

## Um Estudo de Recomendações de Séries e Filmes A Study of TV Series and Movie Recommendations

Mozart da Silva Monteiro<sup>1</sup>, Fernando José Antonio<sup>2</sup>

### RESUMO

Com o tempo e a crescente popularização da internet, a tecnologia foi aprimorada e passou a oferecer uma série de outras funcionalidades, algumas das quais são amplamente utilizadas no dia a dia, como as redes sociais e os serviços de streaming. Diante do surgimento de diversas plataformas de transmissão online, uma dúvida se faz presente: "qual seria o melhor serviço disponível?" Nesse contexto, a pesquisa desenvolvida visa realizar uma análise dos principais serviços de streaming e, para isso, foram empregadas abordagens estatísticas para analisar os dados públicos. Examinaram-se as avaliações médias e como elas variam ao longo do tempo. A partir desse método, foi possível observar comportamentos, padrões, tendências, diferenças entre as plataformas e um possível serviço que tenha uma qualidade superior aos outros.

**PALAVRAS-CHAVE:** Acervo de filmes e séries; análise de dados; padrão de qualidade; plataformas de *streaming*; linhas de tendência.

### ABSTRACT

Over time and with the increasing popularization of the internet, technology has been improved and started to offer a range of additional functionalities, some of which are widely used in daily life, such as social networks and streaming services. Faced with the emergence of various online streaming platforms, a question arises: 'what would be the best available service?' In this context, the conducted research aims to perform an analysis of the leading streaming services, and to achieve this, statistical approaches were employed to analyze public data. Average ratings were examined, as well as how they vary over time. Through this method, it was possible to observe behaviors, patterns, trends, differences between the platforms, and a possible service that may have superior quality compared to the others.

**KEYWORDS:** Data analysis; quality standard; streaming platforms; trend lines.

## INTRODUÇÃO

A internet foi criada na década de 1960 pelo Departamento de Defesa dos Estados Unidos, exclusivamente para fins militares, durante o período da Guerra Fria, quando ocorria uma corrida bélica e tecnológica entre os Estados Unidos e a antiga União Soviética. O objetivo era estabelecer uma rede de comunicação de computadores em pontos estratégicos para descentralizar informações valiosas, evitando que fossem destruídas caso estivessem concentradas em apenas um único local. No entanto, a tecnologia só ganhou destaque no final dos anos 90, quando foi aberta para uso civil (LEMONAKI, 2020). Com a popularização da Internet nos anos 2000, surgiram inúmeros meios de entretenimento. Entre esses, destacam-se os serviços de *streaming*, que, futuramente viriam a substituir as saudosas locadoras de filmes. Porém, no conforto da sua residência, com bastante praticidade para assistir o conteúdo em qualquer dispositivo e com acesso a uma vasta quantidade de conteúdos por um preço acessível. A

<sup>1</sup> Voluntário do Programa Institucional de Voluntariado em Iniciação Científica. Universidade Tecnológica Federal do Paraná, Cornélio Procópio, Paraná, Brasil. E-mail: mozartsilvamonteiro@yahoo.com.br. ID Lattes: <http://lattes.cnpq.br/2665594968947941>.

<sup>2</sup> Docente no Departamento Acadêmico de Ciências da Natureza. Universidade Tecnológica Federal do Paraná, Cornélio Procópio, Paraná, Brasil. E-mail: [fjantonio@utfpr.edu.br](mailto:fjantonio@utfpr.edu.br). ID Lattes: <http://lattes.cnpq.br/2833172211868473>.

plataforma de *streaming* Netflix foi uma das pioneiras que popularizaram a forma de entretenimento. Com o sucesso da empresa, surgiram também concorrentes, no entanto, oferecendo catálogos de filmes e séries distintos entre si. Isso suscita a pergunta: “Qual plataforma melhor atende às necessidades do cliente?”. A análise para determinar o melhor serviço de *streaming* faz-se relevante porque auxilia e torna os consumidores capazes de fazer escolhas informadas, promove a concorrência na indústria e contribui para uma melhor experiência de entretenimento. Até o momento, não foi encontrada nenhuma análise similar à pesquisa realizada.

Neste trabalho, com base nesse questionamento, bancos de dados públicos foram usados para qualificar os serviços de streaming de acordo com a avaliação pública para diferentes conteúdos presentes na plataforma, com o intuito de auxiliar os consumidores a escolher um serviço que mais agradará.

## MATERIAIS E MÉTODOS

Para realizar esta pesquisa, foi necessário dominar a linguagem de programação Python (Python, 2023) a partir do zero. Isso se mostrou essencial para a realização da análise de dados, um procedimento crucial para filtrar e estruturar com eficiência grandes conjuntos de dados, visando a obtenção de informações valiosas. Python é uma linguagem de programação gratuita que encontra aplicações em análise de dados em diversas áreas, incluindo engenharia. Especificamente na engenharia mecânica, a análise de dados pode ser usada para otimizar o design de peças e componentes, reduzindo o tempo e os custos do processo de fabricação, além de contribuir para a compreensão de fenômenos de vibração em estruturas. As bibliotecas do Python utilizadas nesta pesquisa incluem pandas, numpy, matplotlib, seaborn e plotly express. Além disso, foi fundamental adquirir conhecimento e experiência em *bootstrapping*, que é uma técnica estatística de amostragem usada para calcular intervalos de confiança de funções por meio da reamostragem de dados existentes.

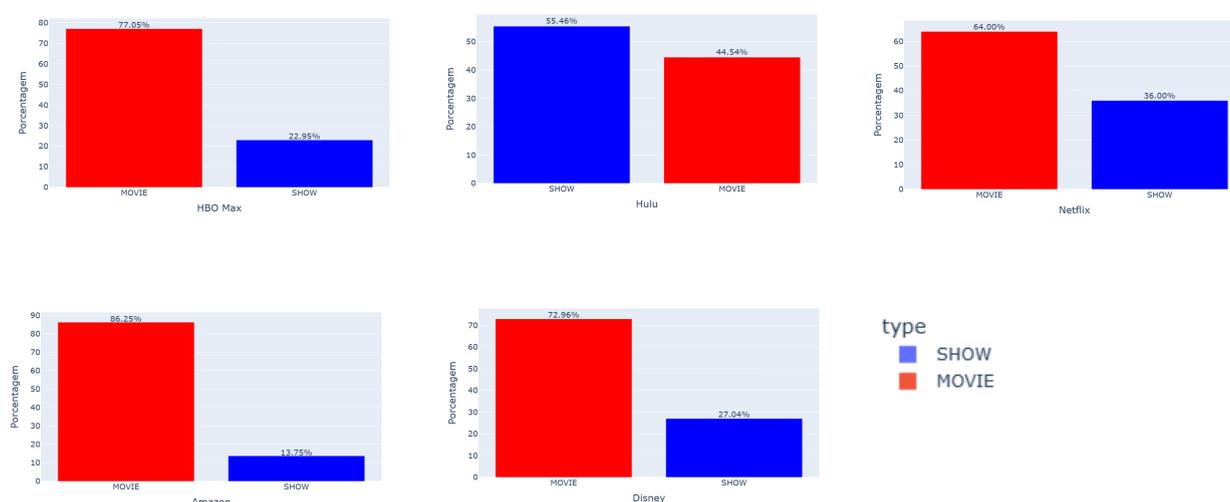
Os dados coletados e analisados foram adquiridos por meio da plataforma Kaggle (KAGGLE, 2023: <https://www.kaggle.com/>), amplamente utilizada tanto para fins de aprendizado em ciência de dados quanto como uma comunidade dedicada a assuntos relacionados à ciência de dados. Outros serviços de *streaming*, como a Apple TV+, não foram incluídos neste trabalho, pois não foram encontrados dados sobre eles. Os dados coletados são séries temporais, ou seja, informações que são observadas ao longo do tempo. Nestes bancos de dados, as entradas são classificadas em filmes ou séries de TV identificados por ID, para evitar duplicações, título, tipo, descrição, ano de lançamento, classificação indicativa, duração, gênero, países que produziram o conteúdo, quantidade de temporadas de cada série, ID no IMDb, nota no IMDb. É importante ressaltar que os dados originais sobre as notas do IMDb provêm de uma fonte confiável e amplamente utilizada na indústria do entretenimento. Além disso, é crucial destacar que todos os códigos usados para obter e analisar esses dados são de código aberto, a fim de garantir transparência e acessibilidade para qualquer interessado. Outros serviços de streaming, como a Apple TV, não foram incluídos neste trabalho, pois não foram encontrados dados sobre eles.

O método de análise consistiu em investigar e qualificar as variáveis isoladamente para a caracterização dos dados. Em seguida, buscou-se a evolução da avaliação média dos produtos ao longo do tempo.

## RESULTADOS E DISCUSSÕES

A primeira análise dos dados consistiu em caracterizá-los para diferenciar as variações entre os serviços oferecidos. De acordo com a Figura 1, a cor azul representa as séries (*SHOW*) e a cor vermelha representa os filmes (*MOVIE*) e observa-se que o serviço com a maior proporção de filmes em relação a séries é a Amazon Prime, enquanto o serviço que possui uma proporção maior de séries em relação a filmes é o Hulu. No entanto, quando se considera que as séries são compostas tipicamente por vários episódios por temporada, o universo de episódios de séries torna-se mais proeminente.

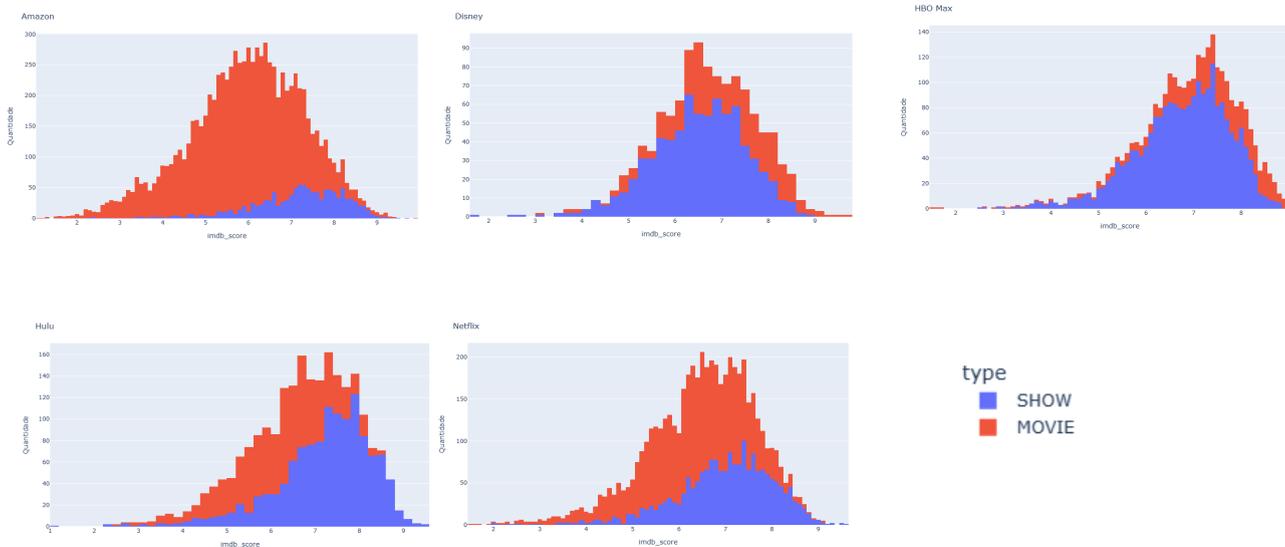
Figura 1 - Caracterização do catálogo de cada plataforma de *streaming*.



Fonte: Autoria Própria(2023)

O segundo ponto analisado foi investigar a distribuição das notas dos conteúdos disponíveis em cada serviço. Essa análise auxiliará na resposta à pergunta: 'Seria correto afirmar que quantidade se traduz em qualidade?'. Para isso, foi construído um histograma para avaliações de filmes e séries em cada plataforma. A análise visual dos histogramas (Figura 2) revela diferenças nas avaliações médias pelo *IMDb*, destacando que a qualidade varia entre os serviços. Na figura, a cor azul identifica as séries (*SHOW*) e a cor vermelha identifica os filmes (*MOVIE*). Além disso, a variância nas avaliações também não é sempre consistente. Se assumirmos que a avaliação média no *IMDb* pode ser associada ao padrão de qualidade de uma produção, a primeira conclusão obtida desses gráficos é que a Disney Plus e a Hulu mantêm o mesmo padrão de qualidade tanto para filmes quanto para séries de TV, enquanto a Amazon Prime apresenta uma grande diferença nas avaliações de filmes e séries, indicando inconsistência. A Tabela 1 detalha as médias e os desvios-padrão das avaliações, mostrando que a Amazon Prime tem o menor desempenho em filmes, com alta variabilidade, enquanto a HBO Max lidera em qualidade, com a menor variação nas avaliações de filmes.

Figura 2 - Histograma da avaliação no IMDb do catálogo de cada plataforma de *streaming*.



Fonte: Autoria Própria (2023)

Tabela 1 - Médias e Desvios padrão do catálogo de cada plataforma de *streaming*.

Plataformas	Filmes		Séries	
	Média das avaliações	Desvio padrão das avaliações	Média das avaliações	Desvio padrão das avaliações
Amazon Prime	5,80	1,28	7,12	1,17
Disney Plus	6,44	1,01	6,94	1,11
HBO MAX	6,77	0,99	7,26	1,18
Hulu	6,13	1,07	7,15	1,12
Netflix	6,25	1,13	6,98	1,08

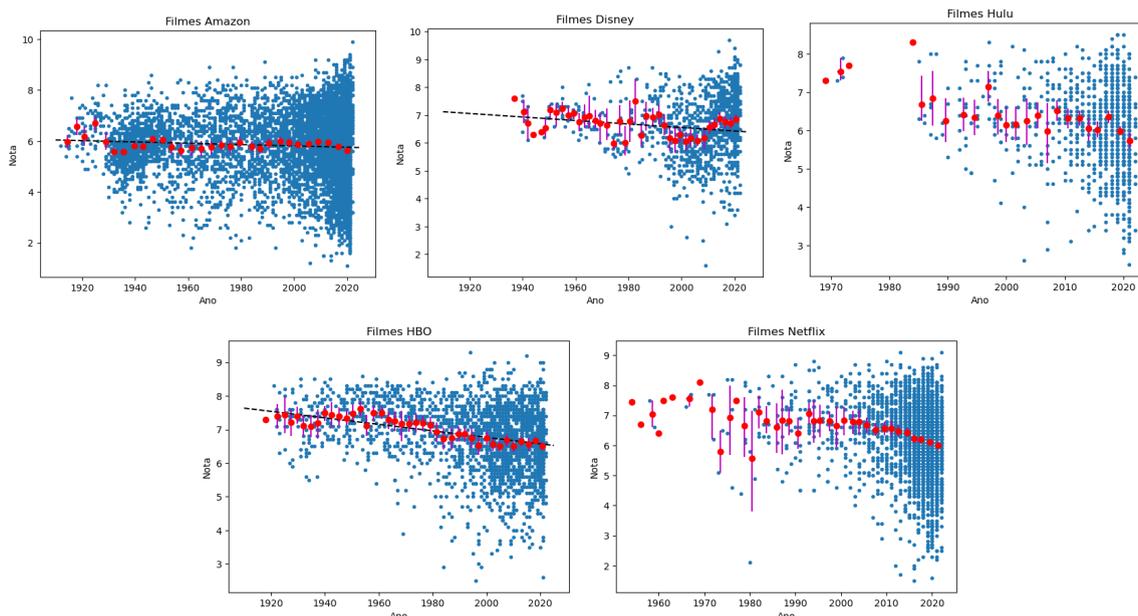
Fonte: Autoria Própria(2023)

Partindo dessa primeira parte de caracterização, investigou-se como a avaliação média de cada produção de filmes e séries dos serviços de *streaming* evoluiu ao longo do tempo. Para isso, construiu-se a Figura 3. Para a investigação, os dados foram separados em janelas temporais equidistantes. A avaliação média foi computada dentro de cada uma dessas janelas, sendo considerado um intervalo de confiança de 95% calculado pelo método de *bootstrap*.

Através da Figura 3, verificou-se que as produções cinematográficas no passado, de modo geral, obtinham avaliações melhores. Além disso, é perceptível que todas as plataformas mostram uma tendência sistemática de queda na nota da avaliação média de suas produções cinematográficas, com exceção da Disney Plus. Isso talvez possa ser atribuído à expansão do mercado, que exige cada vez mais produtos em um menor período de tempo. Apesar de a Disney Plus ter apresentado uma tendência de queda nas avaliações de filmes após 2010, o padrão se alterou para um padrão de crescimento. Embora a Amazon Primos, as concorrentes têm médias de nota no *IMDb* maiores,

mostrando assim que o serviço da Amazon Prime não demonstra um cenário de possível melhora futura de seu acervo de longas-metragens. Também se mostrou o serviço com menor tendência de queda com o passar dos anos, notável que a avaliação média dos filmes do HBO Max tem caído ao longo dos anos, sendo mais críticos a partir da década de 1980. Embora os serviços de streaming não existissem na época em que algumas produções cinematográficas foram lançadas, o conteúdo está atualmente acessível na plataforma, portanto, mesmo que tenha sido avaliado antes de estar disponível no streaming, sua qualidade permanece a mesma.

Figura 3 - Evolução da nota no IMDb do catálogo de cada plataforma de *streaming* para filmes (pontos azuis) destacando médias em janelas equidistantes no eixo horizontal (pontos vermelhos).



Fonte: Autoria Própria(2023)

## CONCLUSÃO

Através dos dados coletados e resultados obtidos, foi possível caracterizar os serviços de streaming de acordo com a avaliação de seus catálogos. Em particular, tornou-se perceptível que a Amazon Prime apresenta uma proporção maior de filmes em relação às séries em comparação com os outros serviços disponíveis, enquanto o Hulu tem a maior proporção de séries em relação a filmes em comparação com as outras plataformas. No entanto, a opção que oferece filmes com a melhor média de nota no *IMDb* e uma qualidade mais consistente é a HBO Max, apresentando também o menor desvio padrão entre os filmes oferecidos, garantindo um padrão de qualidade esperado para seus consumidores em relação aos filmes. Por outro lado, embora a Amazon Prime apresente uma tendência constante na nota do *IMDb* ao longo dos anos, ela apresenta a menor média de nota no *IMDb* e o maior desvio padrão, ou seja, pode-se afirmar que a Amazon Prime oferece o serviço de *streaming* de filmes com qualidade inferior em relação aos concorrentes e também não assegura um padrão de qualidade consistente para seus clientes. Quanto às séries oferecidas por cada plataforma, notou-se que novamente a HBO Max obteve destaque devido ao fato de ser o serviço de vídeo online

com a maior média em relação aos outros. No entanto, ela também tem o maior desvio padrão, fornecendo, de modo geral, um conteúdo de qualidade superior aos outros, porém, não garantindo um padrão de serviços uniforme ao consumidor. Por fim, a Netflix apresenta a segunda pior média; no entanto, tem o menor desvio padrão das notas do *IMDb*. No que se refere às séries, ela oferece um padrão de qualidade mais consistente para seus usuários do que as demais. Portanto, de acordo com a média de nota do *IMDb*, a HBO Max seria a plataforma de transmissão de conteúdo mais bem avaliada e que, dentre a concorrência avaliada neste estudo, ela possivelmente será a que mais agradará aos seus consumidores. Por outro lado, a Amazon Prime e a Disney Plus possuem, respectivamente, as menores médias de nota no *IMDb* em relação a filmes e seriados, podendo ser classificadas como as que provavelmente menos cativarão seus clientes, além de ambas possuírem um alto desvio-padrão nas categorias citadas, não conseguindo manter um padrão de qualidade e, provavelmente, serão as que mais descontentarão os seus consumidores.

## Agradecimentos

Agradecemos à UTFPR pela infraestrutura disponibilizada.

## Conflito de interesse

Não há conflito de interesse.

## REFERÊNCIAS

- ARCOVERDE, Daniel. Streaming: o que é, como funciona e quais as principais plataformas no Brasil? Disponível em: <https://netshow.me/blog/mas-o-que-e-streaming/>. Acesso em: 12/04/2023.
- EFRON, B.; TIBISHIRANI, R. J. **An introduction to the bootstrap**. New York: John Wiley & Sons, 1993. p. 642.
- FORBES. Q1 2020 Proves Streaming Is Essential To Consumers And To The Future Of Media Companies. Disponível em: <https://www.forbes.com/sites/forrester/2020/04/27/q1-2020-proves-streaming-is-essential-to-consumers-and-to-the-future-of-media-companies/?sh=6ef7d07072d8>. Acesso em: 05/02/2022.
- HOSCH, William L. Netflix American company. Disponível em: <https://www.britannica.com/topic/Netflix-Inc>. Acesso em: 15/05/2023.
- KAGGLE. Kaggle: Your Home for Data Science. Disponível em: <https://www.kaggle.com/>. Acesso em: 25/05/2023.
- LAMARE, Amy. How streaming started: YouTube, Netflix, and Hulu's quick ascent. Disponível em: <https://www.businessofbusiness.com/articles/a-brief-history-of-video-streaming-by-the-numbers/>. Acesso em: 16/06/2023.
- LEMONAKI, Dionysia. A Brief History of the Internet – Who Invented It, How it Works, and How it Became the Web We Use Today. Disponível em: <https://www.freecodecamp.org/news/brief-history-of-the-internet/>. Acesso em: 15/03/2023.
- Python Software Foundation. Python.org. Disponível em: <https://www.python.org/>. Acesso em: 22/05/2023
- VANGA, Manohar. PyBootstrap. Disponível em: <https://pypi.org/project/pybootstrap/#files>. Acesso em: 23/05/2023.