



Análise das Músicas Mais Reproduzidas no Spotify (2024)

Daniel dos Santos da Silva, Enzo Ferroni, Hudson Carvalho de Jesus, Vinícius de Souza Sabiá

{10720767@mackenzista.com.br, 10417100@mackenzista.com.br,
10721459@mackenzista.com.br, 10721475@mackenzista.com.br}

<https://github.com/EnzoFerroni/ProjetoAplicadoI>



Sumário

1. Introdução	3
2. Glossário	Erro! Indicador não definido.
3. Contexto do Estudo	4
4. Objetivo do Estudo	4
5. Apresentação de Solução Analítica	5
6. Referências do Dataset	6
7. Descrição do Dataset	7
8. Apresentação dos Metadados	7
9. Análise Exploratória de dados.....	10
10. Conclusão.....	10



1. Glossário

Seção técnica para padronizar termos

Termo	Definição	Relevância para o Estudo	Exemplo no Dataset
Streams	Número de vezes que uma música foi reproduzida no Spotify	Métrica central de sucesso	1, 200, 00 streams
Track Score	Nota dada a música conforme vários fatores (19,4-725)	Métrica de sucesso	20, 120, 230, 450
Playlist Count	Número de playlists que a música está	Métrica de sucesso	30,764, 40,564, 40,415
Playlist Reach	Número de playlist	Métrica de sucesso	196,631,588 79,944,5681
All time rank	Ranque de todos os tempos das músicas	Métrica de sucesso	1, 31, 515, 3510
Popularity	Nota dada a música conforme o spotify (1-96)	Métrica de sucesso	2, 46, 95

2. Introdução

Nos últimos anos, o streaming revolucionou a forma como consumimos música, e o Spotify se tornou uma das principais plataformas nesse cenário. A possibilidade de acessar milhões de faixas a qualquer momento fez com que os hábitos dos ouvintes mudassem drasticamente, criando tendências e redefinindo o conceito de sucesso musical.



Este projeto tem como objetivo analisar os dados das músicas mais reproduzidas no Spotify em 2024 para entender quais fatores influenciam seu desempenho. A partir do dataset disponível no Kaggle, faremos uma exploração detalhada para identificar os gêneros musicais mais populares, a relação entre as músicas mais escutadas, os artistas e gravadoras que mais aparecem na lista e etc.

Com essa análise, esperamos obter insights relevantes sobre o comportamento dos ouvintes e as tendências da indústria musical no contexto atual.

3. Contexto do Estudo

A indústria da música tem passado por transformações intensas nos últimos anos, e as plataformas de streaming desempenham um papel central nesse processo. Diferente da era dos CDs ou do download digital, hoje o sucesso de uma música é medido principalmente pelo número de reproduções que ela recebe.

O Spotify, sendo uma das plataformas mais influentes, pode revelar muito sobre as preferências do público global. Por isso, analisar os dados das músicas mais reproduzidas em 2024 nos permite entender melhor os fatores que contribuem para o sucesso de um artista ou faixa.

Essa análise pode ser útil não apenas para pesquisadores e profissionais de dados, mas também para artistas, produtores e gravadoras que buscam insights sobre o mercado e as preferências do público.

4. Objetivo do Estudo

Com base no dataset *Most_Streamed_Spotify_Songs_2024.csv*, o objetivo é identificar os fatores que impulsionam o sucesso de músicas no Spotify em 2024, utilizando métricas como:

- Spotify Streams (principal indicador de popularidade),
- Spotify Playlist Reach (alcance em playlists),
- Spotify Playlist Count (número de playlists que incluem a música),
- Spotify Popularity (pontuação de popularidade da plataforma),
- YouTube Views e TikTok Views (engajamento em outras plataformas),

Metodologia:

1. Análise Descritiva:
 - Ranking das músicas mais streamed e correlação entre variáveis (ex: streams × playlists).
 - Comparação de médias (ex: músicas explícitas vs. não explícitas).
2. Análise Preditiva (se aplicável):
 - Identificar variáveis com maior impacto nos streams (ex: playlists têm mais peso que YouTube?).
3. Visualização:



- Gráficos de dispersão (streams × playlist reach), heatmaps de correlação e rankings.

5. Apresentação da empresa e problema de pesquisa

O Spotify é uma das maiores plataformas de streaming de áudio do mundo, fundada em 2006 na Suécia com a missão de conectar artistas e ouvintes através de uma experiência musical personalizada. Com mais de 574 milhões de usuários ativos e 31% de market share global, a empresa se destaca por sua abordagem data-driven, utilizando algoritmos de recomendação como o Discover Weekly e o Spotify Wrapped para engajar seu público. Seus valores de inovação, diversidade e paixão pela música guiam iniciativas como o Spotify for Artists e parcerias com IA generativa, como o DJ AI, que aprimoram a descoberta de conteúdo e a retenção de usuários.

O principal desafio da Spotify é melhorar a personalização para reduzir a taxa de cancelamento, utilizando análise de dados e machine learning para entender padrões de escuta e preferências musicais. Projetos como a análise de sentimento em letras e modelos preditivos de sucesso musical reforçam seu compromisso com a inovação. Como líder do setor, a empresa continua investindo em tecnologias avançadas para manter sua vantagem competitiva, provando que os dados são essenciais para transformar a maneira como o mundo consome música.

A análise dos dados do Spotify revela padrões consistentes no comportamento dos usuários, como preferências musicais por horário, dia da semana e até mesmo estado emocional. Esses padrões seguem tendências previsíveis, como o aumento do consumo de músicas animadas durante as manhãs e finais de semana, ou a preferência por canções calmas à noite. Além disso, eventos culturais e lançamentos de artistas geram picos de consumo específicos, demonstrando que há uma forte correlação entre contexto externo e hábitos de escuta. O Spotify já utiliza esses insights em seus algoritmos de recomendação, como o Discover Weekly, que combina filtragem colaborativa e análise de conteúdo para personalizar playlists.

O desenvolvimento de novos algoritmos pode otimizar ainda mais a experiência do usuário e aumentar a retenção. Uma abordagem eficaz seria implementar modelos híbridos que integrem técnicas de deep learning para análise de áudio e letras com reinforcement learning para ajustar recomendações em tempo real. Ferramentas como TensorFlow e PySpark permitiriam processar grandes volumes de dados e identificar padrões complexos, enquanto testes A/B validariam a eficácia das melhorias. Com isso, o Spotify não apenas entenderia melhor os hábitos dos usuários, mas também anteciparia suas preferências, transformando dados em vantagem competitiva e fortalecendo sua posição no mercado de streaming.



6. Referências do Dataset

Para realizar este estudo, utilizaremos o dataset "Most Streamed Spotify Songs 2024", disponível no Kaggle.

Fonte: [Kaggle - Most Streamed Spotify Songs 2024](#)

Restrições de Uso: Os dados são de acesso público e podem ser utilizados para fins educacionais e acadêmicos.

Período da Coleta: 2024



7. Descrição do Dataset

O dataset "Most Streamed Spotify Songs 2024" foi criado e disponibilizado por Nel Giri Yewithana na plataforma Kaggle. Ele foi construído a partir de dados extraídos do Spotify, utilizando a API oficial da plataforma para coletar informações sobre as músicas mais reproduzidas no ano de 2024.

O objetivo da criação desse dataset foi fornecer um recurso abrangente para pesquisadores, analistas de dados e entusiastas da música, permitindo a exploração de padrões musicais, tendências de streaming e características sonoras das faixas mais populares.

Como os dados foram extraídos diretamente da API do Spotify, eles refletem informações oficiais da plataforma, incluindo métricas de popularidade, características acústicas das músicas e estatísticas de reprodução. Dessa forma, o dataset serve como uma fonte confiável para análises sobre o comportamento dos ouvintes e o desempenho das faixas no mercado musical global.

8. Apresentação dos Metadados

Atributos	Definição	Relevância para o estudo	Exemplo no Dataset
Track	Nome da música (String, sem limites).	Identifica a música no ranking.	MILLION DOLLAR BABY
Album Name	Nome do álbum (String, sem limites).	Relaciona a música ao seu álbum de origem.	Million Dollar Baby - Single
Artist	Nome do artista (String, sem limites).	Permite análise de popularidade por artista.	Tommy Richman
Release Date	Data de lançamento (String, formato DD/MM/AAAA).	Ajuda a identificar tendências ao longo do tempo.	26/04/2024
ISRC	Código único da música (String, sem limites).	Importante para rastrear a música em diferentes plataformas.	QM24S2402528
All Time Rank	Posição no ranking global (Inteiro, 1 - 4600).	Indica a popularidade da música.	1



UNIVERSIDADE PRESBITERIANA MACKENZIE
Faculdade de Computação e Informática
Projeto Aplicado I – Curso Ciência de Dados



Track Score	Pontuação da música (Float, 0 - 725.4).	Mede a influência da música no cenário musical.	725,4
Spotify Streams	Número de reproduções no Spotify (Inteiro, 0 - 2,031,280,633).	Indica o sucesso da música na plataforma mais usada de streaming.	390470936
Spotify Playlist Count	Número de playlists no Spotify (Inteiro, 0 - 269,802).	Mede a aceitação da música em curadorias.	30716
Spotify Playlist Reach	Alcance total das playlists no Spotify (Inteiro, 0 - 211,607,669).	Estima a exposição da música dentro da plataforma.	196631588
Spotify Popularity	Índice de popularidade no Spotify (Float, 0 - 100).	Avalia a relevância da música na plataforma.	92
YouTube Views	Visualizações no YouTube (Inteiro, 0 - 1,096,100,899).	Mede o impacto da música na plataforma de vídeo.	84274754
YouTube Likes	Curtidas no YouTube (Inteiro, 0 - 10,629,796).	Mede o engajamento na plataforma.	651565900
TikTok Views	Visualizações no TikTok (Inteiro, 0 - 14,603,725,994).	Representa a exposição da música na rede.	5332281936
YouTube Playlist Reach	Alcance das playlists no YouTube (Inteiro, 0 - 3,351,188,582).	Mede a disseminação via curadoria.	150597040
Apple Music Playlist Count	Número de playlists na Apple Music (Float, 0 - 394).	Avalia a aceitação da música entre usuários da Apple.	210



UNIVERSIDADE PRESBITERIANA MACKENZIE
Faculdade de Computação e Informática
Projeto Aplicado I – Curso Ciência de Dados



AirPlay Spins	Execuções em rádios (Inteiro, 0 - 1,474,799).	Mede a presença da música nas rádios.	40975
SiriusXM Spins	Execuções na SiriusXM (Inteiro, 0 - 2,182).	Mede a popularidade em rádios premium.	684
Deezer Playlist Count	Número de playlists na Deezer (Float, 0 - 264).	Mede a inclusão na plataforma.	62
Deezer Playlist Reach	Alcance das playlists da Deezer (Inteiro, 0 - 36,321,847).	Mede a disseminação na plataforma.	17598718
Amazon Playlist Count	Número de playlists na Amazon Music (Float, 0 - 210).	Mede a aceitação na Amazon Music.	114
Pandora Streams	Reproduções no Pandora (Inteiro, 0 - 190,260,277).	Mostra a presença da música na plataforma.	18004655
Pandora Track Stations	Estações do Pandora com a música (Inteiro, 0 - 203,384).	Indica a curadoria da música.	22931
Soundcloud Streams	Reproduções no SoundCloud (Inteiro, 0 - 7,208,651).	Mede a popularidade na plataforma.	4818457
Shazam Counts	Identificações no Shazam (Inteiro, 0 - 11,822,942).	Indica o interesse do público em descobrir a música.	2669262
TIDAL Popularity	Popularidade da música no TIDAL (Float, 0 - 100).	Mede o impacto da música entre usuários do TIDAL.	N/A
Explicit Track	Indica se a música tem conteúdo explícito (Inteiro, 0 = Não, 1 = Sim).	Importante para análise de restrições e públicos-alvo.	0



9. Análise Exploratória de dados

10. Conclusão

A análise dos dados das músicas mais reproduzidas no Spotify em 2024 permite compreender as tendências e preferências do público na era do streaming. A partir do dataset disponível no Kaggle, será possível identificar padrões que influenciam o sucesso de uma faixa, incluindo gêneros musicais, artistas e gravadoras mais populares. Esses insights são valiosos tanto para a indústria musical quanto para pesquisadores e profissionais de dados. Assim, espera-se que, o estudo contribua para uma visão mais ampla sobre o impacto do streaming no consumo de música e na definição do sucesso musical.