

# FindMe FM

---

Final Project: Group 3

# Project Outline

Topic: Can machine learning predict what songs a user will enjoy based on the audio features of a song they like?

Reason for topic: Interest in the use of Spotify API and other available spotify datasources.

Data Source: Kaggle dataset - Spotify Dataset 1922-2021 ~600k tracks

- Contains info on the audio features of each song (danceability, acousticness, tempo, etc,)
- Dataset is created using the Spotify API

# Data Structure: Tracks

## Primary:

- ID

## Numerical

- acousticness (ranges from 0 to 1)
- danceability (ranges from 0 to 1)
- energy (ranges from 0 to 1)
- duration\_ms (ranges from 0 to 1)
- instrumentalness (ranges from 0 to 1)
- valence (ranges from 0 to 1)
- popularity (ranges from 0 to 1)
- tempo (ranges from 0 to 1)
- liveness (ranges from 0 to 1)
- loudness (ranges from 0 to 1)
- speechiness (ranges from 0 to 1)

## Boolean

- mode (0 = Minor, 1 = Major)
- explicit (0 = No explicit content, 1 = explicit content)

## Categorical

- key (all keys on octave encoded as values ranging from 0 to 11, starting C as 0, C# as 1 and so on...)
- timesignature (the predicted timesignature, most typically 4)
- artists (the artist(s) who made this song)
- artists\_ids (the ids for each artist)
- release\_date (date of when the song was released)
- name (title of the song)

# Questions We Hope to Answer with Data

# Descriptions of the data exploration phase of the project

# Description of the analysis phase of the project

Technologies, languages, tools, and algorithms

