

**Universidade do Minho**  
Escola de Engenharia

# Mineração de Dados

## Trabalho prático – Fase 2

António Silva - PG47033  
Benjamim Costa - PG50258  
Diogo Sambento – PG49999

Ano letivo 2022/2023

# Alterações do tema

## **Tema anterior:**

Analisar a relação que a música pode ter com o estado mental das pessoas.

## **Novo tema:**

Análise dos atributos de uma música de forma a recomendar músicas semelhantes



# Fontes de dados

kaggle



# Visualização e Exploração dos Dados

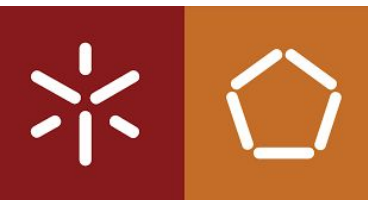
**Dataset:** spotify.csv

**Dimensão do dataset:** 114000 linhas e 20 colunas



## Features:

- track\_id
- track\_artist
- album\_name
- track\_name
- popularity
- duration\_ms
- explicit
- danceability
- energy
- key
- loudness
- mode
- speechiness
- acousticness
- instrumentalness
- liveness
- valence
- tempo
- time\_signature
- track\_genre



# Tipo de dados



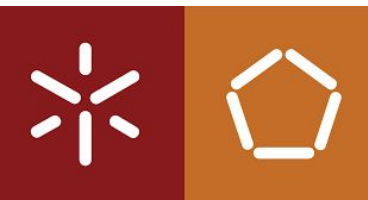
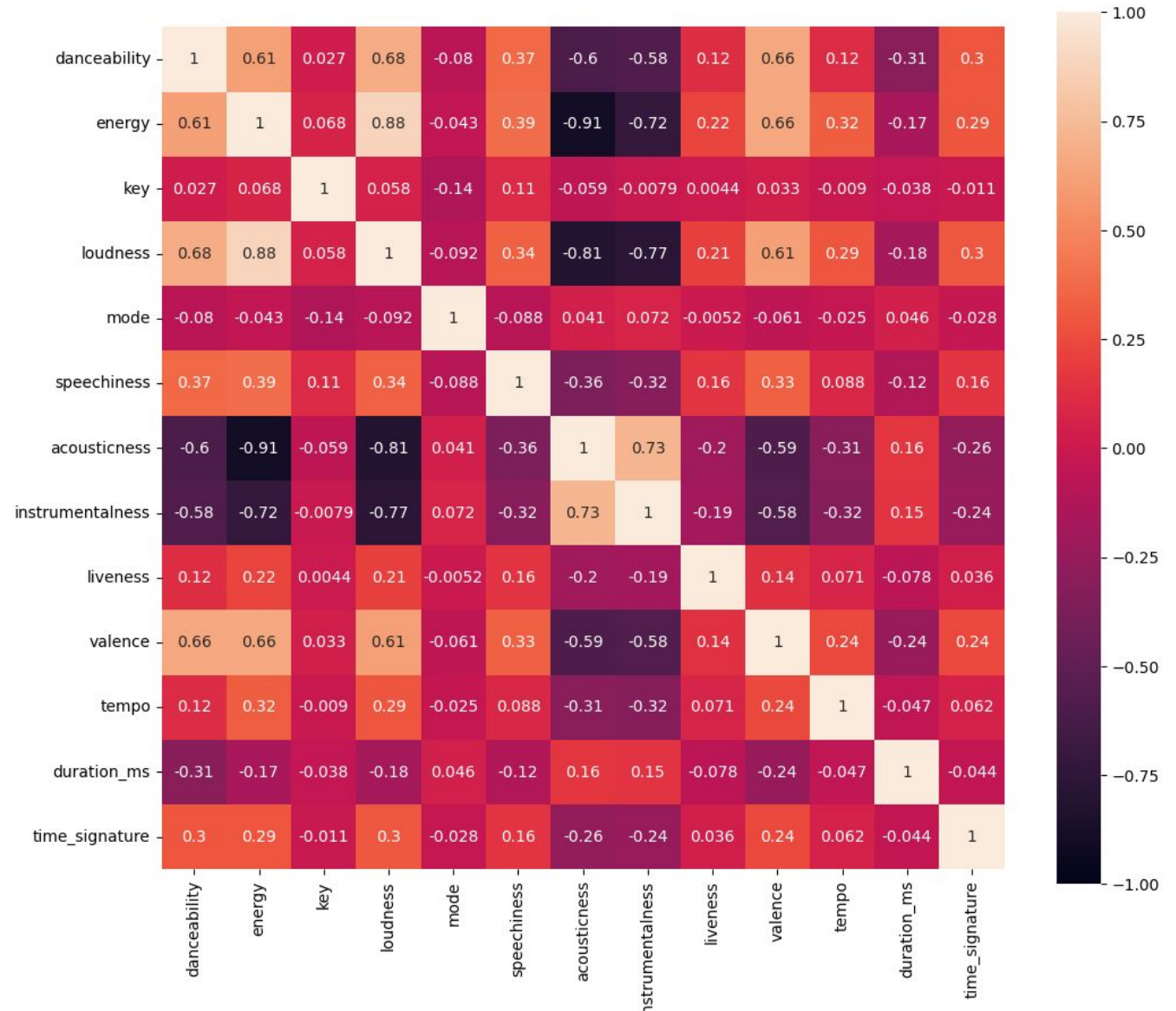
## Features:

- **track\_id** - String
- **track\_artist** - String
- **album\_name** - String
- **track\_name** - String
- **popularity** - Integer entre 0 e 100
- **duration\_ms** - Float
- **explicit** - FALSO ou VERDADEIRO
- **danceability** - Float entre 0.0
- **energy** - Float entre 0.0 e 1.0
- **time\_signature** - Integer entre 3 e 7
- **key** - Integer
- **loudness** - Float entre -60 dB e 0 dB
- **mode** - Integer
- **speechiness** - Float entre 0 e 1
- **acousticness** - Float entre 0 e 1
- **instrumentalness** - Float entre 0 e 1
- **liveness** - Float entre 0 e 1
- **valence** - Float entre 0 e 1
- **tempo** - Float
- **track\_genre** - String



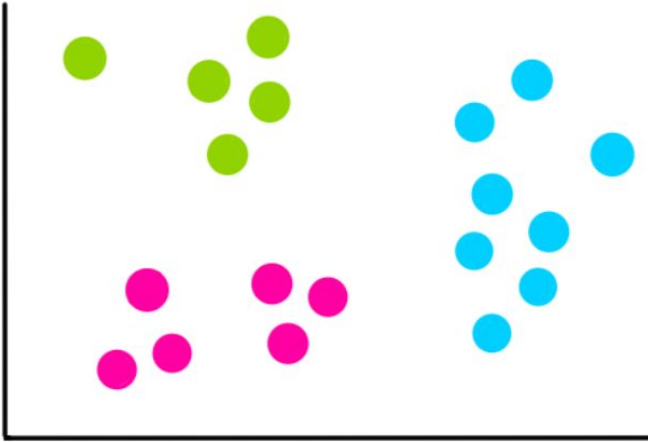
# Preparação e tratamento de dados

- Missing values
- Feature selection
- Dados duplicados

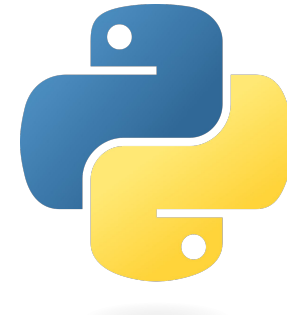


# Modelos e ferramentas utilizadas

- K-means



- Python
- Pandas
- Scikit-learn
- Seaborn
- Matplotlib

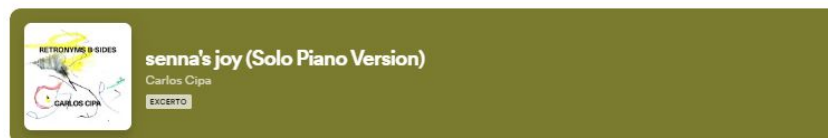


# Página HTML

Se gostas desta música:



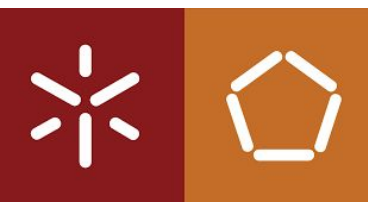
Então experimenta estas: (a ordem é irrelevante)



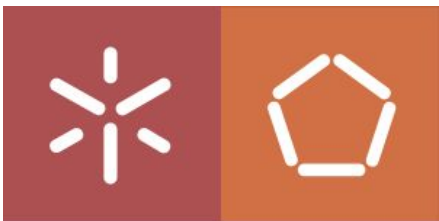
Se gostas desta música:



Então experimenta estas: (a ordem é irrelevante)







**Universidade do Minho**  
Escola de Engenharia

# Mineração de Dados

## Trabalho prático – Fase 2

António Silva - PG47033  
Benjamim Costa - PG50258  
Diogo Sambento – PG49999

Ano letivo 2022/2023