SPOTIFY DATA SCIENCE PROJECT

NEKHA DURAISAMY

GEETHIKA THOTA

PURPOSE

- OUR GOAL IS TO FIND OUT WHAT ATTRIBUTES OF A SONG MAKE IT POPULAR
- THIS WAY, ARTISTS CAN SEE IF THERE'S ANYTHING THEY COULD ADD/CHANGE TO THEIR TRACKS TO INCREASE THEIR CHANCES OF MAKING A POPULAR SONG

PREDICTIONS

- ENERGY WILL BE THE MOST CORRELATED WITH TRACK POPULARITY
- SHORTER SONGS WILL BE MORE POPULAR
- SONGS WITH A HIGHER DANCEABILITY WILL HAVE A HIGHER TRACK POPULARITY

CLEANING DATA

1. WE FOCUSED ON TRACK POPULARITY SO WE GOT RID OF OUTLIERS
2. WE THEN SMOOTHED THE DATA
3. ENDED WITH 32,833 DATA POINTS

ANALYSIS

PART 1. CORRELATION MATRIX

PART 2. DURATION VS. TRACK POPULARITY

PART 3. INSTRUMENTALNESS VS. TRACK POPULARITY

PART 4. ENERGY VS. TRACK POPULARITY

PART 5. SUMMARY

CORRELATION MATRIX

```
track_popularity                0.740655
danceability                    0.066879
energy                         -0.118525
key                            -0.003058
loudness                        0.061112
mode                            0.015152
speechiness                     0.008755
acousticness                    0.108448
instrumentalness               -0.168792
liveness                       -0.056735
valence                         0.039659
tempo                          -0.006633
duration_ms                    -0.191834
smoothed_track_popularity       1.000000
Name: smoothed_track_popularity, dtype: float64
```

- NOT MUCH CORRELATION IN GENERAL
- DURATION, INSTRUMENTALNESS, ENERGY HAD THE HIGHEST CORRELATION
- WE FOCUSED OUR PLOTS ON THOSE ATTRIBUTES

VARIABLES

DURATION_MS

DURATION OF SONG IN MILLISECONDS

INSTRUMENTALNESS

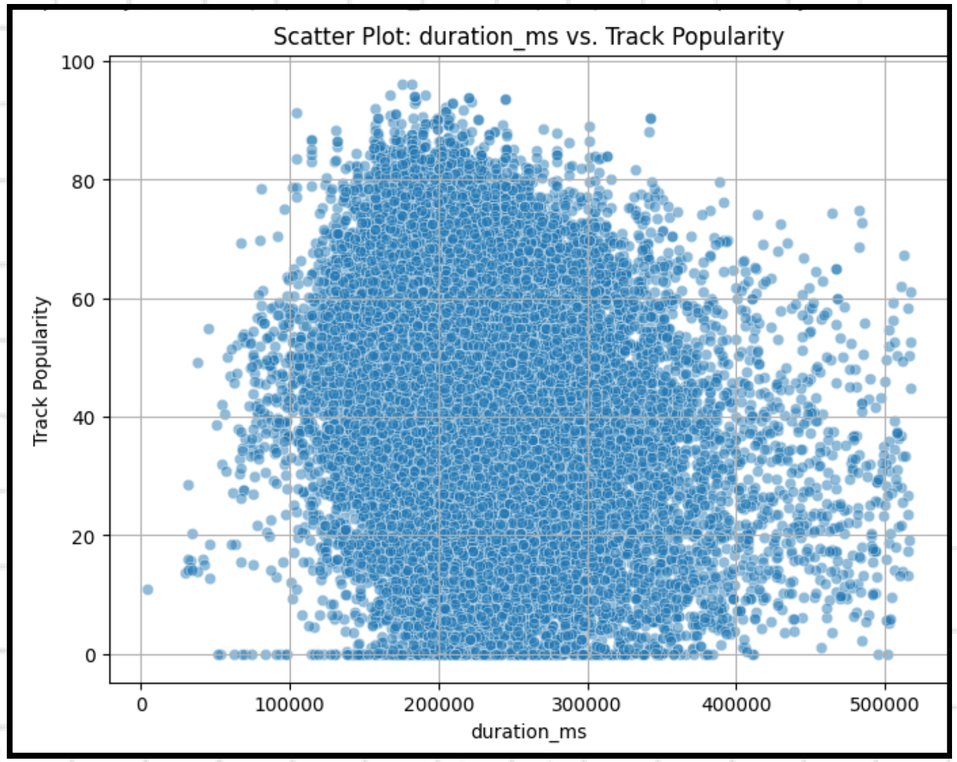
PREDICTS WHETHER A TRACK CONTAINS NO VOCALS SUCH AS ‘OOH’ AND ‘AAH’ SOUNDS; THE CLOSER TO 1.0, THE TRACK CONTAINS NO VOCAL CONTENT

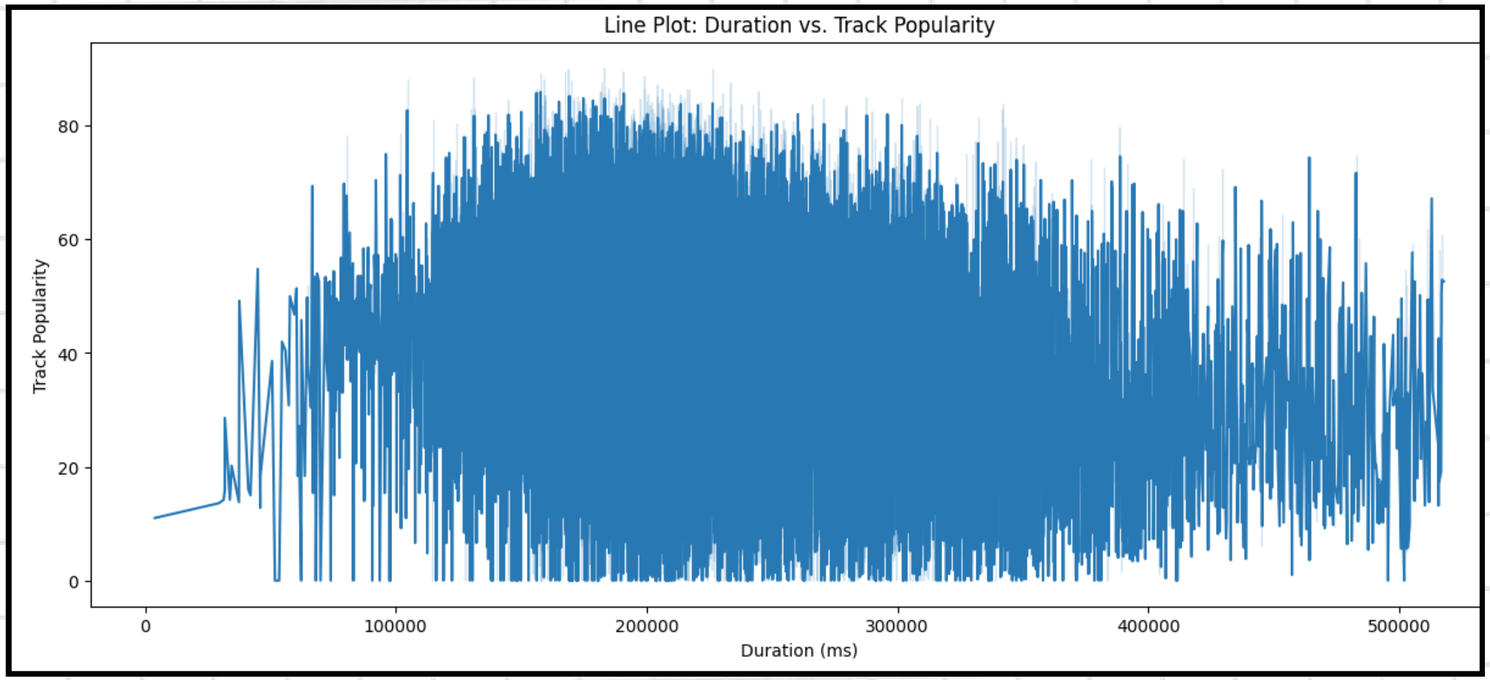
ENERGY

ENERGY REPRESENTS A PERCEPTUAL MEASURE OF INTENSITY AND ACTIVITY SUCH DYNAMIC RANGE, PERCEIVED LOUDNESS, TIMBRE, ETC.

DURATION VS. TRACK POPULARITY

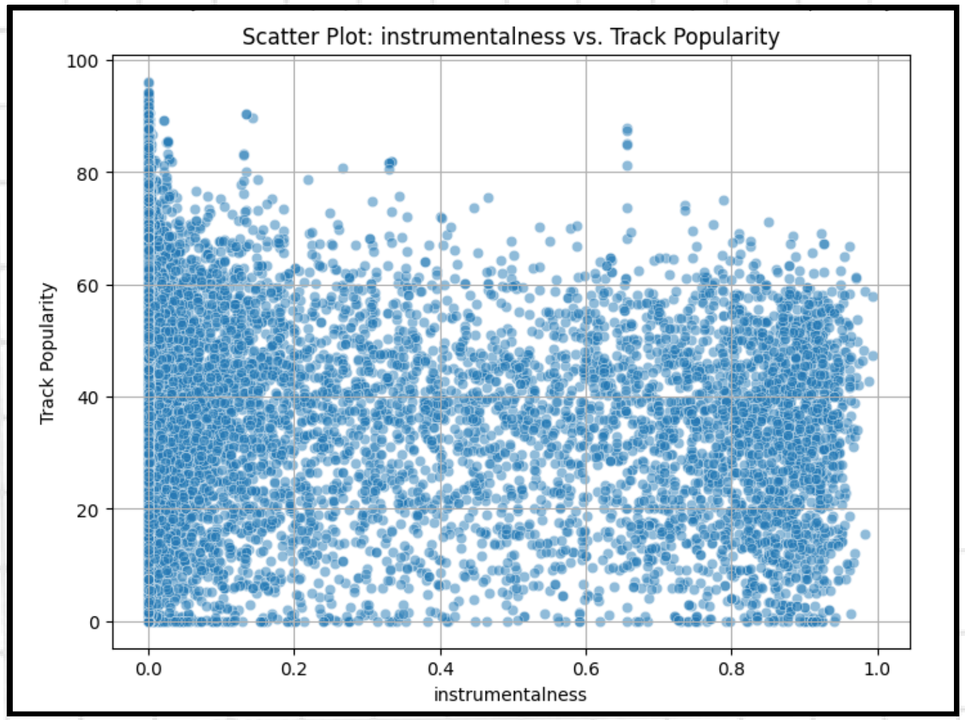
TRACKS WERE ALSO MORE POPULAR WHEN THE DURATION WAS BETWEEN 100,000 AND 300,000 MS

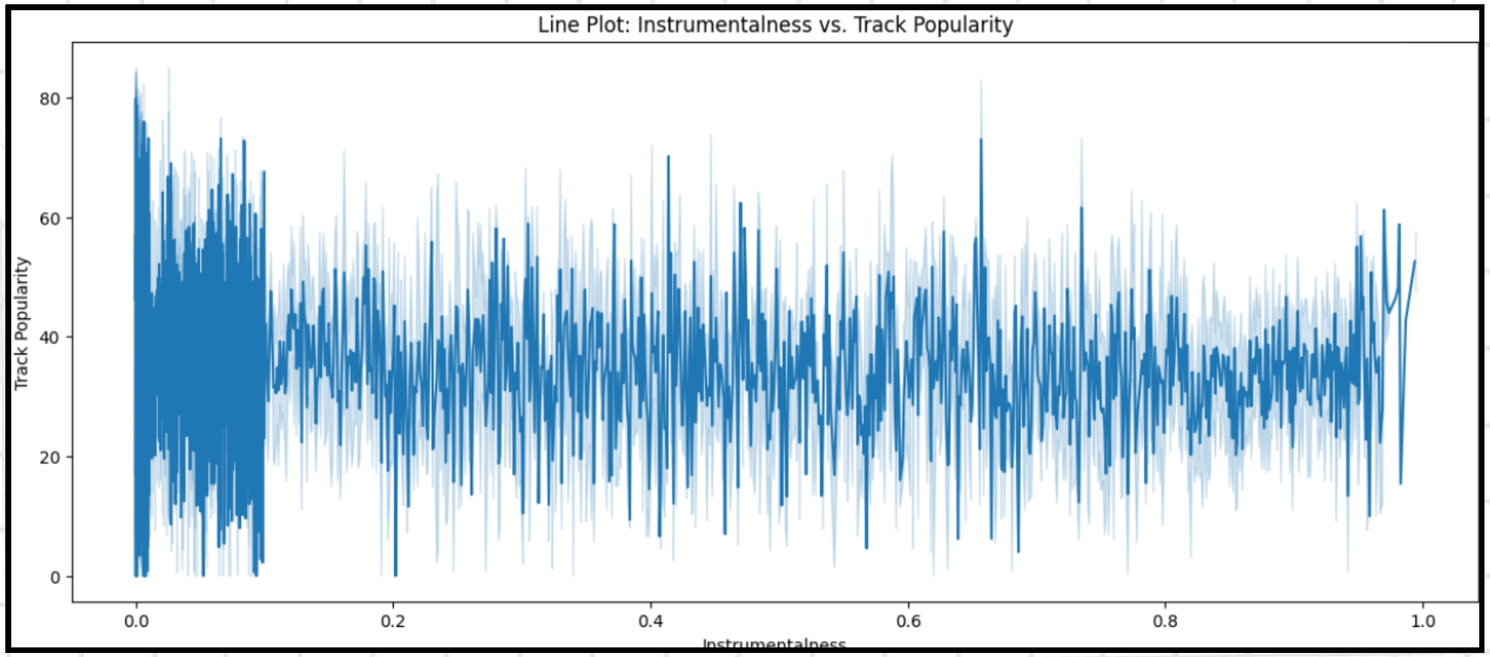


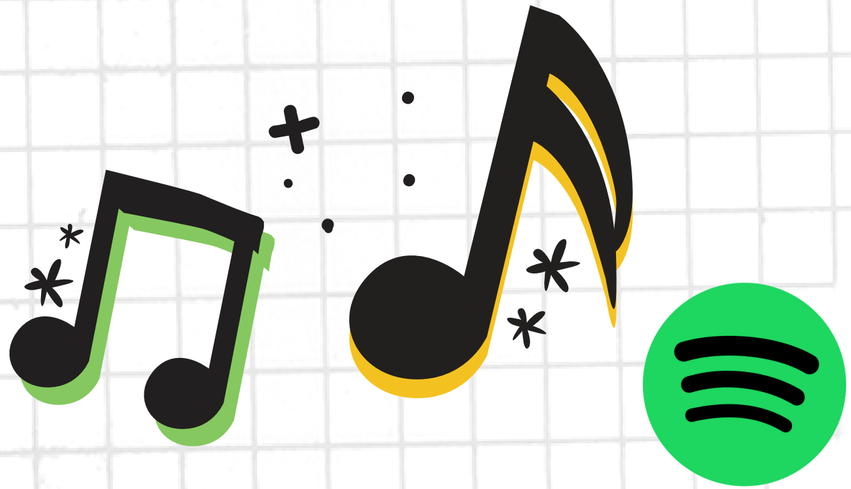


INSTRUMENTALNESS VS. TRACK POPULARITY

SONGS WITH INSTRUMENTALNESS LOWER THAN 0.1 HAVE SHOWN TO BE MORE POPULAR

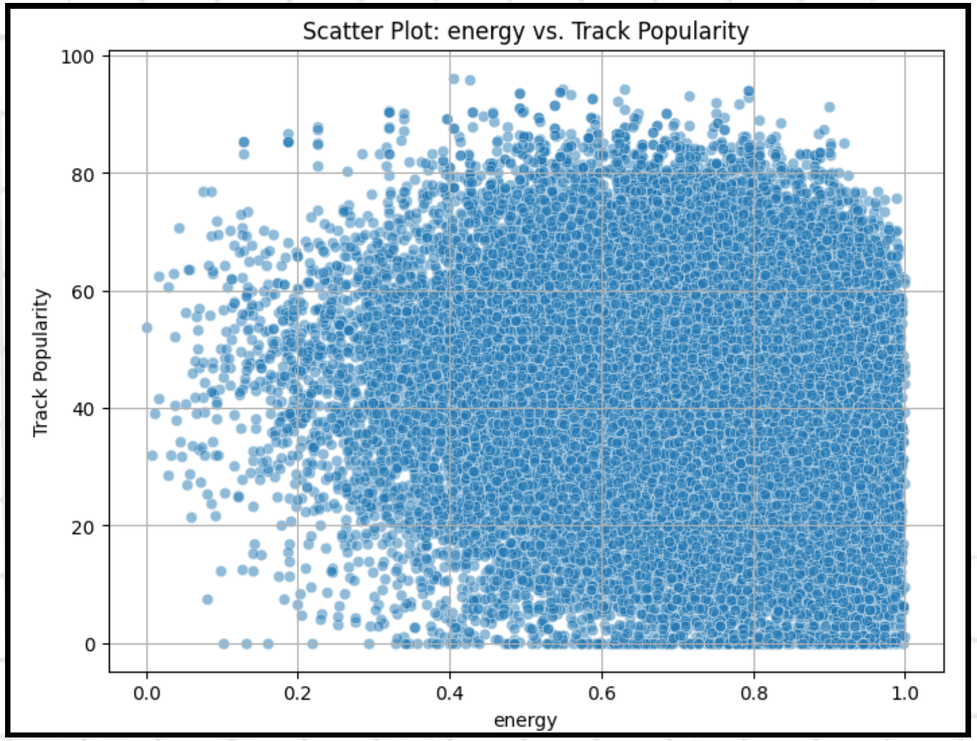


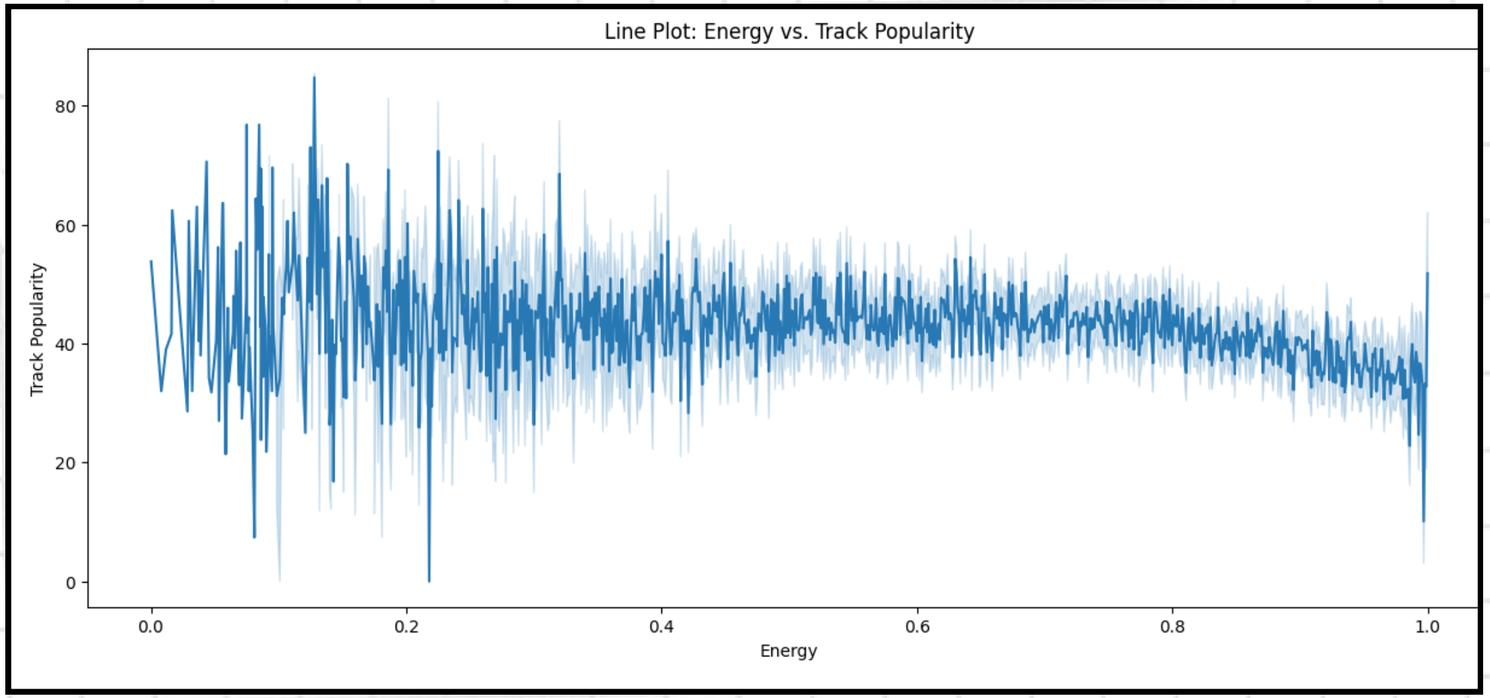




ENERGY VS. TRACK POPULARITY

WE NOTICED THAT TRACKS HAD MORE POPULARITY WHEN ENERGY WAS BETWEEN 0.1 AND 0.25





RESULTS

- DURATION: BETWEEN 100,000-300,000 MS
- INSTRUMENTALNESS: CLOSE TO 0.1
- ENERGY: BETWEEN 0.1 AND 0.25

SUMMARY

- OUR DATA HELPS IN UNDERSTANDING WHAT SPOTIFY USERS ENJOYS WHEN LISTENING TO MUSIC, WHICH CAN HELP ARTISTS BETTER CATER TO USER PREFERENCES
- OUR DATA APPLIES TO ALL TYPES OF MUSIC GENRES, SO INSIGHTS FROM OUR DATA IS NOT LIMITING TO JUST ONE
- WE WERE SURPRISED THAT DANCEABILITY HAD ALMOST NO CORRELATION WITH TRACK POPULARITY


AND THAT'S A WRAP!
THANK YOU