

André Palhares Costa

Uma análise de dados sobre músicas ao longo dos anos (1923-2020)

Proposta de Pesquisa Científica

Universidade Federal de Minas Gerais
Instituto de Ciências Exatas
Departamento de Ciência da Computação

Orientador: Flávio Vinícius Diniz de Figueiredo

Belo Horizonte, Minas Gerais
2021

SUMÁRIO

Introdução	2
1.1. Objetivo Geral	2
1.2. Objetivos Específicos	2
Referencial Teórico	3
Metodologia	5
Resultados Esperados	7
Etapas e Cronograma	8
Referências Bibliográficas	9

1. Introdução

A música é um dos pilares da cultura de uma civilização, sendo muitas vezes utilizada por um povo como forma de se expressar, contribuindo para a criação de uma identidade coletiva. Enquanto algumas músicas marcam um determinado período na história, outras transcendem gerações.

Com o trabalho atual, deseja-se expandir a investigação feita na primeira parte da monografia, analisando atributos musicais, fornecidos pela API Web do Spotify, de faixas não só brasileiras, mas lançadas em todo o planeta e pertencentes aos anos entre 1923 e 2020.

Ao expandir a amostra de dados quanto ao país e ano de lançamento, será possível analisar padrões musicais de diversas épocas através de uma visão global. Espera-se identificar padrões nas presenças dos atributos musicais entre as faixas selecionadas de cada ano e determinar se existe alguma semelhança com o padrão obtido no trabalho anterior para as músicas brasileiras de 1980 a 2020.

Serão utilizados atributos musicais de cerca de 1.2 milhão de faixas obtidas através do site MusicBrainz, que fornece um extenso catálogo musical. Ao analisar uma quantidade maior de músicas, será possível observar padrões ao longo dos 97 anos analisados e também de períodos específicos na história.

1.1. Objetivo Geral

- Analisar atributos musicais das faixas lançadas ao redor do mundo entre 1923 e 2020 e sua evolução ao longo dos anos para identificar características da música global em diferentes períodos.

1.2. Objetivos Específicos

- Identificar padrões nos atributos musicais das faixas já coletadas, analisando-as década a década
- Analisar características das músicas como um todo em períodos específicos do tempo
- Identificar rupturas em padrões musicais em determinados períodos históricos

2. Referencial Teórico

O estudo expandirá as análises feitas na primeira monografia e ainda será pautado nos atributos musicais fornecidos pela API pública do Spotify. Para entender esta análise, é necessário conhecer a definição destes atributos. São eles:

Atributo musical	Descrição
<i>Duration_ms</i>	Duração, em milissegundos, da faixa.
<i>Danceability</i>	Valor numérico que indica quão adequada uma faixa é para dançar. É obtido através da união de vários elementos como batimentos por minuto, intensidade da batida etc.
<i>Energy</i>	Valor percentual que indica o quão intensa, ou energética, uma faixa é. É medido através de elementos como range dinâmico, timbre, altura etc.
<i>Key</i>	Valor numérico que representa a nota musical geral da faixa. 0=C, 1=C#, 2=D etc.
<i>Loudness</i>	Valor em decibéis da faixa que indica o quão alta ela é. Estes valores variam de -60 a 0.
<i>Speechiness</i>	Valor percentual que indica a frequência de palavras faladas em uma música.
<i>Acousticness</i>	Valor percentual que indica quão acústica uma faixa é.
<i>Instrumentalness</i>	Valor percentual que indica a frequência de sons vocais em uma faixa. Músicas que possuem esse atributo maior que 0,5 tendem fortemente a ser instrumentais.
<i>Liveness</i>	Valor percentual que indica a presença de sons de uma plateia na faixa, sugerindo que seja uma gravação ao vivo.
<i>Valence</i>	Valor percentual que indica a positividade de uma faixa. Valores altos tendem a representar músicas que transmitem um sentimento positivo.
<i>Tempo</i>	Valor numérico que indica a quantidade de batidas por minuto de uma faixa.
<i>Time_Signature</i>	Valor numérico que indica o compasso de uma faixa.

O artigo “Quaranstream - how COVID-19 restrictions influenced song features on Spotify TOP-200 charts?” de Jukka-Pekka Forsberg estuda como o período do início da quarentena em razão da Covid-19 (março e abril de 2020) influenciaram no consumo musical das pessoas da Finlândia, Noruega e Suécia, utilizando os atributos fornecidos pela API do Spotify. Pretendemos expandir essa análise para outros períodos da história a nível mundial.

Outro artigo sobre o assunto é “Evacuate the Dancefloor: Exploring and Classifying Spotify Music Listening Before and During the COVID-19 Pandemic in DACH Countries” de Abteilung Musik et. al. aborda o mesmo assunto tratando de uma comparação entre as músicas consumidas no período anterior ao início da quarentena e o período posterior nos países Austria, Suíça e Alemanha.

O artigo “Using Spotify Audio Features to Study the Evolution of Pop Music” de Elena Georgieva e Blair Kaneshiro para a Stanford University tratam do assunto de forma restrita a uma análise de músicas populares em um específico mês nos Estados Unidos e as compara ao longo do tempo de 10 em 10 anos, sem estudar os anos intermediários ou diferentes países.

Outro artigo que aborda parcialmente o tema é o “Demystifying Musical Preferences at Turkish Music Market Through Audio Features of Spotify Charts” de Fatih Pinarbas para o Turkish Journal Of Marketing usa dos mesmos atributos musicais que serão objeto de estudo neste trabalho para identificar padrões no consumo de música especificamente na Turquia e em um período de tempo de somente 6 meses.

A pesquisa “The evolution of popular music: USA 1960–2010” de Matthias Mauch, Robert M. MacCallum, Mark Levy e Armand M. Leroi publicado na Royal Society Open Science evidencia como a presença de determinados gêneros ou estilos musicais ao longo das décadas, estritamente nos Estados Unidos.

3. Metodologia

A análise deste trabalho será feita em 2 etapas: visualização dos dados dos atributos musicais já coletados e sua posterior classificação. Abaixo estão brevemente detalhados cada um dos passos.

1º passo: Gerar visualizações de dados

Inicialmente, pretende-se montar visualizações de dados que possam exibir as informações contidas nos atributos musicais das faixas selecionadas para permitir uma fácil identificação de padrões. Por exemplo a existência de atributos que tiveram maior presença nas músicas de uma década específica ou aqueles que se mantiveram estáveis ao longo do tempo.

Serão utilizadas variadas técnicas de exibição de dados e apresentadas possíveis interpretações sobre os resultados obtidos. Além disso, deseja-se observar os dados em períodos específicos relevantes para a história mundial dentro do período analisado.

2º passo: Utilização de algoritmo de classificação

Depois, de forma semelhante ao trabalho anterior, buscaremos analisar se uma faixa pode facilmente ser classificada como pertencente a uma década ou não, dispondo agora de uma quantidade maior de dados. O uso de algoritmos de classificação permitirá inferir se as músicas são facilmente rotuladas como pertencentes a uma década - demonstrando assim que músicas de uma determinada década são mais parecidas entre si -- ou se os atributos estão tão bem distribuídos ao longo dos anos que essa conclusão não pode ser tomada.

Foi possível observar no estudo passado que quando uma faixa é classificada como pertencente a uma década erroneamente, esse erro tende a ser uma década anterior ou posterior à década real. Será analisado também se esse comportamento ocorre também analisando uma quantidade maior de faixas, de países variados e de um período também maior.

Busca-se identificar também possíveis impactos de períodos históricos nos padrões musicais das faixas analisadas. Para isso, serão utilizados métodos de regressão descontínua, a fim de analisar como as características musicais foram alteradas a partir de determinados acontecimentos na história.

4. Resultados Esperados

Ao final do trabalho espera-se obter uma série de visualizações de dados que permitam observar padrões nos atributos musicais de faixas lançadas ao redor do mundo no período analisado. Pretende-se também analisar década a década a fim de observar características predominantes em uma década específica, outras que não tiveram tanta expressividade e aquelas que se mantiveram estáveis. Por fim, serão analisados períodos específicos de tempo para identificar possíveis características que possam ser influenciadas por acontecimentos históricos de um período.

Além disso, deseja-se também utilizar de um algoritmo de classificação para auxiliar na breve análise de tendência das características das músicas analisadas. A partir do resultado obtido com o algoritmo, será possível identificar quais características presentes nas músicas melhor definem o estilo de música de uma década.

Partindo das análises da primeira monografia, será possível comparar se características que fortemente influenciaram a classificação das faixas populares no Brasil apresentam o mesmo comportamento se analisadas músicas lançadas ao redor do mundo e em um período mais extenso. Caso contrário, identificar quais atributos possuem mais influência se analisadas músicas sob uma visão global.

5. Etapas e Cronograma

De acordo com os passos descritos na seção 3, a organização das tarefas a nível de semanas será feita da seguinte forma:

Semana	Tarefa
24/11	Iniciar exploração dos dados, através de visualizações
01/12	Geração de visualizações de dados que evidenciem correlação temporal, tendências, rupturas e/ou padrões
08/12	Criar visualizações de dados para períodos específicos da história
15/12	Confecção do pitch para apresentação parcial
05/01	Utilização de um algoritmo de classificação para determinar se as músicas são facilmente classificadas de acordo com sua década
12/01	Utilização de um algoritmo de classificação para determinar se as músicas são facilmente classificadas de acordo com sua década
19/01	Utilização de métodos de regressão descontínua para identificar alterações nos padrões musicais
26/01	Produzir um compilado de comparações dos resultados obtidos neste trabalho com aqueles da monografia anterior
02/02	Confecção do pitch para apresentação final
09/02	Finalização do relatório
16/02	Finalização do relatório

6. Referências Bibliográficas

FORSBERG, Jukka-Pekka. Quarantstream: how COVID-19 restrictions influenced song features on Spotify TOP-200 charts?. Aalto University. Espoo, Finlândia, 2021. Disponível em: <https://aaltodoc.aalto.fi/handle/123456789/109669>. Acesso em: 18 nov. 2021.

K. KALUSTIAN, Kework; RUTH, Nicolas. Evacuate the Dancefloor: Exploring and Classifying Spotify Music Listening Before and During the COVID-19 Pandemic in DACH Countries. Jahrbuch Musikpsychologie. 2021. Disponível em: <https://jbdgm.psychopen.eu/index.php/JBDGM/article/view/95/95.pdf>. Acesso em: 18 nov. 2021.

USING Spotify Audio Features to Study the Evolution of Pop Music. In: Elena Georgieva e Blair Kaneshiro. [S. l.]: WiMIR 1st Annual Workshop, 2018. Disponível em: <https://ccrma.stanford.edu/~egeorgie/documents/WiMIR2018.pdf>. Acesso em: 18 nov. 2021.

MAUCH, Matthias; MACCALLUM, Robert; LEVY, Mark; LEROI, Armand. The evolution of popular music: USA 1960?2010. The evolution of popular music: USA 1960?2010, [s. l.], 2015. Disponível em: <https://royalsocietypublishing.org/doi/pdf/10.1098/rsos.150081>. Acesso em: 18 nov. 2021.

PINARBASI, Fatih. Demystifying Musical Preferences At Turkish Music Market Through Audio Features Of Spotify Charts: Multidisciplinary Approaches to Ethics in the Digital Era. Demystifying Musical Preferences At Turkish Music Market Through Audio Features Of Spotify Charts, [s. l.], dezembro 2019. Disponível em: https://www.researchgate.net/publication/338181236_DEMYSTIFYING_MUSICAL_PREFERENCES_AT_TURKISH_MUSIC_MARKET_THROUGH_AUDIO_FEATURES_OF_SPOTIFY_CHARTS. Acesso em: 18 nov. 2021.