

Project Proposal: Spotify Artist Network

Steven Canon-Almagro, Nathanael Huh

1. Nodes and Links

For this project the nodes are song artists. The edges are created between artists through songs and albums. Artists work together through featured tracks and productions on Albums. This network draws inspiration from the actor network.

2. Dataset

We will be using the Spotify API to get artist and song data. Spotify uses the OAuth 2.0 framework. To gain an access token we need a Client ID and Client Secret, this can be obtained by creating an account with Spotify for Developers. The access token then allows us access to the other end points in the API where we can then query the Spotify database and get data in return. Using the query function. We plan to retrieve the top 500 artists, featured songs, and albums. The artists will act as the initial set of artists in the network. Featured songs and albums will act as both an edge guide and any nodes that aren't present in the first dataset.

Artists like Kanye may pose a problem as they don't put features in the track name. Now it's still unknown how exactly Spotify obtains their information, if the artist provides the information, going through the record label or any other method.

We are trying to keep the number of artists low as the data sets will get extremely large because artists have many songs. For example, Drake has around 10-14 albums with 10-25 songs in each, taking the most conservative number Drake alone would have 100 songs just for himself. That is also why we also just want to scrape tracks that have a feature.

If an artist works with another artist multiple times, we can make that edge worth more.

My plan is to get the data and then create a quick program to save that to a csv file.

The csv file then will be used in create the network.

3. Expected size of the network

As of 2022 there are 11 million artists on Spotify. Our goal is to scrape the top 500 artists and map out the connection they have with their collaborators. We are only scraping 500 artists from Spotify, but there will be much more information than that as the artist can work with multiple artists outside the top 500.

4. Questions you plan to ask and why we care

Is the music industry just as connected as the movie industry. What is the most well-connected artist on the Spotify platform? Does the music industry have their own Kevin Bacon? Do the top 500 artists collaborate amongst themselves more than with artists outside the top 500?