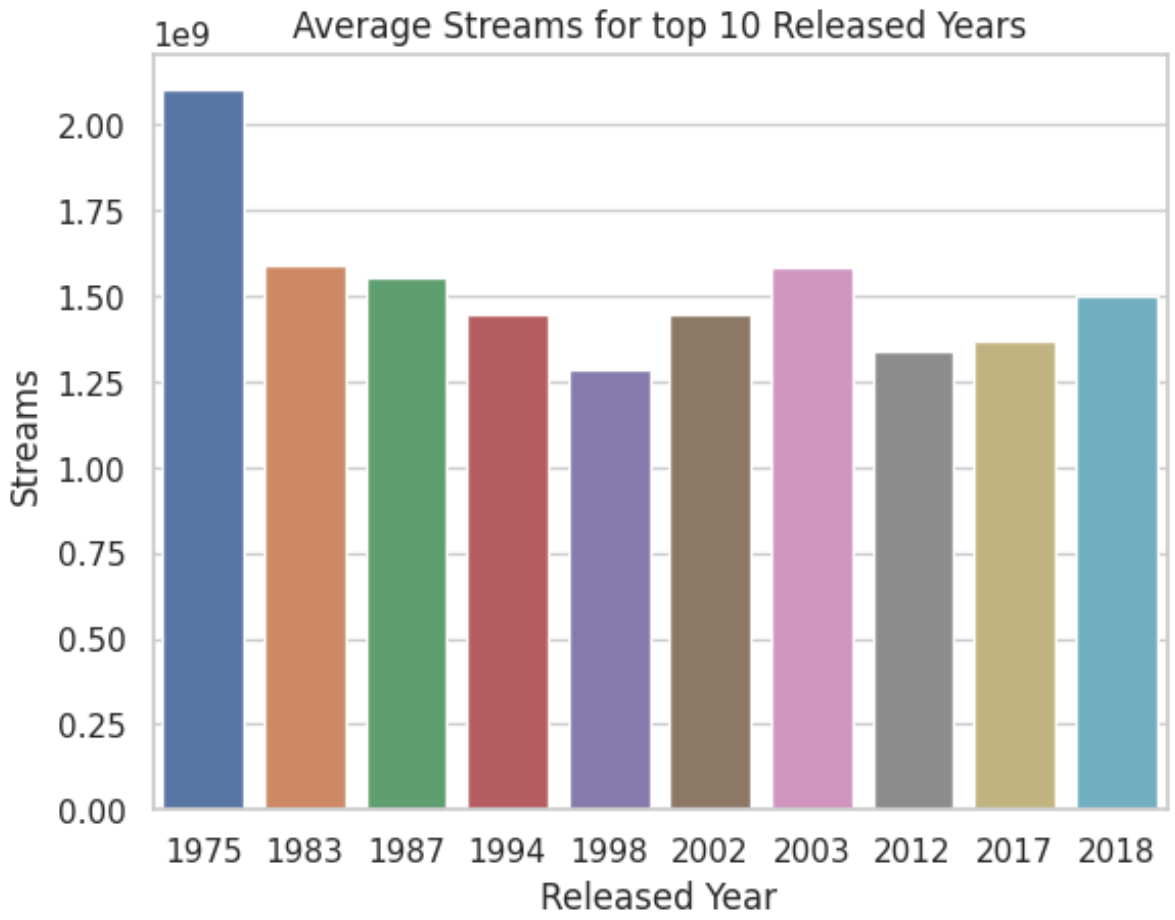
Spotting Spotify Trends

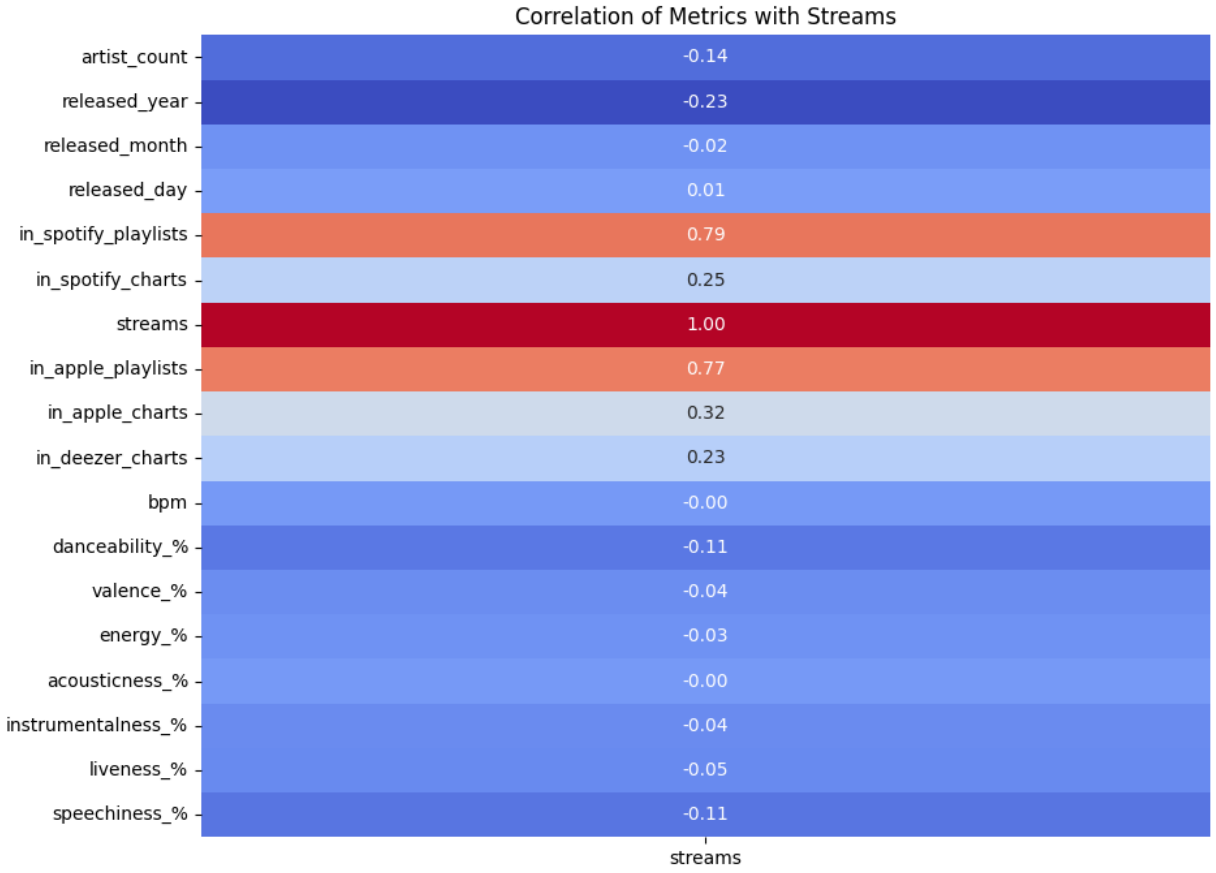
A look into what defines popular music in 2023

What Songs are we Listening to Today?



Top 10 Tracks on Spotify

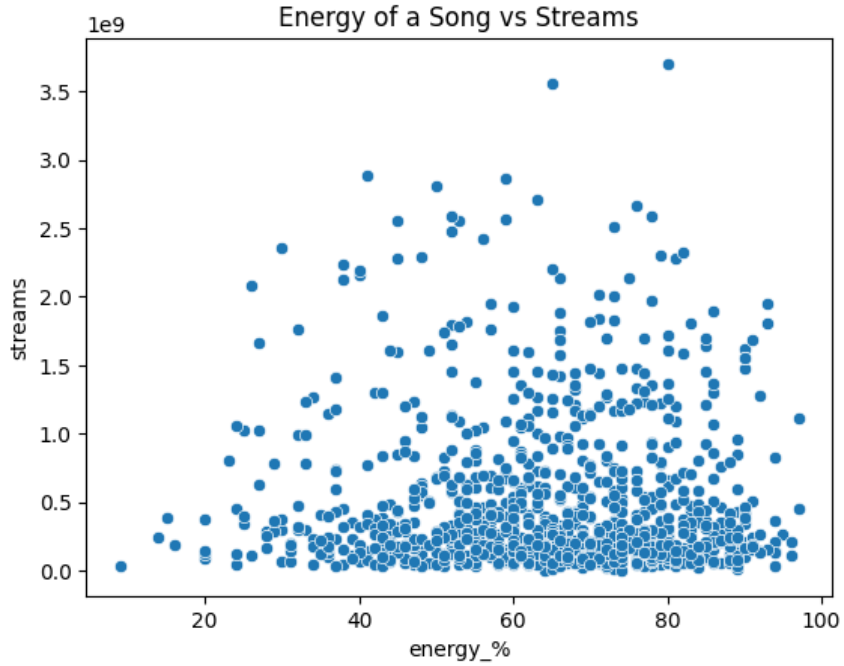
	track_name	artist(s)_name	streams
0	Blinding Lights	The Weeknd	3703895074
1	Shape of You	Ed Sheeran	3562543890
2	Someone You Loved	Lewis Capaldi	2887241814
3	Dance Monkey	Tones and I	2864791672
4	Sunflower - Spider-Man: Into the Spider-Verse	Post Malone, Swae Lee	2808096550
5	One Dance	Drake, WizKid, Kyla	2713922350
6	STAY (with Justin Bieber)	Justin Bieber, The Kid Laroi	2665343922
7	Believer	Imagine Dragons	2594040133
8	Closer	The Chainsmokers, Halsey	2591224264
9	Starboy	The Weeknd, Daft Punk	2565529693

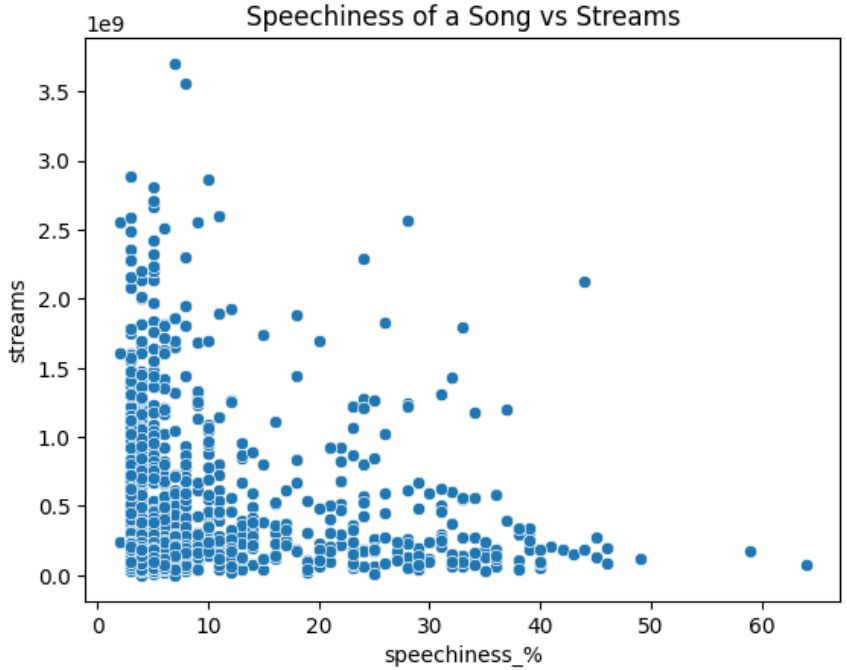


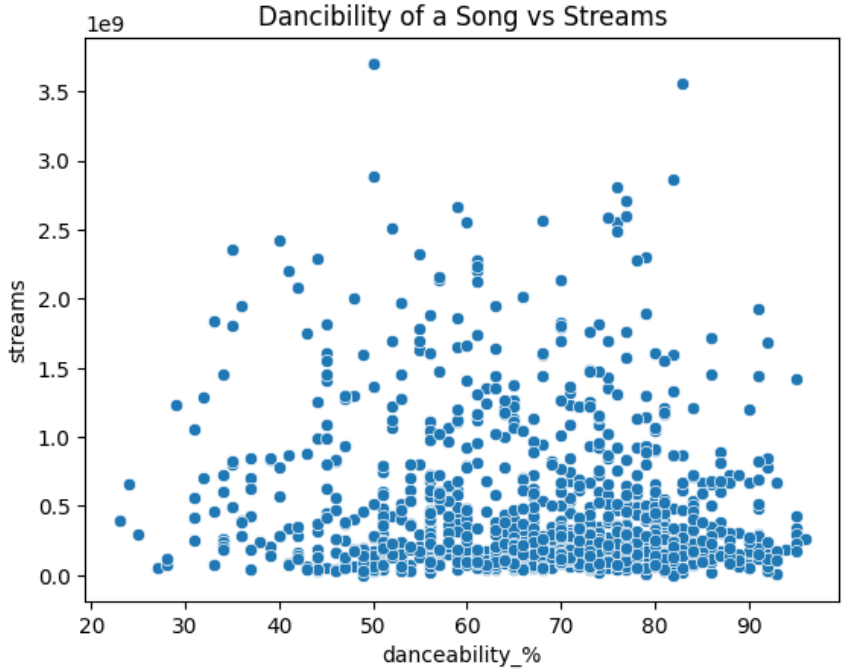
Simpler Songs Perform Better

The Essence of Minimalism

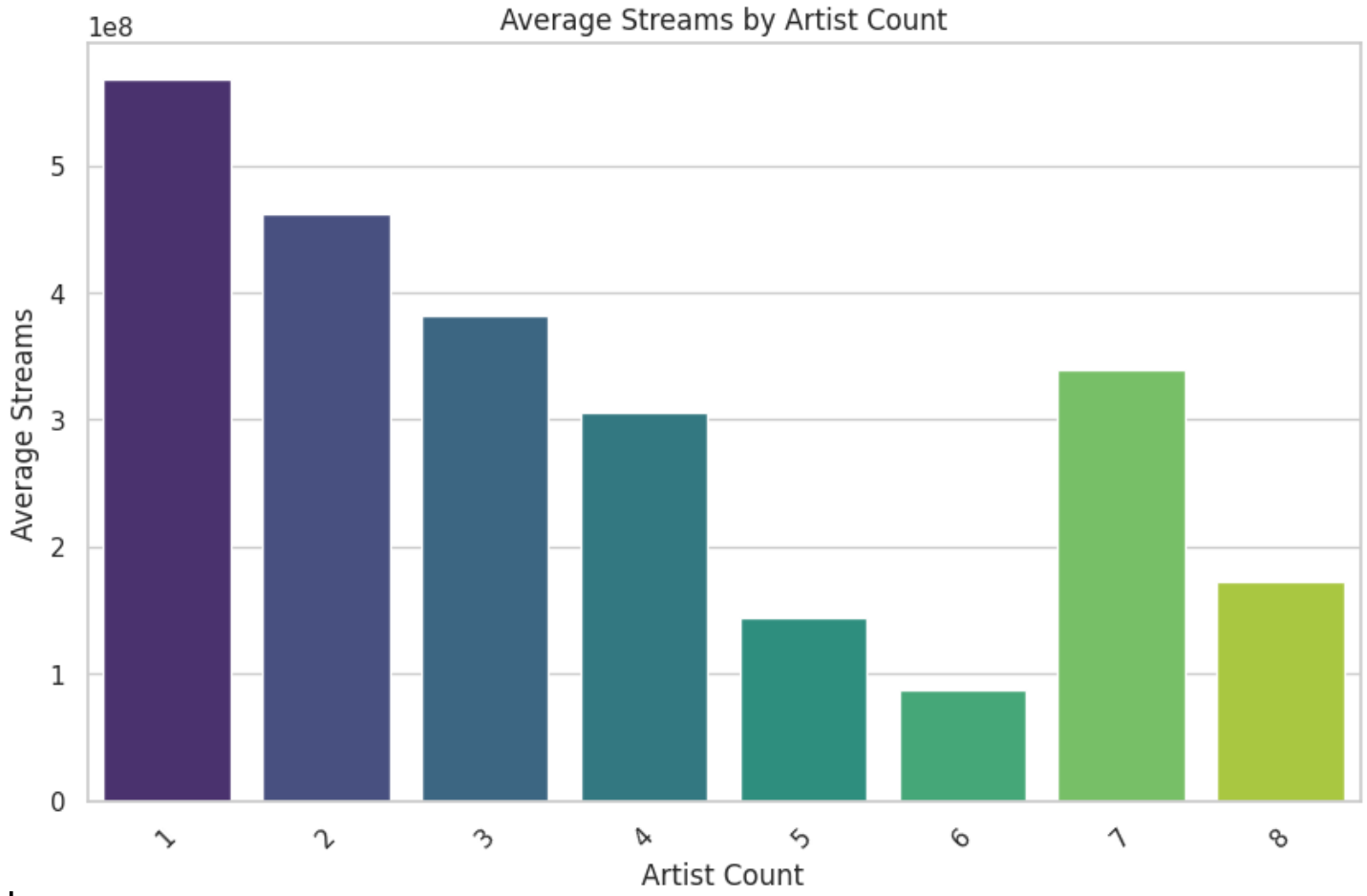
Higher Song Metrics Negatively, or Not, Correlated with Streams





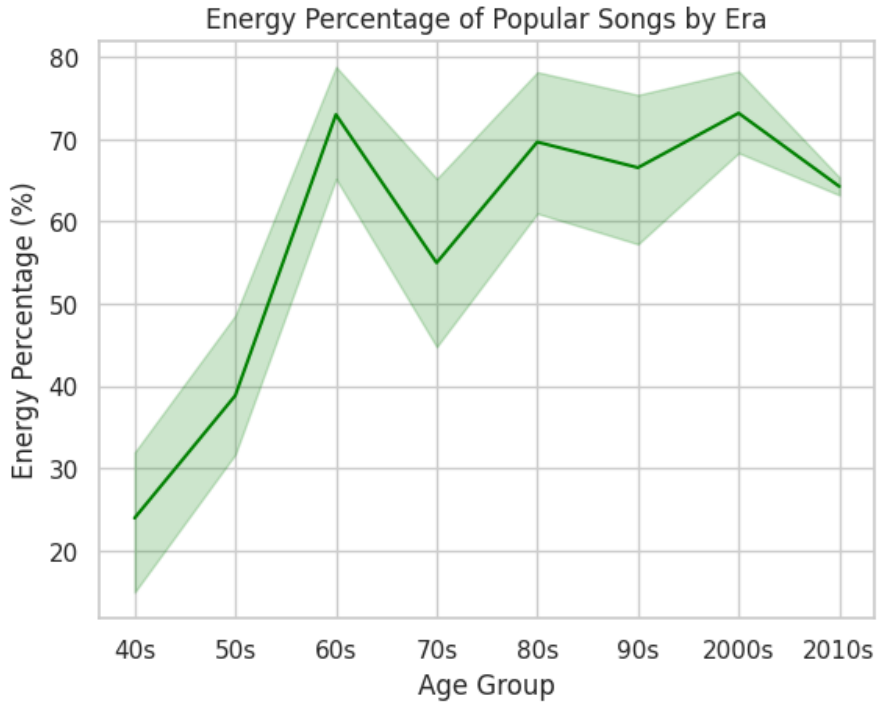


Flying Solo a Recipe for Success?

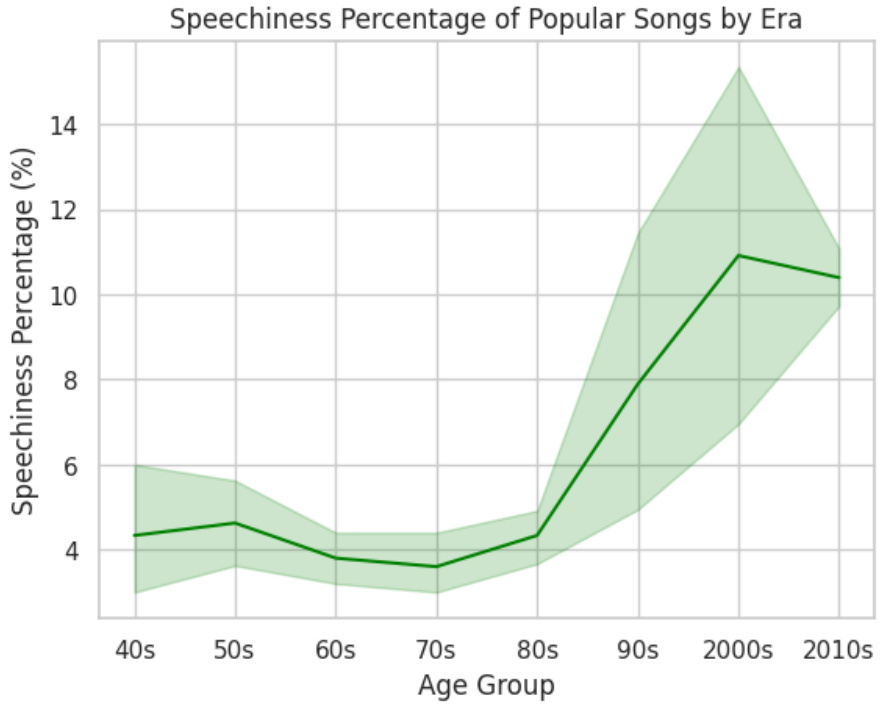


Low sample size for high artist counts:
Only 2 songs for 7

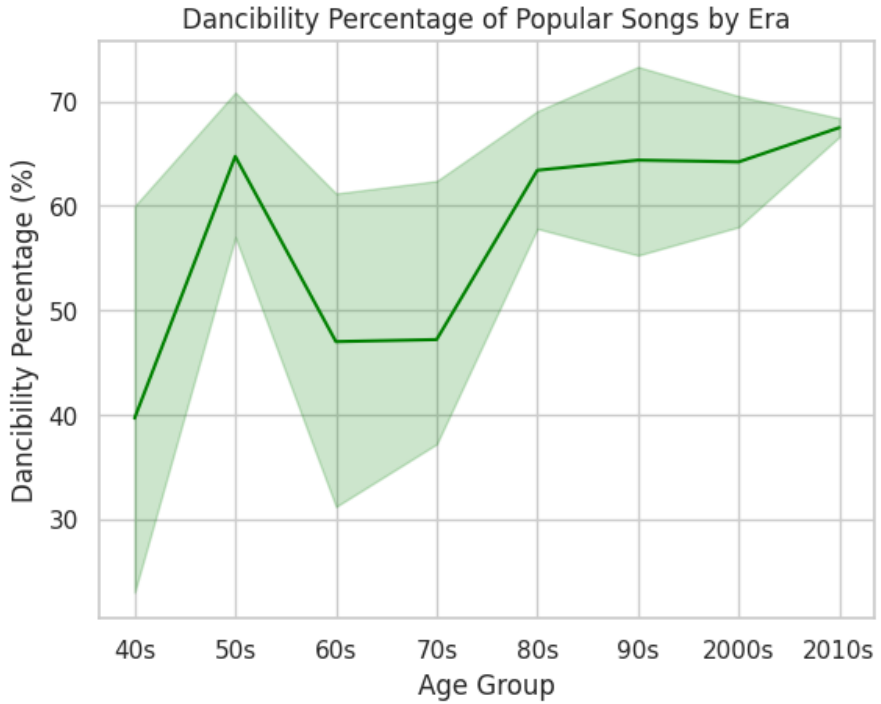
The popular songs from the 60s and 2000s are high energy, while the popular songs from the 40s, 70s, and 90s are lower energy



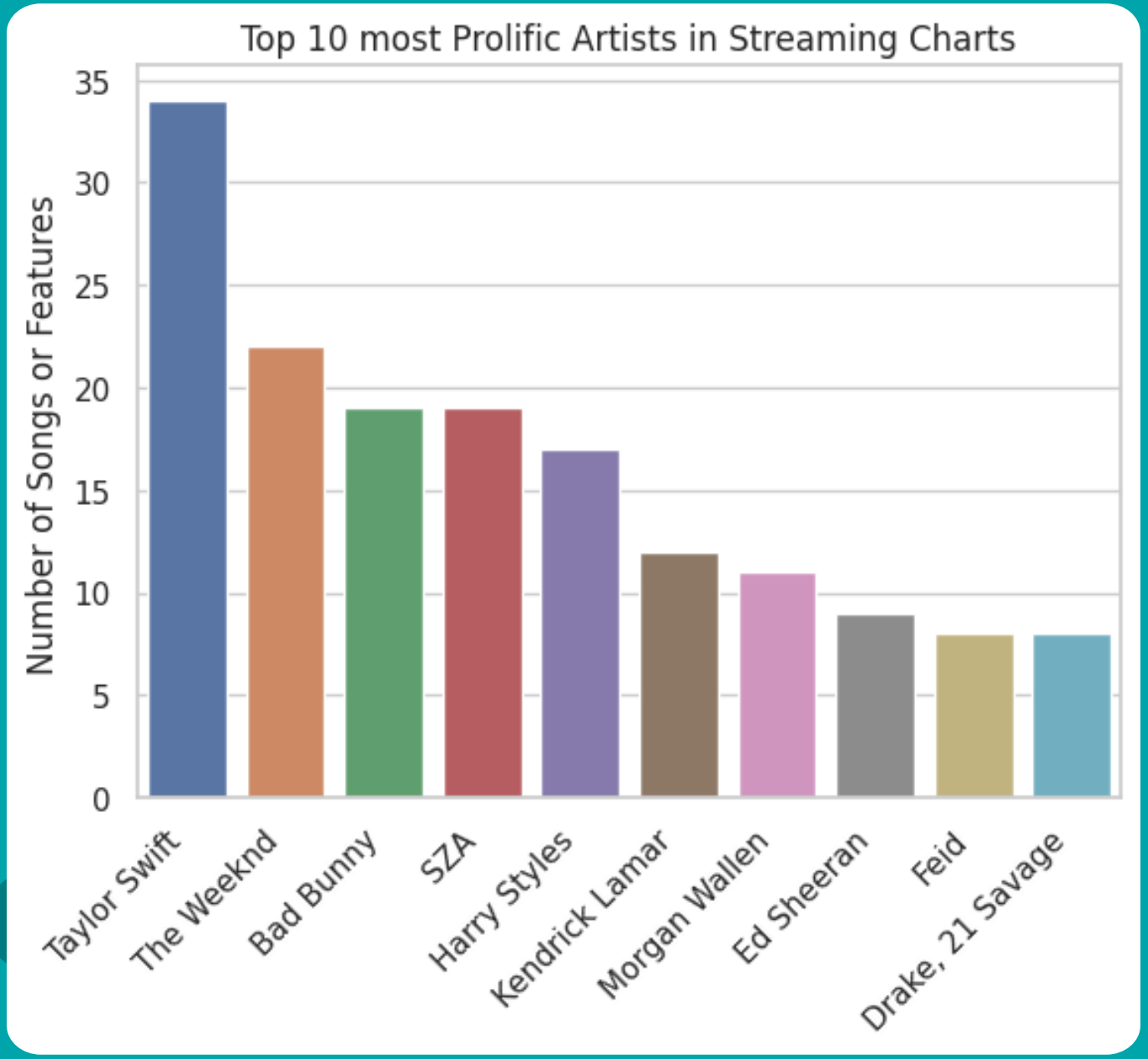
Popular songs from the 2000s are very speechy, while popular songs from the 70s are less so



Popular songs from the 50s and 2010s are very danceable, while the 40s, 60s, and 70s popular songs are less danceable

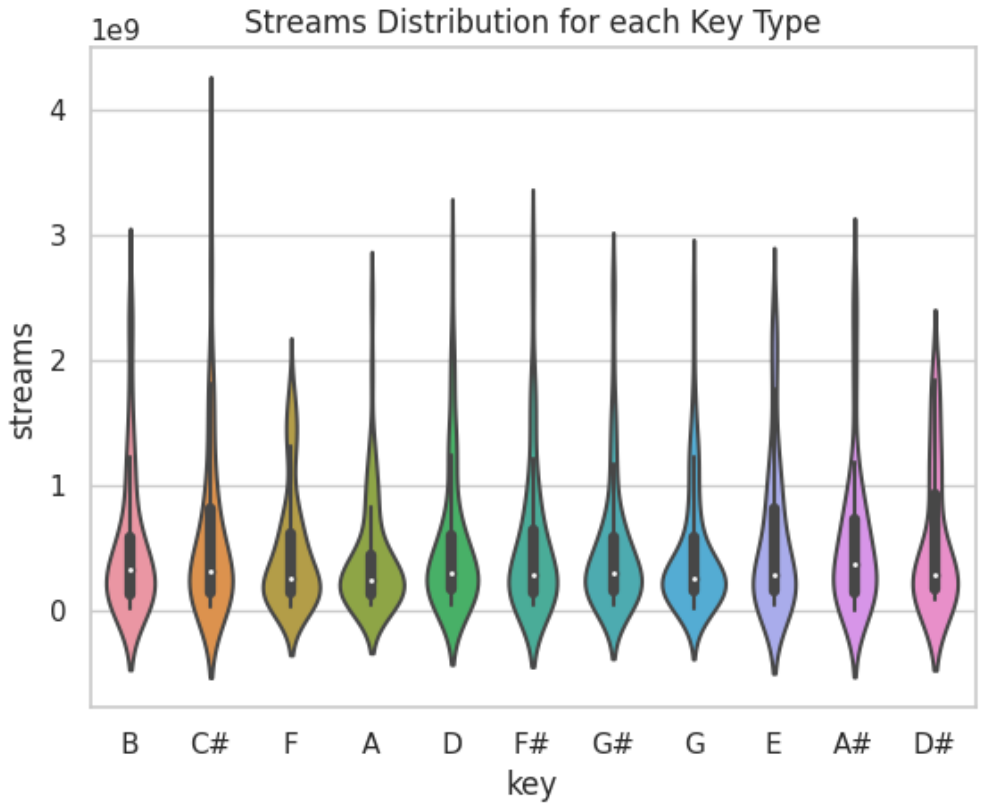


Which Artist Wears the Crown?

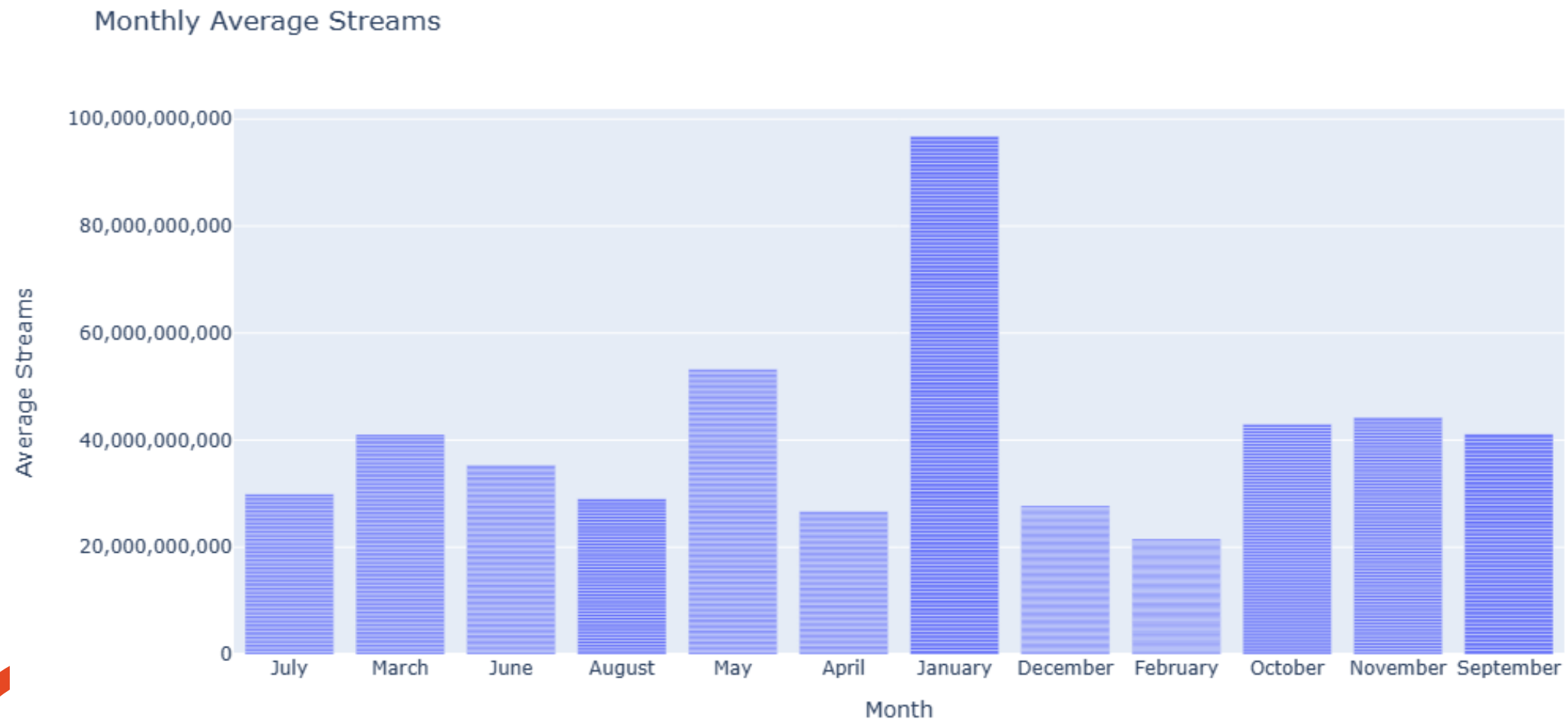


Keys for Clout

Want a song with high potential? Consider making it in C#



Which Month are we Bumping? (Hint: It's January)



Challenges

- Encoding Differences: Dataset not encoded in Default UTF-8, had to research specifications. Only 1.4% of websites use ISO-8859-1 encoding

```
[4]  #Using a different encoding value to ensure accuracy
     df = pd.read_csv('spotify-2023.csv', encoding='ISO-8859-1')
```

- Splitting Artists: List of Artists in one Column

```
[40] #Using explode function to find raw counts of artists in data frame
     df = df.explode('artist(s)_name')
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 953 entries, 0 to 952
Data columns (total 24 columns):
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   track_name            953 non-null    object
 1   artist(s)_name        953 non-null    object
 2   artist_count          953 non-null    int64
 3   released_year         953 non-null    int64
 4   released_month        953 non-null    int64
 5   released_day          953 non-null    int64
 6   in_spotify_playlists  953 non-null    int64
 7   in_spotify_charts     953 non-null    int64
 8   streams               953 non-null    object
 9   in_apple_playlists    953 non-null    int64
 10  in_apple_charts       953 non-null    int64
 11  in_deezer_playlists   953 non-null    object
 12  in_deezer_charts      953 non-null    int64
 13  in_shazam_charts      903 non-null    object
 14  bpm                   953 non-null    int64
 15  key                   858 non-null    object
 16  mode                  953 non-null    object
 17  danceability_%        953 non-null    int64
 18  valence_%             953 non-null    int64
 19  energy_%              953 non-null    int64
 20  acousticness_%        953 non-null    int64
 21  instrumentalness_%    953 non-null    int64
 22  liveness_%            953 non-null    int64
 23  speechiness_%         953 non-null    int64
dtypes: int64(17), object(7)
memory usage: 178.8+ KB
```

Challenges

- Interpreting Columns. What is "Liveness" or "Speechiness"?

- **bpm**: *Beats per minute, a measure of song tempo*
- **key**: *Key of the song*
- **mode**: *Mode of the song (major or minor)*
- **danceability_%**: *Percentage indicating how suitable the song is for dancing*
- **valence_%**: *Positivity of the song's musical content*
- **energy_%**: *Perceived energy level of the song*
- **acousticness_%**: *Amount of acoustic sound in the song*
- **instrumentalness_%**: *Amount of instrumental content in the song*
- **liveness_%**: *Presence of live performance elements*
- **speechiness_%**: *Amount of spoken words in the song*

Source

- Most Streamed Spotify Songs 2023 (kaggle.com) "This dataset contains a comprehensive list of the most famous songs of 2023 as listed on Spotify. It provides insights into each song's attributes, popularity, and presence on various music platforms"