PRACTICAL NO 10

MINI PROJECT

❖ *GROUP MEMBERS :-*

D-03 Arya Harde

D-04-Devyani Kalamkar

D-10 Muskan Setiya

D-15 Sanskriti Gupta

- **Name of the dataset** - Top Hits Spotify from 2000-2019

- **Problem Statement –**

Analyzing the dataset and performing data cleaning and handling and visualization using tableau

- **About Dataset**

This dataset contains audio statistics of the top 2000 tracks on Spotify from 2000-2019. The data contains about 18 columns each describing the track and it's qualities.
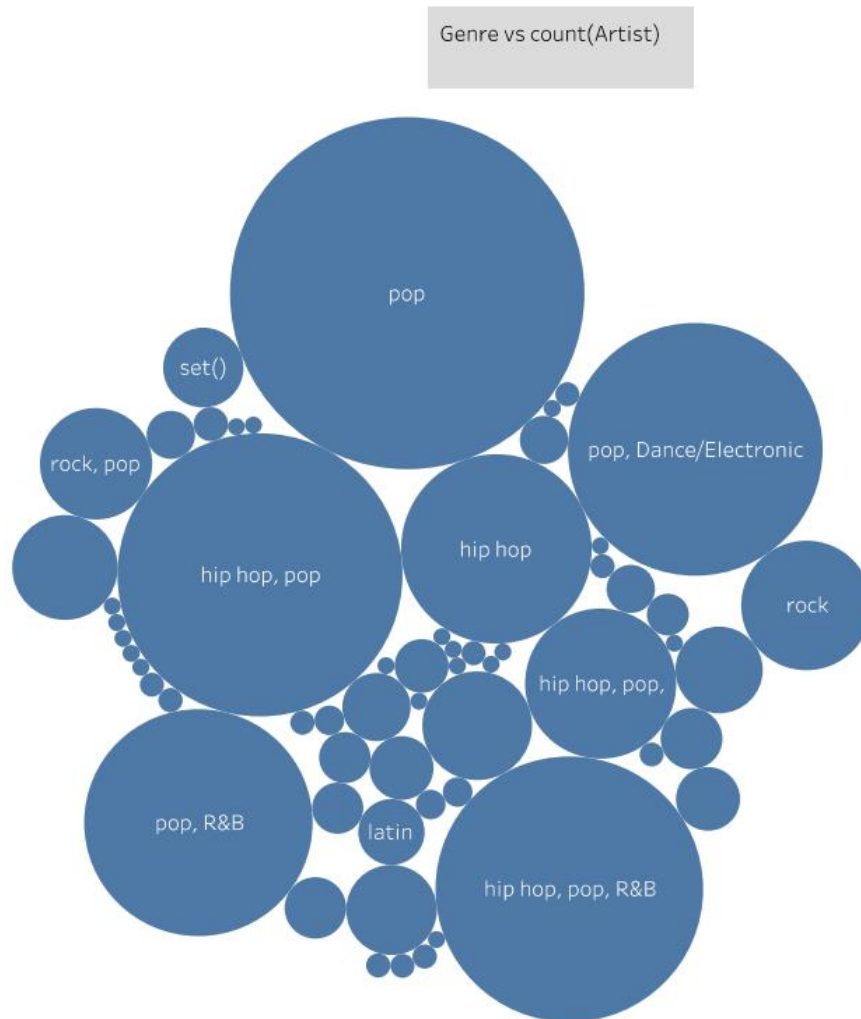
Content

- **artist**: Name of the Artist.
- **song:** Name of the Track.
- **duration_ms:** Duration of the track in milliseconds.
- **explicit**: The lyrics or content of a song or a music video contain one or more of the criteria which could be considered offensive or unsuitable for children.
- **year**: Release Year of the track.
- **popularity:** The higher the value the more popular the song is.
- **danceability:** Danceability describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable.
- energy: Energy is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity.
- **key:** The key the track is in. Integers map to pitches using standard Pitch Class notation. E.g. 0 = C, 1 = C♯/D♭, 2 = D, and so on. If no key was detected, the value is -1.

- **loudness**: The overall loudness of a track in decibels (dB). Loudness values are averaged across the entire track and are useful for comparing relative loudness of tracks. Loudness is the quality of a sound that is the primary psychological correlate of physical strength (amplitude). Values typically range between -60 and 0 db.

- **mode:** Mode indicates the modality (major or minor) of a track, the type of scale from which its melodic content is derived. Major is represented by 1 and minor is 0.

- **speechiness:** Speechiness detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audio book, poetry), the closer to 1.0 the attribute value. Values above 0.66 describe tracks that are probably made entirely of spoken words. Values between 0.33 and 0.66 describe tracks that may contain both music and speech, either in sections or layered, including such cases as rap music. Values below 0.33 most likely represent music and other non-speech-like tracks.

- **acousticness:** A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic.

- **instrumentalness**: Predicts whether a track contains no vocals. "Ooh" and "aah" sounds are treated as instrumental in this context. Rap or spoken word tracks are clearly "vocal". The closer the instrumentalness value is to 1.0, the greater likelihood the track contains no vocal content. Values above 0.5 are intended to represent instrumental tracks, but confidence is higher as the value approaches 1.0.

- **liveness:** Detects the presence of an audience in the recording. Higher liveness values represent an increased probability that the track was performed live. A value above 0.8 provides strong likelihood that the track is live.

- **valence**: A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry).

- **tempo:** The overall estimated tempo of a track in beats per minute (BPM). In musical terminology, tempo is the speed or pace of a given piece and derives directly from the average beat duration.

- **genre:** Genre of the track.

- Usability info :10.00

- License :Other (specified in description)

- Expected update frequency : Never

Link of the dataset :- https://www.kaggle.com/datasets/paradisejoy/top-hits-spotify-from 20002019

e

Story 1

Genre vs count(Artist)

From this graph we conclude that on comparing popularity vs artist plot and enegy vs artist. It is analysed that it is not necessary factor that energy is proportional to popularity.

Story 3

Popularity of Top 10 artists
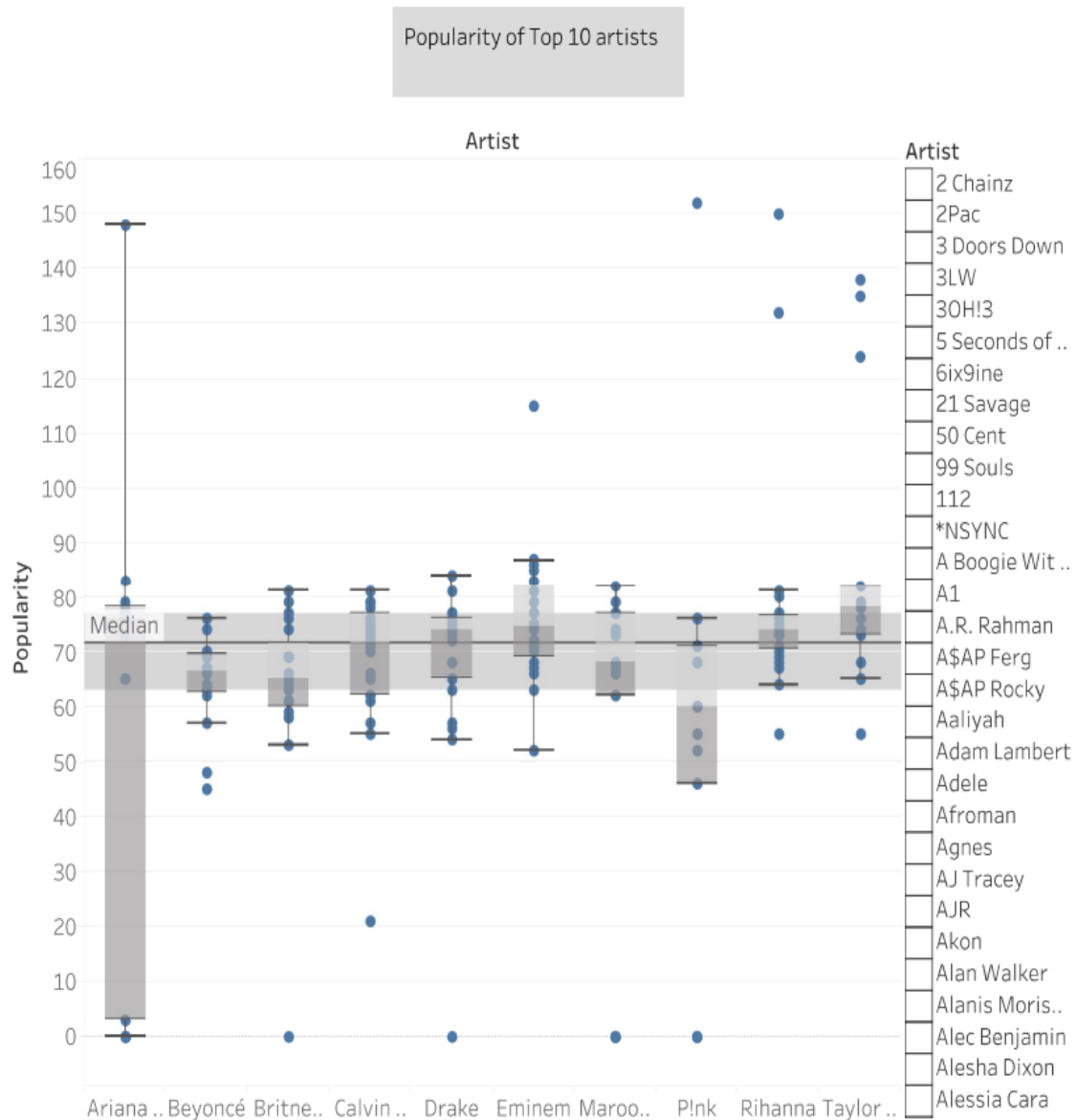
The above graph compares the performance of most popular artist on the basis of there songs. It has been observed that the media popularity of sum of there songs stands at 71.5
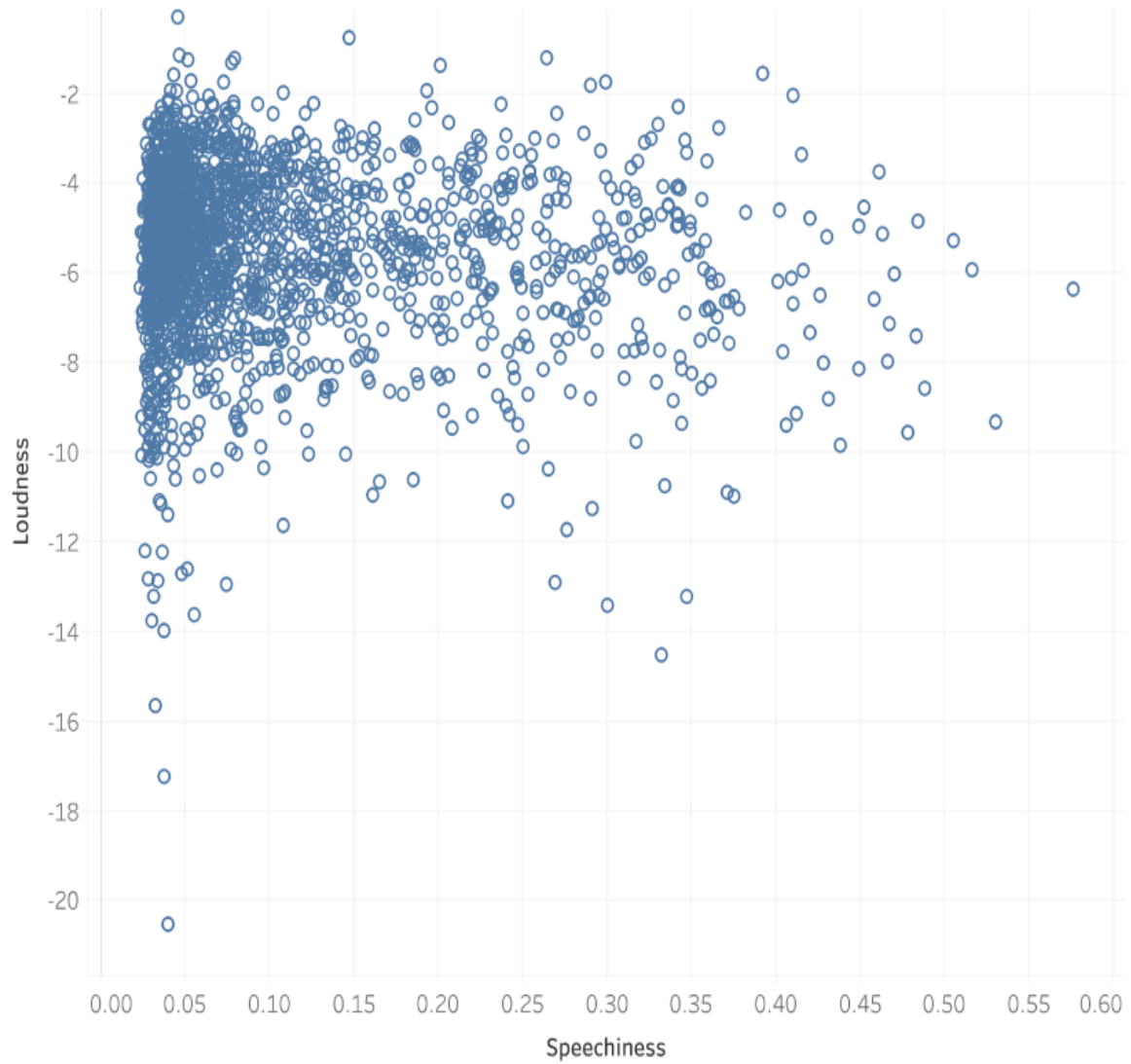
Story 4

Explicit
- False
- True

Count of Artist
2,000
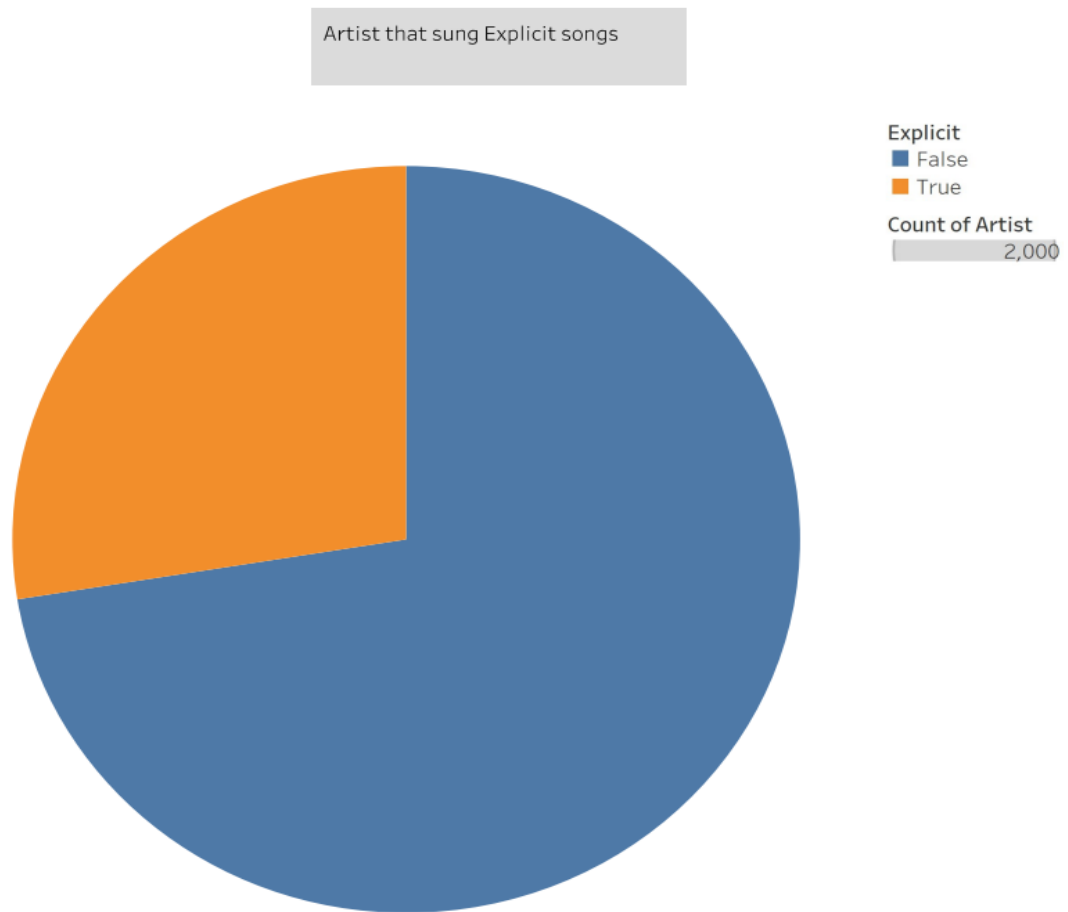
About 27.55% of artist sing explicit songs i.e (explicit ki defination) . We can conclude that these songs can be prohibited or warned to the listeners below the age 16.
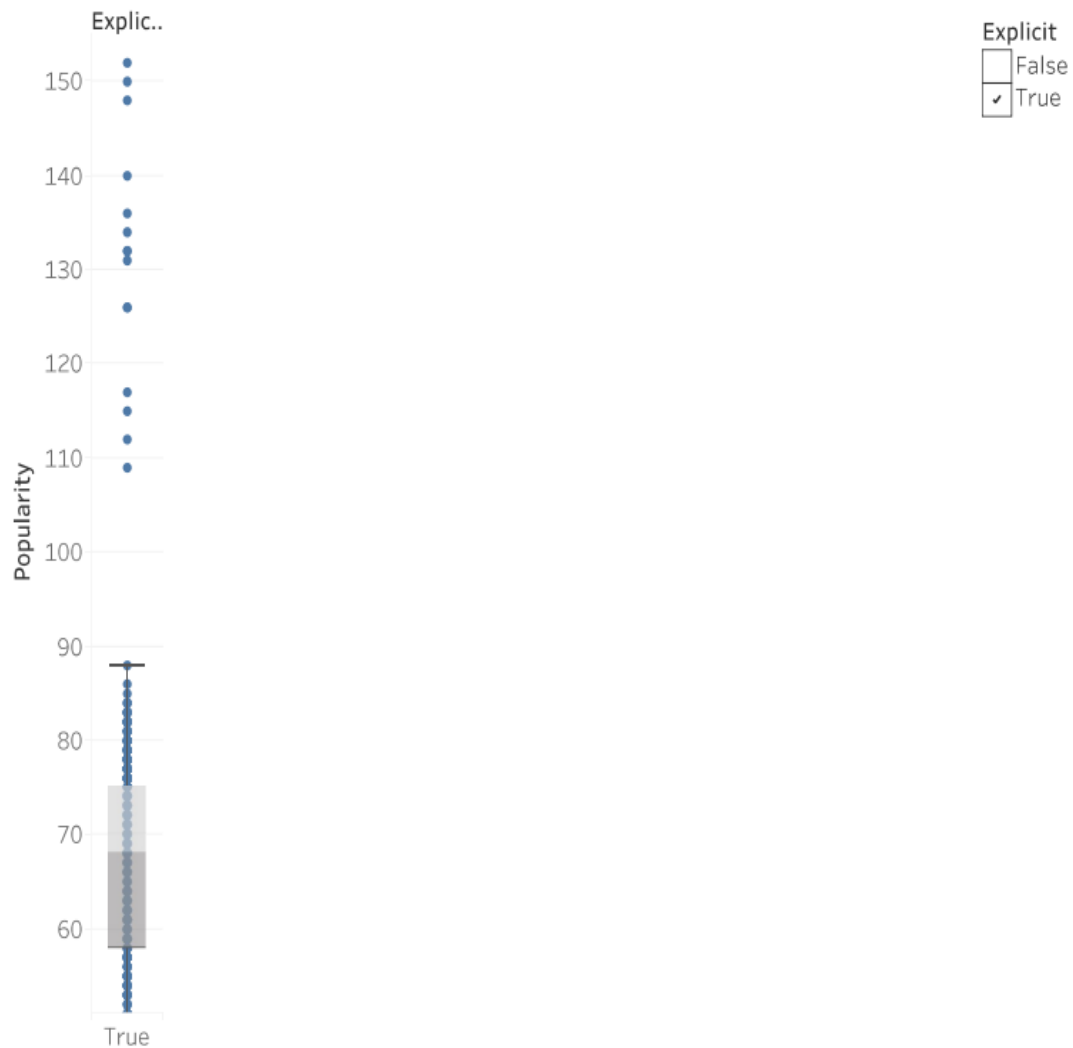
Story 5

Loudness vs Speechiness

Story 6

Expliciteness vs Popularity of a song

Explic..

Popularity

150
140
130
120
110
100
90
80
70
60

True

Explicit
☐ False
☑ True

Explicit:	True
Song:	Whatever You Like
Popularity:	148

Explicit:	True
Song:	Live Your Life
Popularity:	150

Explicit:	True
Song:	Chill Bill (feat. J. Davi$ & Spooks)
Popularity:	152

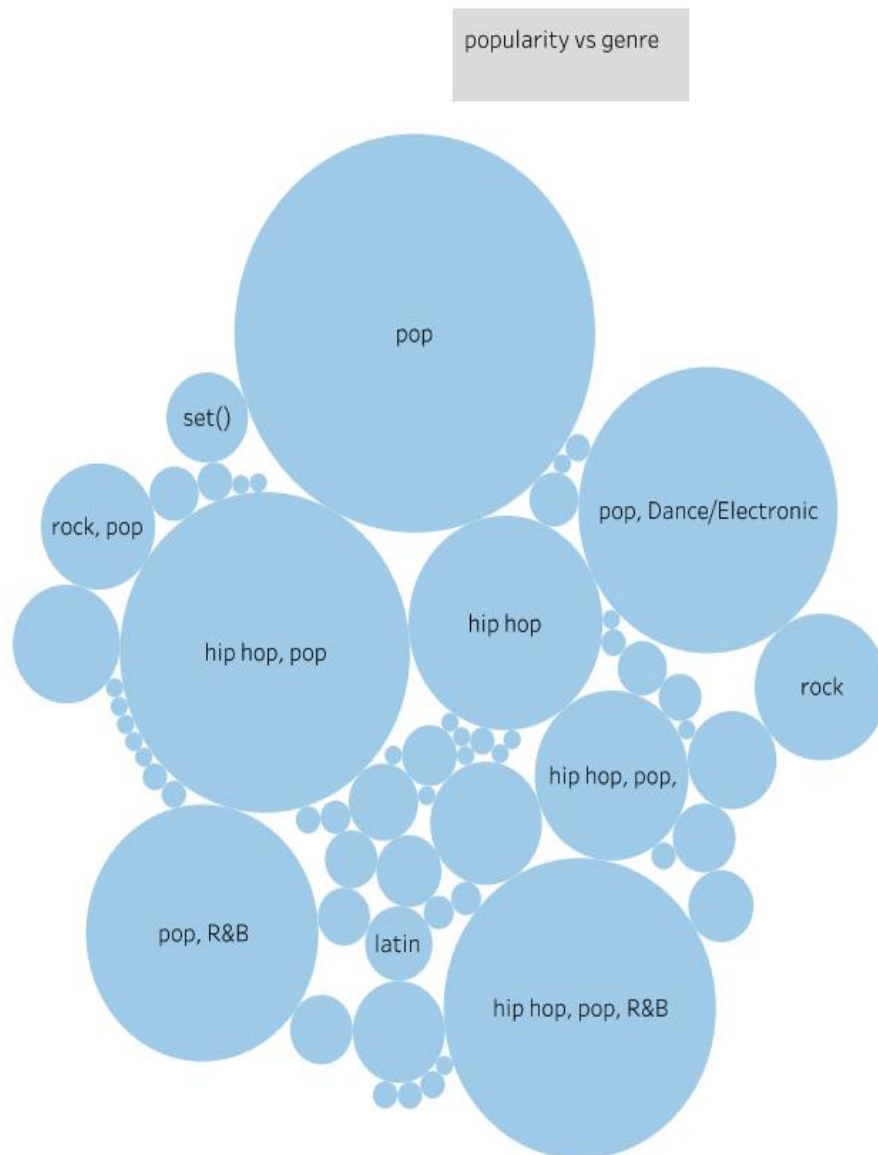On comparing the explicit songs with the popularity, though they contain harsh or sexual words, still there were songs that got hit.

Number of song VS Year
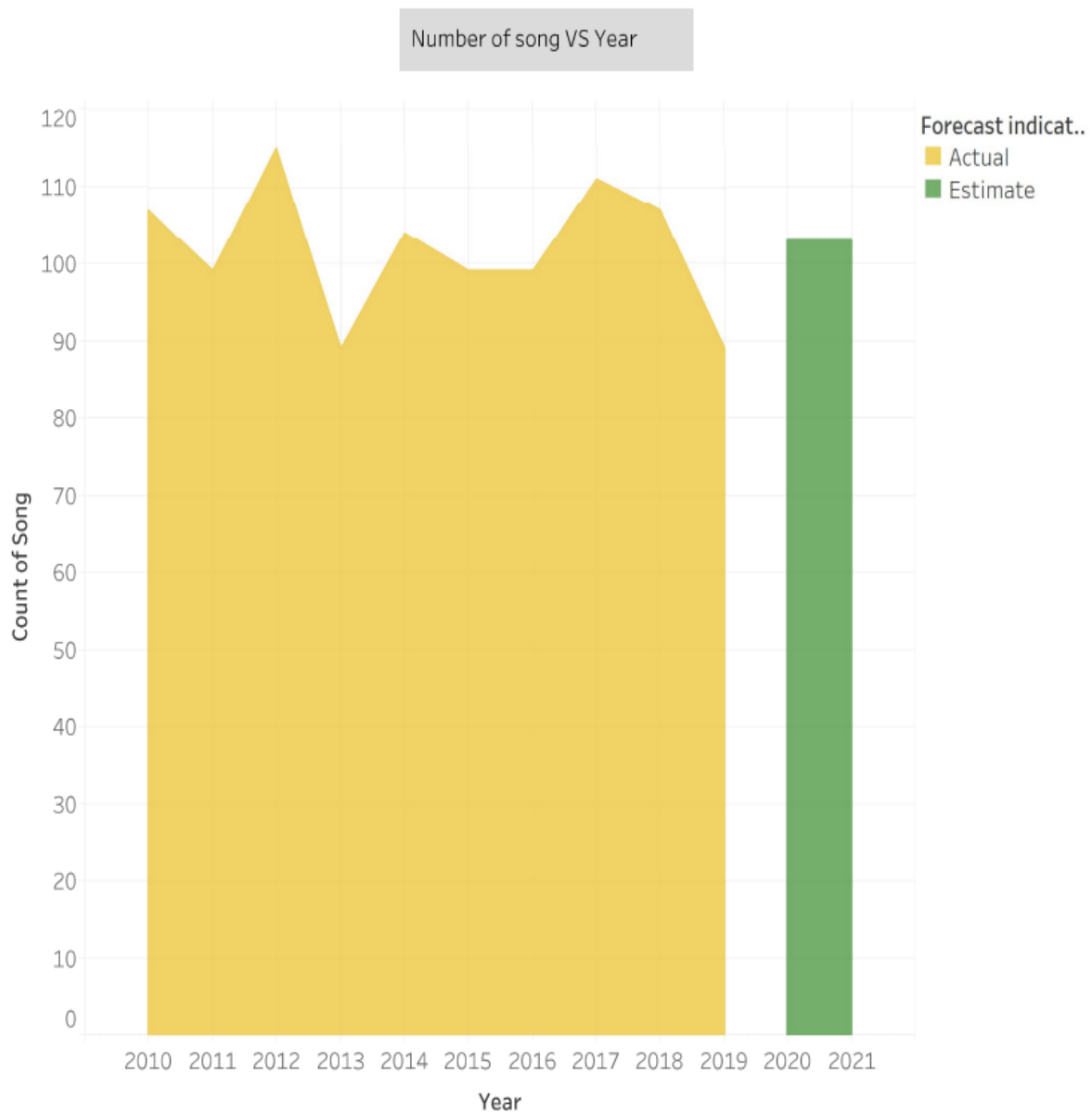
from this graph it is observed that in year 2012 maximum number of songs were released and in year 2013 least number of songs were released.

Danceability vs Duration
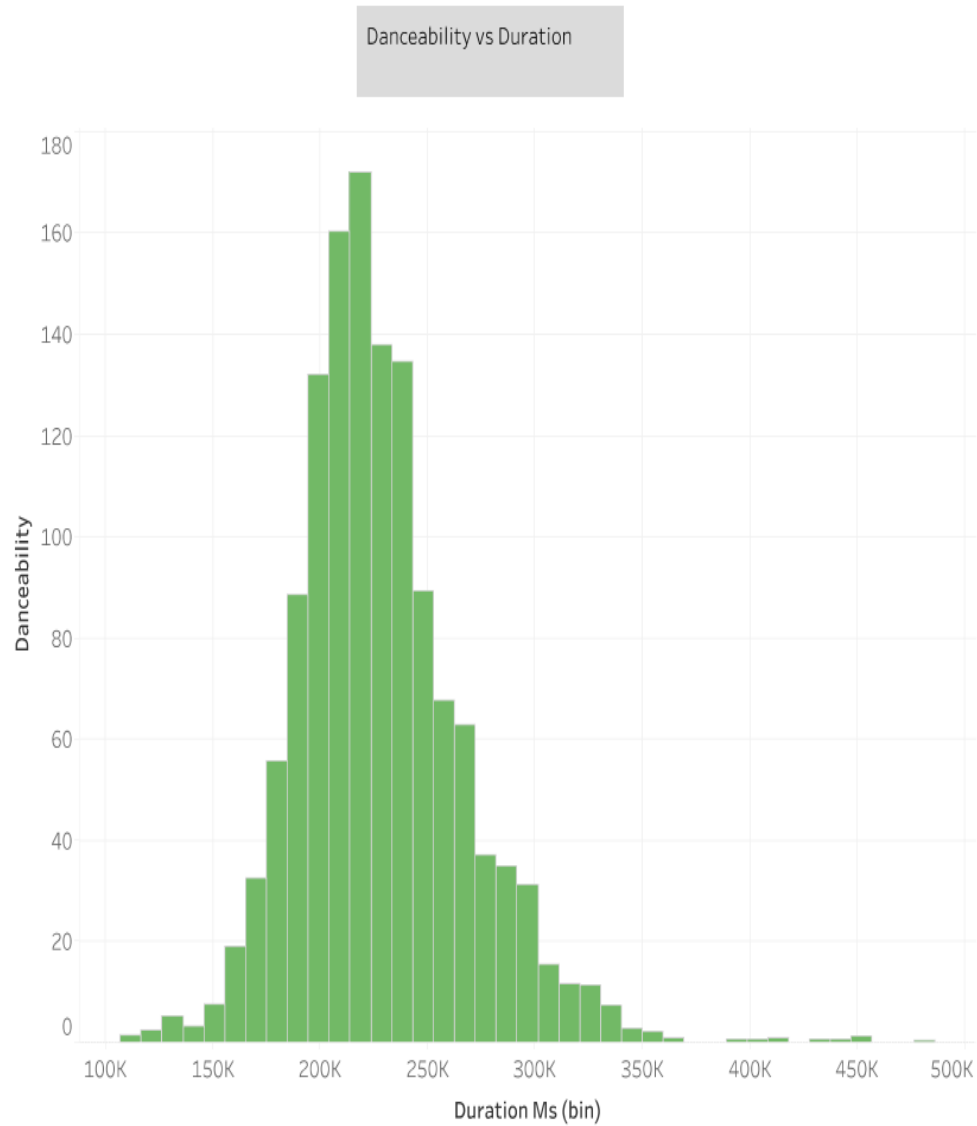
This is the graph between danceability and duration ...it is analysed that people don't enjoy the song if they have maximum duration..they prefered the song with average duration .

popularity vs genre

pop

set()

rock, pop

pop, Dance/Electronic

hip hop, pop

hip hop

rock

hip hop, pop,

pop, R&B

latin

hip hop, pop, R&B

Popularity vs Genre graph shows which genre is liked by most of the people and it's clearly observable that pop music have highest popularity.