

## Dataset Source:

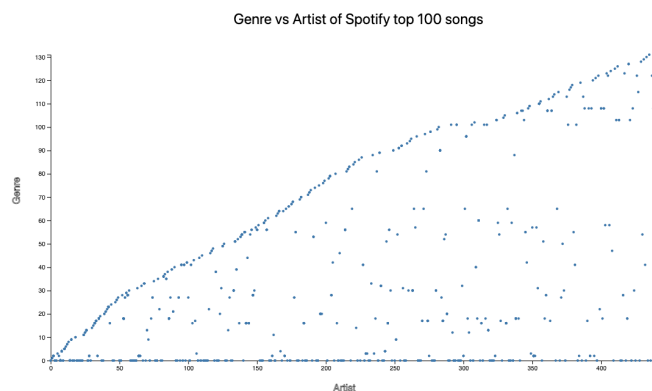
<https://www.kaggle.com/datasets/muhmores/spotify-top-100-songs-of-20152019?select=Spotify+2010+-+2019+Top+100+Songs.xlsx>

## Attributes:

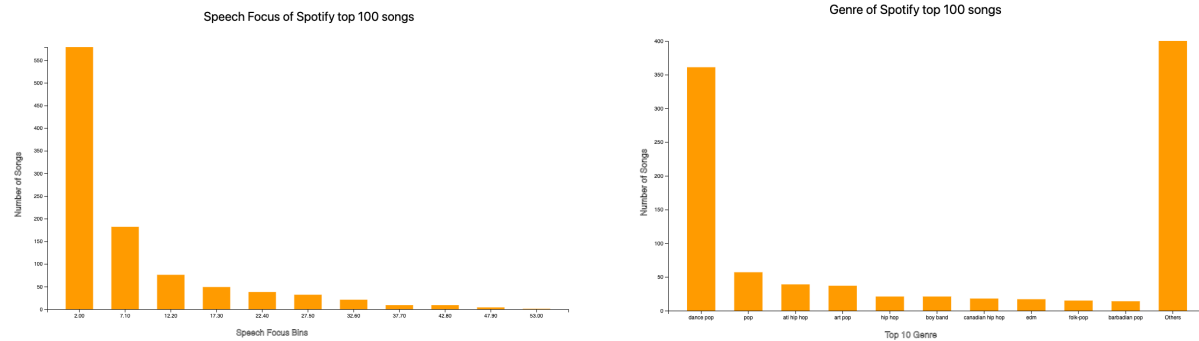
|                 |   |
|-----------------|---|
| Artist          | Song's artist                                 |
| Genre           | Genre of song                                 |
| Released        | Year the song was released                    |
| Tempo           | Beats Per Minute - The tempo of the song      |
| Energy          | How energetic the song is                     |
| Danceability    | How easy it is to dance to the song           |
| Intensity       | Decibel - How loud the song is                |
| Live Likelihood | How likely the song is a live recording       |
| Positiveness    | How positive the mood of the song is          |
| Duration        | Duration of the song                          |
| Acoustic        | How acoustic the song is                      |
| Speech Focus    | The more the song is focused on spoken word   |
| Popularity      | Popularity of the song (not a ranking)        |
| Top Year        | Year the song was a top hit                   |
| Artist Type     | Tells if artist is solo, duo, trio, or a band |

## Why this dataset is interesting

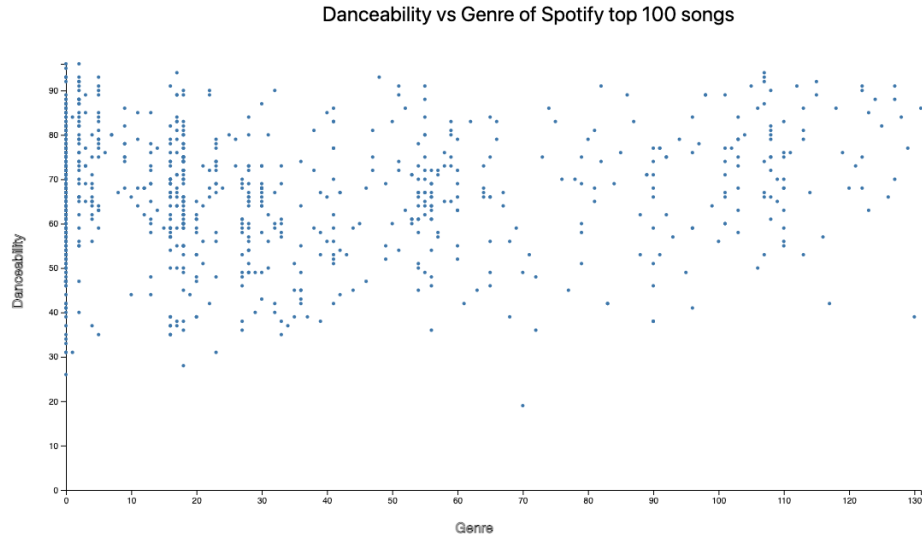
I personally love singing and making music. So this dataset was fascinating to me. I found the concept of analyzing what makes the mass populace tick in terms of music consumption very intriguing. This dataset presents a quantifiable picture of what raw features make a song popular among people. For example, the dataset contains popular artists and popular genres that made it to the top 100 spots over the years on Spotify. But visualizing them together in a scatter plot tells us which genre of songs is more popular among which artist. This is a much more valuable insight to understand the impact of the genre as well as artists on how popular a song is.



Additionally, it also allows us to visualize the impact of raw acoustic features on how popular a song is. For example, we can see that most songs that made it to the top 100 focus less on speech and more on instrumentation and musical composition. This is consistent with another observation that the most popular genre of songs in the top 100 spots is dance pop. Dance pop genre usually comes with a high effort toward instrumentation and musical composition to make the song more dance-worthy.



This can be extended to see the genre vs danceability scatter plot which tells us that certain genres have much more dance-worthy songs than others. In the adjacent plot, genre 0 is dance-pop and we can clearly see that the songs in the dance-pop genre have high danceability. This is consistent with the previous observations.



Overall, this dataset is a very cool way to draw correlations between features that make a song popular among the mass population and also to visualize how impactful each feature is.