

Atividade 03 - Projeto

Equipe

Nome: Donda

Repositório GitLab da equipe: https://gitlab.com/jr_waine/csb-53

Membros:

Waine Barbosa de Oliveira Junior, 1905120,
jr_waine, waine@alunos.utfpr.edu.br, Eng. Comp., UTFPR
Eduardo Yoshio da Rocha, 1508733,
eduardo2798, eduardo.yoshio@outlook.com, Eng. Comp., UTFPR

Tema

Analisar os fatores determinantes para indicação de melhor álbum no Grammy Awards.

Motivação

É muito comum ouvir críticas acerca do Grammy e suas indicações para melhor música e álbum do ano, alegando que os álbuns são indicados não pela qualidade, mas sim porque foram populares ou sucesso de vendas.

Uma análise criteriosa dos dados de álbuns, como nota da crítica, número de vendas/ouvintes do álbum, deve ajudar a checar os argumentos dessa discussão e seu respaldo na realidade.

Perguntas de pesquisa

- Quais são as características em comum dos álbuns indicados e vencedores do Grammy? Popularidade, relevância/influência, crítica positiva, gênero musical?
- Quais são as semelhanças e diferenças entre álbuns indicados ao Grammy e aqueles aclamados pela crítica?
- Os álbuns indicados ao Grammy são os mais relevantes/influentes do ano?

Hipóteses

- Os álbuns indicados e vencedores do Grammy tem muito mais relação com sua popularidade que sua qualidade técnica (crítica positiva)
- Os gêneros dos álbuns indicados ao Grammy têm pouca diversidade com relação aos álbuns aclamados pela crítica

Dados e modelos

Para as métricas necessárias, coletamos dados de 3 fontes diferentes, [Billboard](#), [Metacritic](#) e dos álbuns indicados ao [Grammy](#). Para todas as fontes foram utilizados os anos de 2000 até 2020. Na [Billboard](#) buscamos os top 200 álbuns mais vendidos dos EUA no ano. No [Metacritic](#) buscamos os top 100 álbuns melhores avaliados do ano e também os álbuns indicados ao Grammy. Para o [Grammy](#), os dados foram obtidos manualmente, sendo obtido o ano, artista, álbum e se o álbum venceu ou não.

As principais métricas que devemos usar são: popularidade, através da colocação na Billboard em um breve período de tempo (até 2 anos depois do lançamento); relevância, através da colocação na Billboard em um longo período de tempo (+3 anos depois do lançamento); crítica, através das notas no [Metacritic](#).

Para responder as perguntas, pretendemos usar principalmente gráficos, comparando os devidos álbuns e devidos anos, para observar quais são as tendências. Podemos usar modelos de clusterização para detectar quais gêneros musicais aparecem mais nos álbuns em destaque, e técnicas de detecção de outliers para detectar álbuns com notas muito baixas e que são populares, ou álbuns com gêneros não muito populares.

Cronograma

1	Estudo das técnicas de modelagem a serem utilizadas
2	Criação dos modelos buscando resultados que evidenciem as hipóteses
3	Obtenção de mais dados caso necessário, para alimentar os modelos
4	Visualização parcial dos resultados gerados a partir dos modelos criados
5	Refinamento dos modelos, buscando representar melhor os dados existentes
6	Exibição dos resultados por meio de gráficos ou outras formas de interesse