



Anterior

Apresentação

Próximo

Infográfico



## Desafio



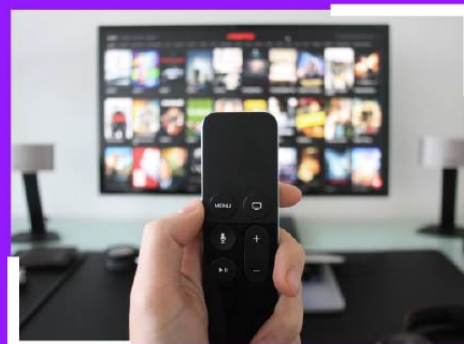
No cenário atual, onde existe uma explosão de conteúdo *stream*, empresas de conteúdo digital estão cada vez mais interessadas em analistas de dados, para o entendimento do perfil dos seus consumidores.



# STREAMING DE FILMES

Você foi contratado por uma empresa de conteúdo digital baseado em *stream* de filmes que está interessada em analisar quais são os 30 filmes mais parecidos com um filme (FM) específico.

Além disso, o proprietário da empresa de conteúdo *stream* também está interessado em identificar os filmes considerado "galinha dos ovos de ouro", que são os 4 filmes que têm correlação com todos os outros a partir da lista de filmes correlacionados a FM.



Baseado em seus conhecimentos, descubra:

- Dado um filme *FM* que um respectivo usuário gostou, quais 30 títulos de filmes você pode recomendá-lo?
- Considerando o conceito da *galinha dos ovos de ouro*, diga quais são os 4 filmes que apresentam maior correlação com todos os outros, a partir da lista de filmes correlacionados a *FM* selecionado pelo usuário?

Nos anexos a seguir, seguem dois arquivos importantes ao trabalho:

O arquivo *id\_filmes\_titulos* contém o ID do filme com seu respectivo título.

[Clique aqui para acessar](#)

Já o arquivo *filmes\_ranking.data* é o conjunto completo de dados de usuário com 100.000 avaliações de 943 usuários em 1.682 filmes. Cada usuário classificou pelo menos 20 filmes. Usuários e filmes são numerados consecutivamente a partir de 1. Os dados são ordenados aleatoriamente. Esta é uma lista separada por tabulação de ID do usuário | ID do filme | classificação | *timestamp*.

[Clique aqui para acessar](#)

Lembre-se de explicar como chegou ao resultado.





Anterior

**Apresentação**

Próximo

**Infográfico**



Anexe seus arquivos:

Extensões permitidas: .png, .txt, .pdf, .doc, .docx, .jpg, .xls, .xlsx, .zip, .rar  
Tamanho máximo de arquivo: 2MB

Enviar resposta

