



## Demografia e criminalidade: um estudo baseado em métodos quantitativos aplicados a “crimes de rua”

Carlos Roberto Souza Carmo<sup>1</sup>

### Resumo

Ao problematizar um possível relacionamento entre certa categoria de crimes de natureza violenta com variáveis demográficas da população de Uberlândia-MG, essa pesquisa teve por objetivo geral analisar como fatores relativos ao gênero e à idade da população daquela cidade poderiam caracterizar como possíveis determinantes das ocorrências relacionadas aos roubos consumados a transeuntes, ocorridos no ano de 2010. A partir da análise de regressão linear múltipla, pelo método *stepwise*, foi identificado que três faixas etárias e o gênero masculino caracterizaram-se como possíveis determinantes dos “crimes de rua” analisados nesse estudo. Além de comparar os resultados desse estudo com os resultados de estudos realizados em dois outros países da América Latina, entre outros achados, pôde-se inferir que a modelagem analítica identificada nessa pesquisa poderia ser utilizada como uma função preditiva daqueles crimes, se admitida uma margem de erro de 8,70%, para um intervalo de confiança de 95%.

**Palavras-chave:** Métodos quantitativos aplicados. Criminalidade. Demografia.

---

<sup>1</sup> Mestre em Ciências Contábeis pela Pontifícia Universidade Católica de São Paulo e Professor da Universidade Federal de Uberlândia-UFU. End.: Av. João Naves de Ávila, 2121, Bloco F - Sala 1F 215, Santa Mônica, Uberlândia-MG – Brasil, CEP 38400-902. E-mail: carlosjj2004@hotmail.com. Recebimento: 21/11/2013 • Aceite: 18/12/2013

## **Demography and crime: a study based on quantitative methods applied to "street crimes"**

### **Abstract**

To discuss a possible relationship between certain category of crimes of a violent nature with demographic population of Uberlândia-MG, this study aimed at analyzing factors related to gender and age of the population of that city could characterize as possible determinants of related events to consummate the theft passersby, occurred in 2010. From the multiple linear regression analysis, stepwise, it was identified that the three age groups and males were characterized as potential determinants of "street crimes" analyzed in this study. In addition to comparing the results of this study with the results of studies conducted in two other countries of Latin America, among other findings, it was inferred that the analytical modeling identified in this study could be used as a predictive function of those crimes, if admitted margin error of 8.70%, for a confidence interval of 95%.

**Keywords:** Quantitative methods applied. Crime. Demographics.

### **Introdução**

A violência sempre foi caracterizada como um fenômeno social inerente a todas as sociedades humanas ao longo de toda a sua história. A despeito dessa constatação, nos dias atuais, experimenta-se uma realidade nunca vista antes, em que, a escalada da violência não tem limites, consolidando-se assim como um fenômeno social do mundo contemporâneo (CAMARGO, 2002).

Diante da disseminação da criminalidade, enquanto um dos muitos meios de violência, a sociedade em geral padece dominada pelo pânico e pela sensação de mal-estar e desequilíbrio social (CALDEIRA, 2002; ZALUAR, 2002).

Posto esse quadro, observa-se que apesar de toda a evolução econômica constatada no Brasil ao longo da última década, os crescentes índices de

criminalidade nas cidades brasileiras fazem com que o assunto seja a pauta de vários debates políticos e estudos (BEATO; REIS, 2000; CANO, 2006; SAPORI, 2007).

Nesse contexto, observa-se que os crimes de natureza violenta já não são mais uma característica das grandes capitais brasileiras. Mais especificamente, na cidade de Uberlândia-MG, cuja população era composta por 604.013 habitantes, segundo o Censo Demográfico de 2010 (IBGE, 2013), foram observadas 3.227 ocorrências policiais relativas a crimes considerados violentos (roubos, homicídios, sequestros, estupro e outros assemelhados), e, ainda, 10.133 ocorrências relativos a crimes de diversas outras naturezas.

Ao considerar que, daquelas 3.227 ocorrências policiais relativas a crimes considerados violentos, mais de 25% ( $[817/3.227] \times 100 \approx 25,32\%$ ) referiam-se a roubos consumados a transeuntes (817 ocorrências), a presente pesquisa teve por objetivo geral analisar como fatores demográficos relativos ao gênero e à idade da população de Uberlândia-MG poderiam caracterizar como possíveis determinantes de crimes violentos ocorridos nas ruas daquela cidade. Para tanto, este processo investigativo foi conduzido a partir da seguinte problemática: como faixa etária e gênero, ambos referentes à população de Uberlândia-MG e identificadas a partir dos dados do Censo demográfico de 2010, podem caracterizar como possíveis determinantes das ocorrências relativas a crimes violentos ocorridos nas ruas daquela cidade, no que se refere aos roubos consumados contra transeuntes, naquele mesmo ano (2010)?

Para responder àquele questionamento direcionador, foram cumpridas cinco etapas básicas que podem ser consideradas os objetivos específicos dessa pesquisa de caráter empírico-quantitativo e que, por sua vez, deram origem às quatro outras seções que compõem esse artigo, além da presente introdução. Inicialmente, foi constituído um breve referencial teórico para caracterização da temática relacionada à criminalidade urbana e seu possível relacionamento com as variáveis demográficas escolhidas para esse estudo. Sendo que, essa etapa deu origem à seção dois deste artigo.

Em seguida, foram descritas as fontes dos dados utilizados nessa pesquisa e, ainda, pesquisado o ferramental analítico capaz de identificar possíveis relacionamentos entre a variável de estudo, (roubos consumados contra transeuntes) e as variáveis demográficas consideradas como possíveis determinantes daquela modalidade de crime urbano (Faixa etária e gênero). Essa etapa foi detalhada na seção três deste trabalho.

Depois de constituída a plataforma teórica e identificados os dados e a metodologia analítica deste estudo, na seção quatro, procedeu-se à análise dos dados e à apresentação dos respectivos resultados.

Finalmente, na quinta e última seção deste trabalho, são apresentadas as considerações finais acerca de todo esse processo de investigação científica, bem como, suas limitações e sugestões para futuros trabalhos científicos sobre a temática em estudo nessa pesquisa.

### **Referencial Teórico**

Ao comparar a evolução da criminalidade com o processo de concentração populacional urbana, Paixão (1983) afirma que a industrialização da produção provocou o crescimento da urbanização que, em função da desorganização social que sempre lhe foi característica, trouxe consigo o crescimento da pobreza e o descontrole social que, por sua vez, alavancaram a criminalidade violenta.

Conforme já dito neste trabalho, a despeito de todo o crescimento econômico experimentado pela sociedade brasileira nas últimas décadas, o mesmo não acontece na ordem pública. Ou seja, conforme observa Saporì (2007, p. 98), “[...] as duas últimas décadas de democracia assistiram a uma considerável deterioração da capacidade do poder público para controlar a criminalidade e a violência”. Ou ainda, nas palavras de Zaluar (2002, p. 55):

[...] pode-se ver que a extensão e a complexidade do problema vêm consumindo grandes esforços de uma sociedade perplexa e desorganizada para o enfrentamento

do problema e que se vê drenada pelo aparato concreto da violência, que afeta de forma avassaladora a vida das pessoas em diversos níveis e de diversas formas.

Com relação específica às ocorrências de crimes urbanos, fatores como a exposição a riscos e a proximidade entre vítima e agressor, entre outros, são determinantes da natureza dos delitos para os quais o ofensor se encontra motivado (BEATO FILHO *et al*, 2002).

Beato Filho *et al* (2002) argumentam que a circulação e o ato de frequentar lugares públicos aumentam a proximidade entre vítima e agressor, e, ainda, características socioeconômicas e atributos como idade e gênero podem constituir em fatores direcionadores da ocorrência de crimes urbanos, uma vez que este tipo de característica humana influencia diretamente os hábitos de uma população.

Com relação à exposição, Beato Filho *et al* (2002) também argumentam que os fatores idade e gênero, por determinarem a frequência dos indivíduos em lugares públicos, pode influenciar também a sua exposição à ocorrência de delitos urbanos. Pois, conforme observa Félix (2002), os crimes com violência são mais frequentes nas regiões centrais das áreas urbanas, uma vez que a aglomeração de pessoas favorece tanto a subtração do bem alheio quanto a fuga do criminoso.

Ao analisar os resultados de várias pesquisas sobre roubos, Gottfredson e Hirschi (1990) afirmam que 70% desses eventos ocorrem nas ruas. Sendo que, em cerca de 50% desse tipo de ocorrência ocorre o uso de armas, e, ainda, em mais de 60% dos casos, esses crimes são cometidos por mais de uma pessoa, e, um fator muito relevante, tais criminosos escolhem suas vítimas nas proximidades dos locais onde residem (GOTTFREDSON; HIRSCHI, 1990).

Esteves (1995) destaca que a idade e o gênero de uma pessoa relacionam-se de forma muito estreita com a maneira com que os indivíduos realizam seus deslocamentos, e, isso por sua vez, pode caracterizar como um possível determinante da possibilidade de tornar uma vítima, ou não, de “um crime de rua”.

O relacionamento dos fatores como idade e gênero e a ocorrência de crimes urbanos apontados por Esteves (1995) é corroborado por Cohen e Felson (1979), e, ainda, por Clarke e Felson (1998).

Cruz, Arguello e González (2001) investigaram os determinantes da vitimização na região metropolitana de El Salvador. Os autores constataram que os homens representam a maioria das vítimas em casos de crimes violentos, sendo que, com relação à idade, foi observado que quanto mais jovens os indivíduos, maior probabilidade de serem vítimas deste tipo de violência urbana.

Ao realizarem um estudo semelhante ao de Cruz, Arguello e González (2001), porém, na região metropolitana de Lima, no Peru, Eyzaguirre e Puga (2001) constataram que mulheres na faixa etária entre 18 e 24 anos, entre outros fatores, têm maior probabilidade de serem vítimas de roubo. Ainda segundo Eyzaguirre e Puga (2001), as evidências coletadas por eles permitiram confirmar sua teoria acerca do fato que pessoas mais jovens se expõem mais e que indivíduos com menor capacidade de proteção são mais atrativos, aumentando a probabilidade de sofrerem crimes.

O Código Penal Brasileiro (BRASIL, 1940), em seu título II, trata dos crimes contra o patrimônio. No seu capítulo 1, aquele Código define furto como o ato de subtrair, para si ou para outrem, coisa alheia móvel (BRASIL, 1940). No seu capítulo 2, o roubo é definido como o ato de subtrair coisa móvel alheia, para si ou para outrem, mediante grave ameaça ou violência à pessoa.

De acordo com a Diretriz Auxiliar de Operações 01/94 da Polícia Militar de Minas Gerais (MINAS GERAIS, 1994), as ocorrências criminais são divididas por naturezas e classificadas a partir de códigos que, entre outros fatores, têm por objetivo facilitar a confecção de banco de dados

Ainda segundo a Diretriz Auxiliar de Operações 01/94 da Polícia Militar de Minas Gerais (MINAS GERAIS, 1994), os crimes referentes à subtração da coisa alheia móvel para si ou para outrem, mediante grave ameaça ou violência à pessoa, são codificados sob o número “C 05027”, cuja denominação técnica é “ROUBO CONSUMADO A TRANSEUNTE”. Esse tipo de crime é executado sem

a utilização de armas, sendo o modo ação caracterização, principalmente, pelo emprego de força física (MINAS GERAIS, 1994).

Desse modo, parece razoável admitir que as variáveis relacionadas ao gênero e à idade da população podem caracterizar-se como possíveis determinantes das ocorrências desse tipo de crime contra os usuários das vias públicas dos bairros da cidade alvo do presente estudo.

### **Procedimentos Metodológicos**

De um total de 817 ocorrências registradas em 83 zonas urbanas da cidade Uberlândia-MG, essa investigação teve como variável de estudo as 760 observações relativas às ocorrências de roubos consumados contra transeuntes, em 53 bairros daquela cidade. Sendo que, essas informações foram obtidas mediante consulta às planilhas cujos dados foram tabulados pelo destacamento da Polícia Militar de Minas Gerais, atuante naquela cidade, conforme descrito no item “i” do Quadro 1.

As possíveis variáveis explicativas utilizadas nesse estudo, portanto, a quantidade de habitantes pertencentes aos gêneros masculino e feminino e, ainda, as quantidades habitantes pertencentes às 19 faixas etárias, também pertencentes a 53 bairros daquela cidade, foram obtidas junto ao site da Prefeitura Municipal de Uberlândia, cuja tabulação teve como base o Censo Demográfico de 2010 realizado pelo IBGE (UBERLÂNDIA, 2013), conforme descrito do item “ii” ao item “xxii” do Quadro 1.

Sendo que, ainda com relação as dados e variáveis utilizados nesse estudo, cabe observar que não foram analisadas todas as 817 ocorrências registradas ao longo do ano de 2010, pois, os relatórios disponibilizados no site da Prefeitura Municipal de Uberlândia (UBERLÂNDIA, 2013) não continham informações completas relativas às variáveis explicativas (gênero e idade) de todos os bairros onde elas ocorreram.

Contudo, há que considerar que as 760 observações relativas aos roubos consumados contra transeuntes que compuseram a amostra dessa pesquisa

representam 93,02% ( $=\frac{760}{817} \times 100 \approx 93,02\%$ ) das ocorrências dessa natureza registradas em 2010, naquela cidade.

**Quadro 1** – Descrição das variáveis utilizadas no estudo

Identificação no banco dados		Unidade	Fonte
i	ROUBO_CONSUMADO_A_TRANSEUNTE	Ocorrência	Planilhas da PMMG (2010)
ii	HabGenMasc	Quantidade de habitantes (unidade)	Planilhas disponibilizadas no site da Secretaria Municipal de Planejamento Urbano - Diretoria de Pesquisas Integradas (2010)
iii	HabGenFem		
iv	Idade_ate_1_ano		
v	Idade_01_a_04		
vi	Idade_05_a_09		
vii	Idade_10_a_14		
viii	Idade_15_a_19		
ix	Idade_20_a_24		
x	Idade_25_a_29		
xi	Idade_30_a_34		
xii	Idade_35_a_39		
xiii	Idade_40_a_44		
xiv	Idade_45_a_49		
xv	Idade_50_a_54		
xvi	Idade_55_a_59		
xvii	Idade_60_a_64		
xviii	Idade_65_a_69		
xix	Idade_70_a_79		
xx	Idade_80_a_89		
xxi	Idade_90_a_99		
xxii	Idade_100_anos_ou_mais		

**Fonte:** elaborado pelo autor com base nos dados da pesquisa.

Com relação ao ferramental estatístico utilizado para análise do possível relacionamento entre a variável de estudo (roubos a transeuntes) e suas respectivas variáveis explicativas (gênero e idade da população), foi



empregada a análise de regressão linear múltipla, pelo método *stepwise*, com o auxílio do pacote estatístico para as ciências sociais *SPSS (Statistical Package for the Social Sciences)* versão 15.0.

Para Stevenson (1986), a análise de regressão linear tem por objetivo identificar como uma variável (variável de estudo ou independente) se comporta em função de alterações ocorridas no comportamento de uma ou mais variáveis independentes ou explicativas.

Sanz (2010) complementa a explicação de Stevenson (1986) ao afirmar que a análise de regressão linear produz uma função matemática analítica e explicativa ( $\hat{Y}$ ) do comportamento de uma variável dependente ( $Y$ ), a partir das informações de um conjunto de variáveis explicativas ( $X$ ), conforme descrito pela Formulação 1.

$$\hat{Y} = b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k + u \quad (1)$$

Na formulação 1, " $b_1, b_2, \dots, b_k$ " são coeficientes que representam os efeitos de cada uma de as variáveis explicativas " $X_1, X_2, \dots, X_k$ " sobre a variável dependente ou variável de estudo ( $Y$ ) (SANZ, 2010). Já o coeficiente " $b_0$ " representa o termo constante, ou independente, do modelo ( $\hat{Y}$ ) pesquisado pela regressão linear (SANZ, 2010). Ao final daquela fórmula, o " $u$ " representa o termo de erro ( $\hat{Y} - Y$ ) do modelo analítico ( $\hat{Y}$ ) em relação aos valores reais ( $Y$ ) das observações referentes à variável estudada (SANZ, 2010).

Com relação ao método *stepwise*, Fávero *et al* (2009) explicam que ele consiste em realizar a inclusão e a exclusão de cada uma das variáveis explicativas ( $X_1, X_2, \dots, X_k$ ), de forma gradativa, até que se identifiquem somente aquelas que melhor se adaptem à modelagem explicativa ( $\hat{Y}$ ) do comportamento da variável dependente ( $Y$ ).

Para validação da modelagem explicativa ( $\hat{Y}$ ) do comportamento das ocorrências relativas aos roubos consumados contra transeuntes que compuseram a amostra dessa pesquisa, além das análises envolvendo coeficiente de correlação (R), coeficiente de determinação (R<sup>2</sup>), estatística "F",

estatística “t”, foram realizados testes estatísticos para diagnóstico da presença de autocorrelação de resíduos (estatística “dw” de Dubin-Watson), multicolinearidade (estatísticas VIF - *variance inflation factor* - e Tolerância - tolerance) e heterocedasticidade (teste de Pesarán-Pesarán)

O coeficiente de correlação (R) avalia o grau de associação do comportamento da variável de estudo (Y) em relação ao comportamento das variáveis explicativas ( $X_1, X_2, \dots, X_k$ ), e, coeficiente de determinação (R<sup>2</sup>) traduz a capacidade explicativa do modelo pesquisado ( $\hat{Y}$ ) (FÁVERO *et al*, 2009).

Segundo Cunha e Coelho (2011), a estatística “f” avalia se a combinação linear das variáveis explicativas ( $X_1, X_2, \dots, X_k$ ) exerce influência sobre a variável de estudo (Y). Sendo que, para que seja descartada a possibilidade das variáveis explicativas conduzirem a um R<sup>2</sup> igual a zero, a modelagem pesquisada a partir da regressão linear ( $\hat{Y}$ ) deve apresentar valores significativamente distantes de 0,05 (Sig. da Estat. “f” < 0,05) (CUNHA; COELHO, 2011).

A estatística “t” tem por finalidade avaliar a possibilidade dos coeficientes da modelagem matemática explicativa do comportamento da variável de estudo ( $b_1, b_2, \dots, b_k$ ) serem diferente de zero, para tanto, ela deve apresentar um valor parâmetro com significância inferior a 0,05 (Sig. da Estat. “t” < 0,05) (CUNHA; COELHO, 2011).

Ao abordar os pressupostos básicos de modelagens explicativas baseadas na análise de regressão linear ( $\hat{Y}$ ), Hair Júnior *et al* (1998) destacam três problemas que normalmente podem invalidar os modelos identificados por meio dessa técnica estatística: autocorrelação dos resíduos, multicolinearidade e heterocedasticidade.

Em relação à autocorrelação dos resíduos ( $\hat{Y} - Y$ ), Fávero *et al* (2009) explicam que esse tipo de problema surge quando variáveis explicativas relevantes não foram incluídas na modelo pesquisado, o que faz com que resíduos ( $\hat{Y} - Y$ ) incorporem os efeitos dessas variáveis, apresentando, assim, correlação indesejada com a variável dependente (Y). Fávero *et al* (2009) observam que tais problemas podem ser diagnosticados mediante o uso da estatística “dw”

ou teste de Durbin-Watson, cujo valor desejável deve ser próximo de 2, mais especificamente, entre 1,77 e 2,23, considerando a quantidade e a natureza das variáveis desse estudo.

Fávero *et al* (2009) afirmam que a multicolinearidade surge quando as variáveis explicativas ( $X_1, X_2, \dots, X_k$ ) apresentam correlação elevada entre si. Para diagnóstico de problemas dessa natureza devem ser utilizadas as estatísticas VIF (*variance inflation factor*) e Tolerância (*tolerance*), cujos valores parâmetros desejáveis devem ser inferiores a 5 para a estatística VIF, e, maiores que 0,20 para estatística de Tolerância (FÁVERO *et al*, 2009).

Os problemas envolvendo a heterocedasticidade surgem quando os resíduos gerados a partir de uma modelagem pesquisada com base na regressão linear múltipla ( $\hat{Y} - Y$ ) apresentam correlação com uma ou mais variáveis explicativas ( $X_1, X_2, \dots, X_k$ ) (FÁVERO *et al*, 2009).

Cunha e Coelho (2011) afirmam que problemas dessa natureza podem ser diagnosticados a partir da aplicação do teste de Pesarán-Pesarán, cuja operacionalização consiste na regressão do quadrado dos resíduos padronizados ( $ZRE^2$ ) em função do quadrado dos valores estimados pela modelagem pesquisada ( $ZPE^2$ ), sendo que, o “Sig. do valor-p” da estatística “F” dessa regressão deve apresentar valor superior a 0,05 para que seja descartada a presença de heterocedasticidade, ou, ainda, a existência de correlação entre os resíduos ( $\hat{Y} - Y$ ) e uma, ou mais, variáveis explicativas ( $X_1, X_2, \dots, X_k$ ).

Diante da problemática proposta para essa pesquisa, da natureza dos dados utilizados e dos procedimentos aplicados para análise, esse trabalho pode ser classificado como uma pesquisa do tipo empírico-analítica (FACHIN, 2001).

### **Análise dos Dados e Resultados**

Ao iniciar o processo de análise das variáveis demográficas que poderiam caracterizar como possíveis determinantes da criminalidade relacionada às 760 observações relativas às ocorrências de roubos consumados contra

transeuntes, registradas ao longo do ano de 2010 em 53 bairros da cidade de Uberlândia, foram identificadas 4 modelagens explicativas, conforme evidenciado na Tabela 1. Sendo que, em todas elas o modelo pesquisado continha um termo constante ou intercepto.

Entretanto, ao analisar os modelos pesquisados inicialmente, foi observado que, independente das variáveis relacionadas à faixa etária da população e gênero ali identificadas apresentarem combinação linear satisfatória (Sig. da Estat. “f” < 0,05) e, ainda, relevantes coeficientes de correlação (R) e de determinação (R<sup>2</sup>), a análise da estatística “t” dos termos constantes daqueles modelos mostrou-se insatisfatória. Pois, os respectivos valores parâmetros apresentaram significância estatística superior a 0,05 (Sig. da Estat. “t” < 0,05), o que não permite descartar a hipótese daquele termo (intercepto ou constante) ser estatisticamente igual a zero, conforme destacado em vermelho na antepenúltima coluna da Tabela 1.

**Tabela 1** – Análise dos coeficientes das modelagens pesquisadas pela regressão linear múltipla pelo método *stepwise*, **com a presença de termo constante ou intercepto (b)**<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-,029	2,346		-,012	<b>,990</b>		
	Idade_80_a_89	,119	,015	,749	8,068	,000	1,000	1,000
2	(Constant)	-4,260	2,386		-1,786	<b>,080</b>		
	Idade_80_a_89	,082	,016	,519	5,022	,000	,644	1,554
	Idade_20_a_24	,009	,002	,385	3,720	,001	,644	1,554
3	(Constant)	2,088	1,920		1,088	<b>,282</b>		
	Idade_80_a_89	,054	,012	,338	4,339	,000	,574	1,743
	Idade_20_a_24	,063	,008	2,711	8,032	,000	,031	32,750
	HabGenMasc	-,011	,002	-2,262	-7,062	,000	,034	29,484
4	(Constant)	,713	1,818		,392	<b>,697</b>		
	Idade_80_a_89	,079	,014	,496	5,678	,000	,386	2,590
	Idade_20_a_24	,068	,007	2,955	9,240	,000	,029	34,767
	HabGenMasc	-,021	,003	-4,244	-6,122	,000	,006	163,424
	Idade_35_a_39	,053	,017	1,676	3,158	,003	,010	95,747

<sup>a</sup>. Dependent Variable: ROUBO\_CONSUMADO\_A\_TRANSEUNTE

**Fonte:** elaborado pelo autor a partir das análises realizadas no *SPSS*, com base nos dados da pesquisa.

Diante dessa primeira evidência, o processo de análise foi reiniciado, porém, contemplando a estimação de modelagens explicativas sem o termo constante ou intercepto, pois, conforme observa Sanz (2010), em alguns casos, desde que os pressupostos teóricos utilizados no processo de investigação assim demandarem, podem ser pesquisadas equações explicativas de determinada variável de estudo sem que seja estimado o respectivo termo constante ou independente ( $b_0$ ). Isso faria com que fosse gerada uma equação explicativa com a notação descrita pela Formulação 2.

$$\hat{Y} = b_1X_1 + b_2X_2 + \dots + b_kX_k + u \quad (2)$$

Nesse sentido, ao pesquisar as possíveis variáveis demográficas que pudessem compor um modelo analítico-explicativo da modalidade criminal analisada nesse estudo, a aplicação da análise de regressão linear identificou quatro modelagens, conforme demonstrado na Tabela 2.

Aqui, vale lembrar que o processo regressão linear pelo método *stepwise* promove a inserção das variáveis explicativas de forma hierárquica até que seja identificada a melhor modelagem explicativa da variável de estudo, ou seja, aquele modelo com maior coeficiente de correlação (R) e, por consequência, maior poder explicativo, que é medido pelo respectivo coeficiente de determinação (R<sup>2</sup>).

Assim, conforme demonstrado na Tabela 2, o melhor modelo explicativo das ocorrências relativas a roubos consumados contra transeuntes em estudo nessa investigação foi o modelo 4 (R= 0,959, e, R<sup>2</sup>=0,919). Sendo que a estatística “dw” (Dubin-Watson) daquele modelo foi de 1,927, portanto, entre 1,77 e 2,23, o que permite descartar a existência de problemas relacionados à autocorrelação dos resíduos.

**Tabela 2** – Resumo das modelagens pesquisadas pela regressão linear múltipla pelo método *stepwise*, sem a presença de termo constante ou intercepto ( $b_0$ )<sup>1,4</sup>

Model	R	R Square <sup>a</sup>	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	,866 <sup>b</sup>	,750	,746	11,01115	
2	,890 <sup>c</sup>	,792	,784	10,14900	
3	,949 <sup>d</sup>	,901	,895	7,08017	
4	,959 <sup>e</sup>	,919	,913	6,44089	1,927

a. For regression through the origin (the no-intercept model), R Square measures the proportion of the variability in the dependent variable about the origin explained by regression. This CANNOT be compared to R Square for models which include an intercept.

b. Predictors: Idade\_80\_a\_89

c. Predictors: Idade\_80\_a\_89, Idade\_20\_a\_24

d. Predictors: Idade\_80\_a\_89, Idade\_20\_a\_24, HabGenMasc

e. Predictors: Idade\_80\_a\_89, Idade\_20\_a\_24, HabGenMasc, Idade\_35\_a\_39

f. Dependent Variable: ROUBO\_CONSUMADO\_A\_TRANSEUNTE

g. Linear Regression through the Origin

**Fonte:** elaborado pelo autor a partir das análises realizadas no *SPSS*, com base nos dados da pesquisa.

Ao dar continuidade ao processo analítico empregado nessa investigação, os dados da análise de variância dos modelos pesquisados, descritos na Tabela 3, permitem inferir que a combinação linear das variáveis explicativas, identificadas a partir da análise de regressão linear com exclusão do termo constante ou intercepto, foi considerada satisfatória, pois, a significância dos valores parâmetros das respectivas estatísticas “F” foram todos menores que 0,05, conforme pode ser visto na última coluna da Tabela 3.

**Tabela 3** – Tabela ANOVA (Análise de Variância) das modelagens pesquisadas pela regressão linear múltipla pelo método *stepwise*, **sem a presença de termo constante ou intercepto** (*b*)<sup>f,g</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	18945,243	1	18945,243	156,255	,000 <sup>a</sup>
	Residual	6304,757	52	121,245		
	Total	25250,000 <sup>b</sup>	53			
2	Regression	19996,884	2	9998,442	97,070	,000 <sup>c</sup>
	Residual	5253,116	51	103,002		
	Total	25250,000 <sup>b</sup>	53			
3	Regression	22743,563	3	7581,188	151,234	,000 <sup>d</sup>
	Residual	2506,437	50	50,129		
	Total	25250,000 <sup>b</sup>	53			
4	Regression	23217,233	4	5804,308	139,913	,000 <sup>e</sup>
	Residual	2032,767	49	41,485		
	Total	25250,000 <sup>b</sup>	53			

a. Predictors: Idade\_80\_a\_89

b. This total sum of squares is not corrected for the constant because the constant is zero for regression through the origin.

c. Predictors: Idade\_80\_a\_89, Idade\_20\_a\_24

d. Predictors: Idade\_80\_a\_89, Idade\_20\_a\_24, HabGenMasc

e. Predictors: Idade\_80\_a\_89, Idade\_20\_a\_24, HabGenMasc, Idade\_35\_a\_39

f. Dependent Variable: ROUBO\_CONSUMADO\_A\_TRANSEUNTE

g. Linear Regression through the Origin

**Fonte:** elaborado pelo autor a partir das análises realizadas no *SPSS*, com base nos dados da pesquisa.

A análise da estatística “t” dos coeficientes das modelagens pesquisadas sinaliza que todos eles são significativamente diferentes de zero, pois, com base nas informações resumidas na Tabela 4, pode-se observar que todos os valores parâmetros daquela estatística apresentaram significância inferior a 0,05 (Sig. da Estat. “t” < 0,05).

Ao analisar as estatísticas de VIF e Tolerance, disponíveis nas duas últimas colunas da Tabela 4, percebe-se que ambas não se apresentaram inferiores a 5 no caso da estatística VIF, e, superiores a 0,20 no caso da estatística Tolerance. Dessa forma, pode-se afirmar que as variáveis explicativas consideradas significativamente relevantes para explicar o comportamento da variável em estudo nessa investigação apresentam multicolinearidade, ou, ainda, estão correlacionadas entre si. Contudo, conforme afirmam Blanchard (1967), Neter e Wasserman (1974), Kenedy (2003), Gujarati (2006) e, ainda, Gujarati e Porter (2011), a multicolinearidade não é um problema considerado grave quando não afeta a variância do preditor, e, algumas situações, como é o caso dessa pesquisa, ela está em grande parte relacionada ao tamanho e à

natureza da amostra de pesquisa (BLANCHARD, 1967; GUJARATI, 2006; KENEDY, 2003).

**Tabela 4** – Análise dos coeficientes das modelagens pesquisadas pela regressão linear múltipla pelo método *stepwise*, **sem a presença de termo constante ou intercepto** ( $b_0$ )<sup>a,b</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	Idade_80_a_89	,119	,009	,866	12,500	,000	1,000	1,000
2	Idade_80_a_89	,075	,016	,548	4,627	,000	,291	3,435
	Idade_20_a_24	,007	,002	,378	3,195	,002	,291	3,435
3	Idade_80_a_89	,058	,012	,427	5,072	,000	,280	3,569
	Idade_20_a_24	,060	,007	3,287	8,186	,000	,012	81,236
	HabGenMasc	-,011	,001	-2,828	-7,402	,000	,014	73,498
4	Idade_80_a_89	,081	,012	,591	6,519	,000	,200	5,001
	Idade_20_a_24	,068	,007	3,723	9,610	,000	,011	91,345
	HabGenMasc	-,021	,003	-5,726	-6,187	,000	,002	521,484
	Idade_35_a_39	,055	,016	2,339	3,379	,001	,003	291,580

a. Dependent Variable: ROUBO\_CONSUMADO\_A\_TRANSEUNTE

b. Linear Regression through the Origin

**Fonte:** elaborado pelo autor a partir das análises realizadas no *SPSS*, com base nos dados da pesquisa.

Ou seja, ao se levar em consideração que a população de uma cidade, por mais numerosa que ela seja, tende a apresentar características demográficas aproximadamente semelhantes, ainda que estratificada por bairros, o fenômeno da multicolinearidade relacionado às variáveis demográficas gênero e faixa etária pode ser considerado inerente à natureza da amostra dessa pesquisa, e, assim, aceitável, conforme preceituado pela literatura relacionada ao processo de análise por meio da regressão linear múltipla (BLANCHARD, 1967; GUJARATI, 2006; GUJARATI; PORTER, 2011; KENEDY, 2003; NETER; WASSERMAN, 1974).

**Tabela 5** – Tabela ANOVA (Análise de Variância) da modelagem pesquisada pela regressão linear desenvolvida para operacionalização do teste de Pesarán-Pesarán<sup>b</sup>



Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1,266	1	1,266	,767	,385 <sup>a</sup>
	Residual	84,130	51	1,650		
	Total	85,395	52			

a. Predictors: (Constant), ZPR\_\_2

b. Dependent Variable: ZRE\_\_2

**Fonte:** elaborado pelo autor a partir das análises realizadas no SPSS, com base nos dados da pesquisa.

Uma vez que a significância do valor parâmetro da estatística “f” da regressão operacionalizada para realização do teste de Pesarán-Pesarán foi superior a 0,05, pode-se afirmar que os resíduos do modelo pesquisado são homocedásticos, portanto, não apresentam problemas relacionados à heterocedasticidade, conforme pode ser observado na Tabela 5.

Diante do exposto até esse ponto, pode-se afirmar que a modelagem explicativa do comportamento da variável em estudo nessa investigação foi estatisticamente válida e pode ser representada pela Formulação 3.

$$\begin{aligned}
 & \textit{Quantidade de roubos consumados a transeuntes}_{(\textit{por bairro, em 2010})} = \\
 & + 0,081.(\textit{número de habitantes com idade entre 80 e 89 anos}) + \\
 & 0,068.(\textit{número de habitantes com idade entre 20 e 24 anos}) - \quad (3) \\
 & 0,021.(\textit{número habitantes do gênero masculino}) + 0,55(\textit{número} \\
 & \quad \textit{de habitantes com idade entre 35 e 39 anos})
 \end{aligned}$$

O estudo dos sinais dos coeficientes que integraram a modelagem explicativa da quantidade de roubos consumados a transeuntes na cidade de Uberlândia, no ano de 2010, representada pela Fórmula 3, permitiu inferir que o número de habitantes, por bairro, pertencentes as faixas etárias “de 80 a 89 anos”, “20 a 24 anos” e “de 35 a 39” anos são direcionadores das ocorrências policiais registradas pela polícia militar daquela cidade. Sendo que, a quantidade de habitantes nessas faixas etárias apresentou o mesmo comportamento (sinal positivo) que a quantidade de ocorrências relativas àquela modalidade criminal, ou seja, nos bairros em que esse número de habitantes é maior

observou-se um maior número de ocorrências. Da mesma forma em que, naqueles bairros em aquelas quantidades são menores, foi observado um menor número de registros de ocorrências relativas a roubos consumados a transeuntes.

Apesar desse estudo não ter utilizado como variáveis explicativas o número de vítimas, mas sim, os dados demográficos do Censo de 2010 do IBGE, ajustados pela prefeitura daquela cidade, a identificação da faixa etária de habitantes com idade entre “20 e 24 anos”, como um dos possíveis determinantes da variável de estudo nessa investigação, permite corroborar parcialmente os achados de Cruz, Arguello e González (2001), que afirmaram que quanto mais jovens os indivíduos, maior probabilidade de serem vítimas deste tipo de violência urbana.

Por outro lado, o estudo dos sinais dos coeficientes das variáveis explicativas detectadas neste estudo também indicou que “número habitantes do gênero masculino” tem comportamento inverso (sinal negativo) ao comportamento da variável de estudo. Isso, por sua vez, pelo menos no que se refere aos “crimes de rua” (roubos consumados a transeuntes) vai contra os achados de Cruz, Arguello e González (2001), que constataram que os homens representam a maioria das vítimas em casos de crimes violentos.

Uma vez que a “quantidade de habitantes do gênero feminino” não foi considerada significativamente relevante, enquanto possível determinante da variável de estudo, este estudo não corrobora com os achados de Eyzaguirre e Puga (2001), que constataram que mulheres na faixa etária entre 18 e 24 anos, entre outros fatores, têm maior probabilidade de serem vítimas de roubo. Por outro lado, ao admitir que essa investigação não estratificou a quantidade de habitantes em cada faixa etária de acordo com seu gênero, poder-se-ia presumir que parte dos achados de Eyzaguirre e Puga (2001) estaria corroborada por este estudo, uma vez que a faixa etária entre “20 e 24 anos” é uma das variáveis explicativas da modalidade criminal investigada nessa pesquisa.

Outra evidência relevante identificada nessa investigação diz respeito ao elevado coeficiente de determinação ( $R^2=91,9\%$ ) e, ainda, ao significativo coeficiente de determinação ajustado ( $R^2$  Ajustado=91,3%) da modelagem explicativa identificada a partir da análise de regressão linear múltipla (Fórmula 3). Ou seja, ao admitir que o  $R^2$  Ajustado (*Adjusted R Square*) fornece uma medida do quanto à modelagem pesquisada pode ser generalizada para a população de dados, ao invés de apenas à respectiva amostra, (FIELD, 2009), pode-se inferir que, se admitida uma margem de erro de 8,70% (100% -  $R^2$  Ajustado), aquela equação explicativa poderia ser usada no processo de identificação dos bairros em que existiria maior probabilidade de ocorrências da mesma natureza que aquela estudada nessa investigação (roubos consumados a transeuntes), se considerado também um intervalo de confiança de 95%.

Adicionalmente, essa última evidência relatada, ou seja, a possível utilização daquela modelagem analítica para fins preditivos, torna-se mais relevante ao se considerar que nesse processo de investigação, aquela modelagem descrita com o auxílio da Fórmula 3 não possui um termo, ou coeficiente, constante. Logo, ao lembrar que o termo constante de uma regressão representa o valor da variável de estudo quando as variáveis explicativas assumem valor igual a zero, ou, ainda, na ausência de variáveis explicativas, o intercepto (ou constante) representa o valor da média esperada para variável de estudo, fica evidente que a combinação linear das variáveis explicativas identificadas nessa pesquisa constitui-se em importantes determinantes das ocorrências relativas ao objeto de estudo dessa investigação.

### **Considerações Finais**

Ao admitir que a violência consolidou-se como um fenômeno social do mundo contemporâneo, esta pesquisa abordou a temática relacionada a uma das modalidades de crimes de natureza violenta mais comum nos dias atuais, ou seja, roubos consumados contra transeuntes.

Nesse sentido, foi problematizado o possível relacionamento entre aquele tipo de criminalidade e algumas das variáveis demográficas relativas ao gênero e às faixas etárias da população de Uberlândia-MG.

Ao final de todo esse processo de investigativo-científico foi evidenciado que três faixas etárias e o gênero masculino caracterizaram-se como possíveis determinantes dos “crimes de rua” analisados nesse estudo.

Cabe observar que o ferramental analítico-quantitativo utilizado nessa pesquisa, além de validar a modelagem matemática explicativa da quantidade ocorrências observada na cidade alvo dessa investigação, permitiu inferir que a equação pesquisada poderia ser usada no processo de identificação dos bairros em que existiria maior probabilidade de ocorrências referentes a roubos consumados a transeuntes.

Ainda foi observado que, ao contrário do que o conhecimento popular demonstra e, ainda, alguns estudos sinalizam a quantidade de habitantes do gênero feminino não foi considerada significativamente relevante para explicar as ocorrências estudadas nessa pesquisa.

Por outro lado, a variável demográfica referente ao “número habitantes do gênero masculino” apresentou comportamento inverso ao comportamento da variável de estudo.

Apesar das contribuições prestadas por essa investigação, destaca-se com sua primeira limitação a utilização de variáveis demográficas de apenas duas naturezas distintas, isto é, gênero e faixa etária. Por outro lado, cabe ressaltar que se tomou o cuidado de se utilizar variáveis que tiveram uma mesma base de referência temporal, ou seja, as ocorrências policiais do ano de 2010 e os dados do Censo realizado nesse mesmo ano.

Ainda com relação às limitações dessa investigação, observa-se que o ideal seria utilizar como variáveis explicativas os dados referentes às vítimas dos crimes aqui estudados. Contudo, além da dificuldade de se obter tais informações, destaca-se que os dados gerais da população permitiram atingir percentuais estatísticos tão significativos que a modelagem analítica identificada nesse estudo poderia ser utilizada como uma função preditiva.

A despeito das limitações destacadas, espera-se que os resultados dessa investigação possam ser somados aos resultados de outros estudos de natureza correlata e, assim, contribuir para o debate e a pesquisa relacionada à criminalidade urbana.

## Referências

- BEATO, C; REIS, I.. Desigualdade, desenvolvimento socioeconômico e crime. In: HENRIQUES, R. (org.). **Desigualdade e pobreza no Brasil**. Rio de Janeiro: IPEA, 2000.
- BEATO FILHO, C. C. *et al.* **Atlas da criminalidade de Minas Gerais**. Belo Horizonte: Centro de Estudos de Criminalidade e Segurança Pública/UFMG, 2002.
- BLANCHARD, O. J.. Comment. *Journal of Business and Economics Statistics, American Statistical Association and American Statistical Association*, Alexandria-VA, v. 5, p. 449-451, 1967.
- BRASIL. Ministério da Justiça. **Código Penal Brasileiro**. Decreto-Lei 2.848, de 07 de dezembro de 1940. Publicado no Diário Oficial no 31/12/1940, p. 2391. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/decreto-lei/del2848.htm](http://www.planalto.gov.br/ccivil_03/decreto-lei/del2848.htm). Acesso em: jan.2012.
- CALDEIRA, T. P. R. Violência, direitos e cidadania: relações paradoxais. **Ciência e Cultura**, Campinas, ano 54, n.1, p. 44-46, jul./set. 2002.
- CAMARGO, A. B. M. **Mortalidade por causas externas no estado de São Paulo e suas regiões**. 2002. Tese (Doutorado) - Faculdade de Saúde Pública da Universidade de São Paulo, São Paulo, 2002.
- CANO, I. Políticas de segurança pública no Brasil: tentativas de modernização e democratização versus a guerra contra o crime. **SUR - Revista Internacional de Direitos Humanos**, São Paulo, ano 3, n. 5, 2006.
- CLARKE, R. V.; FELSON, M. Routine activity and rational choice. In: HENRY, S.; EINSTADTER, W. (Eds.). **The criminology theory reader**. New York: New York University Press, 1998.
- COHEN, L. E.; FELSON, M. Social change and crime rates trends: a routine activity approach. **American Sociological Review**, Washington, DC, v. 44, n. 4, p. 588-608, Aug./1979.

CRUZ, J. M., ARGUELLO, A. T. GONZÁLEZ, F. Factores sociales y económicos asociados al crime violento em El Salvador. In: FAJNZYLBER, P., LEDERMAN, D., LOAYZA, N. (Eds.) **Crimen y violencia en América Latina**. Bogotá: Alfaomega, 2001.

CUNHA, J. V. A. da; COELHO, A. C.. Regressão linear múltipla. In: CORRAR, L. J.; PAULO, E.; DIAS FILHO, J. M. (Coordenadores). **Análise multivariada: para os cursos de administração, ciências contábeis e economia**. São Paulo: Atlas, 2011.

ESTEVES, A.. **A criminalidade urbana e a percepção do espaço na cidade de Lisboa: uma geografia da insegurança**. 1995. Dissertação (Mestrado) - Faculdade de Letras da Universidade de Lisboa, Lisboa, 1995. Disponível em: <https://alpha.sib.uc.pt/?q=content/criminalidade-urbana-e-percep%C3%A7%C3%A3o-do-esp%C3%A7o-na-cidade-de-lisboa-uma-geografia-da-inseguran%C3%A7>. Acesso em: 21/08/2013.

EYZAGUIRRE, H., PUGA, Y. La violencia criminal en Lima Metropolitana: los factores de riesgo y las políticas de prevención y represión. In: FAJNZYLBER, P., LEDERMAN, D., LOAYZA, N. (Eds.) **Crimen y violencia en América Latina**. Bogotá: Alfaomega: 2001.

FACHIN, O. **Fundamentos da metodologia**. 3. ed. São Paulo: Saraiva 2001.

FÁVERO, L. P. *et al.* **Análise de dados: modelagem multivariada para tomada de decisões**. Rio de Janeiro: Elsevier, 2009.

FELIX, S.A. **Geografia do crime: interdisciplinaridade e relevâncias**. Marília: Unesp Publicações, 2002.

FIELD, A.. **Descobrimo a estatística usando SPSS**. 2. ed. Porto Alegre: Artmed, 2009.

GOTTFREDSON, M. R.; HIRSCHI, T. **A general theory of crime**. Stanford: Stanford University Press, 1990.

GUJARATI, D. N. **Econometria básica**. 4. ed. Rio de Janeiro: Campus Elsevier, 2006.

GUJARATI, D. N.; PORTER, D.C. **Econometria básica**. 5. ed. Porto Alegre: AMGH, 2011.

HAIR JÚNIOR., J. F. *et al.* **Multivariate data analysis**. Englewood Cliffs: Prentice Hall, 1998.

IBGE – Instituto Brasileiro de Geografia e Estatística. **IBGE cidades@**. Consulta de dados realizada em 20/09/2013. Disponível em: <<http://cidades.ibge.gov.br>

/xtras/perfil.php?lang=&codmun=317020&search=minas-gerais|uberlandia .  
Acesso em: 20/09/2013.

KENNEDY, P. A. **A guide to econometrics**. 5. ed. Cambridge: MIT Press, 2003.

MINAS GERAIS - Polícia Militar. **Diretriz auxiliar das operações DIAO-1/94-CG**: conceituação e codificação das ocorrências na PMMG. Belo Horizonte: PMMG, 1994.

NETER, J.; WASSERMAN, W. **Applied linear statistical models**. Homewood: Richard D. Irwin, 1974.

PAIXÃO, A. L.. Crimes e criminosos em Belo Horizonte, 1932-1978. In: PINHEIRO, P. S. (Org.). **Crime, violência e poder**. São Paulo: Brasiliense, 1983.

UBERLÂNDIA - Prefeitura Municipal de Uberlândia (Secretaria Municipal de Planejamento Urbano: Diretoria de Pesquisas Integradas). **População por bairros**: 2010. Uberlândia: Prefeitura Municipal de Uberlândia, 2010.  
Disponível em:  
[http://www.uberlandia.mg.gov.br/uploads/cms\\_b\\_arquivos/1460.pdf](http://www.uberlandia.mg.gov.br/uploads/cms_b_arquivos/1460.pdf) . Acesso em: 25/08/2013.

SANZ, P. V.. SPSS17 - **Extracción del conocimiento a partir del analyses de datos**. México (DF): Alfaomega Grupo Editor, 2010.

SAPORI, L. F.. **Segurança pública no Brasil: desafios e perspectivas**. Rio de Janeiro: Editora da FGV, 2007.

STEVENSON, William J. **Estatística aplicada à administração**. São Paulo: Harbra, 1986.

ZALUAR, A. A guerra sem fim em alguns bairros do Rio de Janeiro. **Ciência e Cultura**, Campinas, ano 54, n.1, p.32-38, jul./set. 2002.