

## Objective

EXPLORE addresses the gap in current music recommendations paradigms by providing users with high-quality **interpretable recommendations** as well as accessibility to **control** them as per their **mood**

## Approach

### Data

We are using the MLHD dataset consisting of track listening event history of **1k** users and **900k tracks** amounting to **2.5 GB**. For new users, we use the Spotify API to extract their public playlists.

### Recommendation Algorithm

Employing **collaborative filtering**, we are able to find the sweet spot of recommendation performance and interpretability. Given the sheer number of songs, we chose user-user collaborative filtering to simplify computations. For visibility into our recommendations, we show the **closest neighbour profiles** to help users better understand our recommendation

Additionally, since we use the Music Listening History Dataset we use **inference to create user-song ratings**. The rating is based on the following: Song Frequency, Song Popularity, Recency Bias and User-Platform affinity.

### Integration

Our product starts its lifecycle from pinging the Spotify API and ends with creating visualization in **Tableau**

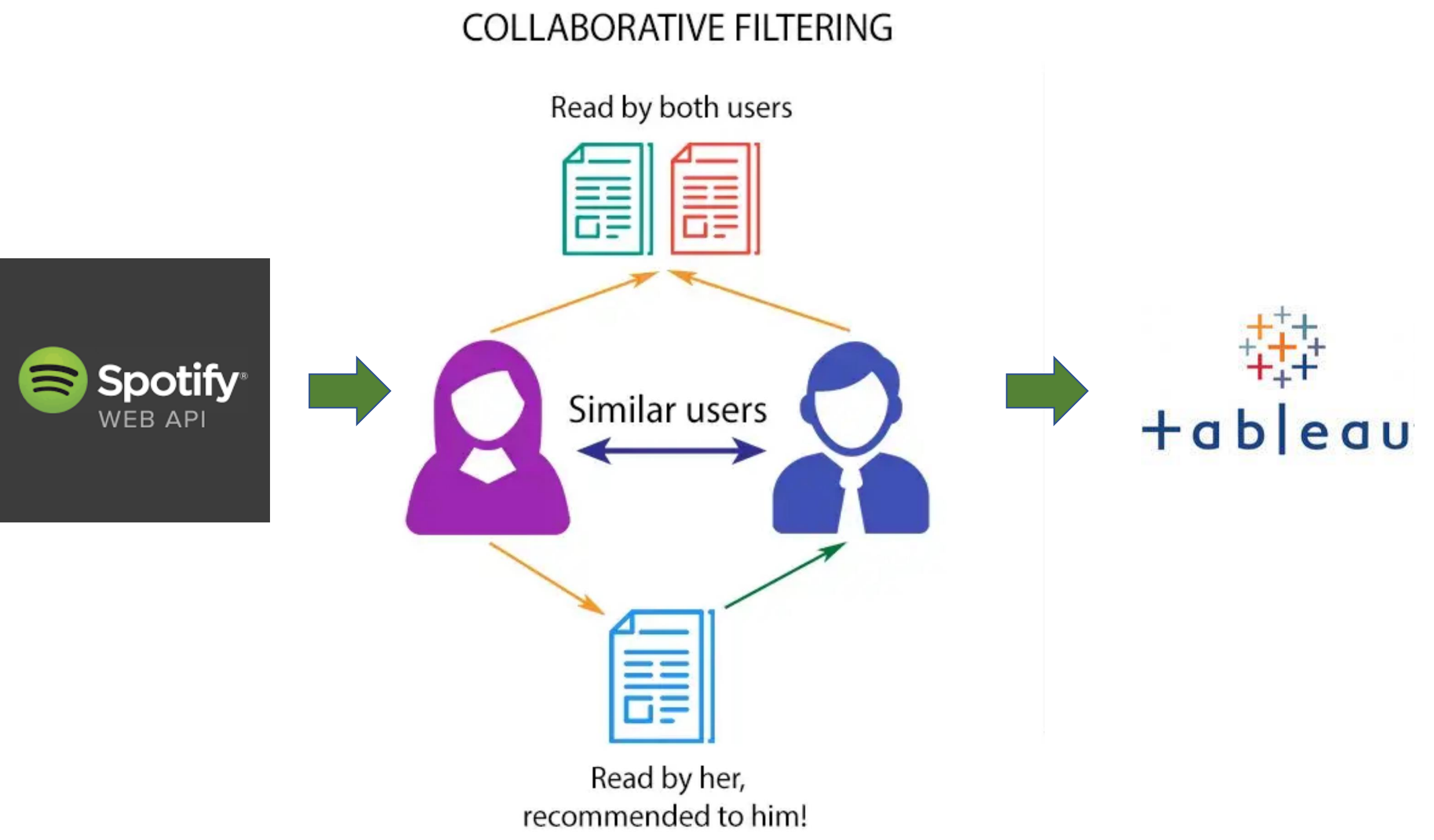


Fig 1: Project Workflow

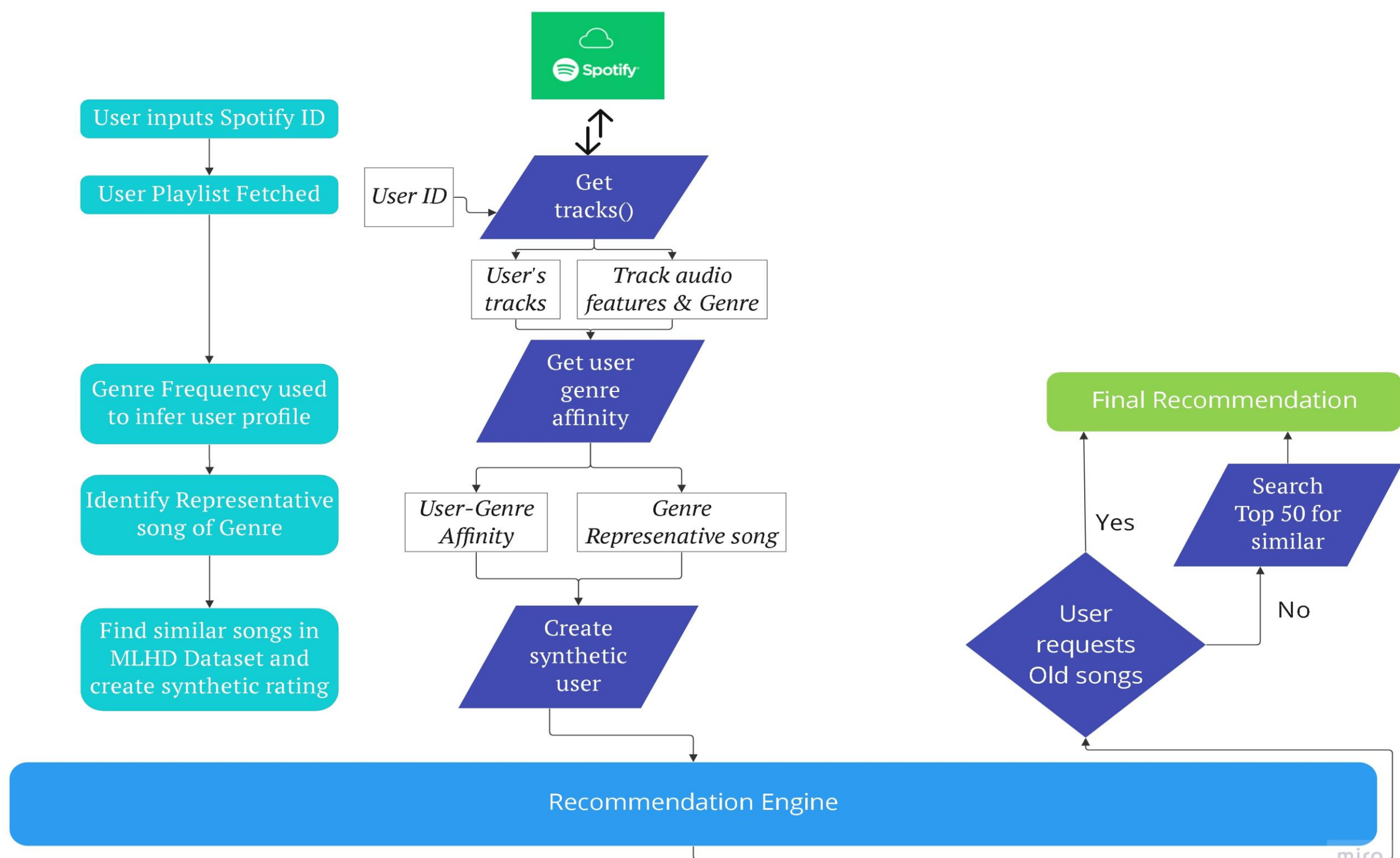


Fig 2: Data Workflow: Using User's Spotify ID to generate personalized recommendations

## Product Visualization

Tableau Dashboard takes Spotify User ID as an input and generates the following:

- Your Personalized Playlist: **Recommended Playlist**
- Your Personalized List Looks like: **Characteristics of Suggestions**
- Mood Controls: **Toggles for Mood Control**

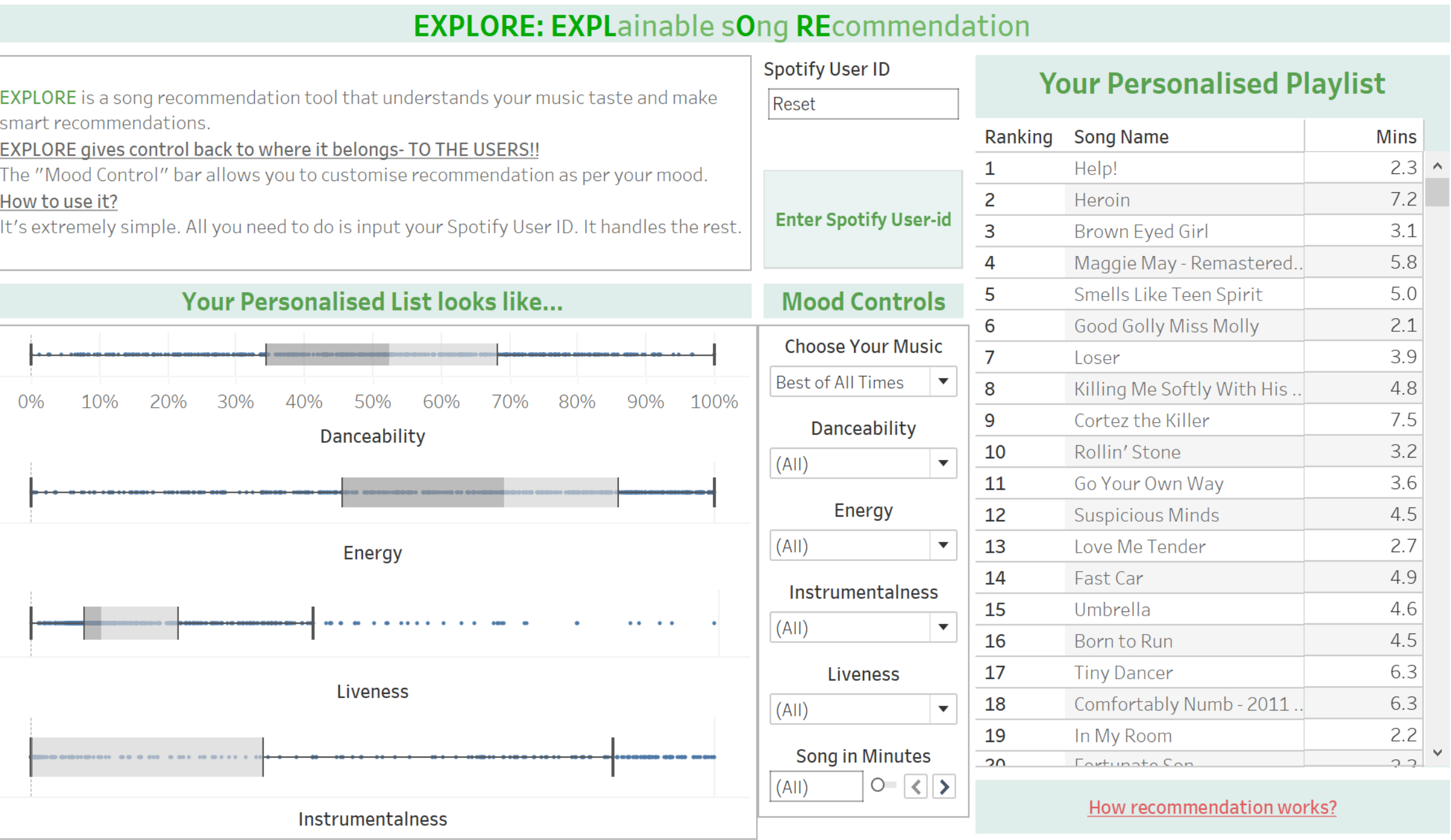


Fig 3: EXPLORE UI: Personalized recommendations with playlist attribute visualization and mood controls

The User Neighbour Network Graph **projects neighbors spatially based on similarity**. Clicking on a neighbor node, the user gets immediate access to compare their music taste to that of the neighbor

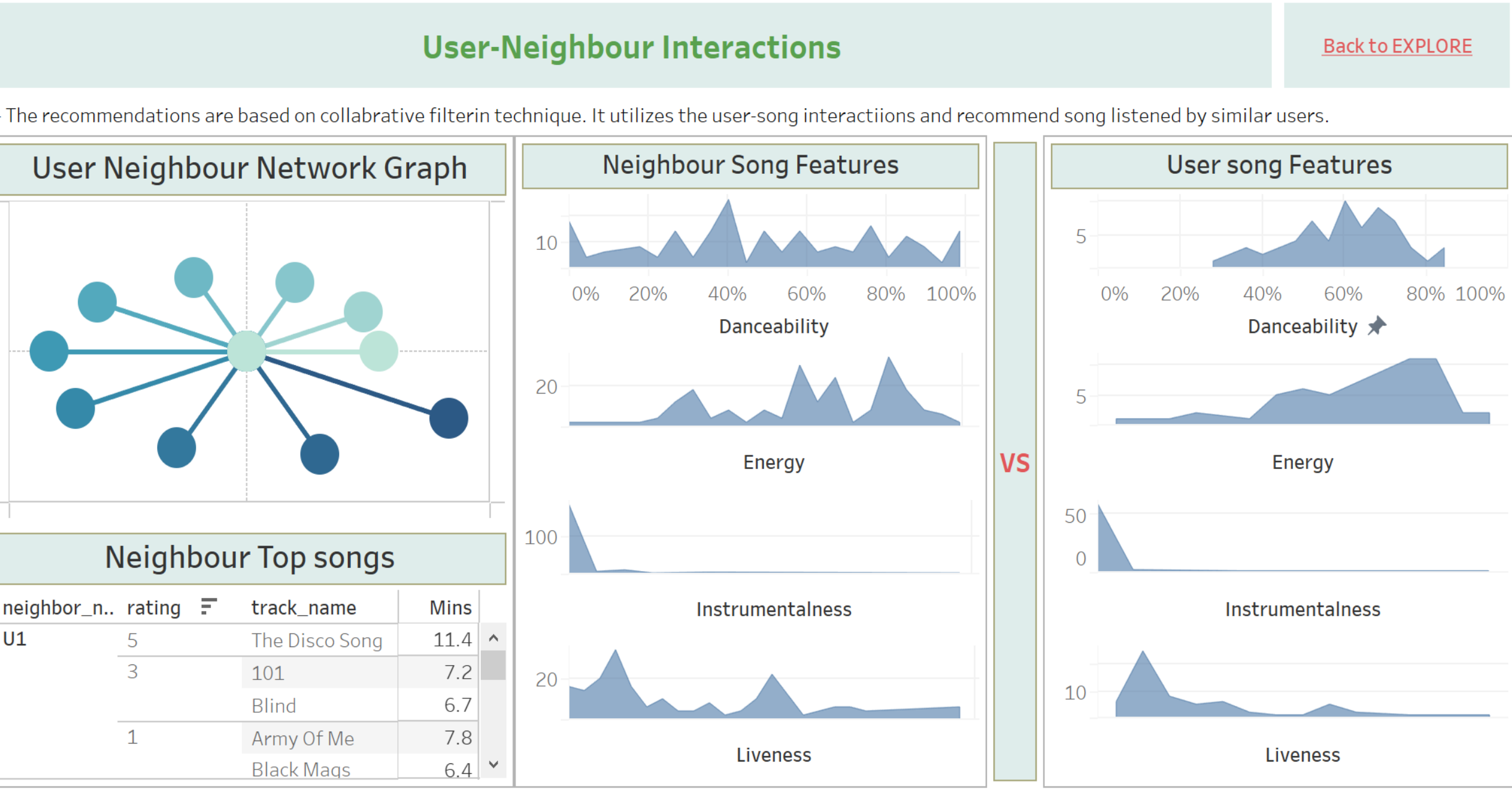


Fig 4: User Neighbour Interactions: Visualizing closest neighbors of a user and comparing their music taste

## Evaluation

- Metrics:** 1. **Deviation** (MSE and RMSE – RMSE ranges from 1 to 4, 1 being the ideal value) 2. **Ranking Based** (MAP@ k and NDCG)
- Train, Test Split (80%, 20%):** Stratified vs Random Sampling
- Results:**

### 1. Deviation Measures:

Train Test Split	Mean Squared Error (Test)	Root Mean squared Error (Test)
Stratified	2.407	1.551
Random Sampling	2.538	1.593

### 2. Ranking Measures: Stratified sampling for ranking metrics

Train Test Split	MAP@K (K = 3)	Mean NDCG
Stratified	0.773	0.873

Our recommendation does a fine job in not only predicting ratings but also in fetching the most relevant songs emphasized by MAP@k and mean NDCG score.