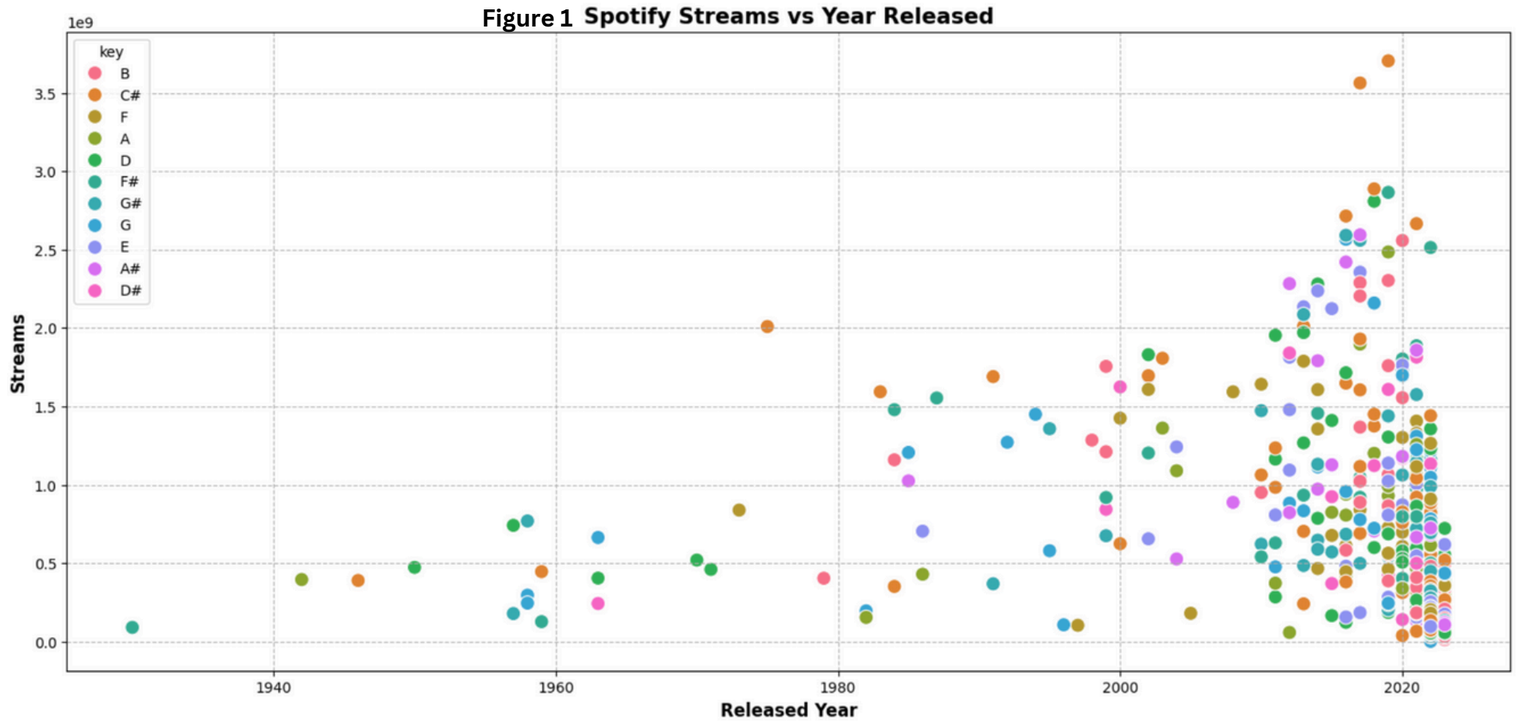
Decoding Music Trends on Spotify

By Yiming Tian
Information Visualization Final Report

Figure 1 Spotify Streams vs Year Released



In figure 1, the scatter plot here shows Spotify streams over the years songs that were released. Each color of the data represents different keys of the songs. The more crowded spot represent a more popular year(s). We can conclude from the figure that the songs that have the most streams on Spotify have the release year in 2020 in the given time span.

1. Popularity Analysis

Figure 2

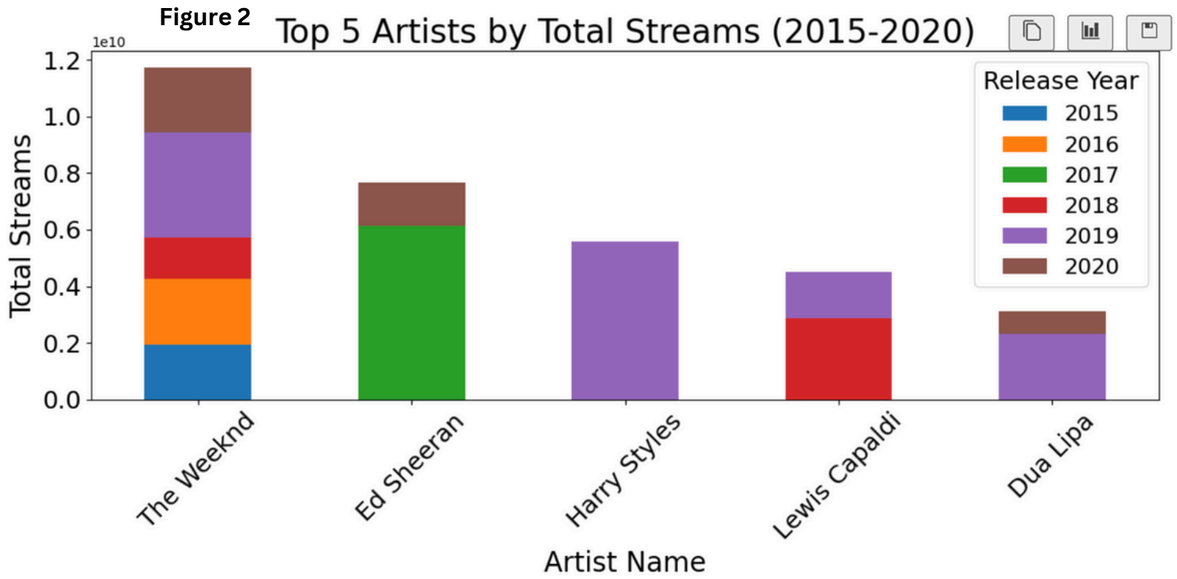
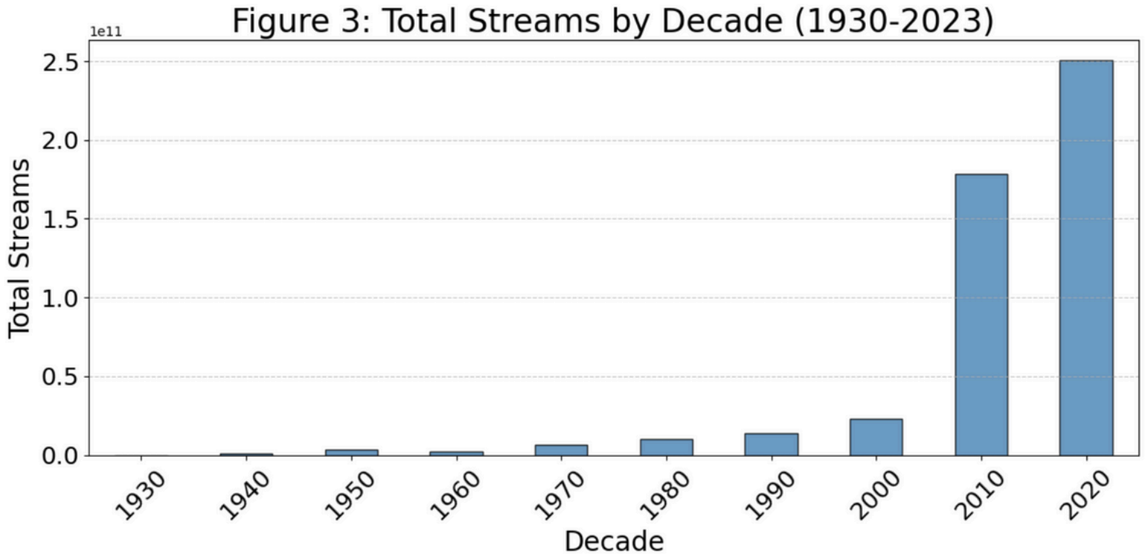
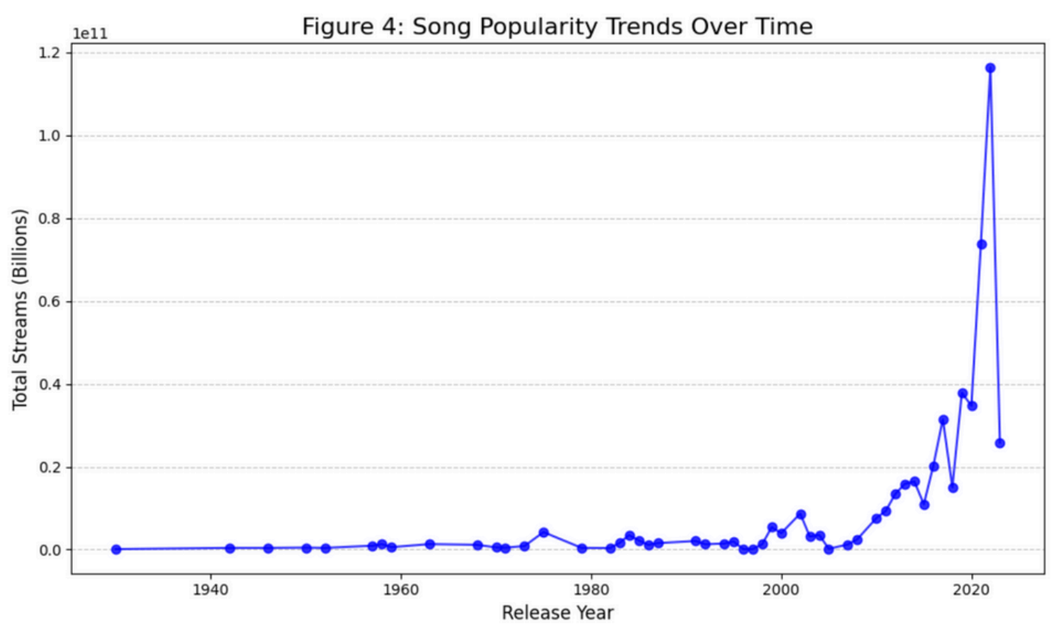


Figure 2 is a bar plot with the top 5 artists: The Weekend, Ed Sheeran, Harry Styles, Lewis Capaldi, and Dua Lipa among the years between 2015 and 2020. With their released year as hue in every bar.
We can also see that 2019 plays a bigger factor is their streams comparing to any other years, since the purple section represents 2019 and it is more prominent.

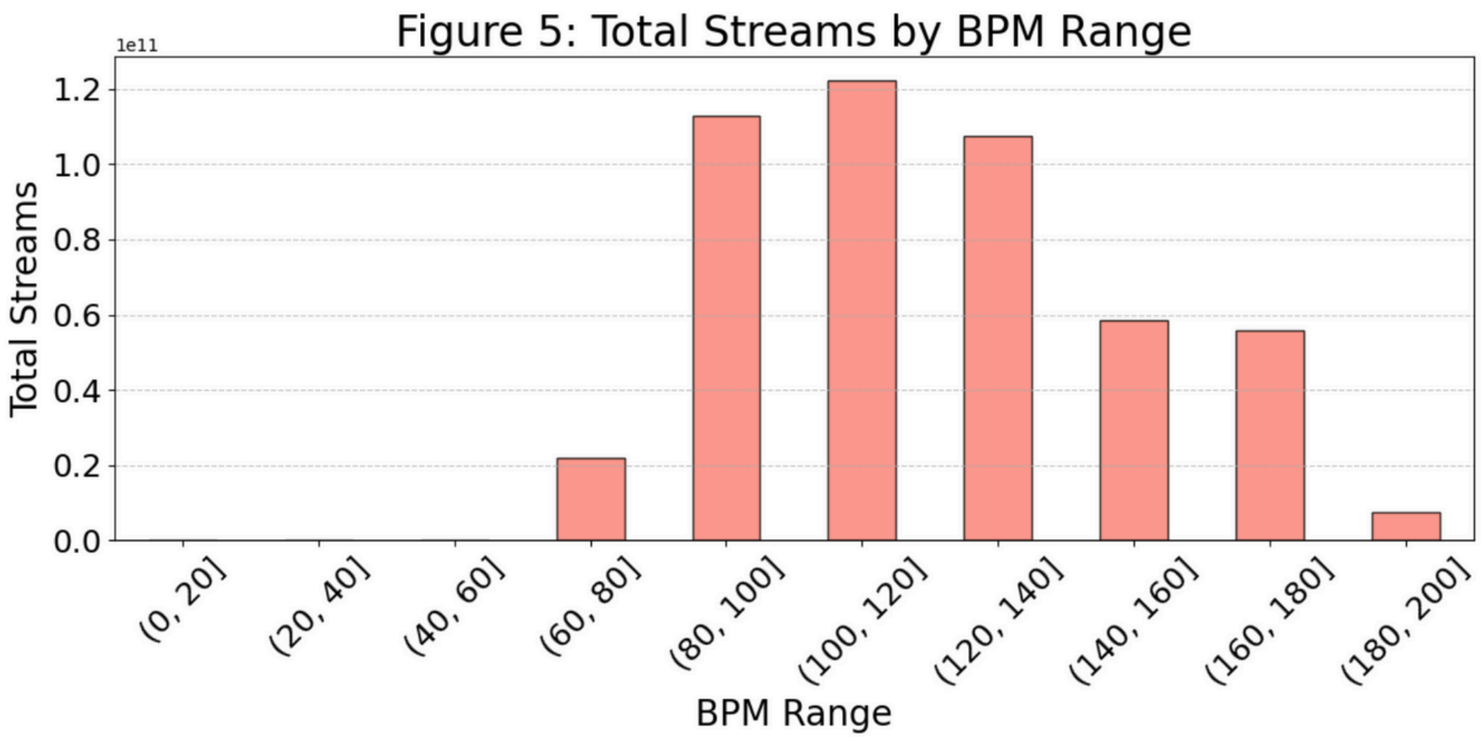


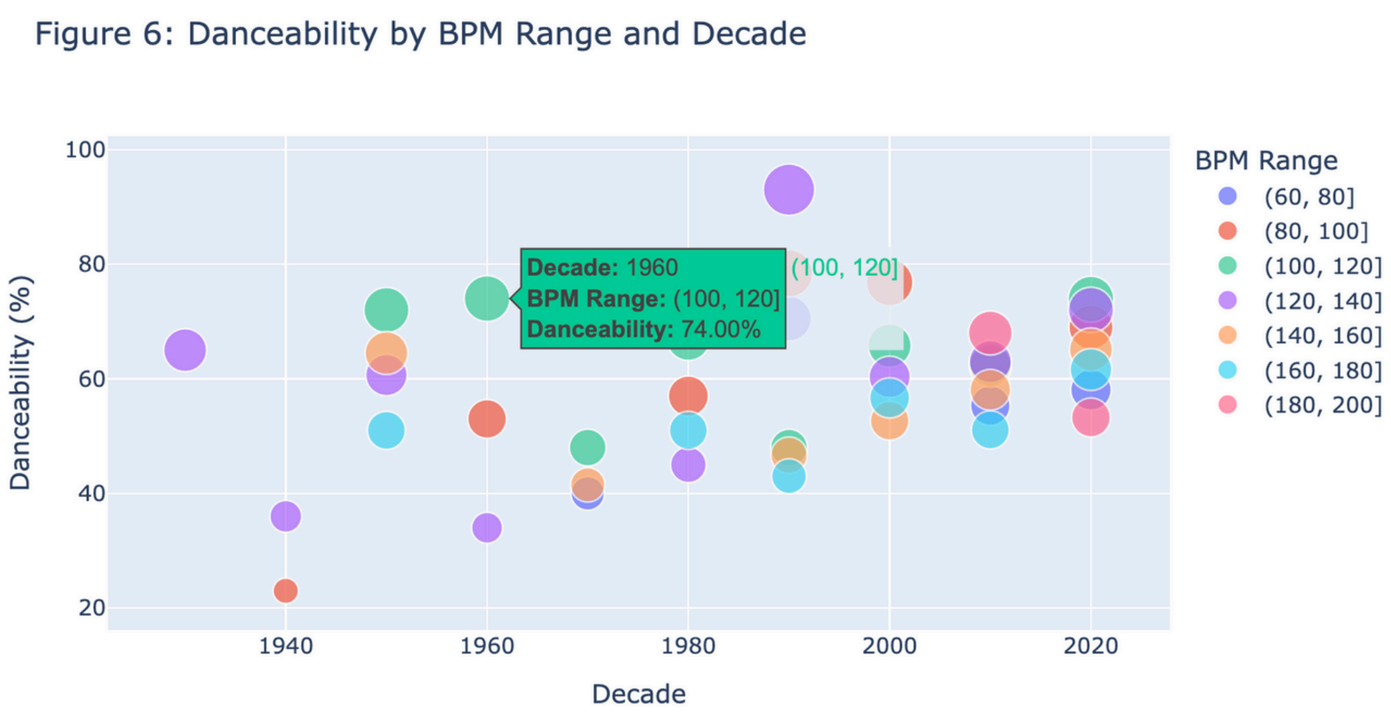


In the two figures above, figure 3 and figure 4:

- **Emergence of a Streaming Boom Around 2000:**
  The sharp rise in total streams begins in the early 2000s, coinciding with the rise of MP3s, iTunes, and eventually, global streaming platforms.
  The 2010s mark a steeper upward trajectory, highlighting the widespread acceptance of streaming as the primary music consumption method.
- **2020s Peak:**
  A massive spike is observed in the 2020s, which could be attributed to:
  - Global growth of streaming services.
  - Increased availability of music catalogs online.
  - Pandemic-related shifts in consumer behavior, with more people turning to digital entertainment.
- **Sustained Interest in Older Music:**
  While recent years dominate in streams, the steady growth from earlier decades shows sustained interest in older songs, possibly due to remasters, covers, or viral trends on social media platforms like TikTok.

2. **BPM Analysis**





From figure 5 and figure6, we can conclude that:

- **Popularity of Mid-Range BPM Songs:**
  Songs in the BPM ranges of 100–120 and 120–140 dominate in terms of total streams.
  This aligns with typical tempos found in pop, dance, and electronic music, which are generally more energetic and widely appealing.
  Low and High BPM Songs Are Less Popular

- **BPM Range vs. Danceability:**
  The BPM range of 100–120 is strongly associated with high danceability percentages, suggesting that this range is optimal for creating danceable tracks.

- **Modern Music Trends:**
  From the 2000s onward, there is a clustering of high danceability songs across various BPM ranges, reflecting the dominance of danceable pop and electronic music in the digital streaming era.

3. **Top Artists and Keys analysis**

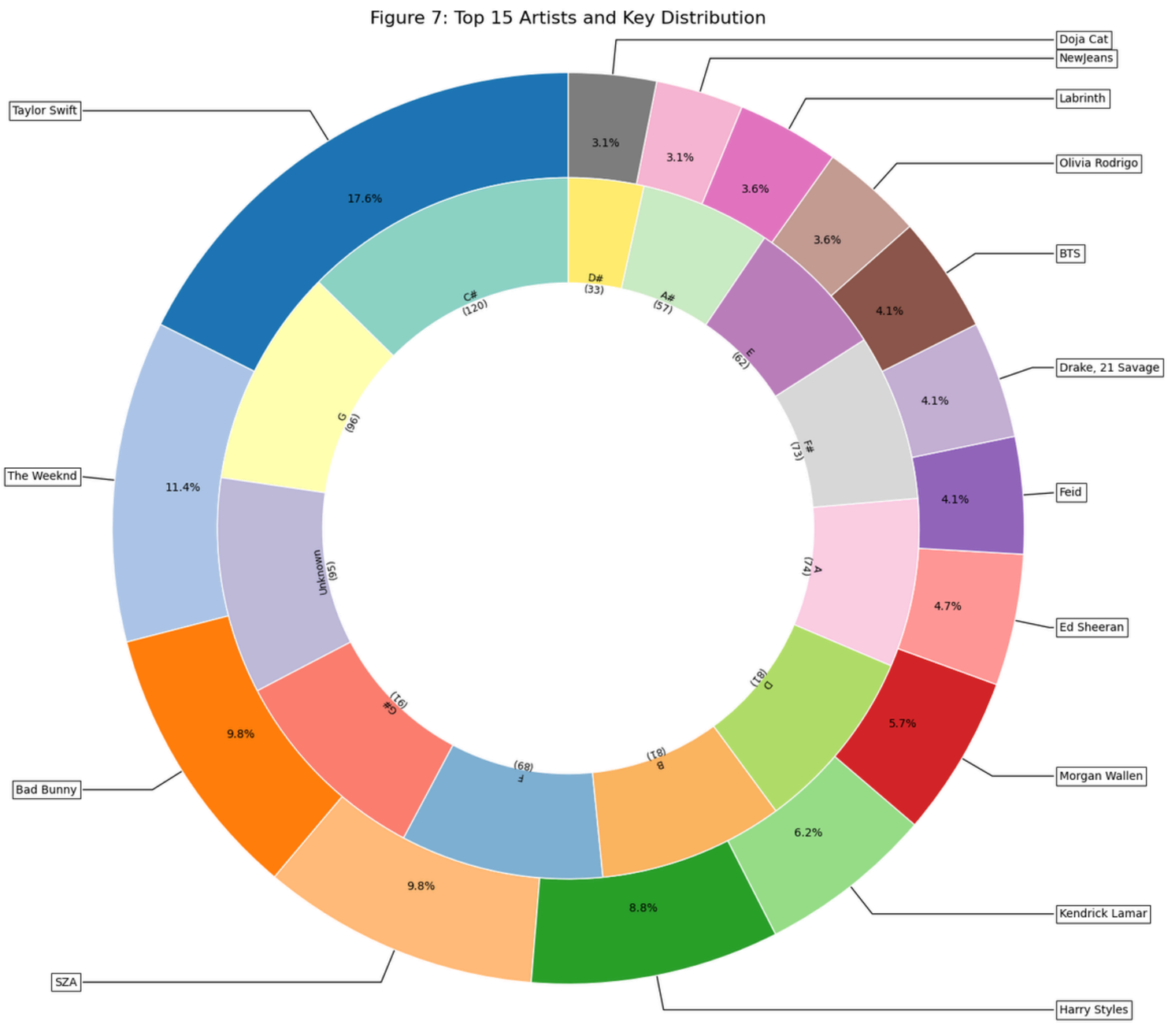


Figure 7: Top 15 Artists and Key Distribution

We can draw a conclusion from figure 7 that:

- **Artist Dominance**

Taylor Swift and SZA have the largest presence in the dataset, likely due to their popularity, the dataset's origin (e.g., streaming trends or personal preferences), or the data collection method.

- **C Major Prevails**

C Major is the most frequent key, possibly due to its ease of playability, common use in popular genres, or historical and cultural influences.

- **Diverse Representation**

The chart showcases a wide range of artists and keys, reflecting diverse musical tastes and the richness and complexity of the music within the dataset.

Importance:

As Spotify wrapped is getting more and more attention by the end of each year recently, by analyzing a Spotify dataset using a variety of visualization techniques, including histograms, scatter plots, interactive plots, and the double-layer donut chart, we can gain valuable insights into broader musical trends and preferences. Histograms can reveal the distribution of streams across different decades and BPM ranges, shedding light on the evolution of popular music. Scatter plots can explore correlations between various musical features, such as danceability and energy, helping us understand which characteristics are often associated with popular songs. Interactive plots allow for dynamic exploration of the data, enabling us to filter and sort by different criteria to identify patterns and trends. Finally, the donut chart effectively summarizes the top artists and their key distributions, providing a concise overview of the musical landscape within the dataset. This multi-faceted approach allows for a comprehensive and insightful analysis of the broader music trends and preferences reflected in the Spotify data.

Data & Method:

The data 'spotify-most-streamed-songs.csv' is from Kaggle, which is an open platform for data scientists. The project is written in python with jupyter notebook. Some libraries that were used to generate the visualizations in python are seaborn, matplotlib, and plotly.

Github:

https://github.com/Yiming-Tian/final-report

Dataset(From Kaggle):

https://www.kaggle.com/datasets/abdulszz/spotify-most-streamed-songs

Logo Citation:

https://loodibee.com/spotify/spotify-logo-black-white/