

Análise da violência doméstica no município de Curitiba

Manoel Flavio Leal¹, Rita Berardi¹, Nádia P. Kozevitch¹

¹Universidade Tecnológica Federal do Paraná (UTFPR) - Curitiba, PR – Brasil
manoelleal@alunos.utfpr.edu.br, {ritaberardi, nadiap}@utfpr.edu.br

Abstract. *Data from recent years shows that there has been an increase in crime indicators, especially domestic. The aim of this study is to analyze historical data related to domestic violence against females, in the Brazilian city of Curitiba between the years 2014 and 2018. To improve the study, demographic data about the city were included. Through analysis and data visualization, several insights have been identified, such as the concentration of occurrences after 6pm on any day of the week, an increase of 50.63% on weekends, identification of four types of crime that represent 82.63% of occurrences, and the region and neighborhoods where most victims are concentrated.*

Resumo. *Ao se observar os últimos anos, nota-se o incremento dos indicadores de criminalidade, com destaque para violência doméstica. O objetivo deste estudo é analisar os dados históricos sobre violência doméstica em Curitiba no período entre 2014 e 2018, com vítimas do gênero feminino. Para enriquecer o estudo, foram incluídos dados demográficos sobre o município. Por meio da análise e visualização dos dados, foram identificados diversos insights, como concentração de ocorrências após as 18:00 em todos os dias da semana, aumento de 50,63% nos fins de semana e quatro naturezas criminais que representam 82,63% das ocorrências, além da identificação da região e bairros onde se concentra o maior número vítimas.*

1. Introdução

A violência não é um problema apenas das grandes metrópoles. Todas as cidades sofrem com o aumento dos índices de criminalidade. Como mostra [Cerqueira et al. 2019], o indicador de violência: taxa de homicídios, entre os anos de 2007 e 2017, aponta uma variabilidade entre as regiões brasileiras. Nos anos de 2016 e 2017, nota-se uma redução dos crimes nas regiões Sudeste e Centro-Oeste, certa estabilidade na região Sul e crescimento acentuado no Norte e no Nordeste.

Entre os vários tipos de crimes e situações de violência, podemos destacar a violência doméstica. Esta, por definição, é tida como toda agressão ocorrida no contexto familiar, ou seja, o agressor é um membro da família. Esta agressão pode ser categorizada como física, sexual, psicológica, moral e patrimonial, sendo que na maioria dos casos as vítimas são mulheres [Gregoli et al. 2018]. De acordo com o levantamento da ONU¹ - Organização das Nações Unidas, existe uma pandemia global de violência contra mulher. Atualmente, 35% das mulheres em todo mundo já sofreram com algum

¹ <https://nacoesunidas.org/violencia-contra-as-mulheres-e-pandemia-global-diz-chefe-da-onu/>

tipo de agressão física ou sexual. Isto demonstra a impunidade e como os países estão distantes do objetivo de acabar com a violência contra a mulher.

De acordo com informações do IBGE² - Instituto Brasileiro de Geografia e Estatística, o município de Curitiba, capital do Paraná, tem uma população estimada de 1.917.185 habitantes. Desses 52,33% são do gênero feminino. O salário médio dos trabalhadores formais é de 4 salários mínimos e IDHM - Índice de Desenvolvimento Humano Municipal, de 0,823, um dos maiores entre os municípios brasileiros. Além disso, a capital paranaense ocupa o vigésimo quarto lugar no ranking de violência entre as 27 capitais do país [Cerqueira et al. 2019].

Os dados sobre criminalidade e violência em Curitiba podem ser obtidos por meio dos registros já cadastrados pelas entidades policiais nos boletins de ocorrência. Estes dados são também disponibilizados para consulta pública, apenas em relatórios consolidados, no site da SESP - Secretaria da Segurança Pública e Administração Penitenciária³. Existe o Portal de Dados Abertos⁴ da Prefeitura de Curitiba, onde é possível consultar e fazer o download dos dados. Porém, as ocorrências registradas pela Guarda Municipal que estão disponibilizadas não possuem informações sobre violência doméstica ou da vítima.

O termo ciência de dados abrange elementos variados que contemplam desde a inteligência artificial, análise e processamento de dados, bem como a visualização e não podendo esquecer de técnicas de estatística e fundamentos da ciência da computação [Silva 2018]. Sendo assim, o objetivo deste trabalho é realizar uma análise nos dados históricos dos boletins de ocorrências sobre violência doméstica no município de Curitiba. Para isso serão utilizadas técnicas de análise e visualização de dados.

O restante do trabalho está organizado da seguinte maneira: a Seção 2 apresenta trabalhos relacionados, seguido pela descrição dos dados na Seção 3. Posteriormente, são apresentadas a análise e a interpretação na Seção 4, seguida pela conclusão, na Seção 5.

2. Trabalhos Relacionados

O acesso aos dados sobre violência contra as mulheres no Brasil, bem como a sua forma de coleta e compartilhamento, ainda está longe do ideal. De acordo com os estudos de [Gregoli et al. 2018], ainda não foi possível implementar o art. 38 da Lei nº 11.340, de agosto de 2006 (Lei Maria da Penha), sobre dados e estatísticas destas ocorrências. Além disso, o estudo deixa clara a dificuldade de se obterem dados estaduais referentes aos registros policiais sobre violência doméstica de forma sistemática e comparável nacionalmente. Porém, cita que, ao obter-se acesso aos dados do Disque 180, foi possível conhecer o percentual de vítimas de violência psicológica, moral, física, sexual e patrimonial. Além deste, alguns números podem ser destacados no estudo de violência contra a mulher no Brasil realizado por [Engel 2019] e publicado pelo Ipea⁵ - Instituto

² <https://cidades.ibge.gov.br/brasil/pr/curitiba/panorama>

³ <http://www.seguranca.pr.gov.br/CAPE/Estatisticas>

⁴ <https://www.curitiba.pr.gov.br/dadosabertos/busca/>

⁵ <http://www.ipea.gov.br>

de Pesquisa Econômica Aplicada. Quando se trata do local da agressão, 49,58% das ocorrências aconteceram na casa da vítima. Ao se analisar a faixa etária, destaca-se que mulheres adolescentes, jovens e adultas são as mais afetadas por agressão física.

Muitos órgãos governamentais nos últimos anos têm disponibilizado em seus sites dados para que o cidadão possa consultar e até mesmo produzir suas próprias análises. Em [Eberhardt and Silveira 2018], esta afirmação é confirmada e destaca a tendência em disponibilizar ferramentas de visualização no próprio site. Entre as informações divulgadas, estão tráfego de veículos, infraestrutura, finanças, demografia, saúde e, em último lugar, criminalidade. Já as formas de visualização mais utilizadas são mapas, *bubble charts* e *treemaps*.

Por outro lado, existe pouca divulgação da existência de dados abertos sobre criminalidade ou violência no Brasil. O estudo realizado por [Montovani et al. 2019], sobre portais de segurança pública dos estados brasileiros, relata-se que, mesmo seguindo os princípios de dados governamentais abertos, na sua maioria os dados são incompletos e confusos, têm curtos intervalos de tempo (poucos meses) e relatórios fixos (sem interatividade), além da divulgação de resultados de ações pontuais.

Outro estudo utilizando dados governamentais e técnicas de ciência de dados é demonstrado em [Flores Junior et al. 2017], no qual é analisada a base de acidentes de trânsito disponibilizada pelo governo do Rio Grande do Sul. O uso de mapas e dados abertos de Curitiba é apresentado em [Bichibichi 2018].

O uso de dados para identificar características sobre criminalidade é apresentado no estudo de [Oliveira et al. 2018], no qual, por meio da análise espaço-tempo, utilizando dados de 12 cidades do EUA, como Nova York, Chicago e Seattle e por meio de séries temporais e *wavelets*, comprova a existência de sazonalidade no comportamento furto e roubo durante o ano na região analisada. Porém, ao se dividir em subregiões, é possível identificar a movimentação das “ondas de crime”, ou seja, os criminosos mudam de região ao longo do tempo.

3. Descrição dos dados

A base de dados utilizada neste artigo é composta de dados oriundos de ocorrências de violência doméstica (não são considerados casos de feminicídio), extraídos do boletim de ocorrência unificado (BOU) disponibilizado pelo CAPE⁶ - Centro de Análise, Planejamento e Estatística e registrados pela Polícia Militar ou Polícia Civil do estado do Paraná, entre os anos de 2014 e 2018, contendo 62.267 registros e 8 atributos. Cada registro da base corresponde a uma ocorrência contendo dados do fato e da vítima, como apresentado na Tabela 1.

⁶<http://www.seguranca.pr.gov.br/CAPE>

Tabela 1. Dados coletados no BOU

Refrente	Dado	Exemplo
Fato	Data da ocorrência	06/01/2017 23:00:00
	Natureza	LESAO CORPORAL
	Bairro	TATUQUARA
Vítima	Idade	24
	Gênero	FEMININO
	Grau de Instrução	SEGUNDO GRAU INCOMPLETO
	Estado Civil	SOLTEIRA
	Cor da pele	BRANCA

O atributo **Data da ocorrência** possui o formato dd/mm/aaaa hh:mm:ss. Todos os 62.267 registros possuem a data preenchida de forma correta. Não foi identificada a existência de valores nulos, brancos ou datas inválidas. O domínio do campo compreende o intervalo entre 01/01/2014 00:00:00 e 31/12/2018 23:00:00. A Figura 1 mostra a distribuição das ocorrências agrupadas por ano. Note-se que no ano de 2017 houve um crescimento de 30,71% em relação ao ano de 2016, depois seguido de uma queda de -10,62% em relação ao ano de 2018.

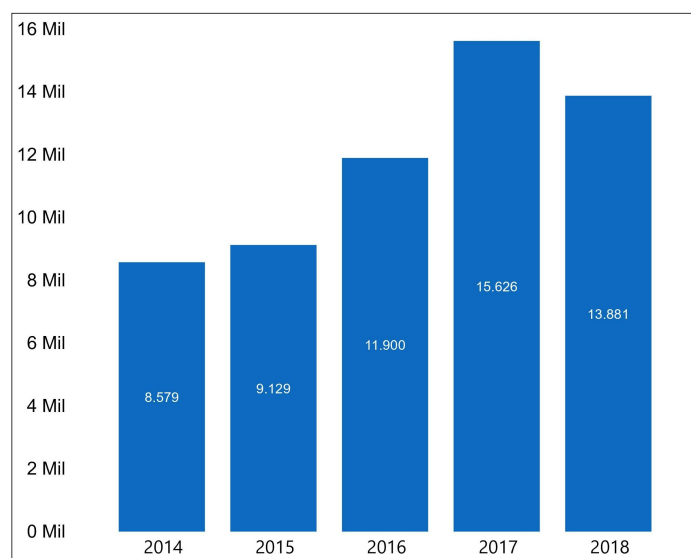


Figura 1. Quantidade de ocorrências registradas por ano.

O atributo **Natureza** tem valores alfanuméricos que contêm um domínio com a descrição do tipo penal do crime. Foram identificados 171 valores diferentes, sendo que Ameaça, Injúria, Lesão Corporal e Vias de Fato representam 82,63% dos registros. A Figura 2 apresenta a distribuição das naturezas.

O atributo **Bairro** apresenta valores no formato alfanumérico. Foram identificados registros em branco e 86 bairros diferentes. Porém, o município de Curitiba possui apenas 75 bairros oficializados⁷. Ao analisá-los, foram identificados

⁷ http://curitibaemdados.ippuc.org.br/Curitiba_em_dados_Pesquisa.htm

erros de digitação e a utilização de apelidos no lugar do nome oficial, como CIC no lugar de Cidade Industrial.

O atributo **Idade** é calculado utilizando a data de nascimento da vítima no momento do preenchimento do boletim. São 62.240 (99,97%) registros com idade entre 0 e 97 anos, 11 (0,02%) com valor em branco e 5 (0,01%) com idade maior que 180 anos. Estas duas últimas ocorrências foram consideradas como erros.

O atributo **Gênero** tem valores no formato alfanumérico, sendo 59.115 (94,94%) dos registros com o gênero feminino, 3.141 (5,04%) com o masculino e 11 (0,02%) com valor não informado.

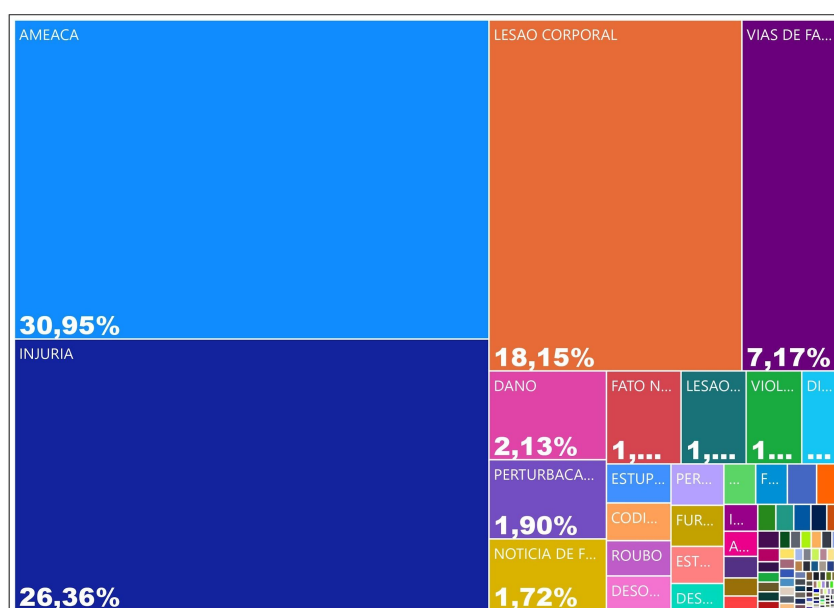


Figura 2. Distribuição das naturezas das ocorrências.

O atributo **Grau de instrução** apresenta valores no formato alfanumérico, sendo 7.418 (11,91%) dos registros com valores em branco, 6.529 (10,49%) não informados, 680 (1,09%) como não alfabetizado, 4.913 (7,89%) com primeiro grau completo, 6.130 (9,84%) com segundo grau completo, 4.626 (7,43%) com terceiro grau completo, 9.089 (15,6%) com primeiro grau incompleto, 14.748 (23,69%) com segundo grau incompleto e 8.134 (13,06%) com terceiro grau incompleto.

O atributo **Estado civil** tem valores no formato alfanumérico, sendo 7.105 (11,41%) dos registros com valores em branco, 3.444 (5,53%) não informados, 416 (0,67%) outros, 20.814 (33,43%) solteiro, 11.616 (18,66%) casado, 9.954 (15,99%) convivente, 4.511 (7,24%) divorciado, 2.746 (4,41%) separado e 1.661 (2,67%) viúvo.

O atributo **Cor da pele** tem valores no formato alfanumérico. Sendo 14.707 (23,62%) dos registros com valores não informados, 38.853 (57,58%) branca, 9.502 (15,26%) parda, 1.862 (2,99%) preta e 342 (0,55%) amarela.

Para este estudo, foram extraídos do site do IPPUC⁸ - Instituto de Pesquisa e Planejamento Urbano de Curitiba, os dados **área em km², tamanho da população, renda média mensal e número de residências próprias**, referentes ao censo de 2010 para cada um dos 75 bairros de Curitiba. No site foi possível realizar o download do arquivo com os dados no formato CSV (*Comma Separated Values*), conforme amostra apresentada na Tabela 2.

Tabela 2. Amostra de Atributos coletados do IPPUC.

bairro	area	populacao	renda	perc_propria
ABRANCHES	4,32	11165	1009,67	0,2824
AGUA VERDE	4,76	49866	3332,67	0,3418

4. Análise e interpretação dos dados

Para realizar a análise e interpretação dos dados coletados neste artigo, foram utilizadas as ferramentas de visualização e análise de dados *Power BI Desktop*⁹ Versão 2.74.5619.621 64-bit (outubro de 2019) e *Python 3.7* distribuição Anaconda 2019.10¹⁰.

Antes de se iniciar a análise dos dados, foram filtradas apenas as ocorrências com vítimas do gênero feminino, o que resultou em 94,94% dos registros, conforme apresentado na Seção 3 deste artigo. Para obter uma visão completa, os dados coletados foram analisados em 4 etapas distintas, tendo cada uma delas o objetivo de explorar uma dimensão deste conjunto de dados. O primeiro recorte foi realizado com dados referentes à ocorrência; o segundo, às vítimas; o terceiro sobre as naturezas, e o quarto e último trata-se do cruzamento destes recortes.

No primeiro recorte, com os atributos **Data da ocorrência** e **Natureza**, foi possível analisar a distribuição das ocorrências entre janeiro de 2014 e dezembro de 2018. A Figura 3 permite identificar que as ocorrências no tempo se comportam de forma não estacionária (apresenta tendências de alta ou de baixa). Também é possível identificar que as ocorrências nos anos de 2014, 2015 e 2016 estão predominantemente aquém da média (33 ocorrências por dia). Em contrapartida, os anos de 2017 e 2018 ultrapassam os limites da média. Outra análise realizada neste recorte é o isolamento dos dados da série temporal para cada um dos anos. No entanto, 2015 não apresenta picos de crescimento nos meses de março e setembro. O ano de 2014 apresenta um comportamento particular, com pico de crescimento entre os meses de abril e julho.

⁸ <http://infocuritiba.ippuc.org.br/imp/>

⁹ <https://powerbi.microsoft.com/pt-br/downloads/>

¹⁰ <https://www.anaconda.com/distribution/>

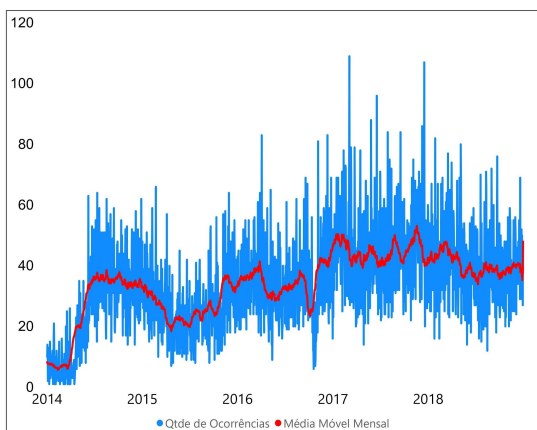


Figura 3. Série temporal de ocorrências.

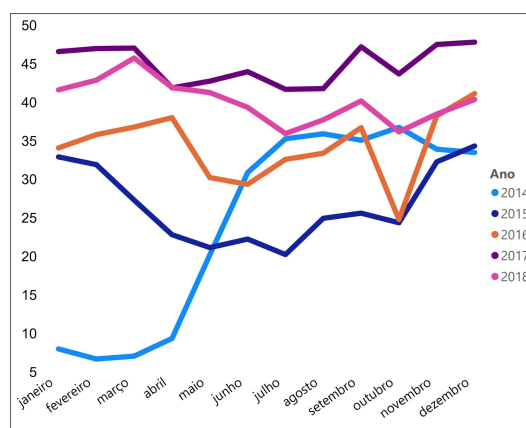


Figura 4. Média mensal de ocorrências.

Para se obter informações mais precisas, as ocorrências também foram analisadas por dia da semana e horário. Para isso, foram agrupadas por dia da semana e também pelo horário, em intervalos de uma hora. Ao se analisar esses dados, constatou-se que houve um aumento de 50,63% nas ocorrências durante o fim de semana em relação à quinta-feira. Na Figura 5, é possível identificar que de segunda-feira a quinta-feira as ocorrências permanecem sem grandes variações, mas a partir sexta-feira até domingo ocorre um aumento significativo das ocorrências. Ao se analisar o horários das ocorrências, identificou-se uma maior concentração após as 18:00 independente do dia da semana como ilustrado a Figura 6.

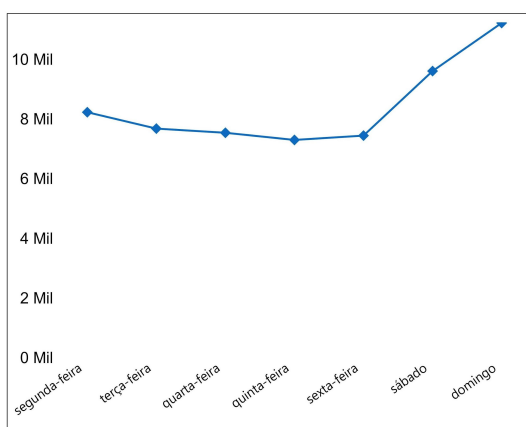


Figura 5. Ocorrências por dia da semana.

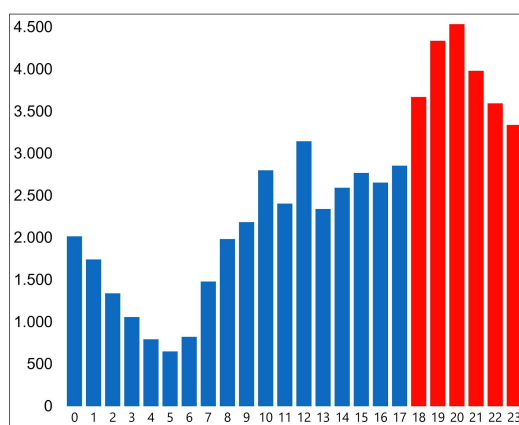


Figura 6. Ocorrências por hora.

No segundo recorte, com atributo **Idade**, foi possível a identificar a faixa etária com maior incidência de vítimas. Para esta análise, foi construído um histograma de frequência, no qual se identificou, como mostra a Figura 7, que a vítimas estão concentradas na faixa de 20 e 40 anos.

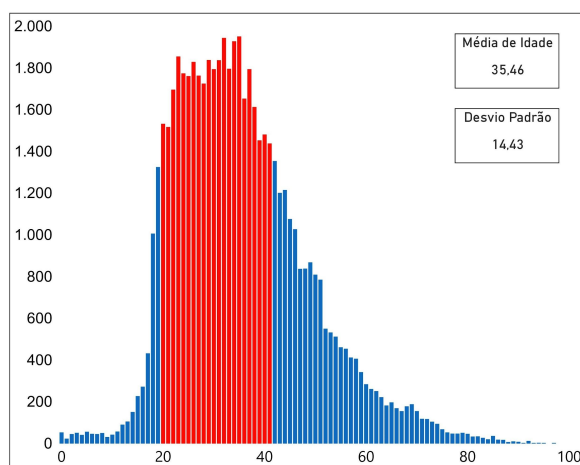


Figura 7. Ocorrências por idade.

No terceiro recorte, com os atributos **Natureza e Idade**, foram identificados quais os tipos de violência têm maior incidência por faixa etária (classes utilizadas pela CAPE) das vítimas. Com isso, identificou-se no cruzamento entre as quatro principais naturezas e a idade das vítimas que existe uma concentração de ocorrências de Lesão Corporal e Vias de Fato, consideradas violências físicas, entre crianças e adolescentes. É possível notar na Figura 8 que existe uma concentração de ocorrências registradas como Lesão Corporal ainda maior na faixa etária entre 01 e 11 anos.

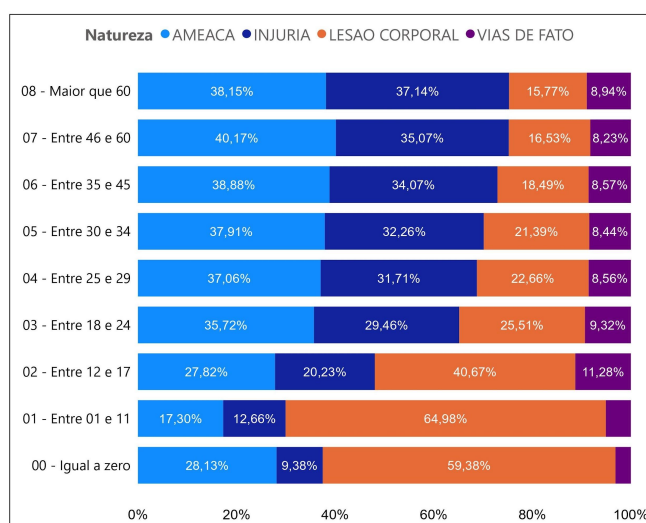


Figura 8. Natureza da ocorrência por faixa etária.

No quarto e último recorte, com os atributos **Bairro e População**, este coletado do IPPUC, foi possível identificar os bairros Cidade Industrial (8,32%), Sítio Cercado (7,72%), Cajuru (6,77%), Uberaba (5,01%) e Tatuquara (4,39%), nesta ordem, como mais violentos, representando 32,21% do total de ocorrências. O Tatuquara é o décimo bairro mais populoso, enquanto os demais estão entre os 5 mais populosos. Ao se utilizar a densidade, ou seja, dividindo o número de ocorrências pela população, obteve-se como resultado os seguintes bairros como mais violentos: Campo do Santana (0,1546), Ganchinho (0,0876), Pinheirinho (0,0721), Caximba (0,0695) e Tatuquara (0,0693). Tatuquara, como é possível notar, aparece nas duas análises. Na Figura 9, os pontos representam cada um dos bairros de Curitiba. Observando-se o gráfico de *box*

plot, é possível identificar os *outliers* (os pontos que estão fora dos limites para cada ano). São eles os bairros mais violentos de Curitiba, de acordo com a densidade. O ponto mais alto no gráfico no ano de 2017 é o bairro do Campo do Santana. Na Figura 10, é possível notar que os bairros mais violentos, apresentados em vermelho, estão localizados na região sul de Curitiba.

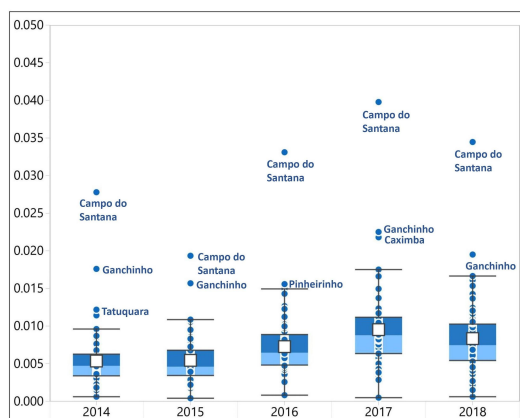


Figura 9. Densidade de Ocorrências por bairro e ano.

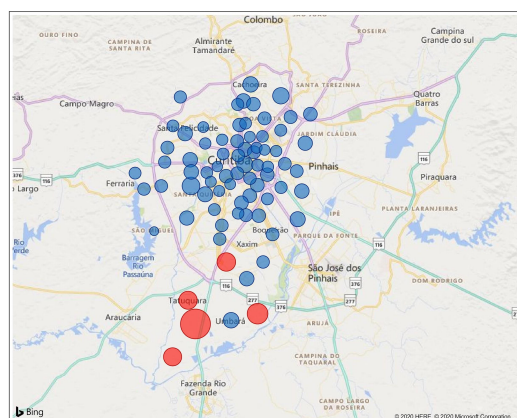


Figura 10. Densidade de ocorrências por localidade

5. Conclusões

A violência doméstica, infelizmente, está presente no cotidiano do brasileiro. Todos os dias são apresentados novos casos, geralmente envolvendo mulheres, em jornais, rádio, redes sociais e TV. Neste estudo, foi apresentada uma das possibilidades para tratar este problema utilizando-se uma das grandes fontes de dados da polícia, o boletim de ocorrência. Com essas informações analisadas, será possível ao poder público entender o cenário da situação atual da violência doméstica em Curitiba.

Toda análise deste estudo foi realizada utilizando ocorrências com vítimas do gênero feminino. Como resultado, pode-se destacar a sazonalidade nos meses de março, setembro e dezembro, quem apresentam um crescimento no número de ocorrências. Ao se analisar as ocorrências por dia da semana, identificou-se um aumento de 50,63% nas ocorrências durante o fim de semana em relação à quinta-feira, voltando a cair na segunda-feira. Além disso, identificaram-se uma concentração e o aumento das ocorrências de violência doméstica após as 18:00 horas, independente do dia da semana. As vítimas têm, em sua grande maioria, entre 20 e 40 anos e cor branca. Quanto ao grau de instrução, as informações não foram conclusivas, porém existe uma leve tendência de que mulheres com ensino superior representam um número menor entre as vítimas - aproximadamente 37% dos casos. Sobre o estado civil da vítima, a distribuição não aponta nenhuma tendência entre mulheres solteiras ou casadas. Ao se cruzar os dados sobre o tipo de violência e a faixa etária das vítimas, foi identificado o maior número de ocorrências de violência física, 78,77% (Lesão Corporal e Vias de Fato), na faixa etária entre 01 e 11 anos. Ao se analisar a localidade das ocorrências, foi possível identificar que existe uma maior concentração nos bairros da região sul de Curitiba, com destaque para Campo do Santana, Ganchinho, Pinheirinho, Caximba e Tatuquara.

Neste estudo, foi possível encontrar vários pontos de atenção, comprovando que a análise e visualização de dados podem trazer informações relevantes, capazes de

auxiliar os gestores públicos na tomada de decisão e na criação de políticas públicas para combater, neste caso, a violência doméstica.

Como trabalhos futuros, pode-se citar a inclusão de informações sobre vítimas de feminicídio, a inclusão de outras bases de dados (como 156¹¹ e dados de hospitais) e sua extensão para outras regiões.

Referências

- Bichibichi, Y., Kozievitch, N., e Carvalho, R. (2018). Análise de evolução de emissão de alvarás próximos a dois shoppings em Curitiba, In Anais da XIV Escola Regional de Banco de Dados, Porto Alegre: SBC.
- Cerqueira, D.; Lima, R.S.; Coelho D.; Alves, P.P.; Reis, M.; Cypriano, O.; Armstrong, K. (2019). Atlas da violência retratos dos municípios brasileiros 2019, <http://www.ipea.gov.br/atlasviolencia/download/21/atlas-da-violencia-dos-municipios-brasileiros-2019>.
- Eberhardt, A.; Silveira, M. (2018). Show me the data!: a systematic mapping on open government data visualization. In Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age (dg.o '18). Association for Computing Machinery, New York, NY, USA, Article 33, 1–10.
- Engel, C. L. (2019). A violência contra a mulher. http://www.ipea.gov.br/retrato/pdf/190215_tema_d_a_violencia_contra_mulher.pdf, Secretaria de Políticas Públicas para as Mulheres.
- Flores Junior, J. A. F.; Steffenello, L. C.; Winck, A. T. (2017). Mapeamento de Padrões de Acidentes de Trânsito com Vítimas Fatais a partir de Dados Públicos do Governo do Estado do Rio Grande do Sul. In XIII Escola Regional de Banco de Dados, p. 91-94.
- Gregoli, R.; Silva, V. e Ribeiro H. M. (2018). Desafios para o acesso e sistematização dos dados de violência contra as mulheres no Brasil: a experiência de implantação do observatório da mulher contra a violência. Núcleo de Estudos e Pesquisa / CONLEG / Senado, Boletim Legislativo, nº 70.
- Mantovani, I. do Couto; Santos, M. J.; Carniello, M. F.; Rodrigues, M. S. (2019). A Segurança Pública no Brasil e seus Portais como Princípios de Transparência e Dados Abertos para Sociedade. In 42º Congresso Brasileiro de Ciências da Comunicação, Pará. <http://portalintercom.org.br/anais/nacional2019/resumos/R14-1085-1.pdf>.
- Oliveira, Marcos and Ribeiro, Eraldo and Bastos-Filho, Carmelo and Menezes, Ronaldo. (2018). Spatio-temporal variations in the urban rhythm: the travelling waves of crime". EPJ Data Sci. 7, 29
- Silva, F.A.R; (2018). Analytical Intelligence in Processes: Data Science for Business. IEEE Latin America Transactions, vol. 16, no.8, 2018, pp 2240-2247.

¹¹ Central de atendimento da Prefeitura de Curitiba.