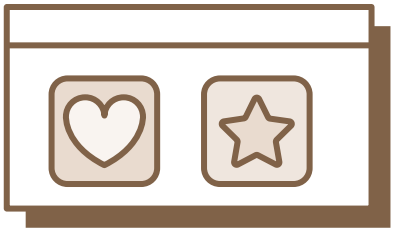
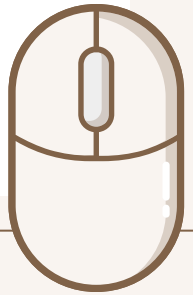




Decoding Popularity

A competitive analysis of Taylor Swift and BTS on Spotify

Ong See Gek Cheryl, Keertana Arun Vasan, Li Yufei

Two artists, two different corners of the world, two distinct music cultures....

Taylor Swift

American Singer and songwriter - 14 Grammy, 39 Billboard music awards, and Many more

BTS

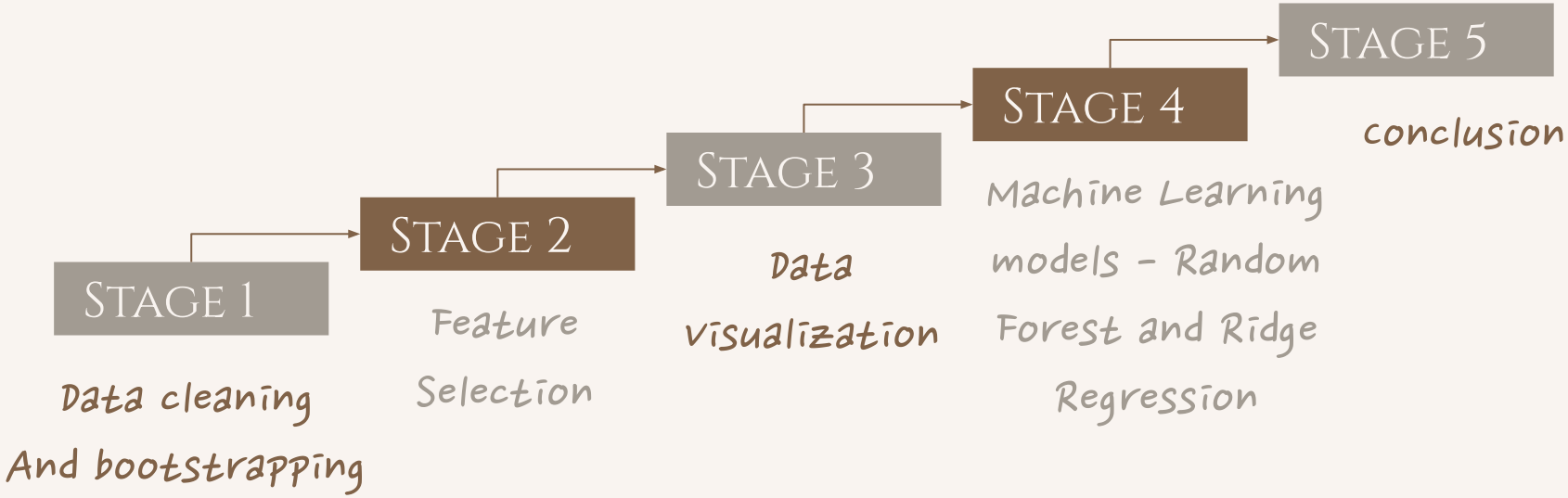
South Korean boy band - achieved numerous firsts for Korean artists, has topped Billboard charts

Problem Statement

How do the musical characteristics of Taylor Swift and BTS on Spotify differ, and what insights can these differences provide into their respective popularity and audience preferences?

Let's see!

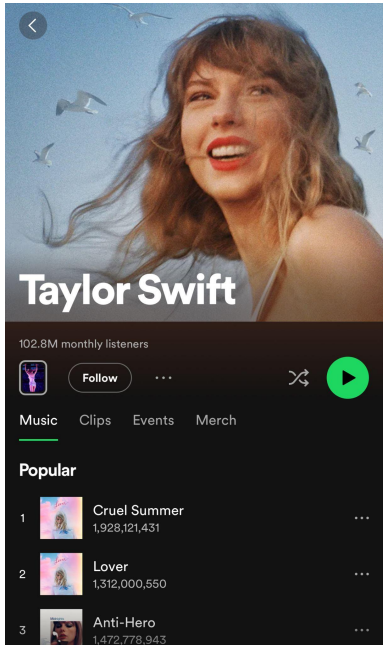
NAVIGATING THE PROJECT PIPELINE

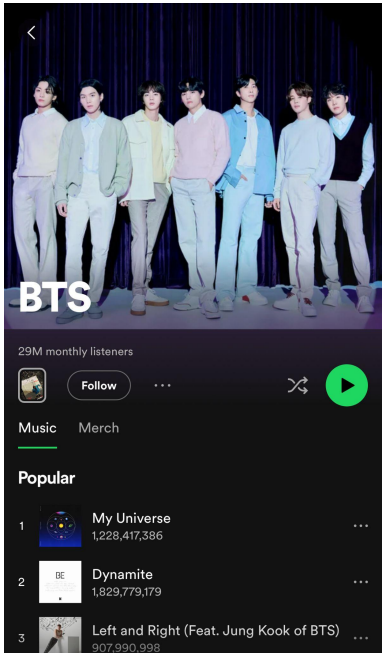


DATA COLLECTION AND PREPARATION

About our Datasets

consists of data from Spotify's API on all albums listed on Spotify for Taylor Swift and BTS ; collected from Kaggle.





kaggle

	artist_name	album_title
0	BTS	2 Cool 4 Skool
1	BTS	2 Cool 4 Skool
2	BTS	2 Cool 4 Skool
3	BTS	2 Cool 4 Skool
4	BTS	2 Cool 4 Skool
...	...	...
525	BTS	O!RUL8,2?
526	BTS	Skool Luv Affair (Special Addition)
527	BTS	Skool Luv Affair (Special Addition)
528	BTS	Wings
529	BTS	Skool Luv Affair (Special Addition)

Result

Data Cleaning

our datasets were quite clean with no outliers or missing values

we noticed a disparity in the number of rows between the two datasets

BOOTSTRAPPING - Randomly sample rows from the BTS dataset and add duplicates or slight variations to increase the overall number of rows of BTS dataset

```
Data columns (total 15 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   artist_name       530 non-null    object
 1   track_name        530 non-null    object
 2   album_name        530 non-null    object
 3   acousticness      530 non-null    float64
 4   danceability      530 non-null    float64
 5   energy            530 non-null    float64
 6   instrumentalness  530 non-null    float64
 7   liveness          530 non-null    float64
 8   loudness          530 non-null    float64
 9   speechiness       530 non-null    float64
 10  tempo             530 non-null    float64
 11  valence           530 non-null    float64
 12  popularity        530 non-null    int64
 13  duration_ms       530 non-null    int64
 14  year              530 non-null    int64
```

```
Data columns (total 17 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   artist_name       224 non-null    object
 1   album_title       224 non-null    object
 2   track_name        224 non-null    object
 3   track_popularity  224 non-null    int64
 4   danceability      224 non-null    float64
 5   energy            224 non-null    float64
 6   key               224 non-null    int64
 7   loudness          224 non-null    float64
 8   mode              224 non-null    int64
 9   speechiness       224 non-null    float64
 10  acousticness      224 non-null    float64
 11  instrumentalness  224 non-null    float64
 12  liveness          224 non-null    float64
 13  valence           224 non-null    float64
 14  tempo             224 non-null    float64
 15  duration_ms       224 non-null    int64
 16  year              224 non-null    int64
```

FEATURE SELECTION

1 Speechiness

2 Loudness

3 Acousticness

▲ For taylor swift

Process where we automatically select those features in our data that contribute most to the prediction variable or output in which we are interested

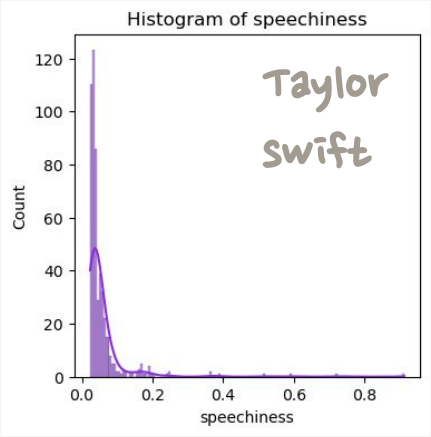
ExtraTreesclassifier from sklearn.ensemble

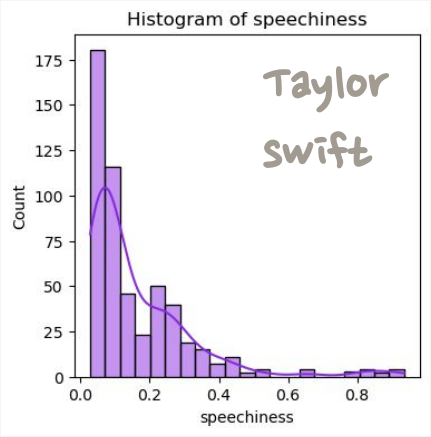
Speechiness 1

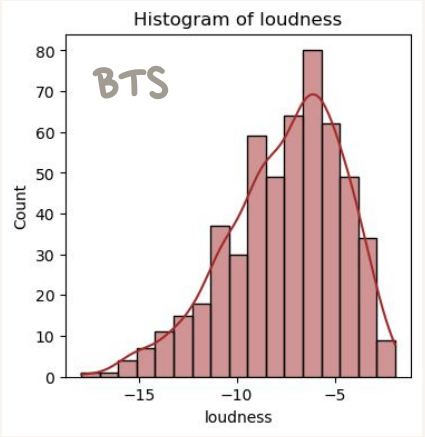
Instrumentalness 2

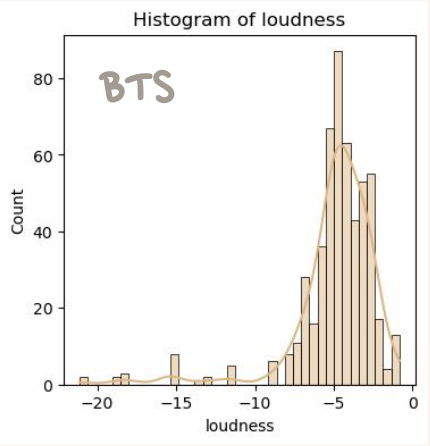
Loudness 3

▲ For BTS





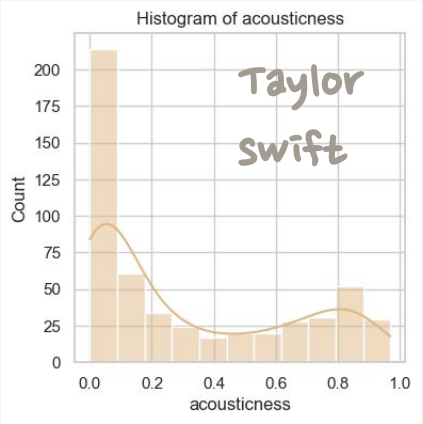


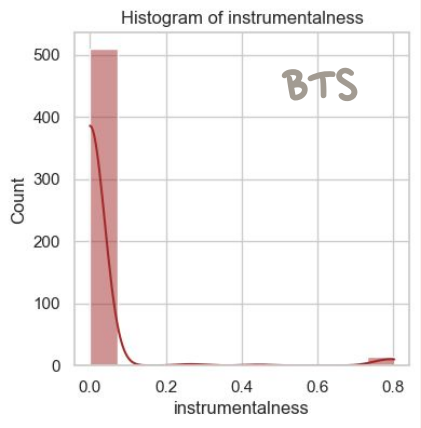


Taylor Swift's music is characterized by low speechiness, consistent loudness, and a mix of low and high acousticness. Her louder songs tend to be more energetic, while her acoustic tracks are less loud and energetic. On the other hand, BTS's music features moderate speechiness, loudness that correlates strongly with energy, and a tendency for tracks with more spoken words to be more acoustic. Their tracks with higher instrumental content also tend to be more danceable.

Some comparative visualizations between Taylor Swift and BTS

For detailed observations and inferences, refer to ReadMe file

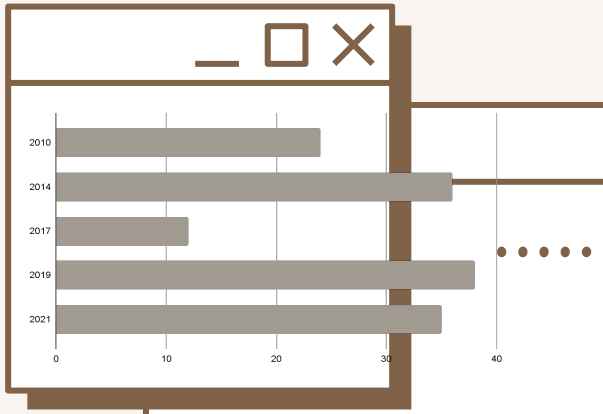




Instrumentalness is not a defining feature for Taylor Swift's music based on the provided analysis. Taylor Swift's music shows a significant portion of low acousticness tracks, but also a minor peak around 0.6, indicating the presence of more acoustic tracks.BTS's music doesn't highlight acousticness as a defining feature.

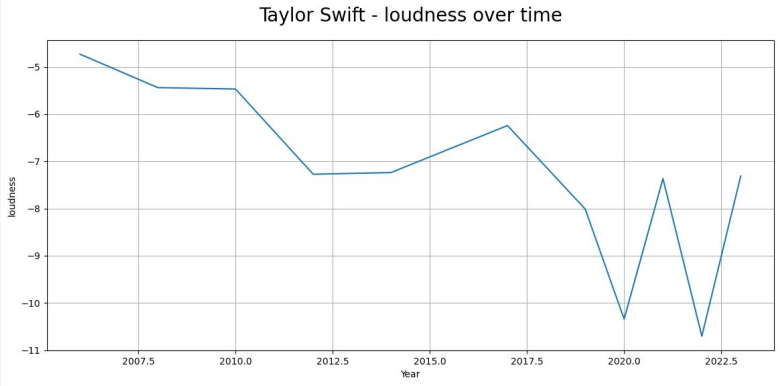
Some comparative visualizations between Taylor Swift and BTS

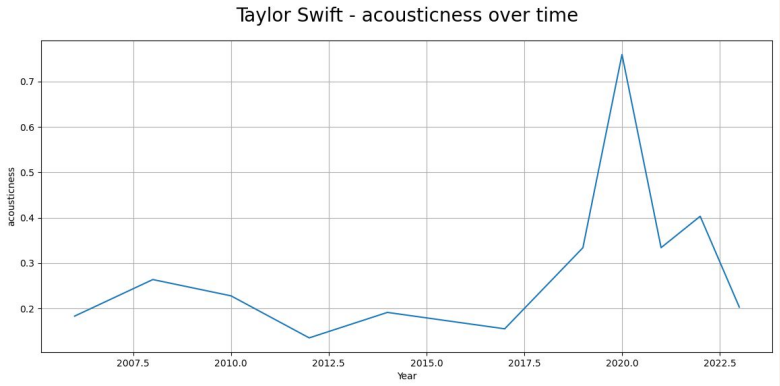
For detailed observations and inferences, refer to ReadMe file

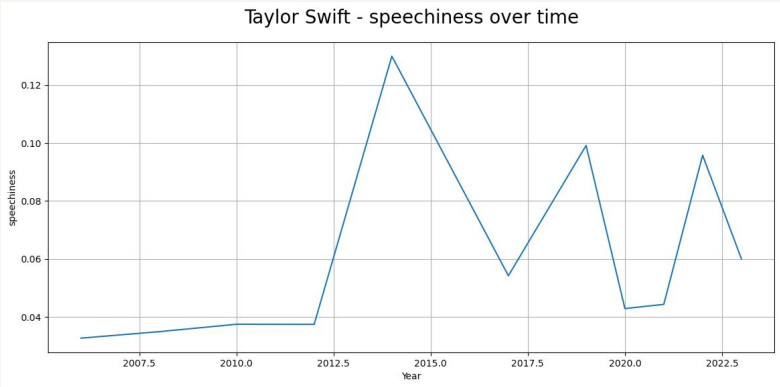


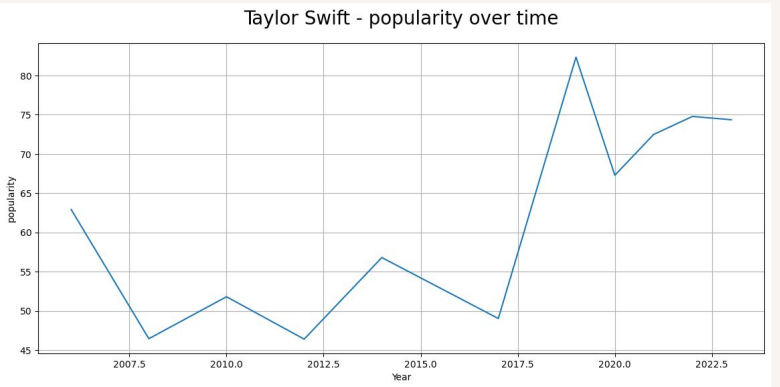
TREND ANALYSIS

How have the musical attributes of Taylor Swift's and BTS's songs have changed over time?

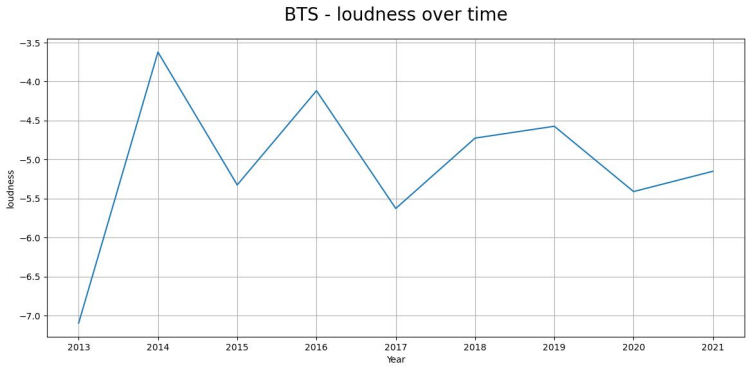


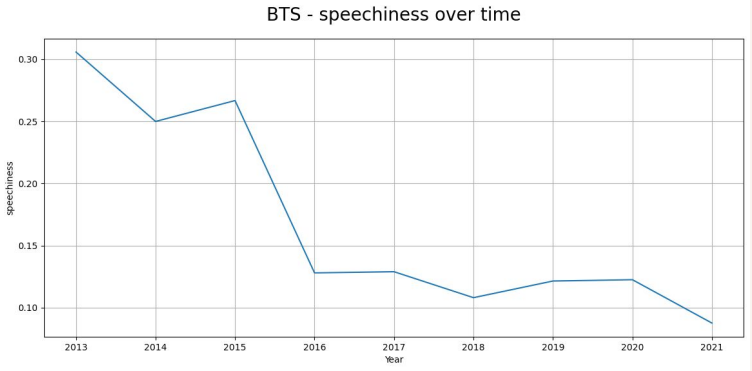


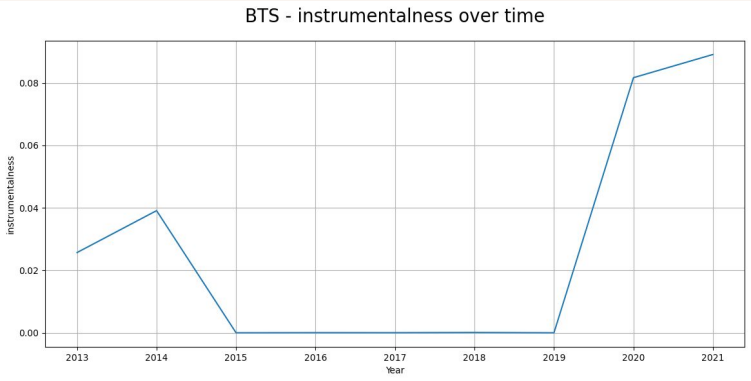


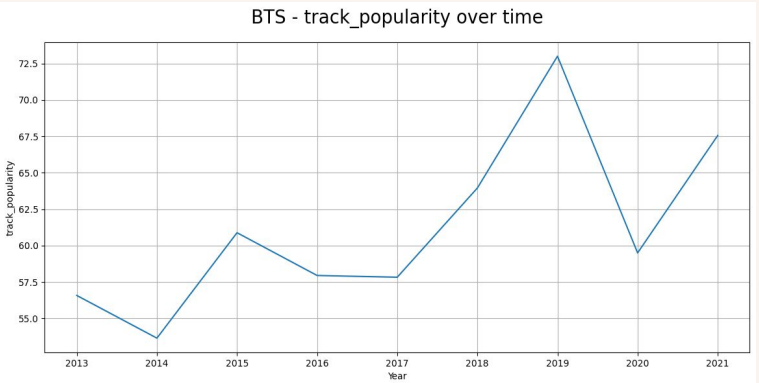


Some of the trend analysis Graphs....









Some of the trend analysis Graphs....

INSIGHTS FROM TREND ANALYSIS

1

Taylor Swift's music has seen a gradual decrease in energy, popularity, key, liveliness, loudness, danceability, valence, and tempo, suggesting a shift towards more introspective music.

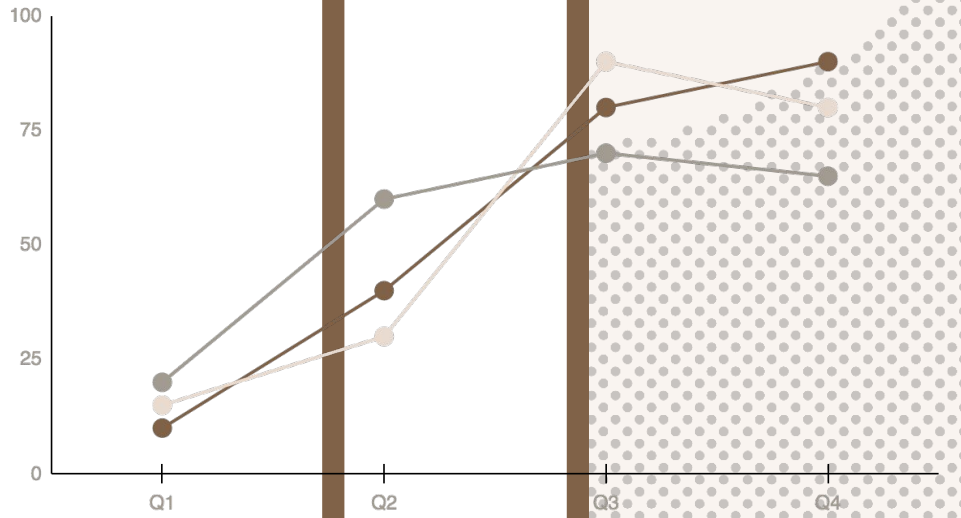
However, her music has seen an increase in danceability, speechiness, acousticness, instrumentalness,& duration, indicating a shift towards more complex song structures & lyrics.

2

BTS's music has seen a decrease in danceability, energy, and speechiness, reflecting their shift from a more hip-hop and rap-centric style to a more pop-centric style.

However, their music has seen an increase in popularity, loudness, mode, and instrumentalness, indicating a greater emphasis on musical elements and production techniques.

MACHINE LEARNING MODELS



We've implemented 3 machine learning models

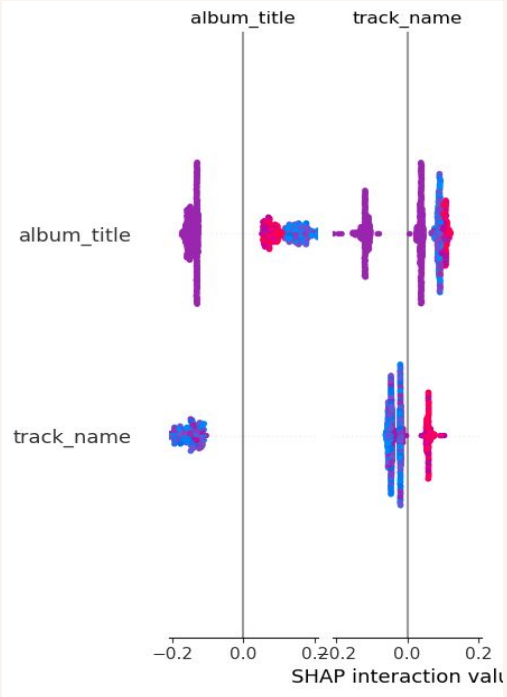
Random Forest classifier was able to perfectly predict the artist of each song in the test set, with an accuracy, precision, and recall of 1.0. Also plotted SHAP Summary Plot

Ridge Regression to predict the 'energy' of a song based on various features.

R squared value: 0.81

Adjusted R squared value: 0.79

Adjusted R >> R squared

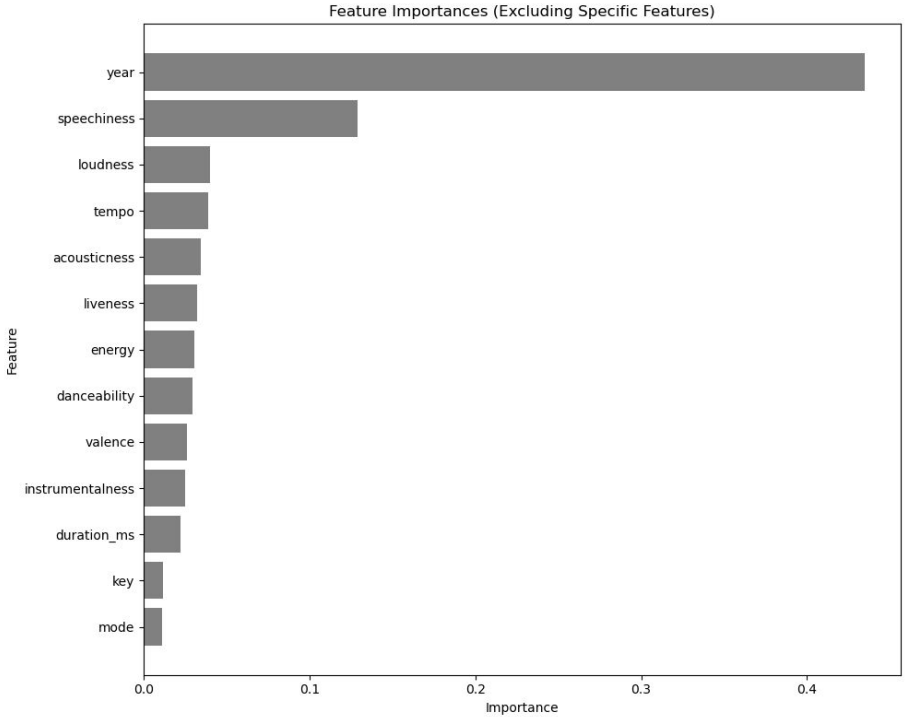


Random Forest Regressor for predicting artists' track popularity performs best among these models.

	R^2 score	RMSE
Linear Regression	0.34	11.39
Decision Tree	0.50	8.47
Random Forest Regressor	0.60	7.57

From Random Forest Regressor , we select features that are most significant in predicting a song's popularity.

Year, Speechness, Loudness



We use quartiles to identify three levels of popularity: High, Medium, and Low and try to see which type of song received relatively higher popularity.

popularity_category	loudness	speechiness	year
low	-5.475544	0.127470	2015.630208
medium	-8.817278	0.072419	2020.111111
high	-8.029961	0.060408	2020.683594

1. Newly-released songs recieve higher popularity.
2. Songs that struck a harmonious balance between speechness and loudness tended to resonate more, receiving higher popularity.

CONCLUSION

The comparative analysis of Taylor Swift and BTS on Spotify reveals that Taylor Swift's music is appealing to audiences who favor **clear vocals and energetic tunes**. While BTS's tracks are catering to those who enjoy **danceable and vibrant music**.

Through analysis, it helps us to understand why Taylor Swift and BTS are popular: they **adapt their music styles to what listeners enjoy**, which includes **releasing fresh tracks** and **focusing on the right mix of music elements** to keep their fans engaged (loudness and speechness especially).

INSIGHTS

Release date is an essential feature. Music producers and artists may consider **recycling old songs** by taking into account some **audio feature composition** that makes music highly popular, e.g. speechness, loudness, tempo.