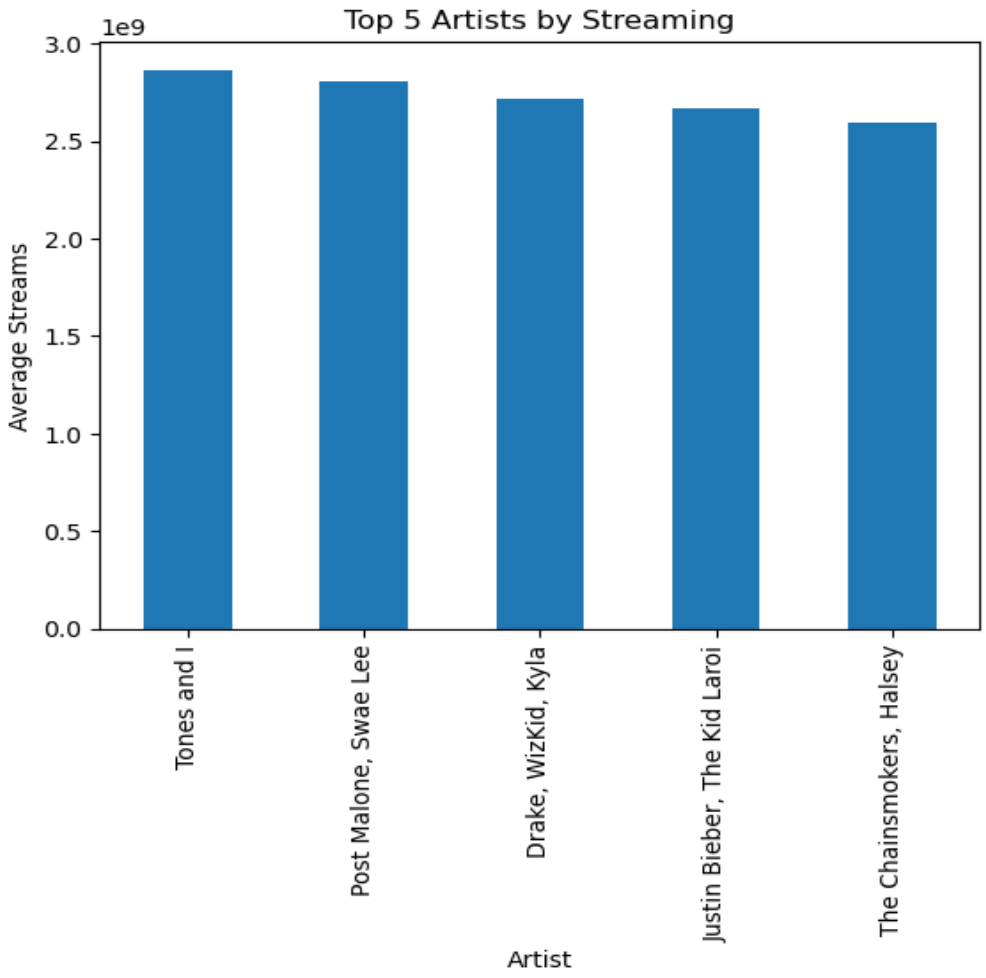
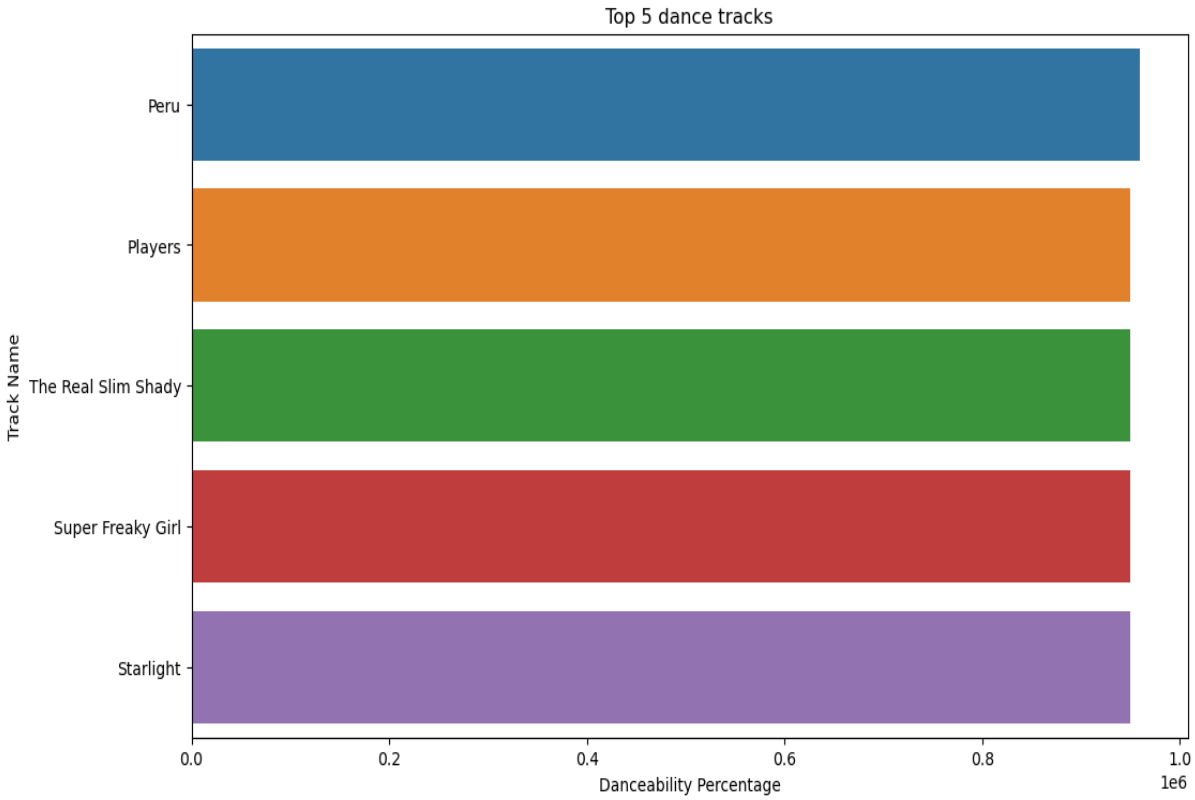
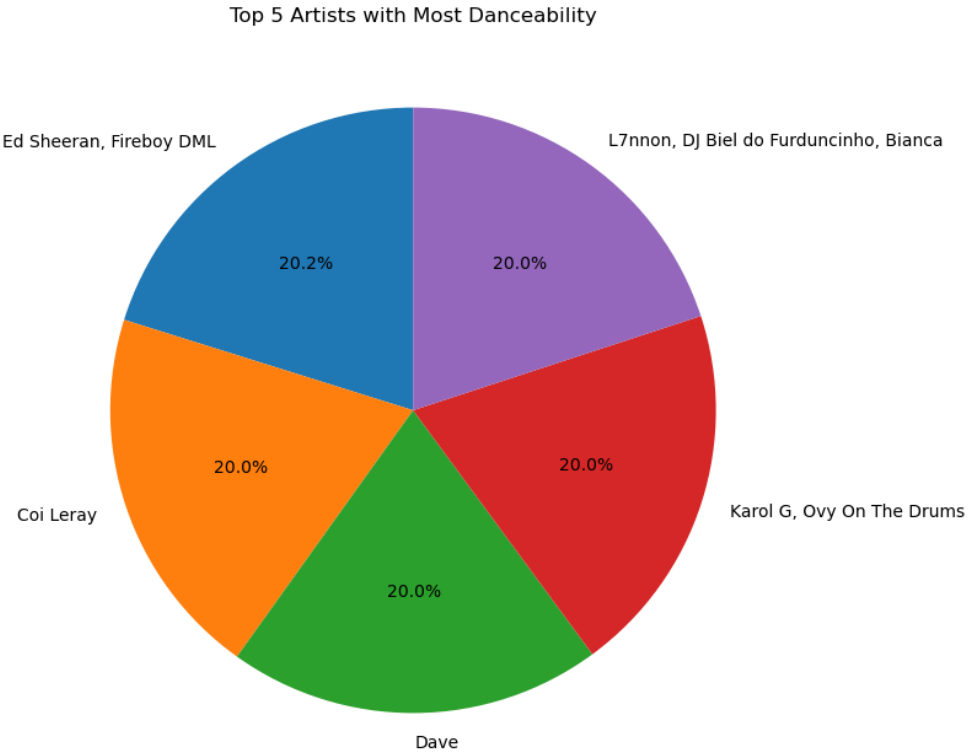
SPOTIFY 2023 SONG ANALYSIS

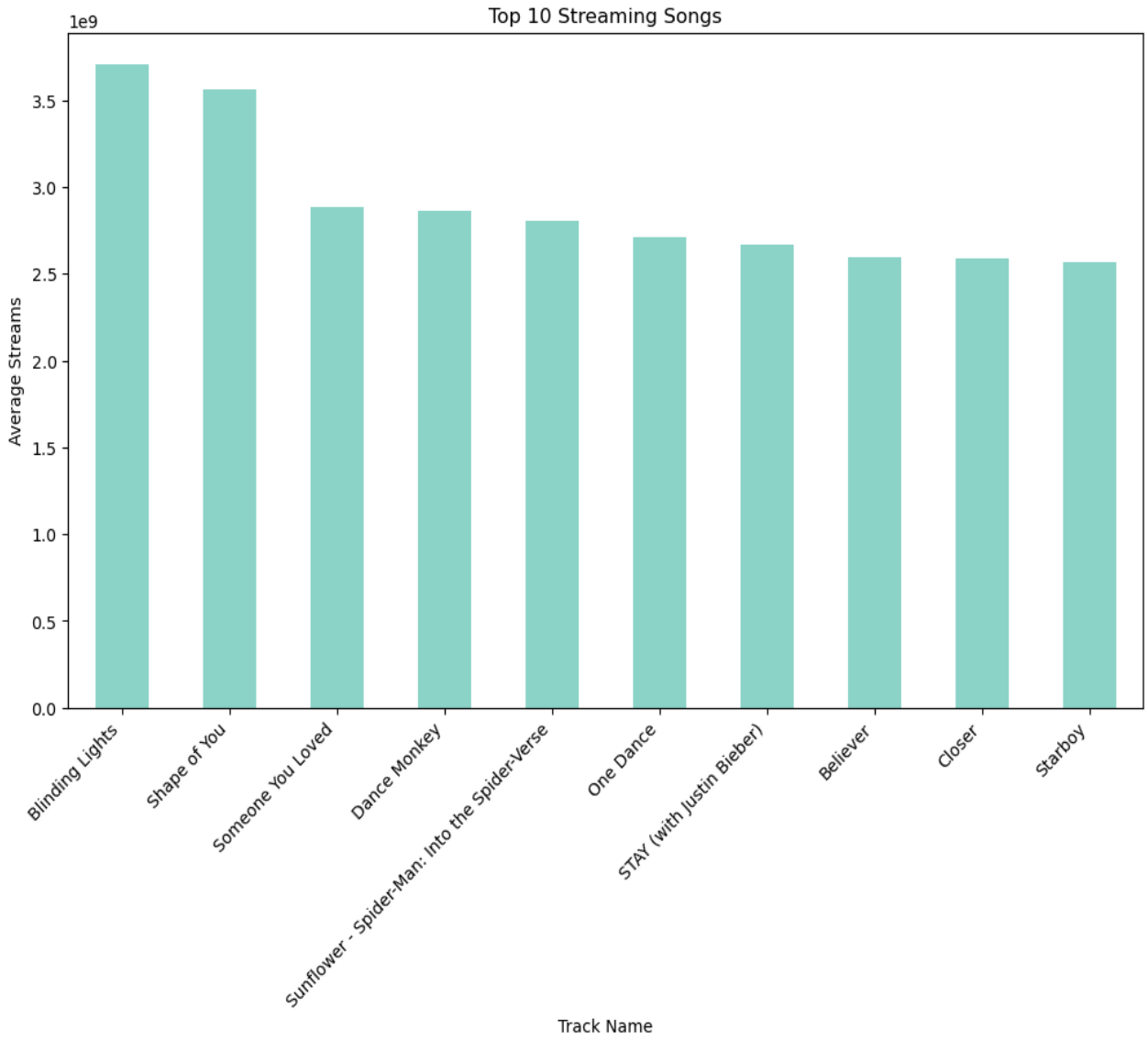


Most streamed song

Blinding Lights(track name) by The Weekend(artist).


Top 5 dance tracks
Peru
Players
The Real Slim Shady
Super Freaky Girl
Starlight
Track Name
0.0
0.2
0.4
0.6
0.8
1.0
1e6
Danceability Percentage

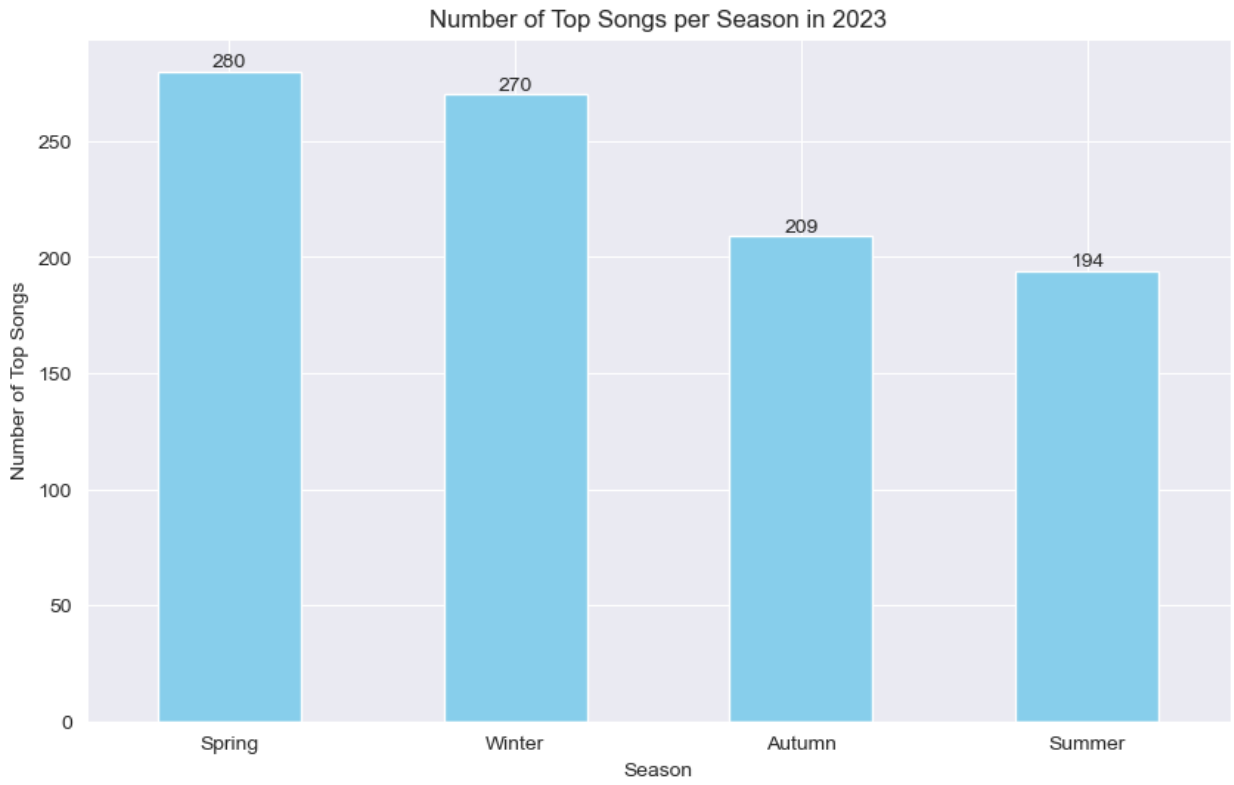

Top 5 Artists with Most Danceability
Ed Sheeran, Fireboy DML
L7nnon, DJ Biel do Furduncinho, Bianca
20.2%
20.0%
20.0%
20.0%
Coi Leray
Karol G, Ovy On The Drums
20.0%
Dave



Correlation Heatmap of Song Characteristics

It shows that **valence** is highly co-related with **danceability** and **energy.**

Seasonal Hits:



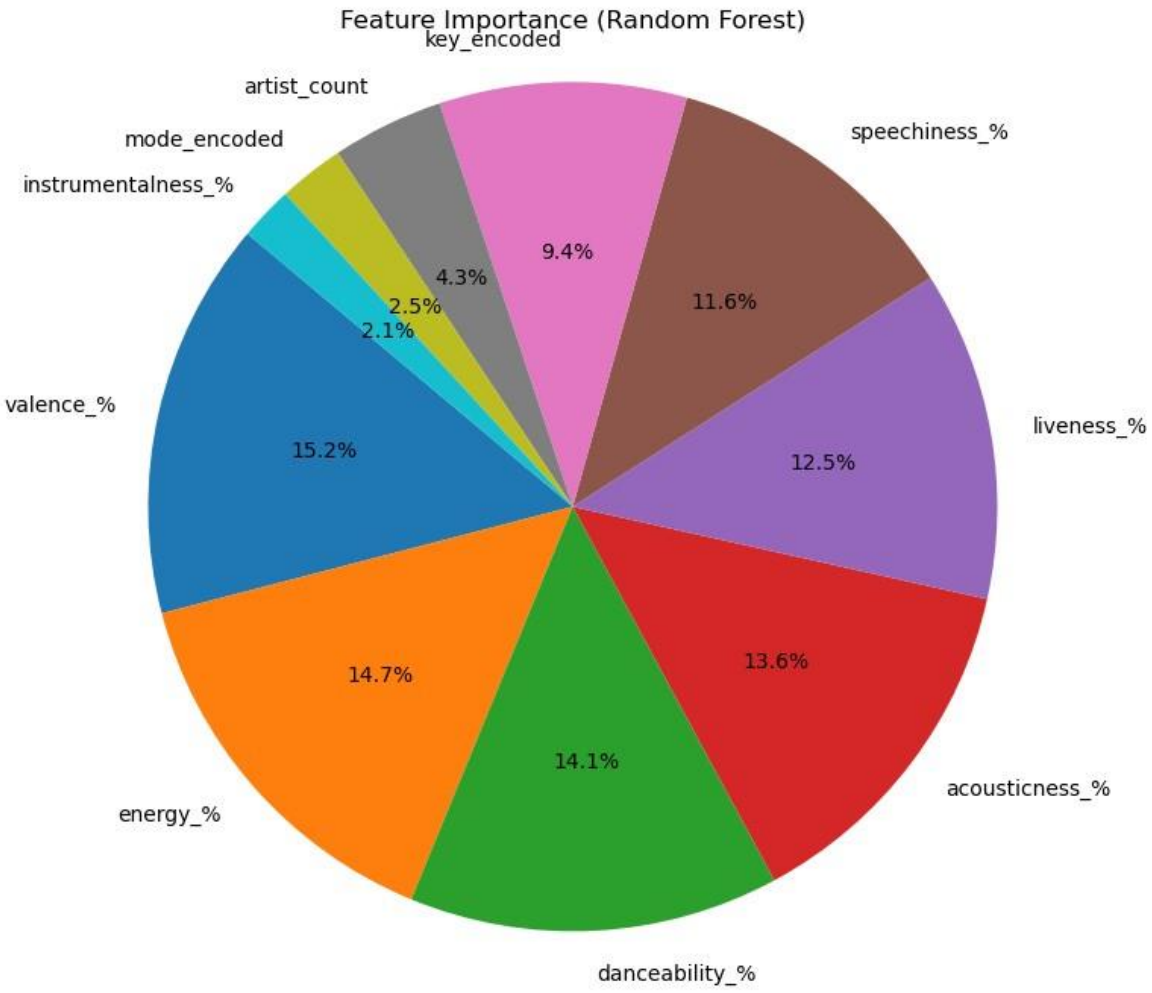
Then different classification model is applied to know which feature of the song play a big role on the number of the streaming of a song.

Classification Model	Accuracy
Decision Tree	0.5654
Cross Validation	0.51

Random Forest	0.5812
SVM	0.5445
Naive Bayes (Gaussian)	0.5384
Logistic Regression	0.5692

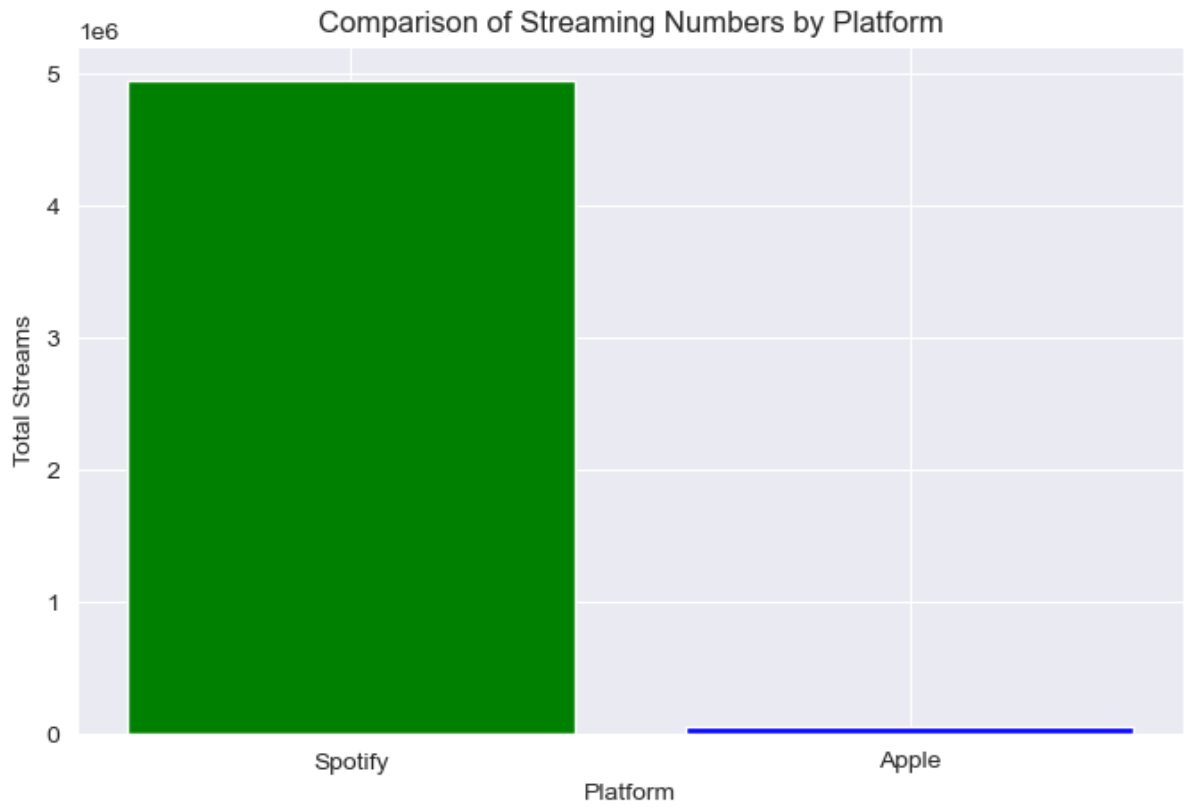
Though the accuracy scores are fluctuated within the different parameters, **Logistic Regression** and **Random Forest** are usually giving highest accuracy among them (almost 0.6).

So, we take the random forest classification model as best model.



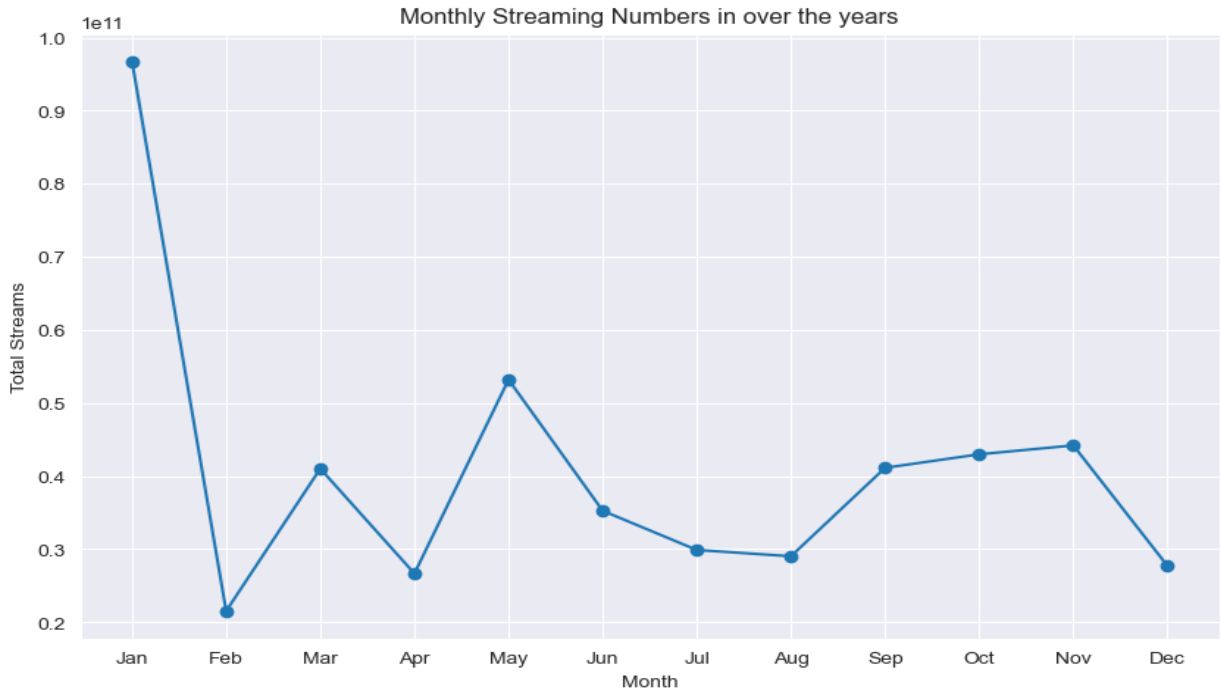
So, we see **valence** and **energy** are the most crucial features for a song in terms of number of streaming of song.

Platform comparison between spotfy and apple



Here we see that most people use spotify as their song app.

Monthly streaming numbers over the years(Mostly 2023)

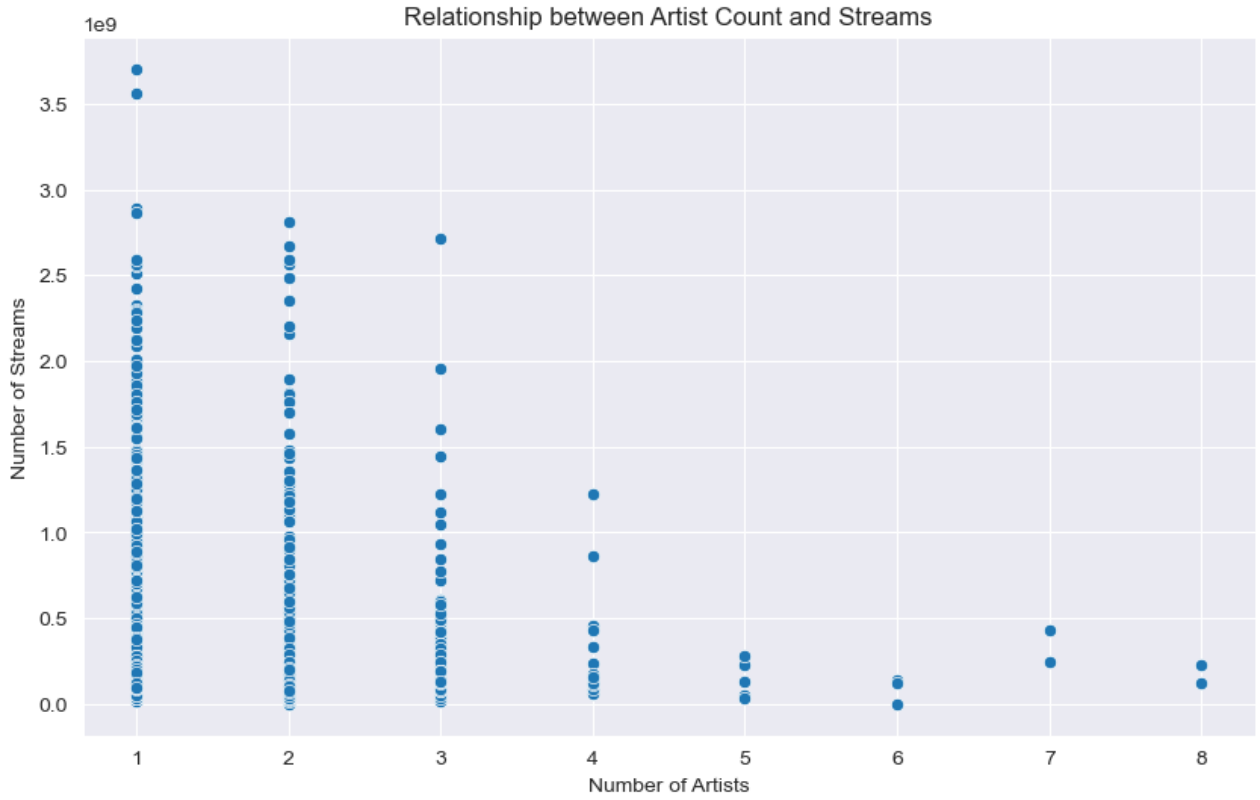


Sentiment Analysis

We use **textblob** to analyse the sentiment .

	track_name	track_name_sentiment	artist(s)_name	artist_name_sentiment
0	Seven (feat. Latto) (Explicit Ver.)	0.0	Latto, Jung Kook	0.0
1	LALA	0.0	Myke Towers	0.0
2	vampire	0.0	Olivia Rodrigo	0.0
3	Cruel Summer	-1.0	Taylor Swift	0.0
4	WHERE SHE GOES	0.0	Bad Bunny	-0.7

Artist Influence on streaming number



Artist name and aggregating the streaming numbers

In [35]:
```
# Grouping the data by artist name and aggregating the streaming numbers
artist_streams = df_tracks.groupby('artist(s)_name')['streams'].sum().sort_values(ascending=False)

# Displaying the top 10 artists by total streaming numbers
for artist, streams in artist_streams.head(10).items():
    print(f"{artist}: {streams}")
```

```
The Weeknd: 14185552870.0
Taylor Swift: 14053658300.0
Ed Sheeran: 13908947204.0
Harry Styles: 11608645649.0
Bad Bunny: 9997799607.0
Olivia Rodrigo: 7442148916.0
Eminem: 6183805596.0
Bruno Mars: 5846920599.0
Arctic Monkeys: 5569806731.0
Imagine Dragons: 5272484650.0
```