In [1]:
```
### Called Dataset
```

In [3]:
```
import os
from IPython.display import Image, display

# Get the current working directory
directorio_actual = os.getcwd()

# Specify the relative path from the current directory
imagen = os.path.join('..', 'data', '01_raw', 'assets', 'img', 'Spotify Imagen.jpg')

# Show image
display(Image(filename=imagen, width=500))
```

Dataset explanation

The dataset contains the main fields of the songs, genres and measurable values to guide the user to listen to more music that he likes.

The goal is to create one or many models that allow recommendations to be delivered to the user based on their preferences. To do this, columns related to energy, tonality, duration, loudness, etc. will be taken into account.

To explore data more effectively, we will answer these questions in the second point called 'Data Understanding'.

1) How many songs are there per music genre?

2) What is the average length of each song and what is the most common genre?

3) Who is the artist with the most songs on the platform?

4) What is the most popular song according to the musical genre?

- track_id: The Spotify ID for the track.
- artists: The names of the artists who performed the track. If there is more than one artist, they are separated by a ;.
- album_name: The name of the album the track appears on.
- track_name: Name of the track.
- popularity: A track's popularity is a value between 0 and 100, with 100 being the most popular. Popularity is calculated by an algorithm and is mostly based on the total number of plays the track has had and how recent those plays are. Generally speaking, songs that are being played a lot now will have a higher popularity than songs that were played a lot in the past. Duplicate tracks (for example, the same track from a single and an album) are rated independently. Artist and album popularity are mathematically derived from the popularity of the tracks.
- duration_ms: The duration of the track in milliseconds.

- explicit: Indicates whether the track has explicit lyrics (true = yes; false = no or unknown).
- danceability: Danceability describes how suitable a track is for dancing, based on a combination of musical elements such as tempo, rhythm stability, pulse strength and overall regularity. A value of 0.0 is the least danceable and 1.0 is the most danceable.
- energy: Energy is a measurement from 0.0 to 1.0 that represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and strong. For example, death metal has high energy, while a Bach prelude scores low on the scale.
- key: The key of the track. Integers are assigned to pitches using standard Pitch Class notation. For example, 0 = C, 1 = C♯/D♭, 2 = D, and so on. If no key is detected, the value is -1
- loudness: The overall loudness of a track in decibels (dB).
- mode: The mode indicates the modality (major or minor) of a track, the type of scale from which its melodic content is derived. Major is represented by 1 and minor by 0.
- speechiness: Speechiness detects the presence of spoken words in a track. The more exclusively spoken the recording is (e.g. talk show, audiobook, poetry), the closer the attribute value is to 1.0. Values greater than 0.66 describe tracks that are likely to be composed entirely of spoken words. Values between 0.33 and 0.66 describe tracks that may contain both music and speech, either in sections or in layers, including cases such as rap music. Values less than 0.33 likely represent music and other non-spoken tracks.
- acousticness: A confidence measure from 0.0 to 1.0 on whether the track is acoustic. 1.0 represents high confidence that the track is acoustic.
- instrumentalness: Predicts whether a track contains no vocals. “Ooh” and “aah” sounds are considered instrumental in this context. Rap or spoken word tracks are clearly “vocal.” The closer the instrumental value is to 1.0, the more likely the track is to contain no vocals.
- liveness: Detects the presence of an audience in the recording. Higher liveness values represent a higher probability that the track was performed live. A value greater than 0.8 indicates a high probability that the track was recorded live.
- valence: A measure from 0.0 to 1.0 that describes the musical positivity conveyed by a track. Tracks with a high valence sound more positive (e.g. happy, joyful, euphoric), while tracks with a low valence sound more negative (e.g. sad, depressed, angry).
- tempo: The estimated overall tempo of a track in beats per minute (BPM). In musical terminology, tempo is the speed or pace of a given piece and is derived directly from the average bar length.
- time_signature: An estimated time signature. A time signature is a notational convention for specifying how many beats are in each measure. A time signature ranges from 3 to 7, indicating 3/4 to 7/4 time signatures.
- track_genre: Genre to which the track belongs