**1 Import cleaned data and required libraries**

<class 'pandas.core.frame.DataFrame'>

Index: 1600 entries, 0 to 1609

Data columns (total 18 columns):

# Column Non-Null Count Dtype

--- ------ -------------- -----

0 number\_row 1600 non-null int64

1 name 1600 non-null string

2 album 1600 non-null string

3 release\_date 1600 non-null datetime64[ns]

4 track\_number 1600 non-null int64

5 id 1600 non-null string

6 uri 1600 non-null string

7 acousticness 1600 non-null float64

8 danceability 1600 non-null float64

9 energy 1600 non-null float64

10 instrumentalness 1600 non-null float64

11 liveness 1600 non-null float64

12 loudness 1600 non-null float64

13 speechiness 1600 non-null float64

14 tempo 1600 non-null float64

15 valence 1600 non-null float64

16 popularity 1600 non-null int64

17 duration\_ms 1600 non-null int64

dtypes: datetime64[ns](1), float64(9), int64(4), string(4)

memory usage: 237.5 KB

None

number\_row int64

name string[python]

album string[python]

release\_date datetime64[ns]

track\_number int64

id string[python]

uri string[python]

acousticness float64

danceability float64

energy float64

instrumentalness float64

liveness float64

loudness float64

speechiness float64

tempo float64

valence float64

popularity int64

duration\_ms int64

dtype: object

number\_row release\_date track\_number acousticness \

count 1600.000000 1600 1600.000000 1600.000000

mean 806.173125 1992-03-26 00:45:54 8.616875 0.248336

min 0.000000 1964-04-16 00:00:00 1.000000 0.000009

25% 403.750000 1970-09-04 00:00:00 4.000000 0.058025

50% 807.000000 1986-03-24 00:00:00 7.000000 0.182000

75% 1209.250000 2017-12-01 00:00:00 11.000000 0.401000

max 1609.000000 2022-06-10 00:00:00 47.000000 0.939000

std 465.232161 NaN 6.566928 0.224416

danceability energy instrumentalness liveness loudness \

count 1600.000000 1600.000000 1600.000000 1600.000000 1600.000000

mean 0.469292 0.793088 0.162978 0.490834 -6.924214

min 0.104000 0.242000 0.000000 0.021900 -18.104000

25% 0.362750 0.674000 0.000218 0.153000 -8.964250

50% 0.458000 0.846000 0.013550 0.378000 -6.512000

75% 0.579000 0.945000 0.176250 0.893000 -4.601750

max 0.887000 0.999000 0.996000 0.998000 -1.014000

std 0.141980 0.178300 0.275519 0.348999 2.882639

speechiness tempo valence popularity duration\_ms

count 1600.000000 1600.000000 1600.000000 1600.000000 1600.000000

mean 0.068433 126.192089 0.584046 20.833750 257274.625625

min 0.023200 46.525000 0.000000 0.000000 21000.000000

25% 0.036500 107.443000 0.405750 13.000000 190842.750000

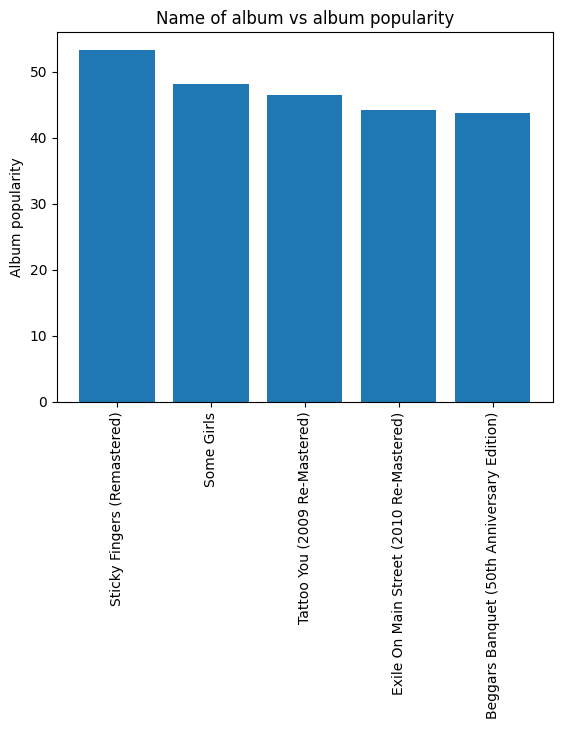
50% 0.051050 124.501500 0.585000 20.000000 243319.500000

75% 0.085925 142.420000 0.780000 27.000000 295213.250000

max 0.335000 216.304000 0.974000 80.000000 831960.000000

std 0.046909 29.243959 0.230305 12.443488 103810.838821

# 2. Recommend albums and songs based on popularity



\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Based on the figure, two recommended albums are:

first album: Sticky Fingers (Remastered)

second album: Some Girls

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Based on popularity of songs, five recommended songs are:

name popularity

1403 Paint It, Black 80

862 Start Me Up - Remastered 2009 76

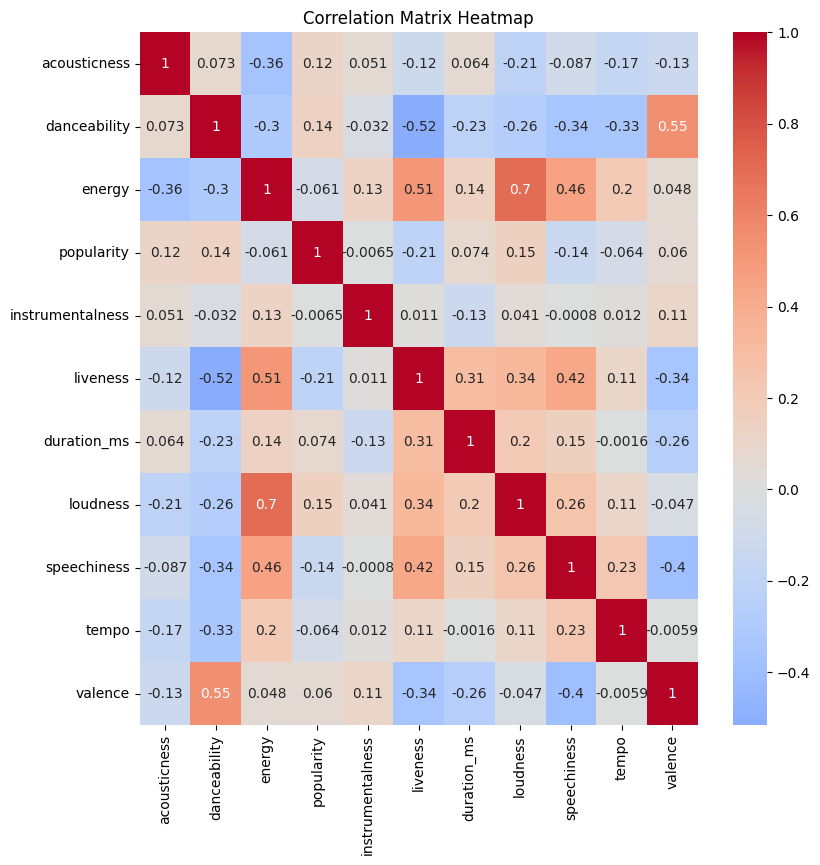
1248 Gimme Shelter 76

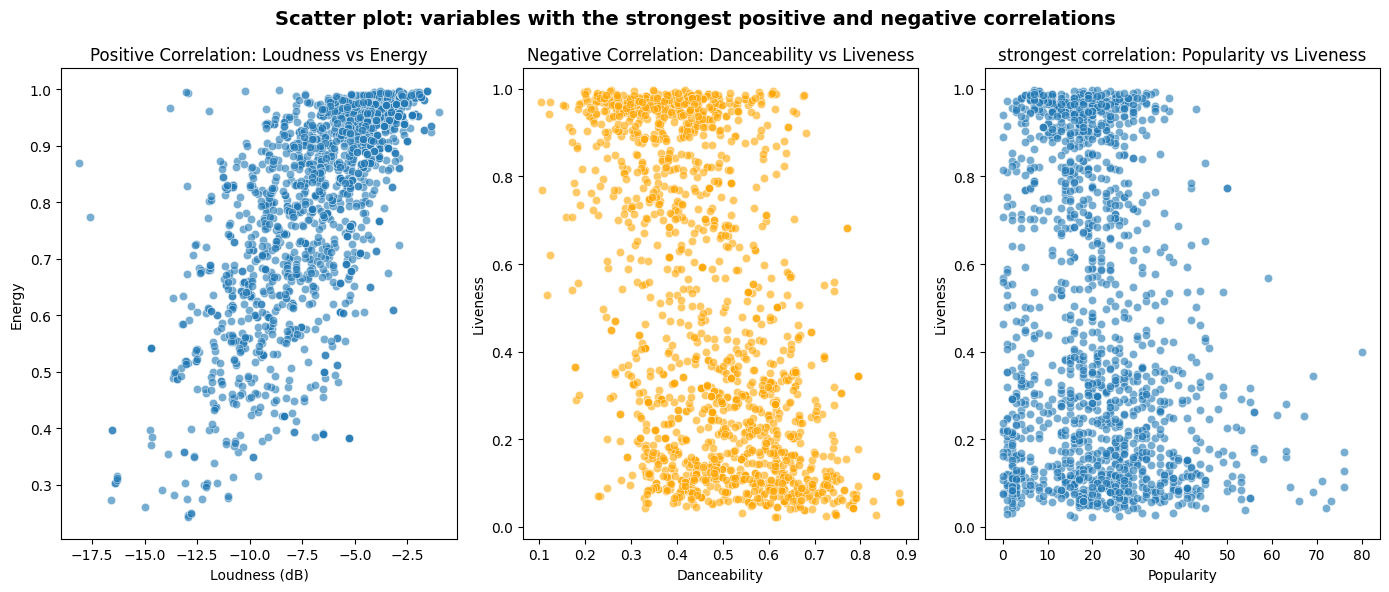
1472 (I Can't Get No) Satisfaction - Mono Version 76

1257 Sympathy For The Devil - 50th Anniversary Edition 73

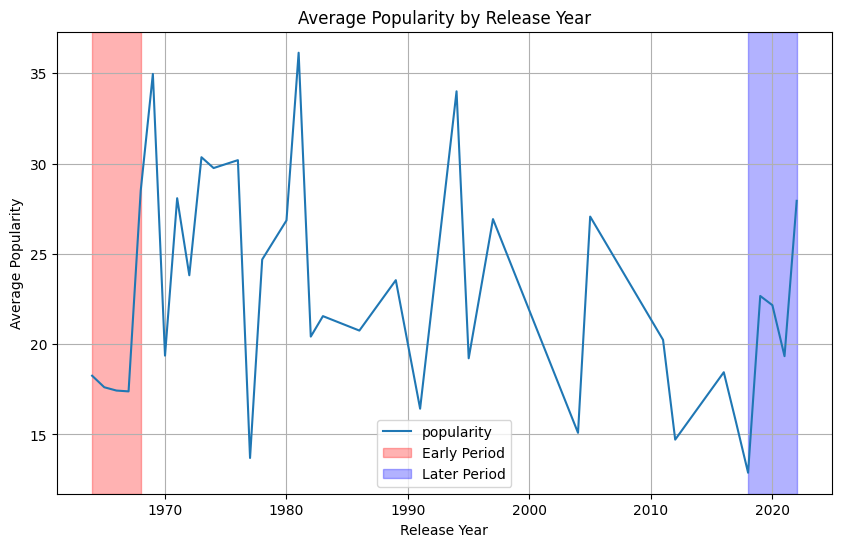
\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

# 3. Examine correlations between song features, especially popularity





# 3. Analyze popularity trends over time



Early period (1964-1968): 19.83

Later period (2018-2022): 20.99

Difference: 1.16