

Análise de Dados de Acidentes de Trânsito em Porto Alegre

Douglas Soares, Jônatas Clementino e Valdemiro Vieira

Centro de Informática
Universidade Federal de Pernambuco

1. Introdução

Segundo estudo feito pelo Observatório Nacional de Segurança Viária, cerca de 400 mil pessoas por ano, são afetadas por acidentes em trânsito no Brasil, dentre as quais 47 mil morrem. Sendo assim, acidentes de trânsito se mostra um tópico essencial a ser analisado e explorado.

2. Objetivo

O conjunto de dados principal que será analisado é o *dataset* que tem informações de acidentes de trânsito em Porto Alegre. Além dele, outras informações relacionadas ao trânsito em Porto Alegre serão coletadas de outros *datasets* como rede de ciclovias e localização de semáforos para poder ter uma noção maior de onde ocorram determinados acidentes e qual a relação desses equipamentos com a quantidade de acidentes. Seguem os links dos *datasets* que pretendemos utilizar:

- Acidentes de Trânsito em Porto Alegre de 2016:
<http://datapoa.com.br/dataset/acidentes-de-transito>
- Endereços dos Semáforos em Porto Alegre:
<http://datapoa.com.br/dataset/enderecos-dos-semaforos>
- Ciclovias Implantadas em Porto Alegre:
<http://datapoa.com.br/dataset/ciclovias-implantadas>
- Endereços das Lombadas em Porto Alegre:
<http://datapoa.com.br/dataset/enderecos-das-lombadas>
- Endereços dos Pardais em Porto Alegre:
<http://datapoa.com.br/dataset/enderecos-dos-pardais>

3. Metodologia

A partir da análise exploratória de todos os conjuntos de dados, serão avaliados algumas correlações entre eles, para inferir padrões e verificar hipóteses levantadas pelo grupo. Tais hipóteses são:

- Número de acidentes reduzem no final de semana.
- O centro da cidade é a região que ocorre mais acidentes.
- Existe uma relação entre a natureza do acidente com a quantidade de vítimas.
- Locais com equipamentos de fiscalização possuem menor índice de acidentes.
- Motos contribuem para o aumento de feridos.

- Ruas com mais equipamentos possuem menor quantidade de acidentes.

Dentre as informações que os conjuntos de dados escolhidos possuem, pode-se destacar as seguintes:

- Endereço, bairro, complemento e número do local do acidente;
- Veículos envolvidos no acidente (moto, automóvel, bicicleta, caminhão...);
- Data do acidente;
- Quantidade de vítimas envolvidas no acidente;
- Tipo de acidente (colisão, capotamento, engavetamento...);
- Localização de ciclovias e ciclofaixas, de semáforos e de equipamentos de fiscalização de trânsito.

Todos os dados foram coletados de *datasets* que a prefeitura de Porto Alegre disponibiliza. Como seria necessário ter informações de latitude e longitude dos acidentes, será utilizada alguma biblioteca que recebe o endereço descrito e retorna uma localização aproximada.

Na parte de análise exploratória de dados, será construído alguns gráficos e visualizações e serão calculadas medidas estatísticas para se ter uma noção maior do que está acontecendo com os dados e validar ou refutar as hipóteses propostas anteriormente. Pode-se citar como objetivos dessa etapa descobrir coisas como:

- Quantidade de acidentes por bairros;
- Quantidades de cada tipo de veículos envolvidos nos acidentes;

Além disso, para ter uma visão melhor dos dados, será construída uma visualização mostrando de maneira geográfica os dados dos *datasets* informados anteriormente para ver se pode observar algum comportamento que não foi percebido apenas olhando para os dados de maneira crua. Como todos os acidentes contém apenas endereços informando ruas e bairros, será feita alguma aproximação para descrever esses locais no mapa.

Também serão implementadas técnicas de aprendizagem com o objetivo de extrair algum tipo de padrão nos dados, ou relacionar variáveis que exercem graus de influência sobre o decréscimo da ocorrência de acidentes.