

# Tarefa 1

## Tarefa 1 - Visualização de Dados

Aluno: Daniel de Miranda Almeida, Matrícula: 241708065

### Pergunta

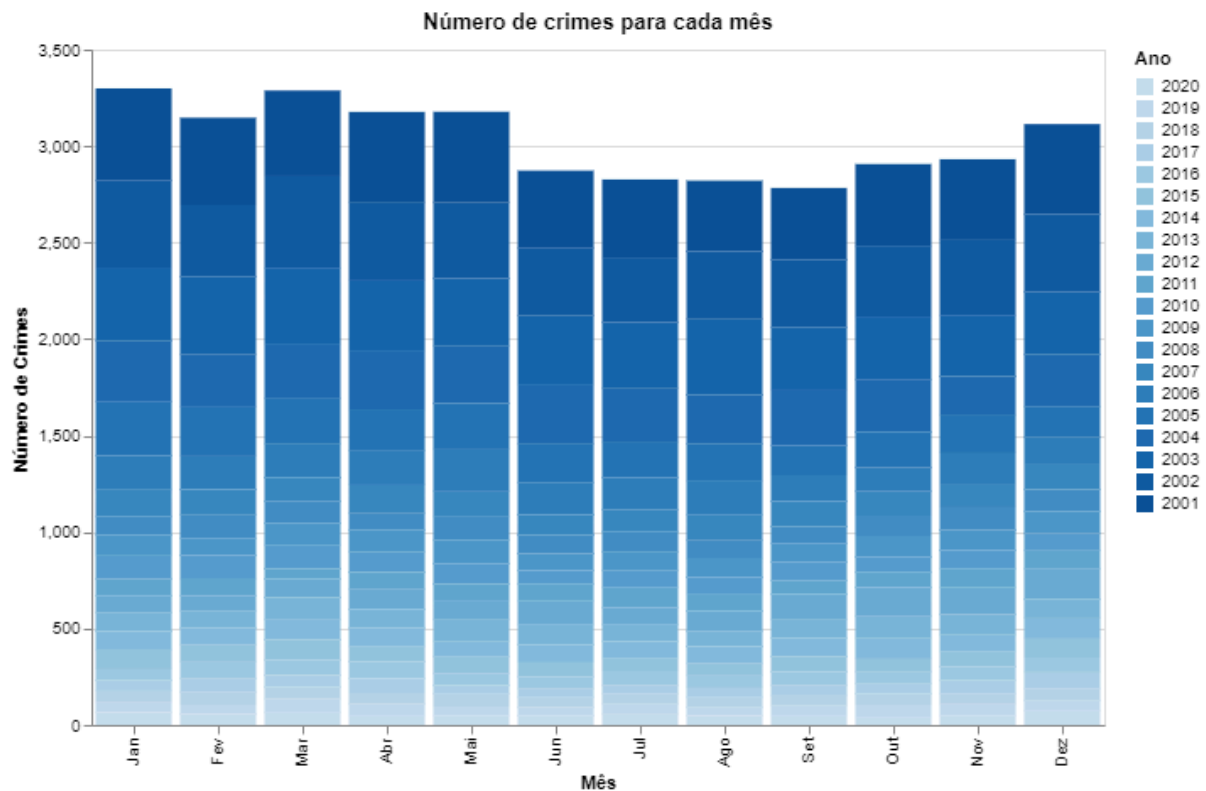
O objetivo nessa tarefa era de responder à seguinte pergunta: Algum mês é menos violento que os outros? Para isso, pensei em dois gráficos principais: um que plotasse os valores agregados de crimes para cada mês e para todos os anos e um que trouxesse as distribuições dos números de crimes para cada mês de maneira agregada para todos os anos.

### Processo

Apliquei algumas pequenas transformações nos dados para facilitar sua manipulação e percebi que os números totais de crimes para os últimos 5 meses de 2021 estavam nulos para todas as cidades, de forma que, para não influenciar nos meus gráficos (fazendo com que os 7 primeiros meses tenham mais dados), limitei a base da dados aos registros até o ano de 2020.

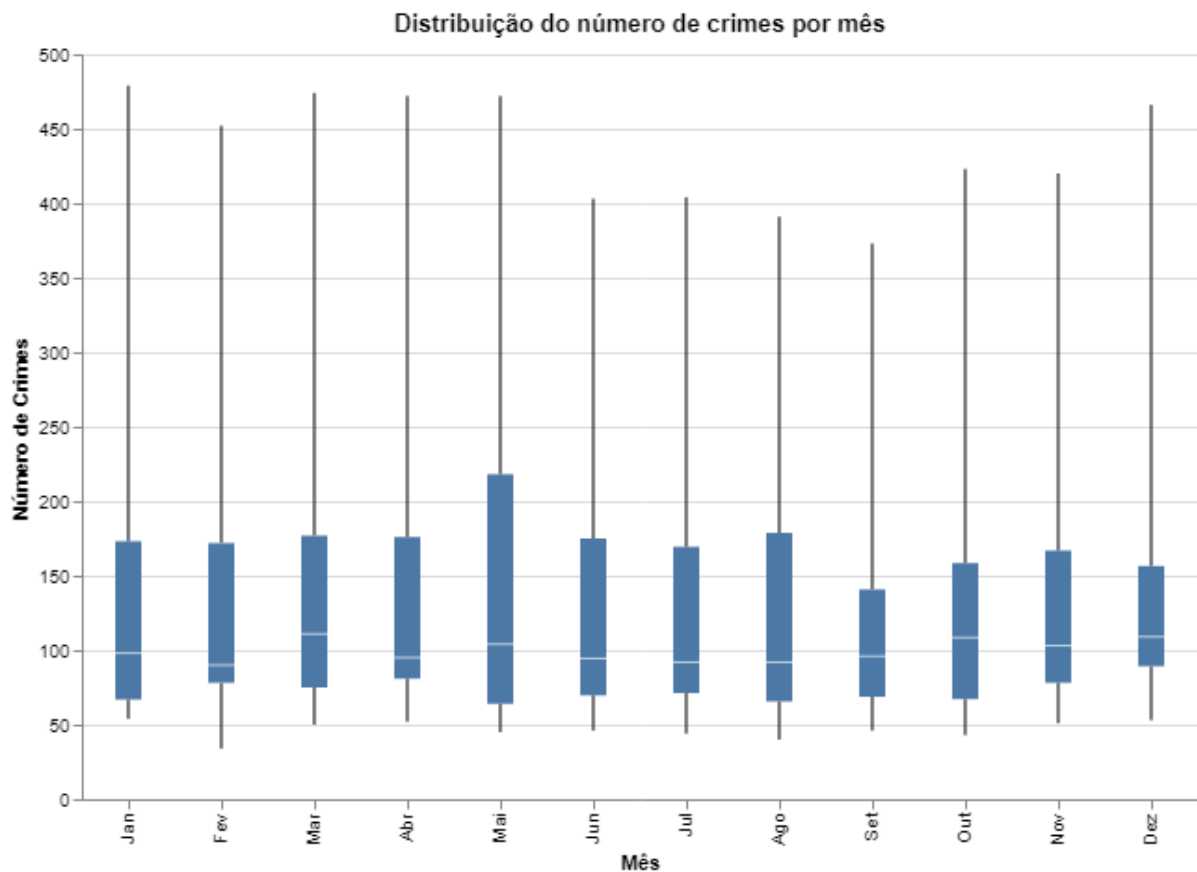
Houve outro problema: a cidade de São Paulo é a maior do Brasil e acaba tendo números desproporcionais em relação às outras. Por isso, escolhi como abordagem reduzir o escopo da minha pergunta apenas para a cidade de São Paulo.

### Gráficos



O primeiro gráfico traz a soma do número de crimes para cada mês agregada para todos os anos (com exceção de 2021). Nele podemos perceber que, somando os números de crimes para cada mês ao longo de todos os anos, os meses de julho, agosto e setembro apresentam os menores totais de crimes, com o menor deles sendo setembro, com um total de 2783 crimes, contra 3299 crimes no mês de janeiro, uma diferença significativa de 516 crimes entre os dois meses, ou 25 crimes por ano.

Além disso, podemos perceber pelo gráfico também que os números de crimes num geral foram diminuindo ao longo dos anos, com os comprimentos dos últimos anos sendo bem menores que os dos primeiros anos.



O segundo gráfico traz agora boxplots para mostrar a distribuição do número de crimes para cada mês, onde podemos ver, pela distribuição para o mês de setembro, que este tem uma tendência a ter menos crimes, uma vez que todas as distribuições tem medianas bastante parecidas mas o terceiro o quartil e maior valor maiores do que os respectivos valores para o mês de setembro.

## Justificativa das escolhas para cada gráfico

Para o primeiro gráfico, escolhi optar para uma representação em barras dos gráficos, com o número de crimes sendo representado pelo comprimento das barras, me alinhando às ideias em uma das *lectures* da semana 2, que defende que mudanças em dados uni-dimensionais (número de crimes) devem ser traduzidas em mudanças no design também uni-dimensionais (comprimento das barras). Além disso, a escolha de um gradiente de cor azul indo de uma cor mais clara para uma mais escura ajuda a destacar no gráfico os comprimentos dos retângulos ao longo dos anos, o que explicita o fato de que os números de crimes no geral foram diminuindo com o passar dos anos.

Já para o segundo gráfico, optei por um boxplot por se tratar de um gráfico muito efetivo e que comunica muito rapidamente a ideia de como são os

números de crimes para diferentes meses. Apesar de ser um gráfico que demanda mais conhecimento específico para o seu entendimento, como o público da minha visualização são pessoas com alto grau de instrução no assunto de visualização de dados, julguei ser um gráfico adequado para transmitir minhas ideias.

## **Comparação com a tarefa 0**

A primeira tarefa acabou não tendo muito influência nessa, visto que optei por responder uma pergunta diferente. Ainda assim, a escolha do boxplot e do gráfico de barras foi igual nas duas tarefas, e de fato, com o que aprendi fazendo a primeira tarefa, junto dos textos lidos, reforcei minha noção da importância desses dois tipos de gráficos na transmissão de informações de forma clara.