

Fontes de dados utilizadas

- **Rolling Stone's 500 Greatest Albums of All Time**

<https://www.kaggle.com/datasets/joebeachcapital/rolling-stones-500-greatest-albums-of-all-time>

Fonte: Kaggle

Frequência de Atualização: Anual

Sobre o dataset: O RS 500 foi elaborado pelos editores da Rolling Stone com base nos resultados de duas pesquisas abrangentes. Em 2003, a Rolling Stone pediu a um painel de 271 artistas, produtores, executivos da indústria e jornalistas que escolhessem os melhores álbuns de todos os tempos. Em 2009, pedimos a um grupo semelhante de 100 especialistas que escolhessem os melhores álbuns dos anos 2000. A partir desses resultados, a Rolling Stone criou esta nova lista dos melhores álbuns de todos os tempos.

- **Spotify User Behavior Dataset**

<https://www.kaggle.com/datasets/meeraajayakumar/spotify-user-behavior-dataset>

Fonte: Kaggle

Frequência de Atualização: Sem frequência

Sobre o dataset: O Conjunto de Dados de Comportamento do Usuário do Spotify é uma coleção abrangente de dados anonimizados que oferece insights valiosos sobre os padrões de comportamento e preferências dos usuários do Spotify. Este conjunto de dados primário foi projetado especificamente para analisar interações do usuário, hábitos de consumo de música e métricas de engajamento na plataforma de streaming de música Spotify. Ele fornece a pesquisadores, analistas de dados e cientistas de dados um recurso valioso para explorar e compreender o comportamento do usuário em profundidade.

- **Data on Songs from Billboard 1999-2019**

<https://www.kaggle.com/datasets/danield2255/data-on-songs-from-billboard-19992019>

Fonte: Kaggle

Frequência de Atualização: Sem frequência

Sobre o dataset: Alguns desses dados foram coletados para um projeto de pesquisa na Universidade Politécnica Estadual da Califórnia, em San Luis Obispo. Esses dados são ótimos para diversas rotas de análise exploratória de dados, análise de texto, engenharia de recursos ou aprendizado de máquina.

Desafios encontrados

O maior trabalho se dá no entendimento dos dados, transformações e formatações (adequação de layout). Quanto a descoberta dos dados, foi algo simples pois o kaggle já entrega todas as formas de acesso.

Estratégia para resolução dos desafios encontrados

A estratégia foi executar pré-processamentos para adequação dos dados para o algoritmo.