

ISTA 131



# Final Project: Spotify

Claire Baker, Megan Straffon, Erin Berridge,  
Jessica Shaffer, Briley Winge



# Introduction

For our final project, we chose to analyze various datasets containing information from Spotify. We wanted to analyze Spotify data to make more descriptive figures in order to better summarize and make predictions from the data.

# Spotify Artist Stats CSV

spotify_artist_data							
	Artist Name	Lead Streams	Feats	Tracks	One Billion	100 Million	Last Updated
1	Drake	50,162,292,808	19,246,513,666	262	6	130	19.09.22
2	Bad Bunny	44,369,032,140	5,391,990,975	163	5	118	20.09.22
3	Ed Sheeran	38,153,682,361	2,791,278,201	240	10	62	19.09.22

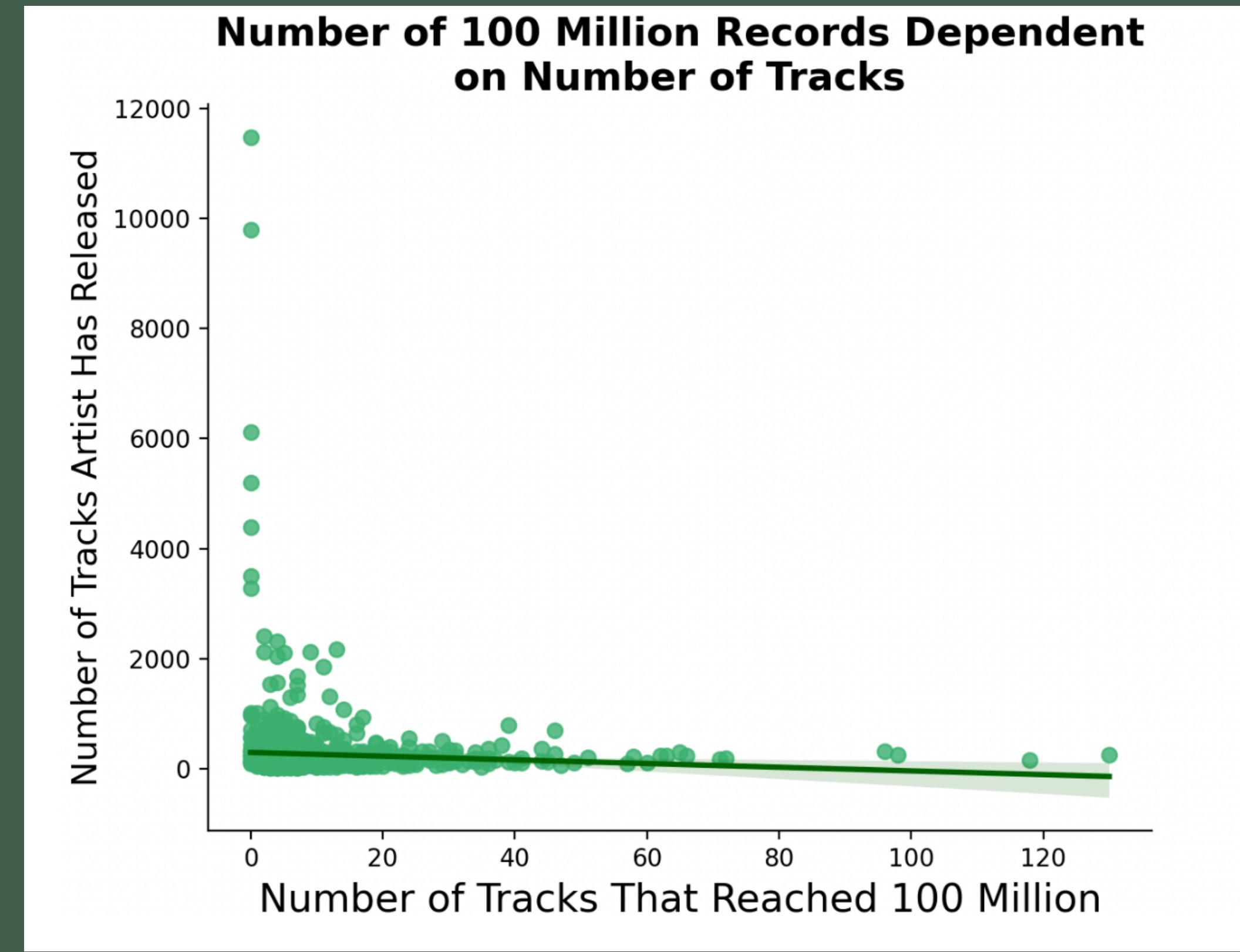
- Statistics for the current Spotify top artists
- Sorted in descending order of artists lead streams
- Data such as number of streams, feats, number of tracks, number of tracks that reached one billion and 100 million, and the date it was last updated

# Spotify Song Attributes CSV

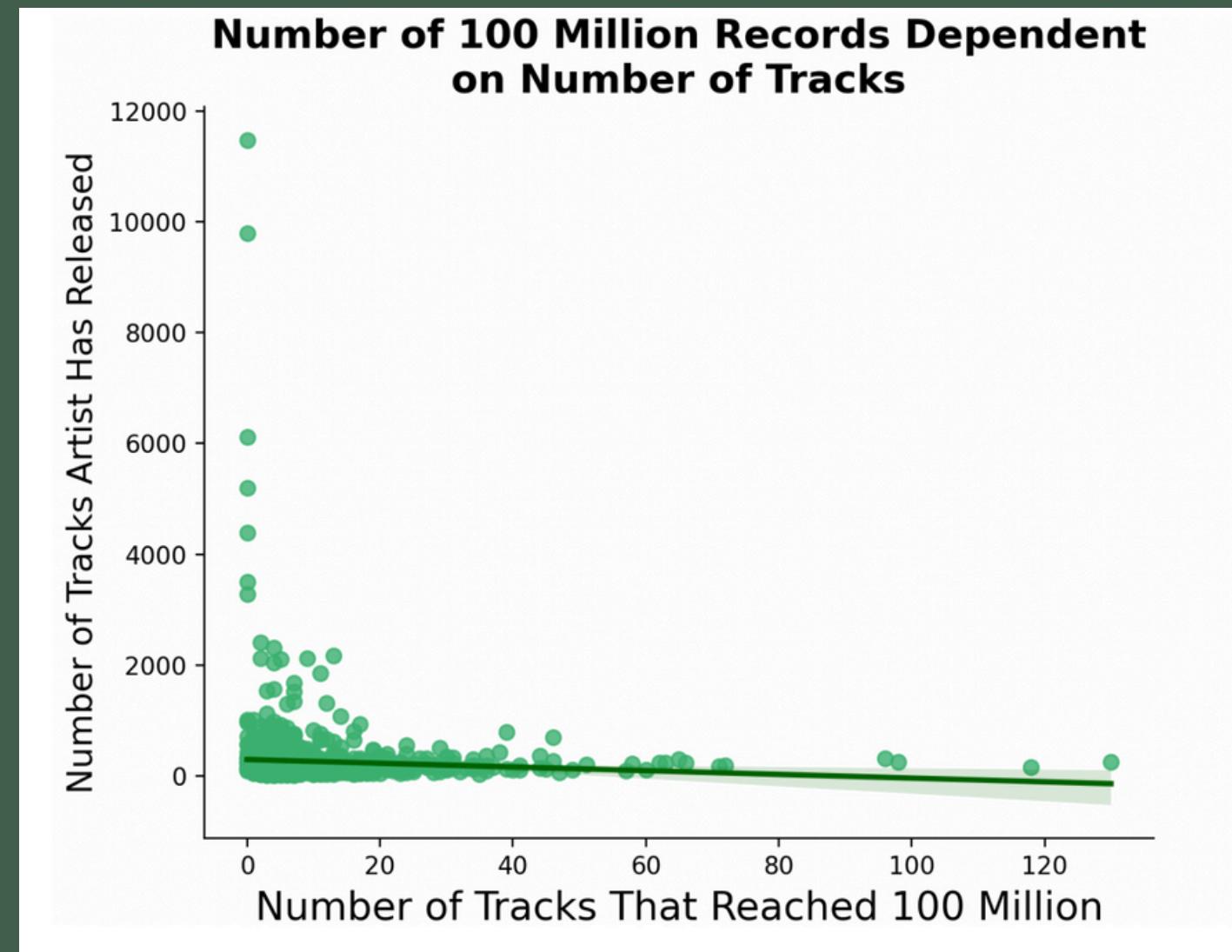
acousticness	danceability	duration_ms	energy	instrumentalness	key	liveness	loudness	mode	speechiness
0.0102	0.833	204600	0.434	0.0219	2	0.165	-8.795	1	0.431
0.199	0.743	326933	0.359	0.00611	1	0.137	-10.401	1	0.0794
0.0344	0.838	185707	0.412	0.000234	2	0.159	-7.148	1	0.289
0.604	0.494	199413	0.338	0.51	5	0.0922	-15.236	1	0.0261

- A dataset of 2017 songs with attributes from Spotify's API
- Contains track attributes, song name, and artist name
- The 13 track attributes: acousticness, danceability, durationms, energy, instrumentalness, key, liveness, loudness, mode, speechiness, tempo, timesignature, valence

# 100 Million Records vs. Tracks

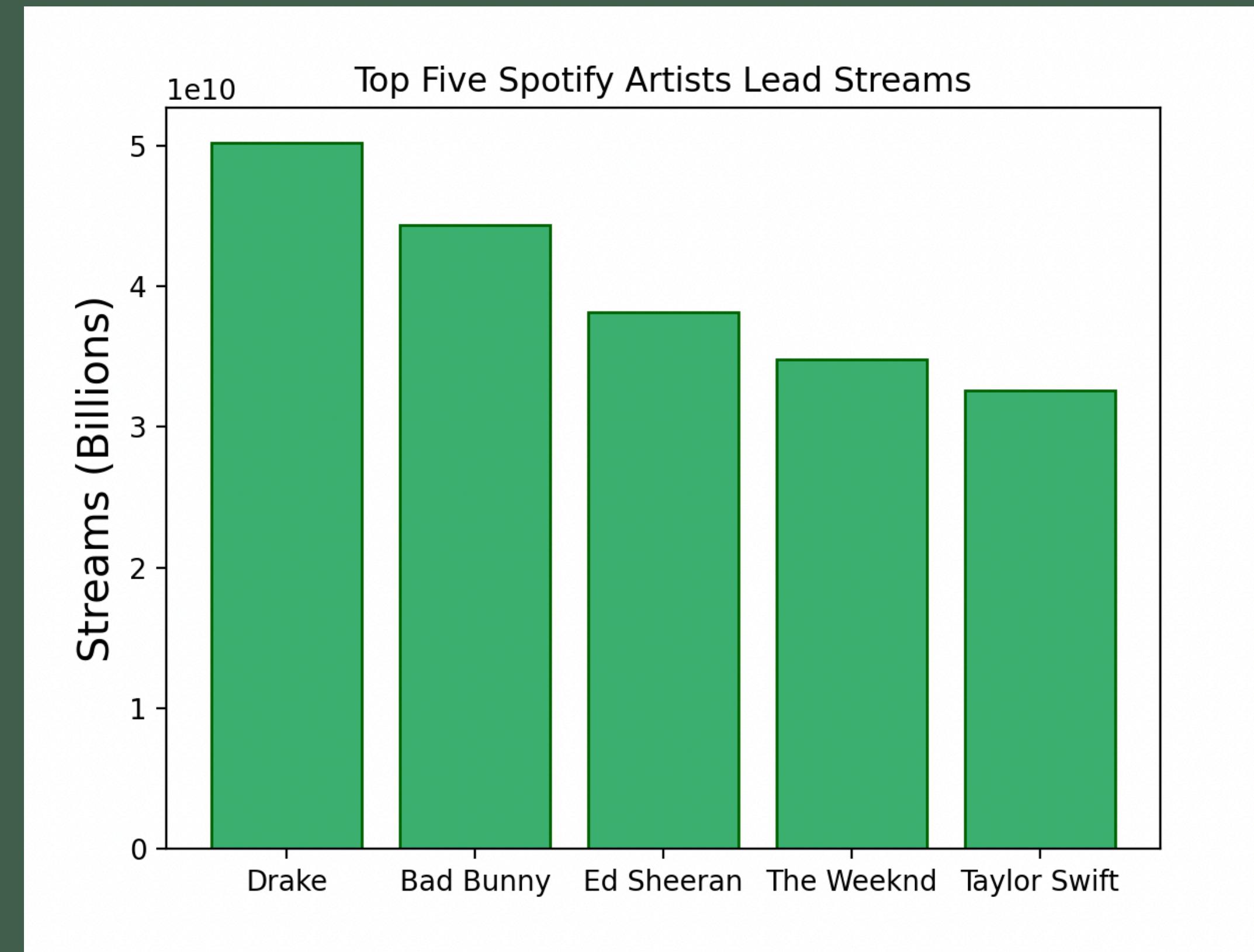


# 100 Million Records vs. Tracks

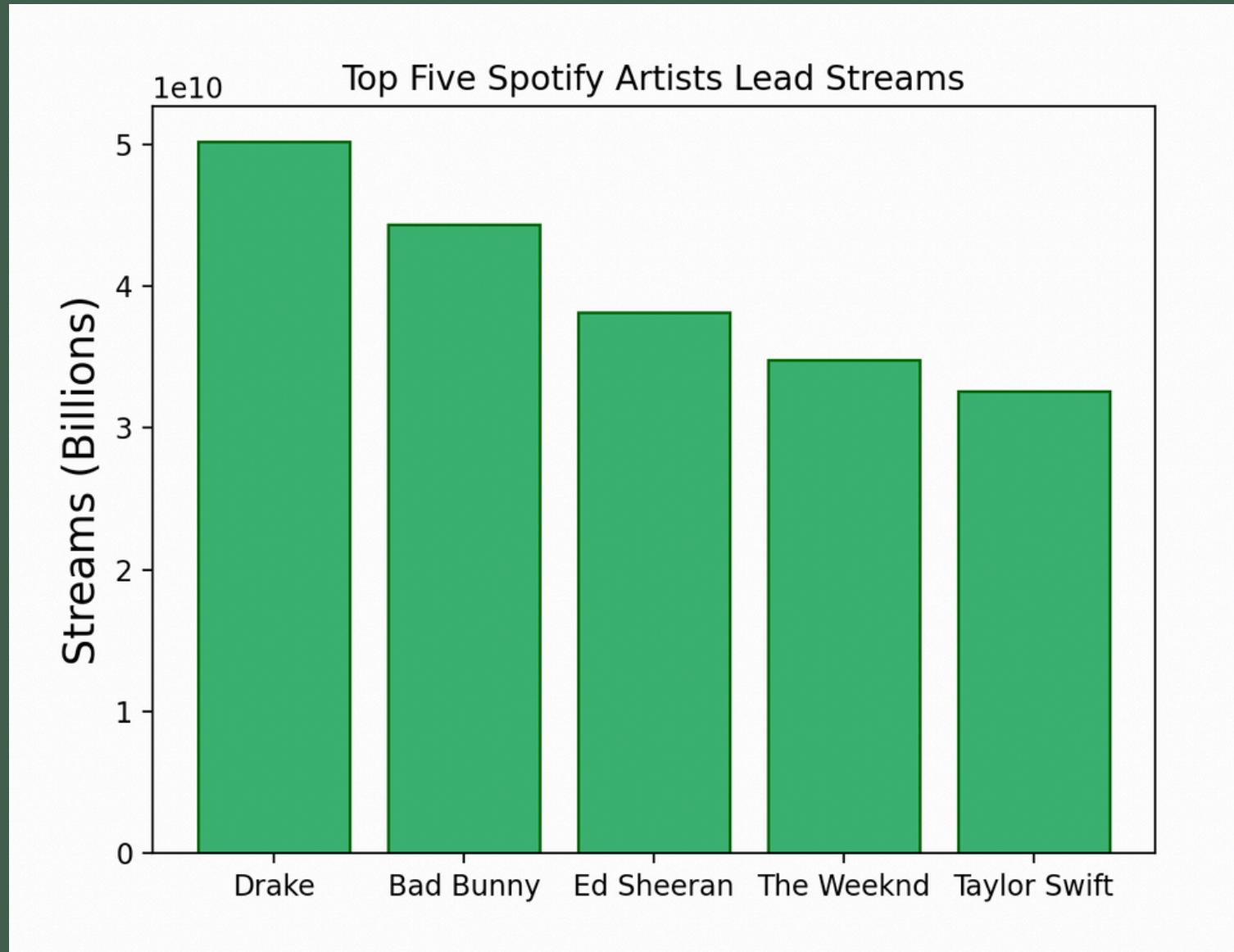


- Scatterplot with a regression line
- Demonstrates the number of tracks that reached 100 Million streams correlated with the number of tracks an artist has released
- The more tracks an artist has released, the less likely they are to have a 100 million streams on them

# Top Five Artists' Number of Streams

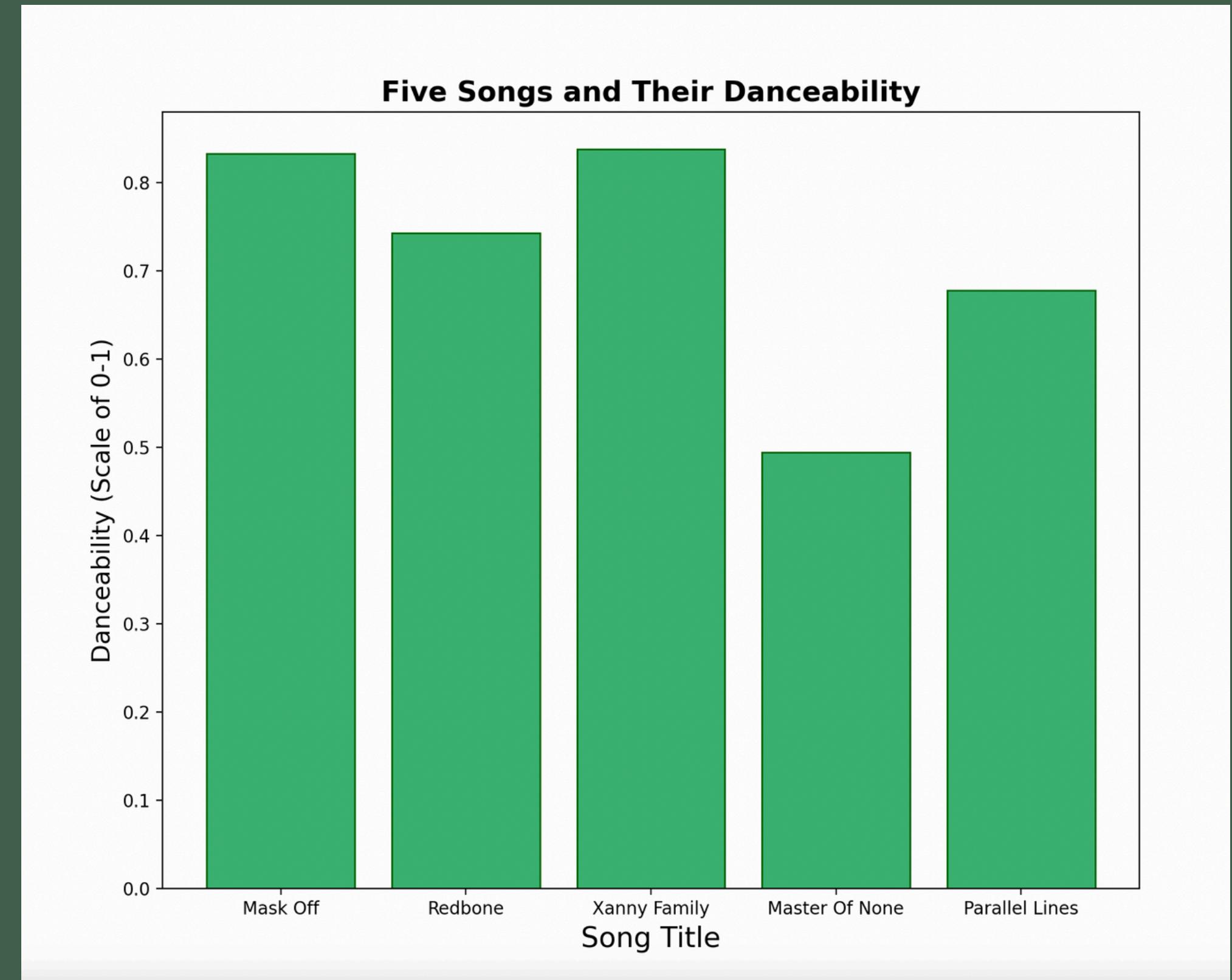


# Top Five Artists' Number of Streams

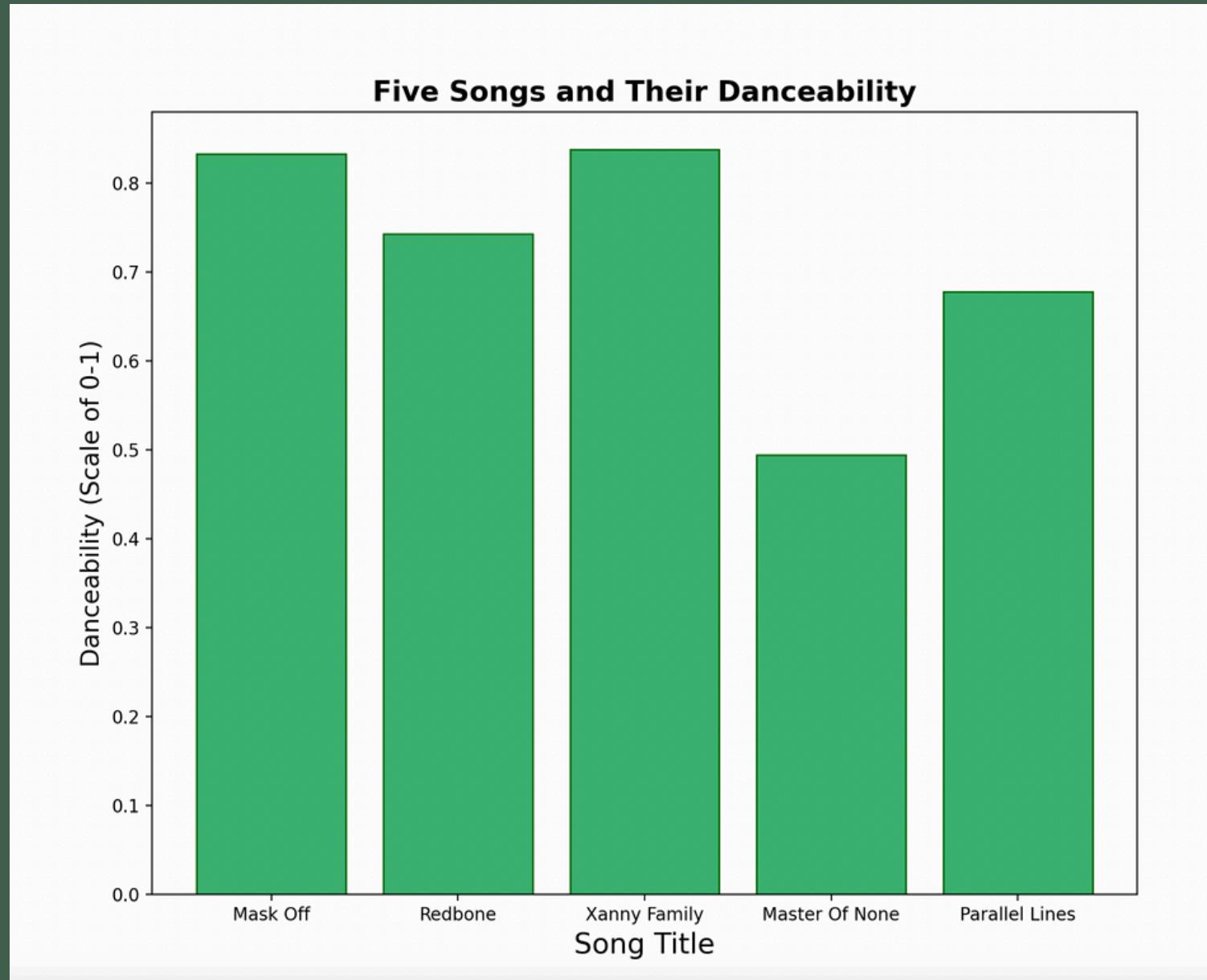


- Barplot
- Demonstrates the top five spotify artists with the most streams
- Drake has the most streams out of any artist, with 50,162,292,808 billion. The number of streams is a negative slope as you move across the x-axis

# Danceability of Five Songs



# Danceability of Five Songs



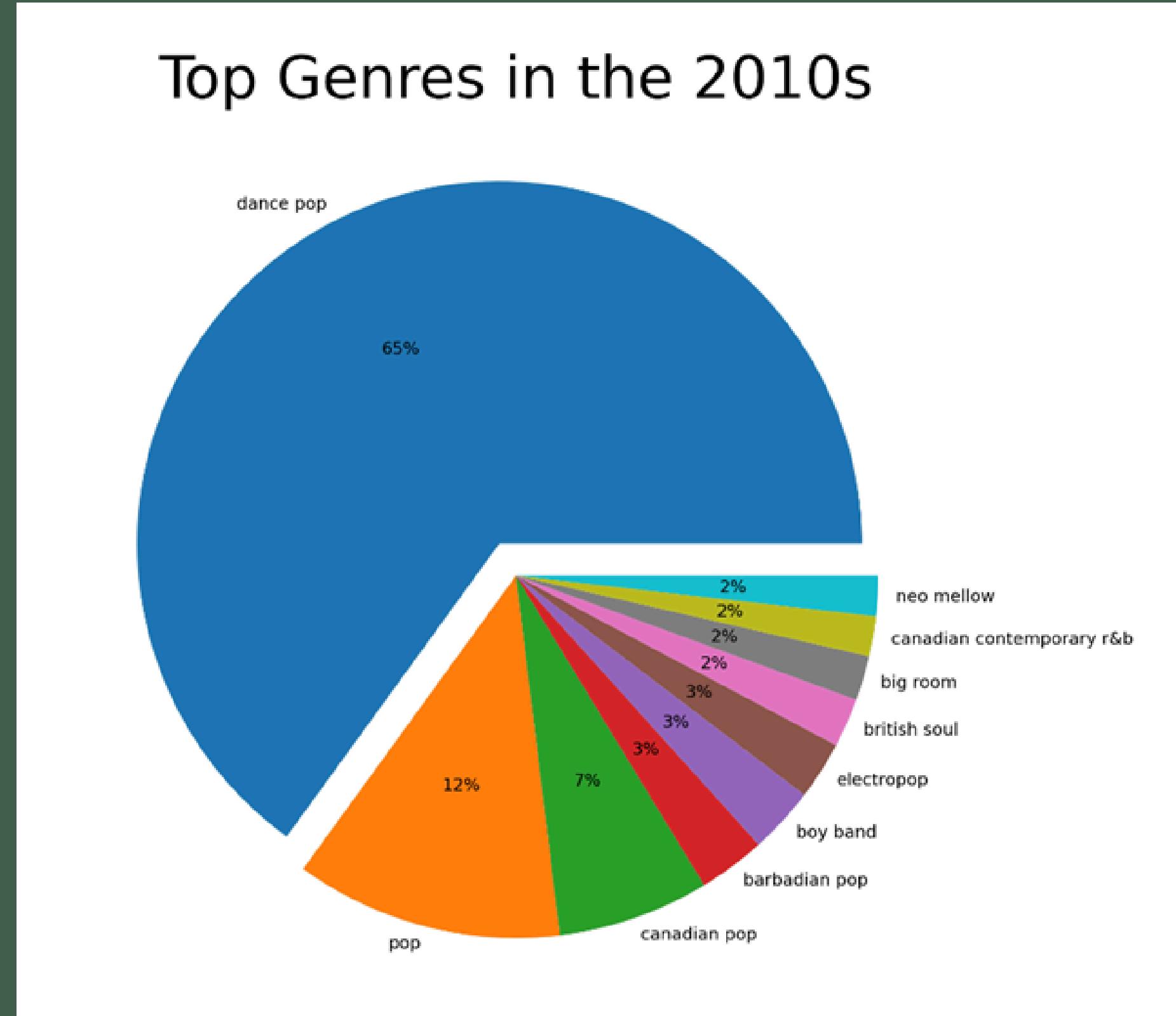
- Barplot
- For this image, I took five very popular songs from different genres and plotted their danceability
- The song with the highest danceability is "Xanny Family" by the rapper Future. Coming in last is "Master of None" by Beach House.

# Top 50s of the 2010s CSV

	title	artist	top genre	year	bpm	nrgy	dnce	dB	live	val	dur	acous	spch	pop
1	Hey, Soul Sister	Train	neo mellow	2010	97	89	67	-4	8	80	217	19	4	83
2	Love The Way You Lie	Eminem	detroit hip hop	2010	87	93	75	-5	52	64	263	24	23	82
3	TiK ToK	Kesha	dance pop	2010	120	84	76	-3	29	71	200	10	14	80
4	Bad Romance	Lady Gaga	dance pop	2010	119	92	70	-4	8	71	295	0	4	79
5	Just the Way You Are	Bruno Mars	pop	2010	109	84	64	-5	9	43	221	2	4	78

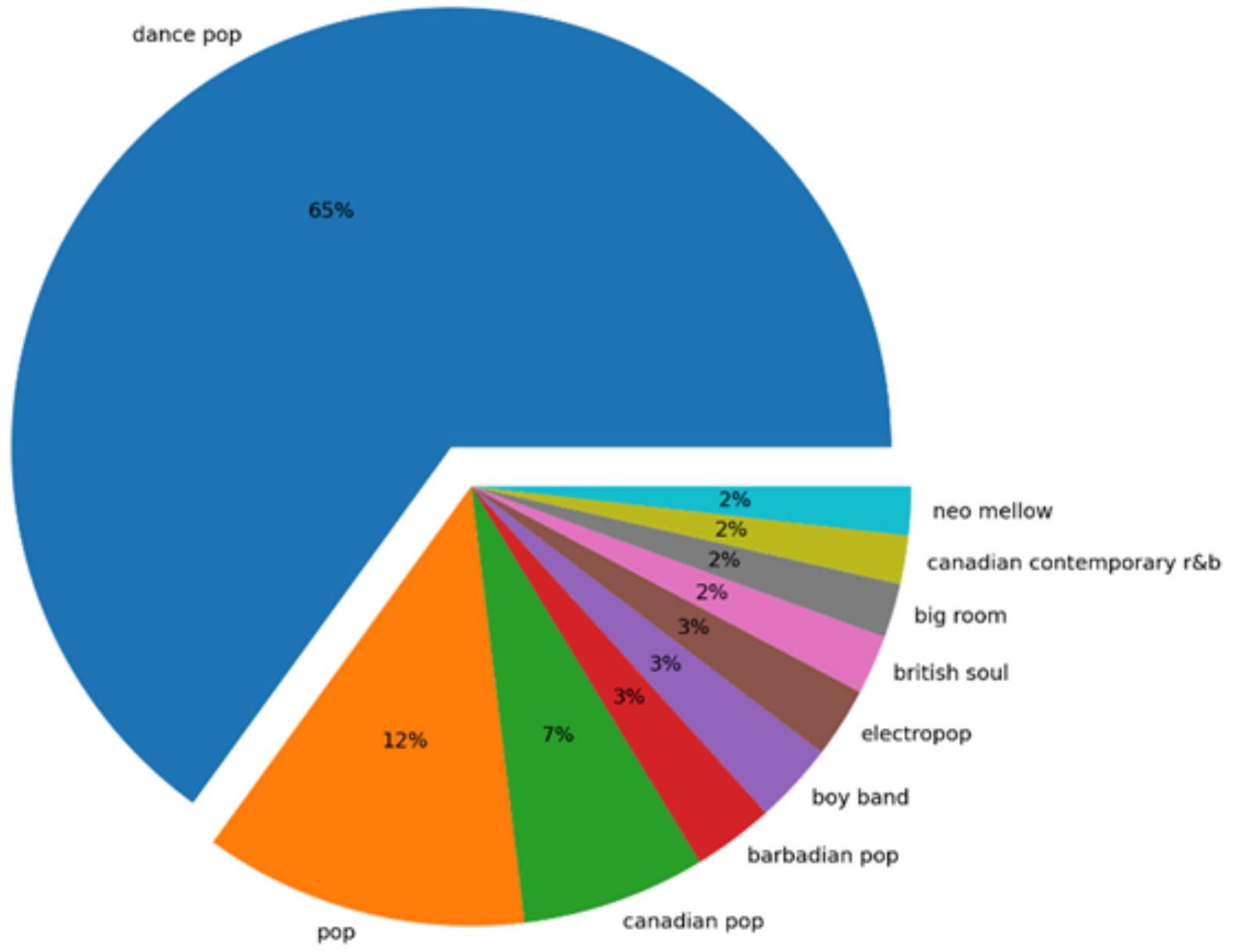
- Top 50 most popular songs on Spotify for each year 2010-2019
- Track title, artist name, and top genre recorded for each song
- Data such as energy, danceability, valence, etc. is also provided in the form of scores

# Top Genres in the 2010s



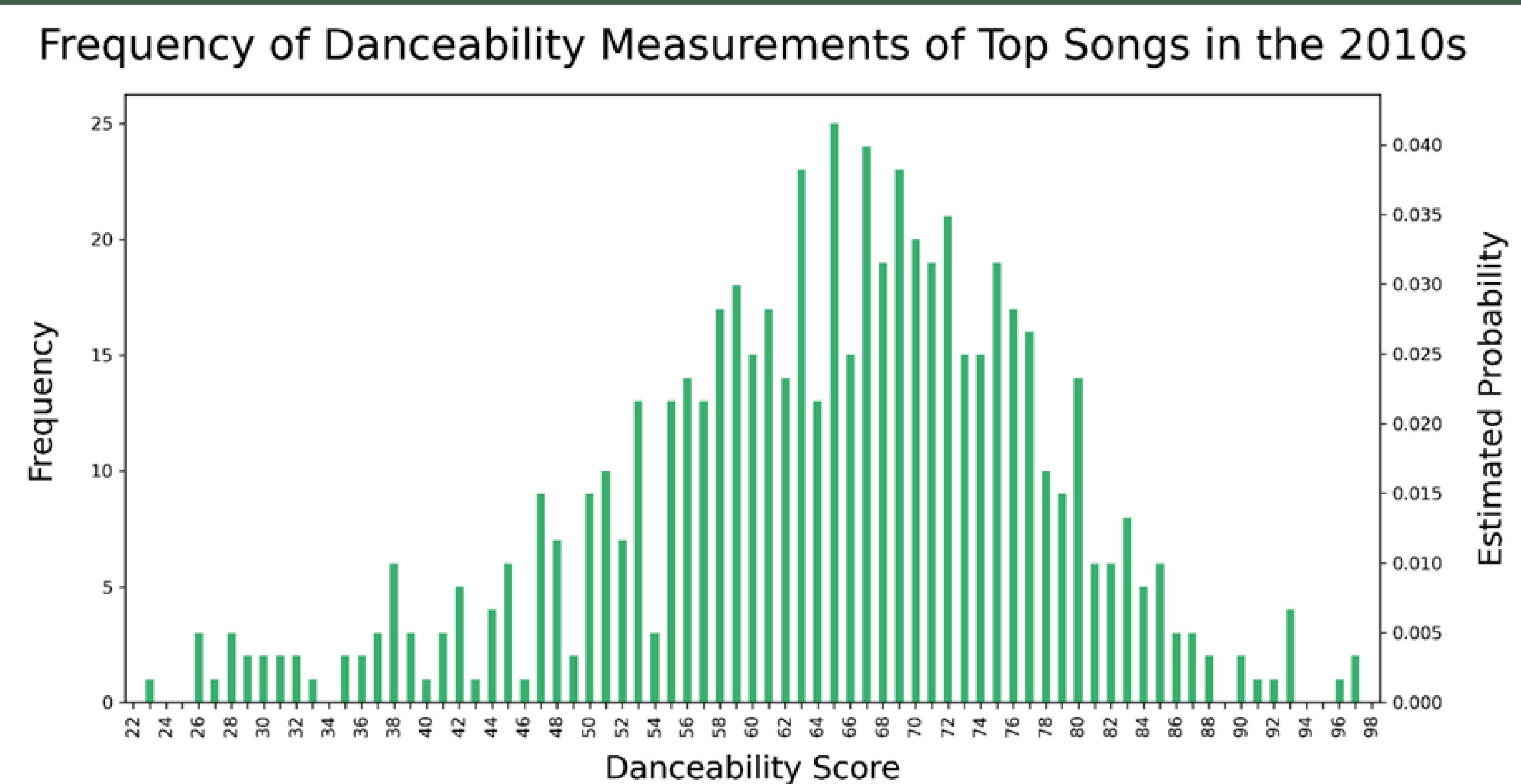
# Top Genres in the 2010s

Top Genres in the 2010s

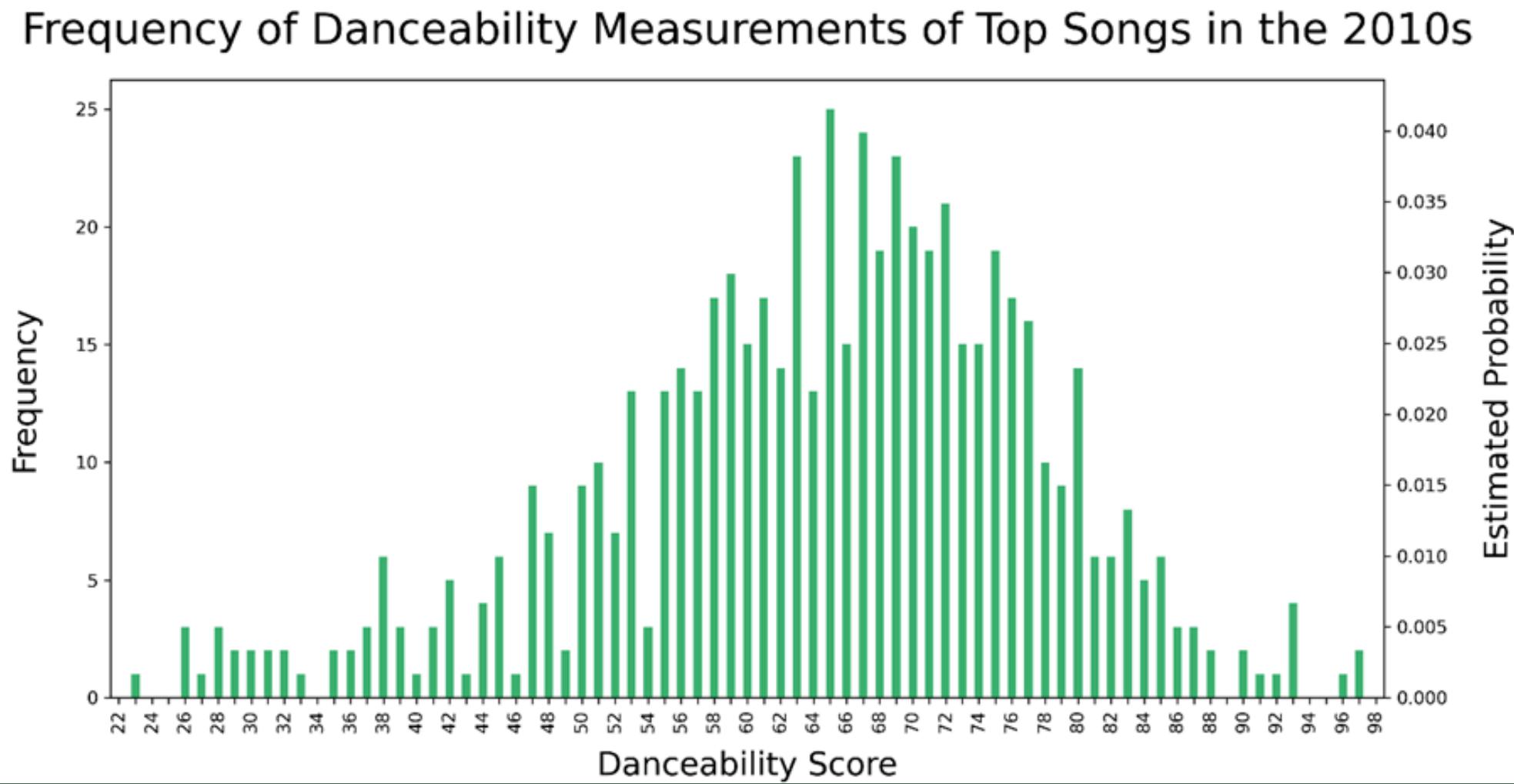


- Pie Chart demonstrating genre popularity
- Most popular genre among popular songs in the 2010s: Dance Pop
- Second, third, and fourth most popular genres also fall under the pop umbrella genre

# Frequency of Danceability Measurements

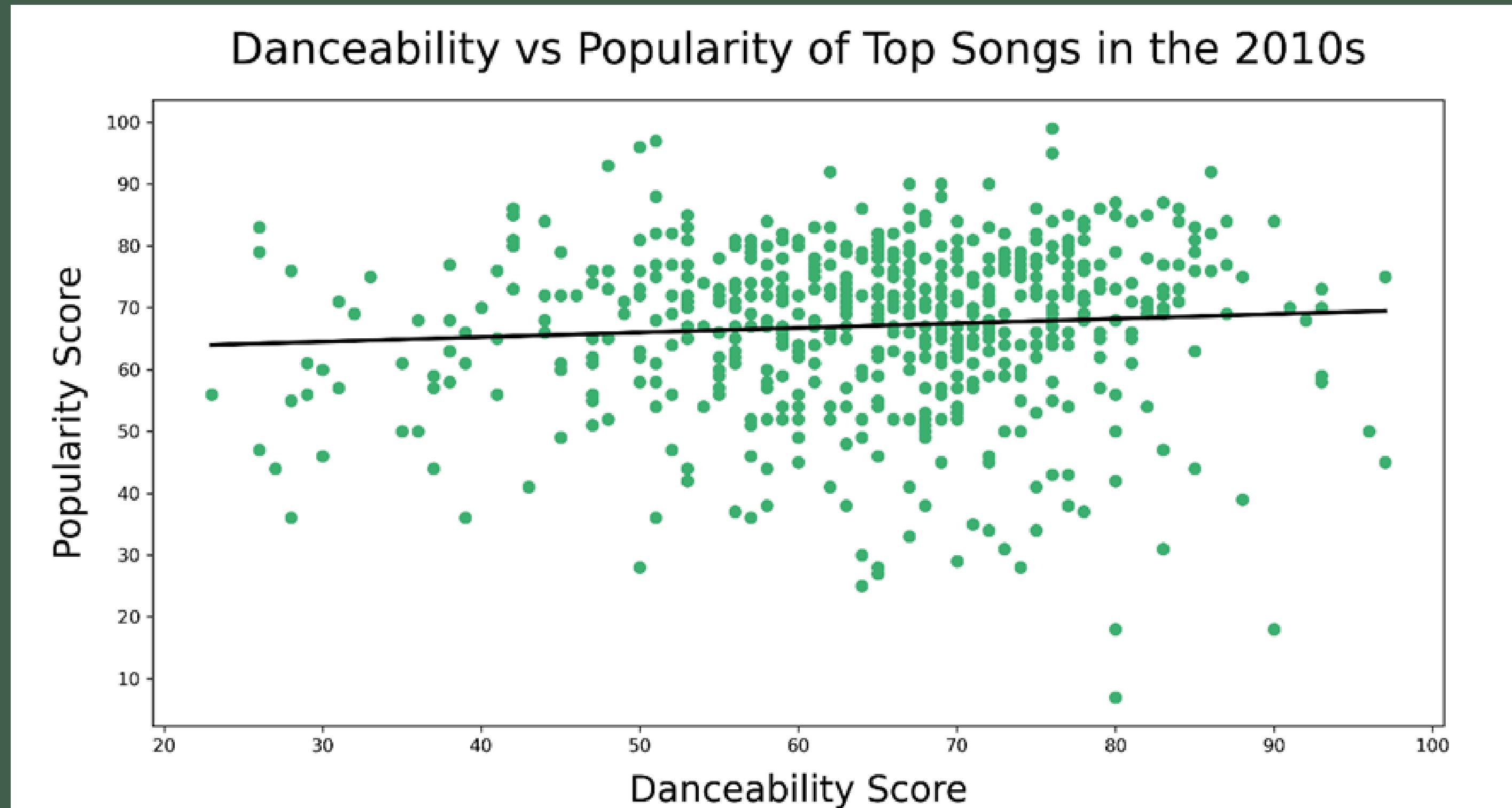


# Frequency of Danceability Measurements



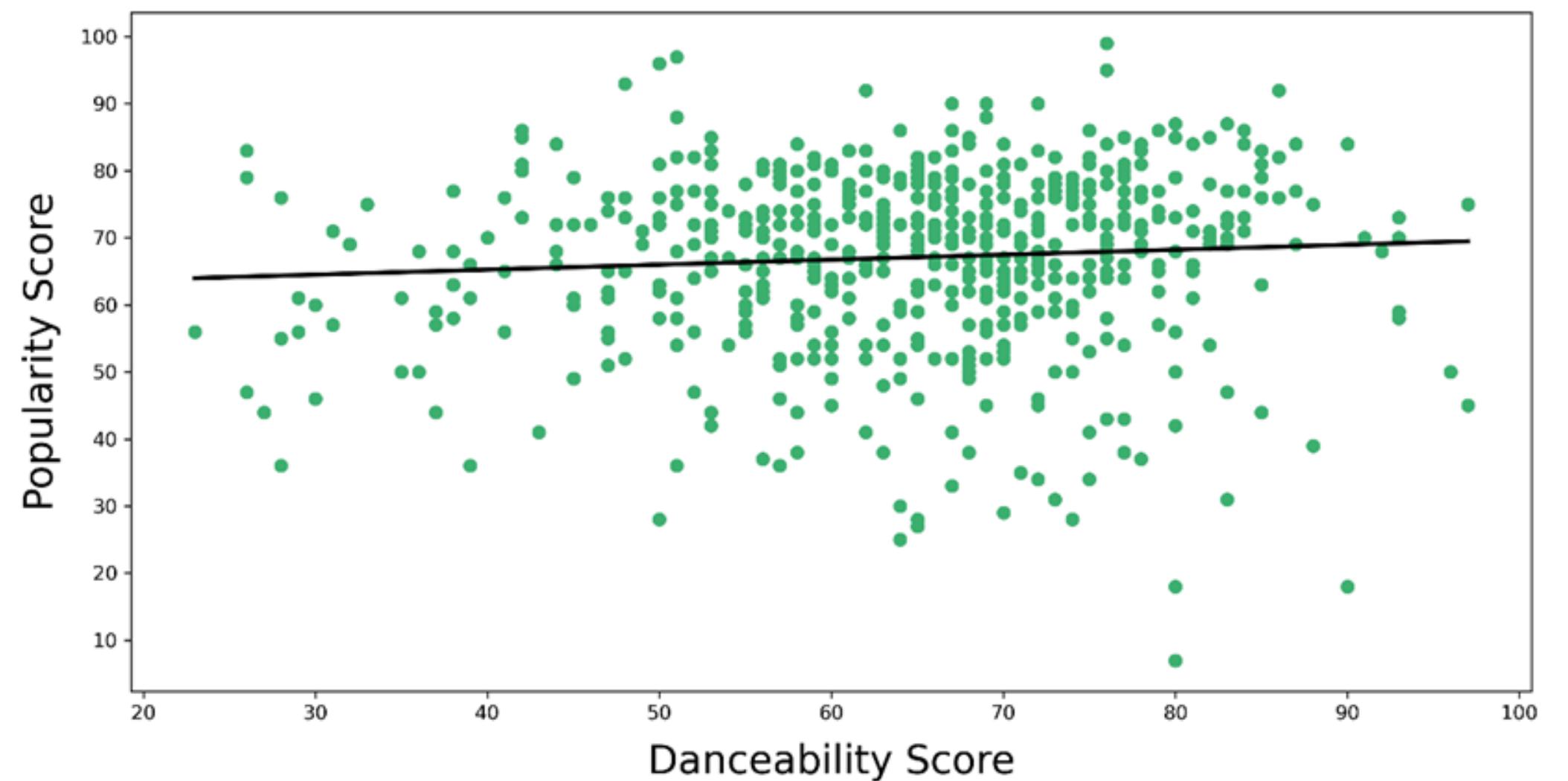
- Frequency chart of Danceability
- Danceability of each song is given a measure from 0 (not danceable) to 100 (very danceable)
- Most frequent scores fall around 65-70 range

# Danceability vs Popularity



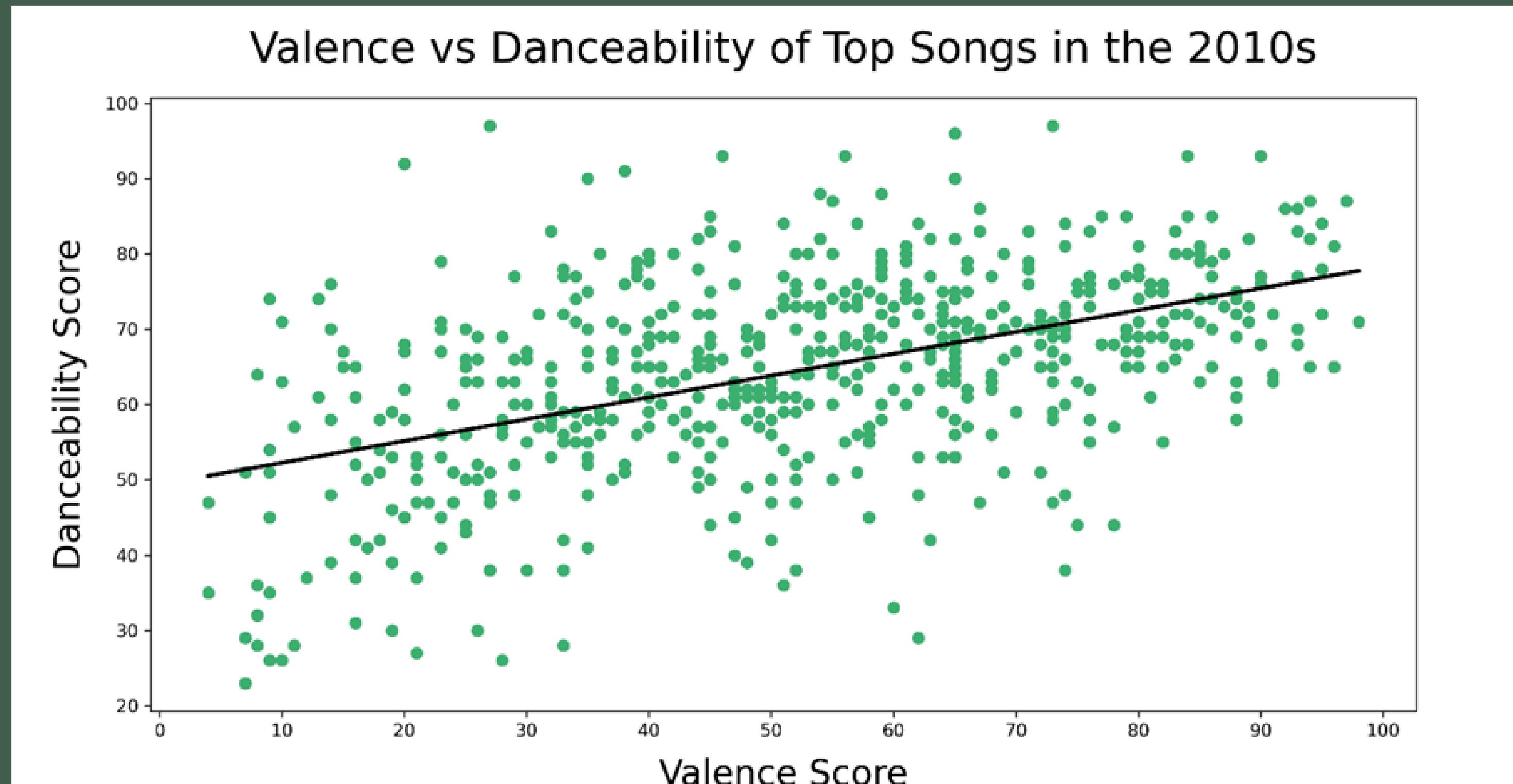
# Danceability vs Popularity

Danceability vs Popularity of Top Songs in the 2010s

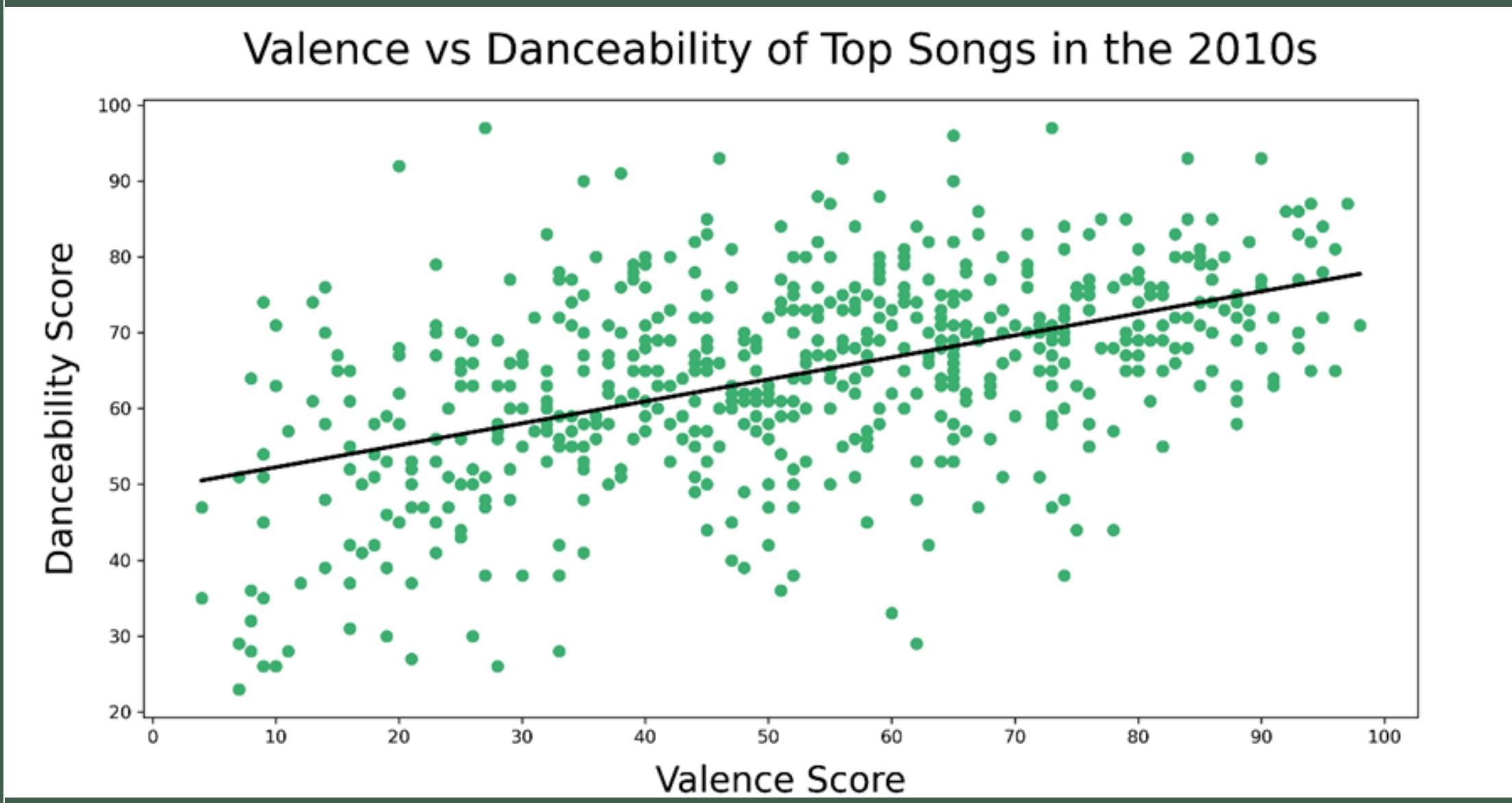


- Danceability vs popularity scatterplot
- Popularity score: 0 (least popular) to 100 (most popular)
- Slight positive correlation: 0.07

# Valence vs Danceability



# Valence vs Danceability



- Valence vs Danceability scatterplot
- Valence: musical positiveness
- Valence score: 0 (least positive) to 100 (most positive)
- More notable positive correlation: 0.34

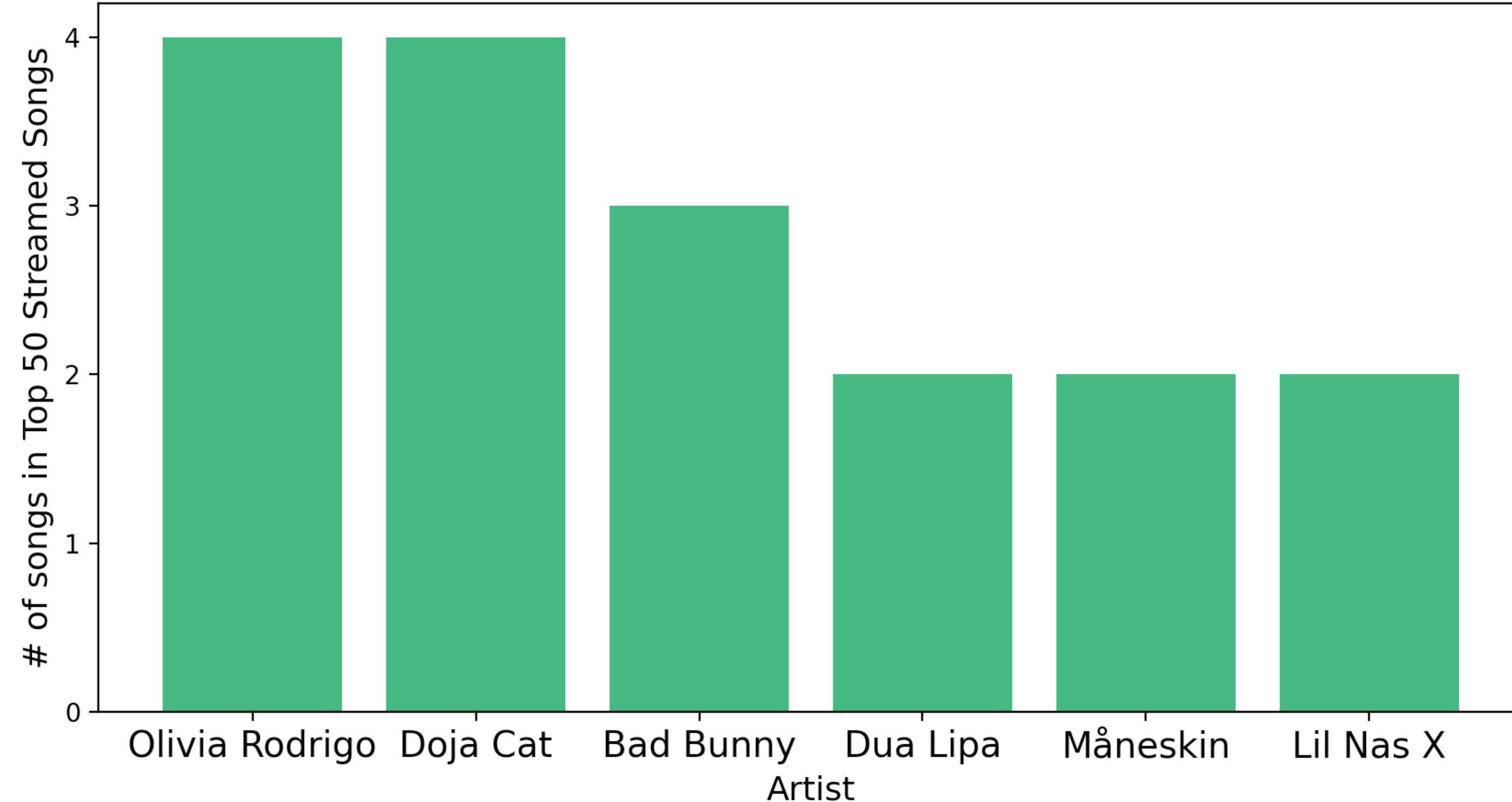
# Spotify Top 50 2021 CSV

<b>id</b>	<b>artist_name</b>	<b>track_name</b>	<b>track_id</b>	<b>popularity</b>	<b>danceability</b>	<b>energy</b>	<b>key</b>	<b>loudness</b>	<b>mode</b>	<b>speechiness</b>
<b>1</b>	Olivia Rodrigo	drivers license	5wANPM4fQCJwkGd4rN57mH	92	0.561	0.431	10	-8.81	1	0.0578
<b>2</b>	Lil Nas X	MONTERO (Call Me By Your Name)	1SC5rEoYDGUK4NfG82494W	90	0.593	0.503	8	-6.725	0	0.22
<b>3</b>	The Kid LAROI	STAY (with Justin Bieber)	5PjdY0CKGZdEuNab3yDmX	92	0.591	0.764	1	-5.484	1	0.0483
<b>4</b>	Olivia Rodrigo	good 4 u	4ZtFanR9U6ndgddUvNcjcG	95	0.563	0.664	9	-5.044	1	0.154
<b>5</b>	Dua Lipa	Levitating (feat. DaBaby)	5nujrmhLynf4yMoMtj8AQF	89	0.702	0.825	6	-3.787	0	0.0601

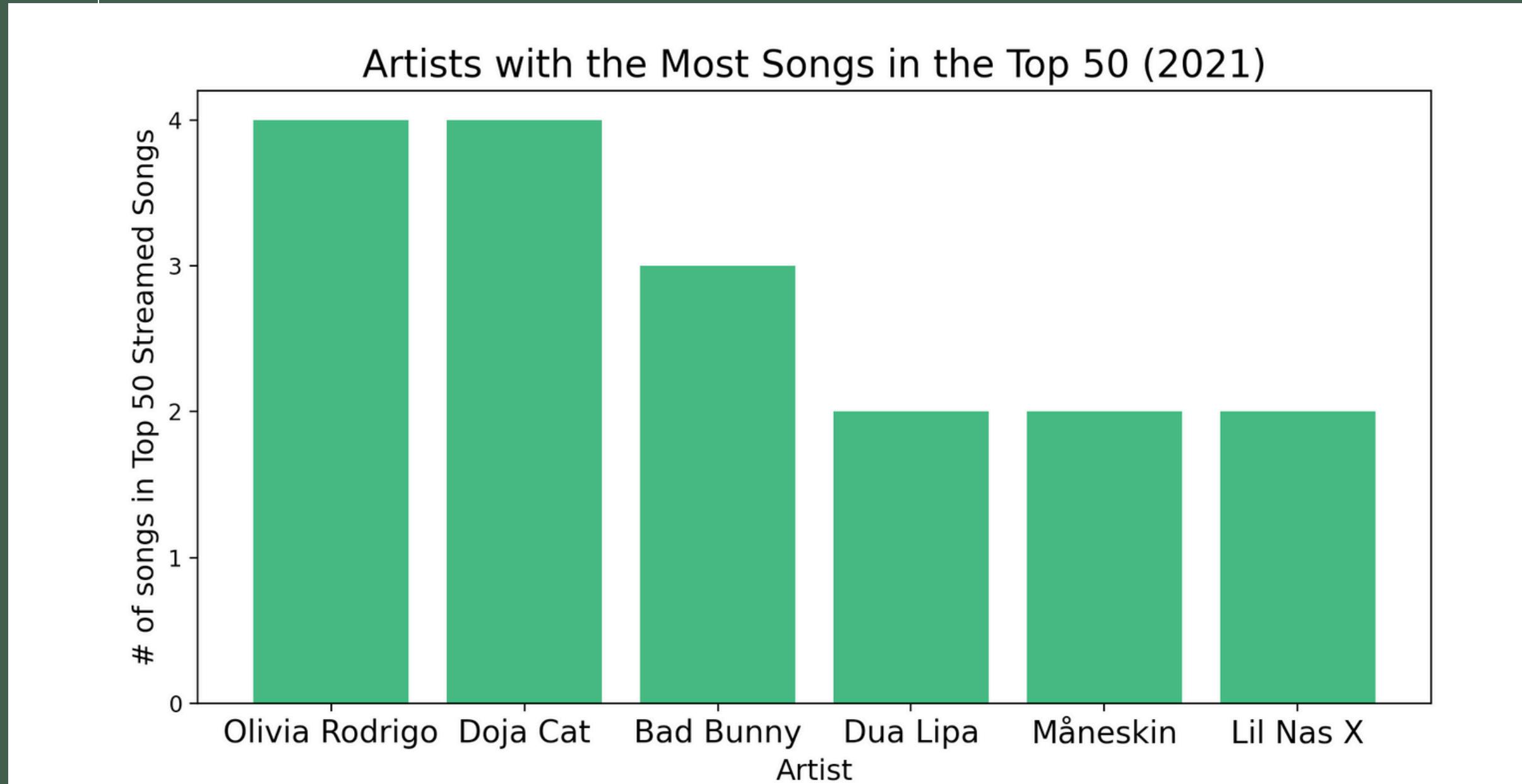
- A dataset of the top 50 streamed songs in 2021 with attributes from Spotify's API
  - Contains song name, track information, and song attributes
  - The 14 track attributes: popularity, danceability, energy, key, loudness, mode, speechiness, acousticness, instrumentalness, liveness, valence, tempo, duration (ms), time signature

# Top Artists of 2021

Artists with the Most Songs in the Top 50 (2021)

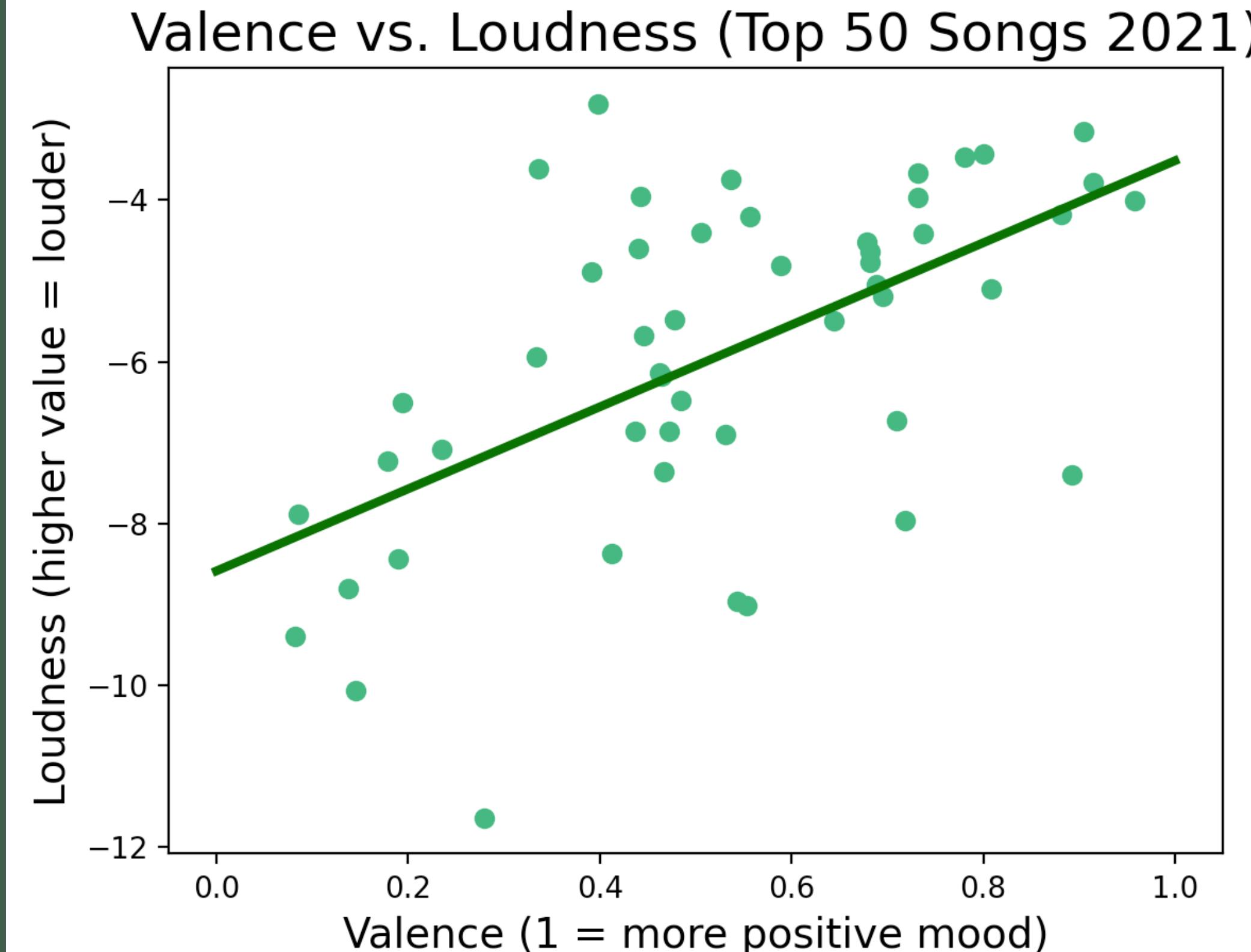


# Top Artists of 2021

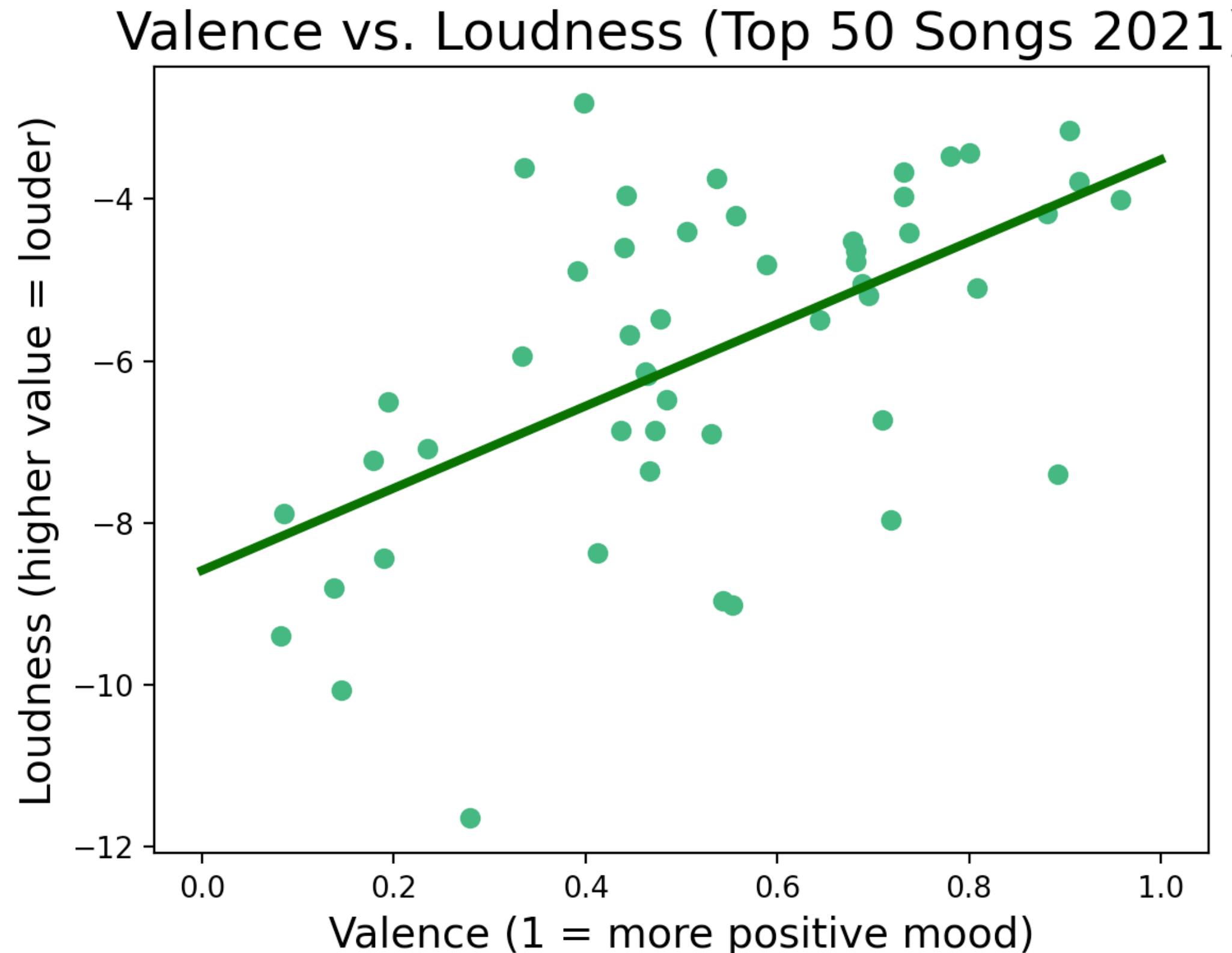


- Top 6 Artists with the most songs in the Top 50 (2021)
- Top 3 Artists account for 22% of songs in the top 50
- Top 6 Artists account for 34% of songs in the top 50

# Valence vs. Loudness (2021)



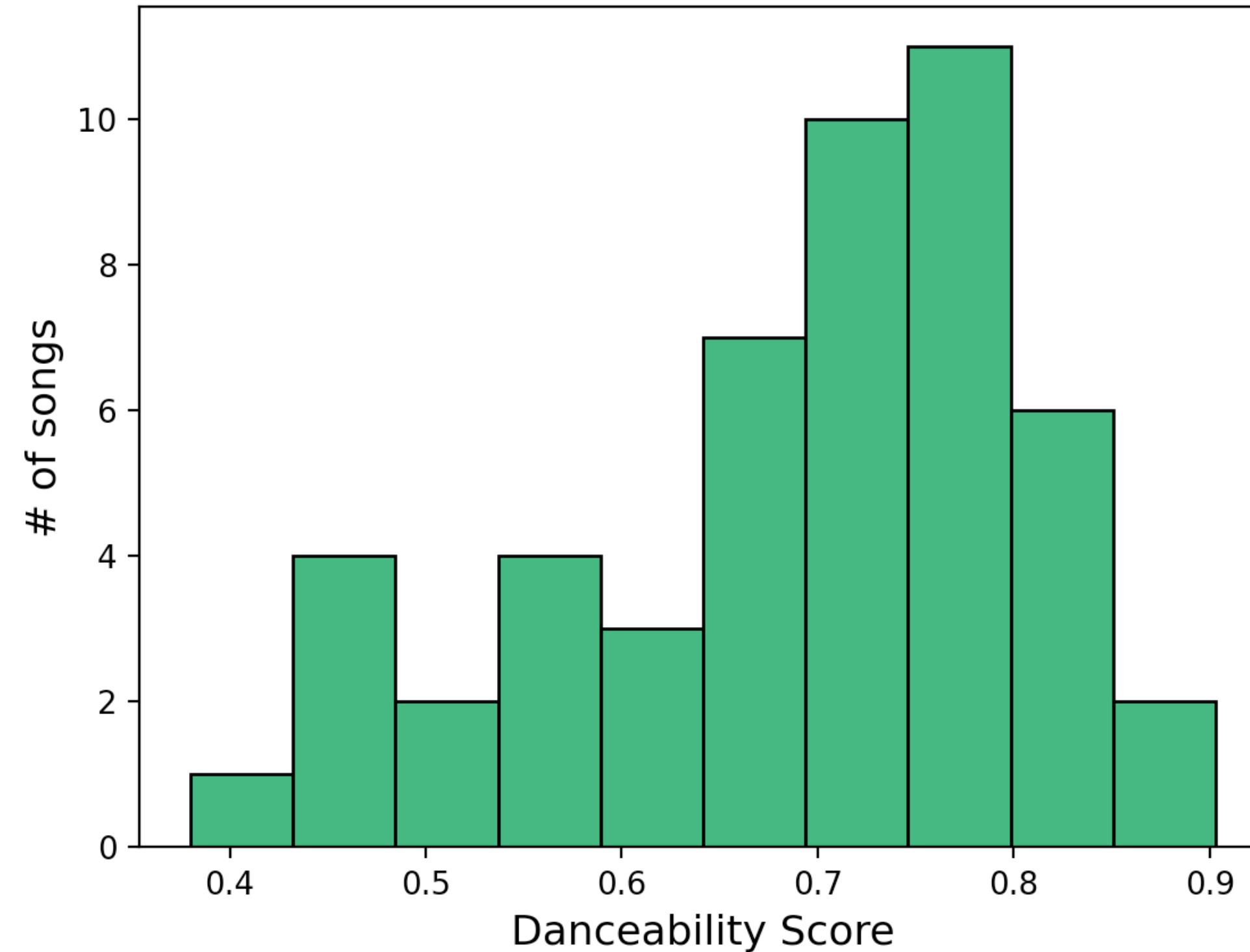
# Valence vs. Loudness (2021)



- Valence: measure of a song's positivity (mood)
- Valence score closer to 1 indicates a 'positive' song
- Moderate positive linear correlation ( $r^2$ : 0.628)

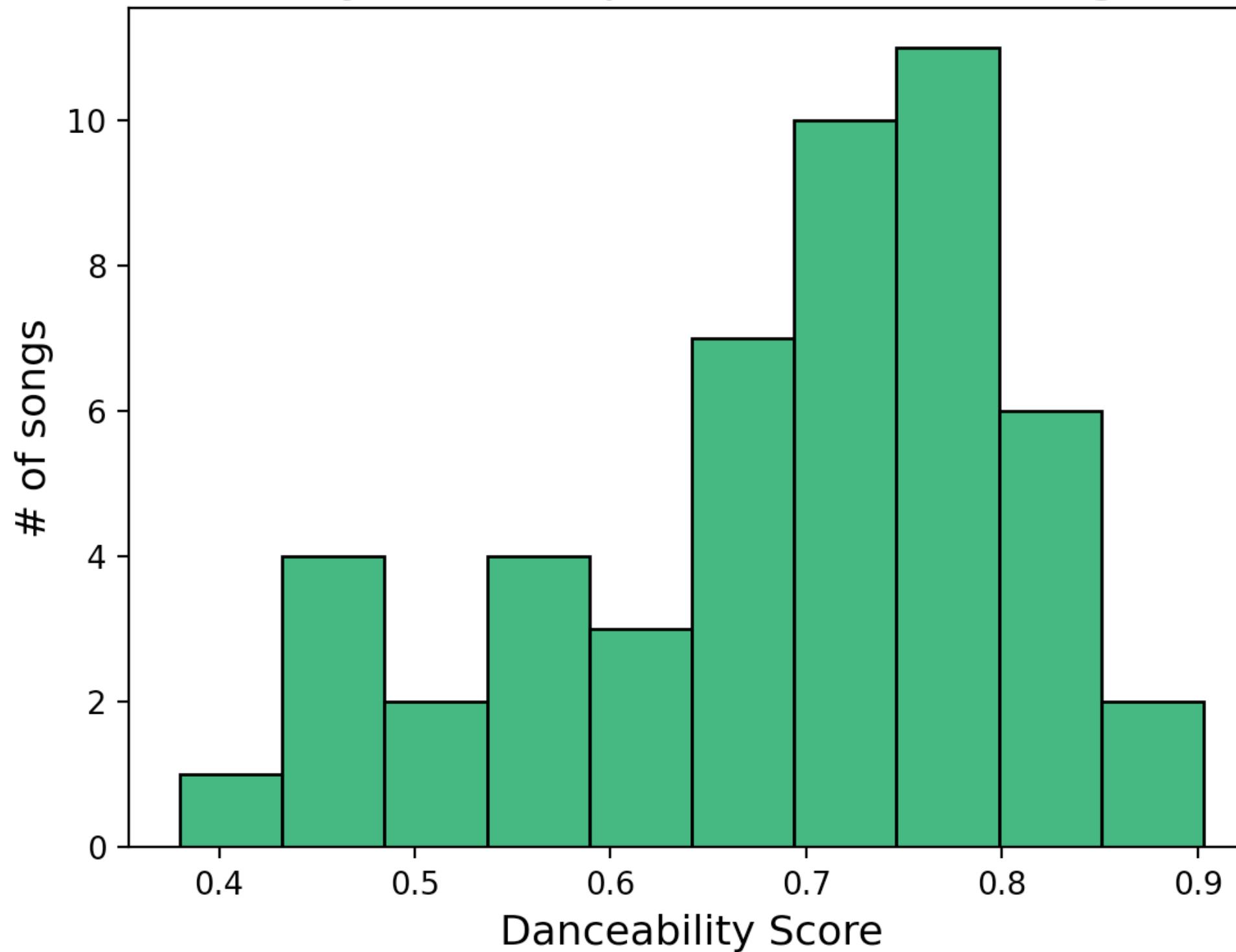
# Danceability Distribution (2021)

Danceability of the Top 50 Streamed Songs (2021)



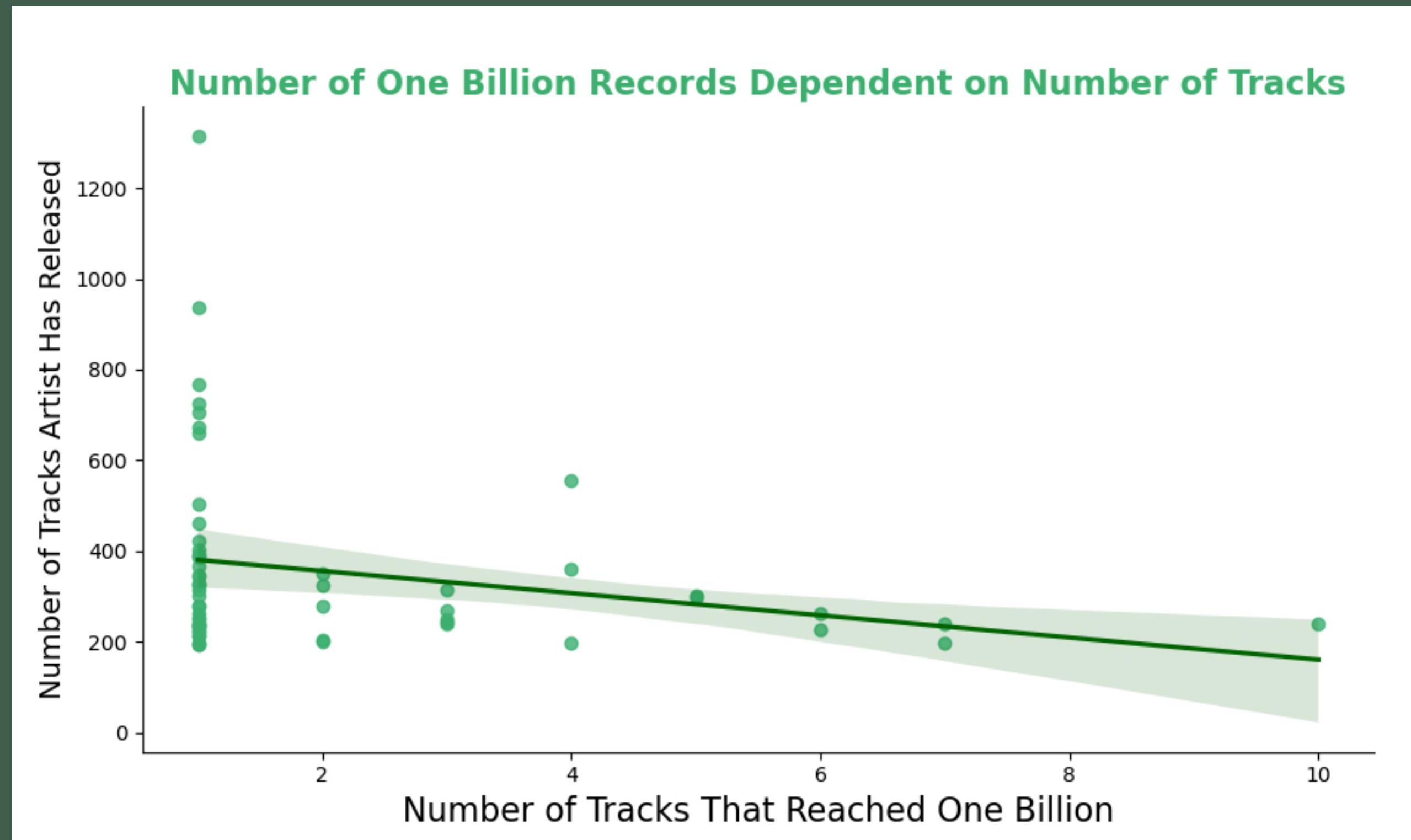
# Danceability Distribution (2021)

Danceability of the Top 50 Streamed Songs (2021)

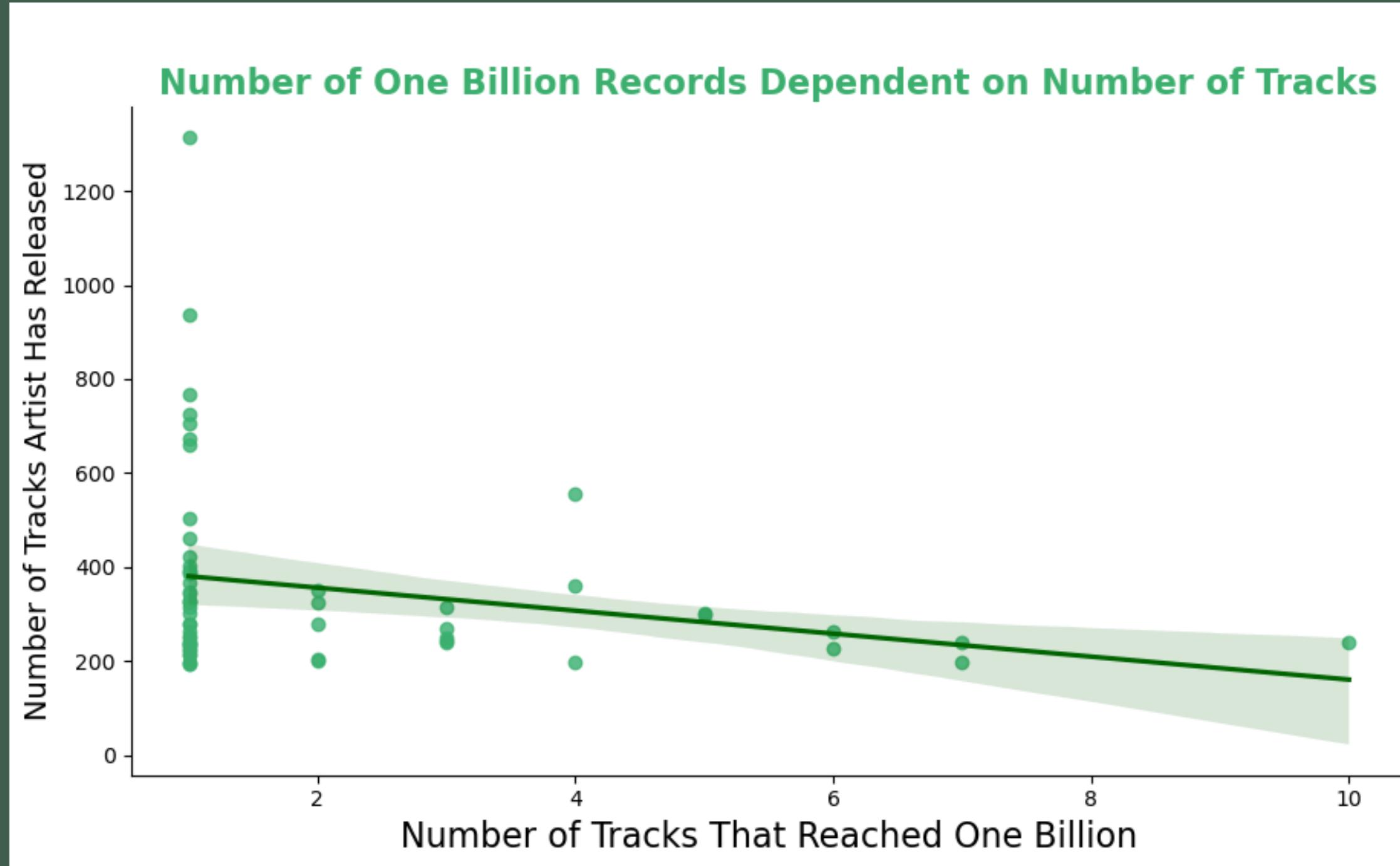


- A danceability score closer to 1 means a song is more danceable
- Left skewed
- Most songs in the top 50 have a relatively high danceability score

# One Billion vs Tracks

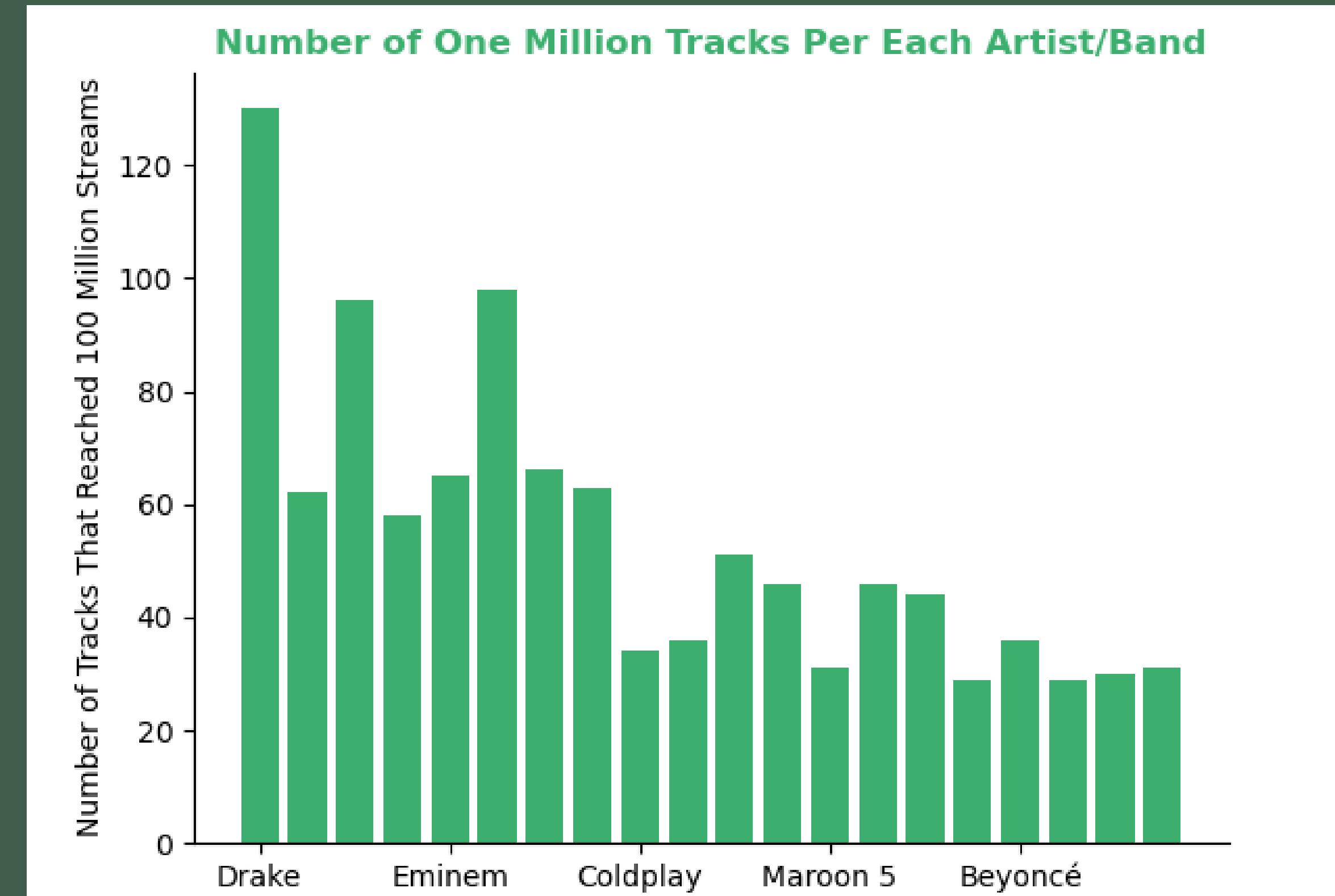


# One Billion vs Tracks

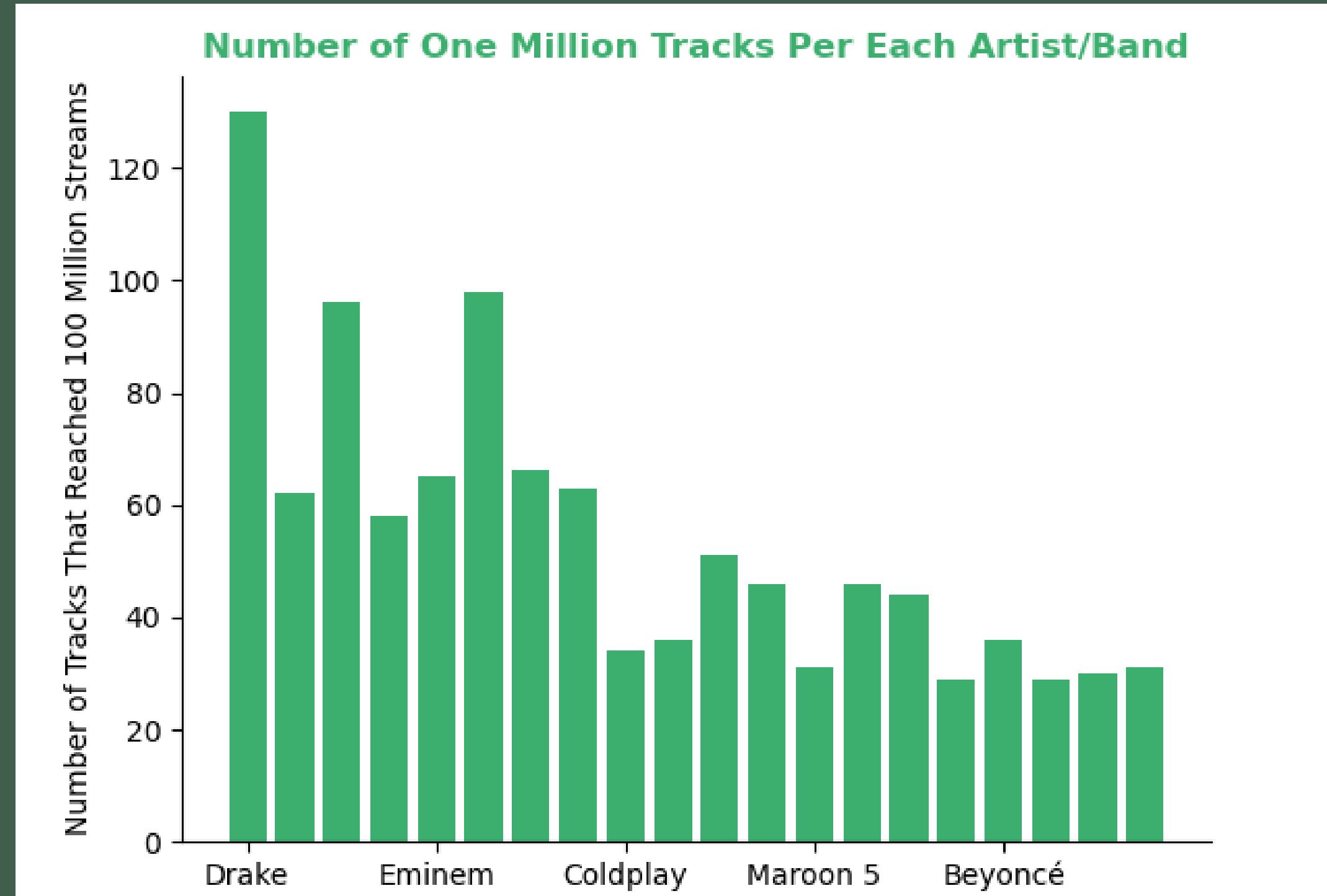


- Scatterplot of number of tracks that reach One Billion streams vs how many songs an artist has produced.
- Modified data for better graph
- Negative trend line
- Correlation Is -0.20
- Ed Sheeran and David Bowie are two outliers

# Number of Tracks that reached 100 Million

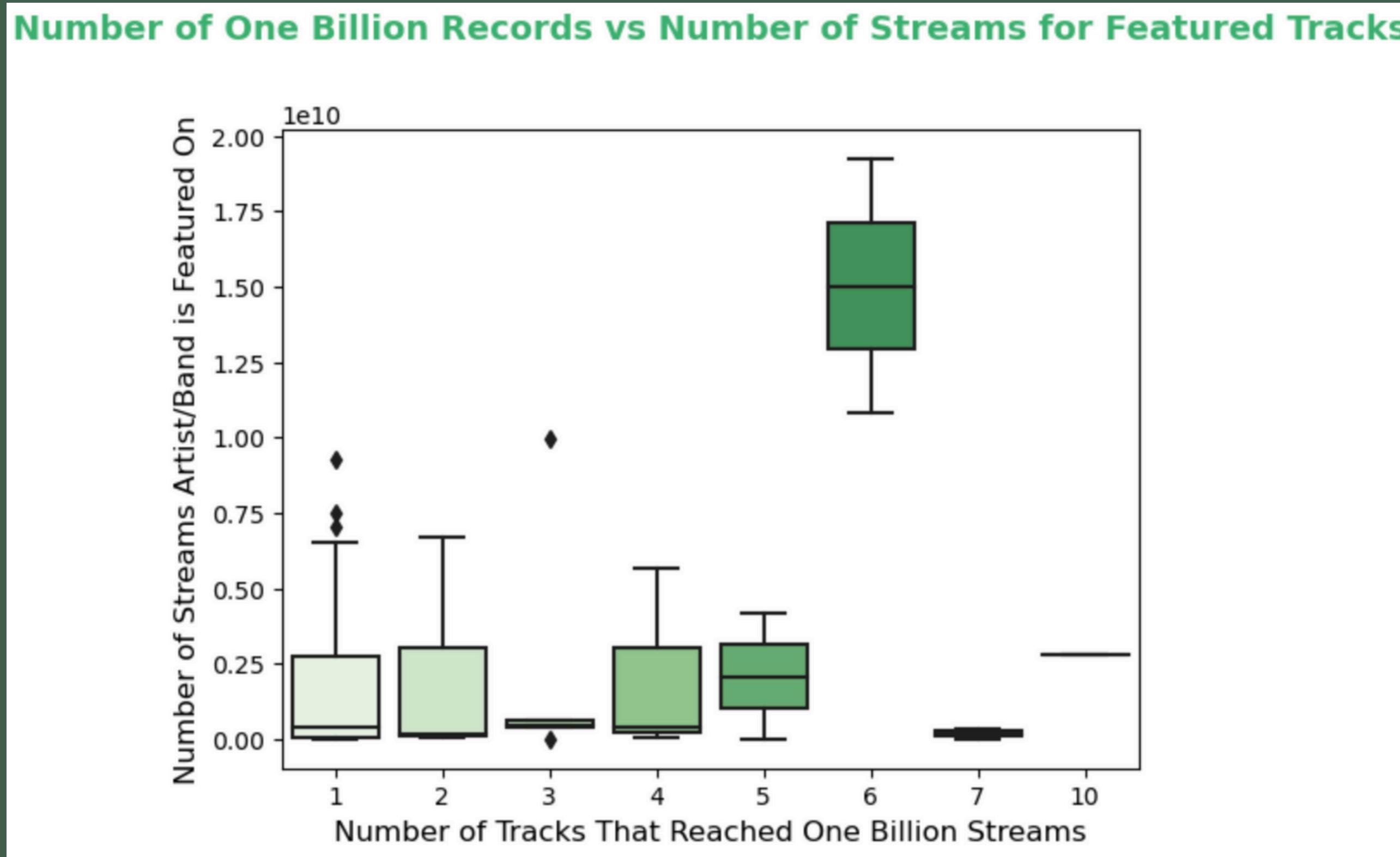


# Number of Tracks that reached 100 Million

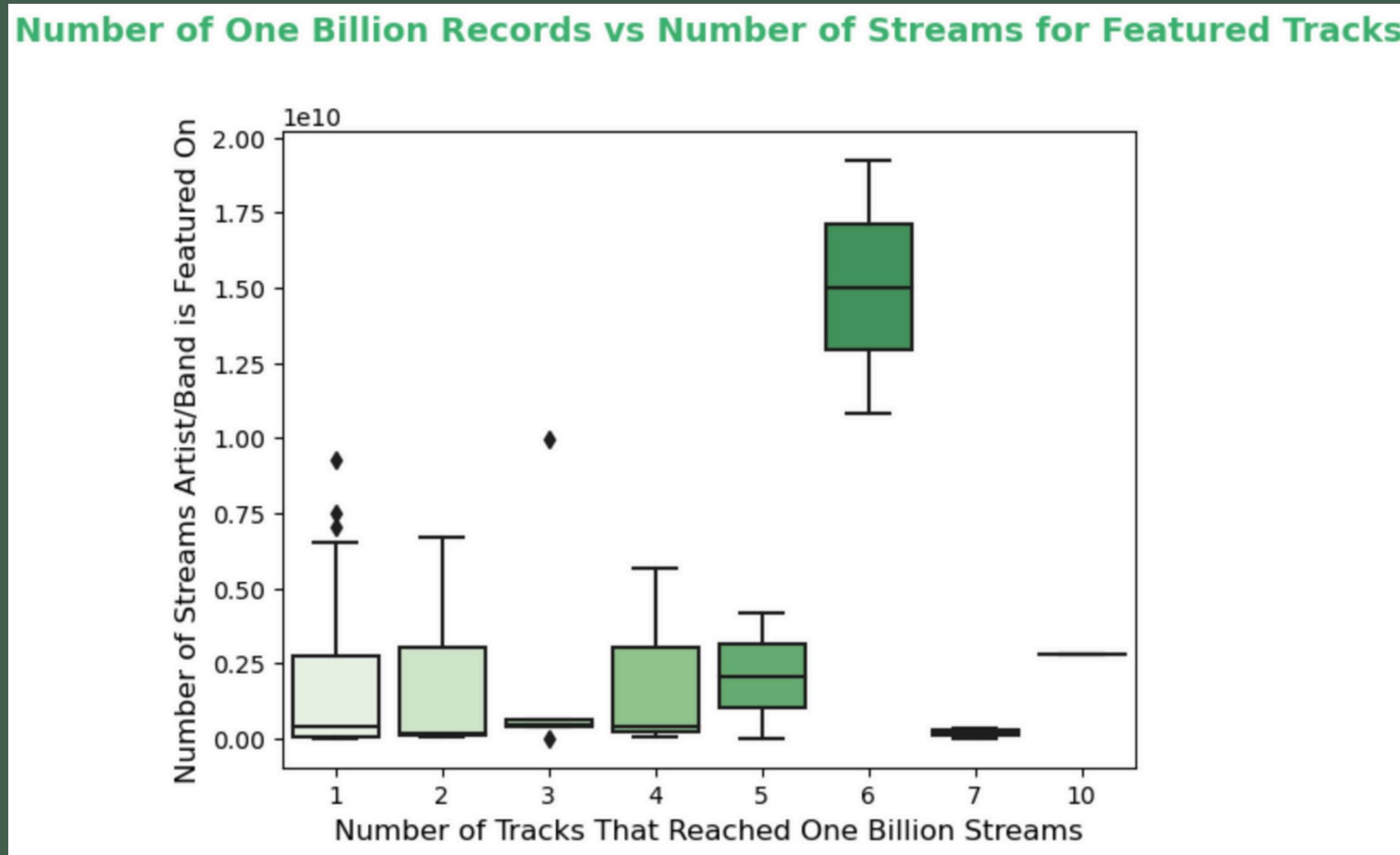


- Barplot showing how many tracks an artist has that has reached 100 Million streams
- Modified data for better graph
- Negative correlation of -0.06
- Drake leads artist (again) with 130 tracks that have passed 100M streams

# One Billion vs Features

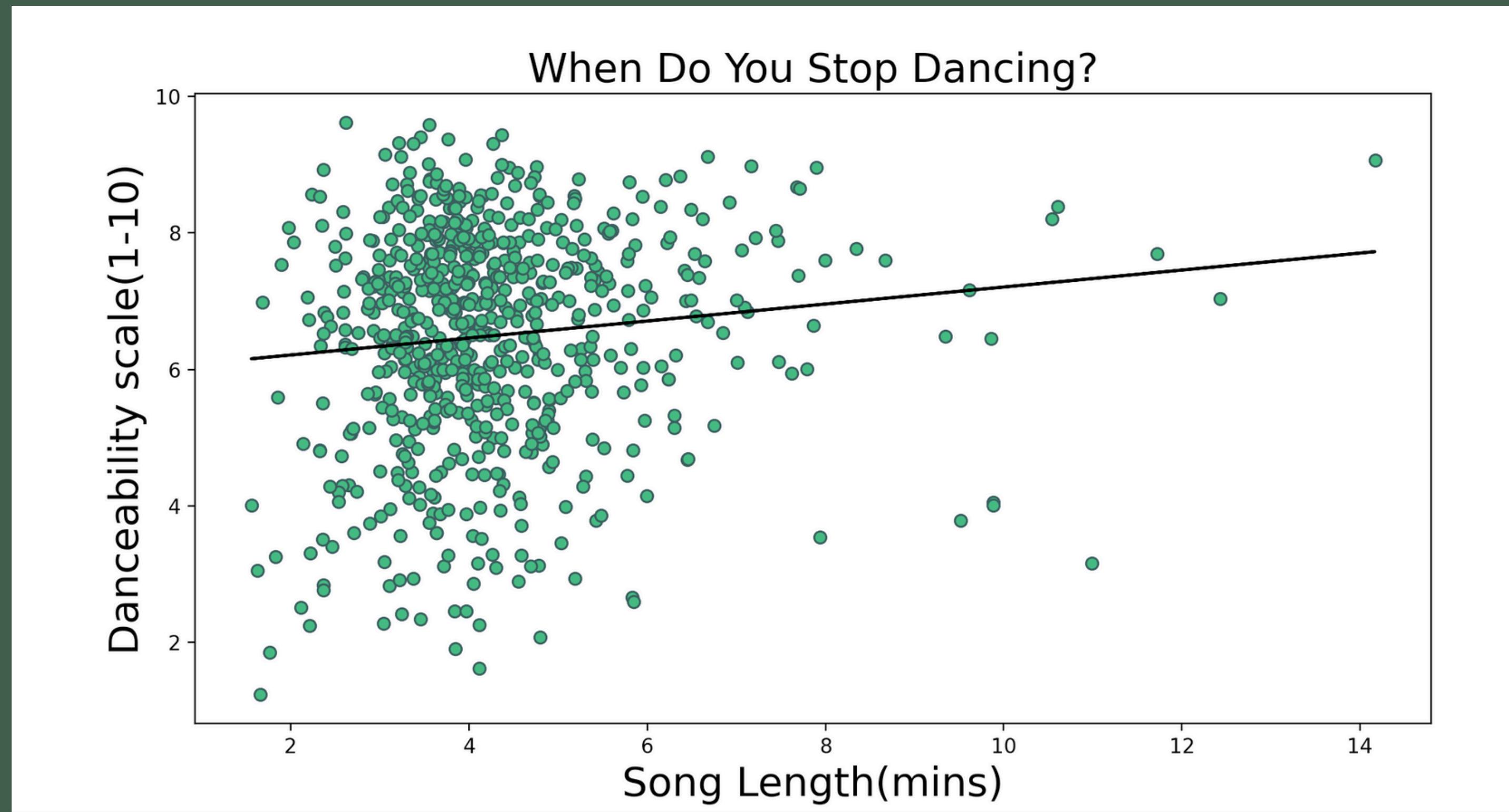


# One Billion vs Features

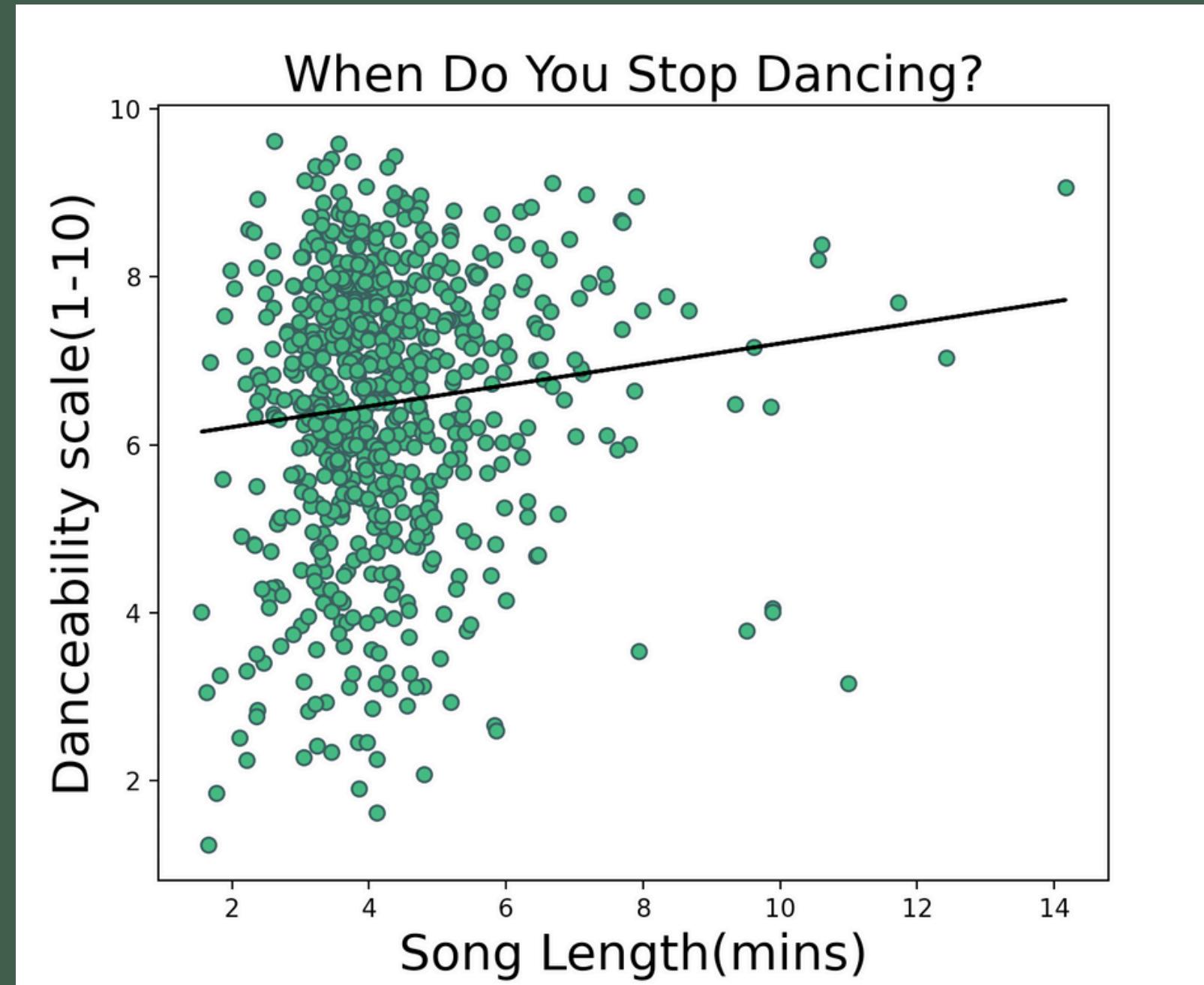


- Box plot tracking the number of records that reached One Billion streams compared to the total amount of streams these featured songs are listened to.
- Y-Axis Is measured In the value of ten billion
- Positive correlation of 0.06

# Danceability vs Length



# Danceability vs Length



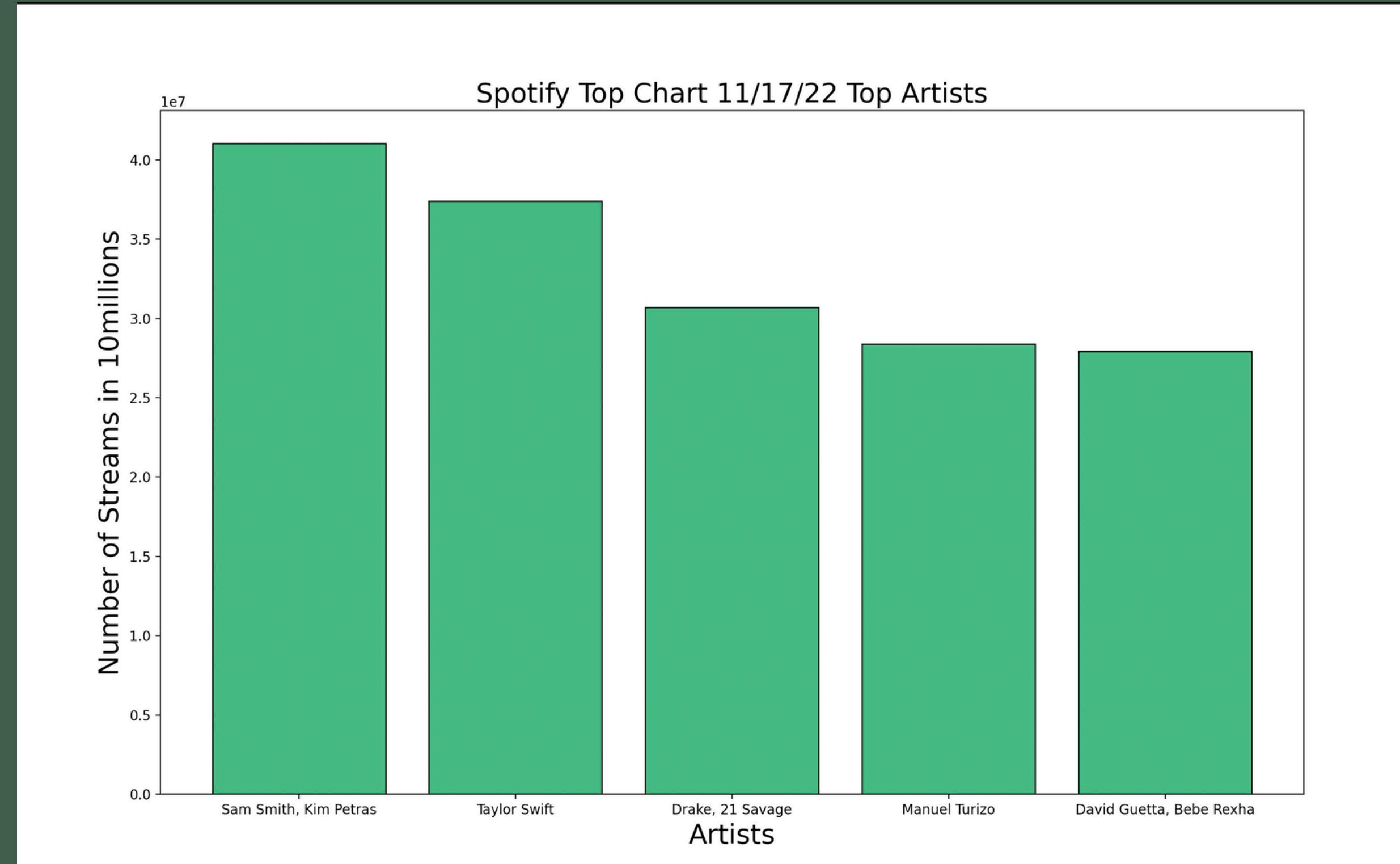
- Scatterplot of Danceability vs song length
- Data from Song Attributes CSV
- Danceability was scored 1-10
- Song length was measured in minutes
- Positive trend line
- Longest=16.64 mins
- Shortest=16 seconds

# Spotify Chart Data CSV

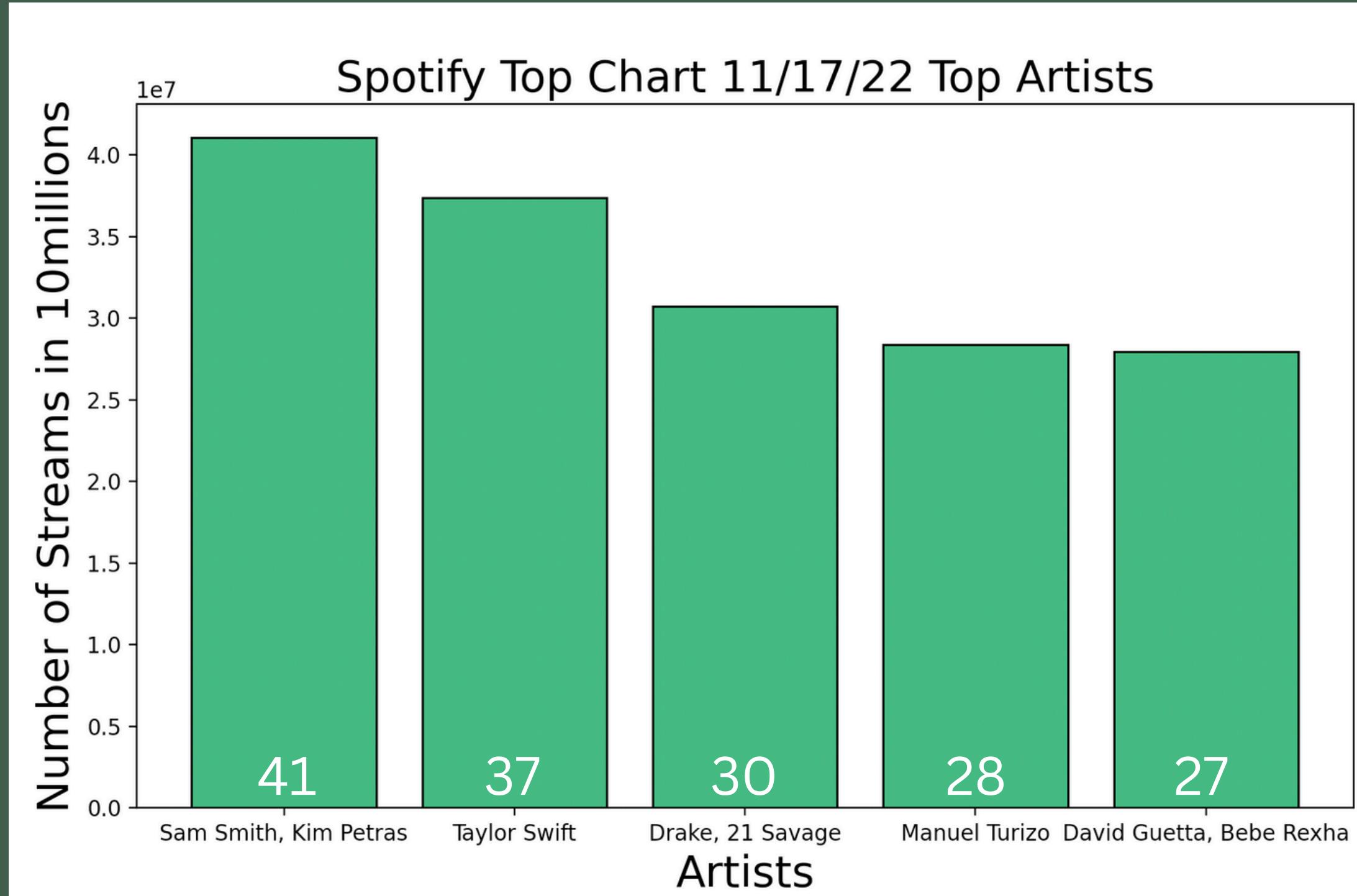
artist_names	track_name	source	peak_rank	previous_rank	weeks_on_chart	streams
Sam Smith, Kim Petras	Unholy (feat. Kim Petras)	EMI	1	1	9	41043843
Taylor Swift	Anti-Hero	Taylor Swift	1	2	5	37394162
Drake, 21 Savage	Rich Flex	OVO / Republ	2	3	3	30709125
Manuel Turizo	La Bachata	Sony Music L	4	4	21	28388964
David Guetta, Bebe Rexha	I'm Good (Blue)	Parlophone L	2	5	13	27930483
Harry Styles	As It Was	Columbia	1	6	34	27653264
Bizarrap, Quevedo	Quevedo: Bzrp Music Sessions, Vol. 52	DALE PLAY R	1	7	20	25922349
Meghan Trainor	Made You Look	Epic	8	18	5	21847336
Clara Benin	Hold The Light	Clara Benin	6	9	10	21620546

- Data comes from the Spotify Charts website
- Updates every Thursday
- Chart is from the week of 11/17/22.
- Has data such as rank, artist name, track name, streams, and more

# Top Artist Streams 11/17

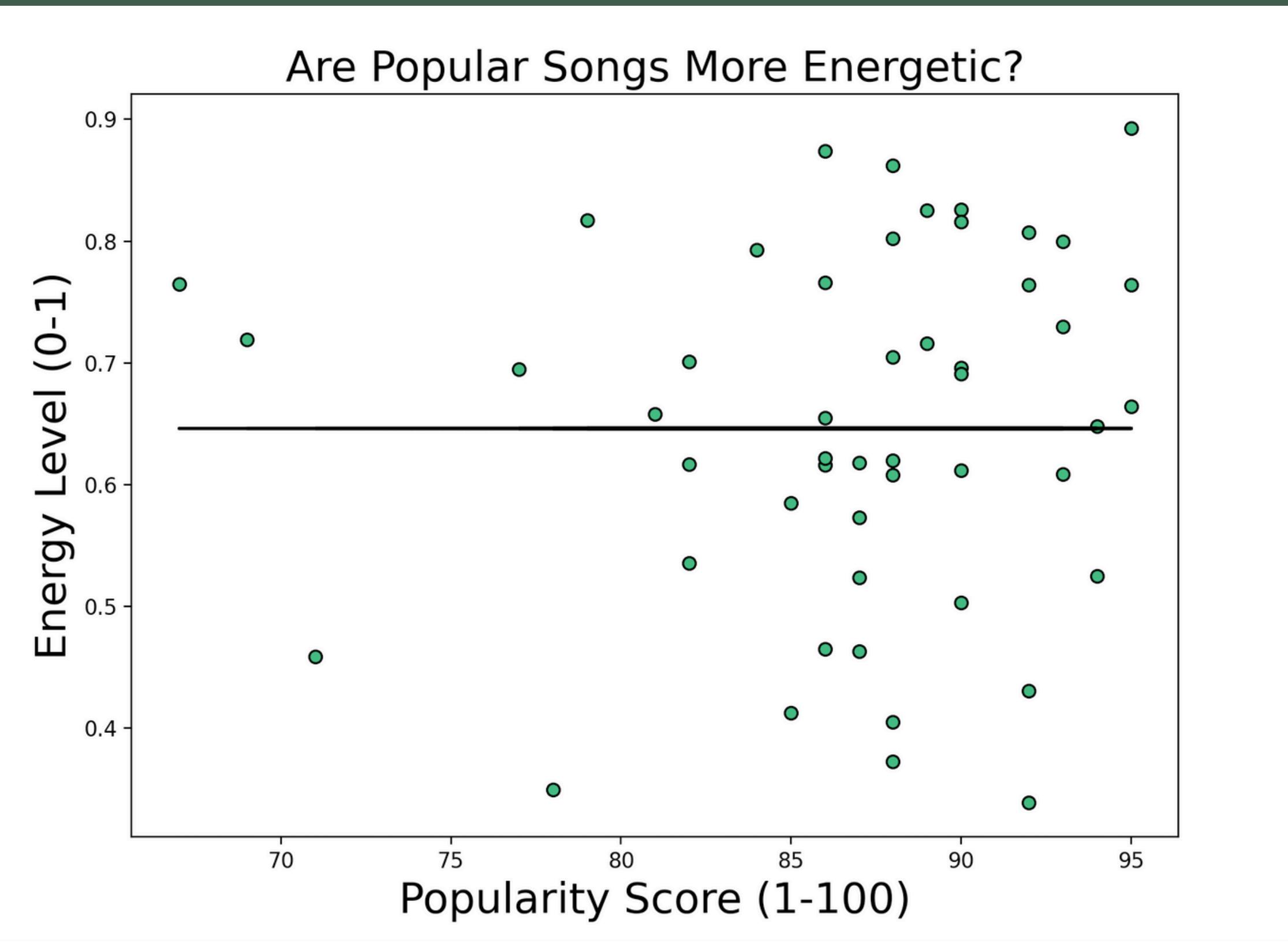


# Top Artist Streams 11/17

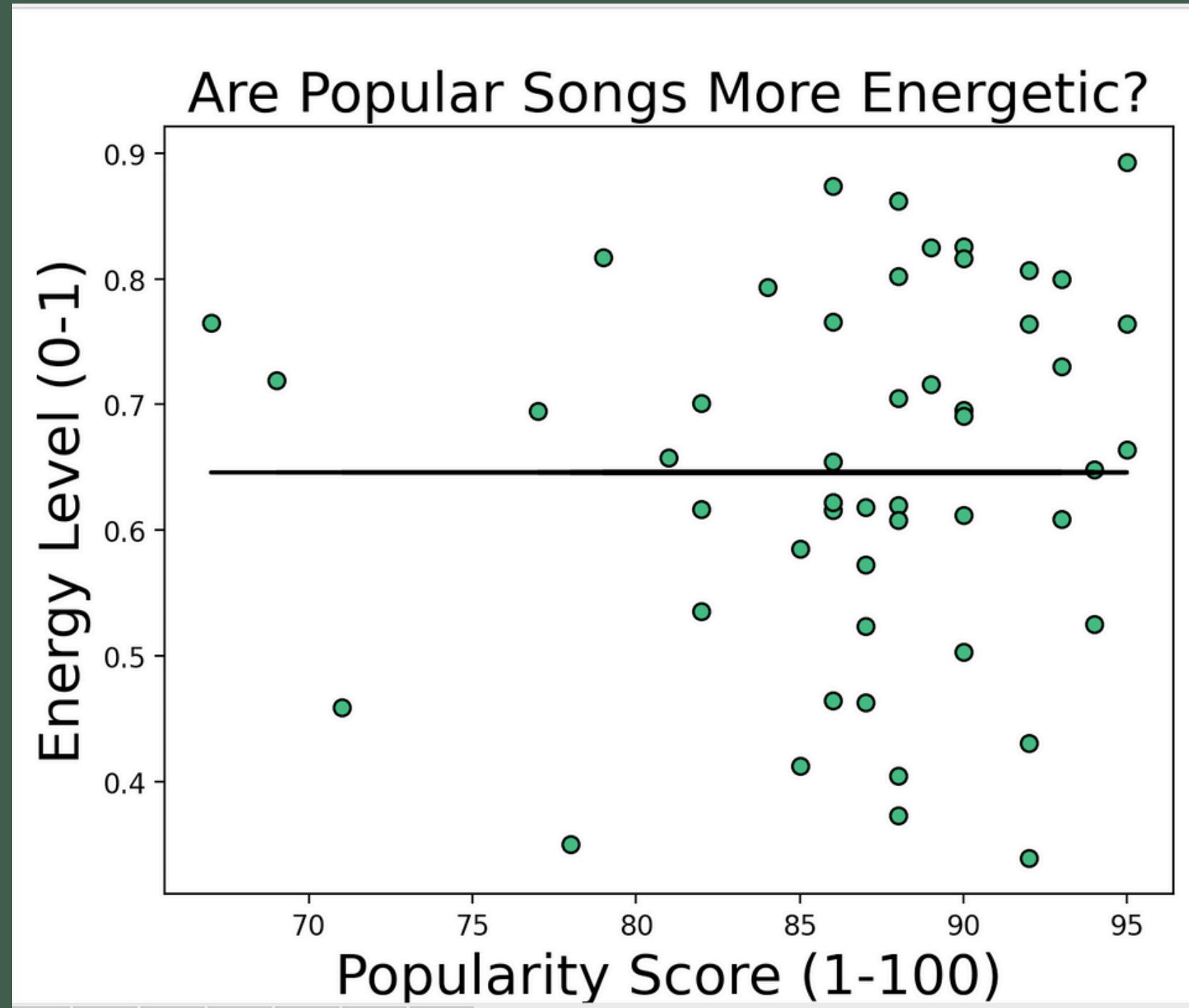


- Bar style chart
- Data came from Spotify Weekly Top Songs Global Chart
- Analyzes top 5 artists and their streams
- Streams are not total, just the streams from that week alone.

# Energy of Popular Songs



# Energy of Popular Songs



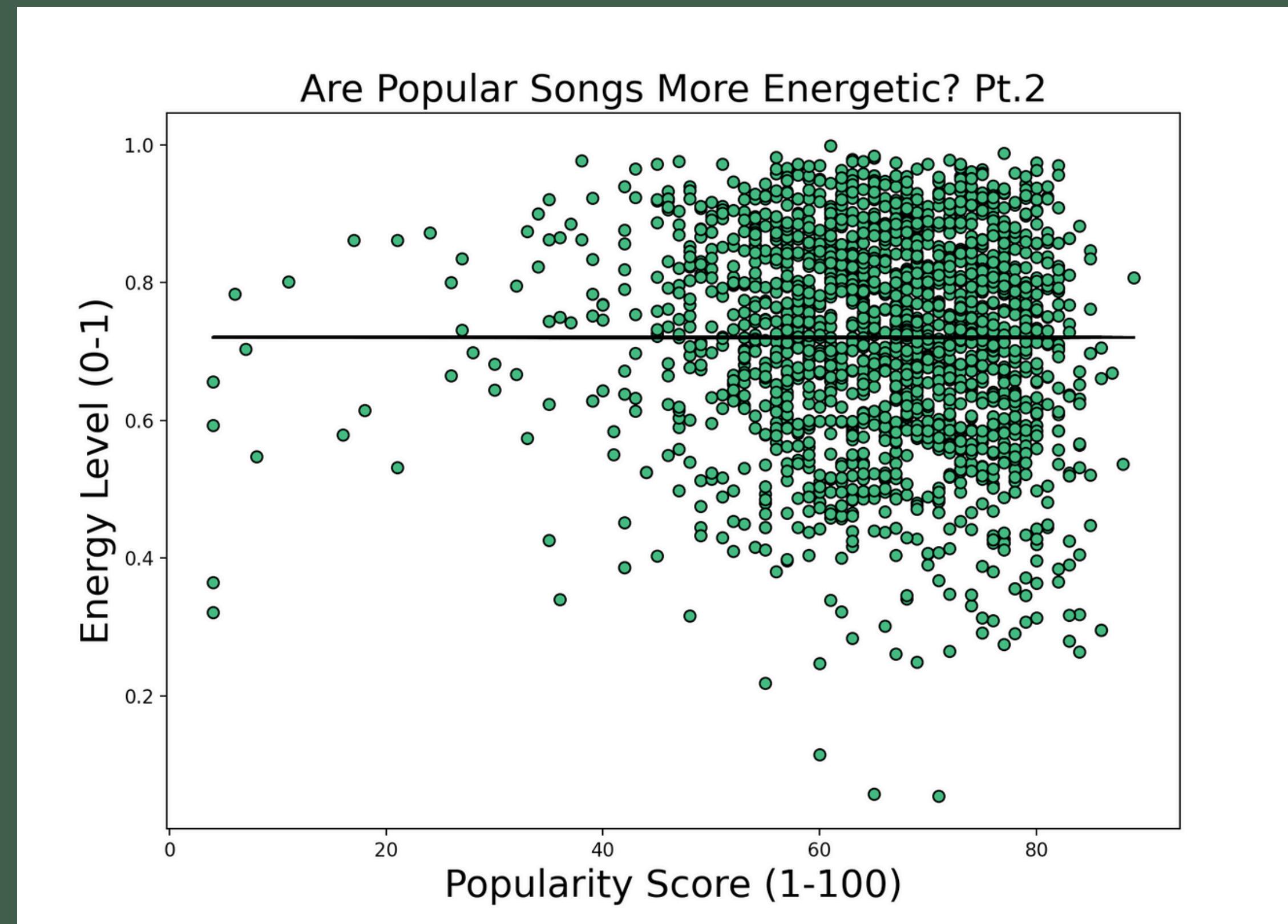
- Scatter plot
- Data from Song Attributes CSV
- Analyzes the energy level of popular songs
- Popularity is measured by a score of 1-100
- Energy level is on a scale of 0-1
- Average energy level between .6-.7
- Only 50 data points

# Top Hits Spotify from 2000-2019 CSV

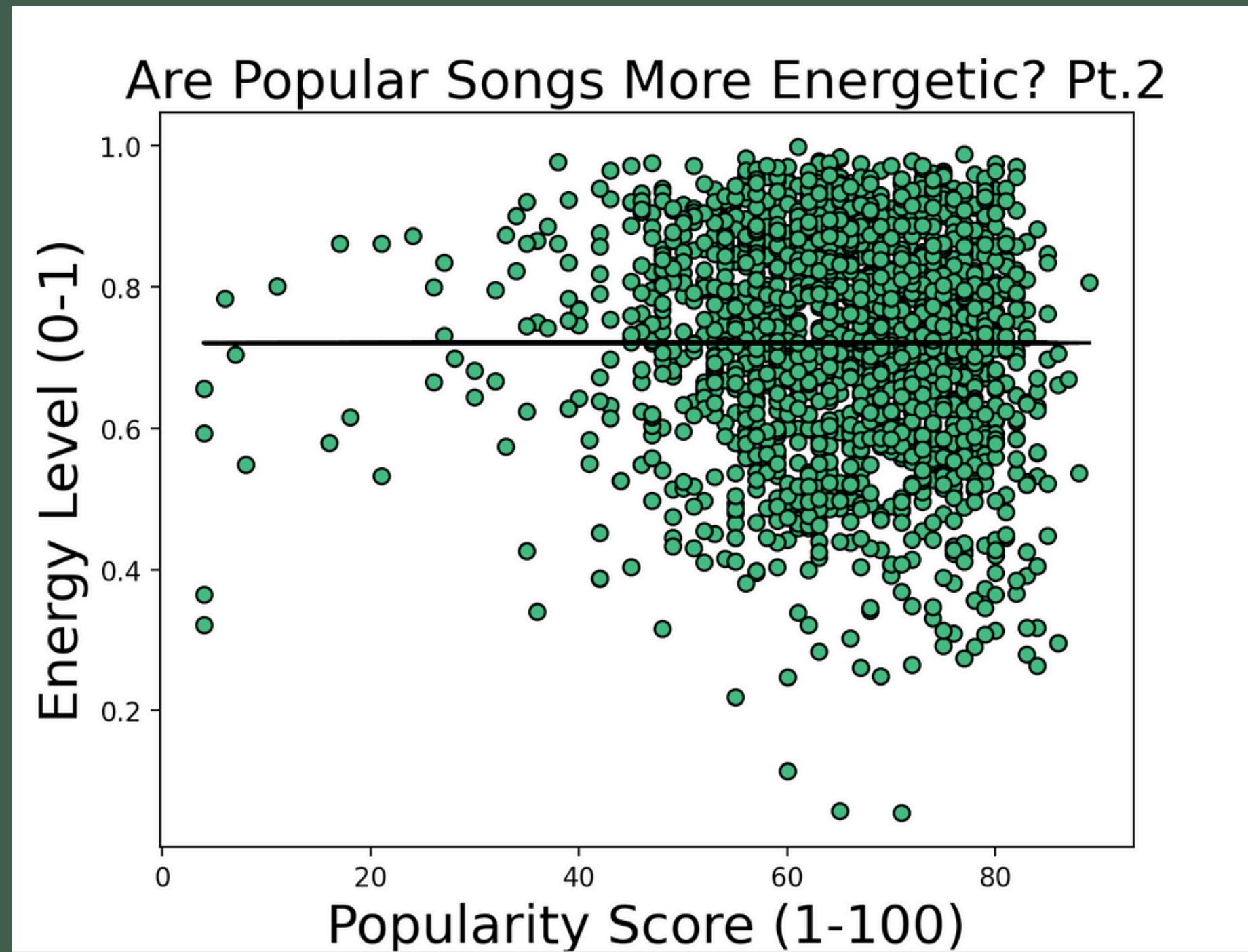
artist	song	duration_ms	explicit	year	popularity	danceability
Britney Spea	Oops!...I Did It	211160	FALSE	2000	77	0.751
blink-182	All The Small Things	167066	FALSE	1999	79	0.434
Faith Hill	Breathe	250546	FALSE	1999	66	0.529
Bon Jovi	It's My Life	224493	FALSE	2000	78	0.551
*NSYNC	Bye Bye Bye	200560	FALSE	2000	65	0.614
Sisqo	Thong Song	253733	TRUE	1999	69	0.706
Eminem	The Real Slim Shady	284200	TRUE	2000	86	0.949
Robbie Williams	Rock DJ	258560	FALSE	2000	68	0.708
Destiny's Child	Say My Name	271333	FALSE	1999	75	0.713

- Data comes from Kaggle
- Top hits from every year listed
- Has data such as artist, song, year, and song attributes similar to the other CSV's used in this project
- Includes songs from 1999, but they were hits in 2000.

# Energy of Popular Songs cont.



# Energy of Popular Songs cont.



- Scatter plot
- Analyzes the energy level of popular songs
- Popularity is measured by a score of 1-100
- Energy level is on a scale of 0-1
- About 1800 data points



## Conclusion

Overall, Spotify does a very thorough job of collecting user data. This data ranges from the way they categorize each song, to the charts and rankings each week. We hope our descriptive figures gave insight to Spotify data collection and user listening trends.

# Questions?

