**Overview:**

With the ever-increasing digitalizing of music from creation to distribution and streaming platforms, the amount of music available has exploded in recent years. This has made it more challenging to manage and navigate through the vast collection of music available.

To create a better and easier user experience, companies like Spotify, Apple Music, Soundcloud, and others have utilized machine learning to generate recommendation and search systems to provide users with a more personalized and relevant experience based on their preferences. These companies use music genre classifiers as part of their recommendation system.

**Data Used:**

We have used 2 datasets in our project:

1. **GTZAN Music Classification Dataset -** A collection of around 10000 30-second-long audio files split into 10 genres. The genres are - blues, classical, country, disco, hip-hop, jazz, metal, pop, reggae, and rock.

We have certain numerical features for 3-second segments of the audio files. In addition to this, we used the audio files to generate 3 kinds of image data:

* Waveshow plots
* MFCC plots
* Mel Spectrogram

1. **Spotify Data –** We created a dataset of features from 500 songs for the 10 genres using the SpotiPy library. The features in this dataset are different from those in the GTZAN dataset.

**Exploratory Data Analysis:**