

# Violência Urbana: Armas de fogo

Edymir Semedo, Joao Luiz, Kessia Nepomuceno, Lucas Biagini

*Universidade Federal de Pernambuco*

---

## Abstract

O Brasil está entre os países com maior taxa de homicídios por armas de fogo no mundo com 25,2 casos para cada 100 mil pessoas [1]. Sabe-se também que em toda a América Latina 66% dos homicídios são por armas de fogo [1].

O Estatuto do Desarmamento, lei federal que entrou em vigor no dia 23 de Dezembro de 2003, aparentemente não teve o efeito desejado porque a violência urbana continua aumentando no país. Isso levou com que nos últimos anos houvesse maior preocupação com políticas desarmamentistas. No entanto, o comércio de armas e munições no Brasil não é proibido, mesmo que ainda possua restrições devido às políticas desarmamentistas em vigor.

Para este projeto nós pretendemos descobrir fatores que contribuem para a violência urbana, mais precisamente crimes com armas de fogo, que poderiam ser levadas em conta em futuras políticas no Brasil. Nossa proposta inicial era fazer um estudo sobre o número de homicídios em Pernambuco nos anos 2000 e 2010 por microrregiões do estado de Pernambuco e dos censos demográficos produzidos pelo IBGE nos anos de 2000 e 2010, porém tendo em vista reduzir o escopo do projeto, decidimos por realizar a análise apenas para o ano 2000.

*Keywords:* Violência Clustering

---

## 1. Obtendo os dados

Os dados utilizados nessa análise foram obtidos através das plataformas do IBGE para tabelas referentes aos censos demográficos e através da plataforma DATASUS para os dados sobre homicídios. A plataforma do DATASUS fornece  
5 os homicídios ocorridos entre 1996 e 2015 em todos os municípios brasileiros,

porém iremos usar apenas os dos anos 2000 e 2010. Outra fonte de dados para o nosso estudo será uma tabela da Wikipedia que relaciona as microrregiões de Pernambuco por municípios e dessa forma, poderemos cruzar os dados do DATASUS com os dados do IBGE.

## 10 **2. Pré Processamento**

Em seguida tratamos todos os dados utilizando o esforço manual e também scripts python para simplificação e junção de informações. Como estávamos trabalhando com planilhas robustas e que geralmente tinham bastante informação hierárquica em seus cabeçalhos, conseguimos sair de planilha com cabeçalhos  
15 complexos, para simples colunas contendo as propriedades necessárias para nossa análise sendo cada coluna uma possível característica de nossa análise que poderia mais a frente se mostrar uma boa característica influenciadora sobre a taxa de homicídio. Foi definido que iríamos focar esforços nos dados de 2000, sendo que também possuíamos informações para o ano de 2010, porém  
20 por questão de tempo e a fim de obter alguma conclusão mais rápida, decidimos utilizar apenas os dados do ano 2000. Com a estruturação das novas planilhas, existiam agora colunas que era em taxa, apresentando dados percentuais, e planilhas que representavam o valor bruto. Em conversa com o professor da disciplina foi recomendado exercitar uma normalização dos dados que levou  
25 em consideração. Para dados brutos que não fossem percentuais foi feito uma minimização por 100 mil habitantes e depois normalizado em cima do máximo e no mínimo apresentado. Já para dados percentuais simplesmente foi dividido por 100 cada valor das colunas, tendo assim dados prontos para a próxima etapa, que seria a clusterização.

## 30 **3. Clusterização**

Com os dados em mãos, fizemos uma seleção das melhores características através de regressões lineares entre cada uma das variáveis obtidas do censo 2000 com relação à taxa de homicídios dos municípios de Pernambuco e filtramos pelo

valor do *pvalue* de cada modelo, pegando todas as variáveis de *pvalue* menor  
35 que  $10^{-9}$ .

As variáveis escolhidas foram:

- (1) *peessoas de 10 anos ou mais de idade grupos de anos de estudo 1 a 3 anos*: Representa todas as pessoas com 10 anos ou mais que estudaram de 1 a 3 anos no momento da pesquisa;
- 40 (2) *peessoas de 10 anos ou mais de idade empregadas no trabalho principal da semana de referencia subgrupo do emprego no trabalho principal trabalhadores domesticos categoria do emprego no trabalho principal sem carteira de trabalho assinada*: Representa todos os trabalhadores domésticos de 10 anos ou mais que se encontravam empregados  
45 no momento da pesquisa sem carteira de trabalho assinada;
- (3) *peessoas de 10 anos ou mais de idade empregadas no trabalho principal da semana de referencia subgrupo do emprego no trabalho principal demais empregados categoria do emprego no trabalho principal outros sem carteira de trabalho assinada (1)*: Representa  
50 todos os demais trabalhadores de 10 anos ou mais que se encontravam empregados no momento da pesquisa sem carteira de trabalho assinada;
- (4) *valor do rendimento nominal mediano mensal do trabalho principal das pessoas de 10 anos ou mais de idade ocupadas na semana de referencia com rendimento no trabalho principal (r\$) posicao na  
55 ocupacao no trabalho principal empregados categoria do emprego no trabalho principal outros sem carteira de trabalho assinada (1)*: Representa o valor do rendimento mediano mensal do trabalho principal dos demais trabalhadores de 10 anos ou mais que estavam empregados no momento da pesquisa sem carteira de trabalho assinada.

60 Estas variáveis foram então usadas para clusterizar os municípios através do algoritmo KMeans, tendo seu melhor agrupamento sido com `n_clusters=3` e produzindo a silhueta da Figura 1. Testamos também muitas outras combinações de variáveis diferentes das 4 selecionadas no intuito de melhorar a clusterização

obtida, porém não obtivemos sucesso.

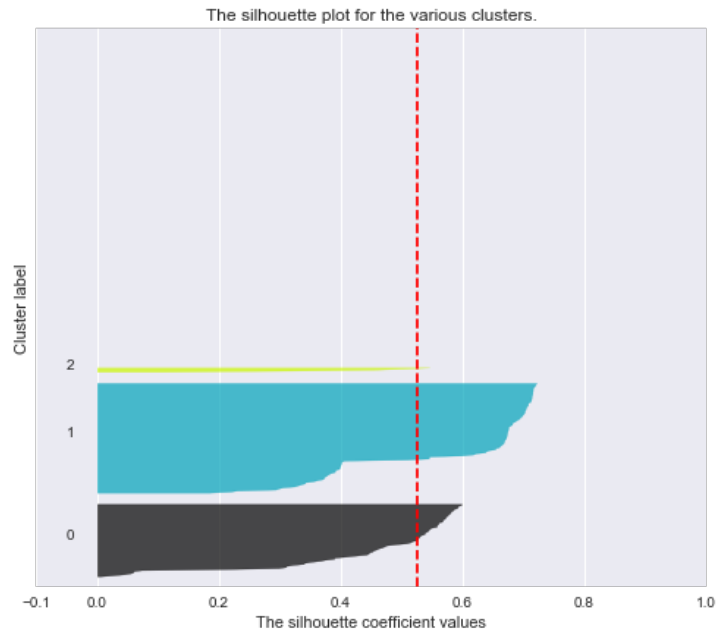


Figura 1: Métrica silhueta da clusterização

65 Usar o algoritmo de Random Forest da biblioteca sklearn para fazer uma regressão e prever quais variáveis socioeconômicas têm uma maior influência nas taxas de homicídios fazia parte da proposta inicial desse projeto, porém também não obtivemos êxito. Neste momento, foi decidido fazer uma análise dos clusters obtidos para poder entender melhor as diferenças entre esses grupos de municípios. Para isso, geramos a Figura 2 em que os valores 0, 1, 2 e 3  
70 representam as 4 variáveis obtidas acima respectivamente.

#### 4. Resultados

Inicialmente notamos que o cluster 2, o agrupamento que possui as maiores taxas em comparação com os outros dois clusters, é um pequeno grupo composto  
75 pelas cidades de Caruaru, Jaboatão dos Guararapes, Olinda, Paulista e Petrolina. Nessas cidades, a densidade populacional contribui para essas taxas, que por sua

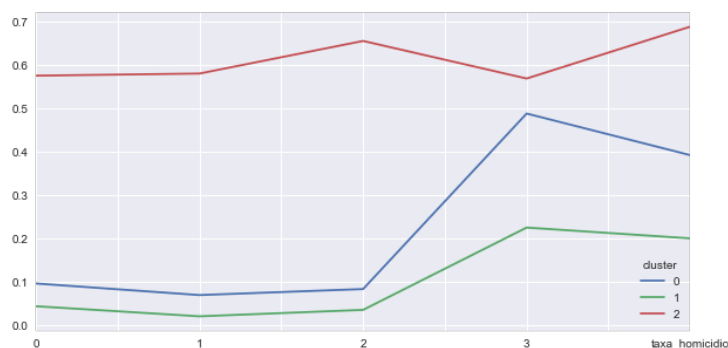


Figura 2: Análise dos cluster

vez também indicam que há uma maior desigualdade social entre as pessoas que moram nessas cidades do que entre as pessoas que moram nos municípios do cluster 0 e 1.

Os clusters 0 e 1, que agrupam municípios do interior do estado em sua maioria, possuem valores para as taxas analisadas menores do que os valores para o cluster 2 como era de se esperar, pois a densidade populacional nesses municípios é menor do que no cluster 2.

Curiosamente, os clusters 0 e 1 seguem parecidos quando se trata de educação e trabalhadores sem carteira assinada, porém se distanciam quando no valor mediano recebido por trabalhadores sem carteira assinada. No caso, o cluster 0 os trabalhadores sem carteira assinada recebem mais do que os do cluster 1 e também a taxa de homicídios do cluster 0 é maior do que a do cluster 1. Essa diferença entre salários pode indicar uma maior desigualdade social presente no cluster 1 que o faz ter uma taxa de homicídios maior.

Analisando separadamente cada uma das variáveis, modelamos os municípios através de uma regressão linear que indica uma forte correlação com a taxa de homicídios. Em todos os casos, podemos sugerir políticas públicas que foquem em alfabetizar jovens e adultos que estudaram de 1 a 3 anos. Para os municípios do cluster 0, poderiam ser aumentadas as fiscalizações dos empregadores que não assinam as carteiras de seus empregados, com o intuito de diminuir a quantidade de pessoas

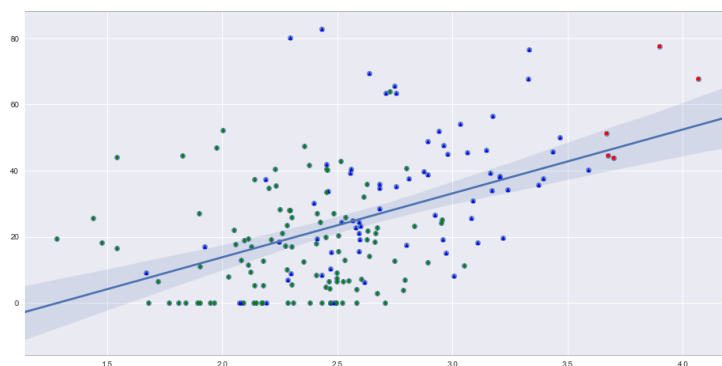


Figura 3: Taxa de homicídios x Pessoas de 10 anos ou mais de idade empregadas no trabalho principal da semana de referência subgrupo do emprego no trabalho principal trabalhadores domesticos categoria do emprego no trabalho principal sem carteira de trabalho assinada



Figura 4: Taxa de homicídios x Pessoa com mais de 10 anos de idade e que possuem de 1 a 3 anos de estudo



Figura 5: Taxa de homicídios x Pessoas de 10 anos ou mais de idade empregadas no trabalho principal da semana de referência subgrupo do emprego no trabalho principal demais empregados categoria do emprego no trabalho principal outros sem carteira de trabalho assinada



Figura 6: Taxa de homicídios x Valor do rendimento nominal mediano mensal do trabalho principal das pessoas de 10 anos ou mais de idade ocupadas na semana de referência com rendimento no trabalho principal posição na ocupação no trabalho principal empregados categoria do emprego no trabalho principal outros sem carteira de trabalho assinada

## 5. Conclusão

A nossa intenção é aprofundar a pesquisa, principalmente no cluster composto por Caruaru, Jaboatão dos Guararapes, Olinda, Paulista e Petrolina. O próximo passo é saber qual a motivação dos homicídios e qual a relação com as grandes cidades pernambucanas. Sabemos já que a nível nacional, “motivo fútil” e

tráfico de drogas são as principais causas dos homicídios e seria interessante estudar isso a nível do estado.

## 105 **References**

- [1] U. N. O. on Drugs, Crime, Global study on homicide 2013: trends, contexts, data, UNODC, 2013.