



Universidade Presbiteriana Mackenzie

**PROJETO DE ANÁLISE DE DADOS
DE
MÚSICA NO SPOTIFY**

ALEX JUNIOR MOURA DA SILVA

ENZO VEMADO

JOAO PEDRO BARRETO DE MELO

LUCIANO GUIMARAES COSTA

Trabalho apresentado como critério de avaliação da disciplina

PROJETO APLICADO III

São Paulo

2024



SUMÁRIO

1. INTRODUÇÃO	3
1.1. CONTEXTO DO TRABALHO	3
1.2. MOTIVAÇÃO E JUSTIFICATIVA	3
1.3. OBJETIVOS.....	4
1.3.1. OBJETIVO GERAL:.....	4
<i>1.3.2. OBJETIVOS ESPECÍFICOS:</i>	<i>4</i>
1.4. BASE DE DADOS:.....	4
<i>1.4.1. DESCRIÇÃO DA BASE DE DADOS:</i>	<i>4</i>
<i>1.4.2. LIMITAÇÕES DA BASE DE DADOS:</i>	<i>5</i>
<i>1.4.3. ORIGEM DOS DADOS:.....</i>	<i>5</i>
CRONOGRAMA:	6

1. INTRODUÇÃO

1.1. CONTEXTO DO TRABALHO

A música, além de ser uma expressão cultural, é uma indústria em constante evolução, influenciada por tendências sociais, avanços tecnológicos e preferências de consumo. Com o avanço das plataformas de streaming, como o Spotify, houve uma mudança significativa na forma como a música é consumida e distribuída. Essas plataformas não apenas ampliaram o acesso à música, mas também geraram uma grande quantidade de dados que podem ser explorados para entender melhor os padrões dos usuários.

1.2. MOTIVAÇÃO E JUSTIFICATIVA

A relevância do tema se manifesta na sua capacidade de fornecer insights sobre o comportamento dos consumidores e o impacto das tendências musicais em escala global. A escolha do tema foi motivada por interesses pessoais e profissionais em análise de dados e na música. A análise desses dados pode revelar padrões comportamentais e preferências.

Desenvolver este projeto oferece diversas vantagens e benefícios. Primeiramente, ele contribuirá para aumentar o conhecimento sobre o comportamento. Em segundo lugar, o projeto pode contribuir para a aplicação de técnicas de análise de dados em um contexto prático, permitindo a aplicação de novos algoritmos e abordagens analíticas.

Esse desenvolvimento poderá contribuir significativamente para o campo de estudo ao oferecer uma visão detalhada do comportamento de consumo de música e propor modelos preditivos robustos baseados em dados acústicos e demográficos.

1.3. OBJETIVOS

1.3.1. OBJETIVO GERAL:

Analisar e modelar dados musicais do Spotify, explorando características acústicas e padrões de consumo para identificar tendências e propor um sistema de recomendação personalizado.

1.3.2. OBJETIVOS ESPECÍFICOS:

1. Analisar as características acústicas das faixas, como dançabilidade, energia, valência e instrumentalidade.
2. Avaliar o impacto de variáveis como popularidade, ano de lançamento e letras explícitas no consumo musical.
3. Desenvolver modelos preditivos para recomendar músicas com base no perfil de consumo dos usuários.
4. Explorar o comportamento de diferentes grupos demográficos em relação ao consumo de músicas específicas.

1.4. BASE DE DADOS:

1.4.1. DESCRIÇÃO DA BASE DE DADOS:

A base de dados utilizada para este projeto foi extraída da API pública do Spotify, conforme descrito na documentação oficial da plataforma. A base abrange uma série de variáveis que descrevem tanto as faixas musicais quanto o comportamento dos ouvintes. Entre as principais variáveis estão:

- **Acousticness (Acústica):** Mede a confiança de que a faixa é acústica, em uma escala de 0,0 a 1,0.
- **Danceability (Dançabilidade):** Avalia quão adequada uma faixa é para dançar, com valores entre 0,0 e 1,0.
- **Energy (Energia):** Mede a intensidade da faixa, em uma escala de 0,0 a 1,0.
- **Valence (Valência):** Indica a positividade emocional de uma faixa, variando de 0,0 a 1,0.



- **Instrumentalness (Instrumentalidade):** Prediz a ausência de vocais em uma faixa.
- **Tempo:** Refere-se ao ritmo da música em batidas por minuto (BPM).
- **Loudness (Volume em dB):** Representa o volume médio da faixa em decibéis.
- **Popularity (Popularidade):** Avalia a popularidade da música, com valores entre 0 e 100.
- **Explicit (Explícito):** Indica se a faixa contém ou não letras explícitas.
- **Key (Chave):** Representa a tonalidade da faixa.
- **Mode (Modo):** Define se a música está em escala maior (1) ou menor (0).

1.4.2. LIMITAÇÕES DA BASE DE DADOS:

Podemos ter algumas limitações na base, por exemplo as métricas de popularidade podem variar de acordo com o número de reproduções recentes, o que pode introduzir viés temporal. Além disso, a previsão de atributos como instrumentalidade ou vivacidade pode não ser precisa em todos os casos, especialmente para faixas com características ambíguas.

1.4.3. ORIGEM DOS DADOS:

Os dados foram extraídos diretamente da API do Spotify e disponibilizados em formato como tabela no Github (de onde consultamos eles). Cada faixa é identificada por um ID único fornecido pela plataforma, permitindo o rastreamento e a análise de todas as suas características associadas.

CRONOGRAMA:

Etapa	Descrição	Início	Término
ETAPA 1 – APRESENTAÇÃO DO PROJETO E CONCEPÇÃO DO PRODUTO	Definição do Projeto: Escolha da base de dados e definição dos objetivos	08/08/2024	07/09/2024
ETAPA 2 – DEFINIÇÃO DO PRODUTO	Coleta de Dados: Extração dos dados da API do Spotify	04/09/2024	05/10/2024
ETAPA 2 – DEFINIÇÃO DO PRODUTO	Limpeza e Preparação: Tratamento de dados ausentes, normalização, etc.	04/09/2024	05/10/2024
ETAPA 3 – METODOLOGIA	Análise Exploratória (EDA): Análise inicial dos dados (distribuição, correlações)	04/09/2024	02/11/2024
ETAPA 3 – METODOLOGIA	Modelagem Preditiva: Desenvolvimento	04/09/2024	02/11/2024

		de modelos de recomendação		
ETAPA 4 – RESULTADOS E CONCLUSÃO		Visualização de Resultados: Criação de gráficos e relatórios com os insights obtidos	03/09/2024	23/11/2024
ETAPA 4 – RESULTADOS E CONCLUSÃO		Conclusão e Escrita: Redação final do relatório e revisão	03/09/2024	23/11/2024
ETAPA 4 – RESULTADOS E CONCLUSÃO		Submissão do Projeto: Envio do projeto final	23/11/2024	23/11/2024