12	1094
FV	59	12	4	1	0	1	0	0	77
Active Margin	1080	176	39	95	66	31	18	16	1521

Fonte: Adaptado de (SSP, 2018)

**Figura 7:** Tabela de contingências para as ocorrências e sub-regiões 5

Ocorrências	Cidade				
	CTA	UBA	SS	IBA	Active Margin
TRO	638	316	194	31	1179
RO	634	313	193	31	1171
RV	34	15	5	0	54
FO	2027	1335	1040	514	4916
FV	152	43	81	21	297
Active Margin	3485	2022	1513	597	7617

Fonte: Adaptado de (SSP, 2018)

Utilizando os dados das Figuras 3-7, pode-se determinar a estatística  $\chi^2$  (Chi-Quadrado Calculado) e as dimensões 1 e 2 da inércia principal total, levando em consideração que o valor-P precisa ser menor que 5%, entre as duas variáveis categóricas. Se o valor-P do for menor que 5%, então aceita-se a hipótese:  $H_1$  = ambas variáveis categóricas não são relacionadas estatisticamente de forma aleatória. Todos os testes  $\chi^2$  e a inércia principal total podem ser observados no Apêndice A.

Na sub-região 1, O valor-P do  $\chi^2 > 0.05$  (Sig.  $\chi^2 = 0,000$ ). Pode-se concluir que as dimensões 1 e 2 esclarecem, respectivamente (85,8% e 10,7%), a da inércia principal total.

Na sub-região 2, O valor-P do  $\chi^2 < 0.05$  (Sig.  $\chi^2 = 0,000$ ). As dimensões 1 e 2 esclarecem, respectivamente (86,7% e 11,9%), a da inércia total principal.

Na sub-região 3, O valor-P do  $\chi^2 < 0.05$  (Sig.  $\chi^2 = 0,000$ ). As dimensões 1 e 2 esclarecem, respectivamente (72,4% e 18,4%), a da inércia total principal.

Na sub-região 4, O valor-P do  $\chi^2 < 0.05$  (Sig.  $\chi^2 = 0,028$ ). As dimensões 1 e 2 esclarecem, respectivamente (66,8% e 21,1%), a da inércia total principal.

Na sub-região 5, O valor-P do  $\chi^2 < 0.05$  (Sig.  $\chi^2 = 0,000$ ). As dimensões 1 e 2 esclarecem, respectivamente (88,1% e 11,7%), a da inércia total principal.

Todos os biplots representando as variáveis categóricas de cada sub-região estão disponíveis no Apêndice B.

#### **4.1 DISCUSSÃO DOS RESULTADOS**

O objetivo específico desta pesquisa foi apresentar a estatística ANACOR como método alternativo de estudo de variáveis categóricas. Ao reunir uma grande quantidade de dados, torna-se difícil a visualização do que realmente é vantajoso para o estudo e, muitas vezes, dados importantes são desprezados. Por meio desse método, consegue-se reunir muitos dados e organizá-los de modo a estudar suas principais similaridades e determinar as principais associações das variáveis categóricas e, dessa maneira, definir as variáveis que mais contribuem.

O estudo promove uma associação profunda das variáveis categóricas que muitas vezes são ignoradas. Além disso, é fornecido também informações em formas de mapas,

facilitando a visualização da proximidade das variáveis categóricas, quais são as duas dimensões mais relevantes que explicam a inércia principal total, as dimensões mais representativas de cada categoria e se a associação estatística entre as variáveis se formou aleatoriamente ou não.

Vale lembrar que a técnica empregada objetiva analisar as significâncias das similaridades, ou seja, ao analisar, por exemplo, as cidades da sub-região 1 e as ocorrências, foram determinadas as coordenadas do biplot com base nas distribuições dos dados.

As dimensões principais das abscissas e ordenadas de cada variável categórica estão disponíveis no Apêndice A.

Na sub-região 1 do Vale do Paraíba (São José do Campos, Jacareí, Caçapava, Paraibuna, Santa Branca, Jambeiro, Monteiro Lobato e Iguaratá), os resultados foram: São José dos Campos possui a maior associação com Total de Roubos – Outros, Roubos – Outros e Furto – Outros. Das cidades analisadas, em Jacareí possui a maior associação com Roubo de Veículos. Caçapava associa-se com a categoria Furto – Outros. Na proporção da Inércia total, os dados contribuíram em 85,8% para primeira dimensão e 10,7% na segunda dimensão. É preciso observar que a soma acumulativa dos dados não foi 100%. Na categoria Ocorrências policiais, Furto – Outros apresentou maior contribuição com os dados na primeira dimensão da inércia, com 56,2% dos dados. Na segunda dimensão, a maior contribuição foi Roubo de Veículos, com 69,6%. Na categoria Cidade, Paraibuna contribuiu com maiores índices na primeira dimensão com 28,8%. Por outro lado, Jacareí apresentou maior contribuição para a segunda dimensão, com 32,6%.

Na sub-região 2 (Taubaté, Pindamonhangaba, Campos do Jordão, Tremembé, São Bento do Sapucaí, São Luiz do Paraitinga, Santo Antônio do Pinhal, Natividade da Serra, Lagoinha e Redenção da Serra), as associações são: Pindamonhangaba associa-se fortemente com as categorias Total de Roubos – Outros, Roubos – Outros e Roubo de Veículos. Taubaté possui associação com a categoria Furto – Outros e a categoria Tremembé associa-se com a categoria Furto de Veículos. Na proporção da Inércia total, os dados contribuíram com 86,7% para primeira dimensão e 11,9% na segunda dimensão, cuja soma acumulativa dos dados não foi 100%. Na categoria Ocorrências policiais, Furto – Outros apresentou maior contribuição com os dados na primeira dimensão da inércia, com 47,6% dos dados. Na segunda dimensão, a maior contribuição foi Roubo de

Veículos, com 82,6%. Na categoria Cidade, Campos do Jordão contribuiu com maiores índices na primeira dimensão com 35,3%. Por outro lado, Pindamonhangaba apresentou maior contribuição para a segunda dimensão, com 38,8%.

Na sub-região 3 (Guaratinguetá, Lorena, Aparecida, Cachoeira Paulista, Potim, Cunha, Piquete, Roseira e Canas), Lorena possui uma associação forte com Total de Roubos – Outros e Roubos – Outros. Aparecida associa-se com Furto – Outros e Piquete a Furto de veículos. Na proporção da Inércia total, os dados contribuíram em 72,4% para primeira dimensão e 18,4% na segunda dimensão, no entanto, a soma acumulativa dos dados não foi 100%. Na categoria Ocorrências policiais, Furto – Outros apresentou maior contribuição com os dados na primeira dimensão da inércia com 31,1% dos dados. Na segunda dimensão, a maior contribuição foi Roubo de Veículos, com 95,8%. Na categoria Cidade, Lorena contribuiu com maiores índices na primeira dimensão com 29,0%. Por outro lado, Lorena novamente apresentou maior contribuição para a segunda dimensão, com 29,1%.

Na sub-região 4 (Cruzeiro, Queluz, Bananal, Lavrinhas, Silveiras, São José do Barreiro, Areias e Arapeí), Lavrinhas possui associação com Total de Roubos – Outros e Roubos – Outros. Cruzeiro e Queluz possuem associações com Furto – Outros. Na proporção da Inércia total, os dados contribuíram em 66,8% para primeira dimensão e 21,1% na segunda dimensão, contudo, a soma acumulativa dos dados não foi 100%. Na categoria Ocorrências policiais, Total de Roubos – Outros apresentou maior contribuição com os dados na primeira dimensão da inércia com 37,9% dos dados. Na segunda dimensão, a maior contribuição foi Roubo de Veículos, com 62,6%. Na categoria Cidade, Lavrinhas contribuiu com maiores índices na primeira dimensão com 42,2%. Por outro lado, Areias apresentou maior contribuição para a segunda dimensão, com 32,5%.

Na sub-região 5 (Caraguatatuba, Ubatuba, São Sebastião e Ilhabela), Caraguatatuba associa-se com as categorias Total de Roubos – Outros, Roubos – Outros e Roubos de Veículos. Ubatuba, por outro lado, é associada a Furto – Outros. Por fim, São Sebastião apresenta forte associação com a categoria Furto de Veículos. Na proporção da Inércia total, os dados contribuíram em 88,1% para primeira dimensão e 11,7% na segunda dimensão, entretanto, a soma acumulativa dos dados não foi 100%. Na categoria Ocorrências policiais, Total de Roubos – Outros apresentou maior contribuição com os dados na primeira dimensão da inércia com 31,3% dos dados. Na segunda

dimensão, a maior contribuição foi Roubo de Veículos, com 95,0%. Na categoria Cidade, Ilhabela contribuiu com maiores índices na primeira dimensão com 61,2%. Por outro lado, Ubatuba apresentou maior contribuição para a segunda dimensão, com 60,4%.

## **CONSIDERAÇÕES FINAIS**

Mediante as análises foram demonstradas, por meio da ANACOR, as maiores associações dimensionais entre as ocorrências de criminalidade nas sub-regiões do Vale do Paraíba e Litoral Paulista.

O crescimento populacional e a exposição dessas pessoas em lugares que oferecem o mínimo ou nenhuma estrutura à educação, segurança, saúde, qualidade de vida e oportunidades de desenvolver habilidades profissionais para engajar no mercado contribuem para que fiquem vulneráveis a práticas criminais, aumentando exponencialmente a exposição de muitos jovens ao cenário da criminalidade.

Nas cidades que possuem maior associação com furtos e roubos (subtração de bens alheios, roubos de cargas, entre outros), as lojas poderiam investir em sistemas tecnológicos para minimizar as ocorrências, como por exemplo, etiquetas eletrônicas que acionam alarmes, aumentar a segurança através das câmeras, zerando as possibilidades de pontos “cegos” nas lojas. Os roubos e furtos de celulares são alvos constantes da criminalidade. É importante elaborar um boletim de ocorrência, pois caso a polícia militar aborde um indivíduo, conseguirá identificar o objeto roubado; e não comprar celulares que não possuem notas fiscais, prática muito comum atualmente nas redes sociais. Atualmente existem aplicativos que ajudam a polícia militar a encontrar os celulares roubados ou furtados por meio do IMEI (número exclusivo do aparelho móvel), que deverá ser fornecido para a polícia no momento que for realizado o boletim de ocorrência.

Apesar dos roubos e furtos em estabelecimentos comerciais serem relevantes, atualmente os celulares são os objetos de cobiça dos criminosos. Algumas orientações de prevenção são: não utilizar o celular nos bolsos de trás das roupas, guardar o aparelho eletrônico em partes isoladas nas bolsas, evitar a exposição constante em lugares onde há muita concentração de pessoas e, caso for registrar algum fato em público, observar ao redor e procurar um local seguro antes de expor o aparelho.

Os índices de roubos e furtos de veículos são grandes, principalmente nas cidades São José dos Campos e Jacareí. O principal motivo dessas ocorrências é para alimentar o mercado clandestino de peças, facilitando a compra por um preço mais barato de peças subtraídas de outrem. É essencial um aumento na circulação de viaturas pelas ruas, buscando minimizar os índices destes incidentes.

Os valores encontrados utilizando a ANACOR mostram uma visão geral sobre ocorrências nas cidades do Vale do Paraíba e Litoral tratadas nesta pesquisa, de modo a orientar a sociedade sobre os altos índices de criminalidade e mostrar que existem métodos preventivos e corretivos para recorrer em tais situações de furto ou roubo.

Sugestões futuras para novas pesquisas na área da análise de correspondência:

- a) Separar as cidades da região do Vale do Paraíba e Litoral Note por número de habitantes (acima de 100 mil habitantes; acima de 50 mil habitantes a 100 mil habitantes; acima de 10 mil habitantes a 50 mil habitantes e abaixo de 10 mil habitantes);
- b) Analisar um estudo multivariado, com base na Análise de Correspondência Múltipla (ACM), incluindo uma ou mais variáveis ao estudo;
- c) Considerar maiores períodos, por exemplo, de 2012 a 2018 e;
- d) Considerar todas as associações, fortes ou fracas, das variáveis relacionadas ao estudo.

## REFERÊNCIAS

- BRADU, D.; GABRIEL, K. R. The biplot as a diagnostic tool for models of two- way tables. *Technometrics*, [s. l.], v. 20, n. 1, p. 47–68, 1978.
- EMPRESA PAULISTA DE PLANEJAMENTO METROPOLITANO. Região metropolitana do vale do paraíba e litoral norte. 2016. Disponível em: <https://www.emplasa.sp.gov.br/RMVPLN>. Acessado em: 08/03/2019.
- FÁVERO; L.P. BELFIORE, P. Análise de dados, técnicas multivariadas exploratórias. Rio de Janeiro: Elsevier, 2015.
- KHANGAR, N. V.; KAMALJA, K. K. Multiple Correspondence Analysis and its applications. *Electronic Journal of Applied Statistical Analysis*, v. 10, n. 2, p. 432–462, 2017.
- KOTHARI, C. *Research methodology: methods and techniques*. New Age: 2004
- KAHN, T. Estatística de criminalidade: manual de interpretação. Coordenadoria de Análise e Planejamento (CAP) - Secretara da Segurança Pública, 2005. Disponível em: <http://www.ssp.sp.gov.br/Estatistica/download/manual.pdf>. Acessado em: 08/03/2019.
- SCHNEIDER, I. *The Mathematical Intelligencer* (2005) 27: 83. <https://doi.org/10.1007/BF02984818>
- SILVA, R. R.; BENIN, G. Análises Biplot: conceitos, interpretações e aplicações. *Ciência Rural*, [s.l.], v. 42, n. 8, p. 1404–1412, 2012.
- SECRETARIA DA SEGURANÇA PÚBLICA. Dados estatísticos do Estado de São Paulo. 2018. Disponível em: <https://www.ssp.sp.gov.br/Estatistica/Pesquisa.aspx>. Acessado em: 17/03/2019.

TORRES-SALINAS, D.; ROBINSON-GARCÍA, N.; JIMÉNEZ-CONTRERAS, E.; HERRERA, F.; LÓPEZ-CÓZAR, E. D. On the use of biplot analysis for multivariate bibliometric and scientific indicators. *Journal of the American Society for Information Science and Technology*, [s. l.], v. 64, n. 7, p. 1468–1479, 2013.

YAN, W.; TINKER, N. A. Biplot analysis of multi-environment trial data: Principles and applications. *Canadian Journal of Plant Science*, [s. l.], v. 86, n. 3, p. 623–645, 2006.