W205 team project: When pop song meets big data

We believe that it is possible to predict if a song will be a hit song or not based on data about the song.

# **Background research:**

In the past, researchers have tried to study if a song will be popular based on the acoustic characteristic about the song, such as song length, wave pattern, pitch pattern and so on [ref<http://scoreahit.com/TheHitEquation> ]. We believe a more accurate result will be obtained if more data about the song are included.

# **Data sources**

We’ve researched places where we can obtain data about songs, here are a list of sources we’ve looked at:

Sound Cloud(<https://developers.soundcloud.com/docs/api/reference#comments> )

YouTube(<https://developers.google.com/youtube/v3/docs/videos> )

Twitter(<https://dev.twitter.com/rest/reference/get/search/tweets> )

iTunes(<https://www.apple.com/itunes/affiliates/resources/documentation/itunes-store-web-service-search-api.html> )

lastfm(<http://www.last.fm/api> )

facebook likes for singer or band(<http://stackoverflow.com/questions/9728279/getting-the-facebook-like-share-count-for-a-given-url> )

After considering the scope of the project, we decided to obtain the following data:

## **Data used as input:**

Song entity:

Name, released date, format, genre, length, label, writers, producers, album, singers, YouTube Views

Singer entity:

Name, Birth date, Genres, Years active, Labels

## **Data used as output:**

Awards, Nominations, Weekly charts position, Year-end charts position

\* Note: definition of a hit song is from<https://en.wikipedia.org/wiki/Hit_single>

A hit single is a recorded song or instrumental released as a single that has become very popular. Though it sometimes means any widely played or big-selling song, the term "hit" usually refers to a single that has appeared in an official [music chart](https://en.wikipedia.org/wiki/Record_chart) through repeated [radio airplay](https://en.wikipedia.org/wiki/Airplay_(song)) or significant commercial sales.[[1]](https://en.wikipedia.org/wiki/Hit_single#cite_note-dcf-229-1)