
Proposta do projeto final de Data Science - Violência Armada e suas tendências futuras nos EUA.

Lerisson Florêncio de Freitas
lff3@cin.ufpe.br,
Matheus Raz de Oliveira Leandro
mrol@cin.ufpe.br



UNIVERSIDADE
FEDERAL
DE PERNAMBUCO



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CIN - CENTRO DE INFORMÁTICA

Lerisson Florêncio de Freitas
lff3@cin.ufpe.br,
Matheus Raz de Oliveira Leandro
mrol@cin.ufpe.br

Proposta do projeto final de Data Science - Violência Armada e suas tendências futuras nos EUA.

Proposta de projeto levantada com o objetivo de construir um projeto como parte dos requisitos para a conclusão da disciplina IF1015 Intro a Ciência dos Dados.

Área de concentração: Ciência da Computação

Orientador: Dr. Renato Vimieiro

Proposta

1.1 Base

Nossa base de estudo foi coletada no repositório online Kaggle para aprendizagem de máquina: Dados de violência armada. Registro abrangente de mais de 260 mil incidentes de violência armada nos EUA entre 2013-2018

1.2 Informações sobre a base

Registrou-se mais de 260 mil incidentes de violência armada, com informações detalhadas sobre cada incidente, disponíveis no formato CSV. O arquivo CSV contém dados de todos os incidentes registrados de violência armada nos EUA entre janeiro de 2013 e março de 2018, inclusive.

Coleta dos dados:

- ❑ **Fase 1:** para cada data entre 1/1/2013 e 31/03/2018, um script Python consultou todos os incidentes que ocorreram naquela data em particular, depois digitalizou os dados e os escreveu em um arquivo CSV. Cada mês tem seu próprio arquivo CSV, com exceção de 2013, já que não foram registrados muitos incidentes a partir de então.
- ❑ **Fase 2:** cada entrada foi aumentada com dados adicionais que não podem ser visualizados diretamente na página de resultados da consulta, como informações do participante, dados de geolocalização etc.
- ❑ **Fase 3:** as entradas foram classificadas em ordem crescente e depois mescladas em um único arquivo CSV.

1.3 Proposta

Atualmente, faltam quantidades grandes e facilmente acessíveis de dados detalhados sobre a violência armada em geral. Para tal resolvemos focar na análise do contexto Norte Americano por ser um país que permite a regulamentação do uso de armas em alguns estados, gerando assim um índice de violência que demanda uma maior incidência de crimes cometidos com o uso delas.

Como mostrado na descrição da base estudada, relatórios como este e este mostram que Nikolas Cruz exibiu muitos sinais de alerta nas redes sociais antes do tiroteio, por exemplo. Para esse fim, a partir dessa base de dados, podemos identificar e levantar hipóteses que nos levem a entender melhor a necessidade de um reforço em tais áreas com maior índice de ocorrências desse viés, tais quais:

Hipóteses:

1. Em que cidades ou condados têm maior probabilidade de ocorrer um crime à mão armada?
2. Qual faixa etária de pessoas está presente na participação de tais crimes?
3. Que tipo de armas possuem maior incidência nos casos de violência?
4. Relacionar em quais distritos de votação para senadores ocorrem mais violência armada.