

Assignment 01: Projeto de Visualização

Visualização de Dados

Felipe Marques Esteves Lamarca

No Assignment 0, questionávamos se havia um padrão de variação na criminalidade ao longo dos meses do ano. Na ocasião, foi feito um compilado de 3 gráficos. Um apresentava, em linha, o número total de homicídios ao longo dos meses, e outros dois em barra que apresentavam a variação de um mês em relação ao anterior — o primeiro considerando os dados gerais, e o segundo separando as cidades em blocos individuais (facets). A ideia que justificava o plot dos percentuais de variação era que pontos mais acentuados poderiam significar, na prática, mudanças na dinâmica da criminalidade ao longo do ano.

A questão, agora, se mantém a mesma: **há um padrão de variação na criminalidade ao longo dos meses?** No caso do gráfico estático desenvolvido para este assignment, mantive a utilização dos gráficos que representam variação em relação ao mês anterior, partindo do fato de que essa é uma boa forma de observar mudanças de tendência ao longo dos meses. No entanto, o *dataset* da tarefa anterior dava conta dos dados de apenas 6 cidades do estado de São Paulo, o que permitia a elaboração e análise de facets sem grandes dificuldades. Agora, com a apresentação de 11 cidades diferentes, optei por representar na parte superior da visualização, ao invés de facets, um gráfico de barras que utilizasse alguma medida dos dados — no caso, a média de variação percentual.

Além disso, foram adicionadas outras três layers ao gráfico, que adicionam, para cada mês, um ponto para a maior variação de acréscimo, um ponto para a maior variação de decréscimo e uma linha que conecta os dois pontos. Com essas layers adicionais, é possível observar que há cidades em que a variação (positiva ou negativa) foi muito maior (em módulo) do que o que foi observado na média. Assim, possibilita a inferência de informações gerais, mas também particulares dos dados¹.

Já no caso da parte inferior da visualização, optou-se por uma abordagem mais geral dos dados do *dataset*. Conforme indicam os títulos dos eixos, a variação percentual foi calculada em cima do número total de homicídios por mês, desconsiderando informações particulares das cidades. O objetivo é justamente comparar informações gerais e particulares, e facilmente permite avaliar que a média de variação de homicídios nas cidades acompanha a variação geral na maioria dos meses. Uma marca visual que optei por reforçar no caso dessa visualização foi a linha mais escura demarcando o

¹ Naturalmente, como se trata de uma visualização estática, só seria possível saber quais cidades tiveram maior variação de acréscimo ou decréscimo se fosse adicionada uma camada textual ao gráfico. Em uma visualização iterativa, essas informações poderiam ser apresentadas de forma mais simples.

$y = 0$, facilitando que o leitor identifique quando a variação em relação ao mês anterior foi positiva ou negativa. Se essa informação é apreendida de forma praticamente imediata no gráfico de barras, esse adicional é bem-vindo para facilitar a leitura das variações no gráfico de linha.

