

MODELO PARA GERAÇÃO DE MAPAS CRIMINAIS: UMA ANÁLISE A PARTIR DE ESTUDO DE CASO DE DADOS ABERTOS

*Antônio Gomes Vieira Neto**

*André Marcelo Panhan***

RESUMO: O presente artigo propõe um modelo para geração de mapas criminais, utilizando ferramentas gratuitas. Para a criação do modelo foram utilizados os dados abertos do portal da transparência da Secretaria de Segurança Pública do Estado de São Paulo (SSP-SP). O estudo de caso teve como objeto de estudo o município de Bragança Paulista no Estado de São Paulo. O modelo tem como objetivo prover uma visualização melhor dos dados criminais e ajudar as instituições de segurança pública no processo de tomada de decisão. Como resultado do modelo foram gerados dois tipos de mapa, um que mostra o ponto exato onde as ocorrências ocorreram e outro um mapa de calor que exibir em cores mais quentes os locais onde se tem maior ocorrência de crimes.

Palavras-chave: Segurança Pública. Tecnologia. ETL. Mapas Criminais. Python.

* Cursando Pós-Graduação *Lato Sensu* (Especialização) em Gestão Estratégica de Tecnologia da Informação (GETI) pelo Instituto Federal de São Paulo (IFSP), Bragança Paulista, São Paulo. Tecnólogo em Análise e Desenvolvimento de Sistemas pela Unifaat, Atibaia, São Paulo. Artigo apresentado como Trabalho de Conclusão de Curso.

** Doutor em Engenharia Elétrica pela Universidade Estadual de Campinas (2011). Mestre em Engenharia Elétrica pela Universidade Estadual de Campinas (2001). Tecnólogo em Processamento de Dados pela Faculdade de Tecnologia de Sorocaba (1999). Atualmente é Pesquisador Colaborador na Universidade Estadual de Campinas e docente do Instituto Federal de São Paulo.

1. INTRODUÇÃO

Com o aumento da densidade demográfica e da criminalidade, os órgãos de segurança pública buscaram empregar novas técnicas de policiamento para conter o aumento nas taxas de criminalidade, partindo de ações que consistiam basicamente na distribuição aleatória do efetivo policial, até a análise estatística das ocorrências.

Novos modelos de gestão da segurança pública como, por exemplo, o Policiamento Comunitário, Policiamento Orientado pela inteligência (POI) e o Policiamento Preditivo surgiram com objetivo de melhorar o processo de tomada de decisão dos gestores da segurança pública e suprir o aumento da demanda policial.

Desde 1995, a Secretaria da Segurança Pública de São Paulo divulga de forma pioneira no Brasil os indicadores criminais trimestralmente no Diário oficial. Em decorrência da evolução da tecnologia e o amplo uso da internet, desde 2011 esses dados passaram a ser divulgados em seu site mensalmente por região, município e unidade policial.

Apesar dos dados estarem disponíveis eles não contam com uma apresentação visual amigável e que facilite a tomada de decisão do gestor que está analisando as informações.

O presente artigo propõe um modelo para geração de mapas criminais, utilizando ferramentas gratuitas, baseado em processo de extração, transformação e carga – ETL (do inglês *Extract, transform, load*) –, que possibilita a criação de mapas criminais, para tanto foram utilizados esses dados abertos do portal da transparência da Secretaria de Segurança Pública do Estado de São Paulo (SSP-SP).

O modelo tem como objetivo prover uma visualização melhor dos dados criminais e ajudar as instituições de segurança pública no processo de tomada de decisão, permitindo aos gestores da segurança pública dos municípios ações preventivas, através da identificação da incidência de crimes por regiões.

O procedimento metodológico foi o estudo de caso realizado na base de dados referentes ao município de Bragança Paulista, localizado no Estado de São Paulo.

2 TECNOLOGIAS DA INFORMAÇÃO NA ÁREA DE SEGURANÇA PÚBLICA

A princípio as tecnologias na área de segurança pública eram voltadas à atividade bélicas e, muitas vezes, tinham características ligadas a letalidade. No entanto, com o tempo, essas tecnologias começaram a ter um foco em aspectos mais preventivos e reativos, de forma menos letal.

Com o avanço tecnológico, as Tecnologias da Informação e Comunicação (TICs) passaram a dar apoio à Segurança Pública em diversas áreas, como Telecomunicações, Processo Informatizado, Sistemas de Informação Geográfica (SIG) e Monitoramento.

2.1 TELECOMUNICAÇÕES

A comunicação é essencial nas atividades de segurança pública, fazendo-se uso de tecnologias de comunicação por voz e de comunicação por dados.

A comunicação por voz é a mais comum entre os agentes de segurança pública. Nesse tipo de comunicação podemos destacar os Sistemas de comunicação por Rádio Frequência e os Sistemas Troncalizados.

Os aparelhos de rádio frequência podem transmitir informações de voz, vídeo e dados para uma frequência específica, onde outros aparelhos sintonizados irão receber as informações.

Apesar de alguns sistemas de rádio frequência ter transmissão de dados, eles são de alta latência e baixa taxa de transmissão, sendo assim são mais indicados para troca de mensagens curtas.

Os Sistemas Troncalizados, também conhecidos *Trunking*, são muito similares aos sistemas de telefonia convencional. Segundo Müller, “O *Trunking* é o compartilhamento automático de canais

em um sistema de múltiplos repetidores. As principais vantagens desse sistema são o menor tempo de acesso (...) e o aumento da capacidade dos canais para uma dada qualidade de serviço” (MÜLLER LOHN, 2012).

Para a comunicação de dados são utilizadas tecnologias de transmissão de dados em redes de celular como HSPA, LTE e WiMax ou, até mesmo, rede Wi-Fi.

No Brasil, o *Long Term Evolution* (LTE) é a tecnologia mais comum utilizada na 4ª geração de telefonia móvel. Como essa tecnologia faz a transmissão de dados de forma mais eficiente, possibilita a transmissão de imagens em alta definição, videoconferências, etc.

2.2 PROCESSO INFORMATIZADO

Tendo em vista o grande volume de dados produzidos por todos os órgãos de segurança pública e outros órgãos relacionados, surgiu a necessidade de informatizar e integrar os processos. No Brasil, com essa finalidade de integração foi criado o INFOSEG, que integra informações de Segurança Pública, Justiça e fiscalização através da internet.

“O INFOSEG disponibiliza dados de Segurança Pública, Justiça e Fiscalização por meio da Internet, utilizando um Índice Nacional, que possibilita acessar dados básicos de indivíduos, armas de fogo, veículos, condutores, CPF e CNPJ” (MANUAL REDE INFOSEG (Versão 4.0), 2009).

2.3 SISTEMAS DE INFORMAÇÃO GEOGRÁFICA (SIG)

Os Sistemas de Informação Geográfica (SIG), também conhecidos como *Geographic Information Systems* (GIS), é um conjunto de sistemas de georreferenciamento, voltados para confecção de mapas, provendo seu processamento e visualização.

Um sistema de informação geográfica (SIG) é um sistema de informação que é concebido para trabalhar com dados referenciados por coordenadas geográficas ou espaciais. Em outras palavras, um SIG é um sistema de bases de dados com capacidades específicas para lidar com dados espacialmente referenciados, bem como um conjunto de operações para trabalhar com dados. De certo modo, um SIG pode ser pensado como mapa de ordem superior (STAR e ESTES, 1990, p. 2-3, apud MÜLLER LOHN, 3000).

Os SIG podem ser usados em diversas áreas que necessitem de informações geográficas, por exemplo, na Segurança Pública pode ser utilizado no mapeamento da criminalidade, utilizando a análise espacial.

Dentre os tipos de análises temos: análise de zonas quentes de criminalidade, análise de redes viárias, planejamento de barreiras policiais, localização rápida de viaturas, mapeamento do tempo, mapeamento do espaço, mapeamento por características registradas, entre outros.

2.4 MONITORAMENTO

O Monitoramento é uma ferramenta essencial no apoio ao policiamento preventivo. Por exemplo, através da distribuição de câmeras pela cidade, os agentes de segurança pública podem monitorar grandes áreas sem a necessidade de se expor aos riscos da profissão. Além do monitoramento, existe a possibilidade de gravação e recuperação das imagens, que podem gerar evidências em processos criminais.

As utilizações dessas tecnologias acabam servindo para inibir alguns tipos de delitos e também na identificação de criminosos.

3 TRABALHOS RELACIONADOS

Nesta seção apresentaremos algumas experiências realizadas no Brasil e referentes à temática.

3.1 SIAPP: UM SISTEMA PARA ANÁLISE DE OCORRÊNCIAS DE CRIMES BASEADO EM APRENDIZADO LÓGICO-RELACIONAL

O Sistema de Apoio ao Policiamento Preditivo (SiAPP) utiliza técnicas de aprendizado de máquina para dar apoio à análise e previsão de ocorrências criminais e é alimentado com dados colaborativos de forma automatizada, criando uma visualização geográfica dos padrões descobertos.

O SiAPP é uma aplicação web, desenvolvida com linguagem *Python* e *Prolog*, rodando em um ambiente de nuvem e em servidor *Apache Tomcat*. Permite a visualização dos dados criminais em duas formas, na primeira forma são apresentados os pontos onde aconteceram as ocorrências e na segunda forma é apresentado um modelo preditivo do mesmo mapa.

Os resultados serão considerados satisfatórios se nos experimentos das previsões se obtiver um resultado de acurácia média maior que 83%.

3.2 EMPREGANDO O POLICIAMENTO PREDITIVO: CONSTRUÇÃO DE UM MODELO DE RISCO DO TERRENO PARA CRIMES CONTRA O PATRIMÔNIO DOS CORREIOS

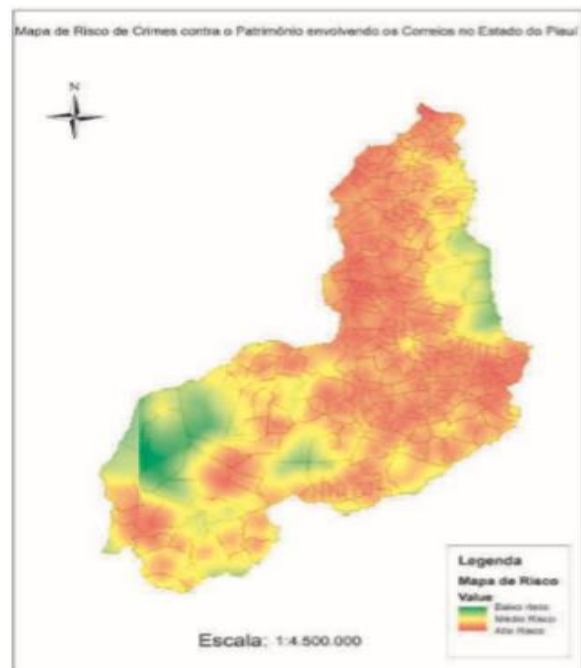
Essa experiência propôs o desenvolvimento de um modelo de risco de terreno, com base em técnicas de Policiamento Preditivo. O modelo emprega técnicas de modelagem em um Sistema de Informações Geográficas (SIG), para criar um mapa de risco de terreno – RTM (do inglês *Risk Terrain Model*) – dos crimes contra o patrimônio dos correios do estado do Piauí no ano de 2013.

O trabalho tem como base variáveis levantadas a partir uma consulta ao Serviço de Análise de Dados e Inteligência Policial (SADIP/DPAT) da Divisão de

Repressão aos Crimes contra o Patrimônio da Polícia Federal (DPAT/DICOR/DPF).

O trabalho teve como resultado o mapa da figura 1, onde as cores cumprem o objetivo de prever os locais de maior risco de ocorrências desse tipo de crime.

Figura 1. Mapa de risco de crimes contra o Patrimônio dos Correios do Piauí.



Fonte: Porcino Silva (2016).

Para validar os dados foi feito um comparativo, da previsão do mapa, foi feito um comparativo entre o resultado e as ocorrências do ano de 2014.

Os resultados foram satisfatórios, no ano seguinte 94,7 dos crimes dessa natureza foram cometidos nas áreas do mapa indicadas como altíssimo e alto.

3.3 O AMBIENTE CONSTRUÍDO E A OCORRÊNCIA DE CRIMES: UMA ANÁLISE EM ESTACIONAMENTOS DE CAMPUS UNIVERSITÁRIO

O trabalho de Soares et al (2017, 102-116) tem foco na relação entre a forma urbana e arquitetônica e as ocorrências de crime em ambientes universitários. Os Autores fazem uma análise de quatro

estacionamentos do campus principal da Universidade Federal de Santa Catarina (UFSC), para identificar se determinados fatores influenciam nas ocorrências de crimes nesses locais.

Foi feita uma análise comparativa com base em levantamentos em loco. Como resultado do estudo foram gerados mapas, gráficos com os resultados relacionados às variáveis utilizadas, como visibilidade, permeabilidade, quantidade de pessoas transitando pelo local, variedade de usos do solo, iluminação e recursos de vigilância.

Com esse estudo pode-se concluir que a ocorrência de crimes é mais frequente em áreas que apresentam configuração espacial com menor campo de visão, com barreiras visuais e menor Intervisibilidade entre as pessoas. E que locais com maior permeabilidade, melhor iluminação e com maior controle com recursos de vigilância tende a contribuir para um número menor ocorrências.

4 PROCESSO

4.1 DADOS ABERTOS

De acordo com a *Open Knowledge Internacional* (2019), dados abertos são aqueles que podem ser usados, reutilizados e distribuídos livremente por qualquer pessoa, exigindo apenas que a fonte seja citada ou que os dados sejam divulgados com a mesma licença.

O conceito de dados abertos está diretamente ligado aos Dados Governamentais Abertos (DGA) que se trata da disponibilização de informações e dados governamentais através da internet de forma publica de maneira que qualquer pessoa possa acessa-la.

Foram apresentados por Eaves (2009) as três leis dos Dados Governamentais Abertos:

- O dado não existe se não for indexado ou encontrado na web.

- O dado não pode ser aproveitado se não estiver aberto e disponível em formato compreensível.
- O dado é inútil se houver algum instrumento legal que não permita a sua replicação.

Os usos de Dados abertos governamentais são importantes para transferência dos dados governamentais e para pesquisas científicas que utilizam dados públicos.

4.2 PROCESSO DE ETL

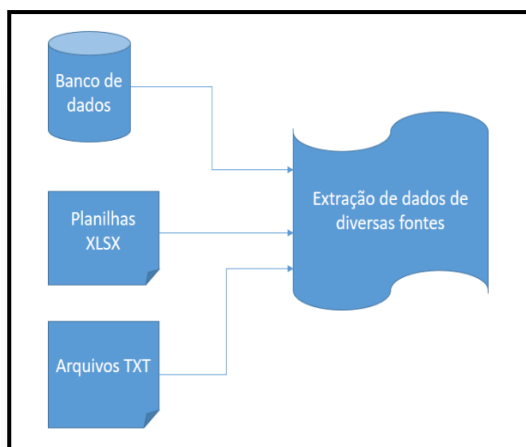
A sigla ETL vem da língua inglesa e quer dizer, *Extract, Transform and Load* (Extração, transformação e carga.). ETL é um processo baseado em ferramentas computacionais de extração, transformação e carga de dados. Os dados podem vir de uma ou mais fontes de dados, assim como o resultado do processo pode ser uma ou mais fonte de dados ou até mesmo um data *warehouse*.

ETL é bastante utilizado em *Business Intelligence*, aplicações de *data warehouse*, *data marts* e migração de dados para sistemas de informação.

4.2.1 Extração

O primeiro passo no processo de ETL é a extração de dados, os dados brutos originados de um ou mais fontes de dados.

Figura 2. Extração de dados de diversas fontes.



Fonte: autoria própria (2019)

Os dados podem ser extraídos de dados estruturados como bancos de dados relacionais, *web services* ou de dados não relacionais como planilhas eletrônicas, arquivos texto, e-mails, etc.

4.2.2 Transformação

Nessa etapa do processo os dados são tratados e transformados de acordo com as regras de negócio da aplicação na qual se destinam os dados.

Durante esta etapa é comum serem encontrados problemas de integridade, dados inválidos, atributos obrigatórios nulos ou em branco, etc.

Segundo Cielo (2010) quaisquer divergências encontradas devem ser solucionadas nessa etapa, com o objetivo garantir a integridade e confiabilidade dos dados.

Outra preocupação é a padronização dos dados já que podem ter sido originados de diversas fontes. Então, problemas relacionados à precisão de campos numéricos, diferenciação de letras maiúsculas e minúsculas, ponto no lugar de vírgulas, entre outros, devem aparecer nessa etapa.

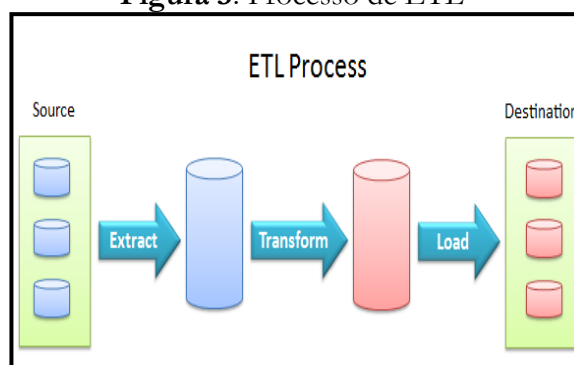
Estes múltiplos formatos de origem devem ser transformados no formato padrão da base de dados destino.

4.2.3 Carga

No processo de carga, os dados extraídos, tratados e transformados nas etapas anteriores são gravados em uma ou mais fontes de dados de destino.

A Figura 1 apresenta a transformação dos dados conforme as regras da base de dados destino.

Figura 3. Processo de ETL



Fonte: datasciencecentral.com (2010)

4.3 APLICAÇÕES

O *Python*¹ é uma linguagem de programação de alto nível com as seguintes características: interpretada, de script, imperativa, orientada a objetos, funcional, dinamicamente e fortemente tipada.

*Pandas*² é uma biblioteca *open source*, licenciada pela BSD, que fornece estruturas de dados de alto desempenho e fáceis de usar e ferramentas de análise de dados para a linguagem de programação *Python* (pandas.pydata.org).

*Folium*³ é uma biblioteca do *Python* baseada na biblioteca *Leaflet.js* que permite a visualização de mapas (pypi.org/project/folium).

1 Disponível em: <python.org>. Acesso em: 4 jun. 2019.

2 Disponível em: <<https://www.pandas.pydata.org>>. Acesso em: 4 jun. 2019.

3 Disponível em: <<https://www.pypi.org/project/folium>>. Acesso em: 4 jun. 2019.

Para analisar os dados foi utilizada a linguagem de programação *Python*, na sua versão 3.7 com as bibliotecas *Pandas* e *Folium*.

5 ESTUDO DE CASO

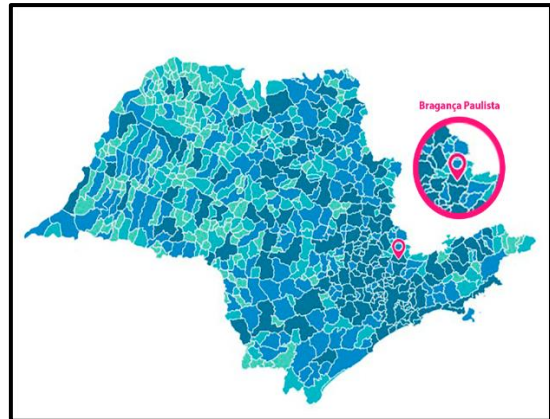
5.1 SELEÇÃO DOS DADOS

O Governo do estado de São Paulo, possui um portal da transparência com os dados abertos da Secretaria de Segurança Pública do Estado de São Paulo (SSP-SP)⁴, os dados começaram a ser divulgados através da internet desde 2011 e é o maior portal de informações sobre Segurança Pública do país.

O portal fornece taxas de homicídio, registros de óbito do Instituto Médico Legal (IML) e boletins de ocorrência de diversas naturezas como homicídio doloso, latrocínio, morte decorrente de intervenção policiais, roubo e furtos de veículos, roubo e furtos de celulares e entre outros. Esses dados são disponibilizados mensalmente e separados por região, município e unidade policial, o que permite a exportação dos dados em formato de planilhas eletrônicas.

A área de estudo escolhida foi o município de Bragança Paulista no estado de São Paulo. Segue o mapa indicando sua localização no estado de São Paulo.

Figura 4. Delimitação da área de objeto de estudo.



Fonte: Adaptado de IBGE (2019)

Dentre os dados disponíveis foram escolhidos os de furto e roubo de veículos do ano de 2018, especificamente da Delegacia Seccional de Bragança Paulista que engloba todos os departamentos de polícia da região em estudo.

Em relação ao formato dos dados, eles são disponibilizados no formato MS *Excel* (XLS). Sua estrutura é bem simples: cada linha do arquivo representa um boletim de ocorrência e cada campo é separado pelas colunas e a descrição do campo no cabeçalho do arquivo. Qualquer editor de planilhas eletrônicas pode ser usado para visualização e manipulação dos arquivos XLS, sem exigir do usuário conhecimentos de programação.

As planilhas contêm os dados detalhando o quantitativo dos boletins de ocorrência de furtos e roubos de veículo. Os boletins fornecem diversos dados, mas para a geração do modelo foram utilizados apenas os seguintes campos: CIDADE, LATITUDE, LONGITUDE, DESCR_MARCA_VEICULO.

4 Disponível em: <http://www.ssp.sp.gov.br/transparenciassp>. Acesso em: 22 jul. 2019

5.2 PROCESSO DE ETL

Os dados do Portal da Transparência passaram pelo processo de ETL, com objetivo de unificar os dados das diferentes fontes, padronizar o formato dos conteúdos e garantir a qualidade e consistências dos dados.

Na etapa de extração, foram extraídos do portal 12 (doze) arquivos referentes aos meses do ano de 2018 para os crimes de furto e roubo de veículos totalizando 24 (vinte e quatro) arquivos.

Na etapa de Transformação, os arquivos passaram por um processo de limpeza, onde foram apagadas as linhas onde a coluna CIDADE não se referia à cidade objeto do estudo e as linhas em que as colunas LATITUDE e LONGITUDE estavam vazias.

Durante a elaboração do *script* para geração dos mapas mostrou-se necessário ajustar alguns dados para adequá-los à biblioteca *Folium*.

Na Etapa de Carga, os arquivos foram unificados através de um macro do Excel, então foi gerado um arquivo único para cada tipo de crime. Os arquivos foram salvos no formato CSV para facilitar a manipulação dos dados pela linguagem de programação *Python*.

5.3 APLICAÇÃO

A partir da base de dados gerada pelo processo de ETL, foi desenvolvida aplicação utilizando a linguagem *Python* e as bibliotecas *Folium* e *Pandas*.

A biblioteca *Pandas* se encarregou de transformar os dados do arquivo CSV em um *Data Frame*. O *Data Frame* é uma estrutura de dados que cria uma tabela com linhas e colunas e armazena em memória. Foram gerados 3 (três) *Data Frames*, um para cada tipo de crime e um com a junção de ambos.

Alguns dados dos *Data Frames* precisaram ser padronizados. Os campos de LATITUDE e LONGITUDE foram padronizados com 9 (nove) caracteres. Foi

feita a remoção dos pontos, seguido da correção da casa decimal.

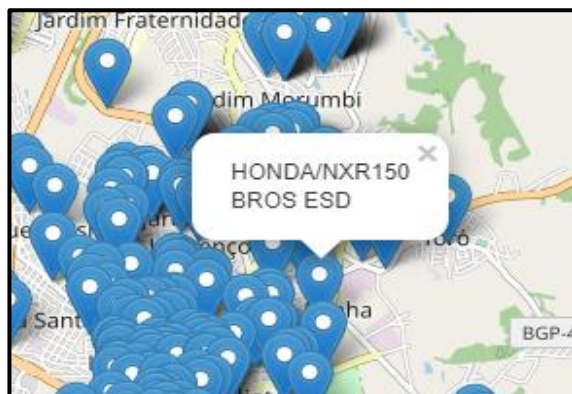
Após a padronização, foi utilizada a biblioteca *Folium* para a geração dos mapas.

5.4 GERAÇÃO DOS MAPAS

Como resultado, foram gerados 2 tipos de mapa, um mapa de ocorrências e um mapa de calor.

O mapa de calor indica as áreas com maior ocorrência dos crimes, por cores mais quentes. E o mapa de ocorrência exibe uma marcação onde cada crime ocorreu, sendo possível configurar um texto para ser exibido ao clicar na marcação. No exemplo abaixo, ao clicar em uma das marcações é exibido o veículo envolvido na ocorrência.

Figura 5 – Exemplo de texto da marcação



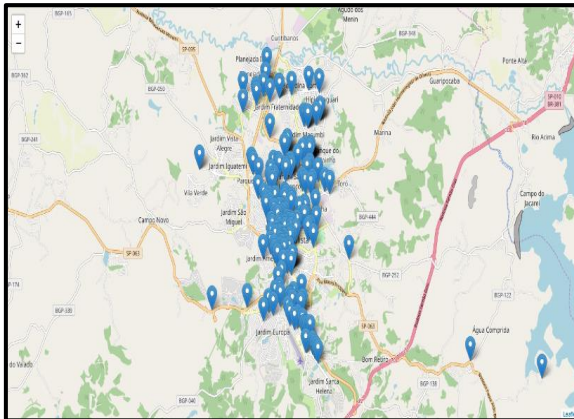
Fonte: autoria própria (2019)

5.4.1 Mapa de Ocorrências

O mapa de ocorrência mostra precisamente em que ponto do mapa os crimes aconteceram e ao clicar em algum ponto exibido do mapa será mostrado o veículo envolvido no crime naquele ponto específico.

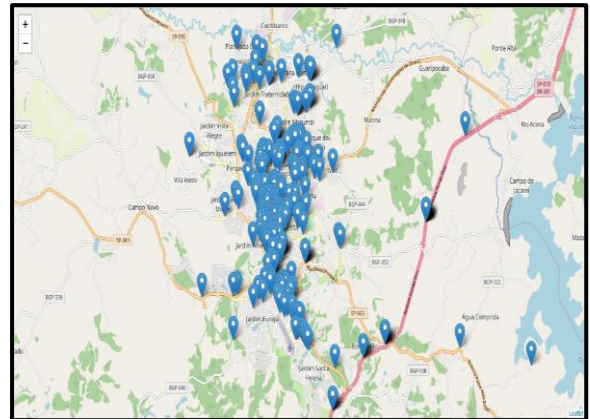
Foram gerados 3 (três) mapas, um de furtos, um de roubo e um terceiro mapa com os dois tipos de crimes.

Figura 6 – Mapa de Ocorrências Furtos 2018



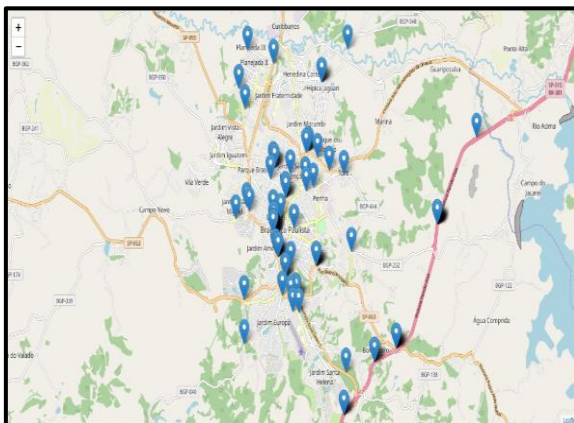
Fonte: autoria própria (2019)

Figura 8 - Mapa de Ocorrências Furtos e Roubos 2018



Fonte: autoria própria (2019)

Figura 7 - Mapa de Ocorrências Roubos 2018

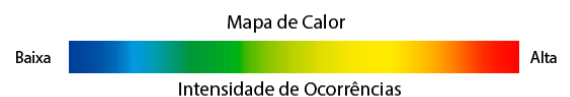


Fonte: autoria própria (2019)

5.4.2 Mapa de Calor

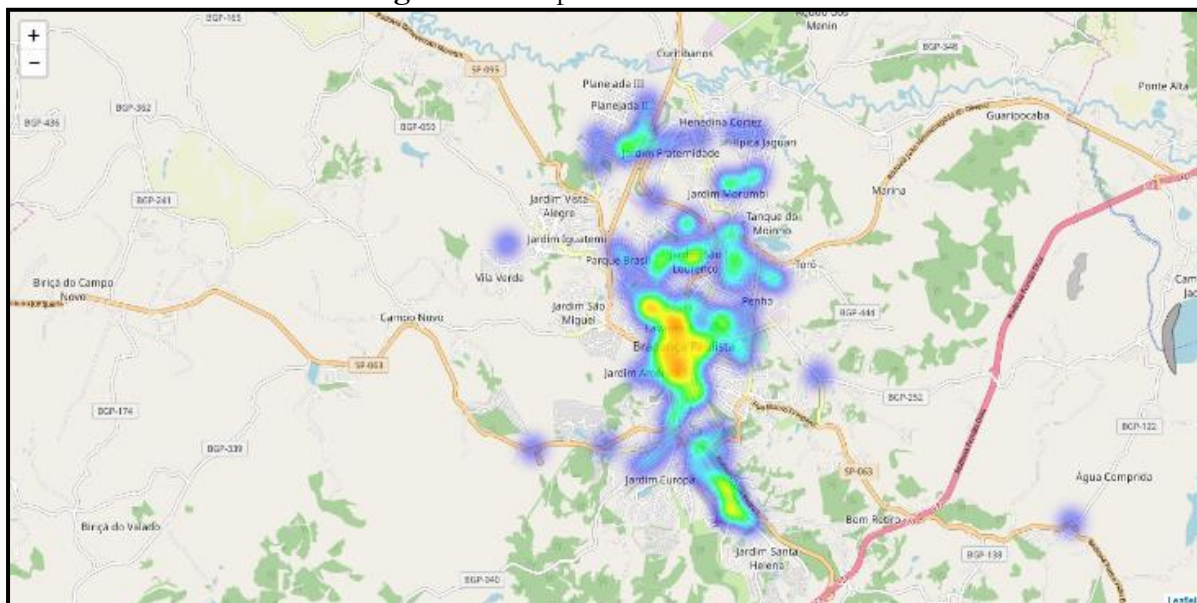
O mapa de calor destaca as áreas onde os delitos acontecem com maior intensidade. As cores mais quentes indicam maior incidência de crimes, conforme a imagem abaixo.

Figura 9 - Mapa de Calor Furtos 2018

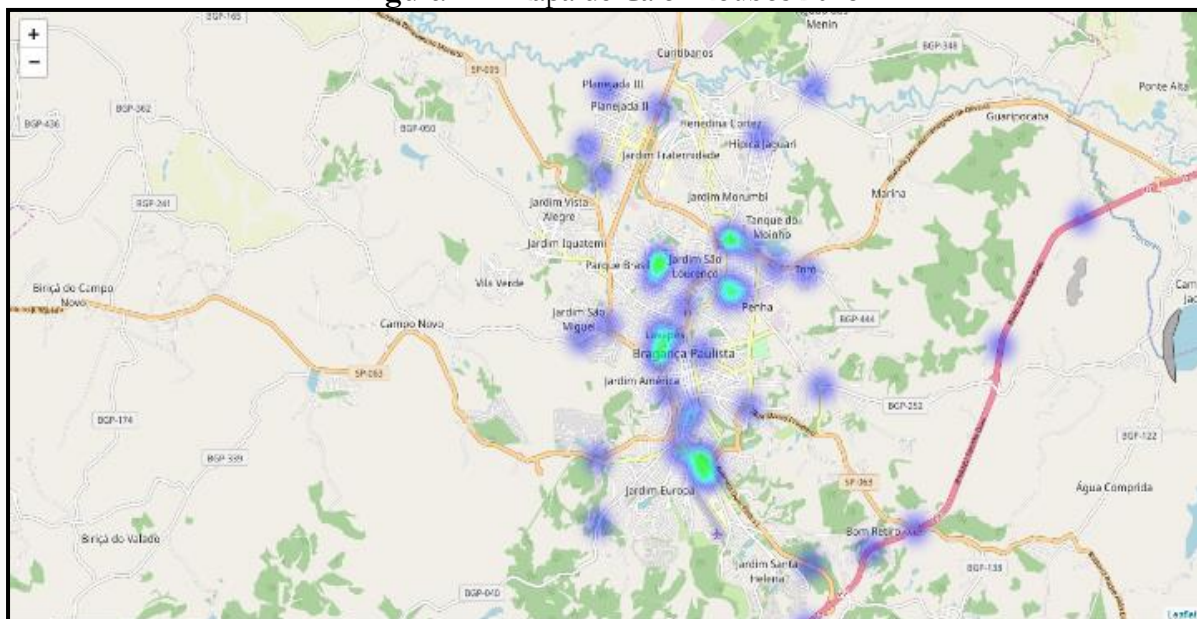


Fonte: o autor

Foram gerados 3 (três) mapas das tipologias criminais, sendo um de furto, um de roubo e um terceiro mapa com os dois tipos de crimes.

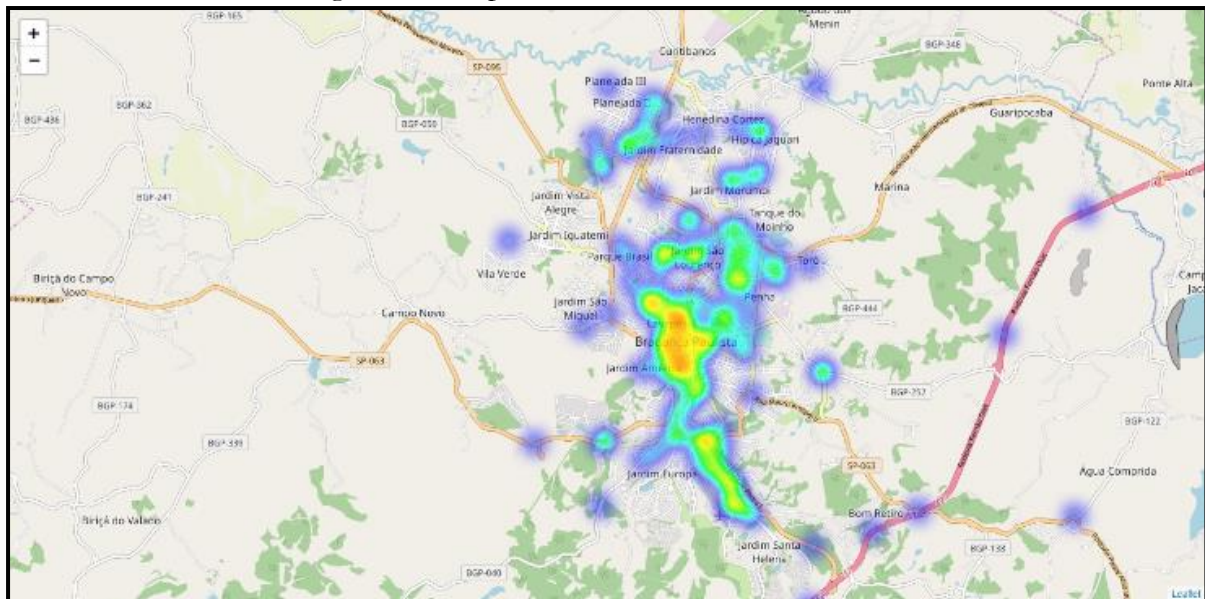
Figura 10 - Mapa de Calor Furtos 2018.

Fonte: autoria própria (2019)

Figura 11 - Mapa de Calor Roubos 2018

Fonte: autoria própria (2019)

Figura 12 - Mapa de Calor Furtos e Roubos 2018



Fonte: autoria própria (2019)

6 AVALIAÇÃO DO MODELO

O modelo trouxe resultados interessantes, com uso de tecnologias gratuitas, algumas até mesmo *open source*⁵ como é o caso do *Python* e da biblioteca *Pandas*. E também a utilização do formato CSV (*Comma-separated values*), compatível com a maioria dos editores de planilha eletrônica do mercado.

Com o preenchimento correto de uma planilha eletrônica e a execução do *script* em uma máquina com o *Python* instalado, os mapas podem ser gerados em questão de segundos.

O modelo pode ser aplicado independente da instituição e do tipo de crime, basta ter apenas as informações geográficas preenchidas na planilha.

⁵ Segundo a Indiana University (2019), open source é um termo que descreve um meio de desenvolver e distribuir software que garante que o software esteja disponível para uso, modificação e redistribuição por qualquer pessoa.

Pode-se inclusive utilizar dados públicos, como foi o caso dos dados utilizados neste estudo.

7 CONCLUSÃO

A utilização de uma ferramenta de visualização de dados permite uma análise mais rápida destes. Os 24 arquivos utilizados puderam ser visualizados de forma intuitiva, recurso este que pode facilitar a tomada de decisão pelos gestores das instituições de Segurança Pública.

O *Python* é uma linguagem de programação *open source* com uma sintaxe simplificada, o que pode facilitar o processo de aprendizagem, além de ser considerada uma das principais linguagens do mercado com um vasto suporte da comunidade. Não estamos aqui, em hipótese alguma realizando propaganda do referido produto, mas sim o tomamos como exemplo de possibilidades e estratégias *free* que podem ser utilizadas para

otimização das tomadas de decisões. Assim como o modelo desenvolvido e demonstrado neste artigo, muitos outros podem ser pensados.

A análise de dados demonstrou que os furtos acontecem como uma maior frequência em relação aos roubos, mas independente disso ele ocorrem praticamente nos mesmos lugares.

Através dos mapas gerados os dados podem ser visualizados de uma maneira mais amigável, com mais clareza e facilidade de interpretação, possibilitando assim uma análise orientada dos dados para tomada de decisão pelas instituições policiais.

REFERÊNCIAS

AZEVEDO, Ana Luísa Vieira de *et al.* **A utilização das estatísticas criminais no planejamento da ação policial:** Cultura e contexto organizacional como elementos centrais à sua compreensão. Brasília - DF: Ci. Inf., 2011. 1-13 p.

BRASIL. Ministério da Justiça. **Manual rede Infoseg (versão 4.0).** Brasília: 2009. Disponível em: < <http://sti.tjrr.jus.br/arqpdf/suporte/infocis/infoseg/INFOSEG-Manual.pdf>> . Acesso em: 31 ago. 2019.

CAMPOS, Sandir Rodrigues. **Validação de dados em sistemas de data warehouse através de índice de similaridade no processo de ETL e mapeamento de trilhas de auditoria utilizando indexação ontológica.** Brasília, DF, p. 30-38, 11 jan. 2013.

DATA SCIENCE CENTRAL. **10 Open Source ETL Tools.** [S. l.]: 2015. Disponível em: < <https://www.datasciencecentral.com/profiles/blogs/10-open-source-etl-tools> >. Acesso em: 29 ago. 2019.

EAVES, D. **The three laws of open government.** Disponível em: < <http://eaves.ca/> >. Acesso em: 12 ago. 2019.

FOLIUM. [S. l.]. Disponível em: < <https://www.pypi.org/project/folium> >. Acesso em: 4 jun. 2019.

FREITAS, Lauro Soares de *et al.* **A institucionalização do modelo de gestão CompStat na Polícia Militar de Minas Gerais sob a perspectiva teórica do Translation e Trabalho Institucional.** Viçosa - MG: [s.n.], 2013. 76-106 p.

INDIANA UNIVERSITY. **About open source software and the Open Source Initiative.** [S. l.]: 2019. Disponível em: < <https://kb.iu.edu/d/annr> >. Acesso em: 06 nov. 2019.

JOH, Elizabeth E. **Policing by numbers:** Big data and the fourth amendment. 2014. ed. [S.l.]: Washington Law Review, 2014. 1-34 p.

LAMAS, João Paulo Campolinas. **Predição de Crimes e Otimização de Ações de Segurança Pública para Cidades de Pequeno Porte Utilizando Geotecnologias.** Viçosa - MG: Universidade Federal de Viçosa, 2013.

LOHN, Josemar Müller. **Tecnologias aplicadas à segurança pública:** livro didático / Josemar Müller Lohn; design instrucional Marina Melhado Gomes da Silva. – Palhoça: UnisulVirtual, 2012.

PANDAS. [S. l.]. Disponível em: < <https://www.pandas.pydata.org> >. Acesso em: 4 jun. 2019.

PORCINO SILVA, Wellington Clay. **Empregando o Policiamento Preditivo:** Construção de um Modelo de Risco do Terreno para Crimes contra o Patrimônio dos Correios. Brasília - DF: Revista Brasileira de Ciências Policiais, 2016.

PYTHON. [S. l.]. Disponível em: < <https://www.python.org> >. Acesso em: 4 jun. 2019.

RIBEIRO, Alexandre Lima; OLIVEIRA, Edemar Costa. **Processos de implantação e migração de dados com utilização de ETL para um ERP comercial**. Canoas, RS, p. 2-4, 29 nov. 2010.

SAISSE, Renan. **Big Data contra o crime**: Efeito Minority Report. 2017. Disponível em: < <http://direitoeti.com.br/artigos/big-data-contra-o-crime-efeito-minority-report> >. Acesso em: 15 maio 2018.

SÃO PAULO. TRANSPARÊNCIA SSP-SP. São Paulo: 2019. Disponível em: < <http://www.ssp.sp.gov.br/transparenciassp/> > . Acesso em: 22 jul. 2019.

SOARES, Mariana et al.. O ambiente construído e a ocorrência de crimes: uma análise em estacionamentos de campus universitário. **PARC Pesquisa em Arquitetura e Construção**, Campinas, SP, v. 8, n. 2, p. 102-116, jun. 2017. ISSN 1980-6809. Disponível em: < <https://periodicos.sbu.unicamp.br/ojs/index.php/parc/article/view/8649893> >. Acesso em: 14 dez. 2017. doi:<http://dx.doi.org/10.20396/parc.v8i2.8649893>.

THE OPEN DEFINITION. [S. l.], 2019. Disponível em: < <https://opendefinition.org> >. Acesso em: 22 jul. 2019.

VAZ, José Carlos; RIBEIRO, Manuella Maia; MATHEUS, Ricardo. **Dados governamentais abertos e seus impactos sobre os conceitos e práticas de transparência no Brasil**. Salvador (BA): Cadernos PPG-AU/UFBA, 2010, p. 45-62.

MODEL FOR CRIMINAL GENERATION MAPS

ABSTRACT: This article proposes a role model for criminal generation map, using free tools. For the creation of this role model, transparency portal's open data of the São Paulo's Security Public Secretary (SSP-SP) has been used. The case study had as object of study the city of Bragança Paulista in the State of São Paulo. The model itself aims to provide a better visualization of criminal data and to assist public security institutions in the decision making process. As a result of the model, two map types have been generated, one shows the exact spot where the occurrences have occurred and the other one is a heat map whose warm colors places where crimes prevalent the most.

Keywords: Public Security. Technology. ETL. Criminal Maps. Python.

Recebido em 19 de setembro de 2019

Aprovado em 14 de novembro de 2019