
Spotify Song Attributes

Dados da plataforma para competição de Data Science: **Kaggle**

CONTEÚDO:

Cada linha da base de dados representa uma música (aqui também chamada de faixa).

Existem 14 colunas (variáveis). Uma coluna indica o ID da música, uma para o nome da música, uma para o artista, uma coluna chamada *liking*¹ que indica preferência da pessoa pela música e mais 10 colunas que indicam atributos quantitativos de cada faixa.

Esses atributos quantitativos da faixa medem características como: acústica, se é dançante, duração, energia, instrumentalidade, se é gravada ao vivo, qualidade do som, presença de palavras, positividade, entre outras.

A Tabela 1 apresenta os códigos das variáveis utilizadas na base de dados e seus respectivos significados.

¹ **Atenção:** em particular, a variável *liking* não representa fielmente os dados disponibilizados pelo [George McIntire](#) no Kaggle. Logo, toda conclusão quanto ao perfil de *playlist* dele é fictícia.

Tabela 1. Código e significado das variáveis da base de dados

Nome da variável	Significado da variável
musica_id	indica um número (identificador) da faixa.
acousticness	medida de confiança de 0,0 a 1,0 que indica se a faixa é acústica. Valor próximo de 1,0, maior a confiança da faixa ser acústica.
danceability	medida de confiança de 0,0 a 1,0 que indica o quanto a faixa é adequada para dançar com base em uma combinação de elementos musicais, ritmo, estabilidade de ritmo, força de batida e regularidade geral. Valor próximo de 0,0 é menos dançável e de 1,0, mais dançável.
duration_ms	duração da faixa, em milissegundos.
energy	medida de confiança 0,0 a 1,0 que é perceptiva de intensidade e atividade da faixa. As faixas energéticas parecem rápidas, altas e barulhentas. Por exemplo, death metal tem alta energia (valor mais próximo a 1), enquanto um prelúdio de Bach é baixo na escala (valor mais próximo de zero).
instrumentalness	medida de confiança de 0,0 a 1,0 que indica se a faixa não contém vocais. Os sons "Ooh" e "aah" são tratados como instrumentais neste contexto. As faixas de rap ou de palavras faladas são claramente "vocais". Quanto mais próximo de 1,0, maior a confiança da faixa não conter conteúdo vocal. No caso, valor acima de 0,5 tem a intenção de representar uma faixa instrumental.
liveness	medida de confiança de 0,0 a 1,0 que indica se a gravação foi feita ao vivo ou não (por exemplo, gravação em show). Valor mais próximo de 1,0 indica que a faixa foi gravada ao vivo. No caso, valor acima de 0,8 oferece uma forte probabilidade de ter sido gravada ao vivo.
loudness	medida do volume geral de uma faixa, em decibéis (dB). Este valor de volume representa a qualidade de um som em termos de força (amplitude) da música. Os valores variam tipicamente entre -60 dB e 0 dB e quanto maior, mais qualidade.
speechiness	medida de confiança de 0,0 a 1,0 que indica presença de palavras faladas em uma faixa. Valor acima de 0,66 descreve que faixa provavelmente é composta inteiramente por palavras faladas. Valor entre 0,33 e 0,66 descreve que faixa pode conter música e fala, em seções ou em camadas, incluindo casos como música rap. Valor abaixo de 0,33 sugere que faixa não tem fala.
tempo	tempo geral estimado de uma faixa em batimentos por minuto (BPM). Na terminologia musical, tempo é a velocidade ou ritmo de uma faixa que deriva diretamente da duração média da batida.
valence	medida de confiança de 0,0 a 1,0 que descreve a positividade musical transmitida por uma faixa. Valor mais próximo a 1,0 soa mais positiva (alegre, comemorativa, eufórica). Valor próximo de 0,0 soa mais negativa (triste, deprimida, zangada).
song_title	indica título da faixa (música)
artist	indica nome do artista
liking	indica se gosta da faixa (Like) ou se não gosta (Dont Like).

Fonte: Textos extraídos e/ou adaptados das seguintes referências:

<https://www.kaggle.com/jonathanbouchet/another-take-on-song-s-popularity-prediction#comments>

<https://medium.com/data-hackers/an%C3%A1lise-de-dados-de-uma-playlist-do-spotify-com-python-e-power-bi-b3149c2bd633>

<https://towardsdatascience.com/is-my-spotify-music-boring-an-analysis-involving-music-data-and-machine-learning-47550ae931de>