

André Palhares Costa

Uma análise de dados sobre a música popular nacional ao longo dos anos

Proposta de Pesquisa Científica

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

Orientador: Flávio Vinicius Diniz de Figueiredo

Belo Horizonte, Minas Gerais
2021

SUMÁRIO

Introdução	2
1.1. Objetivos Gerais	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 uma das partes fundamentais 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. Ao mesmo tempo que definem a cultura de um determinado período, ela pode transcender gerações.

Com este trabalho, deseja-se investigar como a música popular nacional tem evoluído ao longo dos anos, utilizando-se de uma análise de dados sobre os atributos musicais das faixas mais populares entre os anos de 1980 e 2021 fornecidos pela API do aplicativo Spotify.

Nossa cultura é marcada em grande parte pela música popular e analisar a evolução de atributos dessas faixas é uma forma de entender melhor a cultura musical do Brasil. Deseja-se investigar a frequência com que esses atributos estão presentes nas músicas ao longo dos anos, quais destes mantiveram uma frequência média mais estável ou aqueles que tiveram um pico em algum determinado momento e depois já não estavam mais tão presentes nas faixas mais populares.

1.1. Objetivos Gerais

- Analisar atributos musicais das músicas populares nacionais e sua evolução ao longo dos anos, para identificar características marcantes da nossa cultura.

1.2. Objetivos Específicos

- Selecionar uma lista de músicas nacionais populares ao longo das últimas 3 décadas
- Obter informações sobre as músicas, como seus atributos musicais
- Identificar possíveis padrões nesses atributos em períodos específicos
- Comparar padrões ao longo dos anos

2. Referencial Teórico

O estudo que será apresentado é pautado na análise de atributos musicais fornecidos pela API pública do Spotify. Para entender esta análise, é necessário conhecer a definição destes atributos. São eles:

- *Duration_ms*: Duração, em milissegundos, da faixa.
- *Danceability*: Valor que indica quão adequada para dançar esta faixa é. É obtido através da união de vários elementos como batimentos por minuto, intensidade da batida etc.
- *Energy*: Valor percentual que indica o quão intensa uma faixa é. É medido através de elementos como range dinâmico, timbre, altura etc.
- *Key*: Nota musical geral da faixa. 0=C, 1=C#, 2=D etc.
- *Loudness*: Valores em decibéis da faixa. Variam de -60 a 0.
- *Speechiness*: Indica a frequência de palavras faladas em uma música.
- *Acousticness*: Revela quão acústica uma faixa é. Valores entre 0 e 1.
- *Instrumentalness*: Corresponde a frequência de sons vocais na faixa. Músicas com este atributo acima de 0.5 tendem a ser instrumentais.
- *Liveness*: Mede a presença de sons de uma plateia na faixa, indicando que pode ser uma gravação ao vivo.
- *Valence*: Indica a positividade da música. Valores altos tendem a representar músicas que transmitem um sentimento positivo.
- *Tempo*: Batidas por minuto.
- *Time_Signature*: Valor geral indicando o compasso de uma música.

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 compara anos com espaçamento de 10 anos entre eles, sem mencionar aqueles intermediários.

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 também serão objeto de estudo neste trabalho para identificar padrões no consumo de música especificamente na Túrquia 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 nos Estados Unidos.

3. Metodologia

O problema será abordado em 3 etapas: obtenção e tratamento dos dados, visualização dos mesmos e sua posterior classificação. Abaixo estão brevemente detalhados cada um dos passos.

1º passo: Buscar as músicas que serão analisadas distribuídas por décadas

Para obter as músicas que serão objeto de estudo deste trabalho, utilizaremos o site MaisTocadas, que reúne uma lista de músicas mais ouvidas nas rádios a cada ano. Neste site, serão selecionadas apenas as músicas brasileiras mais tocadas de 1980 até o ano de 2016.

Também deve-se buscar playlists que reúnam as músicas mais populares de cada década (1980 adiante) disponíveis no serviço do Spotify. Então uma nova playlist será criada unindo as músicas listadas no site MaisTocadas e as músicas obtidas através de outras playlists já existentes no Spotify. Uma playlist será criada para cada década: anos 80, 90, 00 e 10 (contando de 2010 a 2016).

Por fim, a lista de músicas mais executadas no Spotify Brasil diariamente desde o ano de 2017 até 01/06/2021, disponibilizada no site SpotifyCharts, será adicionada a uma nova playlist referente às músicas atuais.

2º passo: Obter os atributos musicais das músicas presentes em cada uma das playlists

Tendo sido criadas as 5 playlists, sendo 4 delas com músicas separadas por década e 1 com as músicas mais populares de 2017 a 2021, serão obtidos atributos musicais, os chamados *audio features*, de cada uma das músicas através da API pública do Spotify. Para cada uma das playlists de década, serão removidas as músicas cujo atributo referente ao ano de lançamento não for correspondente à década esperada.

Por fim, os dados selecionados serão armazenados em DataFrame do Python e convertidos em um arquivo .csv.

3º passo: Gerar visualizações de dados

Nesse passo, pretende-se exibir as informações obtidas no passo anterior em visualizações de gráficos e/ou tabelas que permitam 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 investigadas variadas técnicas de exibição de dados e buscar interpretações sobre os resultados obtidos, através de explicações sobre estilos musicais das décadas passadas, tendências da música popular no serviço Spotify, preferências musicais do ouvinte brasileiro e etc.

4º passo: Utilizar um algoritmo de classificação

Por último, será utilizado um algoritmo de classificação que tentará identificar em qual década uma música foi lançada, tomando como base seus demais atributos. O uso desse algoritmo ajudará a entender se as músicas são facilmente identificáveis -- demonstrando a existência de características mais particulares a cada década -- ou se os atributos estão tão bem distribuídos ao longo dos anos que essa conclusão não pode ser tomada.

4. Resultados Esperados

Ao final do trabalho espera-se a obtenção de uma série de visualizações de dados que permita a identificação de possíveis padrões nos traços da música popular nacional ao longo dos anos, bem como evidenciar rupturas no tempo ou até mesmo tendências para os próximos anos. Além disso, prover uma análise geral da evolução da presença de atributos musicais nas faixas mais populares ano a ano.

Além disso, deseja-se também utilizar de um algoritmo de classificação para auxiliar na breve análise de tendência das características das músicas nacionais mais populares de cada ano. A partir do resultado obtido com o algoritmo, será possível identificar se características presentes nas músicas populares do passado se assemelham com aquelas das faixas mais ouvidas atualmente.

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
28/06	Buscar lista de músicas nacionais populares de cada ano de 1980 a 2021 usando WebScraping e o site MaisTocadas;
05/07	Obter atributos musicais para a lista selecionada no passo anterior usando a API pública do Spotify
12/07	Geração de visualizações de dados que evidenciem correlação temporal, tendências, rupturas e/ou padrões
19/07	Confecção do pitch para apresentação parcial;
26/07	Geração de visualizações de dados que evidenciem correlação temporal, tendências, rupturas e/ou padrões
02/08	Análise exploratória dos dados e das visualizações geradas explicando os resultados obtidos
09/08	Utilização de um algoritmo de classificação para determinar se as músicas são facilmente classificadas de acordo com sua década
16/08	Utilização de um algoritmo de classificação para determinar se as músicas são facilmente classificadas de acordo com sua década
23/08	Confecção do pitch para apresentação final;
30/08	Finalização do texto
06/09	Finalização do texto

6. Referências Bibliográficas

Mauch M, MacCallum RM, Levy M, Leroi AM. 2015 The evolution of popular music: USA 1960–2010. *R. Soc. open sci.* 2: 150081. <http://dx.doi.org/10.1098/rsos.150081>