Decoding What Makes a Number One Song

Jasmine Lau, Jane Li, Kimsean Pen, YueFeng Xue

DATASCI 200 — 15 December 2023

Introduction

Spotify is the global music streaming service. Much like the services of Netflix or Hulu that serve the purpose of streaming movies or TV shows, Spotify primarily serves as a platform to stream podcasts and music. Spotify's accessibility through its web platform, app, and compatibility with a variety of smart home products has made it the most popular music streaming service globally. We chose Spotify as the platform to study due to it having the largest user base.

To find out more about what people are listening to and what is enjoyed by the general public, we want to dive deeper into the music trends of 2023. We will be using a dataset containing data from the top songs of 2023 according to Spotify's charts. Our main goal is to find what attributes make a top song, and our other goals include finding if music popular on Spotify and ranks highly on the Spotify charts, is popular on other platforms' charts such as Apple Music or Deezer, to see if that song is truly popular for a larger audience and not just popular for Spotify users. Then we would like to gauge how the release time and the number of artists collaborating on a song would impact overall streaming. Through these questions and the analysis we will perform, we can help artists improve their overall success in terms of streams and connect better with their audience through music.

Dataset & Key Features

The dataset is obtained from Kaggle and contains a comprehensive list of the most popular songs of 2023 that is listed on Spotify. It provides insights into each of the song's attributes, popularity, and presence on various music streaming platforms. The data set contains information on the track name, artist(s) name, release date, Spotify playlists and charts, streaming statistics, Apple Music presence, Deezer presence, Shazam charts, and various audio features of a song. There are 24 attributes and 943 unique values for songs. Refer to Appendix Chart 1.1 for a comprehensive breakdown of the attributes incorporated in our analysis.

Focus Questions

1. **Streaming and Chart Performance** - How do the most streamed songs on Spotify compare in terms of streams and chart ranking with its performance on other platforms such as Apple Music and Deezer, and are there significant discrepancies in song performance across these various music streaming services?
2. **Tops Songs (Ranking - weekly, monthly streams and clicks) VS Most Streamed Songs (total stream time over the year) -** What are the shared musical characteristic among the most streamed songs over the year and the top ranking songs and is there a correlation between a song's ranking in terms of streams and clicks and its overall streaming performance across the years?
3. **Features of Top Songs -** Do features (BPM, perceived danceability, energy, and valence) of the song matter?
4. **Artist Collaboration -** How does artist collaboration impact the streamings of a song with a focus on determining if solo artist or collaboration between artists are more prevalent in charts and playlist, and what is the average number of contributing artists in the most streamed songs? Given the artist count and characteristics of top songs, which artist should be selected for collaboration to maximize song appeal and potential for streaming success?
5. **Release Trends in Top Songs** - What are the prevailing trends in number one songs, such as release timing and age, and how do these factors correlate with their specific current streaming success and likelihood of being released at specific times of the year?

Data Cleaning Process and Exploration

Data issues we solved:

Encoding

How exactly the data was collected is unclear to us, however we can assume that it was scraped from various sources due to the data set having information for several different platforms. As a result we needed to better understand the structure of our data and start the cleaning process. Initially, we were not able to read the dataset using the default UTF-8 encoder, despite it being universal. When we initially encountered this issue, we tried to change the encoding by saving the .csv and specifying which encoding method to use, such as UTF-8, UTF-16, both with and without BOM, but all resulted in the columns, 'track_name' and 'artist(s)_name' still having instances of 'ï¿½'. Additionally, using UTF-8 or UTF-16 resulted in us being unable to even read in the dataset. When attempting to read the dataset in using the default UTF-8 encoder, Python threw a UnicodeDecodeError, so we opted to decode the .csv from Kaggle using ISO-88591-1.

Track and Artist Name (*ï¿½,* �)

Even after explicitly declaring our encoding language, ISO-88591-1, there were still several cases of where the more categorical of columns, the 'track_name' and the 'artist(s)_name' columns had characters, such as an apostrophe, replaced by ï¿½, and in some cases, a repeated ï¿½, for example, the phrase 'Taylor's Version' appeared as 'Taylorï¿½ï¿½ï¿½s Version'. Since these characters that were replaced by 'ï¿½' were so varied, we decided to go into the dataset and replace them with the correct characters. We referenced Spotify for both, the correct 'track_name' and the correct 'artist(s)_name'. In total, there were 106 entries, 75 'track_name' entries and 31 'artist(s)_name' entries that were corrected.

Wrong value type in 'streams'

There was a song that had:
"*BPM110KeyAModeMajorDanceability53Valence75Energy69Acousticness7Instrumentalness0Liveness17Speechiness3*" in its 'streams' column. We manually corrected it by obtaining its stream count from Spotify.

Accuracy of Release Dates

An examination of the release dates within our data set revealed an unusually high occurrence of January 1st entries. This pattern prompted an inquiry, given the improbability of a collective release on this particular date, suggesting it may be employed as default value in instances of ambiguous release timing. The dataset was systematically segmented into four subsections, with each team member conducting a validation check of approximately 240 data entries. This audit determined that 51 of the 65 songs with January 1st as their release date, were misdated, necessitating manual correction to align with the current release dates. Our efforts to authenticate release dates encountered challenges due to incomplete records such as older release songs, conflicting sources, and the complexity inherent in discerning whether the 'track_name' was part of an album or single release. A preference was established for utilizing the Spotify platform's release dates, except where external evidence definitively suggested an earlier release as an individual single.

To evaluate the accuracy of the overall dataset release date, a stratified random sampling strategy was employed with each team member analyzing 55 songs release date from the 949 instances, exclusively against Spotifiy's platform. Other dates from secondary sources were excluded in the comparison to maintain the dataset fidelity to its Spotify origins. This process revealed a 76.73% accuracy rate with the data for release date. In the absence of scrapping our own data and without granular insights into the data aggregation methodology of the Kaggle dataset, the precision of our release date accuracy remains capped by these constraints.

Song Duplications
Following this, we also started to see a handful of songs that had the same 'track_name, but different artists, which made us consider if there were songs with the same track name and the same artists in our dataset. Our best guess as to why this happens, where there are several cases of the exact same song appearing in this dataset more than once, is due to the song making to the top of the charts more than once or are the same, aside from being released as a single or as part of an album. We came to this conclusion because of the four songs that had entries where the 'track_name' and 'artist(s)_name' were the same, they had different values for their other attributes. For example, for some of the cases of the same song, they had differing release dates, and a different number of streams. To consolidate the eight cases (4 songs), we decided to take the song with the earliest release date, take the entry with the greatest number of streams, and average all other values: chart presence, 'danceability_%', 'valence_%', etc.

General cleaning:
NaN Values
Upon closer inspection of the dataset, we observed that certain columns, notably 'key' and 'in_shazam_charts' columns had missing values. Specifically the 'key' column had 95 NaN values, but because this column provides significant information in addressing one of our research questions, we decided to manually fill in the missing values. To address this, we identified the rows with missing 'key' values and utilized song track names to determine what key the song is in. Tunebat served as our primary reference, enabling us to accurately update the missing values.

Zero values
As part of the data completeness check, we scanned the zero values in all columns and found 7 columns containing zero values including: in_spotify_charts: 401, in_apple_playlists: 22 in_apple_charts: 97 in_deezer_playlists: 24 , in_deezer_charts: 554 , acousticness_%: 60 instrumentalness_%: 863. For data completeness concern, "acousticness_%" and "instrumentalness_%" are excluded in the music feature analysis of Question 4.

Dropping Columns
Part of our objective is to find out the correlation between most streamed Spotify and the top songs in other music platforms' playlists and charts. The Spotify 2023 data set we have includes data of Spotify, Apple, Deezer's playlists and charts and data. Shazam's chart fits our needs but after we researched all 4 music platforms, we found out that Shazam is an app that's primarily being used as a song identifier (to find out which song it is by playing part of the song). Based on this information, our team decided to drop the 'in_shazam_charts' column not only because the use of the app is not the same as other major music platforms like Spotify, Apple Music, and Deezer, but also because the column consists of 95 NaN values which will not be able to give us enough insights when doing data analysis. Also we dropped the column "instrumentalness_%" as it contains 867 zeros out of the total 949 records.

Approach
Our analysis will begin by examining the relationship between streaming figures and chart success across multiple music platforms, informing our approach to collaborative song creation. We will then explore whether the number of artists on a track influences streaming behavior, hypothesizing that collaborations might unite fan base and enhance stream counts. Further, our inquiry will delve into the musical attributes of the most streamed songs to discern prevalent features that resonate with listeners.

We will conduct an analysis of music features inherent in top performing songs, aiming to identify any persistent trends in music features over time. Such trends could inform our strategy for crafting songs with attributes that are currently within trend. By understanding popular elements of successful songs, we can strategically partner with artists whose creative output or discography aligns with these specific features, thereby leveraging their song production ability in the creation process. Lastly, we will identify the optimal release date that can potentially augment a song's reception and success.

Analysis

Q1. Streaming and Chart Performance

As our dataset includes information about how a song charts on different platforms, namely Spotify, Apple Music, and Deezer, we decided to delve deeper into whether or not songs placed similarly on the charts of different platforms to answer our question of whether or not there are significant discrepancies in song performance across various music streaming services. To address this question, we decided to further clean up our dataset by isolating the songs that had values greater than 0 in all three of the chart attributes: ‘in_spotify_charts’, ‘in_apple_charts’, and ‘in_deezer_charts’. This left us with 324 songs to compare. We assumed that the zeroes indicated that a song did not place on the charts of the platform, as we were not sure what a zero on the charts meant in relation to the other data. We then compared the absolute value of the difference in the values of the chart columns to determine which songs had the biggest discrepancies between where they placed on different platforms’ charts.

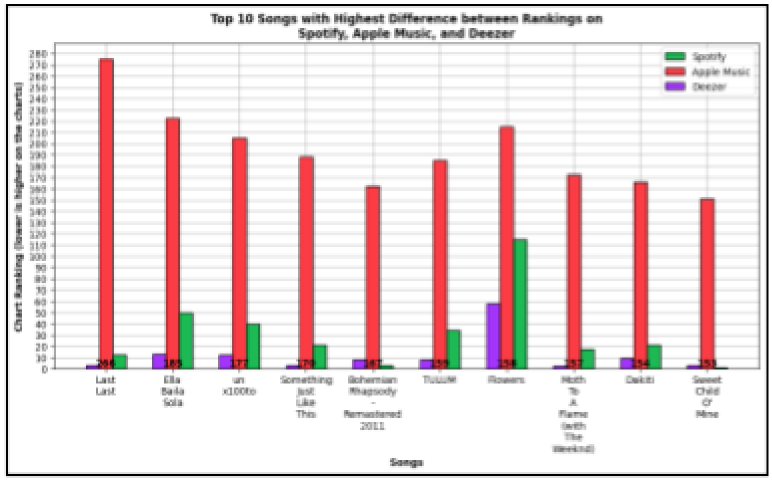

Figure 1.1 Distribution of the rankings of the top 10 songs that have the highest discrepancies between platform rankings, the bolded number is the difference between a song’s chart rank on the different platforms

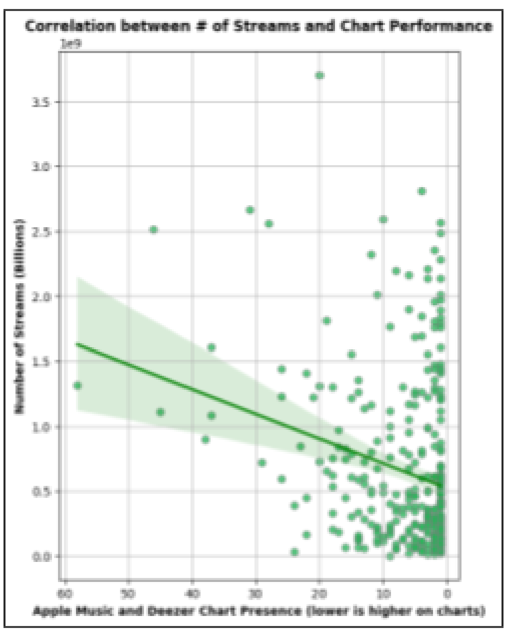

Figure 1.2 Correlation between number of streams and chart presence

To best compare the success of songs on the individual streaming platforms, we decided to look at how a song ranks on a platform. So, to answer this question, we can look at Figure 1.1 and see that even though this figure shows the songs with the highest degree of difference with how a song ranks on the charts of these platforms, there are many instances where a song ranks at the same level across all platforms. However, these large differences in performance, especially on the Apple Music charts is something worth noting. How a song ranks on Deezer and Spotify tend to be much closer in terms of number in the rank when compared to Apple Music. Songs that do relatively well on both Spotify and Deezer, meaning they place highly on the charts on these platforms, do much more poorly on Apple Music. One possible explanation for this is that Apple Music users tend to listen more to music that is not considered ‘mainstream’. When we consider how many more users Spotify and Deezer have combined, these users and platforms are more likely to define what is mainstream, than Apple Music users.

To further understand what it takes to make a number one song, we also looked into if the number of streams a song has can be an indicator of how well a song performs on the charts. To do this, we took the song’s ‘streams’ column which is the number of streams a song has on Spotify, and compared it to where it places on the charts of Apple Music or Deezer, whichever platform ranks the song higher on its charts. Looking at Figure 1.2, we see a regression model of the number of streams and the chart ranking a song has. For this regression analysis it is can be noted that the negative correlation between the number of streams a song has on Spotify and chart performance on Apple Music and Deezer suggests that popularity in terms of streams on Spotify does not necessarily translate to similar levels of popularity on Apple Music and Deezer in terms of how a song places on their charts.This is counterintuitive, since we believe that the higher number of streams a song has, the more successful it is and the more likely it is to place very highly on the charts or other platforms. For example, in Figure 1.2, there is a song that has well over 3.5 billion streams, and yet places on the charts at 20th place. Out of the 324 usable songs, this 3.5 billion streams is by far the most streamed song. Based on Figure 1.2, there is a cluster of songs with fewer than 500 million streams, and they all yet well in the top 10 of Apple Music or Deezer charts. As the number of streams accumulates over the years, after a song is released, even though a song may not be the most popular at the moment, it does indicate its popularity in terms of longevity, where people will continue to stream that song for years to come. The conclusion that we can draw from this visualization is that the number of streams a song has cannot be an indicator of its chart presence, and vice versa, the chart presence of a song cannot be an indicator of its popularity in relation to the number of streams a song has.

Our dataset is not descriptive enough in terms of how the ranking of a song is collected. Whether or not the song's rank is included in the dataset, the song's peak chart rank, or if that is where the song ranked at the time of data collection, was not mentioned in the description of this dataset, but we believe that this information would have been meaningful to our analysis.

Q2. Top Songs (Ranking - weekly, monthly streams and clicks) v. Most Streamed Songs (total stream counts)

To find out if the top songs are also the most streamed songs, we did a comparison analysis in this case. We filtered and compared the top 5 most streamed Spotify songs with the presence of top 5 songs in Spotify, Deezer, and Apple's charts and playlists which will enable us to identify whether there are commonalities between them. We decided to use bar charts to visualize and show the result here.

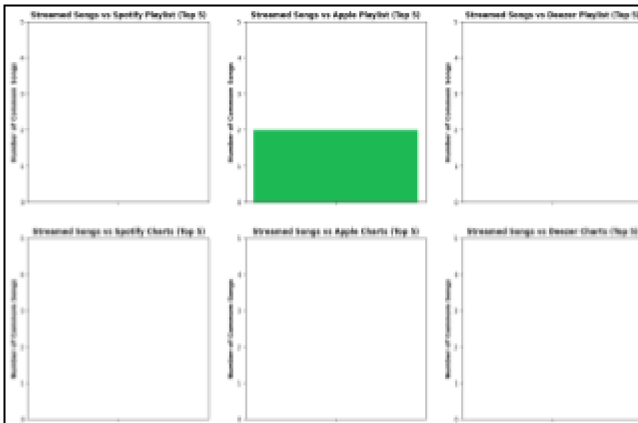

Figure 2.1 Cross-platform similarity in top 5 songs

According to Figure 2.1, we can see that within all the comparisons, only Apple playlist's top 5 songs and top 5 most streamed Spotify songs have songs in common and even with that, that's only 2 songs. This shows that top songs do not necessarily align with the top streamed song and that might be most likely caused by the difference between apps and interests.

The top 5 most streamed Spotify songs only show a shared interest of people on Spotify and considering people tend to focus on using one music app daily, it makes sense why different apps' playlists and charts do not align with Spotify's most streamed songs. In terms of why it also does not align with Spotify's Playlist and Charts, it is simply because the accounting variable is different. Most streamed songs only depend on the number of times a song is streamed at all times but for playlists and charts, it can have other factors like user favoritism, the number of streams within a certain time, etc.

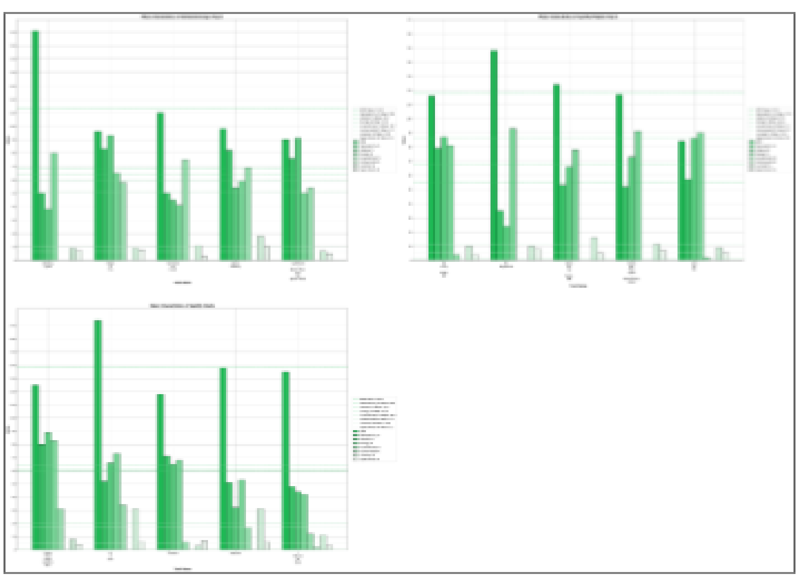

Figure 2.2 Song attributes of the top 5 songs on Spotify

Next we want to see if the most streamed songs on Spotify share similar music features as those in the top song. We began by comparing the top 5 most streamed songs in Spotify 2023 and the top 5 songs in Spotify's playlist and charts, as shown in Figure 2.2. According to the graph, the only similarity is that all those 5 songs across Spotify's streams, charts and playlists share a similar percentage in valence. The top 5 most streamed songs on Spotify has a mean of 64.2 'valence_%', the top 5 songs in Spotify playlist has a mean of 67.2 in 'valence_%' and the top 5 songs in Spotify charts has a mean of 59.2 in 'valence_%'.

Q3. Features of Top Songs

To answer the question of how music features influence song stream performance on Spotify, we looked at all the songs that have over 100 million streams (797 songs in total) to extract a "feature profile" describing what the most streamed songs look like in features. The profile analysis focuses on four main features - bpm, energy, danceability and valence. The reason why focusing on these four is that they have 100% data completeness, so no zero values. Also they are more influential to how a song is perceived as they have a bigger impact on the mood and intensity level of a song.

The feature profile value is defined as the most frequently occurred value. From the histograms of these features (figure 3.1), it's observed that BPM and danceability is clustered around a specific value, 120 for BPM, and 70 for danceability. So the cluster is taken as their profile value as they are considered as the most frequently occurred value.. On the other side, valence and energy do not have a visible cluster as their value is more spread out. So the median is taken as their profile value. To ensure the histogram reflects the accurate distribution, the bin number is set as 100 to capture the frequency distribution at each individual value of the features.

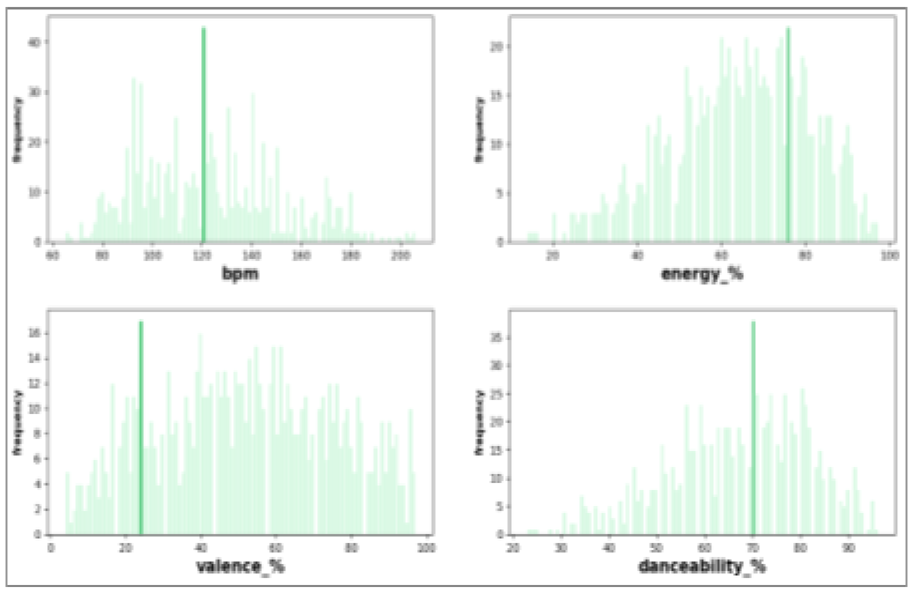

Figure 3.1 Histograms of music features (bpm, energy, valence, danceability) on over 100 million streams songs

Here is the feature profile (Table 3.1) and their interpretation. The most streamed songs on Spotify usually have the bpm around 120, relatively fast, bright beat, the music good for a light cardio; energy level at 66, music is engaging but not too loud; Valence at 52, a balanced emotion tone; Danceability at 70, great music to dance with! An example that aligns well with the profile is Counting Stars from OneRepublic. It has over two billion streams!

Most streamed songs "feature profile"		
Main features	**Value**	**Interpretation**
BPM	120	Moderately fast and bright beat
Energy	66	Lively and engaging, but not too loud or noisy
Valence	52	Balanced and emotional tone
Danceability	70	Good to dance with

Table 3.1 Most streamed songs "feature profile"

Next, we want to know how features trend over time among the 100 million + streams of songs. A time-series line chart visualizes the median of energy and valence features over time is performed to answer the question. The songs are aggregated by release year and the median of energy valence is taken from all songs released each year.

From the time series (Figure 3.2), it can be observed an upward trend in both energy and valence since 2018, suggesting happier, energetic songs gained more popularity. Dip in 1992: the deep dip in 1992 is [Creep]() from Radiohead, an utterly sad song but remains a high popularity since it was released in 1992. It has over 1.27 billion streams. Drop in valence since 1987: valence appears to be visibly higher prior to the 90s, particularly in the 50s and 60s, when Jazz music was really popular. Some of the happiest songs among the most streamed songs are Christmas Jazz music from the 50s and 60s and they continue to be popular now.

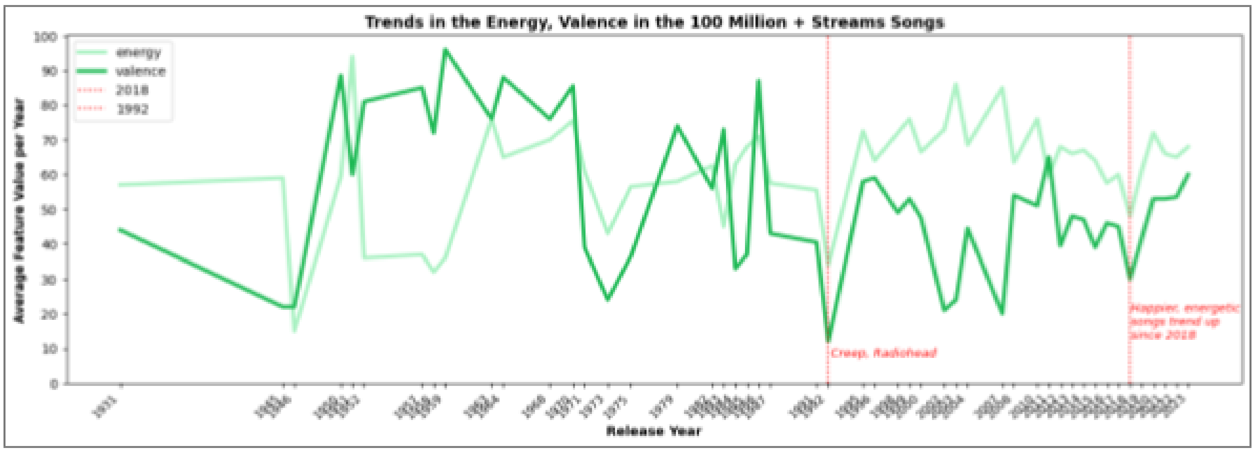

Figure 3.2. Trends in energy, valence of all over 100 million streams songs

Note: The assessed accuracy rate of release date is 76.73%. So the trend analysis can be affected by it as figure 3.2 is based on release date.

Lastly, we are curious about the trend of features in the most streamed songs of more recent years. To be able to compare features across the years, all the songs are grouped by the release year, and then within each year, songs are ranked based on the streams to identify the top 10 most streamed songs for each year. Different from the previous analysis looking at top songs of all year, we now look at top songs within each year. From 2019-2023, we picked the annual top 10 streamed songs and plotted box charts for four main features (bpm, valence, danceability, and energy) to visualize the distribution of the features each year (Figure 3.3).

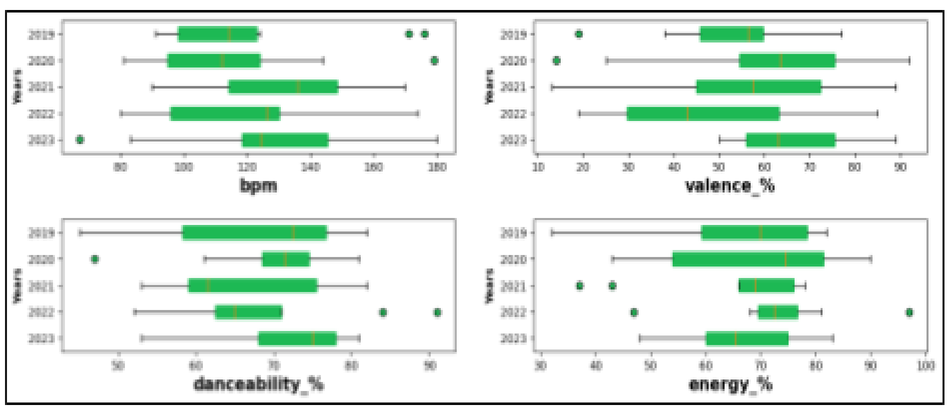

Figure 3.3: Distribution of song attributes through the years

Following the median on the box-and-whisker on Figure 3.3, we observed happier songs are more popular in 2023, 2021: A relatively higher median of valence is observed in 2021, 2023, indicating a happier level of the top 10 compared with the three other years. High energy songs are more popular in 2020, 2022: Electronic/dance, dance pop songs made to the top 10 of 2020, 2022, contributing a higher energy feature compared to the other three years. [I'm Good (Blue)]() from Bebe Rexha & David Guetta, an electronic dance song, is the No.9 of 2022, having an energy level as high as 97. [Bad Habits]() from Ed Sheeran, a dance pop song, is the No.8 of 2020, having an energy level of 92. High bpm level in 2021: The top 10 in 2021 has a higher bpm median compared to other years, suggesting high beats songs gained great popularity within the year. Stay (with Justin Bieber) from Justin Bieber & The Kid Laroi is the No.1 in 2021 and has a bpm of 170. High danceability level in 2023: The median of danceability in 2023 is the highest among the time period from 2019 to 2023. Songs good for dancing achieved massive success in 2023. Five of the top 10 songs in 2023 are Latino music (TQG, Shakira: Bzrp Music Sessions, Vol. 53 from Shakira, La Bebe (Remix), PRC from Peso Pluma, and Ella Baila Sola from Eslabon Armado). All have great danceability.

Q4. Artist Collaboration:
From Figure 4.1 we plotted all streams from Spotify to reveal that the majority of streams are concentrated around songs with a single artist, as indicated by the dense clustering of data points at artist count of one. We also see a declining trend where as the artist count increases, there seems to be a general decline in the number of streams which could imply that collaboration between multiple artists do not perform as well in terms of streaming numbers compared to solo artists, on average. Overall artist count that's above one, the maximum stream is lower than for solo artists, and the number of high streaming outliers also decreases. This initial observation suggests that solo artist songs tend to perform better than ones with features or collaboration.

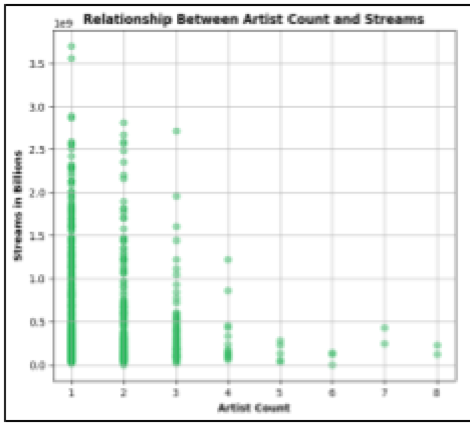

Figure 4.1 Relationship between artist count and streams

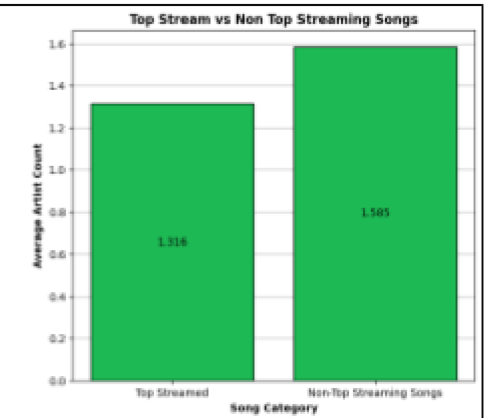

Figure 4.2 Average number of artists for top / non top songs

Furthermore the assumption can be supported by taking the average number of artists for songs that are in the top 10 percentile to distinguish as “Top Streaming Songs” in comparison to Non Top Streaming Songs. Figure 4.2 displays that both categories, top streams and other songs have similar average artist counts with non-top streaming songs being slightly higher. This indicates that collaboration involving multiple artists may not significantly influence the total number of streams a song receives. We also wanted to look at the impact if artists would contribute to their performance in charts. The reason is that with multiple artists on a song, it could lead to more promotion on the song leading to an increase in charts from other music streaming platforms outside of spotify such as Apple and Deezer.

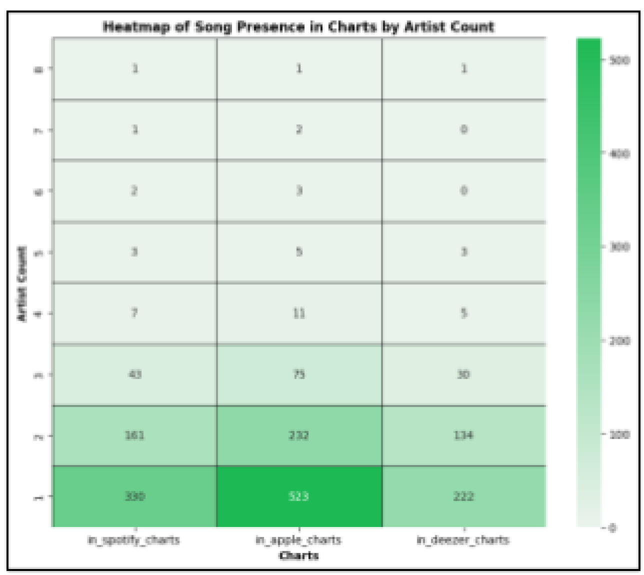

Figure 4.3: Heat map of song presence in charts by artist count

Looking at Figure 4.3, we see that songs with single artists have the highest presence in charts across all three platforms with apple music charts having the highest numbers. As the number of artists decrease, the number of charted songs starts to increase as observed in the increase in intensity of light green to dark green as you move down the rows. With four or more artists, the chart presence significantly diminishes having very few or no songs charted for Deezer charts. This concludes that collaboration between artists have no strong presence or influence in music streaming platform charts. Given our analysis we can conclude that collaboration on music does not affect its popularity so we will continue to look at other characteristics that would affect song popularity.

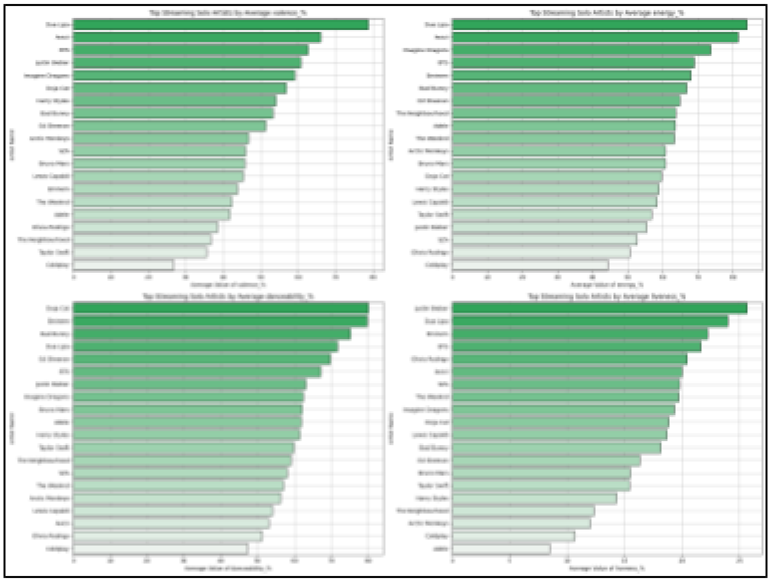

Figure 4.4 Top streaming artist by average song features (features - top left: valence, top right: energy, bottom left: danceability, bottom right: liveliness)

The analysis indicates a trend where solo artists generally achieve higher streaming numbers, and the presence of featured artists appears to have a negligible influence on streaming success and chart performance. Figure 4.4 focuses on solo artists which reveals certain musical attributes prevalent among the most streamed artists. The visualization stratifies top - streaming artists according to the average attributes of valence, energy, danceability, and liveliness in their music repertoire.

For instance, Dua Lipa, predominantly known for her pop genre contributions, exhibits the highest average scores in valence and energy, suggesting her music often carries a positive and energetic tone. Doja Cat, whose music style spans hip hop, pop R&B, and pop rap, typically produces tracks that scores higher on danceability, reflecting on rhythmic and engaging nature suitable for dancing. In terms of live performance, Justin Bieber ranks the highest, indicating a strong presence of live performance elements. These insights allow the opportunity to collaborate strategically with artists, aligning production efforts with their distinctive musical traits to craft songs that capitalize on desired attributes.

Q5. Release Trend in Top Songs:

Looking at release trends, we look at if there are any patterns with top songs' age based on year and month of release. To begin we reorganize and list out the top songs of each year then calculate the age of songs based on the year 2023 with Figure 5.2 displaying our result. According to the line of regression in Figure 5.1, we can observe that the younger the top song is, the more streams it is likely to have. This is justifiable because the advancement of technology makes it easier for people to access songs. With those top songs, it naturally accumulates more streams.

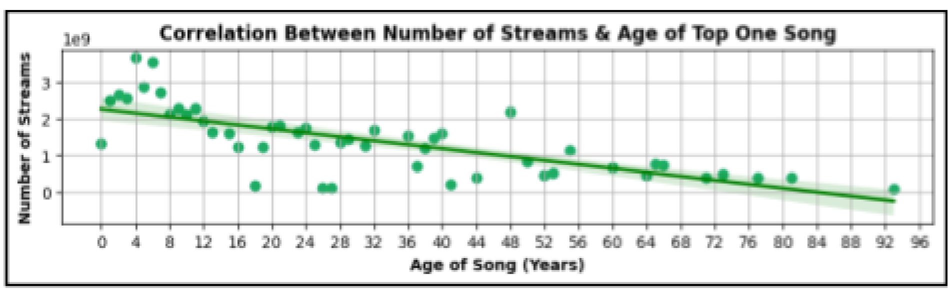

Figure 5.1 Correlation between number of streams and age of top song

It is also interesting to see those centralized outliers in songs around 4 - 6 years old because those are the COVID-19 years. During the pandemic, people were forced to work at home which gave people more time and chances to listen to music. Spotify as one of the most popular music apps became the top choice during the pandemic which is why the top songs of those years are getting more streams compared to those that are 1-3 years old. This can also be illustrated by the top 10 number 1 streamed songs as shown below in Figure 5.2.

From Figure 5.2 we can see that within the top 10 of the #1 streamed songs by year, the top 7 of them are from exactly the year 2016 - 2022 which includes the years where digital transformation is booming and the years of the pandemic. In terms of which months are #1 streamed songs by years more likely to be released, we compared the months of those songs. In Figure 5.3 the chart clearly shows that #1 streamed songs are commonly released during the middle of the year, from April to August. This can be most likely caused by how those music festivals are hosted during the middle of the year which releasing songs during those months can have more free marketing and traffic.

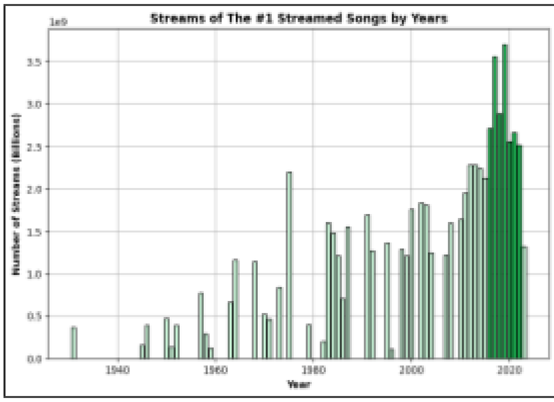

Figure 5.2: Stream of top song by year

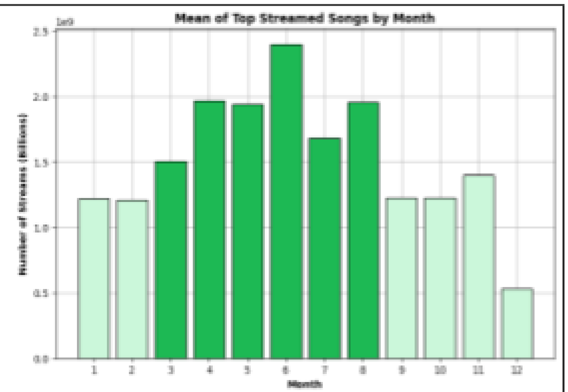

Figure 5.3 Mean of top streamed song by month

Conclusion:

We analyzed these factors on the streaming performance of a song on Spotify from various perspectives:

- Charts performance on Spotify, Apple Music and Deezer
- Playlist presence in Spotify, Apple Music and Deezer
- Music features including: bpm, energy, valence, danceability, acousticness, liveness and speechness
- Artists collaboration vs solo artist
- Age and Release month

Based on our analysis, we can conclude that songs released in the most recent years, after 2019, are likely to achieve a higher stream number compared to songs released prior to 2019 as songs released from 2019 - 2023 account for 76.9% of all songs with over 100 million streams on Spotify. Songs released in the middle of the year (April to August) are likely to gain a higher streaming performance as songs released in these months have a higher average streams compared to songs released in other months. Also songs in alignment with the feature profile has the potential to achieve higher streaming popularity as the profile encapsulates the most common types of features observed in all time top streaming songs. In our analysis, no visible influence on streaming performance has been identified in the relationships with a song's chart performance, playlist presence and artists collaboration.

Our analysis has a few limitations, which can impact the accuracy of our conclusion:

- 76.73% release dates are accurate
- 949 unique entries
- Questions as to how data was collected
- Questions about the Spotify, Apple Music and Deezer placement chart

Appendix

Chart 1.1: Attributes, description, data types and sample values for the variables used in our analysis.

Attributes Used in Analysis	Description	Sample Value	Data Type
track_name	Name of the song	Save Your Tears	object
artist(s)_name	Name of the artist(s) of the song	The Weeknd	object
artist_count	Number of artists contributing to the song	1	int64
released_year	Year when the song was released	2020	int64
released_month	Month when the song was released	3	int64
in_spotify_charts	Presence and rank of the song on Spotify charts	13	int64
streams	Total number of streams on Spotify	1591223784	object
in_apple_charts	Presence and rank of the song on Apple Music charts	115	float64
in_shazam_charts	Presence and rank of the song on Shazam charts	200	int64
bpm	Beats per minute, a measure of song tempo	118	int64
key	Key of the song	C	object
mode	Mode of the song (major or minor)	Major	object
danceability_%	Percentage indicating how suitable the song is for dancing	68	int64
valence_%	Measure of happiness conveyed by a track, high valence (cheerful), low valence (sad, angry)	61	int64
energy_%	Measure of energy level. High energy tracks sound loud, and noisy	82	int64
acousticness_%	Amount of acoustic sound in the song	2	int64
liveness_%	Presence of live performance elements	50	int64
speechiness_%	Amount of spoken words in the song	3	int64
in_deezer_playlists	Number of Deezer playlists the song is included in	10	series
in_apple_playlists	Number of Apple playlists the song is included in	43	int64
in_spotify_playlists	Number of Spotify playlists the song is included in	553	int64

Data set link: [Spotify_2023]