# **Spotify and Youtube**

Dataset of songs of various artist in the world and for each song is present:

* Several statistics of the music version on spotify, including the number of streams;
* Number of views of the official music video of the song on youtube.

# **Content**

It includes 28 variables for each of the songs collected from spotify. These variables are briefly described next:

* **Track**: name of the song, as visible on the Spotify platform.
* **Artist**: name of the artist.
* **Url\_spotify**: the Url of the song.
* **Album**: the album in wich the song is contained on Spotify.
* **Album\_type**: indicates if the song is relesead on Spotify as a single or contained in an album.
* **Uri**: a spotify link used to find the song through the API.
* **Danceability**: describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable.
* **Energy**: is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy. For example, death metal has high energy, while a Bach prelude scores low on the scale. Perceptual features contributing to this attribute include dynamic range, perceived loudness, timbre, onset rate, and general entropy.
* **Key**: the key the track is in. Integers map to pitches using standard Pitch Class notation. E.g. 0 = C, 1 = C♯/D♭, 2 = D, and so on. If no key was detected, the value is -1.
* **Loudness**: the overall loudness of a track in decibels (dB). Loudness values are averaged across the entire track and are useful for comparing relative loudness of tracks. Loudness is the quality of a sound that is the primary psychological correlate of physical strength (amplitude). Values typically range between -60 and 0 db.
* **Speechiness**: detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audio book, poetry), the closer to 1.0 the attribute value. Values above 0.66 describe tracks that are probably made entirely of spoken words. Values between 0.33 and 0.66 describe tracks that may contain both music and speech, either in sections or layered, including such cases as rap music. Values below 0.33 most likely represent music and other non-speech-like tracks.
* **Acousticness**: a confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic.
* **Instrumentalness**: predicts whether a track contains no vocals. "Ooh" and "aah" sounds are treated as instrumental in this context. Rap or spoken word tracks are clearly "vocal". The closer the instrumentalness value is to 1.0, the greater likelihood the track contains no vocal content. Values above 0.5 are intended to represent instrumental tracks, but confidence is higher as the value approaches 1.0.
* **Liveness**: detects the presence of an audience in the recording. Higher liveness values represent an increased probability that the track was performed live. A value above 0.8 provides strong likelihood that the track is live.
* **Valence**: a measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry).
* **Tempo**: the overall estimated tempo of a track in beats per minute (BPM). In musical terminology, tempo is the speed or pace of a given piece and derives directly from the average beat duration.
* **Duration\_ms**: the duration of the track in milliseconds.
* **Stream**: number of streams of the song on Spotify.
* **Url\_youtube**: url of the video linked to the song on Youtube, if it have any.
* **Title**: title of the videoclip on youtube.
* **Channel**: name of the channel that have published the video.
* **Views**: number of views.
* **Likes**: number of likes.
* **Comments**: number of comments.
* **Description**: description of the video on Youtube.
* **Licensed**: Indicates whether the video represents licensed content, which means that the content was uploaded to a channel linked to a YouTube content partner and then claimed by that partner.
* **official\_video**: boolean value that indicates if the video found is the official video of the song.

Music And Mental Health Survey

Music therapy, or MT, is the use of music to improve an individual's stress, mood, and overall mental health. MT is also recognized as an evidence-based practice, using music as a catalyst for "happy" hormones such as oxytocin.

However, MT employs a wide range of different genres, varying from one organization to the next.

The **MxMH** dataset aims to identify what, if any, correlations exist between an individual's music taste and their self-reported mental health. Ideally, these findings could contribute to a more informed application of MT or simply provide interesting sights about the mind. The dataset contains 33 variables.

# **Interpreting data**

1. **Timestamp-**Date and time when form was submitted
2. **Age-**Respondent's age
3. **Primary streaming service-**Respondent's primary streaming service
4. **Hours per day-**Number of hours the respondent listens to music per day
5. **While working-**Does the respondent listen to music while studying/working?
6. **Instrumentalist-**Does the respondent play an instrument regularly?
7. **Composer-**Does the respondent compose music?
8. **Fav genre-**Respondent's favorite or top genre
9. **Exploratory-**Does the respondent actively explore new artists/genres?
10. **Foreign languages-**Does the respondent regularly listen to music with lyrics in a language they are not fluent in?
11. **BPM-**Beats per minute of favorite genre
12. **Frequency [Classical]-**How frequently the respondent listens to classical music
13. **Frequency [Country]-**How frequently the respondent listens to country music
14. **Frequency [EDM]-**How frequently the respondent listens to EDM music
15. **Frequency [Folk]-**How frequently the respondent listens to folks music
16. **Frequency [Gospel]-**How frequently the respondent listens to Gospel music
17. **Frequency [Hip hop]-**How frequently the respondent listens to hip hop music
18. **Frequency [Jazz]-**How frequently the respondent listens to jazz music
19. **Frequency [K pop]-**How frequently the respondent listens to K pop music
20. **Frequency [Latin]-**How frequently the respondent listens to Latin music
21. **Frequency [Lofi]-**How frequently the respondent listens to lofi music
22. **Frequency [Metal]-**How frequently the respondent listens to metal music
23. **Frequency [Pop]-**How frequently the respondent listens to pop music
24. **Frequency [R&B]-**How frequently the respondent listens to R&B music
25. **Frequency [Rap]-**How frequently the respondent listens to rap music
26. **Frequency [Rock]-**How frequently the respondent listens to rock music
27. **Frequency [Video game music]-**How frequently the respondent listens to video game music
28. **Anxiety-**Self-reported anxiety, on a scale of 0-10
29. **Depression-**Self-reported depression, on a scale of 0-10
30. **Insomnia-**Self-reported insomnia, on a scale of 0-10
31. **OCD-**Self-reported OCD, on a scale of 0-10
32. **Music effects-**Does music improve/worsen respondent's mental health conditions?
33. **Permissions-**Permissions to publicize data

Students Exam Scores Dataset

#### **Description-**

This dataset includes scores from three test scores of students at a (fictional) public school and a variety of personal and socio-economic factors that may have interaction effects upon them.

This datasets are **fictional** and should be used for **educational purposes only**

#### **Column Descriptions-**

1. **ID-**Student ID
2. **Gender-**Student Gender
3. **EthnicGroup-** Ethnicity of the student
4. **ParentEduc-**Educational Background of the Student
5. **LunchType-**Lunch Type the student has at school
6. **TestPrep-**Had the student test preparation course
7. **ParentMaritalStatus-** Marital Status of the parents
8. **PracticeSport-**Does the student practice sport or not
9. **IsFirstChild-** Is first child of the parents
10. **NrSiblings-**Number of siblings the student has