

Taxas de Ocorrências Criminais

Proposta de Indicadores e Dashboards

Fonte: <http://dados.gov.br/dataset/sistema-nacional-de-estatisticas-de-seguranca-publica>

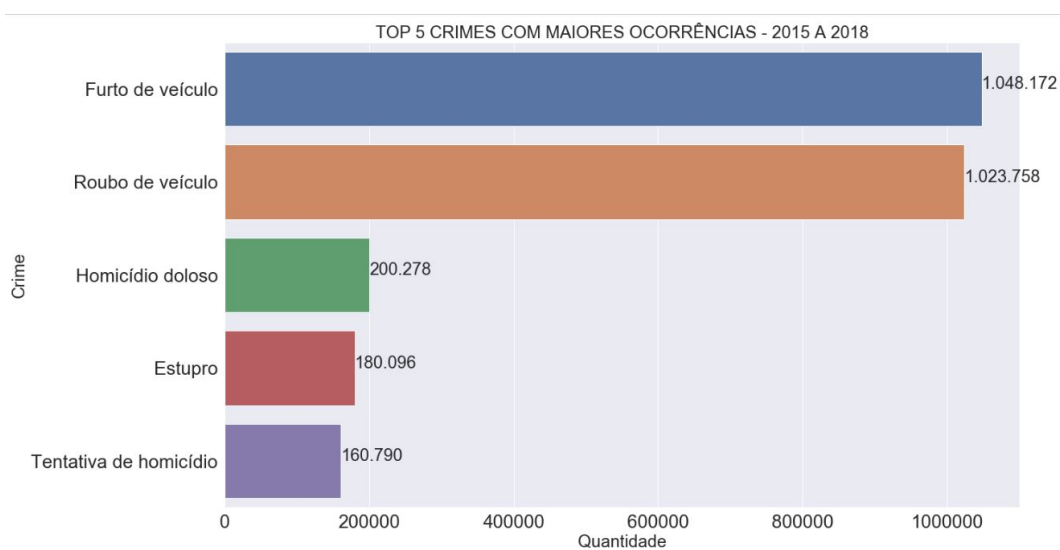
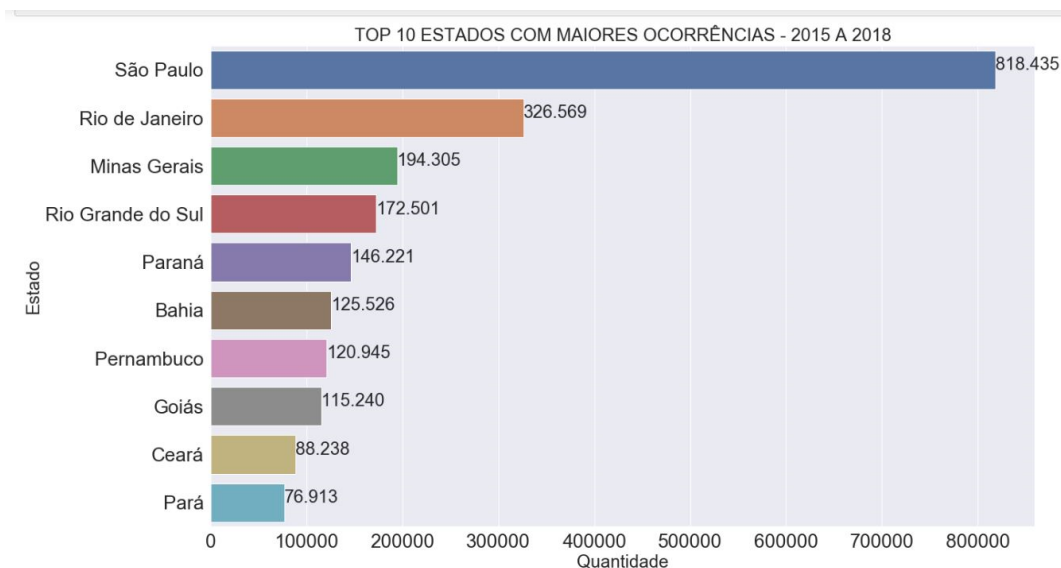
Descrição do documento	
Contém a especificação dos passos realizados para análise exploratória e proposta de criação de gráficos que demonstre as ocorrências criminais no período de 2015 a 2018 em todos os estados brasileiros.	
Dados da demanda	
Projeto:	Ocorrências Criminais
Analista:	Raquel Lima Guimarães
Data:	07/01/2020

Este documento visa demonstrar o tratamento de dados realizado, sugestão de indicadores contendo as estatísticas de crimes nos Estados, como também apontar possíveis erros de cadastro e alimentação de dados de forma manual, visando conscientizar as unidades responsáveis para a automatização da coleta de dados, visando a agilidade na publicação de dados mais recentes.

Conforme publicação no Portal Brasileiro de Dados Abertos da referida fonte de dados, a partir de 10 de janeiro de 2015 os dados de múltiplas fontes foram consolidados e disponibilizados em uma única Base Nacional. Desta forma, podemos desconsiderar dados de anos anteriores a 2015, por não apresentarem dados suficientes para nossa análise, conforme demonstrado nas consultas no item "Estatísticas de Ocorrências" no arquivo de script. Na parte inferior da página do Portal, em "Informações Adicionais", consta a última atualização de dados em 30 de março de 2019, portanto, podemos desconsiderar também os dados de 2019 por estarem incompletos.

Todo o tratamento de dados, bem como a criação dos gráficos foram desenvolvidos utilizando a linguagem Python e suas bibliotecas.

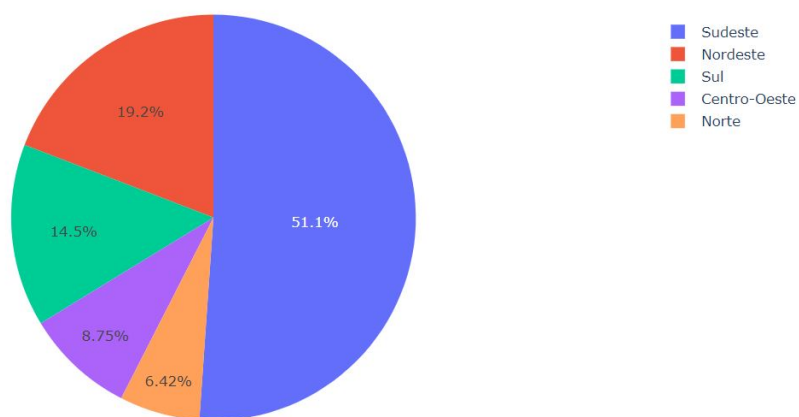
Sugestão de Indicadores



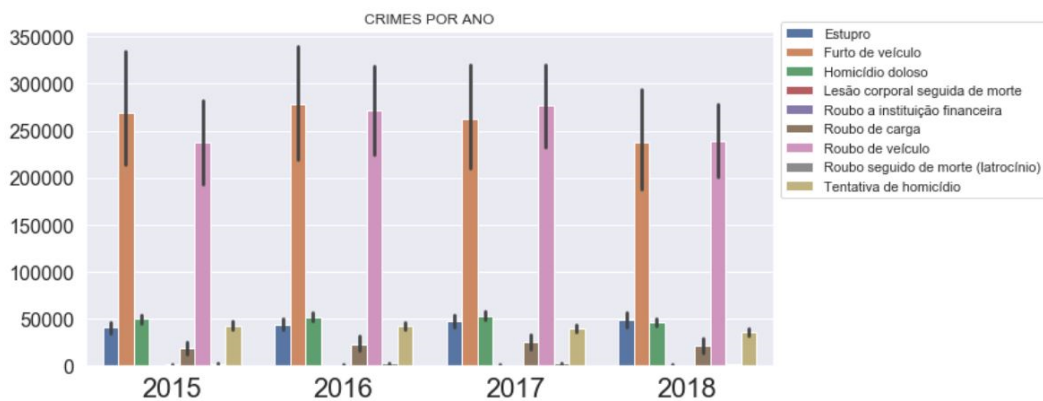
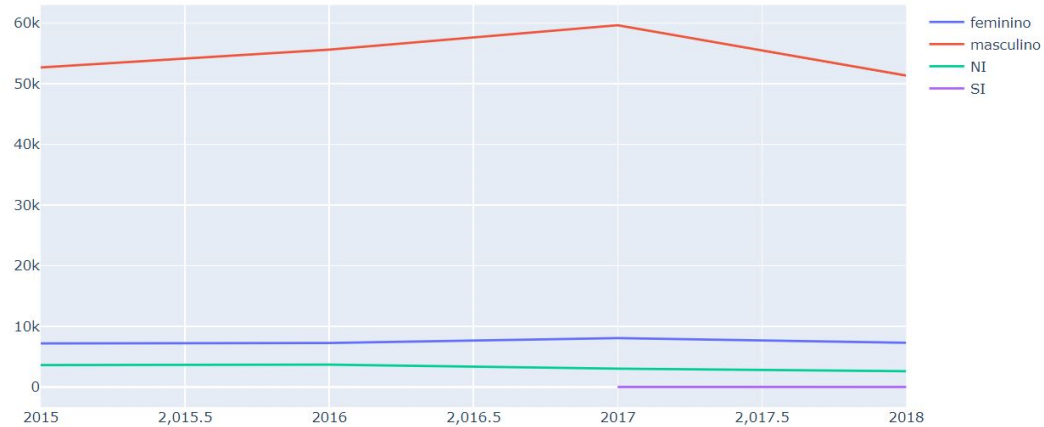
Total de Crimes no Período



Ocorrências Por Região



Vítimas Por Sexo



Resumo dos passos realizados no tratamento dos dados

- 1- ler arquivos;
- 2- converter dataset em um dataframe do pacote Pandas;
- 3- ler dados da abas "Ocorrências" e da aba "Vítimas" de forma separada;
- 4- verificar a quantidade de registros e registros nulos;
- 5- selecionar as colunas de interesse;
- 6- tratar coluna "mês/ano" separando em duas colunas;
- 7- renomear colunas;
- 8- ler dataset de região para adicionar o nome da região aos dataframes;
- 9- verificar nomes de região;

- 10- renomear colunas;
- 11- comparar nomes de estados com os estados dos dataframes;
- 12- fazer o merge para adicionar região aos dataframes;
- 13- criar coluna de ID do mês para os meses ficarem ordenados nos gráficos;
- 14- criar coluna com nomes dos meses;
- 15- tratar coluna do valor considerado (quebra de linha, valores nulos, ponto, traço) de cada dataframe;
- 16- verificar estatísticas de anos e estados;
- 17- criar ID de crimes distintos, caso seja necessário exportar para banco de dados;
- 18- verificar crimes duplicados fora do padrão;
- 19- Foi verificado que no estado da Paraíba existem registros duplicados para o mesmo crime, mês e ano. Esses valores foram somados e as linhas fora do padrão foram deletadas.
- 20- criar ID de crime distinto no dataframe de vítimas;
- 21- verificar duplicidade de crimes distintos. Foi verificada duplicidade no estado do Mato Grosso do Sul. Atualizar a descrição colocando o 'd' minúsculo para ficar no padrão dos demais estados.
- 22- criar dataframe para os top 10 estados com maiores ocorrências;
- 23- criar dataframe para os top 5 crimes com maiores ocorrências;
- 24- criação dos gráficos