**Project Title: Using ETL to determine the most popular artist and songs of 2018**

**Team Members:**

Addison Brown

Ashley Cleveland

Jessica Jordan

**Project Description/Outline:**

This project is to explore one of the most popular music streaming platforms, Spotify. We will be extracting, transforming, and loading data around artist popularity, song popularity, and most popular genres among Spotify users.

**Sources of Data:**

Kaggle: Spotify Track DB

Kaggle: Top Spotify Tracks of 2018

**Data Cleanup & Analysis:**

1. Extract data files into a dataframe using Python. (JJ)
2. Transform the dataframes (JJ)
   1. Create a copy of the file
   2. Select the desired columns
   3. Rename the columns
   4. Clean up the data (remove any duplicates)
3. Load the Data: (AB)
   1. Using Postgres, create the database (spotify\_db) and raw tables for data to be loaded into.
   2. Run the python file to create the connection, confirm the tables and load the dataframes into the database.
4. Create a final technical report with the above information and steps required to reproduce your ETL process. (AC)

**\*\*\*\*\*\*\*Please use the below guidelines for the final report (this section can be deleted upon completion\*\*\*\*\*\*\*\*\*\*\*\***

***Project Report***

***At the end of the week, your team will submit a Final Report that describes the following:***

* ***Extract: your original data sources and how the data was formatted (CSV, JSON, pgAdmin 4, etc).***
* ***Transform: what data cleaning or transformation was required.***
* ***Load: the final database, tables/collections, and why this was chosen.***

***Please upload the report to Github and submit a link to Bootcampspot.***