

By Euodia Ajonye

Spotify Dataset 2000-2019

This project was done using a Spotify dataset from Kaggle at:

<https://www.kaggle.com/datasets/paradisejoy/top-hits-spotify-from-20002019/code>.

Information On The Dataset

Danceability: describes how suitable a song is for dancing

Energy: is a measure from 0.0 to 1.0 and represents perpetual measure of intensity and activity

Speechiness: detects the presence of spoken words in a song

Acousticness: This measures from 0.0 to 1.0 whether the song is acoustic

Liveness: Higher liveness values represent an increased probability that the track was performed live

Valence: This describes the musical positiveness conveyed by a track from 0.0 to 1.0

To carry out this analysis, I renamed some columns, change the column names to uppercase

Summary Of Findings

1. The Neighbourhood's Sweater Weather is an Indie Rock Pop(Rock,pop is a sub-genre of Pop) band, which makes sense because Rock Pop is the most popular genre, the other artists on this list however are Pop artist.
2. Rock, pop is the most popular genre
3. Pop is the most repetitive genre.
4. 0.8 is the most popular energy frequency
5. Rihanna, Drake and Eminem are among the most frequent artists is the most frequent artist
6. Camille Jones has the least popular song
7. The duration of songs varies from song to song .
8. There is some correlation between energy and acousticness.
9. As the popularity of the song increases, the danceability score for that song also increases.
10. The year with the highest song collection is 2012.
11. The artist with the most songs is Rihanna and the lowest is Camille Jones
12. The 2010's had the highest number of songs
13. The 2010s had the songs with the highest number of explicit content
14. Songs with explicit content are less popular than songs without explicit content .
15. Only 27.6% of songs have explicit content.

Key Insights For Presentation

I picked three visualizations

1. Correlation plot
2. Pie chart
3. Countplot