

ESTUDO DE HIT'ABILIDADE

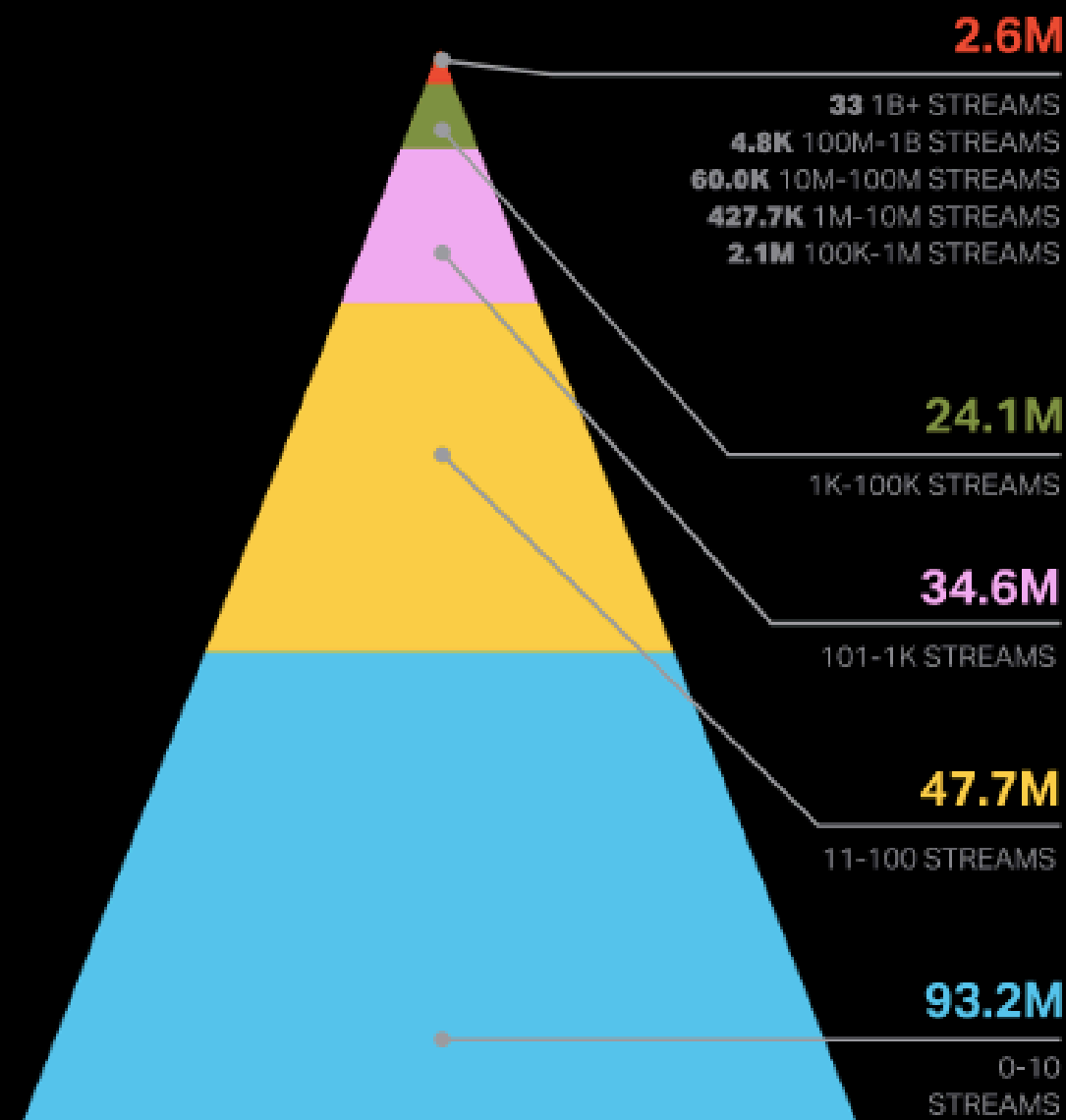
UMA ANÁLISE DE ELEMENTOS
HARMÔNICOS E VERBAIS DAS
INDICAÇÕES AO GRAMMY DE MELHOR
ALBUM POP EM LÍNGUA PORTUGUESA



- **Introdução**
- **Base Teórica**
- **Metodologia**
- **Resultados**
- **Conclusões**

SUMÁRIO

CONTEXTO



- Em 2024, o mercado fonográfico global movimentou US\$ 29,6 bilhões
- Atualmente, estima-se que 100.000 novas músicas são lançadas nas plataformas de streaming diariamente
- Das mais de 200 milhões de músicas disponíveis nos Spotify:
 - Quase 50% obteve menos de 10 reproduções
 - Cerca de 15% obteve mais de 1k reproduções
 - Apenas 1,5% obteve mais de 100k reproduções

Problema central: O quanto (e como) as características estruturais impactam na popularidade das músicas nos apps de streaming? Existe uma receita para o sucesso (ou pelo menos uma medida para estimá-lo)?

JUSTIFICATIVA

- Lacuna entre as áreas da gestão e da música:
 - Decisões estratégicas (ou de marketing) e decisões artísticas são constantemente consideradas como antagônicas, concorrentes ou inconciliáveis
 - Músicos e artistas acabam por desconsiderar completamente as métricas e possibilidades que a área da gestão oferece

- Investigar quantitativamente correlações entre elementos musicais – duração, nº de acordes, quantidade de palavras diferentes, entre outras – e o sucesso comercial das faixas.
- Fornecer dados e insights para que artistas e produtores musicais tomem menos decisões ao acaso (ou seja, sem intencionalidade artística nem intencionalidade estratégica)

OBJETIVOS

QUADRO TEORICO

TERMO	DEFINIÇÃO	AUTOR
Melodia	conjunto de sons dispostos em ordem sucessiva	Med, Bohumil. Teoria da Música. Brasilia: MusiMed, 2017. pag. 8
Harmonia	ciência que estuda os acordes e as relações entre eles;	Med, Bohumil. Teoria da Música. Brasilia: MusiMed, 2017. pag. 8
Acordes	Combinação de três ou mais sons simultâneos diferentes	Med, Bohumil. Teoria da Música. Brasilia: MusiMed, 2017. pag. 56
Escala	Conjunto de notas disponíveis num determinado sistema musical	Med, Bohumil. Teoria da Música. Brasilia: MusiMed, 2017. pag. 38

TIPO DE PESQUISA

- **Natureza:**

- Observacional
- Quantitativa
- Descritiva

- **Método:**

- Dados secundários
- Amostra aleatória de 50 músicas entre 636 indicadas
- Análise de correlação, clusterização e regressão

POPULAÇÃO

O QUE

Músicas dos álbuns indicados ao

ONDE

Grammy Latino de Melhor Album
Pop em Língua Portuguesa

QUANDO

Entre os anos de 2014 e 2024

INSTRUMENTOS DE PESQUISA

- **Variável dependente:** n° de reproduções no YouTube Music → N° de reproduções no YouTube Music (log)
- **Variáveis independentes:**
 - Duração da faixa (seg) → **Taxa de Tensão Harmônica**
 $TTH = (\text{Acordes FE} / \text{Acordes diferentes})$
 - N° de acordes diferentes → **Índice de Densidade Verbal**
 $IDV = (\text{Palavras} / \text{Duração}) * 60$
 - N° de acordes fora da escala → **Índice de Densidade Verbal**
 $IDV = (\text{Palavras} / \text{Duração}) * 60$
 - Total de palavras na letra → **Taxa de Diversidade Lexical**
 $TDL = (\text{Palavras diferentes} / \text{Total de palavras})$
 - N° de palavras diferentes → **Taxa de Diversidade Lexical**
 $TDL = (\text{Palavras diferentes} / \text{Total de palavras})$

- **Fontes:**
 - YouTube Music
 - Letras.mus.br
 - CifraClub
 - Moises.ai
- **Técnicas:**
 - Amostragem aleatória (validação por Teste de Runs)
 - Gráficos Boxplot e cálculo do IQR para a análise de outliers

PROCEDIMENTO DE COLETA

TÉCNICAS E MODELOS DE ANÁLISE

- **Sumarização:** Estatísticas descritivas (média, mediana, moda, amplitude, desvio-padrão, etc.)
- **Agrupamento:** K-means com 4 clusters
- **Associação:** Correlação de Pearson
- **Regressão:** Regressão linear múltipla em duas aplicações (com e sem variáveis não significativas)

SUMÁRIO

STATISTICS

	Reproduções	Duração	Acordes	TTH	Palavras /min	TDL
Media	528.179,85	196,69	6,36	26,12%	69,86	39,13%
Mediana	942.000,00	197,00	6,00	20,00%	70,57	36,57%
DP	0	44,47	2,87	23,20%	30,82	16,98%
Amplitude	291.998.400,00	161,00	13,00	71,43%	116,02	74,49%

- Taxa de Tensão Harmônica (TTH) = (Acordes fora da escala / Total de Acordes diferentes)
- Taxa de Diversidade Lexical (TDL) = (Palavras diferentes / Total de Palavras)

AGRUPAMENTO

	Reproduções	Duracao	Acordes	TTH	Palavras / min	TDL	
Cluster 0	52.764,85	202,71	8,35	40,10%	45,84	49,15%	PERDIDOS NO CONCEITO
Cluster 1	4.669.073,13	168,38	5,38	12,56%	99,17	36,84%	SUCESSO COMO PRÓPRIO CONCEITO
Cluster 2	1.084.758,18	214,40	4,93	22,02%	71,69	29,77%	UM POUCO DE SUCESSO, UM POUCO DE CONCEITO
Cluster 3	29.781,98	284,00	8,80	42,85%	57,38	41,07%	O CONCEITO DE PERDIDO (OUTLIERS)

AGRUPAMENTO

	Reproduções	Duracao	Acordes	TTH	Palavras / min	TDL	
Cluster 0	52.764,85	202,71	8,35	40,10%	45,84	49,15%	PERDIDOS NO CONCEITO
Cluster 1	4.669.073,13	168,38	5,38	12,56%	99,17	36,84%	SUCESSO COMO PRÓPRIO CONCEITO
Cluster 2	1.084.758,18	214,40	4,93	22,02%	71,69	29,77%	UM POUCO DE SUCESSO, UM POUCO DE CONCEITO
Cluster 3	29.781,98	284,00	8,80	42,85%	57,38	41,07%	O CONCEITO DE PERDIDO (OUTLIERS)

AGRUPAMENTO

	Reproduções	Duracao	Acordes	TTH	Palavras / min	TDL	
Cluster 0	52.764,85	202,71	8,35	40,10%	45,84	49,15%	PERDIDOS NO CONCEITO
Cluster 1	4.669.073,13	168,38	5,38	12,56%	99,17	36,84%	SUCESSO COMO PRÓPRIO CONCEITO
Cluster 2	1.084.758,18	214,40	4,93	22,02%	71,69	29,77%	UM POUCO DE SUCESSO, UM POUCO DE CONCEITO
Cluster 3	29.781,98	284,00	8,80	42,85%	57,38	41,07%	O CONCEITO DE PERDIDO (OUTLIERS)

AGRUPAMENTO

	Reproduções	Duracao	Acordes	TTH	Palavras / min	TDL	
Cluster 0	52.764,85	202,71	8,35	40,10%	45,84	49,15%	PERDIDOS NO CONCEITO
Cluster 1	4.669.073,13	168,38	5,38	12,56%	99,17	36,84%	SUCESSO COMO PRÓPRIO CONCEITO
Cluster 2	1.084.758,18	214,40	4,93	22,02%	71,69	29,77%	UM POUCO DE SUCESSO, UM POUCO DE CONCEITO
Cluster 3	29.781,98	284,00	8,80	42,85%	57,38	41,07%	O CONCEITO DE PERDIDO (OUTLIERS)

Associação

	Reproduções (log)	Duração	Acordes	% acordes fora do CH	Palavras/ min
Duração	-0,3474				
Acordes	-0,4982	0,2915			
% acordes fora do CH	-0,4878	0,3012	0,5110		
Palavras/min	0,6450	-0,4717	-0,3422	-0,4847	
% palavras diferentes	-0,4756	0,1192	0,2183	0,2863	-0,6907

REGRESSÃO

	1ª aplic	2ª aplic
<i>Adjusted R Square</i>	0,4638	0,4911
<i>Significance F</i>	0,00000 4311	0,000000 5116

Modelo de regressão linear múltipla com 2 variáveis independentes e 1 variável dependente.

Modelo de regressão linear múltipla com 2 variáveis independentes e 1 variável dependente.

	<i>Coef. 1</i>	<i>P-value 1</i>	<i>Coef. 2</i>	<i>P-value 2</i>
Intercept	5,9535	0,0076	5,1350	0,0000
Duração	-0,0012	0,7918		
Acordes	-0,1182	0,0401	-0,1382	0,0075
TTH	-0,6063	0,3901		
Palavras/ min	0,0083	0,0404	0,0105	0,0001
TDL	-0,5885	0,6389		

Modelo de regressão linear múltipla com 2 variáveis independentes e 1 variável dependente.

CONSIDERAÇÕES FINAIS

	Geral	Cluster 1	Cluster 2	Resultado
Acordes	6,36	5,38	4,93	6,00
TTH	26,12%	12,56%	22,06%	20,00%
Palavras/min	69,86	99,17	71,69	70,00
TDL	39,13%	38,84%	29,77%	36,81%
Reproduções	588.303	4.669.073,13	1.084.758,18	x

- Músicas com menos palavras e mais repetitividade tendem a ter mais reproduções
- Complexidade harmonica tende a reduzir audiência
- Duração menor tem correlação fraca com o número de reproduções

Limitações: fatores externos (marketing, contexto social) não foram analisados

Próximos passos: incluir dados qualitativos, outras categorias e análises temporais

CONSIDERAÇÕES FINAIS

	Geral	Cluster 1	Cluster 2	Resultado
Acordes	6,36	5,38	4,93	6,00
TTH	26,12%	12,56%	22,06%	20,00%
Palavras/min	69,86	99,17	71,69	70,00
TDL	39,13%	38,84%	29,77%	36,81%
Reproduções	588.303	4.669.073,13	1.084.758,18	x

- Músicas com menos palavras e mais repetitividade tendem a ter mais reproduções
- Complexidade harmonica tende a reduzir audiência
- Duração menor tem correlação fraca com o número de reproduções

Limitações: fatores externos (marketing, contexto social) não foram analisados

Próximos passos: incluir dados qualitativos, outras categorias e análises temporais

OBRIGADO !