**Análise da energia musical usando regressão logística no Spotify**

*Carla Kim Gaieski, Giulia Braga Passarelli, Rebeca Moreno dos Santos*

**Introdução**

Imagine que você é um músico e quer compor uma música, porém quer que ela seja calma, qual a probabilidade de você usar o tom menor e qual nota provavelmente usaria? O objetivo desse projeto é responder esta pergunta prevendo qual seria a decisão que você tomaria nesta situação. Assim, temos que comparar os tons e as notas com o quão dançável as músicas são, relacionando isso com o grau de energia das mesmas.

Nesse caso, o método escolhido foi a regressão pois é uma das ferramentas mais utilizadas para investigar a relação entre variáveis na estatística. Ela opera equacionando a variação de uma variável dependente por conta da alteração de uma independente. Para isso, há dois modelos que são mais comuns: a regressão linear e a regressão logística. A primeira é considerada quando a variável de saída é quantitativa, pois o gráfico apresenta um comportamento mais linear sendo, geralmente, mais fácil de encaixar uma função em cima. Já o segundo é próprio para a análise de uma variável de saída categórica, que não se porta da mesma forma que o último. Como queremos analisar as notas e os tons musicais (variáveis qualitativas), elas devem ser a saída em uma regressão logística.

A regressão logística possui uma função para a probabilidade de um ponto que pode ser definida como: .

**Procedimento:**

O Spotify é um serviço de música digital com um acervo de milhões de músicas, de forma a permitir o acesso, a qualquer momento, de uma coleção imensa de canções. Por conta disso, o projeto utiliza a base de dados do Spotify para realizar essa análise. Coletamos uma amostra contendo 100 músicas por meio do dataset “Spotify Song Attributes” do site kaggle.com, utilizando as colunas: “danceability”, “energy”, “key” e “mode”.

Podemos relacionar tom e notas musicais para realizar a análise pois cada nota possui uma tonalidade correspondente que retorna um certo grau enrgético. E, assim, como estamos considerando a energia das músicas, as duas variáveis possuem uma dependência nesse aspecto.

Abaixo está uma tabela com a frequência de cada nota para diferentes tonalidades (tom maior ou tom menor):

Uma imagem contendo captura de tela

Descrição gerada com muito alta confiança