

Relatório Final de Análise de Dados - Plataforma Start-Flix

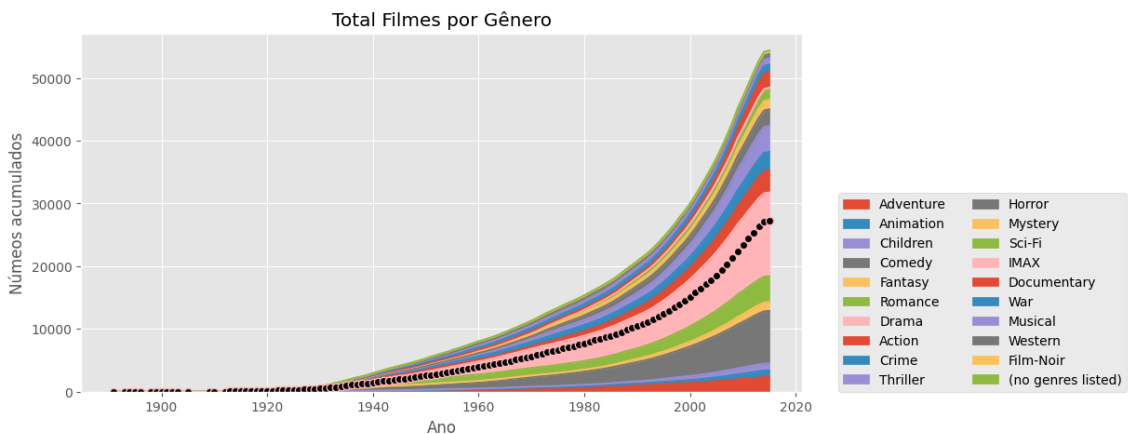
Introdução

A Start-Flix, uma plataforma emergente de streaming, enfrenta desafios críticos relacionados à personalização de seu sistema de recomendação, à ausência de insights aprofundados sobre preferências de usuários e à análise deficiente de desempenho de filmes. Para enfrentar esses obstáculos, foi conduzido um projeto de análise exploratória utilizando o conjunto de dados MovieLens, uma base amplamente reconhecida que contém milhões de avaliações de usuários, dados sobre títulos, gêneros, e comportamento de consumo de conteúdo cinematográfico.

O objetivo deste relatório é apresentar os principais achados obtidos por meio da exploração e visualização dos dados, propor insights estratégicos com base nos padrões identificados e sugerir ações que contribuam diretamente para a melhoria da experiência do usuário na plataforma.

Execução e Análise dos Gráficos

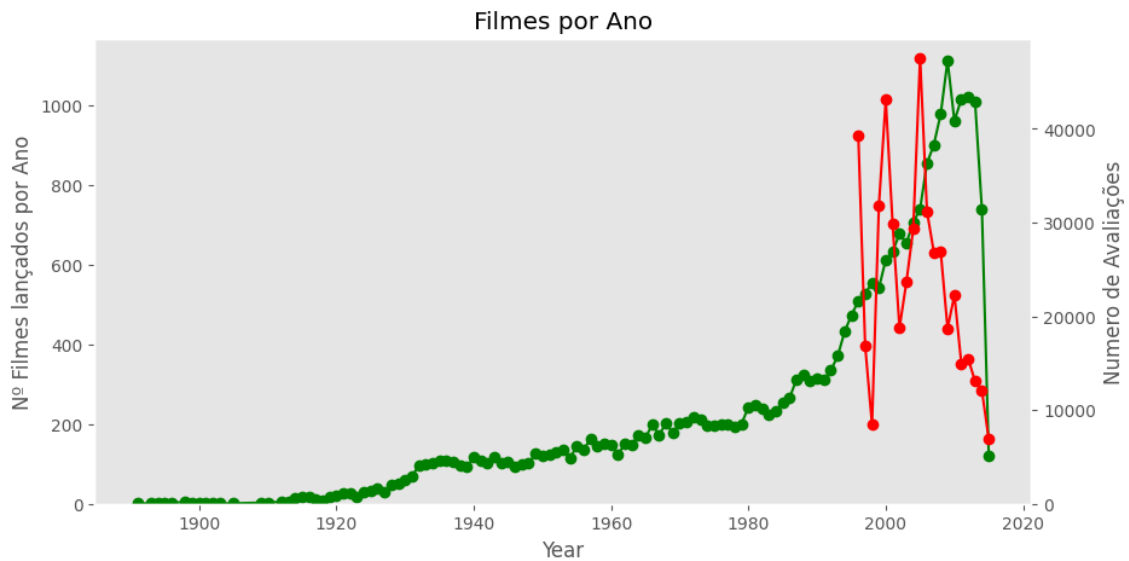
1. Total de Filmes por Gênero (Gráfico de Área Acumulada)



Análise: Este gráfico evidencia a evolução acumulada da produção cinematográfica por gênero ao longo do tempo. Observa-se um crescimento exponencial da quantidade de filmes após a década de 1980, com destaque para os gêneros **Drama**, **Comédia**, **Ação** e **Terror**. Isso sugere uma expansão significativa da indústria cinematográfica e uma diversificação dos gêneros para atender a diferentes nichos de público.

Insight Estratégico: Gêneros com alta produção indicam maior competição e oferta. É essencial que o sistema de recomendação avalie o nível de engajamento individual em cada gênero, para evitar a superexposição a conteúdos comuns e propor sugestões mais equilibradas e personalizadas.

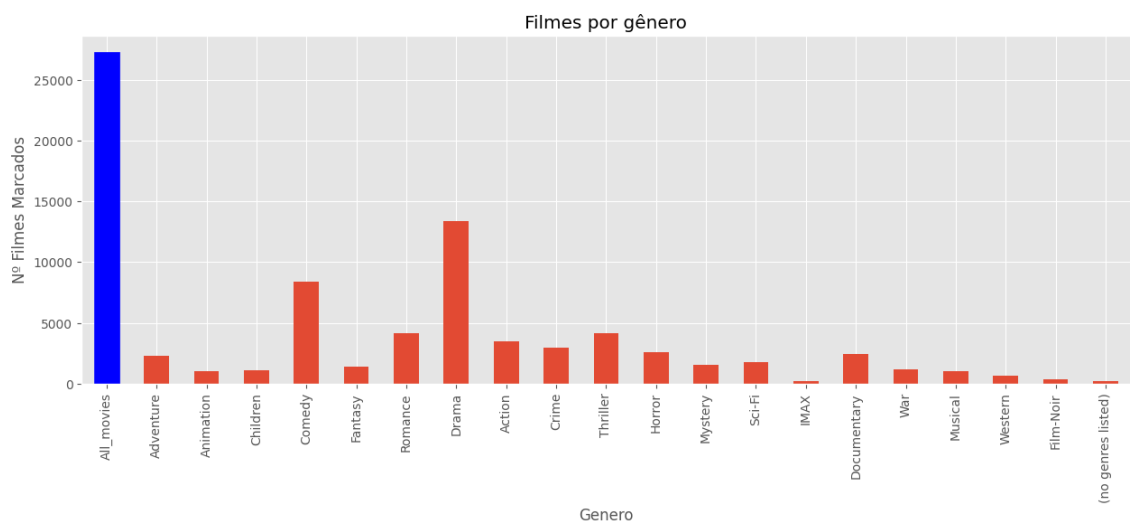
2. Filmes por Ano vs. Avaliações por Ano (Eixo Duplo)



Análise: A linha verde mostra o número de filmes lançados por ano, enquanto a linha vermelha representa a quantidade de avaliações recebidas. Nota-se uma forte correlação entre o aumento de lançamentos e o número de avaliações a partir dos anos 1990, com pico por volta de 2010. Após esse pico, observa-se uma queda nas avaliações, apesar da continuidade de lançamentos.

Insight Estratégico: A queda nas avaliações pode estar relacionada à dispersão de conteúdo e fadiga do usuário diante de tantas opções. A Start-Flix pode explorar estratégias como curadoria automatizada e prompts inteligentes para estimular avaliações e coleta de feedback.

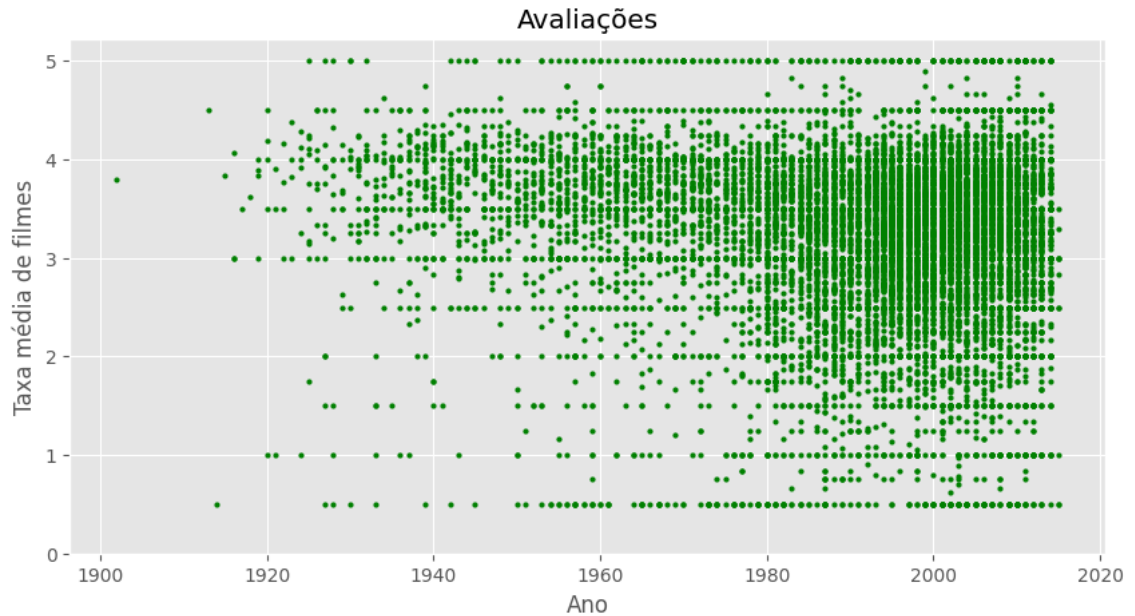
3. Quantidade de Filmes por Gênero (Gráfico de Barras)



Análise: O gênero **Drama** lidera em volume de produção, seguido por **Comédia**, **Thriller**, **Ação** e **Romance**. A predominância de drama pode indicar uma tendência da indústria ou preferência histórica dos produtores e distribuidores.

Insight Estratégico: A análise de preferências individuais por gênero deve considerar a abundância relativa de cada tipo. Gêneros com menor volume, como **Musical** e **Film-Noir**, podem ser utilizados para recomendações diferenciadas e exclusivas, agregando valor à experiência personalizada.

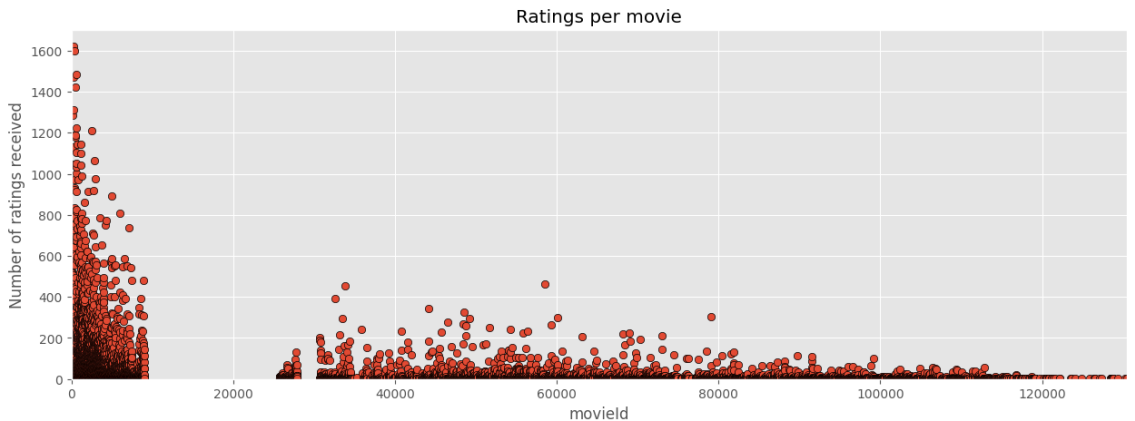
4. Avaliações Médias por Ano (Dispersão Temporal)



Análise: A dispersão dos dados mostra que a maioria dos filmes tende a receber uma média de avaliação entre 3.0 e 4.5, independentemente do ano de lançamento. Há uma maior densidade de filmes avaliados após 1980, o que reflete o aumento no consumo e avaliação digital.

Insight Estratégico: Como há uma concentração de avaliações em notas médias, o sistema pode utilizar técnicas como normalização ou z-score para identificar outliers de alta performance, evitando o viés de popularidade.

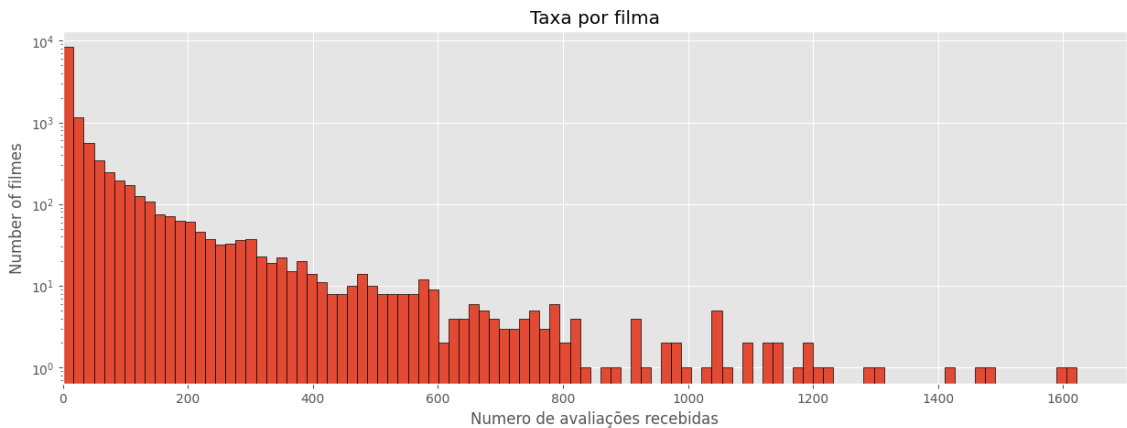
5. Distribuição de Avaliações por Filme (Scatterplot)



Análise: A maioria dos filmes recebe poucas avaliações, com poucos títulos acumulando grandes volumes (distribuição de cauda longa). Isso reforça a existência de um pequeno grupo de filmes altamente populares e uma grande massa de títulos pouco explorados.

Insight Estratégico: Recomendar exclusivamente os títulos mais populares reforça a desigualdade de exposição. A Start-Flix pode implementar técnicas de **diversificação em sistemas de recomendação** para expor usuários a filmes menos avaliados, mas potencialmente relevantes com base em gostos específicos.

6. Frequência de Avaliações por Filme (Histograma Logarítmico)



Análise: A distribuição é altamente assimétrica. A maioria absoluta dos filmes possui menos de 100 avaliações, indicando baixa visibilidade e engajamento. A escala logarítmica evidencia a disparidade entre os filmes mais e menos avaliados.

Insight Estratégico: A plataforma deve empregar mecanismos para aumentar a exposição e coleta de feedback para esses filmes de baixa avaliação, como notificações personalizadas ou campanhas de recomendação direcionada a nichos específicos.

Conclusão Final

A partir das análises realizadas sobre o dataset da plataforma Start-Flix, foi possível extrair uma série de insights estratégicos que devem orientar melhorias substanciais tanto no sistema de recomendação quanto na tomada de decisões de produto e marketing.

1. Recomendações Ineficientes

As visualizações mostram que a maioria dos filmes possui um número muito baixo de avaliações, o que indica uma dispersão significativa no engajamento dos usuários. A análise de avaliações por filme (últimos dois gráficos) evidencia que uma pequena parcela dos títulos recebe a maioria das avaliações, enquanto a maior parte é pouco avaliada.

Ação recomendada:

- Implementar um sistema híbrido de recomendação (colaborativo + baseado em conteúdo), priorizando filmes com menos interações para aumentar sua visibilidade.

- Aplicar técnicas de *cold-start* para novos filmes e usuários, usando metadados como gênero, ano de lançamento e sinopse para melhorar a recomendação inicial.

2. Falta de Insights sobre Preferências

A distribuição por gênero mostra que “Drama”, “Comedy” e “Action” são os gêneros mais populares, mas não necessariamente os mais bem avaliados. Além disso, muitos filmes estão classificados em múltiplos gêneros, o que exige um sistema de recomendação capaz de lidar com essa multietiquetagem.

Ação recomendada:

- Criar perfis dinâmicos de usuários com base em seus históricos de avaliações, focando em seus gêneros favoritos e padrões temporais de consumo.
- Utilizar algoritmos de clusterização (como K-Means) para segmentar usuários com comportamentos semelhantes, facilitando recomendações personalizadas por grupo.

3. Desempenho de Filmes

Observa-se um aumento contínuo na produção de filmes ao longo dos anos, com picos a partir dos anos 2000. No entanto, o número de avaliações não cresce proporcionalmente, indicando uma possível saturação de conteúdo. Além disso, a análise de média de avaliações por ano mostra que a qualidade percebida dos filmes se manteve relativamente estável, apesar do volume crescente.

Ação recomendada:

- Desenvolver dashboards em tempo real com métricas de desempenho por filme (avaliações, média de nota, tempo médio até a primeira avaliação, etc.) para alimentar decisões editoriais e de marketing.
- Priorizar campanhas promocionais e posicionamento na interface para filmes com bom desempenho recente e baixa visibilidade geral.

Recomendações Estratégicas Adicionais:

- **Big Data & Engenharia de Dados:** Criar pipelines robustos para ingestão contínua de dados de avaliações, interações e cliques dos usuários, estruturando-os em um data lake escalável e disponível para análises futuras.
- **Looker Studio & Visual Analytics:** Desenvolver painéis interativos no Looker Studio para que diferentes áreas da empresa (marketing, produto, conteúdo) possam monitorar preferências e tendências em tempo real.
- **Modelos de Machine Learning:** Investir em modelos de aprendizado supervisionado e não supervisionado para prever filmes com alta probabilidade de engajamento e avaliar o risco de churn de usuários com baixa interação.

Em síntese, a análise dos dados aplicados à realidade da Start-Flix demonstra que há um grande potencial de melhoria nas recomendações e na experiência do usuário. A chave está na personalização baseada em dados, na integração de sistemas inteligentes e na construção de um ecossistema analítico sustentável e orientado à tomada de decisão.

Dessa forma, a plataforma poderá não apenas reter seus usuários atuais, mas também expandir sua base de forma mais eficiente e estratégica.