

# Análise de Dados com SQL

## Filmes Lançados Entre 1900-2025



Por: Fábio dos Santos

FONTE: <https://www.kaggle.com/datasets/amitksingh2103/trending-movies-over-the-years>

## -- QUANTIDADE DE FILMES ANALISADOS

```
SELECT COUNT(title) AS Qtd_Filmes  
FROM tabelafilmes_csv;
```

Qtd_Filmes
10000

## -- 10 FILMES MAIS VOTADOS.

```
SELECT title AS Titulo, sum(vote_count) AS Qtd_Votos  
FROM tabelafilmes_csv  
GROUP BY title  
ORDER BY Qtd_Votos DESC  
LIMIT 10;
```

Titulo	Qtd_Votos
The Avengers	63744
Fight Club	60572
The Martian	40338
Logan	39248
Gone Girl	38048
Inception	37480
Interstellar	37167
The Maze Runner	34696
La La Land	34582
The Dark Knight	33867

## -- 5 FILMES COM MAIOR POPULARIDADE.

```
SELECT title AS Titulo, SUM(popularity) AS Popularidade,  
strftime('%Y', release_date) AS Ano_Lançamento, round(vote_average, 2) AS  
Nota_Media  
FROM tabelafilmes_csv  
GROUP BY title, Ano_Lançamento  
ORDER BY Popularidade DESC LIMIT 5;
```

Titulo	Popularidade	Ano_Lançamento	Nota_Media
A Minecraft Movie	694.9673	2025	6.53
Final Destination Bloodlines	535.0346	2025	7.08
Lilo & Stitch	508.3123	2025	7.06
Mission: Impossible - The Final Reckoning	325.1914	2025	7.03
A Working Man	314.9315	2025	6.67

-- IDIOMAS COM MAIS FILMES (EM ORDEM DECRESCENTE DE QUANTIDADE).

```
SELECT original_language AS Linguagem_Original, count(title) AS Qtd_Filmes
FROM tabelafilmes_csv
GROUP BY original_language
ORDER BY Qtd_Filmes DESC LIMIT 10;
```

Linguagem_Original	Qtd_Filmes
en	6471
ja	658
fr	480
ko	365
zh	311
cn	233
es	221
de	198
it	184
hi	87

### Quantidade de Filmes por Idioma



Created with Datawrapper

-- NOTA MÉDIA DOS FILMES EM CADA DÉCADA.

SELECT

CASE

WHEN Ano BETWEEN '2015' AND '2025' THEN '2015-2025'

WHEN Ano BETWEEN '2004' AND '2014' THEN '2004-2014'

WHEN Ano BETWEEN '1993' AND '2003' THEN '1993-2003'

WHEN Ano BETWEEN '1982' AND '1992' THEN '1982-1992'

WHEN Ano BETWEEN '1971' AND '1981' THEN '1971-1981'

WHEN Ano BETWEEN '1960' AND '1970' THEN '1960-1970'

WHEN Ano BETWEEN '1949' AND '1959' THEN '1949-1959'

WHEN Ano BETWEEN '1938' AND '1948' THEN '1938-1948'

WHEN Ano BETWEEN '1927' AND '1937' THEN '1927-1937'

WHEN Ano BETWEEN '1916' AND '1926' THEN '1916-1926'

WHEN Ano BETWEEN '1900' AND '1915' THEN '1900-1915'

ELSE 'Fora do Intervalo'

END AS Decada,

COUNT('titles') AS QTD\_filmes, ROUND(AVG(vote\_average), 3) AS Media\_voto

FROM (

SELECT STRFTIME('%Y', release\_date) AS Ano, vote\_average

FROM tabelafilmes\_csv

WHERE vote\_average > 0

) AS subquery

GROUP BY Decada

HAVING Decada  $\neq$  'Fora do Intervalo'

ORDER BY Decada ASC;

Decada	QTD_filmes	Media_voto
1900-1915	10	4.038
1916-1926	13	5.958
1927-1937	49	6.55
1938-1948	81	6.729
1949-1959	142	6.674
1960-1970	249	6.41
1971-1981	441	6.374
1982-1992	823	6.216
1993-2003	1331	6.376
2004-2014	2112	6.234
2015-2025	4567	5.629

## Média de Notas por Década



Created with Datawrapper

-- QTD DE VOTOS EM CADA DÉCADA.

SELECT

CASE

WHEN Ano BETWEEN '2015' AND '2025' THEN '2015-2025'

WHEN Ano BETWEEN '2004' AND '2014' THEN '2004-2014'

WHEN Ano BETWEEN '1993' AND '2003' THEN '1993-2003'

WHEN Ano BETWEEN '1982' AND '1992' THEN '1982-1992'

WHEN Ano BETWEEN '1971' AND '1981' THEN '1971-1981'

WHEN Ano BETWEEN '1960' AND '1970' THEN '1960-1970'

WHEN Ano BETWEEN '1949' AND '1959' THEN '1949-1959'

WHEN Ano BETWEEN '1938' AND '1948' THEN '1938-1948'

WHEN Ano BETWEEN '1927' AND '1937' THEN '1927-1937'

WHEN Ano BETWEEN '1916' AND '1926' THEN '1916-1926'

WHEN Ano BETWEEN '1900' AND '1915' THEN '1900-1915'

ELSE 'Fora do Intervalo'

END AS Década,

COUNT('title') AS Quantidade\_de\_Filmes

FROM (

SELECT

STRFTIME('%Y', release\_date) AS Ano, vote\_average

FROM tabelafilmes\_csv

WHERE vote\_average > 0

) AS subquery

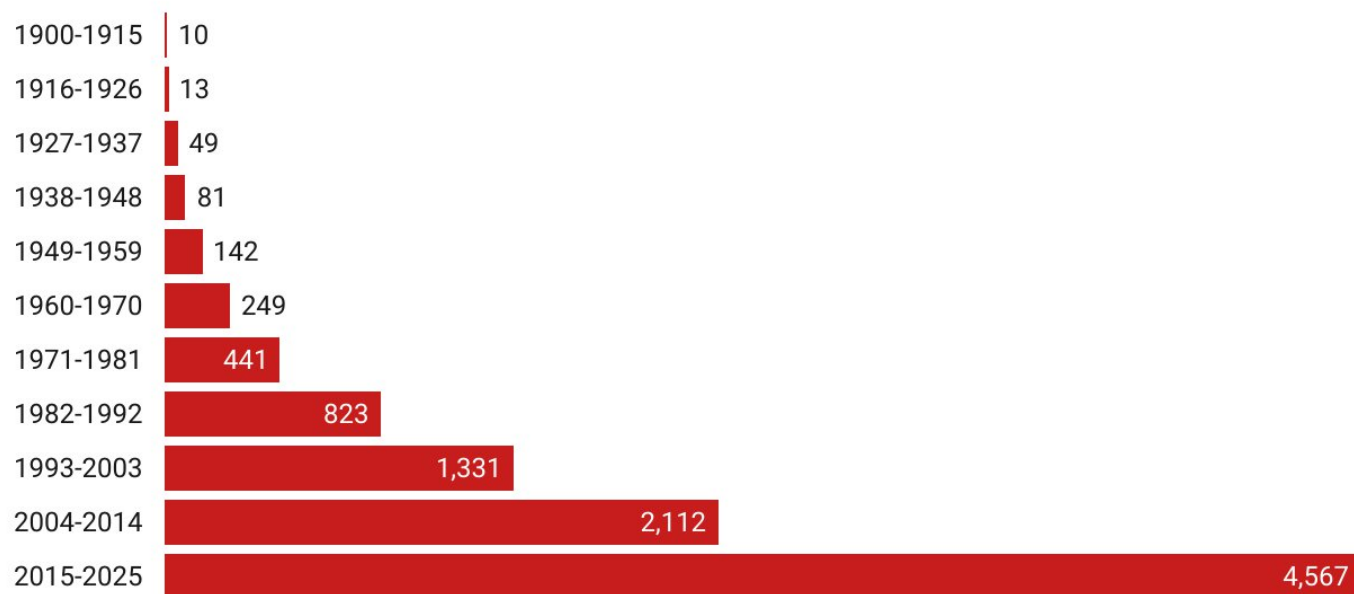
GROUP BY Década

HAVING Década  $\neq$  'Fora do Intervalo'

ORDER BY Década ASC;

Década	Quantidade_de_Filmes
1900-1915	10
1916-1926	13
1927-1937	49
1938-1948	81
1949-1959	142
1960-1970	249
1971-1981	441
1982-1992	823
1993-2003	1331
2004-2014	2112
2015-2025	4567

## Quantidade de Filmes Lançados



Created with Datawrapper

## – CONCLUSÕES.

### Quantidade Total de Filmes

- Foram analisados **10.000 filmes** no total.

### Filmes Mais Populares

- Todos os 5 filmes mais populares são de **2025**, indicando forte expectativa do público por lançamentos futuros

### Idiomas Mais Comuns

- **Inglês** é dominante com 6.471 filmes (quase 65% do total).
- Outras línguas com destaque: Japonês (658), Francês (480), Coreano (365), Chinês (cn+zh somados: 544).

### Nota Média por Década

- Décadas com melhores médias de nota:
  - o 1938–1948: 6.729
  - o 1927–1937: 6.550
  - o 1949–1959: 6.674
- Década com a **pior média**:
  - o 2015–2025: 5.629

**Conclusão:** Apesar do avanço tecnológico, a média de avaliação dos filmes **diminuiu nas últimas décadas**, o que pode indicar:

- Saturação de mercado
- Excesso de produções comerciais em detrimento da qualidade
- Mudança no perfil dos votantes

### Volume de Produções por Década

- A década de **2015–2025** tem o maior número de filmes (4.567), mostrando um **crescimento exponencial** na produção cinematográfica.
- Esse crescimento é impulsionado pela digitalização, streaming e plataformas independentes.

### Conclusões Gerais

1. **Mais filmes não significam melhores filmes:** a qualidade média caiu, mesmo com o aumento da produção.
2. **A popularidade está concentrada em lançamentos recentes e futuros**, o que reforça o papel do marketing e hype.
3. **A indústria cinematográfica é dominada pelo inglês**, mas há uma crescente diversidade linguística.
4. **Clássicos antigos** tendem a ter **notas mais altas**, possivelmente por serem mais seletivamente lembrados e votados.