

Análise Geográfica do Roubo de Veículos no Município de São Paulo de 2011 a 2017: Um Enfoque utilizando Estatística Espacial

Autoria

Vinicius Attie Buainain Georges - vinicius.georges@gmail.com

Bacharelado em Administração Pública/FGV/EAESP - Fundação Getulio Vargas/Esc de Admin de Empresas de São Paulo

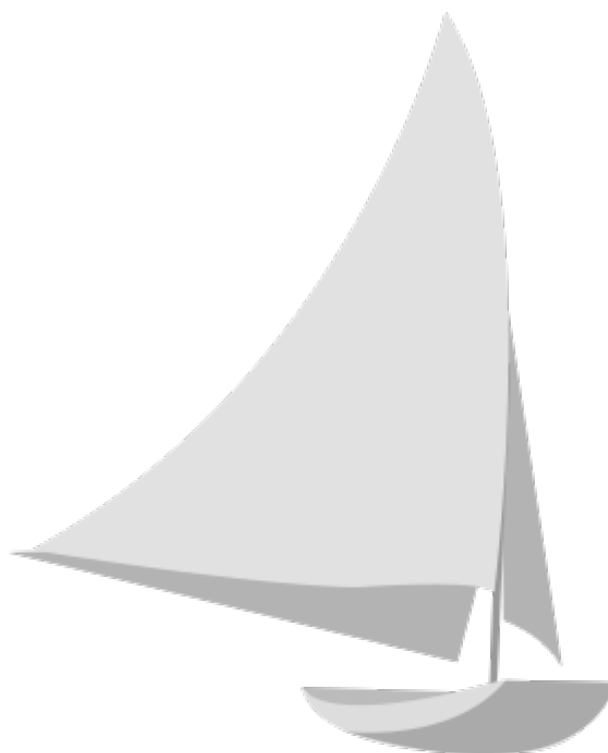
Eduardo de Rezende Francisco - eduardo.francisco@fgv.br

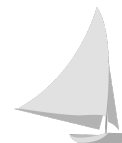
Prog de Mestr Prof em Comportamento do Consumidor - MPCC/ESPM - Esc Sup de Prop e MKT de São Paulo/Ass Esc
Sup de Prop e MKT

Mestr Prof em Gestão e Políticas Públicas - MPGPP/FGV/EAESP - Fundação Getulio Vargas/Esc de Admin de Empresas
de São Paulo

Resumo

Esta pesquisa explora a influência do espaço geográfico em variáveis sócio-econômico-demográficas na dinâmica de roubo de veículos no município de São Paulo de 2011 a 2017. Utilizam-se, como lentes teóricas, a teoria criminológica da Escola de Chicago e o mapeamento criminal, através de Estatística Espacial. 1.050.595 registros de boletins de ocorrência foram extraídos da base da Secretaria de Segurança Pública de São Paulo, filtrados e consistidos, obtendo o total de 276.842 ocorrências válidas de roubos de veículos em São Paulo, agregadas pelas 310 áreas de ponderação do Censo Demográfico 2010. Tem-se como resultado uma prática criminal que tende a se agrupar espacialmente (I de Moran 0,47), apresentando significativa dependência espacial. As regressões da taxa de roubos de veículos aumentaram o R^2 de 21,41% (linear) para 52,70% (espacial). Conclui-se que a variável “espaço” tem grande poder de explicação da dinâmica do roubo de veículos no município de São Paulo.





Análise Geográfica do Roubo de Veículos no Município de São Paulo de 2011 a 2017: Um Enfoque utilizando Estatística Espacial

RESUMO

Esta pesquisa explora a influência do espaço geográfico em variáveis sócio-econômico-demográficas na dinâmica de roubo de veículos no município de São Paulo de 2011 a 2017. Utilizam-se, como lentes teóricas, a teoria criminológica da Escola de Chicago e o mapeamento criminal, através de Estatística Espacial. 1.050.595 registros de boletins de ocorrência foram extraídos da base da Secretaria de Segurança Pública de São Paulo, filtrados e consistidos, obtendo o total de 276.842 ocorrências válidas de roubos de veículos em São Paulo, agregadas pelas 310 áreas de ponderação do Censo Demográfico 2010. Tem-se como resultado uma prática criminal que tende a se agrupar espacialmente (I de Moran 0,47), apresentando significativa dependência espacial. As regressões da taxa de roubos de veículos aumentaram o R^2 de 21,41% (linear) para 52,70% (espacial). Conclui-se que a variável “espaço” tem grande poder de explicação da dinâmica do roubo de veículos no município de São Paulo.

Palavras-chave: Aplicações em Políticas Públicas; Estatística Espacial; Mapeamento Criminal; Roubo de Veículos; Segurança Pública.

1 INTRODUÇÃO

A cidade de São Paulo se construiu de forma espaçada e segregada - classes mais ricas se estabelecendo em localidades geograficamente mais altas (como a região da Av. Paulista) e as classes pobres nas regiões de várzea, perto do rio Tamanduateí ou Anhangabaú. A tipologia arquitetônica foi ao encontro dessa segregação. Empreendimentos com muros altos, grades, seguranças privados e totalmente voltados ao interior e à homogeneização e não à rua e à vida pública - onde o diferente acontece. Caldeira (2000) chama-os de "enclaves fortificados". Essa organização da cidade contribuiu para a escolha na utilização do transporte privado - o veículo.

O presente trabalho visa compreender a situação do roubo de veículo no município de São Paulo - em 2017 foram 32.230 roubos de veículo. A proposta é explorar como esse crime está disposto no espaço da cidade, e verificar se há elementos específicos de cada área que influenciam - ou ao menos, têm relação com - esse crime.

Uma das formas de compreender a segurança pública e a dinâmica criminal é por meio de informação. As estatísticas e análises criminais podem ser um instrumento de *accountability* para segurança pública e administração pública; podem ser uma ferramenta de transparência e controle público, o que fortalece a democracia; e podem servir de insumos para políticas públicas de segurança.

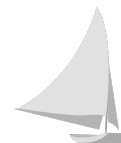
Toma-se como hipótese inicial de trabalho:

Há dependência espacial da distribuição de roubos de veículos em São Paulo?

Assim, assume-se que essa prática criminal não está disposta aleatoriamente no território, e os elementos explorados têm relação com o roubo de veículos no município. Busca-se, outrossim, identificar elementos que permitam mostrar a influência do espaço geográfico nesse fenômeno. Como produto, temos uma das maneiras de realizar uma análise criminal - o mapeamento criminal. O trabalho é uma análise exploratória e espacial do roubo de veículos no município de São Paulo, utilizando-se de técnicas de Estatística Espacial.

2 MAPEAMENTO CRIMINAL E ESTATÍSTICA ESPACIAL

A utilização da perspectiva espacial na análise criminal é principalmente feita por meio do mapeamento das ocorrências. Além de melhorar a visualização dos dados e informações,



oferece uma maior compreensão das dinâmicas espaciais das ocorrências criminais. O mapeamento criminal vem sendo utilizado no auxílio do planejamento operacional das polícias, e melhor direcionamento de recursos para programas e ações. Sua vantagem na utilização do mapeamento criminal é representar de forma simples, visual (mapa) e georreferenciada o que está sendo analisado.

Os mapas coropléticos na segurança pública são utilizados “para médias de área, como taxas de criminalidade, densidade populacional e porcentagens, bem como para informações em escala nominal como uso de área” (Harries, 2000). E por fim, os mapas isolíneos, no campo estudado, representam a densidade da ocorrência de um crime em uma região, encontra-se os famosos *hot-spots*.

Uma vez configurado o mapeamento criminal, é possível trazer diversas ferramentas da estatística espacial, as quais auxiliam na forma de analisar os dados em relação ao seu espaço geográfico. Uma das primeiras explorações na estatística espacial é a ideia de descrever e averiguar os dados georreferenciados, ou *Exploratory Spatial Data Analysis* (ESDA). Francisco (2010) entende que essa ferramenta se dá em quatro passos: a manipulação - envolve o processo de limpeza de dados; a exploração - descrição e visualização de distribuição no espaço, observações extraordinárias; e a confirmação - modelos de estimação e validação a fim de construir análises espaciais multivariadas.

A fim de explorar a distribuição das variáveis no espaço, utiliza-se a autocorrelação espacial, construída a partir de uma matriz de vizinhança (W), $n \times n$, criada pelo pesquisador, para expressar a influência da geografia nas regiões estudadas. A forma mais usual é construí-la por contiguidade, quando estabelece-se o peso 1 para as regiões adjacentes e peso 0 para as não adjacentes. A vizinhança pode se estabelecer por pelo menos um ponto de interseção (tipo *queen*) ou por pelo menos uma aresta (tipo *rook*).

Definida a matriz de vizinhança, é possível calcular a autocorrelação espacial. A forma mais utilizada de autocorrelação espacial é o Índice I de Moran - utilizado para identificar padrões na dispersão, na *clusterização* e na aleatoriedade. Trata-se da correlação da variável com a média da variável em seus vizinhos, conforme vizinhança estabelecida pela matriz W . Assim como o índice de correlação, ele varia entre -1 e +1, sendo 0 o indicativo de aleatoriedade no espaço, o +1 indicativo de *clusterização* e o -1 de dispersão.

A regressão é um instrumento estatístico/econométrico que possibilita explorar a relação entre variáveis independentes (preditoras) e uma variável dependente (y). A equação da regressão linear, *Ordinary Least Squares*, agrega as variáveis independentes (X) multiplicadas por seus parâmetros β (que são estimados pelo instrumento). A diferença entre o valor previsto e o valor real é o erro residual (ε) - conforme ilustrado pela Equação I.

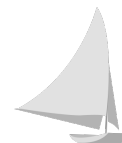
$$y = X\beta + \varepsilon \quad (I)$$

A fim de incluir a influência do espaço das variáveis na explicação da regressão, adiciona-se um termo espacial auto-regressivo (Wy), que captura a influência da geografia na variável dependente. A fórmula da Equação II é conhecida como *Spatial Lagged Auto-Regressive Model*.

$$y = X\beta + \rho Wy + \varepsilon \quad (II)$$

3 CRIMINOLOGIA E O ESPAÇO URBANO

A Escola de Chicago foi de suma importância para a análise criminal, pois rompe com o paradigma de que o crime está atrelado à patologia do homem e leva à análise criminal o estudo de circunstâncias sociais, como entorno e as pessoas que ali habitam. O foco da Escola de Chicago, então, foi estudar o desenvolvimento da grande cidade, principalmente nas cidades norte-americanas, com industrialização, imigração, e conflitos culturais.



Shaw e McKay (1942), da Escola de Chicago descrevem o impacto dos bairros urbanos e suas comunidades na influência do comportamento criminal – a teoria criminológica da Desorganização Social. Essa teoria tem uma relação muito forte com os estudos da Ecologia Humana. Uma das conclusões de Park (1925) é de que a distribuição espacial é definida pelas condições socioeconômicas da população, conforme encontrado em muitas cidades americanas, e está relacionada à delinquência com o espaço. Shaw e McKay (1942) afirmam que características físicas, econômicas, e demográficas estão atreladas à ocupação do espaço e à delinquência. Cabe ressaltar que deve haver um importante cuidado com as correlações entre os indicadores socioeconômicos e a criminalidade – não se pode ser determinista, ou seja, se mora em uma região pobre, com baixos índices socioeconômicos, portanto é criminoso.

Assim como Chicago, São Paulo também passou por um crescimento populacional, urbanização, migração e industrialização ao longo dos séculos XIX e XX. Nesse contexto de industrialização, crescimento e urbanização, São Paulo se expandiu para todos os cantos da cidade, entretanto, seletivamente, a zona oeste abrigava as classes mais abastadas ao passo que a leste foi deixada para as classes populares (Villaça, 2004). E com o passar do tempo, as elites foram se expelindo do centro histórico – ao passo que o mesmo foi perdendo seu caráter de centro financeiro, econômico e de serviços; de modo ao desenvolvimento urbano ter o sentido sudoeste, com novos centros como a Avenida Paulista e a Berrini (Scarlato, 2004). Isto posto, as classes mais ricas passaram a ocupar a zona oeste e a classe mais pobre expelida para as periferias. Como resultado desse fenômeno, observa-se uma cidade segregada, onde encontramos distritos onde moram classes mais abastadas, com melhores indicadores socioeconômicos – geralmente o centro expandido, e distritos onde moram os mais pobres – as regiões periféricas.

4 METODOLOGIA

O foco deste trabalho é explorar o universo de roubo de veículos no município de São Paulo. Optou-se pela abordagem quantitativa, descrita a seguir. Foi levantado esse referencial teórico ligado ao mapeamento criminal e a teorias da prática criminal discutidas na Escola de Chicago, principalmente por se tratar de temáticas utilizadas nas análises de roubo de veículo. Foram analisadas informações dos boletins registrados e suas informações (exceto informações protegidas por lei, como o endereço da vítima e o nome das testemunhas). As informações disponíveis compreendem o número do Boletim de Ocorrência (BO), o modelo (e consequentemente tipo) e ano do veículo roubado e latitude e longitude da localização onde ocorreu a prática criminal, entre outros. Os dados estão disponíveis para todos os 645 municípios de São Paulo, publicados em planilhas mensais; e estão localizados na sessão de transparência da SSP/SP. Utilizou-se o pacote "brccrimR" do *software* R para a carga e consolidação dos dados por meio do *webscrapping*. Registros duplicados e boletins principais e complementares foram posteriormente consolidados para a obtenção da contagem correta de roubo de veículos. Ao final do processo, um *dataframe* com 1.050.595 foi criado – referente ao período de 2011 a 2017. Dados anteriores (de 2003 a 2010, também publicados) não apresentam dados de latitude e longitude das ocorrências e foram, pois, não analisados.

Os dados foram convertidos para formato geográfico (*shapefile*) e preservadas apenas as ocorrências do município de São Paulo (variável "cidade" igual ao valor "S.PAULO"). Por fim, os *softwares* QGIS e ArcView GIS foram utilizados para a filtragem das ocorrências no polígono do município de São Paulo (pela continência através da latitude e longitude) e consolidação de ocorrências de mesma informação temporal e espacial. A base de dados para as análises consolidou, pois, 276.842 observações, de 2011 a 2017.

Em seguida, foi realizado o *spatial join* (junção espacial) das observações por área de ponderação do município de São Paulo, utilizando-se o *software* ArcView GIS. Buscou-se



consolidar a quantidade de ocorrências de roubo de veículo por área de ponderação – variável dependente (y) dos modelos preditivos explorados nesta pesquisa.

Analisou-se as práticas criminais por área de ponderação e não por distritos policiais ou por setor censitário. Anselin (1998) sugere que se analise o fato com a menor área disponível, devido ao *Modifiable Area Unit Problem* (problema da unidade da área modificável). A granularidade de setor censitário (menor unidade territorial disponível nos dados censitários) foi descartada, pois em muitos casos essa unidade refere-se apenas a um edifício (por conter diversos respondentes) e sabe-se também que a maioria do roubo de veículo acontece em via pública, prejudicando assim a análise.

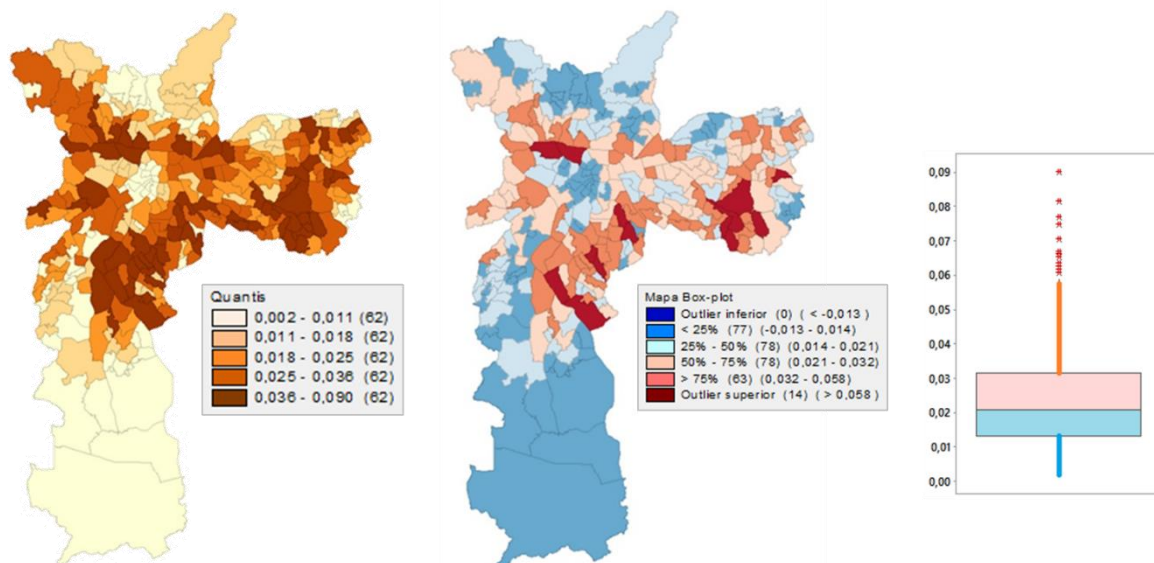
Dados de equipamentos públicos de diversas naturezas (segurança, educação, cultura etc.) foram extraídos da plataforma GeoSampa (2015) e contabilizados por área de ponderação. Ainda, utilizou-se dados do Censo Demográfico 2010, como informações sobre educação e renda média mensal. O Índice Paulista de Vulnerabilidade Social (IPVS) da Fundação SEADE (2010) também foi utilizado – sua granularidade original por setor censitário deu origem à média ponderada (por população) por área de ponderação.

Modelos de regressão linear multivariada da Taxa de roubo de veículos (y) foram aplicados aos dados, utilizando-se o método *stepwise* para a retirada das variáveis não significantes. Após a seleção do melhor modelo linear, o modelo foi promovido a *Spatial Lagged Auto-Regressive Model*, com a incorporação da variável Wy. As ferramentas Minitab 18 e GeoDA 1.12 (2018) foram utilizadas para a execução dos modelos de regressão.

5 ROUBO DE VEÍCULOS EM SÃO PAULO: PORTAL DA TRANSPARÊNCIA

A sessão atual visa analisar os dados da Secretaria de Segurança Pública publicados na seção "transparência" (São Paulo, 2018). A sessão publica os boletins de ocorrência – principais e complementares – e suas informações registradas no distrito policial; desse modo, pode-se verificar informações sobre o período da ocorrência; informações sobre o tipo de veículo roubado (motociclo, automóvel, caminhão etc.) e seu ano de fabricação. As principais informações para a presente seção são as de latitude e longitude; assim, é possível explorar o crime por meio da geolocalização; essas informações só passaram a ser divulgadas a partir de 2011, desse modo, a presente análise concentra seus esforços nesse período.

Figura 1. Mapa de Quantis (esquerda), Mapa *Box-plot* (centro) e *Box-plot* (direita) da Taxa de roubo de veículos por área de ponderação em São Paulo (2011-2017)



Fonte: SSP/SP (2018). Elaboração própria no software GeoDA 1.12 (2018)

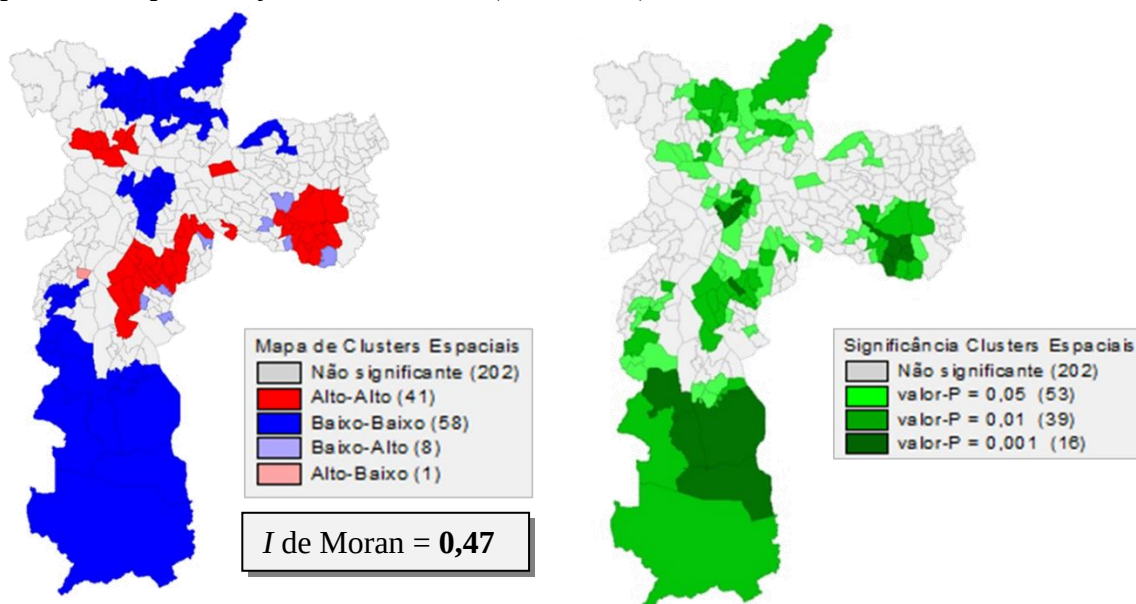
Nota: n = 310 áreas de ponderação



A Figura 1 apresenta um panorama maior das localizações das ocorrências criminais entre 2011 a 2017 - a variável analisada é da taxa de roubo de veículos (a relação entre o total de ocorrências pelo número de pessoas da área). Quanto mais avermelhada a área de ponderação, maior a taxa. É possível verificar um centro expandido apresentando os menores números de ocorrências, além do extremo sul e norte - regiões com abundante natureza. Chama-se atenção o volume de roubo de veículos na região sul e leste da cidade; principalmente a zona limítrofe com os municípios do ABCD (Santo André, São Bernardo do Campo, São Caetano e Diadema) - municípios que apresentam altas taxas de roubo de veículo.

Ainda na análise exploratória espacial dos dados (ESDA), pode-se trazer as áreas de ponderação que excedem a distribuição padronizada dos roubos de veículo no município e se encontram como "pontos fora da curva" - ou *outliers*. É possível visualizar, na Figura 1, o *boxplot* das taxas de roubo de veículos com apenas "valores discrepantes" acima do limite superior. O maior valor encontrado - de aproximadamente 100 ocorrências para cada 1000 habitantes - é da área de ponderação de São Rafael. Os polígonos com cores mais azuladas representam áreas com menor taxa de roubo de veículos, ao passo que os que apresentam cores avermelhadas possuem uma maior taxa de roubo de veículos. As 14 áreas de ponderação vermelhas são os *outliers* superiores.

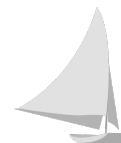
Figura 2. Mapa de Auto-Correlação Espacial Local (*LISA map*) da Taxa de roubo de veículos por área de ponderação em São Paulo (2011-2017)



Fonte: SSP/SP (2018). Elaboração própria no software GeoDA 1.12 (2018)

Nota: $n = 310$ áreas de ponderação. Vizinhança do tipo *queen* de primeira ordem. I de Moran significativa para 5%.

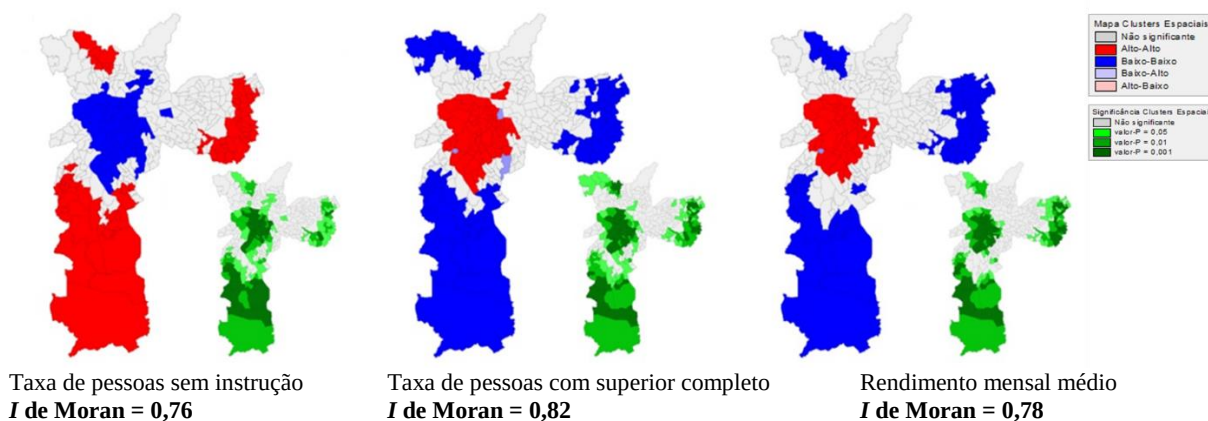
Avaliou-se, de forma exploratória, a influência da geografia na distribuição da taxa de roubo de veículos por área de ponderação no município de São Paulo. O índice de auto-correlação espacial I de Moran obtido, para uma matriz de vizinhança por contiguidade de primeira ordem e do tipo *queen*, foi 0,47, significativo para valor-P < 5%. A Figura 2 apresenta o I de Moran local para cada área de ponderação, destacando os *clusters* espaciais e suas respectivas significâncias. A cor azul mostra distritos que têm baixo número de roubos de veículos e seus distritos vizinhos também; a cor vermelha mostra distritos que possuem alto número de roubos de veículos e seus vizinhos também; o azul e o vermelho mais claros (mais raros) mostram a divergência entre os distritos e seus vizinhos. Os diversos grupos azuis e



vermelhos espalhados pelo município corroboram o padrão de dependência espacial significativa encontrado.

As variáveis independentes apuradas também foram analisadas quanto a seus padrões de dependência espacial. A Figura 3 destaca as variáveis Taxa de pessoas sem instrução (I de Moran = 0,76), Taxa de pessoas com superior completo (I de Moran = 0,82) e Rendimento mensal médio (I de Moran = 0,78), todas com elevada auto-correlação espacial significativa (valor- P abaixo de 5%).

Figura 3. Mapa de Auto-Correlação Espacial Local (*LISA map*) de algumas variáveis independentes por área de ponderação em São Paulo (2011-2017)



Fonte: SSP/SP (2018). Elaboração própria no software GeoDA 1.12 (2018)

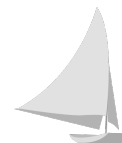
Nota: $n = 310$ áreas de ponderação. Vizinhança do tipo *queen* de primeira ordem. I de Moran significativa para 5% para as três variáveis.

Realizou-se, então, a regressão linear OLS da taxa de roubos de veículos com apenas variáveis independentes que são equipamentos – em sua grande maioria, públicos (retirando dessa categoria apenas clubes e instituições de ensino privados). Obteve-se um modelo final de regressão com 5 variáveis: (i) Centros culturais, (ii) Clubes esportivos, (iii) Escolas de ensino fundamental e médio, (iv) Árvores e (v) Semáforos, e com R^2 de 15,73%. O modelo apresenta relação positiva entre semáforos, árvores e escolas públicas do ensino fundamental e médio com a variável dependente. Já a relação entre centros culturais e clubes esportivos com a taxa de roubo de veículos é negativa – ou seja, quanto maior a quantidade desses equipamentos, menor a incidência do roubo de veículos.

Realizando a *Spatial Auto-Regressive model* com as variáveis significantes do modelo anterior o R^2 aumentou para 48,63% (as variáveis clubes esportivos e árvores deixaram de ser significantes).

Modelos lineares considerando apenas indicadores sócio-econômico-demográficos foram explorados. Um modelo obtido apresentou R^2 de 12,44% e possui 3 variáveis significantes: (i) densidade demográfica, (ii) proporção de pessoas sem instrução e (iii) rendimento mensal médio. Todas elas possuem uma relação inversa com a variável dependente e uma certa multicolinearidade, tipicamente inerente aos fenômenos dessa natureza em São Paulo. O modelo promovido a espacial (incorporação da variável W_y) apresentou R^2 de 51,04%.

Por fim, realizou-se um modelo de regressão contendo todas as variáveis coletadas – tanto de equipamentos como de indicadores socioeconômicos. Obteve-se R^2 de 21,41% e apenas duas variáveis com uma relação positiva com a variável dependente - o número de semáforos e de árvores na área; são 7 variáveis significantes das 23 coletadas. Sendo elas: (i) Densidade habitacional, (ii) Número de Semáforos, (iii) Número de árvores, (iv) Clubes esportivos, (v) Centros educacionais, (vi) Taxa de pessoas sem instrução e (vii) Rendimento mensal médio.



Realizou-se, por fim, a regressão espacial para todas as variáveis, a partir do modelo alcançado. Algumas variáveis - (clube, centros culturais e árvores) - não significantes (para o modelo espacial) foram retiradas; possuindo, assim, uma regressão com 4 variáveis, além da constante e da matriz de vizinhança. O R^2 aumentou para 52,70%, indicando que a variável “espaço” – calculada a partir da matriz de vizinhança (explicada na seção de estatística espacial) – mantém-se muito importante na explicação da taxa de roubo de veículos no município.

6 RESULTADOS E CONSIDERAÇÕES FINAIS

Deve-se levar em consideração, *a priori*, na leitura das informações deste estudo, a existência da “cifra oculta”, ou subnotificação de ocorrências. Caso a ocorrência não seja registrada, o crime não é divulgado nas estatísticas criminais, gerando a chamada “cifra oculta”, a qual distorce as análises criminais do local. A possibilidade de registro digital de roubos, segundo Diógenes Lucca (Estado de São Paulo, 2015), aumentou o número de ocorrências registradas de Roubo de Veículo; em São Paulo, o registro eletrônico de roubos só passou a ser possível a partir de 2013.

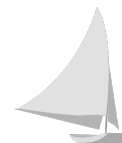
Por fim, é interessante apontar alguns elementos que o trabalho não levou em consideração para construção do modelo, o que possibilita o aprimoramento da abordagem. Foi deixado de lado na construção empírica o impacto do crime organizado no roubo de veículos, a existência de desmanches também. Outro ponto que poderia ter sido incorporado no modelo é a localização dos veículos encontrados – para tanto, seria necessário um entendimento da dinâmica do fluxo do local onde foi roubado e do local onde foi encontrado. Destacando isso, o modelo conseguiu atingir um resultado satisfatório, explicando pouco mais de 50% do fenômeno estudado a partir dos dados disponíveis pela Secretaria de Segurança Pública de São Paulo.

O roubo de veículos é um problema encontrado no município de São Paulo. A presente seção visa compilar o que foi encontrado nas seções de desenvolvimento do trabalho – tanto das estatísticas oficiais como do enfoque aprofundado e georreferenciado – e analisar esses resultados a partir do que foi levantado na seção bibliográfica.

Tangerino (2006) traz esses conceitos da Escola de Chicago para São Paulo. A cidade se expandiu seletivamente – o centro expandido abriga as classes mais abastadas ao passo que a zona leste e a periferia abrigam as classes mais pobres. Com a expansão da elite do centro histórico para o eixo sudoeste da cidade, as pessoas que ali viviam foram “expulsas” e partiram para periferias. O que implica também na condição dos indicadores socioeconômicos da região. A Teoria da Desorganização Social discute que regiões sem organização social são boas para propagação da “cultura delinquente” – indicadores socioeconômicos estariam associados com a ocupação do espaço e a capacidade de organização social da área.

Alguns indicadores socioeconômicos tendem à segregação e à clusterização, principalmente as variáveis de educação (taxa de pessoas com superior completo e sem nenhuma instrução) e variáveis de rendimento (rendimento mensal médio) isso é verificado pela magnitude e significância da autocorrelação espacial pelo Índice I de Moran próximo a 0,8; o centro expandido e o sudoeste da cidade possuem taxas altíssimas de pessoas com superior completo (o mesmo para rendimento mensal médio) ao passo que áreas periféricas possuem altas taxas de pessoas sem nenhuma instrução educacional.

O mesmo procedimento analítico foi realizado para a variável dependente - a taxa de roubo de veículo. O Índice I de Moran é moderado - 0,47; o que mostra leve tendência de clusterização, ou pelo menos pode-se afirmar que a distribuição da prática criminal pelo espaço não é aleatória. Há uma maior incidência do roubo de veículos na região sul, leste e sudeste da cidade - essa última região é limítrofe do município e próxima dos municípios do ABCD paulista. Entretanto, regiões adjacentes do centro expandido também sofrem desse problema.



Então, é necessário compreender quais variáveis conseguem melhor explicar a taxa de roubo de veículos visto que a nossa variável dependente não é aleatória em sua distribuição no espaço. Isso foi feito por meio das regressões – tanto linear (*Ordinary Least Squares*) como espacial (*Spatial Auto-Regressive model*).

Três diferentes regressões foram realizadas: (i) uma contendo apenas variáveis de equipamentos públicos e privados; (ii) contendo apenas indicadores socioeconômicos e (iii) contendo todas as variáveis coletadas. Para essas regressões, foi rodada tanto a *Ordinary Least Squares* como a *Spatial Auto-Regressive model*; deixando nos modelos finais apenas variáveis que possuem um nível de significância inferior a 5%.

Em primeiro momento, chama atenção a diferença entre o coeficiente de determinação - R^2 - calculado a partir da regressão linear e do calculado a partir da regressão espacial. Todos os três modelos tiveram um salto significativo no R^2 ; o que mostra uma grande importância da vizinhança (W_y) no modelo regressivo, ou seja, a variável “espaço” importa de forma bastante significativa para explicação da taxa de roubo de veículos no município de São Paulo.

Retomando a hipótese desse estudo (“Há dependência espacial da distribuição de roubos de veículos em São Paulo?”), pode-se rejeitar a ideia de que não é necessária uma variável “espaço” para o entendimento dessa prática criminal – o espaço interfere de forma significativa no fenômeno roubo de veículos.

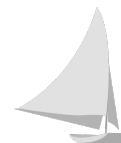
Outra questão que chama atenção é a ausência dos equipamentos de segurança pública (Polícia Militar e Polícia Civil) e da segurança urbana (Guarda Civil Metropolitana) nos modelos finais de regressão, o que mostra uma inexistência de significância (um valor-P maior que 0,05) dessas variáveis independentes.

Deixando de lado o tamanho dos coeficientes (estimadores β), pode-se ressaltar seus sinais; ou seja, qual é a relação entre as variáveis independentes significantes e a taxa de roubo de veículo. Apenas a quantidade de semáforos e o termo espacial auto-regressivo possuem uma relação positiva com a variável dependente. Já as outras variáveis, como densidade demográfica, rendimento e proporção de pessoas sem instrução, possuem uma relação negativa. Chama a atenção, em específico, a relação entre a taxa de pessoas sem instrução educacional e a taxa de roubo de veículos; retoma-se, então, que as áreas de ponderação que possuem maiores frações de pessoas sem formação estão localizadas no extremo leste, extremo sul e no extremo norte, esses dois últimos citados possuem baixa quantidade de roubo de veículos. A equação abaixo consolida a última regressão espacial apresentada da taxa de roubos de veículos, que testou todas as variáveis explicativas utilizadas.

$$\hat{y} = 0,02833 + 0,6897 W_y - 2,92001 \cdot 10^{-6} \text{ Rendimento} \\ - 0,028 \text{ TaxaInstrução} + 0,00013 \text{ Semaforos} - 5,31361 \cdot 10^{-7} \text{ DensHab} \quad (\text{III})$$

Em suma, o presente trabalho traz uma análise estatística espacial do roubo de veículos no município de São Paulo e permite concluir, a partir das regressões lineares e espaciais, que a perspectiva espacial é bastante relevante na explicação dessa prática criminal. O coeficiente de determinação - R^2 - aumenta de forma significativa quando passamos de uma regressão linear, na qual não incluímos a variável espaço na análise, para uma regressão espacial. A estatística espacial, então, esteve presente por toda análise dessa prática criminal.

Nos três modelos de explicação da taxa de roubos de veículos, contendo diferentes tipos de variáveis explicativas. Um modelo utilizou apenas equipamentos públicos e privados, em que passamos de um R^2 de 15,73% (linear) para um R^2 de 48,63% (espacial). A outra regressão envolvendo indicadores socioeconômicos, na qual o R^2 passa de 12,44% (linear) para 51,04% (espacial). E o modelo final com todas as variáveis coletadas, que passa de 21,41% (linear) para 52,70% (espacial). Ressalta-se, novamente, a importância de incluir o espaço no desenvolvimento de análises criminais sobre o roubo de veículos no município de São Paulo.

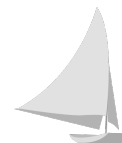


Vale destacar a possibilidade de uso variado do presente trabalho. Ele pode servir como auxílio para formuladores de políticas públicas, ou para áreas de planejamento de instituições da segurança pública e para empresas de seguro de carro, pois traz informações compiladas e analisadas sobre a dinâmica do roubo de veículos em São Paulo.

Outro achado desta pesquisa é sua preocupação e operação da metodologia de coleta, filtragem, limpeza, consistência e consolidação dos dados por área de ponderação. E, por fim, ele também pode ser examinado sob as lentes criminológicas, isso porque relaciona uma grande teoria da criminologia – Escola de Chicago - com dados divulgados pela Secretaria de Segurança Pública de São Paulo.

Desse modo, a incorporação da variável espaço no entendimento do roubo de veículo no município é importante e traz de relevante uma das principais teorias criminológicas que envolve o crime e o espaço, que é a Escola de Chicago, com sua teoria da desorganização social. A ocupação do espaço está relacionada à disputa de dominação e sucessão entre os humanos; e retomamos que no desenvolvimento urbano do município, as classes mais abastadas se estabeleceram no eixo sudoeste da cidade, ao passo que as mais pobres foram se organizando nas periferias e na zona leste. As maiores taxas de roubo de veículo são justamente encontradas nas zonas sul, leste e sudeste da cidade, apesar desse problema estar espalhado pelo município com outras áreas de ponderação sofrendo roubos de veículo, conforme Figura 1.

O trabalho pode ser aprimorado e aprofundado de diversas maneiras. É possível, por exemplo, fazer o mesmo processo de *webscrapping* para coleta de dados georreferenciados de outras práticas criminais, como homicídio doloso, latrocínio ou estupro. Com a metodologia já consolidada, pode-se explorar diversas relações tanto entre os equipamentos e indicadores como entre práticas criminais: Há mais roubo de veículos quando há mais homicídio? Futuros estudos poderão endereçar questões como esta, à luz da Estatística Espacial.



REFERÊNCIAS BIBLIOGRÁFICAS

- ANSELIN, L.. (1998), "Exploratory Spatial Data Analysis in a Geocomputational ArcView GIS, version 3.2: Desktop GIS Software. Redlands, CA: ESRI Inc., 2002.
- CORRÊA, F. BRCRIMR (2017). Disponível em: <<http://curso-r.com/blog/2017/09/13/2017-09-13-brccrimr/>>. Acesso em: 03 mar. 2018.
- FRANCISCO, E. R. (2010), Indicadores de renda baseados em consumo de energia elétrica: Abordagens domiciliar e regional na perspectiva da estatística espacial. 2010. 382 f. Tese (Doutorado em Administração de Empresas), Fundação Getúlio Vargas, São Paulo.
- HARRIES, K. (2000), Mapping Crime: principle and practice. Crime Mapping Research Center: National Institute of Justice. Disponível em: <<http://www.crisp.ufmg.br/livro.html>>. Acesso em: 04 abr. 2018.
- GeoDA 1.12 (2018) Software for Exploratory Spatial Data Analysis by Luc Anselin. Urbana-Champaign, IL: Spatial Analysis Laboratory, University of Illinois, 2018. Disponível em: <<http://geodacenter.asu.edu/software/downloads>>. Acesso em: 26 fev. 2018.
- IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. (2010), Censo 2010. Rio de Janeiro. Disponível em: <<http://www.censo2010.ibge.gov.br/>>. Acesso em: 25 maio 2018.
- JESUS FILHO, J. SSP SP Scraper. Disponível em: <<https://gist.github.com/jjesusfilho/3b0f99f80a0bce019bda3fce54486b9d>>. Acesso em: 03 mar. 2018.
- OPEN SOURCE GEOSPATIAL FOUNDATION. (2018), QGIS 3.0. Disponível em: <https://www.qgis.org/pt_BR/site/index.html>. Acesso em: 02 abr. 2018.
- PARK, R. E. (1925), Ecologia Humana in PIERSON, Donald. Estudos de Ecologia Humana (org.) São Paulo: Martins, 1925, pp. 47 e 48.
- R Studio 1.1. (2018). Disponível em: <<https://www.rstudio.com/products/rstudio/>>. Acesso em: 29 jan. 2018.
- SÃO PAULO. PREFEITURA DE SÃO PAULO. (2015), Geosampa. Disponível em: <http://geosampa.prefeitura.sp.gov.br/PaginasPublicas/_SBC.aspx>. Acesso em: 02 abr. 2018.
- SÃO PAULO. SECRETARIA DE SEGURANÇA PÚBLICA. (2018), Portal da Transparência. Disponível em: <<http://www.ssp.sp.gov.br/transparenciassp/Default.aspx>>. Acesso em: 08 maio 2018.
- SÃO PAULO. SECRETARIA DE SEGURANÇA PÚBLICA. (2014), Mapa de crimes. Disponível em: <<http://www.ssp.sp.gov.br/acoes/leAcoes.aspx?id=33833>>. Acesso em: 08 maio 2018.
- SÃO PAULO. FUNDAÇÃO SEADE. (2010), Índice Paulista de Vulnerabilidade Social. São Paulo. Disponível em: <http://indices-ilp.al.sp.gov.br/view/pdf/ipvs/principais_resultados.pdf>. Acesso em: 12 abr. 2018.
- SCARLATO, F. C. (2004), Busca do centro – o reencontro com a cidade. São Paulo, Contexto.
- SHAW, C., e MCKAY, H. (1942), Juvenile Delinquency in Urban Areas. Chicago, IL: University of Chicago Press.
- TANGERINO, D. P. C. (2007), Crime e Cidade: Violência urbana e a Escola da Chicago. São Paulo: Lumen Juris.
- VILLAÇA, F. (2004), Espaço intra-urbano no Brasil. São Paulo: Nobel, 2004, p.117

ⁱ Pacote criado por Fernando Corrêa, ex-pesquisador do Núcleo de Estudos de Violência da USP (NEV/USP), em que foi agregada uma função de raspagem de dados feita pelo programador José de Jesus Filho.