

# SOCIAL WEB ANALYTICS: SPOTIFY PROJECT

Presented by Group 11:

- Xuan Bach Tran
- Thanh Thong Tran
- Celine Nguyen
- Dinh Tuan Tran



# Project goals

- K-Means Clustering for tracks genres within a playlist.
- Hypothesis testing for the correlation between track popularity and tracks' acousticness.
- Spotify Network Graph for tracks' genres and collaboration network between artists.

# Data structure

## Variables (audio features)

**VALENCE**  
(numeric)

**DANCEABILITY**  
(numeric)

**LIVENESS**  
(numeric)

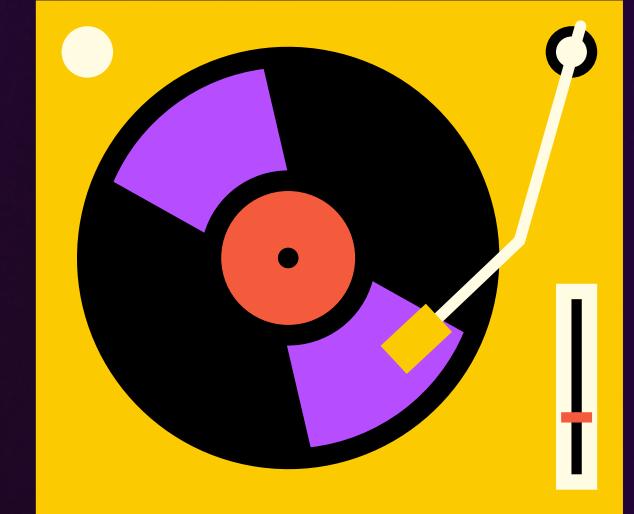
**INSTRUMENTALNESS**  
(numeric)

**SPEECHINESS**  
(numeric)

**ACOUSTICNESS**  
(numeric)

**ENERGY**  
(numeric)

**TEMPO**  
(numeric)

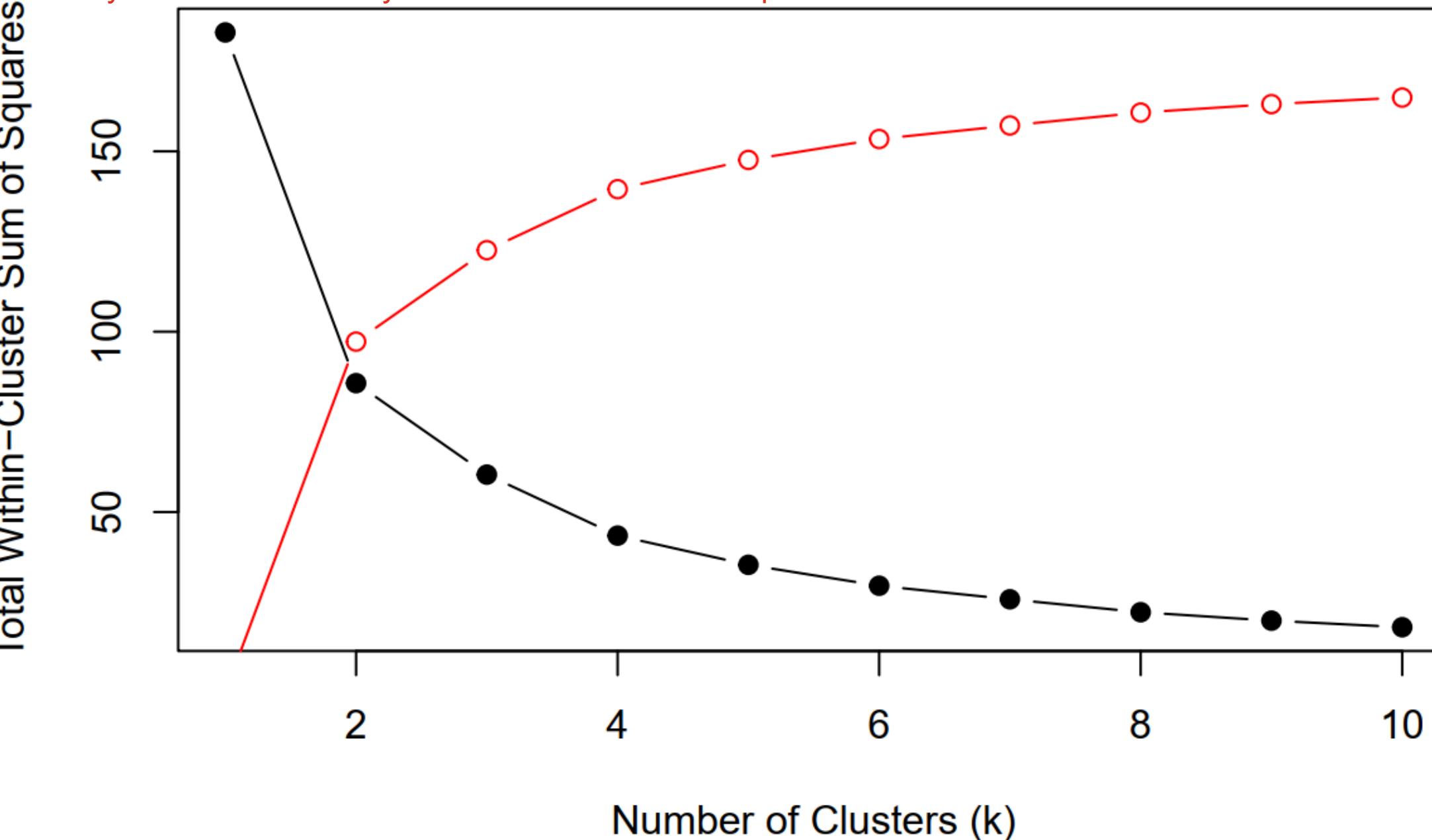


# K-Means Clustering

## Finding the optimal number of initial clusters

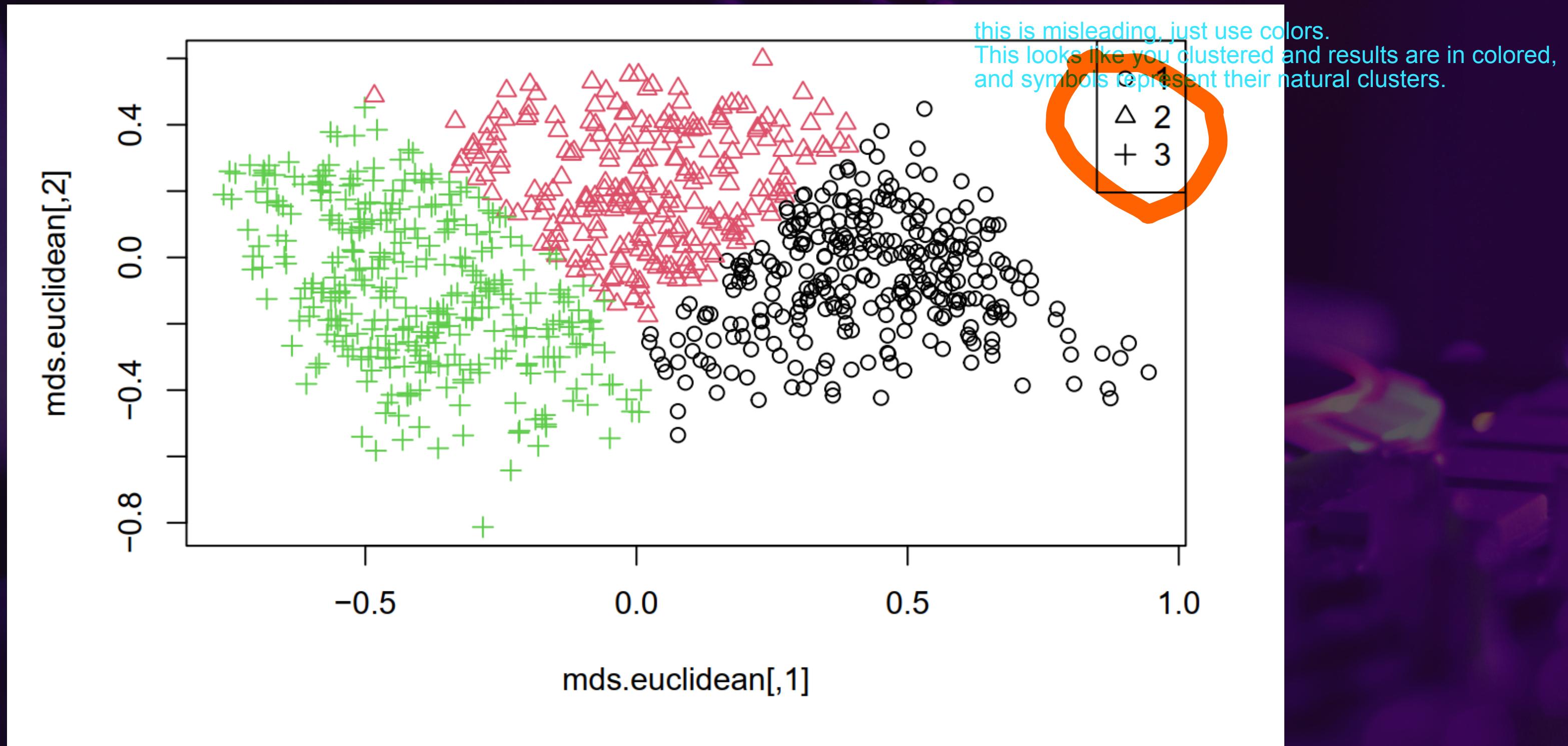
### Elbow Method

This slide is not necessary, elbow is only to decide how many clusters, you can mention in your talk but slide is not required.



# K-Means Clustering

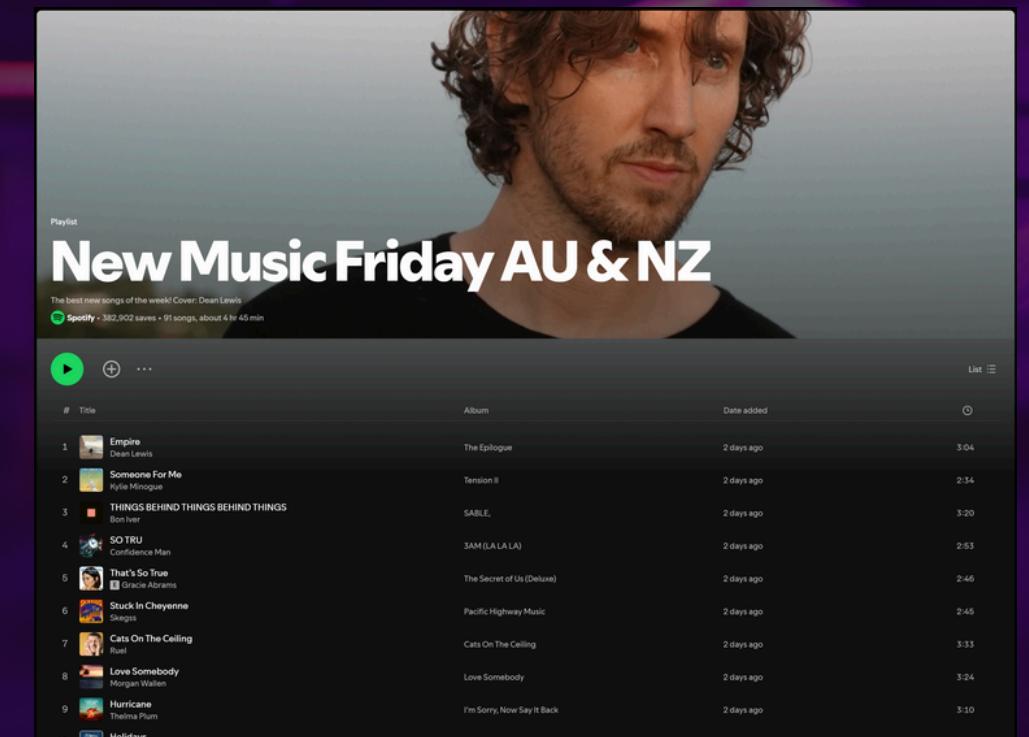
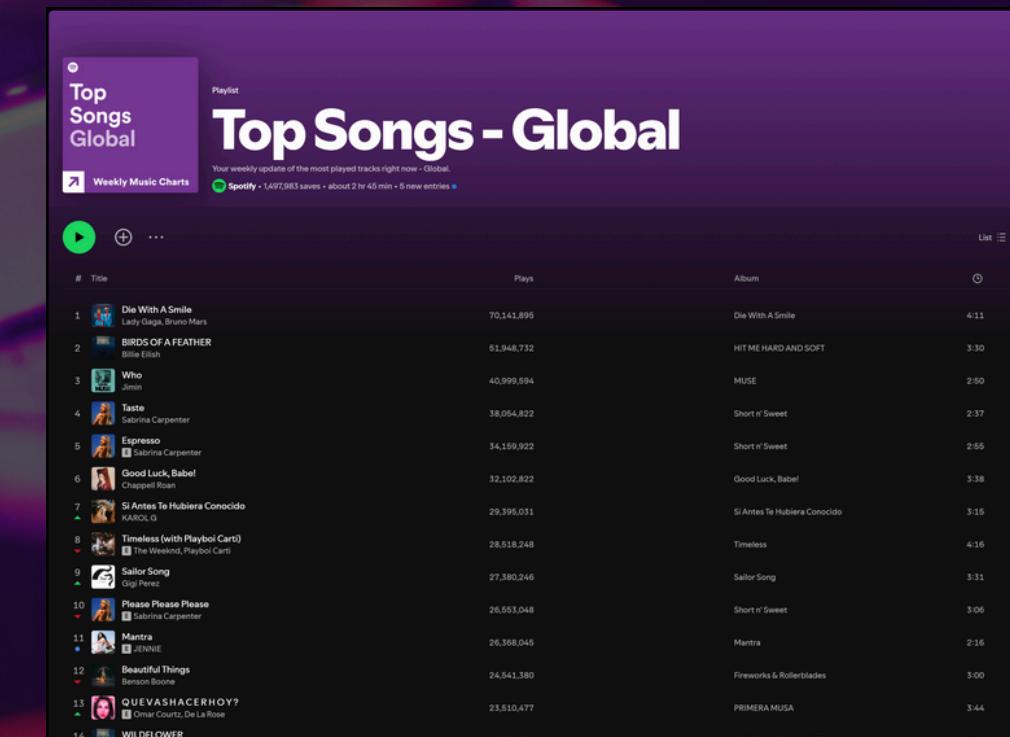
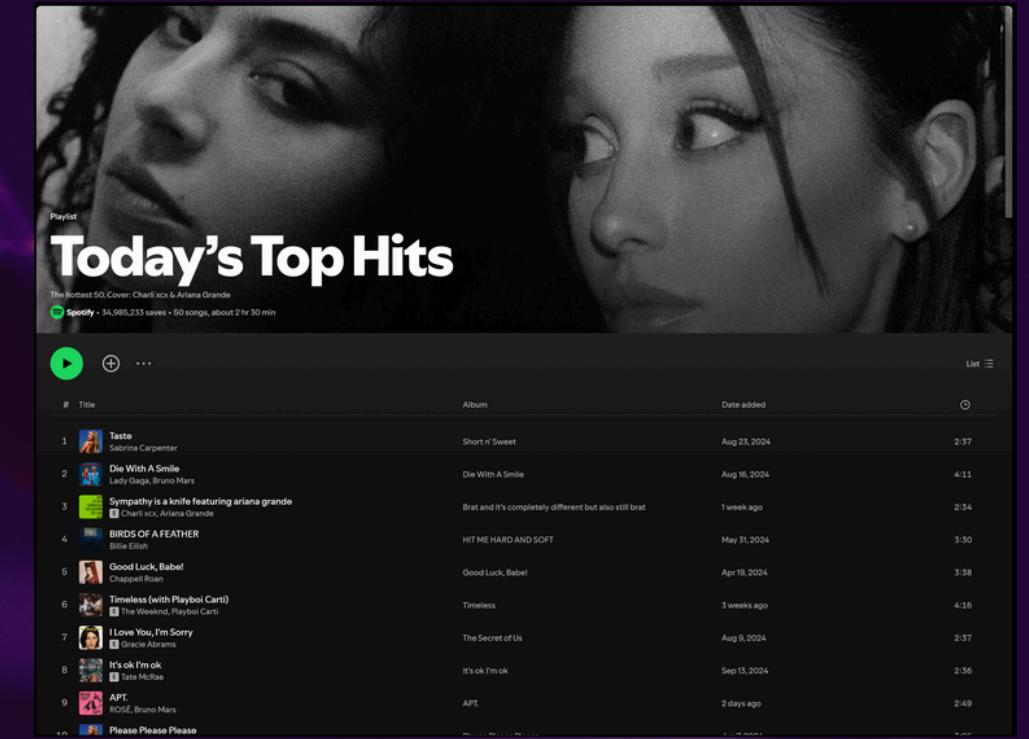
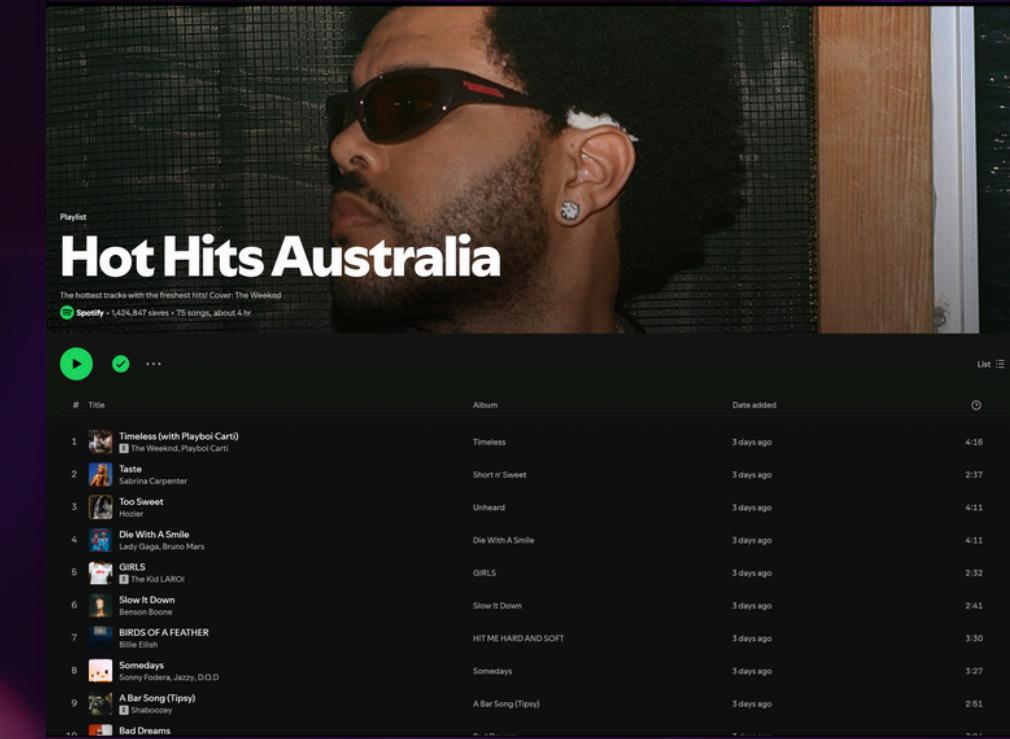
Performing K-Means Clustering with the determined k



# Artists and Genres Network analysis

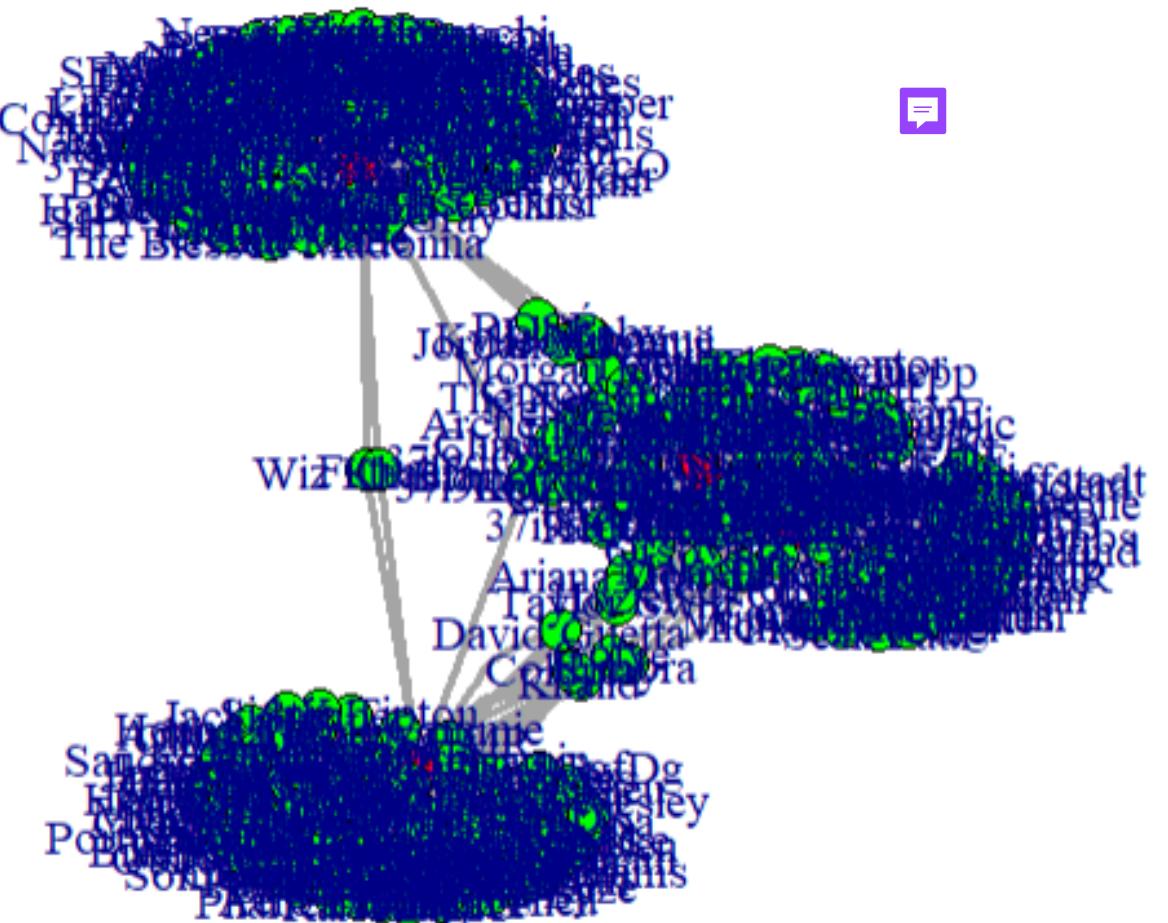
## PLAYLISTS DATA

How artists and genres shape the playlist ecosystem on Spotify?



# ARTISTS & PLAYLISTS

# Spotify Playlist to Artist Network



**Artists** and **Playlists** are represented as nodes

**Artists** act as bridges that connect playlists.

**Artists** with high betweenness scores act as bridges or connectors between multiple playlists





# Hypothesis Test

**Null hypothesis ( $H_0$ ):** There is no linear relationship between the acousticness and popularity

**Alternative hypothesis ( $H_1$ ):** There is a linear relationship between the acousticness and popularity

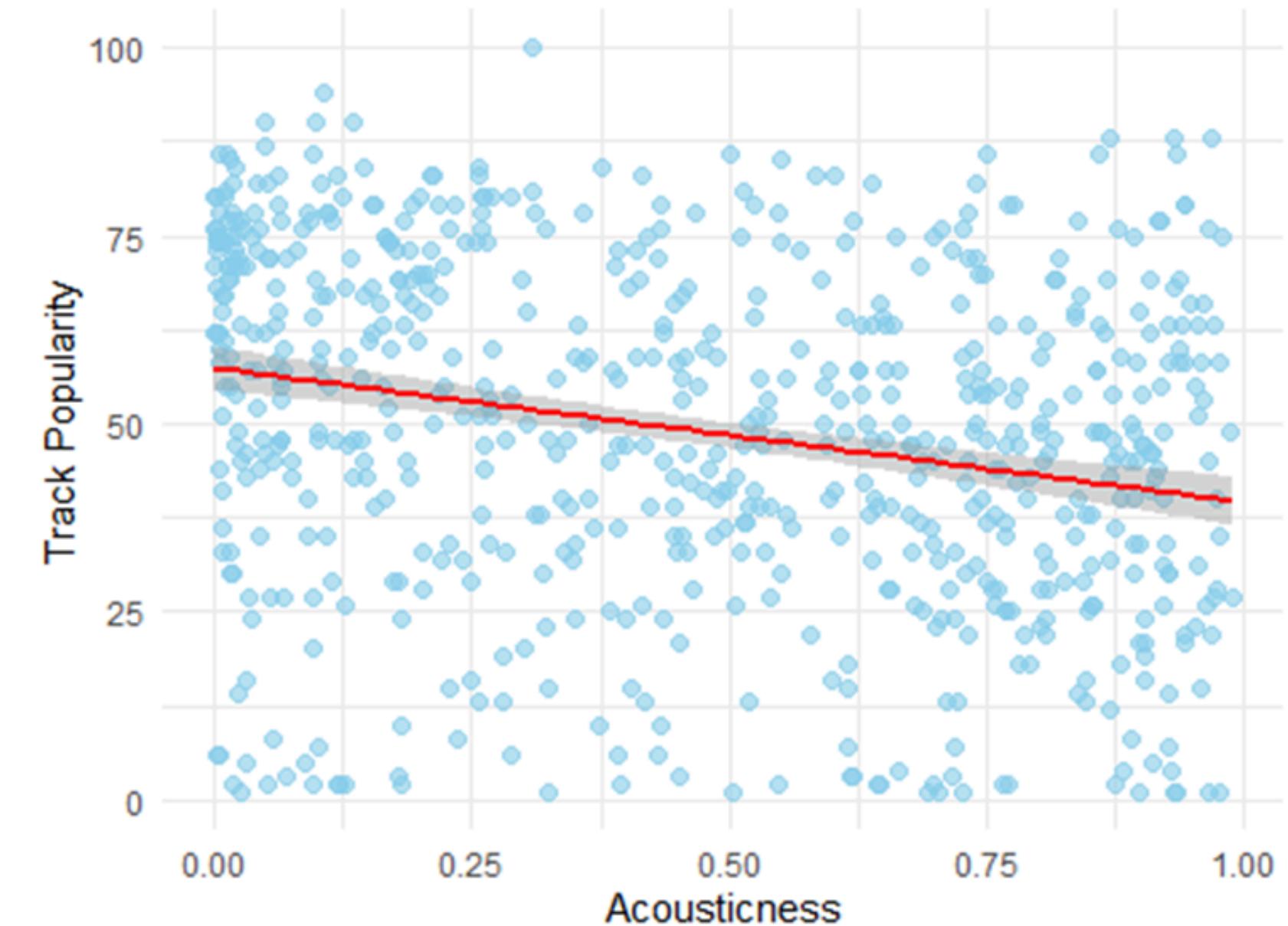
# Key findings



Tracks with higher acousticness scores are generally less popular.

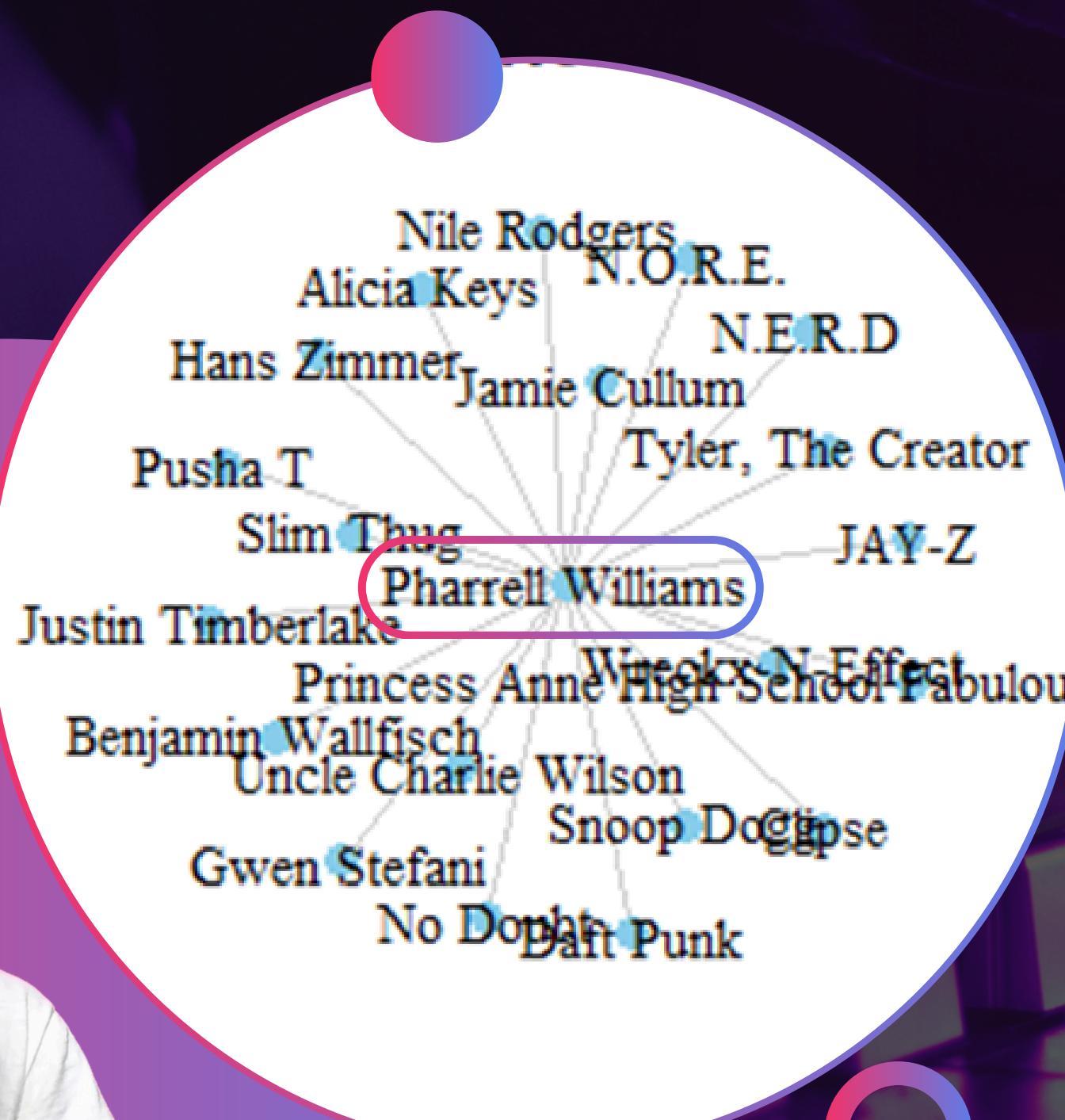


Track Popularity vs Acousticness





# Artist Collaboration Network



## Phase 1

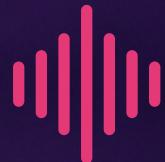
Start from a defined artist and collect his 20 collaborative artists by extracting from his albums information



## Phase 2

Then broadening the network as the same approach to figure out the insight

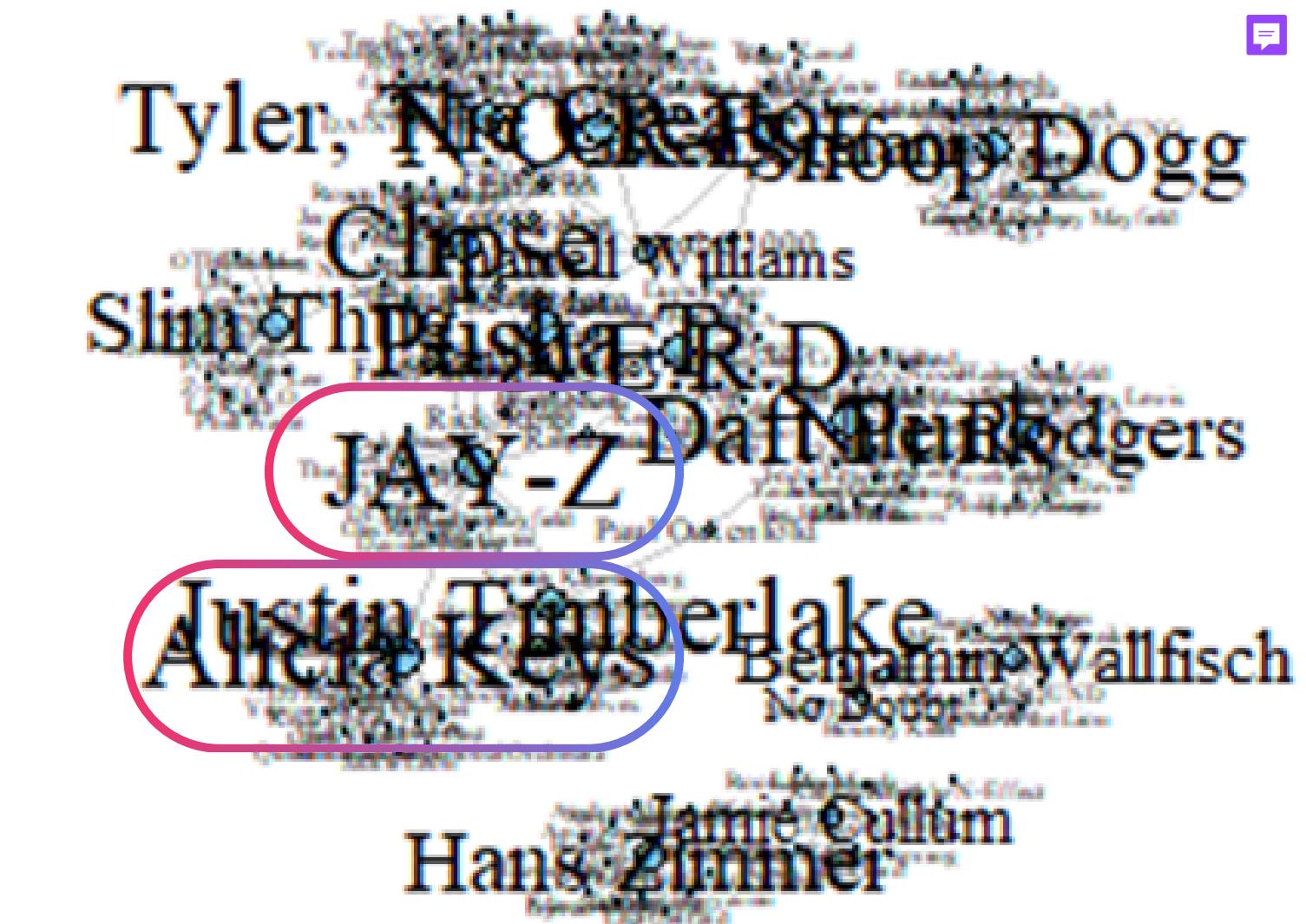
# Key Findings



Alicia Keys and Jay-Z have the most connections, indicating their influence and popularity through frequent collaborations



Those with more connections can potentially leverage these collaborations for even broader audience reach.





# THANK YOU

For Your Attention