



# Spotify Exploratory Data Analysis

*By Favour Ubadike*



Welcome



About the Data



My Objectives



The Top 10



Contributing Factors



Days



Duration



Danceability



Energy



Other



Recommendations



Contact Me



Introduction

# Welcome

- Interests: Cosmetics, Beauty, Gaming, Music and Healthcare
- Diploma in Pharmaceutical Science
- PL 300 Power BI Associate
- Naturally inquisitive

[Play](#)[Follow](#)

• • •

[Welcome](#)[About the Data](#)[My Objectives](#)[The Top 10](#)[Contributing Factors](#)[Days](#)[Duration](#)[Danceability](#)[Energy](#)[Other](#)[Recommendations](#)[Contact Me](#)

Background

# About The Data

+8 ...

- Data is merged from 2 Datasets ' Spotify top chart songs 2022' and 'Spotify Top 10000 Streamed Songs'
- The dataset ranks song names by Total streams, including energy or loudness
- SQL, Excel, Python, PowerBI used to clean, visualise, and gain insight
- A scenario is created to answer a business questions



Welcome



About the Data



My Objectives



The Top 10



Contributing Factors



Days



Duration



Danceability



Energy



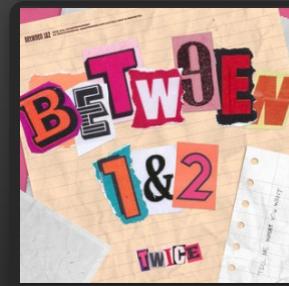
Other



Recommendations



Contact Me



TWICE

# About The Data

+8 ...

TWICE (A Korean girl group) previously released a song 'The Feels', however none of the ranked within top 10 and peaked at 147th. Their artists are planning to release a single and want to know what they can do to make sure that the new song hits top 10.

## Their Request

They request to understand the current top 10 ranking and how they can improve their global streaming number

```
-- Looking for Twice Song(s) in Merged Data set
SELECT Position, Artist_Name, Song_Name, Peak_Position,
weeks_on_chart, Total_Streams, Top_10__xTimes_, danceability,
energy, key, loudness, mode, speechiness, acousticness,
instrumentalness, liveness, tempo, duration_ms
FROM `niyo-sql-project.Spotify.Short`
INNER JOIN `niyo-sql-project.Spotify.Analysis`
ON `niyo-sql-project.Spotify.Short`.track_name =
`niyo-sql-project.Spotify.Analysis`.Song_Name
WHERE Artist_Name = 'TWICE ';
```



Welcome



About the Data



My Objectives



The Top 10



Contributing Factors



Days



Duration



Danceability



Energy



Other



Recommendations

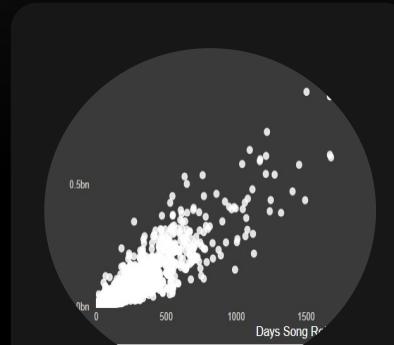


Contact Me

# My Objectives



What are the top 10 songs?



What Factors Contribute to streaming numbers?



How can TWICE reach top 10?



Welcome



About the Data



My Objectives



The Top 10



Contributing Factors



Days



Duration



Danceability



Energy



Other



Recommendations



Contact Me

P	Q	R	S	T	U	V	W	X
punctiveness	instrumentalness	liveness	tempo	time_signature	duration_ms			
0.349	0	0.34	83.903	4	239836			
0.324	0	0.0534	100.028	4	292799	COUNT	147	
0.0847	0	0.149	130.049	4	243837	MAX	248	
0.00146	0.0000954	0.0897	171.005	4	200040	XLOOKUP	Heat Waves	
0.469	0.00000413	0.297	134.021	4	119133			

Objective 1

# The Top 10



• • •

#

Title

1 Ranking Based on Number of Streams

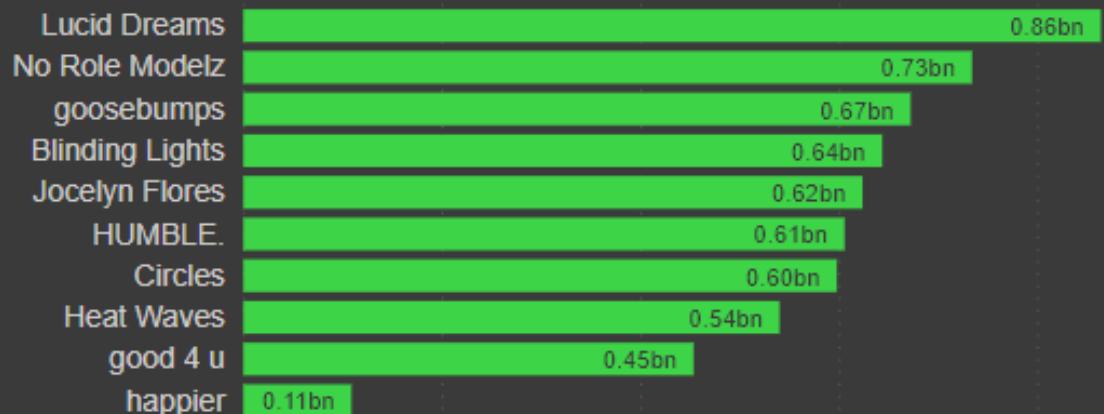
2 Power BI graph to right lists top 10 songs

3 Most streamed song – Lucid Dreams Juice Wrld

4 Least Streamed song has 0.11 billion streams

5 Excel Shows that out of 147 songs, Heat Wave, by Glass Animals, has been in the top 10 the most times

## Top 10 Songs by Streaming Number



0.0bn

0.2bn

0.4bn

0.6bn

0.8bn

Total Streams (Billions)

[Upgrade](#)

Favour U ▾



Welcome



About the Data



My Objectives



The Top 10



Contributing Factors



Days



Duration



Danceability



Energy



Other



Recommendations



Contact Me

# Contributing Factors

These are factors that have been found to Affect Streaming Numbers

**01**

Number of Days Song released

**04**

Acousticness and Speechiness

**02**

Duration of the song

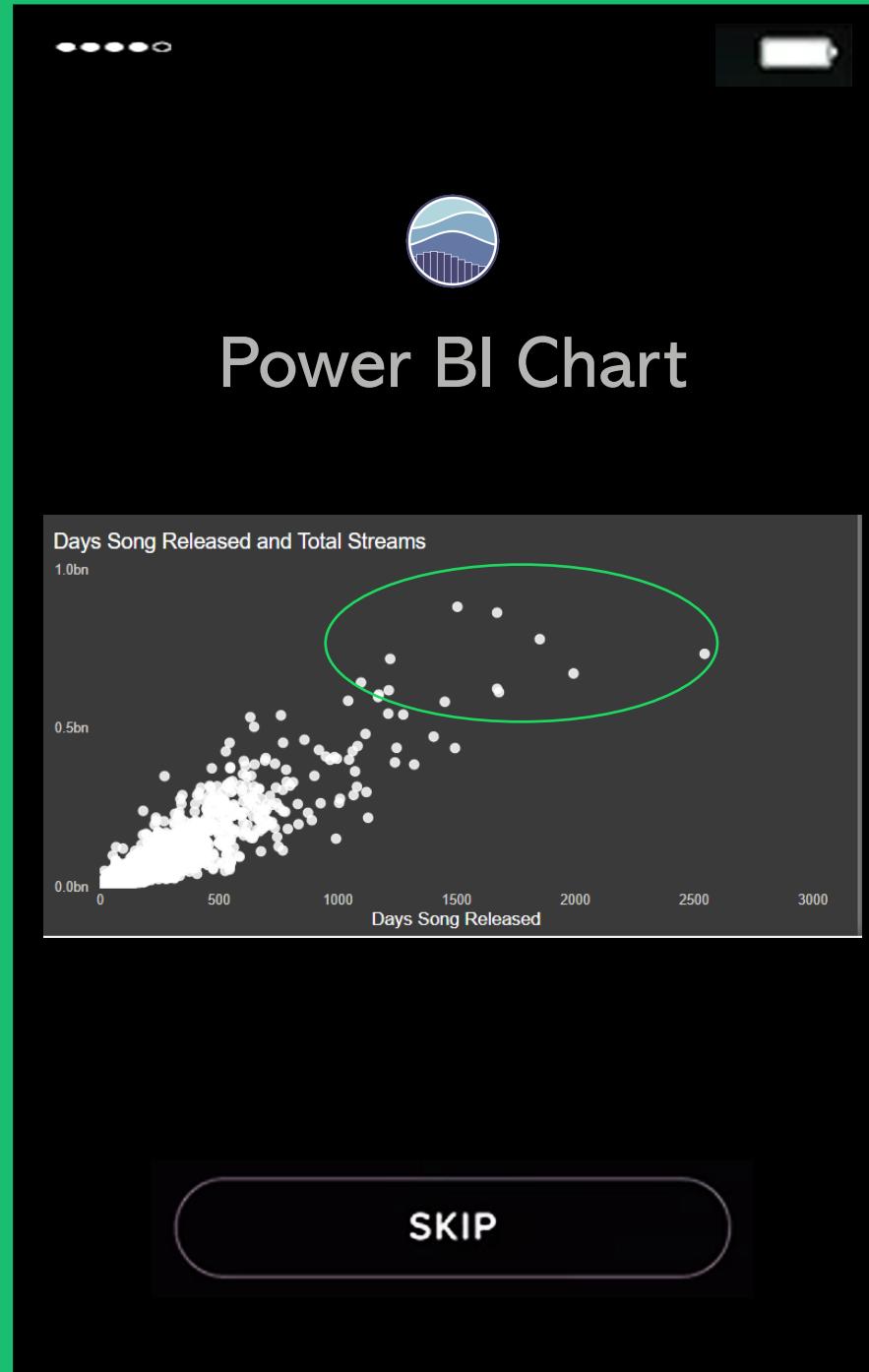
**05**

Energy

**03**

Danceability

- Positive correlation between Number of days and Total streams
- Songs in the top 10 have been released for more than 1000 days
- SQL, compared 2 songs. Song with double streams, had been released double the time
- Expected as more people (who may not be fans) have time to discover the songs



## ORDER BY Days

```
-- Does the number of days the song has been released have any impact on number of streams
SELECT Position, Artist_Name, Song_Name, Days, Total_Streams
FROM `niyo-sql-project.Spotify.Short`
INNER JOIN `niyo-sql-project.Spotify.Analysis`
ON `niyo-sql-project.Spotify.Short`.track_name =
`niyo-sql-project.Spotify.Analysis`.Song_Name
WHERE Position = 2 OR Position = 53
ORDER BY Days Desc;
-- Can see that although the 2nd position has around double the streams, it has been released for around double the amount of time, which may be a contributing factor to its high streams and therefore position
```

Song_Name	Days	Total_Streams
Lucid Dreams	1673	864832399
Shape of You	618	348283656

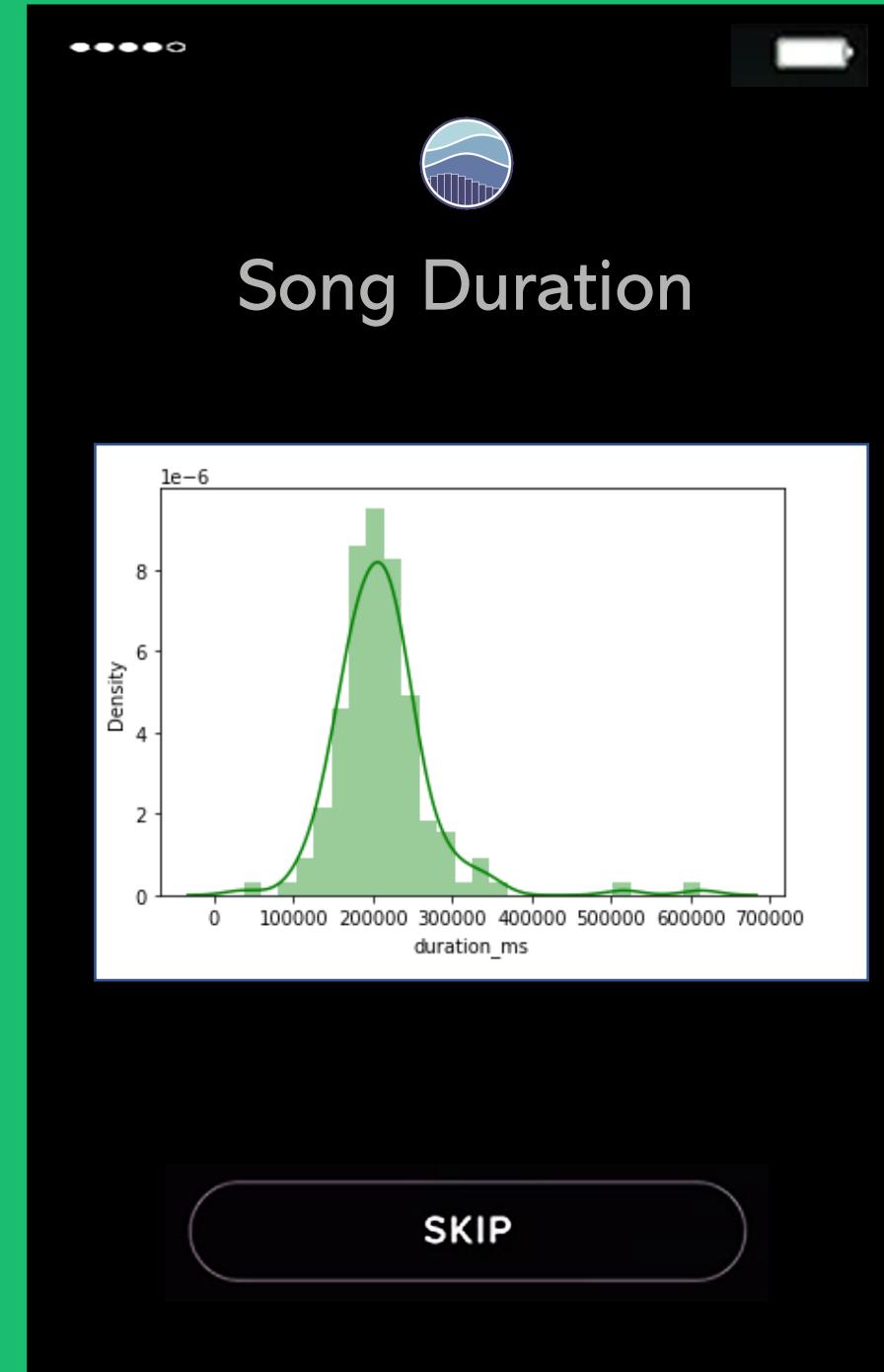
SKIP

- Graphs were found by looking only at songs that have peaked within the top 10
- Most songs in the data have a duration between 2 and 4 minutes
- Expected due to attention span and being radio-friendly

## Python code:

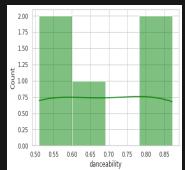
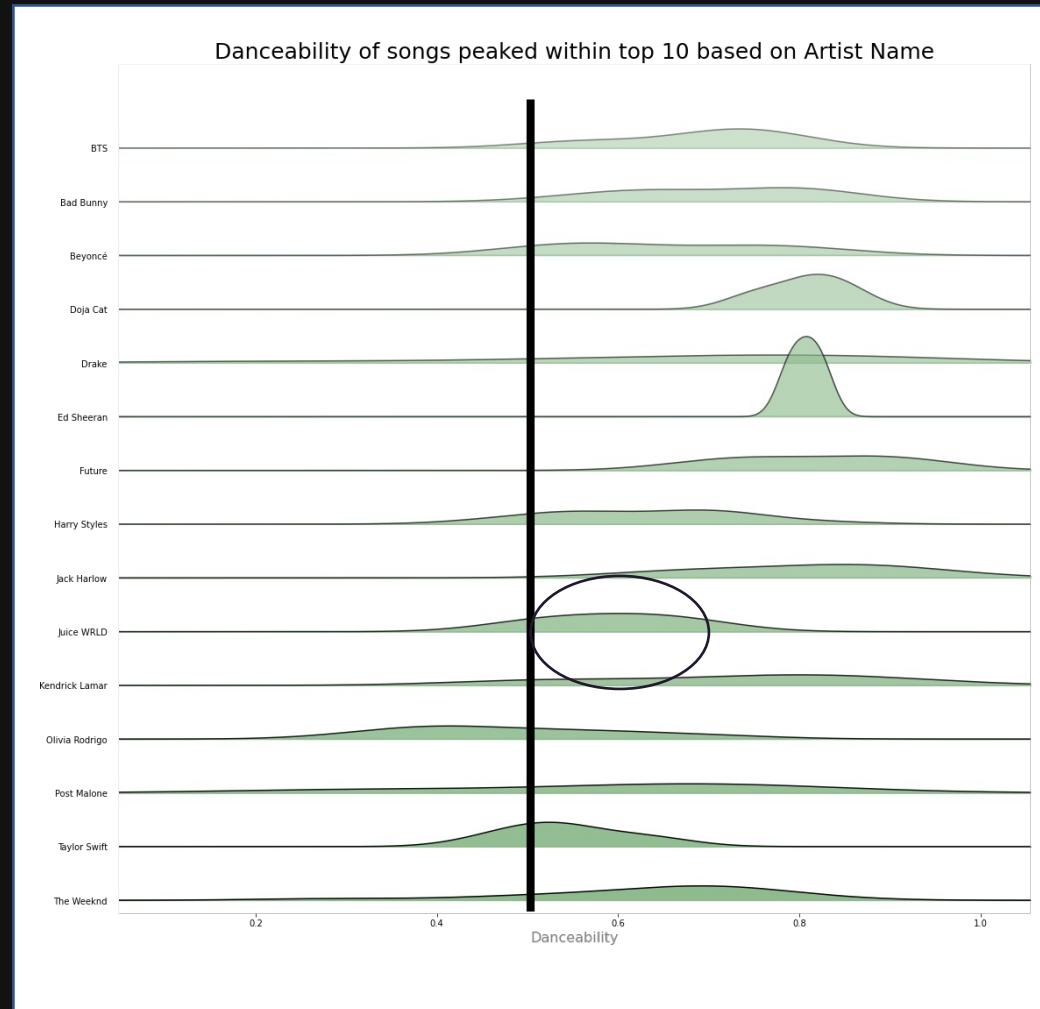
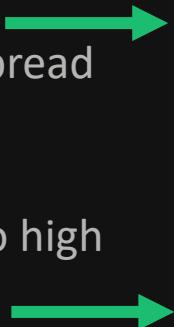
```
# Only looking at songs within top 10
resultdf.drop(resultdf.loc[resultdf['Peak Position'] > 10].index, inplace = True)
resultdf

#Song duration Distribution for songs that have peaked within top 10
sns.distplot(resultdf["duration_ms"], color = 'green')
```



# Danceability

- Ridgeline plot shows Songs that have ranked within top 10, categorized by the Artist name
- Post Malone has a largely spread out danceability of songs
- Majority of have a middle to high Danceability (above 0.5)
- Rise of apps such as TikTok makes danceable songs more popular



Danceability Distribution  
Box Plot



0:23



0:23 - 3:25

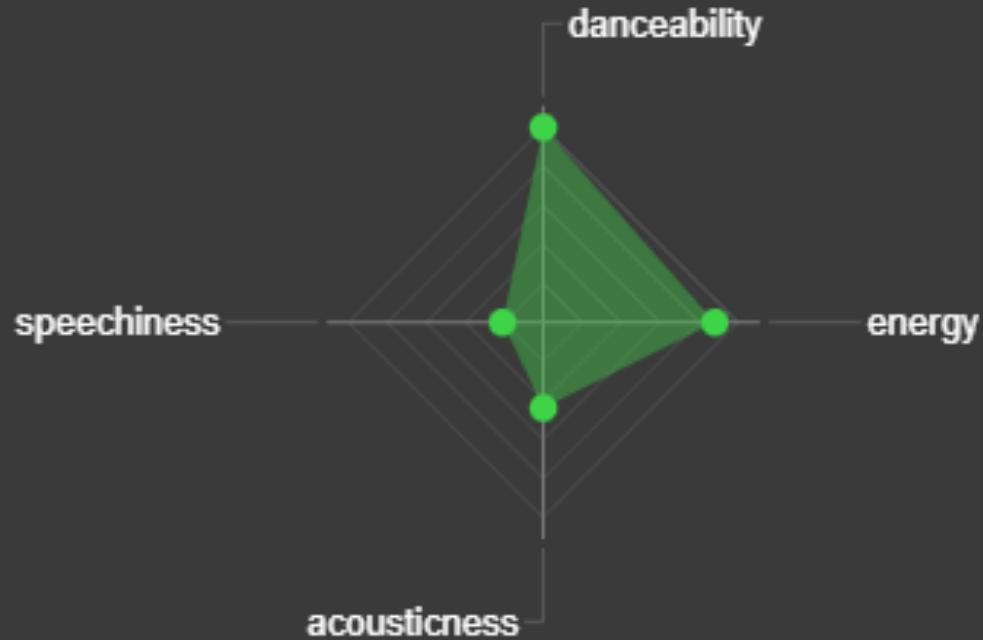


- Radar chart showing 3 songs found in top 10
- Danceability is higher than the other features
- Acousticness and Speechiness kept below 0.5
  - Acousticness: how likely the song is to be acoustic
  - Speechiness: detects presence of spoken words in a song



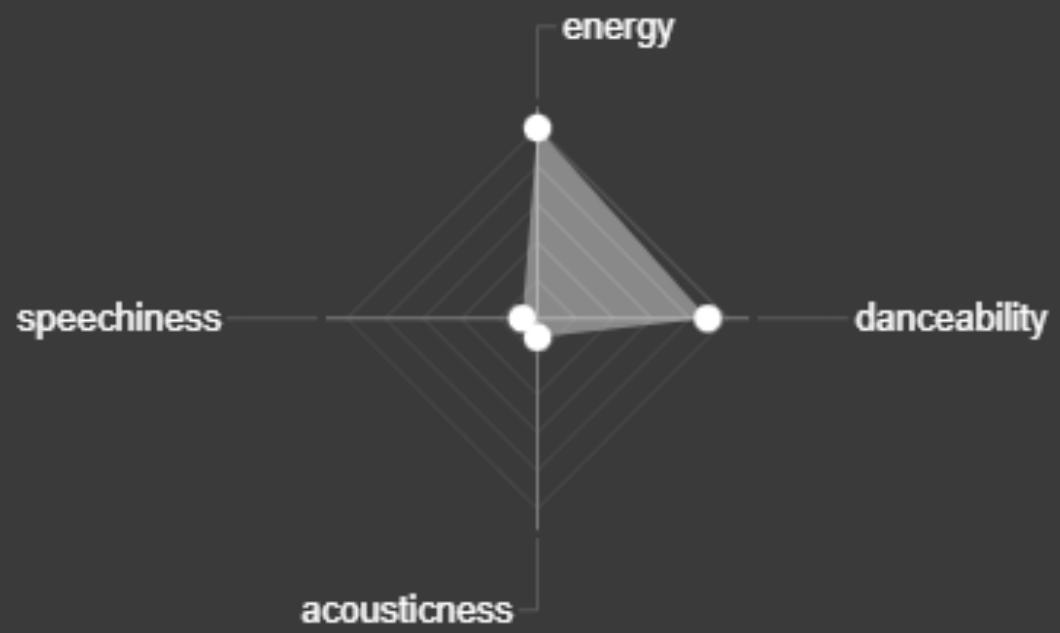
## Attributes of Top 10 Songs

● Sum of Value



## Attributes of Song: The Feels - TWICE

● Sum of Value



- Very Similar Danceability, Energy and Speechiness Values
- Acousticness also kept less than 0.5
- Likely a different reason for low ranking e.g. Korean group not as recognized in global world

[Upgrade](#)

Favour U ▾



Welcome



About the Data



My Objectives



The Top 10



Contributing Factors



Days



Duration



Danceability



Energy



Other



Recommendations



Contact Me

# Recommendations

TWICE should make sure that:

**2 - 4**

Their new song  
should be between  
2 and 4 minutes

**1000+**

They should aim  
for top 10 after  
1000 days of their  
song release

**High**

The Song has high  
danceability

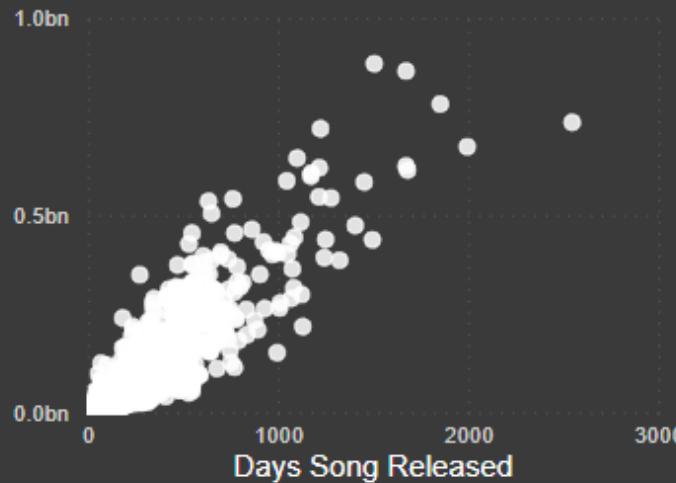
**Low**

The song has low  
acousticness and  
speechiness

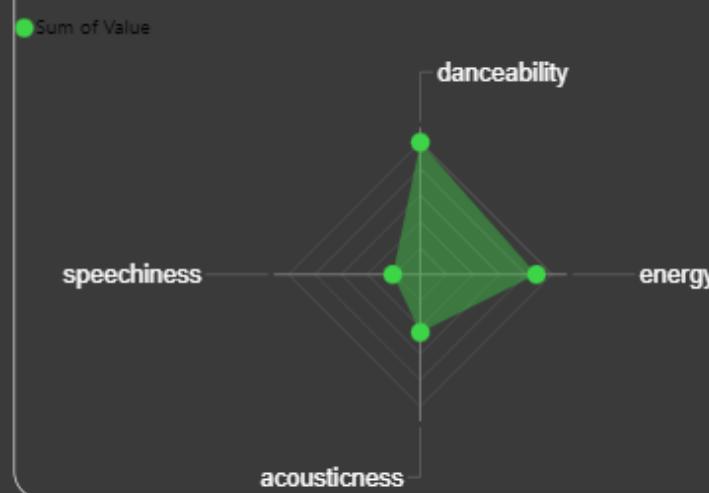


# Spotify EDA Dashboard: Ranking within the Top 10

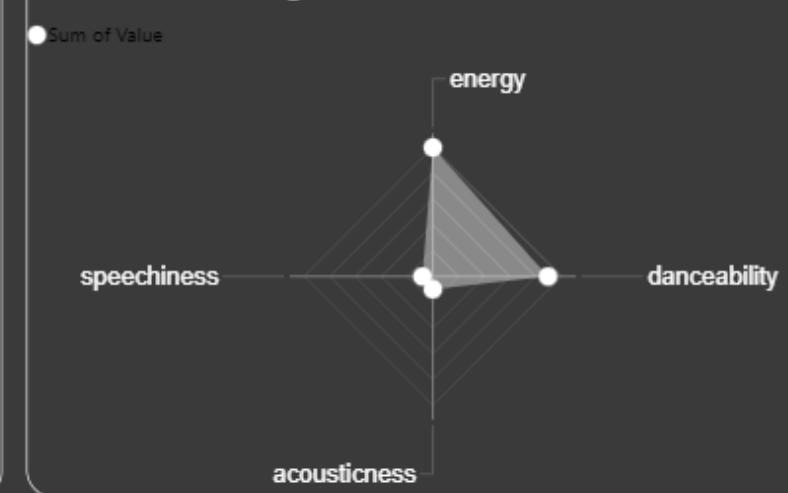
## Days Song Released by Total Streams



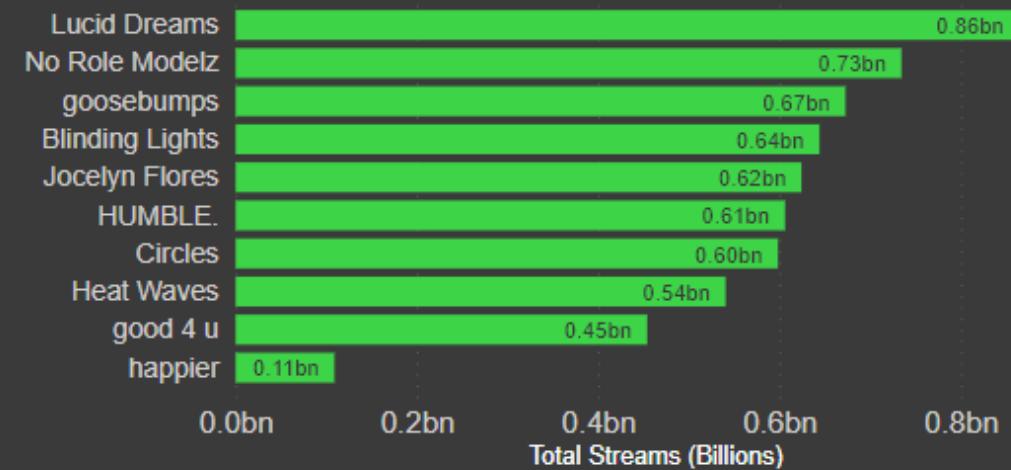
## Attributes of Top 10 Songs



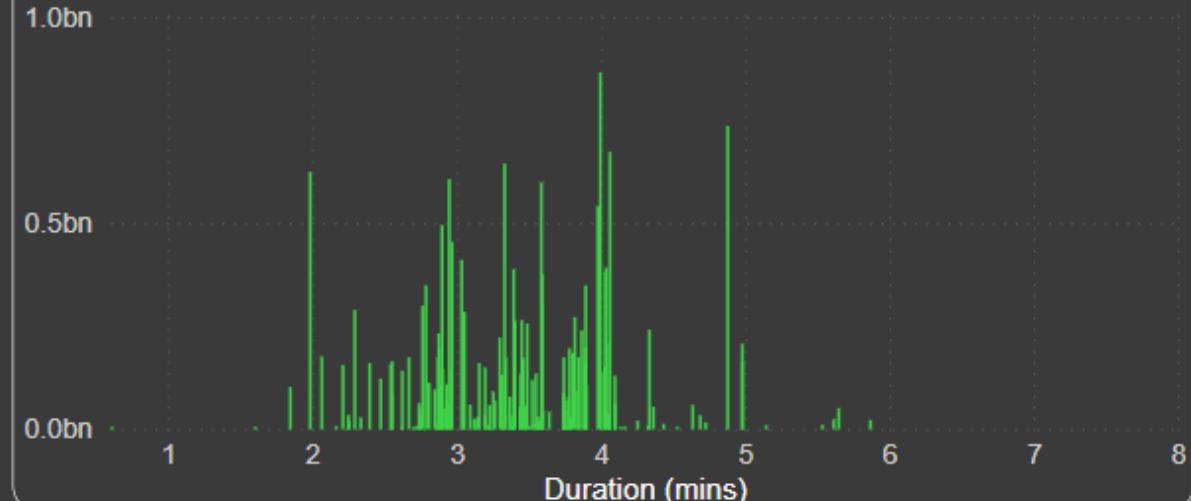
## Attributes of Song: The Feels - TWICE



## Top 10 Songs by Streaming Number



## Duration Density



# Thank You

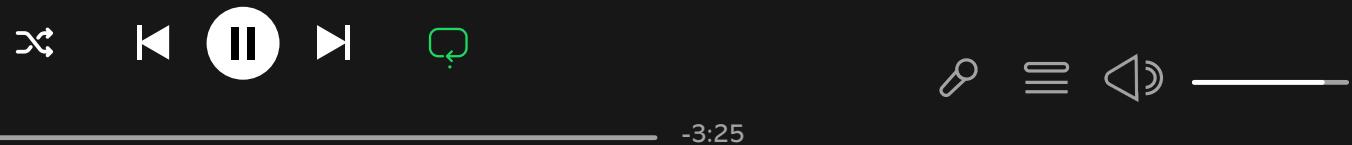


Very much

For listening to my presentation



0:23





Welcome



About the Data



My Objectives



The Top 10



Contributing Factors



Days



Duration



Danceability



Energy



Other



Recommendations



Contact Me

# Contact Me

## Email:

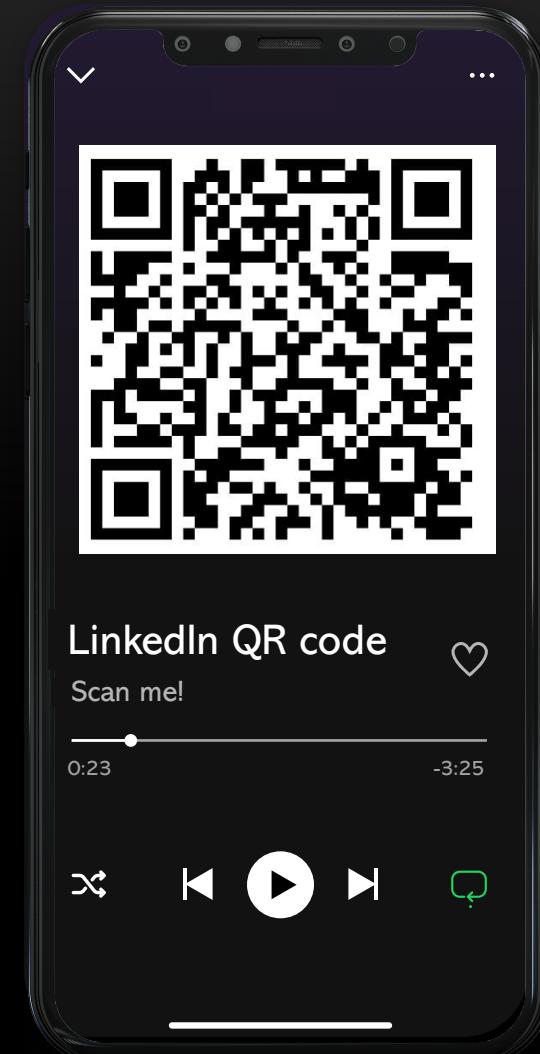
Favourubadike@yahoo.com

## Github:

<https://github.com/FavourU>

## LinkedIn:

<https://www.linkedin.com/in/favourubadike/>



Slides have been created using:



Slide Chef