

MC536 & CHILL

Um estudo sobre as mídias disponíveis em
plataformas de streaming selecionadas

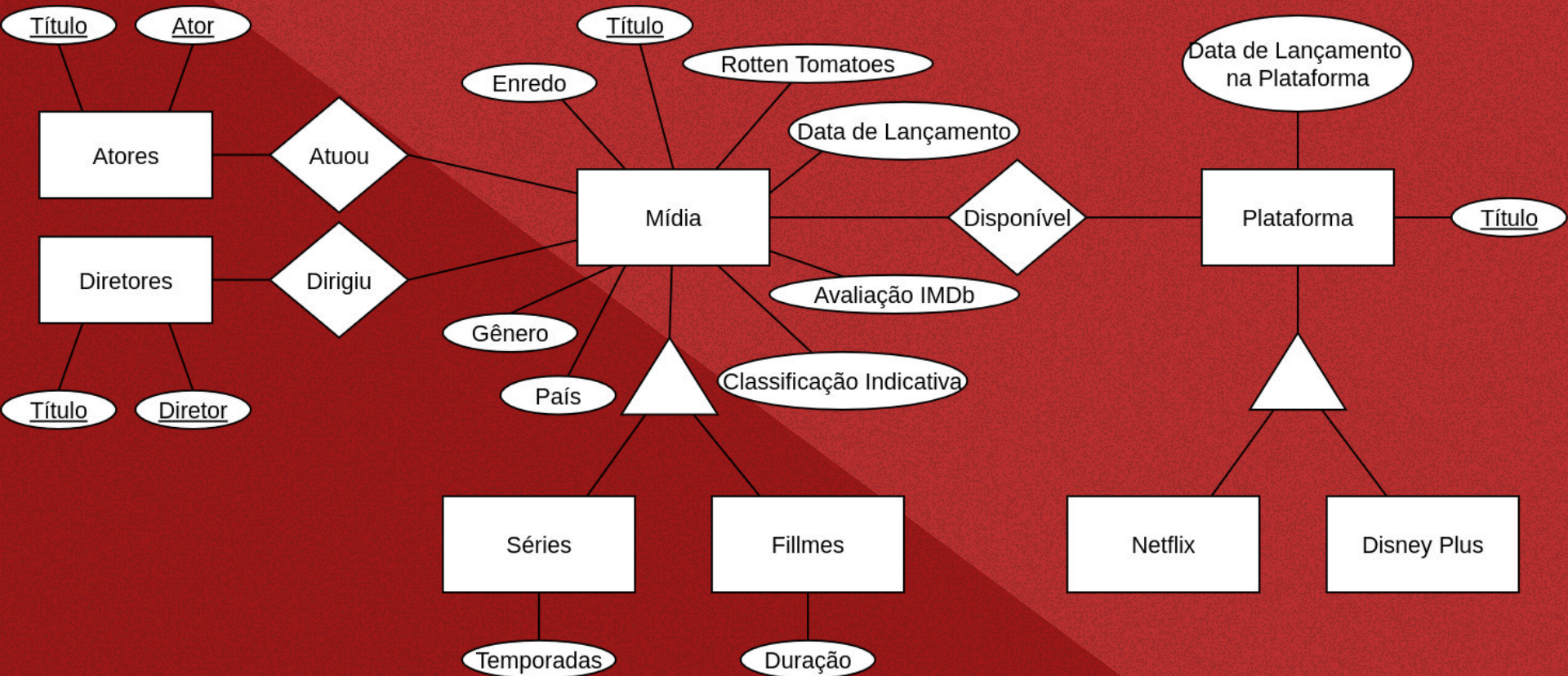
DESCRIÇÃO DO TEMA

Com a grande difusão dos serviços de streaming o acesso a filmes e séries foi amplamente democratizado. Em período de pandemia, em que as pessoas passam mais tempo em casa, esses serviços se popularizaram ainda mais, e estão cada vez mais presentes nas vidas das pessoas. Tendo em vista as inúmeras obras disponíveis nessas plataformas, pretendemos montar um dataset com informações de filmes e séries, com dados provenientes de diferentes fontes e que passarão por um tratamento.

Esperamos disponibilizar informações como: título, ano de lançamento, classificação indicativa, avaliação IMDb, Rotten Tomatoes, data de adição em cada plataforma e em quais plataformas a obra está disponível. Ainda, podemos disponibilizar informações diretamente relacionadas a obra, como: atores, diretores, enredo, gênero e duração.

A princípio pretendemos analisar a disponibilidade na Netflix e Disney Plus, podendo adicionar outras plataformas de acordo com a evolução do trabalho .

ESBOÇO DO MODELO CONCEITUAL DO DATASET



ESBOÇO DO MODELO RELACIONAL DO DATASET

Filmes(título, atores, diretores, enredo, gênero, data de lançamento, duração, plataforma, país, classificação indicativa, avaliação IMDb, Rotten Tomatoes)

Séries(título, atores, diretores, enredo, gênero, data de lançamento, temporadas, plataforma, país, classificação indicativa, avaliação IMDb, Rotten Tomatoes)

Atores(título, ator)

Diretores(título, diretor)

Netflix(título, data de lançamento na plataforma)

Disney Plus(título, data de lançamento na plataforma)

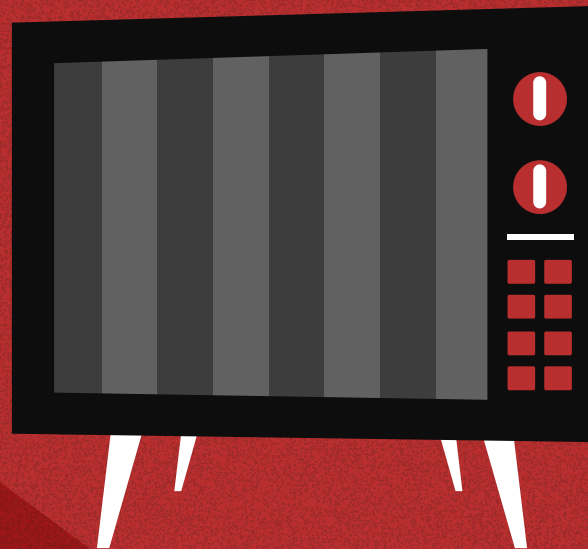
MODELOS LÓGICOS

Pretendemos explorar um modelo relacional (contendo as informações em geral) e um modelo de grafos (para as relações de filme e ator)



PERGUNTAS A SEREM RESPONDIDAS

- Quais filmes têm melhor/pior avaliação
- Qual tempo entre a estréia do filme e o lançamento na plataforma?
- Quais atores têm as melhores avaliações nos filmes em que participaram?
- Quais diretores têm as melhores avaliações em seus filmes?
- Existe alguma relação entre duração e avaliação?
- Existe alguma relação entre gênero e avaliação?
- Quantos filmes/séries foram feitos em cada país?
- Quais atores participaram de mais filmes?



FONTES DE DADOS

Candidatos:

- [TV shows on Netflix, Prime Video, Hulu and Disney+ | Kaggle](#)
- [Movies on Netflix, Prime Video, Hulu and Disney+ | Kaggle](#)
- [Disney Plus Movies and TV Shows | Kaggle](#)
- [Netflix Movies and TV Shows | Kaggle](#)
- [Amazon Prime Movies | Kaggle](#)

Operações:

- Iremos extrair dados desses datasets
- Vamos tratar esses dados somente com as informações úteis e das plataformas que, de fato, iremos utilizar
- Vamos integrar esses dados em tabelas novas, com informações de pelo menos 2 datasets
- Pretendemos tratar os atributos multivalorados criando novas tabelas
- Pretendemos transformar parte desses dados em grafos
- Pode ser necessário tratar dados faltantes

