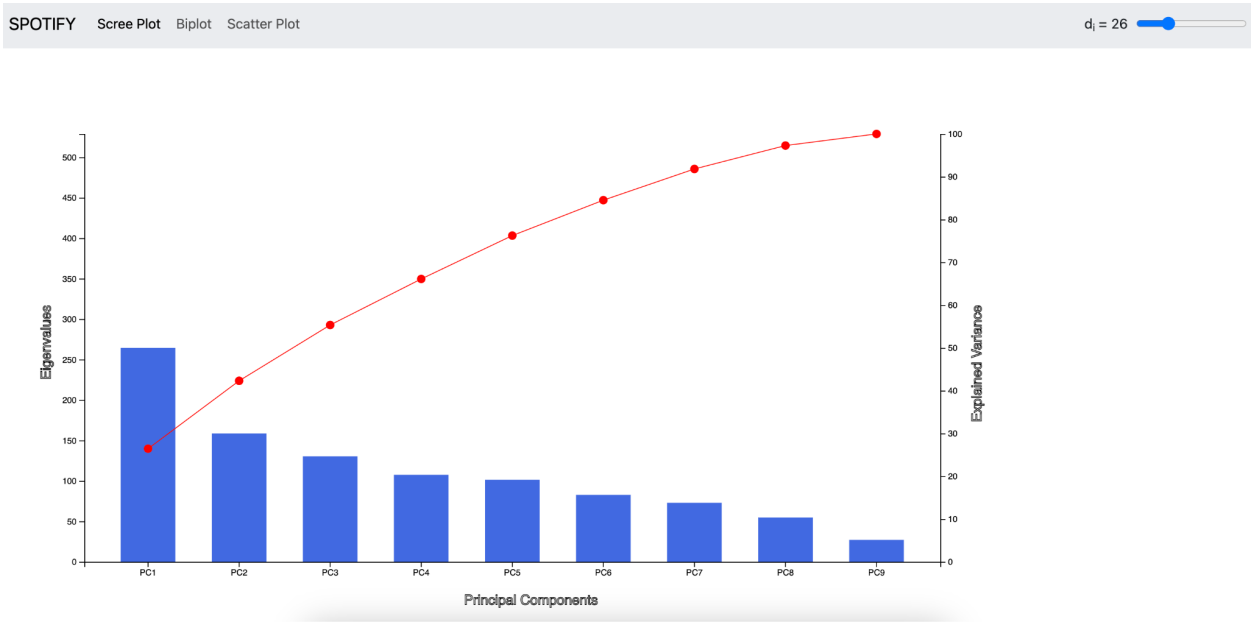
Dataset Source:
https://www.kaggle.com/datasets/muhmores/spotify-top-100-songs-of-20152019?select=Spotify+2010+-+2019+Top+100+Songs.xlsx

Attributes:

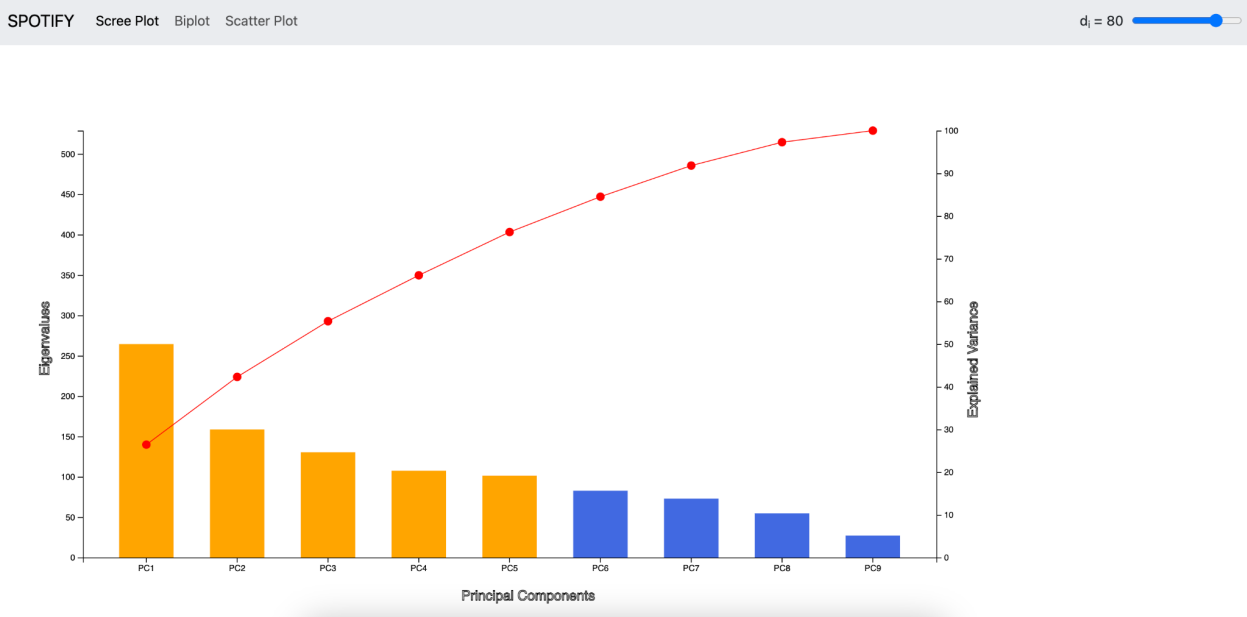
Artist	Song's artist
Genre	Genre of song
Released	Year the song was released
Tempo	Beats Per Minute - The tempo of the song
Energy	How energetic the song is
Danceability	How easy it is to dance to the song
Intensity	Decibel - How loud the song is
Live Likelihood	How likely the song is a live recording
Positiveness	How positive the mood of the song is
Duration	Duration of the song
Acoustic	How acoustic the song is
Speech Focus	The more the song is focused on spoken word
Popularity	Popularity of the song (not a ranking)
Top Year	Year the song was a top hit
Artist Type	Tells if artist is solo, duo, trio, or a band

Scree Plot Observation

The first and most interesting observation is that no one component contributes excessively to the song's popularity. We can see that the first principal component only explains around 27% of the data. This means it is not a predominant factor.

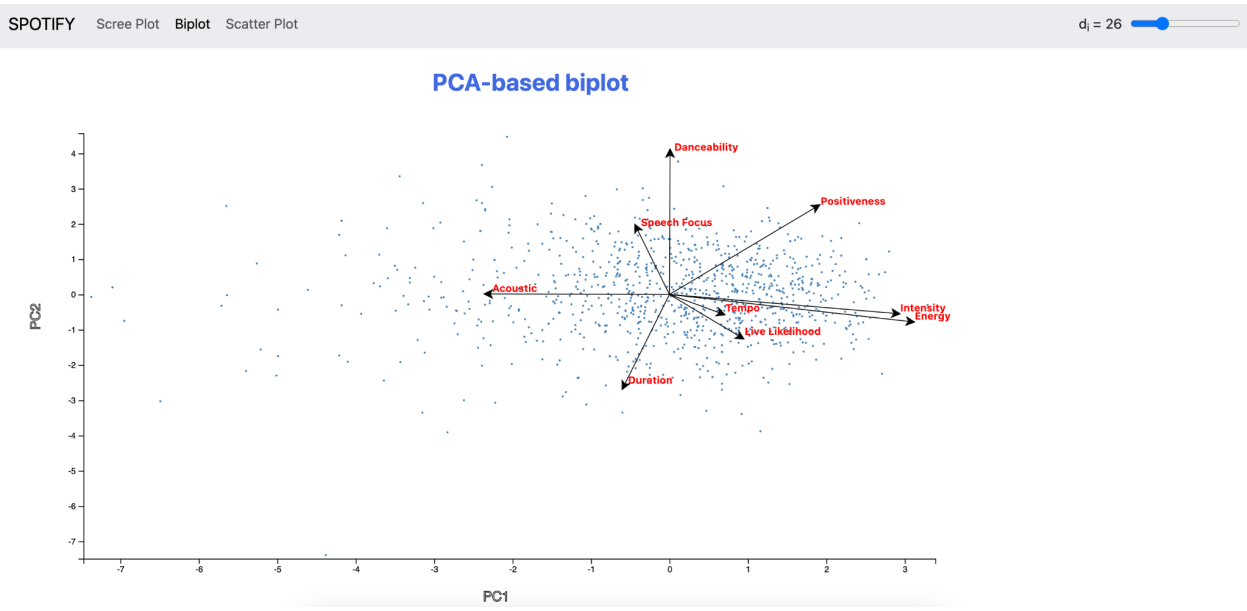


If we increase the minimum desired explainability (using the slider) to 80%, we can see that a total of five principal components are required to explain the data.



Biplot Observations

We can observe from the biplot that the attributes “Acoustic”, “Intensity” and “Energy” have high loading on principal component 1, while “Danceability”, “Speech Focus” and “Duration” have high loading on principal component 2.



Scatterplot Observations

We can see from the scatter plot that the four attributes with the highest squared sum of PCA loadings are “Speech Focus”, “Tempo”, “Energy”, and “Danceability”.