Dataset URL: https://www.kaggle.com/datasets/nelgiriyewithana/top-spotify-songs-2023/

Related Topic Web Link: https://www.cnet.com/tech/services-and-software/best-music-streaming-service/ **(Used in Visualization 2)**

**Data Description**

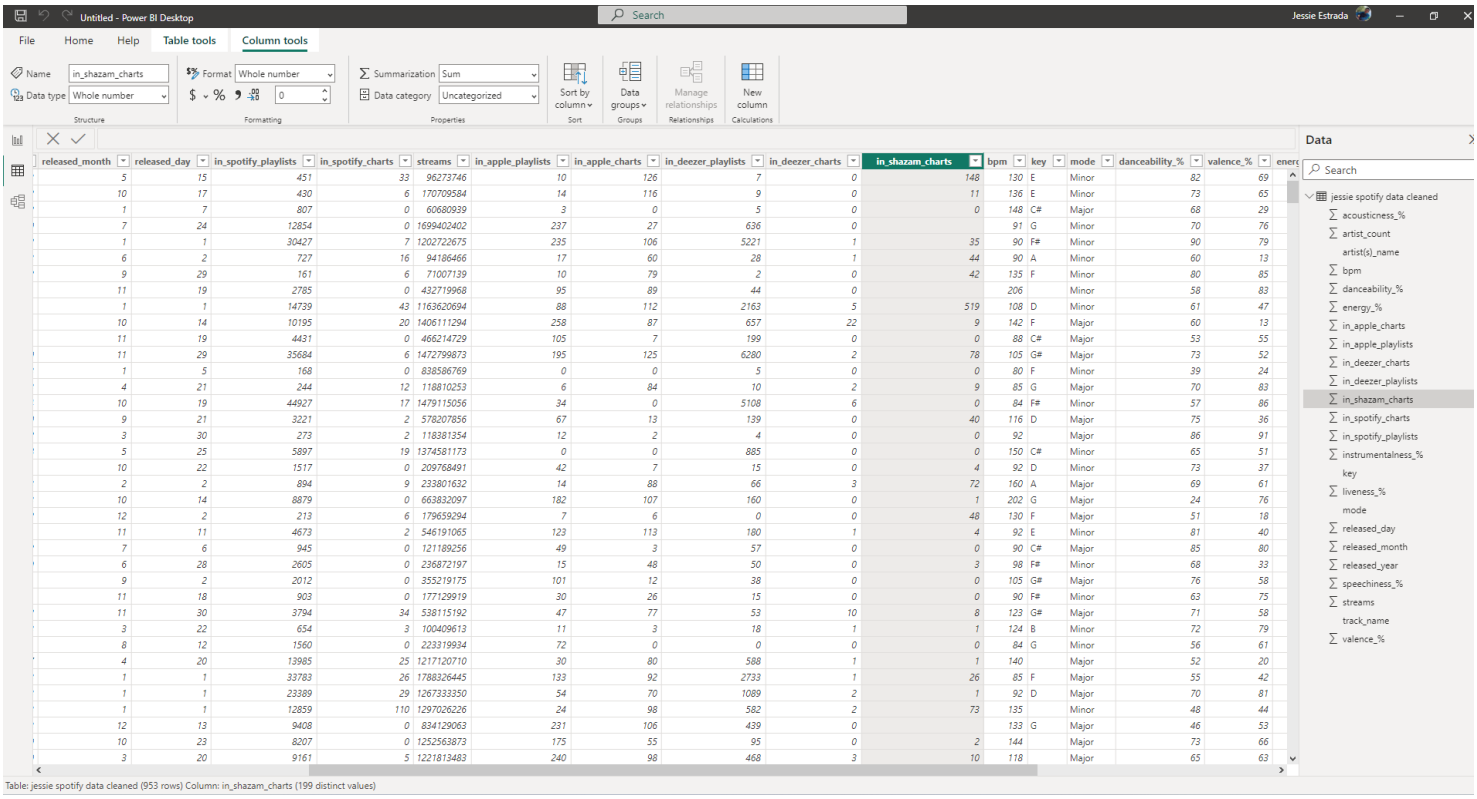
| Data Field | Description of Field | Example Value |
|---|---|---|
| track_name | Name of the song | "Viva La Vida" |
| artist(s)_name | Name of the artist(s) | "Taylor Swift" |
| artist_count | Number of artists contributing to the song | "2" |
| released_year | Year when the song was released | "2023" |
| released_month | Month when the song was released | "7" |
| released_day | Day of the month when the song was released | "14" |
| in_spotify_playlists | Number of Spotify playlist the song is included in | "12211" |
| in_spotify_charts | Presence and rank of the song on Spotify charts | "40" |
| streams | Total number of streams on Spotify | "2,513,188,493" |
| in_apple_playlists | Number of Apple Music playlists the song is included in | "48" |
| in_apple_charts | Presence and rank of the song on Apple Music charts | "222" |
| in_deezer_playlists | Number of Deezer, the song is included in | "45" |
| in_deezer_charts | Presence and rank of the song on Deezer charts | "5" |
| in_shazam_charts | Presence and rank of the charts | "1,021" |
| bpm | Beats per minute, a measure of tempo | "125" |
| key | Key of the song | "A" |
| mode | Mode of the song (major or minor) | "Major" |
| danceability_% | Percentage, indicating how suitable the song is for dancing | "80" |

| | | |
|---|---|---|
| valence_% | Positivity of the musical content | “83” |
| energy_% | Perceived, energy level of the | “82” |
| acousticness_% | Amount of acoustic sound in the song | “7” |
| instrumentalness_% | Amount of instrumental content in the song | “63” |
| liveness_% | Presence of live performance elements | “10” |
| speechiness_% | Amount of spoken words in the song | “4” |

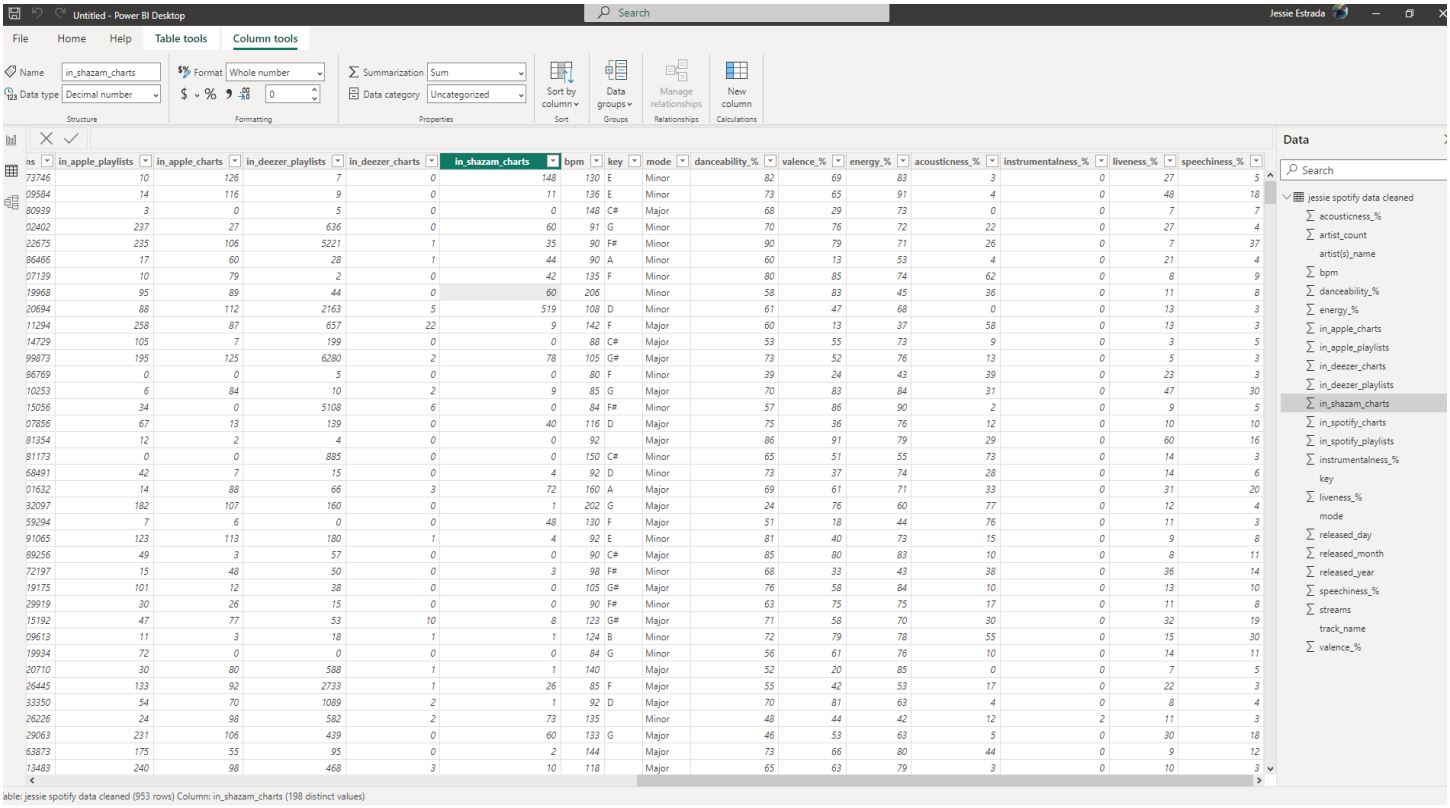
## Data Modeling/Cleaning

1) **Data Cleaning Category Name:** Missing Values (Replace with Mean)

**Pre-Cleaning Screenshot:**



**Post-Cleaning Screenshot:**

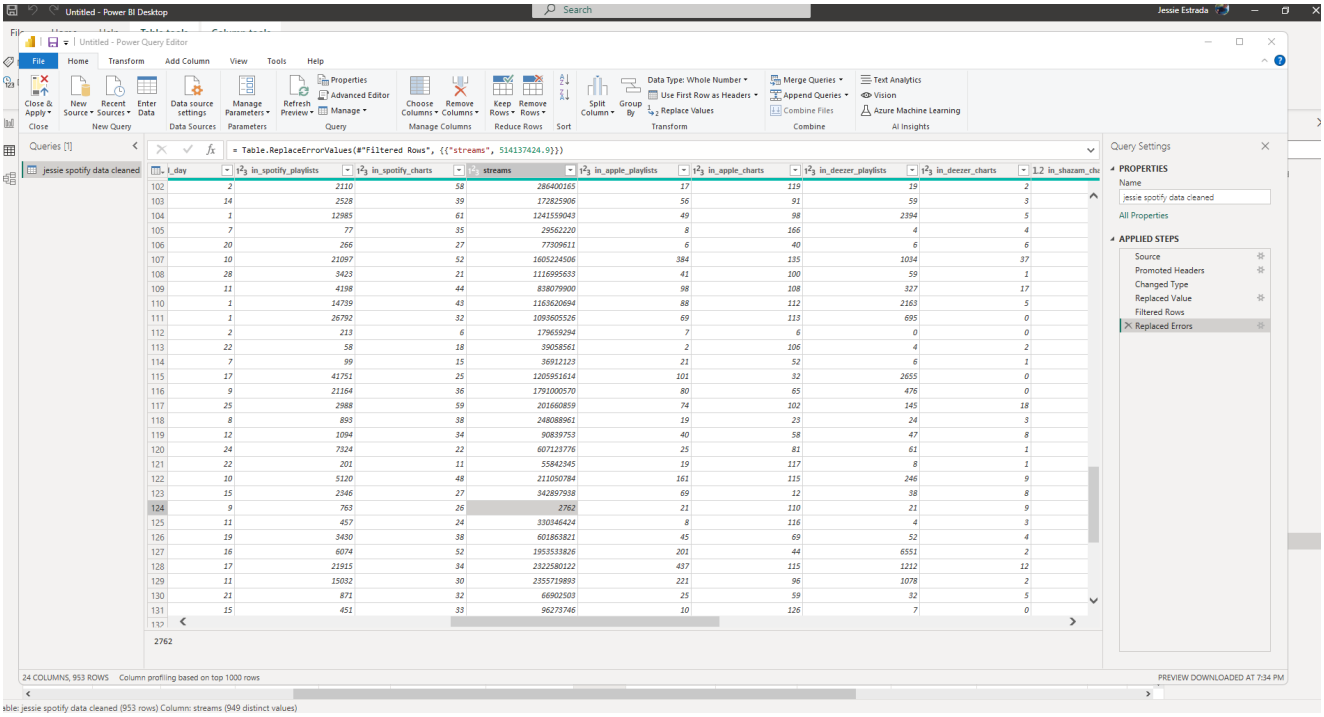



**Explanation of what was done:**

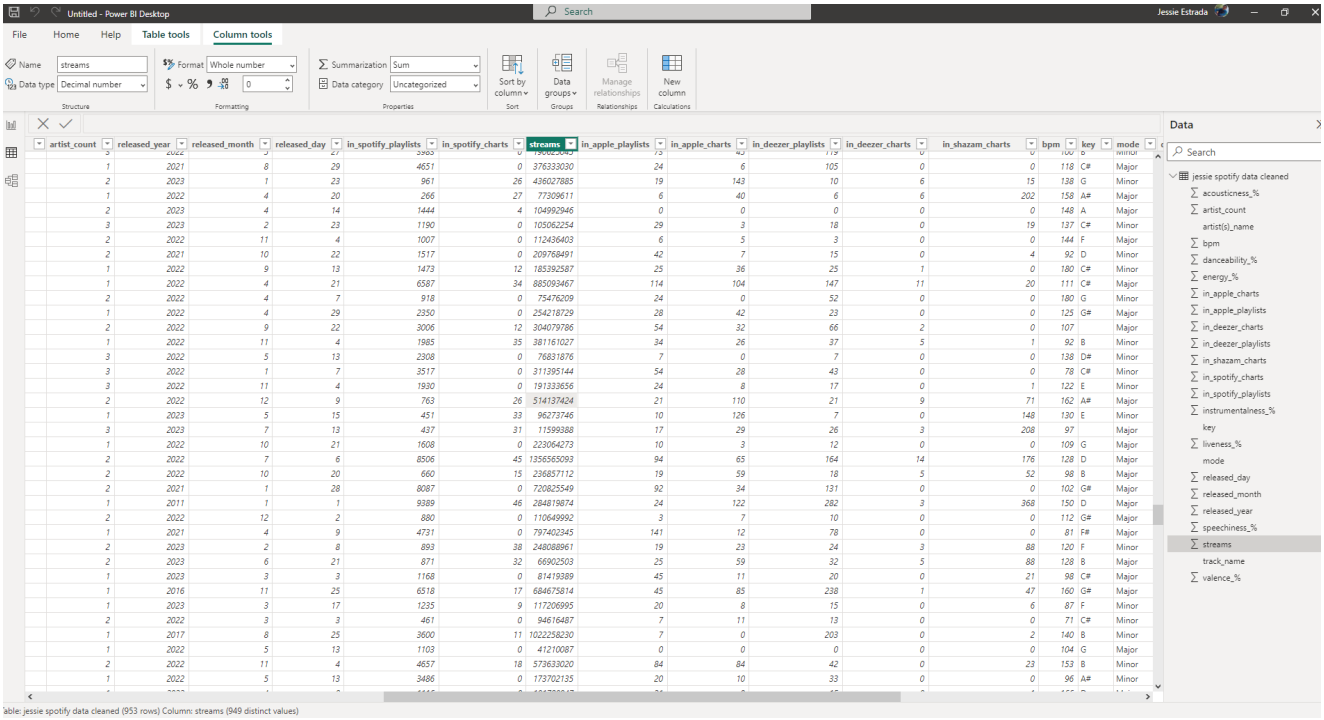
I had to eliminate the blanks from the column for this process. This “in_shazam_charts” column can significantly impact my results because Shazam is an important application where much music is discovered, and this column can be used to see the number of people that listen to these songs and find out what the track name and artist name is. As such, to implement the Missing values data cleaning concept, I had to analyze the column and calculate the mean of the values. Then, I took that mean and replaced the missing values with that calculated mean.

2) **Data Category Name:** Data Outlier (Replace with Mean)

## Pre-Cleaning Screenshot:



## Post-Cleaning Screenshot:

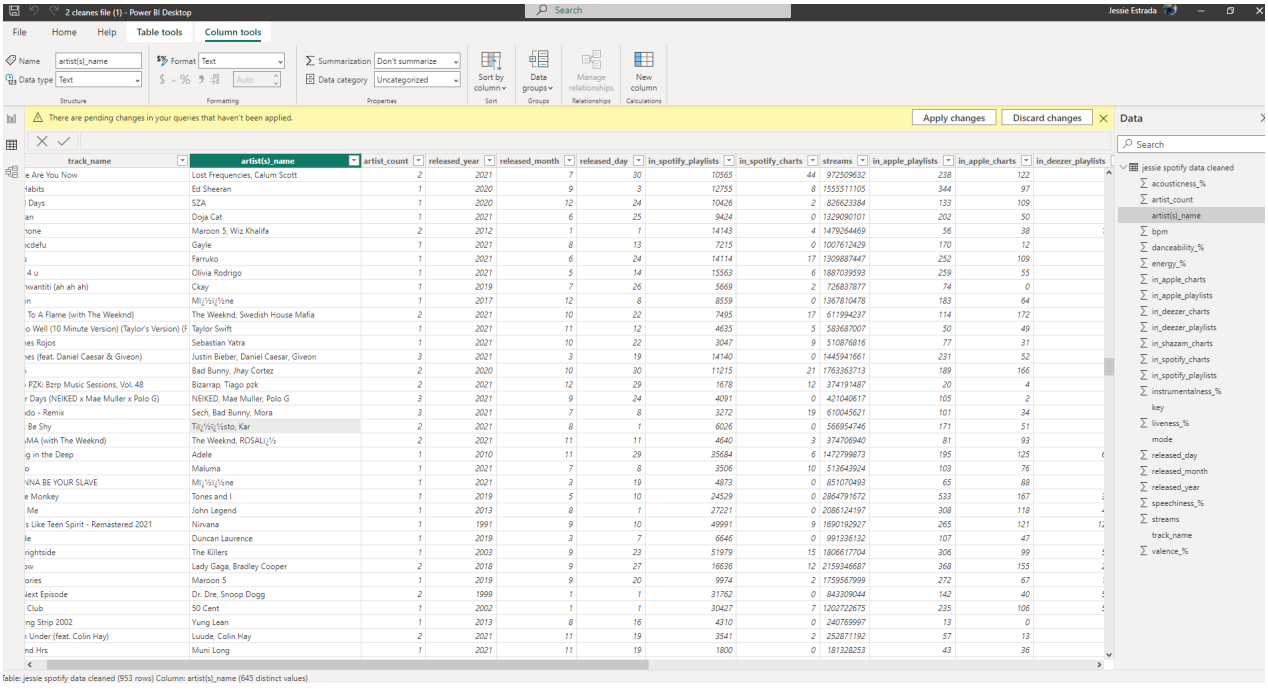


**Explanation of what was done:**

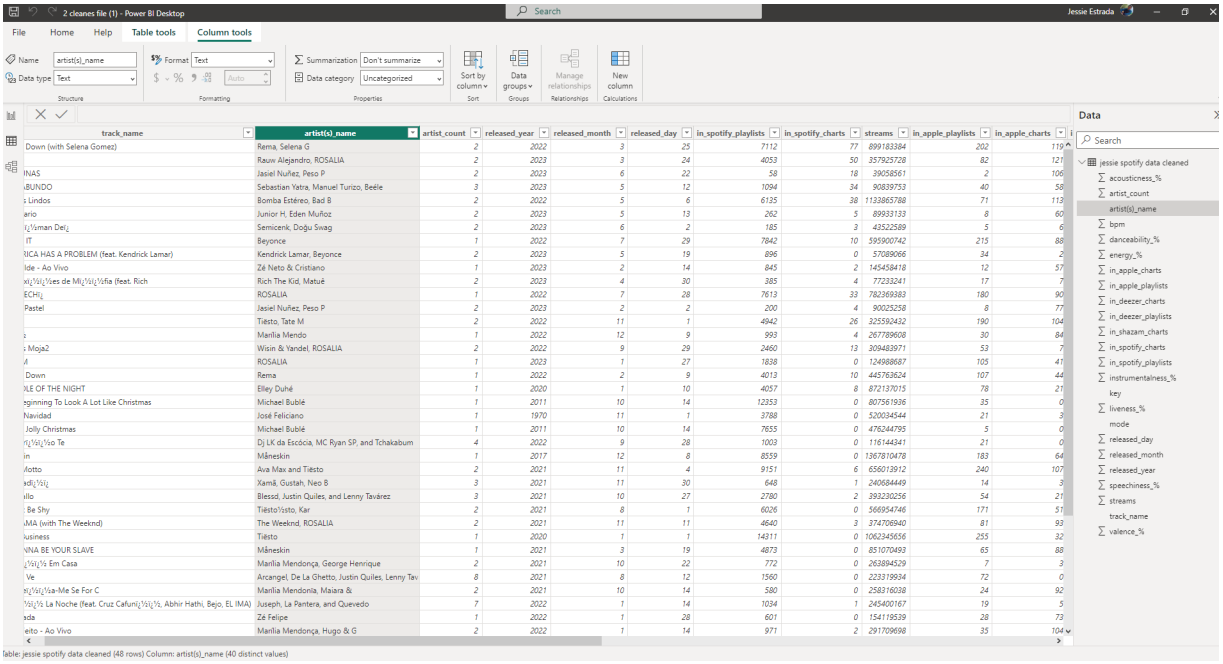
For this process, I had to ensure that all entries for this column made sense and were positively impactful to the analysis. After reviewing the "streams" column, it was evident that it contained data outliers. This column is highly impactful because this column will be used to calculate which songs/artists were the most popular throughout 2023. As such, I needed to ensure that the column was cleaned and prepared for my analysis. In other words, it had entries that did not match up with all other entries in the column. As such, it was necessary to calculate the mean of this column and replace the data outliers with this value. Through this process, we ensure that all entries in this column are meaningful and correlate with one another.

3) **Data Category Name:** Data Typos (Replace Typos)

**Pre-Cleaning Screenshot:**



**Post-Cleaning Screenshot:**




| track_name | artist(s)_name | artist_count | released_year | released_month | released_day | in_spotify_playlists | in_spotify_charts | streams | in_apple_playlists | in_apple_charts |
|---|---|---|---|---|---|---|---|---|---|---|
| Down (with Selena Gomez) | Rema, Selena G | 2 | 2022 | 3 | 25 | 7112 | 77 | 899183384 | 202 | 119 |
| | Rauw Alejandro, ROSALÍA | 2 | 2023 | 3 | 24 | 4053 | 50 | 357925728 | 82 | 121 |
| INAS | Jasiel Nuñez, Peso P | 2 | 2023 | 6 | 22 | 58 | 18 | 39058561 | 2 | 106 |
| iBUNDO | Sebastian Yatra, Manuel Turizo, Beéle | 3 | 2023 | 5 | 12 | 1094 | 34 | 90839753 | 40 | 58 |
| s Lindos | Bomba Estéreo, Bad B | 2 | 2022 | 5 | 6 | 6135 | 38 | 1133865788 | 71 | 113 |
| ario | Junior H, Eden Muñoz | 2 | 2023 | 5 | 13 | 263 | 5 | 89933133 | 8 | 60 |
| ï¿½man Delï¿ | Semicenk, Doğu Swag | 2 | 2023 | 6 | 2 | 185 | 3 | 43522589 | 5 | 6 |
| IT | Beyonce | 1 | 2022 | 7 | 29 | 7842 | 10 | 595900742 | 215 | 88 |
| ICA HAS A PROBLEM (feat. Kendrick Lamar) | Kendrick Lamar, Beyonce | 2 | 2023 | 5 | 19 | 896 | 0 | 57089066 | 34 | 2 |
| lde - Ao Vivo | Zé Neto & Cristiano | 1 | 2023 | 2 | 14 | 845 | 2 | 145458418 | 12 | 57 |
| xï¿½ï¿½es de Mï¿½ï¿½fia (feat. Rich | Rich The Kid, Matuê | 2 | 2023 | 4 | 30 | 385 | 4 | 77233241 | 17 | 7 |
| ECHï¿ | ROSALÍA | 1 | 2022 | 7 | 28 | 7613 | 33 | 782369383 | 180 | 90 |
| Pastel | Jasiel Nuñez, Peso P | 2 | 2023 | 2 | 2 | 200 | 4 | 90025258 | 8 | 77 |
| | Tiësto, Tate M | 2 | 2022 | 11 | 1 | 4942 | 26 | 325592432 | 190 | 104 |
| | Marília Mendo | 1 | 2022 | 12 | 9 | 993 | 4 | 267789608 | 30 | 84 |
| Mojaï2 | Wisin & Yandel, ROSALÍA | 2 | 2022 | 9 | 29 | 2460 | 13 | 309483971 | 53 | 7 |
| /I | ROSALÍA | 1 | 2023 | 1 | 27 | 1838 | 0 | 124988687 | 105 | 41 |
| Down | Rema | 1 | 2022 | 2 | 9 | 4013 | 10 | 445763624 | 107 | 44 |
| XLE OF THE NIGHT | Elley Duhé | 1 | 2020 | 1 | 10 | 4057 | 8 | 872137015 | 78 | 21 |
| ginning To Look A Lot Like Christmas | Michael Bublé | 1 | 2011 | 10 | 14 | 12353 | 0 | 807561936 | 35 | 0 |
| Navidad | José Feliciano | 1 | 1970 | 11 | 1 | 3788 | 0 | 520034544 | 21 | 3 |
| Jolly Christmas | Michael Bublé | 1 | 2011 | 10 | 14 | 7655 | 0 | 476244795 | 5 | 0 |
| ï¿½ï¿½so Te | Dj LK da Escócia, MC Ryan SP, and Tchakabum | 4 | 2022 | 9 | 28 | 1003 | 0 | 116144341 | 21 | 0 |
| in | Måneskin | 1 | 2017 | 12 | 8 | 8559 | 0 | 1367810478 | 183 | 64 |
| /lotto | Ava Max and Tiësto | 2 | 2021 | 11 | 4 | 9151 | 6 | 656013912 | 240 | 107 |
| ïdï¿½ï¿ | Xamã, Gustah, Neo B | 3 | 2021 | 11 | 30 | 648 | 1 | 240684449 | 14 | 3 |
| illo | Blessd, Justin Quiles, and Lenny Tavárez | 3 | 2021 | 10 | 27 | 2780 | 2 | 393230256 | 54 | 21 |
| Be Shy | Tiësto?ssto, Kar | 2 | 2021 | 8 | 1 | 6026 | 0 | 566954746 | 171 | 51 |
| MA (with The Weeknd) | The Weeknd, ROSALÍA | 2 | 2021 | 11 | 11 | 4640 | 3 | 374706940 | 81 | 93 |
| lusiness | Tiësto | 1 | 2020 | 1 | 1 | 14311 | 0 | 1062345656 | 255 | 32 |
| NNA BE YOUR SLAVE | Måneskin | 1 | 2021 | 3 | 19 | 4873 | 0 | 851070493 | 65 | 88 |
| ï¿½ï¿½ Em Casa | Marília Mendonça, George Henrique | 2 | 2021 | 10 | 22 | 772 | 0 | 263894529 | 7 | 3 |
| Ve | Arcangel, De La Ghetto, Justin Quiles, Lenny Tav | 8 | 2021 | 8 | 12 | 1560 | 0 | 223319934 | 72 | 0 |
| ï¿½ï¿½a-Me Se For C | Marília Mendonla, Maiara & | 2 | 2021 | 10 | 14 | 580 | 0 | 258316038 | 24 | 92 |
| ï¿½ï¿½ La Noche (feat. Cruz Cafuneï¿½ï¿½, Abhir Hathi, Bejo, EL IMA) | Juseph, La Pantera, and Quevedo | 7 | 2022 | 1 | 14 | 1034 | 1 | 245400167 | 19 | 5 |
| ada | Zé Felipe | 1 | 2022 | 1 | 28 | 601 | 0 | 154119539 | 28 | 73 |
| eito - Ao Vivo | Marília Mendonça, Hugo & G | 2 | 2022 | 1 | 14 | 971 | 2 | 291709698 | 35 | 104 |

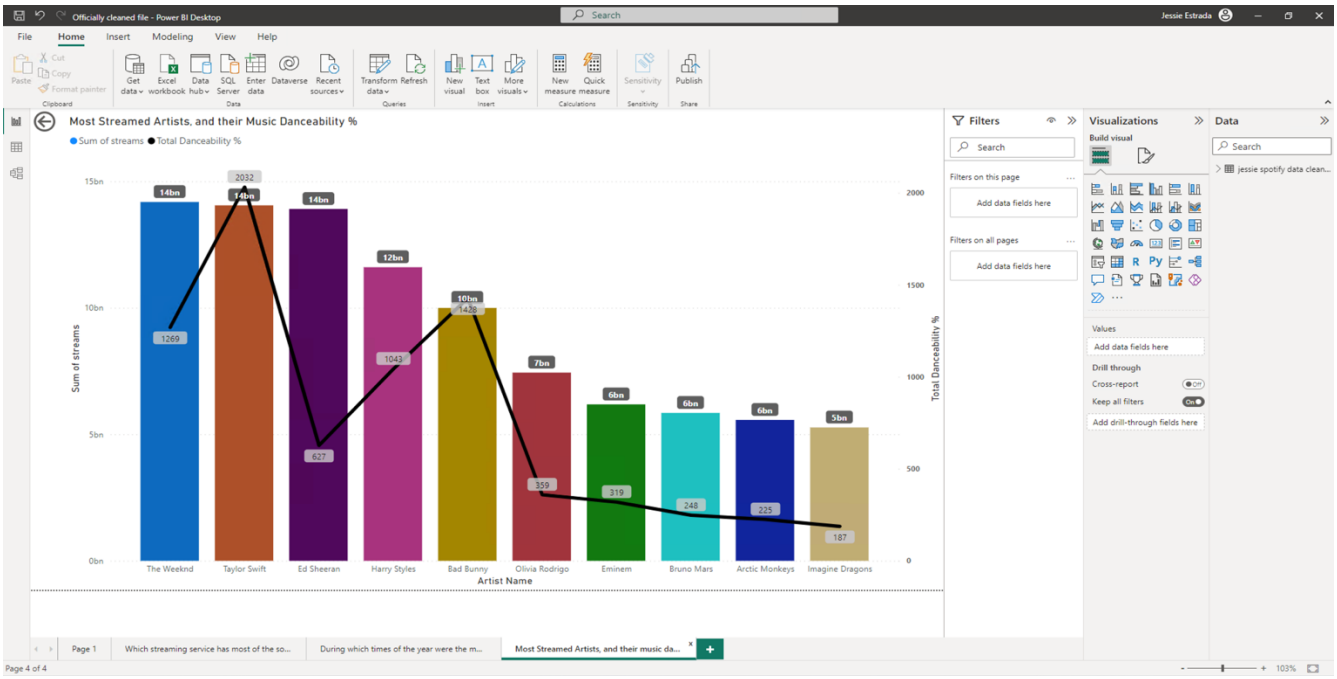


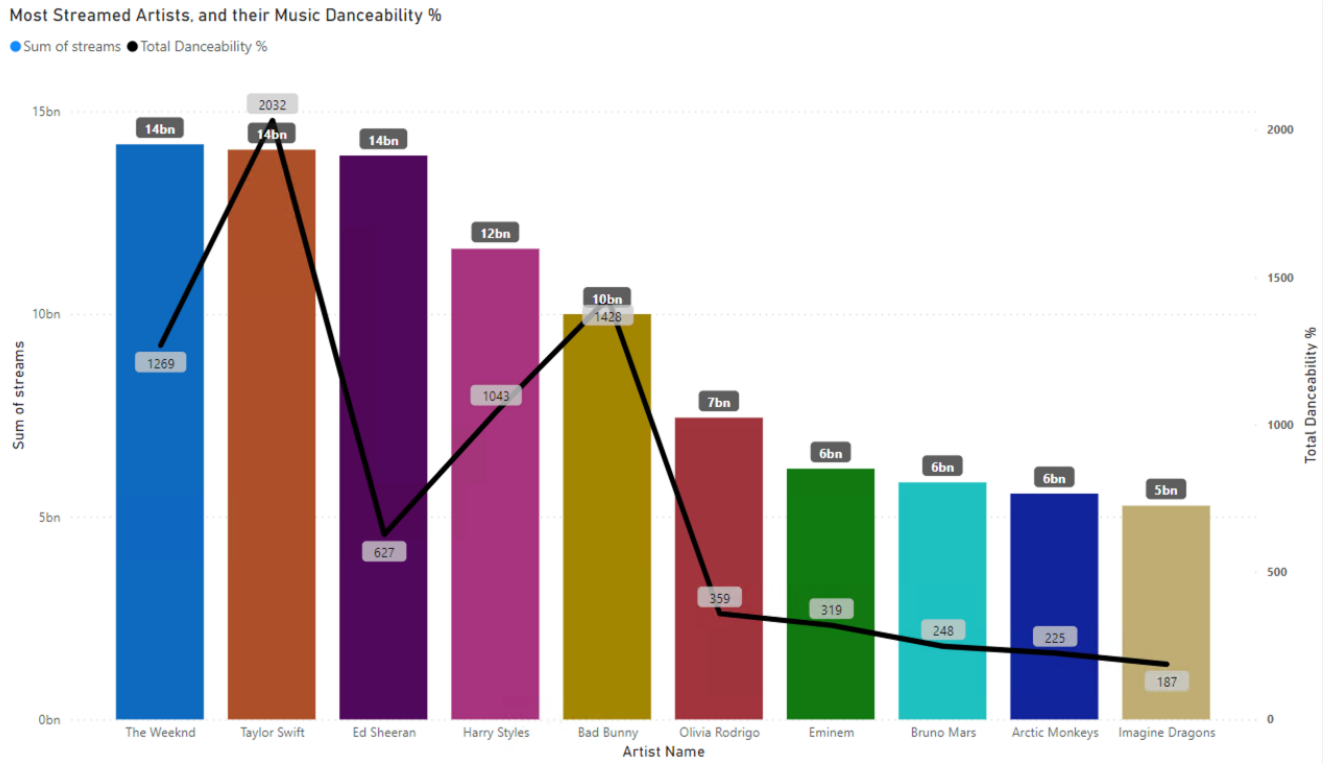
**Explanation of what was done:**

For this data-cleaning concept, I had to analyze the column of “artist(s)_name” and ensure that all artist names were in the correct format. After a review of this column, it was evident that multiple entries contained typos. These typos all occurred with artist names that contained items such as a “tilde” or any special accents. Due to this, the transferring of data into Power BI resulted in these data typos. So, to ensure our data did not contain any typos, it was necessary to locate these entries and replace them with the correct special characters. As a result, our column did not contain any typos, and analysis will now be better due to our more profound understanding of the column since it does not contain any typos.

# Visualizations

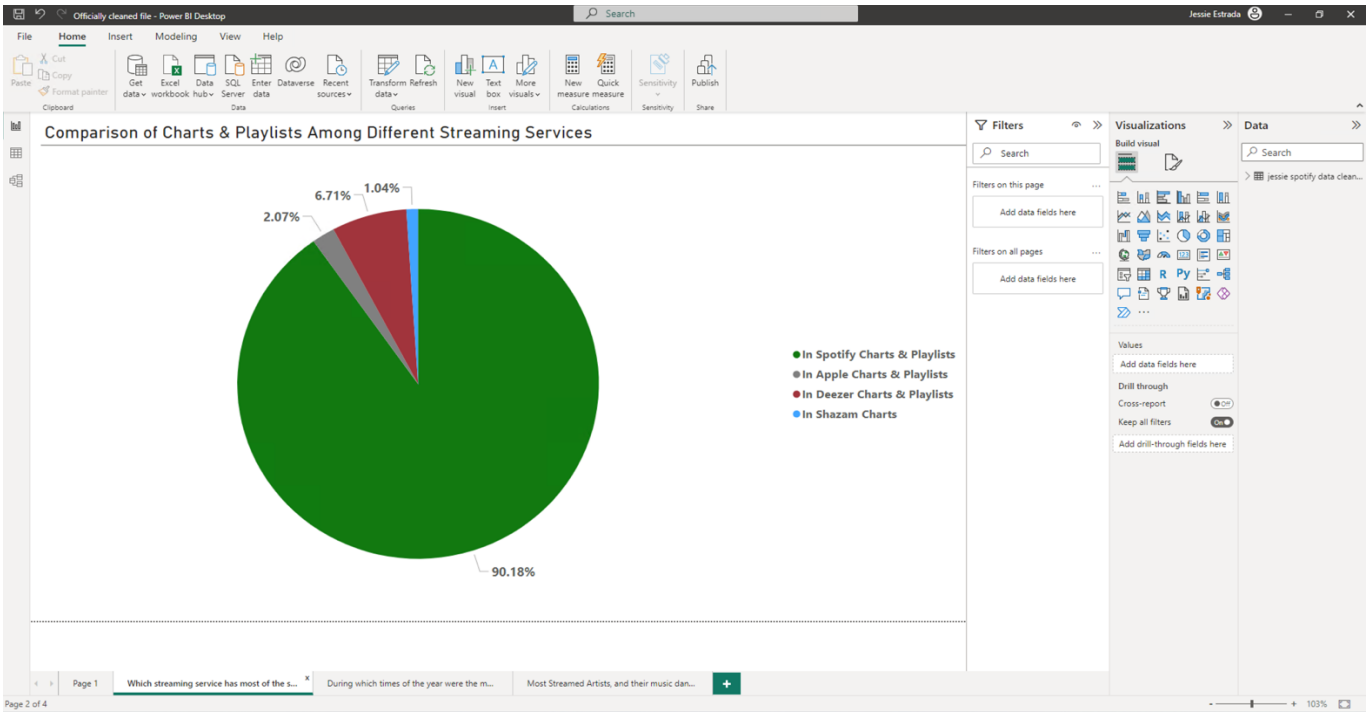
## 1) Which artist has the most appearances in the most streamed Spotify songs of 2023?

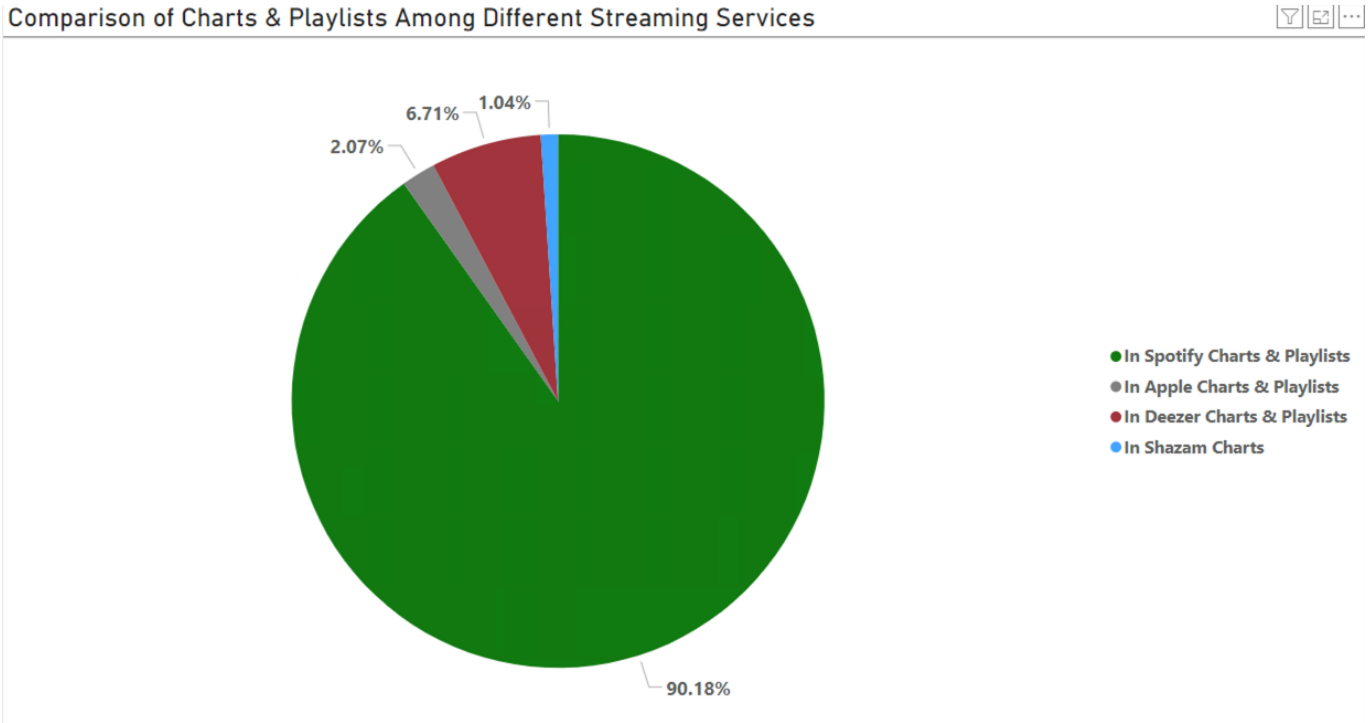


Through this analysis and the creation of this visualization, I came to the conclusion that The Weeknd was the most streamed artist of 2023, with a total of about 14 billion streams! In the visualization, we can see the top 10 most streamed artists, and I wanted to see if there was any common factor in song popularity. In the visualization, there is a y-axis line displaying the total danceability percentage of each artist's song. It is clear that the more streamed artists have a higher percentage of danceability in their music. In the article "The Evolution of Dance Music," Laura Gao states, "Some attributes that make a song more dance-worthy are timeless. Generally, the happier and more energetic a song is, the more likely it will be a dance hit." This correlates with our results because it is clear that artists who tend to add more danceability to their music will, in turn, have more streams on their tracks.

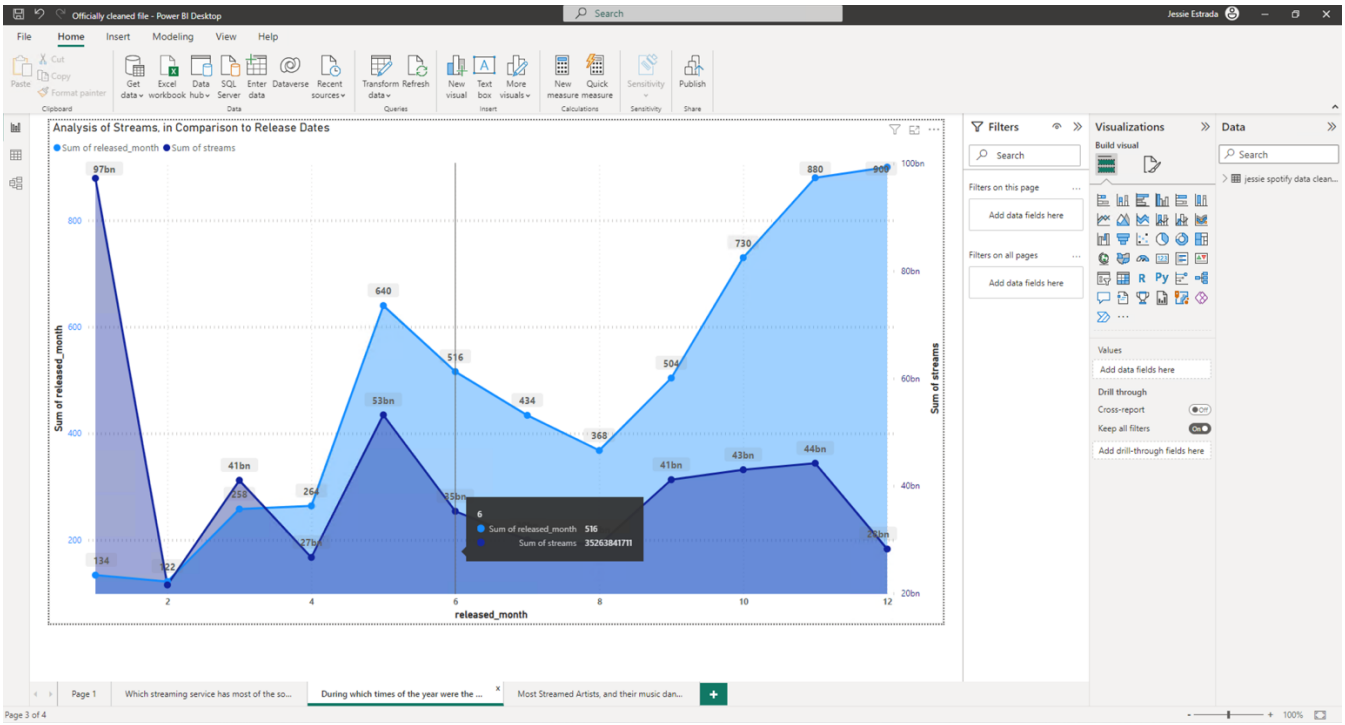
**2) Which streaming service has most of the songs in their charts and playlists?**

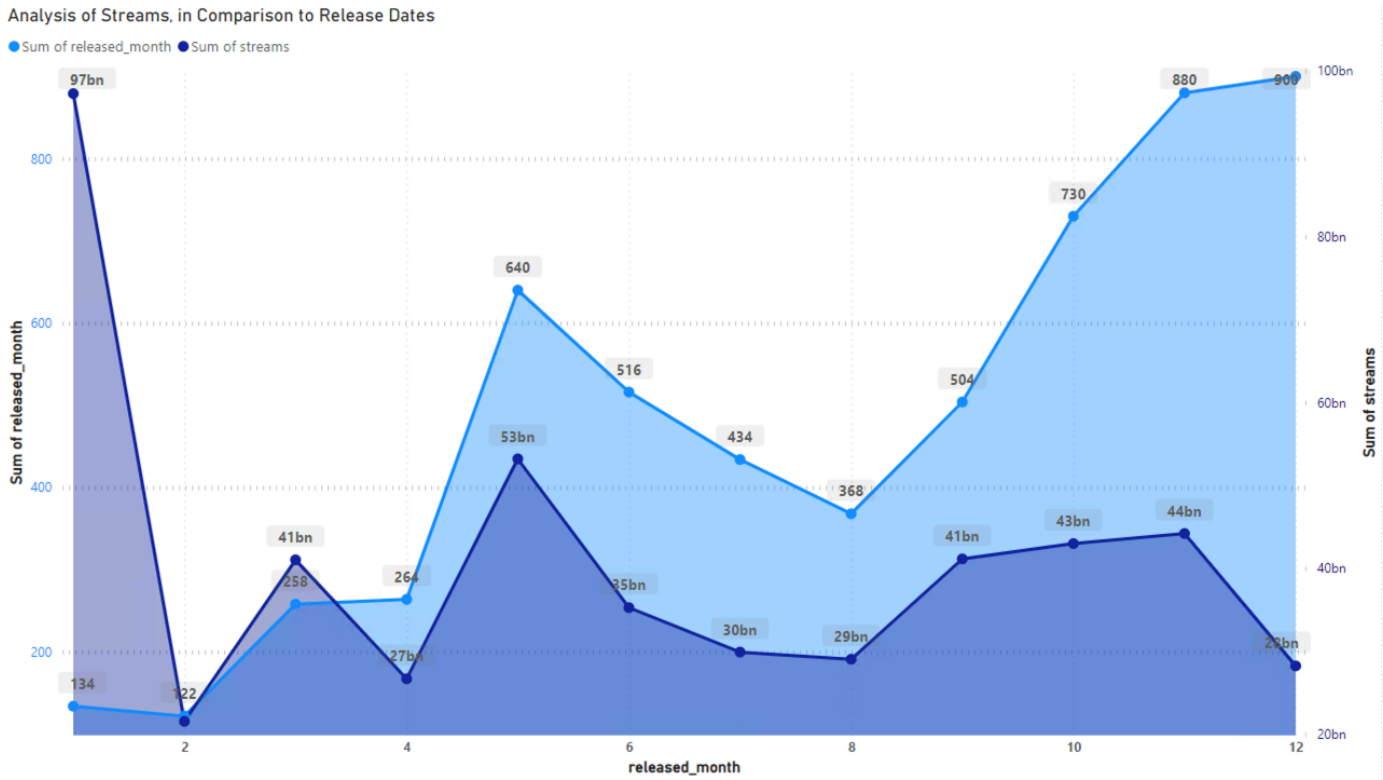





This visualization showcases the percentage of total songs in Spotify, Apple, Deezer, and Shazam charts and playlists. It is very clear that Spotify is the most popular music streaming service. With a total of 90% of all songs in this data being on Spotify charts and playlists, it is clear which streaming service is the most popular. Together, Apple, Deezer, and Shazam only hold 10% of all songs in this data set. In the article, "Best Music Streaming Service of 2023," Ty Pendlebury states, "But there is a clear winner. Spotify is the best streaming service. It offers the best mix of features, including deep community integration, plus the excellent Spotify Connect for streaming to all your devices." In other words, this supports our visualization which states that Spotify is the most popular streaming service, and offers the most variety of music to users.

3) During which times of the year were the most streamed songs released? Are any trends noticed?




This visualization shows an analysis of music streams throughout the year. We can see January – December (1-12), and we can see the variations of different streams throughout each

month (Dark Blue). We can also see the number of releases made throughout the year (Light Blue). In this visualization, it is easy to see how an increase in music releases can lead to an increase in overall streams. In January, we can see that there is a very high amount of streams (97 billion) but we can see a very low amount of releases (134). In December, we can see how there is a very high amount of releases (900) but a very low amount of streams. This is mot likely happening because all of the releases being made in December showcase their streams in January since the year is going to end. In the article, "Best Time To Release Music," Nadav Peleg states, "The most challenging months to release an album are typically November and December. Since all established and major independent labels release their newest tracks, the public and the media frequently pay attention to established acts…Of course, this is a fantastic time to release a song that is related to the holidays, the New Year, or both." This supports our visualization because we were also able to find that Towards the end of the year, there tends to be much more releases from major companies, which in turn, leads to streams increasing into the beginning months of the following year.

Works Cited

"8 Effective Data Cleaning Techniques for Better Data." *MonkeyLearn Blog*, 18 Oct. 2021, monkeylearn.com/blog/data-cleaning-techniques.

"The Evolution of Dance Music — Laura Gao." *Laura Gao*, www.lauragao.com/the-evolution-of-dance-music.

Peleg, Nadav. "Best Time to Release Music - SoundCampaign." *SoundCampaign*, 29 Aug. 2023, soundcamps.com/blog/best-time-to-release-music.

Pendlebury, Ty. "Best Music Streaming Service of 2023." *CNET*, 24 Oct. 2023, www.cnet.com/tech/services-and-software/best-music-streaming-service.