**Project Title: Spotify Ratings vs. Ticket Sales**

**Team Members:**

* Ray Speakmore
* Brent Lee
* DeWitt Tsai
* Arthur Adjamoglian

**Project Description/Outline:**

Scenario: A particularly new record label firm is running into an issue that has been plaguing the music industry: it is difficult to project an accurate expected head-count for a venue selection for new artists.

To expand on this issue, a large challenge that they’re continuously facing is booking a venue that is either too small for a large turn-out or too large for a small turn-out. Either scenario results in losing a large amount of profits as well as damaging the artist’s and the company’s reputation. While the record label firm acknowledges that certain scenarios can turn out to be a positive result, they’d like to request appropriately sized venues for concerts that result in the smallest possible difference between “total available tickets” to “population count of people interested in participating.”

To determine if a solution is available, the firm has used their budget to commission a third person research company to research if there is a correlation, for a new artist, between their Spotify popularity to their expected ticket turn-out rate.

The team has determined a preliminary plan: For a specific artist(s), gather data from Spotify API for their follower amount, popularity, and genre, and then compare that data to reports that publicize the total tickets sold for their concert(s). After gathering enough data, the goal is to run an analysis to see if there are any insightful correlations that can be theorized between the data sets.

The research company has provided that the ultimate goal is to provide insight that can help the record label company to estimate an accurate projection of ticket sales for their artists.

**Research Questions to Answer:**

* While it can be expected that a popular artist on Spotify will sell a large amount of tickets, what does this mean for different genres or for non Top 500 music artists?
* For artists will a low following and/or popularity, is the ratio to the ticket turn-out consistent or unpredictable?
* For artists, is there a certain level of popularity on Spotify that has a more consistent turn-out rate for ticket sales?

**Datasets to be Used:**

* Spotify API
* Pollstar data results from 2018 ratings
* Other data sources that contain ticket sales for specific artists

**Rough Breakdown of Tasks:**

1. First figure out how to grab data from Spotify API
2. They use artist ID, but how do we get an artist's ID?
3. Google search their name with Spotify. Scrap the first listing URL which contains the artist ID at the end.
4. Spotify's guide specifically gives examples in HTML only. We use curl.trillworks.com to convert it into a Python format.
5. We determine which artists to research based off of data pulled from websites that list the concert ticket sales. We would pull ticket counts from popular to unpopular artists, then find their artist ID on Spotify, and then find their popularity rating on Spotify.