



**Universidade do Minho**  
Escola de Engenharia

Universidade do Minho  
Mestrado Integrado em Engenharia Informática

**Perfil EC**

# **Análise de Dados**

## **TP 4**

Novembro 2020

Tânia Filipa Amorim Rocha  
A85176

# 1 Introdução

Este relatório foi desenvolvido no âmbito do quarto guião prático da UC análise de dados. Será discutido e apresentado detalhadamente o processo de construção de um modelo mySQL e implementações extra com base em dados fornecidos num ficheiro csv sob o contexto de séries e filmes adicionados à plataforma de streaming Netflix. Para além disto, será também feita uma análise com modelos estatísticos fornecidos pelo software Tableau.

## 2 Análise e Construção

### 2.1 Modelo

Tendo em conta os dados descritos no csv, foi optado por haver uma tabela de factos que retrata as informações gerais do modelo, isto é, contem um id relativo ao *show* da Netflix e tem informações sobre o mesmo, informações IMDB, datas pertinentes e informações adicionais.

O modelo desenvolvido tem o seguinte aspecto:

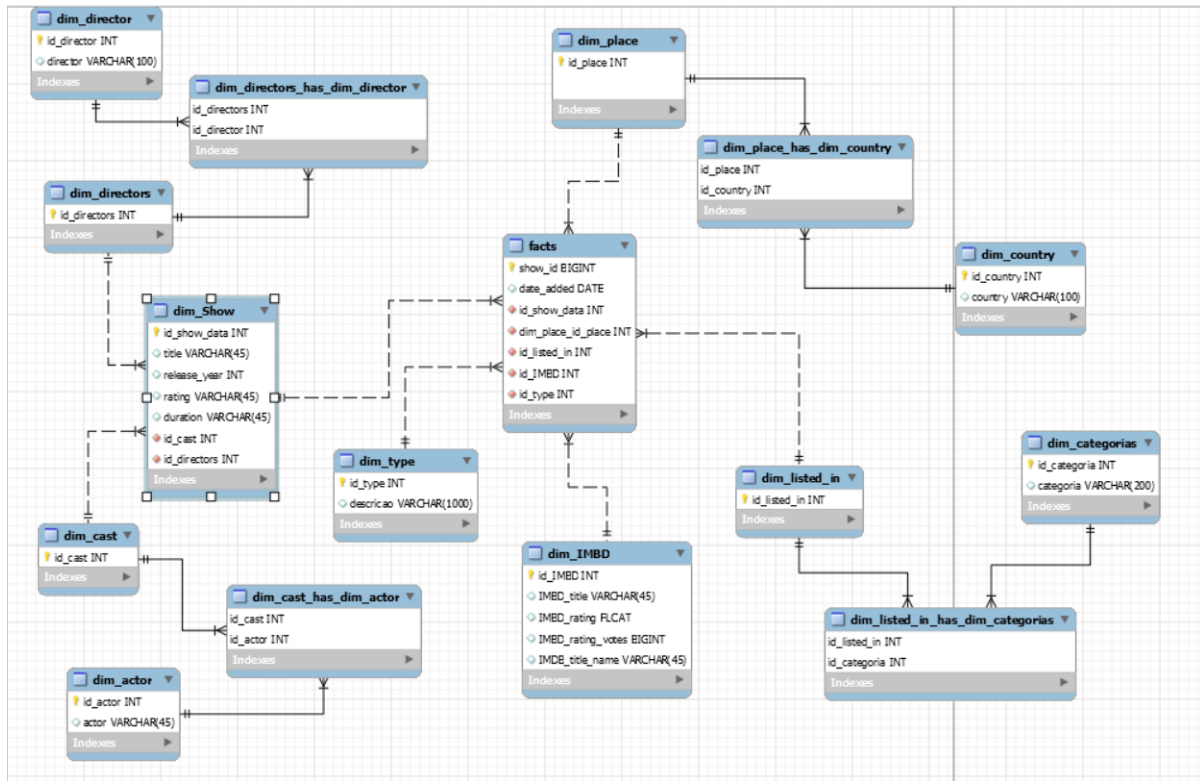


Figura 1: Modelo

Posteriormente foi efetuado o povoamento deste mesmo modelo.

## 2.2 Tableau

Com o objetivo de familiarizar com o tipo de software que permitem análise de dados, foi utilizado o Tableau recorrendo aos dados povoados no modelo em questão. Primeiramente foram feitas as conexões entre as tabelas ficando com o seguinte aspecto:

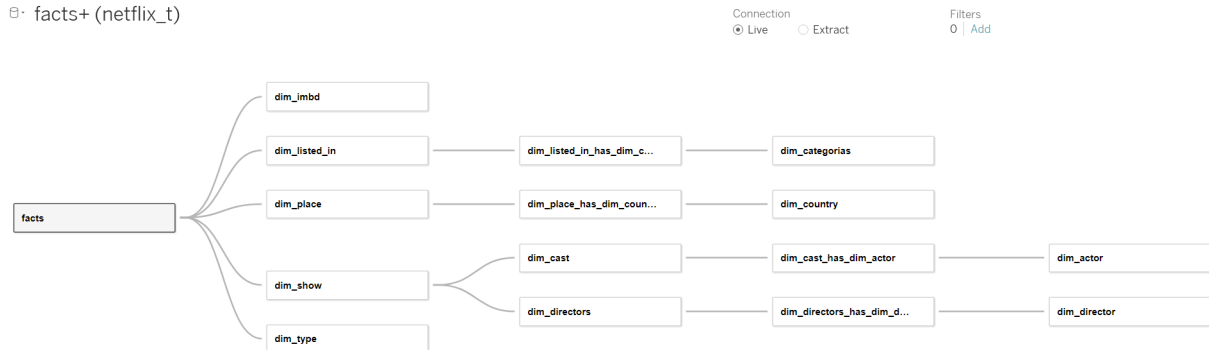


Figura 2: Modelo snowflake

Após uma pequena análise do funcionamento do software foram desenvolvidas as seguintes queries:

- Quantidade de filmes e séries por data de estreia:

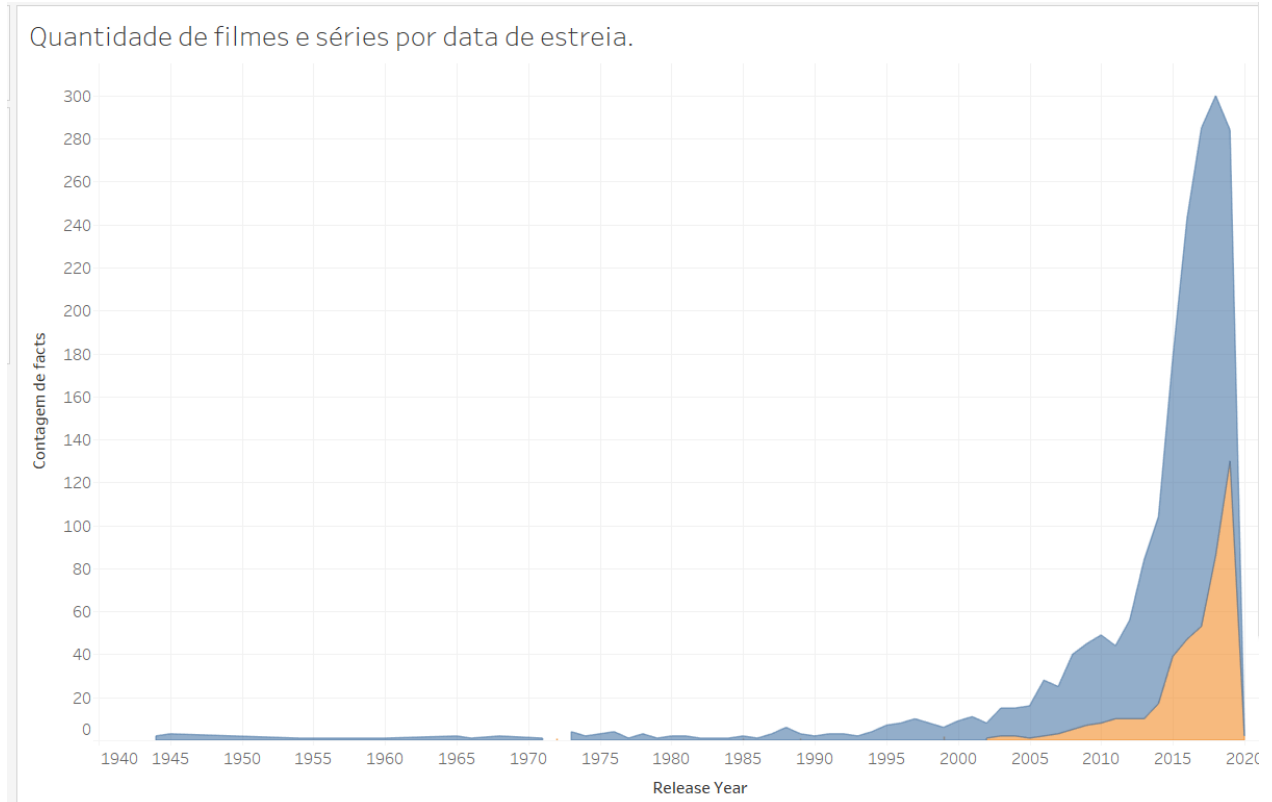


Figura 3: Resultados

- Número de votos ratingIMBD por categoria:

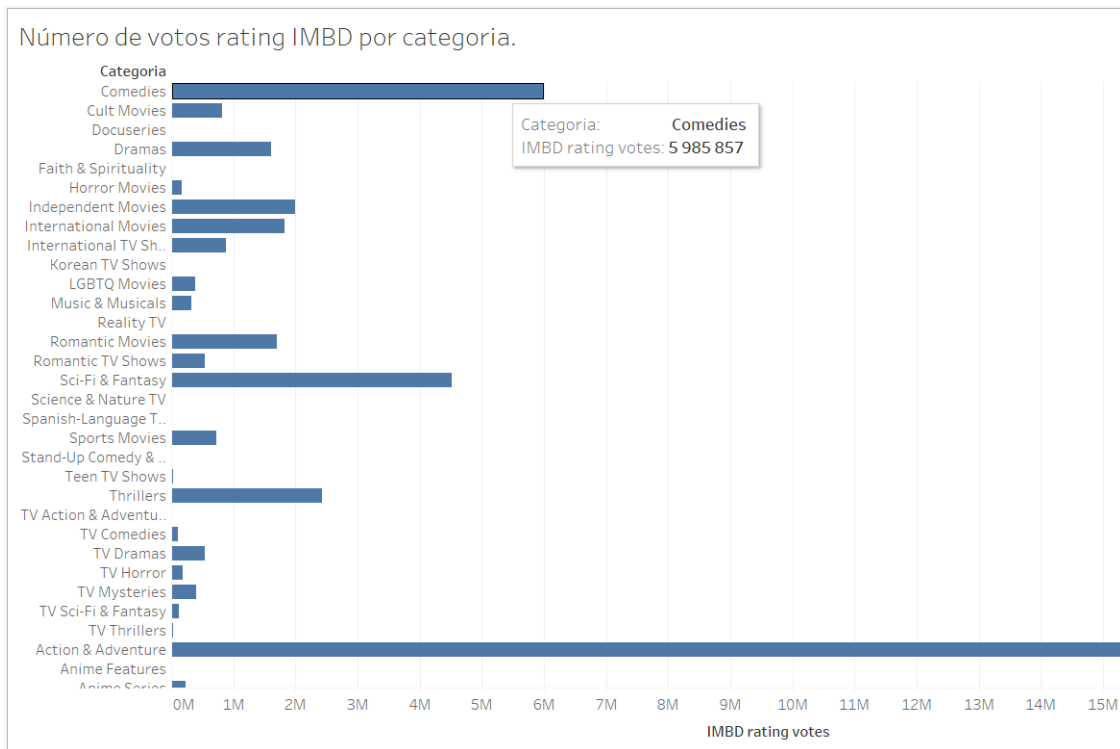


Figura 4: Resultados

- Categorias e Ratings por País:

Categorias e ratings por País.

Country	Rating	Categoria														
		Come..	Cult..	Docu..	Dram..	Faith..	Horr..	Indep..	Inter..	Inter..	Kore..	LGBT..	Musi..	Reali..	Roma..	Sci-Fi..
Argentina	R	Abc														
	TV-MA								Abc							
Australia	NR	Abc														
	PG-13	Abc														
	R	Abc														
	TV-14	Abc														
	TV-MA															Abc
Austria	TV-PG								Abc							
Belgium	PG	Abc	Abc						Abc							
	R	Abc	Abc						Abc							
	TV-14	Abc							Abc							
	TV-MA	Abc	Abc						Abc						Abc	
Botswana	TV-PG	Abc														
Brazil	PG	Abc														
	R							Abc								
Bulgaria	TV-PG	Abc														
Cambodia	PG	Abc	Abc													
Canada	NR	Abc														
	PG	Abc	Abc	Abc					Abc							
	PG-13	Abc	Abc	Abc					Abc							
	R	Abc			Abc			Abc								Abc
	TV-14	Abc	Abc										Abc		Abc	
	TV-G	Abc	Abc													
	TV-MA	Abc													Abc	
	TV-PG	Abc	Abc													Abc
	TV-Y	Abc														
	TV-Y7	Abc														
	UR	Abc														
Chile	R	Abc														
	TV-14	Abc														
China	NR	Abc														
	PG	Abc	Abc													

Figura 5: Resultados

- Número de filmes/séries por país:

Número de filmes/séries por país.

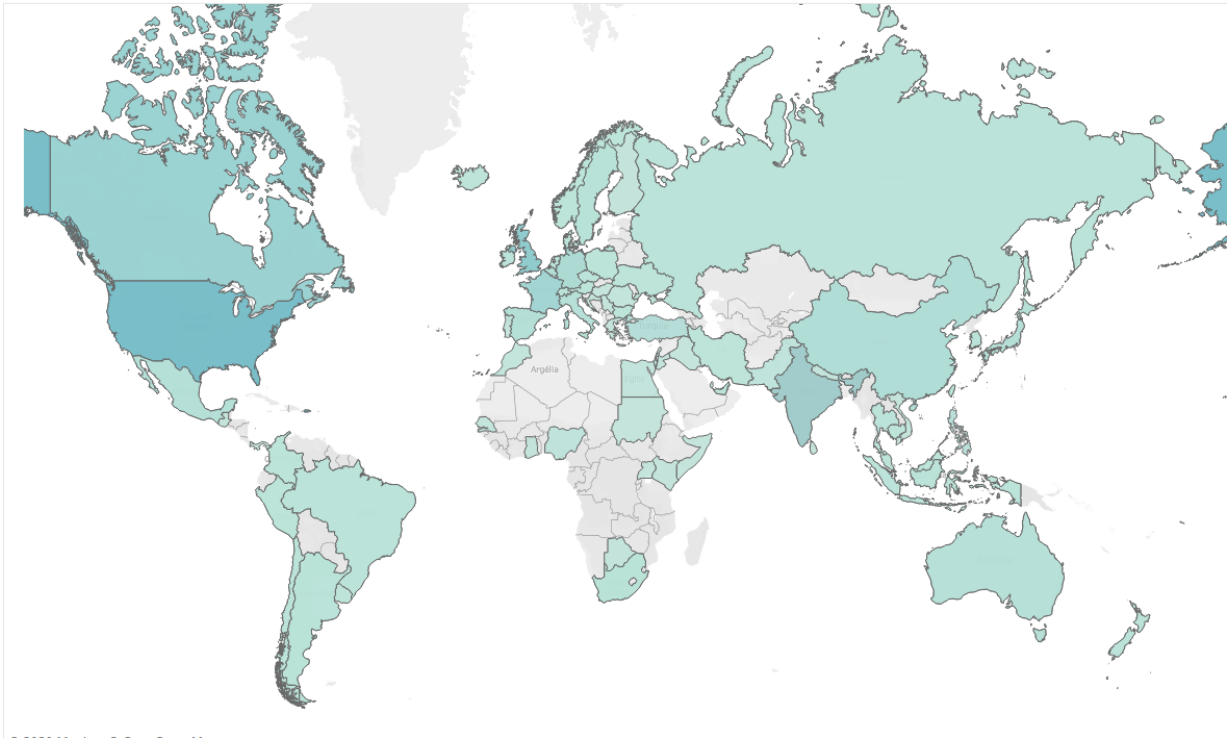


Figura 6: Resultados

- Duração de filmes e Séries:

Duração de Séries e filmes.

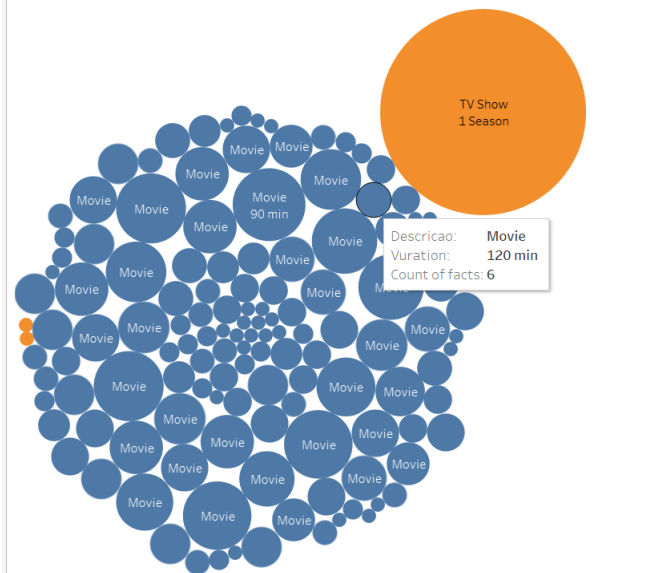


Figura 7: Resultados- Laranja Séries , Azul Filmes

- Número de filmes/series por ator:

Número de filmes/series por ator.

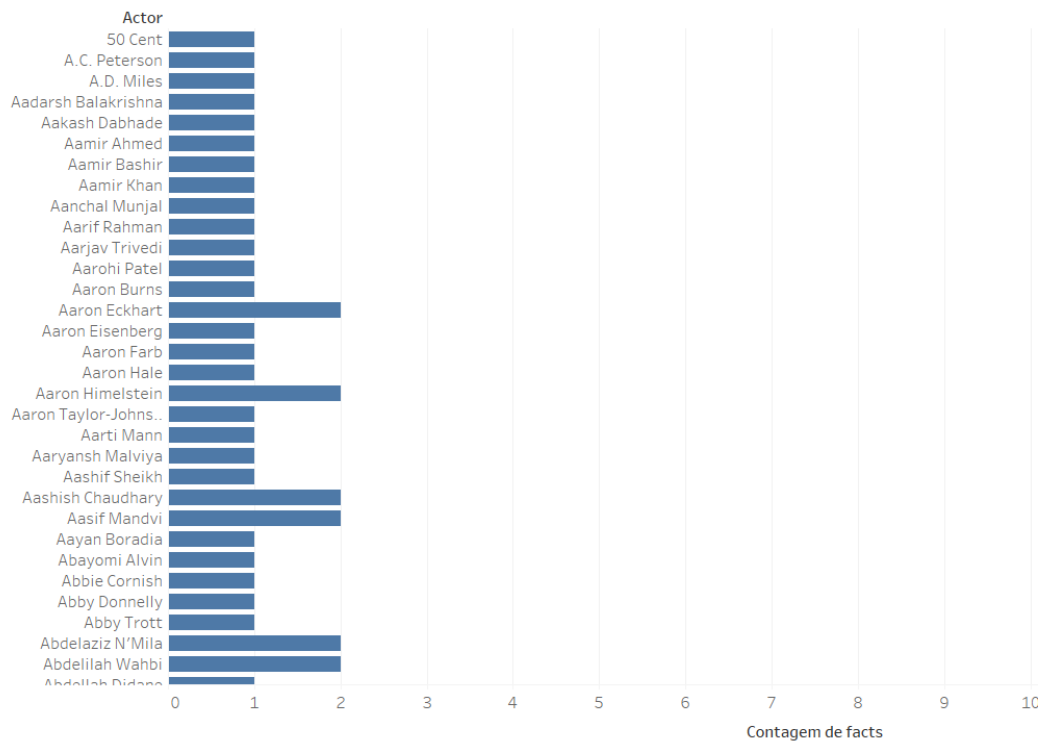


Figura 8: Resultados

### 3 Conclusão

É assim concluído este relatório de descrição de construção de um modelo MySQL com dados fornecidos em formato csv e respectivas análise estatísticas realizadas a partir do software Tableau.