**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

# 2. Scale features for clustering

First five rows of df after scaling:

[[-0.73964281 -0.04433332 1.12156646 3.02440969 1.26448377 -2.07818522

0.88639398 -0.28018266 -2.40558757 0.97802594 -2.01038598]

[ 0.8409549 -1.00955628 0.96447843 0.25422465 1.34760451 0.73608852

0.1592254 0.180022 -1.15554953 1.05841438 -0.03952292]

[ 0.74734928 -0.5868309 0.98691958 0.860543 1.33327335 0.68993582

0.99301694 0.13251016 -1.17726666 1.05841438 0.05671098]

[ 1.42041824 -0.70660309 1.07668416 -0.59132804 1.15843317 0.48207516

2.6563351 0.23266471 -1.8982754 0.89763749 0.46835736]

[ 0.67603071 -1.17160101 0.98691958 -0.38876325 1.36193567 0.63371975

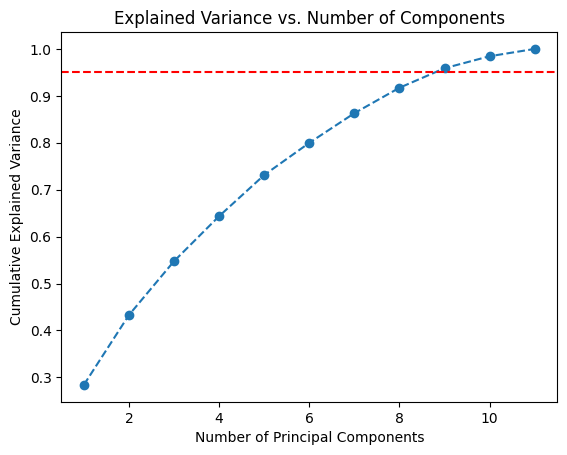
0.52387592 0.14848427 -1.64201325 0.89763749 0.46089916]]

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

Number of features of the dataset to be scaled: 11

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

# 3. Dimension reduction and visualization

****

Components needed for 95% variance: 9

Original shape: (1600, 11)

Reduced shape: (1600, 9)

