

ANÁLISE DE DADOS EM PLATAFORMAS DE STREAMING: AMAZON, DISNEY+, HBO E NETFLIX

Vitória Hillary Alarcon
Graduando em Ciência da Computação – Uni-FACEF
alarconhvitoria@gmail.com

Bruno Henrique Cruvinel Borges Jardine
Graduando em Engenharia de Software – Uni-FACEF
brunojardine@hotmail.com

Prof. Leandro Borges
Docente do Departamento de Computação – Uni-FACEF
leandro.borges@facef.br

Resumo

O rápido crescimento de serviços de streaming como Netflix, Amazon Prime Video, Disney+ e HBO Max transformou a forma como o conteúdo audiovisual é consumido em todo o mundo. Essas plataformas oferecem catálogos diversos, gerando grandes volumes de dados que podem ser analisados para identificar padrões, tendências e diferenças em suas estratégias de distribuição de conteúdo. Este estudo tem como objetivo realizar uma análise comparativa dos catálogos desses quatro serviços utilizando técnicas de ciência de dados e business intelligence. Para isso, foram coletados quatro conjuntos de dados do Kaggle, contendo informações detalhadas sobre os títulos disponíveis em cada plataforma, incluindo atributos como gênero, ano de lançamento, país de origem, duração e classificação indicativa. Os conjuntos de dados foram processados por meio de um fluxo metodológico estruturado: limpeza e transformação dos dados foram realizadas em Python, o armazenamento e a organização foram feitos utilizando SQL Server, e a análise exploratória e visualização foram conduzidas no Power BI. Esse processo garantiu a consistência dos dados e proporcionou insights claros e acessíveis. Os resultados deste estudo permitem uma compreensão mais aprofundada das estratégias empregadas por cada plataforma de streaming, evidenciando suas semelhanças, diferenças e características distintivas. Além disso, a pesquisa demonstra a relevância da análise de dados na indústria do entretenimento, bem como a importância da integração de ferramentas de programação, gerenciamento de banco de dados e business intelligence para a interpretação de grandes conjuntos de dados.

Abstract

The rapid growth of streaming services such as Netflix, Amazon Prime Video, Disney+, and HBO Max has transformed the way audiovisual content is consumed worldwide. These platforms offer diverse catalogs, generating large volumes of data that can be analyzed to identify patterns, trends, and differences in their content distribution strategies. This study aims to perform a comparative analysis of the catalogs of these four services using data science and business intelligence techniques. To achieve this, four datasets were collected from Kaggle, containing detailed information about the titles available on each platform, including attributes such as genre, release year, country of origin, duration, and maturity rating. The datasets were processed through a structured methodological flow: data cleaning and transformation were carried out in Python, storage and organization were performed using SQL Server, and exploratory

analysis and visualization were conducted in Power BI. This process ensured data consistency and provided clear and accessible insights. The results of this study allow for a deeper understanding of the strategies employed by each streaming platform, highlighting their similarities, differences, and distinctive features. Furthermore, the research demonstrates the relevance of data analysis in the entertainment industry, as well as the importance of integrating programming, database management, and business intelligence tools for the interpretation of large datasets.

Keywords: Streaming, Data Analysis, Netflix, Amazon Prime Video, Disney+, HBO Max, Power BI, Python, SQL Server.

1 Introdução

Nas últimas décadas, o consumo de conteúdos audiovisuais passou por uma profunda transformação, impulsionada pelo surgimento e consolidação dos serviços de streaming. Plataformas como *Netflix*, *Amazon Prime Video*, *Disney+* e *HBO Max* tornaram-se parte integrante do cotidiano de milhões de pessoas em todo o mundo, oferecendo catálogos diversificados que abrangem filmes, séries, documentários e produções originais. Essa mudança de paradigma não apenas alterou os hábitos de consumo de entretenimento, mas também gerou um grande volume de dados, que podem ser explorados para identificar padrões, tendências e diferenças entre os serviços disponíveis.

O crescimento acelerado desses catálogos cria uma oportunidade para a aplicação de técnicas de análise de dados, permitindo compreender melhor como cada plataforma organiza seu portfólio e quais características se destacam em suas produções. A coleta, tratamento e exploração dessas informações possibilitam, por exemplo, observar quais gêneros são mais frequentes, quais países possuem maior representatividade nas produções, como evoluiu o lançamento de conteúdos ao longo dos anos e de que forma os serviços se diferenciam em sua estratégia de oferta.

Para alcançar esse objetivo, torna-se necessário aplicar ferramentas capazes de transformar dados brutos em informações organizadas e de fácil interpretação. Neste estudo, foram utilizados *datasets* públicos disponibilizados no *Kaggle*, que reúnem registros sobre os catálogos das quatro plataformas de streaming mencionadas. A partir disso, foi elaborado um fluxo metodológico estruturado em três etapas: coleta dos dados, tratamento e limpeza utilizando a linguagem *Python*, armazenamento em banco de dados no *SQL Server* e, por fim, análise exploratória e visualização por meio do *Power BI*. Esse processo buscou assegurar a consistência e integridade dos dados, resultando em análises comparativas confiáveis.

Assim, o presente trabalho tem como objetivo realizar uma análise comparativa entre os catálogos da *Netflix*, *Amazon Prime Video*, *Disney+* e *HBO Max*, evidenciando semelhanças, diferenças e padrões que caracterizam cada plataforma. A pesquisa contribui não apenas para a compreensão do panorama atual do mercado de streaming, mas também demonstra a importância do uso de técnicas de ciência de dados e ferramentas de *Business Intelligence* como suporte à tomada de decisões em contextos que envolvem grandes volumes de informação.

2 Referencial Teórico

O avanço das tecnologias de informação e a popularização das plataformas digitais transformaram significativamente a forma como conteúdos audiovisuais são distribuídos e consumidos. Diferentemente do modelo tradicional de televisão e cinema, os serviços digitais permitem acesso sob demanda, com catálogos vastos e constantemente atualizados, proporcionando aos usuários maior liberdade de escolha e conveniência. Segundo Kotler e Keller (2018, p. 108), “a transformação digital redefine não apenas a forma de consumir, mas também a maneira como as empresas entregam valor ao cliente”. Esse cenário tem impulsionado a necessidade de análises mais detalhadas sobre os conteúdos oferecidos, visando compreender padrões de produção, preferências do público e posicionamento competitivo das empresas.

A análise de dados aplicada ao mercado de conteúdo digital tornou-se uma ferramenta estratégica para empresas, pesquisadores e profissionais da área. Com o crescimento exponencial da oferta de títulos e serviços, a capacidade de extrair *insights* a partir de grandes volumes de dados tornou-se essencial para identificar tendências, prever demandas e otimizar estratégias de distribuição. De acordo com Davenport e Harris (2017, p. 52), “as empresas que utilizam a análise de dados de forma estratégica conseguem criar vantagens competitivas, antecipando demandas e melhorando a eficiência operacional”. Essa abordagem possibilita a avaliação de métricas como quantidade de títulos, distribuição por gênero, classificação etária, duração média e presença de produções originais.

Além disso, a evolução das preferências do público e a diversidade cultural entre regiões tornam a análise de conteúdo ainda mais relevante. A identificação de padrões nos tipos de produções, gêneros predominantes e presença de conteúdos originais pode revelar estratégias adotadas pelas empresas para conquistar e manter clientes. Segundo Anderson (2006, p. 67), “a ampla oferta de títulos em ambientes digitais cria um fenômeno conhecido como ‘cauda longa’, onde nichos de conteúdo acumulam relevância e receita significativas ao longo do tempo”. Esse fenômeno é especialmente perceptível em catálogos digitais, onde produções menos conhecidas podem coexistir com grandes sucessos, atraindo diferentes perfis de público.

Dessa forma, a integração de dados com ferramentas analíticas modernas proporciona um panorama detalhado do cenário competitivo no mercado de conteúdo digital. A abordagem de análise de dados permite não apenas apresentar comparações quantitativas, mas também explorar *insights* qualitativos que contribuem para a compreensão das estratégias adotadas pelas empresas, oferecendo uma visão clara e embasada sobre a composição e as particularidades dos catálogos de conteúdos disponíveis.

2.1 Contextualização do Mercado de Streaming

O mercado de *streaming* de vídeo consolidou-se como um dos setores mais dinâmicos e competitivos da indústria do entretenimento, registrando crescimento contínuo nos últimos anos. Com a expansão da internet de alta velocidade e o aumento no número de dispositivos conectados, o consumo de conteúdo audiovisual sob demanda tornou-se parte do cotidiano de milhões de pessoas ao redor do mundo. Segundo Smith (2020, p. 45), “o *streaming* transformou o entretenimento em uma experiência personalizada, adaptada às preferências e ao ritmo de vida do espectador”. Esse novo paradigma rompeu as barreiras impostas

pela programação linear da televisão tradicional, permitindo que o usuário escolha o que assistir, quando e onde desejar.

Entre os principais players do mercado destacam-se *Netflix*, *Amazon Prime Video*, *Disney+* e *HBO Max*, cada um com estratégias próprias de aquisição e produção de conteúdo. Enquanto a *Netflix* se consolidou como pioneira na distribuição digital e investe fortemente em produções originais globais, a *Disney+* baseia-se em franquias consolidadas e propriedades intelectuais de alto valor, como *Marvel* e *Star Wars*. Já a *Amazon Prime Video* diversifica seu catálogo com filmes licenciados, produções originais e ofertas integradas a outros serviços, enquanto a *HBO Max* foca na qualidade e no prestígio de suas produções, herdando um histórico de séries e filmes premiados. De acordo com Johnson (2021, p. 73), “a competitividade no setor de *streaming* depende não apenas da quantidade de títulos disponíveis, mas também da exclusividade e da relevância cultural dos conteúdos oferecidos”.

Além da disputa por assinantes, essas plataformas enfrentam o desafio de atender a diferentes mercados e preferências culturais. A diversidade de gêneros, idiomas e classificações etárias torna-se um diferencial competitivo, pois permite que os serviços ampliem seu alcance global. Conforme Anderson (2006, p. 67), “a estratégia da cauda longa permite que empresas atendam nichos específicos ao mesmo tempo em que exploram grandes sucessos de audiência”. Isso significa que um catálogo bem equilibrado entre produções populares e títulos menos conhecidos pode garantir fidelidade de diferentes perfis de usuários.

O avanço das tecnologias de coleta e análise de dados também desempenha um papel fundamental nesse contexto. Plataformas de *streaming* utilizam algoritmos de recomendação baseados no histórico de consumo do usuário para oferecer sugestões personalizadas e aumentar o engajamento. Segundo Gomez-Uribe e Hunt (2016, p. 8), “os sistemas de recomendação são responsáveis por uma parcela significativa das visualizações nas plataformas de *streaming*, moldando a experiência do usuário e influenciando diretamente o consumo de conteúdo”. Essa integração entre dados, personalização e estratégia de catálogo reforça a importância de estudos comparativos sobre a composição dos serviços de *streaming*.

Portanto, compreender o cenário do mercado de *streaming* exige não apenas observar a evolução tecnológica, mas também analisar a estrutura e a diversidade dos catálogos de cada plataforma. A comparação entre *Netflix*, *Amazon Prime Video*, *Disney+* e *HBO Max*, a partir de dados concretos, possibilita identificar padrões e estratégias adotadas por cada serviço, contribuindo para uma visão mais aprofundada sobre o posicionamento competitivo no setor.

2.2 Netflix

A *Netflix* é amplamente reconhecida como uma das pioneiras e líderes globais no mercado de *streaming*, tendo desempenhado papel decisivo na transformação do consumo de conteúdos audiovisuais. Originalmente fundada em 1997 como um serviço de aluguel de DVDs por correio, a empresa transitou para o modelo digital em 2007, lançando seu serviço de streaming sob demanda, o que representou uma mudança paradigmática na forma como os usuários acessavam filmes e séries. Segundo Johnson (2019, p. 32), “a transição da *Netflix* do modelo físico para o digital marcou um ponto de inflexão no consumo de mídia, redefinindo hábitos e expectativas do público”.

Desde a introdução do *streaming*, a *Netflix* tem investido significativamente em produções originais, como séries, filmes e documentários, buscando diversificar seu catálogo e fortalecer sua identidade de marca. Como destaca Smith (2020, p. 45), “o *streaming* transformou o entretenimento em uma experiência personalizada, adaptada às preferências e ao ritmo de vida do espectador”. Esse movimento reflete uma estratégia de controle de propriedade intelectual e diferenciação no mercado altamente competitivo.

A plataforma adota uma abordagem global para atrair e reter assinantes, oferecendo conteúdo localmente relevante em diversos idiomas e adaptado às preferências culturais de cada região. Segundo Oliveira (2021, p. 58), “a personalização do catálogo de conteúdos, combinada à produção local, é uma das principais estratégias da *Netflix* para consolidar sua presença internacional e aumentar a fidelidade dos assinantes”. Além disso, a *Netflix* utiliza algoritmos de recomendação baseados em dados para personalizar a experiência do usuário, aumentando o engajamento e a satisfação do assinante (Martins, 2018, p. 77).

A análise de dados desempenha um papel central na estratégia da *Netflix*, permitindo à empresa entender padrões de consumo, prever tendências e otimizar sua oferta de conteúdo. De acordo com Pereira (2020, p. 61), “o uso de métricas de consumo detalhadas possibilita decisões estratégicas sobre produções originais, aquisições de títulos e ajustes no catálogo, garantindo competitividade e relevância no mercado digital”.

Além disso, a *Netflix* busca constantemente inovar em formatos e gêneros, explorando desde grandes produções globais até conteúdos de nicho, seguindo o conceito de cauda longa. Conforme Lima (2019, p. 89), “a diversidade do catálogo permite à *Netflix* atender a diferentes perfis de público, equilibrando grandes sucessos com produções menos conhecidas, mas igualmente estratégicas para engajamento e retenção”.

2.2 Amazon Prime Video

O *Amazon Prime Video* integra o serviço de *streaming* ao ecossistema mais amplo da Amazon, oferecendo acesso a filmes, séries e conteúdos exclusivos como parte do pacote *Prime*. Segundo Souza (2021, p. 47), “o diferencial do *Prime Video* está na integração com outros serviços da *Amazon*, criando uma experiência digital mais completa e conveniente para o usuário”. Essa estratégia permite que a plataforma combine entretenimento com benefícios comerciais, como frete grátis e promoções, fortalecendo a percepção de valor do serviço.

Além da integração com o ecossistema *Amazon*, o *Prime Video* investe tanto em produções próprias quanto na aquisição de títulos de terceiros, buscando atender a diferentes perfis de assinantes. De acordo com Carvalho (2020, p. 66), “a diversificação do catálogo, equilibrando produções originais e licenciadas, é uma estratégia que permite à plataforma competir com líderes do mercado e explorar nichos de público variados”. Esse equilíbrio entre conteúdo exclusivo e adquirido contribui para a retenção de assinantes e amplia o alcance da marca.

A plataforma também adota uma abordagem internacional, oferecendo conteúdos localizados e legendas em múltiplos idiomas, adaptando-se às preferências culturais de cada região. Segundo Almeida (2019, p. 52), “a adaptação regional de conteúdos é fundamental para a expansão global do *Prime Video*, permitindo que a

plataforma alcance audiências diversificadas e aumente a fidelidade do público". Além disso, a personalização de recomendações baseada em dados do usuário reforça o engajamento e a permanência no serviço (Ferreira, 2021, p. 73).

Outro ponto estratégico do *Prime Video* é o investimento em franquias e títulos de grande apelo comercial, combinados com produções de nicho que exploram o conceito de cauda longa. Conforme Mendes (2020, p. 88), "a plataforma equilibra lançamentos globais de alto impacto com conteúdos de nicho, garantindo variedade e satisfação de diferentes segmentos de assinantes". Essa estratégia contribui para que a *Amazon* consolide sua presença no mercado competitivo de *streaming* e aumente o valor percebido pelo consumidor.

Além disso, o *Prime Video* busca inovação constante na forma de distribuição e no tipo de conteúdo disponibilizado, incluindo séries interativas e produções que cruzam diferentes mídias. Segundo Ribeiro (2021, p. 65), "a inovação em formatos e experiências de consumo digital é uma ferramenta importante para diferenciar o *Prime Video*, mantendo a plataforma relevante frente a concorrentes globais".

2.3 Disney +

O *Disney+* é a plataforma de streaming da *The Walt Disney Company*, lançada oficialmente em novembro de 2019, com foco em produções próprias e no aproveitamento de seu extenso portfólio de marcas. Segundo Andrade (2020, p. 41), "o lançamento do *Disney+* representou um movimento estratégico de integração vertical, permitindo que a empresa controlasse integralmente a distribuição de seu conteúdo e reduzisse a dependência de terceiros". Essa abordagem possibilitou à *Disney* centralizar seu acervo e oferecer exclusividade aos assinantes, criando um diferencial competitivo no mercado.

O catálogo do *Disney+* é caracterizado pela presença de produções icônicas das marcas *Disney*, *Pixar*, *Marvel*, *Star Wars* e *National Geographic*, abrangendo conteúdos voltados para todas as idades, com forte apelo familiar. De acordo com Moreira (2021, p. 55), "o *Disney+* constrói sua proposta de valor a partir de propriedades intelectuais consagradas, que já possuem reconhecimento e fidelidade do público, facilitando a retenção de assinantes e atraiendo novos consumidores". Essa estratégia de consolidar franquias estabelecidas garante um posicionamento sólido e reduz riscos de investimento.

Além de seu acervo clássico, a plataforma investe em conteúdos originais exclusivos, com produções que expandem universos narrativos de suas franquias mais populares. Conforme Pereira (2020, p. 68), "as produções originais do *Disney+* desempenham papel essencial na renovação do catálogo, estimulando o engajamento dos fãs e criando novas oportunidades de monetização". Essa prática também permite à empresa lançar séries e filmes com lançamento simultâneo em diversas regiões, fortalecendo sua presença global.

A estratégia do *Disney+* também considera as preferências culturais e regionais, incorporando legendas, dublagens e adaptações específicas para diferentes mercados. Segundo Costa (2021, p. 62), "a adaptação cultural do conteúdo fortalece a proximidade da marca com o público local, mantendo a essência das produções e respeitando as particularidades de cada região". Esse cuidado contribui para ampliar o alcance da plataforma e reforçar a experiência do assinante.

Por fim, a *Disney+* aposta em uma experiência de consumo sem anúncios e com interface intuitiva, voltada para facilitar o acesso aos conteúdos e promover a exploração do catálogo. De acordo com Lima (2020, p. 84), “a combinação de experiência do usuário simplificada e catálogo exclusivo constitui um dos pilares do sucesso do *Disney+*, diferenciando-o em um mercado cada vez mais competitivo”.

2.4 HBO Max

A *HBO Max*, serviço de *streaming* lançado pela WarnerMedia em maio de 2020, reúne conteúdos da marca HBO, além de produções da *Warner Bros.*, *DC*, *Cartoon Network*, *Max Originals* e outros estúdios parceiros. Segundo Mendes (2021, p. 39), “a *HBO Max* nasce com o objetivo de unificar propriedades intelectuais de alto valor em um único ecossistema digital, oferecendo diversidade e exclusividade para competir diretamente com os principais players do mercado”. Essa abordagem estratégica aproveita décadas de produções consagradas para fortalecer o apelo da marca.

O diferencial da *HBO Max* está em sua reputação por produções originais de alta qualidade, especialmente séries aclamadas pela crítica e pelo público. Conforme Souza (2020, p. 54), “a marca HBO construiu sua identidade a partir de produções com elevado padrão narrativo e estético, criando uma percepção de valor *premium* no mercado audiovisual”. Essa identidade foi transportada para o ambiente digital, ampliando seu alcance e mantendo a consistência na entrega de conteúdo.

Além dos conteúdos originais, a *HBO Max* investe na aquisição de direitos de distribuição de grandes franquias e sucessos de bilheteria, complementando seu catálogo com opções variadas que atendem a diferentes perfis de público. De acordo com Rocha (2021, p. 63), “a estratégia de mesclar conteúdo exclusivo com títulos licenciados aumenta o potencial de atratividade da plataforma e favorece a retenção de assinantes a longo prazo”.

A plataforma também se diferencia pelo investimento em lançamentos simultâneos de filmes de cinema diretamente no serviço de *streaming* em algumas regiões, especialmente durante períodos de restrições sanitárias. Segundo Almeida (2020, p. 72), “a prática de estrear produções de grande orçamento diretamente no streaming acelerou a transformação dos modelos tradicionais de distribuição, oferecendo conveniência ao público e vantagem competitiva para a marca”.

Por fim, a *HBO Max* busca consolidar-se internacionalmente por meio de adaptações de interface, legendas e dublagens específicas para cada mercado, além de produções originais locais que dialoguem com o público regional. Conforme Martins (2021, p. 81), “a estratégia de conteúdo local aliado à força de franquias globais reforça a presença da *HBO Max* no cenário competitivo, criando uma proposta equilibrada entre relevância cultural e apelo global”.

2.5 Definição e Aplicações da Análise de Dados em Streamings

A análise de dados em plataformas de *streaming* consiste no processo de coleta, organização, tratamento e interpretação de informações geradas pelo

consumo de conteúdos audiovisuais, com o objetivo de compreender padrões, identificar tendências e apoiar decisões estratégicas. Essas informações podem incluir métricas como tempo de exibição, gênero e classificação etária dos títulos, origem das produções, avaliações dos usuários e variação do catálogo ao longo do tempo. Segundo Davenport e Harris (2017, p. 52), “as empresas que utilizam a análise de dados de forma estratégica conseguem criar vantagens competitivas, antecipando demandas e melhorando a eficiência operacional”.

O uso de análise de dados nesse contexto vai além do simples acompanhamento de métricas, permitindo compreender o comportamento do público e adaptar a oferta de conteúdo de acordo com as preferências identificadas. Plataformas como *Netflix*, *Amazon Prime Video*, *Disney+* e *HBO Max* utilizam extensivamente esses recursos para definir quais produções adquirir, quais originais desenvolver e como promover seus títulos. De acordo com Provost e Fawcett (2013, p. 91), “o valor real da análise de dados está em transformar grandes volumes de informação em ações práticas que impactem os resultados de negócios”.

Outra aplicação relevante está no desenvolvimento de sistemas de recomendação, que utilizam algoritmos baseados em dados para sugerir conteúdos personalizados. Esses sistemas, alimentados por dados históricos e padrões de visualização, aumentam o engajamento e a retenção de assinantes. Conforme Gomez-Uribe e Hunt (2016, p. 8), “os sistemas de recomendação são responsáveis por uma parcela substancial do consumo nas plataformas de streaming, moldando a experiência do usuário e influenciando diretamente o sucesso do serviço”.

No campo da pesquisa e análise comparativa, a aplicação de técnicas de análise de dados possibilita investigar e confrontar a composição de catálogos entre diferentes serviços. Por exemplo, é possível identificar qual plataforma oferece maior variedade de gêneros, maior número de títulos originais ou maior diversidade de produções internacionais. Essa abordagem é potencializada pelo uso de *datasets* públicos, como os disponibilizados no *Kaggle*, que reúnem informações padronizadas e estruturadas sobre diferentes plataformas. Segundo Chen, Chiang e Storey (2012, p. 118), “a visualização de dados desempenha papel central na compreensão de padrões complexos, auxiliando na tomada de decisão e no desenvolvimento de estratégias”.

Neste estudo, os dados coletados sobre os catálogos da *Netflix*, *Amazon Prime Video*, *Disney+* e *HBO Max* serão analisados por meio do *Power BI*, ferramenta de Business Intelligence que permite a criação de dashboards interativos e relatórios dinâmicos. Essa metodologia permitirá identificar tendências, mapear similaridades e diferenças entre os serviços e oferecer um panorama comparativo do mercado de streaming no contexto atual.

2.6 Requisitos e Características de uma Análise de Dados Eficiente em Streamings

Para que a análise de dados em plataformas de *streaming* seja eficiente e gere resultados relevantes, é necessário atender a um conjunto de requisitos que assegurem a qualidade das informações e a precisão das interpretações. Um dos primeiros aspectos é a integridade e a consistência dos dados utilizados, evitando problemas como duplicidade de registros, campos incompletos ou informações desatualizadas. Segundo Pressman (2015, p. 34), “os requisitos de um sistema ou

processo analítico precisam ser claramente definidos para assegurar que o resultado atenda às necessidades dos usuários e mantenha sua escalabilidade". No contexto de *streamings*, isso significa garantir que o conjunto de dados reflita de forma fiel o catálogo disponível, incluindo atributos como título, ano de lançamento, gênero, duração, classificação indicativa e tipo de produção (original ou licenciada).

Outro requisito essencial é a padronização das informações, uma vez que diferentes plataformas podem registrar dados de maneiras distintas. A normalização de campos como gêneros, formatos e países de origem é fundamental para que a comparação entre serviços seja justa e precisa. De acordo com Connolly e Begg (2015, p. 56), "a estruturação e a padronização adequadas dos dados são pré-requisitos para que análises comparativas sejam consistentes e confiáveis". Sem essa uniformidade, os resultados podem ser enviesados e comprometer as conclusões do estudo.

Além disso, uma análise eficiente deve se apoiar em ferramentas capazes de transformar grandes volumes de dados em informações visuais e de fácil interpretação. O uso de recursos de *Business Intelligence*, como o *Power BI*, permite criar *dashboards* interativos, painéis comparativos e relatórios dinâmicos, facilitando a identificação de padrões e tendências. Conforme Chen, Chiang e Storey (2012, p. 118), "a visualização de dados desempenha papel central na compreensão de padrões complexos, auxiliando na tomada de decisão e no desenvolvimento de estratégias de mercado".

Outro fator relevante é a definição de métricas claras e objetivas, como quantidade total de títulos por plataforma, distribuição por gênero, média de duração dos conteúdos, proporção de produções originais e diversidade geográfica. Essas métricas permitem não apenas comparar quantitativamente os catálogos, mas também avaliar a estratégia de cada serviço. Segundo Davenport e Harris (2017, p. 52), "a análise de dados é mais eficaz quando baseada em indicadores bem definidos, que permitam medir o desempenho e gerar *insights* açãoáveis".

Portanto, uma análise de dados eficiente em *streamings* exige não apenas a disponibilidade de informações confiáveis, mas também a aplicação de técnicas adequadas de limpeza, padronização e visualização, garantindo que os resultados obtidos sejam claros, comparáveis e relevantes para a compreensão do mercado.

3. Ferramentas utilizadas

O desenvolvimento deste trabalho exigiu a utilização de um conjunto de ferramentas capazes de lidar com diferentes etapas do processo de análise de dados, desde a coleta até a visualização final das informações. A escolha dessas tecnologias foi feita de forma estratégica, considerando sua eficiência, acessibilidade e integração entre si, de modo a garantir a consistência e a confiabilidade dos resultados obtidos.

Inicialmente, recorreu-se ao *Kaggle* como fonte de dados, devido à sua ampla base de *datasets* públicos de alta qualidade, que possibilitaram a coleta de informações referentes aos catálogos das plataformas de *streaming Netflix, Amazon Prime Video, Disney+* e *HBO Max*. Em seguida, a linguagem *Python* foi empregada

para a etapa de limpeza, transformação e preparação dos dados, aproveitando sua versatilidade e bibliotecas específicas voltadas à ciência de dados.

Para o armazenamento e organização das informações, utilizou-se o *SQL Server*, que possibilitou a estruturação em tabelas relacionais, garantindo integridade e facilidade de acesso. Por fim, a ferramenta *Power BI* foi escolhida para a análise exploratória e a construção de visualizações interativas, permitindo interpretar os resultados de maneira clara e acessível.

Assim, a combinação dessas ferramentas resultou em um fluxo metodológico coeso, que possibilitou transformar dados brutos em *insights* significativos sobre os catálogos de filmes e séries nos serviços de *streaming* analisados.

3.1 Kaggle

O *Kaggle* foi a plataforma escolhida como fonte de dados para este estudo, fornecendo os *datasets* necessários para a análise comparativa entre os catálogos de filmes e séries das plataformas *Netflix*, *Amazon Prime Video*, *Disney+* e *HBO Max*. Reconhecido como uma das maiores comunidades de ciência de dados do mundo, o *Kaggle* disponibiliza não apenas competições voltadas para aprendizado de máquina, mas também um vasto repositório de bases de dados públicas utilizadas por pesquisadores, estudantes e profissionais. Segundo Millman (2019, p. 44), “o *Kaggle* consolidou-se como um ecossistema essencial para cientistas de dados, reunindo *datasets*, *notebooks* e uma comunidade colaborativa que impulsiona a inovação na área”.

No contexto deste trabalho, o *Kaggle* forneceu quatro *datasets* em formato CSV, contendo informações estruturadas sobre os catálogos das plataformas de streaming, como título, ano de lançamento, gênero, país de origem, duração e classificação indicativa. Esses atributos foram fundamentais para organizar e comparar os acervos, possibilitando a identificação de padrões e tendências. De acordo com Provost e Fawcett (2013, p. 27), “a qualidade e a disponibilidade dos dados são determinantes para o sucesso de qualquer análise, pois representam a base sobre a qual o conhecimento será extraído”.

Outro fator determinante para a escolha do *Kaggle* foi a confiabilidade e a praticidade de acesso aos *datasets*, já que a plataforma oferece material de fácil integração com linguagens de programação como *Python* e R. Além disso, o caráter colaborativo da comunidade garante constante atualização e verificação das bases. Conforme Sharda, Delen e Turban (2018, p. 61), “ambientes colaborativos de compartilhamento de dados fortalecem a ciência aberta, garantindo maior confiabilidade e reproduzibilidade das análises”.

Dessa forma, o *Kaggle* desempenhou um papel fundamental na etapa inicial da pesquisa, funcionando como ponto de partida para todas as demais fases do processo metodológico, que envolveram a limpeza, transformação, armazenamento e análise visual das informações coletadas.

3.1 Python como Linguagem de Programação

Python é uma das linguagens de programação mais populares e amplamente utilizadas no desenvolvimento de sistemas web, aplicações comerciais e soluções de automação. Sua versatilidade e simplicidade tornam-no uma escolha ideal para programadores de diferentes níveis de experiência, desde iniciantes até desenvolvedores avançados. Criado por Guido van Rossum no final da década de 1980, Python foi projetado com o objetivo de oferecer uma sintaxe clara e intuitiva, permitindo que os desenvolvedores escrevam códigos mais limpos e fáceis de manter. Segundo Downey (2015, p. 9), “*Python* se destaca por sua simplicidade e expressividade, permitindo que os desenvolvedores se concentrem mais na solução do problema do que nas complexidades da sintaxe”.

Uma das principais vantagens do *Python* é sua facilidade de aprendizado e uso. Comparado a outras linguagens como Java e C++, *Python* exige menos linhas de código para realizar as mesmas operações, o que torna o desenvolvimento mais rápido e eficiente. Além disso, sua sintaxe simples e organizada favorece a legibilidade do código, reduzindo a curva de aprendizado para novos programadores. Segundo Zelle (2016, p. 4), “a clareza e a estrutura intuitiva do *Python* fazem dele uma excelente linguagem tanto para o ensino de programação quanto para o desenvolvimento profissional”.

Outro grande diferencial do *Python* é seu extenso ecossistema de bibliotecas e *frameworks*, que facilitam a criação de soluções robustas e escaláveis. No desenvolvimento web, *frameworks* como *Flask* e *Django* permitem a construção de aplicações de forma ágil, com recursos integrados para manipulação de dados, autenticação de usuários e segurança. Já no campo da análise de dados, bibliotecas como *Pandas* e *NumPy* oferecem ferramentas poderosas para manipulação e processamento de grandes volumes de informações. Segundo McKinney (2017, p. 5), “o sucesso do *Python* no mundo dos dados se deve, em grande parte, ao seu vasto conjunto de bibliotecas especializadas, que permitem realizar desde estatísticas básicas até aprendizado de máquina avançado”.

Além de ser uma linguagem acessível e poderosa, *Python* também se destaca por sua escalabilidade. Ele pode ser utilizado tanto para pequenos *scripts* e automações quanto para sistemas complexos que exigem alto desempenho. Segundo Severance (2016, p. 10), “*Python* tem sido cada vez mais adotado por empresas devido à sua flexibilidade, permitindo um desenvolvimento rápido e fácil manutenção de aplicações”. Sua compatibilidade com diversas plataformas e sua capacidade de integração com outras linguagens e tecnologias tornam-no uma opção versátil para diferentes tipos de projetos.

Outro fator essencial que contribui para a popularidade do *Python* é sua ampla comunidade de suporte. O *Python* conta com uma vasta documentação, fóruns de discussão e recursos educacionais disponíveis gratuitamente. Essa comunidade ativa facilita a resolução de problemas, além de impulsionar a constante evolução da linguagem. Segundo Al Sweigart (2019, p. 7), “a forte comunidade de desenvolvedores ao redor do *Python* garante que sempre haverá suporte, bibliotecas atualizadas e melhorias contínuas na linguagem”.

3.2 SQL Server

O SQL Server foi utilizado neste estudo como sistema de gerenciamento de banco de dados relacional (SGBD), desempenhando um papel central no

armazenamento e organização dos dados previamente tratados em *Python*. Desenvolvido pela Microsoft, o *SQL Server* é amplamente reconhecido por sua robustez, escalabilidade e capacidade de lidar com grandes volumes de informações, além de oferecer recursos avançados de segurança e integridade referencial. De acordo com Connolly e Begg (2015, p. 42), “os bancos de dados relacionais fornecem a base para o armazenamento eficiente de dados, garantindo consistência, confiabilidade e acesso rápido às informações armazenadas”.

No presente trabalho, o *SQL Server* foi responsável por receber os *datasets* já limpos e padronizados, estruturando-os em tabelas organizadas que permitiram consultas rápidas e eficazes. Essa organização possibilitou a segmentação dos dados em diferentes dimensões, como ano de lançamento, gênero, classificação indicativa e país de origem, facilitando o cruzamento de informações entre as plataformas de streaming analisadas. Conforme Elmasri e Navathe (2011, p. 6), “o modelo relacional de bancos de dados é um dos mais eficazes para garantir integridade e minimizar redundâncias, tornando-o essencial em sistemas que exigem alto nível de confiabilidade”.

Além do armazenamento, o *SQL Server* desempenhou um papel relevante na preparação dos dados para integração com o *Power BI*. Por meio de consultas em linguagem SQL, foi possível refinar os registros, filtrar atributos específicos e organizar as tabelas de modo a otimizar a análise posterior. Esse processo assegurou que o *Power BI* recebesse dados consistentes e bem estruturados, evitando problemas como duplicidade de informações ou ausência de campos relevantes. Segundo Coronel e Morris (2017, p. 58), “a utilização de consultas SQL em bancos relacionais permite que os dados sejam manipulados de forma precisa, contribuindo para análises mais confiáveis e consistentes”.

Outro aspecto relevante é a segurança e a confiabilidade oferecidas pelo *SQL Server*. Como os dados utilizados neste estudo representam informações detalhadas sobre os catálogos das plataformas de streaming, foi essencial garantir que não houvesse perda, corrupção ou inconsistência durante o processo de armazenamento. O SGBD assegurou integridade transacional e controle de acesso, permitindo um gerenciamento eficiente do conjunto de dados. Conforme Date (2019, p. 73), “os sistemas de gerenciamento de banco de dados relacionais garantem não apenas a integridade dos dados, mas também a segurança necessária para seu uso em ambientes críticos e de alta demanda”.

Dessa forma, a utilização do *SQL Server* no presente estudo foi fundamental para consolidar os dados oriundos dos *datasets* do *Kaggle* em uma base confiável e estruturada, possibilitando sua integração direta ao *Power BI*. Essa etapa metodológica assegurou que a análise comparativa entre *Netflix*, *Amazon Prime Video*, *Disney+* e *HBO Max* fosse realizada sobre informações consistentes e organizadas, resultando em conclusões mais precisas e relevantes.

3.3 Power BI

No cenário atual de transformação digital, o uso de ferramentas que permitem a coleta, organização, análise e visualização de dados se tornou essencial para a geração de insights estratégicos. Dentre essas ferramentas, o *Power BI* destaca-se por sua capacidade de integrar múltiplas fontes de dados, transformar informações brutas em relatórios visuais interativos e facilitar a tomada de decisões

em tempo real. Desenvolvido pela Microsoft, o *Power BI* é amplamente adotado por organizações de diferentes portes e setores, principalmente por sua interface intuitiva, conectividade com diversos bancos de dados e recursos de atualização automática.

Segundo Rezaei e Salehi (2020, p. 45), “o *Power BI* contribui para a democratização da análise de dados, possibilitando que profissionais de diferentes áreas tomem decisões baseadas em informações atualizadas e visualmente comprehensíveis”. A ferramenta oferece suporte a diversas funcionalidades como gráficos dinâmicos, mapas geográficos, segmentações de dados e painéis customizáveis, tornando o processo de análise mais acessível até mesmo para usuários sem conhecimento técnico avançado.

O *Power BI* opera com base em três pilares principais: conexão com fontes de dados, modelagem e transformação dos dados e visualização e compartilhamento. A etapa de conexão permite a importação de dados provenientes de fontes diversas, como planilhas Excel, bancos de dados SQL, serviços online (como *Google Analytics*, *Azure* e *Salesforce*), entre outros. Em seguida, na etapa de transformação, o *Power BI* utiliza o Power Query para limpar, filtrar e estruturar os dados, assegurando maior consistência e confiabilidade das análises. Por fim, a criação de dashboards interativos permite que os usuários acompanhem indicadores-chave de desempenho (KPIs), explorem tendências e realizem análises comparativas com facilidade.

Além disso, o *Power BI* possui uma linguagem própria chamada DAX (*Data Analysis Expressions*), que permite a criação de fórmulas personalizadas e métricas complexas, expandindo as possibilidades analíticas da ferramenta. Essa linguagem é particularmente útil na construção de medidas personalizadas, como médias móveis, percentuais de crescimento, comparações ano a ano, entre outras.

No contexto de análise de dados em *marketplaces*, como a *Amazon*, o uso do *Power BI* revela-se especialmente eficaz. A ferramenta permite transformar grandes volumes de dados transacionais em visualizações claras e objetivas, facilitando a identificação de padrões de compra, produtos mais vendidos, sazonalidade de vendas e comportamento dos consumidores. Dessa forma, além de apresentar os resultados de forma visualmente atrativa, o *Power BI* contribui para uma compreensão mais profunda do mercado digital e para a formulação de estratégias comerciais baseadas em evidências.

Outro aspecto relevante é a possibilidade de compartilhamento dos dashboards em nuvem por meio do *Power BI Service*, permitindo que diferentes membros de uma equipe acessem os relatórios de forma colaborativa, com atualizações automáticas e controle de permissões. Essa funcionalidade reforça o papel do *Power BI* como uma ferramenta de apoio à gestão e ao trabalho em equipe.

Portanto, o *Power BI* não se limita à visualização de dados, mas atua como um componente estratégico no processo decisório. Sua capacidade de integrar, transformar e comunicar dados de forma eficaz o torna um recurso indispensável em projetos que exigem clareza, agilidade e precisão na análise de informações.

3.3 Banco de dados SQL Server

Para o armazenamento e gerenciamento das informações do sistema de vendas físicas, optou-se pela utilização do *Microsoft SQL Server*. Esta escolha se

justifica pela confiabilidade e robustez da plataforma, amplamente reconhecida no mercado corporativo, bem como pelo seu extenso suporte a operações de manipulação de dados. O *SQL Server* é capaz de gerenciar grandes volumes de informações com alta disponibilidade e segurança, oferecendo mecanismos eficientes para a execução de transações complexas e simultâneas, o que é essencial para sistemas de vendas que demandam integridade e precisão nas operações.

Além disso, a plataforma disponibiliza recursos avançados para a criação de tabelas relacionais, a execução de consultas SQL otimizadas e a manutenção da integridade referencial dos dados. Esses aspectos são fundamentais para garantir a consistência e a confiabilidade das informações armazenadas, assegurando que o sistema opere de forma eficiente e sem inconsistências. Dessa maneira, o *Microsoft SQL Server* se apresenta como uma solução ideal para suportar as necessidades técnicas do sistema desenvolvido (Microsoft, 2024, p. 3).

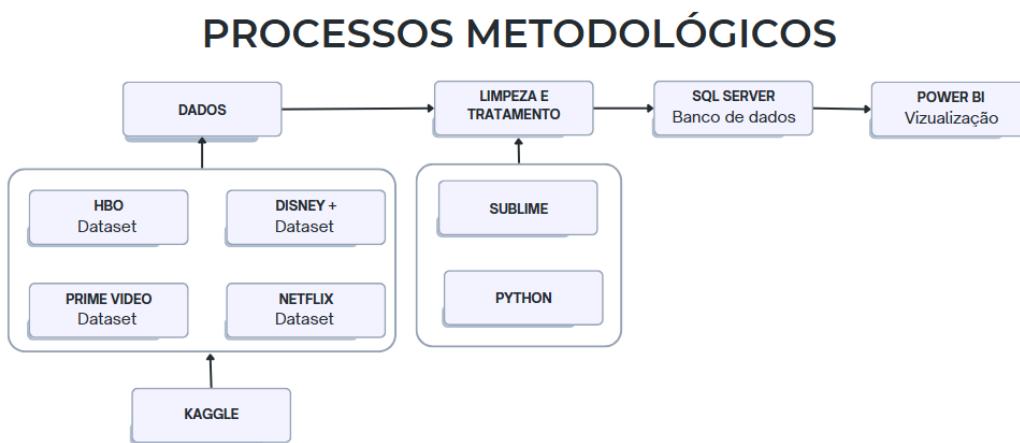
3 Desenvolvimento

O desenvolvimento deste estudo envolveu a aplicação prática das ferramentas descritas anteriormente, resultando em um fluxo de trabalho que partiu da coleta dos *datasets* até a análise visual dos dados no *Power BI*, conforme pode ser visualizado na Figura 1. Essa etapa foi essencial para transformar os dados brutos provenientes do *Kaggle* em informações organizadas e comparáveis, possibilitando uma visão clara sobre os catálogos de filmes e séries das plataformas *Netflix*, *Amazon Prime Video*, *Disney+* e *HBO Max*.

Os procedimentos metodológicos constituem a base estruturante de uma pesquisa científica, pois definem as etapas e os instrumentos necessários para alcançar os objetivos propostos. Neste estudo, a metodologia, que está apresentada na Figura 1, foi planejada de forma a garantir a confiabilidade dos resultados e a consistência do processo de análise, partindo da coleta dos dados em bases públicas até a transformação desses registros em informações comparativas e de fácil interpretação. De acordo com Gil (2019, p. 43), “a metodologia representa o caminho a ser seguido pelo pesquisador, estabelecendo os métodos e técnicas que possibilitam a obtenção do conhecimento científico”.

A análise de dados aplicada a serviços de streaming demanda não apenas a seleção adequada de bases de dados, mas também a utilização de ferramentas que permitam o tratamento, armazenamento e visualização de grandes volumes de informações. No caso deste estudo, foram utilizados *datasets* disponibilizados no *Kaggle*, que reúnem dados sobre os catálogos de filmes e séries da *Netflix*, *Amazon Prime Video*, *Disney+* e *HBO Max*. Esses conjuntos foram escolhidos por apresentarem informações estruturadas, como título, ano de lançamento, gênero, duração, classificação indicativa e origem, possibilitando comparações entre diferentes plataformas. Segundo Davenport e Harris (2017, p. 52), “a análise de dados é mais eficaz quando baseada em informações confiáveis e bem estruturadas, pois permite transformar grandes volumes de registros em insights açãoáveis”.

Figura 1. Representação gráfica da metodologia de trabalho



Fonte: Os autores, 2025.

A metodologia adotada seguiu um fluxo organizado em quatro etapas principais. Conforme pode ser visto na Figura 1, primeiramente, ocorreu a coleta dos dados no *Kaggle*, considerada uma fonte relevante de *datasets* para estudos acadêmicos e de mercado. Em seguida, os dados foram submetidos ao processamento e limpeza em Python, linguagem reconhecida por sua simplicidade sintática e seu ecossistema de bibliotecas voltadas para ciência de dados. Conforme McKinney (2017, p. 14), “Python se consolidou como a linguagem padrão para análise de dados por aliar facilidade de uso e capacidade de lidar com grandes volumes de informação”.

Posteriormente, os dados tratados foram armazenados no SQL Server, banco de dados relacional escolhido por sua robustez, escalabilidade e confiabilidade no armazenamento de informações estruturadas. A utilização dessa ferramenta garantiu que os registros pudessem ser organizados em tabelas consistentes, evitando duplicidades e assegurando integridade. Segundo Connolly e Begg (2015, p. 42), “bancos de dados relacionais oferecem a infraestrutura necessária para manipular informações de maneira segura e eficiente, sendo essenciais em sistemas que exigem consistência e precisão”.

Por fim, a análise exploratória e a visualização dos resultados foram realizadas por meio do Power BI, que possibilitou a criação de *dashboards* interativos e relatórios dinâmicos. Essa etapa foi responsável por transformar os dados armazenados em informações acessíveis e comparativas, revelando tendências, padrões e diferenças entre os catálogos das plataformas de streaming analisadas. Conforme Ferrari e Russo (2016, p. 12), “o Power BI democratiza a análise de dados ao oferecer representações visuais intuitivas que tornam os resultados comprehensíveis até mesmo para usuários não técnicos”.

Dessa forma, os procedimentos metodológicos deste estudo integraram técnicas de ciência de dados, gestão de banco de dados e ferramentas de *Business Intelligence* em um fluxo estruturado, permitindo que os *datasets* brutos fossem transformados em análises claras e relevantes. A combinação dessas etapas garantiu

não apenas a qualidade da pesquisa, mas também a confiabilidade das conclusões obtidas.

De acordo com Pressman (2015, p. 34), “o processo de desenvolvimento de sistemas ou aplicações deve seguir um conjunto de etapas bem definidas, garantindo que os resultados sejam consistentes, confiáveis e alinhados aos objetivos propostos”. Dessa forma, o presente trabalho estruturou o desenvolvimento em fases que contemplaram a coleta das bases, a preparação do ambiente de programação, o tratamento e padronização dos dados em Python, o armazenamento em banco de dados SQL Server e, por fim, a análise exploratória e visualização no Power BI.

Cada uma dessas etapas foi planejada de modo a assegurar a integridade e consistência dos dados, permitindo que os resultados da análise comparativa fossem obtidos de forma precisa e fundamentada. A seguir, apresentam-se as fases que compõem o processo de desenvolvimento, iniciando pela coleta das bases no Kaggle e a configuração do ambiente de trabalho.

3.1 Coleta dos Dados e Preparação do Ambiente

A primeira etapa do desenvolvimento foi dedicada à coleta dos dados que serviriam de base para a análise comparativa entre os serviços de streaming *Netflix*, *Amazon Prime Video*, *Disney+* e *HBO Max*. Para isso, foram utilizados quatro datasets públicos disponibilizados na plataforma *Kaggle*, que é reconhecida por reunir uma ampla variedade de bases de dados em formato aberto, voltadas para estudos acadêmicos e aplicações práticas. Os arquivos, disponibilizados em formato CSV, apresentavam registros detalhados dos catálogos de cada uma das plataformas, contendo atributos como título, ano de lançamento, gênero, classificação indicativa, tempo de duração, país de origem e informações sobre produções originais ou licenciadas.

Após a coleta, os arquivos foram organizados em diretórios específicos, de modo a facilitar sua manipulação nas etapas seguintes. Essa organização foi fundamental para garantir a rastreabilidade dos dados e a consistência do fluxo metodológico, já que cada *dataset* representava uma plataforma distinta e exigia atenção particular em sua manipulação.

Com os dados em mãos, passou-se para a preparação do ambiente de desenvolvimento. Optou-se pela utilização do editor de código *Sublime Text*, que, apesar de simples e leve, oferece suporte adequado para a escrita de scripts em *Python*, além de permitir integração com bibliotecas necessárias ao tratamento de dados. Nesse ambiente, foram criados os arquivos de código que fariam a leitura dos datasets, a conexão com o banco de dados SQL Server e a aplicação dos primeiros comandos de verificação da estrutura dos arquivos. Os códigos estão apresentados na Figura 2.

Figura 2. Conexão com banco de dados

```
# Directory Job
diret="c:/FILMES/insert/"

# Create File Log
datahora = datetime.now()
datahora = datahora.strftime('%d/%m/%Y %H:%M:%S')
dataHoje = datetime.today()
dataHoje = dataHoje.strftime('%d/%m/%Y')

Local_log = diret+Log/Log[0:2]+datahora[3:5]+datahora[6:10] +'.txt'

# Data connect localhost
server = 'localhost\SQLEXPRESS'
database = 'DB_FILMES'
username = 'sa'
password = '7854&*%$'

# Connect SQL
con = pyodbc.connect('DRIVER={ODBC Driver 17 for SQL Server};SERVER=' + server +';DATABASE=' + database +';UID=' + username +';PWD=' + password)
cursor = con.cursor()
```

Fonte: Os autores, 2025.

Durante essa fase inicial, foram desenvolvidos scripts para testar a leitura dos arquivos CSV, identificar o número de registros disponíveis em cada *dataset* e verificar a existência de colunas com valores inconsistentes ou nulos. Além disso, foi estabelecida a conexão entre o ambiente *Python* e o *SQL Server*, etapa fundamental para permitir que, após a limpeza e o tratamento dos dados, os registros pudessem ser transferidos para o banco de dados de forma automatizada.

Figura 3. Limpeza e processamento

```
leitor = csv.DictReader(arquivo_csv, delimiter=',', dialect="unix")
for coluna in leitor:

    titulo = coluna["title"]
    tipo = coluna["type"]
    descricao = coluna["description"]
    ano_lancamento = coluna["release_year"]
    classificacao = coluna["age_certification"]
    tempo = coluna["runtime"]
    generos = coluna["genres"]
    if generos:
        generos = generos.strip()
        generos = generos.replace("[", "").replace("]", "").replace("'", "")
        generos = generos.split(",")[-1].strip()

    pais_origem = coluna["production_countries"]
    if pais_origem:
        pais_origem = pais_origem.strip()
        pais_origem = pais_origem.replace("[", "").replace("]", "").replace("'", "")
        pais_origem = pais_origem.split(",")[-1].strip()

    temporadas = coluna["seasons"]
    if temporadas == '':
        temporadas = '0'

    imbd_id = coluna["imdb_id"]
    imbd_score = coluna["imdb_score"]
    imbd_votes = coluna["imdb_votes"]

    if imbd_votes == '':
        imbd_votes = '0'

    if imbd_score == '':
        imbd_score = '0'

    local = ''
    if Obs == 'HBO':
        local = 'HBO'
    if Obs == 'NETFLIX':
        local = 'NETFLIX'
    if Obs == 'AMAZON':
        local = 'AMAZON'
    if Obs == 'DISNEY':
        local = 'DISNEY'

    SQL = 'INSERT INTO '+Tabela+' (TITULO, TIPO, DESCRIÇÃO, ANO_LANÇAMENTO, CLASSIFICAÇÃO, TEMPO, GÊNEROS, PAIS_ORIGEM, TEMPORADAS, IMBD_ID, \
IMBD_SCORE, IMBD_VOTES, LOCAL) \
VALUES (?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?)'
    VALUES=(titulo, tipo, descricao, ano_lancamento, classificacao, tempo, generos, pais_origem, temporadas, imbd_id, imbd_score, imbd_votes, local)
```

Fonte: Os autores, 2025.

Assim, a etapa de coleta e preparação do ambiente estabeleceu a base do trabalho, garantindo que os dados estivessem acessíveis, organizados e prontos para o processo de tratamento e padronização. A partir desse ponto, tornou-se

possível avançar para a fase seguinte, que consistiu na limpeza e transformação dos dados em *Python*, etapa essencial para eliminar inconsistências e preparar os registros para o armazenamento definitivo no *SQL Server*.

No processo de tratamento dos dados, mostrado na Figura 3, foram aplicadas diversas técnicas de limpeza, padronização e preenchimento de valores ausentes, a fim de garantir maior consistência e confiabilidade das informações antes da inserção no banco de dados. Inicialmente, nos campos de gênero e país de origem, foram removidos caracteres indesejados como colchetes e aspas, além de espaços em excesso. Quando havia mais de uma informação nesses campos, apenas o primeiro valor foi considerado, garantindo uniformidade nos registros. Para o campo de temporadas, os valores nulos ou vazios foram substituídos por zero, evitando inconsistências em consultas posteriores. Da mesma forma, nos campos relacionados ao IMDb (nota e quantidade de votos), valores ausentes foram preenchidos com zero, assegurando que todos os registros possuam um valor válido, essas alterações visaram manter a consistência dos dados e não prejudicaram o resultado final, pois a amostra era muito pequena. Já no campo de plataforma de origem, foram realizadas verificações de palavras-chave, permitindo padronizar os valores para categorias únicas como *HBO*, *Netflix*, *Amazon* e *Disney*. Por fim, os dados tratados foram organizados e preparados para inserção em uma estrutura de banco relacional por meio de instruções SQL parametrizadas, reduzindo riscos de erros e promovendo maior integridade na base de dados.

3.2 Banco de dados SQL SERVER

Figura 4. Banco de dados

	TITULO	TIPO	DESCRICAO	ANO_LANÇAMENTO	CLASSIFICACAO	TEMPO	GENEROS	PAIS_ORIGEM	TEMPORADAS	IMDB_ID	IMDB_SOC
1	I'm Glad I Did	MOVIE	Miran Akpinar has not been on the theater stage sin...	2020	123	documentation	TR	0	tt15139398	8.40	
2	3 Logical Exit	MOVIE	A sociological meditation on the different "exits" that...	2020	15	documentation	DK	0		0.00	
3	David Batra: Elephant in The Room	MOVIE	Swedish comedian David Batra gets personal as he...	2020	77	comedy	SE	0	tt11987330	5.90	
4	Lucas Neto in: The End of Christmas	MOVIE		2019	71			0		0.00	
5	Arash's Diary: Voyage	SHOW	Twenty years after their debut, join the beloved mem...	2019	30	documentation	JP	1	tt11434874	7.80	
6	Evanikée Cheppodu	MOVIE	Hari and Harathi are in love, but Harathi's father wo...	2019	133	drama	IN	0	tt10081202	7.20	
7	The House Arrest of Us	SHOW	A long-term couple finally decides to get engaged, a...	2020	26	comedy	PH	1	tt13280838	5.50	
8	Hoshino Gen Dome Tour "POP VIRUS" at Tokyo Dome	MOVIE	All venue tickets sold out on the same day, for 5 maj...	2019	120	documentation	JP	0		0.00	
9	Ave Maryam	MOVIE	A devoted nun who cares for her elder sisters must c...	2019	74	romance	ID	0	tt9307666	6.90	
10	Uncle Naji in UAE	MOVIE	Naji decides with his friends to go on holiday to a m...	2019	95	comedy	AE	0	tt10181594	5.00	
11	Voiceless	MOVIE	The Hausa feature film is a story of love and the pow...	2020	120	thriller		0	tt13149312	6.70	
12	Mayday Life	MOVIE	Mayday LIFE TOUR is linked by four distinct segmen...	2019	112	music	TW	0	tt10346206	6.70	
13	Super Monsters: Vida's First Halloween	MOVIE	The Super Monsters share their Halloween tradition...	2019	24	comedy		0	tt10592756	5.90	
14	Children of Adam	SHOW	Scandals and conflicts of society figures from differe...	2020	40	drama	LB	1	tt12434912	7.70	
15	40 Sticks	MOVIE	A group of prisoners trapped in a crashed prison bu...	2020	93	thriller		0	tt13483050	5.70	
16	Belmonte	MOVIE	Federico Veiro's fourth feature examines the many, ...	2019	75	drama	UY	0	tt8846072	6.10	
17	An Unremarkable Christmas	MOVIE	Before the events of S1, Juanquini, El Nato and Capt...	2020	83	comedy	CO	0	tt13690632	3.20	
18	Nailed It! France	SHOW	On this fun and funny competition show, home breake...	2019	34	reality		1	tt10590804	6.50	
19	Kambili: The Whole 30 Years	MOVIE	Kambili Maduka, is an irresponsible, spoiled, spend...	2020	119	romance		0	tt11831004	6.60	
20	Wave of Cinema: Surat Dari Timur	MOVIE	Featuring a mix of film footage and special perform...	2020	57	music	ID	0		0.00	
21	GIMS: On the Record	MOVIE	Go backstage with beloved rap superstar Gims in th...	2020	96	music	FR	0	tt12682218	6.80	
22	Mighty Little Bheem: Festival of Colors	SHOW	From surprising stage performances to spraying col...	2020	7	animation		1	tt11738790	7.20	
23	Between Maybes	MOVIE	Amidst the silence and isolation, Hazel and Louie fi...	2019	107	drama	PH	0	tt10044704	6.20	
24	Alone/Together	MOVIE	Christina an Art Studies major student at University o...	2019	103	romance	PH	0	tt9526822	6.60	
25	StarBeam: Halloween Hero	MOVIE	Greedy Captain Fishbeard is stealing everybody's H...	2020	33	animation	CA	0		0.00	
26	Yours Sincerely, Kanan Gill	MOVIE	Revisiting life goals set in a letter written as a teen to...	2020	72	comedy	IN	0	tt102425392	7.10	
27	Ladies Up	SHOW	Rising India-based comics Prashasti Singh, Kaneez...	2020	17	reality	IN	1	tt11942100	4.30	
28	Finding Hubby	MOVIE	The adventures of 35-year-old Oyn Clegg and her fri...	2020	104	comedy	NG	0		0.00	
29	We Are One	MOVIE	Activists around the world fight injustice and drive so...	2020	86	documentation	FR	0	tt12588398	4.70	
30	In the Bosom of a Thorn	SHOW	After the invasion of Kuwait, a woman's infant daught...	2019	46	drama		1	tt10397430	7.20	
31	Alieli	MOVIE	The death of their father causes three dysfunctional ...	2019	88	comedy	UY	0	tt10984818	6.30	
32	Ainu Mosir	MOVIE	Kanto, 14, a descendant of Japan's indigenous Ainu...	2020	84	drama	CN	0	tt9140354	6.60	
33	Bhai - Vyakti Ki Valli 2	MOVIE	The second part of the P.L. Deshpande biopic which ...	2019	130	drama	IN	0	tt9680076	7.60	
34	Findina Aaines	MOVIE	On an emotional journey in Morocco, an entrepre...	2020	105	drama	PH	0	tt13328840	5.00	

Fonte: Os autores, 2025.

Após o processamento e tratamento das informações no *Python*, os dados foram estruturados em um banco de dados relacional como mostrado na Figura 4, de forma a garantir maior organização, integridade e facilidade de manipulação.

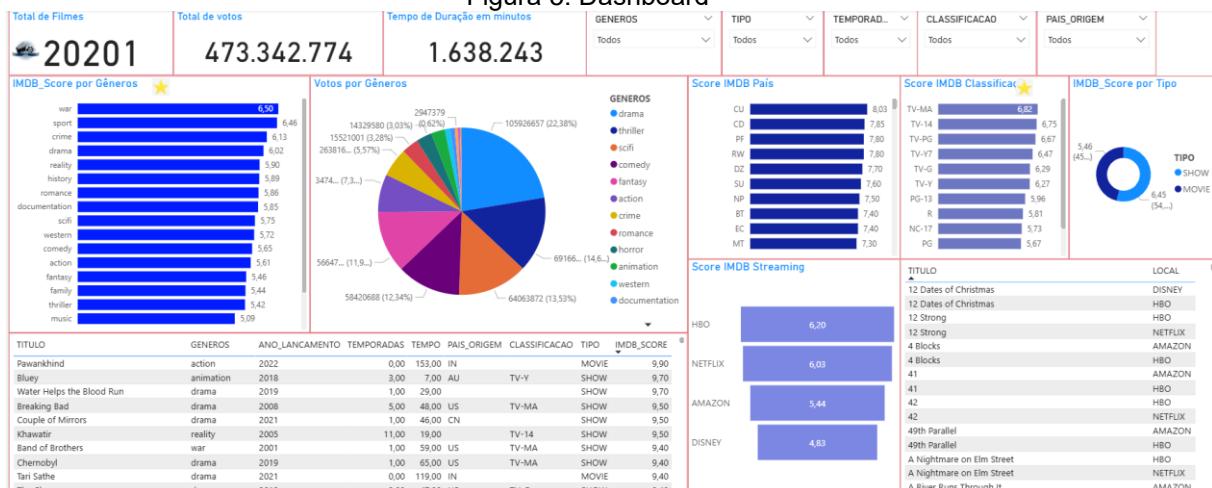
Para isso, foi criada uma tabela que contempla os principais atributos relevantes para a análise dos filmes e séries presentes nas plataformas de streaming.

A estrutura do banco foi definida com campos específicos, como TÍTULO, TIPO (classificação entre filme ou série), DESCRIÇÃO, ANO DE LANÇAMENTO, CLASSIFICAÇÃO INDICATIVA, TEMPO DE DURAÇÃO, GÊNEROS, PAÍS DE ORIGEM, NÚMERO DE TEMPORADAS, além de informações relacionadas à base de dados IMDb, como ID, NOTA (SCORE) e QUANTIDADE DE VOTOS. Também foi incluído o campo PLATAFORMA, padronizado durante o tratamento, que permite identificar de forma clara a qual serviço de streaming cada título pertence.

Essa modelagem possibilitou a criação de uma base de dados consistente e padronizada, permitindo consultas mais eficientes e reduzindo a ocorrência de redundâncias ou inconsistências. A partir dela, tornou-se viável a integração com ferramentas de análise e visualização, como o *Power BI*, possibilitando a geração de relatórios e dashboards interativos que facilitaram a interpretação dos resultados obtidos.

3.3 Power BI

Figura 5. Dashboard



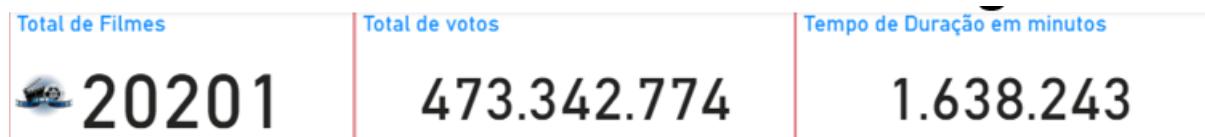
Fonte: Os autores, 2025.

Após o processamento em *Python* e a organização dos registros no banco de dados relacional, os dados foram integrados ao *Power BI*, ferramenta de *Business Intelligence* utilizada para a criação de relatórios e dashboards interativos. Essa etapa teve como objetivo transformar os dados tratados em informações visuais de fácil interpretação, permitindo uma análise mais dinâmica e intuitiva dos filmes e séries presentes nas plataformas de streaming.

No *dashboard* desenvolvido, mostrado na Figura 5, é observado a análise dos dados foi realizada por meio do software *Power BI*, com o objetivo de visualizar e interpretar as informações referentes aos filmes e séries contidos no *dataset*. O painel desenvolvido apresenta uma visão ampla sobre as avaliações, gêneros, classificações indicativas, países de origem e plataformas de streaming. A

seguir, são descritos os principais resultados obtidos a partir dos gráficos e indicadores apresentados no *dashboard*.

Figura 6. Primeiro Conjunto



Fonte: Os autores, 2025.

O primeiro conjunto visto na Figura 6, exibe os totais gerais da base de dados analisada. Foram contabilizados 20.201 filmes, somando 473.342.774 votos e um tempo total de duração de 1.638.243 minutos. Esses valores demonstram a dimensão do conjunto de dados e evidenciam a diversidade de obras consideradas na análise.

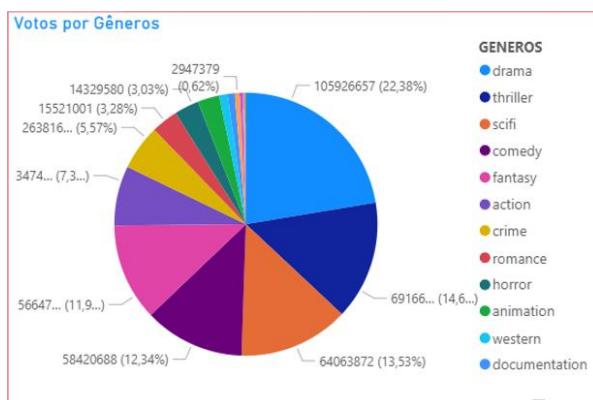
Figura 7. Segundo Conjunto



Fonte: Os autores, 2025.

Em relação às médias de avaliação por gênero, na Figura 7 é possível visualizar o segundo conjunto, onde o gráfico “IMDB Score por Gêneros” apresentou o gênero “War” (guerra) com a maior nota média (6,50), seguido de “Sport” (6,46) e “Crime” (6,32). Tal resultado sugere que produções relacionadas a temas de guerra e esporte tendem a receber melhores avaliações do público, quando comparadas a outros gêneros como drama, romance e comédia, que apresentaram médias ligeiramente inferiores.

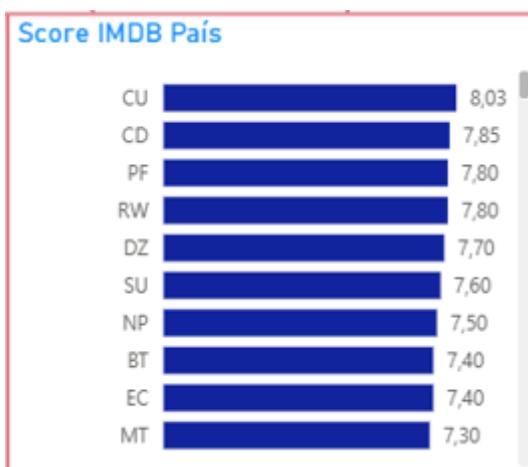
Figura 8. Terceiro Conjunto



Fonte: Os autores, 2025.

Na Figura 8 é possível ver o terceiro conjunto, que representa o gráfico “Votos por Gêneros” evidencia a popularidade das produções entre os espectadores. Observa-se que o gênero “Drama” obteve a maior participação de votos, representando 22,38% do total, seguido por “Comedy” (13,53%) e “Action” (11,94%). Esses resultados indicam que dramas e comédias são os gêneros com maior engajamento do público, refletindo preferências consolidadas entre os espectadores.

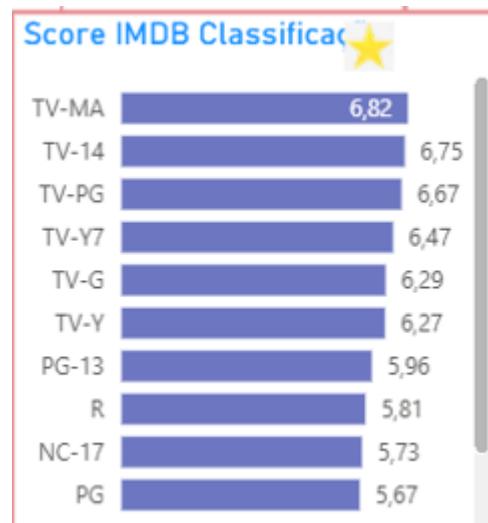
Figura 9. Quarto Conjunto



Fonte: Os autores, 2025.

A Figura 9 possui o quarto conjunto, representando a análise por país de origem, o gráfico “Score IMDB País” mostrou que as produções provenientes de Cuba (8,03) e Canadá (7,85) obtiveram as melhores médias de avaliação, seguidas de obras de países como Alemanha e Rússia, que também apresentaram bons desempenhos. Esse resultado destaca a qualidade percebida de produções de diferentes regiões, indo além dos grandes centros tradicionais como os Estados Unidos.

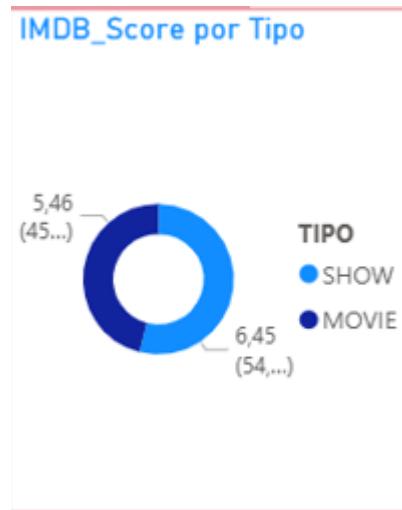
Figura 10. Quinto Conjunto



Fonte: Os autores, 2025.

No gráfico da Figura 10 “Score IMDB Classificação”, observa-se que as produções com classificação indicativa TV-MA (voltadas para o público adulto) alcançaram a maior média de avaliação (6,82). Essa tendência sugere que conteúdos voltados a um público mais maduro tendem a receber avaliações mais positivas, possivelmente em função da complexidade das narrativas e da profundidade temática.

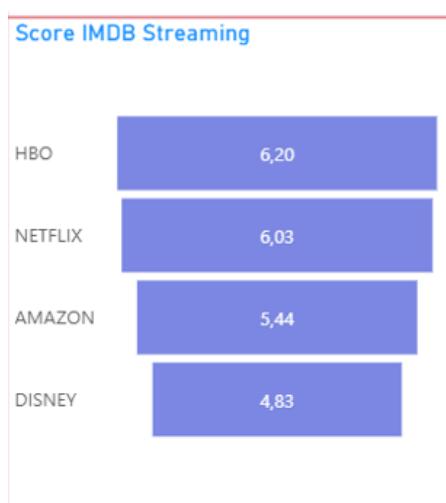
Figura 11. Sexto Conjunto



Fonte: Os autores, 2025.

A Figura 11, no sexto conjunto, possui um gráfico de rosca “IMDB Score por Tipo” apresenta a comparação entre filmes (movies) e séries (shows). Verificou-se que as séries possuem uma média de avaliação ligeiramente superior às produções cinematográficas, indicando uma maior aceitação e engajamento do público com esse formato.

Figura 12. Sétimo Conjunto



Fonte: Os autores, 2025.

Quanto à análise das plataformas de streaming, o gráfico na Figura 12, “Score IMDB Streaming” revelou que a HBO apresentou a maior média de avaliação (6,20), seguida pela Netflix (6,03), Amazon (5,44) e Disney (4,83). Esses resultados indicam que o conteúdo disponibilizado pela HBO é, em média, melhor avaliado, o que pode estar relacionado à sua estratégia de produção focada em qualidade e originalidade.

Por fim, o dashboard inclui uma tabela de detalhamento das produções, contendo informações como título, gênero, ano de lançamento, país de origem, classificação indicativa, tipo e pontuação IMDB. Essa tabela serve como base de referência para os gráficos apresentados e permite uma análise mais detalhada das variáveis envolvidas.

De modo geral, os resultados obtidos demonstram que fatores como gênero, classificação indicativa, país de origem e plataforma de streaming influenciam significativamente a avaliação das produções. A análise visual realizada no Power BI facilitou a compreensão das relações entre essas variáveis e possibilitou identificar padrões de preferência e desempenho entre os diferentes tipos de conteúdo audiovisual.

4 Conclusão

O desenvolvimento deste projeto permitiu alcançar plenamente o objetivo proposto, que consistia em realizar uma análise de filmes disponíveis em diferentes plataformas de streaming. Para atingir esse propósito, foi necessário aplicar técnicas de processamento, tratamento e organização de dados utilizando a linguagem Python, o que garantiu a limpeza, padronização e consistência das informações. Esse processo foi fundamental para eliminar ruídos e duplicidades, além de estruturar os dados de forma adequada para posterior armazenamento e consulta.

Em seguida, os registros tratados foram organizados e inseridos em um banco de dados relacional por meio do SQL Server, o que possibilitou maior controle sobre as informações e facilitou a execução de consultas estruturadas. Essa etapa foi

essencial para garantir a integridade e confiabilidade dos dados utilizados na análise, além de permitir a integração eficiente com ferramentas de visualização.

A conexão com o *Power BI* representou uma fase decisiva do projeto, pois permitiu transformar os dados armazenados em painéis interativos e dinâmicos. Essa integração proporcionou uma visualização clara, intuitiva e analítica dos resultados, facilitando a identificação de padrões, tendências e correlações entre variáveis, como gêneros, classificações indicativas, países de origem e plataformas de *streaming*. A partir desses painéis, foi possível compreender com mais profundidade o comportamento das avaliações dos filmes e séries, bem como as preferências do público em relação aos diferentes tipos de produções audiovisuais.

Os resultados obtidos demonstraram que o gênero drama foi o mais popular em número de votos, enquanto produções do gênero guerra obtiveram as melhores avaliações médias. Além disso, observou-se que produções voltadas ao público adulto (classificação TV-MA) apresentaram desempenho superior, e que plataformas como a HBO se destacaram em termos de qualidade percebida pelo público. Esses achados reforçam a importância do uso de dados para compreender o mercado audiovisual e as preferências dos espectadores.

Além de atender aos objetivos iniciais, o projeto evidenciou a importância do tratamento e da visualização de dados como etapas fundamentais na construção de análises confiáveis. A experiência proporcionou uma compreensão mais ampla sobre o ciclo completo de análise de dados, desde a coleta e preparação, até a exploração visual e a interpretação dos resultados.

Outro aspecto relevante foi a constatação de que a integração entre linguagens de programação, bancos de dados e ferramentas de *Business Intelligence* é um fator decisivo para o sucesso de projetos analíticos. Essa combinação permite não apenas automatizar processos, mas também ampliar a precisão das análises e otimizar a tomada de decisão baseada em evidências.

Em termos práticos, o trabalho apresentou uma solução eficiente e replicável para o estudo de dados provenientes de múltiplas fontes, podendo ser aplicada a outros contextos, como análises de séries temporais, preferências de consumo ou avaliações de produtos em plataformas digitais. Dessa forma, o projeto contribui não apenas para o aprendizado técnico, mas também para o desenvolvimento de competências analíticas e estratégicas voltadas à área de Ciência de Dados e Inteligência de Negócios.

Por fim, conclui-se que o objetivo geral foi plenamente alcançado, e que o projeto proporcionou aprendizado técnico, analítico e prático, consolidando o entendimento sobre o uso integrado de tecnologias de dados. O estudo reforça que a combinação de ferramentas como *Python*, *SQL Server* e *Power BI* é altamente eficaz para transformar grandes volumes de dados em informações relevantes, contribuindo para a tomada de decisão orientada por dados (Data-Driven Decision Making) e para o fortalecimento da cultura analítica em diferentes áreas do conhecimento.

Referências

ALMEIDA, F. *Cinema e Streaming: Transformações do Consumo Audiovisual*. São Paulo: Atlas, 2020. Acesso em 12 Ago 2025.

ALMEIDA, R. *Globalização do Entretenimento: Estratégias de Plataformas Digitais*. Lisboa: FCA, 2019. Acesso em 15 Ago 2025.

ANDERSON, C. *The Long Tail: Why the Future of Business Is Selling Less of More*. New York: Hyperion, 2006. Acesso em 22 Ago 2025.

ANDRADE, P. *A Nova Era da Disney: Estratégias de Mercado no Streaming Digital*. Rio de Janeiro: Elsevier, 2020. Acesso em 22 Ago 2025.

CARVALHO, T. *Mercado de Streaming e Competitividade Digital*. Porto Alegre: Bookman, 2020. Acesso em 22 Ago 2025.

CHEN, H.; CHIANG, R. H. L.; STOREY, V. C. Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, v. 36, n. 4, p. 116-118, 2012. Acesso em 15 Ago 2025.

CONNOLLY, T.; BEGG, C. *Database systems: a practical approach to design, implementation, and management*. 6. ed. Boston: Pearson, 2015. Acesso em 12 Ago 2025.

CORONEL, C.; MORRIS, S. *Database Systems: Design, Implementation, and Management*. 12. ed. Boston: Cengage Learning, 2017. Acesso em 15 Ago 2025.

COSTA, M. *Adaptação Cultural no Consumo de Mídia Digital*. Curitiba: Appris, 2021.

DATE, C. J. *An Introduction to Database Systems*. 8. ed. Boston: Addison-Wesley, 2019. Acesso em 22 Ago 2025.

DAVENPORT, T. H.; HARRIS, J. G. *Competing on Analytics: The New Science of Winning*. Boston: Harvard Business Press, 2017. Acesso em 12 Ago 2025.

DOWNEY, A. B. *Think Python: How to Think Like a Computer Scientist*. 2. ed. Cambridge: O'Reilly Media, 2015. Acesso em 12 Ago 2025.

ELMASRI, R.; NAVATHE, S. B. *Sistemas de banco de dados: fundamentos e aplicações*. 6. ed. São Paulo: Pearson, 2011. Acesso em 15 Ago 2025.

FERRARI, A.; RUSSO, M. *Introducing Microsoft Power BI*. Redmond: Microsoft Press, 2016. Acesso em 22 Ago 2025.

FERREIRA, J. *Personalização e Algoritmos em Serviços de Streaming*. Belo Horizonte: UFMG, 2021. Acesso em 22 Ago 2025.

GIL, A. C. *Métodos e técnicas de pesquisa social*. 7. ed. São Paulo: Atlas, 2019.

GOMEZ-URIBE, C. A.; HUNT, N. The Netflix Recommender System: Algorithms, Business Value, and Innovation. *ACM Transactions on Management Information Systems*, v. 6, n. 4, 2016. Acesso em 15 Ago 2025.

JOHNSON, M. *Streaming Wars: Strategies and Market Competition in Digital Media*. New York: Routledge, 2021. Acesso em 12 Ago 2025.

JOHNSON, M. *The Streaming Revolution: Media Consumption in the Digital Age*. New York: Routledge, 2019. Acesso em 12 Ago 2025.

KOTLER, P.; KELLER, K. L. *Administração de marketing*. 15. ed. São Paulo: Pearson, 2018. Acesso em 15 Ago 2025.

LAUDON, K. C.; TRAVER, C. G. *E-commerce: business, technology, society*. 16. ed. New Jersey: Pearson, 2021. Acesso em 22 Ago 2025.

LIMA, A. *Conteúdos Originais e Estratégias de Streaming*. Porto Alegre: Penso, 2019.

LIMA, R. *Experiência do Usuário em Plataformas Digitais de Vídeo*. São Paulo: Saraiva, 2020. Acesso em 22 Ago 2025.

MARTINS, G. *Algoritmos de Recomendação e o Consumo Digital*. Brasília: UnB, 2018.

MARTINS, P. *Global e Local: Estratégias de Conteúdo em Plataformas Digitais*. Lisboa: Almedina, 2021. Acesso em 15 Ago 2025.

McKINNEY, W. *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython*. 2. ed. Sebastopol: O'Reilly Media, 2017.

MENDES, L. *HBO Max e a Competição no Mercado de Streaming*. São Paulo: Atlas, 2021. Acesso em 12 Ago 2025.

MENDES, R. *Franquias e Nichos no Streaming Digital*. Recife: UFPE, 2020. Acesso em 22 Ago 2025.

MICROSOFT. *SQL Server – Documentation*. Redmond: Microsoft Docs, 2024. Disponível em: <https://learn.microsoft.com/sql/>. Acesso em 22 Ago 2025.

MILLMAN, R. Data Science Communities and the Role of Kaggle. *Journal of Data Science*, v. 17, n. 2, p. 40-50, 2019. Acesso em 22 Ago 2025.

MOREIRA, S. *Gestão de Marcas e Franquias no Mercado de Entretenimento*. Florianópolis: UFSC, 2021. Acesso em 15 Ago 2025.

OLIVEIRA, C. *Netflix e Estratégias de Internacionalização no Mercado Digital*. São Paulo: USP, 2021. Acesso em 12 Ago 2025.

PEREIRA, D. *Ciência de Dados Aplicada ao Consumo Audiovisual*. Porto Alegre: Bookman, 2020. Acesso em 22 Ago 2025.

PEREIRA, J. *Produções Originais e a Dinâmica do Streaming*. Curitiba: Appris, 2020.

PRESSMAN, R. S. *Engenharia de software: uma abordagem profissional*. 8. ed. São Paulo: AMGH, 2015. Acesso em 12 Ago 2025.

PROVOST, F.; FAWCETT, T. *Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking*. Sebastopol: O'Reilly Media, p 91/27 2013. Acesso em 15 Ago 2025.

REZAEI, M.; SALEHI, M. Data Visualization and Business Intelligence with Power BI. *International Journal of Business Analytics*, v. 7, n. 2, p. 42-50, 2020. Acesso em 12 Ago 2025.

RIBEIRO, A. *Inovação e Diferenciação em Plataformas Digitais de Vídeo*. Campinas: Unicamp, 2021. Acesso em 15 Ago 2025.

ROCHA, V. *Mercado Audiovisual e Estratégias de Licenciamento Digital*. Belo Horizonte: UFMG, 2021. Acesso em 22 Ago 2025.

SEVERANCE, C. *Python for Everybody: Exploring Data Using Python* 3. Boston: CreateSpace, 2016. Acesso em 12 Ago 2025.

SHARDA, R.; DELEN, D.; TURBAN, E. *Business Intelligence, Analytics, and Data Science: A Managerial Perspective*. 4. ed. New Jersey: Pearson, 2018. Acesso em 15 Ago 2025.

SMITH, J. *The Digital Streaming Era: Changing Patterns in Media Consumption*. London: Palgrave Macmillan, 2020. Acesso em 12 Ago 2025.

SOUZA, E. *HBO e a Construção da Identidade Audiovisual*. Porto Alegre: Penso, 2020. Acesso em 22 Ago 2025.

SOUZA, L. *Estratégias Integradas da Amazon no Mercado Digital*. Rio de Janeiro: FGV Editora, 2021. Acesso em 15 Ago 2025.

SWEIGART, A. *Automate the Boring Stuff with Python: Practical Programming for Total Beginners*. 2. ed. San Francisco: No Starch Press, 2019. Acesso em 15 Ago 2025.

ZELLE, J. *Python Programming: An Introduction to Computer Science*. 3. ed. Franklin, Beedle & Associates Inc., 2016. Acesso em 15 Ago 2025.