**Title -** title of the song.

**Rank -** rank of the song, determined by number of streams. This is a numerical value.

**Date -** the date in which the data was published (2017-01-01 format).

**Artist -** the artist that made the song.

**Url -** each URL is a link to the song on spotify.

**Region -** region corresponds to the songs region in which the song was charted(ranked). This is a string of characters..

**Chart** - spotify has multiple charts for ranking songs, all of this data set includes the ‘Top 200’ songs by rank on the platform as well as ‘Viral 50’ which tracks trending songs.

**Trend -** trend is the relative position of the song (moving up, moving down, same position, or new entry). This can change depending on streaming numbers (which is a continuously changing variable). Categorical variable.

**Streams -** streams are what determine the song's ranking. Each numerical value of a song's streaming value corresponds to 1 play of the song by a Spotify user. This is a numerical value.

**Track\_id -** every song in the data set has a corresponding track identifier, which is unique to each song, this is how spotify keeps track of each unique track in their database. This is a string of characters.

**Album -** corresponds to the album in which the track belongs to. An album can comprise of many songs, all made by the same artist.

**Popularity -** this is a value from 0-100 that rates popularity of a song. This could be from liking a song on spotify and calculating the average value amongst all likes.

**Duration\_ms -** this is the length of the song in milliseconds.

**Explicit -** determinesif a song has any explicit language. This is a boolean variable (can either be true or false). Also can be classified as a categorical variable.

**Release\_date -**  the date in which each track was released. Some of the entries only have a year (no day or month).

**Available\_markets -** specifies each region that the song is available for streaming in. Each country is denoted with a two character abbreviation.

**Af\_danceability -** describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable. The duration of the track in milliseconds. - from spotify

**Af\_energy -** Energy is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy. - from spotify

**Af\_key -** In music, a key is the primary major or minor scale that a piece of music centers around. A song in a major key is constructed from a major scale, while a song in a minor key is built from a minor scale. These values are scored from 0.0-11.0. - from spotify

**Af\_loudness -** The overall loudness of the track in decibels (dB). - from kaggle

**Af\_mode -** The modality of the track (1 = major, 0 = minor). - from kaggle

**Af\_speechiness -** The presence of spoken words in the track. Numerical value. - from kaggle

**Af\_acousticness -** The acousticness of the track. Numerical value. - from kaggle

**Af\_instrumentalness -** The probability of the track being instrumental. - from kaggle

**Af\_liveness -** A measure of the presence of a live audience in the track. - from kaggle

**Af\_valence -** The musical positiveness or happiness conveyed by the track. - from kaggle

**Af\_tempo -** The tempo of the track in beats per minute (BPM). - from kaggle

**Af\_time\_signature -** An estimated time signature. The time signature (meter) is a notational convention to specify how many beats are in each bar (or measure). The time signature ranges from 3 to 7 indicating time signatures of "3/4", to "7/4". - from spotify developers