

### **Título do Projeto:** O que faz uma música de sucesso?

Por: Ana Carolina Almeida Costa e Vanessa Fermino

### **Objetivo**

O presente projeto foi desenvolvido em dupla e tem por objetivo identificar o que faz uma música obter sucesso em termos de streams. Para alcançar esse objetivo buscamos responder às hipóteses levantadas pela gravadora através da análise dos dados fornecidos. As hipóteses avaliadas incluem:

1. Músicas com BPM (Batidas Por Minuto) mais altas fazem mais sucesso em termos de número de streams no Spotify.
2. As músicas mais populares no ranking do Spotify também possuem um comportamento semelhante em outras plataformas, como a Deezer.
3. A presença de uma música em um maior número de playlists está correlacionada com um maior número de streams.
4. Artistas com um maior número de músicas no Spotify têm mais streams.
5. As características da música influenciam o sucesso em termos de número de streams no Spotify.

### **Ferramentas e Tecnologias**

Para tratamento e limpeza dos dados foi utilizada a ferramenta de SQL. O Power BI foi utilizado para facilitar a visualização das informações tratadas através de tabelas dinâmicas e gráficos, culminando na montagem de um dashboard. Utilizamos o cálculo da correlação de Pearson para verificar a relação entre diferentes variáveis com o total de streams das músicas e avaliar se existe alguma conexão entre elas.

### **Processamento e análises**

→ Os dados fornecidos foram carregados na ferramenta BigQuery e a partir disso foram realizadas várias consultas que permitiram identificar inconsistências nos dados, informações duplicadas e dados fora do escopo da análise;

- Após realizadas as consultas os dados foram limpos e compilados em uma nova planilha onde também foram criadas novas variáveis, como a data do lançamento, informações sobre o total de músicas por artista;
- Os valores nulos encontrados na coluna `in_shazam_charts` foram substituídos por zero e mantidos;
- As músicas duplicadas foram mantidas, pois tinha informações diferentes como data de lançamento ou total de playlists e foram consideradas como versões diferentes da mesma música. Como ainda não sabíamos a influência das características técnicas para o sucesso, optamos por manter. E de qualquer forma, eram apenas 4 músicas, representando um percentual irrisório do total de músicas. Sendo assim, consideramos que manter as duplicatas não afetaria os resultados finais;
- A coluna `key` foi removida por contar com grande quantidade de nulos e não ser considerada uma informação chave para nossas análises.
- A música *Love Grows* de Edison Lighthouse foi removida pois na coluna `streams` continha uma informação divergente que na verdade era um compilado das informações das outras colunas e não foi possível extrair dela a informação do número de streams para a música;
- Foi utilizado o comando `SAFE_CAST` para remover os dados não numéricos da coluna `streams`;
- Foram identificados e limpos os caracteres especiais em `track_name` e `artist_name`, mantendo as informações;
- Foram criados quartis para bpm e todas as características das músicas e os mesmos foram classificados em “muito baixo”, “baixo”, “médio” e “alto”;
- Foram realizadas análises de quantidade de streams por classificação de quartis para cada uma das características das músicas e para bpm.

## **Resultados e Conclusões**

As hipóteses 1, 4 e 5 mencionadas no início foram refutadas, enquanto as hipóteses 2 e 3 foram corroboradas através da correlação de Pearson. Dessa forma, podemos concluir que bpm e informações técnicas das músicas, bem como o total de músicas de um artista não configuram fatores relevantes quando se trata da quantidade de streams de uma música. O número de streams de uma música se mostrou diretamente relacionado à quantidade de playlists nas quais a mesma está incluída e também podemos concluir que o sucesso das músicas se comporta de maneira semelhante nas três plataformas avaliadas (apple, deezer e spotify). A seguir seguem os cálculos de correlação para cada uma das hipóteses:

Soma de corr_spotify_apple <b>0,71</b>	Soma de corr_spotify_bpm <b>-0,02</b>	Soma de corr_spotify_deezer <b>0,80</b>	Soma de corr_streams_acousticness__ <b>0,00</b>
Soma de corr_streams_bpm <b>0,00</b>	Soma de corr_streams_danceability__ <b>-0,11</b>	Soma de corr_streams_energy__ <b>-0,03</b>	Soma de corr_streams_playlists <b>0,78</b>
Soma de corr_streams_speechiness__ <b>-0,11</b>	Soma de corr_streams_total_musicas_por_artista <b>-0,01</b>	Soma de corr_streams_valence__ <b>-0,04</b>	

Através de nossas análises ficou constatado que BPM ou características da música propriamente ditas não são um ponto decisivo para o sucesso. Por outro lado, o número de playlists em que a música está incluída se mostrou um fator determinante do seu sucesso. Mais playlists, mais streams.

Diante desse cenário, podemos focar em estratégias voltadas para o lançamento da música e ações que garantam a inclusão da música em um grande número de playlists.

1. A inclusão de uma música em uma playlist oficial das plataformas depende de uma descrição assertiva do gênero, instrumentos, clima e público-alvo. Essa descrição é fundamental para facilitar o trabalho da curadoria e dos algoritmos em alocar uma música na playlist adequada. Cadastro antecipado da música nas plataformas digitais, permitindo que curadores editoriais tenham acesso ao seu material com antecedência.
2. Outras ações importantes são relacionadas à interação com o público-alvo. Criar o chamado buzz, soltar pequenos trechos e cliques curtos antes do lançamento aumenta a expectativa e o engajamento do público. Assim, quando a música for lançada já haverá um grande público interessado.
3. Além dos fãs, o relacionamento com curadores independentes também é uma excelente alternativa. Essas pessoas possuem grande influência e possuem suas próprias playlists que são amplamente compartilhadas. Estar na playlist de um curador independente traz grandes chances de aumentar o número de streams de uma música.

Espera-se que a implementação conjunta das ações recomendadas acima promova o aumento significativo da probabilidade que uma música tem de fazer sucesso.

## **Limitações/Próximos Passos**

Talvez fosse interessante ter somado os números de streams das músicas repetidas ao invés de considerá-las individualmente.

Sabemos que obter correlação positiva entre duas variáveis não necessariamente indica que possui uma relação de causalidade entre os eventos. Portanto, precisamos considerar que a validação e refutação das hipóteses apresentadas é frágil e necessita de uma análise mais robusta que possa de fato estabelecer a conexão entre os eventos. Uma análise de regressão linear ou regressão logística é recomendada.

## **Links de interesse**

Link para o relatório Power BI:  
[https://drive.google.com/file/d/10EiZyfRmvz74vkVMFyHPct1PlwjY8XV/view?usp=drive\\_link](https://drive.google.com/file/d/10EiZyfRmvz74vkVMFyHPct1PlwjY8XV/view?usp=drive_link)

## Legendas das planilhas:

### Trackinspotify

- track\_id: Identificador exclusivo da música. É um número inteiro de 7 dígitos que não se repete.
- track\_name: Nome da música.
- \*\*artist(s)\_name\*\*: Nome do(s) artista(s) da música.
- artist\_count: Número de artistas que contribuíram na música.
- released\_year: Ano em que a música foi lançada.
- released\_month: Mês em que a música foi lançada.
- released\_day: Dia do mês em que a música foi lançada.
- inspotifyplaylists: Número de listas de reprodução do Spotify em que a música está incluída
- inspotifycharts: Presença e posição da música nas paradas do Spotify
- streams: Número total de streams no Spotify. Representa o número de vezes que a música foi ouvida.

### Trackincompetition

- track\_id: Identificador exclusivo da música. É um número inteiro de 7 dígitos que não se repete.
- inappleplaylists: número de listas de reprodução da Apple Music em que a música está incluída.
- inapplecharts: Presença e classificação da música nas paradas da Apple Music.
- indeezerplaylists: Número de playlists do Deezer em que a música está incluída.
- indeezercharts: Presença e posição da música nas paradas da Deezer.
- inshazamcharts: Presença e classificação da música nas paradas da Shazam.

### Tracktechnicalinfo

- track\_id: Identificador exclusivo da música. É um número inteiro de 7 dígitos que não se repete.
- bpm: Batidas por minuto, uma medida do tempo da música.
- key: Tom musical da música.
- mode: Modo de música (maior ou menor).
- danceability\_%: Porcentagem que indica o quão apropriado a canção é para dançar
- valence\_%: Positividade do conteúdo musical da música.
- energy\_%: Nível de energia percebido da música.
- acusticness\_%: Quantidade de som acústico na música.
- instrumentality\_%: Quantidade de conteúdo instrumental na música.
- liveness\_%: Presença de elementos de performance ao vivo.
- speechiness\_%: Quantidade de palavras faladas na música.