# Trabalho 1

INF-0614-010 - Visualização de Informação

Grupo:

Evandro Santos Rocha

Laíssa Pacheco de Oliveira

Rafael Dantas de Moura

O objetivo deste trabalho é usar a ferramenta de visualização para analisar dados multivariados.

**1. Explorem o conjunto de dados indicado usando técnicas de visualização. Neste momento, a intenção é conhecer o conjunto de dados, suas tabelas e variáveis. (Nada a entregar sobre esta tarefa).**

**2. Assumam o papel de usuário interessado em obter informações relevantes deste conjunto de dados. Elaborem perguntas que vocês teriam interesse em responder, baseadas nesses dados.**

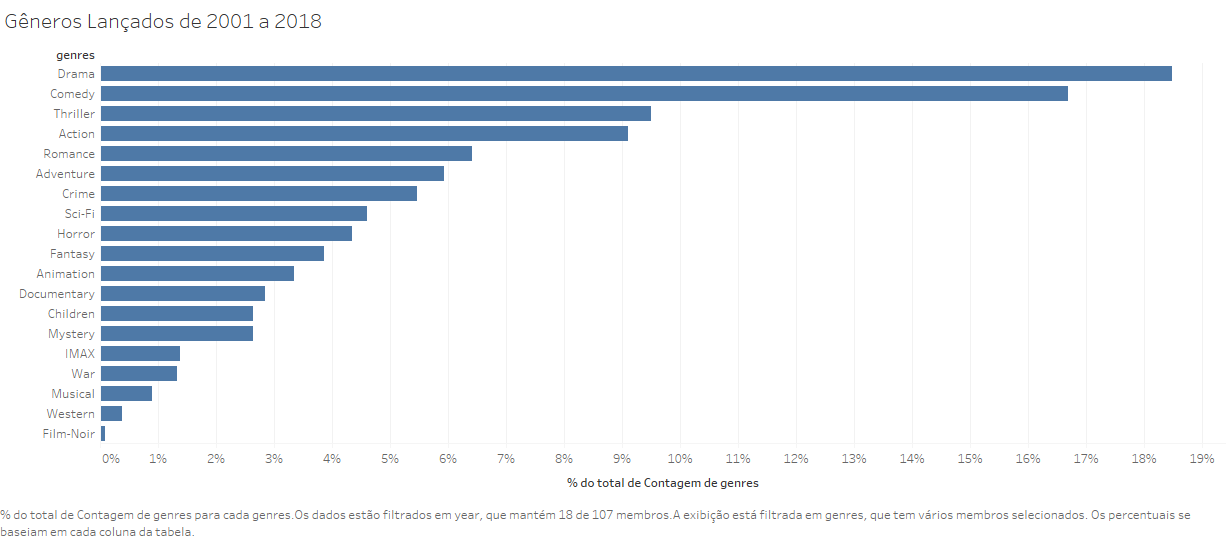
* Quais gêneros de filmes mais produzidos (por ano de produção) nas últimas duas décadas (2001-2018)?
* Dos 3 três gêneros dos filmes mais produzidos, qual foi a média de avaliação? Qual a média de avaliação para os 3 gêneros menos produzidos?
* Será que os gêneros menos produzidos possuem notas médias baixas? Seria um possível motivo para número baixo de produções?

**3. Escolham e usem 3 gráficos interativos para representar o conjunto de dados, explicitando que informações sobre os dados esses gráficos trazem com relação a essas perguntas. Para cada gráfico escolhido:**

**(a) Indiquem o tipo de gráfico e os mapeamentos visuais que foram feitos nele. P. ex.: Eixo X -> Alunos, Eixo Y -> Disciplinas, Cor -> Notas, Listbox -> Ano\_de\_ingresso.**

**(b) Expliquem as informações conseguidas na análise do gráfico. Usem uma ou mais figuras para comprovar e apresentar essas informações.**

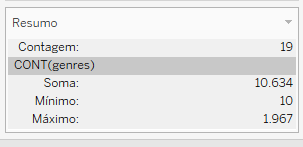
Filtrando os dados para ano de produção do filme entre 2001 e 2018 e removendo todos que não há gênero, podemos observar que os gêneros mais produzidos nas últimas duas décadas foram drama com 18,50% seguido por comédia com 16,69%. O terceiro lugar, gênero thriller, possui metade da representatividade em comparação ao primeiro (drama), com 9,50%.

****

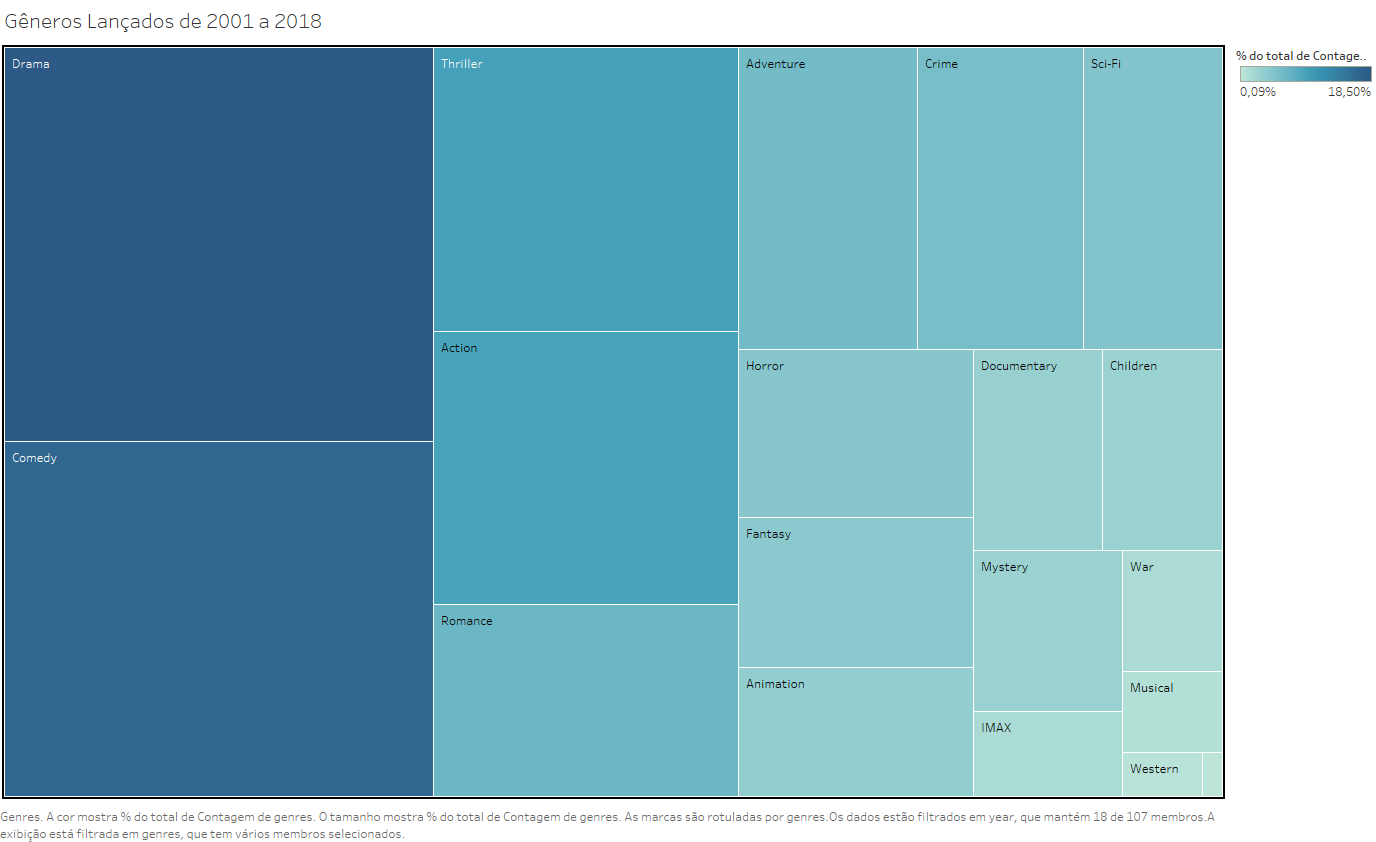
Já os 3 gêneros menos representativos foram:

* Film-Noir (0,09%)
* Western (0,39%)
* Musical (0,89%)

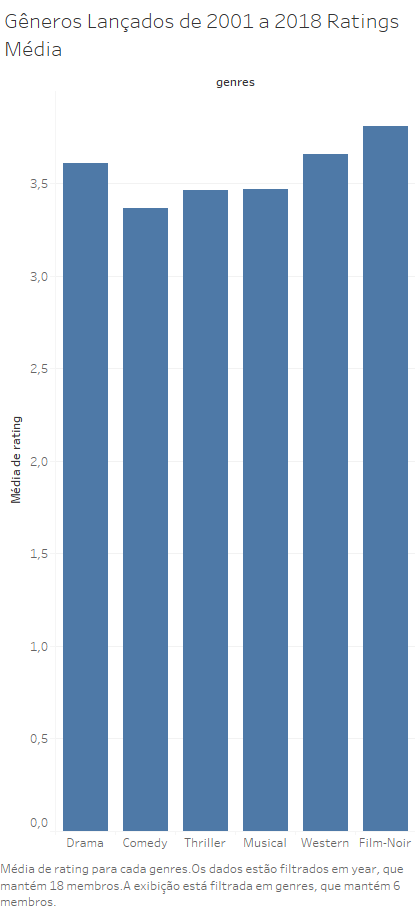
O conjunto de filmes entre os anos citados e excluindo os que não possuem gênero, possui 10.634 exemplos, conforme podemos observar no resumo abaixo:



Na imagem abaixo podemos visualizar a quantidade de filmes proporcionalmente ao conjunto selecionado dos dados:



Separamos o conjunto de filmes com os 3 gêneros mais produzidos e os 3 gêneros menos produzidos para avaliar o rating médio. Queríamos observar se havia uma relação entre os gêneros de filmes com menor produção com avaliações mais baixas em comparação com os de maiores números de produção. Por exemplo: o gênero film-noir representou apenas 0,09% do total de produção de filmes entre 2001 e 2018, queríamos observar se a avaliação média para esse gênero era baixa. Entretanto não era.



Observamos que o gênero Film Noir teve uma nota média superior ao gênero Drama. Film noir é uma expressão francesa designada a um sub gênero de filme policial, derivado do romance de suspense influenciado pelo expressionismo alemão, o qual teve o seu ápice nos Estados Unidos entre os anos 1939 e 1950. Embora o conjunto selecionado seja para as últimas duas décadas (2001-2018), as notas para esse gênero foram relativamente altas. Então podemos supor que embora haja uma produção de filmes bem menor para esse gênero, o seu público é de nicho, o que pode estar mantendo o rating alto.